

Dr. Ing.

## PhD Thesis

# Discontinuous propagating fronts: linear and nonlinear systems

Submitted by  
Hassan Yousefi

BAUHAUS-UNIVERSITÄT WEIMAR  
Faculty of Civil Engineering  
Institute of Structural Mechanics

Supervisor: Prof. Dr.-Ing. Timon Rabczuk





# DISCONTINUOUS PROPAGATING FRONTS: LINEAR AND NONLINEAR SYSTEMS

Dissertation

As required for the conferral of the academic degree

Dr. Ing.

at the Faculty of Civil Engineering  
of the  
BAUHAUS-UNIVERSITÄT WEIMAR

Submitted by

Dr. Hassan Yousefi

Date of birth: September 1978

Place of birth: Tehran, Iran

Weimar

Supervisor: Prof. Dr. Ing. Timon Rabczuk,  
Bauhaus-Universität Weimar

Assessors:

1. Prof. Dr. rer. nat. habil. Klaus Gürlebeck,  
Bauhaus-Universität Weimar
2. Prof. Dr. Stephane Bordas, Université du Luxembourg

Date of defence: September 19th, 2022

## Declaration of Authorship

I hereby solemnly declare that the following dissertation is my own work and that no impermissible assistance from others or references other than those specifically cited were used in its making. Data and/or concepts taken directly or indirectly from other sources have been properly referenced.

I affirm that no individuals were involved in producing the content of this dissertation. Furthermore, no placement or consulting services (promotion consultants or other persons) were paid to assist me in any way. I affirm that no one received direct or indirect pecuniary compensation or payment in kind for work conducted in connection with the content of this dissertation.

This dissertation has not been previously submitted in the same or similar form to any other examination authority in Germany or abroad.

I certify that, to the best of my knowledge, the declaration above is absolutely true and nothing has been concealed.

Place, Date Weimar, September 21th, 2022

Signature



# Acknowledgement

I would like to express my special appreciation and thanks to my supervisor Prof. Dr.-Ing. Timon Rabczuk for his kind support of my Ph.D. study and related research, for his patience, immense knowledge, and confidence towards me. His encouragement always motivated me to advance, and his valuable advice, excellent guidance and help in every stage of my progress are indispensable to the finish of my dissertation.

I am also indebted to the members of my dissertation committee, Professor Klaus Gürlebeck and Professor Stephane Bordas for generously offering their time, and also for their critical review of my dissertation, insightful comments and detailed remarks.

A special thanks to my family. I would like to express my deepest gratitude to my mother and father.

Hassan Yousefi  
Weimar, Germany  
September 2022

# Abstract

The aim of this study is controlling of spurious oscillations developing around discontinuous solutions of both linear and non-linear wave equations or hyperbolic partial differential equations (PDEs). The equations include both *first-order* and *second-order* (wave) hyperbolic systems. In these systems even smooth initial conditions, or smoothly varying source (load) terms could lead to discontinuous propagating solutions (fronts). For the first order hyperbolic PDEs, the concept of *central high resolution schemes* is integrated with the *multiresolution-based adaptation* to capture properly both discontinuous propagating fronts and effects of fine-scale responses on those of larger scales in the multiscale manner. This integration leads to using central high resolution schemes on non-uniform grids; however, such simulation is unstable, as the central schemes are originally developed to work properly on uniform cells/grids. Hence, the main concern is stable collaboration of central schemes and multiresolution-based cell adapters. Regarding central schemes, the considered approaches are: 1) Second order central and central-upwind schemes; 2) Third order central schemes; 3) Third and fourth order central weighted non-oscillatory schemes (central-WENO or CWENO); 4) Piece-wise parabolic methods (PPMs) obtained with two different local stencils. For these methods, corresponding (nonlinear) stability conditions are studied and modified, as well. Based on these stability conditions several limiters are modified/developed as follows: 1) Several second-order limiters with total variation diminishing (TVD) feature, 2) Second-order uniformly high order accurate non-oscillatory (UNO) limiters, 3) Two third-order nonlinear scaling limiters, 4) Two new limiters for PPMs. Numerical results show that adaptive solvers lead to cost-effective computations (e.g., in some 1-D problems, number of adapted grid points are less than 200 points during simulations, while in the uniform-grid case, to have the same accuracy, using of 2049 points is essential). Also, in some cases, it is confirmed that fine scale responses have considerable effects on higher scales. In numerical simulation of nonlinear first order hyperbolic systems, the two main concerns are: *convergence* and *uniqueness*. The former is important due to developing of the spurious oscillations, the numerical dispersion and the numerical dissipation. Convergence in a numerical solution does not guarantee that it is the physical/real one (the uniqueness feature). Indeed, a nonlinear systems can converge to several numerical results (which mathematically all of them are true). In this work, the convergence and uniqueness are directly studied on non-uniform grids/cells by the concepts of *local numerical truncation error* and *numerical entropy production*, respectively. Also, both of these concepts have been used for cell/grid adaptations. So, the performance of these

concepts is also compared by the multiresolution-based method. Several 1-D and 2-D numerical examples are examined to confirm the efficiency of the adaptive solver. Examples involve problems with convex and non-convex fluxes. In the latter case, due to developing of complex waves, proper capturing of real answers needs more attention. For this purpose, using of method-adaptation seems to be essential (in parallel to the cell/grid adaptation). This new type of adaptation is also performed in the framework of the multiresolution analysis.

Regarding second order hyperbolic PDEs (mechanical waves), the regularization concept is used to cure artificial (numerical) oscillation effects, especially for high-gradient or discontinuous solutions. There, oscillations are removed by the regularization concept acting as a post-processor. Simulations will be performed directly on the second-order form of wave equations. It should be mentioned that it is possible to rewrite second order wave equations as a system of first-order waves, and then simulated the new system by high resolution schemes. However, this approach ends to increasing of variable numbers (especially for 3D problems). The numerical discretization is performed by the *compact finite difference* (FD) formulation with desire features; e.g., methods with *spectral-like* or *optimized-error* properties. These FD methods are developed to handle high frequency waves (such as waves near earthquake sources). The performance of several regularization approaches is studied (both theoretically and numerically); at last, a proper regularization approach controlling the Gibbs phenomenon is recommended. At the end, some numerical results are provided to confirm efficiency of numerical solvers enhanced by the regularization concept. In this part, shock-like responses due to local and abrupt changing of physical properties, and also stress wave propagation in *stochastic*-like domains are studied.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.0.1	Outline of the thesis . . . . .	5
<b>2</b>	<b>Literature review and general descriptions</b>	<b>7</b>
2.1	Multiresolution (wavelet) based simulation of partial differential equations (PDEs) . . . . .	7
2.2	First-order nonlinear hyperbolic (wave) equations . . . . .	10
2.2.1	A practical approach for using high resolution schemes on non-uniform grids . . . . .	12
2.2.2	Central high resolution schemes and multiresolution analysis . . .	13
2.2.3	Numerical errors, numerical entropy and grid-adaptation . . . . .	14
2.2.4	Non-convex fluxes . . . . .	15
2.2.5	High-order central high resolution schemes . . . . .	15
2.2.6	Non-polynomial reconstruction over each cell . . . . .	17
2.2.7	Conclusion . . . . .	17
2.3	Second order waves . . . . .	19
<b>3</b>	<b>Multiresolution analysis and grid adaptation</b>	<b>24</b>
3.1	Multiresolution representation of data . . . . .	24
3.1.1	Multiresolution representation of 1-D grids . . . . .	24
3.1.2	Multiresolution representation of 2-D grids . . . . .	28
3.2	Post-processing adapted grids based on the MRA . . . . .	30
3.2.1	1-D grid modification . . . . .	30
3.2.2	2-D grid modifications . . . . .	31
3.2.3	Post-processing 2-D adapted grids . . . . .	33
<b>4</b>	<b>The KT central scheme on non-uniform grids</b>	<b>34</b>
4.1	The KT central scheme on non-uniform grids . . . . .	35
4.1.1	Godunov-type schemes . . . . .	35

Contents

4.1.2	Central high resolution schemes . . . . .	38
4.1.3	Central high-resolution schemes on non-centered non-uniform 1-D cells . . . . .	42
4.1.4	Formulation for the 2-D semi-discrete form . . . . .	46
4.1.5	Systems with source and diffusion terms . . . . .	48
4.2	Nonlinear stability: the TVD condition . . . . .	49
4.2.1	Local TVD conditions . . . . .	52
4.2.2	Irregularity effects on slope limiters . . . . .	56
4.2.3	Slope limiters on cell-centered non-uniform grids . . . . .	59
4.2.4	The performance of the generalized MINMOD limiter on non- uniform grids . . . . .	61
4.3	Wavelet-based adaptive-grid method to solve PDEs . . . . .	70
4.4	Examples . . . . .	71
4.4.1	1-D second order hyperbolic PDEs . . . . .	91
<b>5</b>	<b>Special topics</b>	<b>110</b>
5.1	Central and central-upwind second order schemes . . . . .	110
5.2	Numerical entropy productions . . . . .	112
5.3	Local truncation errors . . . . .	114
5.3.1	One dimensional systems . . . . .	114
5.3.2	Two-dimensional problems . . . . .	117
5.4	Other second-order slope limiters . . . . .	118
5.4.1	Slope limiters on uniform cells: the TVD ones . . . . .	119
5.4.2	Slope limiters on non-uniform cells: the TVD ones . . . . .	121
5.4.3	The UNO limiter on non-uniform cells: cell-centered and non-cell- centered ones . . . . .	129
5.4.4	The Jameson slope limiter, a TVB limiter . . . . .	135
5.5	Non-convex problems . . . . .	138
5.6	Numerical examples . . . . .	138
5.6.1	Non-convex example: a scalar system . . . . .	150
<b>6</b>	<b>Higher order central and CWENO schemes</b>	<b>158</b>
6.1	Central third order high resolution schemes on non-uniform grids . . . . .	159
6.1.1	The fully-discrete form . . . . .	159
6.1.2	The semi-discrete form . . . . .	163
6.1.3	The preliminary construction . . . . .	165

6.1.4	Designing of a limiter for the third-order central high resolution scheme . . . . .	170
6.2	CWENO schemes on non-uniform grids . . . . .	173
6.2.1	The three point stencil on non-uniform grids . . . . .	174
6.2.2	The five point stencil on non-uniform grids . . . . .	176
6.3	The piecewise parabolic method (PPM) . . . . .	178
6.3.1	Formulation for the first approach of PPM (PPM-1) . . . . .	178
6.3.2	Formulation for the second approach of PPM (PPM-2) . . . . .	183
6.4	Complexity . . . . .	184
6.4.1	The complexity of high-order central high resolution schemes . . . . .	185
6.4.2	Complexities of the different third-order central high resolution schemes on adapted grids/cells . . . . .	186
6.5	Reconstruction Error . . . . .	188
6.5.1	The first test problem: the smooth-high gradient function . . . . .	188
6.5.2	The second test problem: a function including a discontinuity . . . . .	190
6.6	Numerical examples . . . . .	192
6.6.1	<b>The 1-D Burgers' equation</b> . . . . .	194
6.6.2	<b>Euler system of equations</b> . . . . .	195
6.6.3	<b>The 2-D scalar Burgers' equation</b> . . . . .	207
6.6.4	<b>2-D Euler systems</b> . . . . .	210
<b>7</b>	<b>Discontinuous solutions of second-order wave equations and the regularization concept</b>	<b>223</b>
7.1	Regularization approaches . . . . .	224
7.2	Numerical simulation of 1-D wave problems via regularization methods . . . . .	227
7.3	Higher order finite difference methods and different filtering approaches . . . . .	232
7.4	Studying the Thikhonov regularization with different constraints by the SVD and GSVD decompositions . . . . .	237
7.5	Implementation algorithms . . . . .	242
7.6	Errors and convergence rates . . . . .	246
7.7	Conservation in Tikhonov-based smoothing . . . . .	250
7.8	Global and local smoothing . . . . .	253
7.9	Tikhonov regularization and correspondence with filtering . . . . .	256
7.10	Solution algorithm for stress wave equations . . . . .	257
7.11	Numerical examples . . . . .	258



<b>8 Conclusion and future works</b>	<b>269</b>
8.1 Conclusion . . . . .	269
8.1.1 First order nonlinear hyperbolic (wave) systems . . . . .	269
8.1.2 Second order nonlinear wave (hyperbolic) systems . . . . .	271
8.2 Future works . . . . .	272
8.2.1 First order waves . . . . .	272
8.2.2 Second order waves . . . . .	273
<b>A Central schemes on uniform grids</b>	<b>274</b>
A.1 Central High Resolution Schemes . . . . .	274
A.1.1 Continuous assumption for the state variable and flux . . . . .	275
A.1.2 Discontinuous variation of state variables and fluxes . . . . .	277
<b>B The NVSF upwind scheme</b>	<b>297</b>
<b>C The TVD condition, Entropy functions and E-schemes</b>	<b>299</b>
C.1 The TVD conditions and limiters . . . . .	299
C.1.1 Global TVD conditions . . . . .	299
C.1.2 Local TVD conditions . . . . .	302
C.1.3 Designing of flux/slope limiters based on the global TVD conditions	303
C.1.4 Relationship between slope and flux limiters . . . . .	306
C.2 Entropy functions and entropy conditions . . . . .	307
C.3 E-schemes and slope limiters . . . . .	308
<b>D Numerical methods for stress waves</b>	<b>311</b>
D.1 The Runge-Kutta $4^{th}$ order for stress wave problems: second order systems	311
D.2 The generalized $\alpha$ -dissipative time integration method . . . . .	312
D.3 The time discontinuous Galerkin method . . . . .	312
D.4 Taylor-Galerkin discretizations . . . . .	313
<b>Bibliography</b>	<b>315</b>
<b>List of Figures</b>	<b>333</b>
<b>List of Tables</b>	<b>343</b>

# Chapter 1

## Introduction

Many of physical phenomena are described by nonlinear hyperbolic PDEs (conservation laws); some of which are: transport phenomena; stress waves in solids (including linear or nonlinear waves); magneto-hydrodynamics (MHD); dam break; Maxwell's equations; gas dynamics; oil reservoir modeling (known as the Buckley-Leverett equations); ice shield and glacier growth modeling; and polymer model. The *transport* phenomenon can describe different physical phenomena, such as: vehicular flow, computer networks, internet traffic, pedestrian dynamics, and transport (traffic) in biology of different scales, including inside a cell-level, cell-level and organism level (e.g., tumor cells, organism colonies, like bacteria & ants) [1, 2]. Even in biological systems the *traffic* phenomenon can occur; it can happen in a cell or between cells and this may lead to different sicknesses; this makes such studying interesting to understand different illness sources. In general, nonlinear conservation laws act as an *universal tool* to describe many physical and industrial phenomena from the molecule (atomistic) level to the continuum scale, as well as multiscale problems [3]

All of the above-mentioned systems have a common feature: *the disturbance which is observed in a certain place in space must propagate with a finite velocity to some other place in this space; as a rule, the process must be close to oscillatory, if it is observed in time* [4]. This moving property (advection feature) is mainly due to presence of hyperbolic characteristic in the systems. In contrast, the parabolic PDEs deal mainly with diffusion problems and elliptic systems express steady state solutions of either parabolic or hyperbolic PDEs [5].

In the hyperbolic systems, localized moving high-gradient fronts develop commonly during propagation. In these problems, discontinuous propagating fronts can easily develop even for smooth initial conditions or smoothly varying sources (loading terms). This is mostly due to the non-linear feature of such systems.

Another important issue in numerical simulation of nonlinear hyperbolic equations is the *length-scale* effects. They may be summarized as follows:

1. A new source of uncertainty in numerical results can be the length-scale effect which is inherent in the nature of nonlinear systems of conservation laws. For instance, in manufacturing of aircraft engines, the length-scale effect in numerical simulations may have the same order of error deduced by manufacturing [6]. To cure these numerical errors, adaptive simulations would be helpful,
2. Another length-scale effects reveal itself in problems with different physical descriptions in different scales, such as models including continuum-atomistic descriptions. Re-expressing such multiscale models as systems of conservation laws (for all scales), the *time-scale coupling* feature can properly be captured between the continuum scale and the atomistic level [3]. This challenging problem exists in many famous multiscale methods like the quasi-continuum (a coarse graining approach) [7], the domain decomposing method (with the atomistic and continuum regions and some interface matching methods) [8] and energy-based schemes (with the drawback that free energy is not accessible by atomistic simulations) [3]. Such proper multiple-scale computational methods pursue understanding the *process-structure-property* relationships of materials specially those are developed recently [9].

Developing of such localization and fine-scale effects motivates using of *adaptive solvers*; and since most of the governing equations can be described as nonlinear hyperbolic/hyperbolic-parabolic PDEs, using of high-order accuracy *non-linear solvers* are essential. These approaches may be summarized as:

1. In the nonlinear solvers, stencil of numerical algorithms are changed based on solutions; so the algorithm is solution dependent and nonlinear. This is in contrast with commonly used finite difference, finite volume, or finite element method using constant stencil everywhere.
2. Regarding adaptive solutions, either method or grid (mesh) adaptation is possible. In the method adaptation, different schemes (or simulating parameters) are used in different regions. For the grid (mesh) adaptation, different approaches have been developed. One of them is based on multiscale approach: multiresolution-based grid (mesh) adaptation. In these approaches, a problem is investigated at

different resolutions, simultaneously. In this regard, overall and localized features of solutions are distinguished and considered in numerical solutions.

In this work, the hyperbolic PDEs are simulated by central high resolution schemes with proper integration of them with wavelet-based adaptation procedures. The accuracy and the effectiveness of numerical solutions can then be obtained by such incorporation. In numerical simulation of first order hyperbolic systems, several important and challenging features exist, as: the lack of inherent (natural) dissipation [10], the forming of artificial (numerical) dissipation and dispersion [11, 12]. Inherent dissipation in a system improves both adaptation procedure and numerical stability. The numerical dispersion leads to developing wiggles in the front and behind of propagating waves, due to the distortion of different phases of propagating waves (fronts). The numerical dissipation has a tendency to flat discontinuities in numerical solutions [11, 12]. Discontinuous solutions are commonly formed in non-linear first-order hyperbolic systems, where controlling both the numerical dissipation and dispersion are challenging. Despite of hyperbolic systems, the inherent dissipation exists in elliptic and parabolic PDEs [13]. These systems are not so sensitive to small perturbations in their numerical solutions. The perturbations or errors dissipate during numerical simulations. Based on this feature, wavelet-based adaptive schemes have successfully been employed in the modeling of the elliptic [14, 15], parabolic problems [16–18] and parabolic-hyperbolic systems [19].

To guarantee solution stability of high-resolution schemes on non-uniform grids, some conditions must be satisfied. Such assurance can be obtained by using the normalized variable and space formulation (NVSF) criterion [20]. In the NVSF criterion, the identifying of propagating directions is necessary which is a complex procedure, especially for 2-D and 3-D problems [21]. To remedy this disadvantage, in this work, central and central-upwind high resolution schemes are considered (like the Kurganov and Tadmor method) [22–24]. Central/central-upwind high resolution schemes have simple and straightforward formulations and act as a black-box solver [25].

Central/central-upwind methods, however, are *sensitive for cell irregularity*; our investigations show that numerical instabilities appear rapidly in adaptive solutions. This is because, these methods do not originally satisfy the NVSF condition. To guarantee the numerical stability, two features should be studied: 1) The performance of slope/flux limiters on non-uniform grids designed with different construction criteria; 2) The effects of grid density variation on numerical solutions and adaptation procedures. Most of slope limiters have been developed for working properly on uniform grids. Using of such limiters on irregular grids can lead to unstable solutions.

Abrupt changing of grid densities can also lead to instability, due to ill-posedness feature of irregular sampled data. To prevent this kind of instability, density variation of adapted grids should be checked. For achieving this purpose, adapted grids are locally rechecked/modified in different resolution levels in the vicinity of high-gradient zones in the wavelet spaces.

Cell-adaptive central/central-upwind/CWENO high resolution schemes with second, third and fourth order accuracy are restudied in this work. Corresponding accuracy, stability conditions and corresponding limiters are provided over wavelet-based adapted cells, as well.

Convergence analysis of numerical solutions of nonlinear conservation laws are also studied directly over non-uniform cells by the concept of the local truncation error [26, 27], re-evaluated on irregular 1-D and 2-D cells. The numerical results confirm that the local truncation error can be used for the convergence studies of 1-D or 2-D systems with convex or even with non-convex fluxes. The uniqueness of numerical solutions is checked by the concept of numerical entropy production. Theoretically, the numerical entropy production is zero in smooth regions while less than zero around shocks and discontinuities [28, 29]. This helps to study the quality of numerical results especially for ones without exact solutions. All calculations will directly be performed on non-uniform grids. The concepts of the local truncation error and numerical entropy production have been used for both grid and method adaptation [30–32]. Different concepts could lead to different adapted results, especially, some concepts may not capture some phenomena. Hence, proper choosing of a adaptation approach would be crucial. In this study, wavelets are used for grid/method adaptation. The adaptation performance of this theory would be compared with the above-mentioned two concepts.

Hyperbolic systems with non-convex fluxes are also studied. In these systems, complex waves develop and so numerical solutions can converge to non-physical results [30]. In this work, wavelets are used to both grid and method adaptations to capture properly physical solutions.

Hyperbolic (wave) problems can be divided into two general families: i) First order waves; 2) Second order waves. Regarding the second order hyperbolic equations, it is possible to simulate them by the high resolution schemes. However, it is essential firstly to re-express them as a first-order hyperbolic system (that is the strain-velocity or stress-velocity representations). This procedure, however, can introduce considerable new degrees of freedom especially for 2-D and 3-D problems. In this regard, it is favorable to simulate second-order waves directly in their original forms. In this study, to

control spurious oscillations in such waves, the regularization approach is used, (e.g., the Tikhonov or total variation (TV) based regularization). In this approach, it is tried to remove noise (spurious oscillations) by the regularization concept by a post-processing stage. For this purpose, in this study, different regularization approaches with different regularization constraints are studied and finally a proper regularization method is suggested to control effectively non-physical oscillations resulted from discontinuities or high-gradient solutions. The main advantageous of this approach is that it can directly be utilized for higher-order waves in their original forms (such as the fourth-order waves in beams) without re-expressing them as a first-order wave equations.

### 1.0.1 Outline of the thesis

In Chapter 2, several literature would be reviewed including: (1) different multiscale simulation approaches of PDEs in the framework of multiresolution analysis; (2) For non-linear conservation laws, the concept of central high resolution schemes on uniform cells and extending of it to non-uniform cells; (3) Some challenging problems in numerical simulation of nonlinear conservation laws, such as: convergence analysis, uniqueness of solutions, problems with non-convex flux; (4) general information about high-order central high resolution schemes; (5) a small review of high-resolution methods using non-polynomial interpolations; (6) challenging problems for simulation second order waves directly in their original form arising from discontinuous solutions or stochastic like variation of properties of computational domain (e.g., heterogeneous materials).

In chapter 3, very briefly interpolating wavelet transform and grid adaptation are presented. It is important to note that both centered and non-centered cells produce by wavelet based grid adaptation. Afterward, it is shown to convert the adapted grids to regularized adapted grids, again in the framework of multiresolution analysis.

In Section 4, firstly, the formulation of the second-order central high-resolution schemes derived on wavelet-based cells for the first time. This formulation is for both centered and non-centered non-uniform cells. For non-linear stability, the global total variation diminishing (TVD) conditions are checked for non-uniform cell. Based on these conditions, three local TVD conditions are restudied and one of them updated to handle non-uniform non-centered cells.

In section 5, to study the convergence of numerical solutions, the truncation errors are derived for non-uniform cells for 1-D and 2-D problems. The performance of the numerical entropy production is checked over non-uniform cells to control uniqueness of solutions.

In section 6, some high-order central high resolution schemes are derived over wavelet-based adapted cells:

1. Third order central high resolution scheme using average values of three successive cells to produce piecewise parabolic polynomial. To preserve monotonicity feature the same shape conditions are updated and proved mathematically over wavelet-based adapted cells. To enforce the monotonicity feature at cell edges, the concept of the nonlinear scaling limiters are used. One of them updated based on the same shape features proved in this subsection,
2. Two new piecewise parabolic methods (PPMs) are introduced in this work for the first time which only use information from only one cells. These can be obtained based on introducing new state variables (obtained by new PDEs resulted from spatial differentiation of original PDEs). To enforce the monotonicity condition inside each cell, a new limiter and a new modification approach are presented as well,
3. We mathematically prove the accuracy order of central WENO schemes with three and five point stencils over non-uniform cells. We prove that regardless of weights, over wavelet-based non-uniform cells, they have the same order of accuracy designed to have on uniform cells,

In chapter 7 The stable simulation of second-order waves in their original form via the Tikhonov-based regularization method is presented. we studied different constraints used for Tikhonov regularization. Recently a new constraint was introduced. We numerically prove its strength for simulation of second-order waves by different benchmarks with discontinuous solutions. For this new constraint: (1) we mathematically prove the convergence property of the corresponding Tikhonov regularization; (2) We present its corresponding filter and compare it with other filters obtained with different constraints. Final, we *extend* the Tikhonov regularization, so it preserve:

1. The conservation property after regularization,
2. The local regularization feature to locally solve the Tikhonov method in a consistent way.

In chapter 8 we conclude the work presented in this study and some suggestions provided for future studies.

# Chapter 2

## Literature review and general descriptions

In this section, numerical simulation of both the first and second order hyperbolic (wave) equations are surveyed.

### 2.1 Multiresolution (wavelet) based simulation of partial differential equations (PDEs)

Multiresolution-based studying has rapidly been developed in many branches of science; one of methods having this powerful feature is the wavelet theory. Wavelet transforms have *multiresolution* feature and act as a mathematical microscope: studying simultaneously a phenomenon with different accuracy [33]. This capability motivates different area of researches, especially numerical simulation of PDEs [33]. Development of this theory was simultaneously performed by scientists, mathematicians and engineers [34]. Wavelets can detect different local features of data separated locally in different resolutions. Wavelets can efficiently distinguish overall smooth variation of a solution from locally high transient ones. These variations exhibit local deviation of the original data from the overall smooth approximation (background data). These localized details should be added to the smooth approximation to reconstruct the original data. In the wavelet theory, the wavelet coefficients (known as detail coefficients) measure the aforementioned localized variations; the magnitude of the detail values is in accordance with variations in data [35]. Wavelet coefficients of large values, thereby, concentrate automatically around high-gradient zones; this can be used as a criterion to identify such zones. There is also a one-to-one correspondence between the wavelet coefficients and



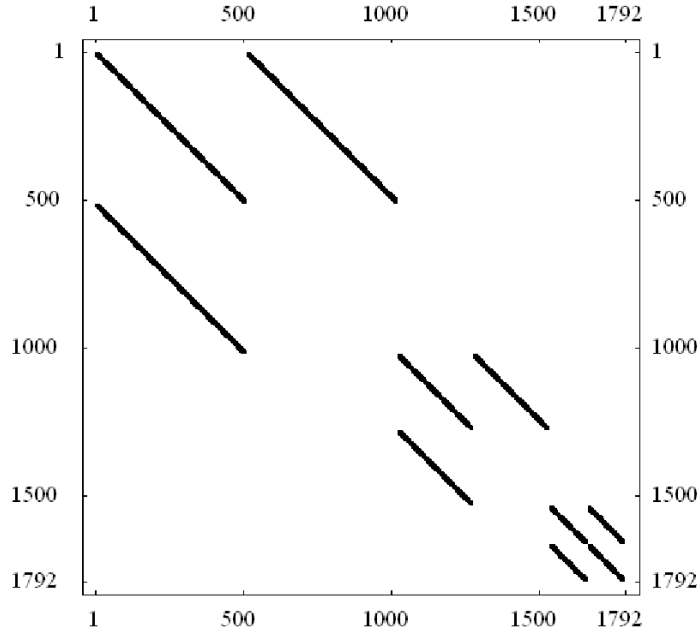


Figure 2.1: The projection of the operator  $d/dx$  into wavelet spaces by the wavelet Db-12.

corresponding spatial locations. Wavelet coefficients of small values reveal the existence of small variation in their corresponding spatial zones. Considering this fact and the one-to-one correspondence feature, the small localized variations and also corresponding grid points can be neglected with an acceptable accuracy and thereby an adaptive solver can be obtained either on adapted grids or by omitting wavelet coefficients of small values.

In general, there are two general approaches for PDE simulations by wavelets: 1) The *projection* method, 2) The *non-projection* one.

In the wavelet-based projection schemes all calculations are performed completely in wavelet spaces. For this case, operators (such as differential or integral ones) can be projected into wavelet spaces [36]. The projection of the most operators are banded in wavelet spaces. The width of the bands can be reduced by defining a *threshold*; the coefficients with smaller values of the threshold are set to zero and so the adaptation is performed in wavelet spaces [36]. For example, in Figure 2.1 the operator  $d/dx$  is projected to wavelet space by the wavelet Daubechies 12 (Db-12) (in the *non-standard* form [36]).

In the projection approach, also, wavelets (and corresponding scaling functions) can be used as shape functions, for instance in wavelet-Galerkin formulations. As there is

a one-to-one correspondence between wavelet functions and grid points, this projection method can be performed on adapted grids, as well (small projection coefficients and corresponding adapted grids are omitted from evaluations). The projection approach were used for simulations of elliptic [37], parabolic and hyperbolic PDEs [33]<sup>1</sup>. For hyperbolic PDEs some projection-based adaptive studies may be summarized as: 1) Adaptive discretization by projection into wavelet spaces in [38–42], 2) Adaptive Petrov-Galerkin scheme using wavelet functions as shape functions [43].

The wavelet-based projection methods, however, have three major drawbacks [18, 44]:

1. Projection of non-linear operators into wavelet spaces,
2. Imposing of different boundary conditions; to handle this case, one approach is developing of wavelets updating their function definition near boundaries based on BCs (known as wavelets on intervals) [37],
3. Modeling the geometry of boundaries.

In non-projection approach, on the other hands, wavelets are used only to adapt grids/cells. Afterward, numerical simulations are performed by other numerical methods. In particular, wavelet-based dynamic cell/grid adaptation methods have successfully been integrated with different solvers of hyperbolic (hyperbolic-parabolic and hyperbolic-elliptic) PDEs having steep moving fronts or sharp transitions in small zones. This approach used for simulation of the elliptic [14, 15], parabolic [16, 18, 43, 45, 46] and hyperbolic [21, 38, 39, 42, 47–54] systems.

In particular, for hyperbolic (hyperbolic-parabolic and hyperbolic-elliptic) PDEs, the above-mentioned studies may be summarized as: integration of the normalized variable and space formulation (NVSF) with MRA-based grid adaptation in [21, 47]; spatio-temporal fully adaptive decompositions performed in the context of finite volume (FV) schemes in [53]; adaptive finite difference methods [48, 49, 54, 55]; adaptive collocation schemes [56, 57]; an adaptive ENO method [58]; an adaptive WENO scheme [59]; stencil adaptations in [51]; integration of a MRA-based space-time adaptive procedure with the AUSM+ FV scheme in [60]; radial basis function (RBF)-based central high resolution schemes on MRA-based adapted grids/cells [61]; RBF-based WENO scheme over MRA-based adapted cells [62]; and integration of multiwavelet-based grid adaptation with discontinuous Galerkin schemes [63].

---

<sup>1</sup>Roughly, the moving property (advection feature) is mainly due to presence of hyperbolic characteristic in the systems. In contrast, the parabolic PDEs deal mainly with diffusion problems and elliptic systems express steady state solutions of either parabolic or hyperbolic PDEs [5].

In these approaches, wavelets are used as a tool to detect and capture steep moving fronts or localized transitions in narrow zones.

## 2.2 First-order nonlinear hyperbolic (wave) equations

Nonlinear conservation laws act as an universal tool to describe many physical and industrial phenomena from the molecule (atomistic) level to the continuum scale, as well as multiscale problems. Hence, to consider the scale-effects, one of the aims of this study is multiresolution-based adaptive simulation of nonlinear wave (front) propagation problems. The governing equation of these systems for 2-D problems can be expressed as:

$$\mathbf{u}_t + \mathbf{F}(\mathbf{u})_x + \mathbf{G}(\mathbf{u})_y = \mathbf{Q}_x^x(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_y) + \mathbf{Q}_y^y(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_y) + \mathbf{S}(\mathbf{u}), \quad x, y \in \Omega, \quad (2.1)$$

where: the vector  $\mathbf{u} \equiv \mathbf{u}(x, y, t)$  denotes the state variables;  $\mathbf{F}(\mathbf{u})$ , and  $\mathbf{G}(\mathbf{u})$  are function of flux vectors in the  $x$  and  $y$  directions, respectively;  $\mathbf{Q}^x(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_y)$  and  $\mathbf{Q}^y(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_y)$  show diffusion terms in the  $x$  and  $y$  directions, respectively;  $\mathbf{S}(\mathbf{u})$  is the source (loading) term;  $\Omega$  indicates the computational domain;  $\mathbf{u}_x := \partial\mathbf{u}/\partial x$  and  $\mathbf{u}_y := \partial\mathbf{u}/\partial y$ . In Eq. (2.1), the initial conditions (ICs) are:  $\mathbf{u}(x, y, t = 0) = \mathbf{u}_0(x, y)$  and the boundary conditions (BCs) are  $\mathbf{u}(x_\Gamma, y_\Gamma, t) = \mathbf{u}_\Gamma$ , where  $\Gamma$  denotes the boundary of the domain  $\Omega$ .

Due to the nonlinearity, discontinuous propagating fronts appear in the approximate solutions, which lead to some numerical drawbacks: the dispersion and dissipation phenomena. To remedy these effects, from early years of 1950s (by profound work of Godunov [64], especially for the Euler equation: gas systems) comprehensive works have been done on numerical solutions of nonlinear hyperbolic systems.

To control dispersion and dissipation phenomena, several nonlinear methods have been developed; also their accuracy and effectiveness have been improved by developing their adaptive versions. Some of these approaches enriched by multiresolution-based adaptations are: 1) Collaboration of point-wise multiresolution adaptation algorithms with high resolution schemes [21, 38, 39, 42, 47, 50–54], 2) Integration of the finite element method with adaptive multiresolution viscosity [65]; 3) Incorporation of the finite volume (and also high-resolution) methods with multiresolution-based adaptation on triangles [66]; 4) Combination of spectral theory with adaptive viscosity approach [67]. These efforts reveal importance of the adaptive solvers.

High resolution schemes use higher order approximations in smooth regions and a

first order one around discontinuities [5, 68]. This procedure is done in a way that spurious oscillations do not grow through time and thereby both numerical dispersion and dissipation phenomena can be controlled effectively. Regarding high resolution schemes two general and popular approaches<sup>2</sup> have been developed, as:

1. The *MUSCL* (Monotone Upstream-Centered Schemes for Conservation Laws) schemes using the concept of flux/slope limiters [70],
2. The *essentially non-oscillatory* (ENO) [71] and *weighted* ENO (WENO) [72, 73] schemes.

The ENO and WENO methods need fundamentally the Riemann solver<sup>3</sup> which is an expensive task. This is because the solver is based on (exact) mathematical evaluations or at least approximating evaluations. To remedy this problem, the central high resolution schemes were developed: the first scheme is the Lax-Friedrich method [74] with the first order accuracy. The higher order central schemes (a type of the MUSCL scheme) have been developed, mainly by Tadmor et al. [22, 25, 75, 76].

Central high resolution schemes have been developed both on staggered and non-staggered grids (meshes). In this central formulation, governing equations could have the *fully-discrete* form (discrete in both time & space) and (or) the *semi-discrete* form (continuous in time and discrete in space). In general, non-staggered formulations have less numerical dissipation and semi-discrete forms can use larger time steps. So, in this work, non-staggered central schemes with the semi-discrete formulation will be used.

In brief, non-staggered central (and also central-upwind) high resolution schemes offer the following benefits: having a simple and straightforward concept; being easy to implement; having less numerical dissipation than ones on the staggered cells, like the Nessyahu and Tadmor (NT) method; offering both semi-discrete and fully-discrete forms;

---

<sup>2</sup>Another approach is using of solution-dependent *artificial viscosity*. In this approach, the artificial viscosity has considerable value around a discontinuity to prevent spurious oscillations [68]. Since the viscosity is solution-dependent, this *nonlinear* scheme has high-order accuracy in smooth areas, and so the method is a high-resolution method. Recently this approach is integrated with entropy solutions as detectors (to identify discontinuities) to reach a high resolution feature, e.g. [69]. Difficulty of this approach is a proper selection of viscosity intensity to preserve only monotonicity without causing unnecessary smearing [68].

<sup>3</sup>A Riemann problem is an initial value problem including a conservation equation together with a piecewise constant data having a single discontinuity. For a conservation equation it can be represented as:

$$u(x, t)_t + f(u(x, t))_x = 0, \text{ where } u(x, t = 0) := \begin{cases} u^l, & \text{for } x \leq 0, \\ u^r, & \text{for } x > 0, \end{cases}$$

where  $u^l$  and  $u^r$  are constants. The solver of the above-mentioned equation is the Riemann solver.

being a Riemann-free solver; no requiring to staggered grid points (as needed in the NT scheme [25]); having comparable second and higher order accuracy with other expensive techniques.

Central/central-upwind schemes [22, 25, 75, 76] are basically developed for uniform cells. Since high resolution schemes can be expensive (especially third and higher-order ones), using of adaptive solvers seems to be necessary. On the other hand, central schemes are sensitive for grid irregularity; hence, in this work, the main concern is improving central schemes for working properly on non-uniform cells.

To have stable upwind, ENO and WENO schemes on non-uniform grids, these approaches are enriched with some stability conditions [21, 47]. One famous condition is the normalized variable and space formulation (NVSF) (see Appendix B) [20, 21, 47, 50]. In these methods, however, the direction of propagation is essential for both the upwind formulation and the stability condition. Such direction detection procedure is indeed complex, cumbersome and expensive, especially in 2-D and higher dimension problems [21]. Some researchers integrated other detection algorithms such as the level set method with these high resolution methods [77]; and this introduces extra computing costs.

As mentioned, central high resolution schemes are also sensitive to grid irregularity. This is because, they do not satisfy any stability criterion for grid irregularity; such as the NVF or NVSF [20, 21, 47, 50].

### 2.2.1 A practical approach for using high resolution schemes on non-uniform grids

As mentioned, central/central-upwind high resolution schemes have theoretically been developed for uniform cells/grids; for such cells/grids, their performance, stability (non-linear stability) and accuracy are studied, as well. In practical/engineering problems, these methods are directly used on non-uniform (adaptive) grids without formulation and/or definition modifications. To preserve numerical stability and accuracy, grid/cell density is let vary gradually. For this purpose a *growth factor* (GF) is recommended. One example of using such GF is illustrated in Figure 2.2 for simulation of flow around a rectangle obstacle [78].

This GF control expanding/growing of grid/cell sizes away from an area with the finest grid/cell size (this fine-scale discretization is essential in the vicinity of discontinuities or zones with large solution gradients). Some recommended GF are 1.2 [79, 80] or 1.3 [81, 82] (smaller values of GF are even reported). One example of an open source

finite volume code is: *OpenFOAM*. It uses the second order high resolution method with several limiters, where the limiters were fundamentally designed on uniform cells. With the grid adaptation approach, explained above, this code is widely used in many practical and engineering problems.

In scientific works one of the popular approaches is the *remapping* technique, where indeed all calculations are done on uniform cells, e.g. [83, 84]. And also in some works limiters are directly used without definition modification, e.g. [85].

## 2.2.2 Central high resolution schemes and multiresolution analysis

The high resolution schemes also have successfully been incorporated with multiresolution-based adaptive methods, [21, 38, 39, 42, 47, 50–54]. These adaptive solvers are mostly based on upwind high resolution schemes [21, 38, 39, 47, 50–53].

As mentioned before, simulation on a non-uniform grid is indeed an ill-posed problem; to guarantee stability of adaptive solution, two general approaches have been suggested: 1) Using of the level-depending thresholding, 2) Using a method satisfying a stability criterion for irregular grids.

Following the first approach, Harten introduced the level-depending thresholding [51]. For data  $f(x)$  belonging to subspace  $V_{J_{max}}$  ( $f(x) \in V_{J_{max}}$ ), the considered threshold for resolution level  $j$  is:  $\epsilon_j = 2^{j-J_{max}-1} \cdot \epsilon_0$  where  $j \in \{J_{min}, J_{min} + 1, \dots, J_{max} - 1\}$  and  $\epsilon_0$  denotes a positive constant [38, 39, 51–53]. By this assumption, thresholds of larger values would be used for coarser resolutions; therefore, it is acceptable to consider that for an adapted point of resolution level  $j$ , there would be sufficient number of surrounding adapted points of the successive coarser resolution level  $j - 1$ . And this will guarantee gradual density-variation of adapted grids. However, our studies show that in numerical simulation of nonlinear propagating fronts, this adapting method does not necessarily lead to a completely stable procedure (for example, adapted results are sensitive for temporal time steps; also long-time behavior can be affected). And so, it is always necessary to be sure about proper distribution of adapted points. It seems, this is due to the nonlinear feature of such problems. In this regard, here, a grid checking-modifying procedure is done after each cell-adaptation by a post-processing stage. This also is performed in the multiresolution framework.

The key idea is to have local and gradual variation of adapted points/cells without any abrupt change of grid density (this can be used for both constant and level-dependent thresholding methods). For this purpose, for a point of a distinct resolution level (say  $j$ ), it should be checked that there are enough number of surrounding points belonging

to both same resolution ( $j$ ) and all the coarser ones ( $\{j - 1, j - 2, \dots, J_{min}\}$ ).

Central/central-upwind methods, as mentioned, are sensitive for cell irregularities; our investigations show that numerical instabilities appear rapidly in adaptive solutions. This is because, these methods do not originally satisfy the NVSF condition (see Appendix B). To guarantee the numerical stability, two features should be studied: 1) The performance of slope/flux limiters on non-uniform cells/grids; 2) The effects of grid density variation on numerical solutions and adaptation procedures. Most of slope limiters have been developed for working properly on uniform grids. Using of such limiters on irregular grids may lead to unstable solutions. In this study, it will be shown that:

1. How to use some limiters without modification of their definitions,
2. Modifying some other limiters for non-uniform grids.

Investigations reveal that the proposed methods lead to stable results comparable with those of NVSF-based schemes.

The Dubuc-Deslauriers (D-D) interpolating wavelets [35, 45, 46] are used here for grid adaptation. This family has simple and straightforward algorithms with physical meaning. All calculations can then be performed in the physical domain. The D-D wavelets use minimal spatial support for data reconstruction (approximation), and this is important, since: larger inter-distance in two sampled data is, smaller correlation between them exists [18, 45, 46].

Semi-discrete PDEs in the spatial domain are solved in time by an explicit TVD integration method, such as the second or third order TVD Runge-Kutta scheme. As all spatio-temporal calculations are done in the physical domain, the method is simple and conceptually straightforward [18].

### 2.2.3 Numerical errors, numerical entropy and grid-adaptation

In nonlinear conservation laws, discontinuous solutions develop typically. Thereby, common error estimation concepts, mainly based on the Taylor expansion, can not be used. For this reason, the concept of the local truncation error is used to assess convergence of solutions to weak ones as a practical approach [26, 27]. Related formulations for such error estimation would be provided for 1-D and 2-D non-uniform cells. This kind of error has direct relationship with the locally defined  $L^1$  norm (the so-called  $Lip'$ -norm) introduced by Tadmor [86, 87] for nonlinear 1-D scalar conservation laws with convex fluxes (also see Sec. 5). The numerical results confirm that the local truncation error can also

be used for convergence study of 1-D or 2-D conservation laws, even with non-convex fluxes.

Uniqueness of numerical solutions is checked by the concept of the numerical entropy production. Theoretically, the numerical entropy production is zero in smooth regions while less than zero around shocks and discontinuities [28, 29]. This helps to study quality of numerical results especially for those without exact solutions. All calculations will directly be done on non-uniform grids.

The concepts of the local truncation error and the numerical entropy production have been used for both grid and method adaptation [30–32]. Different concepts could lead to different adapted results, especially, some concepts may not capture some phenomena. Hence, proper choosing of an adaptation approach would be crucial. In this study, wavelets are used for grid/method adaptation. Adaptation performance of the wavelet theory would be compared with above-mentioned two other concepts.

#### 2.2.4 Non-convex fluxes

Hyperbolic systems with non-convex fluxes would also be studied; these systems need non-classical solvers due to developing of complex waves. These systems can explain important phenomena, such as: Euler equations of gas dynamic with a non-convex flux, polymer system used for simulation of polymer flooding processes in enhanced oil recovery and mechanical wave equations with non-convex fluxes. This latter case, for example, can occur in the stress wave propagation in sediment layers under water [88, 89]. Another important phenomenon including the non-convex fluxes is the *phase transition*; this phenomenon can develop in solid materials and flow of compressible fluids (gas or liquid). Some examples are: i) In solid materials: shape memory alloys [90]; ii) In gas-liquid materials: the evaporation phenomenon [91].

It will be shown that even though numerical solutions converge to the weak form solutions (controlling by the local truncation error), they may not be physical (real) ones due to the existence of complex waves in these problems [30]. In this work, wavelets are used to both grid and method adaptations for capturing properly physical solutions.

#### 2.2.5 High-order central high resolution schemes

In order to have a third-order accuracy for polynomial-based formulations, (in the spatial domain) in central high resolution schemes, a quadratic polynomial should be used in the reconstruction stage over each cell [92, 93]. This reconstruction was performed by



three approaches:

1. Quadratic average-interpolating polynomials were evaluated directly by average solutions of three successive cells with the central formulations [92, 93, 93]. The same shape feature was enforced by a non-linear limiter acting as a scaling limiter [93]. This limiter was later used for other types of high-order high resolution schemes working on uniform cells/grids [94–96]. This type of scaling limiter was successfully extended for unstructured cells [97, 98],
2. Piecewise parabolic polynomials (PPM) [99], determined only by information of each cell could reduce the localized spurious oscillations around discontinuities and the over-smoothing phenomenon over smooth solutions [92], generated by the parabolic reconstruction over three successive cells (the first approach). These effects, however, can be reduced by using more localized information. In this method, the reconstruction stage on the cell  $I_j$  was based on the cell average value  $\bar{u}_j$  and cell-edge values  $u_{j\pm 1/2}$ . For estimation of  $u_{j\pm 1/2}$  values, the average-values of four neighbor cells were employed. To avoid generated numerical dissipation Popov et al. [100] improved the method by a local stencil: edge values  $u_{j\pm 1/2}$  were estimated by the concept of the characteristic lines from solutions of the previous time step in cell  $I_j$ . This method improved numerical accuracy in the expense of using some mathematical computations and information. To remedy the over-smoothing phenomenon around the *smooth extrema*, new limiters [101–103], and hybrid methods were also proposed [104].
3. A quadratic average interpolating polynomial, resulted from a convex and symmetric combination of left, central and right average-interpolating polynomials, known as the central-WENO schemes [105, 106].

In the first and second approaches, the monotonic feature is enforced by slope/flux limiters, while in the third approach, the monotonic property is obtained by proper weighting effects of one-sided and central interpolations.

The central schemes mentioned above have the total variation bounded (TVB) feature, which gives more flexibility in comparison to the total variation diminishing (TVD) property. In the first and second approaches, the TVB feature results from satisfying the *local maximum principle* and a non-oscillatory property [93]. The original TVB third-order high resolution formulation [93] was integrated with the central formulation

on staggered grids [93]. Then, a new nonlinear limiter was proposed to satisfy the TVB feature, followed by 1-D and 2-D semi-discrete forms [92].

The mentioned TVB-based high resolution schemes ([92, 93, 93]) were developed for uniform cells/grids. For the case of nonuniform cells/grids, most works were based on the *remapping* technique: all calculations were performed on uniform cells/grids and then remapped on the original irregular cells/grids [83, 84, 107].

Finally, it should be mentioned that the third-order CWENO reconstruction has been studied on irregular cell-centered meshes [108–110]. Also errors in simulations for the third order CWENO scheme were studied for uniform [111] and cell-centered non-uniform cells [108]. In most of the aforementioned adaptive solvers, numerical entropy was used as a tool for adaptive mesh refinement. A fourth order CWENO scheme with cell-centered five-point stencil on uniform cells was proposed for multidimensional problems [112]. Capdeville [113] developed the fifth-order CWENO reconstruction on constant cell-centered non-uniform grids based on *implicit* and *upwind* formulations, where the method has a five-point stencil.

## 2.2.6 Non-polynomial reconstruction over each cell

In this section, so far, it is assumed that variation of a state variable can be described by a polynomial. Regarding central high resolution schemes, the constant variation over each cell can lead to the first-order Lax-Friedrichs scheme. A linear reconstruction over each cell can lead to the second-order NT (with staggered cells) or second-order KT (with non-staggered cells) scheme. Using of the parabolic reconstruction leads to third-order schemes. By increasing polynomial orders, higher-order schemes can be obtained, especially by the concept of CWENO reconstruction procedure (i.e., using of lower order polynomials to reproduce higher-order ones without oscillations).

Instead of polynomials, other type of variation (reconstructions) can be assumed over each cell, such as: logarithmic [114, 115], hyperbola [116], rational [117], exponential [118, 119], spectral [120], moving least square (MLS) [98] and radial-basis functions [61, 121, 122].

## 2.2.7 Conclusion

Based on the aforementioned discussions, in brief, in this work, it will be shown that to preserve numerical stability of central high resolution schemes on non-uniform cells, it is necessary:

1. To modify central schemes and recheck/update nonlinear stability conditions: here for second and higher order central schemes, the *TVD* (total variation diminishing), *TVB* (total variation bounded) and *UNO* (uniformly high order accurate non-oscillatory) nonlinear stability conditions are studied; on uniform cells, for TVD and TVB criteria see [5, 22, 68, 76, 92, 93], and for second and third-order accuracy UNO definitions see [123, 124]. Here, the second-order adaptive UNO limiters are also reformulated to handle MRA-based adapted cells. Regarding these conditions, central formulations and the performance of limiters are restudied,
2. To adapt properly considering cells: The source of the numerical dissipation and dispersion (as mentioned in the Introduction section) is the truncation error and it introduces a new source of error with significant effect on non-uniform grids. This is because, estimation on an irregular grid points is indeed an ill-posed problem. To solve an ill-posed problem, there are two general approaches: 1) Solving directly the ill-posed problem by a regularization method [48, 49] or a stabilization condition [21, 47]; 2) Firstly replacing the ill-posed problem with a *nearly well-posed* one, and then solving the new problem by common schemes. Here, the second approach is used by a post-processing stage over adapted cells for simulation of the first-order hyperbolic PDEs.

Employing these two steps, different central high resolution schemes will be studied on nonuniform grids, in this work; they are:

1. The second-order central KT scheme (with the TVD, TVB and UNO stability conditions),
2. Second-order central-upwind schemes (with the TVD stability condition),
3. Higher order central schemes: 1) Polynomial-based third-order solver (with the TVB stability condition), enhanced with nonlinear updated scaling limiters to enforce the *max-min* criterion; 2) The third-order accuracy piecewise parabolic method (PPM)
4. Central weighted essentially non-oscillatory schemes (CWENO): third and fourth order accuracy, respectively with three and five point stencils (both with the TVB stability condition).

Furthermore:

1. For cell-adaptive second-order central high resolution schemes, the formulation of local truncation errors on MRA-based adapted cells are developed for 1-D and 2-D problems. Then, these formulations are used to control convergence in some numerical simulations,
2. Some slope limiters are updated or introduced for handling MRA-based adapted cells (cell-centered and non-centered ones),
3. Integration of wavelet-based parameter adaptation (for limiters) for proper simulation of problems with non-convex fluxes,
4. Studying the performance of numerical entropy production on MRA-based adapted cells.

Finally it should be mentioned that to have the non-oscillatory feature in the time domain, the temporal integration schemes with the TVD property are used. In this study, the TVD Runge-Kutta method with second-order or third-order accuracy is used [22, 125].

## 2.3 Second order waves

An effective way for solving second order hyperbolic PDEs is to rewrite them as a system of first order hyperbolic equations and then to simulate them with one of schemes developed for first order wave equations. This will lead to oscillation free results with small numerical dissipation [126]. The second order hyperbolic PDEs can also be solved in their original second order form. For this purpose, several approaches have been developed:

1. Using some artificial dissipation without modifying the governing equations. For this case, dissipation is inherently added in evaluation procedures. In the time domain, the algorithmic dissipative time integration methods have been developed for removing spurious oscillations [127–130]. For the spatial domain, inherent filtering concept is also developed in derivative estimations [131, 132],
2. Adding some artificial viscosity in the governing equations to stabilize the solution. This can be done by using some local artificial viscosity around high-gradient zones in the spatial domain [133, 134]. Hughes [135] showed that this approach damps mainly the middle modes without affecting the lower and higher modes

substantially. As artificial diffusion decreases accuracy of solutions considerably, methods using the artificial diffusion only in high frequency ranges were developed; such as, the spectral viscosity schemes [65, 136, 137]. This approach has been employed for both first and second order hyperbolic systems,

3. Filtering spurious oscillations from numerical solutions in the spatial domain by a post-processing stage [13, 138]. These schemes were successfully used in simulation of hyperbolic systems on uniform grid points [139–141], and non-uniform grids [48, 49, 55]. It should be mentioned that many smoothing schemes working satisfactorily on uniform grids are not suitable for non-uniform ones: leading to unstable or unreliable results [142].

The concept of the high-resolution treatment has recently been advised for handling second-order hyperbolic systems [143]. The performance of this new concept will also be compared with the regularization-based method (proposed in this work). The aim of this work is to solve directly the second order hyperbolic systems. Such approach has the following advantages:

1. Less degrees of freedom are needed (in case of elastodynamic problems, in second-order form three dependent variables exist, while between nine and fifteen variables are needed when they are re-formulated as first-order systems),
2. The solution of the first-order system also satisfies the second-order form: this can be done by imposing/checking some constraints. This is because, the first order form admits more acceptable solutions,
3. Possibility of using larger time-steps in second-order form compared to the first-order one [144].

In this work, the spatially filtering approach is used. This will be done by a post-processing stage to remove properly spurious oscillations from solutions containing discontinuities. However, curing of the numerical dispersion effects is a crucial task. This is because:

1. Estimation of a function and corresponding derivatives from its sampled points can be considerable as a type of (semi) ill-posed problem. This would be more clear by considering the fact that different estimated values (data and corresponding derivatives) can be obtained for different interpolation methods. From mathematical

point of view, this fact may be explainable as follows. Sampling procedure can be considered as convolution of a continuous function  $f(x)$  with the Dirac delta  $\delta(x)$  as:  $f_i = f(x_i) = \int f(x)\delta(x_i - x) dx$ . This integral is known as the first-order Fredholm integral, and corresponding inverse problem could be a (semi) ill-posed problem [145, 146] i.e.: finding  $f(x)$  from corresponding sampled data  $\{f(x_i)\}$ ,

2. For numerical simulations, it is necessary that derivatives of estimated functions also are continuous. This means that the derivatives can be properly estimated up to some order. This derivative estimation (from noisy data) is another ill-posed problem. This is because, effects of small noise amplify considerably derivative values,
3. Spatial adaptation will insert a new source of ill-posedness; most of filtering methods can not properly work on irregular points/meshes.

Different estimation approaches of functions and their derivatives have been developed; such as: finite-difference methods [147], integral-based schemes [148–151], regularization approaches [120, 152, 153], or interpolation methods [154]; for a general overview, the reader is referred to [155, 156]. In this work, the regularization approach will be used, where a variational functional is employed. The key challenge is proper handling of the numerical dispersion (known as edge preserving in image processing problems). Therefore, the total variation (TV) based regularization method was developed [157], for preserving discontinuities. Such results, however, are not smooth enough for numerical simulations. Hence, a proper selection of a functional with sufficient smoothness being free from spurious oscillations (due to the numerical dispersion) is important.

In this study, several regularization approaches with different constraints will be studied. Furthermore, effects of imposing extra information containing extra local information will be investigated (it is known as the *model-base* regularization). By using such local extra information, it is tried to impose local features. This extra information is different from the conventional constraints.

These regularization techniques have several advantages over most other denoising schemes; for the Tikhonov method, they are:

1. It has a unique and stable closed-form solution with fast algorithms; this makes it completely popular,
2. It leads to the smoothest possible results for a distinct value of estimation error among all of the regularization schemes,

3. Proper boundary value estimation can be obtained. Estimation errors around boundaries are known as the Runge phenomenon. The Tikhonov regularization method is one of the single-interval schemes that can effectively handle it (without using an anti-Runge scheme) [158, 159],
4. Noise can directly be removed from irregularly spaced data where standard filtering techniques used in time series analysis become awkward to implement,
5. By proper selection of prior information, derivatives up to some degree can be estimated,
6. It can deal with many types of ill-posed problems such as data irregularity or derivative estimation problems.

Since the Tikhonov method acts as a post processor, it can be integrated with different explicit/implicit higher-order grid based methods, such as: compact finite difference schemes [131, 132, 138]. For the Tikhonov-based regularization, other favorable features can also be considered to improve estimated results; such as: conservative regularization (even for non-Gaussian noise), adaptive and/or local (global) smoothing. These extensions for common Tikhonov methods will also be presented in this work.

One difficulty in regularization methods is proper choosing of regularization parameters. Indeed, most of the existing estimating methods lead to over or under smoothing results [160, 161]. Current experiments reveal that even small amount of regularization can considerably improve numerical results. In general, the trial-and-error method can be recommended to find an optimum range of regularization parameters. Error bounds and convergence rates are studied for commonly used constraints or regularization definitions; e.g., see: [162, 163]. For a recently proposed constraint of the Tikhonov method (which one that can properly handle discontinuity effects), error bounds and convergence rates will also be studied here.

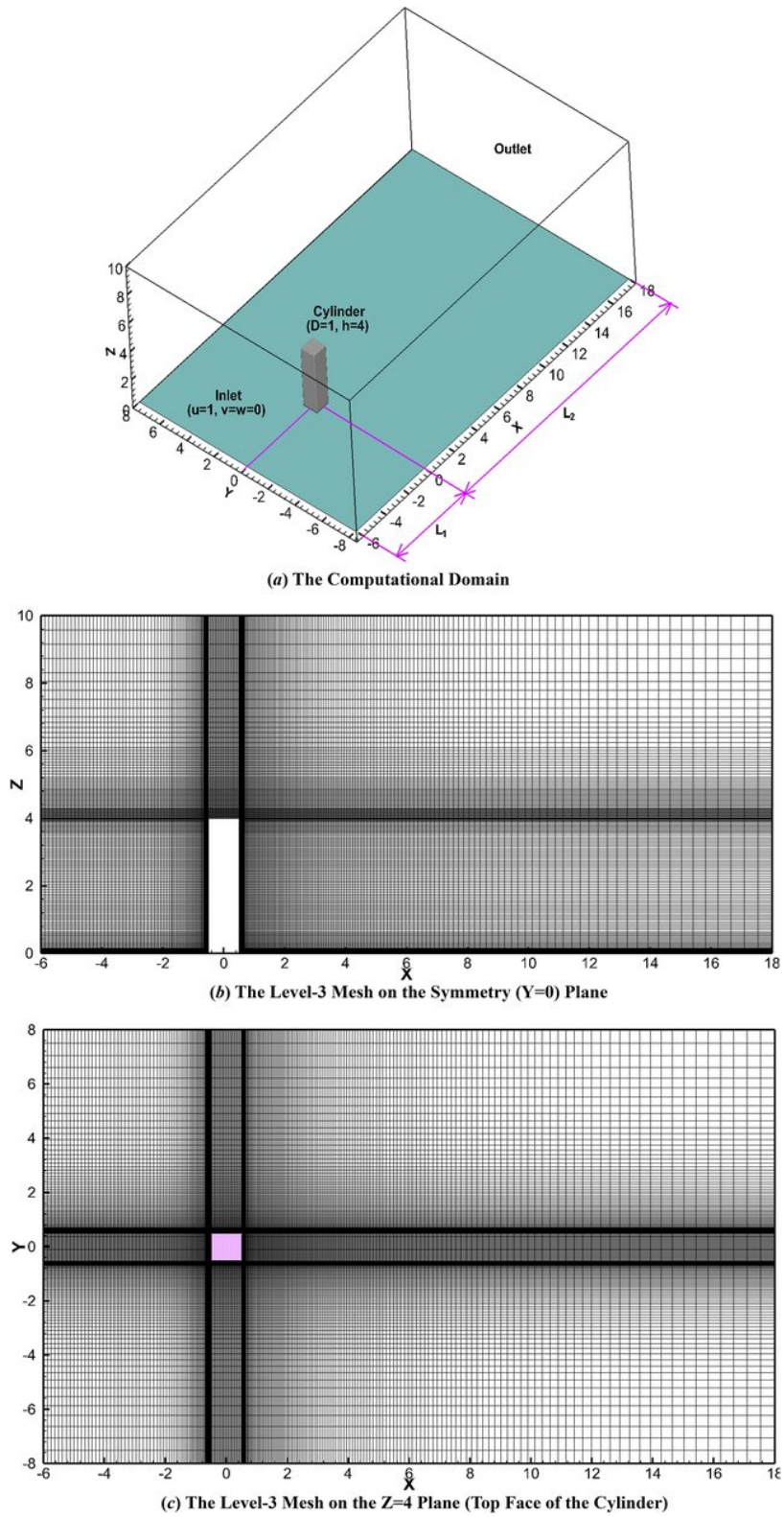


Figure 2.2: Non-uniform cells obtained by a GF (growth factor) for simulation of flow around a rectangle obstacle [78].



# Chapter 3

## Multiresolution analysis and grid adaptation

### 3.1 Multiresolution representation of data

In this section, wavelet-based multiresolution analysis will be described for both 1-D and 2-D data. For 2-D case, the algorithm is based on the 1-D one; this means the dimension-by-dimension algorithm is used for 2-D data.

#### 3.1.1 Multiresolution representation of 1-D grids

A dyadic grid on spatial interval  $[0, 1]$  is assumed as follows [46]:

$$V_j = \left\{ x_{j,k} \in [0, 1] : x_{j,k} = \frac{k}{2^j} \text{ for all } j \in \mathbb{Z}, \text{ and } k \in \{0, 1, \dots, 2^j\} \right\}, \quad (3.1)$$

where  $j$  and  $k$  are the resolution level (corresponding to spatial scale  $1/2^j$ ) and spatial position ( $k/2^j$ ), respectively. This definition of dyadic grid points  $V_j$  is ended to the condition  $x_{j-1,k} = x_{j,2k}$  and the multiresolution representation core: i.e.,  $V_j \subset V_{j+1}$ . The points belonging to  $V_{j+1} \setminus V_j$  are denoted by  $W_j$ ; this subspace can be expressed as:

$$W_j = \left\{ x_{j+1,2k+1} \in (0, 1) : x_{j+1,2k+1} = \frac{2k+1}{2^{j+1}} \text{ for all } j \in \mathbb{Z}, \text{ and } k \in \{0, 1, \dots, 2^j - 1\} \right\}. \quad (3.2)$$

So it can intuitively be concluded that:  $V_j \oplus W_j = V_{j+1}$ . This means the detail subspace  $W_j$  with the approximation subspace  $V_j$  can create (span) the next finer approximation subspace  $V_{j+1}$  with more details. By repeating this decomposition procedure

on  $V_{J_{max}}$ , it is obvious that:

$$V_{J_{max}} = V_{J_{min}} \oplus \sum_{i=0}^{N_d-1} W_{J_{min}+i}, \quad N_d = J_{max} - J_{min}, \quad (3.3)$$

where,  $J_{max}$ ,  $J_{min}$  and  $N_d$  denote the finest resolution, the coarsest resolution and number of decomposing levels, respectively [35].

A continuous function  $f(x)$ , defined on  $V_{J_{max}}$ , is assumed (i.e.,  $x \in V_{J_{max}}$ ). Regarding the multiresolution representation, the function can be decomposed as [35, 46]:

$$\begin{aligned} f(x) &= \sum_{l=0}^{2^{J_{min}}} c_{J_{min},l} \phi_{J_{min},l}(x) + \sum_{j=J_{min}}^{J_{max}-1} \sum_{n=0}^{2^j-1} d_{j,n} \psi_{j,n}(x) \\ &= Pf_{J_{min}} + \sum_{j=J_{min}}^{J_{max}-1} Qf_j, \end{aligned} \quad (3.4)$$

where  $\phi(x)$  and  $\psi(x)$  are scaling and wavelet functions, respectively; the functions  $\phi(x)$  and  $\psi(x)$  measure overall behavior and local fluctuations of data, respectively; sets  $\{\phi_{j,k}\} := \{\phi(2^j x - k)\}$  and  $\{\psi_{j,l}\} := \{\psi(2^j x - l)\}$  denote dilated and shifted versions of  $\phi(x)$  and  $\psi(x)$ , respectively (they used for measuring localized features of data at different resolutions). Coefficients  $c_{j,l}$  and  $d_{j,k}$  are approximation and detail coefficients with resolution  $j$ . The operators  $Pf_j$  and  $Qf_j$  show the approximation and detail information of  $f(x)$ , defined on grid points  $V_j$ , and  $W_j$ , respectively. The approximation on successive finer resolution  $j + 1$ , can then be obtained as:  $Pf_{j+1} = Pf_j + Qf_j$  [35].

In this study, the interpolating D-D wavelet of order  $2M - 1$  with support  $\text{Supp}(\phi) = [-2M + 1, 2M - 1]$  is used. It can be obtained by auto-correlations of Daubechies scaling function of order  $M$  (having  $M$  vanishing moments) [35]. By using this family, the transform coefficients in Eq. (3.4) can totally be evaluated in the physical space with physical meanings. The approximation coefficients ( $c_{J_{min},l}$ ) are equal to sampled values of  $f(x)$  at points  $x_{J_{min},l} \in V_{J_{min}}$ , and the detail coefficients ( $d_{j,n}$ ) are differences between  $f(x_{j+1,2n+1})$  and  $Pf_j(x_{j+1,2n+1})$ , or  $d_{j,n} = f(x_{j+1,2n+1}) - Pf_j(x_{j+1,2n+1})$  (for the D-D wavelets,  $Pf_j(x_{j+1,2n+1})$  can be obtained by the local Lagrange interpolation on points  $V_j$ ); see Figure 3.1 [35, 45, 46].

Implementation of the D-D wavelet transform on 1-D grids can be summarized as follows:

1. Let us assume the finest resolution level to be  $J_{max}$  (with sampling step  $1/2^{J_{max}}$ );

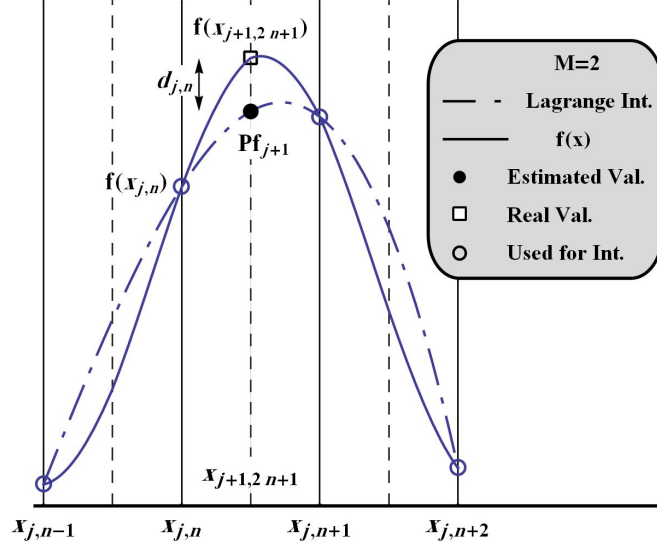


Figure 3.1: Schematic illustration for calculation of the D-D wavelet transform.

the coarsest resolution level is chosen to be  $J_{min}$ , so number of decomposition levels is  $N_d = J_{max} - J_{min}$ ,

2. The approximation coefficients: on the coarsest sub-space, due to interpolation feature of  $\phi$ ,  $c_{J_{min},k}$  are equal to sampled values of  $f(x) \in V_{J_{max}}$  at points  $x_{J_{min},k} \in V_{J_{min}}$ ; i.e.,  $c_{J_{min},k} = f(x_{J_{min},k})$  for  $k = \{0, 1, \dots, 2^{J_{min}}\}$ ,
3. The detail coefficients (see Figure 3.1): as mentioned before, at  $x_{j+1,2k+1}$ , differences of projection and real values are the detail coefficients. For  $x_{j+1,2k+1}$  and for the D-D wavelet of order  $2M - 1$ , local Lagrange interpolation using  $2M$  neighbor points of  $x_{j+1,2k+1}$  belonging to  $V_j$  are employed for estimation of  $Pf_j(x_{j+1,2k+1})$ . For each point  $x_{j+1,2k+1}$  on an infinite signal, the selected grid points are located in symmetric style around  $x_{j+1,2k+1}$ ; they are:
 
$$\{\{x_{j+1,2k+2n}\} : n \in \{-M + 1, -M + 2, \dots, M\}\}.$$
 For a finite signal with sampled points  $\{\{x_i\} : i = \{0, 1, \dots, 2^{J_{max}}\}\}$  and for the case  $M = 2$ , it is easy to show that  $Pf_j(x_{j+1,2k+1})$  would be:

$$Pf_j(x_{j+1,2k+1}) = \frac{1}{16} \begin{cases} 5f(x_{j+1,k}) + 15f(x_{j+1,k+2}) - 5f(x_{j+1,k+4}) + f(x_{j+1,k+6}), & \text{for } k = 0, \\ 9f(x_{j+1,2k}) - f(x_{j+1,2k-2}) + 9f(x_{j+1,2k+2}) - f(x_{j+1,2k+4}), & \text{for } k = 1, \dots, N_j - 2, \\ f(x_{j+1,k-6}) - 5f(x_{j+1,k-4}) + 15f(x_{j+1,k-2}) + 5f(x_{j+1,k}), & \text{for } k = N_j - 1, \end{cases} \quad (3.5)$$

where:  $N_j = 2^j$ . Having  $Pf_j(x_{j+1,2k+1})$  values the detail coefficients  $d_{j,k} = f(x_{j+1,2k+1}) - Pf_j(x_{j+1,2k+1})$  can be evaluated for levels  $j \in \{J_{min}, \dots, J_{max} - 1\}$ .

Having the  $d_{j,k}$  values, uniform grids can be adapted, as below. In the interpolating wavelet theory, transform coefficients and grid points have one-to-one correspondence. This feature leads to a simple 1-D grid adaptation algorithm. For  $f(x) \in V_{J_{max}}$ , a predefined threshold  $\epsilon$  is assumed, then, in each level of resolution  $j \in \{J_{min}, J_{min} + 1, \dots, J_{max} - 1\}$ , points  $x_{j+1,2n+1} \in W_j$  are omitted from original calculating grid points, if corresponding detail coefficients,  $d_{j,n}$  satisfy the condition  $d_{j,n} < \epsilon$ . This means that the function at the point  $x_{j+1,2n+1}$  is smooth enough, so that its contribution in the approximation,  $d_{j,n}\psi_{j,n}(x)$ , can be neglected (see Eq. 3.4). Donoho [164] showed that such truncation error is in accordance with threshold values. Finally it should be mentioned that the predefined threshold can be either level-dependent or not. One popular level-dependent threshold is  $\epsilon_j = \epsilon_0/2^{J_{max}-j-1}$  for  $j \in \{J_{min}, J_{min} + 1, \dots, J_{max} - 1\}$ . This means a smaller threshold is assumed for finer resolution. To use a constant threshold value for all resolution levels, it is better to normalize the detail coefficient of resolution  $j$  as:  $d_{j,n} = d_{j,n}/f_j^{ref}$ . In this study constant pre-defined threshold with normalized factor  $f_j^{ref} = \max\{f(x_j)\}$  is used.

In the following, an example is presented containing both a discontinuity and smooth-high-gradient response, Figure 3.2 (a). Schematic shape of grid distribution at different resolution levels is also shown in Figure 3.2 (b). In figure (a), the solid line and solid points show the original and adapted data, respectively. In figure (b), the solid points and hollow circles represent distribution of multiresolution-base sampling points and adapted points in different resolution levels, respectively. It is clear that: 1) there is a one-to-one correspondence between wavelet coefficients and grid points; 2) points belonging to  $W_j$  of larger resolution levels, concentrate more densely and thereby they can detect finer details (fluctuations).

The presented adaptation procedure was shown to be efficient in the resolution of scalar functions. For resolution of functions in vector system, the previous procedure is modified to reflect the solutions' behaviors of all equations. Namely, the resultant adapted grid is simply superposition of all adapted grids; which each adapted grid corresponds to a functions of the vector system.

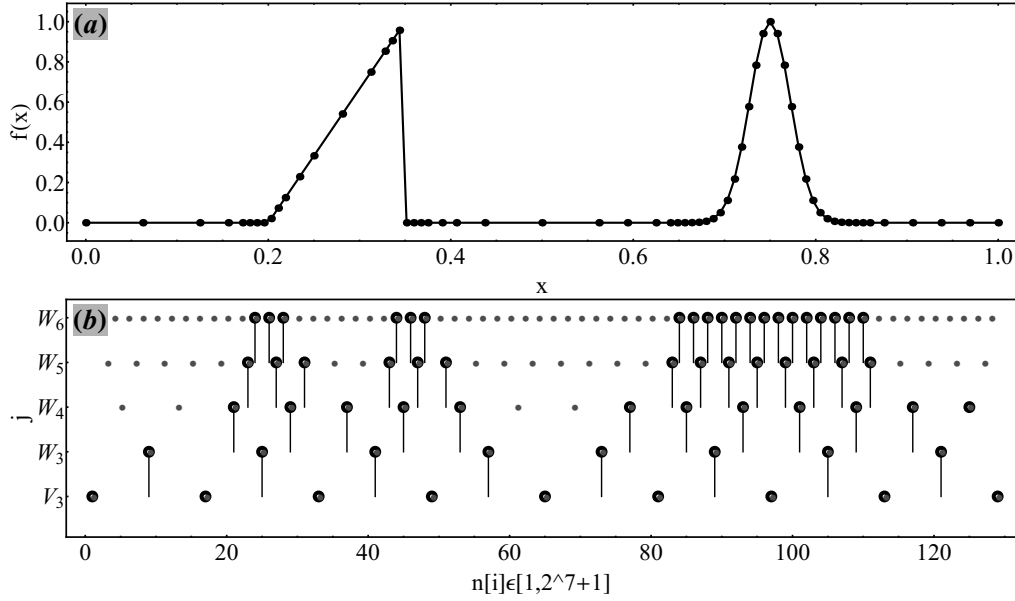


Figure 3.2: Grid adaptation for a discontinuous and a smooth but high-gradient data where  $\epsilon = 10^{-5}$ . a) the considered data (solid line) and adapted points (solid points); b) distribution of adapted points in different resolution levels: solid points are distribution of multiresolution-based sampled points, and hollow circles denotes the adapted points.

### 3.1.2 Multiresolution representation of 2-D grids

Consider a uniform grid of spatial locations  $(x, y) \in [0, 1] \times [0, 1]$ ; a set of points belonging to subspace  $V_j$  can be defined as [21]:

$$V_j = \{(x_{j,k}, y_{j,l}) : x_{j,k} = k2^{-j}, y_{j,l} = l2^{-j}\} \quad \text{for all } j, k, l \in \mathbb{Z}, \quad (3.6)$$

where  $j$  denotes resolution level and  $J_{min} \leq j \leq J_{max}$ ; coefficients  $k$  &  $l$  measure spatial locations. Figure 3.3 illustrates a schematic representation of points belonging to  $V_j$  ( $\bullet \in V_j$ ) and  $V_{j+1}$  ( $\bullet \cup \circ \in V_{j+1}$ ) for case  $j = 2$ .

Regarding Eq. (3.6), it is clear that  $u_{j,(k,l)} = u_{j+1,(2k,2l)}$  where  $u_{j,(k,l)} := u(x_{j,k}, y_{j,l})$ . This is because:  $V_{j+1} \subset V_j$ .

Same as the 1-D case, for 2-D information a detail subspace  $W_j$  belonging to subspace  $V_{j+1} \setminus V_j$  can be defined. These points are shown in Figure 3.3 with hollow circles; i.e.:  $\circ \in W_j$ . In this detail sub-space, three point types  $s_1$ ,  $s_2$  and  $s_3$  can be distinguished, as:

1. Points in even-numbered points in the  $x$  direction and odd-numbered in the  $y$

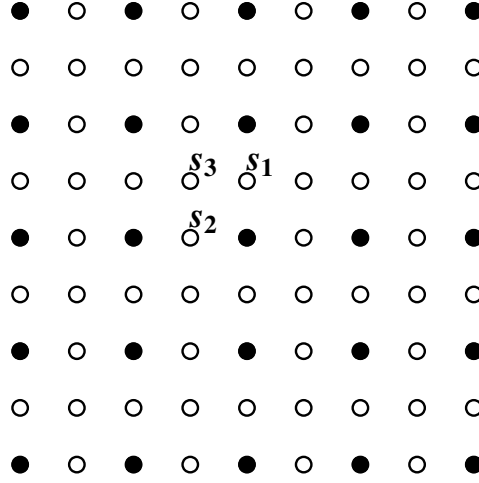


Figure 3.3: Spatial locations of points belonging to subspaces  $V_2$  and  $V_3$ , where  $\bullet \in V_2$  and  $\bullet \cup \circ \in V_3$ , for  $(x_i, y_j) \in [0, 1] \times [0, 1]$ .

direction; i.e., set  $s_1 = \{(x_{j+1,2k}, y_{j+1,2l+1})\}$ ,

2. Points in odd-numbered points in the  $x$  direction and even-numbered in the  $y$  direction; i.e., set  $s_2 = \{(x_{j+1,2k+1}, y_{j+1,2l})\}$ ,

3. Odd-numbered grid points in both directions:  $s_3 = \{(x_{j+1,2k+1}, y_{j+1,2l+1})\}$ .

Based on sub-spaces  $V_j$  and  $W_j$ , a 2-D multiresolution analysis can be obtained as the 1-D case. For point set  $\{(x_{j,k}, y_{j,l}) \in s_i : i \in \{1, 2, 3\}\}$ , four wavelet coefficients  $\{d_{j,\{k,l\}}^i : i \in \{1, 2, 3, 4\}\}$  can be defined as:

1. For points  $s_1$ , wavelet coefficients are  $d_{j,\{k,l\}}^1$ ; they measure local variations in the vertical ( $y$ ) direction,
2. For points  $s_2$ , wavelet coefficients are  $d_{j,\{k,l\}}^2$ ; they measure local variations in the horizontal ( $x$ ) direction,
3. For points  $s_3$ , two sets of wavelet coefficients can be defined as  $d_{j,\{k,l\}}^3$  and  $d_{j,\{k,l\}}^4$ ; they measure respectively local variations in the  $x$  and  $y$  directions.

Detail coefficients  $d_{j,\{k,l\}}^i$  for  $i \in \{1, 2, 3, 4\}$  can be evaluated as:

$$\begin{aligned} \{d_{j,\{k,l\}}^1 &:= u(x_{j+1,2k}, y_{j+1,2l+1}) - P_j^y [u(x_{j+1,2k}, y_{j+1,2l+1})] : k = 0 \text{ to } 2^j, \& l = 0 \text{ to } 2^j - 1\}, \\ \{d_{j,\{k,l\}}^2 &:= u(x_{j+1,2k+1}, y_{j+1,2l}) - P_j^x [u(x_{j+1,2k+1}, y_{j+1,2l})] : k = 0 \text{ to } 2^j - 1, \& l = 0 \text{ to } 2^j\}, \\ \{d_{j,\{k,l\}}^3 &:= u(x_{j+1,2k+1}, y_{j+1,2l+1}) - P_j^x [u(x_{j+1,2k+1}, y_{j+1,2l+1})] : k, l = 0 \text{ to } 2^j - 1\} \text{ and} \\ \{d_{j,\{k,l\}}^4 &:= u(x_{j+1,2k+1}, y_{j+1,2l+1}) - P_j^y [u(x_{j+1,2k+1}, y_{j+1,2l+1})] : k, l = 0 \text{ to } 2^j - 1\}; \end{aligned}$$

where:  $k, l \in \{0, 1, \dots, 2^j - 1\}$ ,  $j \in \{J_{min}, J_{min} + 1, \dots, J_{max} - 1\}$  [21]; operators  $P_j^x$  and  $P_j^y$  denote the projection (the local Lagrange interpolation) operator in the  $x$  and  $y$ -directions, respectively. Definition of these operators are the same as the 1-D case.

For 2-D adaptation of grids, a predefined threshold value  $\epsilon$  is chosen (or a set of level-dependent ones,  $\epsilon_j$ ). Points  $(x_{j+1,2k}, y_{j+1,2l+1})$  or  $(x_{j+1,2k+1}, y_{j+1,2l})$  would be eliminated if corresponding detail coefficient, respectively  $d_{j,\{k,l\}}^1$  or  $d_{j,\{k,l\}}^2$ , is less than the threshold. For adaptation of points  $(x_{j+1,2k+1}, y_{j+1,2l+1})$ , it is necessary both detail coefficients  $d_{j,\{k,l\}}^3$  and  $d_{j,\{k,l\}}^4$  are less than the threshold [21].

## 3.2 Post-processing adapted grids based on the MRA

Once adapted grids are obtained via the aforementioned wavelet-based procedures, grids are modified considering resolution level of each point; this is done to guarantee gradual density variation of grid points.

### 3.2.1 1-D grid modification

The procedure for modification of 1-D adapted grids can be summarized as:

1. Setting the level resolution ( $j$ ) equal to the finest resolution, i.e.:  $j = J_{max} - 1$ ,
2. Considering points belonging to the detail space of resolution  $j$ ; i.e., points:
 
$$\{x_{j+1,2k+1} = (2k + 1)/2^{j+1} \in W_j\},$$
3. Existence controlling of  $N_s$  neighbor-points for each side of the point  $x_{j,k} \in W_j$  at the same level of resolution, i.e., points:
 
$$\{x_{j+1,2(k+i)+1} : i \in \{-N_s, -N_s + 1, \dots, N_s\}, i \neq 0\},$$
4. Existence controlling of  $N_c$  neighbor-points for each side of the point  $x_{j,k} \in W_j$  at the successive coarser resolution (subspace  $W_{j-1}$ ). This step is only done for levels  $j > J_{min}$ ,
5. Adding the extra points from steps 3 and 4 to the adapted grid (updating the modified adapted grid),
6. If  $j > J_{min}$ , set  $j = j - 1$  and then following steps 2 through 6,
7. If  $j = J_{min}$ , consider only steps 2 and 3. These points are added to the updated adapted grid points, as the final stage.

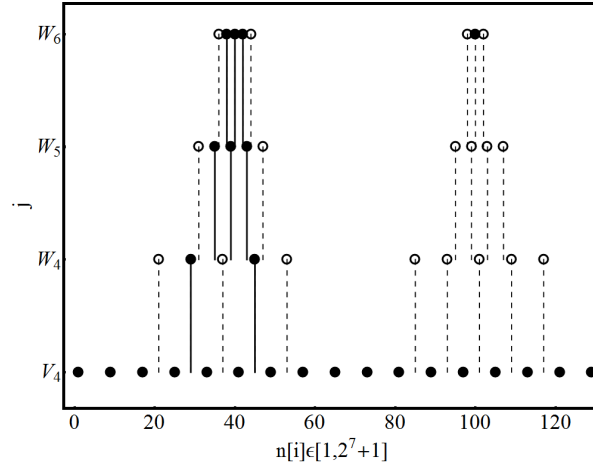


Figure 3.4: Modification of an adapted grid; solid points and hollow circles correspond to adapted and extra added points, respectively. For grid modification stage, we assume:  $N_s = N_c = 1$ .

The already-mentioned post processing procedure is illustrated in Figure 3.4; there, distribution of adapted grid points is shown in different levels of resolution. Solid points correspond to assumed (original) adapted points and the hollow ones associate with points obtained after the post-processing procedure. This modification leads to locally semi-uniform distributions.

### 3.2.2 2-D grid modifications

The 2-D grid modification will be done by controlling and adding new points in the same and coarser resolution levels, as the 1-D grid case. For this purpose, for different points  $s_1$ ,  $s_2$  or  $s_3$  (Figure 3.3), different adding procedures will be considered.

#### Adding extra new points by the multiresolution concept

1. **Adding in the same resolution** Depending on point type ( $s_1$ ,  $s_2$  or  $s_3$ ), different inserting procedures will be considered; different neighbor points will be added for each point  $s_1$ ,  $s_2$  or  $s_3$ , see Figure 3.3. There, for an adapted point belonging to  $W_2$ , new neighbor points of  $W_2$  are locally added. In this figure, hollow squares are the new extra points added around each point  $s_1$ ,  $s_2$  or  $s_3$ . Here, only one row or column of the nearest points is considered for modification stage. In general, more surrounding points of  $W_2$  can be considered for each point  $s_1$ ,  $s_2$  or  $s_3$ .



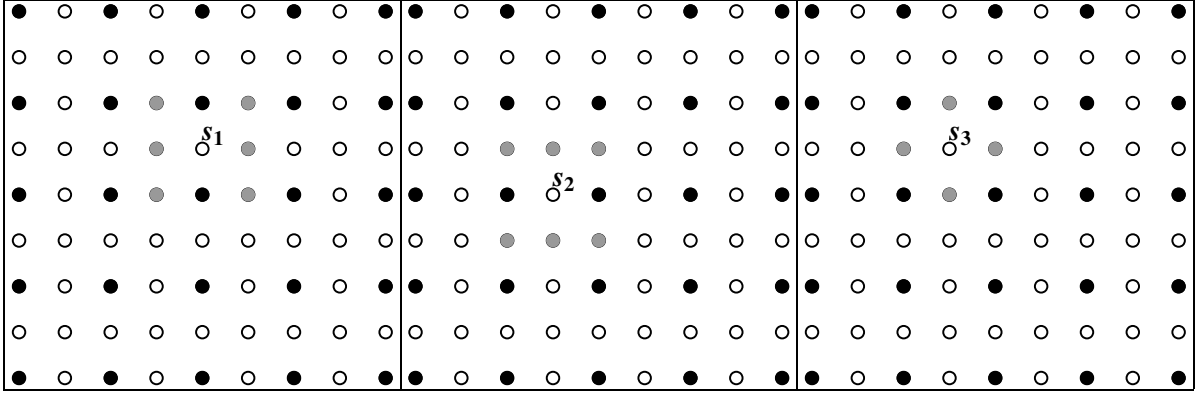


Figure 3.5: Adding procedure in the same resolution for point types  $s_1$ ,  $s_2$  or  $s_3$ , where  $\bullet \in V_2$ ,  $\bullet \cup \circ \in V_3$ ,  $\circ \in W_2$  and bright gray solid points are the added points belonging to  $W_2$ .

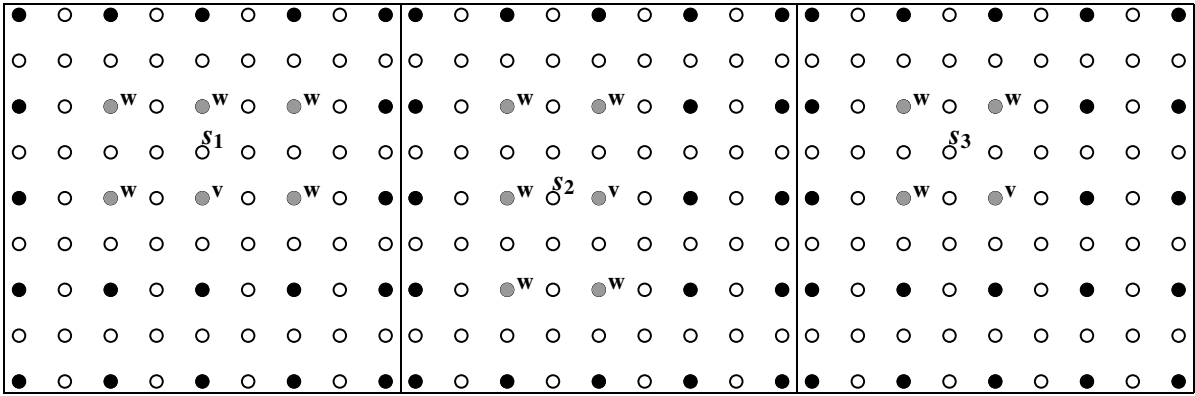


Figure 3.6: Adding procedure in the successive coarser resolution for point types  $s_1$ ,  $s_2$  or  $s_3$ , where  $\bullet \in V_2$ ,  $\bullet \cup \circ \in V_3$ ,  $\circ \in W_2$  and bright gray solid points are the added points belonging to  $V_2$ .

2. **Adding in the successive coarser resolution** In this case, for a point belonging to  $W_j$  some new surrounding points including in  $W_{j-1}$  are added. To guarantee symmetry and gradual concentration of modified adapted points, it may also be necessary to consider some extra points from sub-space  $V_{j-1}$ . For different points  $s_1$ ,  $s_2$  and  $s_3$ , different neighbor points will be considered. Such grid checking/adding procedures are shown in Figure 3.3. In each illustration, points denoted by  $w$  belong to sub-space  $W_{j-1} = W_{2-1}$  and points with name  $v$  belongs to the sub-space  $V_{j-1} = V_{2-1}$ . In these figures, only one row or column of the nearest distance is considered.

### 3.2.3 Post-processing 2-D adapted grids

The modification algorithm for 2-D grids is generally similar to the 1-D case; it can be reviewed as follows:

1. Set the level resolution  $j$  (where  $J_{min} \leq j \leq J_{max} - 1$ ) equal to the finest resolution level; i.e.,  $j = J_{max} - 1$  (with spatial sampling steps  $dx = dy = 1/2^{J_{max}-1}$ ),
2. Consider the set of points corresponding to detail sub-space of resolution  $j$ ; i.e., points  $\{(x_{j,k}, y_{j,l}) \in W_j\}$ ,
3. Control  $N_s$  neighbor-rows or columns of surrounding points for each side of the point  $(x_{j,k}, y_{j,l})$ ; this control is done at the same level of resolution,
4. Control  $N_c$  neighbor-rows or columns of points for each side of the point  $(x_{j,k}, y_{j,l})$ ; this is done at the successive coarser resolution (subspace  $W_{j-1}$ ),
5. Regarding step 3, add the surrounding points in subspace  $W_j$ ,
6. Considering step 4, add the surrounding points in subspace  $W_{j-1}$ ; in this stage, it may also be needed to add some new points in the sub-space  $V_{j-1}$  (as explained before),
7. If  $j > J_{min}$ , set  $j = j - 1$  and go back to step 2, else go to step 8,
8. If  $j = J_{min}$ , consider only steps 2, 3 and 5.

# Chapter 4

## The KT central scheme on non-uniform grids

**Introduction** Wave propagation problems are encountered in many domains of physics and engineering: from gas dynamic problems to mechanical waves. A special class of wave propagation problems are described mathematically by so-called hyperbolic systems of conservation laws. Most hyperbolic systems of conservation laws in physics are non-linear and solving them analytically is often difficult, when not impossible. Besides, such analytical solutions are limited for cases with simple boundary conditions and boundary geometry. Using of numerical methods are essential to approximate these equations. Their non-linear feature often leads to discontinuous solutions: often known as *shocks* or *shock waves*. These type solutions trigger the failure and instability of most of classical numerical methods.

The Godunov approach aims to provide a solution to this problem by including the possible discontinuous character of the solution directly in the numerical method. Godunov-type methods consider the numerical solution as being discontinuous in essence, a continuous profile being a particular case of a discontinuous one. In these methods, space is discretized into volume cells. The numerical solutions are not measured by their values at a set of (discrete) points, but by their average values over the cells. And this type of formulation leads to conservative numerical schemes in each cell by considering the exchange of fluxes at its interfaces with all surrounding cells.

The concept of the piecewise solution with possible discontinuities can prevent spurious oscillations (see Appendix A) and yields to the concept of high-resolution schemes: these methods use higher-order approximations in smooth domains and first or a lower approximation around discontinuities to prevent non-physical oscillations.

In this chapter, the following concepts will be studied:

1. The concept of Godunov approach is reviewed, at first. Based on it, two general schemes, the upwind and central methods are explained,
2. Following the central formulation, the KT scheme is provided for non-uniform cells/grids,
3. Non-linear stability conditions, here TVD, are surveyed for uniform cells at first, and then are modified to include non-uniform cells.

Finally some 1-D and 2-D numerical examples are provided to confirm capability of the adaptive KT solver.

## 4.1 The KT central scheme on non-uniform grids

### 4.1.1 Godunov-type schemes

The aim of this study is numerical simulation of the hyperbolic system of conservation laws:

$$u_t + f(u)_x = 0, \quad (4.1)$$

by Godunov-type schemes. For this, *sliding* average of  $u(., t)$  of Eq. (4.1) is computed over the cell  $I_x$  of length  $\Delta x$ , as:

$$\begin{aligned} \frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} u(\zeta, t)_t d\zeta + \frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} f(u(\zeta, t))_\zeta d\zeta = 0, \Rightarrow \\ \bar{u}_t + \frac{1}{\Delta x} \left[ f \left( u \left( x + \frac{\Delta x}{2}, t \right) \right) - f \left( u \left( x - \frac{\Delta x}{2}, t \right) \right) \right] = 0, \end{aligned} \quad (4.2)$$

where  $I_x = \left\{ \zeta \mid |\zeta - x| \leq \frac{\Delta x}{2} \right\}$ ;  $\bar{u}(x, t) := \frac{1}{|I_x|} \int_{I_x} u(\zeta, t) d\zeta$ .

In the next step, Eq. (4.2) is integrated over time with sufficiently small time-step  $\Delta t$ , i.e.  $t \leq \tau \leq t + \Delta t$ , as:

$$\bar{u}(x, t + \Delta t) = \bar{u}(x, t) - \frac{1}{\Delta x} \left[ \int_t^{t+\Delta t} f \left( u \left( x + \frac{\Delta x}{2}, \tau \right) \right) d\tau - \int_t^{t+\Delta t} f \left( u \left( x - \frac{\Delta x}{2}, \tau \right) \right) d\tau \right]. \quad (4.3)$$

We assume that the solution of Eq. (4.1) at time  $t = t^n$  is a piecewise polynomial in

the form:

$$w(x, t^n) = \sum_j [p_j(x)\chi_j(x)], \quad \chi_j(x) := 1_{I_j}, \quad (4.4)$$

where  $p_j(x)$  is a algebraic polynomial defined on the cell  $I_j(x)$ . The set  $\{p_j(x)\}$  forms a piecewise polynomial. Evaluation of  $w(\cdot, t^n)$  in time by Eq. (4.3) leads to:

$$\bar{w}(x, t + \Delta t^n) = \bar{w}(x, t^n) - \frac{1}{\Delta x} \left[ \int_t^{t+\Delta t} f\left(w\left(x + \frac{\Delta x}{2}, \tau\right)\right) d\tau - \int_t^{t+\Delta t} f\left(w\left(x - \frac{\Delta x}{2}, \tau\right)\right) d\tau \right]. \quad (4.5)$$

This form of representation is known as the *fully-discrete* form, since Eq. (4.1) is discretized both in the time and space.

Based on *sampling* procedure of  $x$  in Eq. (4.5), two general methods can be provided: *Upwind* and *Central* schemes.

### Upwind schemes

Let us sample Eq. (4.5) at the mid-cells,  $x = x_\nu$ ; then the evolved solution is:

$$\bar{w}_\nu^{n+1} = \bar{w}_\nu^n - \frac{1}{\Delta x} \left[ \int_t^{t+\Delta t} f\left(w\left(x_{\nu+\frac{1}{2}}, \tau\right)\right) d\tau - \int_t^{t+\Delta t} f\left(w\left(x_{\nu-\frac{1}{2}}, \tau\right)\right) d\tau \right], \quad (4.6)$$

where  $\bar{w}_\nu^n := \frac{1}{\Delta x} \int_{I_\nu} w(\zeta, t^n) d\zeta$  and  $x_{\nu\pm\frac{1}{2}} := x_\nu \pm \frac{\Delta x}{2}$ . It is clear that, the above equation gives the average solutions on *non-staggered* grids.

Based on aforementioned formulations/descriptions, for utilizing Eq. (4.6), the following steps should be followed:

1. For given cell averages  $\bar{w}_\nu^n$ , *reconstruct* the point values  $w(\cdot, t^n)$  as piecewise polynomial approximation:

$$w(x, t^n) = \sum_j [p_j(x)\chi_j(x)], \quad \bar{p}_\nu(x_\nu) = \bar{w}_\nu^n, \quad (4.7)$$

2. Compute  $\bar{w}_\nu^{n+1}$  by the *evolution* procedure, Eq. (4.6). There,  $w(x_{\nu+\frac{1}{2}}, \tau \geq t^n)$  are determined by the *generalized Riemann* problems:

$$w_t + f(w)_x = 0, \quad t \geq t^n, \quad w(x, t^n) = \begin{cases} p_\nu(x), & x < x_{\nu+\frac{1}{2}}, \\ p_{\nu+1}(x), & x > x_{\nu+\frac{1}{2}}. \end{cases} \quad (4.8)$$

It should be mentioned that due to possible discontinuities at  $x_{\nu+\frac{1}{2}}$  at  $t^n$ , a family of nonlinear waves/shocks can be produced around  $x_{\nu+\frac{1}{2}}$  in cells  $I_\nu$  and  $I_{\nu+1}$ . To handle such solutions, an exact Riemann solver or at least an approximating one is essential to be used for solving such waves.

The original Godunov scheme uses a constant-piecewise reconstruction as:  $w(x, t^n) = \sum_j \bar{w}_j^n \chi_j$  by an exact Riemann solver.

### Central schemes

In this approach, a piecewise polynomial,  $w(x, t^n) = \sum_j p_j(x) \chi_j(x)$ , is used to approximate numerical solutions satisfying the sliding average solution (4.5). Here again, in each cell  $I_\nu$ , a polynomial is defined with possible jumps in cell-interfaces,  $x_{\nu\pm\frac{1}{2}} := x_\nu \pm \frac{\Delta x}{2}$ .

In the central schemes, the sampling points in Eq. (4.5) are cell interfaces,  $x = x_{\nu+\frac{1}{2}}$ , which leads to the relationship:

$$\bar{w}(x_{\nu+\frac{1}{2}}, t^{n+1}) = \bar{w}(x_{\nu+\frac{1}{2}}, t^n) - \frac{1}{\Delta x} \left[ \int_{t^n}^{t^{n+1}} f(w(x_{\nu+1}, \tau)) d\tau - \int_{t^n}^{t^{n+1}} f(w(x_\nu, \tau)) d\tau \right]. \quad (4.9)$$

The evolved solutions (4.9) are obtained on *staggered* grids.

In brief, the solution procedure consists of:

1. *Reconstruct* the piecewise polynomial  $w(., t^n)$  based on cell average values  $\{\bar{w}_\nu^n\}$ , as:

$$w(x, t^n) = \sum_j p_j(x) \chi_j(x), \quad \bar{p}_\nu(x_\nu) = \bar{w}_\nu^n, \quad (4.10)$$

and then compute staggered average values  $\bar{w}_{\nu+\frac{1}{2}}^n$  as:

$$\bar{w}_{\nu+\frac{1}{2}}^n = \frac{1}{\Delta x} \left[ \int_{x_\nu}^{x_{\nu+\frac{1}{2}}} p_\nu(x) dx + \int_{x_{\nu+\frac{1}{2}}}^{x_{\nu+1}} p_{\nu+1}(x) dx \right], \quad (4.11)$$

2. Evaluate the  $\bar{w}_{\nu+\frac{1}{2}}^{n+1}$  values from Eq. (4.9): the *evolution* stage. Let  $\{a_k(u)\}_k$  denotes the eigenvalues of the Jacobian  $A(u) := \frac{\partial f}{\partial u}$ . Then possible jumps in interfaces  $x_{\nu+1/2}$  can not reach cell-centers  $x_\nu$  for sufficiently small values of  $\Delta t$ , if  $\Delta t \leq \frac{1}{2} \Delta x \cdot \max_k |a_k(u)|$ ; this condition is known as the *CFL* (Courant-Friedrichs-Lewy) condition. With this condition, both flux functions  $f(x_\nu, \tau \geq t^n)$  and  $f(x_{\nu+1}, \tau \geq$

$t^n$ ) remain smooth during time interval  $t^n \leq \tau \leq t^{n+1}$ . And a high order quadrature rule can be used for the time integration in Eq. (4.9).

### 4.1.2 Central high resolution schemes

As mentioned in the previous subsection, for utilizing central schemes, two steps are followed for handling Eq. (4.9): 1) *Reconstruction*, and 2) *Evolution*. The resulted average solutions,  $\bar{w}_{\nu+1/2}^{n+1}$ , are obtained on staggered grids. On the other hand, staggered schemes are more dissipative than non-staggered ones. For this reason, another step can be used to *project* the solutions  $\bar{w}_{\nu+1/2}^{n+1}$  on non-staggered grids  $x_\nu$  [75]. For such projection, the evolved solutions  $\bar{w}_{\nu+1/2}^{n+1}$  are first reconstructed as [75]:

$$w(., t^{n+1}) = \sum_j p_{j+1/2}(w) \chi_{j+1/2}(x) \quad (4.12)$$

and then the projected solution  $\bar{w}_\nu^{n+1}$  is [75]:

$$\bar{w}_\nu^{n+1} = \frac{1}{x_{\nu+1/2} - x_{\nu-1/2}} \left\{ \int_{x_{\nu-1/2}}^{x_\nu} p_{\nu-1/2} dx + \int_{x_\nu}^{x_{\nu+1/2}} p_{\nu+1/2} dx \right\}. \quad (4.13)$$

This projection for different methods is presented in Figures. A.4 (for LxF), A.6 (for NT) and A.8 (for KT) in Appendix A.

Following the concept of the central schemes, different methods are developed, some of the most important ones are:

1. Staggered and non-staggered *Lax-Friedrichs* (LxF) schemes (with first-order accuracy) [74, 75],
2. Staggered and non-staggered *Nessyahu-Tadmor* (NT) schemes (with second-order accuracy) [25, 75],
3. The *Kurganov-Tadmor* (KT) scheme (with second-order accuracy) [22],
4. The *Liu-Tadmor* third-order scheme [76].

The details of deriving the LxF, NT and KT schemes are presented on uniform grids in Appendix A.

### The KT central scheme on centered non-uniform grid points

Here, the semi-discretized form of a scalar first-order hyperbolic system,  $u_t + F(u)_x = 0$ , will be provided for the KT method [22] on centered non-uniform grids, with cell centers  $x_i = (x_{i-1/2} + x_{i+1/2})/2$ ; where  $x_i$  is cell center for cell  $I_i$  and  $x_{i\pm 1/2}$  are  $i^{\text{th}}$  cell edges.

In this study, the semi-discrete form of hyperbolic systems  $u_t + F(u)_x = 0$  will be considered. To derive this form, the Reconstruction/Evolution/Projection (REP) concept is used, see Figure 4.1. For the KT method, these steps can be summarized as:

1. **Reconstruction:** In this stage, at time step  $t = t^n$ , the solution  $u(x, t^n)$  is reconstructed over cells in a *linear* piecewise form based on cell average values at cell-centers,  $u(x_i, t^n) = u_i^n$  (In general, in each cell, the reconstructed values can be constant, linear, or higher order one for providing first, second and higher-order accuracy, respectively).
2. **Evolution:** After reconstruction, the average solution  $u(x_i, t^n)$  is evolved in time to obtain solution  $u(x_{i+1/2}, t^{n+1})$  at the next time step (it should be mentioned that since reconstruction is linear the cell-average values are equal to cell-center values),
3. **Projection:** Evolved solutions are obtained on staggered grids  $x_{i+1/2}$ . To project these solutions on the previous original grid, the projection step is done by averaging the new solution  $u(x, t^{n+1})$  on the previous original grid. In this stage, before averaging, the solution  $u(x, t^{n+1})$  is again reconstructed in a piecewise form on staggered grids (cells). For the reconstruction, smooth and non-smooth zones can be distinguished from each other and different piecewise solution (with different variations) can then be considered for different regions (this will be explained in more detail).

Considering the REP procedure, for the  $i^{\text{th}}$  cell with cell-centered middle point  $x_i$  (i.e.,  $x_i = (x_{i+1/2} - x_{i-1/2})/2$ ), it is easy to show that the semi-discrete form of the 1-D scalar hyperbolic equation is:

$$\frac{du_i}{dt} + \frac{F(u_{i+1/2}^*) - F(u_{i-1/2}^*)}{\Delta x_i} = 0, \quad (4.14)$$

where:  $\bar{u}_i \equiv u_i = \left\{ \int_{x_{i-1/2}}^{x_{i+1/2}} u(x, t) dt \right\} / \Delta x_i$  denotes the cell center solution (estimated state values) on  $i^{\text{th}}$  cell;  $F(u_{i\pm 1/2}^*)$  shows a proper combination of left and right recon-



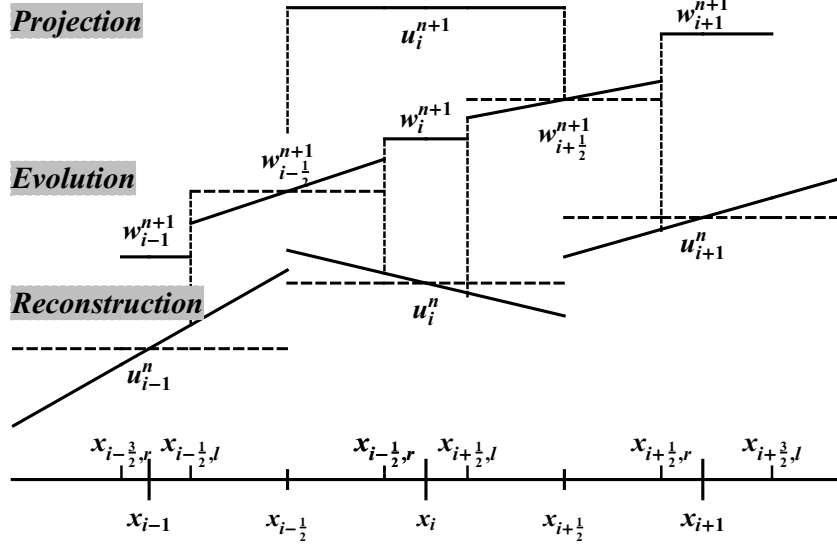


Figure 4.1: The reconstruction/evolution/projection concept: the second order accuracy. In this figure  $x_{i\pm k,m} \equiv x_{i\pm k,m}^n$ , for  $m \in \{l, r\}$ .

structed state values and corresponding fluxes at cell edges  $x_{i\pm 1/2}$ . The left and right reconstructed state values are shown respectively by  $u_{i+1/2}^L$  and  $u_{i+1/2}^R$ ; for spatially second order methods, these values can be evaluated as:  $u_{i+1/2}^L = u_i + (u_x)_i (x_{i+1/2} - x_i)$  and  $u_{i+1/2}^R = u_{i+1} - (u_x)_{i+1} (x_{i+1} - x_{i+1/2})$  where  $(u_x)_i$  shows a limited slope at point  $x_i$ ;  $\Delta x_i = x_{i+1/2} - x_{i-1/2}$  is the  $i^{\text{th}}$  cell length. It should be mentioned that for deriving the semi-discrete form, at first, the fully-discrete form is derived (for evolution from time  $t^n$  to  $t^{n+1} = t^n + \Delta t$ ); the semi-discrete form can be obtained as  $\Delta t \rightarrow 0$  (for details see Appendix A). The resulted equation is continuous in the time but discrete in the spatial domain.

For the KT scheme, the general assumptions are:

1. To estimate non-smooth zones around cell edges  $x_{i+1/2}$  (in the projection step), maximum values of local propagation speeds  $a_{i+1/2}^n = a(u(x_{i+1/2}, t^n))$  are used regardless of propagation directions; i.e., if eigenvalues of  $(\partial F / \partial u)_{i+1/2}$  are  $\lambda_1 < \lambda_2 < \dots < \lambda_N$ , then  $a_{i+1/2} \equiv a_{i+1/2}^n = \max_{j \in \{1, N\}} (|\lambda_j(u_{i+1/2}^L)|, |\lambda_j(u_{i+1/2}^R)|)$ . The estimated non-smooth zone belongs then to region  $[x_{i+1/2,l}^n, x_{i+1/2,r}^n]$ , where:  $x_{i+1/2,l}^n = x_{i+1/2} - a_{i+1/2}^n \Delta t$  and  $x_{i+1/2,r}^n = x_{i+1/2} + a_{i+1/2}^n \Delta t$ . This non-smooth zone is reconstructed by a linear function. For sufficiently small values of  $\Delta t$  (i.e.,  $\Delta t \leq \frac{1}{2} a_{i+1/2}^n \Delta x_i$ ), the possible shock-like solutions do not reach the cell center. So a smooth zone exists in each cell for  $x \in (x_{i-1/2,r}^n, x_{i+1/2,l}^n)$ ; this solution is reconstructed by a constant function, see Figure 4.1.

2. In projection step, to estimate a limited slope  $(u_x)_{i+1/2}^{n+1}$  at center of the non-smooth zone,  $x_{i+1/2}$ , distance of this point from middle of surrounding smooth zones are considered; for point  $x_{i+1/2}$  the left ( $\Delta_L$ ) and right ( $\Delta_R$ ) distances are:  $\Delta_L = \frac{(x_{i+1/2,r}^n - x_{i+1/2,l}^n)}{2} + \frac{(x_{i+1/2,l}^n - x_{i-1/2,r}^n)}{2} = \frac{(x_{i+1/2,r}^n - x_{i-1/2,r}^n)}{2}$  and  $\Delta_R = \frac{(x_{i+1/2,r}^n - x_{i+1/2,l}^n)}{2} + \frac{(x_{i+3/2,l}^n - x_{i+1/2,r}^n)}{2} = \frac{(x_{i+3/2,l}^n - x_{i+1/2,l}^n)}{2}$ . Based on these distances, the limited slope  $(u_x)_{i+1/2}^{n+1}$  can be obtained. For example by the *generalized MINMOD* limiter, the limited slope is:

$$(u_x)_{i+1/2}^{n+1} = \text{MINMOD} \left( \theta \frac{w_{i+1}^{n+1} - w_{i+1/2}^{n+1}}{\Delta_R}, \frac{w_{i+1}^{n+1} - w_i^{n+1}}{\Delta_R + \Delta_L}, \theta \frac{w_i^{n+1} - w_{i-1/2}^{n+1}}{\Delta_L} \right), \quad (4.15)$$

where: the MINMOD function itself is defined as:

$$\text{MINMOD}(z_1, z_2, \dots) := \begin{cases} \min_i(z_i), & \text{if } z_i > 0 \quad \forall i, \\ \max_i(z_i), & \text{if } z_i < 0 \quad \forall i, \\ 0, & \text{if } z_i = 0 \quad \forall i. \end{cases} \quad (4.16)$$

In the generalized MINMOD limiter, the parameter  $\theta$  controls the dissipation in numerical solutions:  $\theta = 1$  and  $\theta = 2$  lead to the most and the least dissipated results, respectively. In this regard, this parameter is chosen such that  $1 \leq \theta \leq 2$  [22].

Regarding these assumptions, it can be shown that:

$$F(u_{i+1/2}^*) = \frac{1}{2} \left\{ [F(u_{i+1/2}^R) + F(u_{i+1/2}^L)] - a_{i+1/2} [u_{i+1/2}^R - u_{i+1/2}^L] \right\}. \quad (4.17)$$

In formulation of the central high resolution schemes, it is not necessary to know the directions of propagating fronts. This makes these methods simple to implement with straightforward algorithms. The first proposed central scheme is the LxF method with one order spatial accuracy; to improve the accuracy, the LxF method was then modified by Nessyahu and Tadmor (the NT scheme) to achieve a second order accuracy scheme. The shortcoming of both the LxF and NT methods is that they use staggered grids. This leads to more dissipative algorithms, generally. For this, Kurganov and Tadmor modified the NT scheme to work on non-staggered grids, the KT scheme; the accuracy

of this new scheme can be of second or higher order, depending of polynomial-order in the construction stage.

### 4.1.3 Central high-resolution schemes on non-centered non-uniform 1-D cells

By using the cell-centered non-uniform cells, the TVD stability condition does not satisfy without altering definition of slope/flux limiters. To preserve the TVD condition without limiter modification, it will be shown that by shifting cell centers in some cells, the TVD condition can be reached. These cells act as transmitting cells connecting two surrounding uniform cells of different lengths. By such cell-center shifting, a new source of the truncation error will be introduced. Effects of this new error will be discussed.

Considering the REP, the full-discrete and semi-discrete forms of the scalar hyperbolic equation  $u_t + F(u)_x = 0$  will be provided on non-centered non-uniform cells. For generality, it is assumed that cell centers  $x_j$  are not located at cell centers; so, the left and right cell-edge positions are:  $x_{j-1/2} := x_j - (1 - p_j)\Delta x_j$  and  $x_{j+1/2} := x_j + p_j\Delta x_j$ , where  $0 < p_j < 1$  (for  $p_j = 1/2$  the cell center  $x_j$  is the middle point of  $j^{\text{th}}$  cell).

**The reconstruction and evolution stages** In the reconstruction stage, the linear piecewise function is used, 4.1. For the evolution stage, the calculations are as follows:

1. The spatio-temporal volume  $\Delta x_{j+1/2} \times \Delta t$  is considered, where  $\Delta x_{j+1/2} := x_{j+1/2,r}^n - x_{j+1/2,l}^n = 2\Delta t a_{j+1/2}^n$ , and  $\Delta t := t^{n+1} - t^n$ . By averaging on this volume and considering the midpoint integration rule in time, the solution  $w_{j+1/2}^{n+1} := w(x_{j+1/2}, t^{n+1})$  can be obtained as:

$$\begin{aligned}
 w_{j+1/2}^{n+1} &= \frac{1}{\Delta x_{j+1/2}} \int_{x_{j+1/2,l}^n}^{x_{j+1/2,r}^n} u(x, t^{n+1}) dx \\
 &= \frac{1}{\Delta x_{j+1/2}} \int_{x_{j+1/2,l}^n}^{x_{j+1/2,r}^n} u(x, t^n) dx - \frac{1}{\Delta x_{j+1/2}} \int_{t^n}^{t^{n+1}} \left[ F(u_{j+1/2,r}^n) - F(u_{j+1/2,l}^n) \right] dt \\
 &= \frac{1}{4} \left( -\Delta t a_{j+\frac{1}{2}}^n (u_x)_j^n + \Delta t a_{j+\frac{1}{2}}^n (u_x)_{j+1}^n \right. \\
 &\quad \left. + 2\Delta x_j p_j (u_x)_j^n + 2\Delta x_{j+1} (p_{j+1} - 1) (u_x)_{j+1}^n + 2u_j + 2u_{j+1} \right) \\
 &\quad - \frac{1}{2a_{j+\frac{1}{2}}^n} \left( F_{j+\frac{1}{2},r}^{n+\frac{1}{2}} - F_{j+\frac{1}{2},l}^{n+\frac{1}{2}} \right),
 \end{aligned} \tag{4.18}$$

where  $F_{j+\frac{1}{2},r}^{n+\frac{1}{2}} := F\left(u\left(x_{j+\frac{1}{2},r}, t^{n+\frac{1}{2}}\right)\right)$  and  $F_{j+\frac{1}{2},l}^{n+\frac{1}{2}} := F\left(u\left(x_{j+\frac{1}{2},l}, t^{n+\frac{1}{2}}\right)\right)$ .

2. The spatio-temporal volume  $\Delta x_j^{l-r} \times \Delta t$  is considered, where  $\Delta x_j^{l-r} := x_{j+1/2,l}^n - x_{j-1/2,r}^n = \Delta x_j - \Delta t (a_{j-1/2}^n + a_{j+1/2}^n)$ . By averaging on this volume and considering the midpoint rule in time, the evolved solution  $w_j^{n+1} := w(x_j, t^{n+1})$  is:

$$\begin{aligned}
 w_j^{n+1} &= \frac{1}{\Delta x_j^{l-r}} \int_{x_{j-1/2,r}^n}^{x_{j+1/2,l}^n} u(x, t^{n+1}) dx \\
 &= \frac{1}{\Delta x_j^{l-r}} \int_{x_{j-1/2,r}^n}^{x_{j+1/2,l}^n} u(x, t^n) dx - \frac{1}{\Delta x_j^{l-r}} \int_{t^n}^{t^{n+1}} \left[ F(u_{j+1/2,l}^n) - F(u_{j-1/2,r}^n) \right] dt \\
 &= \frac{1}{2} \left( \Delta t \left( a_{j-\frac{1}{2}}^n - a_{j+\frac{1}{2}}^n \right) (u_x)_j^n + \Delta x_j (2p_j - 1) (u_x)_j^n + 2u_j \right) - \frac{\Delta t \left( F_{j+\frac{1}{2},l}^{n+\frac{1}{2}} - F_{j-\frac{1}{2},r}^{n+\frac{1}{2}} \right)}{\Delta x_j - \Delta t \left( a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n \right)}. \tag{4.19}
 \end{aligned}$$

3. Regarding volume  $\Delta x_{j-1/2} \times \Delta t$ , where  $\Delta x_{j-1/2} := x_{j-1/2,l}^n - x_{j-1/2,r}^n$ , the evolved solution  $w_{j-1/2}^{n+1} := w(x_{j-1/2}, t^{n+1})$  can be obtained with the similar procedure explained in Eq. (4.18); it is:

$$\begin{aligned}
 w_{j-1/2}^{n+1} &= \frac{1}{\Delta x_{j-1/2}} \int_{x_{j-1/2,l}^n}^{x_{j-1/2,r}^n} u(x, t^{n+1}) dx \\
 &= \frac{1}{\Delta x_{j-1/2}} \int_{x_{j-1/2,l}^n}^{x_{j-1/2,r}^n} u(x, t^n) dx - \frac{1}{\Delta x_{j-1/2}} \int_{t^n}^{t^{n+1}} \left[ F(u_{j-1/2,r}^n) - F(u_{j-1/2,l}^n) \right] dt \\
 &= \frac{1}{4} \left( -\Delta t a_{j-\frac{1}{2}}^n (u_x)_{j-1}^n + \Delta t a_{j-\frac{1}{2}}^n (u_x)_j^n + 2\Delta x_j (p_j - 1) (u_x)_j^n \right. \\
 &\quad \left. + 2p_{j-1} \Delta x_{j-1} (u_x)_{j-1}^n + 2u_{j-1} + 2u_j \right) - \frac{1}{2a_{j-\frac{1}{2}}^n} \left( F_{j-\frac{1}{2},r}^{n+\frac{1}{2}} - F_{j-\frac{1}{2},l}^{n+\frac{1}{2}} \right). \tag{4.20}
 \end{aligned}$$

Considering  $w_{j+1/2}^{n+1}$ ,  $w_j^{n+1}$ , and  $w_{j-1/2}^{n+1}$  definitions in Eqs. (4.18), (4.19) and (4.20), a piecewise-linear approximation on the staggered grid in the evolution stage is (see Figure (4.1)):

$$\tilde{w}(x, t^{n+1}) := \sum_j \left\{ w_{j+1/2}^{n+1} + (u_x)_{j+1/2}^{n+1} (x - x_{j+1/2}) \mathbf{1}_{[x_{j+1/2,l}^n, x_{j+1/2,r}^n]} + w_j^{n+1} \mathbf{1}_{[x_{j-1/2,r}^n, x_{j+1/2,l}^n]} \right\},$$

where  $(u_x)_{j\pm 1/2}^{n+1}$  denote limited derivatives (calculated at time  $t^{n+1}$ ) and  $\mathbf{1}_{[a,b]}$  shows the unit (box) function having unit value on the spatial interval  $[a, b]$  and is zero elsewhere.

**The projection stage** The fully discrete second-order central scheme can be obtained by averaging  $\tilde{w}(x, t^{n+1})$  on the interval  $[x_{j-1/2}, x_{j+1/2}]$ , as:

$$\begin{aligned}
 \bar{u}_j^{n+1} &= \frac{1}{\Delta x_j} \int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{w}(x, t^{n+1}) dx \\
 &= \frac{1}{4\Delta x_j} \left[ 2 \left( \Delta t \left( F_{j-\frac{1}{2},l}^{n+\frac{1}{2}} - F_{j+\frac{1}{2},l}^{n+\frac{1}{2}} + F_{j-\frac{1}{2},r}^{n+\frac{1}{2}} - F_{j+\frac{1}{2},r}^{n+\frac{1}{2}} \right) + \Delta x_j \left( \Delta x_j (2p_j - 1) (u_x)_j^n + 2u_j \right) \right) \right. \\
 &\quad + 2\Delta t a_{j-\frac{1}{2}}^n \left( \Delta x_{j-1} p_{j-1} (u_x)_{j-1}^n - \Delta x_j (p_j - 1) (u_x)_j^n + u_{j-1} - u_j \right) \\
 &\quad + 2\Delta t a_{j+\frac{1}{2}}^n \left( \Delta x_j (-p_j) (u_x)_j^n + \Delta x_{j+1} (p_{j+1} - 1) (u_x)_{j+1}^n - u_j + u_{j+1} \right) \\
 &\quad - \Delta t^2 \left( a_{j-\frac{1}{2}}^n \right)^2 \left( (u_x)_{j-1}^n + (u_x)_j^n - 2(u_x)_{j-1/2}^{n+1} \right) \\
 &\quad \left. + \Delta t^2 \left( a_{j+\frac{1}{2}}^n \right)^2 \left( (u_x)_j^n + (u_x)_{j+1}^n - 2(u_x)_{j+1/2}^{n+1} \right) \right], \tag{4.21}
 \end{aligned}$$

where  $\bar{u}_j^{n+1}$  denotes average of solution  $u(x, t^{n+1})$  on the cell  $I_j \in [x_{j-1/2}, x_{j+1/2}]$ .

**The semi-discrete form** Regarding the full discrete form (Eq. (4.21)), for small enough  $\Delta t$  values, the following approximations can be assumed: 1) ignoring of terms corresponding to  $\Delta t^2$ ; 2)  $F_{j\pm\frac{1}{2},r}^{n+\frac{1}{2}} \approx F_{j\pm\frac{1}{2},r}^n = F_{j\pm 1/2}^R$  and  $F_{j\pm\frac{1}{2},l}^{n+\frac{1}{2}} \approx F_{j\pm\frac{1}{2},l}^n = F_{j\pm 1/2}^L$  (due to the Taylor expansion in smooth zones). By these approximations, Eq. (4.21) can be rearranged and rewritten as:

$$\begin{aligned}
 \frac{\bar{u}_j^{n+1} - u_j^n}{\Delta t} &\approx \frac{1}{4\Delta x_j} \left\{ 2 \left( -F_{j+\frac{1}{2},l}^n - F_{j+\frac{1}{2},r}^n + F_{j-\frac{1}{2},l}^n + F_{j-\frac{1}{2},r}^n \right) \right. \\
 &\quad - 2 \left( -u_{j-1} a_{j-\frac{1}{2}}^n + u_j a_{j-\frac{1}{2}}^n + u_j a_{j+\frac{1}{2}}^n - u_{j+1} a_{j+\frac{1}{2}}^n \right) \\
 &\quad + 2 \left( \Delta x_{j-1} p_{j-1} a_{j-\frac{1}{2}}^n (u_x)_{j-1}^n + \Delta x_j \left( a_{j-\frac{1}{2}}^n - p_j \left( a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n \right) \right) (u_x)_j^n \right) \\
 &\quad \left. + 2 \left( \Delta x_{j+1} (p_{j+1} - 1) a_{j+\frac{1}{2}}^n (u_x)_{j+1}^n \right) \right\} - \left( \frac{1}{2} - p_j \right) (u_x)_j^n \frac{\Delta x_j}{\Delta t}. \tag{4.22}
 \end{aligned}$$

Since  $u_{j+1/2}^L := u_j + p_j \Delta x_j (u_x)_j^n$  and  $u_{j-1/2}^R := u_j - (1 - p_j) \Delta x_j (u_x)_j^n$ , Eq. (4.22) can be rewritten as:

$$\frac{\bar{u}_j^{n+1} - u_j^n}{\Delta t} + \frac{F_{j+1/2}^* - F_{j-1/2}^*}{\Delta x_j} \approx Q_j^n, \tag{4.23}$$

where  $F_{j+1/2}^* := \frac{F_{j+1/2}^R + F_{j+1/2}^L}{2} - \frac{a_{j+1/2}}{2} (u_{j+1/2}^R - u_{j+1/2}^L)$  and  $Q := -\left(\frac{1}{2} - p_j\right) (u_x)_j^n \frac{\Delta x_j}{\Delta t}$ . On the other hand, spatial distance between the midpoint  $\bar{x}_j$  and  $x_j$  is:  $\delta_j := x_j - \bar{x}_j = [x_{j+1/2} - p_j \Delta x_j] - [x_{j+1/2} - \frac{\Delta x_j}{2}] = \left(\frac{1}{2} - p_j\right) \Delta x_j$ . So average solutions  $\bar{u}_j^n$  relate to point value solutions,  $u_j^n$  as:  $u_j^n \approx \bar{u}_j^n + \delta_j \times (u_x)_j^n$ . In this regard, Eq. (4.23) can be written as:

$$\frac{\bar{u}_j^{n+1} - \bar{u}_j^n}{\Delta t} + \frac{F_{j+1/2}^* - F_{j-1/2}^*}{\Delta x_j} \approx 0. \quad (4.24)$$

At the limit  $\Delta t \rightarrow 0$ , this equation becomes:

$$\frac{d\bar{u}_j}{dt} + \frac{F_{j+1/2}^* - F_{j-1/2}^*}{\Delta x_j} = 0. \quad (4.25)$$

So, in the sense of average solutions, the semi-discrete form remains conservative. We can rewrite Eq. (4.23) based on point value solutions,  $u_j^n$ ; for this purpose, we approximate  $u_j^{n+1}$  as  $u_j^{n+1} = \bar{u}_j^{n+1} + \delta_j \times (u_x)_j^{n+1}$ ; so we have:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{F_{j+1/2}^* - F_{j-1/2}^*}{\Delta x_j} \approx Q_j^n - Q_j^{n+1}, \quad (4.26)$$

We conclude this section with the following remarks regarding (Eqs. (4.23) and (4.26)).

1. The term  $Q$  acts as a new source of the truncation error,
2. Effects of  $\Delta t$  in  $Q$  can be controlled by small-enough spatial sampling steps  $\Delta x_j$ . This can be done by a proper grid adaptation procedure around high-gradient zones,
3. If  $x_j$  is in the middle of the  $j^{\text{th}}$  cell, i.e.  $p = 1/2$ , the truncation error  $Q$  vanishes:  $Q = 0$ ,
4. Around discontinuities, since limited slopes  $(u_x)_j^n$  are nearly zero, the error  $Q$  approaches to zero,
5. Average solution on  $j^{\text{th}}$  cell is  $\bar{u}_j^n := (u_j^n)_{ave} = u_j^n + Q \times \Delta t$ .

### Some remarks

1. In this section, the generalized MINMOD limiter is used for estimation of slope  $(u_x)_i$  as [22]:

$$(u_x)_i = \text{MINMOD}\left(\theta \frac{u_i - u_{i-1}}{\Delta_L^n}, \frac{u_{i+1} - u_{i-1}}{\Delta_L^n + \Delta_R^n}, \theta \frac{u_{i+1} - u_i}{\Delta_R^n}\right), \quad \theta \in [1, 2], \quad (4.27)$$

where  $\Delta_R^n := p_i \Delta x_i + (1 - p_{i+1}) \Delta x_{i+1}$ , and  $\Delta_L^n := p_{i-1} \Delta x_{i-1} + (1 - p_i) \Delta x_i$ .

2. The semi-discrete form (Eq. (4.24) or (4.26)) can be evaluated in time by a standard time-integration method, such as the Runge-Kutta (RK), or Adams-Bashford scheme. Here, the time integration is performed by the second-order RK solver having the TVD property [22]; this algorithm is:

$$\begin{aligned} \bar{u}_i^{(1)} &= \bar{u}_i^{(0)} + \Delta t C[u^{(0)}], \\ \bar{u}_i^{(2)} &= \bar{u}_i^{(1)} + \frac{\Delta t}{2} (C[u^{(1)}] - C[u^{(0)}]), \end{aligned} \quad (4.28)$$

where  $C[u] = -\frac{1}{\Delta x_i} (F_{i+1/2}^* - F_{i-1/2}^*)$ , and  $\bar{u}_i^{(0)} := u_i^n$  is the average solution on the  $i^{\text{th}}$  cell at  $t = t^n$ .

#### 4.1.4 Formulation for the 2-D semi-discrete form

Let us assume to have a solution at time  $t^n$ . It is possible to reconstruct the 2-D solution as a non-oscillatory piece-wise linear polynomial, as:

$$u(x, y, t) \approx \sum_{j,k} \left[ u_{j,k}(t) + (u_x)_{j,k} (x - x_j) + (u_y)_{j,k} (y - y_k) \right] \mathbf{1}_{[x_{j-1/2}, x_{j+1/2}] \times [y_{k-1/2}, y_{k+1/2}]}, \quad (4.29)$$

where:  $(u_x)_{j,k}$  denotes a limited approximation of the exact derivative in the  $x$  direction, as:  $(u_x)_{j,k} := \phi_{j,k}^x \left[ \frac{\partial}{\partial x} u_{j,k}(t) \right]$ , wherein  $\phi_{j,k}^x$  shows a limiter at  $(x_j, y_k)$  working at the  $x$  direction; in the same way, for the  $y$  direction, we have:  $(u_y)_{j,k} := \phi_{j,k}^y \left[ \frac{\partial}{\partial y} u_{j,k}(t) \right]$ ;  $u_{j,k}(t) := u(x_j, y_k, t)$ . The piece-wise reconstruction (4.29) with the MINMOD limiters guarantees the non-oscillatory feature in the frame work of the (local) maximum principle, see [165], Theorem 1.

For spatial discretization, since grid points are not in a tensor product style, the simple concept of the *dimension by dimension* discretization is followed; in this regard, the semi-discrete formulation for the scalar 2-D hyperbolic PDE,  $\mathbf{u}_t + \mathbf{F}_x + \mathbf{G}_y = \mathbf{0}$ ,

can be written in terms of average solutions at point  $(x_j, y_k)$ , as:

$$\frac{\Delta \bar{u}_{j,k}}{\Delta t} + \frac{F_{j+1/2,k}^* - F_{j-1/2,k}^*}{\Delta x_j} + \frac{G_{j,k+1/2}^* - G_{j,k-1/2}^*}{\Delta y_k} = 0,$$

where:  $\bar{u}_{j,k} := \bar{u}(x_j, y_k, t)$  is an average solution; reconstructed fluxes  $F_{j+1/2,k}^*$  and  $G_{j,k+1/2}^*$  are:

$$\begin{aligned} F_{j+1/2,k}^* &:= \frac{F_{j+1/2,k}^R + F_{j+1/2,k}^L}{2} - \frac{a_{j+1/2,k}^x}{2} (u_{j+1/2,k}^R - u_{j+1/2,k}^L), \\ G_{j,k+1/2}^* &:= \frac{G_{j,k+1/2}^R + G_{j,k+1/2}^L}{2} - \frac{a_{j,k+1/2}^y}{2} (u_{j,k+1/2}^R - u_{j,k+1/2}^L), \end{aligned}$$

where:  $x_{j-1/2} := x_j - (1 - p_j^x)\Delta x_j$  and  $x_{j+1/2} := x_j + p_j^x\Delta x_j$ ;  $y_{k-1/2} := y_k - (1 - p_k^y)\Delta y_k$  and  $y_{k+1/2} := y_k + p_k^y\Delta y_k$ ;  $u_{j+1/2,k}^R := u_{j+1,k} - (u_x)_{j+1,k}^n (x_{j+1} - x_{j+1/2})$  and  $u_{j+1/2,k}^L := u_{j,k} + (u_x)_{j,k}^n (x_{j+1/2} - x_j)$ ;  $u_{j,k+1/2}^R := u_{j,k+1} - (u_y)_{j,k+1}^n (y_{k+1} - y_{k+1/2})$  and  $u_{j,k+1/2}^L := u_{j,k} + (u_y)_{j,k}^n (y_{k+1/2} - y_k)$ ; the local speeds,  $a_{j+1/2,k}^x$  and  $a_{j,k+1/2}^y$ , can be evaluated as:  $a_{j+1/2,k}^x := \max \rho [\partial F(u_{j+1/2,k}^R) / \partial u, \partial F(u_{j+1/2,k}^L) / \partial u]^4$  and  $a_{j,k+1/2}^y := \max \rho [\partial G(u_{j,k+1/2}^R) / \partial u, \partial G(u_{j,k+1/2}^L) / \partial u]$ .

**The maximum principle for 2-D central schemes** It is well known that exact entropy (physical) solution of a 2-D scalar conservation law satisfies a *maximum principle* [165]; this principle is:

$$\max_{j,k} \{ \bar{u}_{j,k}(t^{n+1}) \} \leq \max_{j,k} \{ \bar{u}_{j,k}(t^n) \}.$$

This principle guarantees that a monotone solution remains monotone through time.

Based on the work by Kurganov and Tadmor, [22], theorem 5.1, with a proper *CFL* condition and MINMOD limiters in the  $x$  and  $y$  directions, the resulting fully-discrete central scheme satisfies the maximum principle. So, a monotone reconstructed solution (based on Eq. (4.29)), will remain monotone through time.

---

<sup>4</sup>Operator  $\rho(\mathbf{A})$  (where  $\mathbf{A} := \partial F(\mathbf{u}) / \partial \mathbf{u}$ ) shows the spectral radius of the matrix  $\mathbf{A}$ , or:  $\rho(\mathbf{A}) := \max_i |\lambda_i(\mathbf{A})|$ , wherein  $\{\lambda_i\}$  are eigenvalues of  $\mathbf{A}$ .



### 4.1.5 Systems with source and diffusion terms

**1-D systems** If diffusion ( $Q(u, u_x)$ ) and source ( $S(u)$ ) terms are included, then the 1-D convection equation can be rewritten as a convection-diffusion system:

$$u_t + F(u)_x = Q(u, u_x)_x + S(u) \quad (4.30)$$

In this case, it is shown that the semi-discretized form of the central scheme is [22]:

$$\frac{du_i}{dt} + \frac{1}{\Delta x_i} (F_{i+1/2}^* - F_{i-1/2}^*) = \frac{1}{\Delta x_i} (P_{i+1/2} - P_{i-1/2}) + S_i, \quad (4.31)$$

where  $S_i = S(u_i)$ ;  $F_{i\pm 1/2}^*$  have the previous definitions and  $P_{i\pm 1/2}$  are:

$$\begin{aligned} P_{i-1/2} &= \frac{1}{2} \left\{ Q(u_{i-1}, \frac{u_i - u_{i-1}}{\Delta x_{i-1}}) + Q(u_i, \frac{u_i - u_{i-1}}{\Delta x_{i-1}}) \right\}, \\ P_{i+1/2} &= \frac{1}{2} \left\{ Q(u_i, \frac{u_{i+1} - u_i}{\Delta x_i}) + Q(u_{i+1}, \frac{u_{i+1} - u_i}{\Delta x_i}) \right\}. \end{aligned} \quad (4.32)$$

The semi-discrete form (4.31) can be evaluated in time by the TVD second-order RK solver as [22]:

$$\begin{aligned} \bar{u}_i^{(1)} &= \bar{u}_i^{(0)} + \Delta t C[u^{(0)}], \\ \bar{u}_i^{(2)} &= \bar{u}_i^{(1)} + \frac{\Delta t}{2} (C[u^{(1)}] - C[u^{(0)}]), \end{aligned} \quad (4.33)$$

where  $C[u] = -\frac{1}{\Delta x_i} (F_{i+1/2}^* - F_{i-1/2}^*) + \frac{1}{\Delta x_i} (P_{i+1/2} - P_{i-1/2}) + S_i$ , and  $\bar{u}_i^{(0)} := u_i^n$  (denotes average solutions on the cell  $i$ ).

**2-D systems** Consider a 2-D convection-diffusion system having diffusion terms and a source term as follows:

$$u_t + F(u)_x + G(u)_y = Q^x(u, u_x, u_y)_x + Q^y(u, u_x, u_y)_y + S(u), \quad (4.34)$$

where  $u = u(x, y, t)$ ;  $F(u)$  and  $G(u)$  denote fluxes in  $x$  and  $y$  directions, respectively;  $Q^x(u, u_x, u_y)$  and  $Q^y(u, u_x, u_y)$  are diffusion terms in  $x$  and  $y$  directions, respectively;  $S(u)$  represents the source term.

Let us assume a cell-centered 2-D rectangle cell  $\Omega \in [x_{i-1/2}, x_{i+1/2}] \times [y_{k-1/2}, y_{k+1/2}]$  with cell widths  $\Delta x$  and  $\Delta y$ . By following the REP stages and considering the bi-linear

variation for the state variable through the rectangle cell (Eq. (4.29)), it is easy to show that the semi-discrete form of the KT formulation is [22]:

$$\begin{aligned} \frac{du_{i,k}}{dt} + \frac{1}{\Delta x_i} (F_{i+1/2,k}^* - F_{i-1/2,k}^*) + \frac{1}{\Delta y_k} (G_{i,k+1/2}^* - G_{i,k-1/2}^*) = \\ \frac{1}{\Delta x_i} (P_{i+1/2,k}^x - P_{i-1/2,k}^x) + \frac{1}{\Delta y_k} (P_{i,k+1/2}^y - P_{i,k-1/2}^y) + S_{i,k}, \end{aligned} \quad (4.35)$$

where:

$$\begin{aligned} F_{i+1/2,k}^* &= \{[F(u_{i+1/2,k}^R) + F(u_{i+1/2,k}^L)] - a_{i+1/2,k}^x [u_{i+1/2,k}^R - u_{i+1/2,k}^L]\}/2, \\ G_{i,k+1/2}^* &= \{[G(u_{i,k+1/2}^R) + G(u_{i,k+1/2}^L)] - a_{i,k+1/2}^y [u_{i,k+1/2}^R - u_{i,k+1/2}^L]\}/2, \\ P_{i+1/2,k}^x &= \{Q^x(u_{i,k}, \frac{u_{i+1,k} - u_{i,k}}{\Delta x_i}, (u_y)_{i,k}) + Q^x(u_{i+1,k}, \frac{u_{i+1,k} - u_{i,k}}{\Delta x_i}, (u_y)_{i+1,k})\}/2, \\ P_{i,k+1/2}^y &= \{Q^y(u_{i,k}, (u_x)_{i,k}, \frac{u_{i,k+1} - u_{i,k}}{\Delta y_k}) + Q^y(u_{i,k+1}, (u_x)_{i,k+1}, \frac{u_{i,k+1} - u_{i,k}}{\Delta y_k})\}/2, \\ S_{i,k} &= S(u(x_i, y_k)). \end{aligned} \quad (4.36)$$

## 4.2 Nonlinear stability: the TVD condition

In this subsection, a nonlinear stability condition for nonlinear conservation laws will be reviewed, which originally is developed for uniform grids. The criterion is the *total variation diminishing* (TVD) condition to prevent developing of spurious oscillations in numerical solutions of nonlinear hyperbolic systems.

The TVD condition leads to a set of *global* requirements, known as the *positivity* conditions. Based on these general conditions, some *local* conditions can be derived. These conditions are important for designing slope/flux limiters. Finally, performance of flux/slope limiters will be studied for non-uniform non-centered grids (cells).

Before providing stability conditions, definitions of *total variation* (TV) and TVD are presented.

**The TV definition and the TVD condition** Total Variation (TV) of a continuous solution  $u(x)$  is defined as:

$$TV(u) := \int |u'(x)| dx, \quad (4.37)$$

where  $u'(x) := du(x)/dx$ . For discrete solutions  $u_i := u(x_i)$  it becomes:

$$TV(u) := \sum_i |u_{i+1} - u_i|, \quad (4.38)$$

A solution is TVD if

$$\sum_i |u_{i+1}^{n+1} - u_i^{n+1}| \leq \sum_i |u_{i+1}^n - u_i^n|, \quad (4.39)$$

or  $TV(u^{n+1}) \leq TV(u^n)$ .

This TVD condition guarantees that a monotone solution, will remain monotone through time and this prevent of forming spurious oscillations (this will be shown).

### Global TVD conditions

To control the development of spurious oscillations in numerical simulation of hyperbolic systems, it is necessary to show that a monotone (non-increasing or non-decreasing) profile remains monotone during time evolution. To achieve high-order of accuracy, a relaxed monotonicity condition — so called Total Variation Diminishing (TVD) is sought [68, 166].

### TVD conditions based on the fully-discrete form of nonlinear conservation laws

Based on the *three-point-stencil* formulations, the fully-discrete form of nonlinear conservation laws can be written as:

$$u_i^{n+1} = u_i^n + C_{i+1/2}^+(u_{i+1}^n - u_i^n) - C_{i-1/2}^-(u_i^n - u_{i-1}^n), \quad (4.40)$$

where:  $u_i^n := u(x_i, t^n)$ ,  $C_{i+1/2}^+$  and  $C_{i-1/2}^-$  are some coefficients. Based on this general form, the TVD condition which is necessary for numerical stability will be derived.

**Theorem 1.** *In order to the fully-discrete form (4.40) is TVD, the following conditions are sufficient:*

$$\begin{aligned} C_{i+1/2}^+ + C_{i+1/2}^- &\leq 1, \\ C_{i+1/2}^- &\geq 0, \\ C_{i+1/2}^+ &\geq 0. \end{aligned} \quad (4.41)$$

**TVD conditions based on the semi-discrete form of nonlinear conservation laws**

Again for three point stencil, the semi-discrete-form can be presented as:

$$\frac{d}{dt}u_i(t) = C_{i+1/2}^+(u_{i+1}^n - u_i^n) - C_{i-1/2}^-(u_i^n - u_{i-1}^n). \quad (4.42)$$

**Theorem 2.** *In order to the semi-discrete form (4.42) is TVD, the following conditions are sufficient:*

$$\begin{aligned} C_{i+1/2}^+ + C_{i+1/2}^- &\leq 1, \\ C_{i+1/2}^- &\geq 0, \\ C_{i+1/2}^+ &\geq 0. \end{aligned} \quad (4.43)$$

The proofs of Theorems 1 and 2 are presented in Appendix C.

Theorems 1 and 2 show that for both fully-discrete and semi-discrete forms the TVD conditions are the same. A general (global) TVD condition derived in [166] and mentioned above is expressed in terms of the above-mentioned global positivity condition [167–169].

**Designing flux/slope limiters, based on the global positivity conditions**

**Flux limiters** Based on the global TVD conditions, a proper flux limiter can be designed. By a flux limiter, a higher-order numerical flux  $F_i^H$  can be divided into two parts:

1. A low-order flux,  $F_i^L$ ; using of this flux leads to non-oscillatory results in expense of low-order accuracy,
2. The residual flux, as:  $F_i^H - F_i^L$ . The higher-order flux can be recovered as:  $F_i^H = F_i^L + \phi_i (F_i^H - F_i^L)$ , where  $\phi_i := \phi(x_i)$  is a flux limiter with values in the range:  $0 \leq \phi_i \leq 1$ . For smooth solutions  $\phi_i$  approaches to the unit value and therefore a higher-order solution is obtained, while around discontinuities it approaches zero, to prevent spurious oscillations.

It is possible to show that a flux limiter  $\phi_i$  satisfying the global TVD conditions yields to the following relationships:

$$\begin{aligned} 0 \leq \frac{\phi_i(r_i)}{r_i} &\leq 2, \\ \phi(r_i) &\leq 2, \end{aligned} \tag{4.44}$$

where  $r_i$  is a smoothness detector (monitor) defined as:  $r_i := \frac{u_i - u_{i-1}}{u_{i+1} - u_i}$ . Also for  $r_i \leq 0$ , we assume that  $\phi(r_i) = 0$  to guarantee that solutions remain TVD; for details of this requirement and deriving of Eq. (4.44), please see Appendix C.

**Slope limiters** The relationship between a slope limiter and a flux limiter is the next question. It is possible to show that performance of a flux limiter is the same as a slope limiter based on the above-mentioned relationships (see Appendix C), so a linear reconstruction can be written as:

$$u(x, t^n) = u_i^n + (u_x)_i(x - x_i), \tag{4.45}$$

where  $(u_x)_i$  is a limited slope:  $(u_x)_i = \phi_i \frac{u_{i+1} - u_i}{x_{i+1} - x_i}$ .

Finally it should be mentioned that to have solutions of second-order accuracy, it is necessary that linear solutions are not limited, and this means:  $\phi(r = 1) = 1$ .

### 4.2.1 Local TVD conditions

*Local* TVD conditions were studied in [170] where it was shown that for the TVD property to hold, it suffice to verify the positivity condition Eq. (4.41) in the extreme cells and their surrounding grid points. In the present context of Eq. (4.25) or Eq. (4.31), these *local* TVD conditions amount to:

1. In extreme points, for both uniform and non-uniform cells, the numerical fluxes  $F_{i\pm 1/2}^*$  satisfy (see [170], Lemma 2.1) (also see Appendix C):

- a) at maximum values  $u_i$ :  $F_{i+1/2}^* \geq F_{i-1/2}^*$ ,

- b) at minimum values  $u_i$ :  $F_{i+1/2}^* \leq F_{i-1/2}^*$ ,

2. Osher [171] showed that in order to satisfy the positivity conditions, and therefore the TVD property, we should have:

$$0 \leq \frac{\Delta x}{\Delta_- u_i} (u_x)_i \leq 1 \quad \text{and} \quad 0 \leq \frac{\Delta x}{\Delta_+ u_i} (u_x)_i \leq 1,$$

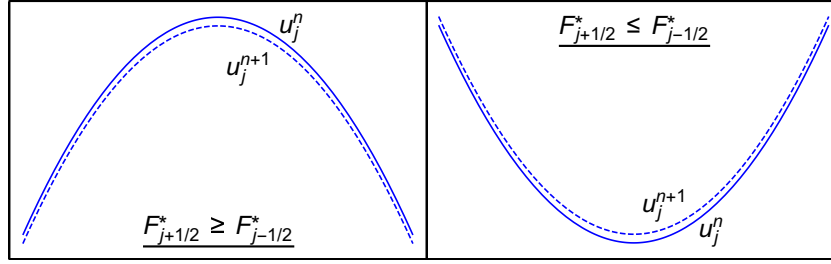


Figure 4.2: The first local TVD condition.

3. For neighbor points of an extreme  $u_i$  (with estimated derivative  $(u,x)_i = 0$ ; this is based on the condition 2), for uniform grids with width  $\Delta x$ , the neighbor derivatives meet conditions (see [170], example 2.4):

$$\frac{1}{2} \left| \frac{\Delta x}{\Delta_- u_i} (u,x)_{i-1} \right| \leq 1 \quad \text{and} \quad \frac{1}{2} \left| \frac{\Delta x}{\Delta_+ u_i} (u,x)_{i+1} \right| \leq 1.$$

In these conditions, we have:  $u_i := u(x_i)$ ;  $\Delta x_i := x_{i+1/2} - x_{i-1/2}$  denotes  $i^{\text{th}}$  cell width; in case of uniform grids:  $\Delta x = \Delta x_i$ ;  $F_{i\pm 1/2}^*$  represent reconstructed fluxes at right and left cell boundaries  $x_{i\pm 1/2}$ ;  $\Delta_{\pm} u_i := \pm(u_{i\pm 1} - u_i)$ ;  $(u,x)_{i\pm 1}$  are the estimated (limited) first derivatives at points  $x_{i\pm 1}$ .

The conditions 1-3 have physical meaning:

1. The first condition states that:
  - a) An extreme solution with maximum value should decrease in time (i.e.,  $u_i^{n+1} \leq u_i^n$ ), while an extreme solution with minimum value should increase through time (i.e.,  $u_i^{n+1} \geq u_i^n$ ); these conditions are presented in Figure 4.2
  - b) Also, a *monotone solution remains monotone during time evolution*, for this see Appendix C, the proof of Lemma 1.
2. From the second condition, it is clear that at extreme points we must have:  $(u,x)_i = 0$ . For instance, let us assume  $u_i$  (at  $x_i$ ) is a maximum point, where  $\Delta_+ u_i < 0$  and  $\Delta_- u_i > 0$ ; from the second condition we have:

$$\begin{aligned} 0 &\geq (u,x)_i \geq \frac{\Delta_+ u_i}{\Delta x} = -C_1, \\ 0 &\leq (u,x)_i \leq \frac{\Delta_- u_i}{\Delta x} = C_2, \end{aligned}$$

where  $C_1$  and  $C_2$  are positive constants. So it is clear these inequalities are satisfied, if and only if:  $(u,x)_i = 0$ . These conditions are presented in Figure 4.3

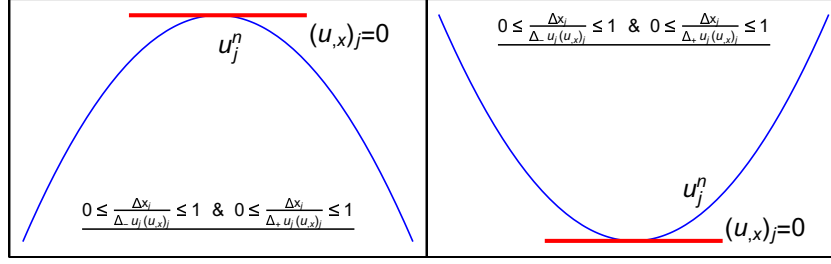


Figure 4.3: The second local TVD condition.

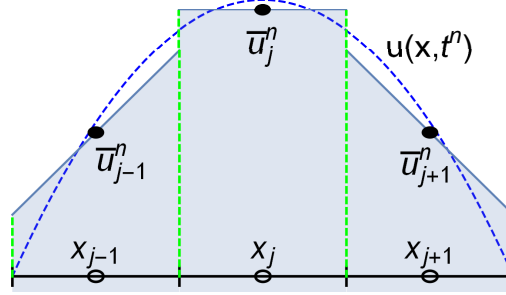


Figure 4.4: The third local TVD condition.

- Due to the third condition, the reconstructed values in cell edges should satisfy the monotonicity feature: magnitude of the reconstructed values should be less than immediately neighbor cell center values. This means, for example, if  $u_{i-1}^n \leq u_i^n \leq u_{i+1}^n$ , then  $u_{i-1}^n \leq u_{i-1/2}^n \leq u_i^n \leq u_{i+1/2}^n \leq u_{i+1}^n$ . This condition is presented in Figure 4.4

The third local TVD condition is for uniform cells. For providing corresponding conditions over non-uniform cells, let us assume the linear advection equation  $u_t + u_x = 0$ . For a E-scheme the following conditions should be satisfied (Appendix C):

$$\begin{aligned} \text{Sign}(u_{i+1} - u_i) &= \text{Sign}(f(u) - F_{i+1/2}^*), \\ \text{Sign}(u_i - u_{i-1}) &= \text{Sign}(f(u) - F_{i-1/2}^*), \end{aligned} \quad (4.46)$$

where  $f$  denotes the flux function and  $F^*$  indicates the numerical flux. Let us assume the numerical flux of the KT scheme as:

$$F_{i+1/2}^* := \frac{1}{2} \left[ (f_{i+1/2}^r + f_{i+1/2}^l) - a_{i+1/2} (u_{i+1/2}^r - u_{i+1/2}^l) \right], \quad (4.47)$$

Also let us assume that  $u_i$  is an extremum point where due to the TVD condition  $(u_x)_i = 0$ . Then the reconstructed values  $u_{i+1/2}^r$  and  $u_{i+1/2}^l$  are:

$$\begin{aligned} u_{i+1/2}^r &= u_{i+1} - (u_x)_{i+1} \frac{\Delta x_{i+1}}{2}, \\ u_{i+1/2}^l &= u_i + (u_x)_i \frac{\Delta x_i}{2} = u_i. \end{aligned} \quad (4.48)$$

Also the reconstructed values  $u_{i-1/2}^r$  and  $u_{i-1/2}^l$  read:

$$\begin{aligned} u_{i-1/2}^r &= u_i - (u_x)_i \frac{\Delta x_i}{2} = u_i, \\ u_{i-1/2}^l &= u_{i-1} + (u_x)_{i-1} \frac{\Delta x_{i-1}}{2}. \end{aligned} \quad (4.49)$$

Regarding  $f(u) = u$ , the numerical fluxes read:

$$\begin{aligned} F_{i+1/2}^* &= \frac{1}{2} \left[ \left( \left( u_{i+1} - (u_x)_{i+1} \frac{\Delta x_{i+1}}{2} \right) + (u_i) \right) - 1 \times \left( \left( u_{i+1} - (u_x)_{i+1} \frac{\Delta x_{i+1}}{2} \right) - (u_i) \right) \right] \\ &= u_i, \\ F_{i-1/2}^* &= \frac{1}{2} \left[ \left( (u_i) + \left( u_{i-1} + (u_x)_{i-1} \frac{\Delta x_{i-1}}{2} \right) \right) - 1 \times \left( (u_i) - \left( u_{i-1} + (u_x)_{i-1} \frac{\Delta x_{i-1}}{2} \right) \right) \right] \\ &= u_{i-1} + (u_x)_{i-1} \frac{\Delta x_{i-1}}{2}. \end{aligned} \quad (4.50)$$

Then based on Eq. (4.46) for non-uniform cells (due to the E-condition (Appendix C)) at extreme point  $x_i$ , we have:

$$\begin{aligned} \text{Sign}(u_{i+1} - u_i) &= \text{Sign} \left( \left\{ \left( u_{i+1} - \frac{\Delta x_{i+1}}{2} (u_x)_{i+1} \right) - (u_i) \right\} \right), \\ \text{Sign}(u_i - u_{i-1}) &= \text{Sign} \left( \left\{ (u_i) - \left( u_{i-1} + \frac{\Delta x_{i-1}}{2} (u_x)_{i-1} \right) \right\} \right). \end{aligned} \quad (4.51)$$

Following Appendix C, from these relationships, TVD preserving derivatives should satisfy:



$$\frac{1}{2} \left| \frac{\Delta x_{i-1}}{\Delta_- u_i} (u_{,x})_{i-1} \right| \leq 1 \quad \text{and} \quad \frac{1}{2} \left| \frac{\Delta x_{i+1}}{\Delta_+ u_i} (u_{,x})_{i+1} \right| \leq 1. \quad (4.52)$$

These two relationships confirm also as before necessity of the monotone reconstruction in cell edges for non-uniform cells.

For deriving details of Eqs. (4.51) and (4.52), please see Appendix C.

## 4.2.2 Irregularity effects on slope limiters

The central/central-upwind schemes (e.g., the KT method) is originally developed for uniform cells. Most of them satisfy the TVD and monotonicity preserving conditions. On irregular cells, these methods do not completely meet these conditions and therefore do not remain necessarily stable. In the following, effects of cell irregularities will be studied by considering the monotonicity preserving necessity and the local TVD conditions. It will be shown that only the grid modification (for gradual variation of grids) is not generally enough to guarantee TVD results: this *only-grid modification* is a common approach in most engineering works.

### Slope limiters on uniform grids

Most of slope limiters meet some necessary conditions, as: 1) the TVD; 2) preserving linear approximation; and 3) symmetric feature. These properties are explained in brief, as follows.

**The TVD property** Considering physical meanings of the local TVD conditions 1-3, the upper limit of a TVD limiter for  $i^{\text{th}}$  cell is (for more details, one can see [172]):

$$u'_i = \min \left\{ \frac{2(\Delta_+ u_i)}{\Delta x}, \frac{2(\Delta_- u_i)}{\Delta x} \right\}, \quad (4.53)$$

where:  $\Delta x = \Delta x_i$ ,  $u_i$  denotes cell center solution on the  $i^{\text{th}}$  cell and the operator  $u'_i$  is the upper limit of a discrete approximation of  $grad(u_i)$  in a way that solutions remain TVD. This can easily be obtained, since at extrema points  $u'_i = 0$ , also the first condition show that a monotone solution remain monotone during time. Now, let us assume that  $x_i$  is located between two extrema; then based on the third condition, Eq. (4.52), we can rewrite Eq. (4.52) for  $(u_{,x})_i$  as:

$$\frac{1}{2} \frac{\Delta x_i}{\Delta_- u_{i+1}} (u_x)_i \leq 1,$$

$$\frac{1}{2} \frac{\Delta x_i}{\Delta_+ u_{i-1}} (u_x)_i \leq 1,$$

but, it is clear that:  $\Delta_- u_{i+1} = \Delta_+ u_i$  and  $\Delta_+ u_{i-1} = \Delta_- u_i$ . So, the above inequalities become:

$$\frac{1}{2} \frac{\Delta x_i}{\Delta_+ u_i} (u_x)_i \leq 1,$$

$$\frac{1}{2} \frac{\Delta x_i}{\Delta_- u_i} (u_x)_i \leq 1,$$

and so Eq. (4.53) can be obtained.

To have TVD solutions, estimated derivative  $(u_x)_i := du(x_i, t)/dx$  should be limited by a slope limiter  $\phi_i := \phi(R_i)$  as  $\phi_i \times (u_x)_i$ . The function  $\phi_i$  is a slope limiter at  $x_i$ , where  $0 \leq \phi_i \leq 1$ . Parameter  $R_i$  measures smoothness by the relative of successive gradients around point  $x_i$ ; for uniform grids its definition is:  $R_i = \frac{\Delta_+ u_i / \Delta_+ x_i}{\Delta_- u_i / \Delta_- x_i} = \frac{\Delta_+ u_i}{\Delta_- u_i}$  (it should be mentioned that  $r_i = 1/R_i$ ).

For all TVD slope limiters, it is necessary that:  $\phi_i(u_x)_i \leq u'_i$  (this upper TVD limiter will be derived on non-uniform cells).

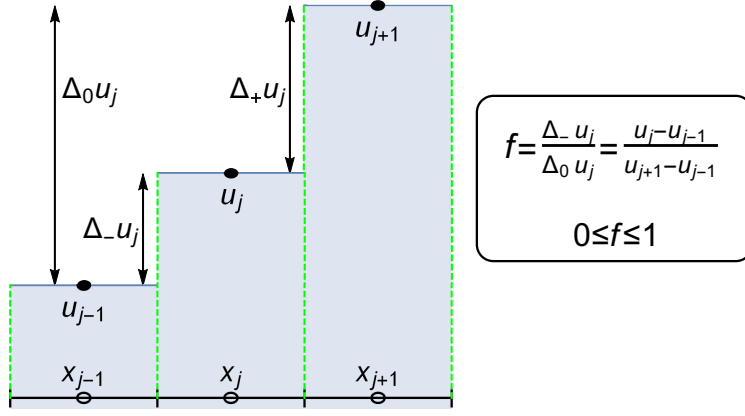
**The linear approximation preserving** For a linear function with slope  $s$  on a uniform grid, it is clear:  $R_i = \Delta_+ u_i / \Delta_- u_i = \{s \cdot (\Delta_+ x_i)\} / \{s \cdot (\Delta_- x_i)\} = 1$ . In this regard, the linear preserving condition is:  $\phi(1) = 1$ . This constraint can also be obtained by considering the MINMOD limiter definition: the linear preserving feature is conserved if the forward and backward derivatives are the same, i.e.,  $R_i = 1$ .

**The symmetric feature** The condition is:  $\phi(R_i) = \phi(1/R_i)$ . This condition assure that slope limiter effects are the same for forward and backward propagating solutions. To more clarify this feature, a new parameter  $f$  can be defined as:

$$f = \frac{\Delta_- u_i}{\Delta_0 u_i}, \tag{4.54}$$

where  $\Delta_0 u_i := u_{i+1} - u_{i-1}$ . This new parameter is illustrated in Figure 4.5.

For monotone solutions,  $u_i$  is always between  $u_{i-1}$  and  $u_{i+1}$ , so  $f$  belongs always to the range  $[0, 1]$ . The parameters  $R$  and  $f$  have then a relationship with each other, as:


 Figure 4.5: The definition of the parameter  $f$ .

$$\begin{aligned}
 R &= \frac{u_{i+1} - u_i}{u_i - u_{i-1}} = \frac{\frac{u_{i+1} - u_i}{u_{i+1} - u_{i-1}}}{\frac{u_i - u_{i-1}}{u_{i+1} - u_{i-1}}} = \frac{u_{i+1} - u_i + (u_{i-1} - u_{i-1})}{u_{i+1} - u_{i-1}} \\
 &= \frac{\frac{u_{i+1} - u_{i-1}}{u_{i+1} - u_{i-1}} - \frac{u_i - u_{i-1}}{u_{i+1} - u_{i-1}}}{\frac{u_i - u_{i-1}}{u_{i+1} - u_{i-1}}} \\
 &= \frac{1 - \frac{\Delta_- u_i}{\Delta_0 u_i}}{\frac{\Delta_- u_i}{\Delta_0 u_i}} \\
 &= \frac{1 - f}{f}.
 \end{aligned} \tag{4.55}$$

The symmetric property is now clear, since:

$$\phi(R) = \phi((1 - f)/f) \equiv \phi(f), \tag{4.56}$$

and

$$\phi(1/R) = \phi\left(\frac{f}{1 - f}\right) = \phi\left(-\frac{-f}{1 - f}\right) = \phi\left(-\frac{-f + 1 - 1}{1 - f}\right) = \phi\left(\frac{1 - (1 - f)}{1 - f}\right) \equiv \phi(1 - f). \tag{4.57}$$

And since  $\phi(1 - f) \equiv \phi(f)$ , then:

$$\phi(f) = \phi(1 - f) \quad \text{for } 0 \leq f \leq 1. \quad (4.58)$$

### 4.2.3 Slope limiters on cell-centered non-uniform grids

A grid with non-uniform cells is considered. The widths of successive cells can be related to each other by coefficients:  $a = \Delta x_{i-1}/\Delta x_i$  and  $b = \Delta x_{i+1}/\Delta x_i$ . A possible jump in a solution between points  $x_{j+1}$  and  $x_{j-1}$  is denoted by  $\Delta_0 u_i$ , where  $\Delta_0 u_i := u_{i+1} - u_{i-1}$ . The forward ( $D_+ u_i$ ), backward ( $D_- u_i$ ) and centered ( $D_0 u_i$ ) differences can then be written as:

$$\begin{aligned} D_{\pm} u_i &:= \frac{\Delta_{\pm} u_i}{\Delta_{\pm} x_i}, \\ D_0 u_i &:= \frac{\Delta_0 u_i}{\Delta_- x_i + \Delta_+ x_i}. \end{aligned} \quad (4.59)$$

Considering the definition of  $f$ , these derivatives can be rewritten as:

$$\begin{aligned} D_+ u_i &= \frac{2(1-f)\Delta_0 u_i}{(1+b)\Delta x_i}, \\ D_- u_i &= \frac{2f\Delta_0 u_i}{(1+a)\Delta x_i}, \\ D_0 u_i &= \frac{2\Delta_0 u_i}{(2+a+b)\Delta x_i}. \end{aligned} \quad (4.60)$$

**The TVD property** This can be obtained by the local TVD conditions 1-3 (see [172]). It has the same definition as the uniform case (Eq. (4.53)):

$$u'_i = \min \left\{ \frac{2\Delta_+ u_i}{\Delta x_i}, \frac{2\Delta_- u_i}{\Delta x_i} \right\}. \quad (4.61)$$

On the other hand, we have  $\Delta_- u_i = f\Delta_0 u_i$  and  $\Delta_+ u_i = (1-f)\Delta_0 u_i$ , so  $u'_i$  can be written as:

$$u'_i = \frac{2\Delta_0 u_i}{\Delta x_i} \min \{f, 1-f\}. \quad (4.62)$$

This TVD-limit is illustrated in Figure 4.6 for cell-centered nonuniform cells; figure

(a) is for the uniform case where  $a = b = 1$ .

**The linear approximation preserving** For non-uniform grids, the linear preserving condition for a slope limiter is [172]:

$$\phi_i D_i (f = f_p) = \frac{2\Delta_0 u_i}{(a + b + 2)\Delta x_i}, \quad (4.63)$$

where  $f_p = \frac{a+1}{a+b+2}$  and  $\phi_i D_i (f_p)$  denotes the limited value of the first derivative at  $f = f_p$ . The condition in Eq. (4.63) can be achieved easily: the linear preserving feature is conserved if the backward and forward derivatives are equal to each other:  $D_+ u_i = D_- u_i$ ; from this equation, the value of  $f = f_p$  can be calculated, as:

$$\begin{aligned} \frac{2(1-f)\Delta_0 u_i}{(1+b)\Delta x_i} &= \frac{2f\Delta_0 u_i}{(1+a)\Delta x_i}, \\ \Rightarrow \frac{1-f}{1+b} &= \frac{f}{1+a}, \\ \Rightarrow (1+b)f &= (1+a) - (1+a)f, \\ \Rightarrow f = f_p &= \frac{1+a}{2+a+b}. \end{aligned} \quad (4.64)$$

By inserting this value in the definition of either  $D_+ u_i$  or  $D_- u_i$ , Eq. (4.63) can be obtained.

Let us assume the definition of the MINMOD limiter as:

$$(\phi_i D_i)_{MM} := \text{MINMOD}(D_- u_i, D_+ u_i) = \text{MINMOD}\left(\frac{2f\Delta_0 u_i}{(1+a)\Delta x_i}, \frac{2(1-f)\Delta_0 u_i}{(1+b)\Delta x_i}\right),$$

Then, the value of  $(\phi_i D_i)_{MM}(f = f_p)$  is the value of the MINMOD limiter at the  $f_p$ .

**The symmetric condition** This feature is not fulfilled, in general.

#### 4.2.4 The performance of the generalized MINMOD limiter on non-uniform grids

The generalized MINMOD (GMINMOD) limiter, can be rewritten as:

$$(\phi_i D_i)_{\text{GMINMOD}} = \text{GMINMOD} \{ \theta D_- u_i, D_0 u_i, \theta D_+ u_i \}, \quad 1 \leq \theta \leq 2, \quad (4.65)$$

or equivalently:

$$(\phi_i D_i)_{\text{GMINMOD}} = \text{GMINMOD} \left\{ \theta \frac{2f \Delta_0 u_i}{(1+a) \Delta x_i}, \frac{2 \Delta_0 u_i}{(2+a+b) \Delta x_i}, \theta \frac{2(1-f) \Delta_0 u_i}{(1+b) \Delta x_i} \right\}, \quad 1 \leq \theta \leq 2, \quad (4.66)$$

In case of non-uniform grids, different definitions of  $D_0 u_i$  can be considered. One is based on the above-mentioned definition:  $D_0 u_i = \Delta_0 u_i / (\Delta_- x_i + \Delta_+ x_i)$ , and the other one can be obtained by the first order least square based estimated slope <sup>5</sup>, as [172]:

$$D_0 u_i = D_{\text{Lsq.}} u_i := \frac{\Delta_+ x_i^2}{\Delta_- x_i^2 + \Delta_+ x_i^2} D_+ u_i + \frac{\Delta_- x_i^2}{\Delta_- x_i^2 + \Delta_+ x_i^2} D_- u_i. \quad (4.67)$$

On uniform grids, both definitions lead to the first order central difference equation:  $(u_{i+1} - u_{i-1}) / (2 \Delta x_i)$ . The performance of these two GMINMOD limiters will be illustrated on non-uniform grids.

Regarding the definitions of  $f$  and  $\Delta_0 u_i$ , the function  $D_{\text{Lsq.}} u_i$  (Eq. (4.67)) can be rewritten as:

$$D_0 u_i = D_{\text{Lsq.}} u_i = \frac{(1+b)(1-f) + (1+a)}{(1+b)^2 + (1+a)^2} \left\{ \frac{f \cdot \Delta_0 u_i}{\Delta x_i} \right\},$$

The performance of the GMINMOD limiters on uniform and non-uniform grids are shown in Figure 4.6. Figure 4.6(a) is for uniform case and Figures 4.6(b,c) are for non-uniform cases. In all figures, center of each cell,  $x_i$ , is in the middle of  $i^{\text{th}}$  cell,

---

<sup>5</sup>Let  $x_{i-1}$ ,  $x_i$  and  $x_{i+1}$  denote the centers of three adjacent cells on an irregular mesh with numerical solutions  $u_{i-1}$ ,  $u_i$  and  $u_{i+1}$ , respectively. The least square problem is to find the slope  $D_L$  in the cell  $i$  to minimize the residual of [172]:

$$\begin{pmatrix} x_{i+1} - x_i \\ x_{i-1} - x_i \end{pmatrix} \cdot D_{\text{Lsq.}} = \begin{pmatrix} u_{i+1} - u_i \\ u_{i-1} - u_i \end{pmatrix}$$

Forming the normal equations and solving for  $D_{\text{Lsq.}}$  results to the above-mentioned equation.

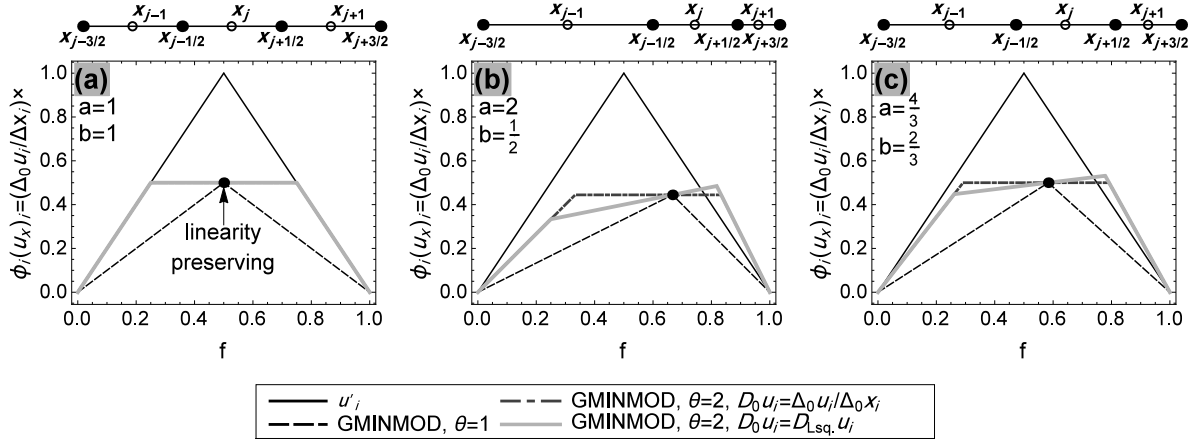


Figure 4.6: The performance of GMINMOD limiters on non-uniform grids. Regarding the top row in each figure, the hollow circles and the solid points represent the center of cells and the cell edges, respectively.

i.e.:  $x_i = (x_{i+1/2} + x_{i-1/2})/2$ . A gradual variation of cell lengths is considered for irregular grids. In Figure 4.6(b) length of cells are:  $\Delta x_{i-1} = 2dx$ ,  $\Delta x_i = dx$ , and  $\Delta x_{i+1} = 0.5dx$ ; and in Figure 4.6(c) the lengths are:  $\Delta x_{i-1} = 2dx$ ,  $\Delta x_i = 1.5dx$ , and  $\Delta x_{i+1} = dx$ . The results confirm that: 1) on gradually varying grids, limiters may not completely remain in the TVD region; 2) the symmetry condition may not satisfy; 3) more gradual variation of grids is, more stability exists; 4) the long-term numerical stability cannot be guaranteed; and 5) the GMINMOD limiter by the direct-definition of central differencing,  $D_0 u_i = (\Delta_0 u_i) / (\Delta_- x_i + \Delta_+ x_i)$ , leads to more symmetric behaviors, so this definition will be considered.

It will be shown that to have a TVD solution (defined on non-uniform grids) without modifying limiter definitions (the GMINMOD limiter, here), the cell middle point,  $x_i$ , should be shifted slightly; and this is only necessary for transmitting cells (a cell between two surrounding uniform cells with different cell lengths). In the following, at first, it will be shown how to choose properly cell centers/edges by the adaptive wavelet transform. It will then mathematically be proved that why such spatial configurations lead to stable and TVD solutions without modifying the MINMOD and GMINMOD limiter definitions.

### Choosing of cell centers and edges by the wavelet-based adaptation algorithm

In this work, at first, cell center positions,  $x_j$ , are evaluated by the adaptive wavelet transform, and then, cell edges are simply assumed to be the middle point of them, as:  $x_{j+1/2} = (x_j + x_{j+1})/2$ .

The interpolating wavelet theory in Chapter 3 uses the pyramid algorithm. In this formulation, distance between detail coefficients in the resolution  $j$  is twice those in the resolution  $j + 1$ . Consider an adapted grid where for every adapted points of resolution level  $j + 1$ , there exist always two surrounding adapted points of resolution level  $j$ . For such adapted grids, inter-distances of successive points increase or decrease gradually by the dyadic pattern. Let us consider an adapted cell-centers as  $\{\dots, 2dx, 2dx, dx, dx, \dots\}$ . This means, for grid points  $\{\dots, x_{j-2}, x_{j-1}, x_j, x_{j+1}, x_{j+2}, \dots\}$ , we have:  $x_{j-1} - x_{j-2} = x_j - x_{j-1} = 2dx$  and  $x_{j+1} - x_j = x_{j+2} - x_{j+1} = dx$ . Such configuration is illustrated in Figure 4.7(a). For these adapted points, it is assumed that there exists always three successive points with equal distances from each other. In this figure, the cell edges are the middle points of two successive cell centers, i.e.:  $x_{j+1/2} = (x_j + x_{j+1})/2$ . By this, the length of created cells are:  $\{\dots, \Delta x_{j-1} = 2dx, \Delta x_j = 1.5dx, \Delta x_{j+1} = dx, \dots\}$ . Note that  $j^{\text{th}}$  cell acts as a transiting cell with a shifted cell center ( $x_j$  is no longer in the middle of cell  $j$ ). In all the surrounding cells, all cell centers remain in the middle of cells. In the next subsection, it will be proved that such cell configuration leads to stable and TVD results. This cell configuration can obtain by the post-processing stage, Chapter 3.

### Stability and TVD conditions on wavelet-based adapted grids

In the proposed wavelet-based grid adaptation, as mentioned, cell center of transmitting cells do not remain in the middle of cells, Figure 4.7(a). In Figure 4.7, three pattern of successive cells are distinguishable: 1)  $\{j - 2, j - 1, j\}$ ; 2)  $\{j - 1, j, j + 1\}$ ; and 3)  $\{j, j + 1, j + 2\}$ . These sets are shown in Figures 4.7 (b) to (d). In the set (2), the transmitting cell  $j$  is in the middle, while in the remaining groups, the transmitting cell  $j$  is the first or the last cell.

In the following, the TVD local conditions are checked and modified for such spatially non-centered cell centers. This will be done for the set (2) (see Figure 4.7(b)), and then it will be checked for the groups (1) and (3). It should be mentioned that for non-centered non-uniform grids, it is possible also to derive the TVD condition based on the conditions 1 to 3, Section 4.2.1 (similar to the procedure leading to Eq. (4.53)). However, in the following, an alternative procedure will be considered, based on solving the advection equation with the upwind formulation.

**i: The TVD condition when the transmitting cell is in the middle of surrounding cells (the set (2))** To provide the TVD condition for the cell set  $\{j - 1, j, j + 1\}$  (see



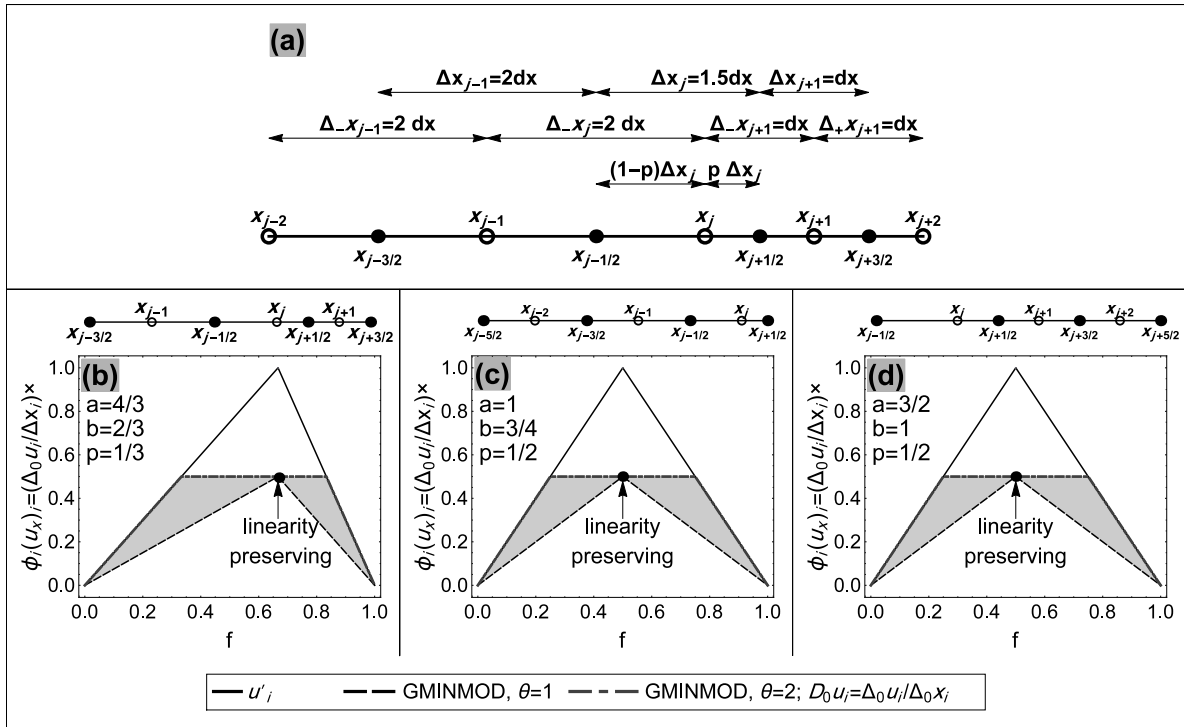


Figure 4.7: Wavelet based adaptive distribution of cell centers  $x_j$ , corresponding edge locations and TVD feature of the GMINMOD limiter on resulted non-uniform grids.

Figure 4.7(b)), a right propagating scalar advection equation is considered as:  $u_t + \bar{a}u_x = 0$  for  $\bar{a} > 0$ . For simplicity, the forward Euler discretization in time will be used. The upwind finite volume method with second order accuracy will be considered for the spatial discretization. The resulted discretized system is:

$$u_j^{n+1} = u_j^n - \frac{\bar{a}\Delta t}{\Delta x_j} (u_{j+\frac{1}{2}}^L - u_{j-\frac{1}{2}}^L). \quad (4.68)$$

The symbol  $L$  represents the upwind flux and therefore the upwind-based reconstructed value of  $u_{j+1/2}^L$  is:

$$u_{j+\frac{1}{2}}^L = u_j + (p\Delta x_j) S_j,$$

where  $S_j := \phi_j D_j$  is a limited slope at point  $x_j$ . Let us assume in the transmitting cell  $j$ , the cell center  $x_j$  locates in a way that  $x_{j+1/2} - x_j = p\Delta x_j$  and  $x_j - x_{j-1/2} = (1-p)\Delta x_j$ . For the wavelet based adapted grids, it is easy to show that  $p \in \{1/3, 1/2, 2/3\}$ .

To show that a method is TVD, it should be confirmed that: 1) a monotone increasing (or decreasing) solution remains monotone increasing (or decreasing) in time (the monotonicity preserving feature: resulted from the positivity condition [170]); 2) if  $u_j^n$  is a local maximum (or minimum), then at the next time step:  $u_j^{n+1} \leq u_j^n$  (or  $u_j^{n+1} \geq u_j^n$ ) (due to the first local TVD condition).

**Controlling of the monotonicity preserving condition** Assume a monotonically increasing solution at time step  $t = t^n$  as:  $u_{j-1}^n \leq u_j^n \leq u_{j+1}^n$ . This feature should satisfy at the next time step, i.e.:  $u_{j-1}^{n+1} \leq u_j^{n+1} \leq u_{j+1}^{n+1}$ .

For the monotone increasing solution  $\{u_j^n\}$ , slopes  $S_{j-1}$  and  $S_j$  are positive, so due to the *upwind* approximation:  $u_{j-1/2}^L = u_{j-1}^n + (\Delta x_{j-1}/2)S_{j-1} \leq u_j^n$  and  $u_{j+1/2}^L = u_j^n + (p\Delta x_j)S_j \geq u_j^n$  (please note that  $I_{j-1}$  and  $I_j$  are the the cell-centered and transmitting cells, respectively). These inequalities are obtained due to the monotone reconstruction constraint: the local TVD condition 3. Hence from Eq. (4.68), we have:

$$u_j^{n+1} \leq u_j^n - \frac{\bar{a}\Delta t}{\Delta x_j} (u_j^n - u_j^n) = u_j^n. \quad (4.69)$$

To complete the proof, we need to estimate a lower bound for  $u_j^{n+1}$ . At first, due

to the local TVD constraint on reconstruction edge values (the condition 3), it is clear that:  $u_j^n - (1-p)\Delta x_j S_j \geq u_{j-1}^n$ , or:  $p\Delta x_j S_j \geq -\left(-u_{j-1}^n + u_j^n - \Delta x_j S_j\right)$ . By considering this relationship, and the conditions  $S_j \geq 0$  and  $S_{j-1} \geq 0$ , from Eq. (4.68), we have:

$$\begin{aligned} u_j^{n+1} &= u_j^n - \lambda_j \left( \left( u_j^n + (p\Delta x_j) S_j \right) - \left( u_{j-1}^n + \left( \frac{\Delta x_{j-1}}{2} \right) S_{j-1} \right) \right) \\ &\geq u_j^n - \lambda_j \left( \left( u_j^n + [u_j^n - u_{j-1}^n - \Delta x_j S_j] \right) - \left( u_{j-1}^n \right) \right) \\ &\geq u_j^n - \lambda_j \left( \left( u_j^n + u_j^n - u_{j-1}^n \right) - \left( u_{j-1}^n \right) \right) = u_j^n (1 - 2\lambda_j) + 2\lambda_j u_{j-1}^n, \end{aligned}$$

so:

$$u_j^{n+1} \geq u_{j-1}^n, \quad (4.70)$$

where  $\lambda_j := \bar{a}\Delta t/\Delta x_j$ . The relationship (4.70) is valid for  $\lambda_j \leq 0.5$ . Eqs. (4.69) and (4.70) lead to the condition  $u_{j-1}^{n+1} \leq u_{j-1}^n \leq u_j^{n+1} \leq u_j^n \leq \dots$ . This means a monotone solution remains monotone through time.

### Controlling of the extreme conditions (the first local TVD condition)

1. **Local maximum:** Let us assume a right propagating wave, that is:  $\bar{a} > 0$  and  $u_{j+1}^n$  is a local maximum on the group cell (2) (Figure 4.7(b)). Due to the extreme condition, we have: 1)  $S_{j+1} = 0$  (the second local TVD condition); 2)  $u_{j+1}^n \geq u_j^n$  and  $u_{j+1}^n \geq u_{j-1}^n$ ; and 3)  $S_j \geq 0$  (due to the monotone reconstruction feature). Rewriting Eq. (4.68) for the maximum point, we have:

$$\begin{aligned} u_{j+1}^{n+1} &= u_{j+1}^n - \frac{\bar{a}\Delta t}{\Delta x_{j+1}} \left( u_{j+\frac{3}{2}}^L - u_{j+\frac{1}{2}}^L \right) \\ &= u_{j+1}^n - \frac{\bar{a}\Delta t}{\Delta x_{j+1}} \left( [u_{j+1}^n + 0.5\Delta x_{i+1} S_{j+1}] - [u_j^n + p\Delta x_j S_j] \right) \\ &= u_{j+1}^n - \frac{\bar{a}\Delta t}{\Delta x_{j+1}} \left( [u_{j+1}^n] - [u_j^n + p\Delta x_j S_j] \right). \end{aligned}$$

To guarantee that  $u_{j+1}^{n+1} \leq u_{j+1}^n$ , the term in the parentheses should be non-negative; hence:

$$0 \leq S_j \leq \frac{u_{j+1}^n - u_j^n}{p\Delta x_j} = \frac{\Delta_+ u_j^n}{p\Delta x_j}. \quad (4.71)$$

2. **Local minimum** Let us assume that the wave propagate to the left direction, or:  $\bar{a} < 0$  and  $u_{j-1}^n$  is a local minimum. This means:  $u_{j-1}^n \leq u_j^n$ ,  $u_{j-1}^n \leq u_{j-2}^n$ ,  $S_{j-1} = 0$  and  $S_j \geq 0$ . By the *downwind* approximation, the spatio-temporal discretization of equation  $u_t + \bar{a}u_x = 0$  reads:

$$u_j^{n+1} = u_j^n - \frac{\bar{a}\Delta t}{\Delta x_j} \left( u_{j+1/2}^R - u_{j-1/2}^R \right).$$

Re-expressing this equation at  $x_{j-1}$ , we get:

$$\begin{aligned} u_{j-1}^{n+1} &= u_{j-1}^n - \frac{\bar{a}\Delta t}{\Delta x_{j-1}} \left( u_{j-\frac{1}{2}}^R - u_{j-\frac{3}{2}}^R \right) \\ &= u_{j-1}^n - \frac{\bar{a}\Delta t}{\Delta x_{j-1}} \left( \left[ u_j^n - (1-p)\Delta x_j S_j \right] - \left[ u_{j-1}^n - 0.5\Delta x_{j-1} S_{j-1} \right] \right) \\ &= u_{j-1}^n - \frac{\bar{a}\Delta t}{\Delta x_{j-1}} \left( \left[ u_j^n - (1-p)\Delta x_j S_j \right] - \left[ u_{j-1}^n \right] \right). \end{aligned}$$

where we should note that  $S_j \geq 0$  and  $S_{j-1} = 0$  ( $x_{j-1}$  is a minimum point so TVD slope is zero). Since  $\bar{a} < 0$  and due to the constraint  $u_{j-1}^{n+1} \geq u_{j-1}^n$  (the local TVD condition 3), we should have:  $\left( \left[ u_j^n - (1-p)\Delta x_j S_j \right] - \left[ u_{j-1}^n \right] \right) \geq 0$ . This means:

$$0 \leq S_j \leq \frac{u_j^n - u_{j-1}^n}{(1-p)\Delta x_j} = \frac{\Delta_- u_j^n}{(1-p)\Delta x_j}. \quad (4.72)$$

From Eqs. (4.71) and (4.72), the TVD constraint on non-uniform grids is:

$$S_j = (u'_j)_{TVD} = \min \left\{ \frac{\Delta_- u_j^n}{(1-p)\Delta x_j}, \frac{\Delta_+ u_j^n}{p\Delta x_j} \right\}. \quad (4.73)$$

Or equivalently, based on the definition of  $f$ , the TVD constraints can be re-expressed as:

$$S_j = (u'_j)_{TVD} = \min \left\{ \frac{f}{1-p} \frac{\Delta_0 u_j^n}{\Delta x_j}, \frac{1-f}{p} \frac{\Delta_0 u_j^n}{\Delta x_j} \right\}. \quad (4.74)$$

- ii: The TVD condition when the transmitting cell is the first or the last cell (the set (1) or (3))** In this case, it is easy to show that the TVD limiter constraint,  $u'_j$  is the same as cell-centered cells, see Eq. (4.61).

### Controlling of the TVD condition for slope limiters on wavelet-based adapted grids

A modified wavelet-based adapted grid is considered; the modification is done with the post-processing stage. It is assumed also there always exists at least three neighbor cell centers of equal distances from each other, as explained in Section 4.2.4 and illustrated in Figure 4.7(a). In this figure, we have:  $x_{j-1} - x_{j-2} = 2dx$ ,  $x_j - x_{j-1} = 2dx$ ,  $x_{j+1} - x_j = dx$  and  $x_{j+2} - x_{j+1} = dx$ . This configuration of grid points leads to stable and TVD solutions, which will be studied later. Depending of the transmitting cell location, three cell sequences are detectable: 1)  $\{j-2, j-1, j\}$ ; 2)  $\{j-1, j, j+1\}$ ; 3)  $\{j, j+1, j+2\}$  (Figure 4.7(a)).

**The cell sequence  $\{j-1, j, j+1\}$ : Transmitting cell  $j$  (with a shifted cell center) is the middle one** In this case, we have  $\Delta_-x_j = [a/2 + (1-p)]\Delta x_j$  and  $\Delta_+x_j = [p + b/2]\Delta x_j$ . So, the backward, forward and central derivatives (in the GMINMOD limiter, Eq. (4.65)) can be rewritten as:

$$D_-u_j = \frac{2f\Delta_0u_j}{[a + 2(1-p)]\Delta x_j}, \quad D_+u_j = \frac{2(1-f)\Delta_0u_j}{[2p + b]\Delta x_j}, \quad D_0u_j = \frac{2\Delta_0u_j}{(2 + a + b)\Delta x_j}, \quad (4.75)$$

where  $a := \Delta x_{j-1}/\Delta x_j = 4/3$  and  $b := \Delta x_{j+1}/\Delta x_j = 2/3$ ;  $p$  measures cell-center shifting and here  $p = 1/3$  (see Figure 4.7(a)-(b)).

The TVD-based-bound  $u'_j$  (Eq. (4.73)) and GMINMOD (Eq. (4.65)) are illustrated in Figure 4.7(b). The comparison offers: 1) the limiter remains completely in the TVD region; 2) the linear preserving feature is satisfied; 3) at the expense of the symmetric feature, the transmitting cell  $j$  acts properly for joining surrounding cells; and 4) shifting of the cell center leads to a TVD result (Figure 4.7(b)), while cell-centered one does not (Figure 4.6(c)).

**The cell sequence  $\{j-2, j-1, j\}$  (see Figure 4.7(c))** In this case, we have:  $\Delta_-x_j = (1+a)\frac{\Delta x_j}{2}$  and  $\Delta_+x_j = [1 + 2b(1-p)]\frac{\Delta x_j}{2}$ . Hence:

$$D_-u_j = \frac{2f\Delta_0u_j}{(1+a)\Delta x_j}, \quad D_+u_j = \frac{2(1-f)\Delta_0u_j}{[1 + 2b(1-p)]\Delta x_j}, \quad D_0u_j = \frac{2\Delta_0u_j}{(2 + a + 2(1-p)b)\Delta x_j}, \quad (4.76)$$

where  $a = 1$ ,  $b = 3/4$  and  $p = 1/3$ . Definition of the TVD-based-bound  $u'_{j-1}$  is the same as the uniform case, Eq. (4.61). Comparison of this limiter with the GMINMOD

is shown in Figure 4.7(c). It is obvious, the GMINMOD limiter remains in the TVD region.

**The cell sequence  $\{j, j+1, j+2\}$  (see Figure 4.7(d))** Here we have:  $\Delta_-x_j = \frac{(1+2ap)\Delta x_j}{2}$  and  $\Delta_+x_j = \frac{(1+b)\Delta x_j}{2}$ . And so:

$$D_-u_j = \frac{2f\Delta_0u_j}{(1+2ap)\Delta x_j}, D_+u_j = \frac{2(1-f)\Delta_0u_j}{(1+b)\Delta x_j}, D_0u_j = \frac{2\Delta_0u_j}{(2+2ap+b)\Delta x_j}, \quad (4.77)$$

where  $a = 3/2$ ,  $b = 1$  and  $p = 1/3$ . The TVD-limit for the first derivative at  $x_{j+1}$ ,  $u'_{j+1}$  is the same as the uniform case. This function is compared with the GMINMOD limiter in Figure 4.7(d), where stable and TVD results are attached.

In General, it can be concluded that on irregular grid points with typical grid configuration illustrated in Figure 4.7(a), numerical solutions will remain TVD and thereby stable.

### A constraint on the cell center adaptation in the wavelet-based algorithm

As mentioned before, it is always assumed that there exist at least three neighbor cell centers of equal distance from each other in non-uniform adapted cell centers. Let assume an adapted cell with a weaker cell-sequence condition:  $x_{j+1} - x_j = 0.5(x_j - x_{j-1})$  or  $x_{j+1} - x_j = 2(x_j - x_{j-1})$ . As an example let consider:  $x_{j-1} - x_{j-2} = 8dx$ ,  $x_j - x_{j-1} = 4dx$ ,  $x_{j+1} - x_j = 2dx$  and  $x_{j+2} - x_{j+1} = dx$  (Figure 4.8(a)). For this case, if cell edges are assumed to be in the middle of cell centers,  $x_{j+1/2} = 0.5(x_j + x_{j+1})$ , then cell lengths are:  $\Delta x_{j-1} = 6dx$ ,  $\Delta x_j = 3dx$  and  $\Delta x_{j+1} = 1.5dx$ . Hence:  $\Delta_-x_j = [(1-p) + ap]\Delta x_j$  and  $\Delta_+x_j = [p + (1-p)b]$ , and so:

$$D_-u_j = \frac{f\Delta_0u_j}{[(1-p) + ap]\Delta x_j}, D_+u_j = \frac{(1-f)\Delta_0u_j}{[p + (1-p)b]\Delta x_j}, D_0u_j = \frac{\Delta_0u_j}{[1 + ap + b(1-p)]\Delta x_j}, \quad (4.78)$$

where  $a = 2$  and  $b = 0.5$ . For all cells, shifting coefficient is equal to  $p = 1/3$  (see Figure 4.8(a)).

The GMINMOD (Eq. (4.65)) and  $u'_j$  (from non-uniform cases: Eq. (4.73)) are compared in Figure 4.8(b). It is clear that the limiter is slightly outside the TVD domain. For this reason, long term stability of numerical solutions could vanish. For cell-centered cases, where  $x_j = (x_{j+1/2} + x_{j-1/2})/2$ , the results are not also TVD, see Figure 4.6(b).

To have stable results, cell centers are located in a way that there always exists at

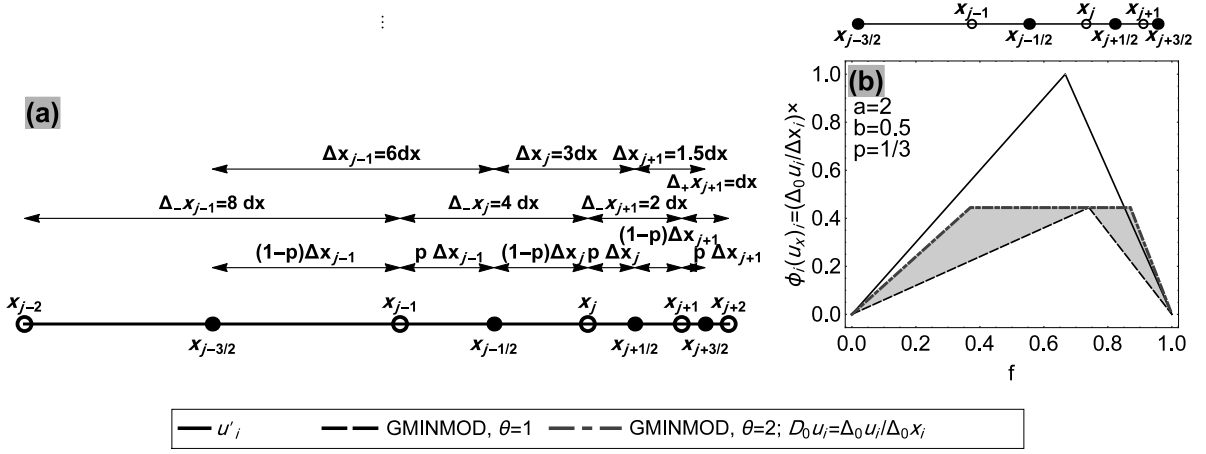


Figure 4.8: Effects of a weaker restriction on wavelet-based grid variation. For this figure, we assume  $x_{j+1} - x_j = 0.5(x_j - x_{j-1})$ ; a) cell configurations; b) the GMINMOD limiter performance.

least three neighbor grids (cell centers) of equal distance from each other. For guarantee this condition, in the grid modification stage (the post-processing stage), it is assumed, at least, to have:  $N_s = N_c = 1$ .

### 4.3 Wavelet-based adaptive-grid method to solve PDEs

Assume the solution of a PDE is  $\mathbf{f}(\mathbf{x}, t)$ . At time step  $t = t_n$ , the procedure for the solution of the PDE in adaptive wavelet-based framework is summarized as a sequence of steps:

1. Determine the adaptive grid from  $\mathbf{f}(\mathbf{x}, t_{n-1})$  (time-step  $n - 1$ ) by using the adaptive wavelet transform. For a point without  $\mathbf{f}(\mathbf{x}, t_{n-1})$  value, the local Lagrange interpolation scheme is used to estimate the missed information. For adaptation, normalized value of functions  $f_i(\mathbf{x}, t_{n-1})$  will be considered; i.e.:

$$\left\{ \frac{f_i(\mathbf{x}, t_{n-1})}{\text{Max}_j |f_j(\mathbf{x}, t_{n-1})|} \right\} < \epsilon.$$

In most cases initial proposed values of pre-defined threshold value could be selected from range  $\epsilon \in \{10^{-5}, 10^{-4}\}$ . In general for each  $f_i(\mathbf{x}, t_{n-1})$  different threshold values can be considered, as well; in this work, however, for all components  $f_i(\mathbf{x}, t_{n-1})$ , one threshold value is used, which leads to a simple algorithm.

2. Modify the adapted grid obtained from the previous step by the post-processing stage; the initial proposed values for  $N_s$  and  $N_c$  would be:  $N_s = 2$  and  $N_c = 1$ ;

however, one should at least consider  $N_s = 1$  and  $N_c = 1$ .

3. Use the semi-discretize form of the KT central high resolution scheme. The TVD based time-stepping method like the TVD-based Runge-Kutta method [22, 125] by taking into account the CFL condition can be used to solve the ODEs at time  $t = t_n$ ; here, the second order Runge-Kutta method is used.
4. Go back to step 1.

In practice, to have a cost-effective computation, the grid is not adapted at each time step. In fact, depending on the velocity of moving fronts, it is adapted after several time steps, for example every ten or twenty time steps.

## 4.4 Examples

The following 1-D and 2-D examples are to study the effectiveness of the proposed method. The main assumptions are: 1- Applying D-D interpolating wavelet of order 3; 2-Using the generalized MINMOD flux limiter; 3- Using the semi-discrete form of the KT scheme; 4- Integrating in time domain by the RK second-order solver having the TVD feature.

In general, Figure 4.9 summarizes the MRA-based simulation by the central schemes (for 1-D problems).

### 1-D first order hyperbolic PDEs

**The linear advection equation** The linear pure advection equation is:

$$u_t + (au)_x = 0, \quad (4.79)$$

where  $a$  is the velocity. It is assumed this equation has the initial condition (IC)  $u(x, t = 0) = u_0(x)$ ; then the exact solution of the Eq. (4.79) at time  $t$  is:

$$u(x, t) = u_0(x - at). \quad (4.80)$$

In this example, the IC is:

$$u_0(x) = H(x - 0.2) - H(x - 0.4), \quad (4.81)$$



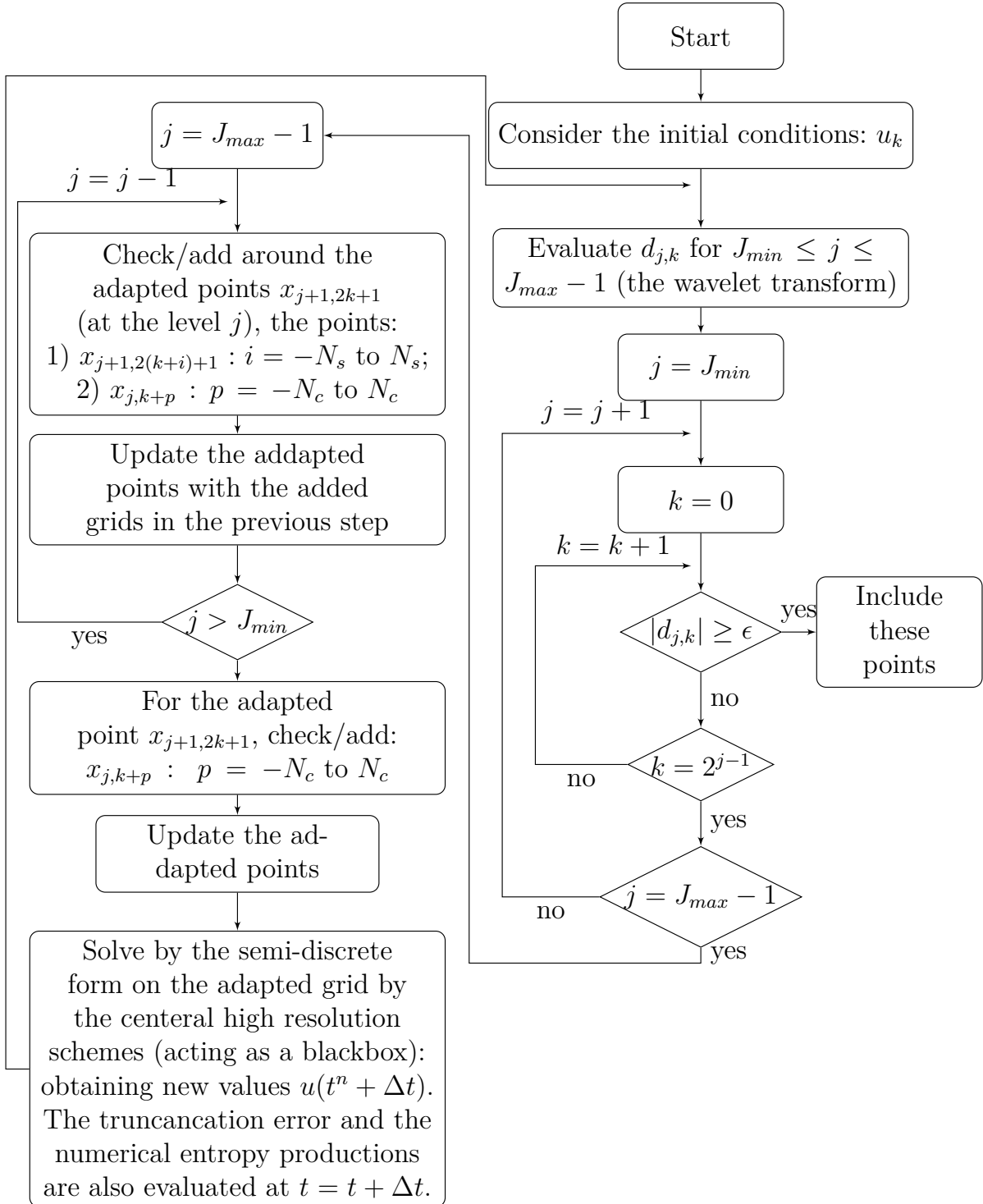


Figure 4.9: The MRA-based algorithm for the simulation of hyperbolic problems.

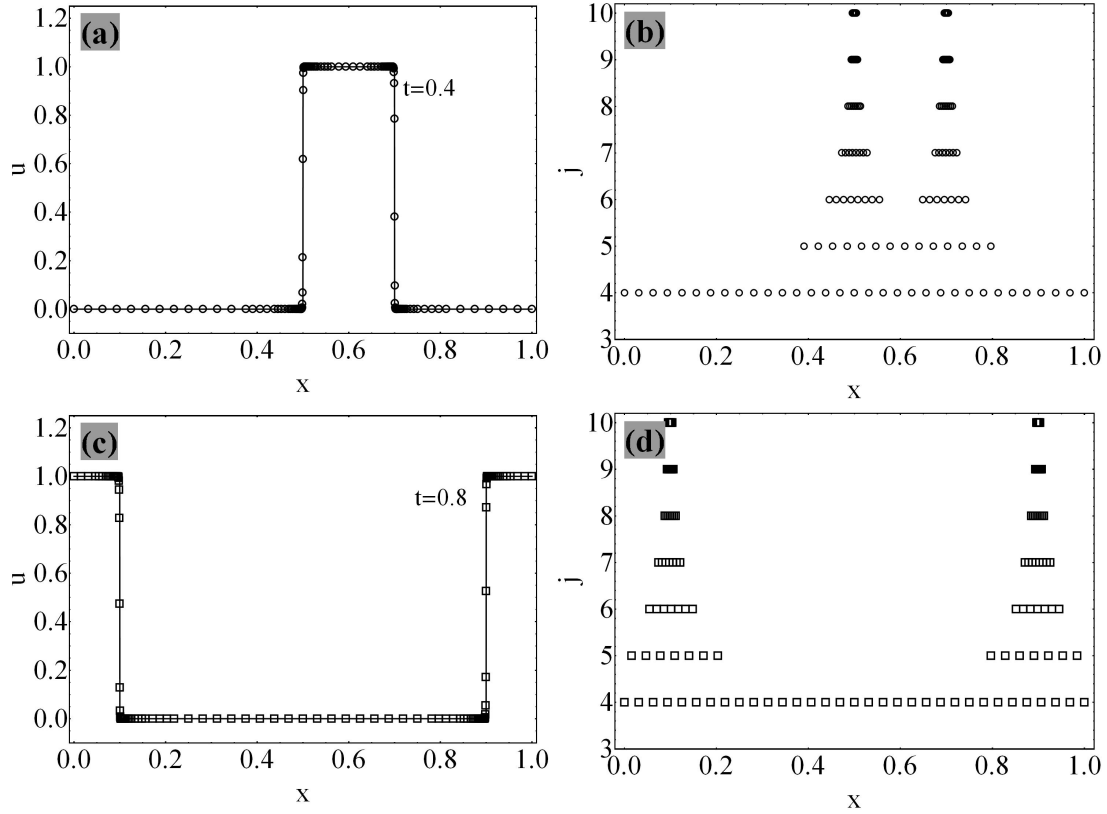


Figure 4.10: Numerical solutions and corresponding adapted grid points represented in different resolution levels; a,b) for  $t = 0.4$ ; c,d) for  $t = 0.8$ .

where  $H(x)$  denotes the Heaviside function, which represents the unit step function, equals to 0 for  $x < 0$  and 1 for  $x \geq 0$ . It is assumed that the boundary conditions (BCs) at  $x = 0$  and  $x = 1$  both are periodic:  $u(x = 0, t) = u(x = 1, t)$ . Eq. (4.79) is simulated numerically with parameters:  $a = 1$ ,  $\epsilon = 10^{-3}$  (threshold),  $\theta = 2$  (the flux limiter parameter),  $N_c = N_s = 2$  (for post-processing stage),  $J_{max} = 11$ , and  $J_{min} = 5$ . Numerical results and corresponding adapted grids are presented in Figure 4.10 at time  $t = 0.4$  and  $t = 0.8$ . Variation of number of adapted grid points ( $N_g$ ) during time is shown in Figure 4.11, where number of uniform grid points at the finest resolution is  $2^{J_{max}+1} = 2049$ . This confirms efficiency of the adaptation procedure. At time  $t = 0.75$  and  $t = 0.85$  discontinuous fronts reach the boundaries and this leads to jumping in  $N_g$  values.

**The Burgers' equation** This is the simplest model to simulate nonlinear advection. It appears in different applied mathematics, for instance nonlinear acoustics, fluid me-

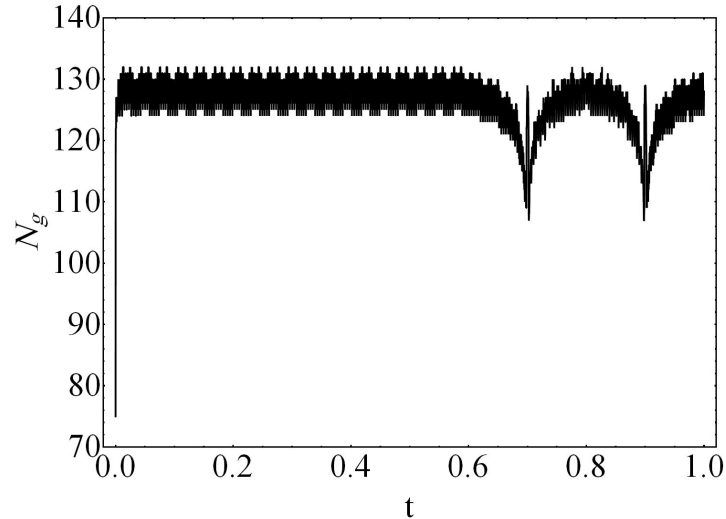


Figure 4.11: Number of grids points,  $N_g$  during time; at times  $t = 0.75$  and  $t = 0.85$  discontinuous fronts reach the boundaries. In the finest resolution, number of grid points is 2049.

chanics and traffic flow. The Burgers' equation is defined as follows:

$$u_t + \frac{1}{2} (u^2)_x = 0, \quad (4.82)$$

where  $u$  is the conserved quantity and its flux is  $F(u) = u^2/2$ . The system is nonlinear, so that discontinuous fronts will develop during front propagation. Here it is assumed that the initial and boundary conditions are:

$$\text{BCs : } u(x = 0, t) = u(1, t) = 0, \quad \text{ICs : } u(x, t = 0) = \sin(2\pi x) + \frac{1}{2}\sin(\pi x). \quad (4.83)$$

For the above mentioned conditions, a discontinuity starts to appear around  $t \approx 0.158$ . This discontinuous front will propagate to the right side after this time. Assumptions for the numerical simulations are:  $\epsilon = 10^{-3}$  (threshold) and  $\theta = 2$  (the flux limiter parameter),  $N_c = N_s = 2$  (for post-processing stage),  $J_{max} = 11$ , and  $J_{min} = 5$ . The numerical results are illustrated in Figure 4.12 at times 0.158, 0.5, and 1. This figure contains numerical results, exact solutions and corresponding adapted grids in different resolutions. The results confirm that adapted points are properly concentrated around propagating fronts.

The number of grid points used during simulation,  $N_g$  is presented in Figure 4.13 where the uniform grid has  $2^{J_{max}+1} = 2049$  points. This confirms the adaptation is

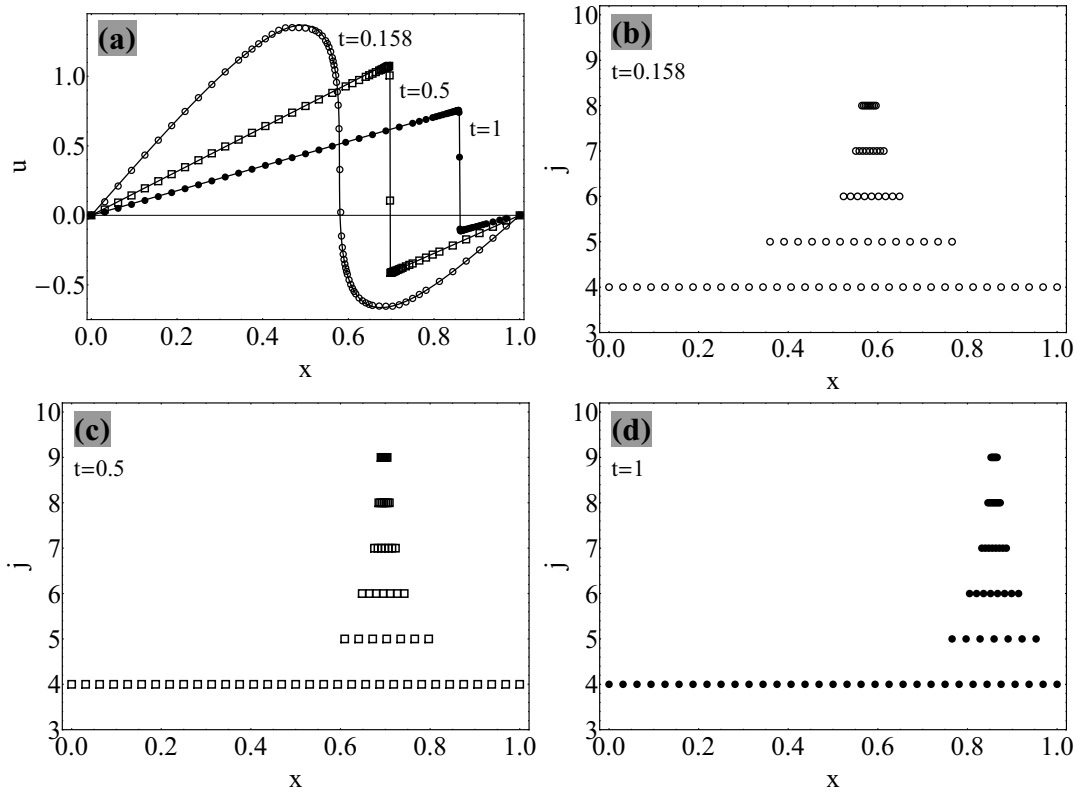


Figure 4.12: Numerical results of the Burgers' equation and corresponding adapted grid points in different resolution levels; in figure (a), exact solutions are presented by the solid line and numerical solutions are illustrated by points and hollow shapes. In the finest resolution, number of grid points is 2049.

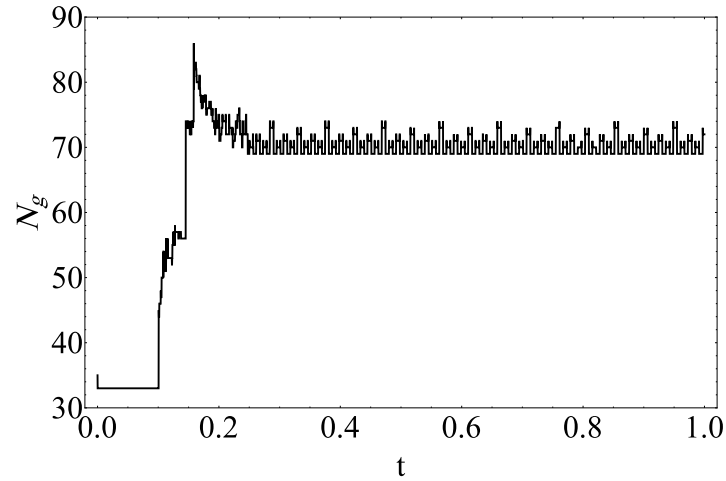


Figure 4.13: Number of grid points  $N_g$  used during the simulation; the uniform grid has 2049 points.

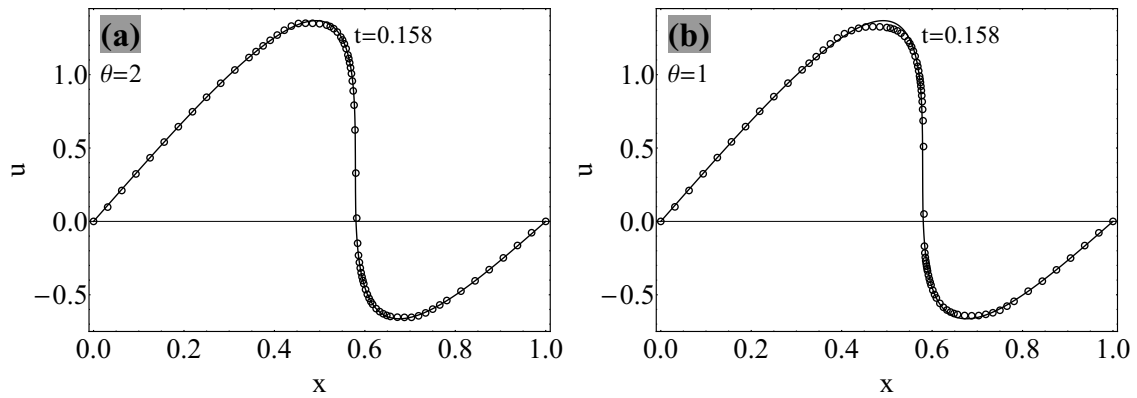


Figure 4.14: Parameter effect of the generalized MINMOD slope limiter; a) result for  $\theta = 2$ ; b) solution for  $\theta = 1$ .

effective and also the Figure 4.13 shows the adapting procedure is stable.

Effect of the flux limiter parameter  $\theta$  is studied in Figure 4.14 at time 0.0158; there two values are selected:  $\theta = 1$  and  $\theta = 2$ . The former corresponds to the most dissipative solution, and the later one associates with the least dissipative result (Figure 4.14).

### Some comparisons with other MRA-based adaptive methods

In the following, the results of the adaptive KT scheme are compared with: i) The adaptive NVSF method [46]; ii) The adaptive essentially nonoscillatory-Roe (ENO-Roe) scheme [173, 174]; iii) The adaptive scheme proposed by Holmström [54] by the integration of central finite difference (FD) schemes with the sparse point representation (SPR)

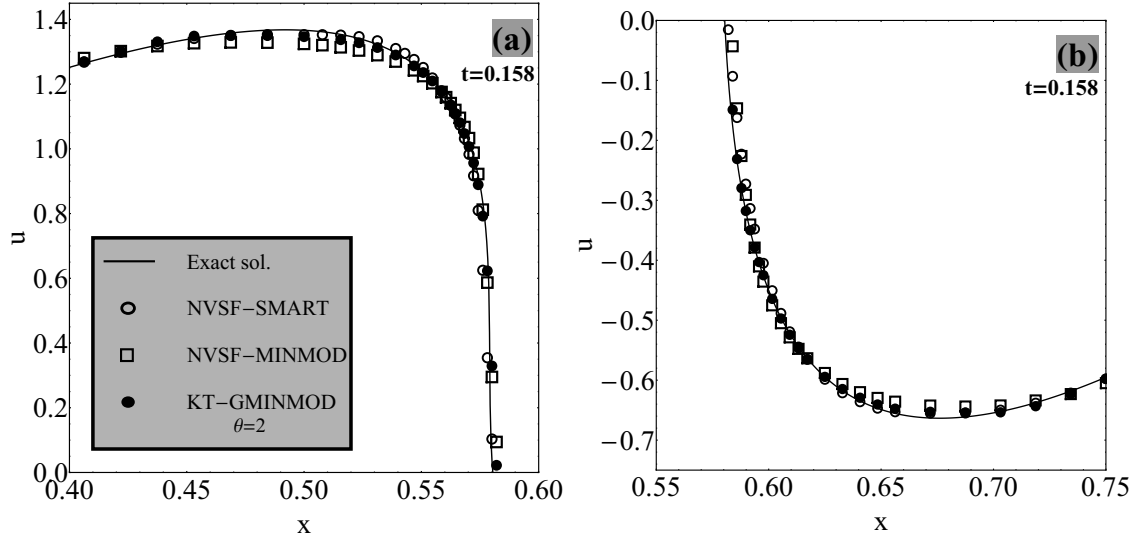


Figure 4.15: Comparison of the proposed adaptive central method and NVSF-based upwind scheme.

of solutions (SPR-FD) .

**i) The adaptive NVSF** The KT central scheme and NVSF-based method [46] (see Appendix B for details) are compared on wavelet-based adapted grids (the NVSF-based method is basically developed for non-uniform grid points). The values of the parameters are:  $\epsilon = 10^{-3}$ ,  $\theta = 2$  (for GMM in the KT scheme),  $N_c = N_s = 2$ ,  $J_{max} = 11$  and  $J_{min} = 5$ . The *second* order TVD Runge-Kutta method is used for the time integration. The results are presented in Figure 4.15 at time 0.158. For the NVSF formulation, two types of flux limiters are considered: the SMART and MINMOD limiters [46]. The results offer that the proposed method is comparable with the schemes originally provided for non-uniform grid points.

**ii) The adaptive ENO-Roe scheme** Both the adaptive KT and ENO-Roe methods are simulated with the parameters:  $\epsilon = 10^{-2}$ ,  $N_c = N_s = 1$ ,  $J_{max} = 12$  and  $J_{min} = 4$ . For the KT scheme, it is assumed:  $\theta = 2$  (for the GMM). For both methods, the time integration is performed by the *third-order* TVD Runge-Kutta method.

Adaptive and exact solutions are presented in Figure 4.16 at  $t = 0.158$ , 0.5 and 1; in these illustrations, markers  $\circ$  and  $\times$ , and the solid line show the solutions obtained by the ENO-Roe, the KT and the exact method, respectively. It is obvious that results from two adaptive methods have a good agreement (the KT method is slightly more dissipative than the ENO-Roe scheme; see the solutions at  $t = 0.158$ ).

**iii) The SPR-FD scheme proposed by Holmström [54]** In this benchmark, a Burgers' equation with a diffusion term is considered, as:

$$u_t + \frac{1}{2} (u^2)_x = Q^x(u, u_x)_x, \quad (4.84)$$

where here  $Q^x(u, u_x) = \mu u_x$  is the diffusion term, in which  $u_x := \partial u / \partial x$  and  $\mu$  is a constant. The BCs and IC are assumed to be:

$$\text{BCs : } u(x=0, t) = u(1, t) = 0, \quad \text{IC : } u(x, t=0) = \sin(2\pi x). \quad (4.85)$$

In this study, the assumed parameters are:  $\mu = 10^{-4}$ ,  $\epsilon = 10^{-5}$ ,  $\theta = 2$ ,  $J_{max} = 14$  and  $J_{min} = 4$  [54]. For the KT scheme, the third-order TVD Runge-Kutta method is used for the time integration; for the SPR-FD method, the common fourth-order Runge-Kutta method is used for the time integration and spatial derivatives are approximated by the centered finite difference method stencil of order  $p = 4$ . A discontinuity develops in the solution of Eqs. (4.84) at  $x = 0.5$  which obtains its maximum at  $t = 0.25$ . The adaptive solutions are presented in Figure 4.17 at  $t = 0.25$ . The zoomed in solutions reveal that both solutions can properly resolve the discontinuity.

**1-D Problems with non-convex fluxes** Here, a relatively difficult and challenging problem containing a high-order non-linear flux will be tested; the problem includes a non-convex flux, i.e.:

$$F(u) = (u^2 - 1)(u^2 - 4)/4. \quad (4.86)$$

Two cases are assumed for the initial condition:

**Case 1** The considered Riemann initial condition is:

$$u(x, t=0) = 3 \text{ Sign}(x). \quad (4.87)$$

In numerical simulations, it is assumed: the threshold,  $\epsilon$ , values are:  $\epsilon \in \{10^{-3}, 0.5 \times 10^{-4}, 10^{-4}\}$ ;  $J_{max} = 11$ ;  $\theta = 2$ ; number of decomposing level is  $N_d = J_{max} - J_{min} = 7$ ;  $dt = 10^{-5}$ . Number of grid points used during simulations,  $N_g$  for the above-mentioned thresholds are presented in Figure 4.18. There, Figure 4.18(a) shows threshold effects

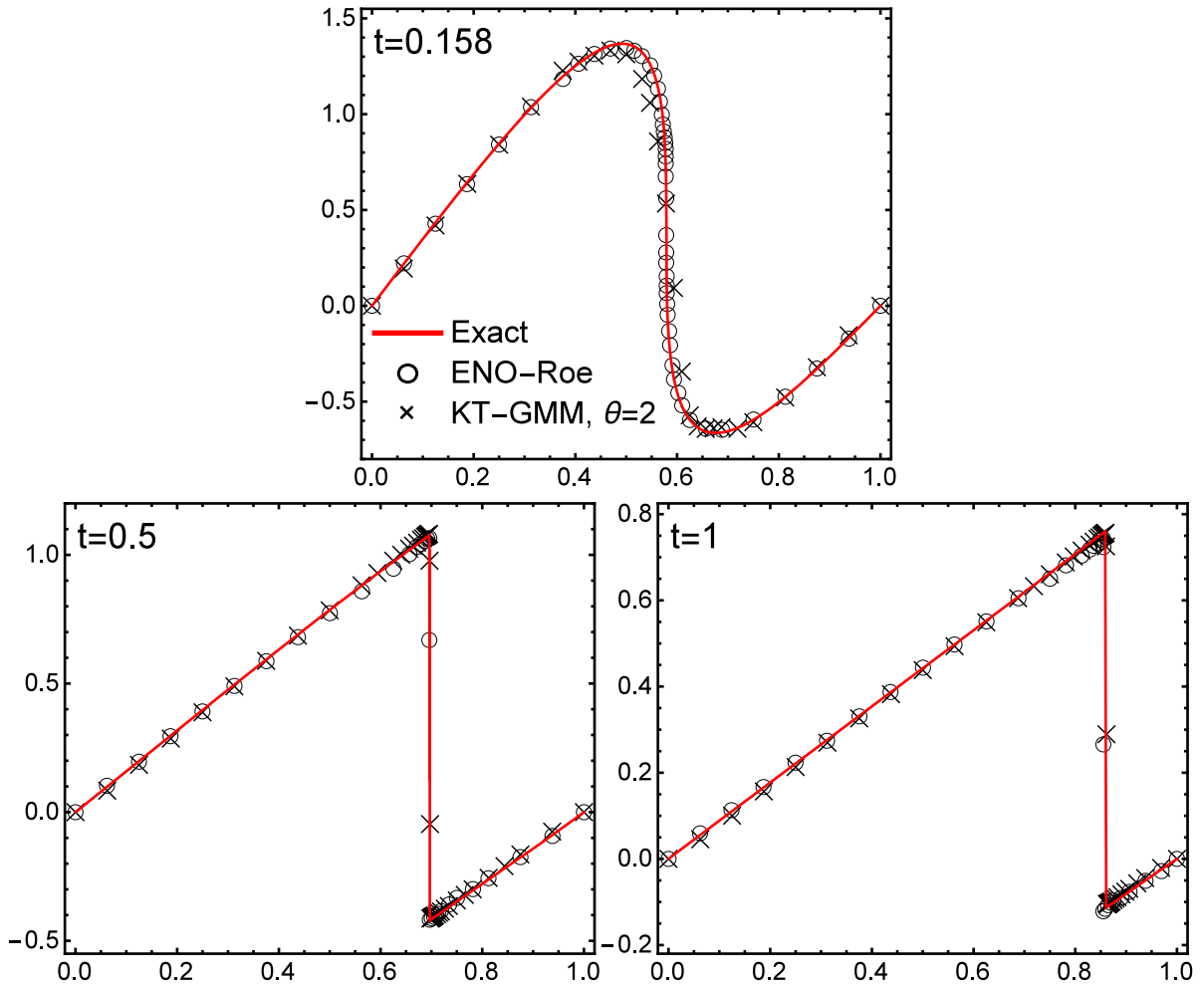


Figure 4.16: The Comparison between the adaptive KT and ENO-Roe [173] schemes on MRA-based adapted grids at  $t = 0.158, 0.5$  and  $1$ .



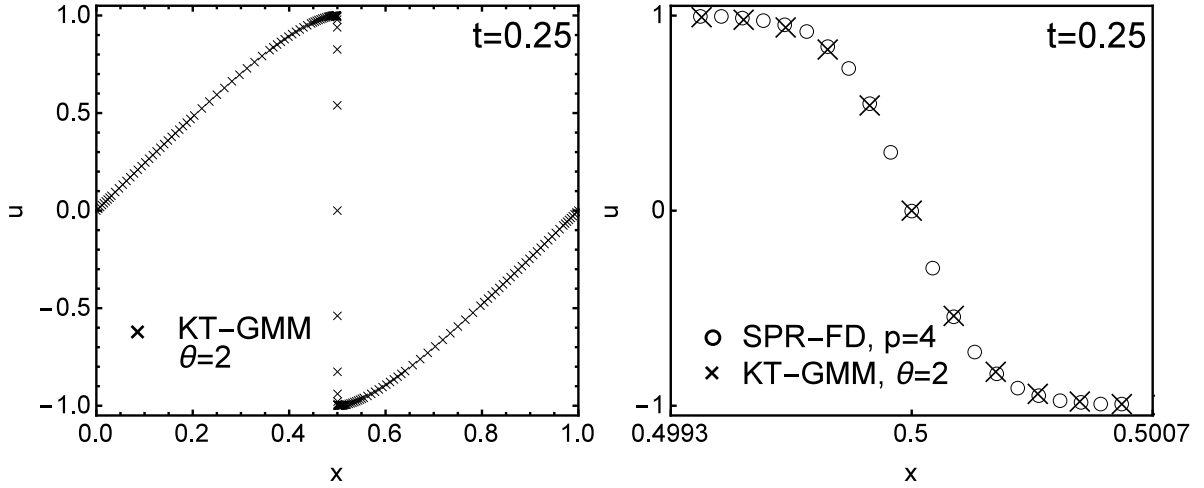


Figure 4.17: The Comparison between the adaptive KT and SPR-FD [54] schemes on MRA-based adapted grids at  $t = 0.25$ , where  $J_{max} = 14$  and  $J_{min} = 4$ .

on number of adapted grids; in Figure 4.18(b) number of uniform grid points without adaptation (with number  $N_g = 2^{J_{max}+1}$ ) is compared with those after grid adaptation in logarithmic scale; the compression capability of the wavelet-based adaptation is clear.

Adapted solutions for different threshold values are illustrated in Figure 4.19 at  $t = 0.04$ ; there, exact and numerical results, as well as distribution of adapted points at different resolutions are presented. The results indicate that: 1) Adapted points are properly concentrated around high-gradient zones; 2) The discontinuity is successfully captured and simulated; 3) By decreasing the  $\epsilon$  values, more grid points are concentrated around high gradient zones.

In this example, to investigate the multiresolution-based grid modification effects, the problem is resolved without grid modification stage. Corresponding results are presented in Figure 4.20. There, in figure (a) The solution and corresponding adapted grid are shown for case  $N_c = N_s = 0$ . In figure (b),  $N_g$  values of solution with and without modification stage are compared with each other; where for the solution with grid modification it is assumed  $N_c = 1$  and  $N_s = 2$ . The results confirm that without employing grid modification stage, spurious oscillations develop quickly; this will lead to instability. Due to this, thereby, number of adapted grid points increases rapidly (Figure 4.20 (b)).

**Case 2** The initial condition is:

$$u(x, t = 0) = -2 \text{Sign}(x). \quad (4.88)$$

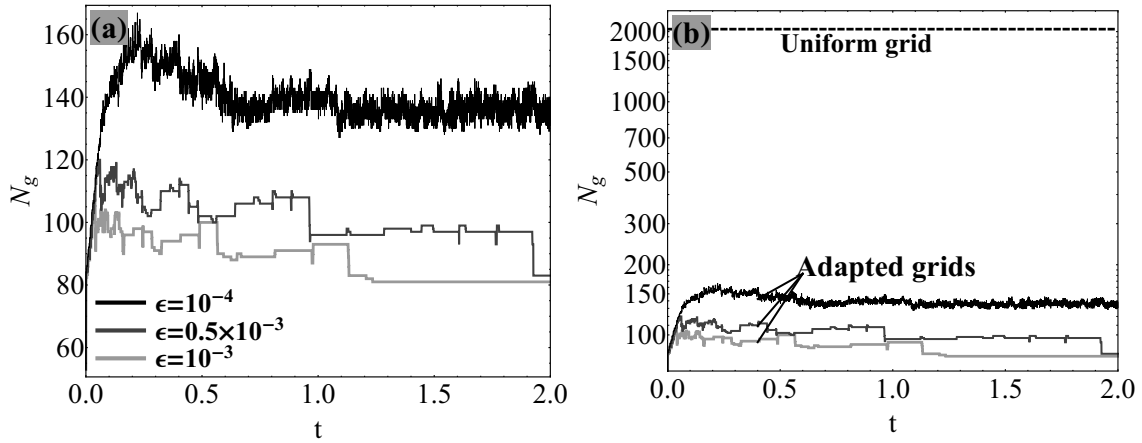


Figure 4.18:  $N_g$  values during simulations for different  $\epsilon$  values; a) adapted grids, b) comparison of  $N_g$  values between the uniform grid and corresponding adapted ones.

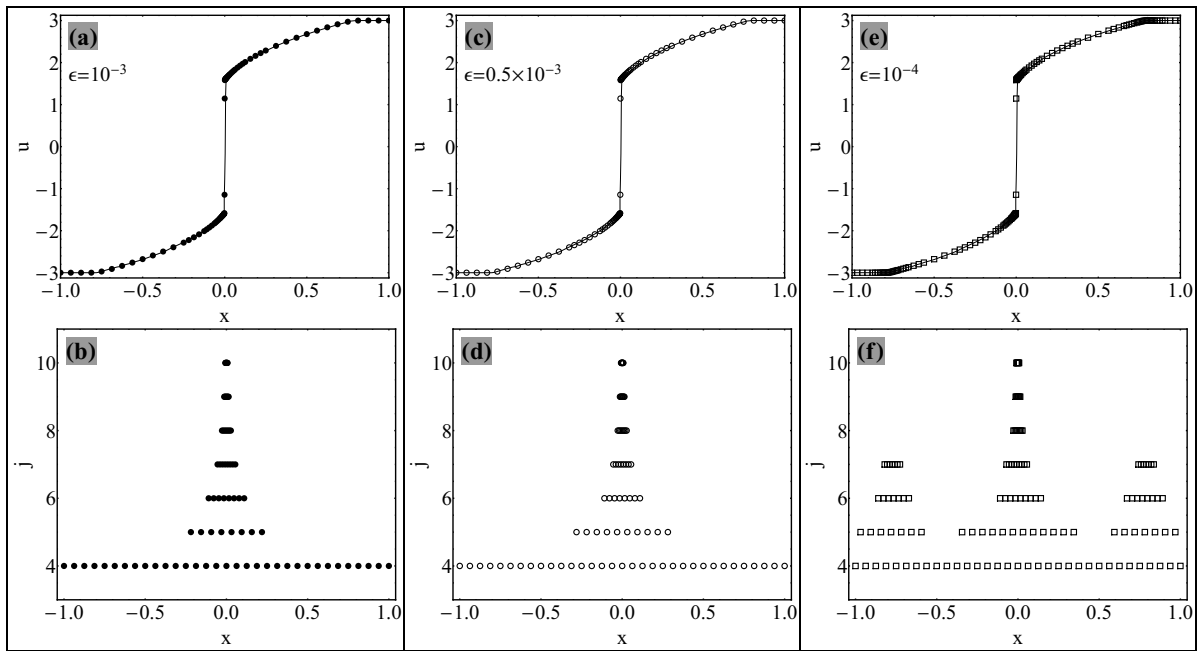


Figure 4.19: Adapted solutions for different threshold values at  $t = 0.04$ ; a, b)  $\epsilon = 10^{-3}$ , c, d)  $\epsilon = 0.5 \times 10^{-3}$ , e, f)  $\epsilon = 10^{-4}$ . Solid line and shapes are the exact solution and adapted ones, respectively. In all calculations it is assumed that:  $N_c = 1$  and  $N_s = 2$ .

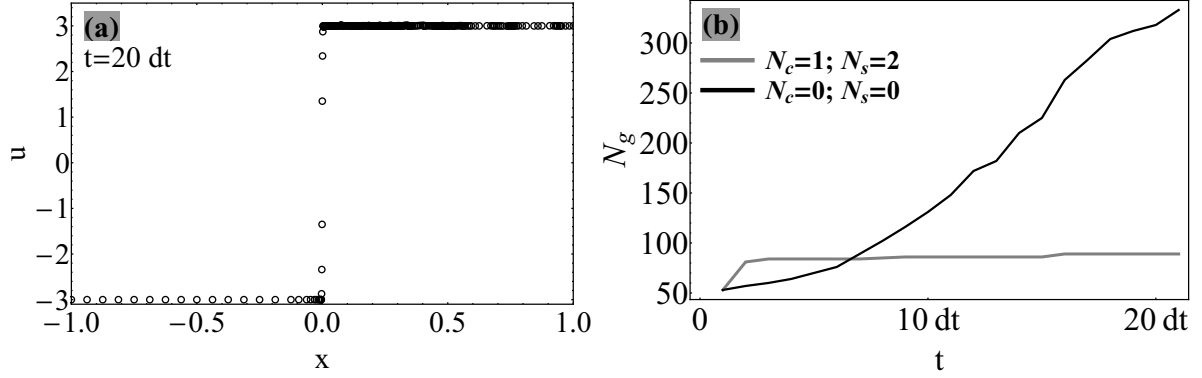


Figure 4.20: Effect of adaptive simulation without grid modification stage. a) Adapted solution and corresponding adapted grid points in case  $\{N_c = N_s = 0\}$ ; b)  $N_g$  values of solutions in two cases: 1) with grid modification ( $\{N_c = 1 \& N_s = 2\}$ ), 2) without grid modification ( $\{N_c = 0 \& N_s = 0\}$ ).

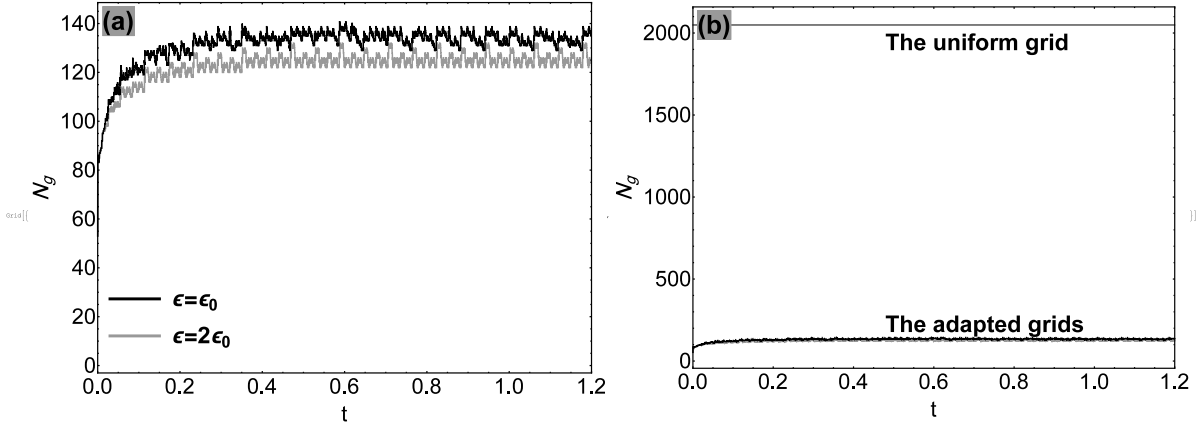


Figure 4.21:  $N_g$  values during simulations for different threshold values; it is assumed  $\epsilon_0 = 5 \times 10^{-4}$ ,  $\theta = 2$ ;  $J_{max} = 11$  and  $N_d = 6$ .

The threshold used for simulations are:  $\epsilon = \{\epsilon_0, 2\epsilon_0\}$ , where  $\epsilon_0 = 5 \times 10^{-4}$ . Also, it is assumed that:  $\theta = 2$ ;  $J_{max} = 11$  and number of decomposition levels is  $N_d = 6$  (i.e.,  $J_{min} = 5$ ). The grid points used during simulations,  $N_g$  is presented in Figure 4.21. It is clear that smaller values of thresholds lead to larger value of adapted points. Numerical results obtained at  $t = 1.2$  are illustrated in Figure 4.22. There, the exact [22] and numerical results, as well as distribution of adapted points at different resolutions are shown. The results indicate that: 1) Adapted points are properly concentrated around high-gradient zones; 2) The discontinuity is successfully captured.

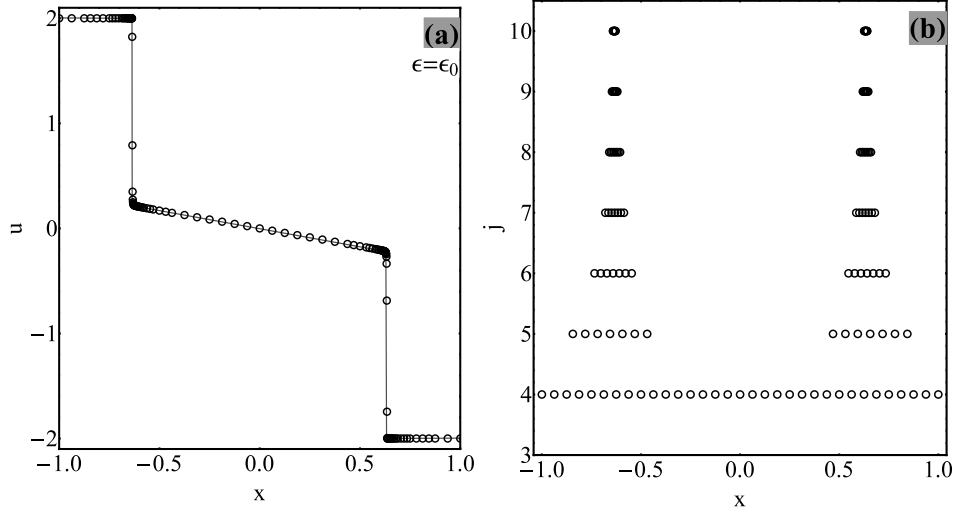


Figure 4.22: The solution at  $t = 1.2$  and corresponding adapted points for  $\epsilon = \epsilon_0$ ; the solid line and hollow circle illustrate exact and numerical solutions, respectively.

**Euler system of equations** The 1-D Euler system of equations appearing in gas dynamic problems can be expressed as:

$$\frac{\partial}{\partial t} \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix} + \frac{\partial}{\partial x} \begin{bmatrix} \rho u \\ \rho u^2 + P \\ u(E + P) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad (4.89)$$

where  $\rho$ ,  $u$  and  $E$  are gas density, velocity and total energy, respectively. The pressure  $P$  is obtained as:

$$P = (\gamma - 1) \left( E - \rho u^2 / 2 \right), \quad (4.90)$$

where  $\gamma$  is the ratio of specific heats, and here it is assumed to be  $\gamma = 1.4$ .

**I: Interaction of two blast waves** One of the challenging problems in the gas dynamic is the two blast waves interaction. The initial conditions are assumed to be:

The Riemann initial condition is [174, 175]:

$$\{\rho, u, P\}|_{t=0} = \begin{cases} \{1, 0, 1000\}, & \text{if } x < 0.1, \\ \{1, 0, 0.01\}, & \text{if } 0.1 < x < 0.9, \\ \{1, 0, 100\}, & \text{if } x > 0.9. \end{cases} \quad (4.91)$$

The boundary conditions are reflecting walls, where the velocity  $u$  is an odd function of distance from the wall,  $\rho$  and  $P$  are even functions respect to the wall in spatial domain [174, 175]. In general, the boundary condition can be satisfied at time step  $n$  by using  $r$  ghost cells for the reflecting wall,

$$\begin{cases} \rho_{-j+1} &= \rho_j, \\ P_{-j+1} &= P_j, \quad \text{for } j = 1, \dots, r. \\ u_{-j+1} &= -u_j, \end{cases} \quad (4.92)$$

In numerical simulations, only one ghost cell is utilized for each reflecting wall, i.e.:  $r = 1$ .

The numerical simulation is performed with  $\epsilon = 10^{-2}$ ,  $dt = 0.2 \times 10^{-5}$ , and  $\theta = 2$ . The results at times 0.01, 0.016, 0.026, 0.028, 0.03, 0.032, 0.034, and 0.038, are illustrated in Figure 6.26. For each time, the numerical solution, the corresponding adapted grid and corresponding distribution of adapted points in different resolution levels are presented. It is clear that points are properly adapted in the vicinity of both high-gradient and discontinuous zones.

**II: The interaction of Mach 3 right-moving front and an entropy wave** In this example, the interaction of an entropy sine wave with a Mach 3 right-moving front will be studied; this problem is also known as the Shu and Osher problem [174]. This problem is important due to the simulation of the shock-turbulence interactions. Amplified high-frequency entropy waves develop after the shock. These high frequency waves can be captured with numerical schemes with low numerical dissipation. Hence, this problem is a challenging benchmark to study the performance of numerical methods in: 1) Simulation of shock-turbulence interactions, 2) Handling of high frequency waves. This challenging problem was developed to reveal capabilities of high order schemes.

The Riemann initial condition is [174, 175]:

$$\{\rho, u, P\}|_{t=0} = \begin{cases} \{3.857143, 2.629369, 10.33333\}, & \text{if } x \leq -4, \\ \{1 + 0.2 \sin(5x), 0, 1\}, & \text{if } x > -4. \end{cases} \quad (4.93)$$

The considered spatio-temporal computational domain is:  $\Omega = (-5, 5) \times (0, T)$ .

The numerical and exact solutions, and corresponding adapted points are illustrated in Figure 4.24 at  $t = 1.8$  for parameters:  $\epsilon = \epsilon_0 = 5 \times 10^{-4}$ ,  $J_{max} = 13$ ,  $J_{min} = 5$  (or  $N_d = 8$ ),  $N_c = 1$ ,  $N_s = 2$ ,  $\theta = 2$ , and  $dt = 0.0002$ . There, the solid line and hollow circles denote reference [175] and numerical solutions, respectively. It is

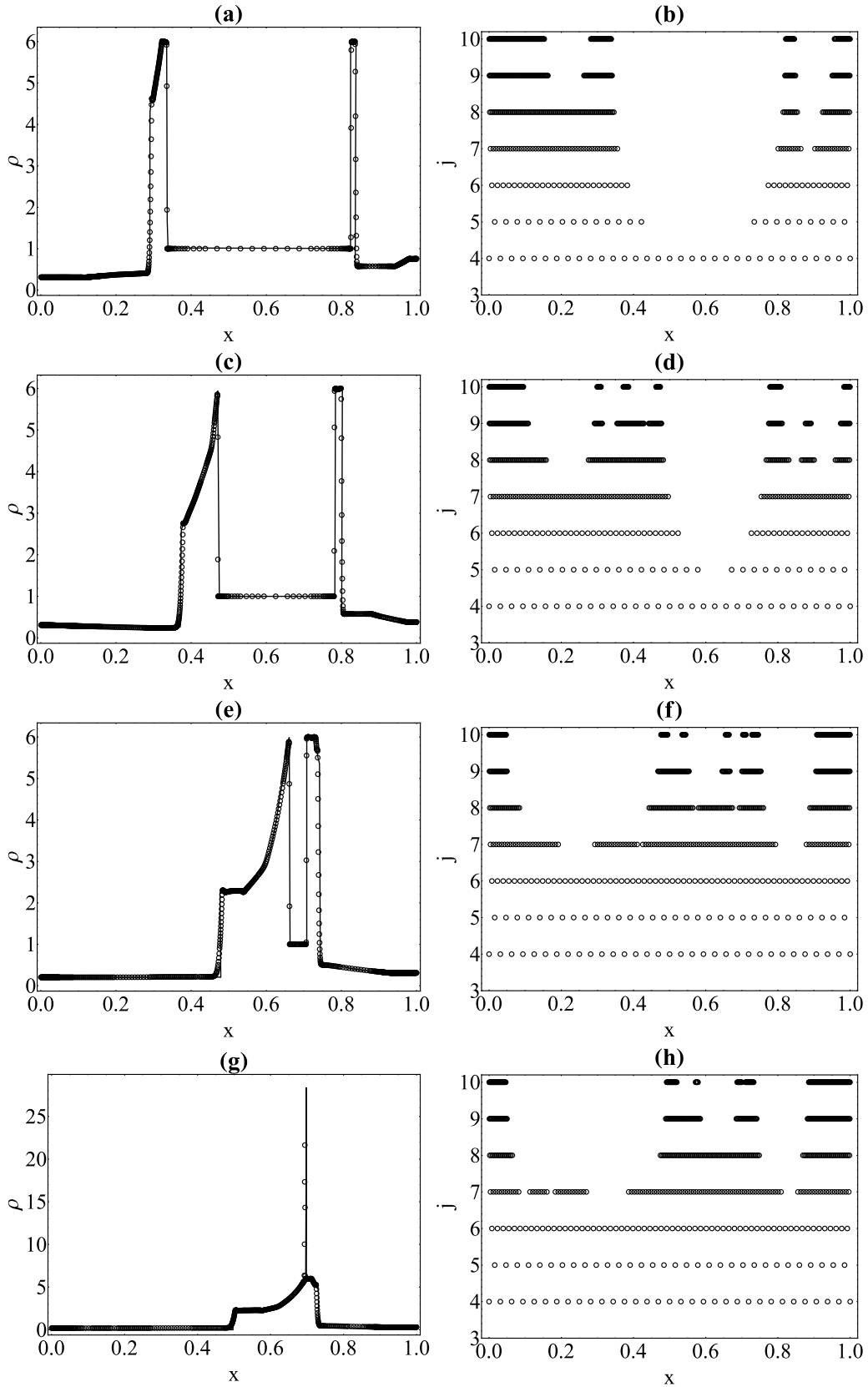


Figure 4.23: Numerical solutions and corresponding adapted grids of interaction of two blast waves with parameters  $\epsilon = 10^{-2}$ , and  $\theta = 2$ ; a-b)  $t = 0.01$ ; c-d)  $t = 0.016$ ; e-f)  $t = 0.026$ ; g-h)  $t = 0.028$ ;

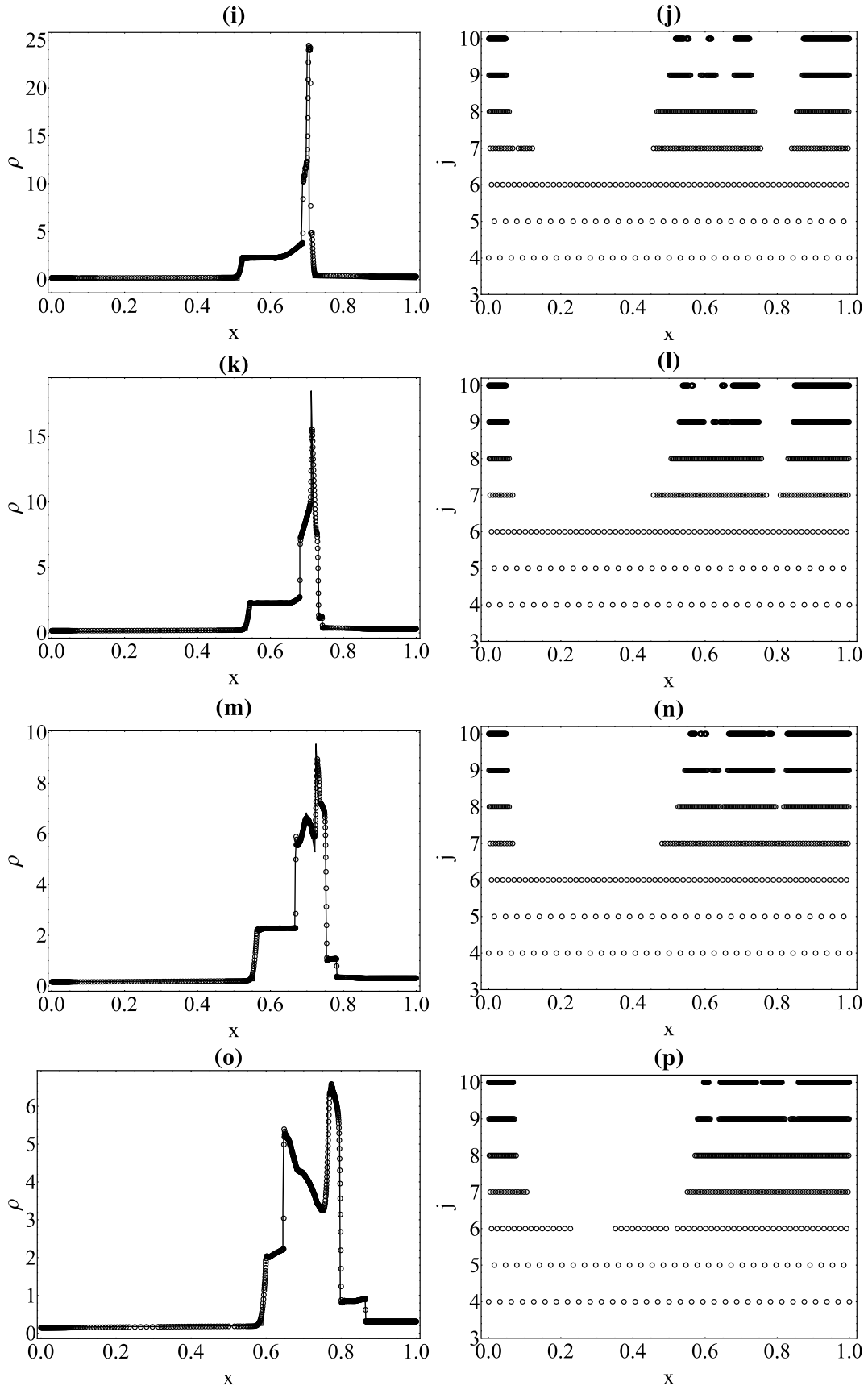


Figure 4.23: (continued) i-j)  $t = 0.03$ ; k-l)  $t = 0.032$ ; m-n)  $t = 0.034$ ; o-p)  $t = 0.038$ .

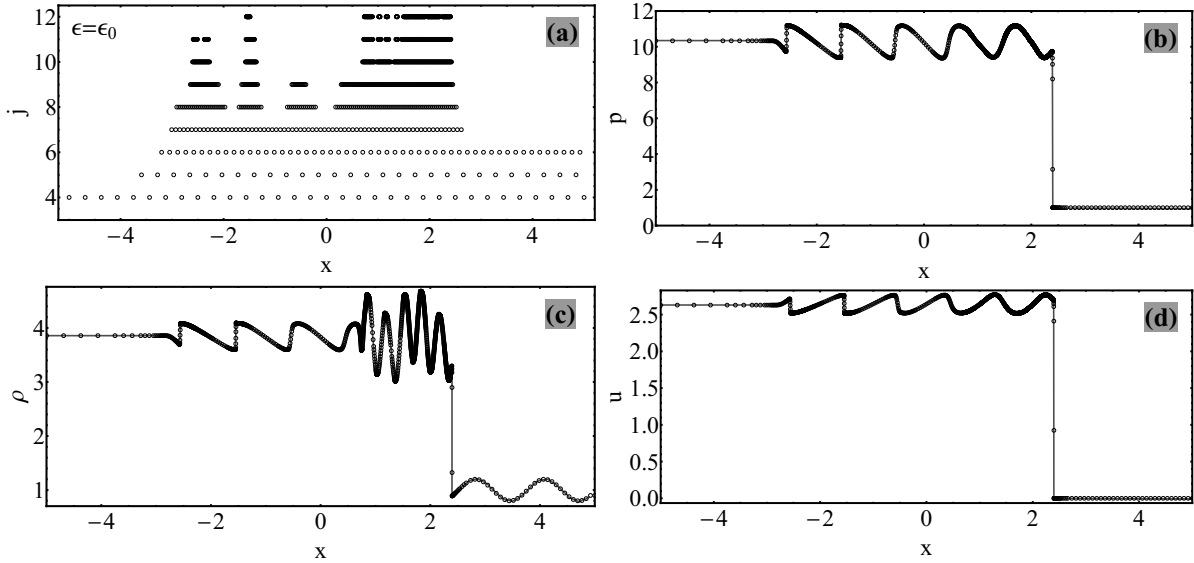


Figure 4.24: Reference and numerical solutions of entropy and shock wave interaction at  $t = 1.8$ . Solid lines and shapes are the reference and adaptive results, respectively. In all calculations it is assumed that:  $N_c = 1$  and  $N_s = 2$ .

obvious that the adapted points concentrate properly around both high-gradient zones and discontinuities, and there are good agreements between numerical and reference solutions.

To investigate threshold value effects, some other simulations are done for threshold values:  $\epsilon \in \{8\epsilon_0, 2\epsilon_0, \epsilon_0, \epsilon_0/2\}$ . In all simulations, it is assumed that the finest and coarsest resolution levels are  $J_{max} = 13$  and  $J_{min} = 5$ , respectively. The results and corresponding adapted grid points are illustrated in Figures 4.25 and 4.26, respectively; there, it is assumed:  $x \in [-3, -2.45]$ , and  $N_c = 1$  and  $N_s = 2$ . The results show that for  $\epsilon$  of large values dissipation phenomenon occurs (Figure 4.25 (a)); by decreasing  $\epsilon$  values, more adapted points concentrate around high-gradient regions (Figures 4.25 & 4.26) resulting in proper capturing of such zones. In this investigation, the number of grid points used during simulations ( $N_g$ ) is shown in Figure 4.27; in Figure 4.27(b) a comparison between the number of uniform grid (with grid number  $N_g = 2^{J_{max}} + 1$ ) and the corresponding adapted grids is also presented to show the compression ratio.

Resolution effect is also investigated; for this, three different problems with different  $J_{max}$  values are considered. The considered finest resolution levels are:  $J_{max} \in \{13, 11, 9\}$ ; in all cases it is assumed that  $J_{min} = 5$  and  $\epsilon = \epsilon_0$ . The solutions and corresponding adapted points are illustrated in Figure 4.28 for  $N_c = 1$  and  $N_s = 2$ . The results indicate that accuracy of solutions is in accordance with the number of resolution



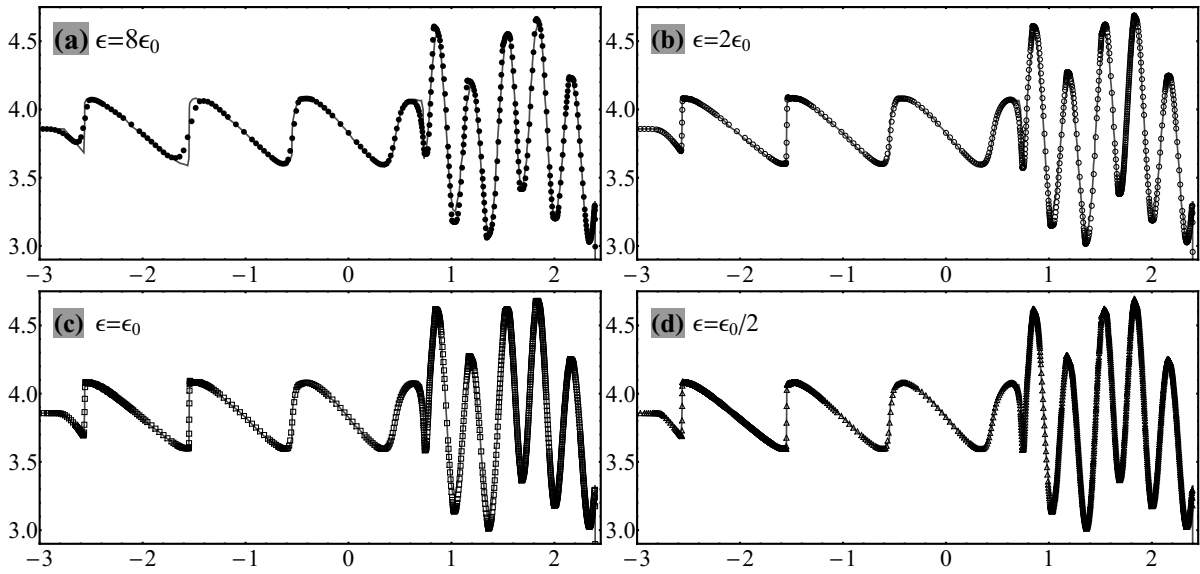


Figure 4.25: Threshold effects on density approximations; it is assumed that  $J_{max} = 13$ ,  $N_c = 1$ , and  $N_s = 2$ .

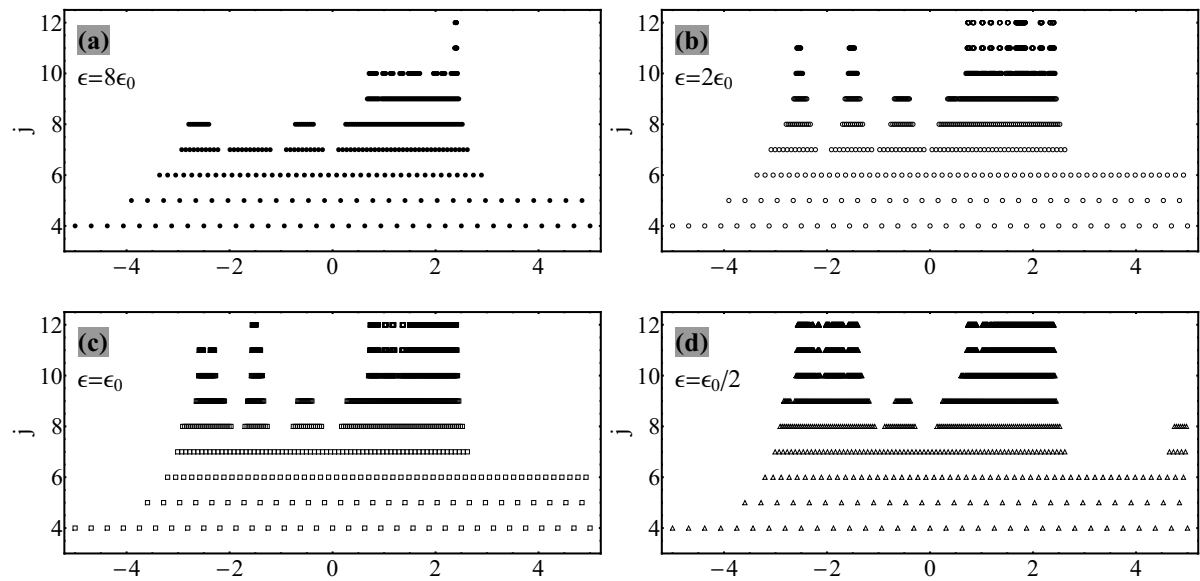


Figure 4.26: Threshold effects on adaptation; it is assumed that  $J_{max} = 13$ ,  $N_c = 1$ , and  $N_s = 2$ .

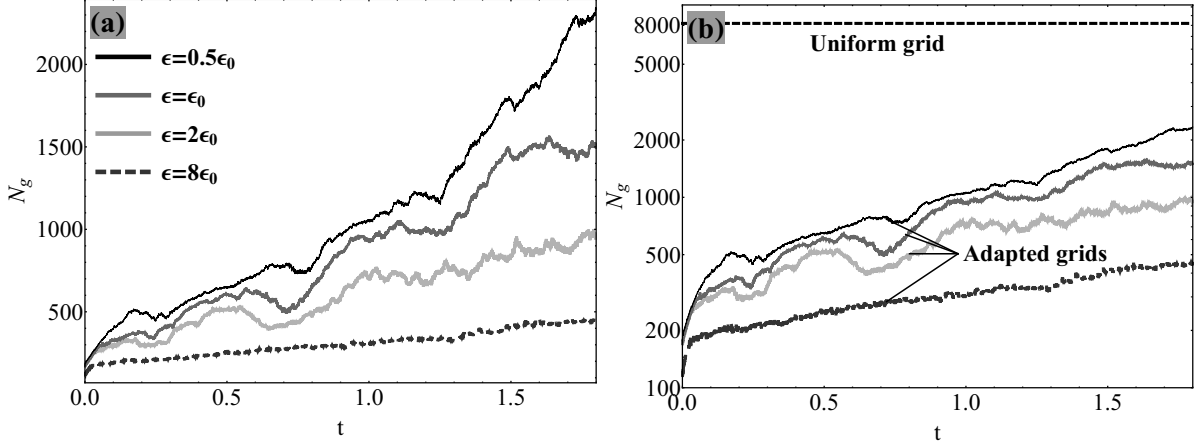


Figure 4.27:  $N_g$  values during simulations for different  $\epsilon$  values; it is assumed  $\epsilon_0 = 5 \times 10^{-4}$  and  $J_{max} = 13$ ; a) adapted grids, b) comparison of  $N_g$  values between the uniform and the adapted grids.

levels.

**The ideal magneto-hydrodynamics problem (MHD)** The conservation form of the MHD system for 1-D flow along x-direction is:

$$\mathbf{u}_t + \mathbf{F}_x = \mathbf{0}, \quad (4.94)$$

where vectors  $\mathbf{u}$  and  $\mathbf{F}$  denote vector of the state variable and the flux function, respectively; they are defined as:

$$\begin{aligned} \mathbf{u} &= \{\rho, \rho v_x, \rho v_y, \rho v_z, B_y, B_z\}^T, \\ \mathbf{F} &= \{\rho v_x, \rho v_x^2 + p^* - B_x^2, \rho v_x v_y - B_x B_y, \rho v_x v_z - B_x B_z, B_y v_x - B_x v_y, \\ &\quad B_z v_x - B_x v_z, (E + p^*)v_x, B_x(B_x v_x + B_y v_y + B_z v_z)\}^T, \end{aligned} \quad (4.95)$$

where:  $v_x, v_y$  and  $v_z$  are velocity components in the  $x, y$  and  $z$  directions, respectively;  $\mathbf{B}$  denotes the magnetic field with components  $B_x, B_y$  and  $B_z$ ; and  $\rho$  is the density. The parameters  $p^*$  and  $E$  are the total (full) pressure and total energy, respectively; they are:

$$\begin{aligned} p^* &= p + \frac{1}{2}(B_x^2 + B_y^2 + B_z^2), \\ E &= \frac{1}{2}(v_x^2 + v_y^2 + v_z^2) + \frac{p}{\gamma - 1} + \frac{1}{2}(B_x^2 + B_y^2 + B_z^2), \end{aligned} \quad (4.96)$$

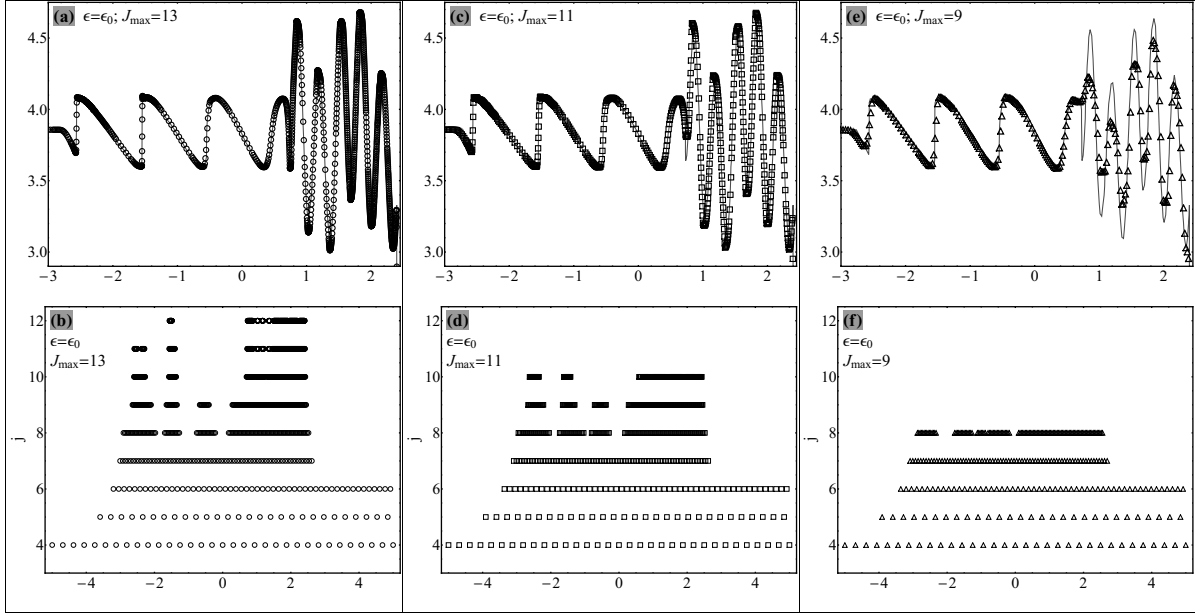


Figure 4.28: Threshold effects on solutions and adaptations; it is assumed that  $J_{min} = 5$ ,  $\epsilon = \epsilon_0$ ,  $N_c = 1$ , and  $N_s = 2$ .

where:  $p$  denotes pressure and  $\gamma$  is the ratio of specific heats. The initial condition is considered to be a Riemann problem. The assumed case is the Brio and Wu shock tube problem [176]:

$$\{\rho, v_x, v_y, v_z, B_y, B_z, E\} = \begin{cases} \{1, 0, 0, 0, 1, 0, 1\}, & \text{if } x < 0.5, \\ \{0.125, 0, 0, 0, -1, 0, 0.1\}, & \text{if } x > 0.5. \end{cases} \quad (4.97)$$

with parameters:  $B_x = 0.75$  and  $\gamma = 2$ .

This system is a standard problem in the MHD analysis to control stability and accuracy of numerical methods. In solution of this system, there are compound waves which are combination of waves: 1) slow; 2) fast; 3) Alfvén<sup>6</sup>. This reveals complexity of the MHD systems in comparison with the hydrodynamic case: the latter problem has only one type of wave; i.e., the sound wave.

In the numerical simulations, it is assumed that:  $J_{max} \in \{10, 11, 12, 13\}$ ,  $J_{min} = 5$ ,  $N_c = 1$ ,  $N_s = 2$ ,  $\epsilon = 5 \times 10^{-4}$  and  $\theta = 2$  (flux limiter parameter). For case  $J_{max} = 11$  (i.e., the finest sampling step is  $dx = 1/2^{11}$ ) the temporal time step is  $dt = 0.0003$  and

<sup>6</sup>In the MHD systems there are three types of waves: 1) fast magneto-sonic (or magneto-acoustic) waves known as fast waves; 2) slow magneto-sonic waves known as slow waves; 3) Alfvén waves. The fast and slow waves are compressive and the Alfvén waves are shear (transverse) waves.

for other values of  $J_{max}$ , the  $dt$  is chosen in a way that the CFL number remains equal to those of case  $J_{max} = 11$ .

The results are shown in Figure 4.29 for different  $J_{max}$  values at  $t = 0.12$ ; it is clear that the adapted points concentrate properly around high gradient zones in different resolutions.

In the Brio-Wu problem, the following components can be detected, from left to right (see Figure 4.29): 1) A fast rarefaction (moving to the left); 2) A slow compound wave including of a slow rarefaction and a slow shock (propagating to the left); 3) A contact discontinuity (propagating to the right side); 4) A slow shock, and 5) a fast rarefaction (the cases 4 & 5 move to the right)<sup>7</sup>.

To investigate these results in more detail, zoomed in solutions of Figure 4.29 are presented in Figure 4.30; the results confirm that: 1) both the Gibb's (dispersion) and dissipation phenomena decrease generally as the number of resolution level increases; 2), increasing of  $J_{max}$  values leads to more localization of spurious oscillations (Figure 4.30(c)); 3) the numerical solutions remain stable despite of existing some non-physical fluctuations and this is because of the TVD feature of the KT scheme.

The  $N_g$  values of adaptive simulations for different  $J_{max}$  values and corresponding number of uniform grids are presented in Figure 4.31; in Figure 4.31(b) scale of presentation is logarithmic. Due to few numbers of localization,  $N_g$  values of adapted grids alter slightly by increasing the  $J_{max}$  values.

#### 4.4.1 1-D second order hyperbolic PDEs

**Wave propagation problems in nonlinear-elastic bars** To solve second order hyperbolic PDEs by the high resolution schemes, it is necessary at first to rewrite them as a first order hyperbolic system. In this context, the 1-D equation of stress wave can be

---

<sup>7</sup>*Shock* and *rarefaction* waves are two types of nonlinear waves arise from abrupt changing in the pressure: a shock wave compresses a domain while a rarefaction wave expands a medium. A *contact-discontinuity* is a surface separating two neighbor media with different densities and temperatures. This surface, however, is in pressure equilibrium and so there is no flow across it

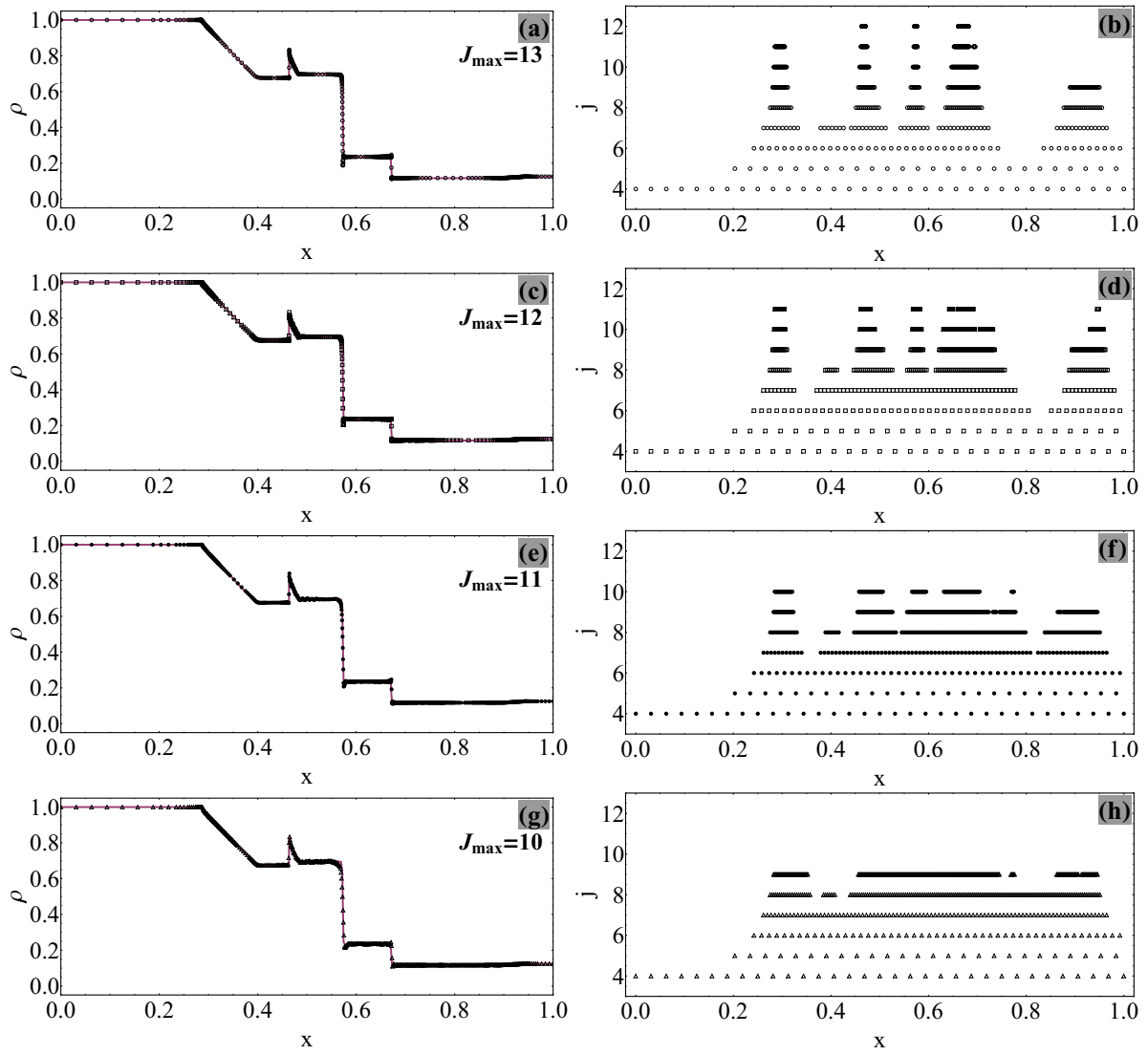


Figure 4.29: Adaptive solutions and corresponding distribution of adapted points in different resolutions for various  $J_{max}$  values at  $t = 0.12$ . In all calculations, it is assumed that:  $N_c = 1$  and  $N_s = 2$ .

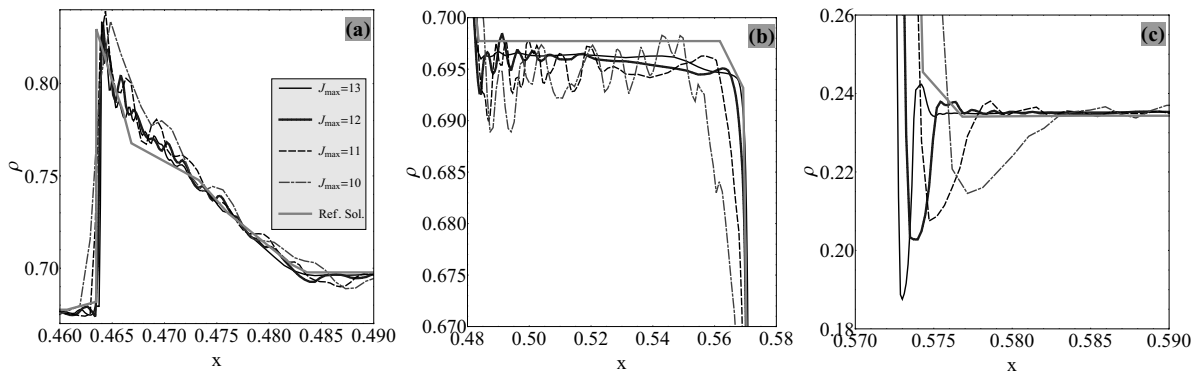


Figure 4.30: Effect of the  $J_{max}$  values regarding dispersion and dissipation phenomena; the illustrations are from zoomed in solutions at  $t = 0.12$ .

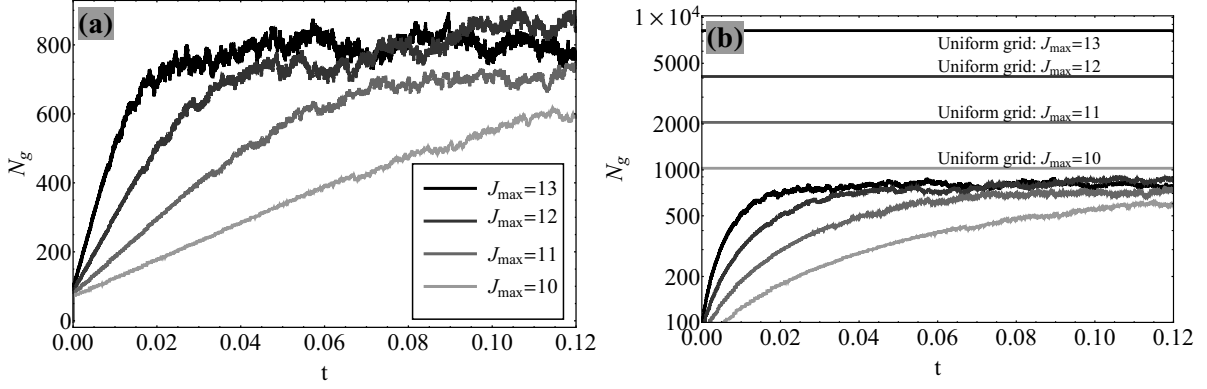


Figure 4.31: Number of grid points used in adaptive solutions for different  $J_{max}$  values; a) adapted grids, b) comparison of  $N_g$  values between the uniform grids with different finest resolutions and corresponding adapted ones.

represented as [126]:

$$\begin{aligned}
 \text{PDEs : } & \begin{cases} \bar{\varepsilon}_t - u_x = 0, \\ (\rho(x)u)_t - \sigma_x = 0, \end{cases} \\
 \text{BCs : } & u(x = 0, t) = \begin{cases} -0.2(1 + \cos(\pi(t - 30)/30)), & \text{if } t \leq 60, \\ 0, & \text{if } t > 60, \end{cases} \\
 & \sigma(x = 0, t) = 0,
 \end{aligned} \tag{4.98}$$

where  $u$ ,  $\bar{\varepsilon}$ ,  $\sigma$  and  $\rho$  denote velocity, strain, stress and density, respectively. The relationship of stress and strain is considered to be:  $\sigma \equiv \sigma(K(x), \bar{\varepsilon}) = K(x)\bar{\varepsilon} + \beta K(x)^2 \bar{\varepsilon}^2$ ; wherein  $K(x)$  denotes the bulk modulus of compressibility and  $\beta$  is a constant. In the numerical simulations it is assumed that:  $\rho(x) = 2$ ;  $K(x) = 2$ ;  $J_{max} = 11$ ;  $J_{min} = 5$ ;  $N_c = 1$ ;  $N_s = 2$ ; the threshold values are:  $\epsilon \in \{10^{-4}, 10^{-5}, 10^{-6}\}$ . To have an elastic-nonlinear wave propagation problem, it is assumed that:  $\beta = 0.3$ .

For using the (KT) central scheme, it is essential to evaluate the maximum speed of propagation at cell edges as:

$a_{j+1/2} = \max\{|a_{j+1/2}^R|, |a_{j+1/2}^L|\}$ , where  $a_{j+1/2}^{R/L} = \pm \sqrt{(d\sigma/d\bar{\varepsilon})/\rho(x)}$  and  $(d\sigma/d\bar{\varepsilon})^{R/L} = K(x) + 2\beta K(x)^2 \bar{\varepsilon}^{R/L}$ . In these formulations, parameter  $\bar{\varepsilon}^{R/L}$  denotes the reconstructed strain values at the left ( $\bar{\varepsilon}^L$ ) or right ( $\bar{\varepsilon}^R$ ) side of cell edges of spatial location  $x_j + \Delta x/2$  (in uniform cell cases).

The  $N_g$  values for different threshold values are shown in Figure 4.32. Numerical and reference solutions [126], and corresponding adapted grid points are presented in Figure 4.33 at  $t = 80$ ,  $t = 160$  and  $t = 240$  for threshold value  $\epsilon = 10^{-6}$ . As the system is non-

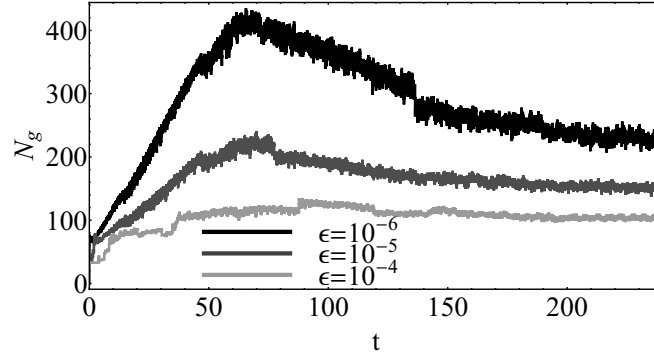


Figure 4.32: Number of grid points ( $N_g$ ) used in the adaptation procedure for different threshold values.

linear, a discontinuous front develops during propagation; it is evident that the adapted points are concentrating in the vicinity of both discontinuities and high-gradient regions.

### 1-D first order advection-diffusion problems

In this subsection, some problems having advection-diffusion feature will be presented.

**A Burgers'-type equation with saturating dissipation equation** This system has the following differential form [177]:

$$u_t + (u^2)_x = \left( \frac{u_x}{\sqrt{1 + u_x^2}} \right)_x. \quad (4.99)$$

In this example the Riemann initial condition is considered as:

$$u(x, t = 0) = \begin{cases} 1/2, & x < 0, \\ -1/2, & x > 0. \end{cases} \quad (4.100)$$

In the numerical simulation, it is assumed that:  $J_{max} = 11$ ,  $J_{min} = 5$ ,  $N_c = 1$ ,  $N_s = 2$ ,  $\epsilon = 6 \times 10^{-2}$ ,  $\theta = 2$  and  $dt = 0.0001$ .

Due to the formation of a subshock, located at  $x = 0$ , it was shown that the numerical solution of this system is quite a challenging problem.

The numerical results and corresponding number of grid points ( $N_g$ ) during adaptive simulation are shown in Figures 4.34 and 4.35, respectively. In Figure 4.34(a), the solid line and circles are the asymptotic [177] and adaptive solutions, respectively. Figure 4.34(b) illustrates distribution of adapted grid points in different levels of resolution. In Figure 4.35(b),  $N_g$  value of the uniform case is compared with those of adapted case in

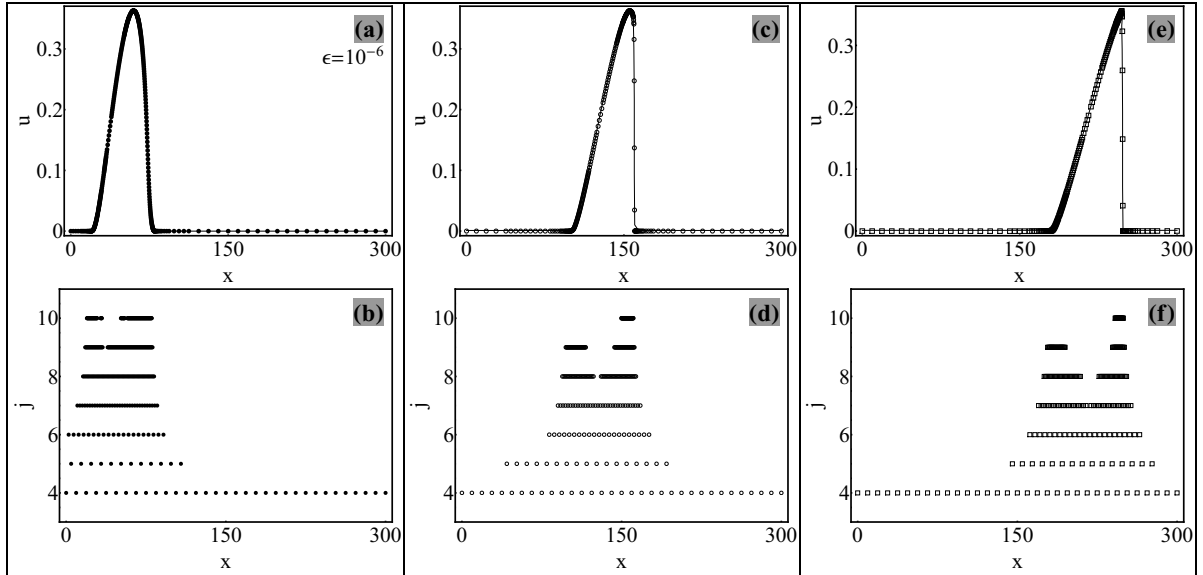


Figure 4.33: Numerical solutions of the non-linear wave propagation problem and corresponding adapted grids where  $\epsilon = 10^{-6}$  and  $\theta = 2$ ; a-b)  $t = 80$ ; c-d)  $t = 160$ ; e-f)  $t = 240$ . In the top row, solid lines are reference solutions, and shapes are numerical ones. In all calculations it is assumed that:  $N_c = 1$  and  $N_s = 2$ .

logarithmic scale. From Figures 4.34 and 4.35, it is clear that with a few number of grid points (less than 70 points) an excellent accuracy is obtained; to achieve such accuracy by uniform cells, more than 800 points are necessary [177].

**The Buckley-Leverett equation without the diffusion term** In this case, the governing equation is a hyperbolic PDE with the flux:

$$F(u) = \frac{u^2}{4u^2 + (1-u)^2}, \quad \Omega \in [-1, 1] \times [0, T], \quad (4.101)$$

with the Riemann initial condition:

$$u(x, t = 0) = \begin{cases} 1, & -1/2 \leq x \leq 0, \\ 0, & \text{elsewhere.} \end{cases} \quad (4.102)$$

In numerical simulations, it is assumed:  $\epsilon_0 = 10^{-3}$ ,  $\theta = 2$ ,  $J_{max} = 11$ ,  $J_{min} = 5$ , and  $dt = 0.0002$ . The numerical result and corresponding adapted grid points are illustrated in Figure 4.36 at  $t = 0.4$ . It is evident that adapted solution is successfully done.



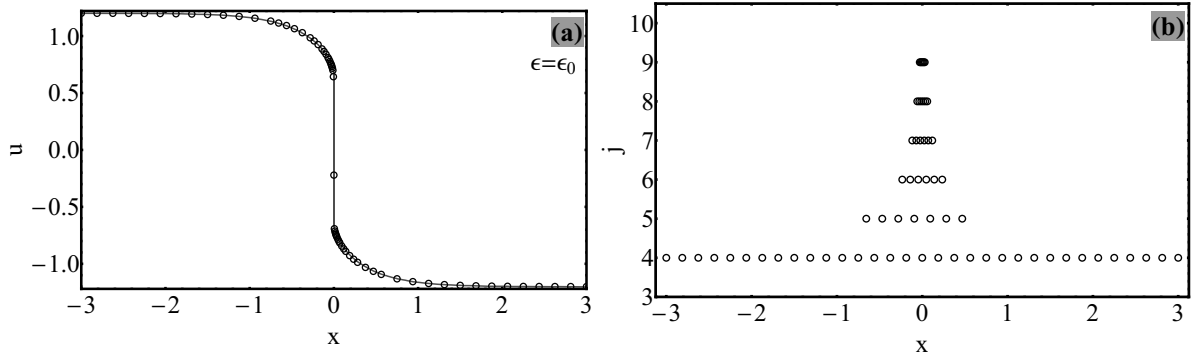


Figure 4.34: Adapted solution for the threshold value  $\epsilon = \epsilon_0 = 6 \times 10^{-2}$  at  $t = 1.5$ . Here, the solid line denotes the asymptotic solution. In all calculations it is assumed that:  $N_c = 1$  and  $N_s = 2$ .

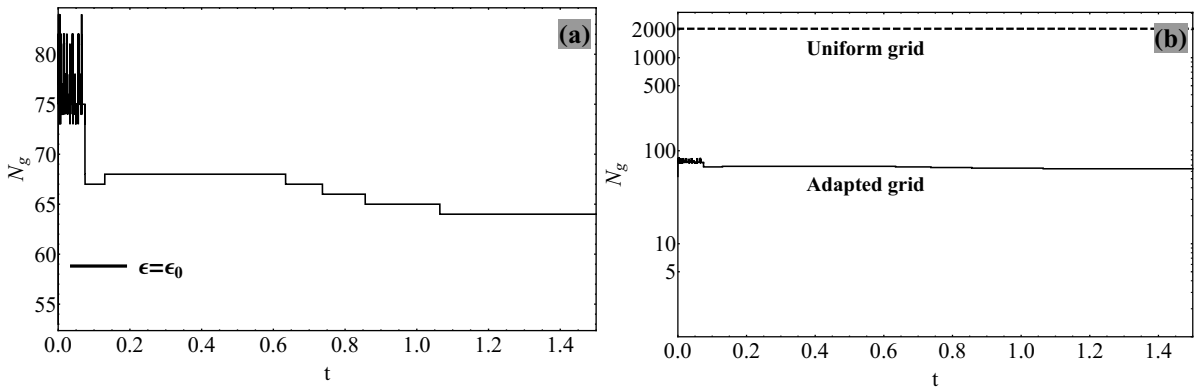


Figure 4.35: Number of grid points ( $N_g$ ) used in the adaptation procedure for the threshold value  $\epsilon = \epsilon_0 = 6 \times 10^{-2}$ ; a) adapted grids, b) comparison of  $N_g$  values between the uniform and the corresponding adapted grids.

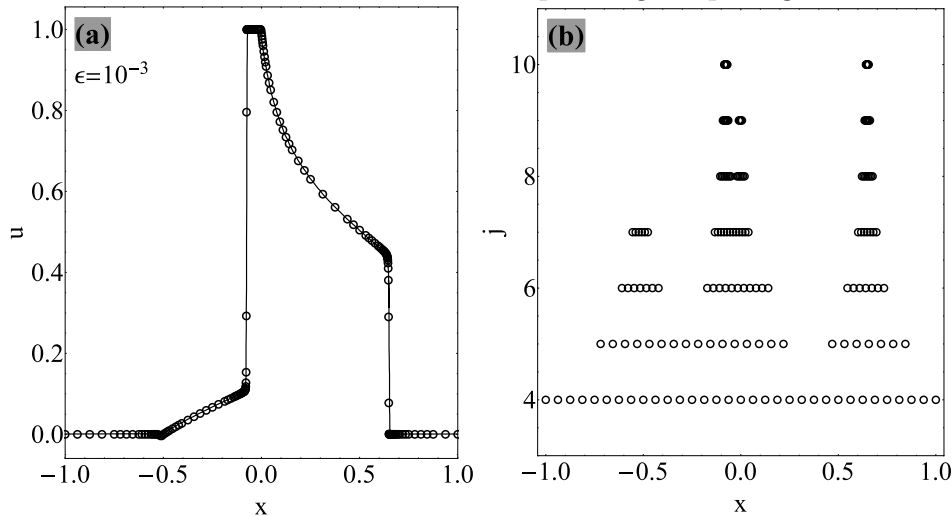


Figure 4.36: The solution at  $t = 0.4$  and corresponding adapted points; a) the exact and numerical results, b) distribution of adapted points in different resolutions. In (a), the solid line and hollow circle illustrate exact and numerical solutions, respectively.

**The diffusive Buckley-Leverett equation with and without gravitational effects**

The governing equation of the Buckley-Leverett system is [22]:

$$u_t + F(u)_x = \alpha Q(u, u_x)_x,$$

where:

$$Q(u, u_x) = (v(u)u_x),$$

$$\alpha v(u) \geq 0,$$

$$F(u) = \frac{u^2}{u^2 + (1-u)^2} : \text{ for a system without gravitational effects,}$$

$$F(u) = \frac{u^2}{u^2 + (1-u)^2} (1 - 5(1-u)^2) : \text{ for a system with gravitational effects.}$$

(4.103)

Here it is assumed:  $v(u) = 4u(1-u)$ ; this means that  $v(u)$  vanishes at  $u = 0$  and  $u = 1$ . The parameter  $\alpha$  denotes a constant and in this example the considered value is  $\alpha = 0.01$ . The assumed initial condition is of the Riemann type, as:

$$u(x, t = 0) = \begin{cases} 0, & 0 \leq x < (1 - (1/\sqrt{2})), \\ -1/2, & (1 - (1/\sqrt{2})) < x \leq 1. \end{cases} \quad (4.104)$$

Numerical results are presented in Figure 4.37 at  $t = 0.2$ ; in calculations it is assumed:  $J_{max} = 11$ ,  $J_{min} = 5$ ,  $N_c = 1$ ,  $N_s = 2$ ,  $\epsilon = 10^{-4}$ ,  $\theta = 2$  and  $dt = 0.00001$ . There, in Fig 4.37(a), solid lines and hollow shapes correspond to reference solutions [22] and adaptive results, respectively. Distribution of adapted points in different resolutions are presented in Figures 4.37(b) and 4.37(c). There figures (b) and (c) are for cases with and without gravity effects. It is obvious that adapted points concentrate around high gradient zones as expected.

Variation of  $N_g$  values during adaptive simulations is shown in Figure 4.38; it is clear that among  $2^{11} + 1$  uniform points in the finest resolution, less than 150 points are needed for the adaptive simulations (Figure 4.38(b)).

**The glacier growth model** In this example, the Glacier growth problem with two different source terms will be presented (considering the source term ends to a non-homogeneous equation). The governing convection-diffusion equation is [22]:

$$h_t + F(h)_x = \alpha(v(h)h)_x + S(x, t, h), \quad (4.105)$$

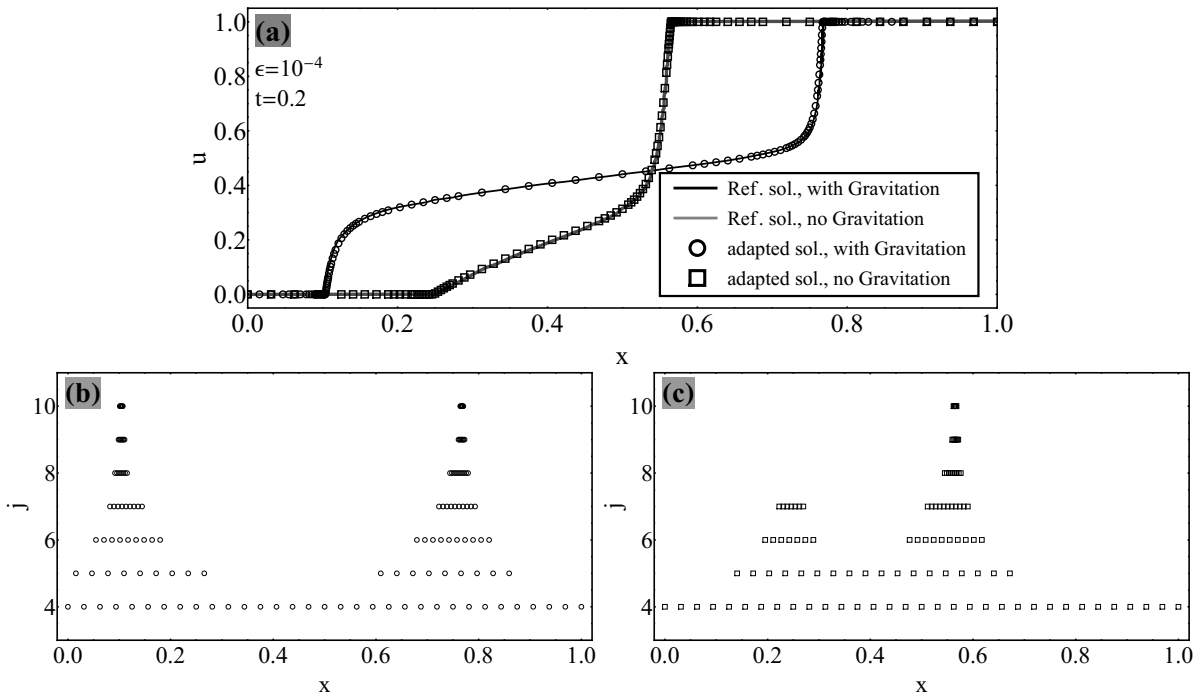


Figure 4.37: Adaptive simulation of the Buckley-Leverett equation with and without the gravitational effects; a) adaptive solutions; b) distribution of adapted points in different resolution levels for the case with the gravity effects; c) distribution of adapted points for the case without the gravity effects. In all calculations it is assumed that:  $N_c = 1$  and  $N_s = 2$ .

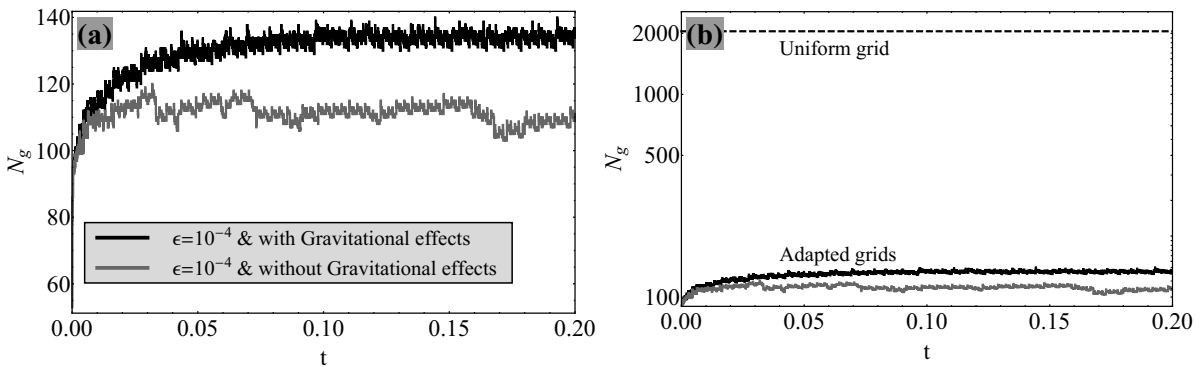


Figure 4.38: Number of grid points,  $N_g$ , used for adaptive simulations; a) adapted grids, b) comparison of  $N_g$  values between the uniform grid and corresponding adapted ones.

where  $h = h(x, t)$  denotes height of a glacier above a flat mountain. In this example, it is assumed:  $F(h) = (h + 3h^6)/4$ ;  $v(h) = 3h^6$ ;  $\alpha = 0.01$ ;  $S(x, t, h)$  denotes the source term with definition:

$$S(x, t, h) = \begin{cases} S_0(x, t), & h(x, t) > 0, \\ \max(S_0(x, t), 0), & h(x, t) = 0. \end{cases} \quad (4.106)$$

In the following, this system with two different source functions,  $S_0(x, t)$  will be presented.

**Case 1** Let  $S_0(x, t)$  and  $h(x, t = 0)$  be:

$$S_0(x, t) = \begin{cases} 0, & x < -0.4, \\ 0.5(x + 4), & -0.4 \leq x \leq 0.2, \\ -0.5x, & x > 0.2, \end{cases} \quad \& \quad h(x, t = 0) = \begin{cases} 1, & x < 0, \\ 0, & x > 0. \end{cases} \quad (4.107)$$

The adapted numerical solutions and corresponding distribution of adapted points in different resolutions are presented in Figure 4.39. There, the shapes are the adapted solutions and the solid lines are the reference ones obtained by the KT second-order semi-discrete scheme, SD2, on 800 uniform cells [22]. The results confirm that the adapted points concentrate properly around high gradient regions. In numerical computations, it is assumed:  $J_{max} = 11$ ,  $J_{min} = 5$ ,  $N_c = 1$ ,  $N_s = 2$ ,  $\epsilon = 10^{-4}$ ,  $\theta = 2$  and  $dt = 0.00003$ . This simple problem describes a right propagating melting ice sheet without considering seasonal variations [22].

Number of grid points used in adaptation procedure is presented in Figure 4.40; in Figure 4.40(b) number of uniform grid and corresponding adapted one are presented. It is clear that due to melting in time the number of adapted points reduces gradually.

**Case 2** Here it is assumed:

$$S_0(x, t) = \begin{cases} 0, & x \leq -5, \\ -0.01x + 0.05 \sin(2\pi t), & x > -5, \end{cases} \quad (4.108)$$

$$h(x, t = 0) = 0.$$

In this source definition, seasonal variations are considered by the term  $0.05 \sin(2\pi t)$ .

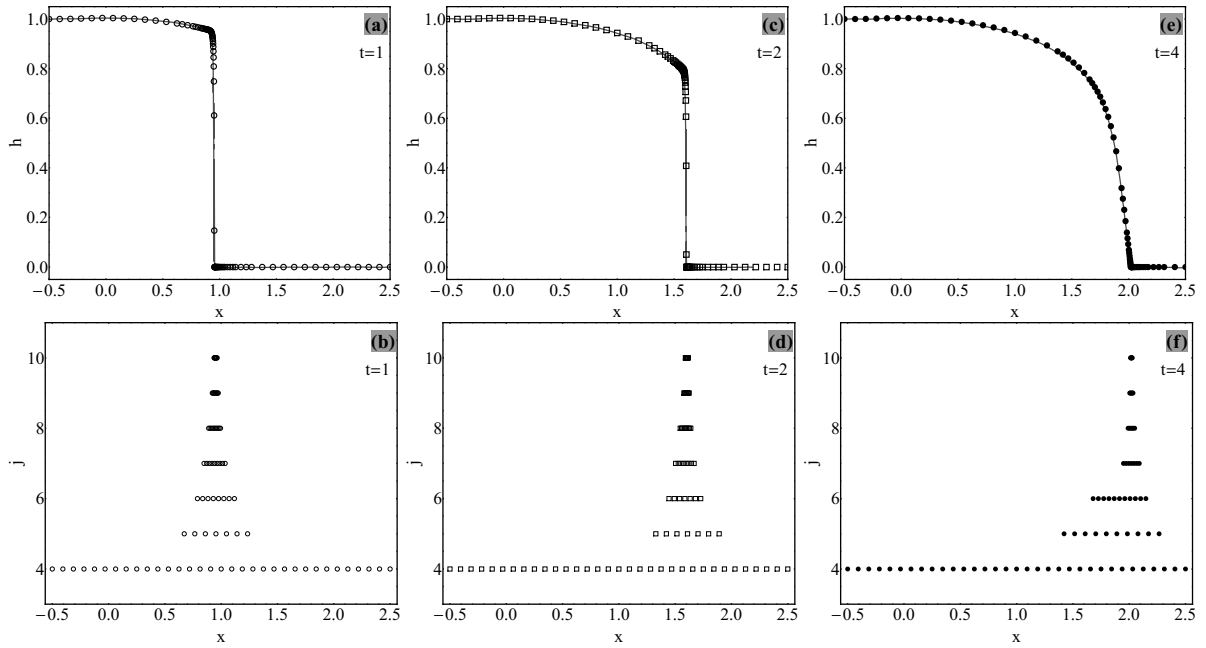


Figure 4.39: Right propagating ice shield simulated by the adaptive method and corresponding distribution of adapted points in different resolution levels; shapes are the adaptive solutions and the solid lines are the reference ones obtained by SD2 method with 800 uniform cells.

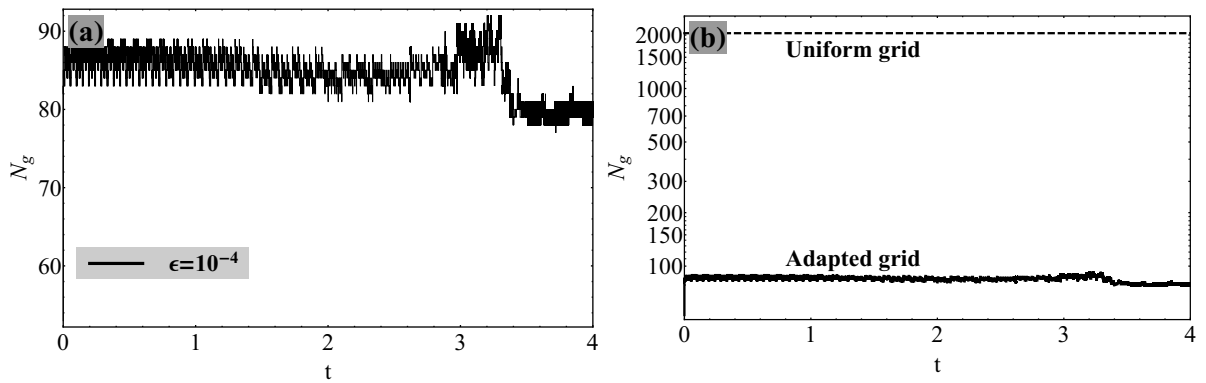


Figure 4.40: Number of grid points  $N_g$  used for adaptive simulation; a) adapted grids, b) comparison of  $N_g$  values between the uniform grid and corresponding adapted one.

The numerical results and the corresponding distribution of adapted grid points in different resolutions are presented in Figure 4.41; for these results it is assumed:  $J_{max} = 11$ ,  $J_{min} = 5$ ,  $N_c = 1$ ,  $N_s = 2$ ,  $\epsilon = 10^{-5}$ ,  $\theta = 2$  and  $dt = 0.001$ . In these figures, the solid lines are the reference solutions obtained by SD2 method with 600 uniform cells [22], and shapes are numerical results.

Comparison of solutions obtained by different thresholds  $\epsilon \in \{10^{-4}, 10^{-5}\}$  are presented in Figure 4.42 at  $t = 15$ . It is obvious that by using smaller values of the threshold, local behaviors of solutions are properly captured. This problem reveals importance of very fine scale effects: the large scale responses are affected by fine scale features. This is mainly due to multiscale feature of the operators in this example. In Figure 4.43,  $N_g$  values during time are presented for different threshold values:  $\epsilon \in \{10^{-4}, 10^{-5}, 5 \times 10^{-6}\}$ ; they are compared with corresponding uniform grid in Figure 4.43(b). The results exhibit that smaller threshold values lead to more adapted grid points.

**The shallow water problem (small dam break)** For a body of water with zero slope angles and zero friction coefficient, the governing equation of motion is [5]:

$$\frac{\partial}{\partial t} \begin{Bmatrix} h \\ hv \end{Bmatrix} + \frac{\partial}{\partial x} \begin{Bmatrix} hv \\ hv^2 + \frac{1}{2}gh^2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}, \quad (4.109)$$

where  $h \equiv h(t, x)$  is the height of water,  $v \equiv v(t, x)$  denotes the velocity, and  $g$  is the gravity acceleration. The problem is a Riemann problem with the following initial conditions,

$$h(0, x) = \begin{cases} 2, & x < 0, \\ 1, & x \geq 0, \end{cases}, \quad v(0, x) = \begin{cases} 0, & x < 0, \\ 0, & x \geq 0. \end{cases} \quad (4.110)$$

It is also assumed that:  $dt = 5 \times 10^{-5}$ ,  $\theta = 2$  and  $\epsilon = \{10^{-3}, 2.5 \times 10^{-4}, 10^{-4}\}$ . The results for three different threshold values at time 1 sec are illustrated in Figure 4.44. The results indicate that by decreasing the threshold value, more points are concentrated around the high gradient zones, and thereby dissipation effect is reduced. Variations of adapted grid points for three different threshold values are illustrated in Figure 4.45. It is clear that the number of grid points has an inverse proportion with threshold values.

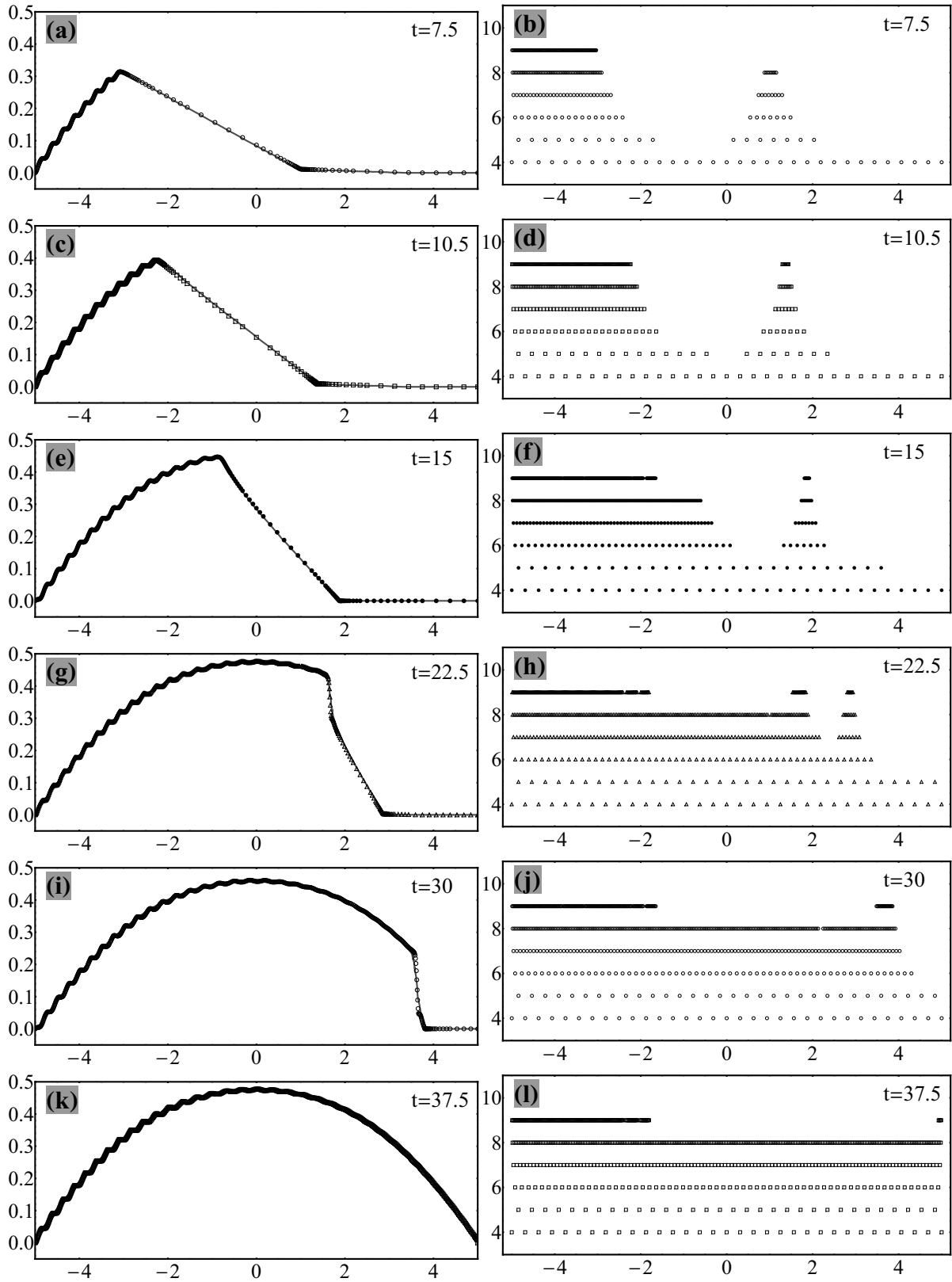


Figure 4.41: Glacier growth; the shapes are numerical solutions and the solid lines are the reference ones; where  $N_c = 1$  and  $N_s = 2$ .

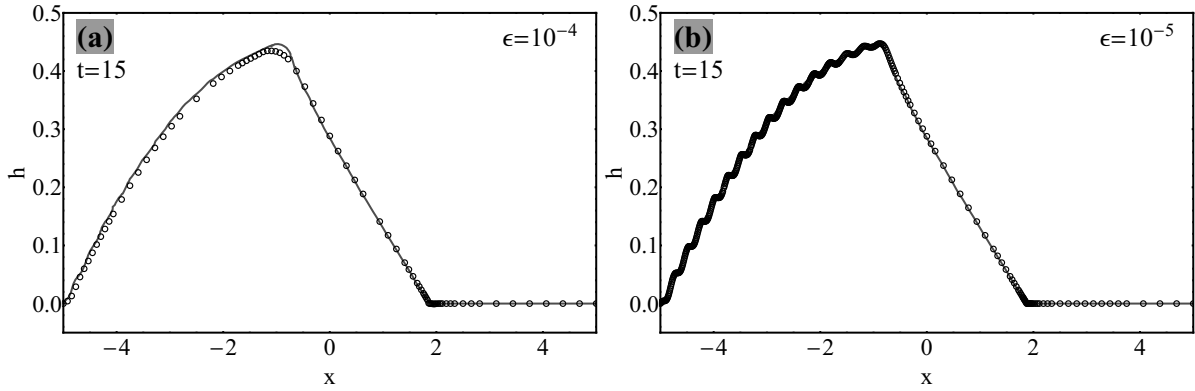


Figure 4.42: Threshold effects on the localized feature detection; hollow circles and solid line are the adaptive and reference solutions, respectively.

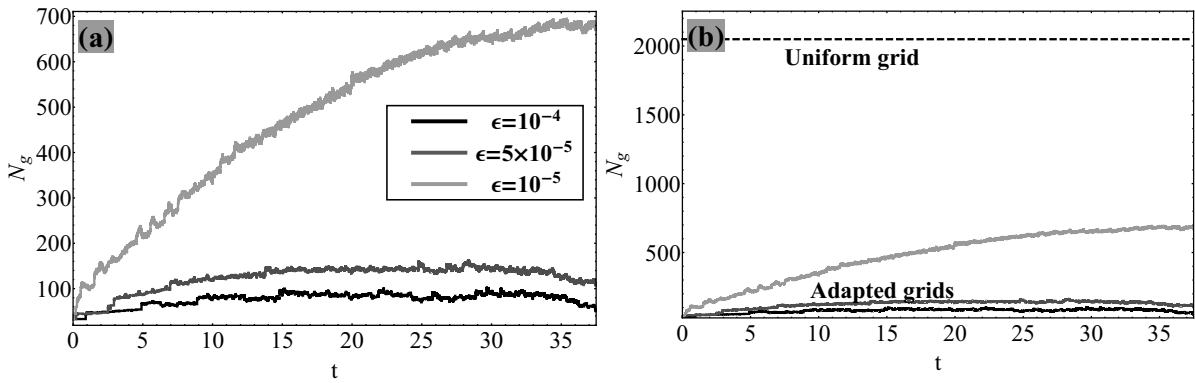


Figure 4.43: Number of grid points during adaptive solutions; a) adapted grids, b) comparison of  $N_g$  values between the uniform grid and corresponding adapted ones.



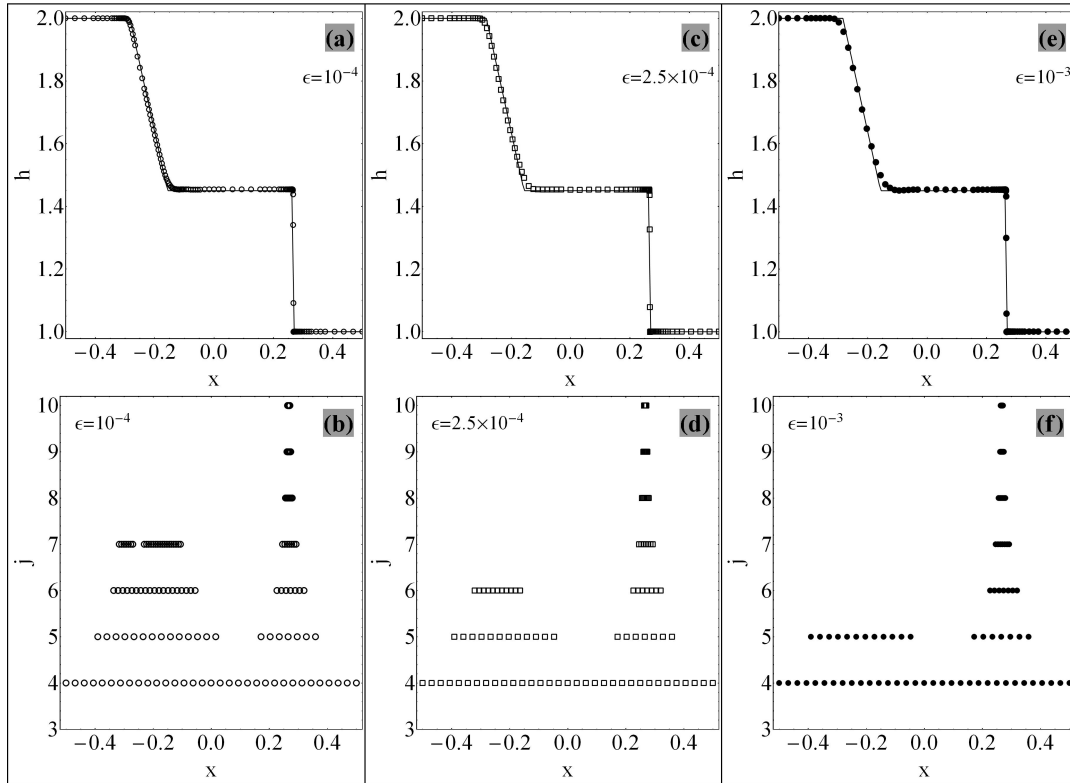


Figure 4.44: Effect of threshold values on adaptive solution and the corresponding grid; a-b)  $\epsilon = 10^{-4}$ ; c-d)  $\epsilon = 2.5 \times 10^{-4}$ ; e-f)  $\epsilon = 10^{-3}$ .

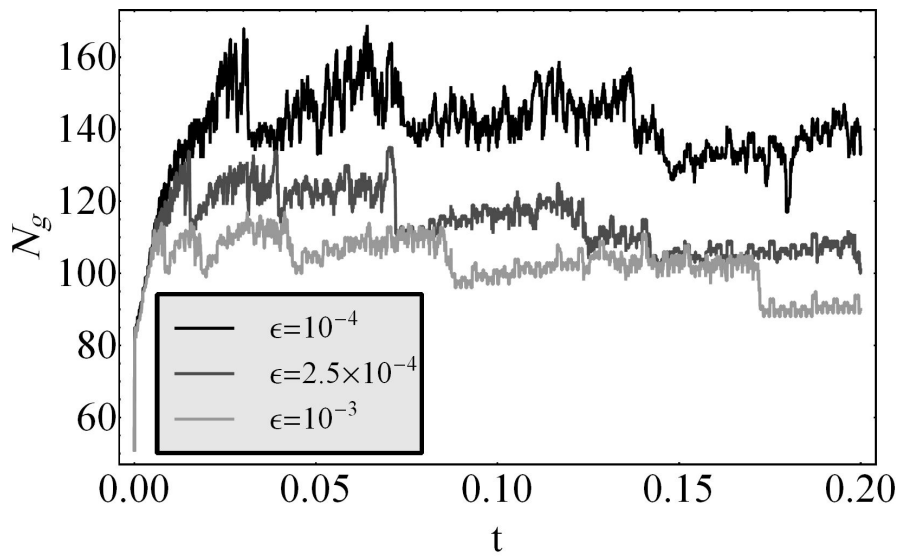


Figure 4.45: Number of grid points ( $N_g$ ) used in the adaptation procedure for different threshold values.

## 2-D examples

In the following, numerical solution of some 2-D hyperbolic and hyperbolic-parabolic systems will be presented.

### The 2-D scalar convection-diffusion Burgers' equation

**Case 1** The first example is a 1-D hyperbolic-parabolic problem solved in a 2-D domain; the 2-D scalar convection-diffusion Burgers'-type equation is [22]:

$$u_t + (u^2)_x + (u^2)_y = \varepsilon'(v(u)u_x)_x + \varepsilon'(v(u)u_y)_y. \quad (4.111)$$

In this example, it is assumed  $\varepsilon' = 0.1$  and

$$v(u) = \begin{cases} 0, & |u| \leq 0.25, \\ 1, & |u| > 0.25. \end{cases} \quad (4.112)$$

The discontinuous function  $v(u)$  implies that Eq. 4.111 has hyperbolic nature for the range  $u \in [-0.25, 0.25]$ . The considered initial condition is:

$$u(x, y, t = 0) = \begin{cases} 1, & -(1/\sqrt{2}) - 0.4 < x < -(1/\sqrt{2}) + 0.4, \\ -1, & +(1/\sqrt{2}) - 0.4 < x < +(1/\sqrt{2}) + 0.4, \\ 0, & \text{otherwise.} \end{cases} \quad (4.113)$$

The spatial square  $\Omega \in [-2, 2] \times [-2, 2]$  is considered as the computational domain. In numerical calculations, it is assumed that:  $\epsilon = 0.25 \times 10^{-3}$  (threshold),  $dt = 0.0002$ ,  $\theta = 2$  (the MINMOD smoothing parameter),  $J_{max} = 7$ ,  $J_{min} = 5$ ,  $N_c = 2$  and  $N_s = 2$ . In the numerical solution, an adapted grid is updated after each ten time steps.

A snapshot of the 2-D solution with corresponding adapted grid is illustrated in Figure 4.46 at  $t = 0.7$ . This 2-D result is compared with the 1-D solution [22] obtained on 200 uniform cells; see Figure 4.47. It is clear that the two results are in a good agreement.

**Case 2** Here the previous 2-D Burgers' equation is resolved with another initial condition having localized feature. The initial condition is equal to -1 and 1 inside two circles with center locations  $(0.5, 0.5)$  and  $(-0.5, -0.5)$ , respectively; radius of both circles are 0.4. In numerical calculations, it is assumed that:  $\epsilon = 0.25 \times 10^{-3}$  (threshold),  $dt = 0.0002$ ,  $\theta = 2$ ,  $J_{max} = 8$ ,  $J_{min} = 5$ ,  $N_c = 2$ ,  $N_s = 2$  and  $\Omega \in [-1.5, 1.5] \times [-1.5, 1.5]$ .

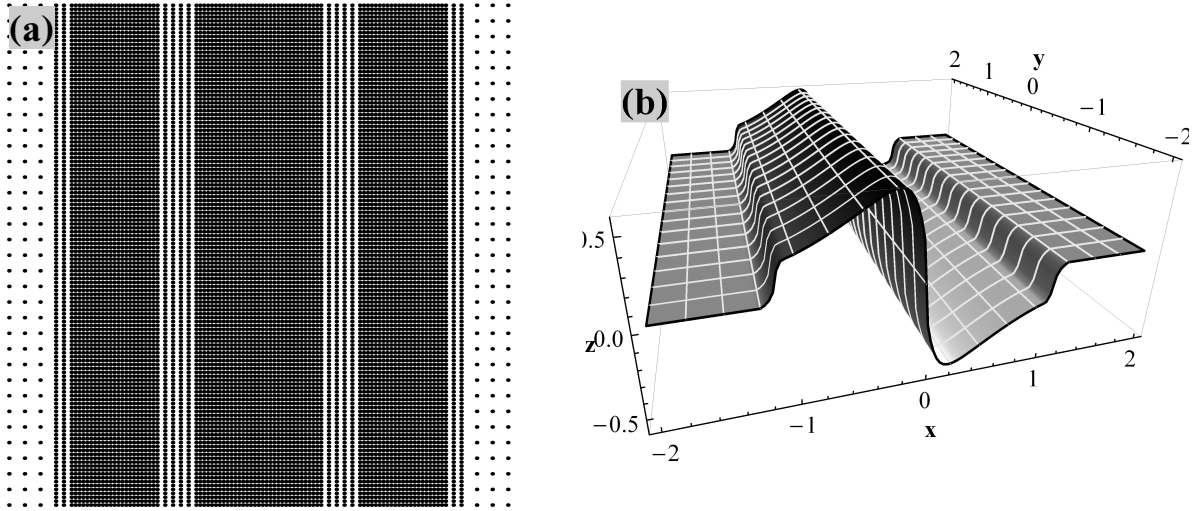


Figure 4.46: 2-D adaptive solution of the Burgers' type equation at  $t = 0.7$ ; a) adapted grid; b) numerical solution. In calculations it is assumed:  $N_c = 2$ ,  $N_s = 2$ ,  $\epsilon = 0.25 \times 10^{-3}$ .

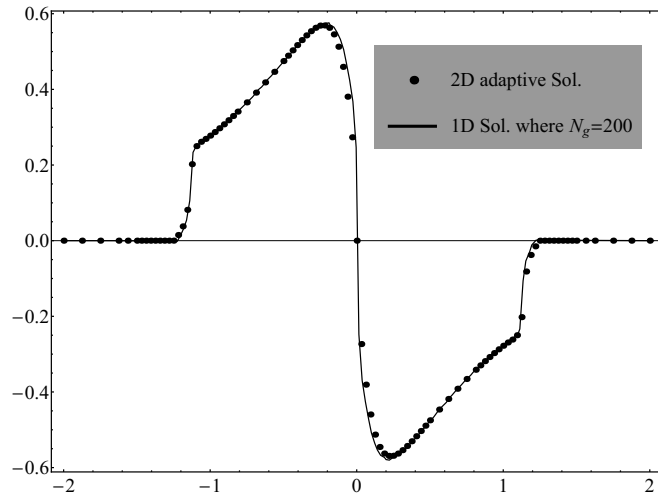


Figure 4.47: Comparison of 2-D adapted solution with 1-D one; the 1-D solution is obtained on a uniform grid of 200 cells.

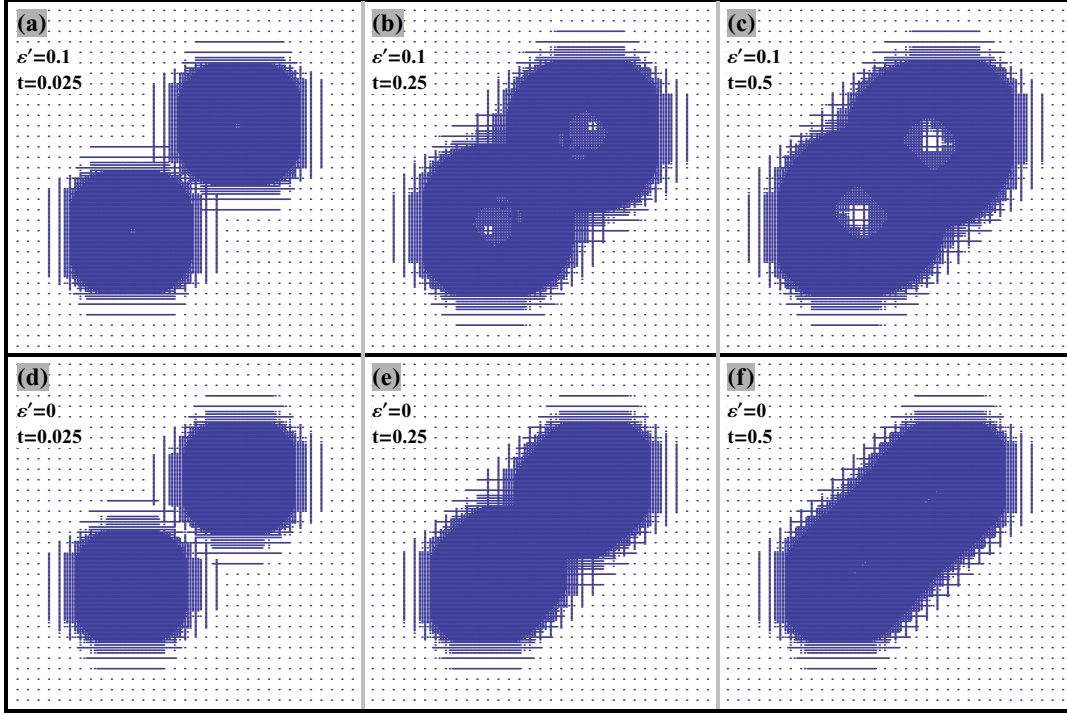


Figure 4.48: Adapted grid points for different  $\varepsilon'$  values; a-c)  $\varepsilon' = 0.1$ ; d-f)  $\varepsilon' = 0$ . In the calculations it is assumed:  $\varepsilon = 0.25 \times 10^{-3}$ ,  $N_c = 2$ ,  $N_s = 2$ .

The results and corresponding adapted points are presented in Figures 4.48 and 4.49, respectively, where  $\varepsilon' \in \{0, 0.1\}$ . For  $\varepsilon' = 0$  and  $\varepsilon' = 0.1$ , Eq. 4.111 turns to be the pure hyperbolic and parabolic-hyperbolic systems, respectively.

It is clear that adapted points properly concentrated around high gradient zones. The  $N_g$  values for different  $\varepsilon'$  values are illustrated in Figure 4.50; they are also compared with those in the uniform case with grid number  $N_g = 2^{J_{max}+1} \times 2^{J_{max}+1}$ . It is evident that due to the diffusion term, more grid points are mobilized in the case  $\varepsilon' = 0.1$ .

**The shallow water problem** In this example a vector form of a 2-D system will be considered; it is the shallow water problem with governing equations:

$$\begin{cases} h_t + (hu)_x + (hv)_y = 0, \\ (hu)_t + (hu^2 + 0.5gh^2)_x + (huv)_y = 0, \\ (hv)_t + (huv)_x + (hv^2 + 0.5gh^2)_y = 0, \end{cases} \quad (4.114)$$

where  $h \equiv h(t, x, y)$  is height of the water;  $u \equiv u(t, x, y)$  and  $v \equiv v(t, x, y)$  denote velocities of water in the  $x$  and  $y$  directions, respectively;  $g$  is the gravitational acceleration. The computational domain is inside the spatial square  $\Omega \in [0, 200] \times [0, 200]$ . The initial

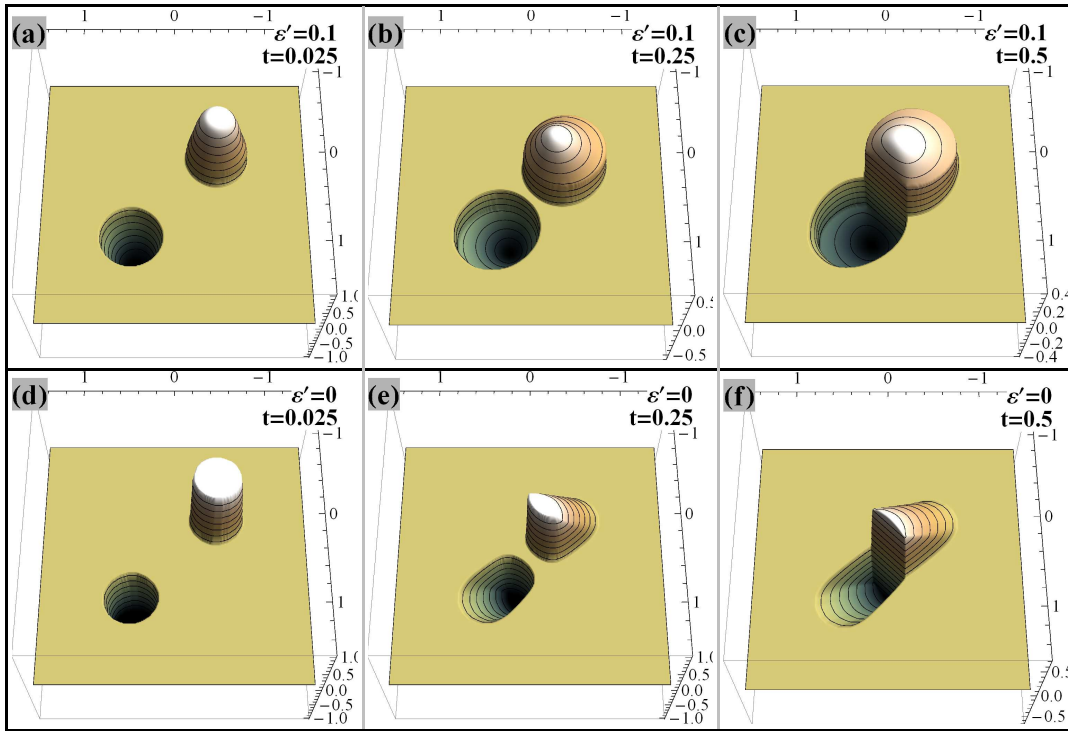


Figure 4.49: Solution of 2-D Burgers' equation for different  $\varepsilon'$  values; a-c)  $\varepsilon' = 0.1$ ; d-f)  $\varepsilon' = 0$ .

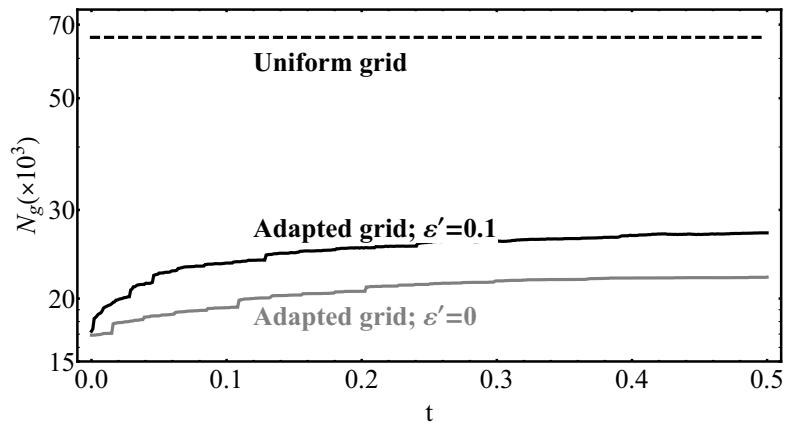


Figure 4.50: Number of grid points used in adaptive solutions for different  $\varepsilon'$  values.

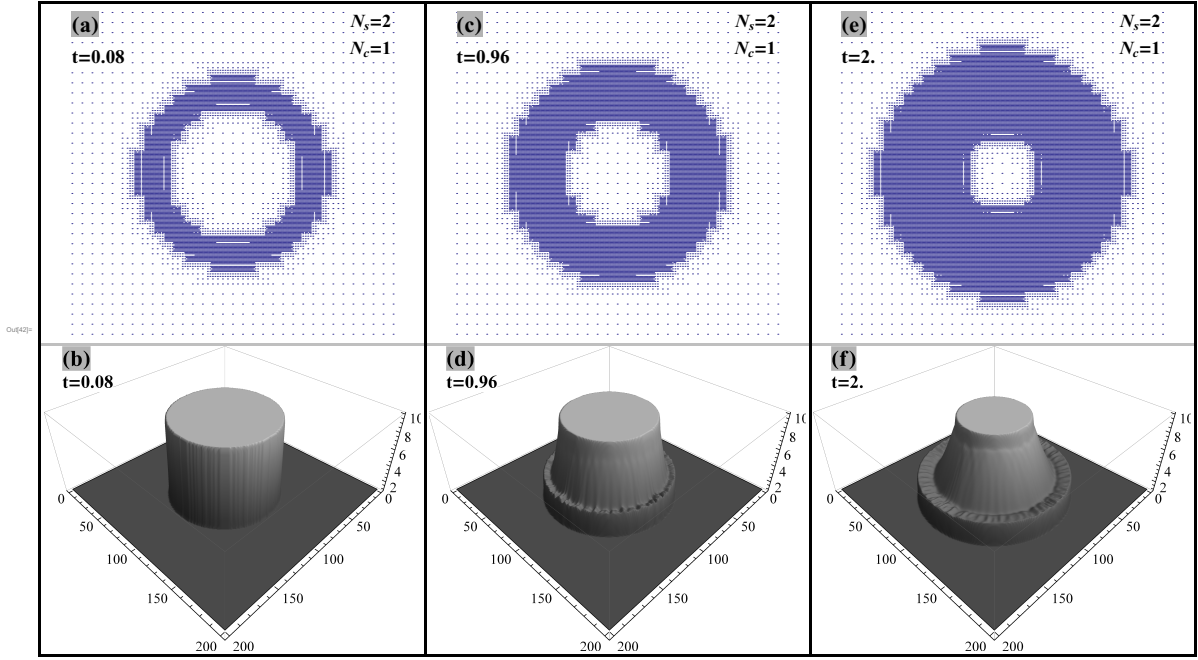


Figure 4.51: Solutions and corresponding adapted grid points of the shallow water problem, where:  $\epsilon = 0.25 \times 10^{-3}$ ,  $N_c = 1$ ,  $N_s = 2$ .

conditions are:

$$h(t = 0, x, y) = \begin{cases} 10, & (x - 100)^2 + (y - 100)^2 \leq 50^2, \\ 1, & (x - 100)^2 + (y - 100)^2 > 50^2, \end{cases} \quad \text{and} \quad u(0, x, y) = v(0, x, y) = 0. \quad (4.115)$$

For numerical simulations, it is assumed that:  $\epsilon = 0.25 \times 10^{-3}$  (threshold),  $dt = 0.004$ ,  $\theta = 2$ ,  $J_{max} = 8$ ,  $J_{min} = 5$ ,  $N_c = 1$  and  $N_s = 2$ ; adaptation procedure is repeated after each 20 time steps. Snapshot of results and corresponding adapted grid points are shown in Figure 4.51; it is evident that the points are concentrated around high gradient zones and both solutions and corresponding adapted grids remain symmetric during time.

# Chapter 5

## Special topics

Regarding second order central high resolution schemes, in this section, some special topics are studied. They are:

1. Central-upwind schemes: for reduction of the numerical dissipation,
2. The concept of the *numerical entropy production*: for checking of numerical *uniqueness*,
3. The *local truncation errors*: to control numerical convergence,
4. Providing other slope limiters for non-uniform grids,
5. The *non-convex* hyperbolic systems and their challenging problems in numerical simulations.

### 5.1 Central and central-upwind second order schemes

Here, the semi-discretized form of a scalar first-order hyperbolic system,  $u_t + F(u)_x = 0$ , will be provided for the KT method [22] and two other improvements of this scheme (to have less numerical dissipation) [23, 24]. As mentioned before (in Eq. (4.14)), by the REP procedure [22] for cell-centered non-uniform cells, the semi-discrete form is:

$$\frac{du_j}{dt} + \frac{F_{j+1/2}^* - F_{j-1/2}^*}{\Delta x_j} = 0, \quad F_{j\pm 1/2}^* := F(u_{j\pm 1/2}^*). \quad (5.1)$$

Other two improvements of the central KT scheme are central-upwind methods, noting here by  $M1$  and  $M2$ . They use two different maximum local propagation speeds for right

and left directions at cell edges  $x_{j+1/2}$ . The speeds are shown by  $a_{j+1/2}^R$  and  $a_{j+1/2}^L$  for right and left directions, respectively. Such speed distinguishing leads to a narrower non-smooth zone around cell edge  $x_{j+1/2}$ , and therefore less dissipative schemes. The scheme  $M2$  is an improvement of the  $M1$  method using a narrower non-smooth zone [23, 24].

**The first less dissipative method: a central-upwind scheme** This central-upwind scheme is the first improvement of the KT central scheme to reduce its numerical dissipation; we denote it by  $M1$ . In this method, propagation directions in cell edges  $x_{i+1/2}$  are important; the boundaries of the non-smooth zone around edge  $i+1/2$  are:  $x_{i+1/2,l}^n = x_{i+1/2} + a_{i+1/2}^L \Delta t$  and  $x_{i+1/2,r}^n = x_{i+1/2} + a_{i+1/2}^R \Delta t$ ; where  $|a_{i+1/2}^L|$  and  $|a_{i+1/2}^R|$  are maximum left and right propagating speeds at  $t = t^n$ , respectively. They are defined as:  $a_{i+1/2}^L = \min \{ \lambda_1(u_{i+1/2}^L), \lambda_1(u_{i+1/2}^R), 0 \}$  and  $a_{i+1/2}^R = \max \{ \lambda_N(u_{i+1/2}^L), \lambda_N(u_{i+1/2}^R), 0 \}$ ; where the set  $\{ \lambda_i \}$  is eigenvalues of the Jacobean  $\partial F(u)/\partial u$  at  $x_{i+1/2}$  in a way that  $\lambda_1 < \lambda_2 < \dots < \lambda_N$ . This distinguishing of the left-right speeds leads to smaller smooth zones and thereby a less dissipative method. The limited slope in the projection step is calculated as the KT scheme; the left ( $\Delta_L$ ) and the right ( $\Delta_R$ ) distances are measured from center of non-smooth zones to center of surrounding smooth regions, Figure 5.1. The limited slope can then be estimated same as the KT scheme.

**The second less dissipative method: a central-upwind scheme** This method is an improvement of the  $M1$  scheme to have less numerical dissipation, referring by  $M2$ . In this method, non-smooth zones are estimated same as the  $M1$  scheme. For limited slope calculations in the projection step, smaller spatial distance is assumed: distance from mid point  $x_{j+1/2}$  from edges  $x_{j+1/2,l}/x_{j+1/2,r}$  (of the non-smooth zone); Figure 5.1 (compare figure (a) with figure (b)). Edges of non-smooth zone from  $x_{j+1/2}$  are measured and then averaged, as:

$$\Delta = \left\{ \left( x_{j+1/2,r}^n - x_{j+1/2} \right) + \left( x_{j+1/2} - x_{j+1/2,l}^n \right) \right\} / 2.$$

Since  $\left( x_{j+1/2,r}^n - x_{j+1/2} \right) = a_{j+1/2}^R \Delta t$  and  $\left( x_{j+1/2} - x_{j+1/2,l}^n \right) = a_{j+1/2}^L \Delta t$ , then  $\Delta = \frac{\Delta t}{2} \left( |a_{j+1/2}^R| + |a_{j+1/2}^L| \right) = \frac{\Delta t}{2} \left( a_{j+1/2}^R - a_{j+1/2}^L \right)$ . The limited slope is then estimated by:

$$(u_x)_i^{n+1} = \text{MINMOD} \left( \theta \frac{\tilde{u}_{j+1/2,r}^{n+1} - w_{j+1/2}^{n+1}}{\Delta}, \frac{\tilde{u}_{j+1/2,r}^{n+1} - \tilde{u}_{j+1/2,l}^{n+1}}{2\Delta}, \theta \frac{w_{j+1/2}^{n+1} - \tilde{u}_{j+1/2,l}^{n+1}}{\Delta} \right);$$

where  $\tilde{u}^{n+1}$  denotes an approximated solution by the Taylor expansion at next time step  $n+1$  at edges of non-smooth solutions,  $x_{j+1/2,r}^n$  and  $x_{j+1/2,l}^n$  (5.1). Values of  $\tilde{u}^{n+1}$  can be estimated as:  $\tilde{u}^{n+1} = u^n + \Delta t \times u_t^n = u^n - \Delta t \times F(u^n)_x$ . This approximation (i.e., using of the Taylor expansion) is valid in smooth zones. Regarding the methods  $M1$



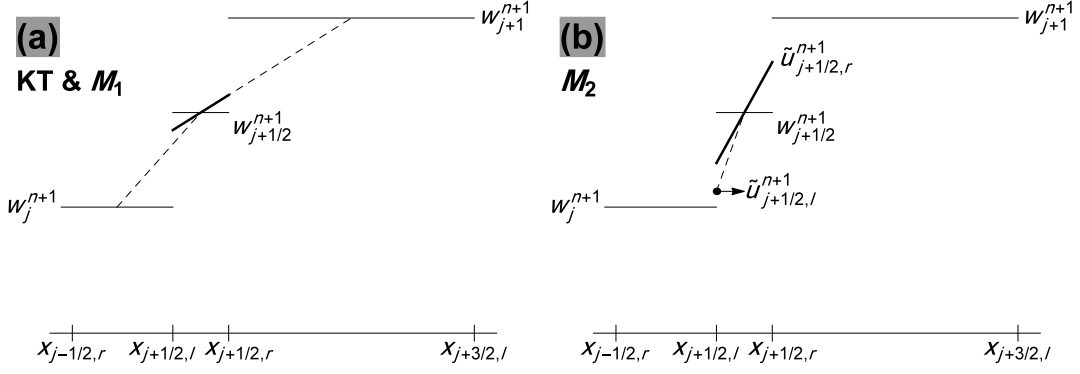


Figure 5.1: Estimation of the slope  $(u_x)^{n+1}$  in different schemes; a) KT and M1, b) M2.

and  $M2$ , for more details see [23, 24].

Expressed in terms of the left and right states  $F_{i+1/2}^{L/R} := F(u_{i+1/2}^{L/R})$ , the reconstructed fluxes for the KT,  $M1$  and  $M2$  schemes are:

1. The central KT scheme:

$$F_{i+1/2}^* := \frac{1}{2} [F_{i+1/2}^R + F_{i+1/2}^L] - \frac{a_{i+1/2}}{2} [u_{i+1/2}^R - u_{i+1/2}^L]. \quad (5.2)$$

2. For the central-upwind method  $M1$ :

$$F_{i+1/2}^* := \frac{a_{i+1/2}^R F_{i+1/2}^L - a_{i+1/2}^L F_{i+1/2}^R}{a_{i+1/2}^R - a_{i+1/2}^L} + a_{i+1/2}^L a_{i+1/2}^R \frac{u_{i+1/2}^R - u_{i+1/2}^L}{a_{i+1/2}^R - a_{i+1/2}^L}. \quad (5.3)$$

3. For the central-upwind method  $M2$ :

$$F_{i+1/2}^* := \frac{a_{i+1/2}^R F_{i+1/2}^L - a_{i+1/2}^L F_{i+1/2}^R}{a_{i+1/2}^R - a_{i+1/2}^L} + \frac{a_{i+1/2}^L a_{i+1/2}^R}{2} \times \frac{u_{i+1/2}^R - u_{i+1/2}^L}{a_{i+1/2}^R - a_{i+1/2}^L}. \quad (5.4)$$

## 5.2 Numerical entropy productions

Nonlinear hyperbolic systems (from conservation laws) can converge to several numerically converged solutions; while one of them is the physical (real) solution. Lax and Wendroff proved that if a sequence of numerical solutions converges to some function  $u(x, t)$  as the computing grid is refined, by some sequence  $\Delta x$  and  $\Delta t \rightarrow 0$ , then this function will be a weak solution of the conservation law.

**Theorem 3. (Lax and Wendroff [68])** Consider a sequence of grids indexed by  $l = 1, 2, \dots$ , with mesh parameters  $\Delta x_l, \Delta t_l \rightarrow 0$  as  $l \rightarrow \infty$ . Let  $\tilde{u}_l(x, t)$  denote the numeri-

cal approximation computed with a consistent and conservative method on the  $l$ th grid. Suppose that  $\tilde{u}_l$  converges to a function  $u$  as  $l \rightarrow \infty$ , in the sense made precise below. Then  $u(x, t)$  is a weak solution of the conservation law.

This theorem states that for the system  $u_t + f(u)_x = 0$ , the converged solution  $\tilde{u}$  only satisfies the weak-form equation:

$$\int_0^\infty \int_{-\infty}^{+\infty} [\phi_t u + \phi_x f(u)] dx dt = - \int_{-\infty}^{+\infty} \phi(x, 0) u(x, 0) dx,$$

where  $\phi(x, t) \in C^1$  is a test function. However, this does not mean a converged solution is the physical solution. This proper convergence can be checked by the entropy condition.

Calculation of numerical entropy production helps one control quality/uniqueness of numerical results especially for ones without exact solutions. This is because, theoretically, the numerical entropy production is zero in smooth regions while less than zero around shocks and discontinuities.

Puppo [28, 29] showed how to estimate the numerical entropy production for staggered central high resolution schemes. After these works, Puppo et al. [31] improved the previous works for non-staggered central/central-upwind schemes.

The entropy function  $\eta(u)$  with flux  $\psi(u)$  satisfies the relationship (for details see Appendix C):

$$\eta_t + \psi_x(u) \leq 0, \quad \eta := \eta(x), \quad \psi' = \eta' \cdot f'.$$

Following works [28, 29], by integrating this relationship in the spatio-temporal volume  $[x_j - \frac{\Delta x_j}{2}, x_j + \frac{\Delta x_j}{2}] \times [t^n, t^{n+1}]$ , and discretizing the inequality by considering the REP concept (used in the KT,  $M1$  or  $M2$  method), we have:

$$\eta_j^{n+1} - \left( \eta_j^n - \frac{1}{\Delta x_j} \int_{t^n}^{t^{n+1}} [\psi_{j+1/2}^* - \psi_{j-1/2}^*] \right) \leq 0, \quad \eta_j^n := \eta(x_j, t^n). \quad (5.5)$$

Based on this inequality, the density of numerical entropy production at  $x_j$ ,  $S_j^n$ , can be defined as:

$$S_j^n = \frac{1}{\Delta t} \left\{ \eta_j^{n+1} - \left( \eta_j^n - \frac{1}{\Delta x_j} \int_{t^n}^{t^{n+1}} [\psi_{j+1/2}^* - \psi_{j-1/2}^*] \right) \right\}, \quad S_j^n := S(x_j, t^n), \quad (5.6)$$

where  $\psi_{j\pm 1/2}^* := \psi_{j\pm 1/2}^* \left( (u_{j\pm 1/2}^L)^n, (u_{j\pm 1/2}^R)^n \right)$ .

In Eqs. (5.5) and (5.6), the terms in parentheses estimate the evolved values of the entropy at the next time step (this can be done with the KT, M1 or M2 method). Theoretically, as mentioned before and in Appendix C, the parameter  $S_j^n$  is zero in smooth regions, however, in numerical solutions, it may be slightly less or more than zero [31]. *This kind of entropy definition cannot efficiently detect some phenomena like contact discontinuities even by increasing number of resolution levels* (see the Sod and the Lax problems in the Euler gas dynamic system). To remedy this, other approaches like the entropy viscosity scheme can be recommended [69].

## 5.3 Local truncation errors

### 5.3.1 One dimensional systems

The aim of this section is to estimate local truncation errors in hyperbolic systems of conservation laws with the governing equation:  $u_t + F(u)_x = 0$ ,  $u(x, t = 0) = u_0(x)$ . This concept is important to check the convergence of numerical solutions.

It is easy to show that solution of this equation is also a weak solution satisfying the relationship:

$$E(u, \phi) = - \int_{t=0}^{\infty} \int_X \{u(x, t)\phi_t(x, t) + F(u(x, t))\phi_x(x, t)\} dx dt + \int_X \{u(x, t)\phi_t(x, t)\} dx = 0, \quad (5.7)$$

where  $\phi(x, t)$  is a test function that  $\phi(x, t) \in C_0^\infty(X \times [0, \infty))$ . One effective and practical way to measure convergence of a numerical solution  $u(x, t)$  is to check how much it fails to satisfy (5.7); this can be measured by evaluating  $E(u, \phi)$  [26, 27]. For a convex scalar hyperbolic system ( $d^2F(u)/du^2 \geq \alpha > 0$ ), the function  $E(u, \phi)$  measures point-wisely real errors. In nonlinear conservation laws, discontinuous solutions develop typically; in these cases standard methods of error estimation are not valid. Such approaches consider the Taylor expansion which is based on the smoothness assumption.

The final point is the relationship of this local truncation error with the weak  $Lip'$ -norm theory studied for convex scalar one dimensional hyperbolic systems. Numerical results confirm that it is also an effective tool for systems of one dimensional PDEs, 2-D problems, and even non-convex systems.

To be sure that  $E(u, \phi)$  measures errors of  $u(x, t)$ , some constraints should be met by the test function  $\phi(x, t)$ : 1) The space spanned by the test functions should have a higher order accuracy, so that the order of error  $E(u, \phi)$  is affected by solutions  $u(x, t)$ ; 2) The test function have continuous derivatives; 3) To estimate locally the error  $E(u, \phi)$ , the test function has to be compact support in spatio-temporal domains.

We now consider the truncation error  $E_j^n := E(u^\Delta, \phi)$  for the piecewise constant approximate solution  $u^\Delta(x, t)$  as:

$$u^\Delta(x, t) = \sum_{j,n} [u_j^n \cdot \chi_{C_j \times T^n}(x, t)], \quad C_j \times T^n := [x_{j-1/2}, x_{j+1/2}] \times [t^{n-1/2}, t^{n+1/2}], \quad (5.8)$$

wher  $\chi_{C_j \times T^n}(x, t)$  denotes the unit function over the spatio-temporal domain  $C_j \times T^n$ ; it is zero elsewhere.

The compact test function  $\phi(x, t)$  is assumed to be  $\phi_j^n(x, t) := B_j(x)B^n(t)$ , where  $B_j(x)$  and  $B^n(t)$  are the quadratic B-splines with center points  $x = x_j$  and  $t = t^n$ . In this case, the supports of  $B_j(x)$  and  $B^n(t)$  belong to  $x \in [x_{j-3/2}, x_{j+3/2}]$  and  $t \in [t^{n-3/2}, t^{n+3/2}]$ , respectively.

**Obtaining higher-order B-splines by the *recurrence feature* between higher-order and lower-order B-splines** Recalling that the B-spline of order  $k$  with centered position  $j$  can be obtained as:

$$B_{j,k} = B_{j,k-1}w_{j,k} + B_{j+1,k-1}(1 - w_{j+1,k}), \quad w_{j,k}(x) := \begin{cases} \frac{x-x_j}{x_{j+k}-x_j}, & \text{if } x_j \neq x_{j+k-1}, \\ 0, & \text{otherwise.} \end{cases} \quad (5.9)$$

where  $B_{j,0} = 1$  for  $j \leq x < j+1$  and  $B_{j,0} = 0$  for  $x < j$  and  $x \geq j+1$ .

B-splines should also satisfy the partition of unity condition  $\sum_j B_{j,k}(x) = 1$ . For the second order spline, Eq. (5.9) leads to:

$$B_{j,2} = B_{j,1}w_{j,2} + B_{j+1,1}(1 - w_{j+1,2}). \quad (5.10)$$

Let us consider three successive cells, with length ratios,  $a := \Delta x_{j-1}/\Delta x_j$  and  $b := \Delta x_{j+1}/\Delta x_j$ , where  $\Delta x_j := x_{j+1/2} - x_{j-1/2}$  denotes  $j^{th}$  cell length with cell middle point  $\bar{x}_j$ . It is easy to show that  $B_j(x)$  is:

$$B_j(x) = \begin{cases} \frac{(2x-2\bar{x}_j+\Delta x+2a\Delta x)^2}{4a(1+a)\Delta x^2}, & \text{if } x_{j-3/2} \leq x \leq x_{j-1/2}, \\ \frac{-4(2+a+b)(x-\bar{x}_j)^2-4(a-b)(x-\bar{x}_j)\Delta x+(2+3b+a(3+4b))\Delta x^2}{4(1+a)(1+b)\Delta x^2}, & \text{if } x_{j-1/2} \leq x \leq x_{j+1/2}, \\ \frac{(-2x+2\bar{x}_j+\Delta x+2b\Delta x)^2}{4b(1+b)\Delta x^2}, & \text{if } x_{j+1/2} \leq x \leq x_{j+3/2}, \\ 0, & \text{otherwise,} \end{cases} \quad (5.11)$$

where:  $\Delta x := \Delta x_j$ ;  $\bar{x}_j = (x_{j-1/2} + x_{j+1/2})/2$ ;  $x_{j-3/2} := \bar{x}_j - \Delta x(1/2 + a)$ ;  $x_{j-1/2} := \bar{x}_j - \Delta x/2$ ;  $x_{j+3/2} := \bar{x}_j + \Delta x(1/2 + b)$ ;  $x_{j+1/2} := \bar{x}_j + \Delta x/2$ .

It is easy to check that on uniform grids, where  $a = b = 1$ , the above-mentioned B-spline definition leads to the B-spline on uniform grids. Since in time-domain, in this study, a constant time-step is used, the definition of  $B^n(t)$  is the quadratic B-spline on uniform grids with step  $\Delta t$ , as:

$$B^n(x) = \begin{cases} \frac{(t-t^{n-3/2})^2}{2\Delta t^2}, & \text{if } t^{n-3/2} \leq t \leq t^{n-1/2}, \\ \frac{3}{4} - \frac{(t-t^n)^2}{\Delta t^2}, & \text{if } t^{n-1/2} \leq t \leq t^{n+1/2}, \\ \frac{(t-t^{n+3/2})^2}{2\Delta t^2}, & \text{if } t^{n+1/2} \leq t \leq t^{n+3/2}, \\ 0, & \text{otherwise,} \end{cases} \quad (5.12)$$

here:  $t^{n\pm\frac{1}{2}} := (t_n \pm \frac{\Delta t}{2})$  and  $t^{n\pm\frac{3}{2}} := (t_n \pm \frac{3}{2}\Delta t)$ .

For spatially non-uniform grid points, it is straightforward to show that the local truncation error can be expressed as:

$$E_j^n = \Delta x_j \mathcal{U}_j^n + \Delta t \mathcal{F}_j^n, \quad (5.13)$$

where  $\mathcal{U}_j^n$  is expressed in terms of the time differences  $\Delta u_\alpha^n := 1/2(u_\alpha^{n+1} - u_\alpha^{n-1})$ ,

$$\mathcal{U}_j^n = \frac{a^2}{3(a+1)} \Delta u_{j-1}^n + \frac{(a(3b+2) + 2b+1)}{3(a+1)(b+1)} \Delta u_j^n + \frac{b^2}{3(b+1)} \Delta u_{j+1}^n,$$

and  $\mathcal{F}_j^n$  is expressed in terms of the time averages  $\mu F_\alpha^n := 1/6(F_\alpha^{n-1} + 4F_\alpha^n + F_\alpha^{n+1})$ ,

$$\mathcal{F}_j^n = -\frac{(b-a)}{(a+1)(b+1)} \mu F_j^n - \frac{a}{a+1} \mu F_{j-1}^n + \frac{b}{b+1} \mu F_{j+1}^n.$$

### 5.3.2 Two-dimensional problems

Let us consider a two dimensional scalar hyperbolic PDE as:

$$u_t + F(u)_x + G(u)_y = 0, \quad u_0(x, y) := u(x, y, t = 0). \quad (5.14)$$

As in the 1-D case, the weak solution of (5.14) satisfies

$$\begin{aligned} E(u, \phi) = & - \int_{t=0}^{\infty} \int_{\Omega} u(x, y, t) \phi_t(x, y, t) + F(u(x, y, t)) \phi_x(x, y, t) + G(u(x, y, t)) \phi_y(x, y, t) \, dx dy dt \\ & + \int_{\Omega} u(x, y, t) \phi_t(x, y, t) \, dx dy = 0. \end{aligned} \quad (5.15)$$

We quantify the truncation error  $E_{j,k}^n := E(u^\Delta, \phi)$  for the piecewise constant approximate solution,  $u^\Delta(x, y, t)$ ,

$$u^\Delta(x, y, t) := \sum_{j,k,n} u_{j,k}^n \chi_{C_{j,k} \times T^n}(x, y, t),$$

where  $C_{j,k} \times T^n := [x_{j-1/2}, x_{j+1/2}] \times [y_{k-1/2}, y_{k+1/2}] \times [t^{n-1/2}, t^{n+1/2}]$ .

The test function  $\phi(x, y, t)$  can also be defined as:  $\phi_{j,k}^n(x, y, t) := B_j(x)B_k(y)B^n(t)$ . Where all functions  $B_j(x)$ ,  $B_k(y)$  (from (5.11)), and  $B^n(t)$  (from (5.12)) denote quadratic B-splines. Let us assume the successive cell length ratios in  $x$  and  $y$  directions to be:  $\left\{ a = \frac{\Delta x_{j-1}}{\Delta x_j}, b = \frac{\Delta x_{j+1}}{\Delta x_j} \right\}$  and  $\left\{ c = \frac{\Delta y_{k-1}}{\Delta y_k}, d = \frac{\Delta y_{k+1}}{\Delta y_k} \right\}$ , where:  $\Delta x = \Delta x_j := x_{j+1/2} - x_{j-1/2}$  and  $\Delta y = \Delta y_k := y_{k+1/2} - y_{k-1/2}$ . By these assumptions, the local truncation error  $E_{j,k}^n$  can be represented as:

$$E_{j,k}^n = \left\{ \Delta x_j \Delta y_k \mathcal{U}_{j,k}^n + \Delta t \Delta y_k \mathcal{F}_{j,k}^n + \Delta t \Delta x_j \mathcal{G}_{j,k}^n \right\}, \quad (5.16)$$

where  $\mathcal{U}_{j,k}^n$  is expressed in terms of the time differences  $\Delta u_{\alpha,\beta}^n := 1/2 (u_{\alpha,\beta}^{n+1} - u_{\alpha,\beta}^{n-1})$

$$\begin{aligned}
 \mathcal{U}_{j,k}^n = & \frac{a^2 d^3}{9(a+1)b(b+1)} \Delta u_{j-1,k+1}^n + \frac{a^2(a(3b+2)+2b+1)u_{j-1,k}^{n-1}}{9(a+1)^2(b+1)} \Delta u_{j-1,k}^n + \\
 & \frac{b^2 c^3}{9a(a+1)(b+1)} \Delta u_{j+1,k-1}^n + \frac{b^2(a(3b+2)+2b+1)}{9(a+1)(b+1)^2} \Delta u_{j+1,k}^n + \\
 & \frac{c^3(a(3b+2)+2b+1)}{9a(a+1)^2(b+1)} \Delta u_{j,k-1}^n + \frac{d^3(a(3b+2)+2b+1)}{9(a+1)b(b+1)^2} \Delta u_{j,k+1}^n + \\
 & \frac{(a(3b+2)+2b+1)^2}{9(a+1)^2(b+1)^2} \Delta u_{j,k}^n + \frac{ac^3}{9(a+1)^2} \Delta u_{j-1,k-1}^n + \frac{bd^3 u_{j+1,k+1}^{n-1}}{9(b+1)^2} \Delta u_{j+1,k+1}^n,
 \end{aligned}$$

and  $\mathcal{F}_{j,k}^n$  and  $\mathcal{G}_{j,k}^n$  are expressed in terms of the time averages  $\mu Z_{\alpha,\beta}^n := 1/6 \left( Z_{\alpha,\beta}^{n-1} + 4Z_{\alpha,\beta}^n + Z_{\alpha,\beta}^{n+1} \right)$ ,

$$\begin{aligned}
 \mathcal{F}_{j,k}^n = & -\frac{c^3}{3(a+1)^2} \mu F_{j-1,k-1}^n + \frac{(a-b)c^3}{3a(a+1)^2(b+1)} \mu F_{j,k-1}^n + \frac{bc^3}{3(a^2+a)(b+1)} \mu F_{j+1,k-1}^n \\
 & -\frac{a(2b+a(3b+2)+1)}{3(a+1)^2(b+1)} \mu F_{j-1,k}^n - \frac{ad^3}{3(a+1)(b^2+b)} \mu F_{j,k}^n + \frac{(a-b)(2b+a(3b+2)+1)}{3(a+1)^2(b+1)^2} \mu F_{j,k}^n \\
 & + \frac{(a-b)d^3}{3(a+1)b(b+1)^2} \mu F_{j,k+1}^n + \frac{b(2b+a(3b+2)+1)}{3(a+1)(b+1)^2} \mu F_{j+1,k}^n + \frac{d^3}{3(b+1)^2} \mu F_{j+1,k+1}^n,
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{G}_{j,k}^n = & -\frac{c^2 b^2}{3(a^2+a)(b+1)} \mu G_{j+1,k-1}^n + \frac{(a-b)b^2}{3(a+1)(b+1)^2} \mu G_{j+1,k}^n + \frac{d^2 b}{3(b+1)^2} \mu G_{j+1,k+1}^n \\
 & -\frac{ac^2}{3(a+1)^2} \mu G_{j-1,k-1}^n + \frac{a^2(a-b)}{3(a+1)^2(b+1)} \mu G_{j-1,k}^n + \frac{a^2 d^2}{3(a+1)(b^2+b)} \mu G_{j-1,k+1}^n \\
 & -\frac{c^2(2b+a(3b+2)+1)}{3a(a+1)^2(b+1)} \mu G_{j,k-1}^n + \frac{(a-b)(2b+a(3b+2)+1)}{3(a+1)^2(b+1)^2} \mu G_{j,k}^n \\
 & + \frac{d^2(2b+a(3b+2)+1)}{3(a+1)(b+1)^2 b} \mu G_{j,k+1}^n.
 \end{aligned}$$

## 5.4 Other second-order slope limiters

So far, it is shown that the families of the MINMOD limiters work well on non-uniform grids. In the following, the performance of other families will be investigated.

### 5.4.1 Slope limiters on uniform cells: the TVD ones

Let us consider uniform cell and the concept of the central formulation for approximation of the first derivative. The edge value for the cell  $I_j$  can be evaluated as:

$$\begin{aligned} u_{i+1/2}^L &= u_i + \frac{\Delta x}{2} \left( \phi(R_i) \frac{\Delta_0 u_i}{2\Delta x} \right), \\ u_{i-1/2}^R &= u_i - \frac{\Delta x}{2} \left( \phi(R_i) \frac{\Delta_0 u_i}{2\Delta x} \right), \end{aligned} \quad (5.17)$$

where  $\Delta_0 u_i := u_{i+1} - u_{i-1}$ ,  $\Delta x = x_{i+1/2} - x_{i-1/2}$  denotes the length of the cell  $I_j$  and  $R = R_i := \frac{\Delta_+ u_i}{\Delta_- u_i}$  measures the subsequent gradients at  $x_j$ , wherein  $\Delta_+ u_i = u_{i+1} - u_i$  and  $\Delta_- u_i = u_i - u_{i-1}$ . In other words, the limited slope at  $x_j$  is:

$$(\phi D)_i = \phi(R_i) \times D_0 u_i, \quad (5.18)$$

wherein  $D_0 u_i := \frac{\Delta_0 u_i}{\Delta_0 x_i}$ . Using this limited slope, different limiters can be provided. Regarding the definition of  $R_i$ , some of common limiters are:

$$\begin{aligned} (\phi_i)_{BJ} &:= \min \left\{ 1, \frac{4}{R+1}, \frac{4R}{R+1} \right\}, \\ (\phi_i)_{VL} &:= \frac{4R}{(R+1)^2}, \\ (\phi_i)_{VA} &:= \frac{2R}{R^2+1}, \\ (\phi_i)_{MM} &:= \min \left\{ \frac{2}{1+R}, \frac{2R}{1+R} \right\}, \end{aligned} \quad (5.19)$$

where  $BJ$ ,  $VL$ ,  $VA$  and  $MM$  stand for the Barth-Jespersen, van Leer, van Albada and MINMOD limiters, respectively.

For instance, let us assume a monotonically increasing solution  $\{u_i\}$  measured over the cell set  $\{I_{j-1}, I_j, I_{j+1}\}$ , in which  $I_j$  is the transmitting cell, that is:  $x_{j+1/2} - x_j = p\Delta x_j$  and  $x_j - x_{j-1/2} = (1-p)\Delta x_j$  wherein  $\Delta x_j = x_{j+1/2} - x_{j-1/2}$ . Then the MINMOD limiter can be re-expressed on  $x_j$  as:



$$\begin{aligned}
 MM \left( \frac{u_{i+1} - u_i}{\Delta x}, \frac{u_i - u_{i-1}}{\Delta x} \right) &= \text{Min} \left( \frac{u_{i+1} - u_i}{\Delta x}, \frac{u_i - u_{i-1}}{\Delta x} \right) \\
 &= \text{Min} \left( \frac{u_{i+1} - u_i}{u_i - u_{i-1}}, 1 \right) \times \frac{(u_{i+1} - u_i)}{\Delta x} = \text{Min} \left( 1, \frac{1}{R} \right) \times \frac{(u_{i+1} - u_i)}{\Delta x} \\
 &= \text{Min} \left( 1, \frac{1}{R} \right) \times \frac{2(u_{i+1} - u_i)}{2\Delta x} \times \frac{u_{i+1} - u_{i-1}}{u_{i+1} - u_{i-1}} \\
 &= \text{Min} \left( 2, \frac{2}{R} \right) \times \frac{u_{i+1} - u_i}{u_{i+1} - u_{i-1}} \times \frac{u_{i+1} - u_{i-1}}{2\Delta x} \\
 &= \text{Min} \left( 2, \frac{2}{R} \right) \times \frac{u_{i+1} - u_i}{u_{i+1} - u_i + u_i - u_{i-1}} \times D_0 u_i \\
 &= \text{Min} \left( 2, \frac{2}{R} \right) \times \frac{1}{1 + \frac{1}{R}} \times D_0 u_i \\
 &= \text{Min} \left( \frac{2R}{1 + R}, \frac{2}{1 + R} \right) \times D_0 u_i.
 \end{aligned} \tag{5.20}$$

Regarding the definition of the parameter  $f := \Delta_- u_i / \Delta_0 u_i$ , the limiters presented in Eq. (5.19) can be re-expressed as:

$$\begin{aligned}
 (\phi_i)_{BJ} &:= \min \{1, 4f, 4(1 - f)\}, \\
 (\phi_i)_{VL} &:= 4f(1 - f), \\
 (\phi_i)_{VA} &:= \frac{2f(1 - f)}{f^2 + (1 - f)^2}, \\
 (\phi_i)_{MM} &:= \min \{2f, 2(1 - f)\}.
 \end{aligned} \tag{5.21}$$

Other TVD limiters can also be defined, such as:

$$(\phi_i)_{\sin} := \sin(\pi f), \tag{5.22}$$

where  $(\phi_i)_{\sin}$  denotes the *sin* limiter. Regarding the TVD constraint for uniform cells, the above-mentioned limiters are presented in Figure 5.2. It is clear the limiters remain in the TVD bound and satisfy the linearity preserving feature (presented by a black point).

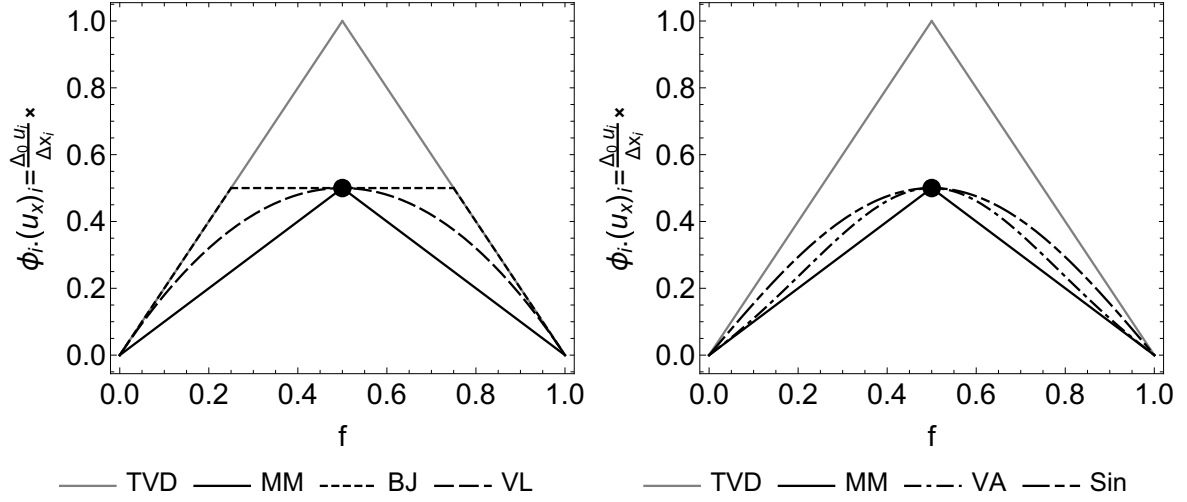


Figure 5.2: The variation of some commonly used limiters in terms of the parameter  $f$ ; the black point represent the linearity preserving location.

### 5.4.2 Slope limiters on non-uniform cells: the TVD ones

In general, the limited central first-derivative,  $(u_x)_j$  can be defined on a transmitting (non-centered) cell as:

$$\begin{aligned}
 (\phi_i D_i) &= (\phi_i) \times \frac{\Delta_0 u_i}{\Delta_0 x_i} \\
 &= (\phi_i) \times \frac{\Delta_0 u_i}{\frac{\Delta x_{j-1}}{2} + \Delta x_j + \frac{\Delta x_{j+1}}{2}} \\
 &= (\phi_i) \times \frac{\Delta_0 u_i}{a \frac{\Delta x_j}{2} + \Delta x_j + b \frac{\Delta x_j}{2}} \\
 &= (\phi_i) \times \frac{\Delta_0 u_i}{\left(\frac{a}{2} + 1 + \frac{b}{2}\right) \Delta x_j} \\
 &= \left\{ \frac{1}{\left(\frac{a}{2} + 1 + \frac{b}{2}\right)} (\phi_i) \right\} \times \frac{\Delta_0 u_i}{\Delta x_j}.
 \end{aligned} \tag{5.23}$$

Regarding non-uniform grids, it is easy to show that the limiters presented in (5.21) and (5.22) do not necessarily remain in the TVD region or pass through the linearity preserving point.

To work properly on wavelet-based adapted grids, it is essential to re-define new limiters or to modify the common ones satisfying the TVD conditions and the linearity preserving feature. In general, such limiters should satisfy some conditions:

1.  $(\phi_j D_j)(f = 0) = 0$ ,

2.  $(\phi_j D_j)(f = 1) = 0$ ,
3. Satisfying the linearity preserving condition:  $(\phi_j D_j)(f = f_p) = \frac{2}{2+a+b} \frac{\Delta_0 u_j}{\Delta x_j}$ , where  $f_p = \frac{2+a-2p}{2+a+b}$ ,
4. Advising to have zero value of the first-derivative at the peak point  $f = f_p$ , that is:  $\left(\frac{d}{df} \phi_j D_j\right)(f = f_p) = 0$ ,
5. Remaining in the TVD region.

In the following, some limiters are updated/defined to handle multiresolution-based adapted cells.

### The van Leer limiter

For monotonic solutions, the original van Leer limiter can be assumed to be the harmonic average of the first derivatives  $D_- u_j$  and  $D_+ u_j$ , as:

$$\frac{1}{(\phi_{VL} D)_j} = \left(\frac{1}{2}\right) \left(\frac{1}{D_- u_j}\right) + \left(\frac{1}{2}\right) \left(\frac{1}{D_+ u_j}\right). \quad (5.24)$$

On uniform cells, since  $D_- u_j := f \frac{\Delta_0 u_j}{\Delta x_j}$  and  $D_+ u_j := (1-f) \frac{\Delta_0 u_j}{\Delta x_j}$ , it is easy to show that the relationship in Eq. (5.21) can be obtained.

Over non-uniform cells, using the harmonic averaging of  $D_- u_j$  and  $D_+ u_j$ , in which  $D_- u_j := \frac{2f}{1+a} \frac{\Delta_0 u_j}{\Delta x_j}$  and  $D_+ u_j := \frac{2(1-f)}{1+b} \frac{\Delta_0 u_j}{\Delta x_j}$ , the limited slope reads:

$$(\phi_{VL}^{NU} D)_j = \frac{4(1-f)f}{(1-f)(1+a) + f(1+b)} \frac{\Delta_0 u_j}{\Delta x_j}, \quad (5.25)$$

where  $a := \Delta x_{j-1}/\Delta x_j$  and  $b := \Delta x_{j+1}/\Delta x_j$ . For a transition (non-centered) cell (e.g.:  $a = 4/3$ ,  $b = 2/3$  and  $p = 1/3$ ) this limiter is presented in Figure 5.3(a). It is obvious that it does not satisfy the linearity preserving feature.

To update the van Leer limiter for non-uniform cells, here, a weighted harmonic averaging is proposed for monotonic solutions, as:

$$\frac{1}{(\phi_{VL}^{WNU} D)_j} = \left(\frac{\Delta_- x_j}{\Delta_0 x_j}\right) \left(\frac{1}{D_- u_j}\right) + \left(\frac{\Delta_+ x_j}{\Delta_0 x_j}\right) \left(\frac{1}{D_+ u_j}\right). \quad (5.26)$$

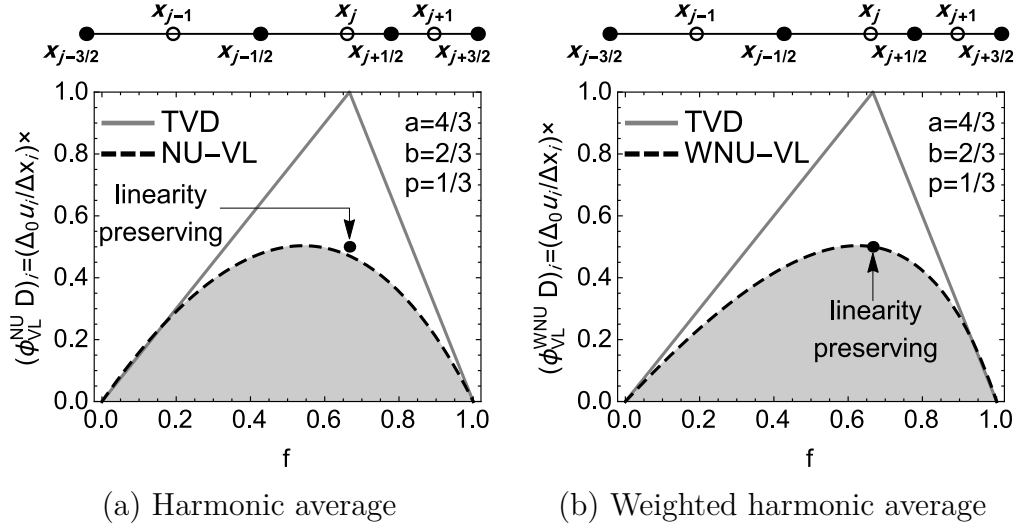


Figure 5.3: The modified van Leer limiters on a transmitting cell.

Re-expressing this equation based on the parameter  $f$ , it reads:

$$(\phi_{VL}^{WNU} D)_j = \frac{-2f(1-f)(a+b+2)}{-a^2(1-f) + a(1-f)(2p-3) - (b+2)f(b+2p-1) - 2(1-p)} \times \frac{\Delta_0 u_j}{\Delta x_j} \quad (5.27)$$

This limiter is illustrated in Figure 5.3(b) on a transmitting cell; it is obvious that it satisfies both the TVD and the linearity preserving properties.

### The van Albada limiter

It is easy to show that the van Albada limiter on uniform cells can be obtained by a proper weighted harmonic averaging of the forward ( $D_+ u_j$ ) and the backward ( $D_- u_j$ ) first derivatives, as:

$$\frac{1}{(\phi_{VAD})_j} = \left( \frac{f}{(f) + (1-f)} \right) \left( \frac{1}{D_+ u_j} \right) + \left( \frac{1-f}{(f) + (1-f)} \right) \left( \frac{1}{D_- u_j} \right). \quad (5.28)$$

Since, on uniform cells,  $a = b = 1$ ,  $D_- u_j := f \frac{\Delta_0 u_j}{\Delta x_j}$  and  $D_+ u_j := (1-f) \frac{\Delta_0 u_j}{\Delta x_j}$ , then:

$$\begin{aligned}
 (\phi_{VA}D)_j &= \frac{f(1-f)}{2f^2-2f+1} \frac{\Delta_0 u_j}{\Delta x_j} \\
 &= \frac{2f(1-f)}{f^2+(1-f)^2} \frac{\Delta_0 u_j}{2\Delta x_j}
 \end{aligned} \tag{5.29}$$

The symmetric feature with respect to  $f$  is clear in Eq. (5.29).

For non-uniform non-centered cell  $I_j$  let us re-express the weighted harmonic average as:

$$\frac{1}{(\phi_{VA}^{WNU}D)_j} = \left( \frac{\alpha f}{(\alpha f) + (1 - \alpha f)} \right) \left( \frac{1}{D_+ u_j} \right) + \left( \frac{(1 - \alpha f)}{(\alpha f) + (1 - \alpha f)} \right) \left( \frac{1}{D_- u_j} \right). \tag{5.30}$$

where  $\alpha$  is an unknown constant. The constraint is the linearity preserving feature, that is:  $(\phi_i D_i)(f_p) = \frac{2}{a+b+2} \frac{\Delta_0 u_i}{\Delta x_i}$ , where  $f_p = \frac{2+a-2p}{2+a+b}$ . Then it is straightforward to show that the coefficient  $\alpha$  reads:

$$\alpha = \frac{b + 2p}{2 + a - 2p}, \tag{5.31}$$

where  $x_{i+1/2} - x_i = p\Delta x_i$ ,  $a := \Delta x_{i-1}/\Delta x_i$  and  $b := \Delta x_{i+1}/\Delta x_i$ .

Finally, based on the modified van Albada formulation, the limited slope reads:

$$\left( \phi_{VA}^{WNU}D \right)_i = \frac{1}{\frac{(1+b)f(b+2p)}{2(-1+f)(2+a-2p)} - \frac{(1+a)\left(-1+\frac{f(b+2p)}{2+a-2p}\right)}{2f}} \times \frac{\Delta_0 u_i}{\Delta x_i}. \tag{5.32}$$

The updated van Albada limiter on a transmitting cell is presented in Figure 5.4 with the parameters  $a = 4/3$ ,  $b = 2/3$  and  $p = 1/3$ . Also, in this figure the TVD boundary and the MM limiter are presented. It is clear the updated van Albada limiter remains between the MM limiter and the TVD constraint, and  $\left( \phi_{VA}^{WNU}D \right)_i$  preserves the linearity preserving condition, as well.

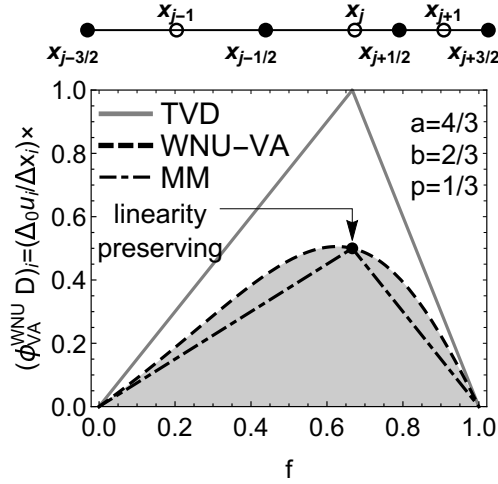


Figure 5.4: The modified van Albada limiter on a transmitting cell, derived from a weighted harmonic average.

### Modified van Albada limiters

On non-uniform con-centered cells, the van Albada limiter defined in Eq. (5.30) can be re-expressed with new weights as:

$$\frac{1}{(\phi_{mVA,n}^{WNU} D)_j} = \left( \frac{\alpha \cdot f^{\frac{1}{n}}}{(\alpha \cdot f^{\frac{1}{n}}) + (1 - \alpha \cdot f^{\frac{1}{n}})} \right) \left( \frac{1}{D_+ u_j} \right) + \left( \frac{(1 - \alpha \cdot f^{\frac{1}{n}})}{(\alpha \cdot f^{\frac{1}{n}}) + (1 - \alpha \cdot f^{\frac{1}{n}})} \right) \left( \frac{1}{D_- u_j} \right), \quad (5.33)$$

where  $\alpha$  and  $n$  are unknown constants.

To satisfy the linearity preserving feature, it is essential that:

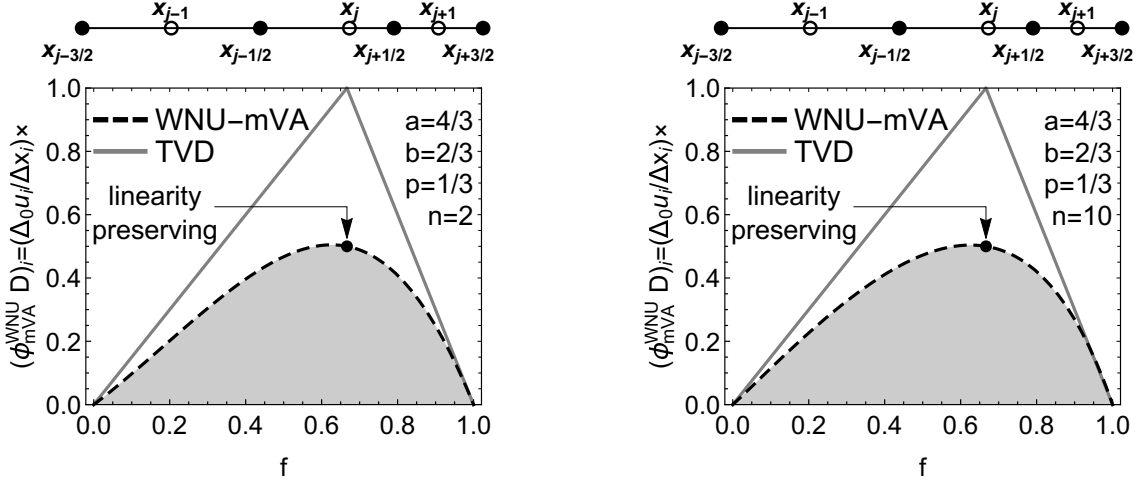
$$\alpha = f_p^{-\frac{1}{n}} \frac{(b + 2p)}{2} \frac{2}{2 + a + b}. \quad (5.34)$$

It is straightforward to show that for  $n = 1$ , Eq. (5.34) is equal to Eq. (5.31).

For two values of  $n$ ,  $n = 2$  and  $n = 10$ , over a transmitting cell  $I_j$  with the parameters  $a = 4/3$ ,  $b = 2/3$  and  $p = 1/3$ , the limiters  $(\phi_{mVA,n}^{WNU} D)_j$  are presented in Figure 5.5. It is obvious both the TVD and linearity preserving features are satisfied.

### The Superbee limiter

Another famous TVD slope limiter is the Superbee limiter introduced by Roe [178], defined as:



(a) The modified Albada limiter with  $n = 2$  (b) The modified Albada limiter with  $n = 10$

Figure 5.5: The modified van Albada limiters on a transmitting cell with different  $n$  values.

$$(\phi_i)_{SB} D_i = \text{MAXMOD} \{ \delta_i^{(1)}, \delta_i^{(2)} \}, \quad (5.35)$$

where:

$$\begin{aligned} \delta_i^{(1)} &= \text{MINMOD} \left\{ \frac{u_{i+1} - u_i}{x_{i+1} - x_i}, 2 \frac{u_i - u_{i-1}}{x_i - x_{i-1}} \right\}, \\ \delta_i^{(2)} &= \text{MINMOD} \left\{ 2 \frac{u_{i+1} - u_i}{x_{i+1} - x_i}, \frac{u_i - u_{i-1}}{x_i - x_{i-1}} \right\}, \end{aligned} \quad (5.36)$$

and:

$$\text{MAXMOD} (a_1, a_2) = \begin{cases} s \cdot \max(|a_1|, |a_2|), & \text{if } \text{sign}(a_1) = \text{sign}(a_2) = s, \\ 0, & \text{otherwise.} \end{cases} \quad (5.37)$$

Again, let us consider the cell set  $\{I_{j-1}, I_j, I_{j+1}\}$ , in which  $I_j$  is the transmitting cell, that is:  $x_{j+1/2} - x_j = p\Delta x_j$  and  $x_j - x_{j-1/2} = (1-p)\Delta x_j$  wherein  $\Delta x_j = x_{j+1/2} - x_{j-1/2}$ . And also the cells  $I_{j-1}$  and  $I_{j+1}$  are the cell-centered cells. In this case for a monotone solution over  $\{I_{j-1}, I_j, I_{j+1}\}$ , Eq. (5.35) can be re-expressed as:

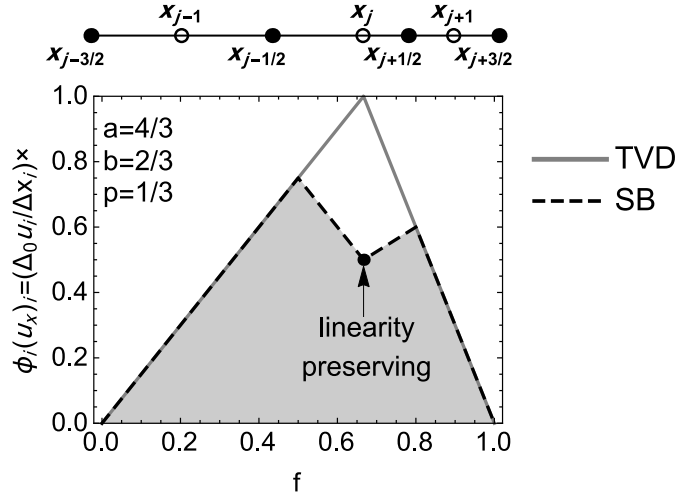


Figure 5.6: The performance of the SuperBee limiter on non-uniform cells.

$$\begin{aligned}
 (\phi_i)_{SB} D_i = & \\
 & \left( \max \left\{ \min \left[ \left( 2 \frac{1-f}{2p+b} \right), 2 \left( 2 \frac{f}{2(1-p)+a} \right) \right], \min \left[ 2 \left( 2 \frac{1-f}{2p+b} \right), \left( 2 \frac{f}{2(1-p)+a} \right) \right] \right\} \right) \\
 & \times \frac{\Delta_0 u_i}{\Delta x_i}
 \end{aligned} \tag{5.38}$$

### Some other new TVD limiters

Let us assume a set of non-uniform cells; designing of some other second-order TVD limiters can be performed by satisfying some conditions. For a transmitting cell, these conditions are:

1. At  $f = f_p = \frac{2+a-2p}{2+a+b}$ , the limited derivative should be  $(\phi_i D_i) = \frac{2}{2+a+b} \frac{\Delta_0 u_i}{\Delta x_j}$ ,
2. For monotone data  $\phi_i(f=0) = 0$ ,
3. For monotone data  $\phi_i(f=1) = 0$ ,
4. For  $0 \leq f \leq 1$ , a second-order limiter must be between the MM limiter and the TVD upper bound, that is:  $(\phi_i)_{MM} \leq (\phi_i) \leq (\phi_i)_{TVD}$ .

Also, it is recommended that at  $f = f_p$ , the first derivative of a limiter with respect to  $f$  is zero, i.e.:  $\frac{d(\phi)_i}{df} = 0$ .



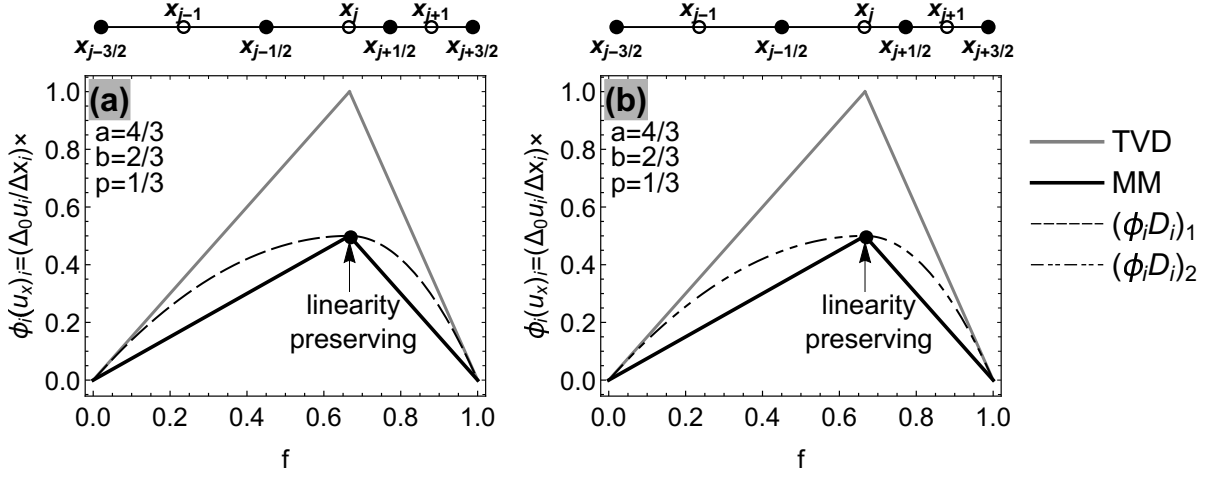


Figure 5.7: The performance of  $(\phi_j D_j)_1$  and  $(\phi_j D_j)_2$  limiters; a) The limiter  $(\phi_j D_j)_1$ ; b) The limiter  $(\phi_j D_j)_2$ .

Two possible examples are as follows:

$$(\phi_i D_i)_1 := \begin{cases} \frac{f}{1-p} \left[ 1 - \frac{a}{2+a-2p} \left( \frac{f}{f_p} \right)^{\frac{2(1-p)}{a}} \right] \times \frac{\Delta_0 u_j}{\Delta x_j}, & \text{for } f \leq f_p, \\ \frac{(1-f)}{p} \left[ 1 - \frac{b}{b+2p} \left( \frac{1-f}{1-f_p} \right)^{\frac{2p}{b}} \right] \times \frac{\Delta_0 u_j}{\Delta x_j}, & \text{for } f > f_p. \end{cases} \quad (5.39)$$

$$(\phi_i D_i)_2 := \begin{cases} \frac{2f_p}{a-2p+2} \left( 1 - \left( 1 - \frac{f}{f_p} \right)^{a+(1+a(\frac{1}{2(1-p)}-1))} \right) \times \frac{\Delta_0 u_j}{\Delta x_j}, & \text{for } f \leq f_p, \\ \frac{2(1-f_p)}{b+2p} \left( 1 - \left( 1 - \frac{1-f}{1-f_p} \right)^{b+(1+b(\frac{1}{2p}-1))} \right) \times \frac{\Delta_0 u_j}{\Delta x_j}, & \text{for } f > f_p. \end{cases} \quad (5.40)$$

Over a transmitting cell (surrounded with two cell-centered cells), these limiters, the TVD boundaries and the linearity preserving points are presented in Figure 5.7, where  $a = 4/3$ ,  $b = 2/3$  and  $1/3$ . It is obvious that  $(\phi_j D_j)_1$  and  $(\phi_j D_j)_2$  are second-order TVD limiters.

On uniform grids (i.e.:  $a = b = 1$  and  $p = 1/2$ ), the limiters  $(\phi_j D_j)_1$  and  $(\phi_i D_i)_2$  recover the *van Leer quadratic* limiter [172].

### 5.4.3 The UNO limiter on non-uniform cells: cell-centered and non-cell-centered ones

Regarding flux/slope limiters in the central high resolution schemes, different construction criteria were developed. One of the most famous ones is the TVD limiters working in a way that total variations (TVs) of numerical solutions do not increase through time. And this guarantees solutions without spurious oscillations. Different TVD limiters were successfully developed (on uniform cells); however, they have some drawbacks:

1. Most of the limiters were basically developed for uniform cells,
2. They lead to the first-order accuracy around extrema and so end to over-smoothed results around them,
3. They have at most the first-order accuracy for 2-D/3-D problems,
4. The construction criteria for limiters are restrictive.

To remedy these shortcomings, different construction approaches were suggested for limiters, such as the total variation bounded (TVB) [179] and UNO [123, 124] ones. Especially to cure the over-smoothing phenomenon around extrema, the UNO limiters were suggested. They try to preserve the same accuracy-order both over smooth and discontinuities responses and at the same time they try to control non-physical oscillations.

The formula of the second-order UNO limiter on uniform grids can directly be derived by the Taylor's series. On non-uniform cells, a modified formula can be derived by using interpolating parabolic polynomials. In the following, three sets of non-uniform cells are studied to derive the UNO-limiter at  $x_i$ .

#### The non-uniform grid $\{x_{i-2}, x_{i-1}, x_i\}$

Let  $P(x)$  denotes the interpolating parabolic polynomial over the discrete data  $\{f_{i-2}, f_{i-1}, f_i\}$  defined on  $\{x_{i-2}, x_{i-1}, x_i\}$ , where  $f_i := f(x_i)$  and  $f(x)$  is a differentiable function. The first and the second derivatives of  $P(x)$  at  $x_i$  can be evaluated as:

$$P'(x_i) = \frac{-(x_{i-2} - x_{i-1})(x_{i-2} + x_{i-1} - 2x_i)f_i + (x_{i-2} - x_i)^2 f_{i-1} - (x_{i-1} - x_i)^2 f_{i-2}}{(x_{i-2} - x_{i-1})(x_{i-2} - x_i)(x_{i-1} - x_i)}. \quad (5.41)$$

And:

$$P''(x_i) = \frac{2 \left( \frac{f_i - f_{i-1}}{x_i - x_{i-1}} - \frac{f_{i-1} - f_{i-2}}{x_{i-1} - x_{i-2}} \right)}{x_i - x_{i-2}}. \quad (5.42)$$

Let a differentiable function  $g(x)$  for  $x \in [a, b]$  is sampled on a non-uniform grid  $\{x_i\}_{i=1}^N$ , as  $\{g_i\}_{i=1}^N$ ; and a polynomial of order  $N$ ,  $P_N(x)$  interpolates the data. The first derivative at  $x_i$  can be estimated by  $P'_N(x_i)$  with the error (denoted by  $E(\cdot)$ ), as [180]:

$$E(g) = g'(x_i) - P'_N(x_i) = \frac{1}{N+1} g^{(N+1)}(\zeta) \cdot \prod_{\substack{j=0 \\ j \neq i}}^N (x_i - x_j), \quad \text{for } \zeta \in [a, b]. \quad (5.43)$$

And so  $g'(x_i) = P'_N(x_i) + \mathcal{O}(h^{N-1})$ , where  $h$  corresponds to the *maximum* local grid spacing. In this regard,  $f'_i := f'(x_i)$  can be estimated on  $\{x_{i-2}, x_{i-1}, x_i\}$  as:

$$f'_i = P'(x_i) + \mathcal{O}(h^2), \quad (5.44)$$

where  $h$  denotes the maximum local grid spacing. Let us set  $z(x) = g'(x)$ , then for the second derivative  $f''_i := f''(x_i)$ ,  $E(z)$  can be estimated as well by Eq. (5.43), and so:

$$f''_i = P''(x_i) + \mathcal{O}(h). \quad (5.45)$$

By some calculations,  $P'(x_i)$  presented in Eq. (5.41) can be re-expressed as:

$$P'(x_i) = \frac{f_i - f_{i-1}}{x_i - x_{i-1}} + \frac{1}{2} (x_i - x_{i-1}) P''(x_i). \quad (5.46)$$

Hence,  $f'_i$  with the second-order accuracy (defined in Eq. (5.44)) reads:

$$f'_i = \frac{f_i - f_{i-1}}{x_i - x_{i-1}} + \frac{1}{2} (x_i - x_{i-1}) P''(x_i) + \mathcal{O}(h^2), \quad (5.47)$$

where  $\frac{f_i - f_{i-1}}{x_i - x_{i-1}} =: D_- f_i$  denotes the first-order estimation of  $f'_i$  obtained by a linear interpolation ( $P_2(x)$ ) over the grid  $\{x_{i-1}, x_i\}$ , that is:

$$f'_i = \frac{f_i - f_{i-1}}{x_i - x_{i-1}} + \mathcal{O}(h). \quad (5.48)$$

In terms of Eq. (5.48),  $f'_i$  with the second-order accuracy, defined in Eq. (5.47), can be re-expressed as:

$$f'_i = D_- f_i + \frac{\Delta_- x_i}{2} \left( \frac{D_- f_i - D_- f_{i-1}}{\frac{\Delta_0 x_{i-1}}{2}} \right) + \mathcal{O}(h^2), \quad (5.49)$$

where:  $\Delta_- x_i := x_i - x_{i-1}$ ;  $\Delta_0 x_i := x_{i+1} - x_{i-1}$ ; and  $D_- f_i := \Delta_- f_i / \Delta_- x_i$  wherein  $\Delta_- f_i := f_i - f_{i-1}$ .

### The non-uniform grid $\{x_{i-1}, x_i, x_{i+1}\}$

Let us assume irregular grid  $\{x_{i-1}, x_i, x_{i+1}\}$  with corresponding sampled data  $\{f_{i+j}\}_{j=-1}^1$ ; using the interpolating parabolic polynomial  $P(x)$  over these data, it is easy to show that:

$$P'(x_i) = -\frac{(\Delta_- x_{i+1})^2 f_{i-1} + (-\Delta_0 x_i)(\Delta_- x_{i+1} - \Delta_- x_i) f_i - (\Delta_- x_i)^2 f_{i+1}}{\Delta_- x_{i+1} \cdot \Delta_0 x_i \cdot \Delta_- x_i}, \quad (5.50)$$

and

$$P''(x_i) = \frac{D_- f_{i+1} - D_- f_i}{\frac{1}{2}(\Delta_0 x_i)}. \quad (5.51)$$

Using the first-order backward formula for the first derivative,  $D_- f_i$ ,  $P'(x_i)$  (Eq. (5.50)) can be re-expressed as:

$$P'(x_i) = D_- f_i + \frac{1}{2}(\Delta_- x_i) P''(x_i) = D_- f_i + \frac{(D_- f_{i+1} - D_- f_i)(\Delta_- x_i)}{\frac{1}{2}(\Delta_0 x_i) \cdot 2}. \quad (5.52)$$

Regarding the relationship for the error in the derivative estimation, Eq. (5.43), the estimation of the first derivative reads:

$$f'(x_i) = P'(x_i) + \mathcal{O}(h^2) = D_- f_i + \frac{(D_- f_{i+1} - D_- f_i)(\Delta_- x_i)}{\frac{1}{2}(\Delta_0 x_i) \cdot 2} + \mathcal{O}(h^2). \quad (5.53)$$

Now let us assume the definition of the first-order forward formula for the first derivative,  $D_- f_{i+1}$ . In this case  $P'(x_i)$  (Eq. (5.50)) can also be re-expressed as:

$$P'(x_i) = D_- f_{i+1} - \frac{1}{2}(\Delta_- x_{i+1}) P''(x_i) = D_- f_{i+1} - \frac{(D_- f_{i+1} - D_- f_i)(\Delta_- x_{i+1})}{\frac{1}{2}(\Delta_0 x_i) \cdot 2}. \quad (5.54)$$

And so:

$$f'(x_i) = D_- f_{i+1} - \frac{(D_- f_{i+1} - D_- f_i) (\Delta_- x_{i+1})}{\frac{1}{2} (\Delta_0 x_i)} \frac{1}{2} + \mathcal{O}(h^2). \quad (5.55)$$

**The non-uniform grid**  $\{x_i, x_{i+1}, x_{i+2}\}$

Let us assume that the discrete data  $\{f(x_{i+j})\}_{j=0}^2$  are sampled on the grid points  $\{x_i, x_{i+1}, x_{i+2}\}$ . Now, the function  $P(x)$  denotes the interpolating parabolic polynomial over these data. It is easy to show that:

$$P'(x_i) = -\frac{\Delta_- x_{i+2} (\Delta_- x_{i+1} + \Delta_0 x_{i+1}) f_i - (\Delta_0 x_{i+1})^2 f_{i+1} + (\Delta_- x_{i+1})^2 f_{i+2}}{\Delta_0 x_{i+1} \cdot \Delta_- x_{i+1} \cdot \Delta_- x_{i+2}}, \quad (5.56)$$

and

$$P''(x_i) = \frac{D_- f_{i+2} - D_- f_{i+1}}{\frac{1}{2} (\Delta_0 x_{i+1})}. \quad (5.57)$$

Considering the definition of  $D_- f_{i+1}$ , it is straightforward to re-express Eq. (5.56) as:

$$P'(x_i) = D_- f_{i+1} - \frac{1}{2} (\Delta_- x_{i+1}) P''(x_i) = D_- f_{i+1} - \frac{D_- f_{i+2} - D_- f_{i+1}}{\frac{1}{2} (\Delta_0 x_{i+1})} \frac{\Delta_- x_{i+1}}{2}, \quad (5.58)$$

and regarding Eq. (5.43),  $f'(x_i)$  reads:

$$f'(x_i) = P'(x_i) + \mathcal{O}(h^2) = D_- f_{i+1} - \frac{D_- f_{i+2} - D_- f_{i+1}}{\frac{1}{2} (\Delta_0 x_{i+1})} \frac{\Delta_- x_{i+1}}{2} + \mathcal{O}(h^2). \quad (5.59)$$

**The limited value of**  $f'(x_i)$

In terms of Eqs. (5.49) and (5.53),  $f'(x_i)$  may be approximated by the averaging of the two candidates as:

$$f'_i \approx D_- f_i + \frac{1}{2} (\Delta_- x_i) \text{MM} \left( \frac{D_- f_{i+1} - D_- f_i}{\frac{1}{2} (\Delta_0 x_i)}, \frac{D_- f_i - D_- f_{i-1}}{\frac{1}{2} (\Delta_0 x_{i-1})} \right), \quad (5.60)$$

and from Eqs. (5.55) and (5.59),  $f'(x_i)$  may also be estimated as:

$$f'_i \approx D_- f_{i+1} - \frac{1}{2} (\Delta_- x_{i+1}) \text{MM} \left( \frac{D_- f_{i+2} - D_- f_{i+1}}{\frac{1}{2} (\Delta_0 x_{i+1})}, \frac{D_- f_{i+1} - D_- f_i}{\frac{1}{2} (\Delta_0 x_i)} \right). \quad (5.61)$$

Based on these two candidates, Eqs. (5.60) and (5.61), the limited value of  $f'(x_i)$  may be evaluated as:

$$f'_i \approx \text{MM} \left\{ D_- f_i + \frac{1}{2} (\Delta_- x_i) \text{MM} \left( \frac{D_- f_{i+1} - D_- f_i}{\frac{1}{2} (\Delta_0 x_i)}, \frac{D_- f_i - D_- f_{i-1}}{\frac{1}{2} (\Delta_0 x_{i-1})} \right), \right. \\ \left. D_- f_{i+1} - \frac{1}{2} (\Delta_- x_{i+1}) \text{MM} \left( \frac{D_- f_{i+2} - D_- f_{i+1}}{\frac{1}{2} (\Delta_0 x_{i+1})}, \frac{D_- f_{i+1} - D_- f_i}{\frac{1}{2} (\Delta_0 x_i)} \right) \right\}. \quad (5.62)$$

### Some Remarks:

1. As mentioned the MM limiter satisfies the TVD condition on centered and non-centered adapted cells Section 4.2; where for the MRA-based adapted cells, for centered and non-centered cells we have respectively  $p = 1/2$  and  $p = 1/3$ . In this regard, the estimation of the first derivative in Eq. (5.62) with the MM limiter would be *stable* and so have the UNO feature,
2. For uniform cells (i.e.:  $\Delta_- x_i = \Delta_- x_{i+1} = h$  and  $\Delta_0 x_{i+1} = \Delta_0 x_i = \Delta_0 x_{i-1} = 2h$ ), the original UNO-limiter can be retrieved (especially when  $h = 1$ ),
3. Other averaging functions may be used instead of the MM function, such as [181]:

$$m(x, y) = \begin{cases} x, & |x| \leq |y|, \\ y, & |x| > |y|. \end{cases} \quad (5.63)$$

4. Some other averaging operators may also be recommended, such as *harmonic* (known as the *harmod* function), *power* and *ENO* limiters [182].

### Reconstruction errors by using the GMM (TVD) and UNO limiters

The performance of the MM, GMM and UNO limiters are studied on adapted grids by measuring reconstruction error at cell edges. The norm at the cell interfaces can be measured as [61, 183–185]:

$$Err^p := \frac{1}{N} \left( \left\| u_{i+1/2}^- - f \left( x_{i+1/2}^- \right) \right\|_{L_{\Delta x}^{p_0}} + \left\| u_{i+1/2}^+ - f \left( x_{i+1/2}^+ \right) \right\|_{L_{\Delta x}^{p_0}} \right), \quad (5.64)$$

where  $N$  denotes the number of cell-interfaces  $\{x_{i+1/2}\}$ ;  $u_{i+1/2}^- = u_i + u'_i \times (x_{i+1/2} - x_i)$  and  $u_{i+1/2}^+ = u_{i+1} - u'_{i+1} \times (x_{i+1} - x_{i+1/2})$  indicate the left and right reconstructed values at  $x_{i+1/2}$ , respectively, in which  $u'_i$  is a proper estimation of the first derivative;

Table 5.4.1: Reconstruction error for  $f_1(x)$  at cell-interfaces for the norms with  $p_0 = 1$ 

$J_{max}$	MM limiter (GMM: $\theta = 1$ )	GMM limiter $\theta = 2$	UNO limiter By MM	UNO limiter By $m(., .)$
7	0.0462394	0.0323892	0.0326315	0.0424423
8	0.0267894	0.0152814	0.0147827	0.0153478
9	0.0140767	0.00721122	0.0070828	0.0070828
10	0.0075057	0.00386423	0.00383173	0.00383173
11	0.00604574	0.00306278	0.00310451	0.00310451

$x_{i+1/2}^\pm := \lim_{\delta x \rightarrow 0} (x_{i+1/2} \pm \delta x)$ . The definition of the norm is:

$$\|(\cdot)_i\|_{L_{\Delta x}^{p_0}} := \left( \sum_{i=1}^N |(\cdot)_i|^{p_0} \right)^{\frac{1}{p_0}}, \quad \|(\cdot)_i\|_{L_{\Delta x}^\infty} := \max_i |(\cdot)_i|. \quad (5.65)$$

In the following, two benchmark functions are studied on MRA-based adapted grids; as: 1) A smooth function including a localized high-gradient variation; and 2) A smooth function with a discontinuity.

**The first test problem** Let us consider a smooth function with localized high gradient variation, defined as [183]:

$$f_1(x) := \text{Sin}(2\pi x) + \text{Exp}\left(-\bar{\beta}(x - 0.5)^2\right) \quad \text{for } x \in [0, 1], \quad (5.66)$$

where  $\bar{\beta} = 20000$ . Let us assume that the coarsest resolution level is  $J_{min} = 5$  (with sampling step  $dx = 1/2^{J_{min}}$ ) and the finest resolution levels belong to  $J_{max} \in \{7, 8, 9, 10, 11\}$ . For the cell adaptation, the threshold value is equal to  $\epsilon = 0.5 \times 10^{-4}$ . The function  $f_1(x)$  and the distribution of adapted points in different resolution levels are presented in Figure 5.8 for different  $J_{max}$  values. In this figure, the solid line and solid points  $\bullet$  represent  $f_1(x)$  and adapted points (in different resolutions), respectively;  $W_j$  denotes the detail space of the resolution level  $j$  and  $V_i$  indicates the approximation space of the resolution level  $i = 5$ . The reconstruction errors at interfaces are measured by  $L_1$  norm, that is  $p_0 = 1$  in Eqs. (5.64) and (5.65). The errors of reconstructions with the MM, GMM and UNO limiters are presented in Table 5.4.1. The results confirm that the performance of the UNO limiters is comparable with the GMM limiter (with the TVD feature).

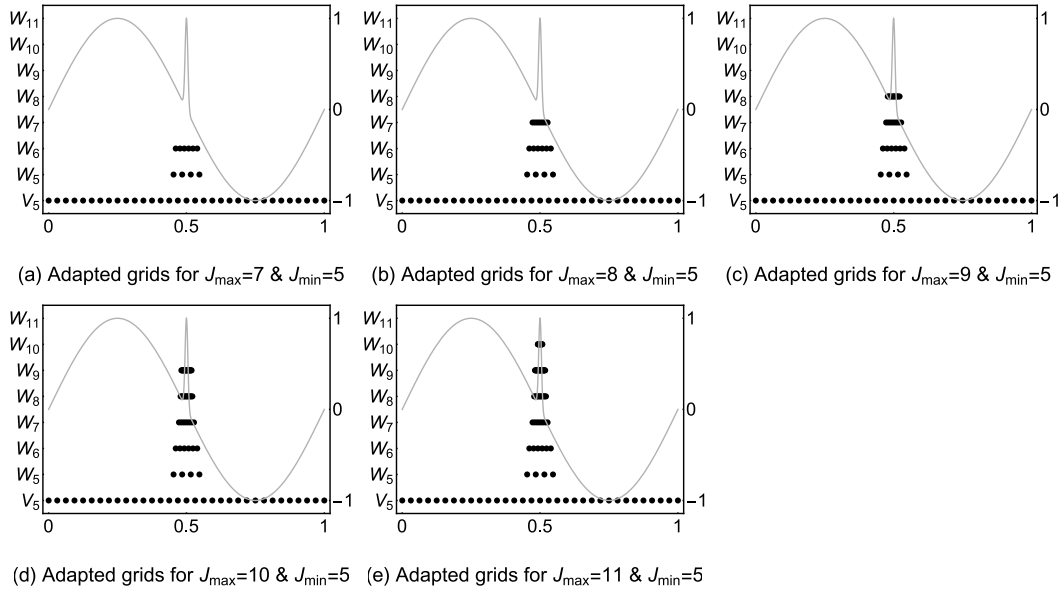


Figure 5.8: The function  $f_1(x)$  (the solid lines) and corresponding adapted grid points (the plot markers  $\bullet$ ) in different resolution levels for different  $J_{\max}$  values; in all illustrations,  $J_{\min} = 5$  and  $\epsilon = 0.5 \times 10^{-4}$ .

**The second test problem** In this problem, a smooth function including a discontinuity is assumed, as [183]:

$$f_2(x) = \text{Sin}(2\pi x) - \text{H}(x - 1/2) + 1/2, \quad \text{for } x \in [0, 1]. \quad (5.67)$$

In this problem, it is assumed that:  $J_{\min} = 5$ ,  $J_{\max} \in \{7, 8, 9, 10, 11, 12\}$  and  $\epsilon = 0.5 \times 10^{-4}$ . The function  $f_2(x)$  and corresponding adapted grid points (at different resolution levels) are presented in Figure 5.9 for different  $J_{\max}$  values. In this figure, the solid line represents  $f_2(x)$  and the plot markers  $\bullet$  indicate the adapted points in different resolution levels. The errors are presented in Table 5.4.2 for  $p_0 = 1$ . The results confirm the reasonable performance of the UNO limiter in comparison to the GMM limiter.

#### 5.4.4 The Jameson slope limiter, a TVB limiter

There are some other limiters with different designing philosophies developed originally for working on non-uniform grids. One example is the *Jameson* slope limiter,  $(\phi_j D_j)_{Ja}$ ; the corresponding limited slope can be defined as:



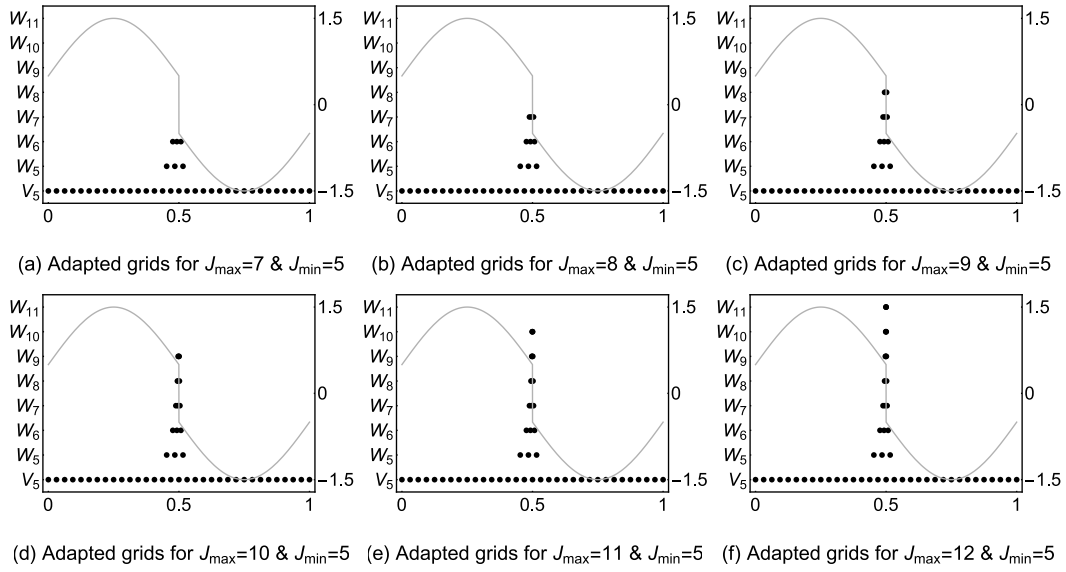


Figure 5.9: The function  $f_2(x)$  (the solid lines) and corresponding adapted grid points (the plot markers  $\bullet$ ) in different resolution levels for different  $J_{\max}$  values; in all illustrations,  $J_{\min} = 5$  and  $\epsilon = 0.5 \times 10^{-4}$ .

Table 5.4.2: Reconstruction error for  $f_2(x)$  at cell-interfaces for the norms with  $p_0 = 1$

$J_{\max}$	MM limiter (GMM: $\theta = 1$ )	GMM limiter $\theta = 2$	UNO limiter By MM	UNO limiter By $m(.,.)$
7	0.0399457	0.0362582	0.033061	0.033061
8	0.0367327	0.0326647	0.0305192	0.0305192
9	0.0339901	0.0299104	0.0283381	0.0283381
10	0.0316237	0.0276812	0.0264477	0.0264477
11	0.0295635	0.0258092	0.0247941	0.0247941
12	0.0277545	0.0241977	0.0233352	0.0233352

$$\left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a} := R(\Delta_-u_{j+2}, \Delta_-u_j) \left[ \frac{(\Delta_-u_{j+2} + \Delta_-u_j) / \Delta x}{2} \right], \quad (5.68)$$

where:  $\Delta_-u_{j+2} := u_{j+2} - u_{j+1}$ ;  $\Delta x$  denotes an uniform sampling step; and the reduction factor  $R$  is:

$$R(\Delta_-u_{j+2}, \Delta_-u_j) := 1 - \left| \frac{\Delta_-u_{j+2} - \Delta_-u_j}{|\Delta_-u_{j+2}| + |\Delta_-u_j|} \right|^q, \quad (5.69)$$

where  $q$  is a positive constant, wherein  $q \geq 2$ . Reconstructed solutions  $u_{j+1/2}^L$  and  $u_{j+1/2}^R$  can be obtained as:  $u_{j+1/2}^L = u_j + (\Delta x/2) \left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a}$  and  $u_{j+1/2}^R = u_{j+1} - (\Delta x/2) \left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a}$ .

In the definition of  $\left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a}$ , indeed, slopes from left and right side of  $x_{j+1/2}$  are averaged (i.e., slopes  $(\Delta_-u_{j+2}/\Delta x)$  and  $(\Delta_-u_j/\Delta x)$ ) and then it is limited by the reduction coefficient  $R$ .

It is possible to use directly Eq. (5.68) for non-uniform grids, where in this case, for  $\Delta x$ , an average value of  $\Delta x_j$  can be used, as:  $\Delta x = [\Delta_-x_{j+2} + \Delta_-x_j] / 2$ . This directly using will lead to a more dissipative but stable solutions. The limiter performance can be improved by using real slope definitions in Eq. (5.68), as:

$$\left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a}^m := R(D_-u_{j+2}, D_-u_j) \frac{D_-u_{j+2} + D_-u_j}{2}, \quad (5.70)$$

where  $D_-u_{j+2} := \Delta_-u_{j+2} / \Delta_-x_{j+2}$ . The reconstructed state values become:  $u_{j+1/2}^L = u_j + (x_{j+1/2} - x_j) \left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a}^m$  and  $u_{j+1/2}^R = u_{j+1} - (x_{j+1} - x_{j+1/2}) \left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a}^m$ .

It is easy to show that both definitions (5.68) and (5.70), preserve the linearity condition even on nonuniform grids; this means solutions will have second order accuracy.

Regarding simulations on adapted grids, both definitions  $\left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a}^m$  and  $\left(\phi_{j+1/2}D_{j+1/2}\right)_{J_a}$  remain stable. For cases without the post-processing stage, however, these limiters also lead to erroneous adapted grid points and non-monotone preserving solutions. The post-processing stage of adapted grids can remedy this problem (this will be studied numerically).

## 5.5 Non-convex problems

In this section, difficulties encountered in numerical simulations of non-convex hyperbolic problems will be reviewed [30]. For convex or concave fluxes,  $F''(u) := \partial^2 F / \partial u^2$  has the same sign everywhere. This feature is important since  $F'(u)$  is varying monotonically as  $u$  varies. And this leads to developing of rarefaction and shock waves separately in different locations. For the non-convex systems, solutions can be more complex and can involve both shock and rarefaction in a specific location [178]. By some numerical examples, it will be shown that even though solutions have convergence (to weak form solutions), due to existence of complex waves in these problems, they are not physical (real) solutions. This kind of systems can explain important phenomena, such as: Euler equations of gas dynamic with a non-convex flux, polymer system used for simulation of polymer flooding processes in enhanced oil recovery, mechanical wave equations with non-convex fluxes and problems involving the phase-transition phenomenon.

In numerical simulation section, it will be illustrated that even spatial adaptation does not guarantee convergence to real solutions. For remedy this problem, using of adaptive limiters (adaptive schemes) are recommended [30]. In this case, most dissipative limiters are used around sharp transitions (discontinuities) while less dissipative limiters in remaining smooth regions. This idea can easily be integrated with the wavelet-based adaptation algorithms: using most dissipative limiters around highly concentrated adapted grid points.

## 5.6 Numerical examples

The following examples are to study the effectiveness of the proposed method concerning nonlinear 1-D and 2-D first order hyperbolic systems. The main assumptions are: 1- Applying the D-D interpolating wavelet of order 3; 2- Using the generalized MINMOD flux/slope limiter in all problems; 3- Repeating re-adaptation processes every time step; 4- Using the semi-discrete form of central and central-upwind schemes; 5- Integrating in time by the TVD Runge-Kutta second-order solver.

**Performance of the TVD second-order limiters by the Burgers' equation** Let us consider the Burgers' equation, defined in the previous section defined on  $x \in [0, 1]$  with IC:  $u(x, t = 0) = \sin(2\pi x) + \sin(\pi x)/2$ . For numerical simulations, it is assumed that:  $J_{max} = 11$ ,  $J_{min} = 4$ ,  $\epsilon = 0.5 \times 10^{-3}$  a  $dt = 0.0002$ . For this problem, the

performance of the VL, VA, modified VA with  $n = 2$  and  $n = 10$ , MaxMod,  $(\phi_j D_j)_1$  and  $(\phi_j D_j)_2$  limiters are studied with the second-order central  $KT$  scheme. The solutions are presented in Figure 5.10 at  $t = 0.159$ . The distribution of the corresponding adapted grids in different resolution levels,  $W_5$  to  $W_{10}$  and  $V_5$ , are presented in Figure 5.11. The results confirm that: 1) the MaxMod limiter leads to the most dissipative result around the discontinuity; 2) Adapted points concentrate properly around the discontinuity and high-gradient zones.

**Euler system of equations** The governing system of equations for the Euler gas dynamic problem is presented in Eq. (6.73). In the following, three different problems with different initial and boundary conditions will be studied. These diverse conditions lead to different bench-mark problems, as: 1) the Sod problem [186]; 2) the Lax problem [187]; 3) Interaction of an entropy sine wave with a Mach 3 right-moving front [71].

In the above mentioned system, the entropy function  $\eta(u)$  and the entropy flux  $\psi(u)$  are defined as:

$$\begin{aligned}\eta(u) &= -\rho \log\left(\frac{\rho \cdot e}{\rho^\gamma}\right), \\ \psi(u) &= -v\eta(u),\end{aligned}\tag{5.71}$$

where  $e$  denotes the internal energy per unit mass, defined as:

$$E = \rho e + \frac{1}{2}\rho v^2.\tag{5.72}$$

**The Sod Problem** This problem was developed for studying the performance of different numerical methods. It is a long gas tube divided into two equal parts with a diaphragm; each part contains gas with different features. Its solution includes, from right to left, a *shock* wave, a *contact discontinuity* (the second discontinuity in the  $\rho(t)$ ) and an *expansion zone or rarefaction wave*. In which, the expansion gas is separated from compressing gas by the contact discontinuity, and the rarefaction wave is a continuous process that the high pressure gas flows to the low pressure domain.

Corresponding initial conditions are:

$$\begin{pmatrix} \rho \\ u \\ P \end{pmatrix}_{t=0} = \begin{cases} \{0, 0, 1\}^T, & x \leq 0.5, \\ \{0.125, 0, 0.1\}^T, & x > 0.5. \end{cases}\tag{5.73}$$

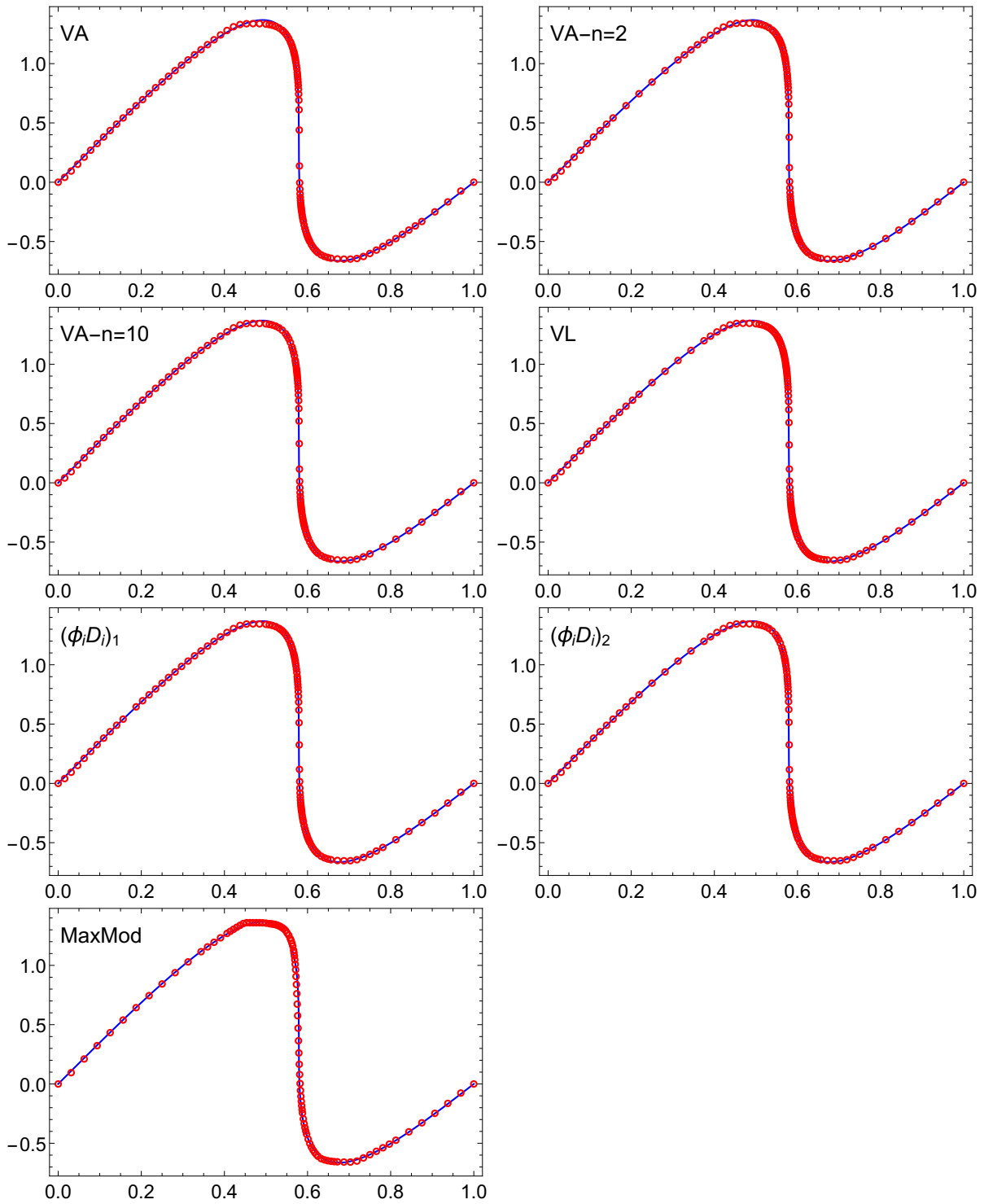


Figure 5.10: Numerical solutions of the Burgers' equation with different TVD second-order limiters at  $t = 0.159$ .

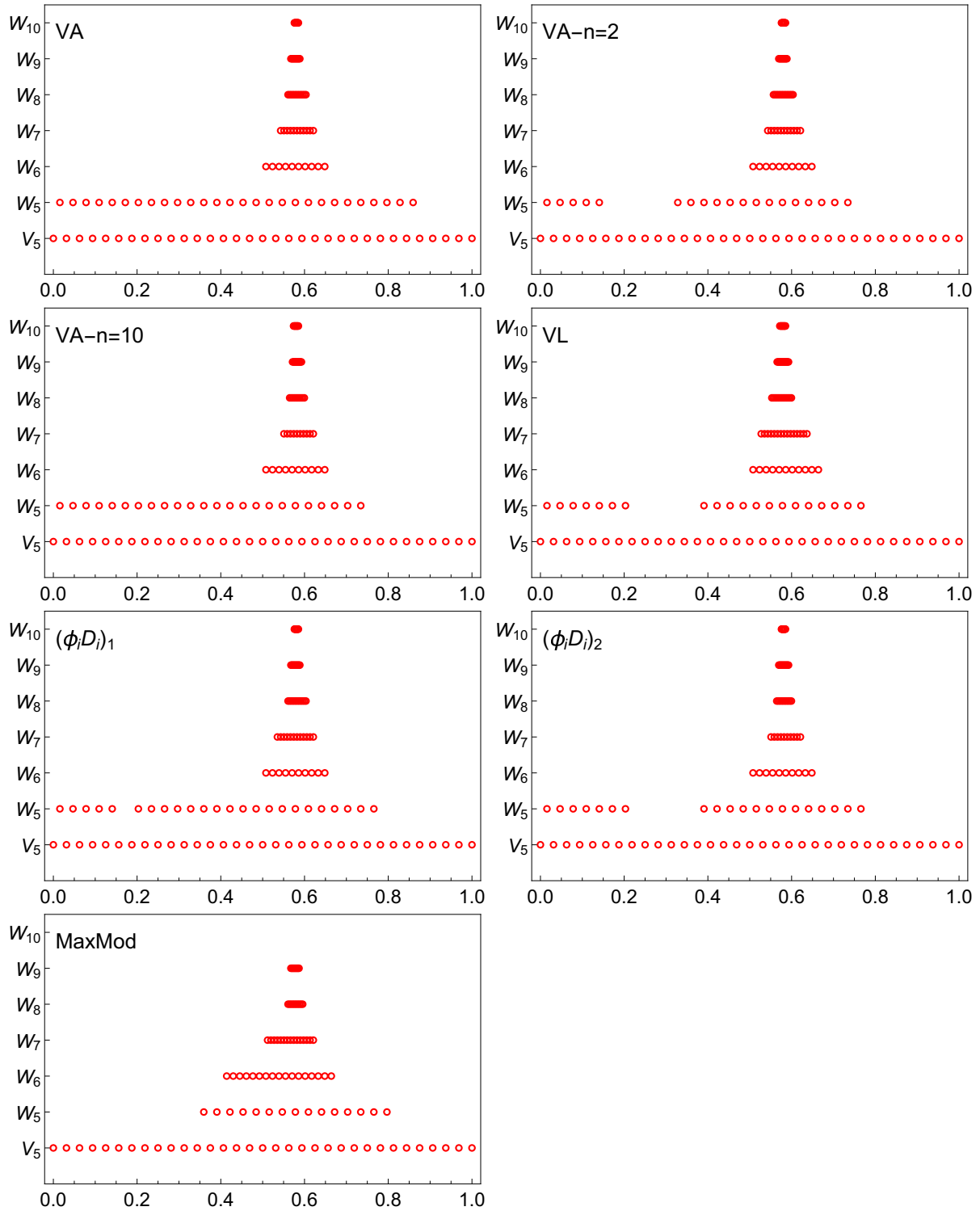


Figure 5.11: Distribution of adapted cells in different resolution levels for the Burgers' equation obtained with different TVD second-order limiters at  $t = 0.159$ .

An unbounded 1-D domain is assumed: a Riemann problem. Assumptions for numerical simulations are:  $J_{max} = 11$ ,  $N_d = 6$ ,  $\epsilon = 10^{-3}$ ,  $\theta = 2$ ,  $N_c = 2$ ,  $N_s = 1$  and  $dt = 10^{-5}$ . The numerical solutions and corresponding adapted grid points are illustrated in Figure 5.12 at time 0.2 with methods KT,  $M1$  and  $M2$ . For the KT method, results are presented in Figures 5.12(a)-(c); these figures contains numerical results, corresponding entropy productions ( $S_j^n$ ), adapted grids and local truncation errors ( $E_j^n$ ). These results for the  $M1$  and  $M2$  schemes are provided respectively in Figures 5.12(d-f) and Figures 5.12(g-i).

Figure 5.12 provides that: 1) Methods  $M1$  and  $M2$  have less numerical dissipation in comparison with the KT scheme; 2) The  $M2$  scheme leads to the smallest dissipation; 3) Using a less dissipative method, more grid points concentrate automatically in different resolutions. This is confirmed by comparing  $N_g$  values of these three methods during simulations, see Figures 5.12(b,e,h) and Figure 5.13.

The numerical entropy production cannot detect the contact discontinuity in this example (Figures 5.12(a),(d) and (g)), even by less dissipative methods with fine enough resolutions. Entropy  $S_j^n$  has small values in the rarefaction zone (for  $0.25 < x < 0.45$ ), but can properly detect shock waves. The local truncation error can capture both the shock wave and contact discontinuity. The local errors  $E_j^n$  have considerable values in rarefaction zones. This zone is not detected by the wavelet theory; as a result, grid points do not adapted there. Considering the wavelet-based adapted points, results of  $S_j^n$ , and  $E_j^n$ , different criteria lead to different adapted grids. In this example, it seems that wavelet-based adaptation method leads to more realistic adapted grids.

Effects of the post-processing stage are investigated by some numerical simulations in the following. At first, effect of considering the post-processing stage is studied. Two simulations with and without the post-processing stage are done and results are presented in Figure 5.14. Figures 5.14(a-b) and Figures 5.14(c-d) include solutions with and without the post-processing step, respectively. The results indicate that post-processing adapted grids have significant effects on solution stability. The numerical instability grows rapidly in absence of the post-processing step.

The post-processing stage contains both grid modification in the same resolution and successive coarser resolution, see Section 3.2. To study effects of them, two types of modifications are considered: full and partial grid modification (by a post-processing). The modifications are: 1) Partial post processing: for a grid point having resolution  $j$ , new points are only added at the corresponding resolution level; here we assume:  $N_s = 2$  and  $N_c = 0$ ; 2) Full post processing: both resolution level  $j$  and  $j + 1$  are controlled;

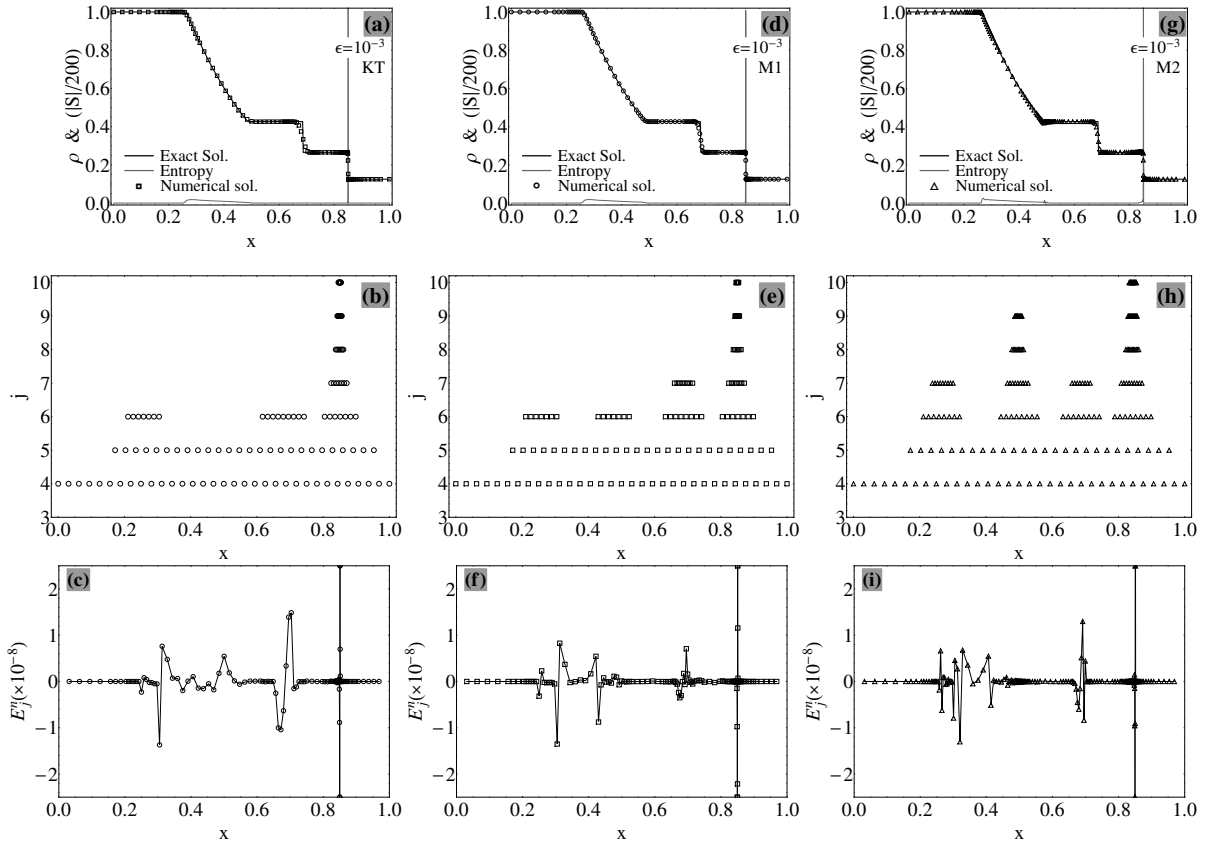


Figure 5.12: Numerical results, corresponding truncation errors, entropy productions, and adapted points in different resolutions for the Sod problem with the KT,  $M1$  and  $M2$  schemes at  $t = 0.2$ ; a-c) the KT scheme; d-f) the  $M1$  method; g-i) the  $M2$  scheme.

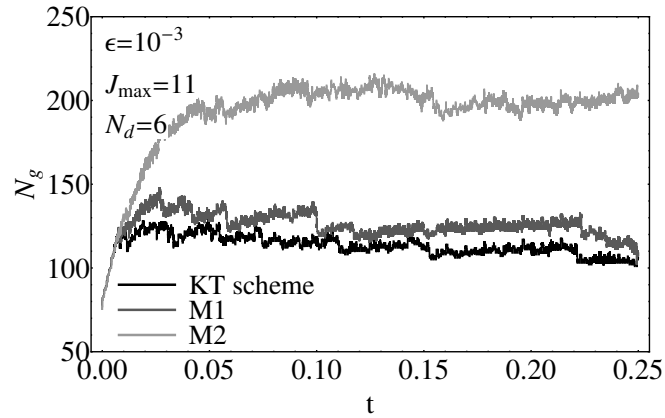


Figure 5.13: Number of adapted grid points  $N_g$  during simulations. In the finest resolution, number of grid points is  $2^{11} + 1$ .



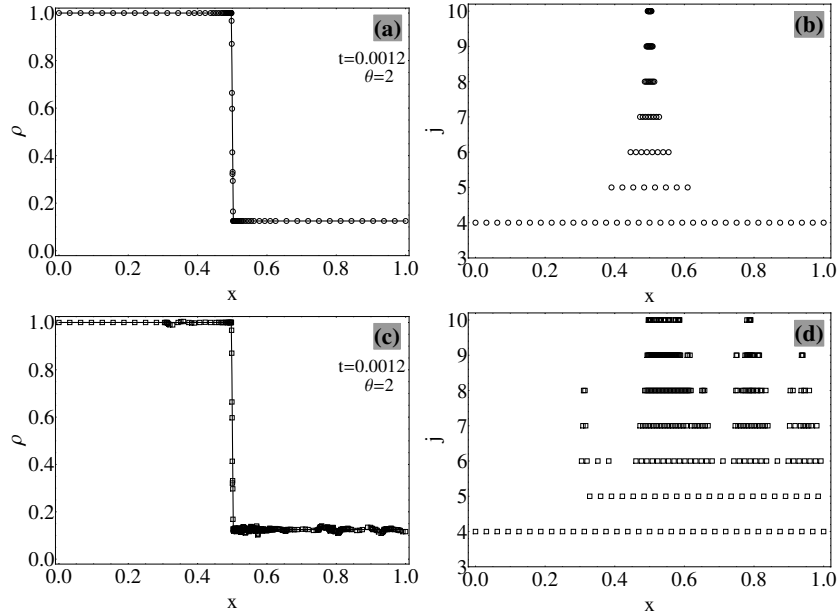


Figure 5.14: Post-processing effects on stability of solutions; a-b) with the post-processing stage; c-d) without the post-processing step. In figures (a) and (c), solid lines and hollow shapes are exact and numerical solutions, respectively.

we choose:  $N_s = 2$  and  $N_c = 1$ . The former, the semi-modification, is frequently used in wavelet-based adaptation procedures. Numerical results are presented in Figure 5.15. Figure 5.15(a) and Figure 5.15(b) correspond to methods using the full and partial post-processing stages, respectively. This figure shows that long-term stability can be obtained in case of having full-modification, 5.15(a).

**The Lax Problem** In the solution of this problem, from right to left, a shock wave, a contact discontinuity and a rarefaction zone are developed. Amplitude of the shock wave and the contact discontinuity are larger than those of the Sod problem. In brief, the Sod and the Lax problems are used as benchmarks with the different values of shock waves, contact discontinuities and rarefaction zones.

The initial conditions are:

$$\begin{pmatrix} \rho \\ u \\ P \end{pmatrix}_{t=0} = \begin{cases} \{0.445, 0.69887, 3.5277\}, & x \leq 0.5, \\ \{0.5, 0, 0.571\}, & x > 0.5, \end{cases} \quad (5.74)$$

and the problem is a Riemann problem. For simulations, it is assumed:  $\epsilon = 10^{-3}$ ,  $\theta = 2$ ,

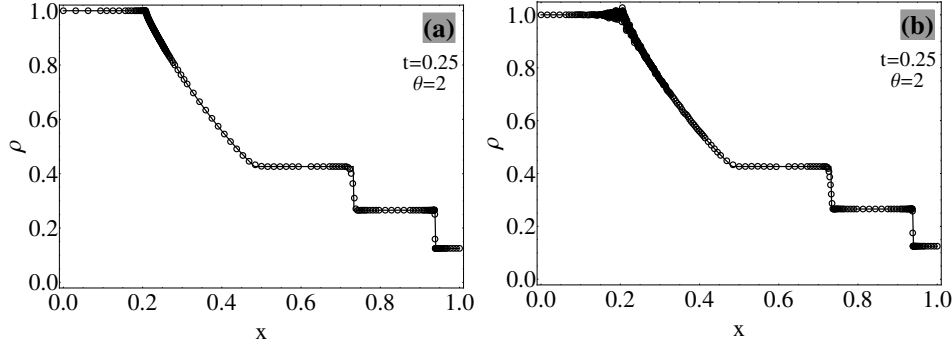


Figure 5.15: Full and semi post-processing effects; a) full post processing,  $N_s = N_c = 2$ ; b) semi-post processing,  $N_s = 2$  and  $N_c = 0$ . Solid lines and hollow shapes are exact and numerical solutions, respectively.

and  $dt = 10^{-5}$ .

For the three methods KT,  $M1$  and  $M2$ , solutions  $\rho$ , corresponding entropies  $S_j^n$ , adapted grids, and truncation errors  $E_j^n$  are presented in Figure 5.16 at time 0.16. Same as the Sod problem, the results offer that: 1) Less dissipative methods mobilize more adapted grid points of fine resolutions; 2) Numerical entropy production of these methods cannot detect the contact discontinuity; 3) The local truncation errors  $E_j^n$  can detect the contact discontinuity zones; 4)  $E_j^n$  can also detect rarefaction zones; 5) The wavelet transform can properly capture all phenomena: shock waves, rarefaction and contact discontinuity zones.

### Interaction of an entropy sine wave with a Mach 3 right-moving front

#### i) The performance of the KT, $M1$ and $M2$ schemes on MRA-based adapted cells

This challenging problem was developed to reveal high order scheme capabilities by Shu and Osher [71]. Here, it is assumed the ratio of specific heats is  $\gamma = 1.4$ . The Riemann initial conditions are presented in Eq. (4.93) [71, 175].

The considered computational domain is:  $\Omega \in (-5, 5) \times (0, T)$ ; Assumed parameters are:  $\epsilon = \epsilon_0 = 5 \times 10^{-3}$ ,  $J_{max} = 11$ ,  $J_{min} = 5$  (or  $N_d = 6$ ),  $N_c = 1$ ,  $N_s = 2$ ,  $\theta = 2$ ,  $dt = 0.00025$ .

The numerical entropy production, numerical and exact solutions are illustrated in Figure 6.23 at  $t = 1.8$ . There, the solid lines and hollow shapes are the reference [175] and numerical solutions, respectively. Regarding numerical entropy productions, it is clear that both the  $M1$  and  $M2$  methods lead to less numerical dissipation in comparison

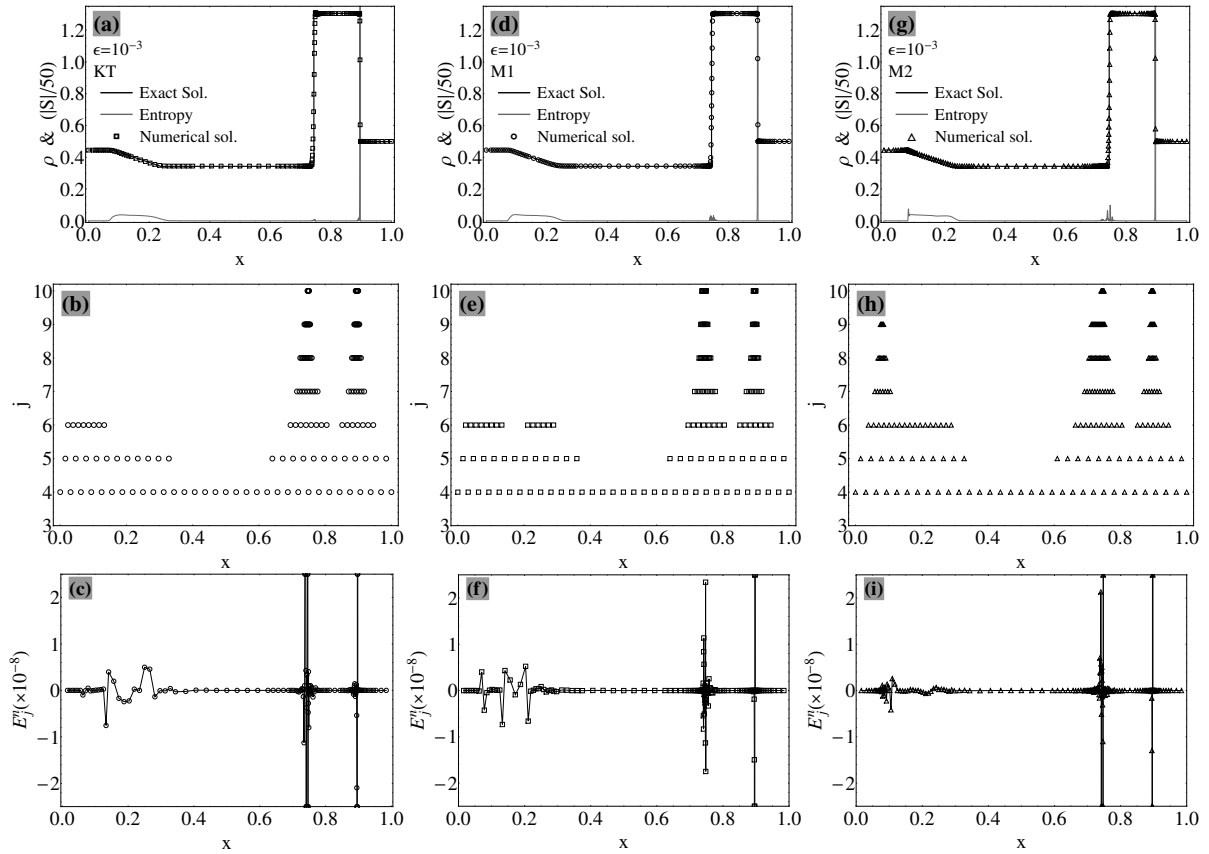


Figure 5.16: Numerical results, corresponding entropy production, truncation errors, and adapted points of different resolutions for the Lax problem at  $t = 0.1$ ; a-c) the KT scheme; d-f) the  $M1$  method; g-i) the  $M2$  scheme.

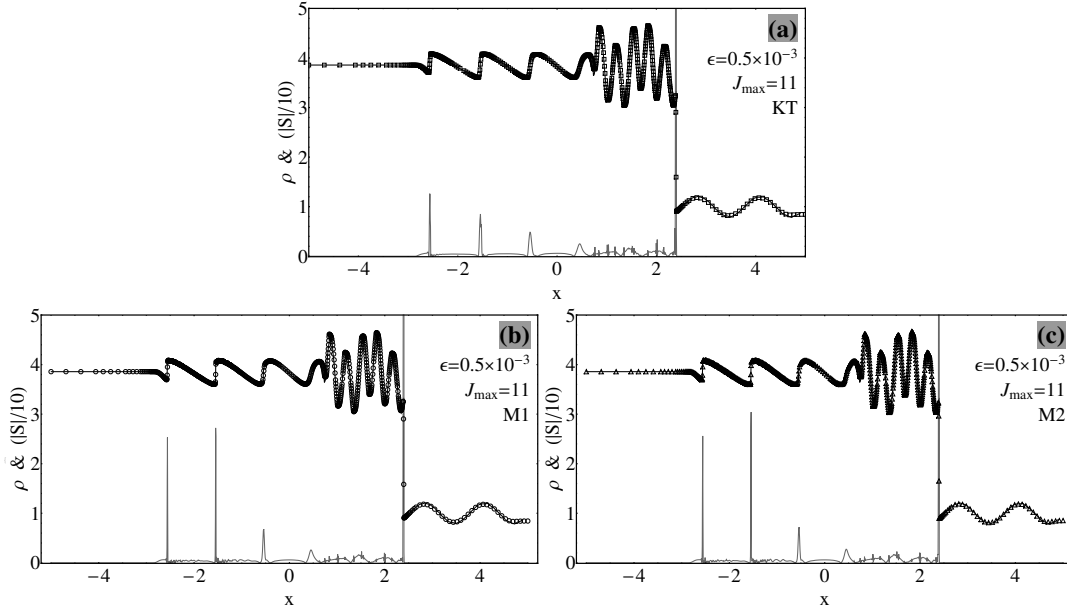


Figure 5.17: Numerical results, and entropy productions for the right propagating front with the KT,  $M1$  and  $M2$  schemes at  $t = 1.8$ . In these figures, solid black lines are the reference solutions, hollow shapes are numerical ones, and the gray solid lines are numerical entropy productions.

to the KT scheme. The  $M2$  scheme leads to the least dissipative results, since magnitude of entropy  $S$  is larger than both the KT and  $M1$  methods. Distribution of adapted points at different resolution levels for these three schemes are shown in Figure 6.24 at  $t = 1.8$ . The methods  $M1$  and  $M2$  lead to more adapted points of high resolutions. In Figure 5.19 local truncation errors for these three methods are presented, which confirm numerical convergence.

**ii) The performance of the KT scheme with the MM-based UNO limiter on MRA-based adapted cells**

This problem is re-simulated with the KT scheme with both GMM (with  $\theta = 2$ ) and the second-order MM-based UNO limiters. The results are presented in Figure 5.20 for the density  $\rho$  at  $t = 1.8$ . In this illustration the solid line and hollow circle markers denote the reference and numerical solutions, respectively. The results confirm that the UNO limiter behaves more smoothly while preserving accuracy. This feature leads to the mobilizing of less adapted cells/grids.

**2-D Euler equation of gas dynamics for ideal gases** The governing equation for the 2-D system is:

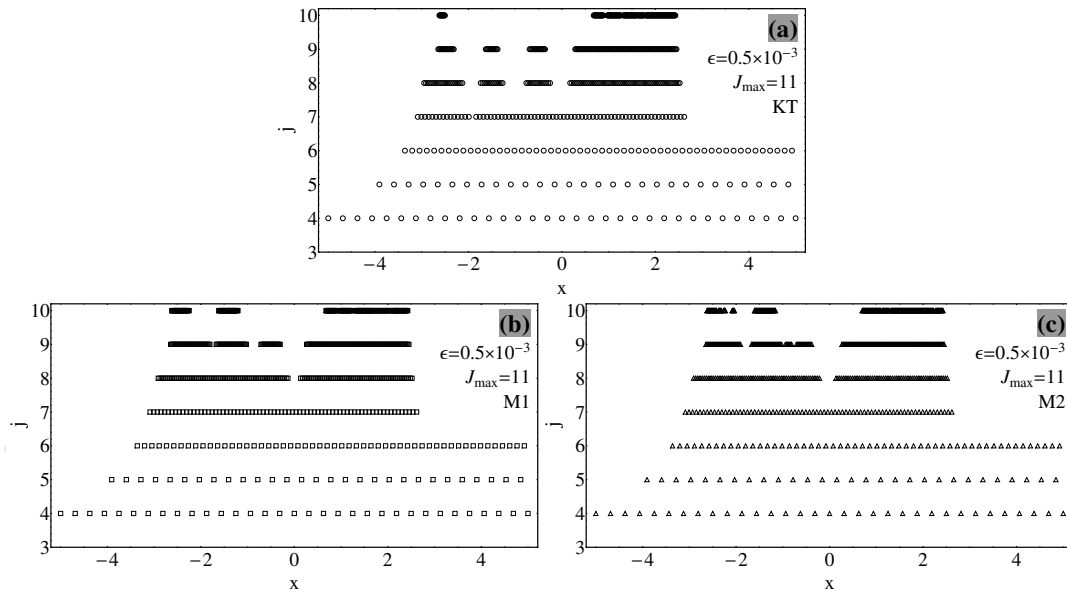


Figure 5.18: Distribution of adapted grid points in different levels of resolution for the KT,  $M1$  and  $M2$  schemes at  $t = 1.8$ .

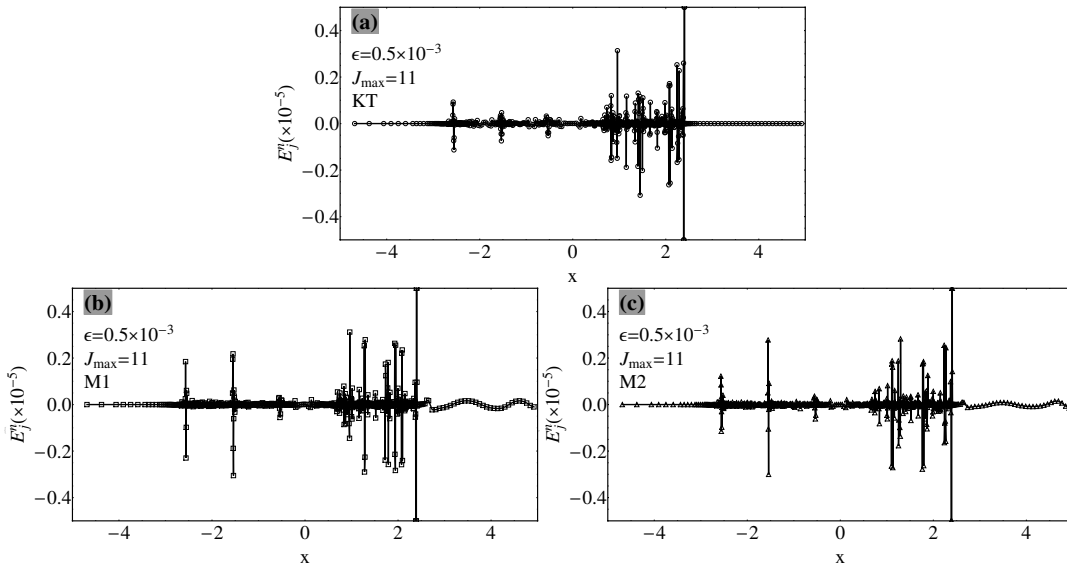


Figure 5.19: Local truncation errors for the right propagating front with the KT,  $M1$  and  $M2$  schemes at  $t = 1.8$ .

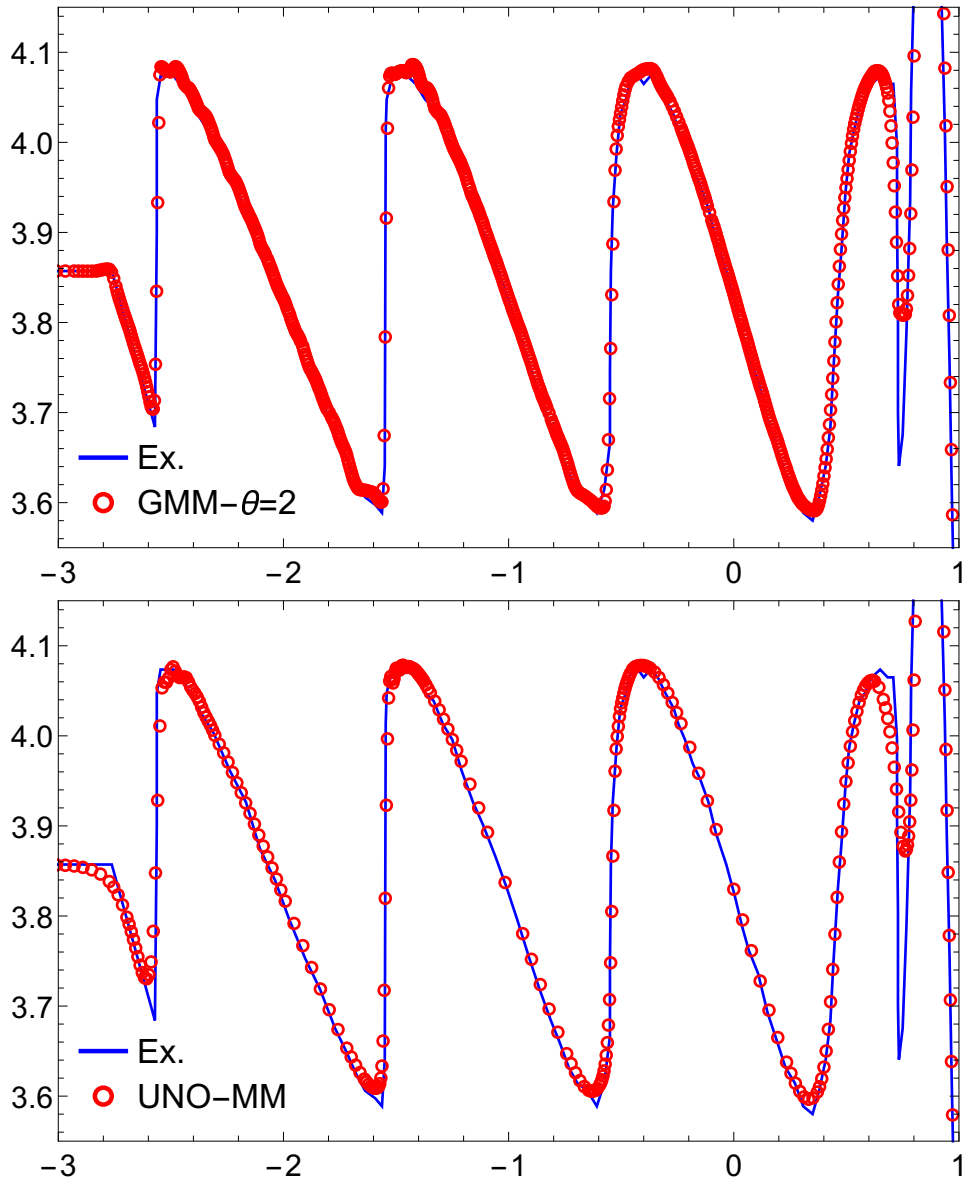


Figure 5.20: Adaptive numerical solutions ( $\rho$ ) for the right propagating front obtained with KT scheme using the GMM and the MM-based UNO limiters at  $t = 1.8$ .

$$\mathbf{u}_t + \mathbf{F}_x + \mathbf{G}_y = \mathbf{0},$$

where state values are  $\mathbf{u} = \{\rho, \rho u, \rho v, E\}^T$ ; the flux vectors in the  $x$  and  $y$  directions are  $\mathbf{F} = \{\rho u, \rho u^2 + P, \rho uv, u(E + P)\}^T$  and  $\mathbf{G} = \{\rho v, \rho uv, \rho v^2 + P, v(E + P)\}^T$ , respectively.

In this example, radially symmetric initial conditions are assumed with respect to the origin which it forms a dense localized high-pressure gas with zero initial velocity. There exists initial higher density and higher pressure inside a circle with radius  $r = 0.4$ ; corresponding values are:  $\{\rho_{in} = 1, \rho_{out} = 0.1\}$  and  $\{P_{in} = 1, P_{out} = 0.1\}$ . Other initial values are:  $u_{in} = u_{out} = v_{in} = v_{out} = 0$  [30, 188]. The computation domain belongs to  $\Omega \in [-1.5, 1.5] \times [-1.5, 1.5]$ . To control the symmetric of solutions and corresponding adapted grids in simulations, total of the computing domain  $\Omega$  is considered. Due to the symmetric initial conditions, the exact solutions are symmetric with respect to the origin. This feature can be used to control the performance of high resolution solvers.

The numerical results are presented in Figure 5.21 at  $t = 0.4$ . Figures 5.21(a) and (b) are from the KT scheme and Figures 5.21(c) and (d) belong to the  $M1$  method. The results offer that: 1) All solutions and corresponding adapted grids are symmetric; 2) Due to numerical dissipation, the KT solver leads to slightly different result from the  $M1$  one. To clarify the numerical dissipation effects, cut of solutions are compared along  $y = 0$ , Figure 5.22. This figure confirms that the KT scheme ends to more dissipative results.

The local truncation errors  $E_j^n$  for the two schemes (KT and  $M1$ ) and corresponding adapted grids are presented in Figure 5.23 at  $t = 0.2$ . It is clear that the errors are properly concentrated in high-gradient zones detected properly by the wavelet transform.

### 5.6.1 Non-convex example: a scalar system

We consider a scalar hyperbolic system with a non-convex flux as:

$$F(u) = \begin{cases} \frac{u(1-u)}{4}, & u < \frac{1}{2}, \\ \frac{u^2}{2} - \frac{u}{2} + \frac{3}{16}, & u \geq \frac{1}{2}. \end{cases} \quad (5.75)$$

Two different initial conditions will be considered.

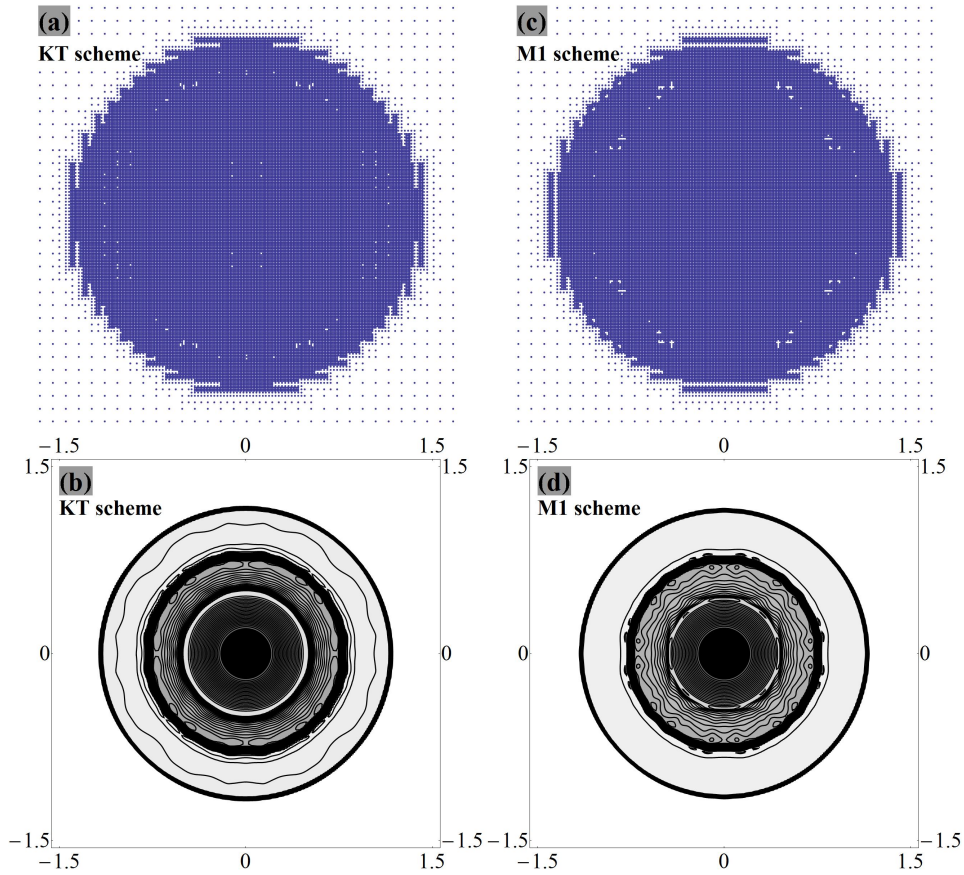


Figure 5.21: The 2-D Euler gas dynamic adaptive solutions with corresponding adapted grids at  $t = 0.4$ .

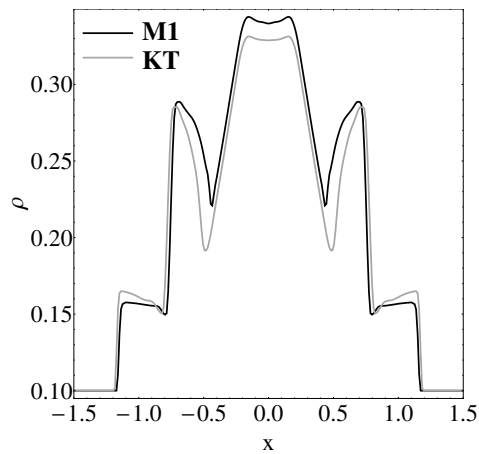


Figure 5.22: Comparison of numerical results obtained by the KT and M1 schemes along  $y = 0$  at  $t = 0.4$ .



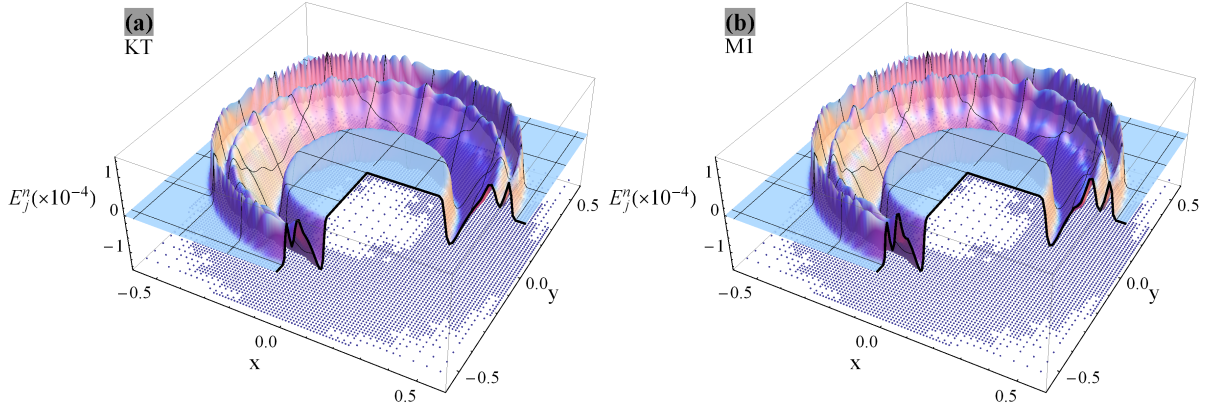


Figure 5.23: Local truncation errors for the KT and  $M1$  schemes at  $t = 0.214$ .

$$u^{(I)}(x, t = 0) = \begin{cases} 0, & x < 0.25, \\ 1, & x \geq 0.25, \end{cases} \quad u^{(II)}(x, t = 0) = \begin{cases} 1, & x < 0.25, \\ 0, & x \geq 0.25. \end{cases} \quad (5.76)$$

In numerical simulations, it is assumed that:  $J_{max} = 11$ ,  $n_d = 6$  (number of decomposition levels),  $\epsilon = 10^{-3}$ ,  $dt = 0.0004$  and  $N_s = N_c = 1$  (for grid modification in the post-processing stage).

For the initial condition  $u^{(I)}(x, 0)$ , the adaptive numerical results and corresponding local truncation errors are shown in Figure 5.24. In Figure 5.24(a), the solution is obtained by the KT scheme with the generalized MINMOD limiter with constant parameter  $\theta = 2$ . It is clear that the adapted solution does not converge to real one; the resulted local truncation errors are presented in Figure 5.24(b).

For Figures 5.24(c) and (d) an adaptive  $\theta$  is assumed; it depends linearly on spatial positions of cell centers  $x_j$ , as:

In numerical simulations two types of  $\theta$  are assumed: constant and adaptive. In adaptive case, in this work, it is assumed  $\theta$  depends linearly on spatial positions of cell centers  $x_j$ , as:

$$\theta(x_j) = 1 + \frac{\{(\Delta x_j + \Delta x_{j+1})/2\} - \Delta x_{min}}{\Delta x_{max} - \Delta x_{min}}, \quad \theta_j := \theta(x_j),$$

where  $\Delta x_j := x_j - x_{j-1}$ ,  $\Delta x_{min} := \min \{\Delta x_j\}$ , and  $\Delta x_{max} := \max \{\Delta x_j\}$ . So, around high-gradient solutions  $\theta \rightarrow 1$  and in smooth regions  $\theta \rightarrow 2$ .

From Figures 5.24(b) and (d), it is clear that the local truncation errors approach zero

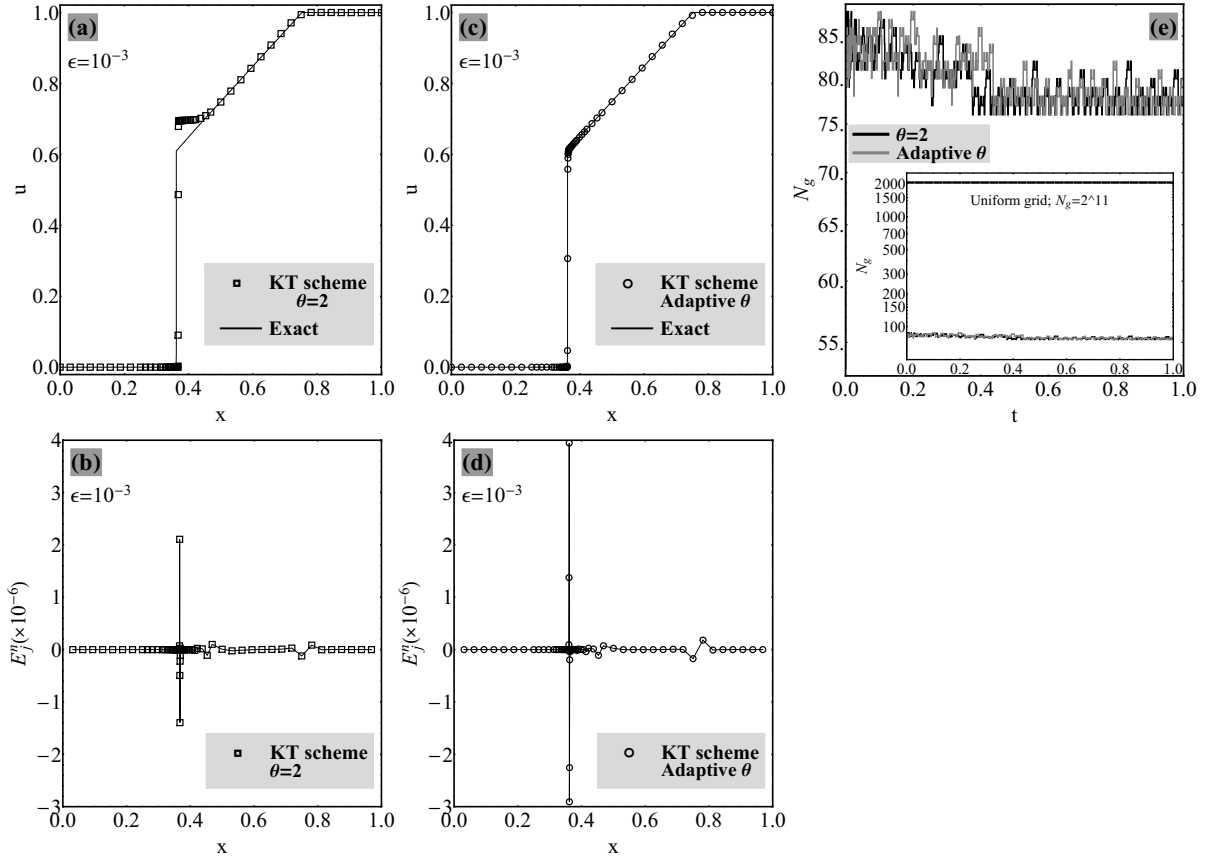


Figure 5.24: Numerical solution of the non-convex conservation law- case 1 with corresponding local truncation error; a, b) constant parameter  $\theta = 2$  and corresponding local truncation error; c,d) with adaptive  $\theta$  and corresponding local truncation error; e) number of adapted grid points during simulations.

in smooth regions; around high gradient zones and discontinuities the local errors are nearly significant. These figures confirm also that the both solutions are converged to weak solutions; however, only the result in Figure 5.24(c) is the physical solution.

It is obvious that using of adaptive smoothing leads to a correct solution. Numbers of adapted grid during time are presented in Figure 5.24(e) for two choice of the limiter parameter.

For the initial condition  $u^{(II)}(x, 0)$ , the numerical results are presented in Figure 5.25; there, Figures 5.25(a,b) and Figures 5.25 (c,d) are for constant and adaptive parameter  $\theta$ , respectively. Numbers of adapted grid points in time are presented in Figure 5.25 (e). Same as the previous example, it is clear that by integrating the adaptive limiter with spatially adaptive grids, an accurate solution can be obtained with small number of grid points. And both solutions are the converged weak solutions, while Figure 5.25 (c)

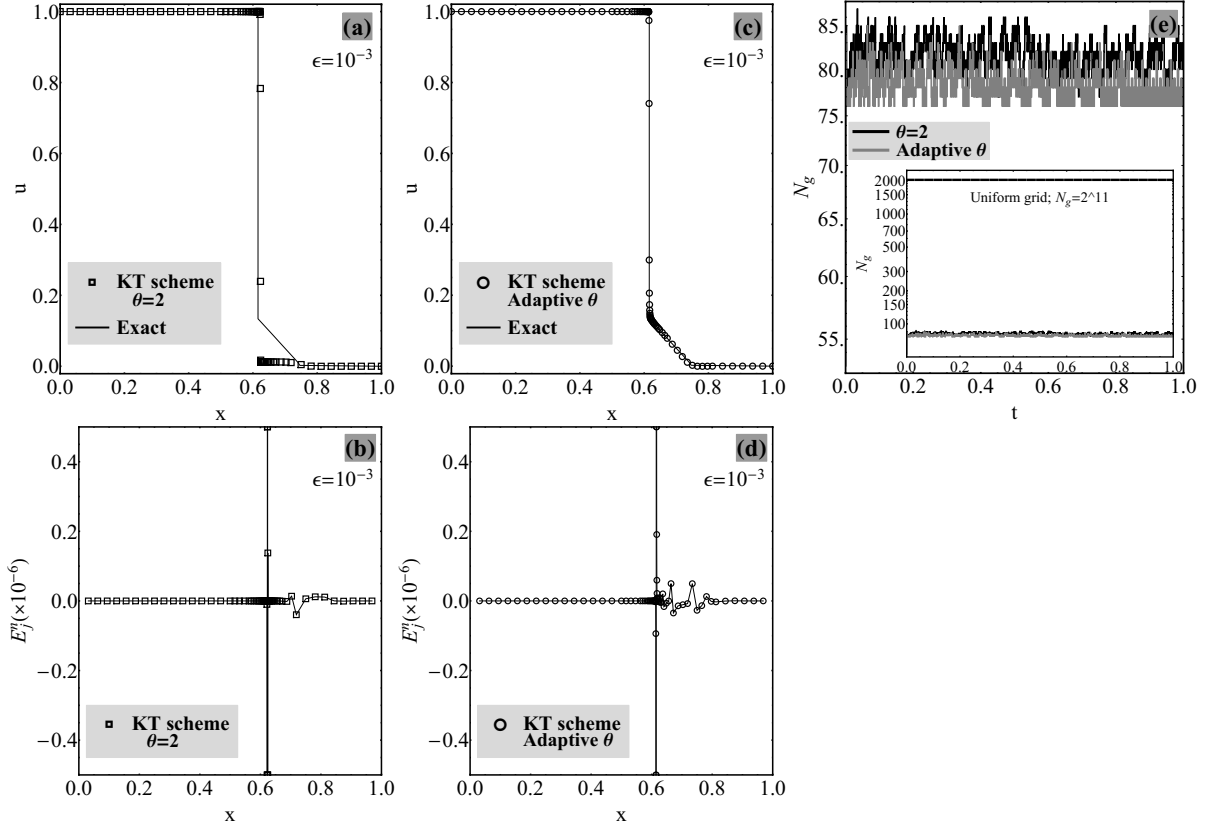


Figure 5.25: Numerical solution of the non-convex conservation law- case 2 with corresponding local truncation error; a, b) constant parameter  $\theta = 2$  and corresponding local truncation error; c,d) with adaptive  $\theta$  and corresponding local truncation error; e) number of adapted grid points during simulations.

represents the physical solution.

**A non-convex example: the polymer system** The governing equation of the polymer system is:

$$\frac{\partial}{\partial t} \begin{pmatrix} s \\ b \end{pmatrix} + \frac{\partial}{\partial x} \begin{pmatrix} f(s, c) \\ cf(s, c) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (5.77)$$

where  $s$  denotes water saturation; parameter  $c$  is the polymer concentration; function  $f := f(s, c)$  presents the fractional flow function of water; parameter  $b$  is function of  $s$  and  $c$  where  $b := b(s, c)$ . Functions  $b$  and  $f$  are assumed to be:  $b(s, c) = sc + a(c)$  and  $f(s, c) = \frac{s^2}{s^2 + (0.5+c)(1-s)^2}$ , where  $a(c)$  denotes the adsorption function and in this example, it is:  $a(c) = \frac{c}{5(1+c)}$ . The exact solution of this problem is complex and contains

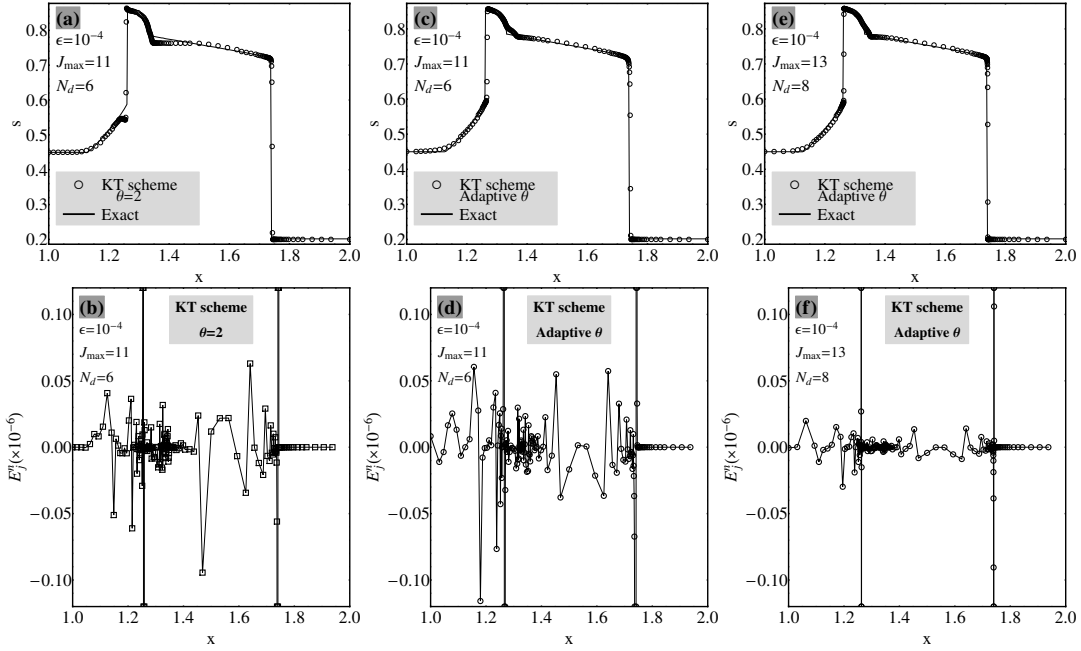


Figure 5.26: Numerical solutions, corresponding adapted grids and local truncation errors for the polymer system; a, b) with  $\theta = 2$ ; c,d) with adaptive  $\theta$ ; e,f) adaptive  $\theta$ .

both simple and composite waves which makes it challenging for numerical solvers [30].

The eigenvalues of the polymer system (Eq. (5.77)) are:  $\lambda_1 = f_s(s, c)$  and  $\lambda_2 = f(s, c)(s + a'(c))^{-1}$ .

For numerical simulations, we assume:  $J_{max} = 11$ ,  $N_d = 6$ ,  $\epsilon = 10^{-4}$ ,  $dt = 0.00025$ , and  $N_s = N_c = 1$  (for grid modification).

To study resolution effect, this example is also re-simulated for a fine resolution with resolution number:  $J_{max} = 13$ . In this case, number of decomposition levels is  $N_d = 8$  and the time step is chosen in such a way that the CFL number does not change.

Numerical results and corresponding local truncation errors for the parameter  $s$  are illustrated in Figure 5.26. The results provide that: 1) For case  $\theta = 2$ , the numerical solution does not converge to real one, even though it is a converged weak solution (Figure 5.26(b)); 2) By using an adaptive scheme with adaptive  $\theta$ , the result is nearly in accordance with the real solution, Figure 5.26(c); 3) By increasing  $J_{max}$  values (or using finer resolutions), the numerical solution approaches to the reference one, Figure 5.26(e); 4) All results are the converged weak solutions due to errors  $E_j^n$ ; 5) using more higher resolution level  $J_{max}$ , smaller local truncation errors are.

**2-D scalar conservation laws with non-convex fluxes** Let us assume a rotating wave with non-convex fluxes, defined as:

$$u_t + \{\sin(u)\}_x + \{\cos(u)\}_y = 0,$$

where  $u := u(x, y, t)$  and the initial condition is:

$$u(x, y, t = 0) = \begin{cases} \frac{14\pi}{4}, & x^2 + y^2 < 1, \\ \frac{\pi}{4}, & x^2 + y^2 \geq 1. \end{cases}$$

This benchmark test, originally proposed in [30], is a challenging 2-D problem for many high resolution schemes due to the developing of 2-D composite waves in its solution.

For numerical simulations, we have:  $J_{max} = 8$ ,  $J_{min} = 5$  (or  $N_d = 3$ ),  $\epsilon = 10^{-4}$ ,  $dt = 0.5 \times 10^{-3}$ , and  $N_s = N_c = 1$  (for modification of adapted grid ). For modeling, two different choices of  $\theta$  are assumed: 1) The constant one with value  $\theta = 2$ ; 2) The adaptive implementation of  $\theta$ . The latter is also based on the 1-D linear interpolation of  $\theta$  on adapted grid points, as:

$$\theta(Z_j) = 1 + \frac{(\Delta Z_{ave})_j - \Delta Z_{min}}{\Delta Z_{max} - \Delta Z_{min}},$$

where  $(\Delta Z_{ave})_j := \{(\Delta Z_j + \Delta Z_{j+1})/2\}$ ,  $\Delta Z_j := Z_j - Z_{j-1}$ ,  $Z_j \in \{x_j, y_j\}$ ,  $\Delta Z_{max} = 1/2^{J_{min}}$ , and  $\Delta Z_{min} = 1/2^{J_{max}}$ . For each direction,  $\theta(Z_j)$  is calculated independently.

The numerical results and corresponding adapted grid points are shown in Figure 5.27 at  $t = 1$ . Figures (a, b) correspond to the  $\theta$ -adaptive results and figures (c, d) are from the constant  $\theta$ . Again the  $\theta$ -adaptive solver converges to proper and physical results [30].

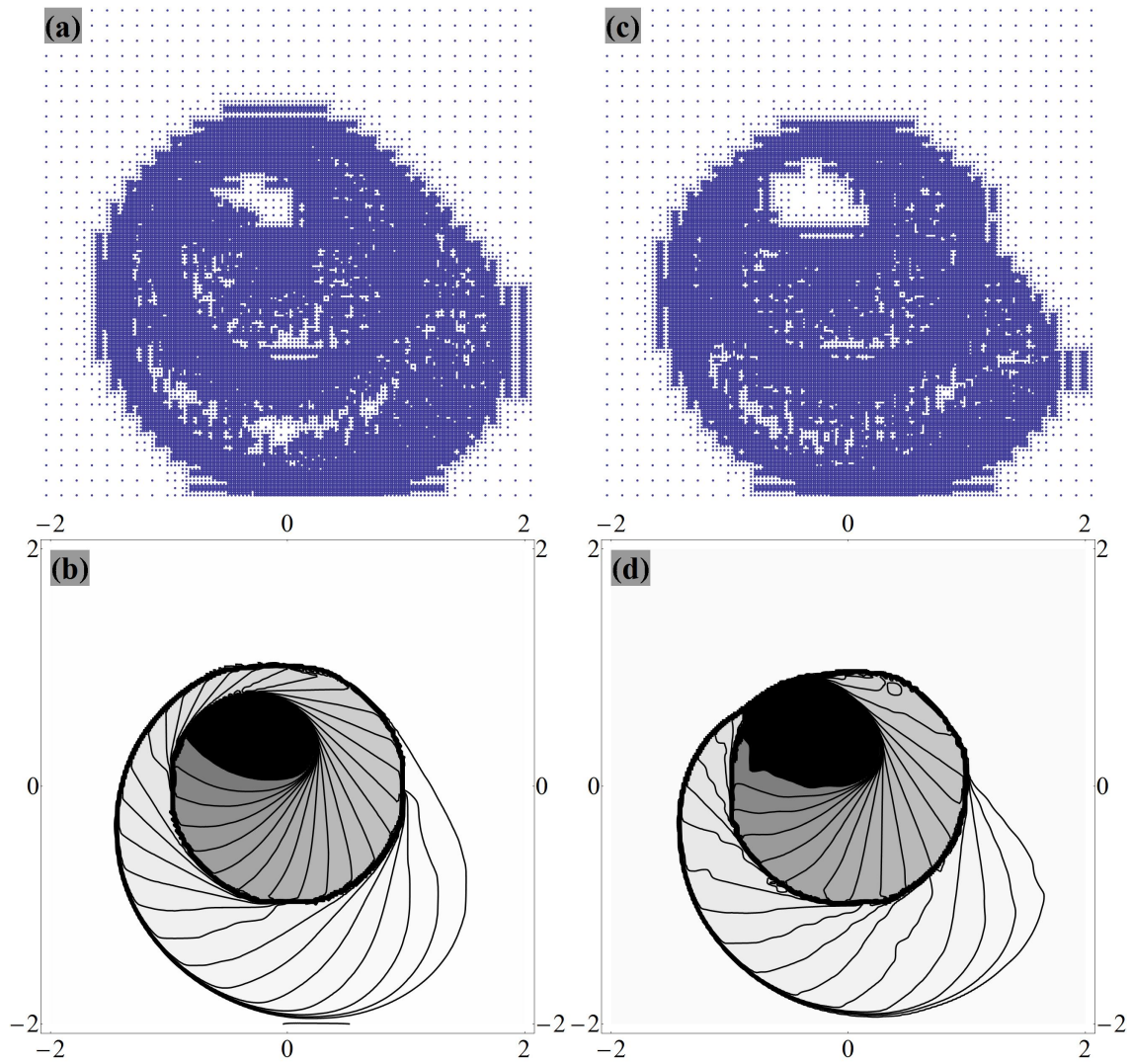


Figure 5.27: Adaptive solutions of 2-D non-convex conservation law system at  $t = 1$ ; a, b) obtained with  $\theta$ -adaptive KT scheme; c, d) based on the KT scheme with  $\theta = 2$ .

# Chapter 6

## Higher order central and CWENO schemes

**Introduction** The aim of this chapter is a stable integration of higher-order high resolution schemes with multiresolution-based adapted grids. Such cooperation is important since higher order high resolution schemes are computationally expensive and adaptive solvers can improve both computing efficiency and numerical accuracy. The nonlinear stability condition is the TVB condition [93]. This TVB feature results from satisfying a local maximum principle and a non-oscillatory property.

This chapter contains:

1. Section 6.1 devotes to third-order central high resolution schemes based on average-interpolating quadratic polynomials, originally developed for uniform cells in [93], [76] and [92]. The TVB non-oscillatory property is enforced by considering a nonlinear limiter [76]. In this section, at the first, the fully-discrete and semi-discrete forms of the third order high resolution schemes on non-uniform grids will be provided. Main features of the average-interpolating quadratic polynomials  $\{q_j(x)\}$  are explained. Local quadratic average interpolating polynomials ( $q_j(x)$ ) are employed to reconstruct a piecewise quadratic polynomial ( $\{q_j\}$ ); they satisfy the conservation law. In this stage, to guarantee the non-oscillatory property in reconstructed piecewise functions, the initial piecewise function  $\{q_j\}$  is locally modified by using a limiter. And in the projection stage, again, quadratic variation interpolation are assumed in non-smooth zones while constant variations are assumed in smooth regions.
2. Section 6.2 presents the third and fourth order CWENO high resolution schemes on non-uniform grids. In this section, formulations are based on Section 6.1 and

the concept of the CWENO interpolation in the reconstruction stage: the reconstruction is based on a convex combination (weighting) <sup>8</sup> of left-side, central and right-side polynomials to have the TVB feature. There, both the *three point stencil* and the *five point stencil* are re-formulated.

3. Section 6.3 describes piecewise parabolic reconstruction, based on parabolic polynomials using local data. In Section 6.1, data of three successive cells  $I_{j-1}$ ,  $I_j$  and  $I_{j+1}$  are used for the reconstruction stage. However, in this section local data of cell  $I_j$  is used: they are cell average value ( $\bar{u}_j$ ) and two derivative values. Two approaches are proposed based on local data: i) the first approach uses the cell average value and the first ( $u'_j$ ) and second ( $u''_j$ ) derivative values at cell center of cell  $I_j$ ; ii) the second approach uses the cell average value and the first derivatives at cell edges ( $u'_{j\pm 1/2}$ ).

Finally performances of these three approaches are demonstrated by some examples.

## 6.1 Central third order high resolution schemes on non-uniform grids

### 6.1.1 The fully-discrete form

In this section, at first, based on a spatially third-order discretization, a fully-discrete form is provided on non-uniform grids. Based on this formulation, in the limiting state ( $\Delta t \rightarrow 0$ ), the semi-discrete form will be derived [189].

For deriving the fully-discrete form, three stages *reconstruction-evolution-projection* (REP) will be followed [22]. The main concept of the REP procedure is presented in Figure 6.1. To reconstruct the average-interpolating parabolic polynomial ( $q_j(x)$ ) over cell  $I_j$ , three average information  $\bar{u}_{j-1}$ ,  $\bar{u}_j$  and  $\bar{u}_{j+1}$  are used, as presented in Figure 6.2.

#### **The reconstruction stage**

A polynomial-based piecewise reconstruction is assumed on cells  $\{I_j\}$  at time  $t^n$ ; for the cell  $I_j$ , we have:

---

<sup>8</sup>Let the set of vectors  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  spans the subspace  $\mathbf{v}$ :  $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \dots + \lambda_n \mathbf{x}_n \in \mathbf{v}$  for all scalars  $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$  such that  $\sum_i \lambda_i = 1$ . With the above assumptions for  $\lambda_i$ , such combination is called a *convex combination* of vectors  $\mathbf{x}_1, \dots, \mathbf{x}_n$ .



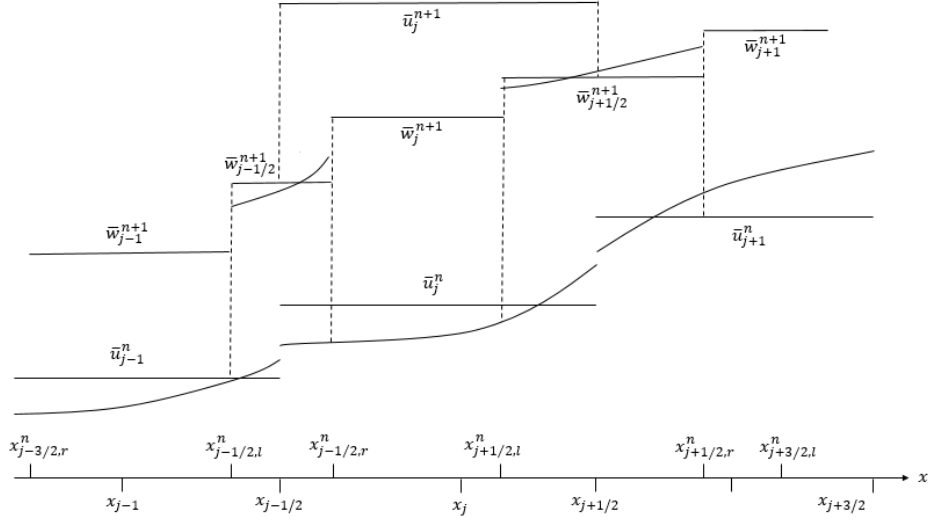
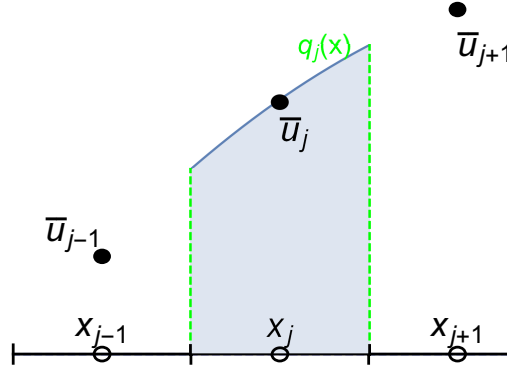


Figure 6.1: The REP procedure for the third-order central scheme.


 Figure 6.2: The three-average solutions used for evaluation of  $q_j(x)$ .

$$P_j(x, t^n) = A_j + B_j(x - x_j) + \frac{1}{2}C_j(x - x_j)^2, \quad x \in (x_{j-1/2}, x_{j+1/2}). \quad (6.1)$$

### The evolution stage

It is assumed that the cell center  $x_j$  is not necessarily located in the middle of the cell  $I_j$ . The location of  $x_j$  can be determined from cell edges  $x_{j\pm 1/2}$  as:  $x_{j+1/2} := x_j + p_j \Delta x_j$  and  $x_{j-1/2} := x_j - (1 - p_j) \Delta x_j$ . Spatial locations  $x_{j\pm 1/2,l}^n$  and  $x_{j\pm 1/2,r}^n$  are also defined as:  $x_{j\pm 1/2,l}^n := x_{j\pm 1/2}^n - a_{j\pm 1/2}^n \Delta t$  and  $x_{j\pm 1/2,r}^n := x_{j\pm 1/2}^n + a_{j\pm 1/2}^n \Delta t$ . Parameter  $a_{j+1/2}^n$  shows the upper bound of propagating speed of a possible discontinuity at cell edge  $x_{j+1/2}$ , as:

$$a_{j+1/2}^n := \max_{u \in \mathcal{C}(u_{j+1/2}^L, u_{j+1/2}^R)} \rho \left( \frac{\partial F(u)}{\partial u} \right), \quad (6.2)$$

where  $\rho(A)$  shows the spectral radius of the matrix  $A$ <sup>9</sup>, that is:  $\rho(A) := \max_i |\lambda_i(A)|$ <sup>10</sup> where  $\lambda_i(A)$  is  $i^{\text{th}}$  eigenvalue of  $A$ ;  $u_{j+1/2}^L$  and  $u_{j+1/2}^R$  denote respectively left and right immediate values of  $u_{j+1/2}$ , i.e.:  $u_{j+1/2}^L := P_j(x_{j+1/2}, t^n)$  and  $u_{j+1/2}^R = P_{j+1}(x_{j+1/2}, t^n)$ ;  $\mathcal{C}(u_{j+1/2}^L, u_{j+1/2}^R)$  represents a curve in phase space connecting  $u_{j+1/2}^L$  and  $u_{j+1/2}^R$  by the Riemann fan.

For genuinely nonlinear or linearly degenerate<sup>11</sup> case,  $a_{j+1/2}^n$  becomes simply:

$$a_{j+1/2}^n := \max \left\{ \rho \left( \frac{\partial F}{\partial u} \left( u_{j+1/2}^L \right) \right), \rho \left( \frac{\partial F}{\partial u} \left( u_{j+1/2}^R \right) \right) \right\}. \quad (6.3)$$

### Computing the integral over the spatio-temporal intervals $[x_{j+1/2,l}^n, x_{j+1/2,r}^n] \times [t^n, t^{n+1}]$

The result of integration is:

<sup>9</sup>The conservation laws  $\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = \mathbf{0}$  can be rewritten as:  $\mathbf{u}_t + \mathbf{A}\mathbf{u}_x = \mathbf{0}$ , where  $\mathbf{A} = \partial \mathbf{f} / \partial \mathbf{u}$ . This form of the conservation laws is known as the *quasilinear* form of the hyperbolic equation and the matrix  $\mathbf{A}$  is the *Jacobian* matrix.

<sup>10</sup>Parameter  $\rho$  denotes the spectral radius; that is, the largest modulus of the eigenvalues of the Jacobian matrix evaluated at cell edges

<sup>11</sup>Let assume a system of nonlinear conservation laws  $\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = \mathbf{0}$ . As mentioned, this system can also be represented as the quasilinear form  $\mathbf{u}_t + \mathbf{A}\mathbf{u}_x = \mathbf{0}$  where  $\mathbf{A} = \mathbf{f}'(\mathbf{u}) := \partial \mathbf{f} / \partial \mathbf{u}$ . Let  $\lambda_i := \lambda_i(\mathbf{A})$  and  $\mathbf{r}_i := \mathbf{r}_i(\mathbf{A})$  denote the  $i$ -th eigenvalue and  $i$ -th eigenvector of  $\mathbf{A}$ , respectively.  $\lambda_i$  defines a characteristic field called  $\lambda_i$ -th or simply  $i$ -th characteristic field.

A  $i$ -th characteristic field is called *linearly degenerate* if:

$\frac{\partial \lambda_i}{\partial \mathbf{u}} \cdot \mathbf{r}_i = 0$  for all  $\mathbf{u}$ .

For linear hyperbolic systems, as their eigenvalues are constant, their fields are linearly degenerate. In the case of nonlinear hyperbolic systems, the contact discontinuity is a discontinuity in a linearly degenerate field: the characteristics are all parallel across this type of discontinuity. This means the characteristic speeds do not change across the discontinuity.

A  $i$ -th characteristic field is called *genuinely nonlinear* if:

$\frac{\partial \lambda_i}{\partial \mathbf{u}} \cdot \mathbf{r}_i \neq 0$  for all  $\mathbf{u}$ .

In this case, characteristic speeds can change across a discontinuity. Nonlinear shocks (converging characteristics) and nonlinear expansions or rarefaction (diverging characteristics) are some examples of genuinely nonlinear characteristic fields.

$$\begin{aligned}
 \bar{\omega}_{j+\frac{1}{2}}^{n+1} &= \frac{1}{2} (A_j + A_{j+1}) \\
 &+ \frac{1}{4} \left[ \Delta t a_{j+\frac{1}{2}}^n (B_{j+1} - B_j) + 2 (B_j \Delta x_j p_j + B_{j+1} (\Delta x_j p_j + x_j - x_{j+1})) \right] \\
 &+ \frac{1}{12} \left[ \Delta t^2 (a_{j+\frac{1}{2}}^n)^2 (C_j + C_{j+1}) + 3 \Delta t a_{j+\frac{1}{2}}^n (C_{j+1} (\Delta x_j p_j + x_j - x_{j+1}) - C_j \Delta x_j p_j) \right. \\
 &+ \left. 3 (C_j \Delta x_j^2 p_j^2 + C_{j+1} (\Delta x_j p_j + x_j - x_{j+1})^2) \right] \\
 &- \frac{1}{2 a_{j+1/2}^n \Delta t} \left[ \int_{t^n}^{t^{n+1}} \left\{ F(u(x_{j+1/2,r}^n)) - F(u(x_{j+1/2,l}^n)) \right\} dt \right].
 \end{aligned} \tag{6.4}$$

**Computing the integral over the spatio-temporal intervals**  $[x_{j-1/2,r}^n, x_{j+1/2,l}^n] \times [t^n, t^{n+1}]$

$$\begin{aligned}
 \bar{\omega}_j^{n+1} &= A_j + \frac{1}{2} B_j \left[ \Delta t (a_{j-\frac{1}{2}}^n - a_{j+\frac{1}{2}}^n) + \Delta x_j (2p_j - 1) \right] \\
 &+ \frac{1}{6} C_j \left[ \Delta t^2 \left( (a_{j-\frac{1}{2}}^n)^2 + (a_{j+\frac{1}{2}}^n)^2 \right) - \Delta t a_{j-\frac{1}{2}}^n (\Delta t a_{j+\frac{1}{2}}^n + \Delta x_j (2 - 3p_j)) \right. \\
 &+ \left. a_{j+\frac{1}{2}}^n \Delta t \Delta x_j (1 - 3p_j) + \Delta x_j^2 (3p_j^2 - 3p_j + 1) \right] \\
 &- \frac{1}{\Delta x_j - (a_{j-1/2}^n + a_{j+1/2}^n) \Delta t} \left[ \int_{t^n}^{t^{n+1}} \left\{ F(u(x_{j+1/2,l}^n)) - F(u(x_{j-1/2,r}^n)) \right\} dt \right].
 \end{aligned} \tag{6.5}$$

**Computing the integral over the spatio-temporal intervals**  $[x_{j-1/2,l}^n, x_{j-1/2,r}^n] \times [t^n, t^{n+1}]$

$$\begin{aligned}
 \bar{\omega}_{j-\frac{1}{2}}^{n+1} &= \frac{1}{2} (A_{j-1} + A_j) \\
 &+ \frac{1}{4} \left[ \Delta t a_{j-\frac{1}{2}}^n (B_j - B_{j-1}) + 2 (B_{j-1} \Delta x_{j-1} p_{j-1} + B_j (\Delta x_{j-1} p_{j-1} + x_{j-1} - x_j)) \right] \\
 &+ \frac{1}{12} \left[ \Delta t^2 (a_{j-\frac{1}{2}}^n)^2 (C_{j-1} + C_j) + 3 \Delta t a_{j-\frac{1}{2}}^n (C_j (\Delta x_{j-1} p_{j-1} + x_{j-1} - x_j) - C_{j-1} \Delta x_{j-1} p_{j-1}) \right. \\
 &+ \left. 3 (C_{j-1} \Delta x_{j-1}^2 p_{j-1}^2 + C_j (\Delta x_{j-1} p_{j-1} + x_{j-1} - x_j)^2) \right] \\
 &- \frac{1}{2 a_{j-1/2}^n \Delta t} \left[ \int_{t^n}^{t^{n+1}} \left\{ F(u(x_{j-1/2,r}^n)) - F(u(x_{j-1/2,l}^n)) \right\} dt \right].
 \end{aligned} \tag{6.6}$$

### The projection stage

After evaluating the average evolved solutions  $\bar{\omega}_{j+\frac{1}{2}}^{n+1}$ ,  $\bar{\omega}_j^{n+1}$  and  $\bar{\omega}_{j-\frac{1}{2}}^{n+1}$  (respectively from Eqs. (6.4)-(6.6)), we can reconstruct a third order average interpolating piece-wise function with the non-oscillatory feature. Such piece-wise functions are denoted as:  $\tilde{\omega}_j^{n+1}(x)$  and  $\tilde{\omega}_{j+\frac{1}{2}}^{n+1}(x)$ ; they are defined as:

$$\begin{aligned}\tilde{\omega}_{j+1/2}^{n+1}(x) &= \tilde{A}_{j+1/2} + \tilde{B}_{j+1/2}(x - x_j) + \frac{1}{2}\tilde{C}_{j+1/2}(x - x_j)^2, \quad \text{for } x \in (x_{j+1/2,l}^n, x_{j+1/2,r}^n), \\ \tilde{\omega}_j^{n+1}(x) &= \bar{\omega}_j^{n+1}, \quad \text{for } x \in (x_{j-1/2,r}^n, x_{j+1/2,l}^n).\end{aligned}\tag{6.7}$$

The projected solution  $\bar{u}_j^{n+1}$  at time  $t^{n+1}$  can be obtained by a projection step, as:

$$\begin{aligned}\bar{u}_j^{n+1} &= \frac{1}{\Delta x_j} \left[ \int_{x_{j-1/2}}^{x_{j-1/2,r}} \tilde{\omega}_{j-1/2}^{n+1}(x) dx + \int_{x_{j-1/2,r}}^{x_{j+1/2,l}} \tilde{\omega}_j^{n+1}(x) dx + \int_{x_{j+1/2,l}}^{x_{j+1/2}} \tilde{\omega}_{j+1/2}^{n+1}(x) dx \right] \\ &= \bar{\omega}_j^{n+1} \left[ 1 - \lambda_j \left( a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n \right) \right] + \lambda_j \left[ \tilde{A}_{j-\frac{1}{2}} a_{j-\frac{1}{2}}^n + \tilde{A}_{j+\frac{1}{2}} a_{j+\frac{1}{2}}^n \right] \\ &\quad + \frac{1}{2} \Delta t \lambda_j \left[ \tilde{B}_{j-\frac{1}{2}} \left( a_{j-\frac{1}{2}}^n \right)^2 - \tilde{B}_{j+\frac{1}{2}} \left( a_{j+\frac{1}{2}}^n \right)^2 \right] + \frac{1}{6} \Delta t^2 \lambda_j \left[ \tilde{C}_{j-\frac{1}{2}} \left( a_{j-\frac{1}{2}}^n \right)^3 + \tilde{C}_{j+\frac{1}{2}} \left( a_{j+\frac{1}{2}}^n \right)^3 \right],\end{aligned}\tag{6.8}$$

where  $\lambda_j := \frac{\Delta t}{\Delta x_j}$ .

### 6.1.2 The semi-discrete form

The limit of the following equation approximates the semi-discrete form:

$$\frac{d}{dt} \bar{u}_j(t) = \lim_{\Delta t \rightarrow 0} \frac{\bar{u}_j^{n+1} - \bar{u}_j^n}{\Delta t},\tag{6.9}$$

or:

$$\begin{aligned} \frac{d}{dt}\bar{u}_j(t) = & \\ \lim_{\Delta t \rightarrow 0} \left\{ \frac{1}{\Delta t} (\bar{\omega}_j^{n+1} - \bar{u}_j^n) - \frac{1}{\Delta x_j} (a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n) \bar{\omega}_j^{n+1} + \frac{1}{\Delta x_j} (\tilde{A}_{j-\frac{1}{2}} a_{j-\frac{1}{2}}^n + \tilde{A}_{j+\frac{1}{2}} a_{j+\frac{1}{2}}^n) \right. & \\ \left. + \frac{1}{2} \lambda_j \left[ \tilde{B}_{j-\frac{1}{2}} (a_{j-\frac{1}{2}}^n)^2 - \tilde{B}_{j+\frac{1}{2}} (a_{j+\frac{1}{2}}^n)^2 \right] + \frac{1}{6} \Delta t \lambda_j \left[ \tilde{C}_{j-\frac{1}{2}} (a_{j-\frac{1}{2}}^n)^3 + \tilde{C}_{j+\frac{1}{2}} (a_{j+\frac{1}{2}}^n)^3 \right] \right\}. & \end{aligned} \quad (6.10)$$

Since as  $\Delta t \rightarrow 0$ , widths of all the Riemann fans approach zero, then:

$$\tilde{A}_{j-\frac{1}{2}} \rightarrow \bar{\omega}_{j-1/2}^{n+1} \quad \& \quad \tilde{A}_{j+\frac{1}{2}} \rightarrow \bar{\omega}_{j+1/2}^{n+1}, \quad (6.11)$$

and also regarding Eq. (6.1), we have:

$$\begin{aligned} u(x_{j+1/2,r}^n, t) & \rightarrow P_{j+1}(x_{j+1/2}, t) \\ & = A_{j+1} - B_{j+1} ((1 - p_{j+1})\Delta x_{j+1}) + \frac{1}{2} C_{j+1} ((1 - p_{j+1})\Delta x_{j+1})^2 := u_{j+1/2}^R, \\ u(x_{j+1/2,l}^n, t) & \rightarrow P_j(x_{j+1/2}, t) \\ & = A_j + B_j (p_j \Delta x_j) + \frac{1}{2} C_{j+1} (p_j \Delta x_j)^2 := u_{j+1/2}^L. \end{aligned} \quad (6.12)$$

By substituting Eqs. (6.4)-(6.6),(6.11) and (6.12) in Eq. (6.10), we have:

$$\begin{aligned} \frac{d}{dt}\bar{u}_j(t) = & - \frac{1}{2\Delta x_j} \left[ F(u_{j+1/2}^R(t)) + F(u_{j+1/2}^L(t)) - F(u_{j-1/2}^R(t)) - F(u_{j-1/2}^L(t)) \right] \\ & + \frac{a_{j+1/2}^n}{2\Delta x_j} \left[ u_{j+1/2}^R(t) - u_{j+1/2}^L(t) \right] - \frac{a_{j-1/2}^n}{2\Delta x_j} \left[ u_{j-1/2}^R(t) - u_{j-1/2}^L(t) \right]. \end{aligned} \quad (6.13)$$

The semi-discrete form (6.13) then can be rewritten as:

$$\frac{d}{dt}\bar{u}_j(t) + \frac{F_{j+1/2}^* - F_{j-1/2}^*}{\Delta x_j} = 0, \quad (6.14)$$

where  $F_{j\pm 1/2}^* := F^*(u_{j\pm 1/2})$  and

$$F_{j\pm 1/2}^* := \frac{F(u_{j\pm 1/2}^R(t)) + F(u_{j\pm 1/2}^L(t))}{2} - \frac{a_{j\pm 1/2}^n}{2} [u_{j\pm 1/2}^R(t) - u_{j\pm 1/2}^L(t)]. \quad (6.15)$$

### 6.1.3 The preliminary construction

Three successive cells  $I_{j-1}$ ,  $I_j$  and  $I_{j+1}$  are considered with cell lengths  $\Delta x_{j-1}$ ,  $\Delta x_j$  and  $\Delta x_{j+1}$ . For these cells, cell-averages are denoted respectively by  $\bar{u}_{j-1}$ ,  $\bar{u}_j$  and  $\bar{u}_{j+1}$ .

Let  $q_j(x)$  denotes a quadratic cell-average interpolating polynomial obtained with the centered stencil on the cells  $\{I_{j-1}, I_j, I_{j+1}\}$ , where:

$$\frac{1}{\Delta x_i} \int_{I_i} q_j(x) dx = \bar{u}_i, \quad i = \{j-1, j, j+1\}. \quad (6.16)$$

It is easy to show that if  $a := \Delta x_{j-1}/\Delta x_j$  and  $b := \Delta x_{j+1}/\Delta x_j$ , then  $q_j(x)$  would be:

$$q_j(x) = a_0(x - \bar{x}_j)^2 + b_0(x - \bar{x}_j) + c_0, \quad (6.17)$$

where  $\bar{x}_j$  denotes the middle point of cell  $I_j$ , i.e.:  $\bar{x}_j = (x_{j-1/2} + x_{j+1/2})/2$  and

$$\begin{aligned} a_0 &= -\frac{3[(u')_j^- - (u')_j^+]}{2\Delta x_j(a+b+1)}, & b_0 &= \frac{(2a+1)(u')_j^+ + (2b+1)(u')_j^-}{2(a+b+1)}, \\ c_0 &= \bar{u}_j + \frac{\Delta x_j [(u')_j^- - (u')_j^+]}{8(a+b+1)} = \bar{u}_j - \frac{\Delta x_j^2}{12} a_0, \end{aligned} \quad (6.18)$$

where  $(u')_j^- := \Delta u_j^- / [\frac{1}{2}(1+a)\Delta x_j]$ ;  $(u')_j^+ := \Delta u_j^+ / [\frac{1}{2}(1+b)\Delta x_j]$ ;  $\Delta u_j^- := \bar{u}_j - \bar{u}_{j-1}$ ; and  $\Delta u_j^+ := \bar{u}_{j+1} - \bar{u}_j$ . The piecewise reconstruction can be obtained as:  $q(x) = \sum_j q_j(x)\chi_j$ , where  $\chi_j$  has unit value on the  $I_j$  and is zero elsewhere.

In each cell  $I_j$ , the function  $q(x)$  satisfies the following *updated* features:

1.  $\frac{1}{\Delta x_j} \int_{I_j} q_j(x) dx = \bar{u}_j$  (the conservation property),
2.  $\forall x \in I_j, q_j(x) - u(x, t_{n+1}) = \mathcal{O}(\Delta x^3)$  (the third order accuracy),
3. The function  $q_j(x)$  has the *same shape* of cell-averages,
4. a) A *maximum* cell: If  $I_j$  is a maximum cell, i.e.:  $\bar{u}_{j-1} < \bar{u}_j > \bar{u}_{j+1}$ , then  $q_j(x)$

satisfies:

$$q_j(x_{j+1/2}) \geq L(x_{j+1/2}) \quad \text{and} \quad q_j(x_{j-1/2}) \geq L(x_{j-1/2}) \quad (6.19)$$

where  $L$  denotes a linear interpolation operator between  $\{\{\bar{x}_i, \bar{u}_i\}\}$  values. This means that Eq. (6.19) can be expressed as:

$$q_j(x_{j+1/2}) \geq \frac{\bar{u}_{j+1} + b\bar{u}_j}{1+b} \quad \text{and} \quad q_j(x_{j-1/2}) \geq \frac{\bar{u}_j + a\bar{u}_{j-1}}{1+a}, \quad (6.20)$$

where  $x_{j\pm 1/2}$  denote cell edges.

b) A *minimum* cell: for a minimum cell, i.e.:  $\bar{u}_{j-1} > \bar{u}_j < \bar{u}_{j+1}$ , then  $q_j(x)$  satisfies:

$$q_j(x_{j+1/2}) \leq L(x_{j+1/2}) \quad \text{and} \quad q_j(x_{j-1/2}) \leq L(x_{j-1/2}) \quad (6.21)$$

or equivalently:

$$q_j(x_{j+1/2}) \leq \frac{\bar{u}_{j+1} + b\bar{u}_j}{1+b} \quad \text{and} \quad q_j(x_{j-1/2}) \leq \frac{\bar{u}_j + a\bar{u}_{j-1}}{1+a}. \quad (6.22)$$

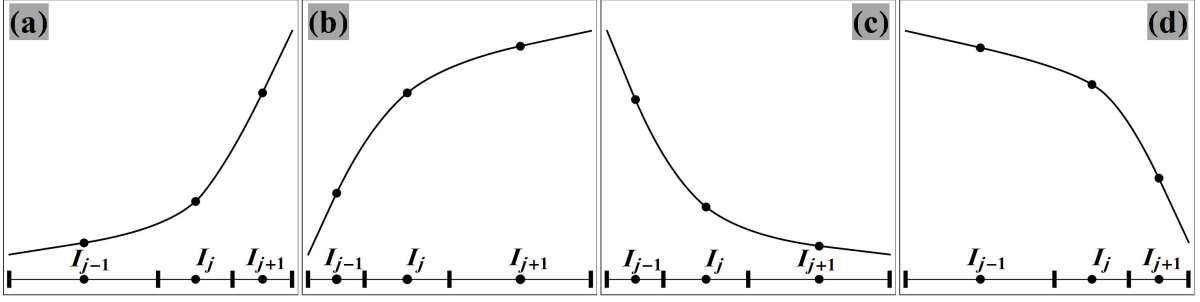
The conditions 1 to 3 are the same as those mentioned in [93] while the fourth condition is updated here.

In the following, properties 1 through 4 are derived which are mentioned in Section 6.1.3. It is clear that features (1) and (2) are satisfied. *Having the same shape means if cell averages are locally monotone,  $q_j(x)$  is also monotone in those cells; e.g.: if  $\bar{u}_{j-1} \leq \bar{u}_j \leq \bar{u}_{j+1}$  then  $q_j(x)$  is increasing and vice-versa. Also, in an extremum cell, the function  $q_j(x)$  has an extremum.* Aforementioned properties will be confirmed in the following. In all calculations, it is assumed that  $0.5 < a < 2$  and  $0.5 < b < 2$ . This is because of the multiresolution-based grid-adaptation and also the post-processing stage of grid adaptations.

### The same shape feature

The first derivative of  $q_j(x)$  ( $q'_j(x)$ ) can be evaluated as:

$$q'_j(x) = \frac{[\Delta x_j (2a + 1) + 6(x - \bar{x}_j)]}{2\Delta x_j (a + b + 1)} \frac{2\Delta u_j^+}{(1+b)\Delta x_j} + \frac{[\Delta x_j (2b + 1) - 6(x - \bar{x}_j)]}{2\Delta x_j (a + b + 1)} \frac{2\Delta u_j^-}{(1+a)\Delta x_j}. \quad (6.23)$$


 Figure 6.3: Different variation patterns over three successive cells  $\{I_j\}$ .

This equation can be written as:

$$q'_j(x) = (g_1(x) + g_2(x)) / (\Delta x_j^2 (a+1)(b+1)(a+b+1)), \quad (6.24)$$

where

$$\begin{aligned} g_1(x) &= (a+1) [\Delta x_j (2a+1) + 6(x - \bar{x}_j)] \Delta u_j^+, \\ g_2(x) &= (b+1) [\Delta x_j (2b+1) - 6(x - \bar{x}_j)] \Delta u_j^-. \end{aligned} \quad (6.25)$$

In the following,  $q_j(x)$  with monotone variations (increasing and decreasing cell averages) and the same shape feature around extremum points will be studied.

**Case I: Increasing cell averages** For the monotone increasing case, i.e.:  $\Delta u_j^\pm \geq 0$ , two variation patterns are considerable: convex (Figure 6.3(a)) and concave increasing cases (Figure 6.3(b)):

1. The convex increasing: In this case (Figure 6.3(a))  $a \geq 1$ ,  $b \leq 1$  and  $\Delta u_j^+ > \Delta u_j^-$ : due to a proper cell adaptation and convex variation of cell averages. It should be shown that for  $x \in I_j$ , we have  $q'_j(x) > 0$ . Since both  $g_1(x)$  and  $g_2(x)$  are linear,  $g_1(x) + g_2(x)$  is also linear. So, if  $q'_j(x)$  at points  $x_{i\pm 1/2}$  are positive, then  $q'_j(x)$  is also positive for  $x \in I_j$ :
  - a) Controlling of  $q'_j(x_{j-1/2})$ : Since  $x_{j-1/2} - \bar{x}_j = -\Delta x_j/2$ ,  $a \geq 1$  and  $b \leq 1$ , then  $g_1(x_{j-1/2}) \geq 0$  and  $g_2(x_{j-1/2}) \geq 0$ ; therefore  $q'_j(x_{j-1/2}) \geq 0$ .
  - b) Controlling of  $q'_j(x_{j+1/2})$ : Since  $x_{j+1/2} - \bar{x}_j = \Delta x_j/2$ ,  $a \geq 1$  and  $b \leq 1$ , then  $g_1(x_{j+1/2}) \geq 0$ ,  $g_2(x_{j+1/2}) \leq 0$  and  $g_1(x_{j+1/2}) \geq |g_2(x_{j+1/2})|$ ; therefore  $q'_j(x_{j+1/2}) \geq 0$ .



So, for the convex increasing of  $\{\bar{u}_i : i = j - 1, j, j + 1\}$ ,  $q_j(x)$  will remain monotone increasing.

2. The concave increasing: For this case (Figure 6.3(b)), we have:  $b \geq 1$ ,  $a \leq 1$  and  $\Delta u_j^+ < \Delta u_j^-$ :

- a) At the point  $x_{j-1/2}$ , it is easy to show that:  $g_1(x_{j-1/2}) \leq 0$ ,  $g_2(x_{j-1/2}) \geq 0$  and  $|g_2(x_{j-1/2})| \geq |g_1(x_{j-1/2})|$ ; hence  $q'_j(x_{j-1/2}) \geq 0$ ,
- b) At the point  $x_{j+1/2}$ :  $g_1(x_{j+1/2}) \geq 0$  and  $g_2(x_{j+1/2}) \geq 0$ ; so,  $q'_j(x_{j+1/2}) \geq 0$ .

Hence,  $q'_j(x) \geq 0$  for  $x \in I_j$ .

**Case II: Decreasing cell averages** In this case, for a monotone decreasing we have  $\Delta u_j^\pm \leq 0$ .

And again convex (Figure 6.3(c)) and concave decreasing (Figure 6.3(d)) cases are possible:

1. The convex decreasing: For this case (Figure 6.3(c)), we have:  $b \geq 1$ ,  $a \leq 1$  and  $|\Delta u_j^-| \geq |\Delta u_j^+|$ :

- a) At the point  $x_{j-1/2}$ , it is easy to show that:  $g_1(x_{j-1/2}) \geq 0$ ,  $g_2(x_{j-1/2}) \leq 0$  and  $|g_2(x_{j-1/2})| \geq |g_1(x_{j-1/2})|$ ; hence  $q'_j(x_{j-1/2}) \leq 0$ .
- b) At the point  $x_{j+1/2}$ :  $g_1(x_{j+1/2}) \leq 0$  and  $g_2(x_{j+1/2}) \leq 0$ ; so  $q'_j(x_{j+1/2}) \leq 0$ .

In this regard,  $q'_j(x) \leq 0$  for  $x \in I_j$ .

2. The concave decreasing: For this case (Figure 6.3(d)), we have:  $a \geq 1$ ,  $b \leq 1$  and  $|\Delta u_j^-| \leq |\Delta u_j^+|$ ; therefore:

- a) At the point  $x_{j-1/2}$ :  $g_1(x_{j-1/2}) \leq 0$  and  $g_2(x_{j-1/2}) \leq 0$ ; so:  $q'_j(x_{j-1/2}) \leq 0$ ,
- b) At the point  $x_{j+1/2}$ :  $g_1(x_{j+1/2}) \leq 0$ ,  $g_2(x_{j+1/2}) \geq 0$  and  $|g_1(x_{j+1/2})| \geq |g_2(x_{j+1/2})|$ ; hence:  $q'_j(x_{j+1/2}) \leq 0$ .

So, in general:  $q'_j(x) \leq 0$  for  $x \in I_j$ .

**Case III: Extrema points** In these points, due to the cell adaptation,  $a \rightarrow 1$  and  $b \rightarrow 1$ ; therefor, it is easy to show that:

$$q'_j(x) = \frac{(u')_j^- [\Delta x_j - 2(x - x_j)] + (u')_j^+ [\Delta x_j + 2(x - x_j)]}{2\Delta x_j}. \quad (6.26)$$

Hence in the cell edges  $x_{j\pm 1/2}$ , the function  $q'_j(x)$  becomes:  $q'_j(x_{j-1/2}) = (u')_j^-$  and  $q'_j(x_{j+1/2}) = (u')_j^+$ . Therefore  $q'_j(x_{j-1/2}) \cdot q'_j(x_{j+1/2}) = (\bar{u}_j - \bar{u}_{j-1})(\bar{u}_{j+1} - \bar{u}_j) / \left[ \frac{1}{4}(1+a)(1+b)\Delta x_j^2 \right] < 0$ . Hence,  $q_j(x)$  has the same shape of cell-averages  $\{\bar{u}_i\}$ .

### The fourth property

1. The maximum cell: At the two edges of  $I_j$ , the fourth property is controlled, where  $a \geq 1$  and  $b \geq 1$ :

- a) Controlling at  $x = x_{j-1/2}$ : The density of an adapted grid increases by approaching to the maximum point  $x_j$  from the left side. As a result:  $a \geq 1$ ,  $b \geq 1$ ,  $\Delta u_j^- \geq 0$  and  $\Delta u_j^+ \leq 0$ ; therefore:

$$\begin{aligned} \delta_{j-1/2}^{(max)} &:= q(x_{j-1/2}) - \frac{\bar{u}_j + a\bar{u}_{j-1}}{1+a} \\ &= \frac{(b+1)(a^2 + (b+1)(a-1))\Delta u_j^- - a(a+1)\Delta u_j^+}{(a+1)(b+1)(a+b+1)}. \end{aligned} \quad (6.27)$$

It is clear that  $\delta_{j-1/2}^{(max)} > 0$ .

- b) Controlling at  $x = x_{j+1/2}$ : For this case, it is obvious that:  $a \geq 1$ ,  $b \geq 1$ ,  $\Delta u_j^- \geq 0$  and  $\Delta u_j^+ \leq 0$ . So,

$$\delta_{j+1/2}^{(max)} := q(x_{j+1/2}) - \frac{\bar{u}_{j+1} + b\bar{u}_j}{1+b} = \frac{b(-(a+1)\Delta u_j^+ + (b+1)\Delta u_j^-)}{(a+1)(b+1)(a+b+1)}. \quad (6.28)$$

It is clear that  $\delta_{j+1/2}^{(max)} > 0$ .

2. The minimum cell: Controlling at  $x = x_{j-1/2}$  and  $x = x_{j+1/2}$ . It is straightforward to show that:

$$\begin{aligned} \delta_{j-1/2}^{(min)} &:= q(x_{j-1/2}) - \frac{\bar{u}_j + a\bar{u}_{j-1}}{1+a} = -\delta_{j-1/2}^{(max)}, \\ \delta_{j+1/2}^{(min)} &:= q(x_{j+1/2}) - \frac{\bar{u}_{j+1} + b\bar{u}_j}{1+b} = -\delta_{j+1/2}^{(max)}. \end{aligned} \quad (6.29)$$

It is clear that  $\delta_{j-1/2}^{(min)} < 0$  and  $\delta_{j+1/2}^{(min)} < 0$ .

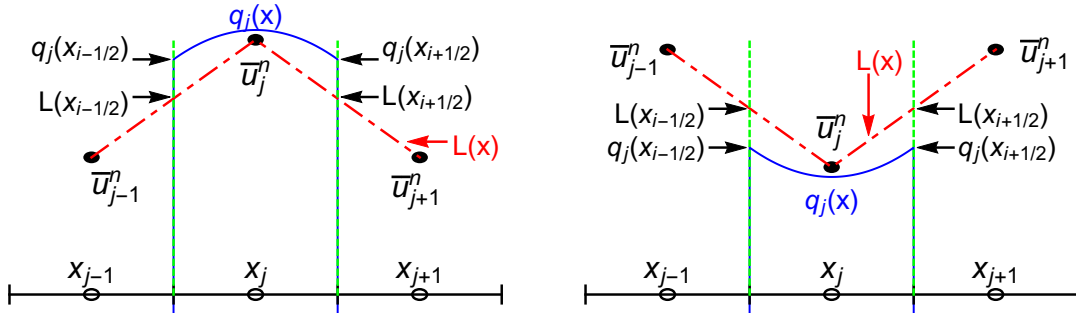


Figure 6.4: The fourth reconstruction feature at max and min cells.

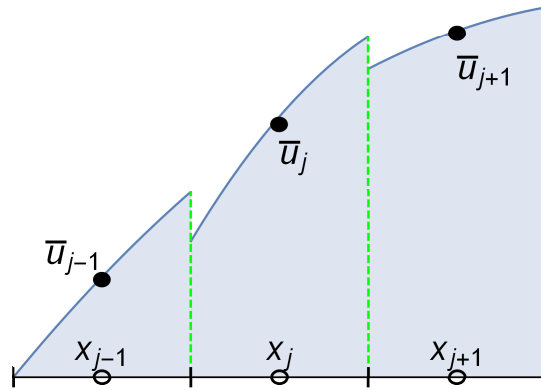


Figure 6.5: Possible forming of non-physical extrema at cell-edges.

These two conditions at max and min cells are presented in Figure 6.4.

### 6.1.4 Designing of a limiter for the third-order central high resolution scheme

In the previous section, the piecewise-parabolic reconstruction,  $q(x) = \sum_j q_j(x)\chi_j$  is presented. It is shown that respect to average solutions  $\{\bar{u}_j\}$ ,  $q(x)$  has the *same shape* feature. *This feature guarantees that no new extrema develop through inside of each cell  $I_j$ . So, it is possible new extrema appear at cell interfaces  $\{x_{j+1/2}\}$* ; for instance see Figure 6.5. To prevent this, the piecewise-parabolic reconstruction is modified by a *limiting* procedure. For adapted grids, two limiting approaches will be studied here.

**The first approach** For cell  $I_j$ , this modification can be done by a convex combination of  $q_j(x)$  and cell average  $\bar{u}_j$ :

$$P_j(x) = \bar{u}_j + \theta_j^{(1)} (q_j(x) - \bar{u}_j), \quad 0 \leq \theta_j^{(1)} \leq 1, \quad (6.30)$$

where:

$$\theta_j^{(1)} := \begin{cases} \min \left\{ \frac{M_{j+1/2} - \bar{u}_j}{M_j - \bar{u}_j}, \frac{m_{j-1/2} - \bar{u}_j}{m_j - \bar{u}_j}, 1 \right\}, & \text{if } \bar{u}_{j-1} < \bar{u}_j < \bar{u}_{j+1}, \\ \min \left\{ \frac{M_{j-1/2} - \bar{u}_j}{M_j - \bar{u}_j}, \frac{m_{j+1/2} - \bar{u}_j}{m_j - \bar{u}_j}, 1 \right\}, & \text{if } \bar{u}_{j-1} > \bar{u}_j > \bar{u}_{j+1}, \\ 1, & \text{otherwise,} \end{cases} \quad (6.31)$$

and:

$$\begin{aligned} M_{j+1/2} &= \max \left\{ \frac{\bar{u}_{j+1} + b\bar{u}_j}{1+b}, q_{j+1}(x_{j+1/2}) \right\}, & M_{j-1/2} &= \max \left\{ \frac{\bar{u}_j + a\bar{u}_{j-1}}{1+a}, q_{j-1}(x_{j-1/2}) \right\}, \\ m_{j+1/2} &= \min \left\{ \frac{\bar{u}_{j+1} + b\bar{u}_j}{1+b}, q_{j+1}(x_{j+1/2}) \right\}, & m_{j-1/2} &= \min \left\{ \frac{\bar{u}_j + a\bar{u}_{j-1}}{1+a}, q_{j-1}(x_{j-1/2}) \right\}, \end{aligned} \quad (6.32)$$

and:

$$M_j = \max_{x \in I_j} q_j(x), \quad m_j = \min_{x \in I_j} q_j(x). \quad (6.33)$$

Since  $q_j(x)$  is monotone, parameters  $M_j$  and  $m_j$  can simply be evaluated as:

$$M_j = \max \left\{ q_j(x_{j-1/2}), q_j(x_{j+1/2}) \right\}, \quad m_j = \min \left\{ q_j(x_{j-1/2}), q_j(x_{j+1/2}) \right\}. \quad (6.34)$$

This limiting concept is presented in Figure 6.6.

**The second approach** This approach leads to less dissipative results. For cell  $I_j$ , this limited reconstruction can be done by a convex combination of  $q_j(x)$  and a linear variation inside the cell:

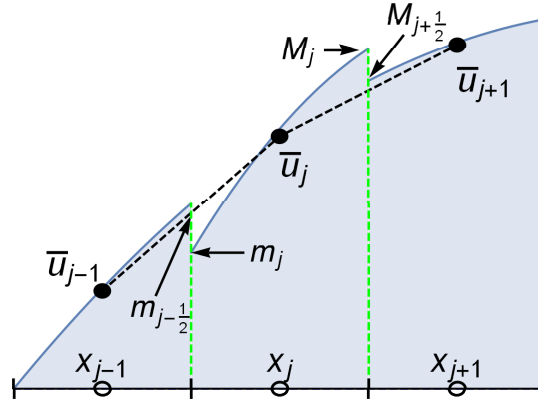


Figure 6.6: Parameters used for scaling limiting at cell-edges for monotone increasing data.

$$P_j(x) = \mathcal{L}_j(x) + \theta_j^{(2)} (q_j(x) - \mathcal{L}_j(x)), \quad 0 \leq \theta_j^{(2)} \leq 1, \quad (6.35)$$

where  $\mathcal{L}_j(x)$  denotes the linear variation; it can be defined as:

$$\mathcal{L}_j(x) = \bar{u}_j + u'_j \times (x - \bar{x}_j), \quad (6.36)$$

where:  $\bar{u}_j := \int_{x_{j-1/2}}^{x_{j+1/2}} u(x) dx / \Delta x_j$ ,  $\bar{x}_j := (x_{j-1/2} + x_{j+1/2})/2$ , and  $u'_j$  denotes a proper estimation of the first spatial derivative.  $u'_j$  is estimated here by the MINMOD limiter (by the TVD feature):

$$u'_j = \text{MINMOD} \left( \frac{\bar{u}_j - \bar{u}_{j-1}}{\bar{x}_j - \bar{x}_{j-1}}, \frac{\bar{u}_{j+1} - \bar{u}_j}{\bar{x}_{j+1} - \bar{x}_j} \right), \quad (6.37)$$

where:

$$\text{MINMOD}(a, b) = \frac{\text{Sign}(a) + \text{Sign}(b)}{2} \min(|a|, |b|). \quad (6.38)$$

And the limiter  $\theta_j^{(2)}$  can be defined as:

$$\theta_j^{(2)} := \begin{cases} \min \left\{ \frac{M_{j+1/2} - \mathcal{L}_j(x_{j+1/2})}{M_j - \mathcal{L}_j(x_{j+1/2})}, \frac{m_{j-1/2} - \mathcal{L}_j(x_{j-1/2})}{m_j - \mathcal{L}_j(x_{j-1/2})}, 1 \right\}, & \text{if } \bar{u}_{j-1} < \bar{u}_j < \bar{u}_{j+1}, \\ \min \left\{ \frac{M_{j-1/2} - \mathcal{L}_j(x_{j-1/2})}{M_j - \mathcal{L}_j(x_{j-1/2})}, \frac{m_{j+1/2} - \mathcal{L}_j(x_{j+1/2})}{m_j - \mathcal{L}_j(x_{j+1/2})}, 1 \right\}, & \text{if } \bar{u}_{j-1} > \bar{u}_j > \bar{u}_{j+1}, \\ 1, & \text{otherwise,} \end{cases} \quad (6.39)$$

where:

$$\begin{aligned} M_{j\pm 1/2} &= \max \left\{ \frac{1}{2} \left( \mathcal{L}_j(x_{j\pm 1/2}) + \mathcal{L}_{j\pm 1}(x_{j\pm 1/2}) \right), q_{j\pm 1}(x_{j\pm 1/2}) \right\}, \\ m_{j\pm 1/2} &= \min \left\{ \frac{1}{2} \left( \mathcal{L}_j(x_{j\pm 1/2}) + \mathcal{L}_{j\pm 1}(x_{j\pm 1/2}) \right), q_{j\pm 1}(x_{j\pm 1/2}) \right\}. \end{aligned} \quad (6.40)$$

## 6.2 CWENO schemes on non-uniform grids

In this section, grid-based adaptive central-WENO (CWENO) schemes will be provided. Here, reconstructions are assumed to be CWENO-based piecewise-parabolic functions. Let assume  $u^n(x)$  is the exact solution at  $t^n$ , and  $\bar{u}_j$  denotes corresponding average solution on the cell  $I_j$ . Using the set  $\{\bar{u}_j\}$ , the reconstructed piecewise function is  $P_j(x, t^n) = \sum_j R_j(x) \chi_j$ ; where:  $\chi_j$  is equal to unit on the interval  $[x_{j-1/2}, x_{j+1/2}]$  and zero elsewhere;  $R_j(x)$  denotes a polynomial of order 2.

Functions  $R_j(x)$  and  $P_j(x)$  should satisfy the following features:

1. Conservation:  $\frac{1}{\Delta x_j} \int_{I_j} R_j(x) dx = \bar{u}_j$ ,
2. Accuracy:
  - a) Point-wise accuracy:

$$\forall x \in I_j, R_j(x) - u(x, t_{n+1}) = \mathcal{O}(\Delta x^r) : \quad r - \text{order accuracy}, \quad (6.41)$$

- b) Average-based accuracy: since in the central formulation, in evolution stage, solutions are obtained on staggered grids, so it is necessary that reconstructed solutions have  $r$  order accuracy for half-cell averages, as well; for the cell  $I_j$ ,

this means:

$$\begin{aligned} \frac{1}{x_{j+1/2} - \bar{x}_j} \int_{\bar{x}_j}^{x_{j+1/2}} P_j(x, t^n) dx &= \frac{1}{x_{j+1/2} - \bar{x}_j} \int_{\bar{x}_j}^{x_{j+1/2}} u(x, t^n) dx + \mathcal{O}(\Delta x^r), \\ \frac{1}{\bar{x}_j - x_{j-1/2}} \int_{x_{j-1/2}}^{\bar{x}_j} P_j(x, t^n) dx &= \frac{1}{\bar{x}_j - x_{j-1/2}} \int_{x_{j-1/2}}^{\bar{x}_j} u(x, t^n) dx + \mathcal{O}(\Delta x^r), \end{aligned} \quad (6.42)$$

where  $\bar{x}_j := (x_{j-1/2} + x_{j+1/2})/2$ .

3. Non-oscillatory reconstruction: following Section 6.1.3 and WENO reconstruction this will be achieved.

Based on the above mentioned properties, two CWENO schemes having three and five points are provided. The three and five point stencils lead to third and (at least) fourth order accuracy in the sense of average solutions, respectively.

### 6.2.1 The three point stencil on non-uniform grids

In Section 6.1.3, it is shown how to construct a parabolic polynomial  $q_j(x)$  in a way that it remains non-oscillatory on properly adapted grids. In Section 6.1.4, it is shown how to control/prevent oscillations by using a limiter definition. Another approach is to use the CWENO concept. In smooth regions,  $P_j(x)$  is a quadratic function ( $P_j(x) = q_j(x)$ ), while around a discontinuity, a linear convex combination of the left-side, central, and the right-side approximations can be used to prevent spurious oscillations. By using the convex combination of different polynomials with different order and stencils, we have:

$$P_j(x) = \sum_i W_i^j P_i^j(x), \quad \sum_i W_i^j = 1, \quad i \in \{L, C, R\}, \quad W_i^j \geq 0, \quad (6.43)$$

where:  $\{W_i^j\}$  are constants;  $P_L^j(x)$  and  $P_R^j(x)$  are left and right linear functions, respectively.  $P_L^j(x)$  and  $P_R^j(x)$  interpolate average values on cell sets  $\{I_{j-1}, I_j\}$  and  $\{I_j, I_{j+1}\}$ , respectively (to have the conservation property). So we have:  $P_L^j(x) = \bar{u}_j^n + \frac{\bar{u}_j^n - \bar{u}_{j-1}^n}{\bar{x}_j - \bar{x}_{j-1}}(x - \bar{x}_j)$  and  $P_R^j(x) = \bar{u}_j^n + \frac{\bar{u}_{j+1}^n - \bar{u}_j^n}{\bar{x}_{j+1} - \bar{x}_j}(x - \bar{x}_j)$ , where  $\bar{x}_j := \frac{1}{2}(x_{j-1/2} + x_{j+1/2})$ . This three-point stencil of CWENO scheme with three sub-stencils is presented in Figure 6.7.

As mentioned, in smooth regions, we desire to have  $P_j(x) = q_j(x)$ ; that is:

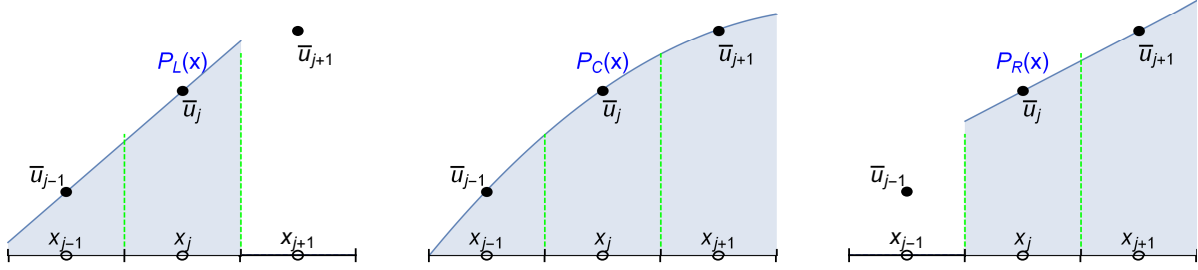


Figure 6.7: CWENO scheme defined over three cells with three sub-stencils.

$$q_j(x) = C_L P_L(x) + C_C P_C(x) + C_R P_R(x), \quad \sum_i C_i = 1, \quad i \in \{L, C, R\}, \quad x \in I_j, \quad (6.44)$$

where  $\{C_i : i \in \{L, C, R\}\}$  are constants. From this equation, central function  $P_C(x)$  can be obtained.

Following the WENO concept, the weight coefficients  $W_i^j$  in Eq. (6.43) can be obtained as:

$$W_i^j = \frac{\alpha_i^j}{\sum_k \alpha_k^j}, \quad \alpha_i^j = \frac{C_i}{(\epsilon + IS_i^j)^\beta}, \quad i, k \in \{L, C, R\}, \quad (6.45)$$

where:  $IS_i^j$  denotes the *smoothness indicator*, designed for detecting discontinuities or high gradients;  $\beta$  is a constant and is assumed to be equal to 2, in this study;  $\epsilon$  is a small constant. It is recommended to be around  $\epsilon = 10^{-6}$ .

Approximating the polynomial  $y_{I_j}(x) = \alpha_0 + \alpha_1(x - x_j) + \alpha_2(x - x_j)^2$  by the reconstructed function  $P_j(x_{j\pm 1/2})$ , for the case  $\epsilon = 0$ , the truncated error in the sense of



average approximations on non-uniform grids would be:

$$\begin{aligned}
 & \frac{1}{\bar{x}_j - x_{j-1/2}} \int_{x_{j-1/2}}^{\bar{x}_j} [y_{I_j}(x) - P_j(x)] dx = \frac{1}{96(a+b+1)(IS_C^2 C_R IS_L^2 + IS_R^2 (IS_C^2 C_L + C_C IS_L^2))} \\
 & - \left\{ 2(3b+5)(a+b+1)C_R IS_L^2 (IS_C - IS_R) (IS_C + IS_R) + \right. \\
 & \quad \left. IS_R^2 \left( (7a+b+4)IS_L^2 - 2(3a+1)(a+b+1)C_L (IS_C - IS_L) (IS_C + IS_L) \right) \right\} \alpha_2 \Delta x_j^3 + \mathcal{O}(\Delta x_j^4), \\
 & \frac{1}{x_{j+1/2} - \bar{x}_j} \int_{\bar{x}_j}^{x_{j+1/2}} [y_{I_j}(x) - P_j(x)] dx = \frac{1}{96(a+b+1)(IS_C^2 C_R IS_L^2 + IS_R^2 (IS_C^2 C_L + C_C IS_L^2))} \\
 & - \left\{ IS_R^2 \left( 2(3a+5)(a+b+1)C_L (IS_C - IS_L) (IS_C + IS_L) + (a+7b+4)IS_L^2 \right) + \right. \\
 & \quad \left. 2(3b+1)(a+b+1)C_R IS_L^2 (IS_C - IS_R) (IS_C + IS_R) \right\} \alpha_2 \Delta x_j^3 + \mathcal{O}(\Delta x_j^4),
 \end{aligned} \tag{6.46}$$

where:  $a := \Delta x_j / \Delta x_{j-1}$  and  $b := \Delta x_j / \Delta x_{j+1}$ . This means that order of approximation in smooth zones is  $\mathcal{O}(\Delta x_j^3)$ .

Based on works [105, 106], for each cell  $I_j$ , the smoothness indicators are defined as:

$$IS_i^j = \sum_{l=1}^2 \int_{x_{j-1/2}}^{x_{j+1/2}} [\Delta x_j]^{2l-1} [P_i^{(l)}]^2 dx, \quad i \in \{L, C, R\}, \tag{6.47}$$

where  $P_i^{(l)} := d^l P_i / dx^l$ . It is straightforward to show that:

$$\begin{aligned}
 IS_L^j &= \frac{4(\Delta u_j^-)^2}{(a+1)^2} = (\Delta x_j (u')_j^-)^2, \\
 IS_R^j &= \frac{4(\Delta u_j^+)^2}{(b+1)^2} = (\Delta x_j (u')_j^+)^2.
 \end{aligned} \tag{6.48}$$

Parameter  $IS_C$  can directly be determined by the definition (6.47).

Same as the uniform grid case, it is easy to show that for non-uniform grids, the symmetric values of  $C_i$  values ( $C_L = C_R$ ) leads to a third order of accuracy in the sense of average values (from Eqs. (6.42)).

## 6.2.2 The five point stencil on non-uniform grids

In this approach, function  $P_j(x)$  ( $x \in I_j$ ) is constructed by a convex combination of  $q_k(x)$ , as:

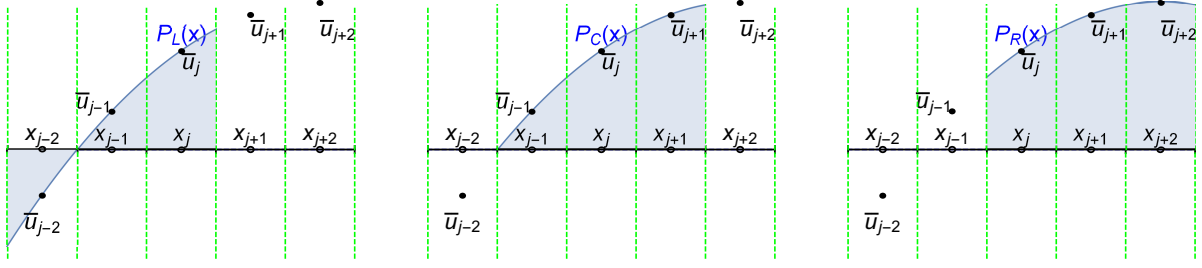


Figure 6.8: CWENO scheme defined over five cells with three sub-stencils.

$$P_j(x) = W_{j-1}^j q_{j-1}(x) + W_j^j q_j(x) + W_{j+1}^j q_{j+1}(x), \quad \sum_{i=j-1}^{j+1} W_i^j = 1, \quad W_i^j \geq 0, \quad x \in I_j. \quad (6.49)$$

Coefficients  $W_i^j$  are defined as Eq. (6.45), i.e.:  $W_i^j = \alpha_i^j / (\alpha_{j-1}^j + \alpha_j^j + \alpha_{j+1}^j)$ . Definitions of  $\alpha_i^j$  and  $IS_i^j$  are the same as ones in Eqs. (6.45) and (6.47).

This five-point stencil of CWENO scheme with three sub-stencils (using parabolic polynomials for each sub-stencil) is presented in Figure 6.8.

Approximating the polynomial  $y_{I_j}(x) = \alpha_0 + \alpha_1(x - x_j) + \alpha_2(x - x_j)^2 + \alpha_3(x - x_j)^3$  by the reconstructed function  $P_j(x_{j\pm 1/2})$ , the truncated error in the sense of average approximations would be:

$$\begin{aligned} \frac{1}{\bar{x}_j - x_{j-1/2}} \int_{x_{j-1/2}}^{\bar{x}_j} [y_{I_j}(x) - R_j(x)] dx &= \frac{1}{64 (IS_C^2 C_R IS_L^2 + IS_R^2 (C_C IS_L^2 + IS_C^2 C_L))} \\ &\quad \left\{ (2a_1 + 1) IS_R^2 \left( (2a_1 + 2a_2 + 1) IS_C^2 C_L - (2b_1 + 1) C_C IS_L^2 \right) + (2b_1 + 1) (2b_1 + 2b_2 + 1) IS_C^2 C_R IS_L^2 \right\} \\ &\quad \alpha_3 \Delta x_j^4 + \mathcal{O}(\Delta x_j^5), \\ \frac{1}{x_{j+1/2} - \bar{x}_j} \int_{\bar{x}_j}^{x_{j+1/2}} [y_{I_j}(x) - R_j(x)] dx &= -\frac{1}{\bar{x}_j - x_{j-1/2}} \int_{x_{j-1/2}}^{\bar{x}_j} [y_{I_j}(x) - R_j(x)] dx, \end{aligned} \quad (6.50)$$

where:  $a_i := \Delta x_{j-i} / \Delta x_j$  and  $b_i := \Delta x_{j+i} / \Delta x_j$  for  $i = \{1, 2\}$ .

Hence, if  $R_j(x)(P_j(x))$  interpolates half-cell averages solutions, Eq. (6.42), on non-uniform grids, the order of accuracy is:  $r = 4$ .

### 6.3 The piecewise parabolic method (PPM)

The concept of the piecewise parabolic method (PPM) was firstly introduced in [99]; it has third order accuracy in the spatial domain. In this method, for the reconstruction stage on the cell  $I_j$ , the cell average value  $\bar{u}_j$  and cell-edge values  $u_{j\pm 1/2}$  are used. For estimation of  $u_{j+1/2}$  values, four neighbor cell average values are employed. Using of this neighbor information, however, can yield more numerical dissipation. Popov et al. [100] improves the method by a local stencil: edge values  $u_{j\pm 1/2}$  are estimated by the concept of the characteristic lines from solutions of the previous time step in the cell  $I_j$ . This method improves numerical accuracy in expense of using some mathematical computations and information. One drawback of the original PPM [99] is that the limiter defined there, can lead to over-smoothing phenomenon around *smooth extrema*; to remedy this drawback, some improvements were proposed for the limiter [101, 102].

The aim of this section is to obtain a semi-discrete third order central high resolution scheme by the piecewise parabolic reconstruction method via a *local stencil*. The previous third order method (Section 6.1) uses three neighbor cells for reconstructions and this yields: i) a type of averaging which leads to smoother numerical solutions for smooth responses, ii) possibility of existence of local spurious oscillations around discontinuities. To prevent these cases, local information will be integrated by two different approaches in the reconstruction stage, as: in each cell  $I_j$

1. using of the cell average value ( $\bar{u}_j$ ) and both the first derivative ( $u'_j$ ) and second ( $u''_j$ ) derivative values: denoted here by the PPM-1
2. using of the cell average value ( $\bar{u}_j$ ) and the first derivatives values at cell-edges ( $u'_{j\pm 1/2}$ ): denoted by PPM-2,

where  $u'_j := du_j/dx$  and  $u''_j := d^2u_j/dx^2$ .

#### 6.3.1 Formulation for the first approach of PPM (PPM-1)

Let us consider the scalar conservation law:

$$u_t + f(u)_x = 0, \quad x \in (-\infty, \infty), \quad t \in (0, \infty), \quad (6.51)$$

where  $u(x, t = 0) = u_0(x)$ . The first spatial derivative of Eq. (6.51) can be represented as:

$$\begin{aligned} v_t + h(u, v)_x &= 0, \quad x \in (-\infty, \infty), \quad t \in (0, \infty), \\ v(x, 0) &= v_0(x), \end{aligned} \quad (6.52)$$

where  $v(x, t) := \partial u / \partial x$  and  $h(u, v) := f(u)_x$ . The flux  $h(u, v)$  can also be written as:  $h(u, v) = f'(u)u_x = f'(u)v$  (where  $f'(u) := \partial f / \partial u$ ). Let us again evaluate the first spatial derivative of Eq. (6.52), as:

$$\begin{aligned} a_t + g(u, v, a)_x &= 0, \quad x \in (-\infty, \infty), \quad t \in (0, \infty), \\ a(x, 0) &= a_0(x), \end{aligned} \quad (6.53)$$

where:  $a(x) := \partial v / \partial x$  and  $g(u, v, a) = h(u, v)_x = h(u, v)_u u_x + h(u, v)_v v_x = h(u, v)_u v + h(u, v)_v a$ .

Let assume non-uniform cells  $\{I_j\}$ , where  $I_j = [x_{j-1/2}, x_{j+1/2}]$ . Cell centers  $x_j$  locate in a way that  $x_{j+1/2} - x_j = \Delta x_j^r$  and  $x_{j-1/2} - x_j = \Delta x_j^l$ . For the cell  $I_j$ , the cell-average is  $\bar{u}_j = (1/\Delta x_j) \int_{x_{j-1/2}}^{x_{j+1/2}} u(x) dx$ ; the first and second derivatives at  $x_j$  are denoted by:  $(u')_j := \partial u(x_j, t) / \partial x$  and  $(u'')_j := \partial^2 u(x_j, t) / \partial x^2$ . Regarding  $\{\bar{u}_j, v_j, a_j\}$  values, a parabolic function  $q_j(x)$  can be interpolated for the region  $[x_{j-1/2}, x_{j+1/2}]$ , as:

$$q_j(x) = \bar{a}_j + \bar{b}_j (x - x_j) + \frac{1}{2} \bar{c}_j (x - x_j)^2, \quad (6.54)$$

where:

$$\begin{aligned} \bar{a}_j &= \bar{u}_j - \frac{[\Delta x_j^l]^2 [- (\Delta x_j^l u''_j + 3u'_j)] + [\Delta x_j^r]^3 u''_j + 3 [\Delta x_j^r]^2 u'_j}{6\Delta x_j}, \\ \bar{b}_j &= u'_j, \\ \bar{c}_j &= u''_j, \end{aligned} \quad (6.55)$$

where  $u'_j := v_j$  and  $u''_j := a_j$ .

Having the parabolic reconstructions  $\{q_j\}$ , a piecewise average interpolating function  $P(x)$  can be obtained as:

$$P(x) = \sum_j [q_j(x) \chi_{I_j}], \quad I_j \in [x_{j-1/2}, x_{j+1/2}], \quad (6.56)$$

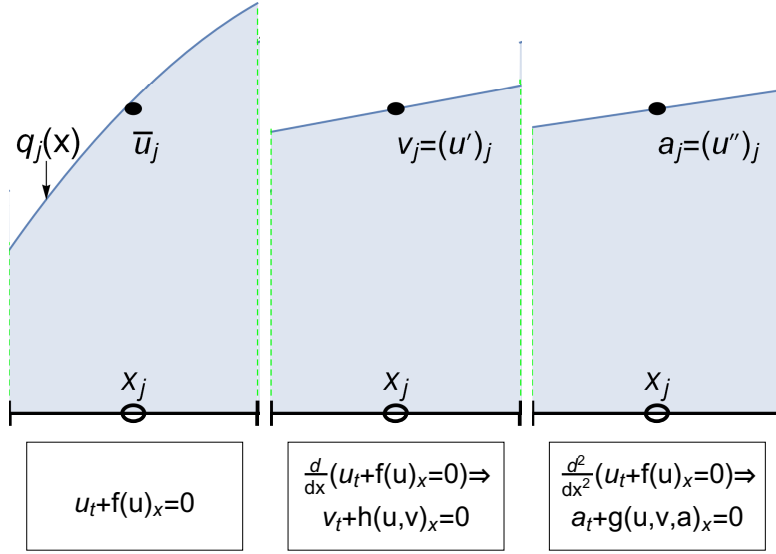


Figure 6.9: Using local stencil for the first PPM method.

where  $\chi_{I_j}$  denotes the unit box function: it has unit value inside  $I_j$  and is zero elsewhere. This type of interpolation is presented in Figure 6.9.

The next step is to guarantee that  $P(x)$  will remain monotone by some modifications. This can be done directly by two approaches:

**The first approach** By using the limiter defined in the Section 6.1, Eq. (6.30) with the limiter in Eq. (6.31) and the parameters defined in Eqs. (6.32) and (6.33).

**The second approach** At first, monotone the piecewise parabolic reconstruction in each cell, and then controlling/modification for satisfying the maximum principle at cell-edges (by the limiter defined in Eq. (6.31)). Following this approach, we have:

1. Inside each cell  $I_j$ : to guarantee the monotonicity feature inside each cell, the concept of *monotonization* will be used by updating  $q_j$  polynomials by utilizing new limiters. Since solutions  $u$ ,  $v$  and  $a$  are independent parameters, this monotone can be done simply, based on these values.
2. At cell edges  $\{x_{j+1/2}\}$ : this will be done by the concept of geometric rescaling, performed in Eq. (6.30) by the limiter and corresponding parameters from Eqs. (6.31)-(6.32) and (6.34) (derived for monotone solutions).

As mentioned, the monotonization procedure is done by defining a limiter to guarantee that a non-decreasing (non-increasing) solution remains non-decreasing (non-increasing). The updated  $q_j(x)$  after limiting is:

$$q_j(x) = \hat{a}_j + \hat{b}_j (x - x_j) + \frac{1}{2} \hat{c}_j (x - x_j)^2, \quad (6.57)$$

where  $\hat{a}$ ,  $\hat{b}$  and  $\hat{c}$  have the same definitions of Eq. (6.55), except that  $u_j''$  is replaced with  $(\phi_j u_j'')$ . Parameter  $\phi_j$  denotes a limiter; it tides solutions  $u_j$ ,  $v_j$  and  $a_j$  (from Eqs. (6.51), (6.52) and (6.53)) together. The limiter definition is:

$$\phi_j = \begin{cases} \min \left( 1, \frac{|u_j'|}{|u_j''| \cdot |\Delta x_j^l|} \right), & \text{if } \text{Sign}(u_j') = \text{Sign}(u_j''), \quad u_j'' \neq 0 \quad \& \quad u_j' \neq 0, \\ \min \left( 1, \frac{|u_j'|}{|u_j''| \cdot |\Delta x_j^r|} \right), & \text{if } \text{Sign}(u_j') \neq \text{Sign}(u_j''), \quad u_j'' \neq 0 \quad \& \quad u_j' \neq 0, \\ 1, & \text{if } u_j'' = 0 \quad \& \quad u_j' = 0, \\ \min \left( 1, \frac{|\epsilon|}{|u_j''| \cdot |\Delta x_j^l|} \right), & \text{if } u_j'' \neq 0 \quad \& \quad u_j' = 0, \end{cases} \quad (6.58)$$

where:  $u_j' := v_j$ ,  $u_j'' := a_j$  and  $\epsilon \ll 1$  denotes a small constant. The case  $\{u_j'' \neq 0, u_j' = 0\}$  shows a local extremum; to cure over-smoothing only at extremum points, the last definition in the limiter (6.58) is done. With this modification, the extrema inside cells are pushed to the edges of cells. In these extremum points, also, the first derivatives are updated as:  $u_j' = \text{Sign}(u_j'') |\epsilon|$ .

After this slope limiting of  $v_j(x) = u_j'(x)$  by limiting the  $a_j(x) = u_j''(x)$ , the modified  $v_j(x)$  does not change its sign over cell  $I_j$ , and so  $u_j(x)$  becomes monotone. This type of modification by limiting approach is presented in Figure 6.10, where the dashed (red) and solid (blue) lines represent the original  $v_j(x)$  and the limited  $v_j(x)$ , respectively.

*It is easy to show that after limiting of  $u_j''$  and modifying of  $u_j'$ , the updated  $q_j(x)$  in Eq. (6.57) remains conservative; this means:  $\int_{x_{i-1/2}}^{x_{i+1/2}} q_j(x) dx / (\Delta x_j) = \bar{u}_j$ .*

The system of Eqs. (6.51)-(6.53) can be rewritten as a system of semi-discrete form:

$$\text{sign}(u'_j) = \text{sign}(u''_j)$$

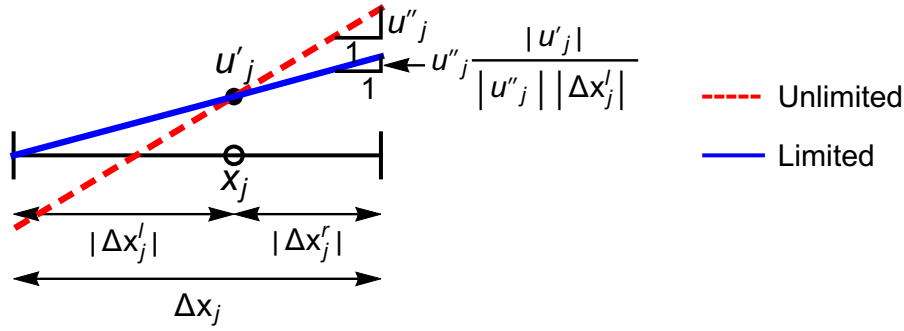


Figure 6.10: Modification of  $u'_j(x) = v_j(x)$  by limiting  $u''(j)$ , to prevent sign changing of  $v_j(x)$ . The original data and the limited one are denoted by the red dashed the solid blue lines, respectively.

$$\begin{aligned} \frac{du_j}{dt} &= -\frac{1}{\Delta x_j} (F_{j+1/2}^* - F_{j-1/2}^*), \\ \frac{dv_j}{dt} &= -\frac{1}{\Delta x_j} (H_{j+1/2}^* - H_{j-1/2}^*), \\ \frac{da_j}{dt} &= -\frac{1}{\Delta x_j} (G_{j+1/2}^* - G_{j-1/2}^*), \end{aligned} \quad (6.59)$$

where  $F^*$ ,  $H^*$  and  $G^*$  are numerical fluxes. The system of Eqs. (6.59) has  $r$ th order accuracy if:

$$\frac{1}{\Delta x_j} (F_{j+1/2}^* - F_{j-1/2}^*) = f(u)_x|_{x=x_j} + \mathcal{O}(\Delta x^r), \quad (6.60a)$$

$$\frac{1}{\Delta x_j} (H_{j+1/2}^* - H_{j-1/2}^*) = h(u, v)_x|_{x=x_j} + \mathcal{O}(\Delta x^{r-1}), \quad (6.60b)$$

$$\frac{1}{\Delta x_j} (G_{j+1/2}^* - G_{j-1/2}^*) = g(u, v, a)_x|_{x=x_j} + \mathcal{O}(\Delta x^{r-2}). \quad (6.60c)$$

Following the concept of central high resolution schemes, for Eq. (6.60a), the third order formulation presented in Eq. (6.15) can be used. For (6.60b), the second-order central/central-upwind schemes can also be used. The performance of second-order central schemes on adapted grids is studied in Sections 4 and 5. In this work, Eq. (6.60c) is discretized by second-order central/central-upwind schemes, as well (a first

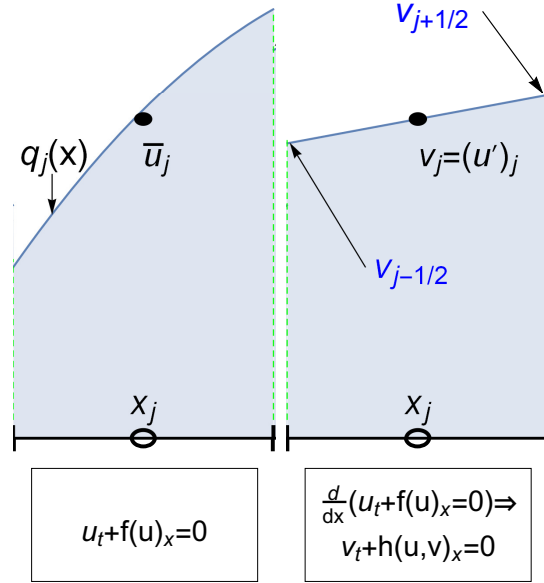


Figure 6.11: Using local stencil for the second PPM method.

order central scheme can also be used, e.g., the Lax-Friedrichs central scheme).

### 6.3.2 Formulation for the second approach of PPM (PPM-2)

The conservation law (6.51) is considered again. Corresponding first spatial derivative can be expressed as (6.52). Based on these equations, a local parabolic reconstruction will be developed by  $\bar{u}_j$  and  $v_{j\pm 1/2}$ . It is necessary that solvers of Eq. (6.51) and Eq. (6.52) have third and second accuracy orders, respectively. Let the local parabolic polynomial  $q_j(x)$  for  $x \in I_j$  is represented by Eq. (6.54). The coefficients  $\bar{a}_j$ ,  $\bar{b}_j$  and  $\bar{c}_j$  can be evaluated as:

$$\begin{aligned}\bar{a}_j &= \bar{u}_j - \frac{2(u'_{j-1/2} - u'_{j+1/2}) \Delta x_j^l \Delta x_j^r + (\Delta x_j^l)^2 (-u'_{j-1/2} - 2u'_{j+1/2}) + (\Delta x_j^r)^2 (2u'_{j-1/2} + u'_{j+1/2})}{6\Delta x_j}, \\ \bar{b}_j &= -\frac{\Delta x_j^l u'_{j+1/2} - \Delta x_j^r u'_{j-1/2}}{\Delta x_j}, \\ \bar{c}_j &= \frac{u'_{j+1/2} - u'_{j-1/2}}{\Delta x_j},\end{aligned}\tag{6.61}$$

where  $u'_j := v_j$ . This type of interpolation is presented in Figure 6.11.

It is necessary to monotoneize the local polynomial  $q_j(x)$  with coefficients (6.61). For



$$\text{sign}(u'_j) = \text{sign}(u''_j)$$

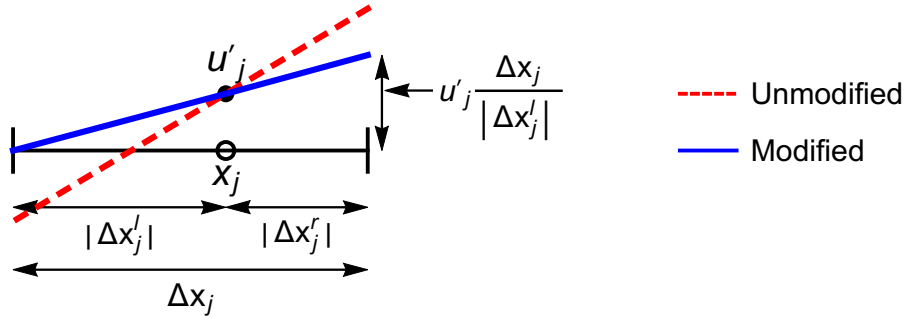


Figure 6.12: Modification of  $v_j(x)$ , the red dashed line, to prevent sign changing of  $v_j(x)$  as the solid blue line.

this,  $u'_{j\pm 1/2}$  values are updated as  $\hat{u}'_{j\pm 1/2}$ :

$$\begin{aligned} \hat{u}'_{j-1/2} &= u'_{j-1/2}, \quad \hat{u}'_{j+1/2} = u'_{j+1/2}, & \text{if } \text{Sign}(u'_{j-1/2}) = \text{Sign}(u'_j) = \text{Sign}(u'_{j+1/2}), \\ \hat{u}'_{j-1/2} &= 0, \quad \hat{u}'_{j+1/2} = u'_j \frac{\Delta x_j}{|\Delta x'_j|}, & \text{if } \text{Sign}(u'_{j-1/2}) \neq \text{Sign}(u'_j) \ \& \ \text{Sign}(u'_{j+1/2}) = \text{Sign}(u'_j), \\ \hat{u}'_{j-1/2} &= u'_j \frac{\Delta x_j}{|\Delta x'_j|}, \quad \hat{u}'_{j+1/2} = 0, & \text{if } \text{Sign}(u'_{j-1/2}) = \text{Sign}(u'_j) \ \& \ \text{Sign}(u'_{j+1/2}) \neq \text{Sign}(u'_j). \end{aligned} \quad (6.62)$$

After this modification, the modified  $v_j(x) = u'_j(x)$  does not change its sign over cell  $I_j$ , and so  $u_j(x)$  becomes monotone. This type of modification is presented in Figure 6.12, where the dashed (red) and solid (blue) lines represent the original  $v_j(x)$  and the modified  $v_j(x)$ , respectively.

*These updating are performed in a way that cell values  $v_j := u'_j$  do not alter.* It should be mentioned that as  $v_j$  should have second order accuracy, a linear reconstruction is used in each cell. So by the above mentioned modifications,  $u'_j = v_j$  does not alter, and the updated piecewise polynomial satisfies the conservation feature.

## 6.4 Complexity

In numerical simulations by high resolution schemes, the computational complexity is directly controlled by cell numbers and the cost can be controlled by effective adaptation procedures. For this purpose, the fast MRA-based cell/grid adaptation is used in this

work. For 1-D data of length  $M$ , wavelet transforms (with the pyramidal algorithm) use  $M$  operations and so they are fast and effective. Hence, the main computational cost is due to the central schemes. In the following, the computational complexity of the proposed central schemes is estimated by some numerical simulations (the numerical approach) [38].

### 6.4.1 The complexity of high-order central high resolution schemes

Following the references [38], the complexity of different high-order central high resolution schemes are evaluated by the numerical approach (by some simulations). Let us assume the Burgers' equation  $u_t + (1/2u^2)_{,x} = 0$ , for  $x \in [0, 1]$  where  $u := u(x, t)$ ; the boundary conditions (BCs) and the initial condition (IC) are: BCs:  $u(0, t) = u(1, t) = 0$  and IC:  $u(x, 0) = \text{Sin}(2\pi x) + 1/2 \times \text{Sin}(\pi x)$ . The parameters used in numerical simulations are:

1. Levels  $J_{max}$  and  $J_{min}$  denote the finest and coarsest resolution levels, respectively,
2. The parameter  $N_{J_{max}} = 2^{J_{max}} + 1$  is the number of uniform cells in the finest resolution level (where the sampling step is  $dx = 1/2^{J_{max}}$ ),
3. The parameter  $N_{dt}$  is the number of time-steps,
4. The ratio  $r_{N,\epsilon}^{J_{max}} = \max(N_g) / N_{J_{max}}$  denotes the minimum reduction (compression) factor in grid numbers during numerical simulations by the MRA-based grid adaptation method with the finest resolution level  $J_{max}$  and the pre-defined threshold  $\epsilon$ ,
5. The parameter  $C_{MS}$  denotes the computational time.

The complexities of the high-order central high resolution schemes are investigated separately without and with the adaptive solver.

#### Complexities of the third-order central high resolution schemes on uniform grids/cells

The main source of computational complexity is the number of cells/grids in the spatial domain, denoted here by  $M$ . In this study, different resolution levels  $j = J_{max}$  are used for simulations with location numbers  $M = N_{J_{max}} = 2^{J_{max}} + 1$ . The time step  $dt$  is assumed to be constant for different uniform grids/cells of different cell-lengths. For the methods  $M_1, M_2, CWENO - 3P, CWENO - 5P, PPM - 1$  and  $PPM - 2$ , results

are presented in Table 6.4.1 after 395 time steps at  $t = 0.158$ , when a right propagating shock wave is developed for the first time. Theoretically, the computational time  $C_{MS}$  decreases asymptotically by a factor  $r_c = (C_{MS})_{J_{max}} / (C_{MS})_{J_{max}-1}$ . Regarding Table 6.4.1, this ratio is about 2 for all central schemes and so the computational complexity is asymptotically  $\mathcal{O}(M)$  operations.

### 6.4.2 Complexities of the different third-order central high resolution schemes on adapted grids/cells

For numerical estimation of the computational complexity of adaptive simulations, the Burgers' equation is re-considered. Assumptions in the simulations are:  $\epsilon = 0.75 \times 10^{-3}$  (the wavelet threshold); the CFL value for all simulations is constant and equal to  $\frac{\Delta t}{(\min_i(\Delta x_i))} \max_j |\lambda_j| = 0.25$ ; the coarsest level  $J_{min}$  is kept constant and equals to 5 for all simulations (with the sampling step  $\Delta x = 1/2^{J_{min}}$ ); to capture the fine resolution effects,  $J_{max}$  is increased one-by-one from 9 to 12; the MRA-based grid adaptation is performed after each time step. Results are presented in Table 6.4.2 at  $t = 0.158$  (the time of developing the first discontinuity in solutions).

The relative number  $r_{(N,\epsilon)}^{J_{max}}$  decreases asymptotically by a factor  $r_j := r_{(N,\epsilon)}^{J_{max}-1} / r_{(N,\epsilon)}^{J_{max}}$  of about 2. The total number of adapted cells increases by a factor  $q_l := \frac{[\max(N_g)]_{J_{max}-1}}{[\max(N_g)]_{J_{max}}} \approx 2r_j$ . The ratio  $r_j$  is around 0.5 and so  $q_l \approx 1$ . This means that the computational complexity in the spatial domain does not vary considerably by adding an extra finer resolution level for capturing fine responses. This reveals the importance of adaptive solvers in the spatial-domain. On the other hand, by adding an additional refinement, the time step decreases by a factor 0.5 due to the CFL condition. Hence, the total computational complexity in the spatio-temporal domains would increase by  $2 \times q_l$ . This can be confirmed by the ratio of  $r_c := (C_{MS})_{J_{max}} / (C_{MS})_{J_{max}-1}$  which is around 2 for each scheme.

Table 6.4.1: The computation cost of different high-order central schemes on uniform cells; for all cases  $J_{min} = 5$ .

$J_{max}$	$N_{J_{max}}$	$M_1$	$M_2$	$CWENO - 3P$	$CWENO - 5P$	PPM-1	PPM-2
		$C_{MS}$	$C_{MS}$	$C_{MS}$	$C_{MS}$	$C_{MS}$	$C_{MS}$
6	65	0.64	0.87	0.69	0.87	1.13	1.04
7	129	1.22	1.72	1.29	1.73	2.13	2.03
8	257	2.44	3.38	2.57	3.39	3.90	3.91
9	513	4.73	6.69	4.97	4.76	7.52	7.43

Table 6.4.2: The computation cost of different high-order central schemes on adaptive cells. The parameter  $C_{MS}$  is measured in Minutes and  $J_{min} = 5$  for all cases.

$J_{max}$	$N_{J_{max}}$	$N_{dt}$	$M_1$	$M_2$	$CWENO - 3P$	$CWENO - 5P$	PPM-1	PPM-2
			$(r_{(N,\epsilon)}^{J_{max}}(\%), C_{MS})$	$(r_{(N,\epsilon)}^{J_{max}}(\%), C_{MS})$	$(r_{(N,\epsilon)}^{J_{max}}(\%), C_{MS})$	$(r_{(N,\epsilon)}^{J_{max}}(\%), C_{MS})$	$(r_{(N,\epsilon)}^{J_{max}}(\%), C_{MS})$	$(r_{(N,\epsilon)}^{J_{max}}(\%), C_{MS})$
9	513	485	(16.29, 0.60)	(19.18, 0.77)	(22.47, 0.75)	(25.15, 0.76)	(18.97, 1.06)	(17.53, 0.84)
10	1025	970	(8.04, 1.29)	(9.59, 1.76)	(10.31, 1.54)	(11.75, 1.51)	(9.48, 2.20)	(9.59, 1.95)
11	2049	1940	(4.69, 2.91)	(4.58, 3.96)	(6.34, 3.09)	(6.23, 3.23)	(5.51, 5.4)	(5.46, 4.29)
12	4097	3880	(2.34, 6.80)	(2.24, 9.47)	(3.10, 7.04)	(3.20, 7.69)	(2.97, 13.19)	(2.68, 10.26)

## 6.5 Reconstruction Error

In this study, the performance of different reconstruction schemes is measured by considering reconstruction errors at cell-interfaces [184, 185]. The error norm proposed in [184, 185] is considered in this study:

$$Err^p := \frac{1}{N} \left( \| u_{i+1/2}^- - f(x_{i+1/2}) \|_{L_{\Delta x}^p} + \| u_{i+1/2}^+ - f(x_{i+1/2}) \|_{L_{\Delta x}^p} \right), \quad (6.63)$$

with the norm definition:

$$\| (\cdot)_i \|_{L_{\Delta x}^p} := \left( \sum_{i=1}^N |(\cdot)_i|^p \right)^{1/p} \quad \text{and} \quad \| (\cdot)_i \|_{L_{\Delta x}^\infty} := \max_i |(\cdot)_i|, \quad (6.64)$$

where  $N$  denotes the number of interfaces  $\{x_{i+1/2}\}$  and  $u_{i+1/2}^-$  and  $u_{i+1/2}^+$  are the left and right reconstructed values at  $x_{i+1/2}$ , respectively.

All of the *test* problems are one dimensional because the dimension-by-dimension method is used for high-dimensional problems. For different reconstruction schemes, errors are evaluated on multiresolution-based adapted grids to study the scale effects. Two test problems are selected; the first one has a smooth-high gradient response and the second one includes a discontinuity. These test problems were proposed by [139] for studying the performance of multiscale (adaptive) solvers.

It should be mentioned that the reconstruction procedure in PPMs include the  $M_1$  scheme for state variables and linear reconstruction by the GMM for the velocity and acceleration of the state variables. So, in the following, the reconstruction performance of the  $M_1$ ,  $M_2$ , linear reconstruction, *CWENO* –  $3P$  and *CWENO* –  $5P$  schemes are studied.

### 6.5.1 The first test problem: the smooth-high gradient function

The test function is defined as [139]:

$$f_1(x) := \text{Sin}(2\pi x) + \text{Exp}\left(-20000(x - 0.5)^2\right). \quad (6.65)$$

In numerical evaluations, the parameters are:  $\epsilon = 0.5 \times 10^{-4}$  (the wavelet threshold);  $J_{min} = 5$  (the coarsest resolution level) and  $J_{max} = \{7, 8, 9, 10, 11\}$  (the finest resolution level). The function  $f_1(x)$  and corresponding adapted grids for different  $J_{max}$

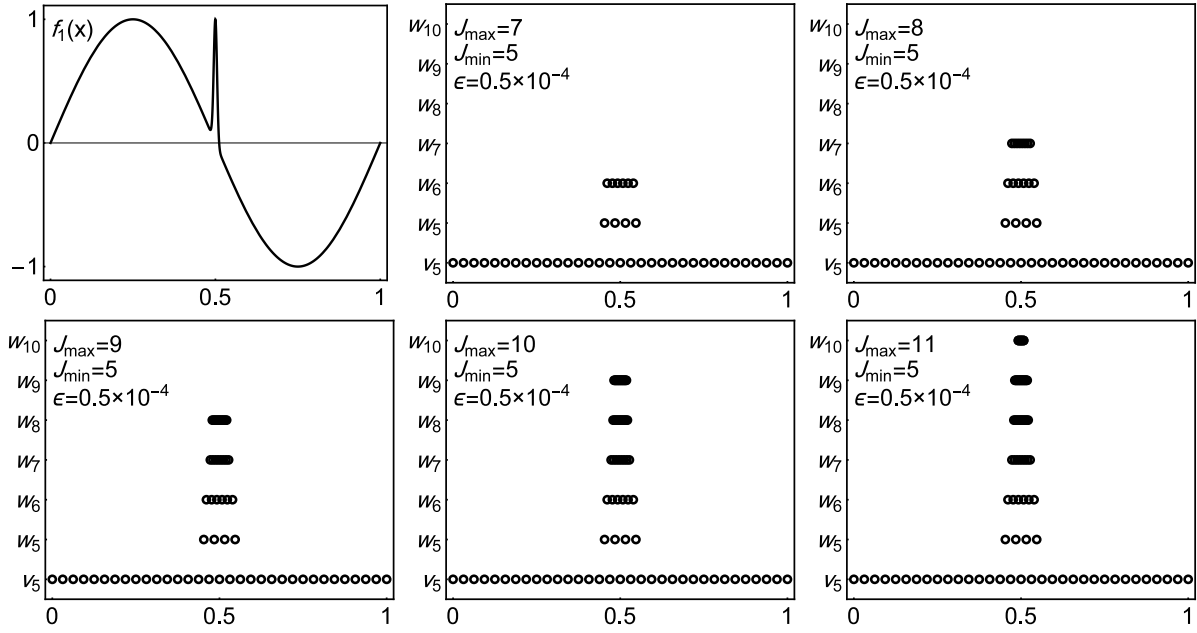


Figure 6.13: The first test function  $f_1(x)$  and distribution of adapted grids in different resolution levels for different  $J_{max}$  values.

values are illustrated in Fig. 6.13. It is clear that by increasing  $J_{max}$  values, more adapted points with finer scales concentrate around the high gradient zone. For these adapted grids, the reconstruction errors with  $p = 1$  with different  $J_{max}$  values are presented in Tables 6.5.1 and 6.5.2. Table 6.5.1 includes methods  $M_1$ ,  $M_2$  and the linear reconstruction by the generalized MINMOD limiter with the parameter  $\theta = 2$ . In Table 6.5.2, the performance of the methods  $CWENO - 3P$  and  $CWENO - 5P$  are provided for different  $C_L$ ,  $C_R$  and  $C_C$  values, where  $C_L = C_R$  and  $C_C = 1 - C_L - C_R$ .

Table 6.5.1: Reconstruction error at interfaces with  $p = 1$  for the first test function.

$J_{max}$	$M_1$	$M_2$	$L - GMM$
7	0.021833	0.0203572	0.031139
8	0.010856	0.0088252	0.0183609
9	0.005245	0.0032685	0.0109897
10	0.0032004	0.0016517	0.0072496
11	0.0047077	0.0022905	0.00957725

Table 6.5.2: Reconstruction error at interfaces with  $p = 1$  for the first test function.

$J_{max}$	<i>CWENO</i> – 3 <i>P</i>			<i>CWENO</i> – 5 <i>P</i>		
	$C_{L,R} = 0.15$	$C_{L,R} = 0.25$	$C_{L,R} = 0.35$	$C_{L,R} = 0.15$	$C_{L,R} = 0.25$	$C_{L,R} = 0.35$
7	0.028974	0.0308683	0.0317729	0.0207154	0.0213235	0.021617
8	0.016985	0.01898	0.019901	0.00667118	0.00636655	0.00583898
9	0.00775609	0.00893157	0.00963636	0.000830034	0.00084366	0.00161227
10	0.00329951	0.00397226	0.00444691	0.000173458	0.000296736	0.000591564
11	0.00240646	0.0029527	0.00333988	0.000137795	0.00024083	0.000464516

### 6.5.2 The second test problem: a function including a discontinuity

This test function is defined as [139]:

$$f_2(x) := \text{Sin}(2\pi x) - \text{H}\left(x - \frac{1}{2}\right) + \frac{1}{2}, \quad (6.66)$$

where the function  $\text{H}(x)$  denotes the Heaviside (the unite step) function. The assumed parameters are:  $\epsilon = 0.5 \times 10^{-4}$  (the wavelet threshold);  $J_{min} = 5$  (the coarsest resolution level) and  $J_{max} = \{7, 8, 9, 10, 11, 12\}$  (the finest resolution level). The function  $f_2(x)$  and corresponding adapted grids for different  $J_{max}$  values are illustrated in Fig. 6.14. Reconstruction errors for  $p = 1$  for different central schemes are presented in Tables 6.5.3 and 6.5.4. Schemes  $M_1$ ,  $M_2$  and the linear reconstruction are presented in Table 6.5.3 and different CWENO schemes with different symmetric weights are compared in Tables 6.5.4.

The results from these two test problems offer that:

1. For smooth high gradient cases, the best performance is for the *CWENO* – 5*P*; however, around a discontinuity its performance reduces due to stencil length in comparison to other schemes,
2. For smooth high gradient solutions, performance of all central schemes is better than the linear reconstruction,
3. The MRA-based adaptation can reduce the reconstruction error in all schemes; indeed, by including an extra finer resolution level, reconstruction errors decrease,
4. All CWENO schemes with different weight values work properly,
5. Presence of a discontinuity considerably reduces the convergence rate.

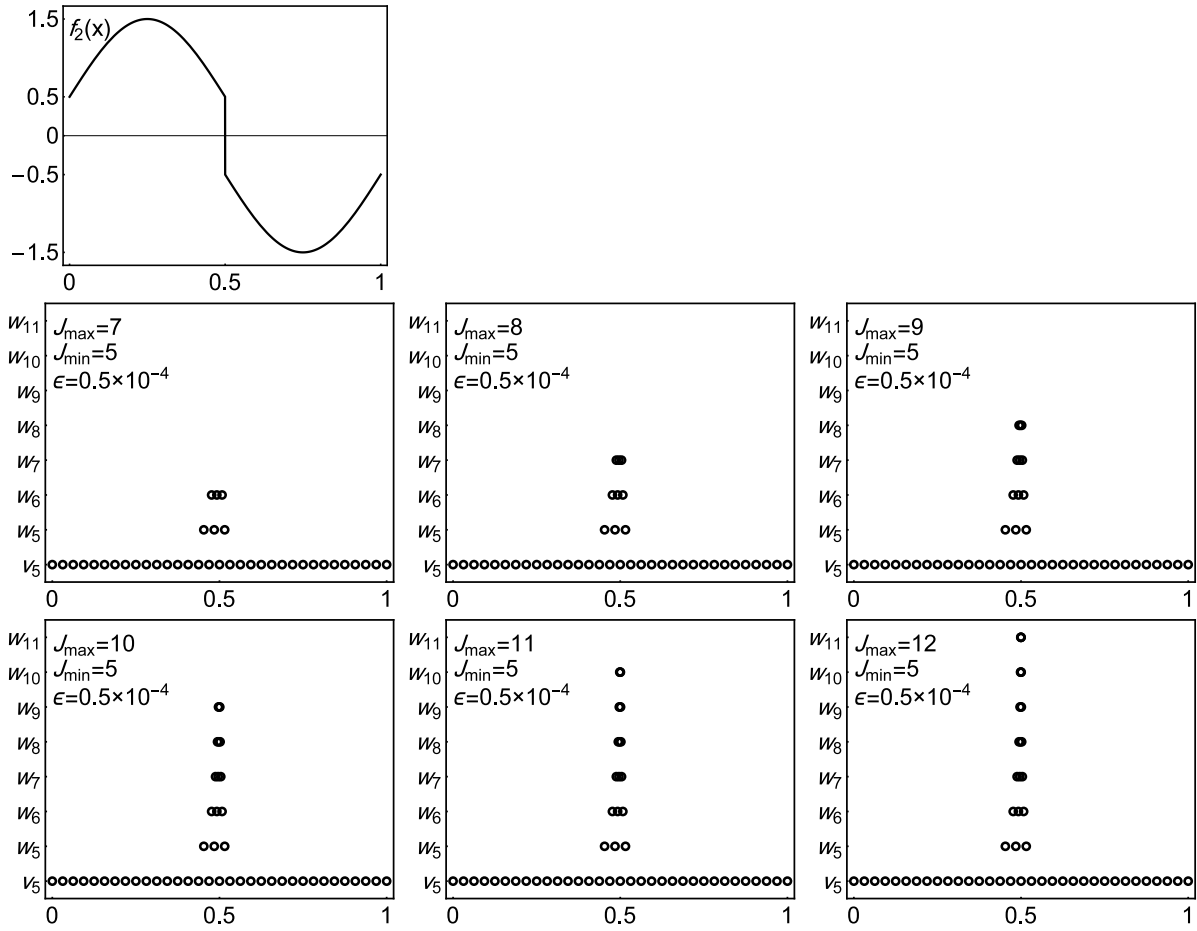


Figure 6.14: The second test function  $f_2(x)$  and distribution of adapted grids in different resolution levels for different  $J_{\max}$  values.



Table 6.5.3: Reconstruction error at interfaces with  $p = 1$  for the second test function.

$J_{max}$	$M_1$	$M_2$	$L - GMM$
7	0.035328	0.036003	0.0504697
8	0.0326298	0.0334207	0.0464694
9	0.0303033	0.0311699	0.0431562
10	0.02831257	0.0291939	0.0403336
11	0.0265889	0.0274487	0.0378817
12	0.0250701	0.0258981	0.0357224

Table 6.5.4: Reconstruction error at interfaces with  $p = 1$  for the second test function.

$J_{max}$	<i>CWENO</i> – 3 <i>P</i>			<i>CWENO</i> – 5 <i>P</i>		
	$C_{L,R} = 0.15$	$C_{L,R} = 0.25$	$C_{L,R} = 0.35$	$C_{L,R} = 0.15$	$C_{L,R} = 0.25$	$C_{L,R} = 0.35$
7	0.0305937	0.0315606	0.0321786	0.0792499	0.0793677	0.0796173
8	0.0283527	0.0292479	0.0298207	0.0734487	0.0735584	0.0737891
9	0.0264191	0.0272533	0.027787	0.0684397	0.0685418	0.0687567
10	0.0247328	0.0255137	0.0260133	0.0640524	0.064148	0.0643492
11	0.0232488	0.0239828	0.0244525	0.0600356	0.0601255	0.0603146
12	0.0219328	0.0226253	0.0230684	0.0554995	0.0555843	0.0557627

## 6.6 Numerical examples

The following 1-D and 2-D examples are adopted to study and compare the effectiveness of the proposed methods for simulation of nonlinear first order hyperbolic systems. The main assumptions are:

1. Applying the D-D interpolating wavelet of order 3,
2. Repeating re-adaptation processes every time step for 1-D problems and every twenty step for 2-D examples,
3. Using the semi-discrete form of central high resolution schemes,
4. Temporal integration is performed by the TVD Runge-Kutta solvers. The third-order Runge-Kutta method is used for the third and fourth order central schemes and the second-order Runge-Kutta scheme for the second-order central schemes.

For the PPMs, for original (Eq. (6.51)) and newly added hyperbolic equations (Eqs. (6.52) and (6.53)), different central solvers with different accuracy are used:

1. The original system by the third-order solver using PPM concepts,

2. For additional equations, the second-order Kurganov-Tadmor (KT) central high resolution scheme (using the generalized MINMOD limiter) [22] is adopted. It should be mentioned that theoretically the first order scheme is sufficient for  $a$ ; however, second order solvers using TVD limiters have first order accuracy around extrema and zero slope at extrema. Also in case of multidimensional problems, TVD-based solvers lead to at most first order accuracy. These conditions lead to the over-smoothing phenomenon. In this regard, to control numerical dissipation, for both velocities and accelerations the second order solvers are used.

The semi-discrete form of the KT method for the scalar conservation law  $u_t + F(u)_x = 0$  (on uniform and non-uniform cells) is [22]:

$$\frac{d\bar{u}_i}{dt} + \frac{F_{i+1/2}^* - F_{i-1/2}^*}{\Delta x_i} = 0, \quad (6.67)$$

where  $\Delta x_i := x_{i+1/2} - x_{i-1/2}$ ;  $\bar{u}_i$  is the average solution on cell  $I_j$  and since variation of  $u(x)$  on  $I_j$  is linear, the average solution  $\bar{u}_i$  and  $u_i$  are equal to each other, that is  $\bar{u}_i = u_i$ ;  $F_{i\pm 1/2}^*$  denotes the numerical flux at cell edges  $x_{i\pm 1/2}$ , defined as:

$$F_{i\pm 1/2}^* := \frac{F_{i\pm 1/2}^R + F_{i\pm 1/2}^L}{2} + a_{i\pm 1/2} \frac{u_{i\pm 1/2}^R - u_{i\pm 1/2}^L}{2}, \quad (6.68)$$

where  $F_{i+1/2}^R$  and  $F_{i+1/2}^L$  are the right and left reconstructed fluxes at the edge  $x_{i+1/2}$ , respectively;  $a_{i+1/2}$  denotes the absolute value of maximum velocity at  $x_{i+1/2}$ .

For the above-mentioned nonlinear conservation law, the Riemann problem can be described by the following initial condition (IC):

$$u(x, t = 0) = \begin{cases} u_L, & x \leq x_0, \\ u_R, & x > x_0. \end{cases} \quad (6.69)$$

For the PPM methods, it is also essential to evaluate initial conditions (ICs) for  $v(x, t = 0) := \partial u(x, 0)/\partial x$  and  $a(x, t = 0) := \partial^2 u(x, 0)/\partial x^2$ , based on  $u(x, t = 0)$ . These ICs can be obtained as:

$$v(x, t = 0) = (u_R - u_L)\delta(x - x_0), \quad a(x, t = 0) = 0, \quad (6.70)$$

where  $\delta(x)$  is the Dirac delta function. This function can be approximated as:

$$\delta(x - x_0) \approx \begin{cases} \frac{1}{\Delta}, & \text{for } x_0 - \frac{\Delta}{2} \leq x \leq x_0 + \frac{\Delta}{2}, \\ 0, & \text{elsewhere.} \end{cases} \quad (6.71)$$

In numerical simulations, a proper choice for  $\Delta$  may be  $\Delta/2 = 2 \times (1/2^{J_{max}})$ .

In results presented in the following, different methods are denoted by:

1.  $M_1$  and  $M_2$ : the third order limited solvers on three successive cells defined respectively in Eqs. (6.30) and (6.35),
2. PPM-1 and PPM-2: the piecewise parabolic methods respectively with extra information  $\{v_j, a_j\}$  and  $\{v_{j\pm 1/2}\}$ ,
3. CWENO-3P and CWENO-5P: central-WENO schemes respectively with three and five-points stencils .

### 6.6.1 The 1-D Burgers' equation

The Burgers' equation is defined as:

$$u_t + \frac{1}{2} (u^2)_x = 0, \quad x \in [0, 1],$$

where  $u$  is the conserved quantity and its flux is  $F(u) = u^2/2$ . The system is nonlinear, so discontinuous fronts will develop during propagation. A smooth IC with fixed boundary conditions (BCs) are assumed as:

$$BCs : u(x = 0, t) = u(1, t) = 0, \quad IC : u(x, t = 0) = \sin(2\pi x) + \frac{1}{2}\sin(\pi x).$$

For this IC, a right-moving discontinuity develops around  $t \approx 0.158$ . Numerical assumptions are:  $\epsilon = 10^{-3}$  (threshold),  $N_c = N_s = 2$  (for grid modification in the post-processing stage),  $J_{max} = 10$ ,  $J_{min} = 5$  and  $\theta = 2$  (for the generalized MINMOD limiter in PPM schemes).

The numerical results are illustrated in Fig. 6.15 at times 0.158, 0.5 and 1. This figure contains numerical results, exact solutions and corresponding adapted grids in different

resolutions. The solid lines and shapes denote the exact and numerical solutions, respectively.

Variations of both the number of grid points ( $N_g$ ) and TV (total variation) in time are presented in Fig. 6.16 for different methods.

The results confirm that:

1. High-order high-resolution methods are properly integrated with the adaptive solver;
2. The adaptation procedure is effective; since, instead of 1025 points (in the finest resolution level) nearly less than 80 adapted points are used for all methods;
3. Adapted points are properly concentrated around the propagating fronts;
4. Solutions are TVB (having bounded TV);
5. Localized spurious oscillations develop around the discontinuity for  $M_1$ ,  $M_2$ , PPM-1 and PPM-2 schemes; the PPM schemes, however, can decrease such localized oscillations;
6. The CWENO-5P method leads to more precise results than CWENO-3P (compare solutions at  $t = 0.158$ );
7. Performance of both PPM-1 and PPM-2 methods are nearly the same.

## 6.6.2 Euler system of equations

In the following, four different problems with different initial and boundary conditions will be studied. The examples are: 1) The Sod problem [186], 2) The Lax problem [187], 3) Interaction of an entropy sine wave with a Mach 3 right-moving front [71, 174], 4) Interaction of two blasts [174, 175, 190].

### The Sod Problem

For this benchmark, the initial conditions are:

$$\begin{pmatrix} \rho \\ u \\ P \end{pmatrix}_{t=0} = \begin{cases} \{0, 0, 1\}^T, & x \leq 0.5, \\ \{0.125, 0, 0.1\}^T, & x > 0.5. \end{cases}$$

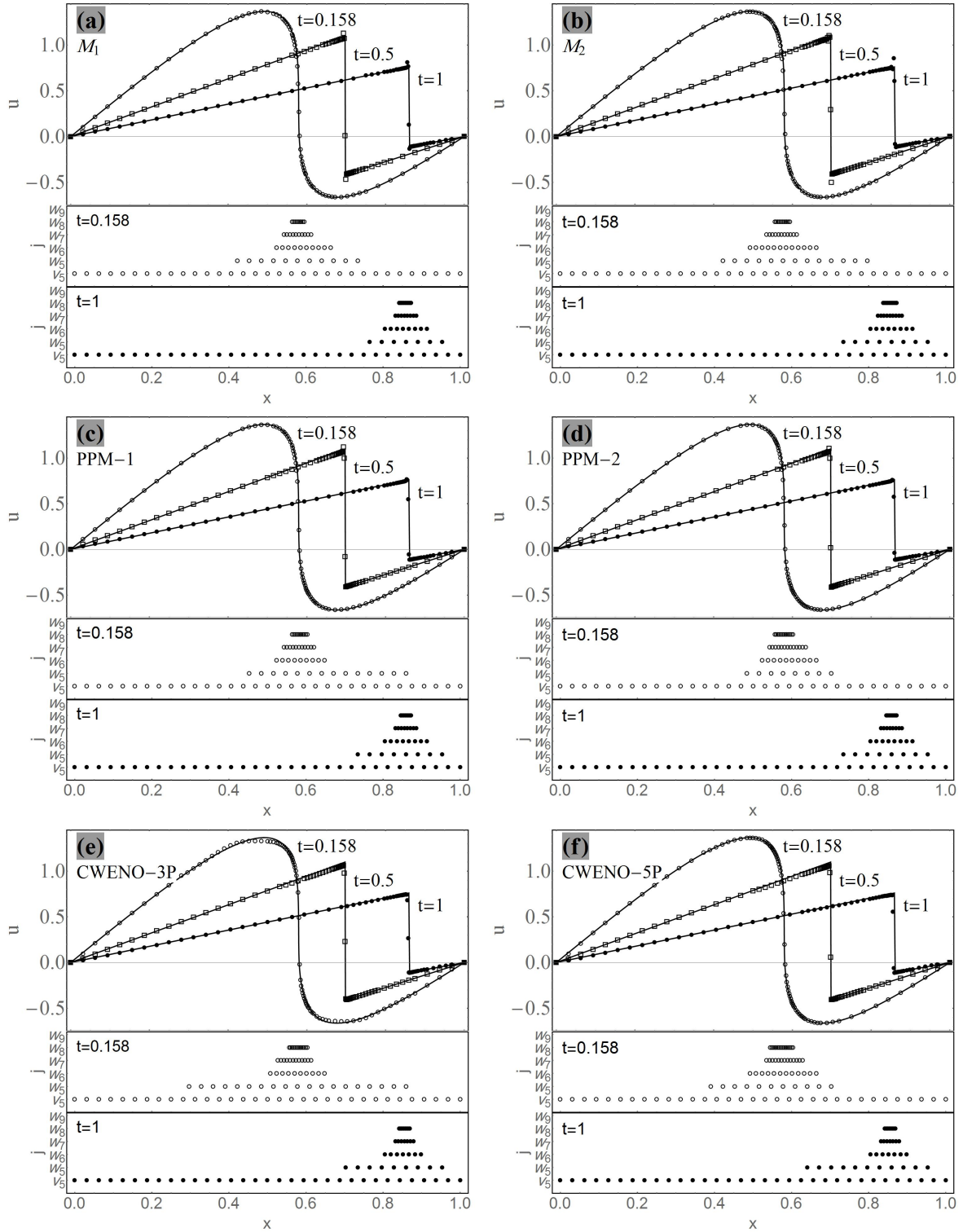


Figure 6.15: Numerical results and corresponding adapted points in different resolutions for the Burgers' problem with different approaches.

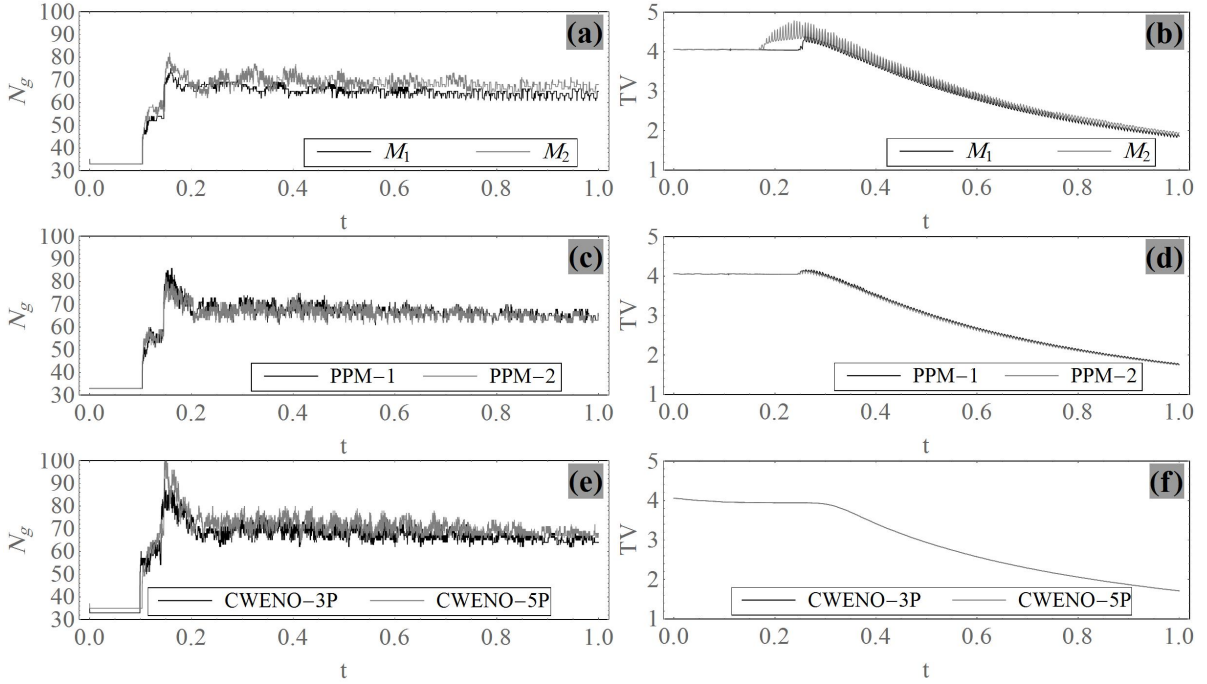


Figure 6.16: Number of adapted grid points  $N_g$  and corresponding  $TV$  during simulations with different approaches. In the finest resolution, the number of grid points is  $2^{10} + 1$ .

An unbounded 1-D domain is assumed: a Riemann problem. Assumptions for numerical simulations are:  $J_{max} = 11$ ,  $N_d = 6$ ,  $\epsilon = 10^{-4}$ ,  $N_c = 2$ ,  $N_s = 1$ ,  $dt = 0.0000244$  and  $\theta = 2$  (in the generalized MINMOD limiter).

The numerical solutions for  $\rho$  and corresponding adapted grid points at time 0.2 are illustrated in Fig. 6.17. The solid lines and shapes denote the exact and numerical solutions, respectively. Variations of both the grid point number ( $N_g$ ) and  $TV$  of  $\rho$  in time are presented in Fig. 6.18 for different methods.

Error in the adaptive simulations measured by  $L^1$  and  $L^\infty$  norms are presented in Table 6.6.1 at  $t = 0.2$ . The definitions of these norms are:

$$\|(\cdot)\|_{L^1} := \frac{1}{N_g} \sum_{i=1}^{N_g} |(\cdot)_i|, \quad \|(\cdot)\|_{L^\infty} := \max_i |(\cdot)_i|, \quad (6.72)$$

where  $N_g$  denotes the number of adapted points.

The  $L^1$  norm of the error as a function of the run-time for  $M_1$ ,  $M_2$ ,  $CWENO - 3P$ ,  $CWENO - 5P$ ,  $PPM - 1$  and  $PPM - 2$  are presented in Fig. 6.19, evaluated for the interval  $0 \leq t \leq 0.2$ . The assumed parameters are:  $J_{max} = 11$ ,  $N_d = 6$  (the

number of decomposition levels),  $N_c = 2$ ,  $N_s = 1$ ,  $dt = 0.0000244$ ,  $C_L = C_R = 0.25$  and the adaptation procedure is performed at each time step (which is expensive). In Fig. 6.19, in each curve, the plot-markers correspond from left to right, respectively, to threshold values  $\epsilon \in \{5 \times 10^{-3}, 10^{-3}, 5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-5}, 10^{-5}, 10^{-6}\}$ . All codes are written in *Mathematica 10* and nearly all functions are compiled in the Mathematica (by the *Compile* function). The results illustrate that the PPMs are the most expensive approach and the *CWENO - 5P* is the most accurate method.

Table 6.6.1: The accuracy of different schemes for solving the Sod problem at  $t = 0.2$ .

Error	$M_1$	$M_2$	<i>CWENO - 3P</i>	<i>CWENO - 5P</i>	<i>PPM - 1</i>	<i>PPM - 2</i>
$L^1$	3.380e-03	3.319e-03	2.759e-03	2.229e-03	3.345e-03	2.821e-03
$L^\infty$	8.326e-02	7.914e-02	8.877e-02	9.061e-02	6.171e-02	6.773e-02

### The Lax problem

For this case, ICs are:

$$\begin{pmatrix} \rho \\ u \\ P \end{pmatrix}_{t=0} = \begin{cases} \{0.445, 0.69887, 3.5277\}, & x \leq 0.5, \\ \{0.5, 0, 0.571\}, & x > 0.5. \end{cases}$$

For numerical simulations, it is assumed:  $\epsilon = 5 \times 10^{-4}$ ,  $J_{max} = 11$ ,  $J_{min} = 5$  (or  $N_d = 6$ ),  $dt = 2 \times 10^{-5}$  and  $\theta = 2$  (in the generalized MINMOD limiter).

Adaptive solutions of  $\rho$  and corresponding adapted grids at different resolution levels at time 0.16 are presented in Fig. 6.20 for different methods. The solid lines and shapes represent the exact and numerical solutions, respectively. Variations of  $N_g$  and TV of  $\rho$  in time are presented in Fig. 6.21 for different methods. Error in the adaptive simulations measured by the  $L^1$  norm is also provided in Table 6.6.2 at  $t = 0.2$ .

The  $L^1$  norm of the error as a function of the run-time for different schemes are presented in Fig. 6.22, evaluated for  $0 \leq t \leq 0.16$ . The assumed parameters are the same as the Sod problem, except for  $dt = 2 \times 10^{-5}$ . The results offer again that the PPMs are the most expensive approach and the *CWENO - 5P* is the most accurate method.

Regarding the Sod and the Lax problems, the results offer that:

1. The wavelet transform can properly capture all phenomena: shock waves, rarefaction and contact discontinuity zones,

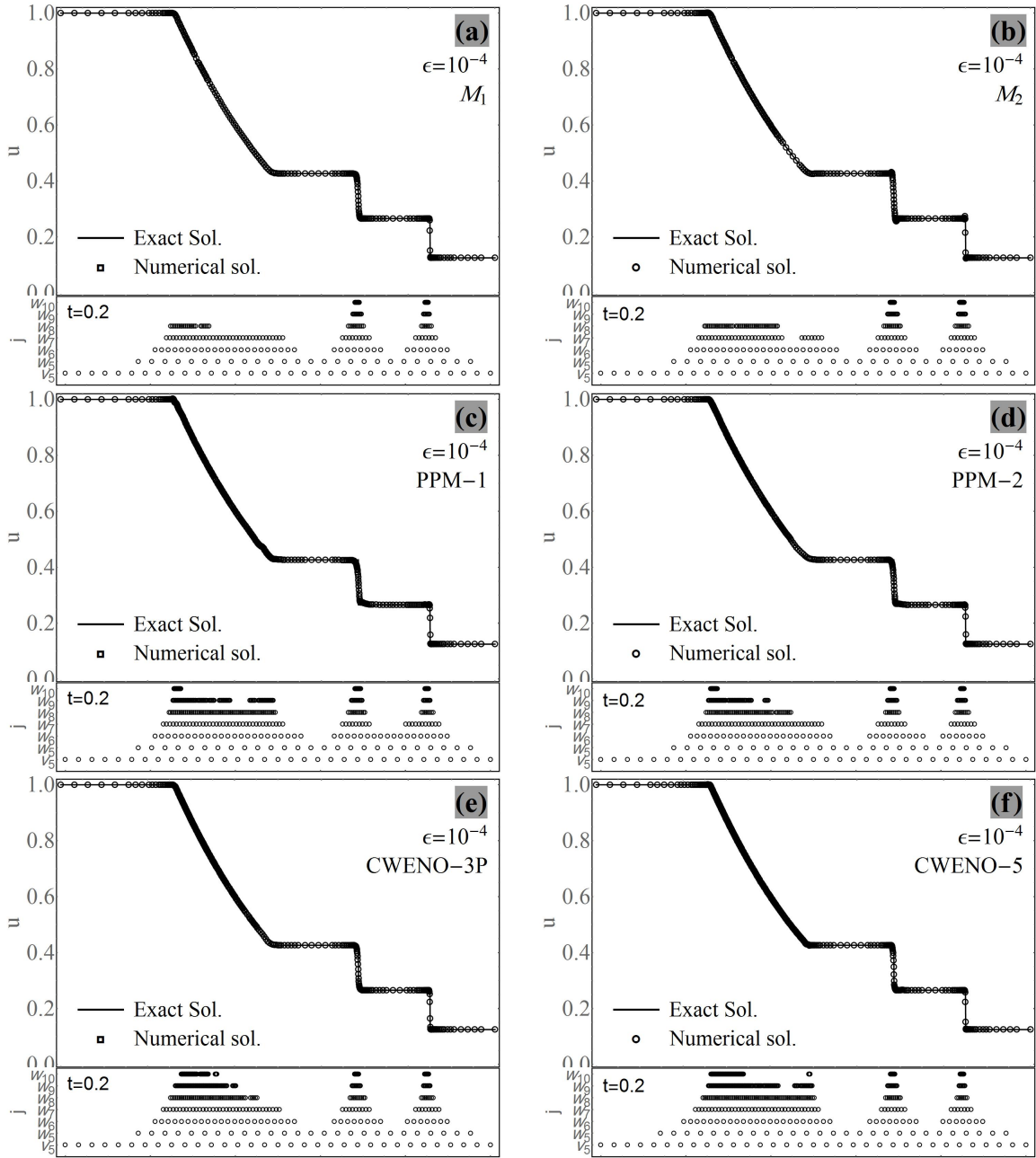


Figure 6.17: Numerical results of  $\rho$  ( $u \equiv \rho$ ) and corresponding adapted points in different resolutions for the Sod problem with different approaches at  $t = 0.2$ . Solid lines and shapes denote exact and numerical solutions, respectively.



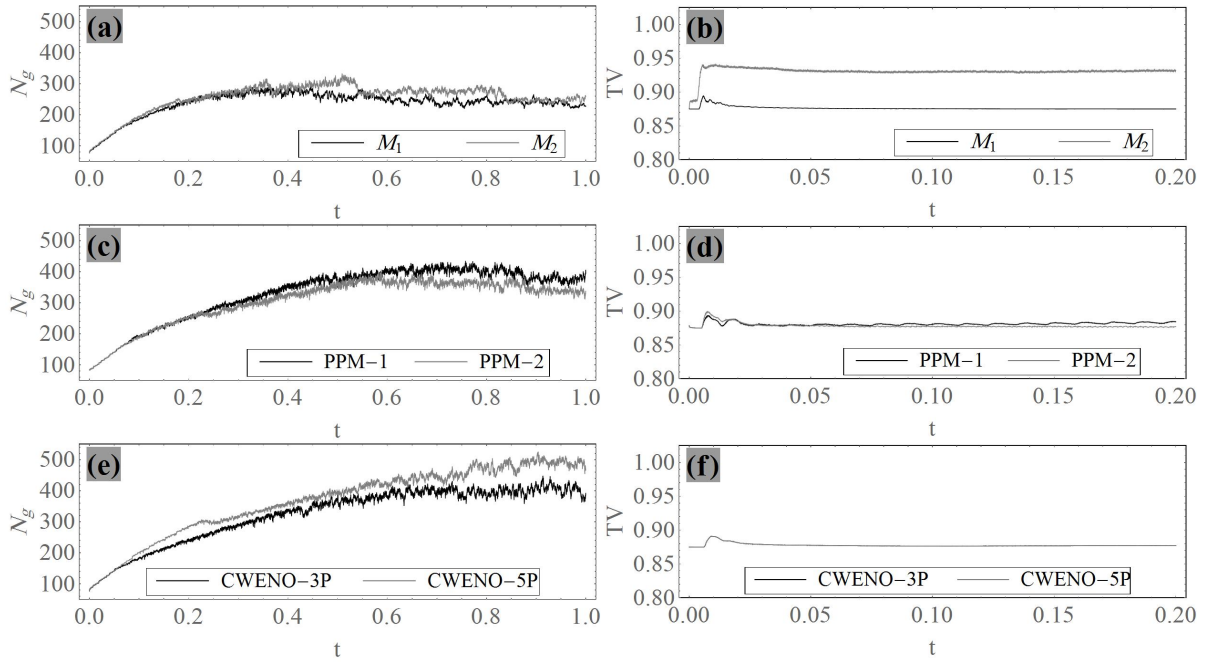


Figure 6.18: Number of adapted grid points  $N_g$  and corresponding total variations  $TV$  of  $\rho$  for the Sod problem during simulations with different approaches. In the finest resolution, number of grid points is  $2^{11} + 1$ .

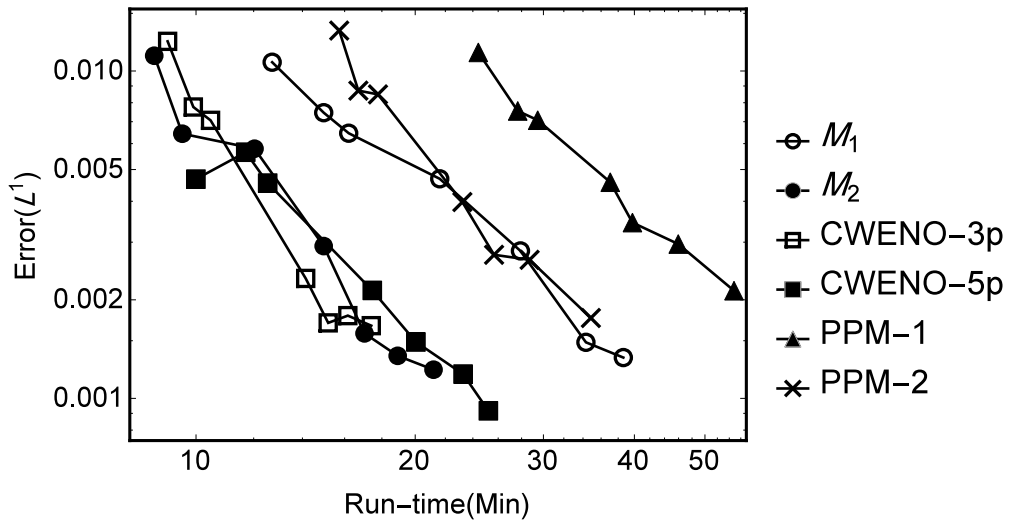


Figure 6.19:  $L^1$  error as a function of run-time for different schemes for the Sod problem, evaluated for  $0 \leq t \leq 0.2$ .

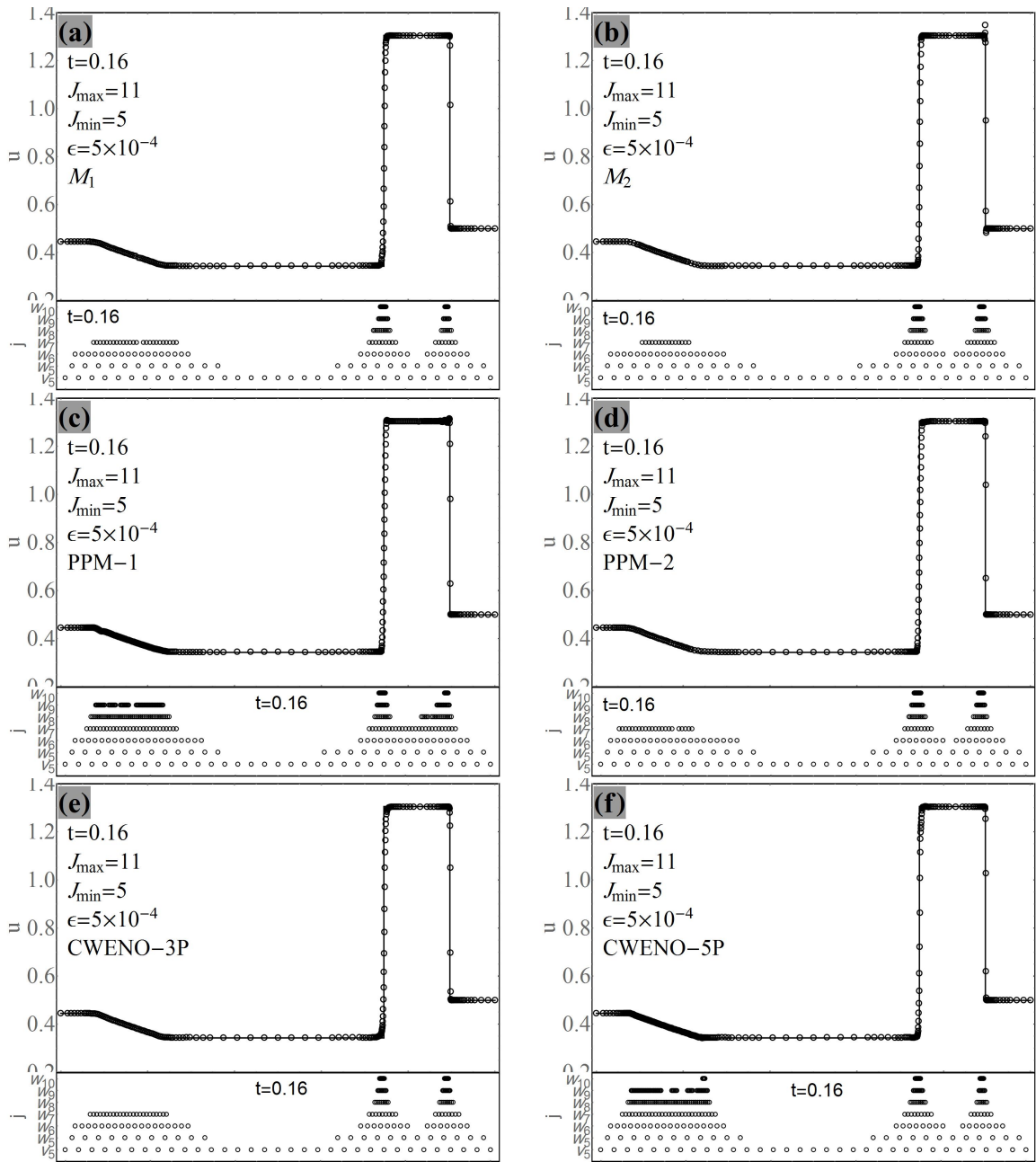


Figure 6.20: Numerical results of  $\rho$  ( $u \equiv \rho$ ) and corresponding adapted points in different resolutions for the Lax problem with different approaches at  $t = 0.16$ .

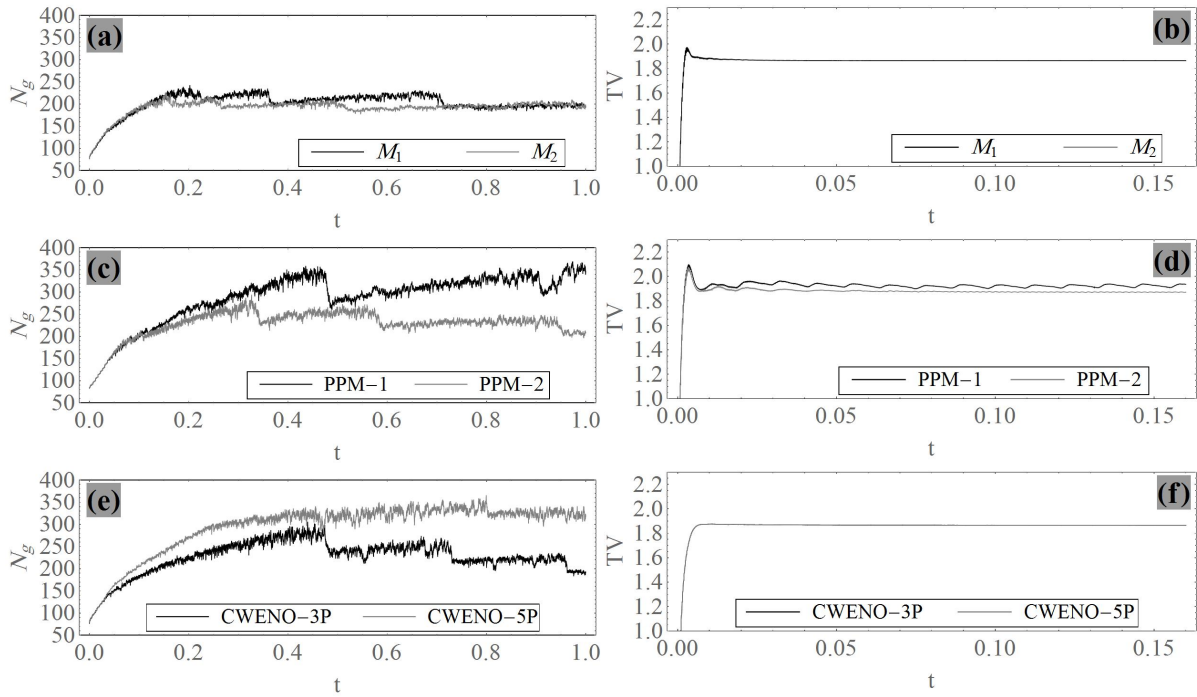


Figure 6.21: Number of adapted grid points  $N_g$  and corresponding  $TV$  of  $\rho$  during simulations for the Lax problem with different approaches. In the finest resolution, the number of grid points is  $2^{11} + 1$ .

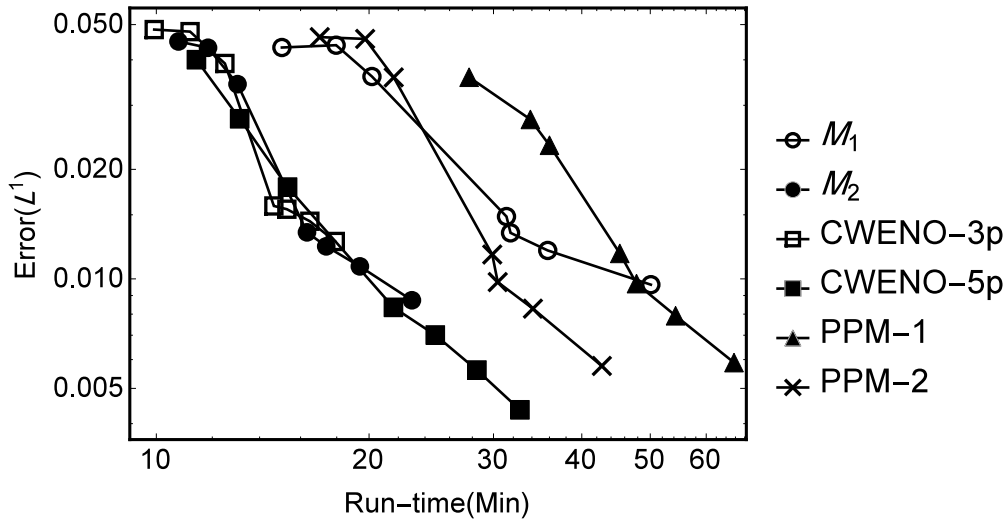


Figure 6.22:  $L^1$  error as a function of run-time for different schemes for the Lax problem, evaluated for  $0 \leq t \leq 0.16$ .

Table 6.6.2: Accuracy for different schemes for solving the Lax problem at  $t = 0.16$ .

Error	$M_1$	$M_2$	$CWENO - 3P$	$CWENO - 5P$	$PPM - 1$	$PPM - 2$
$L^1$	2.907e-02	2.145e-02	3.199e-02	1.637e-02	2.237e-02	3.624e-02

2. Grid points are properly adapted for all schemes,
3. Number of grid points used by adaptation procedures are considerably smaller than corresponding uniform grid in the finest resolution (of size  $2^{11} + 1$ ),
4. The  $M_2$  method leads to less dissipation around discontinuities (this may increase spurious oscillations),
5. Total variations of all solutions are bounded through time: the TVB feature,
6. Schemes  $M_1$  and  $M_2$  use smaller number of  $N_g$  through time in comparison to other methods,
7. CWENO schemes can completely control developing of spurious oscillations and so have the smoothest TV in time,
8. The CWENO-5P method mobilizes more adapted points than the CWENO-3P scheme,
9. The PPM-2 scheme leads smoother results than the PPM-1.

### The Shu and Osher problem

In this example, the interaction of an entropy sine wave with a Mach 3 right-moving front (known as the Shu and Osher problem) [73, 174] are studied. This challenging problem was developed to reveal capabilities of high order schemes.

The Riemann initial condition is [174, 175]:

$$\{\rho, u, P\}|_{t=0} = \begin{cases} \{3.857143, 2.629369, 10.33333\} & \text{if } x \leq -4, \\ \{1 + 0.2 \sin(5x), 0, 1\} & \text{if } x > -4.. \end{cases} \quad (6.73)$$

The considered spatio-temporal computational domain is:  $\Omega = (-5, 5) \times (0, T)$ . In numerical simulations, the assumed parameters are:  $\epsilon = 10^{-3}$ ,  $J_{max} = 12$ ,  $J_{min} = 5$  (or  $N_d = 6$ ),  $N_c = 1$ ,  $N_s = 2$ ,  $\theta = 2$  (for the GMM limiter) and  $dt = 0.000976563$ .

The numerical and exact solutions and corresponding adapted points are illustrated in Figs. 6.23 and 6.24 for the spatial range  $x \in [-3, 2.5]$  at  $t = 1.8$ . In these figures, the

solid line and hollow-circle markers denote the reference [175] and numerical solutions, respectively. In the PPM methods, for simulation of newly added PDEs (including velocities and accelerations), the KT formulation and two central-upwind schemes with less numerical dissipation are used [23, 24]. The numerical fluxes for the central-upwind schemes are defined as:

1. The first approach, denoted by  $C_1$ :

$$F_{i+1/2}^* := \frac{a_{i+1/2}^+ F_{i+1/2}^- - a_{i+1/2}^- F_{i+1/2}^+}{a_{i+1/2}^+ - a_{i+1/2}^-} + a_{i+1/2}^- a_{i+1/2}^+ \frac{u_{i+1/2}^+ - u_{i+1/2}^-}{a_{i+1/2}^+ - a_{i+1/2}^-}, \quad (6.74)$$

2. The second approach, denoted by  $C_2$ :

$$F_{i+1/2}^* := \frac{a_{i+1/2}^+ F_{i+1/2}^- - a_{i+1/2}^- F_{i+1/2}^+}{a_{i+1/2}^+ - a_{i+1/2}^-} + \frac{a_{i+1/2}^- a_{i+1/2}^+}{2} \frac{u_{i+1/2}^+ - u_{i+1/2}^-}{a_{i+1/2}^+ - a_{i+1/2}^-}, \quad (6.75)$$

where  $F_{i+1/2}^-$  and  $F_{i+1/2}^+$  denote fluxes just before and after the interface  $x_{i+1/2}$ , respectively;  $u_{i+1/2}^-$  and  $u_{i+1/2}^+$  denote the reconstructed values just before and after  $x_{i+1/2}$ , respectively; and  $a_{i+1/2}^-$  and  $a_{i+1/2}^+$  denote the maximum left and right propagation speeds, respectively.

The results offer that:

1. For all schemes, adapted points concentrate properly around both high-gradient zones and discontinuities,
2. For the  $M_1$ ,  $M_2$ ,  $CWENO - 3P$  and  $CWENO - 5P$ , there are good agreements between numerical and reference solutions,
3. For the  $PPM - 1$  and  $PPM - 2$  schemes, all responses (obtained by the KT,  $C_1$  and  $C_2$  schemes) include the over-smoothing phenomenon around extrema. However, the  $C_1$  and  $C_2$  central-upwind schemes can generally reduce numerical dissipation. It seems that the main reason is the performance of the GMM limiter around the local extrema: around these points the TVD-limiters have first order accuracy and the estimated slopes at the extrema points are zero [191].

The  $L^1$  norm of the error as a function of the run-time for different schemes are illustrated in Fig. 6.25, evaluated for the interval  $0 \leq t \leq 1.8$ . The assumed parameters are the same as the Sod problem, except for:  $J_{max} = 12$ , and  $dt = 0.0000244$ . The plot-markers in each curve correspond from left to right, respectively, to the threshold

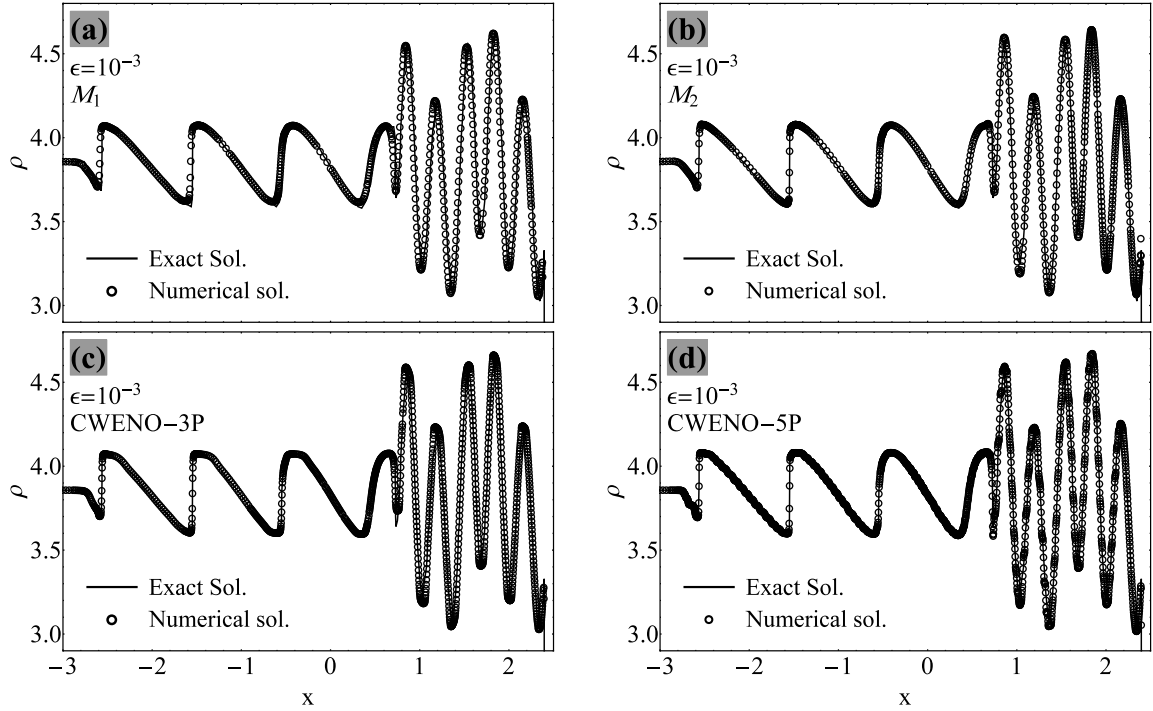


Figure 6.23: Numerical results from adaptive solvers for the entropy and shock wave interaction with  $M_1$ ,  $M_2$  and  $CWENO$  methods at  $t = 1.8$ .

values  $\epsilon \in \{5 \times 10^{-3}, 10^{-3}, 5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-5}, 10^{-5}, 10^{-6}\}$ . The solution of this problem includes several local extrema and due to the TVD limiters used in the PPMs, these schemes have more numerical dissipation (as explained before). This is clear in the error-runtime curves of the PPMs. The results show that the PPMs are the most expensive approach and again the  $CWENO - 5P$  is the most accurate method.

### Interaction of two blast waves

One of the challenging problems in the gas dynamic problem is the two blast waves interaction [190]. For this problem, the Riemann initial condition is [174, 175, 190]:

$$\{\rho, u, P\}|_{t=0} = \begin{cases} \{1, 0, 1000\}, & \text{if } x < 0.1, \\ \{1, 0, 0.01\}, & \text{if } 0.1 \leq x < 0.9, \\ \{1, 0, 100\}, & \text{if } x \geq 0.9. \end{cases} \quad (6.76)$$

The boundary conditions are reflecting walls, where the velocity  $u$  is an odd function of distance from the wall,  $\rho$  and  $P$  are even functions with respect to the wall in the spatial domain [174, 175].

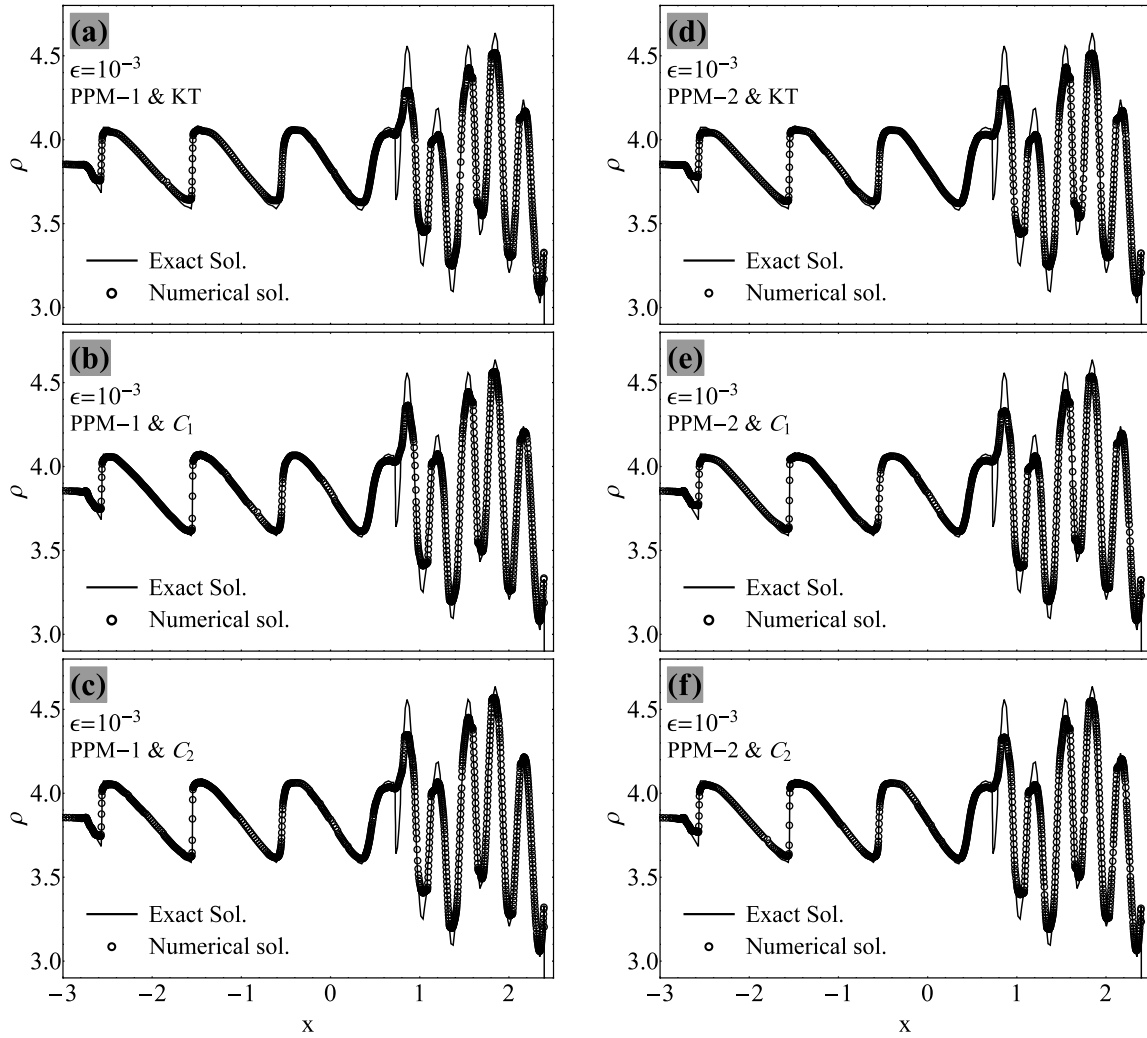


Figure 6.24: Numerical results from adaptive solvers for the entropy and shock wave interaction with  $PPM - 1$  and  $PPM - 2$  methods at  $t = 1.8$ ; the  $KT$ ,  $C_1$  and  $C_2$  second order central high resolution schemes with the GMM limiter (with  $\theta = 2$ ) are used for the evaluations of velocities and accelerations.

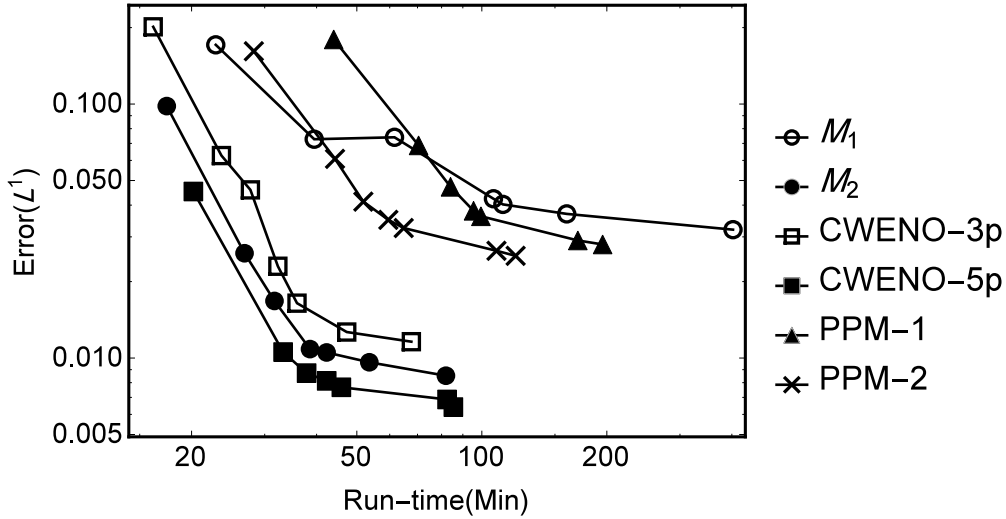


Figure 6.25:  $L^1$  error as a function of run-time for different schemes for the Shu-Osher problem, evaluated for  $0 \leq t \leq 1.8$ .

The numerical simulations are performed with the parameters  $\epsilon = 10^{-2}$ ,  $dt = 0.2 \times 10^{-5}$  and  $\theta = 2$ . The results are illustrated in Fig. 6.26; in each row, solutions are at times 0.016, 0.028 and 0.032, respectively, from left to right. In each illustration, the exact (reference) and numerical solutions are presented by solid lines and hollow-circle markers, respectively. It is clear that points are properly adapted in the vicinity of both high-gradient and discontinuous zones.

### 6.6.3 The 2-D scalar Burgers' equation

The 2-D scalar Burgers' equation is assumed as follows:

$$u_t + (u^2)_x + (u^2)_y = 0, \quad \Omega \in [-2, 2] \times [-2, 2], \quad (6.77)$$

where  $u := u(x, y, t)$  and fluxes in the x and y directions are  $u^2$ .

The IC,  $u(x, y, t = 0)$ , is equal to -1 and 1 inside two circles with center locations  $(0.5, 0.5)$  and  $(-0.5, -0.5)$ , respectively; radius of both circles are 0.4. For both PPM-1 and PPM-2 schemes, ICs of  $v_x := \partial u / \partial x$  and  $v_y := \partial u / \partial y$  can be obtained by the concept of the Dirac delta function, Eq. (6.70) and corresponding approximation in Eq. (6.70).

In numerical calculations, it is assumed that:  $\epsilon = 0.25 \times 10^{-3}$  (threshold),  $dt = 0.0002$ ,  $\theta = 2$ ,  $J_{max} = 8$ ,  $J_{min} = 5$ ,  $N_c = 2$ ,  $N_s = 1$  and  $\theta = 2$ .



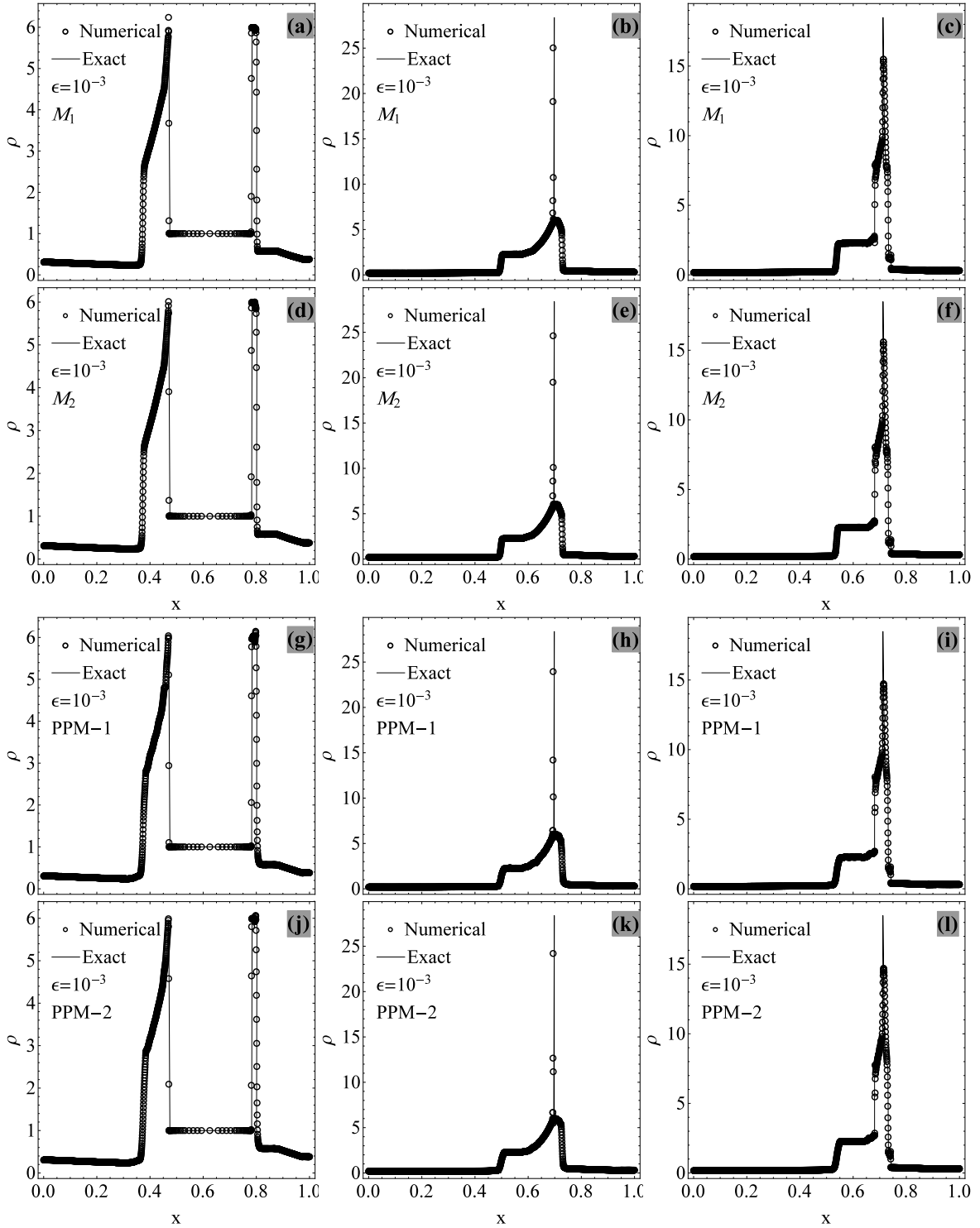


Figure 6.26: Numerical results from adaptive different solvers for the interaction of two shock waves; in each row, solutions are for times 0.016, 0.028 and 0.032 respectively from left to right.

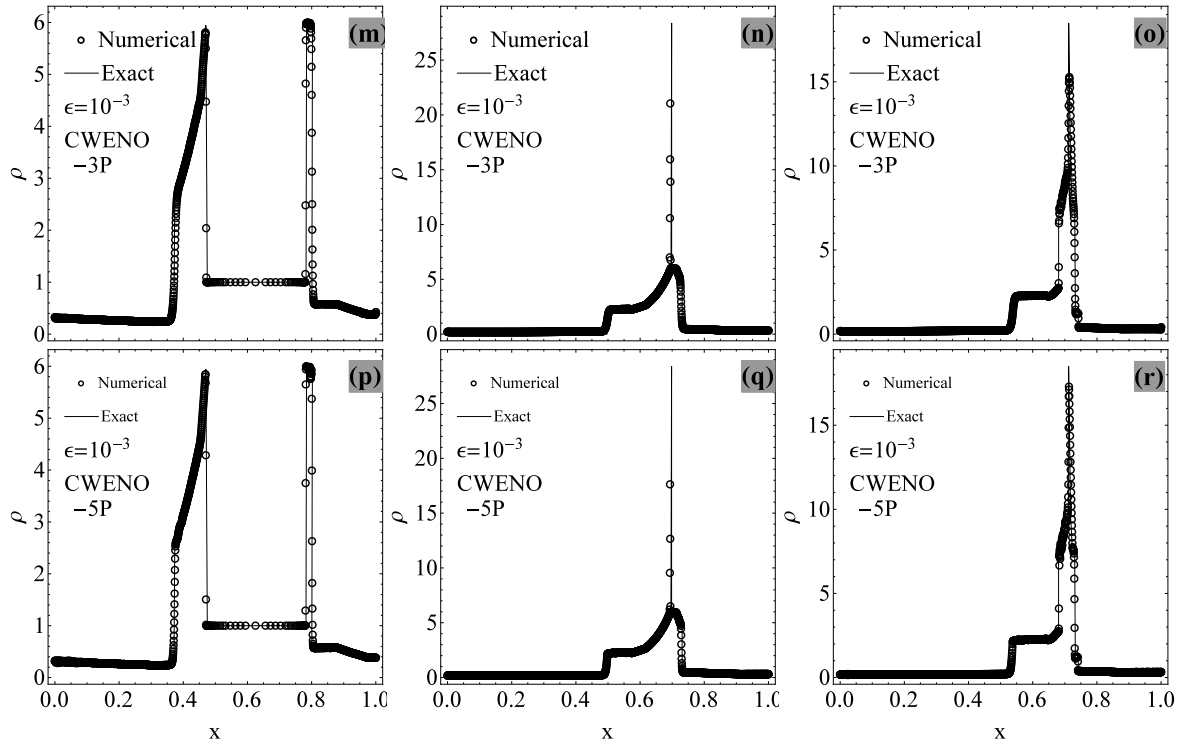


Figure 6.26 (cont.): Numerical results from adaptive different solvers for the interaction of two shock waves; in each row, solutions are for times 0.016, 0.028 and 0.032 respectively from left to right.

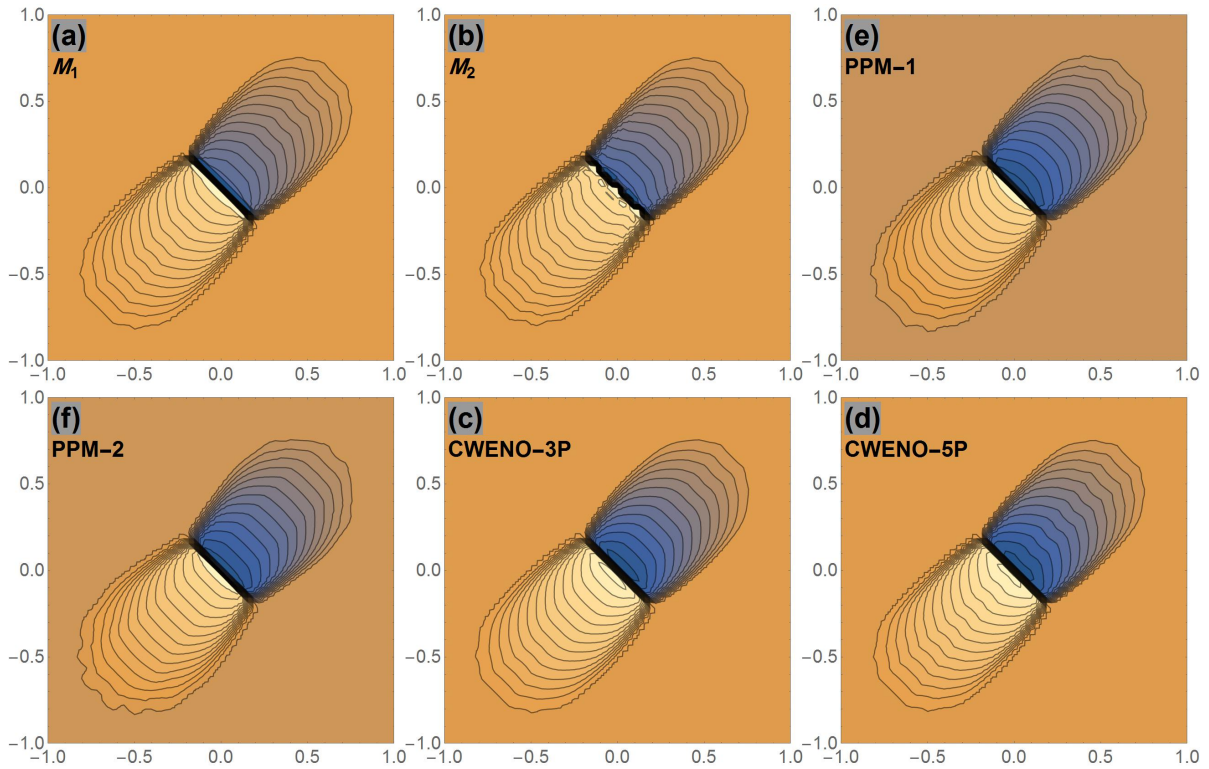


Figure 6.27: Numerical results for the 2-D Burgers' problem with different schemes for  $\Omega \in [-1, 1] \times [-1, 1]$  at  $t = 0.5$ .

For different schemes, contour-plot of numerical results are presented in Fig. 6.27 at  $t = 0.5$  for  $\Omega \in [-1, 1] \times [-1, 1]$ . Their 3-D representations are also illustrated in Fig. 6.28. Adapted grid points are shown at times 0.02, 0.2 and 0.5 in Fig. 6.29 for  $\Omega \in [-2, 2] \times [-2, 2]$ .

Based on results, it is clear that:

1. All schemes can capture the discontinuity,
2. Localized spurious oscillations develop in  $M_1$ ,  $M_2$ , PPM-1 and PPM-2 schemes,
3. PPMs can reduce such oscillations,
4. Central-WENO schemes can effectively control the localized spurious oscillations.

#### 6.6.4 2-D Euler systems

The governing equations for the 2-D Euler system are [192]:

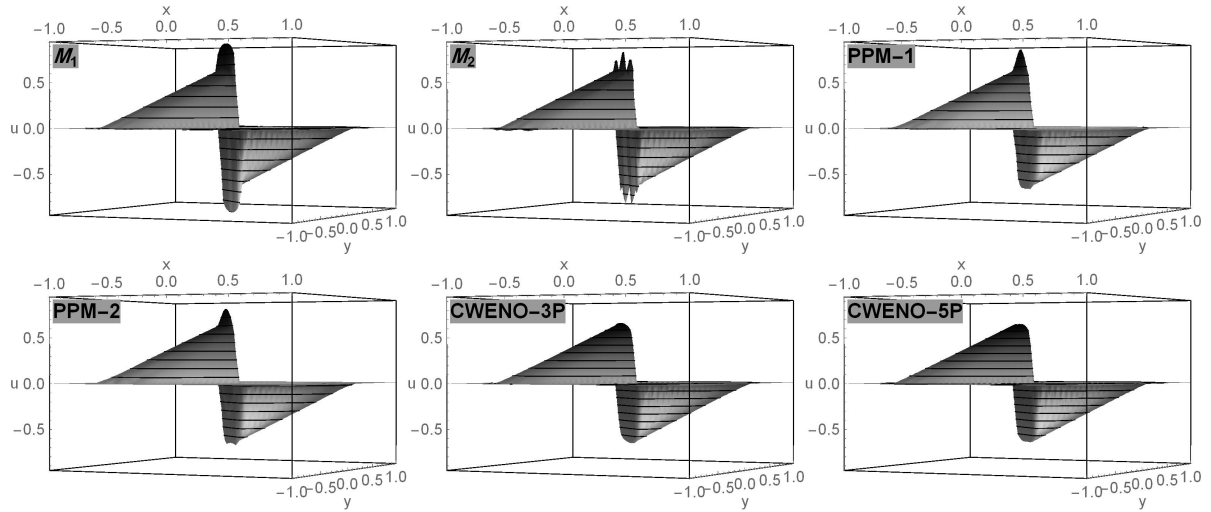


Figure 6.28: The 3-D representations of numerical results for the 2-D Burgers' problem with different schemes for  $\Omega \in [-1, 1] \times [-1, 1]$  at  $t = 0.5$ .

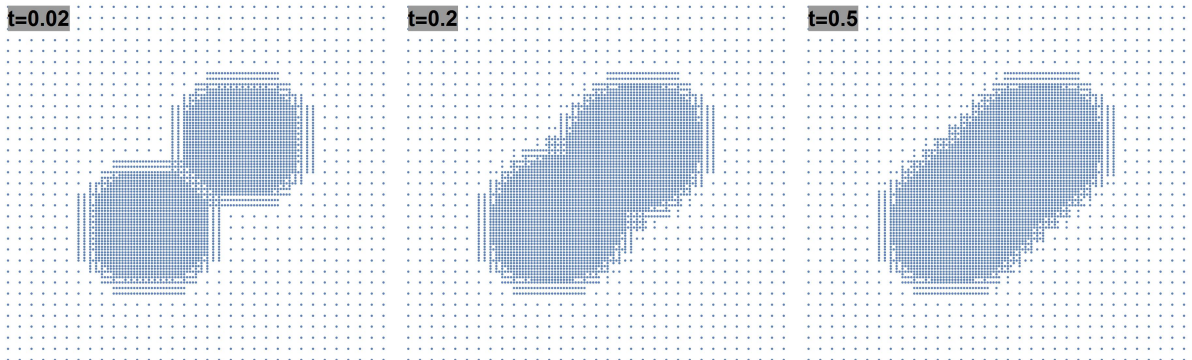


Figure 6.29: Adapted grids for the 2-D Burgers' problem at different times for the  $M_1$  scheme where  $\Omega \in [-2, 2] \times [-2, 2]$ . For adaptation, we have:  $J_{max} = 8$ ,  $J_{min} = 5$  and  $\epsilon = 0.25 \times 10^{-3}$ .

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{0}, \quad \Omega = [0, 4] \times [0, 1], \quad (6.78)$$

where

$$\begin{aligned} \mathbf{U} &= [\rho, \rho u, \rho v, E]^T, \\ \mathbf{F} &= [\rho u, \rho u^2 + p, \rho uv, u(E + p)]^T, \\ \mathbf{G} &= [\rho v, \rho uv + p, \rho v^2 + p, v(E + p)]^T. \end{aligned} \quad (6.79)$$

Here,  $u$  and  $v$  denote the velocity components in the  $x$  and  $y$  directions, respectively;  $\rho$  is the density;  $E$  denotes the total energy; and  $p$  is the pressure. Parameter  $E$  is defined as:

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho(u^2 + v^2), \quad (6.80)$$

with the adiabatic index  $\gamma = 1.4$ .

### Double Mach reflection of a strong shock

This problem was originally introduced by Woodward and Colella [193] as a benchmark for simulation of shock waves in the form of 2-D Euler equations. For the initial conditions, a planar shock is considered having a  $60^\circ$  angle with respect to the  $x$ -axis. The pre- and post-shock conditions are [192, 193]:

$$\begin{cases} \rho = 1.4, & u = 0, & v = 0, & p = 1, & \text{if } x > \frac{1}{6} + \frac{y}{\tan\frac{\pi}{3}}, \\ \rho = 8.0, & u = 8.25\sin\frac{\pi}{3}, & v = -8.25\cos\frac{\pi}{3}, & p = 116.5, & \text{otherwise.} \end{cases} \quad (6.81)$$

The boundary conditions of the domain are: the left and right boundaries are continuous; the lower boundary is wall (reflecting) for  $1/6 \leq x \leq 4$  and has the post-shock conditions for  $0 \leq x < 1/6$ ; the upper boundary has a time-dependent feature, describing the incoming inclined planar shock wave. The challenging problem is the proper imposing of the upper boundary condition as the grid adaptation method is nearly sensitive. In this study, instead of the Dirichlet BCs at the upper boundary, an oblique

first-order extrapolation is used [194]. As the angle of the planar shock to the horizontal axis is  $60^\circ$ , for uniform cells ( $\Delta x = \Delta y$ ), and one ghost cell, this condition becomes [194, 195]:

$$\mathbf{U}_{i,m+1} = \mathbf{U}_{i-1,m-1}, \quad (6.82)$$

where  $m$  is the number of grid cells in the  $y$ -direction.

The assumed parameters in numerical simulations are:  $\epsilon = 10^{-3}$  (the wavelet threshold);  $J_{max} = 6$ ;  $J_{min} = 3$ ;  $dt = 2 \times 10^{-6}$ . For PPM methods extra simulations are performed with  $J_{max} = 7$  and  $J_{min} = 3$  to consider the effect of finer resolution effects.

Numerical results for the  $M_1$ ,  $M_2$ ,  $CWENO - 3P$  and  $CWENO - 5P$  schemes are presented in Fig. 6.30 at  $t = 0.2$  sec.; thirty contours are used for presenting the contour plots. The corresponding MRA-based adapted grids are illustrated in Fig 6.31.

Contour plots of numerical results (with 30 contours) for the  $PPM - 1$  and  $PPM - 2$  methods and their adapted grids are presented in Figs. 6.32 and 6.33, respectively. These results are for two cases:  $J_{max} = 6$  and  $J_{max} = 7$ . From Fig. 6.32, it is obvious that the solution is dissipative; however, by adding a finer resolution level, the performance is improved. The main reason is that the TVD limiters have at most *first-order accuracy* in the 2D/3D problems and this leads to dissipative results.

In general, the results offer that:

1. As the central schemes are TVB, each method leads to different adapted grids (due to presence of some localized spurious oscillations; and this leads to different contour plots),
2. The PPM schemes lead to dissipative responses due to performance of TVD-limiters on multidimensional problems: having at most first-order accuracy. Increasing the resolution levels can improve the performance of the PPM schemes. Also it is possible to use other type of limiters (e.g., TVB); for a review of different limiters, see end of the Introduction section.

### **A Mach three wind tunnel with a step**

This problem was introduced in [193] to investigate the performance of different methods in handling singular solutions developed around the edge of a step [192]. The Euler equations (6.78) are solved in a wind tunnel with a unit length wide and 3 length units long which it includes a step with 0.2 units high, started from 0.6 length units from the

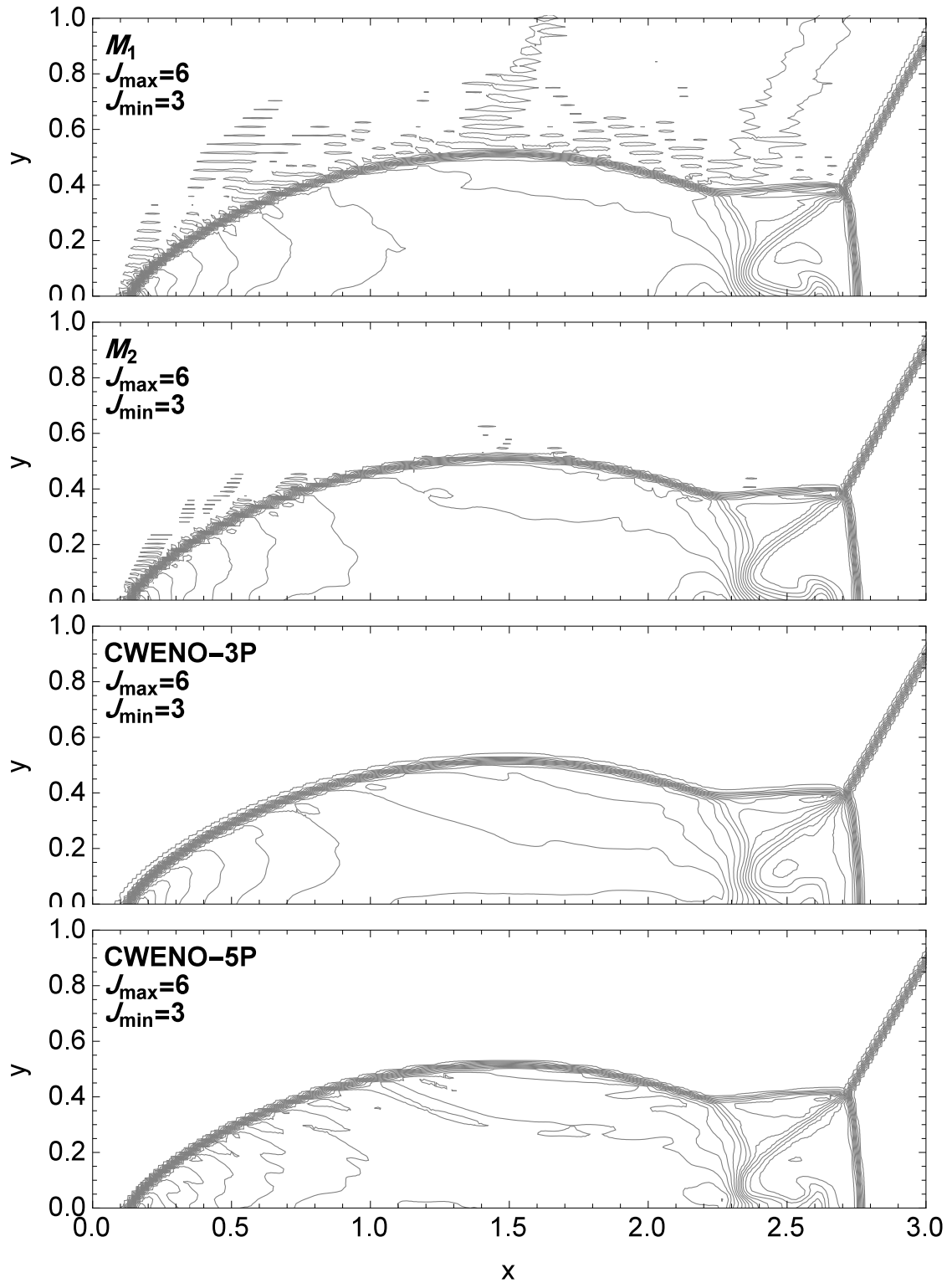


Figure 6.30: Numerical results obtained by the  $M_1$ ,  $M_2$ ,  $CWENO-3P$  and  $CWENO-5P$  schemes at  $t = 0.2$ ; the number of contours is 30.

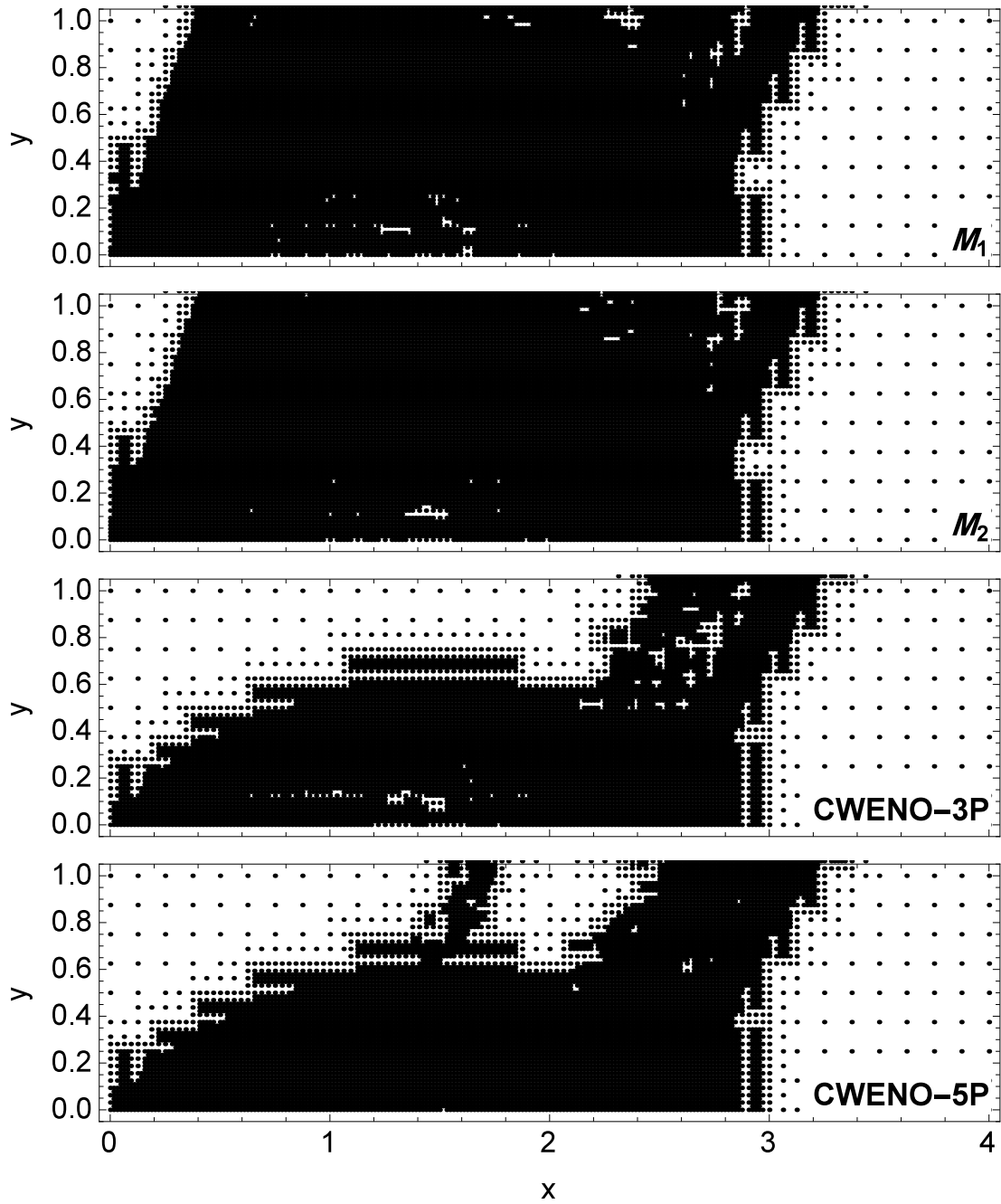


Figure 6.31: Adapted grids for the numerical simulations of the  $M_1$ ,  $M_2$ , CWENO-3P and CWENO-5P schemes at  $t = 0.2$ .



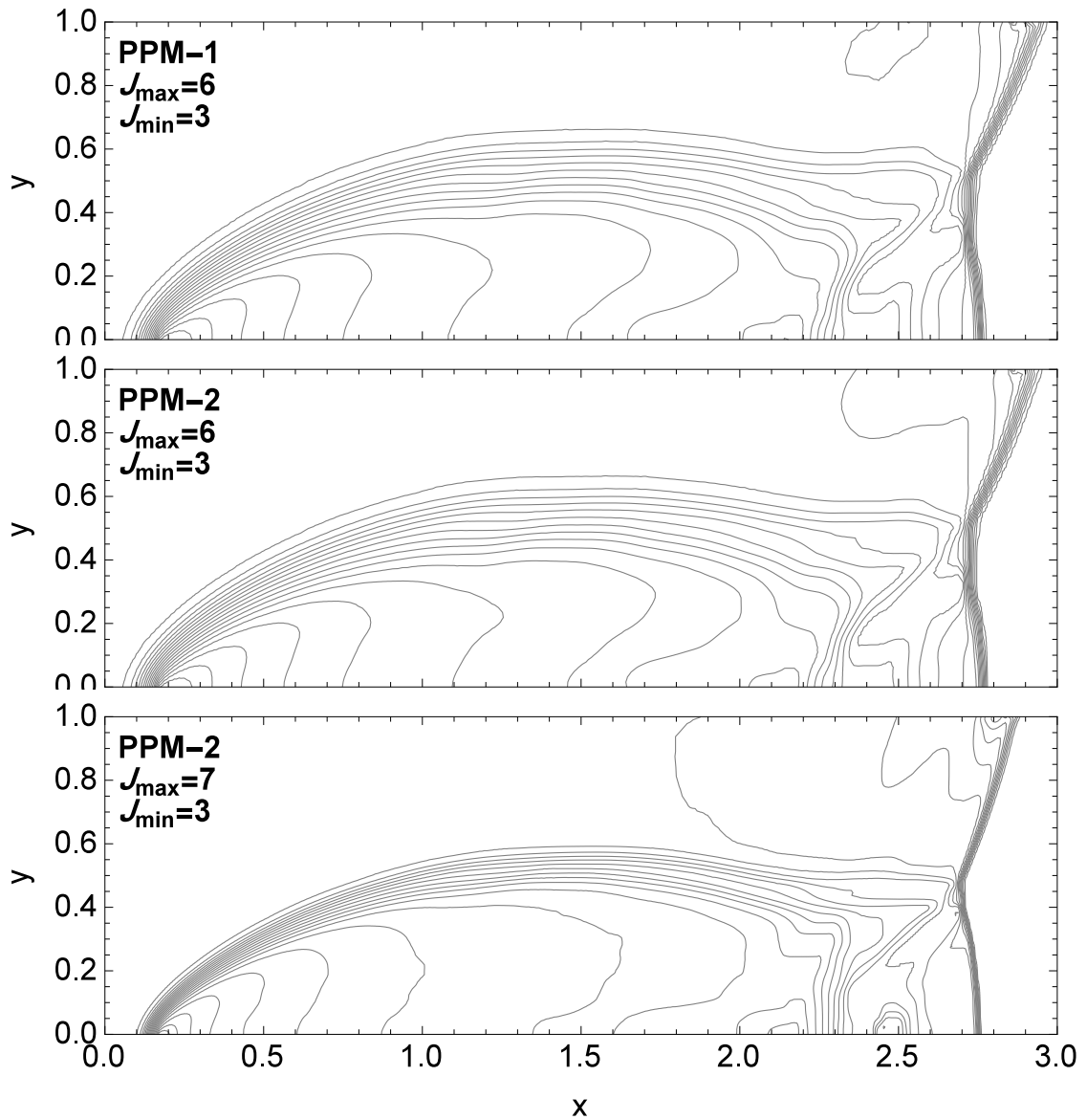


Figure 6.32: Numerical results obtained by the *PPM* – 1 and *PPM* – 2 schemes at  $t = 0.2$ ; the number of contours is 30.

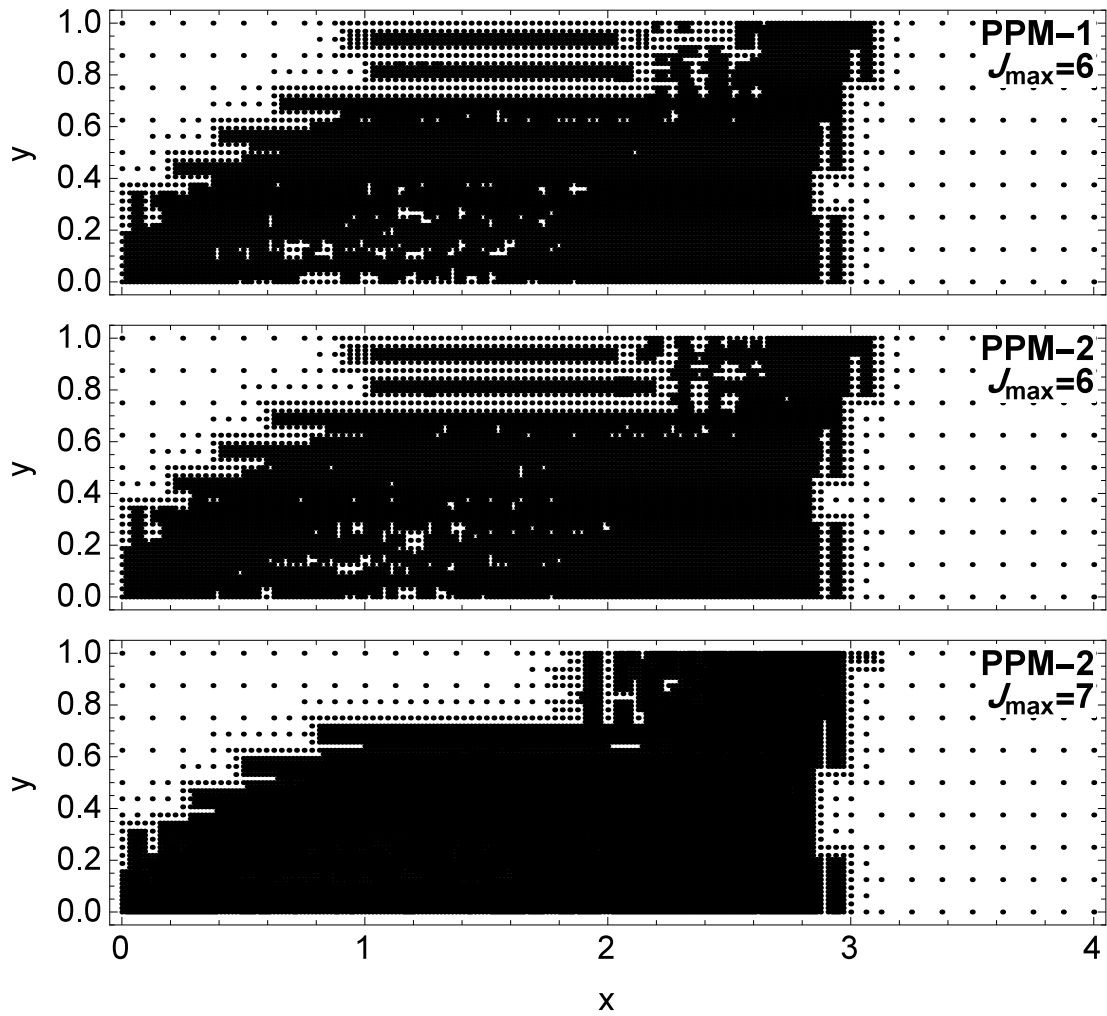


Figure 6.33: Adapted grids for the numerical simulations of the  $PPM-1$  and  $PPM-2$  schemes at  $t = 0.2$ .

left boundary. The left and the right boundaries are inflow (entrance) and outflow (exit) boundaries, respectively. The remaining boundaries at the top and bottom are wall.

Initially, the wind tunnel is filled everywhere with a inviscid gas (where  $\gamma = 1.4$ ) with density  $\rho = 1.4$ , pressure  $p = 1.0$ , horizontal velocity  $u = 3$  and vertical velocity  $v = 0$ . These conditions mean that the gas continuously enters from the inflow (left) boundary.

In numerical simulations, the parameters are:  $\epsilon = 10^{-3}$  (the wavelet threshold),  $J_{max} = 6$ ,  $J_{min} = 3$ ,  $dt = 2 \times 10^{-5}$  and  $\theta = 2$  (for the GMM).

For the  $M_1$ ,  $M_2$ ,  $CWENO - 3P$  and  $CWENO - 5P$  schemes, contour plots with 30 contours are illustrated in Fig. 6.34 at  $t = 3$  sec. The corresponding MRA-based adapted grids are presented in Fig 6.35.

Development of adaptive solutions in time and the corresponding adapted grids are presented in Fig. 6.36 (by the  $CWENO - 3P$  scheme).

Results from PPMs are presented in Fig. 6.37; it is obvious that the solutions are dissipative (due to the behavior of TVD limiters for multidimensional problems) and solution accuracy is in accordance to number of resolution levels.

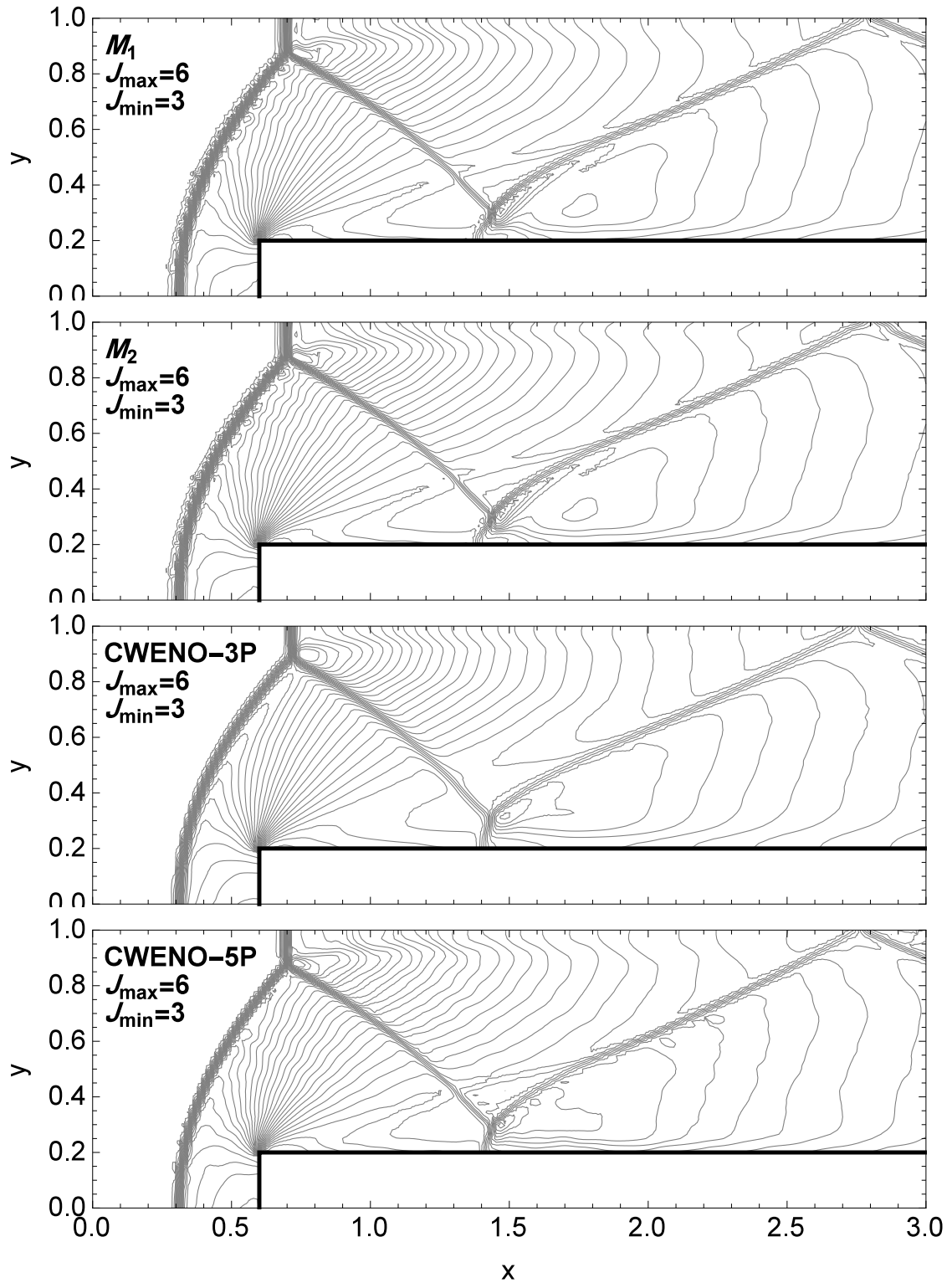


Figure 6.34: Numerical results obtained by the  $M_1$ ,  $M_2$ ,  $CWENO-3P$  and  $CWENO-5P$  schemes at  $t = 3$ ; the number of contours is 30.

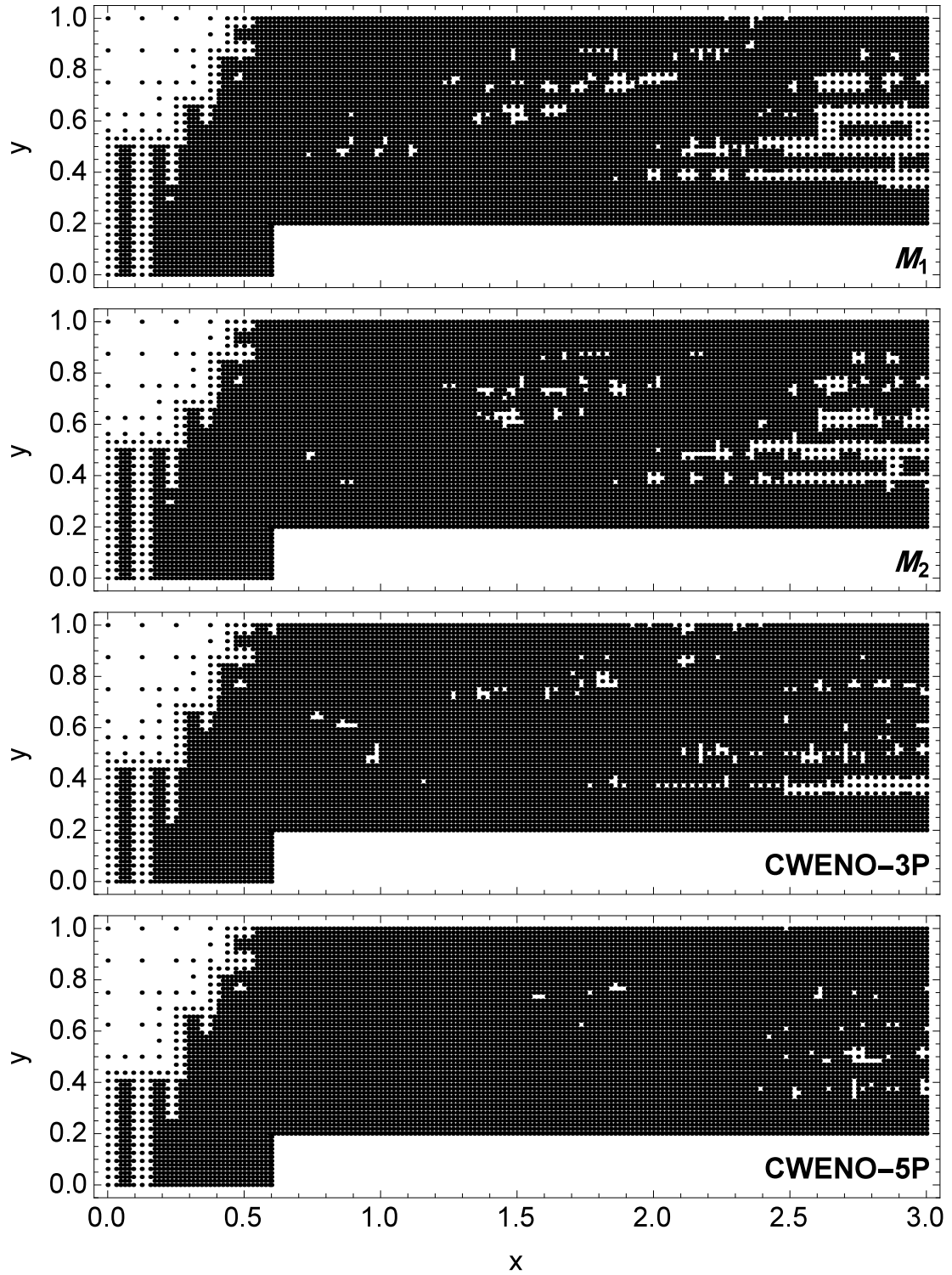


Figure 6.35: Adapted grids for the numerical simulations of the  $M_1$ ,  $M_2$ ,  $CWENO - 3P$  and  $CWENO - 5P$  schemes at  $t = 3$ .

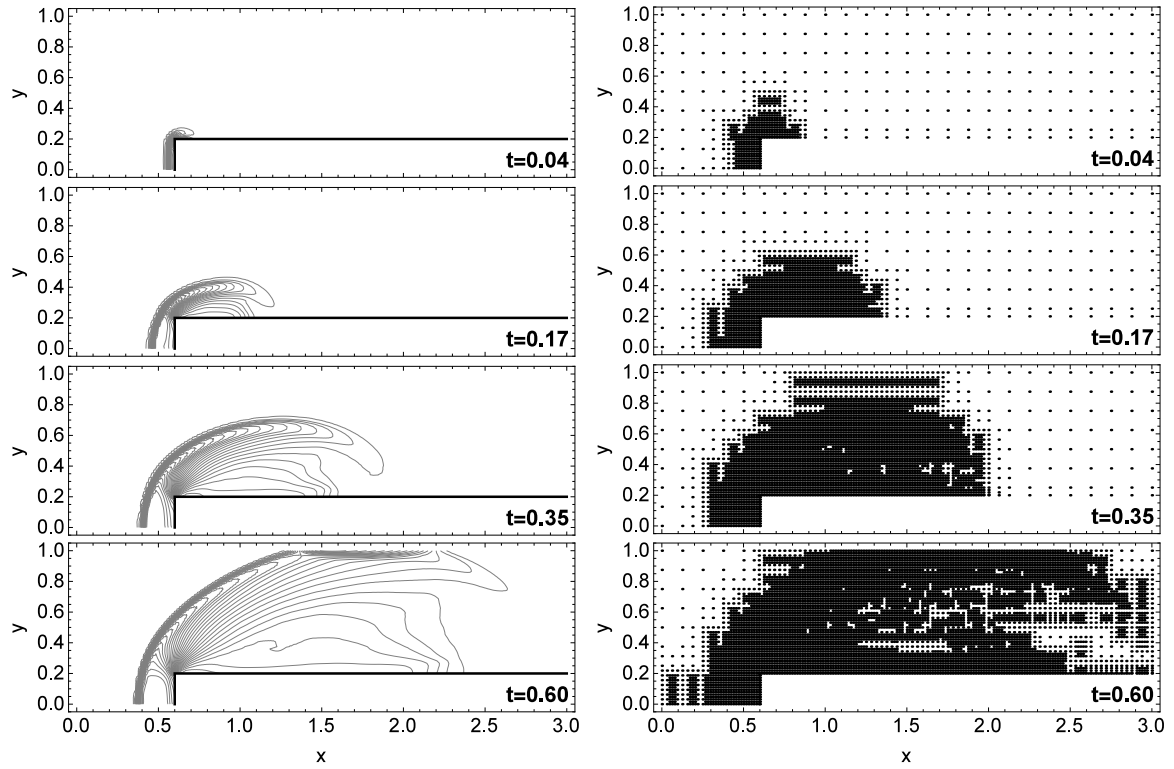


Figure 6.36: Adapted solutions (with 30 contours) and their grids at times 0.04, 0.17, 0.35 and 0.60 obtained with the *CWENO* –  $3P$  scheme.

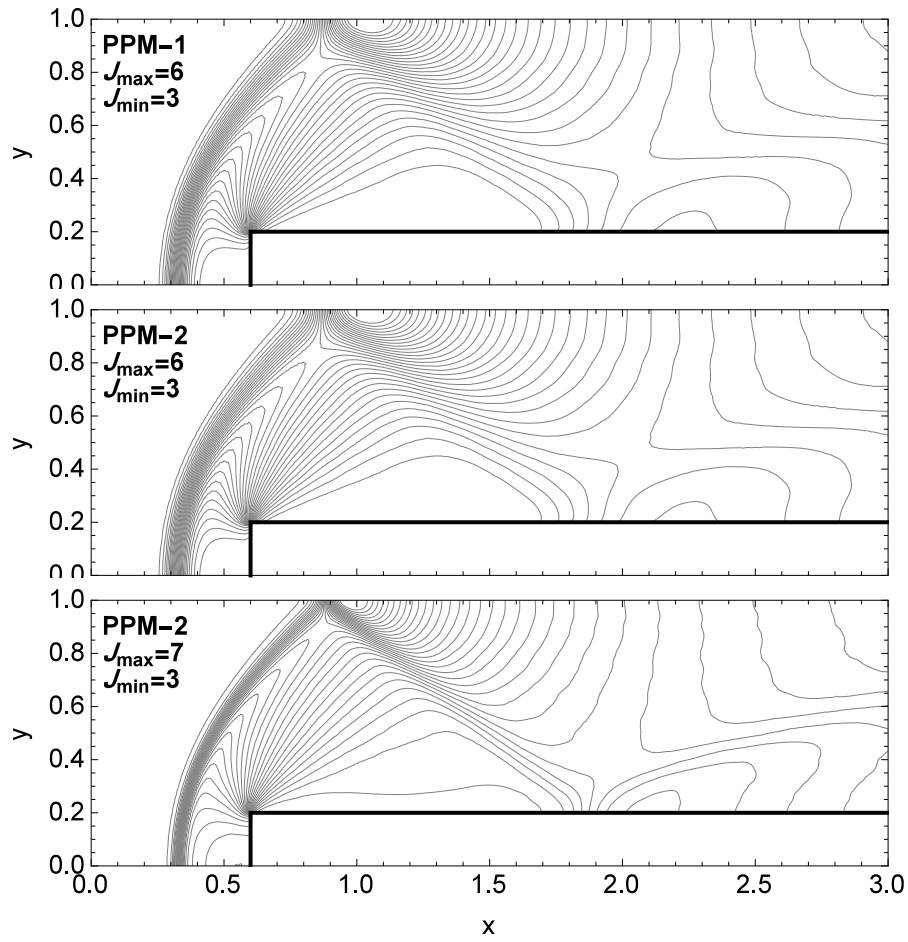


Figure 6.37: Adapted solutions obtained with the PPM schemes with different resolution levels at  $t = 1.8$  with 30 contours.

# Chapter 7

## Discontinuous solutions of second-order wave equations and the regularization concept

**Introduction** In this chapter, numerical simulation of second-order hyperbolic PDEs (mechanical waves) is studied for cases having discontinuous solutions or propagating fronts. The concept of regularization is used for removing spurious oscillations from numerical solutions as a post-processing stage. There are different regularization approaches with different modified/improved versions; they are designed for different problems, hence proper utilization of them for numerical simulations is essential. Regarding numerical method for numerical simulations, here, the compact finite difference method and optimized ones are used. These methods were designed to capture high-frequency waves (those are near earthquake sources), where common FD and FE methods can not handle properly such waves.

This chapter is organized as follows. Section 7.1 describes the main concept of regularization. Regularization-based numerical solutions of 1-D stress wave propagation problems containing discontinuities are studied in Section 7.2. In Section 7.3, the explicit/implicit higher order finite difference methods, developed for high-frequency waves, are explained with corresponding filtering algorithms. In this section, the Tikhonov-based smoothing will be integrated with such higher-order methods. The performance of this collaboration will be studied by both smooth and discontinuous solutions. In Section 7.4, by using the singular value decomposition (SVD) and the generalized one (GSVD), effects of different constraints for the Tikhonov method are studied. The implementation of the Tikhonov method with different constraints are presented in Section 7.5. In Section 7.6, error estimations and convergence rates are investigated. These



concepts are then studied by a benchmark having a discontinuity; there, both of the numerical dispersion and Runge phenomenon are investigated. Section 7.7 devotes to the conservative filtering feature in the Tikhonov method. In Section 7.8, it will be shown how to have a local smoothing with Tikhonov methods by considering extra constraints. Relationship between Tikhonov-based smoothing methods and classical filtering concepts in signal processing schemes will be explained in Section 7.9. Section 7.10 discusses general algorithms for the numerical simulation of second-order hyperbolic systems. Some examples in finite and infinite-periodic domains with fixed and absorbing boundary conditions are tested in Section 7.11; here, effects of the numerical dispersion and artificial dissipation are studied.

## 7.1 Regularization approaches

### Thikhonov regularization method with different constraints

The aim of regularization is to replace an ill-posed problem with a nearly well posed one. Let assume uniformly sampled (noisy) data  $\{\{y_i\}; i \in \{1, 2, \dots, n\}\}$ ; to have a stable estimation of smoothed (denoised) data  $\{\{f_j\}; j \in \{1, \dots, n\}\}$  from  $\{y_i\}$ , a functional  $Q$  resulted from a linear combination of the residual norm  $\rho(\mathbf{y} - \mathbf{f}) = \|\mathbf{y} - \mathbf{f}\|_2 = \sqrt{\sum_{i=1}^n (y_i - f_i)^2}$  and prior information  $\Omega(\mathbf{f})$  is considered, as [145, 146]:

$$Q(\mathbf{f}) = \rho(\mathbf{y} - \mathbf{f})^2 + \lambda \times \Omega(\mathbf{f})^2, \quad (7.1)$$

where  $\lambda$  ( $0 \leq \lambda < \infty$ ) is a penalization factor. The solution  $\{f_j\}$  minimize this functional. Parameter  $\lambda$  controls trade-off between error in estimations and smoothness. The cases  $\lambda \rightarrow 0$  and  $\lambda \rightarrow \infty$  lead to the linear fit and interpolation problems, respectively. Taking advantage of the remapping relation  $\lambda = (1-p)/p$ , the range of  $\lambda$  can be changed to  $[0, 1]$ ;  $p = 1$  results in an interpolation problem, and lower values yield more smoother estimations. It should be mentioned that the commonly used Tikhonov formulation can slightly be modified, for example as:  $Q(\mathbf{f}) = \rho(\mathbf{y} - \mathbf{f})^2 + (\int f(x)dx)^2 + \lambda \times (\|\mathbf{f}\|^2 + \|\mathbf{f}'\|^2)$  (where  $\mathbf{f}' := df/dx$ ) [196].

The stationary point of  $Q(\mathbf{f})$  is the minimizing solution:  $\mathbf{f}_{min} = \underset{\mathbf{f}}{\operatorname{argmin}}(Q(\mathbf{f}))$ . Some constraints  $\Omega^2(\mathbf{f})$  with corresponding features are briefly reviewed in Table 7.1.1. There, constraints with and without models (prior extra information) are presented.

The model-based Tikhonov regularization was introduced by Barakat et al. [197].

It introduces extra information  $f_{model}$  (containing generally localized information) to improve regularization. The modified constraint (measured usually by a (semi) norm) is:  $\Omega_{model}^2 = \|D^{(m)}(f - f_{model})\|_2^2$  (where  $D^{(m)}$  denotes an operator for  $m^{\text{th}}$  derivative). If the model  $f_{model}$  is properly chosen, this modification will lead to a good estimation. For example, if  $f_{model}$  has some discontinuities, this local information of signal can be used in regularization procedures. Even though, the Tikhonov method leads to unique and stable solutions, prior general information like bounded smoothness is not often enough to obtain a suitable solution due to a global character. It will be shown that model-based regularized solutions are also sensitive to the model, itself. This sensitivity is due to existence of high-frequency components in systems. This will be confirmed by the GSVD in Sec. 7.4.

The Tikhonov regularization method leads to a closed form solution. However, the main disadvantage of the Tikhonov method is that it cannot properly handle discontinuities in a considered system.

All of the above mentioned constraints are defined in the  $L_2$  space; there exists some other effective constraints which measure prior information in  $L_1$  space (e.g., with (semi) norm  $\int (|f^{(m)}(x)| dx, m \geq 0$ ; where  $f^{(m)} := d^{(m)}f/dx^{(m)}$ ). They are developed to handle special problems, like the edge preserving restoration (preserving discontinuity). In this case, the functional  $Q$  can generally be defined as:  $Q(\mathbf{f}) = \rho(\mathbf{y} - \mathbf{f})^2 + \lambda \times \Omega(\mathbf{f})$ . Some of such constraint definitions are: 1) the Total Variations (TV) [157]; 2) the incompressibility; 3) local rigidity [198]. For these regularization methods, closed-form solutions can not be provided; they are mainly solved by iterative algorithms.

As mentioned, the main shortcoming of the Tikhonov method is proper handling of discontinuities. The TV-based regularization is originally developed to preserve image edges or data discontinuities. Corresponding constraint can be defined as:  $\Omega_{TV}(f) = \int_{-\infty}^{+\infty} |f^{(m)}(x)| dx, m \in \{0, 1\}$ . Regarding numerical simulations, however, the authors experience shows that this approach does not have enough smoothness; this makes it unsuitable for PDE solutions (this will be studied).

**Some remarks:** considering above mentioned methods, it should be noted that:

1. In the above-mentioned regularization methods, it is assumed that the noise has the white Gaussian feature. This means, the discrete values  $y_i := y(x_i)$  can be written as:  $y_i = f_i + \epsilon_i$ ; where  $i \in \{1, 2, \dots, n\}$ ;  $x_1 \leq x_2 \leq \dots \leq x_n$ ;  $\{\epsilon_i\}$  are random, uncorrelated errors with zero mean and variance  $\sigma^2$ ;  $f(x)$  is the denoised (smooth) function which should be estimated.

Constraint	Feature
$\Omega_0^2(x) = \int_{-\infty}^{+\infty} (f'(x))^2 dx,$	Using a smoothness constraint: having first order continuity [199],
$\Omega_1^2(x) = \int_{-\infty}^{+\infty} (f''(x))^2 dx,$	Using a smoothness constraint: having second order continuity and a physical meaning. It measures flexural energy in a beam [145, 146],
$\Omega_2^2(x) = \int_{-\infty}^{+\infty} (f''(x) + \alpha f'(x))^2 dx,$	Using a smoothness constraint: regularized solutions have second order continuity. This definition was recently proposed [200]. It leads to smoothing splines in tension,
$\Omega_3^2(x) = \int_{-\infty}^{+\infty} (f''(x)^2 + \alpha^2 f'(x)^2) dx,$	Using a smoothness constraint: regularized solutions have second order continuity. This leads to smoothing splines in tension. This definition has a physical meaning: it measures energies of flexural and axial deformations in a beam [201–204],
$\Omega_{model}^2(x) = \ D^{(m)}(f - f_{model})\ _2^2;$ $m \in \{0, 1, 2\},$	The model-based regularization by using extra prior information $f_{model}$ ; it can include several locally imposed features [197],
$\Omega_4^2(x) = \int_{-\infty}^{+\infty} ( f'(x) + 1  - 1)^2 dx,$	Area (or volume) preserving constraint [205].

 Table 7.1.1: Different constraint definitions in the  $L_2$  space for the Tikhonov method.

2. For other types of noise, different modified definitions should be used. For the Poisson noise see , e.g. [206], and for general cases, see [207, 208].
3. Regarding several regularization constraints, it has been noted that their effects can be viewed as a high-pass filter [209, 210]. This will be shown in Sec. 7.9.

### Curing discontinuity effects: the numerical dispersion around discontinuities

To handle properly discontinuities in regularization problems, several approaches are proposed; some of which are:

1. The common Tikhonov method (using common smoothness constraints, e.g.:  $\Omega_1^2$ ) with adaptive weight coefficients [211]. In this approach, The aim is to estimate a function with different smoothness in different spatial locations. However, discontinuity types and corresponding locations should be known, as prior information. This makes this approach, in general, unfeasible.

2. Using a smoothness constraint having a tension term (i.e.,  $\Omega_2^2$  and  $\Omega_3^2$ ) in the Tikhonov approach,
3. The TV-based regularization,
4. The model-based Tikhonov regularization.

The first, second and fourth approaches have closed-form solutions with fast algorithms; this makes such methods appealing for numerical simulation of boundary value problems. The third approach can effectively handle discontinuities. However, it suffers from lack of smoothness and utilizing of iterative solvers.

## 2D regularization methods

Extension of 1D regularization definitions to higher-dimension problems are straightforward: the same concepts and measurements can be used, as well. For example, for the Tikhonov method with a smoothness constraint (if the smoothness is measured by the gradient constraint, like  $\Omega_0^2$  in the 1D case), the corresponding functional can be written as:  $Q = \int_{\Omega} (z(x, y) - f)^2 d\Omega + \lambda \int_{\Omega} |\nabla f|_2^2 d\Omega$  (where  $|\nabla f|_2^2 := \{D_x(f)^2 + D_y(f)^2\}$ , where the operator  $D_i$  is the first derivative definition in the  $i^{\text{th}}$  direction and  $i \in \{x, y\}$ ).

In this work, to have a cost-effective algorithm, higher-order problems are solved based on 1D algorithms. For this purpose, firstly, in each direction, 1D algorithms are independently implemented, and then average of results are considered as a regularized solution. This means, for spatial point  $(x_i, y_j)$ , the approximated solution is:  $f(x_i, y_j) = 0.5(f_1(x_i) + f_2(y_j))$ , where  $f_1$  and  $f_2$  are the regularized solutions obtained independently in the  $x$  and  $y$  directions, respectively.

## 7.2 Numerical simulation of 1-D wave problems via regularization methods

In this section, 1D wave propagation problems including (semi) discontinuity will be studied to reveal effectiveness of different regularization approaches. The problem is longitudinal wave propagation in a bar with a constant cross section area. The governing equation is:  $\partial^2 u(x, t) / \partial t^2 = c^2 \partial^2 u(x, t) / \partial x^2 : 0 \leq x \leq 1$ ; the initial (ICs) and boundary (BCs) conditions are  $\{u(x, t = 0) = u_0 \ \& \ \dot{u}(x, t = 0) = v_0\}$  and  $\{u(0, t) = u(1, t) = 0\}$ , respectively (where  $\dot{u} := du/dt$ ). Here  $u$  denotes the axial deformation and  $c$  is the wave

propagation speed. The velocity  $c$  is assumed to be  $c = \sqrt{E/\rho} = 1$ ; where  $E$  and  $\rho$  are the module of elasticity and density of the bar, respectively.

For numerical simulations, in spatial domain, the higher-order finite difference method is used: the fourth-order explicit central difference approximation for the second derivative. The temporal integration is done with the Runge-Kutta 4<sup>th</sup> order method, see Appendix D. For spatio-temporal discretizations, it is assumed:  $dx = 1/2^9$  and  $dt = 10^{-8}$ . After each five time steps, numerical solutions are regularized (denoised) with different regularization approaches (as a post-processor).

Different imposed initial displacements will be considered; both smooth but high gradient deformations and those having discontinuities. The considered ICs for  $u_0$  are:

1.  $u^{(1)}(x, t = 0) = u_0^{(1)} = \sqrt{\max\left(1 - \left(\frac{x-x_c}{0.5 \times Supp}\right)^2, 0\right)}$ , where  $x_c$  and  $supp$  are center point and support length of the function, respectively. Here, it is assumed:  $x_c = 0.5$  and  $supp = 0.2$ ,
2. A dilated unit box with definition:  $u^{(2)}(x, t = 0) = u_0^{(2)} = \text{UnitBox}(4(x - 0.5))$ , where  $\text{UnitBox}(x) = H(x + 0.5) - H(x - 0.5)$ ; the function  $H(x)$  is the Heaviside function,
3. A compact sawtooth wave function:  $u^{(3)}(x, t = 0) = u_0^{(3)} = \text{UnitBox}\left(\frac{(x-x_c)}{supp}\right) \frac{(x-x_c+0.5)}{supp}$ , where  $x_c = 0.5$  and  $supp = 0.2$ .

In all cases, it is assumed  $v_0 = 0$ .

In the following, numerical results of the wave propagation problem will be studied; they are resulted from different regularization/constraint definitions.

### **Numerical results from the Tikhonov method with different constraints and without a model**

For regularization stage, penalizing (smoothing) parameter is assumed to be:  $p = 0.875$  and  $\alpha = 1$ . Numerical results, denoised by the Tikhonov method with constraints  $\Omega_1^2$ ,  $\Omega_2^2$  and  $\Omega_3^2$ , are shown in Fig. 7.1. The constraint  $\Omega_1^2$  is a commonly used smoothness-based regularization, and the  $\Omega_2^2$  and  $\Omega_3^2$  are those improved by considering a tension term (see Table 7.1.1). Adding this term helps reduce undesirable fluctuations from fitted curves. The constraint  $\Omega_3^2$  are widely used, and the  $\Omega_2^2$  is recently recommended (see Table 7.1.1). Based on the results, it is clear that, even though both of constraints  $\Omega_2^2$  and  $\Omega_3^2$  use the concept of tensioned curves, in all cases the constraint  $\Omega_2^2$  can effectively control the numerical dispersion.

### **The model-based Tikhonov regularization**

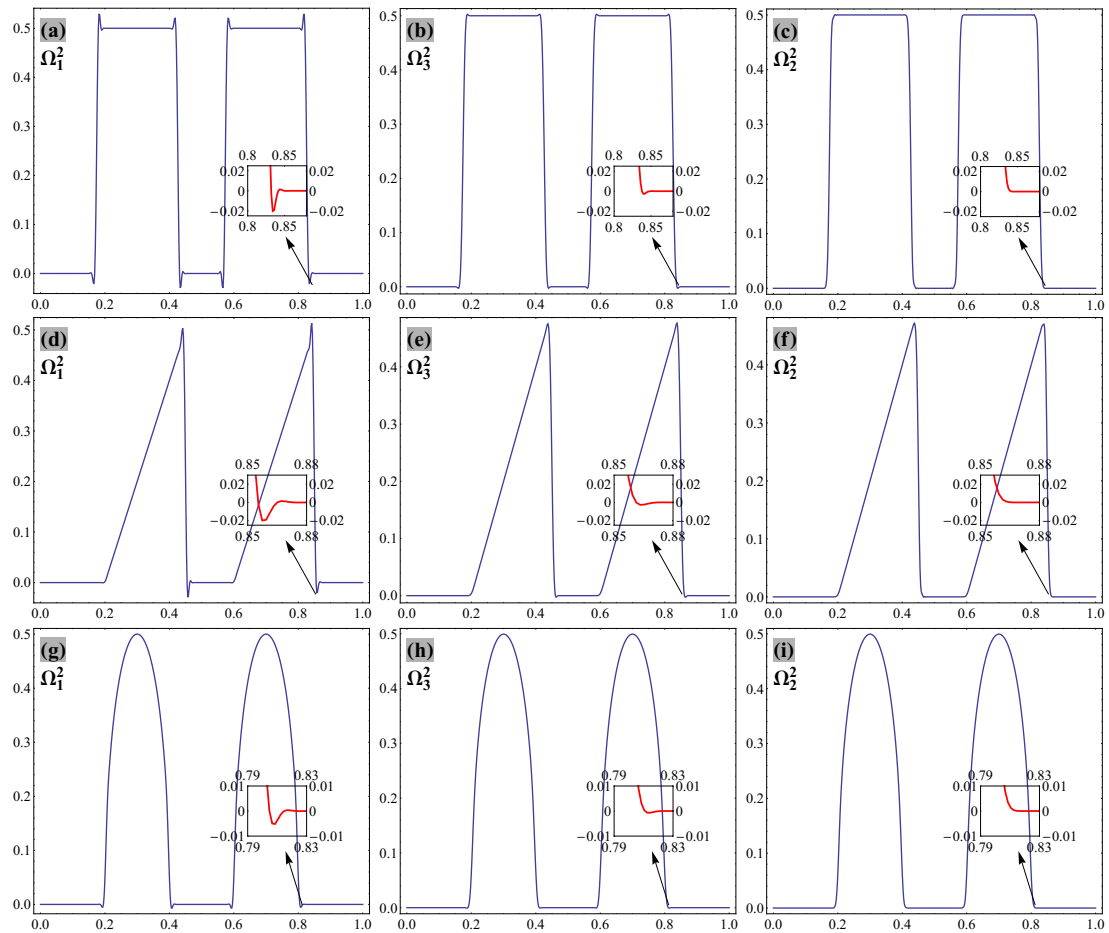


Figure 7.1: Numerical solutions with different Tikhonov methods with parameters  $p = 0.875$  and  $\alpha = 1$  at  $t = 0.2$ . The numerical dispersion appears in solutions obtained by the constraints  $\Omega_1^2$  and  $\Omega_3^2$ .

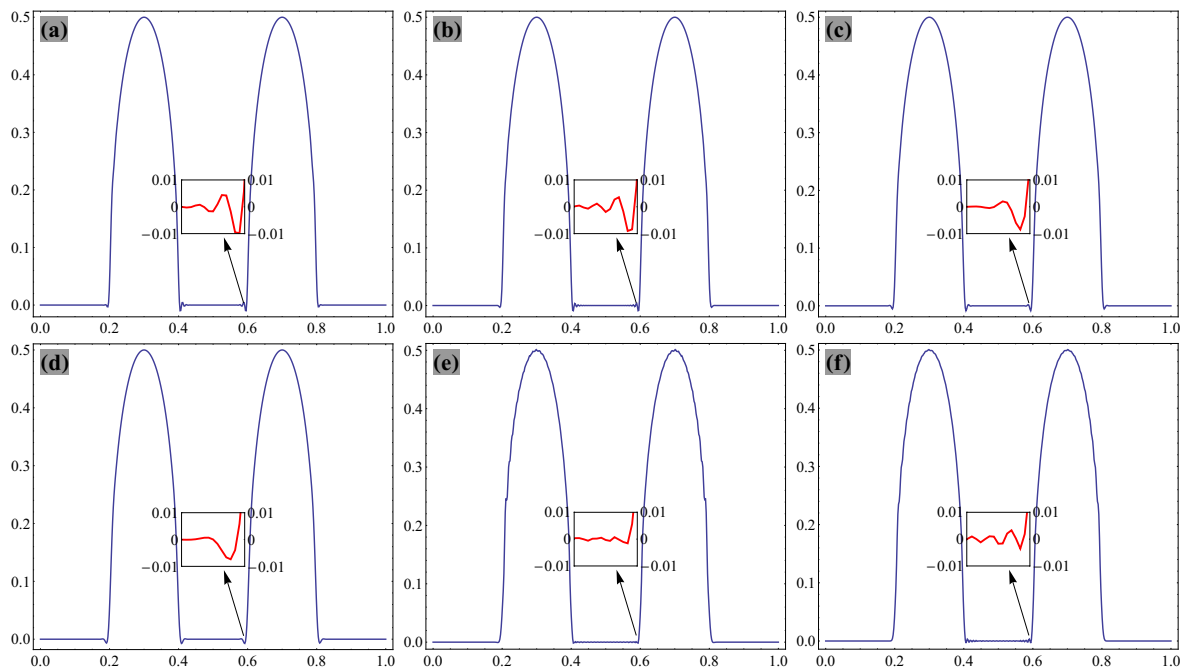


Figure 7.2: Numerical solutions of the wave propagation problem, denoised with the model-based Tikhonov scheme with constraints  $\Omega_1^2$ ,  $\Omega_2^2$  and  $\Omega_3^2$  at  $t = 0.2$ ; for each figure, for pair  $\{\mathbf{f}_{model}, \mathbf{f}_{min}\}$ , corresponding the constraint pairs are: a)  $\{\Omega_1^2, \Omega_1^2\}$ ; b)  $\{\Omega_2^2, \Omega_2^2\}$ , c)  $\{\Omega_3^2, \Omega_3^2\}$ , d)  $\{\Omega_{TV}, \Omega_1^2\}$ , e)  $\{\Omega_{TV}, \Omega_2^2\}$ , f)  $\{\Omega_{TV}, \Omega_3^2\}$ .

In the following, the effects of the model-based Tikhonov regularization approach will be studied by some numerical simulations. The considered constraints for obtaining  $\mathbf{f}_{min}$  are:  $\Omega_1^2$ ,  $\Omega_2^2$  and  $\Omega_3^2$ . To estimate  $\mathbf{f}_{model}$ , a regularization problem is firstly solved at each denoising stage; for this reason, the Tikhonov method with the constraints  $\Omega_1^2$ ,  $\Omega_2^2$ , and  $\Omega_3^2$  are used. The  $\mathbf{f}_{model}$  is also estimated with the TV regularization approach. Regarding the wave propagation problem, the results are presented in Fig. 7.2 with IC  $u^{(1)}(x, t = 0)$  at  $t = 0.2$  (simulation of smooth but high-gradient solutions). There, in figures (a) to (f), for  $\mathbf{f}_{model}$  and  $\mathbf{f}_{min}$ , the following constraints are considered: {figure (a): for  $f_{model}$ :  $\Omega_1^2$  & for  $f_{min}$ :  $\Omega_1^2$ }; {figure (b): for  $f_{model}$ :  $\Omega_2^2$  & for  $f_{min}$ :  $\Omega_2^2$ }; {figure (c): for  $f_{model}$ :  $\Omega_3^2$  & for  $f_{min}$ :  $\Omega_3^2$ }; {figure (d): for  $f_{model}$ :  $\Omega_{TV}$  & for  $f_{min}$ :  $\Omega_1^2$ }; {figure (e): for  $f_{model}$ :  $\Omega_{TV}$  & for  $f_{min}$ :  $\Omega_2^2$ }; {figure (f): for  $f_{model}$ :  $\Omega_{TV}$  & for  $f_{min}$ :  $\Omega_3^2$ }.

It is clear that in all cases spurious oscillations exist, and the model-based regularization can not effectively control the numerical dispersion. This is due to amplification of high-frequency components in the system (this will be shown). This shortcoming is more highlighted in discontinuous cases. To study this, the wave propagation problem

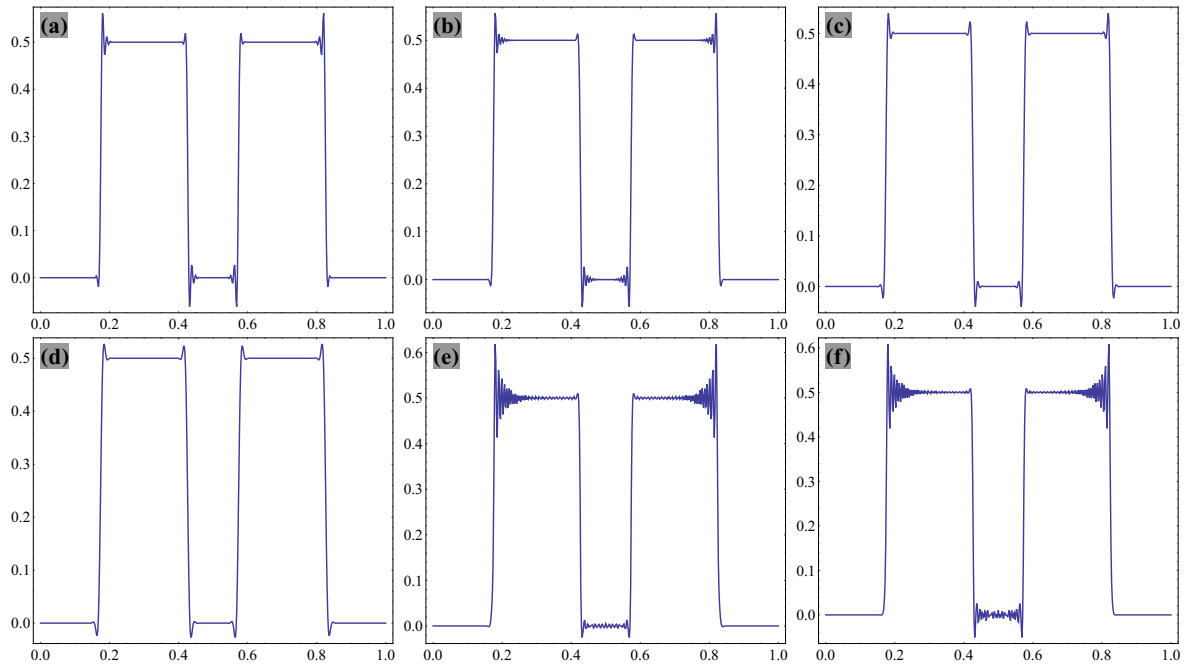


Figure 7.3: Numerical solutions of the wave propagation problem having discontinuous fronts, denoised with the model-based Tikhonov scheme with constraints  $\Omega_1^2$ ,  $\Omega_2^2$  and  $\Omega_3^2$  at  $t = 0.2$ ; for each figure, for the pair  $\{\mathbf{f}_{model}, \mathbf{f}_{min}\}$ , corresponding constraints are: a)  $\{\Omega_1^2, \Omega_1^2\}$ ; b)  $\{\Omega_2^2, \Omega_2^2\}$ , c)  $\{\Omega_3^2, \Omega_3^2\}$ , d)  $\{\Omega_{TV}, \Omega_1^2\}$ , e)  $\{\Omega_{TV}, \Omega_2^2\}$ , f)  $\{\Omega_{TV}, \Omega_3^2\}$ .

with initial condition  $u^{(2)}(x, t = 0)$  is considered. The corresponding results are represented in Fig. 7.3 with the same assumptions of Fig. 7.2. It is obvious that the effects of spurious oscillations are considerable in this case. This is due to existence of more high frequency components in this system in comparison to the previous smooth high-gradient solutions.

To study the estimation of errors in different regularization approaches, the L-curves of solutions are also presented. The L-curve is a graphical log-log representation of a constraint  $\Omega^2$  (or  $\Omega_{TV}$ ) against the estimation error  $\rho_2^2$ . Regarding the 1D wave propagation problem, for the smooth-high gradient IC ( $u^{(1)}(x, t = 0)$ ), corresponding L-curves are presented in Fig 7.4. In the simulations, different regularization parameters are assumed, as:  $p \in \{0.65, 0.7, 0.73, 0.75, 0.8, 0.85, 0.9, 0.92, 0.94, 0.96, 0.98, 0.99\}$ . The results are for time  $t = 0.2$ .

In figure (a), the L-curve of numerical solutions obtained by the Tikhonov-based filtering are compared with each other; the considered constraints are:  $\Omega_1^2$ ,  $\Omega_2^2$ , and  $\Omega_3^2$ . It is clear that the common Tikhonov method (using  $\Omega_1^2$ ) leads to the smoothest possible result. By using the two other constraints, at the expense of error in estimations,



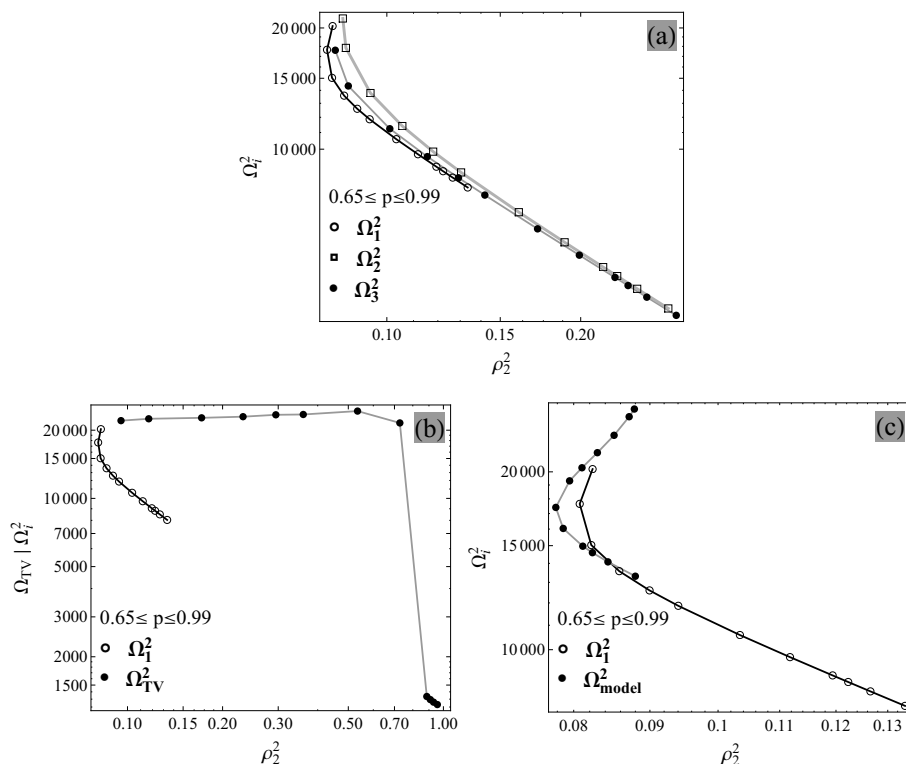


Figure 7.4: The L-curves for numerical solutions of the 1D wave propagation problem at  $t = 0.2$ , denoised with different regularization approaches.

smoother results can be obtained. Also  $\Omega_2^2$  constraint leads to slightly smoother results than those of  $\Omega_3^2$ . In figure (b) TV-based results are compared with those obtained by the common Tikhonov scheme; it is obvious, the TV-based results lead to worse results (this is due to lack of enough-smoothness). Finally, in figure (c), the common Tikhonov method with and without the model are compared with each other; for the model-based filtering, the constraint pair  $\{\mathbf{f}_{model}, \mathbf{f}_{min}\} = \{\Omega_1^2, \Omega_1^2\}$  is used. In the model-based Tikhonov regularization, smooth solutions are only obtained for small  $p$  values; this is due to presence of small high-frequency oscillations in over-smoothed regularized solutions (as seen before).

### 7.3 Higher order finite difference methods and different filtering approaches

Considering linear wave propagation problems, two general higher-order differencing approaches will be reviewed in this section. The approaches are: 1) those using inherent

filtering; 2) ones utilizing filters by a post-processing stage. In the former case, spatial filtering can inherently be considered in corresponding formulations. Performance of these two approaches would then be studied for both smooth-high gradient and discontinuous solutions. Below, above-mentioned approaches and corresponding filters are reviewed.

### Approach 1: Inherent filtering

Here, filtering effects are inserted in definition of spatial derivatives. This inserting can be done to have either a maximum order of accuracy or some optimized quantities [131, 132].

1. Maximum order schemes In this approach, order of accuracy is maximized by considering error terms. For a uniform grid  $x_j := j \Delta x$ , at point  $x_j$ , the first spatial derivative can be approximated as:

$$(\delta_x u)_j = \frac{1}{\Delta x} \{ (d_3 - a_3) u_{j-3} + (d_2 - a_2) u_{j-2} + (d_1 - a_1) u_{j-1} + d_0 u_j + (d_1 + a_1) u_{j+1} + (d_2 + a_2) u_{j+2} + (d_3 + a_3) u_{j+3} \}, \quad (7.2)$$

where:  $u_j := u(x_j)$ ;  $(\delta_x u)_j := d(u(x_j))/dx$ ;  $a_i$  and  $d_i$  are coefficients of anti-symmetric  $((\delta_x^a u)_j)$  and symmetric  $((\delta_x^s u)_j)$  parts, respectively. The symmetric part,  $(\delta_x^s u)_j$  acts as a spatial filter. The maximum order of accuracy for  $(\delta_x u)_j$  without and with filtering effects  $((\delta_x^s u)_j)$  are six and five, respectively. For the six-order accuracy, the coefficients are:  $a_1 = 3/4$ ,  $a_2 = -3/20$ ,  $a_3 = 1/60$ ,  $d_1 = d_2 = d_3 = 0$ ; and for the five-order one, the set  $\{d_i\}$  becomes:  $d_1 = -3d_0/4$ ,  $d_2 = 3d_0/10$ , and  $d_3 = -d_0/20$ . In our study, it is assumed:  $d_0 = 0.1$  [131].

Regarding a system of pure advection,  $\mathbf{u}_t + (\mathbf{A}\mathbf{u})_x = \mathbf{0}$  (where,  $\mathbf{u}_t := d\mathbf{u}/dt$  and  $(\mathbf{A}\mathbf{u})_x := d(\mathbf{A}\mathbf{u})/dx$ ), the spatial first derivative of the flux  $(\mathbf{A}\mathbf{u})$  can be approximated as:

$$(\mathbf{A}\mathbf{u})_x \approx \delta_x^a \mathbf{A}\mathbf{u} + \delta_x^s |\mathbf{A}| \mathbf{u}, \quad (7.3)$$

where:  $|\mathbf{A}| = \mathbf{X} |\boldsymbol{\Lambda}| \mathbf{X}^{-1}$ . Matrices  $\mathbf{X}$  and  $\boldsymbol{\Lambda}$  denote right eigenvectors and eigenvalues of  $\mathbf{A}$ , respectively.

Temporal integration of the advection equation can be done by the six stage explicit method mentioned in [131], as:

$$u_{n+\alpha_i}^{(i)} = u_n + \Delta t \alpha_i f_{n+\alpha_{i-1}}^{(i-1)}, \quad \text{for } i = \{1, 2, \dots, 6\}, \quad (7.4)$$

where:  $i$  denotes integration stage;  $\alpha_0 = 0$ , and  $\alpha_6 = 1$ ;  $f_n^{(0)} = f_n$ ;  $u_{n+\alpha_6}^{(6)} = u_{n+1}$ ;  $u_n := u(t_n)$ ;  $f^{(k)} := du^{(k)}/dt$ . To have a six-order temporal accuracy, in case of linear homogeneous ordinary differential equations (ODEs), the set  $\{\alpha_i\}$  should be:  $\alpha_1 = 1/6$ ,  $\alpha_2 = 1/5$ ,  $\alpha_3 = 1/4$ ,  $\alpha_4 = 1/3$ ,  $\alpha_5 = 1/2$ .

2. The optimized method Definition of  $(\delta_x u)_j$  (from Eq. (7.2)) with filtering effects and six-stage temporal integration (Eq. (7.4)) can be optimized for some desire error behavior or requirements. To resolve more waves of larger wave numbers (or short waves), corresponding optimized solution can be obtained in spatio-temporal domains by coefficients:  $a_1 = 0.7599613$ ,  $a_2 = -0.1581220$ ,  $a_3 = 0.01876090$ ,  $d_1 = -0.07638461$ ,  $d_2 = 0.03228961$ ,  $d_3 = -0.005904994$ ,  $\alpha_1 = 0.168850$ ,  $\alpha_2 = 0.197348$ ,  $\alpha_3 = 0.250038$ ,  $\alpha_4 = 0.333306$ ,  $\alpha_5 = 0.5$  [131, 132].

### Approach 2: Post-processing based filtering

In this approach, at first, derivatives or solutions are obtained and then by a post-processing stage, non-physical oscillations are removed from numerical solutions. In the following, generalized compact finite difference schemes with different features and a general filtering method are reviewed.

**Compact finite difference schemes** In this approach, a linear combination of data values  $\{u_i\}$  are locally used for estimation of derivatives values in each grid point. This combination for the first and second derivatives can be written as [138]:

$$\beta_m u_{i-2}^{(m)} + \alpha_m u_{i-1}^{(m)} + u_i^{(m)} + \alpha_m u_{i+1}^{(m)} + \beta_m u_{i+2}^{(m)} = c_m Z_{i,3}^{(m)} + b_m Z_{i,2}^{(m)} + a_m Z_{i,1}^{(m)}, \quad \text{for } m \in \{1, 2\}, \quad (7.5)$$

where:  $Z_{i,j}^{(1)} := \frac{u_{i+j} - u_{i-j}}{2(j\Delta x)}$ ;  $Z_{i,j}^{(2)} := \frac{u_{i+j} - 2u_i + u_{i-j}}{(j\Delta x)^2}$ ; and  $u_i^{(m)} := d^m u_i / dx^m$ . Eq. (7.5) is also known as the generalized Padé scheme. Relationships between coefficients  $\{a_m, b_m, c_m\}$  and  $\{\alpha_m, \beta_m\}$  can be obtained by matching the Taylor series coefficients. The truncation

Name	Method/order	$\beta_2$	$\alpha_2$	$a_2$	$b_2$	$c_2$
4-C	Central/ $4^{th}$	0	0	$\frac{4}{3}(1 - \alpha_2)$	$\frac{1}{3}(-1 + 10\alpha_2)$	0
4-P	Padé/ $4^{th}$	0	1/10	$\frac{4}{3}(1 - \alpha_2)$	$\frac{1}{3}(-1 + 10\alpha_2)$	0
6-T	Tridiagonal scheme/ $6^{th}$	0	2/11	$\frac{4}{3}(1 - \alpha_2)$	$\frac{1}{3}(-1 + 10\alpha_2)$	0
8-C	Collatz/ $8^{th}$	$\frac{38\alpha_2-9}{214}$	$\frac{344}{1179}$	$\frac{696-1191\alpha_2}{428}$	$\frac{2454\alpha_2-294}{535}$	0
10	$10^{th}$	$\frac{43}{1798}$	$\frac{334}{899}$	$\frac{1065}{1798}$	$\frac{1038}{899}$	$\frac{79}{1798}$
4-S	Spectral-like/ $4^{th}$	0.05569169	0.50209266	0.21564935	1.7233220	0.17659730

Table 7.3.1: Coefficients of generalized Padé approximations for the second derivative with different accuracy and periodic boundary conditions [138].

Name	Method /order	$\hat{\beta}$	$\hat{\alpha}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$
F-6-E	Explicit/ $6^{th}$	3/10	0	1/2	3/4	3/10	1/20
F-4-T-1	Tridiagonal/ $4^{th}$	0	0.4	$\frac{(5+6\hat{\alpha}-6\hat{\beta}+16\hat{d})}{8}$	$\frac{(1+2\hat{\alpha}+2\hat{\beta}-2\hat{d})}{2}$	$\frac{(1-2\hat{\alpha}-14\hat{\beta}+16\hat{d})}{-8}$	0
F-4-T-2	Tridiagonal/ $4^{th}$	0	0.475	$\frac{(5+6\hat{\alpha}-6\hat{\beta}+16\hat{d})}{8}$	$\frac{(1+2\hat{\alpha}+2\hat{\beta}-2\hat{d})}{2}$	$\frac{(1-2\hat{\alpha}-14\hat{\beta}+16\hat{d})}{-8}$	0
F-4-P-1	Pentadiagonal/ $4^{th}$	0.2265509	0.4627507	0.8470630	1.166845	0.3422386	0.02245659
F-4-P-2	Pentadiagonal/ $4^{th}$	0.1702929	0.6522474	0.9891856	1.321180	0.3333548	0.001359850

Table 7.3.2: Filter coefficient values used for filtering the generalized Padé approximations by a post-processing stage for the periodic boundary condition [138].

error can be obtained by the first un-matched coefficient. The coefficients  $\{a_m, b_m, c_m\}$  and  $\{\alpha_m, \beta_m\}$  for case  $m = 2$  (i.e., the second derivative) with periodic boundary conditions are presented in Table 7.3.1 for different spatial accuracy [138]. For estimation of first and second derivatives on a domain with general boundary conditions, please see [138].

Noise can be filtered by a post-processing stage in the spatial domain; following the idea of local differencing, filtered data can be estimated as [138]:

$$\hat{\beta}\hat{u}_{i-2} + \hat{\alpha}\hat{u}_{i-1} + \hat{u}_i + \hat{\alpha}\hat{u}_{i+1} + \hat{\beta}\hat{u}_{i+2} = \hat{d}\hat{Z}_{i,3} + \hat{c}\hat{Z}_{i,2} + \hat{b}\hat{Z}_{i,1} + \hat{a}\hat{Z}_{i,0}, \quad (7.6)$$

where:  $\hat{Z}_{i,j} := \frac{u_{i+j} + u_{i-j}}{2}$ ; and  $\{\hat{u}_i\}$  denotes the filtered data at grid points  $\{x_i\}$ . The coefficients  $\{\hat{a}, \hat{b}, \hat{c}, \hat{d}\}$  and  $\{\hat{\alpha}, \hat{\beta}\}$  are presented in Table 7.3.2 for the periodic boundary condition and for different accuracy.

### Performance of two approaches for smooth and discontinuous solutions

Below, performance of two approaches (1) and (2) will be studied for both smooth and discontinuous solutions by a 1D scalar stress wave propagation problem. There, effects of

different filtering methods, inherent and post-processing ones, will be considered. In case of the post-processing approach, effects of Tikhonov-based smoothing (regularization) will also be studied.

The example is a wave propagation problem in an elastic bar with unit length, unit wave-propagation velocity ( $c = 1$ ) and periodic boundaries. The problem is only subjected to an imposed initial displacement  $u_0(x)$  ( $v_0(x) = 0$ ). Two different smooth and discontinuous initial displacements are considered respectively, as:

$$u_0(x) := \text{Exp}(-500(x - 0.5)^2), \text{ and } u_0(x) := \text{Unitbox}(4(x - 0.5)) \text{ for } 0 \leq x \leq 1.$$

In numerical simulations, it is assumed:

1. Both the  $l_2$  and  $l_1$  norms are used to measure errors; error definitions for  $l_2$  and  $l_1$  norms are  $\|e\|_2 := \frac{1}{N_g} \left\{ \sum_{i=1}^{N_g} (u_i - u_i^{ex})^2 \right\}^{1/2}$  and  $\|e\|_1 := \frac{1}{N_g} \sum_{i=1}^{N_g} |u_i - u_i^{ex}|$ , respectively; where  $u_i := u(x_i)$ ;  $u^{ex}$  denotes the exact solution; and  $N_g$  shows number of grid points,
2. For the time integration, the Runge-kutta 4<sup>th</sup> order is used for the explicit 4<sup>th</sup> order (4-C) and the Padé differencing schemes (4-P) (see Table 7.3.1); the 6<sup>th</sup> order, six stage temporal integration (Eq. (7.4)) is used for remaining spatial differencing operators [132],
3. In all simulations, a fixed Courant number,  $C_{CFL} := \frac{c \Delta t}{\Delta x}$  (known as the CFL condition) is used, as:  $C_{CFL} = 1/3$ ,
4. Smoothing is done at each time step (for both inherent and post-processing approaches).

Convergence rates are studied in Fig. 7.5 for both smooth and discontinuous solutions at  $t = 0.2$ ; slope of lines are measured as  $1 : (-v)$  (i.e.: 1 and  $(-v)$  units in the horizontal and vertical directions, respectively). In this figure, the top and bottom rows belong to the smooth and discontinuous solutions, respectively. Fig. 7.5(a) corresponds to the generalized Padé schemes without filtering stage. The filtered results (by a post-processor) are represented in Fig. 7.5(b) for two different filters: the 6<sup>th</sup> order explicit (“f-6-E”) and 4<sup>th</sup> order implicit pentadiagonal (“f-4-p-1”) methods (see Table 7.3.2). In Fig. 7.5(c), the results belong to the inherent filtering approach; here, the 6<sup>th</sup> order method (Max order-6<sup>th</sup>) does not use any filters. Fig. 7.5(d) is for the generalized Padé schemes using the Tikhonov-based smoothing with constraints  $\Omega_1^2$  and  $\Omega_2^2$  (with parameters:  $p = 0.99$  and  $\alpha = 0.99$ ). For the discontinuous solution, all of the results are

re-presented in Figs. 7.5(e)- 7.5(g). The post-processing (Eq. (7.6)), inherent (Eq. (7.2)) and Tikhonov-based filtering are used for denoising compact difference schemes and their results are presented respectively in Figs. 7.5(e), 7.5(f) and 7.5(g). The results offer that: 1) in smooth solutions (Fig. 7.5(a)), predicted convergence rates can be obtained (except for the 10<sup>th</sup> order one where it seems that a more accurate time integration method should be used); 2) filtering changes convergence rates and this is considerable for discontinuous solutions; in this case, the rates are generally less than one; 3) regarding discontinuous solutions, for the inherent filtering approach, performance of optimized methods is better than those of the maximum order schemes; this is because, optimized methods can detect more waves of large wave numbers (this will be clarified); 3) in smooth solutions, performance of the Tikhonov method with the constraint  $\Omega_1^2$  is better than those with the constraint  $\Omega_2^2$ ; however in discontinuous solutions, the  $\Omega_2^2$  constraint can control the numerical dispersion more effectively and can prevent the numerical dispersion (will be studied).

Convergence rates in the  $l_1$  norm is also presented for the discontinuous solutions in Fig. 7.6. It is clear that convergence rates are less than one for the all cases.

To clarify filtering effectiveness (for controlling the numerical dispersion and dissipation phenomena), solutions  $u(x, t = 0.2)$  are presented in Fig. 7.7 for different FD methods and filtering approaches. It is clear that the Tikhonov method with constraint  $\Omega_2^2$  can properly be integrated with other differencing schemes to control the numerical dispersion.

## 7.4 Studying the Tikhonov regularization with different constraints by the SVD and GSVD decompositions

In this section, at first, the definitions of singular value decomposition (SVD) and generalized SVD (GSVD) methods [146] are briefly reviewed. Thereafter, with these methods, the effects of the constraints  $\Omega_1^2$ ,  $\Omega_2^2$ , and  $\Omega_3^2$  with and without a model-solution ( $\mathbf{f}_{model}$ ) will be studied.

**The SVD decomposition** The SVD decomposition of a rectangular matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is:  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \sum_{i=1}^n \mathbf{u}_i \sigma_i \mathbf{v}_i^T$ . Matrices  $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_m\} \in \mathbb{R}^{m \times m}$  and  $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \in \mathbb{R}^{n \times n}$  are unitary matrices containing respectively the left singular ( $\mathbf{u}_i \in \mathbb{R}^m$ ) and right singular ( $\mathbf{v}_i \in \mathbb{R}^n$ ) vectors. The matrix  $\mathbf{\Sigma}_{m \times n} = \text{diag}(\sigma_1, \dots, \sigma_n)$  is diagonal and contains singular values  $\sigma_i$ , where  $\sigma_i \geq \sigma_{i+1}$  and  $\sigma_i \geq 0$ .

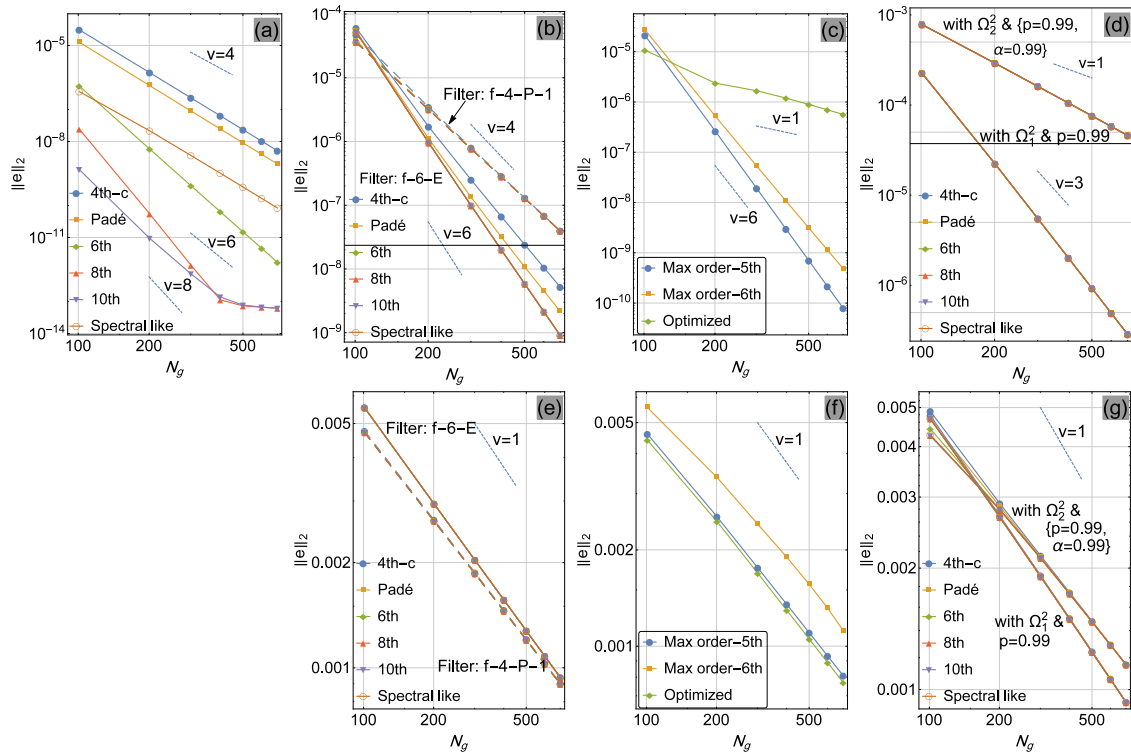


Figure 7.5: Convergence rates in the  $l_2$  norm for different FD methods for smooth (top row) and discontinuous (bottom row) solutions; a) different compact FDs without filtering; b, e) different compact FDs with filtering; c, f) the inherent filtering approach; d, g) Tikhonov based smoothing by a post-processor stage.

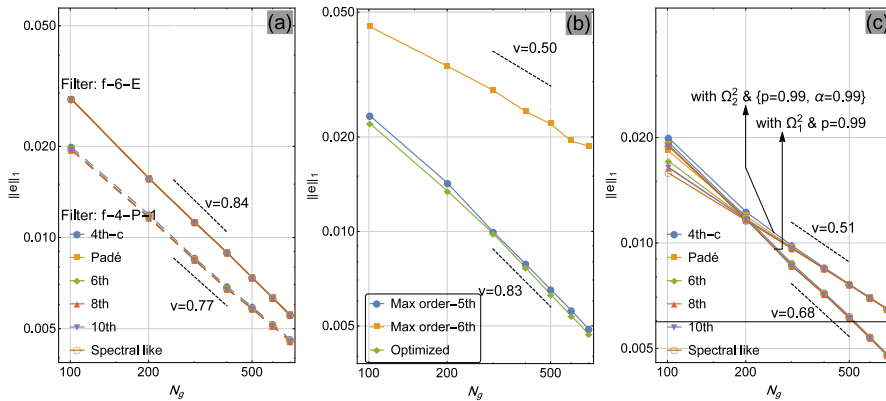


Figure 7.6: Convergence rates in the  $l_1$  norm for different FD methods for discontinuous solutions.

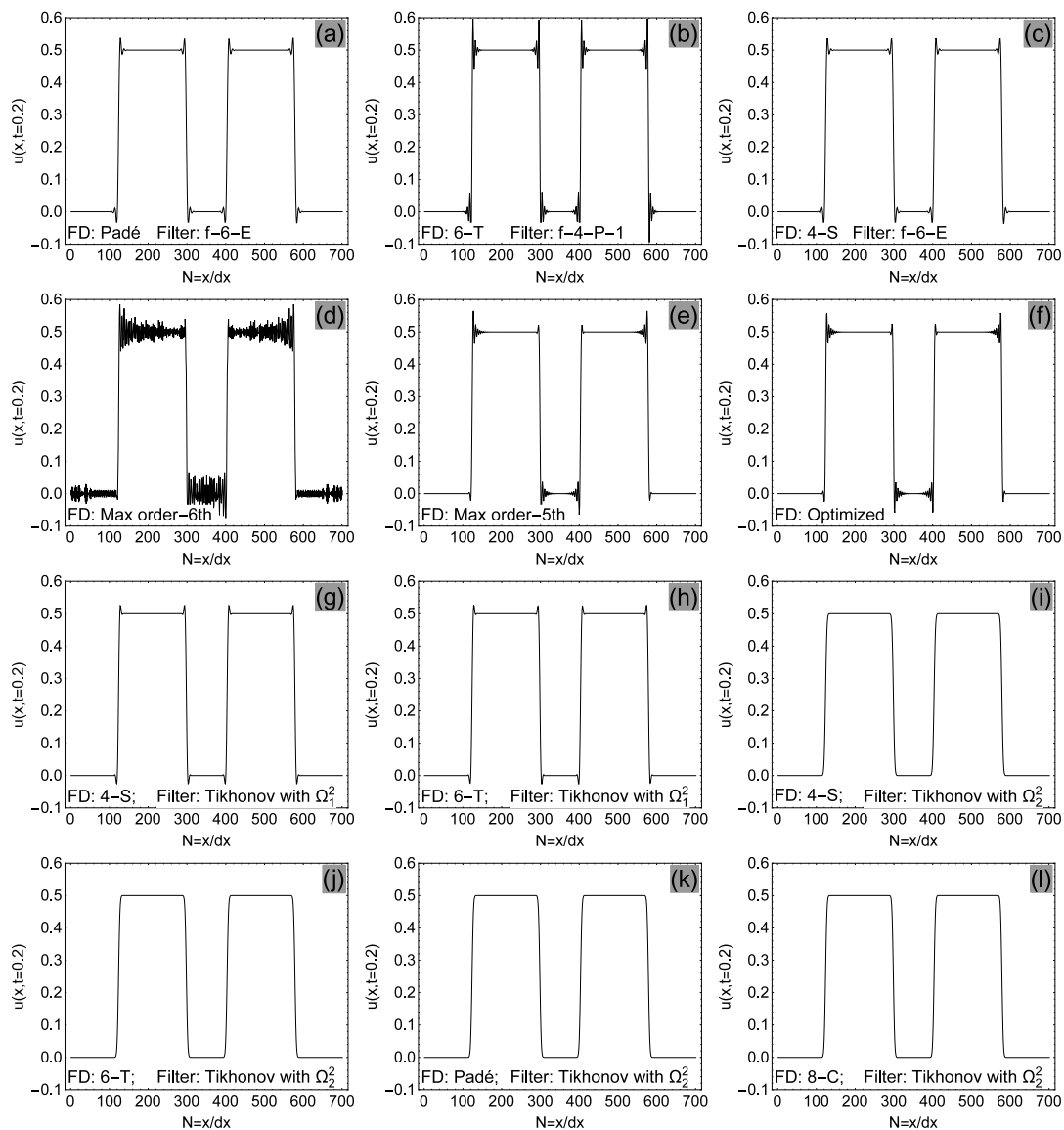


Figure 7.7: Discontinuous solutions  $u(x, t = 0.2)$  for different FD schemes and different filtering approaches.



**The GSVD decomposition** Let us assume:  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and  $\mathbf{L} \in \mathbb{R}^{p \times n}$ , where:  $m \geq n \geq p$  and  $\mathcal{N}(\mathbf{A}) \cap \mathcal{N}(\mathbf{L}) = \{\mathbf{0}\}$  (symbol  $\mathcal{N}$  shows the null space). Then the GSVD of  $(\mathbf{A}, \mathbf{L})$  is [146]:

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-p} \end{pmatrix} \mathbf{X}^{-1}, \quad \mathbf{L} = \mathbf{V} (\mathbf{M}, \mathbf{0}) \mathbf{X}^{-1}, \quad (7.7)$$

where:  $\mathbf{U} \in \mathbb{R}^{m \times n}$ ;  $\mathbf{V} \in \mathbb{R}^{p \times p}$ ;  $\mathbf{X} \in \mathbb{R}^{n \times n}$ . Matrices  $\mathbf{U}$  and  $\mathbf{V}$  have the unitary feature; and  $\mathbf{X}$  is a non-singular matrix. Matrices  $\mathbf{\Sigma}$  and  $\mathbf{M}$  are non-negative diagonal matrices as:  $\mathbf{\Sigma} = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_p\}$  and  $\mathbf{M} = \text{diag}\{\mu_1, \mu_2, \dots, \mu_p\}$ . The singular values have following properties: 1)  $1 \geq \sigma_p \geq \sigma_{p-1} \geq \dots \geq \sigma_1 > 0$ ; 2)  $1 \geq \mu_1 \geq \mu_2 \geq \dots \geq \mu_p > 0$ ; 3)  $\sigma_i^2 + \mu_i^2 = 1$ . The *generalized singular values*  $\gamma_i$  of  $(\mathbf{A}, \mathbf{L})$  is then equal to:  $\gamma_i = \sigma_i / \mu_i$ .

### The Tikhonov method without a model (extra information)

Regarding the linear system  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , functional of the Tikhonov regularization with constraint operator  $\mathbf{L}$  is :  $Q = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda \|\mathbf{L}\mathbf{x}\|_2^2$ . It is easy to show that, the minimizing solution of  $Q$  is:  $\mathbf{x}_{min} = (\mathbf{A}^T \mathbf{A} + \lambda \mathbf{L}^T \mathbf{L})^{-1} (\mathbf{A}^T \mathbf{b})$ . By substituting the SVD/GSVD decomposed forms of the operators, the  $\mathbf{x}_{min}$  can be rewritten as [146]:

$$\begin{aligned} \mathbf{x}_{min} &= \sum_{i=1}^n \frac{\mathbf{u}_i^T \cdot \mathbf{b}}{\sigma_i} \mathbf{v}_i, & \text{for } \lambda = 0 & \text{ (without any regularization),} \\ \mathbf{x}_{min} &= \sum_{i=1}^n F_i \frac{\mathbf{u}_i^T \cdot \mathbf{b}}{\sigma_i} \mathbf{v}_i : & F_i &= \frac{\sigma_i^2}{\sigma_i^2 + \lambda}, & \text{for } \mathbf{L} = \mathbf{I}, & (7.8) \\ \mathbf{x}_{min} &= \sum_{i=1}^p F_i \frac{\mathbf{u}_i^T \cdot \mathbf{b}}{\sigma_i} \mathbf{x}_i + \sum_{i=p+1}^n (\mathbf{v}_i^T \cdot \mathbf{b}) \mathbf{x}_i : & F_i &= \frac{\gamma_i^2}{\gamma_i^2 + \lambda} & \text{for } \mathbf{L} \neq \mathbf{I}. \end{aligned}$$

In the above equations,  $F_i$  acts as a filter; it damps effects of singular values with small values. For solutions obtained without regularization (i.e.:  $\lambda = 0$ ), the solution  $\mathbf{x}_{min}$  would be sensitive for  $\sigma_i$  values of small (or nearly zero) values. They amplify corresponding  $\mathbf{v}_i$  vector effects: the vectors with high fluctuations. Hence a proper inverse solution  $\mathbf{x}_{min}$  is not obtainable.

In the following, performance of  $F_i$  values for the Tikhonov regularization method with different constraints will be studied. The constraints are: 1)  $\Omega_1^2(x) = \int_{-\infty}^{+\infty} (f''(x))^2 dx$ ; 2)  $\Omega_2^2(x) = \int_{-\infty}^{+\infty} (f''(x) + \alpha f'(x))^2 dx$ ; 3)  $\Omega_3^2(x) = \int_{-\infty}^{+\infty} (f''(x)^2 + \alpha^2 f'(x)^2) dx$ .

For numerical study, the Heaviside function,  $H(x)$  is considered for spatial domain:

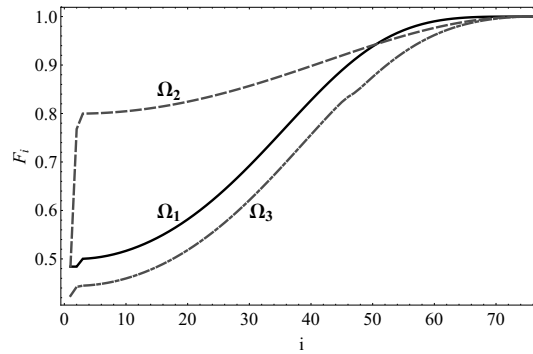


Figure 7.8: Filter coefficients for different constraints in the Tikhonov method, where  $p = 0.8$  &  $\alpha = 1$ .

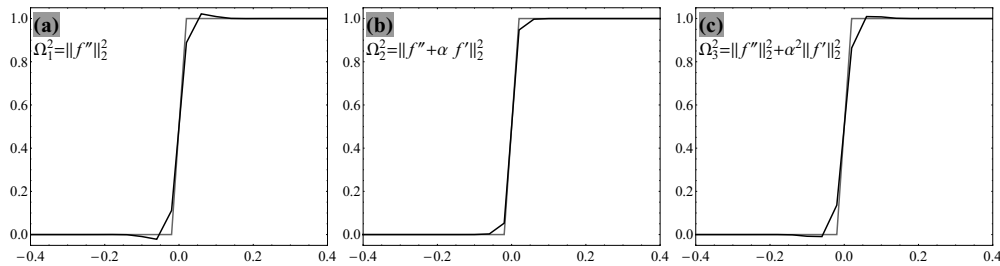


Figure 7.9: Regularized solutions with different constraints, where  $p = 0.8$  &  $\alpha = 1$ .

$-1.5 \leq x \leq 1.5$ . The function is uniformly sampled at 76 points. In Fig. 7.8,  $F_i$  values are presented for these three constraints with parameter  $p = 0.8$  and  $\alpha = 1$ . The figure offers that filters due to the constraint  $\Omega_1^2$  and  $\Omega_3^2$  have the same trend, but the latter leads to smaller filter coefficients. The constraint  $\Omega_2^2$  has a different trend and has larger filter values for small  $i$  values.

Regularized solutions according to Eq. (7.8) are presented in Fig. 7.9; it is clear that the Tikhonov method with  $\Omega_2$  constraint can effectively control the numerical dispersion.

The final stage for this problem is checking the discrete Picard condition; this condition states that to be sure that a regularization scheme works properly, corresponding  $|\mathbf{u}_i \cdot \mathbf{b}|$  values should on average decay toward zero faster than corresponding generalized singular values  $\gamma_i$ . This comparison is done in Fig. 7.10 for the three constraints. It is clear that results from constraints  $\Omega_1$  and  $\Omega_2$  satisfy completely the condition and the constraint  $\Omega_3$ , on average meets the discrete Picard condition.

### The Tikhonov method with a model (extra information)

If it is assumed that we have a model for a solution, like  $(\mathbf{b}_{model})_{p \times 1}$ , and  $\mathbf{L} \neq \mathbf{I}$ , then the corresponding functional is:  $Q = \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda \|\mathbf{L}(\mathbf{x} - \mathbf{b}_{model})\|_2^2$ . It is straightforward

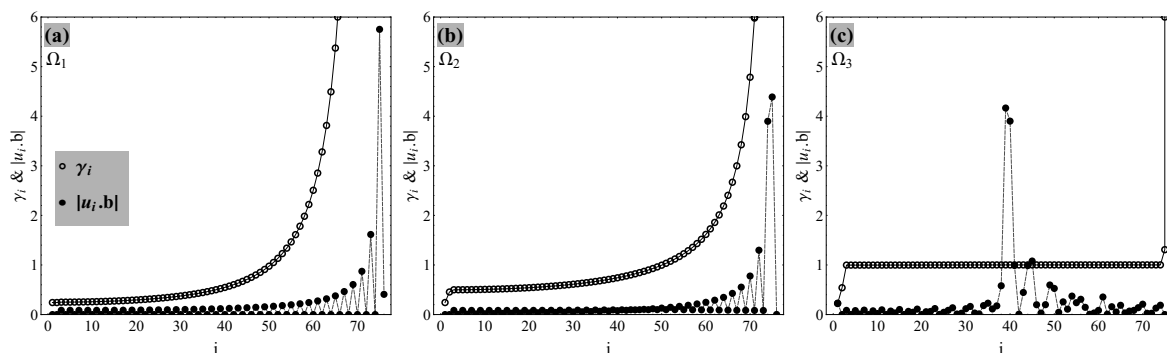


Figure 7.10: Control of the discrete Picard condition for regularized solutions with different constraints, where  $p = 0.8$  &  $\alpha = 1$ .

to show that the solution is:  $\mathbf{x}_{min} = (\mathbf{A}^T \cdot \mathbf{A} + \lambda \mathbf{L}^T \cdot \mathbf{L})^{-1} \{ \mathbf{A}^T \cdot \mathbf{b} + \lambda \mathbf{L}^T \cdot \mathbf{L} \cdot \mathbf{b}_{model} \}$ . If we set  $\mathbf{L} \cdot \mathbf{b}_{model} = \mathbf{b}_{model}^L$  (this setting helps more clarifying behavior of regularized solutions) and using the GSVD, it is easy to show that the solution  $\mathbf{x}_{min}$  can be written as:

$$\mathbf{x}_{min} = \left\{ \sum_{i=1}^p F_i \frac{\mathbf{u}_i^T \cdot \mathbf{b}}{\sigma_i} \mathbf{x}_i + \sum_{i=p+1}^n (\mathbf{u}_i^T \cdot \mathbf{b}) \mathbf{x}_i \right\} + \left\{ \sum_{i=1}^p \alpha_i \frac{\mathbf{v}_i^T \cdot \mathbf{b}_{model}^L}{\mu_i} \mathbf{x}_i \right\}, \quad (7.9)$$

where:  $F_i = \frac{\gamma_i^2}{\gamma_i^2 + \lambda}$  and  $\alpha_i = \frac{\lambda}{\gamma_i^2 + \lambda}$ .

Depending on  $\lambda$  values, two extreme conditions can be happened: 1) if  $\lambda \gg 0$ , then  $\alpha_i \rightarrow 1$ ; 2) if  $\lambda \rightarrow 0$  then  $\{F_i \rightarrow 1 \text{ \& } \alpha_i \rightarrow 0\}$ . For the case (1),  $\mu_i$  coefficients of small values cause oscillations in regularized solutions and in the case (2), filter coefficients  $F_i$  do not filter out oscillations. Hence, it is clear that why oscillation effects always remain in the regulated solutions. This effect will be amplified in discontinuous solutions (as numerically shown in numerical simulation of the 1D stress wave propagation problem, Sec. 7.2).

## 7.5 Implementation algorithms

In this section, implementation algorithms of Tikhonov-based regularization schemes will be presented.

Regarding sampled data  $\{y_i\}$  for  $i \in \{1, 2, \dots, n\}$ , the functional  $Q$  of the Tikhonov method is:

$$Q = \sum_{i=1}^n (y_i - f_i)^2 + \lambda \times \Omega(f)_2^2. \quad (7.10)$$

In a more general form, in case of continuous functions  $y := y(x)$ , the functional  $Q$  can be rewritten as:

$$Q = \int_{-\infty}^{+\infty} (y - f)^2 dx + \lambda \times \Omega(f)_2^2. \quad (7.11)$$

The solution  $f$  is the stationary point (extreme) of functional  $Q$ . For different  $\Omega(f)$  definitions, corresponding implementation algorithms will be provided.

Using the finite difference approximation for discrete values  $\{f(x_i) : i = 1, 2, \dots, n\}$ , the order- $m$  derivative can be approximated as:  $\mathbf{f}^{(m)} \approx \mathbf{D}^{(m)} \cdot \mathbf{f}$ . Derivative Matrices  $\mathbf{D}^{(1)}$  and  $\mathbf{D}^{(2)}$  can be approximated for uniform grids as:

$$\mathbf{D}^{(1)} = \frac{1}{\Delta x} \begin{pmatrix} -1 & 1 & & & \\ -1 & 1 & & & \\ & & \ddots & \ddots & \\ & & & & -1 & 1 \end{pmatrix}_{n \times n}, \quad \mathbf{D}^{(2)} = \frac{1}{\Delta x^2} \begin{pmatrix} 1 & -2 & 1 & & & \\ 1 & -2 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & & 1 & -2 & 1 \\ & & & & 1 & -2 & 1 \end{pmatrix}_{n \times n}, \quad (7.12)$$

where  $\Delta x = x_{i+1} - x_i$  is the uniform sampling step in spatial domain. The  $\mathbf{D}^{(1)}$  and  $\mathbf{D}^{(2)}$  matrices, both have first-order spatial accuracy.

For numerically discretization of an integral (e.g., the integral in Eq. (7.11)), it can be approximated as:

$$\int_{x_1}^{x_n} |f(x)|^2 dx \approx \mathbf{f}^T \cdot \mathbf{B} \cdot \mathbf{f}, \quad (7.13)$$

where  $\mathbf{B}$  denotes the integration rule matrix; for the midpoint integration rule, we have:

$$\int_{x_1}^{x_n} |f(x)|^2 dx \approx \frac{1}{2}(-x_1 + x_2)f_1^2 + \sum_{i=2}^{n-1} \left\{ \frac{1}{2}(-x_{i-1} + x_{i+1})f_i^2 \right\} + \frac{1}{2}(-x_{n-1} + x_n)f_n^2, \quad (7.14)$$

or

$$\int_{x_1}^{x_n} |f(x)|^2 dx \approx \frac{1}{2} f_1(-x_1 + x_2) f_1 + \sum_{i=2}^{n-1} \left\{ \frac{1}{2} f_i(-x_{i-1} + x_{i+1}) f_i \right\} + \frac{1}{2} f_n(-x_{n-1} + x_n) f_n. \quad (7.15)$$

So, the matrix  $\mathbf{B}$  can be obtained as:

$$\begin{aligned} \int_{x_1}^{x_n} |f(x)|^2 dx &\approx \mathbf{f}^T \cdot \left\{ \frac{1}{2} \text{diag}(-x_1 + x_2, -x_1 + x_3, \dots, -x_{n-2} + x_n, -x_{n-1} + x_n) \right\} \cdot \mathbf{f} \\ &= \mathbf{f}^T \cdot \mathbf{B} \cdot \mathbf{f}, \end{aligned} \quad (7.16)$$

where:  $\text{diag}\{a_1, \dots, a_n\}$  shows a  $n \times n$  diagonal matrix with diagonal elements  $a_i$ ; the vector  $\mathbf{f}$  denotes  $\mathbf{f} = \{f_1, \dots, f_n\}^T$ .

Considering above-mentioned derivative and integral approximations, the minimizing solution,  $\mathbf{f}_{min}$ , for the Tikhonov regularization with different constraints will be presented in the following.

**Tikhonov regularization with constraint:**  $\Omega_2^2 = \int_{-\infty}^{+\infty} (f''(x) + \alpha f'(x))^2 dx$ . Firstly, the functional  $Q(\mathbf{f})$  is rewritten as:

$$Q(\mathbf{f}) = \int (y - f)^2 dx + \lambda \times \left\{ \int (f''^2 + \alpha^2 f'^2 + 2\alpha f' f'') dx \right\}. \quad (7.17)$$

The discretized form of the functional  $Q$  is:

$$\begin{aligned} Q(\mathbf{f}) &= (\mathbf{y} - \mathbf{f})^T \cdot \mathbf{B} \cdot (\mathbf{y} - \mathbf{f}) + \lambda \times \\ &\left\{ (\mathbf{D}^{(2)} \mathbf{f})^T \cdot \mathbf{B} \cdot (\mathbf{D}^{(2)} \mathbf{f}) + \alpha^2 (\mathbf{D}^{(1)} \mathbf{f})^T \cdot \mathbf{B} \cdot (\mathbf{D}^{(1)} \mathbf{f}) + 2\alpha (\mathbf{D}^{(1)} \mathbf{f})^T \cdot \mathbf{B} \cdot (\mathbf{D}^{(2)} \mathbf{f}) \right\}. \end{aligned} \quad (7.18)$$

To find minimizing solution, the functional  $Q$  is minimized with respect to  $\mathbf{f}$  as:

$$\begin{aligned} \frac{\partial Q(\mathbf{f})}{\partial \mathbf{f}} = & -2\mathbf{B} \cdot (\mathbf{y} - \mathbf{f}) + \lambda \times \\ & \left\{ 2 \left( \mathbf{D}^{(2)} \right)^T \cdot \mathbf{B} \cdot \left( \mathbf{D}^{(2)} \cdot \mathbf{f} \right) + 2\alpha^2 \left( \mathbf{D}^{(1)} \right)^T \cdot \mathbf{B} \cdot \left( \mathbf{D}^{(1)} \cdot \mathbf{f} \right) + 2\alpha \left( \mathbf{D}^{(2)} \right)^T \cdot \mathbf{B} \cdot \left( \mathbf{D}^{(1)} \cdot \mathbf{f} \right) \right. \\ & \left. + 2\alpha \left( \mathbf{D}^{(1)} \right)^T \cdot \mathbf{B} \cdot \left( \mathbf{D}^{(2)} \cdot \mathbf{f} \right) \right\} = 0. \end{aligned} \quad (7.19)$$

Please note that  $\frac{\partial Q(\mathbf{f})}{\partial \mathbf{f}} = \left( \frac{\partial Q(\mathbf{f})}{\partial \mathbf{f}} \right)^T$  (since it is scalar) and  $\mathbf{B} = \mathbf{B}^T$  (since  $\mathbf{B}$  is diagonal matrix). Hence, the minimizing solution  $\mathbf{f}_{min}$  is:

$$\begin{aligned} \mathbf{f}_{min} = & \left\{ \mathbf{B} + \lambda \left( \mathbf{D}^{(2)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(2)} + \lambda \alpha^2 \left( \mathbf{D}^{(1)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(1)} + \right. \\ & \left. \lambda \alpha \left( \mathbf{D}^{(2)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(1)} + \lambda \alpha \left( \mathbf{D}^{(1)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(2)} \right\}^{-1} \cdot (\mathbf{B} \cdot \mathbf{y}). \end{aligned} \quad (7.20)$$

**Tikhonov regularization with constraint:**  $\Omega_3^2 = \int_{-\infty}^{+\infty} (f''(x)^2 + \alpha^2 f'(x)^2) dx$ . In this case, by following the previous procedure for finding  $\mathbf{f}_{min}$  for  $\Omega_2^2$ , it is easy to show that the solution is the same as Eq. (7.20), but this time without the first and the second derivative interaction terms; i.e.:

$$\mathbf{f}_{min} = \left\{ \mathbf{B} + \lambda \left( \mathbf{D}^{(2)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(2)} + \lambda \alpha^2 \left( \mathbf{D}^{(1)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(1)} \right\}^{-1} \cdot (\mathbf{B} \cdot \mathbf{y}). \quad (7.21)$$

### Some remarks

1. Minimizing solution  $f_{min}(x)$  for the commonly used Tikhonov regularization method with constraint  $\Omega_1^2 = \int_{-\infty}^{+\infty} f''(x)^2 dx$  can simply be obtainable by setting  $\alpha = 0$  in Eq. (7.20) or (7.21) (without any tension effect).
2. For weighted residual norm, i.e.:  $\|W_i(y_i - f_i)\|_2^2$ , with weights  $W_i$ , variable smoothing ( $\lambda = \lambda(x)$ ) and variable tension ( $\alpha = \alpha(x)$ ) parameters, it is easy to show that Eq. (7.20) can be written as:

$$\begin{aligned} \mathbf{f}_{min} = & \left\{ \mathbf{W} \cdot \mathbf{B} + \lambda \cdot \left( \mathbf{D}^{(2)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(2)} + \left( \lambda \cdot \boldsymbol{\alpha} \cdot \boldsymbol{\alpha}^T \right) \cdot \left( \mathbf{D}^{(1)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(1)} + \right. \\ & \left. \left( \lambda \cdot \boldsymbol{\alpha} \right) \cdot \left( \mathbf{D}^{(2)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(1)} + \left( \lambda \cdot \boldsymbol{\alpha} \right) \cdot \left( \mathbf{D}^{(1)} \right)^T \cdot \mathbf{B} \cdot \mathbf{D}^{(2)} \right\}^{-1} \cdot (\mathbf{W} \cdot \mathbf{B} \cdot \mathbf{y}), \end{aligned} \quad (7.22)$$

where:  $\mathbf{W} = \text{diag}\{W_1, W_2, \dots, W_n\}$ ;  $\boldsymbol{\lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ ;  $\boldsymbol{\alpha} = \text{diag}\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ . Corresponding algorithm for constraint  $\Omega_3^2$  can easily be obtained by canceling the first and the second derivative interaction effects in Eq. (7.22). For the common Tikhonov method (with  $\Omega_1^2$  constraint) by setting the tension term to zero, corresponding algorithm is obtainable.

3. For the model-based Tikhonov regularization with the general constraint  $\Omega_2^2 = \int_{-\infty}^{+\infty} (f''(x) + \alpha f'(x))^2 dx$ , it is easy to show that the corresponding numerical algorithm is:

$$\mathbf{f}_{min} = \{\mathbf{B} + \lambda\boldsymbol{\Gamma}\}^{-1} \{\mathbf{B}.\mathbf{y} + \lambda\boldsymbol{\Gamma}.\mathbf{f}_{model}\}, \quad (7.23)$$

where:

$$\boldsymbol{\Gamma} = \left\{ \left(\mathbf{D}^{(2)}\right)^T .\mathbf{B}.\mathbf{D}^{(2)} + \alpha^2 \left(\mathbf{D}^{(1)}\right)^T .\mathbf{B}.\mathbf{D}^{(1)} + \alpha \left(\mathbf{D}^{(2)}\right)^T .\mathbf{B}.\mathbf{D}^{(1)} + \alpha \left(\mathbf{D}^{(1)}\right)^T .\mathbf{B}.\mathbf{D}^{(2)} \right\}. \quad (7.24)$$

4. According to the aforementioned numerical implementation algorithms, since all matrices are sparse and banded, the algorithms are fast and cost-effective.
5. It should be mentioned that regularized results can have grid dependency feature; this means to have the same regularized results on finer or coarser grids, different regularization parameters should be used (i.e.,  $p$  and  $\alpha$ ). To prevent this grid dependency, here, the sampled data are remapped in a way that the new sampling step becomes unit, i.e.:  $\Delta x = 1$ .

## 7.6 Errors and convergence rates

### Convergence rate for the Tikhonov method with different constraints

The Tikhonov method with constraints  $\Omega_1^2$  and  $\Omega_3^2$  are special cases of  $(m, s)$ -splines (for case  $s = 0$ , this family becomes  $D^m$ -splines [163]) and  $(m, l, s)$ -splines (also known as  $L^{m,l,s}$ -splines), respectively [212]. The  $(m, s)$ -splines are a special case of  $(m, l, s)$ -splines for  $l = 0$  [212]. To be exact,  $(2, 0)$ -splines and  $(1, 1, 0)$ -splines are solutions of the 1-D Tikhonov method with constraints  $\Omega_1^2$  and  $\Omega_3^2$ , respectively. For these two general spline families, the error bounds and convergence rates are studied even in presence of noise [162, 163, 212].

The aim of this sub-section is to show roughly that the Tikhonov method with constraint  $\Omega_2^2$  is also bounded and has the convergence property.

Functional of Tikhonov methods with constraints  $\Omega_2^2$  and  $\Omega_3^2$  are respectively denoted by  $Q_2$  and  $Q_3$ ; on finite domain  $x \in [x_1, x_{N_g}]$ , it is easy to show that  $Q_2$  relates to  $Q_3$  as:

$$Q_2 = Q_3 + 2\alpha\lambda \int_{x_1}^{x_{N_g}} f'' f' dx, \quad (7.25)$$

where  $N_g$  denotes number of grid points,  $f' := df/dx$  and  $f'' := d^2f/dx^2$ . Minimizing solution of both  $Q_2$  and  $Q_3$  are L-splines; these splines can be written as bellow on spatial domain  $x_i \leq x \leq x_{i+1} = x_i + \Delta x$ , [200, 202]:

$$f_i(x) = a_i + b_i(x - x_i) + c_i e^{\alpha(x-x_i)} + d_i e^{-\alpha(x-x_i)}, \quad (7.26)$$

where,  $\Delta x$  is an uniform sampling step, and coefficients  $\{a_i, b_i, c_i, d_i\}$  are unknown, real and bounded values. Using Eq. (7.26), functional  $\int f'' f' dx$  on  $x_i \leq x \leq x_i + \Delta x$  leads to:

$$\begin{aligned} \gamma_i &:= \int_{x_i}^{x_i+\Delta x} f'' f' dx \\ &= \frac{1}{2} \alpha e^{-2\alpha\Delta x} (-1 + e^{\alpha\Delta x}) (d_i + c_i e^{\alpha\Delta x}) (2b_i e^{\alpha\Delta x} + (1 + e^{\alpha\Delta x}) (-d_i + c_i e^{\alpha\Delta x})) \alpha. \end{aligned} \quad (7.27)$$

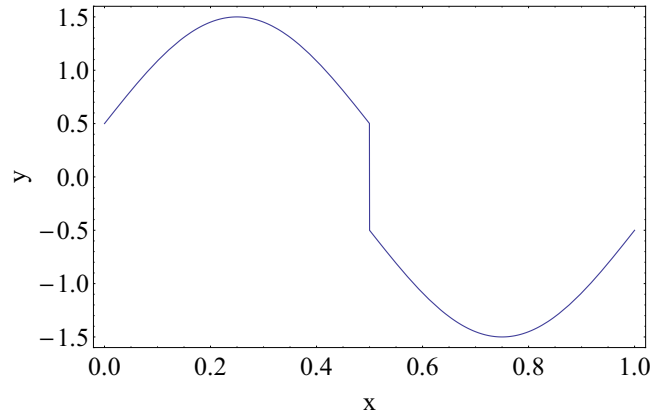
Using the Taylor series,  $\gamma_i$  can be expanded as:

$$\gamma_i = [b_i c_i \alpha^2 + b_i d_i \alpha^2 + c_i^2 \alpha^3 - d_i^2 \alpha^3] \Delta x + \left[ \frac{1}{2} b_i c_i \alpha^3 - \frac{1}{2} b_i d_i \alpha^3 + c_i^2 \alpha^4 + d_i^2 \alpha^4 \right] \Delta x^2 + O[\Delta x^3], \quad (7.28)$$

or:

$$\gamma_i = [b_i (c_i + d_i) + (c_i^2 - d_i^2) \alpha] \alpha^2 \Delta x + O[\Delta x^2]. \quad (7.29)$$




 Figure 7.11: The test function  $y(x)$ .

Since coefficients  $\{a_i, b_i, c_i, d_i\}$  are real and bounded values, so there exists a real and positive value such  $C_i$  where:  $\gamma_i \leq C_i \alpha^2 \Delta x$ . By integration of  $\gamma_i$  on domain  $x \in [x_1, x_{N_g}]$ , the following inequality can be obtained between  $Q_2$  and  $Q_3$ :

$$Q_3 - 2\alpha^3 \lambda \Delta x \left( \sum_{i=1}^{N_g-1} C_i \right) \leq Q_2 \leq Q_3 + 2\alpha^3 \lambda \Delta x \left( \sum_{i=1}^{N_g-1} C_i \right). \quad (7.30)$$

It is clear that,  $\lim_{\Delta x \rightarrow 0} Q_2 = Q_3$ . Hence,  $Q_2$  is bounded and has the convergence property; its convergence rate also approaches to the functional  $Q_3$  as the sampling step approaching zero.

### A benchmark problem

As studied before, the constraint  $\Omega_2^2$  in the Tikhonov method can control the numerical dispersion. In this section, the *error in estimation* and *convergence rate* will numerically be studied by a benchmark problem. The problem has both smooth and discontinuous features. By this example, both the **numerical dispersion** and **Runge** phenomena are investigated. The function is:

$$y(x) = \frac{1}{2} + \text{Sin}(2\pi x) - H(x - 0.5), \quad (7.31)$$

where  $H(x)$  denotes the Heaviside function; the function  $y(x)$  is illustrated in Fig. 7.11.

In Fig. 7.12, the regulated results obtained by constraints  $\Omega_1^2$  (the common Tikhonov method) and  $\Omega_2^2$  are compared with each other. Considered parameters are: 1) for the

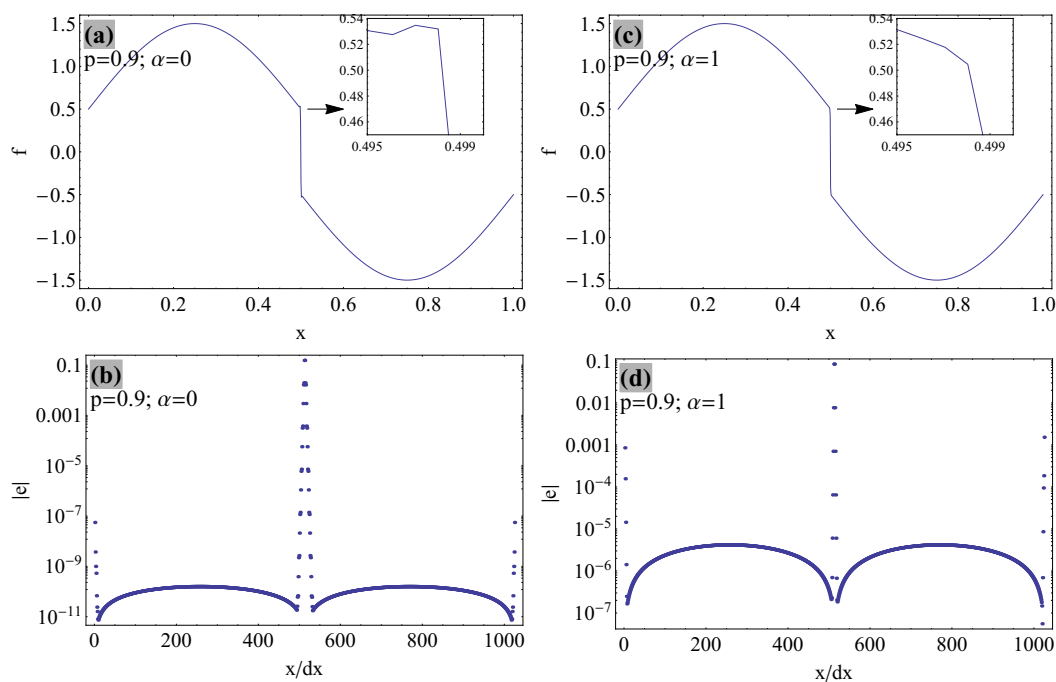


Figure 7.12: Error in estimations, the numerical dispersion and Runge phenomena resulted from the Tikhonov method with different constraints; (a) & (b) regularization with constraint  $\Omega_1^2$ ; (c) & (d) regularization with constraint  $\Omega_2^2$ ;

case  $\Omega_1^2$ :  $p = 0.9$ ; 2) for the case  $\Omega_2^2$ :  $p = 0.9$  &  $\alpha = 1$ . From this figure, it is clear that using the constraint  $\Omega_2^2$  leads to: 1) developing of more localized errors (of large magnitudes) around discontinuous solutions; 2) rising of larger errors in smooth regions; 3) controlling of the numerical dispersion more effectively; 4) increasing of the Runge phenomenon around boundaries. Fig. 7.13 presents convergence rates of different solutions for the two constraints. It is obvious that: 1) convergence rates are near to each other and for all of them, corresponding rates are near to 0.6; 2) with the same  $p$  values, the  $\Omega_2^2$  constraint leads to less numerical errors in  $L_1$  and  $L_2$  senses (even though the errors resulted from the  $\Omega_2^2$  are more than those of  $\Omega_1^2$  in smooth regions). The difference of two estimated errors, resulted from two different constraints  $\Omega_1^2$  and  $\Omega_2^2$ , increases considerably as  $p$  values approach to one.

As mentioned before, the constraint  $\Omega_2^2$  leads to both: 1) larger the Runge phenomenon (around boundaries); 2) larger estimated errors in smooth regions. To cure these drawbacks, one effective way is to use variable tension ( $\alpha$ ) values. The values can be close to zero in smooth regions; it can locally be increased around high gradient or discontinuous zones. Regarding the test function  $y(x)$  (Eq. (7.31)), a Gaussian function for  $\alpha(x)$  is assumed: it is centered around the discontinuity (with spatial location  $x = 0.5$ ). The

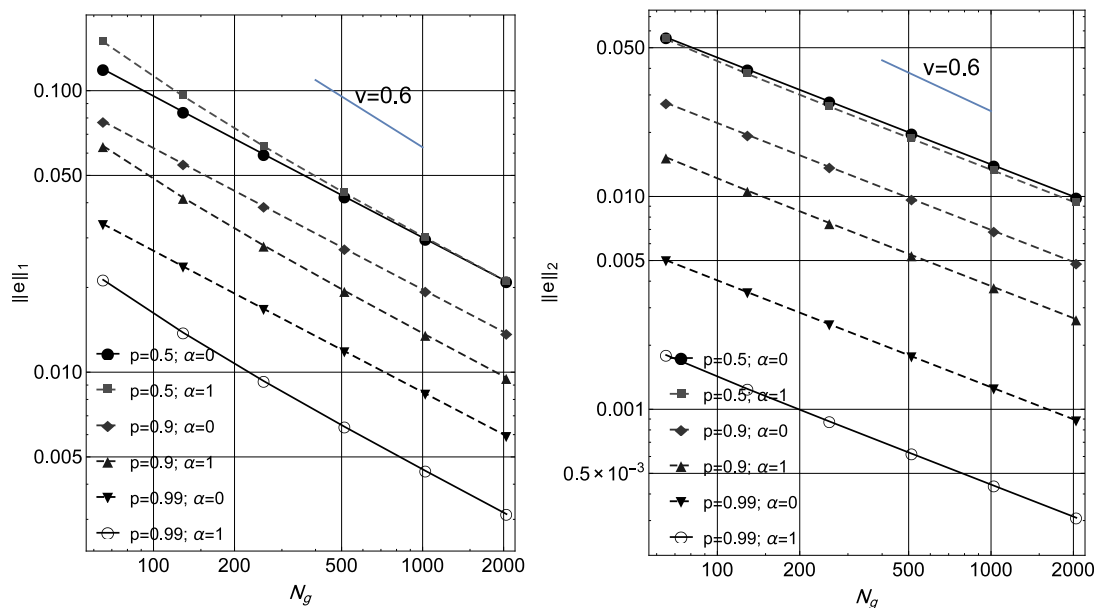


Figure 7.13: Convergence rates for Tikhonov regularization in  $L_1$  and  $L_2$  norms with the constraints  $\Omega_1^2$  and  $\Omega_2^2$  for different  $p$  and  $\alpha$  values.

estimated solution, corresponding estimation errors and convergence rates are presented in Fig. 7.14 for  $p = 0.9$ . Figs. 7.14(a), (b) and (c) present the estimated solution, estimation error, and convergence rate, respectively. The results offer that the non-uniform estimation both improves accuracy of estimations in smooth areas and decreases the Runge phenomenon around boundaries. In this case, the numerical dispersion is successfully controlled, as well.

## 7.7 Conservation in Tikhonov-based smoothing

For sampled data  $\{y_i\}$  with Gaussian noise, Tikhonov-based regularization finds smoothed data  $\{f_i\}$  in such a way that:  $y_i = f_i + \varepsilon_i$ ; where: the set  $\{\varepsilon_i\}$  denotes the noise with zero mean, i.e.:  $\sum_i \varepsilon_i = 0$  [211]. In this case, filtered data remain conservative, since:  $\sum_i y_i = \sum_i f_i$ . In numerical simulations, however, such assumption for noise type is not true and thereby smoothed data do not remain conservative.

Conservative smoothing, however, can be obtained by imposing an extra constraint in the Tikhonov method, as:

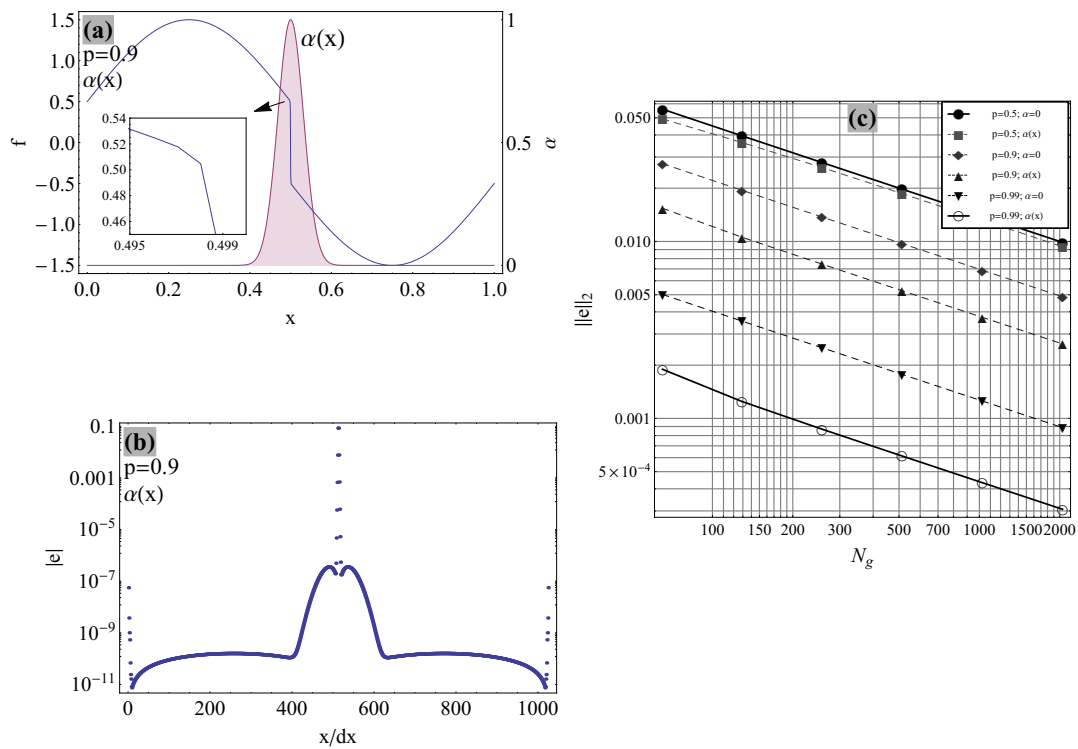


Figure 7.14: Variable  $\alpha$  parameter in Tikhonov-based regularization with the constraint  $\Omega_2^2$ ; a) regularized results ( $f$ ) and  $\alpha(x)$ ; b) estimation errors; c) convergence rates.

$$Q = \sum_{i=1}^n (y_i - f_i)^2 + \lambda \Omega^2(f),$$

$$\text{subjected to: } \frac{1}{n} (\mathbf{I} \cdot \mathbf{y}) = \frac{1}{n} (\mathbf{I} \cdot \mathbf{f}), \quad \left( \text{i.e. : } \sum_i y_i = \sum_i f_i \right)$$
(7.32)

where:  $\mathbf{I}_{1 \times n}$  denotes a single row matrix with unit elements, as:  $\mathbf{I}_{1 \times n} = [1, \dots, 1]_{1 \times n}$ ; vectors  $\mathbf{f}$  and  $\mathbf{y}$  are:  $\mathbf{f} = [f_1, \dots, f_n]^T$  and  $\mathbf{y} = [y_1, \dots, y_n]^T$ . Using the Lagrange multiplier method, the modified functional,  $Q_L$  becomes:

$$Q_L = \sum_{i=1}^n (y_i - f_i)^2 + \lambda \Omega^2(f) + \lambda^L \left\{ \frac{1}{n} (\mathbf{I} \cdot (\mathbf{f} - \mathbf{y}))^T \right\},$$
(7.33)

where  $\lambda^L$  is the Lagrange coefficient. For case  $\Omega^2(f) = \Omega_1^2(f) = \int (f''(x))^2 dx$ , extremes of the functional  $Q_L$  with respect to  $\mathbf{f}$  and  $\lambda^L$  are:

$$\frac{\partial Q_L}{\partial \mathbf{f}} = \mathbf{0} \Rightarrow 2 \underbrace{\left[ \mathbf{B} + \lambda \left\{ (\mathbf{D}^{(2)})^T \cdot \mathbf{B} \cdot \mathbf{D}^{(2)} \right\} \right]}_{\mathbf{\Gamma}} \cdot \mathbf{f} + \frac{\lambda^L}{n} \mathbf{I}^T = 2\mathbf{B} \cdot \mathbf{y},$$
(7.34)

$$\frac{\partial Q_L}{\partial \lambda^L} = 0 \Rightarrow \frac{1}{n} (\mathbf{I} \cdot \mathbf{y}) = \frac{1}{n} (\mathbf{I} \cdot \mathbf{f}).$$

These equations lead to the following linear system:

$$\begin{bmatrix} \mathbf{\Gamma} & \frac{\mathbf{I}^T}{n} \\ \frac{\mathbf{I}}{n} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \lambda^L \end{bmatrix} = \begin{bmatrix} 2\mathbf{B} \cdot \mathbf{y} \\ (\mathbf{I} \cdot \mathbf{y}) / n \end{bmatrix}.$$
(7.35)

Obtaining of such linear systems for other  $\Omega^2(f)$  definitions are straightforward.

**Example:** Linear data  $\{(x_i, y_i)\} = \{(i, i)\}; \quad i = 1, \dots, 7$  is perturbed as:

$$\{(x_i, y_i)\} = \{ \{1, 1\}, \{2, 2.01\}, \{3, 2.97\}, \{4, 3.99\}, \{5, 5.03\}, \{6, 6\}, \{7, 7\} \}.$$

Using the conservative Tikhonov regularization with parameter  $p = 0.9$ , smoothed data will be:

$$\{f_i\} = \{1.005100307, 2.000351, 2.9790275, 3.99291, 5.0189070, 6.005833, 6.9978714\}.$$

In this case, we have:  $\sum_i y_i - \sum_i f_i = 2.07 \times 10^{-12}$  ( $\sum_i y_i = 28$  &  $\sum_i f_i = 28$ ).

For the common Tikhonov regularization (without the constraint of the conservative smoothing), the smoothed result is:

$\{f_i\} = \{1.0054894578, 2.979230914, 3.993112981, 5.01911043, 6.006076, 6.9982605\}$ ;

where:  $\sum_i f_i = 28.0019$ . It is clear the smoothing is not conservative.

The constraint of conservative smoothing can also be imposed by the *penalty* method in the functional  $Q$ , as:

$$Q_P = \sum_{i=1}^n (y_i - f_i)^2 + \lambda \Omega^2(f) + C^P \left\{ \left[ \frac{1}{n} (\mathbf{I} \cdot (\mathbf{f} - \mathbf{y}))^T \right] \cdot \left[ \frac{1}{n} (\mathbf{I} \cdot (\mathbf{f} - \mathbf{y})) \right] \right\}, \quad (7.36)$$

where  $C^P$  is a predefined penalty coefficient.

## 7.8 Global and local smoothing

The common Tikhonov method leads to global smoothing. Two approaches can be followed to have local smoothing (regularization): i) Local regularization (with consistency with surrounding data); ii) Global regularization with variable weights and/or smoothing parameters. Below the first approach will be discussed.

### Local regularization

The main idea is to smooth only some local zones of data. At the boundary points of the local zones, the continuity of both smoothed data and corresponding derivatives (up to some order) should be preserved with surrounding information. To have such local Tikhonov regularization, the concept of smoothing with constraints can be used for local zones of data. These extra constraints guarantee continuity at boundary points. The functional  $Q$  subjected to some new extra constraints can be expressed as:

$$Q = \sum_i (y_i - f_i)^2 + \lambda \times \Omega^2(f), \quad (7.37)$$

subjected to:  $\{\bar{\mathbf{A}}_j \cdot \mathbf{f}^{(j)} = \mathbf{y}_0^{(j)} : j \in \{0, 1, \dots, m\} \quad \& \quad f \in C^m\}$ ,

where:  $(\bar{\mathbf{A}}_j)_{r \times n}$ ,  $(\mathbf{f}^{(j)})_{n \times 1}$ ,  $(\mathbf{y}_0^{(j)})_{r \times 1}$ ,  $\mathbf{f}^{(0)} := \mathbf{f}$ ;  $\mathbf{y}_0^{(0)} = \mathbf{y}_0$ ;  $\mathbf{f}^{(j)} := d^j \mathbf{f} / dx^j$  and  $\mathbf{y}_0^{(j)} := d^j \mathbf{y}_0 / dx^j$ , for  $j \geq 1$ ;  $f_i := f(x_i)$ ;

$\{\mathbf{f} = \{f_i\} : i \in \{1, \dots, n\}\}$  denotes the smooth function needed to be estimated at distinct  $n$  spatial points;  $\mathbf{y}_0^{(j)} = \{(y_0^j)_i\}$  for  $i = \{1, 2, \dots, r\}$  is a vector of data values needed to be interpolated at some predefined points (like the boundary points). For a predefined point  $x_k$  then it is needed to have:  $\{f^{(j)}(x_k) = y_0^{(j)}(x_k) : j \in \{0, 1, \dots, m\}\}$ ;

Matrix  $\bar{\mathbf{A}}_j$  denotes a connection matrix of size  $r \times n$ , where  $n$  and  $r$  are vector lengths of  $\mathbf{f}$  and  $\mathbf{y}_0$ , respectively. At the spatial location  $x_k$  (where  $k \in \{1, 2, \dots, n\}$ ), for  $l^{\text{th}}$  element (where  $k \in \{1, 2, \dots, r\}$ ) of  $\mathbf{y}_0^{(j)}$  (i.e.:  $(\mathbf{y}_0^{(j)}(x_k))_{l,1}$ ), element  $(\bar{\mathbf{A}}_j)_{l,k}$  is equal to one and remaining elements of the  $l^{\text{th}}$  row are zero.

For example let us assume information  $\mathbf{f} = \{f_1, f_2, f_3, f_4\}$  defined at points  $\{x_1, x_2, x_3, x_4\}$ . If we want to locally smooth data over  $x_2$  to  $x_3$ , with only one boundary point at each side of local information (then the left boundary point is  $x_2$  and the right one is  $x_3$ ), then we can conclude that  $n = 4$  and  $r = 2$ . Then the matrix  $\bar{\mathbf{A}}_j$  is

$$\bar{\mathbf{A}}_j = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

and so  $\bar{\mathbf{A}}_j \cdot \mathbf{f} = \{f_2, f_3\} = \mathbf{y}_0$ .

Using the Lagrange multiplier to impose the extra constraints, the new functional  $Q_L$  is:

$$Q_L = \sum_i (y_i - f_i)^2 + \lambda \times \Omega^2(f) + \sum_{j=0}^m \boldsymbol{\lambda}_j^L (\bar{\mathbf{A}}_j \cdot \mathbf{f}^{(j)} - \mathbf{y}_0^{(j)})^T, \quad (7.38)$$

where  $\boldsymbol{\lambda}_j^L$  denotes a Lagrange multiplier vector.

By finding extreme values of  $Q_L$  with respect to  $\mathbf{f}$  and vectors  $\{\boldsymbol{\lambda}_j^L\}$ , one can find both  $\mathbf{f}$  and  $\{\boldsymbol{\lambda}_j^L\}$ . Let assume  $f(x) \in C^2$ , and  $\Omega^2(f) = \Omega_1^2(f) = \int (f''(x))^2 dx$ ; then it is easy to obtain a discretized form of  $Q_L$ , as:

$$Q_L = (\mathbf{y} - \mathbf{f})^T \cdot \mathbf{B} \cdot (\mathbf{y} - \mathbf{f}) + \lambda (\mathbf{D}^{(2)} \cdot \mathbf{f})^T \cdot \mathbf{B} \cdot (\mathbf{D}^{(2)} \cdot \mathbf{f}) + \boldsymbol{\lambda}_0^L (\bar{\mathbf{A}}_0 \cdot \mathbf{f} - \mathbf{y}_0)^T + \boldsymbol{\lambda}_1^L (\bar{\mathbf{A}}_1 \cdot (\mathbf{D}^{(1)} \cdot \mathbf{f}) - \mathbf{y}_0')^T + \boldsymbol{\lambda}_2^L (\bar{\mathbf{A}}_2 \cdot (\mathbf{D}^{(2)} \cdot \mathbf{f}) - \mathbf{y}_0'')^T, \quad (7.39)$$

where  $\mathbf{f}^{(1)}$  and  $\mathbf{f}^{(2)}$  are respectively approximated as,  $\mathbf{D}^{(1)} \cdot \mathbf{f}$  and  $\mathbf{D}^{(2)} \cdot \mathbf{f}$ .

Extreme values of  $Q_L$  can be obtained, as:

$$\begin{aligned}
 \frac{\partial Q_L}{\partial \mathbf{f}} = \mathbf{0} &\Rightarrow 2 \underbrace{\left[ \mathbf{B} + \lambda \left\{ (\mathbf{D}^{(2)})^T \cdot \mathbf{B} \cdot \mathbf{D}^{(2)} \right\} \right]}_{\mathbf{\Gamma}} \cdot \mathbf{f} + \lambda_0^L \cdot \overline{\mathbf{A}}_0^T + \lambda_1^L \cdot (\overline{\mathbf{A}}_1 \cdot \mathbf{D}^{(1)})^T + \lambda_2^L \cdot (\overline{\mathbf{A}}_2 \cdot \mathbf{D}^{(2)})^T = 2\mathbf{B} \cdot \mathbf{y}, \\
 \frac{\partial Q_L}{\partial \lambda_0^L} = \mathbf{0} &\Rightarrow \overline{\mathbf{A}}_0 \cdot \mathbf{f} = \mathbf{y}_0, \\
 \frac{\partial Q_L}{\partial \lambda_1^L} = \mathbf{0} &\Rightarrow (\overline{\mathbf{A}}_1 \cdot \mathbf{D}^{(1)}) \cdot \mathbf{f} = \mathbf{y}_0', \\
 \frac{\partial Q_L}{\partial \lambda_2^L} = \mathbf{0} &\Rightarrow (\overline{\mathbf{A}}_2 \cdot \mathbf{D}^{(2)}) \cdot \mathbf{f} = \mathbf{y}_0''.
 \end{aligned} \tag{7.40}$$

These equations can be represented in a matrix form, as:

$$\begin{pmatrix} \mathbf{\Gamma} & \overline{\mathbf{A}}_0^T & (\overline{\mathbf{A}}_1 \cdot \mathbf{D}^{(1)})^T & (\overline{\mathbf{A}}_2 \cdot \mathbf{D}^{(2)})^T \\ \overline{\mathbf{A}}_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\overline{\mathbf{A}}_1 \cdot \mathbf{D}^{(1)}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\overline{\mathbf{A}}_2 \cdot \mathbf{D}^{(2)}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \cdot \begin{bmatrix} \mathbf{f} \\ \lambda_0^L \\ \lambda_1^L \\ \lambda_2^L \end{bmatrix} = \begin{bmatrix} 2\mathbf{B} \cdot \mathbf{y} \\ \mathbf{y}_0 \\ \mathbf{y}_0' \\ \mathbf{y}_0'' \end{bmatrix}. \tag{7.41}$$

The above left-hand side matrix can be ill-posed, in general. For such cases, the extra constraints can be imposed by the penalty method. In this case, the functional  $Q$  can be modified as:

$$Q_P = \sum_i (y_i - f_i)^2 + \lambda \Omega^2(f) + \sum_{j=0}^m C_j^P (\overline{\mathbf{A}}_j \cdot \mathbf{f}^{(j)} - \mathbf{y}_0^{(j)})^T (\overline{\mathbf{A}}_j \cdot \mathbf{f}^{(j)} - \mathbf{y}_0^{(j)}), \tag{7.42}$$

where  $C_j^P$  denotes predefined penalty coefficients.

For an example, an exponential function is assumed, as:  $y(x) = \text{Exp}(-40(x - 0.5)^2)$ . The aim is the local smoothing on spatial range  $0.30 \leq x \leq 0.98$  in such a way that the smoothed data have  $C^2$  continuity with surrounding data. Edge constraints (on  $y$ ,  $y'$  and  $y''$ ) are only imposed for two boundary end points  $x = 0.30$  and  $x = 0.98$  (more boundary points can be assumed for each edge). Smoothed results are presented in Fig. 7.15, where the smoothing parameter is  $p = 0.7$ . In this figure, discrete values of  $\mathbf{f}'$  and  $\mathbf{f}''$  are computed with the same operators used in the regularization procedure, i.e.:  $\mathbf{f}' = \mathbf{D}^{(1)} \cdot \mathbf{f}$  and  $\mathbf{f}'' = \mathbf{D}^{(2)} \cdot \mathbf{f}$ . It is clear that the local smoothed data and corresponding derivatives have consistency with neighbor data.



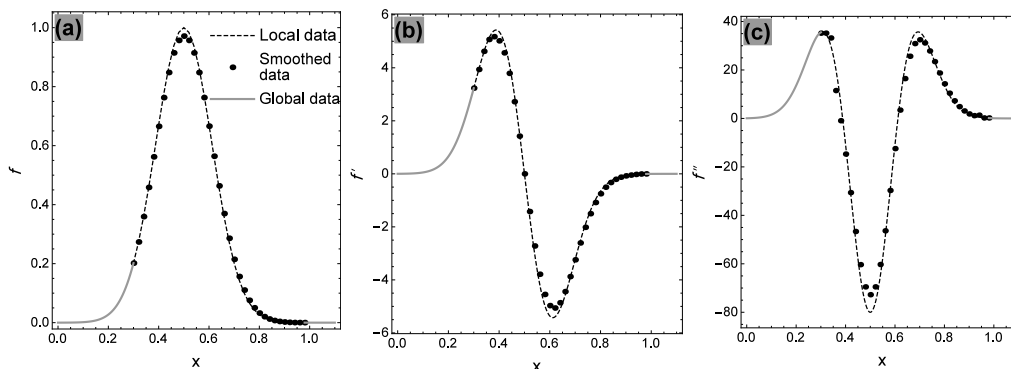


Figure 7.15: Local smoothing with  $C^2$  continuity with surrounding data; value of smoothing parameter is  $p = 0.7$ .

## 7.9 Tikhonov regularization and correspondence with filtering

It is mentioned that the regularization effect is similar to a low-pass filtering [211]. Constraints of the Tikhonov method measure high-frequency components of data; these components are then enforced to approach zero. This filtering feature will be studied in more detail in this section.

Regarding functional  $Q = \int F(x, y, y', y'') dx$ , where  $y := y(x)$ ,  $y' := dy/dx$  and  $y'' := d^2y/dx^2$ , it is easy to show that corresponding Euler-Lagrange differential equation is:  $F_y - \frac{d}{dx}(F_{y'}) + \frac{d^2}{dx^2}(F_{y''}) = 0$ , where:  $F_y := dF/dy$ ,  $F_{y'} := dF/dy'$  and  $F_{y''} := dF/dy''$ .

For the Tikhonov regularization with constraint  $\Omega_2^2(x) = \int (f'' + \alpha f')^2 dx$ , the function  $F$  is:  $F(x, f, f', f'') = (y - f)^2 + \lambda(f'' + \alpha f')^2$ . It is straightforward to show that equivalent Euler-Lagrange equation for constant  $\lambda$  and  $\alpha$  is:

$$-(y - f) - \lambda\alpha^2 f'' + \lambda f^{(iv)} = 0. \quad (7.43)$$

Rewriting this equation in the Fourier space, we have:

$$\mathcal{F}(f) - \mathcal{F}(y) - \lambda\alpha^2(i\omega)^2\mathcal{F}(f) + \lambda(i\omega)^4\mathcal{F}(f) = 0, \quad (7.44)$$

or,

$$\mathcal{F}(f) = \frac{1}{1 + \lambda\alpha^2\omega^2 + \lambda\omega^4} \mathcal{F}(y), \quad (7.45)$$

where:  $\mathcal{F}(f)$  denotes the Fourier transform of the function  $f(x)$  belonging to the Hilbert space  $L_2$  (i.e.:  $\int f(x)^2 dx < \infty$ );  $i^2 = -1$ ; and  $\omega \in [0, \pi]$ . So using the constraint  $\Omega_2^2(x)$  leads to the filter function  $H_2(\omega)$  as:

$$H_2(\omega) = 1 / (1 + \lambda\alpha^2\omega^2 + \lambda\omega^4). \quad (7.46)$$

For constraints  $\Omega_1^2(x)$  and  $\Omega_3^2(x)$ , it is easy to show that corresponding filter functions are:

$$\begin{aligned} H_1(\omega) &= 1 / (1 + \lambda\omega^4), \\ H_3(\omega) &= 1 / (1 + \lambda\alpha\omega^2 + \lambda\omega^4). \end{aligned} \quad (7.47)$$

It is clear that: 1)  $H_j(0) = 1$  and  $H_j(\pi) = 0$  for  $j \in \{1, 2, 3\}$ ; 2) thereby, functions  $\{H_j(\omega)\}$  act as low-pass filters. In fact, they are a type of Butterworth low pass filters [211].

## 7.10 Solution algorithm for stress wave equations

The main idea is to use the regularization stage as a post-processing step in the commonly used finite-difference (or collocation) method with higher-order accuracy. For this reason, consider the solution of a second order hyperbolic system to be  $\mathbf{f}(t)$ . At time step  $t_n$ , the solution procedure can be summarized as the following steps:

1. Approximate spatial derivatives with finite difference discretizations; this can be done by the generalized Padé approximations, compact differencing equations (see Sec. 7.3) [131, 138] or the fast and iterative algorithm proposed by Fornberg [213]. In this work, explicit fourth order spatial accuracy is used for derivative estimations (to have a higher-order solver),
2. Discretize PDEs in the spatial domain and solve resulted semi-discrete systems, i.e., discrete in space and continuous in time. A standard time-stepping method

like the Runge-Kutta method of 4<sup>th</sup> order (see Appendix D) or the explicit six-stage method [131] can be used to solve the resulted ODEs at time  $t = t_n$  to obtain solution at the next time step  $t = t_{n+1}$ ,

3. Denoise spurious oscillations (this step can be done directly on non-uniform grid points, as well); for this purpose the Tikhonov method with constraint  $\Omega_2^2$  is recommended. Either variable or constant  $p$  or  $\alpha$  values can be considered. In this work, constant values are used. The recommended values for  $p$  and  $\alpha$  are:  $0.9 \leq p \leq 0.99$ , and  $\alpha \approx 1$ . For the Tikhonov method other extra conditions can also be considered; such as: conservative (Sec. 7.7) or local (Sec. 7.8) smoothing.
4. Go back to step 1,

In practice, to have a cost-effective computation, the solution is not denoised at each time step. This can be done after some time steps (depending on the wave velocity), for example following ten time steps.

## 7.11 Numerical examples

In this section, one 1-D and several 2-D examples are presented to show effectiveness of the proposed approach. The 2-D examples are: 1) a membrane subjected to an imposed initial discontinuous deformation (to simulate propagating discontinuous fronts); 2) an infinite-periodic domain containing sharp and localized variation of physical property (a narrow fluid-filled crack with finite dimension); 3) wave propagation in a medium with several fluid-filled cavities with stochastic locations (to have stochastic-like simulations) [214, 215].

In simulations, for Tikhonov methods it is assumed: 1) they are not enhanced to be conservative; 2) the global smoothing approach is used.

### Example 1:

In this example, the wave propagation in a 1-D linear bar with the box-shaped initial imposed deformation is re-simulated (see Sec. 7.2) with other commonly-used approaches. They are basically developed to remedy discontinuity effects in elasto-dynamic problems. The methods are:

1. The finite element method with linear spatial shape functions using numerically dissipative time integration scheme ( $\alpha$ -NDTI) [130]. The assumed parameters are:  $\gamma = 0.6$ ;  $\beta = 0.25(0.5 + \gamma)^2$ ;  $\alpha = -0.0683$ ;  $dt = 0.003$ ; number of elements is  $N_e = 256$ .

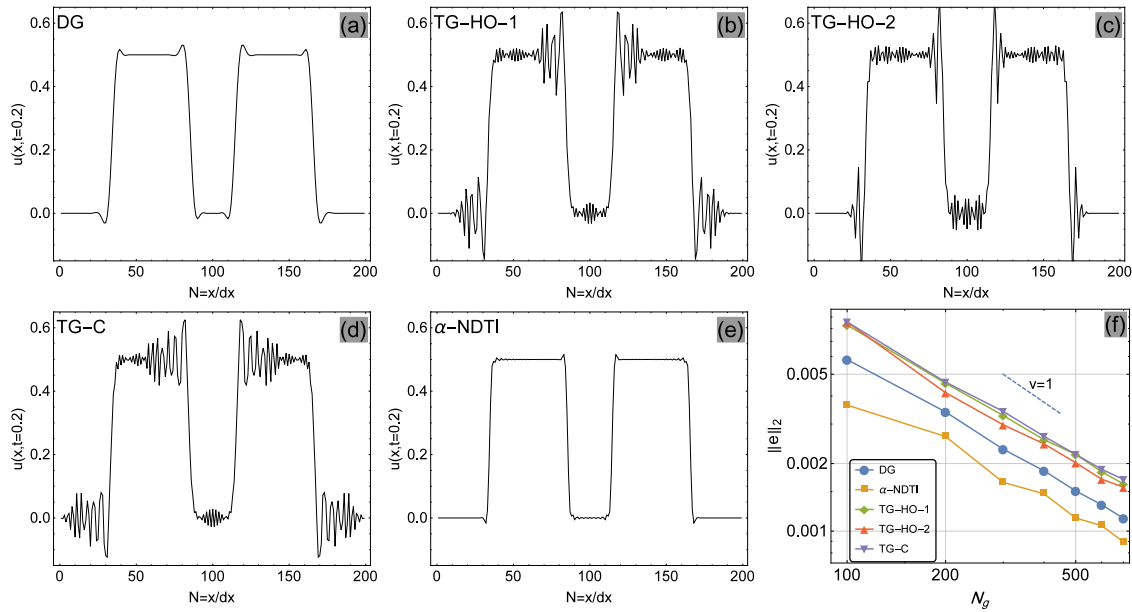


Figure 7.16: Numerical results for wave propagation in a bar due to the box-shaped imposed initial condition at  $t = 0.2$ .

2. Common Taylor-Galerkin method (TG-C) (second order) and higher order one (TG-HO) (third order) [216]; for the TG-C case, assumed parameters are:  $\gamma = 0.5$  and  $\alpha = 0.5$ ;  $dt = 0.002$ ;  $N_e = 256$ . For the TG-HO method, two set of parameters are considered as: i) for TG-HO-1:  $\gamma = 0.5$ ;  $\alpha = 0.5$ ;  $N_e = 256$ ;  $dt = 0.002$ ; ii) for TG-HO-2:  $\gamma = 0.5$ ;  $\alpha = 1.5$ ;  $N_e = 256$ ;  $dt = 0.002$ .
3. Discontinuous time Galerkin method (DTG) [217] (piece-wise linear approximations in the time and linear approximations in the space). The considered parameters are:  $dt = 0.01$  and  $N_e = 256$ .

In each scheme, the considered parameters are the suggested values in the corresponding references, to have a good estimation of the exact solution. In Appendix D, the above-mentioned methods are briefly reviewed. For numerical simulations, properties of the bar are:  $E = 1$  (module of elasticity),  $A = 0.01$  (cross section area), and  $\rho = 1$  (density of the bar). The results are presented in Fig. 7.16. Although these methods lead to stable solutions, they can not effectively control the numerical dispersion.

As mentioned in the Introduction section, the idea of high-resolution schemes is recently extended to second order hyperbolic systems [143]. Bellow the aim is to compare such results with those of our proposed method. For this reason, this example is resolved on spatial domain  $x \in [-1, 1]$  with new initial conditions:  $u_0(x) = \text{Unitbox}(2x)$  and  $v_0(x) = 0$ . For the Tikhonov method it is assumed: 1) smoothing is done at each

time step; 2) The functional  $\Omega_2^2$  is used as the constraint; 3) smoothing parameters are:  $p = 0.99$  &  $\alpha = 1$ .

Numerical results are presented in Fig. 7.17 at  $t = 0.2$ ; symbols  $P-4$ ,  $UW2$ ,  $UW4$ , and  $HR2$  denote the 4<sup>th</sup> order Padé, second order upwind, fourth order upwind and second-order high resolution approximations, respectively. For details of  $UW2$ ,  $UW4$ , and  $HR2$  methods, please see [143]. The results offer that: 1) smoothed  $P-4$  method (using the Tikhonov smoothing method) prevents forming of spurious oscillations in both  $u$  and  $v$  solutions; 2) numerical dissipation of smoothed  $P-4$  is less than those of  $HR2$  and coincides with those of  $UW2$  (Fig. 7.17(b)); 3) numerical dispersion in smoothed  $P-4$  is the least one (Fig. 7.17(a,c)). The results of  $UW2$ ,  $UW4$ , and  $HR2$  are from [143].

**Example 2:** Consider vibration of a rectangular membrane with four fixed sides subjected to an initial imposed displacement; the governing equations are:

$$\begin{aligned} PDE : \quad & c^2 \left( \partial^2 u / \partial x^2 + \partial^2 u / \partial z^2 \right) = \partial^2 u / \partial t^2; \quad \Omega \in [0, 1] \times [0, 1], \\ ICs : \quad & u(x, y, t = 0) = U(x, y) \text{ \& } \partial u / \partial t(x, y, t = 0) = 0, \\ BCs : \quad & u(0, y, t) = u(1, y, t) = u(x, 0, t) = u(x, 1, t) = 0. \end{aligned} \tag{7.48}$$

The initial condition has discontinuities; it is:  $U(x, y) = H \left( \left( \frac{x-0.5}{0.2} \right)^2 + \left( \frac{y-0.5}{0.2} \right)^2 \right)$  (where  $H$  denotes the Heaviside function). The finite difference scheme is used for spatial discretization; the second derivative is approximated with the fourth-order explicit central scheme; number of grid points is:  $(2^8 + 1) \times (2^8 + 1)$ . The time-integration is done by the Runge-Kutta scheme of 4<sup>th</sup> order (see Appendix D). For the post-processing stage, the Tikhonov method with three different constraints are considered: 1)  $\Omega_1^2$  with the parameter  $p = 0.99$ ; 2) the constraint  $\Omega_2^2$  with parameters  $p = 0.99$  &  $\alpha = 1$ ; 3)  $p = 1$ , i.e.: without considering any regularization. Other parameters are assumed to be:  $c = 1$  (wave velocity);  $dt = 0.0015$  (the time integration step). The denoising (regularization) procedure is repeated after each five time steps. The IC ( $U(x, y)$ ) and numerical results are presented in Fig. 7.18 at  $t = 0.24$ . According to the results, it is clear that: 1) without regularization, the solution becomes unstable; 2) small amount of the regularization leads to stable solutions; 3) results obtained by  $\Omega_2^2$  constraint can control more effectively spurious oscillations around propagating discontinuous fronts.

The numerical result obtained by the post-processing stage with constraint  $\Omega_2^2$  and the solution obtained from the modal analysis will also be compared. For the square membrane with dimensions  $(x, y) \in [a \times b]$ , and with four fixed boundaries, the mode

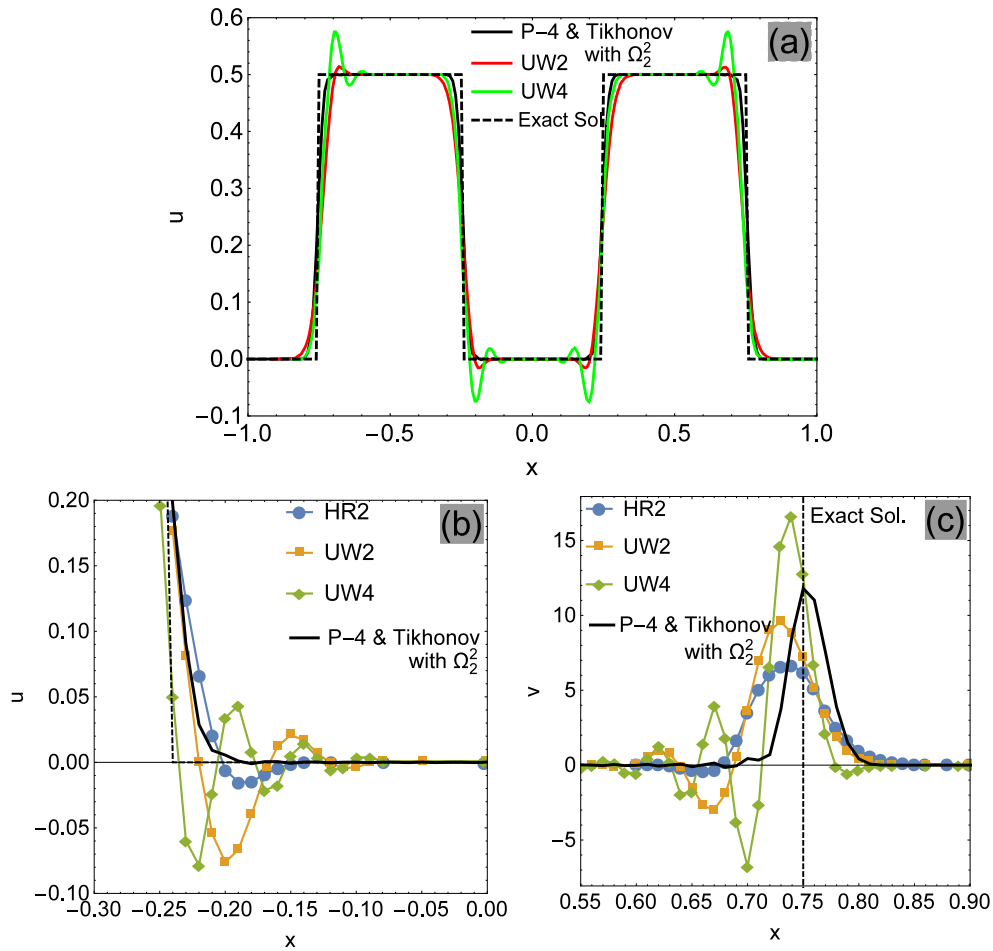


Figure 7.17: Direct simulation of the second order scalar wave propagation problem at  $t = 0.2$ ; symbols  $P-4$ ,  $UW2$ ,  $UW4$  and  $HR2$  denote the 4<sup>th</sup> order Padé, second order upwind, fourth order upwind and second-order high resolution approximations, respectively. The results of  $UW2$ ,  $UW4$  and  $HR2$  are from [143].

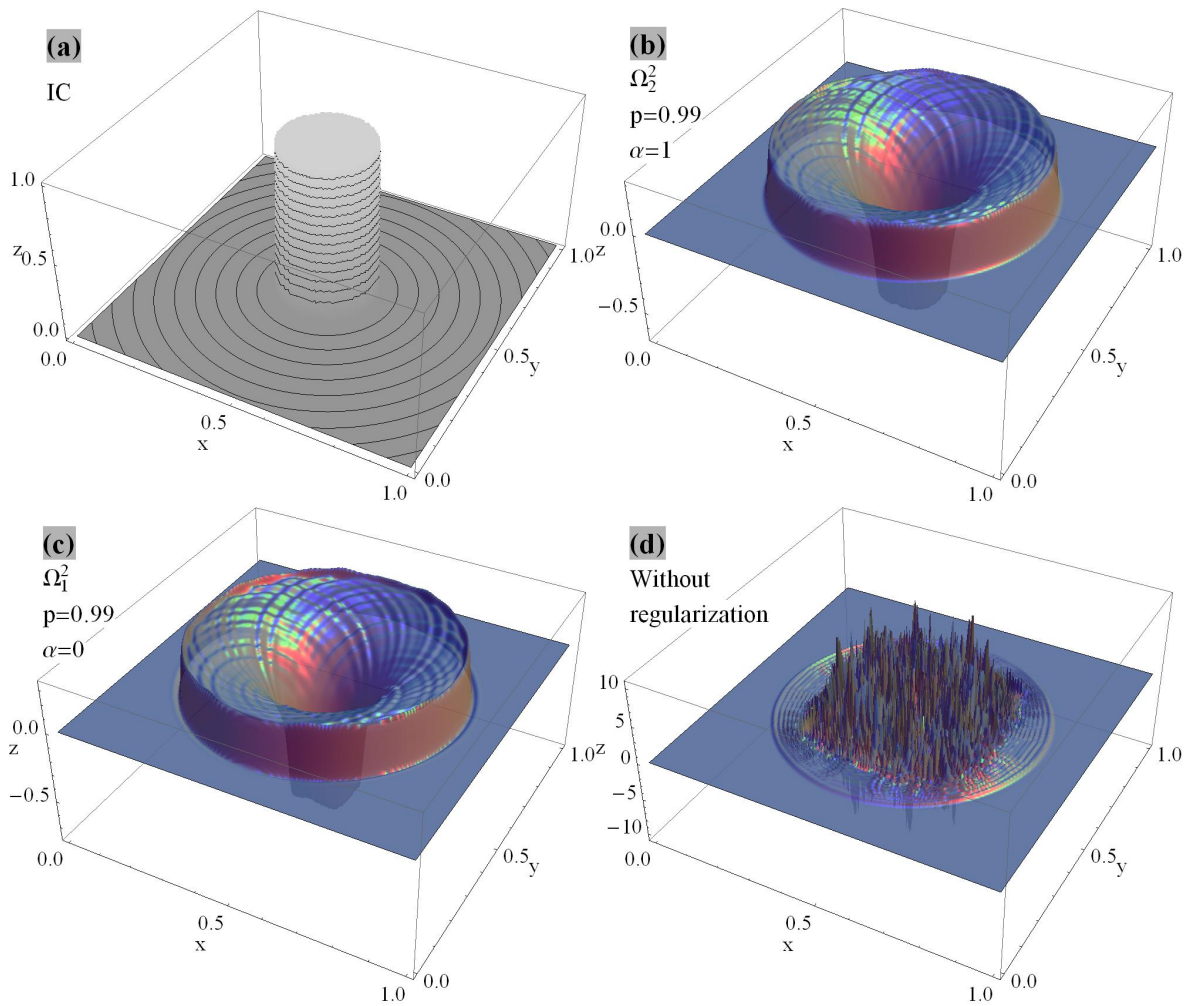


Figure 7.18: Numerical results with different assumptions of regularization constraints at  $t = 0.24$ ; a) the IC condition ( $U(x, y)$ ); b) the solution based on using the constraint  $\Omega_2^2$ ; c) the solution based on using the constraint  $\Omega_1^2$ ; d) the solution without any regularization.

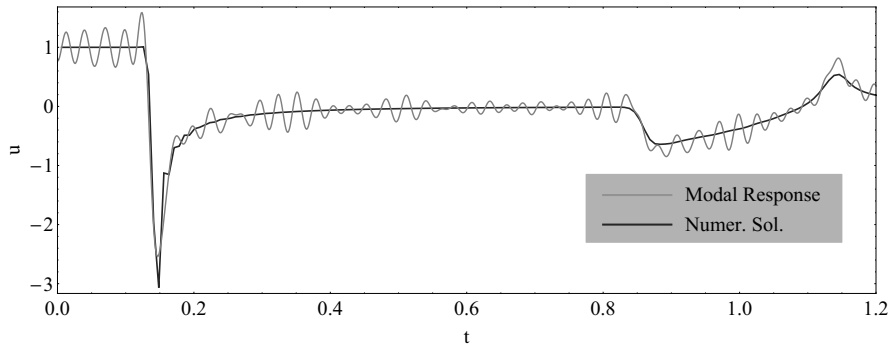


Figure 7.19: Comparison of the numerical method having the post-processing stage with constraint  $\Omega_2^2$ , and the modal analysis based solution (with  $60 \times 60$  first modes) at spatial position  $(x, y) = (0.5, 0.5)$ .

shapes are:

$$M_{n,m}(x, y) = \sin\left(\frac{m\pi x}{a}\right) \times \sin\left(\frac{n\pi y}{b}\right). \quad (7.49)$$

The results are compared in Fig. 7.19 at spatial location  $(x, y) = (0.5, 0.5)$  for time duration  $t \in [0, 1.2]$ . For the modal analysis, it is assumed:  $(n, m) = (60, 60)$  (i.e.,  $60 \times 60$  first modes are considered). Due to discontinuities, the numerical dispersion exists in the modal analysis. Therefore, a stable and oscillation free numerical result can be obtained by considering the post-processing stage (the constraint is:  $\Omega_2^2$ ).

**Example 3:**

In this example, the stress wave propagation problem is studied in a medium including high and abrupt variations in its physical parameters. In such systems, numerical methods which can not handle (semi) discontinuous solutions (caused by the material variation) have stability and accuracy problems around contact zones (e.g., the fluid-solid contact zone). Due to this material variation, speeds of elastic waves are largely different around the contact zone. The incident waves, P or S, can be reflected from interface in the form of P and S waves. The incident P wave is reflected as P (denotes by PPr) and S (shown by PSr) waves; for the incident SV wave the reflected P and S waves are shown by SPr and SSr, respectively. If the second material is water, since only P waves can be transmitted to the fluid layer, the transmitted P wave due to incident P and S waves are denoted by PPt and SPt, respectively. Another phenomenon due to existence of a sharp corner is the diffraction. The P wave is diffracted from crack edges into the solid medium as diffracted P (PPd) and diffracted S (PSd) waves.



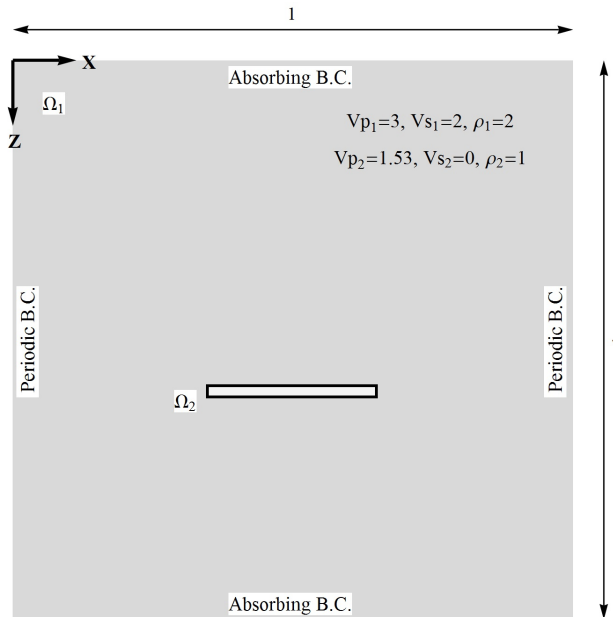


Figure 7.20: Schematic shape of an infinite-periodic medium including a fluid-filled narrow crack.

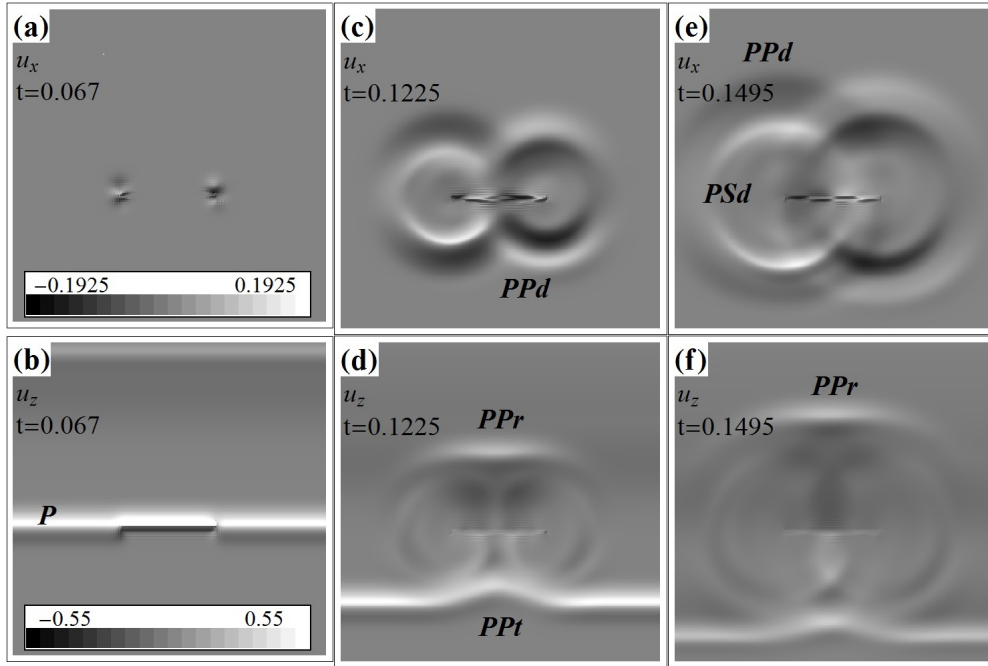
It is assumed that the medium has infinite-periodic boundaries, and it contains a narrow fluid-filled crack of finite length. Due to this crack, each P or S (SV) incident waves are reflected and transmitted. Also due to edge of the crack, the diffraction phenomenon will be happened.

The schematic shape of the medium along with the description of the crack configuration is illustrated in Fig. 7.20. There, position of periodic and absorbing boundary conditions are also illustrated.

Here, to consider infinite boundaries the absorbing boundary concept is used. This boundary is explicitly considered in the governing P-SV equations, as:

$$\begin{aligned} \{(\lambda + 2\mu) u_{x,xx} + \mu u_{x,zz}\} + \{(\lambda + \mu) u_{z,xz}\} &= \rho \{u_{x,tt} + Q(x, z)u_{x,t}\}, \\ \{(\lambda + 2\mu) u_{z,zz} + \mu u_{z,xx}\} + \{(\lambda + \mu) u_{x,xz}\} &= \rho \{u_{z,tt} + Q(x, z)u_{z,t}\}. \end{aligned} \quad (7.50)$$

The absorbing boundaries are commonly used for simulation of infinite boundaries. In this system, absorbing boundary condition is considered explicitly; or equivalently the wave equations are modified by damping term  $Q(x, z)$ . This acts as an attenuation factor. This factor is nearly zero in the computation domain and increases gradually when approaches to the artificial boundaries. This causes incoming waves towards these


 Figure 7.21: Snapshots of solutions  $u_x$  and  $u_z$ .

kind of boundaries to diminish gradually [218]. In general, there is not any absorbing boundary that could absorb all of the incoming energies, so, some small long-period reflections always remain. In this problem, the attenuation function is:

$$Q(x, z) = a_z (\text{Exp}(b_z \cdot z^2) + \text{Exp}(b_z \cdot (1 - z)^2)), \text{ where } a_z = 120 \text{ and } b_z = -100.$$

In the numerical simulation, it is assumed that: 1) the post-processing stage is done by the Tikhonov method with constraint  $\Omega_2^2$ ; 2) the incident wave is a plane P wave, produced by an initial imposed deformation with equation:  $u_z(x, z, t = 0) = \text{Exp}(-700(z - 0.25)^2)$ ; 3) the regularization is performed after each five time steps with parameters  $p = 0.97$  and  $\alpha = 1$ ; 4) time step is:  $dt = 0.001$ ; 5) number of grid points in each direction is:  $2^8 + 1$ ; 5) time-integration is done by the Runge-Kutta 4<sup>th</sup> order method.

The snapshots of  $u_x$  and  $u_z$  are illustrated in Fig. 7.21. There, the reflected, transmitted and diffracted waves are shown. The zoomed-in solution  $u_z$  at incident time,  $t = 0.067$  is illustrated in Fig. 7.22; it is clear that formation of spurious oscillations are prevented due to discontinuous solutions formed around contact zone.

#### Example 4:

Our concern in this example is controlling of artificial dispersion developing in stochastic like solutions. So, a wave propagation problem in a stochastic medium will be studied. In such systems, propagating fronts can not develop due to the diffraction phenomenon.

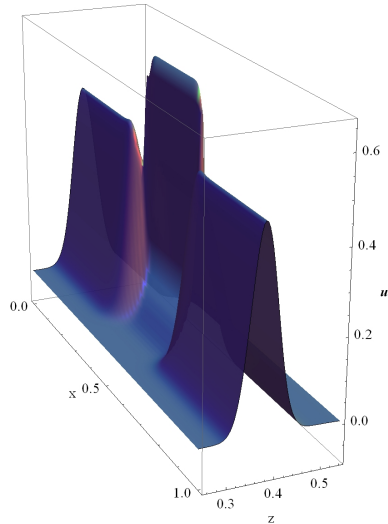


Figure 7.22: Snapshot of solutions  $u_z$  at  $t = 0.067$ .

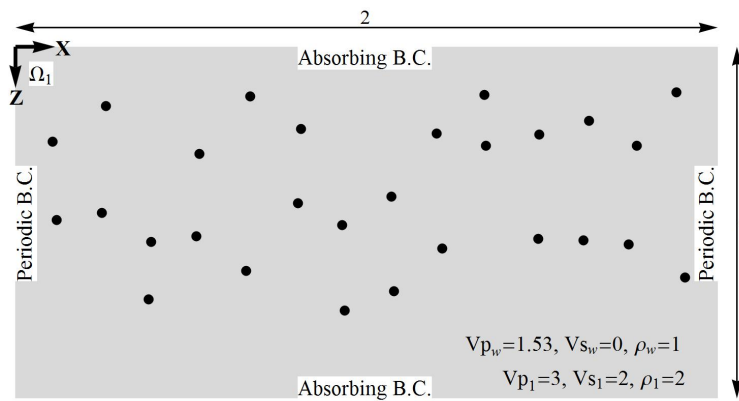


Figure 7.23: A homogeneous medium with stochastic fluid-filled small cavities.

The challenging problem is developing of dominate non-physical dispersive waves.

One practical approach for stochastic media simulation is to consider several stochastic fluid-filled small cavities in homogeneous media [214, 215]. The considered medium is illustrated in Fig. 7.23, where the black disks show the cavities. It is assumed that: 1) the medium has infinite-periodic boundary conditions; 2) a plane P wave propagates in the  $z$  direction.

In the numerical simulations, we have: 1) spatial discretization is done with the explicit central finite difference method of fourth order accuracy; 2) the time integration method is the Runge-Kutta scheme of fourth order with time step  $dt = 0.001$ ; 3) post-processing stage is performed using the Tikhonov method with constraint  $\Omega_2^2$  and parameters  $p = 0.99$  and  $\alpha = 1$ ; and 4) the regularization repeats after each five time steps; 4) number

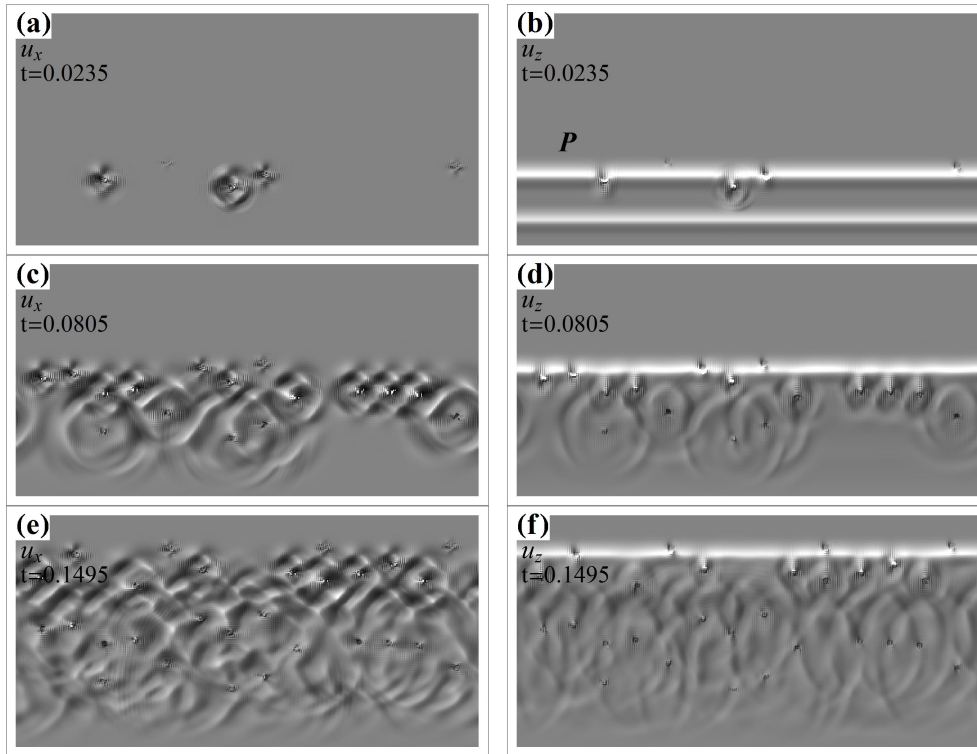


Figure 7.24: Formation of stochastic-like solutions  $u_x$  and  $u_z$  from the fluid-filled cavities due to diffraction of the incident plane  $P$  wave.

of grid points is:  $(2^9 + 1, 2^8 + 1)$ .

Snapshots of results for  $u_x$  and  $u_z$  are presented at Fig. 7.24 for different times. According to the  $u_x$  solutions, it is clear that the diffracted waves propagate and propagating fronts can not be formed in the domain due to stochastic nature of solutions.

Fig. 7.25 illustrates an investigation on the regularization effects. It compares two solutions of  $u_x$  with and without the regularization-stage (with the constraint  $\Omega_2^2$ ) at  $t = 0.0805$ . It is shown that even marginal regularization can prevent occurrence of the artificial dispersion (commonly occurred in stochastic-like numerical solutions).

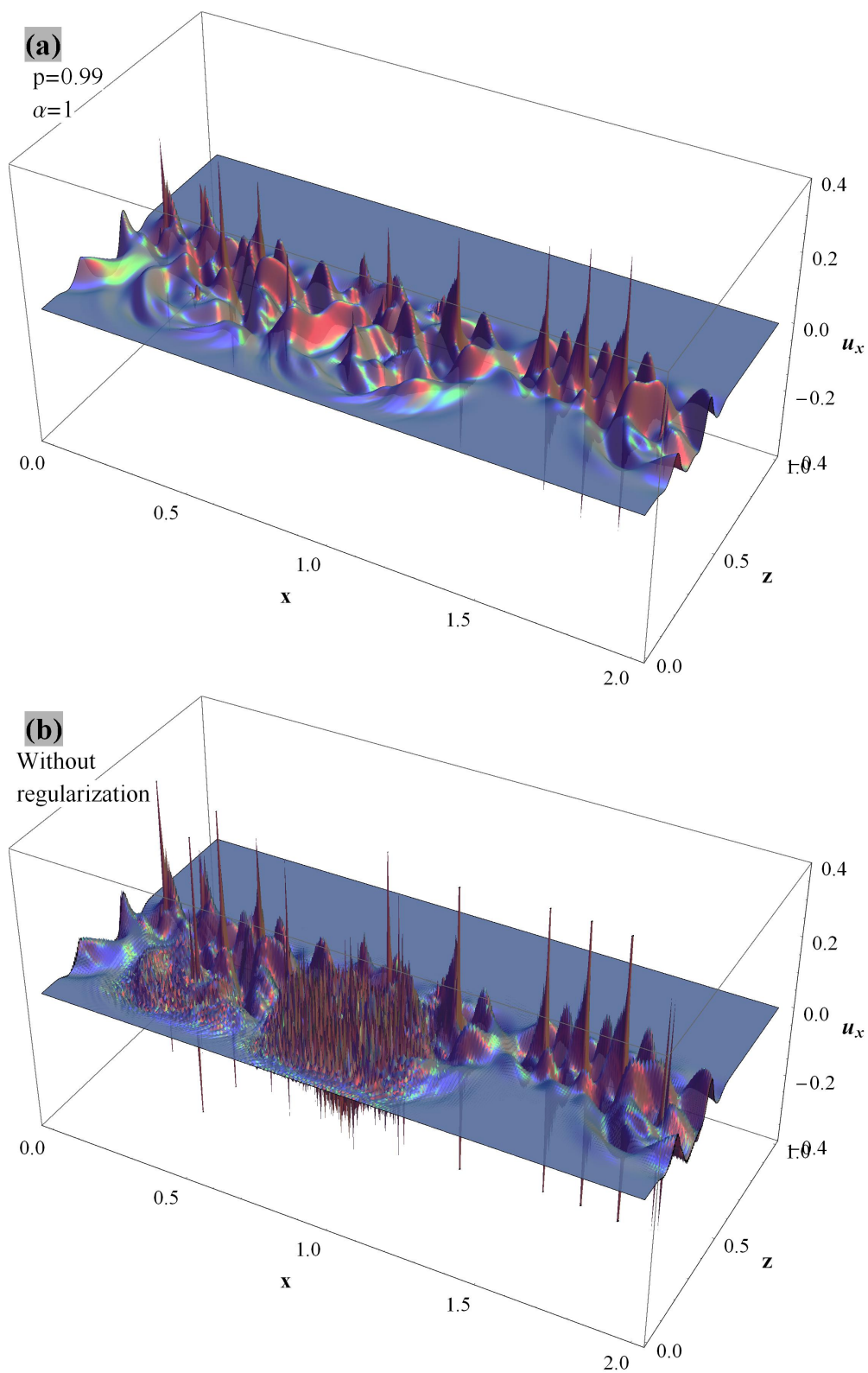


Figure 7.25: Snapshots of the P-SV wave propagation in a medium having fluid-filled cavities with stochastic locations; a) the solution with the regularization step with the constraint  $\Omega_2^2$ ; b) the solution without the regularization step.

# Chapter 8

## Conclusion and future works

### 8.1 Conclusion

In this study, numerical simulation of first and second order wave (hyperbolic) equations are studied. For each one, the following results have obtained.

#### 8.1.1 First order nonlinear hyperbolic (wave) systems

In this study, a wavelet-based adaptation procedure is properly integrated with central, central-upwind, PPM and CWENO high resolution schemes for simulation of first order hyperbolic PDEs. It is shown that central high-resolution schemes become unstable on non-uniform cells even those have gradual variations of grid densities. This is because, the NVSF criterion is not satisfied by central schemes. Since their slope/flux limiters do not remain TVD/TVB on irregular cell-centered cells. Two key ideas are followed to remedy the instability problem: 1) Replacing local irregular grids (resulted from the wavelet-based adaptation) with abrupt changing with grids having gradual variations (replacing an ill-posed problem with a nearly well-posed one); 2) Updating nonlinear stability conditions and restudying the performance of slope limiters on irregular cells. The grid modification stage is also done in the framework of multiresolution analysis; there, proper pattern of density variation of adapted grids is provided and its performance is studied both theoretically and numerically. The TVD, TVB and UNO conditions are reviewed and provided for non-uniform cells. It is shown that on cell-centered non-uniform cells: i) For numerical stability, slope limiter definitions may be modified; ii) The long-term stability feature does not meet. To overcome these challenging problems, non-cell-centered cells are used. These cells act as transmitting cells (cells that connecting surrounding cells locally having uniform sizes); they are located only

at transmitting zones. In this case, we show that: 1) Some of the common limiters can be used without modification (the MINMOD and the Generalized MINMOD limiters); 2) It is possible to modify other limiters in a way that they also preserve the stability conditions. The TVD conditions are derived for both cell-centered and non-cell-centered cells. The proper cell centers in the transmitting cells are derived based on the stability conditions.

The local truncation errors for 1-D and 2-D problems are provided on non-uniform cells for convergence studying. Also the concept of the numerical entropy production is investigated for uniqueness insurance. These two concepts have also been used as possible criteria for grid adaptation. In this regard, the performance of these concepts are compared with the wavelet-based algorithm. Numerically, it is shown that: 1) The numerical entropy production can not detect some phenomena such as contact discontinuities; 2) These concepts can also have sufficient values in rarefaction domains. This can lead to unnecessary concentration of grid points in such regions (this can also be seen in [31, 32]); 3) It seems that the performance of the local truncation errors as a detector may be better than the numerical entropy production; 4) It seems that wavelets can properly detect all of the shock waves, rarefaction regions, and contact discontinuities.

Non-linear hyperbolic systems with non-convex fluxes are studied, as well. In these systems, for capturing physical responses, method-adaptation is also performed again by the multiresolution analysis. It is shown that for both physical and non-physical solutions, local truncation errors of weak solutions have the convergence feature. And so, using of the method-adaptation is essential (e.g., using of different slope limiter in different zones in the spatial domain).

In brief, the studied schemes, slope/flux limiters and stability conditions are as follows:

1. Providing the fully-discrete and semi-discrete formulations of second-order and third-order central high resolution schemes over non-uniform centered and non-centered cells (based on polynomial based reconstruction on each cell),
2. Updating the TVD stability conditions for centered and non-centered cells,
3. For second-order central schemes, studying/providing second order limiters with TVD, TVB and UNO feature,
4. Updating the definition of two non-linear scaling limiters to preserve the conservation, the monotonicity and the same shape properties,

5. Proposing two PPM central high resolution schemes with corresponding limiters,
6. Studying CWENO schemes with three and five point stencils and corresponding accuracy-order on non-uniform cells (that is the third-order and fourth-order schemes).

### 8.1.2 Second order nonlinear wave (hyperbolic) systems

We have presented an approach including a post-processing stage to control the numerical dispersion around discontinuities for numerical simulation of the mechanical wave propagation problems (known as the second order hyperbolic PDEs). Several compact high-order finite-difference methods are successfully integrated with the Tikhonov method as a post-processor. This is performed to control spurious oscillations formed around discontinuities due to the numerical dispersion. For this controlling, the crucial point is a proper choice of a constraint for the Tikhonov-based regularization. Different types of regularization with different constraints are studied. And finally a proper constraint with some possible extensions is advised.

Regarding the Tikhonov method, two general approaches are studied: regularization with and without a model. A model contains extra almost local information (such as discontinuity locations) to improve regulated results. However, it is shown (by numerical studies and GSVD decomposition method), the model-based results are sensitive to the model; if the model is exact, the regulated solution can properly handle discontinuity effects, otherwise results are sensitive and the numerical dispersion affects them. It is shown that the constraint resulted from proper combination of smoothness and the tension concept can properly control the numerical dispersion around discontinuities. This definition is different from the classical one, described as:  $\Omega_3^2 = \int f''(x)^2 + \alpha^2 f'(x)^2 dx$ . The recently proposed definition is:  $\Omega_2^2 = \int (f''(x) + \alpha f'(x))^2 dx$  [200]. For this constraint, corresponding error bounds and convergence rates are qualitatively studied in this work. It is shown for the Tikhonov method with  $\Omega_2^2$  constraint, the regularization concept can effectively control the numerical dispersion around discontinuities. To improve smoothing performance resulted from the Tikhonov methods, it is shown how to add some other favorable features, such as conservative and local smoothing. To clarify smoothing effects of the Tikhonov method, its relationship with the filtering concept is also studied.

The proposed method can easily be extended to higher dimensions and systems of PDEs; since it is totally based on 1-D algorithms. Finally to confirm efficiency of the



proposed method, some 1-D and 2-D examples are presented. In one 1-D example, the performance is compared with those of other commonly used methods developed for stress wave propagation problems. This example confirms that the performance is good and acceptable. According to another 1-D benchmark, by using spatially varying regularization parameters (the adaptive regularization concept), accuracy of solutions can be improved while the Runge phenomenon can be controlled. This can be achieved, for example, by employing non-constant weights, smoothing ( $p$ ) or tension ( $\alpha$ ) coefficients in the Tikhonov method.

Some 2-D examples containing propagating discontinuous fronts are studied. Also, wave propagation in media with localized sharp transition of material properties are considered; e.g., a medium with fluid-filled crack of finite length and a medium with fluid-filled cavities with stochastic spatial locations. The results confirm that existence of even marginal regularization effects can considerably improve stability and accuracy of solutions.

Finally, it should be mentioned that one drawback of the regularization approach is proper estimation of corresponding parameters (here  $p$  and  $\alpha$ ); it seems that the trial and error method is useful to adjust proper values.

## 8.2 Future works

### 8.2.1 First order waves

1. In this study, polynomial-based reconstructions are used in formulation of central, central-upwind and CWENO schemes. Some other possible alternatives could be: Hermite-based reconstructions [219, 220]; exponential polynomials [221]; and spectral-based reconstructions [222],
2. Central/central-upwind high resolution schemes on unstructured finite volumes were developed, e.g. [223, 224]. For unstructured meshing, an important point is proper detection of edges/discontinuities on piecewise information (for proper triangulation of computing domains to preserve sharp edges). Some powerful multiscale/multiresolution based algorithms can be integrated by the central schemes. A powerful multiscale method is presented in [225], and another multiresolution-based one is provided in [226],
3. The concept of constraint minimization to preserve TVD solutions in nonlinear

hyperbolic PDEs is another interesting approach [227]. Its performance on multi-dimensional and irregular grids can be studied.

### **8.2.2 Second order waves**

1. So far, in our study, the Tikhonov method is used on structured grids for denosing of spurious oscillations; this concept can be implemented on unstructured meshes by the multivariate splines, the approach mentioned in [228],
2. The Tikhonov method leads to the smoothing splines or splines-in-tension. For data with high frequency contents, spectral-like accuracy of approximations is favorable by the Tikhonov method. In this regard, the idea of the trigonometric splines may be used [229, 230].

# Appendix A

## Central schemes on uniform grids

### A.1 Central High Resolution Schemes

In this section, the concept of central high resolution schemes will be studied on uniform cells. The considered system is the hyperbolic equation of the general form:

$$u_t + f(u)_x = 0, \quad (\text{A.1})$$

where  $u := u(x, t)$  is the state variable and  $f(u)$  denotes the flux. The discretization will be done by the finite volume method. For this, the spatio-temporal volume  $[x_{i-1/2}, x_{i+1/2}] \times [t^n, t^{n+1}]$  is assumed, where  $\Delta x = x_{i+1/2} - x_{i-1/2}$  and  $\Delta t = t^{n+1} - t^n$ . Integration of this equation over this volume leads to:

$$\int_{t^n}^{t^{n+1}} \int_{x_{i-1/2}}^{x_{i+1/2}} [u_t + f(u)_x] dx dt = 0, \quad (\text{A.2})$$

or:

$$\int_{x_{i-1/2}}^{x_{i+1/2}} [u(x, t^{n+1}) - u(x, t^n)] dx + \int_{t^n}^{t^{n+1}} [f(u(x_{i+1/2}, t)) - f(u(x_{i-1/2}, t))] dt = 0. \quad (\text{A.3})$$

On the other hand:  $\int_{x_{i-1/2}}^{x_{i+1/2}} u(x, t^n) dx = u_i^n \Delta x$  and  $\int_{t^n}^{t^{n+1}} f(u(x_{i+1/2}, t)) dt = F_{i+1/2}^{n+1/2} \Delta t$ , where  $u_i^n$  is the average of  $u(x, t)$  on the domain  $[x_{i-1/2}, x_{i+1/2}]$ , and  $F_{i+1/2}^{n+1/2}$  denotes average of the flux through  $t^n$  to  $t^{n+1}$  at  $x_{i+1/2}$ ;  $F_{i+1/2}$  is also known as the numerical flux. By these relationships, Eq. (A.3) can be written as:

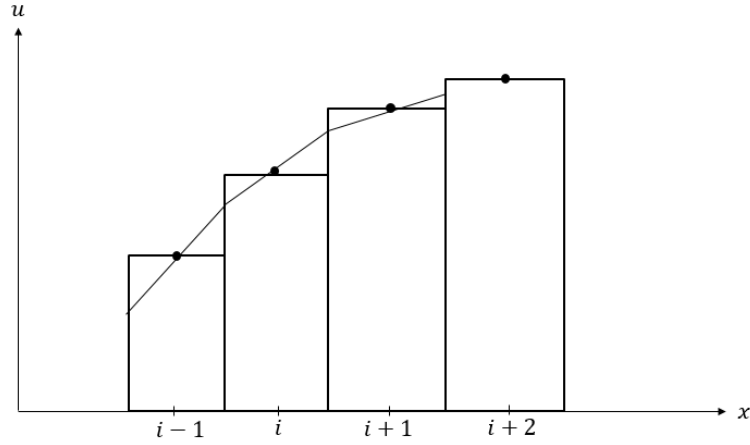


Figure A.1: Continuous variation of the state variable and flux through cells.

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (F_{i+1/2}^{n+1/2} - F_{i-1/2}^{n+1/2}).$$

This discrete representation is known as the fully-discrete form. Different assumptions can be considered for variation of  $u$  (and so for  $f(u)$ ) through cell edges: continuous or discontinuous. In the following, at first, effects and drawbacks of the continuity assumption for the parameter  $u$  will be studied. Then by relaxing this constraint, several central schemes will be derived without such continuity constraint.

### A.1.1 Continuous assumption for the state variable and flux

In this case, a continuous variation is assumed through cell edges (Figure A.1). With this assumption, the average flux ( $F$ ) can be approximated over time  $[t_n, t_{n+1}]$  as:

$$F_{i-1/2}^n = \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} f(u(x_{i-1/2}, t)) dt \approx \frac{1}{\Delta t} (F_{i-1/2}^n \times \Delta t). \quad (\text{A.4})$$

Let also assume linear variation of  $F$  through cells; then:

$$\begin{aligned} F_{i-1/2}^n &= \frac{1}{2}(F_{i-1}^n + F_i^n), \\ F_{i+1/2}^n &= \frac{1}{2}(F_i^n + F_{i+1}^n). \end{aligned} \quad (\text{A.5})$$

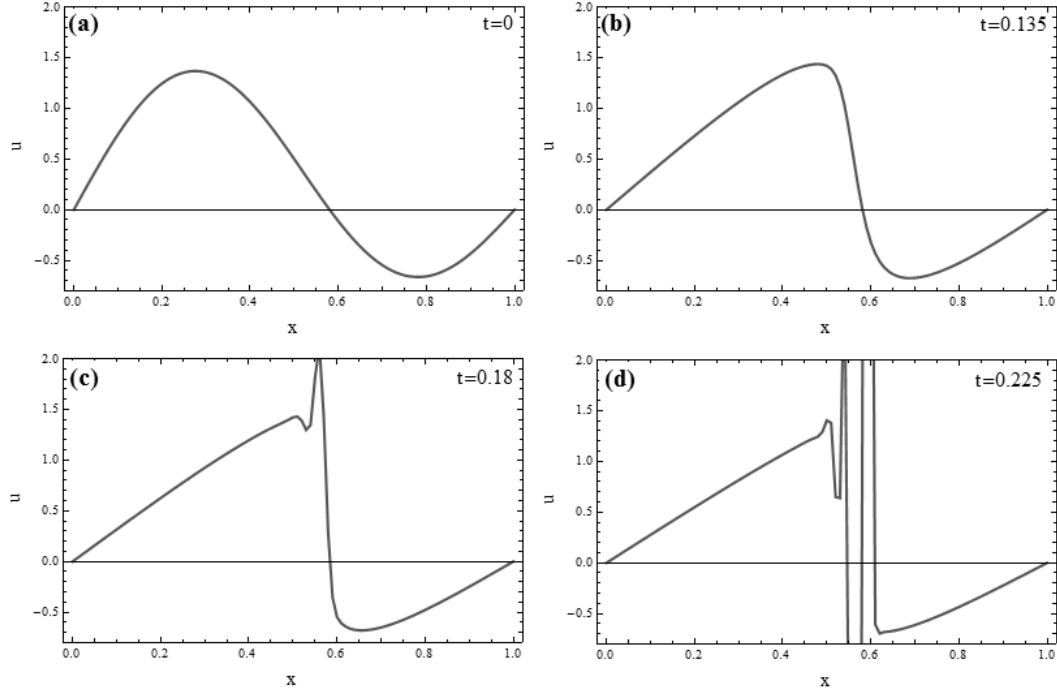


Figure A.2: Simulation of the Burgers' equation with numerical solution using continuous variable and flux.

So the fully-discrete form can be written as:

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (F_{i+1/2}^n - F_{i-1/2}^n). \quad (\text{A.6})$$

Using the two approximations of the fluxes  $F$ s, Eq. (A.5), Eq. (A.6) can be rewritten as:

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{2\Delta x} (F_{i+1}^n + F_i^n - F_i^n - F_{i-1}^n) = u_i^n - \frac{\Delta t}{2\Delta x} (F_{i+1}^n - F_{i-1}^n). \quad (\text{A.7})$$

This method is generally unstable for non-linear hyperbolic problems; it can not be used, even by using a very small time steps.

**Example:** The Burgers' equation ( $\frac{\partial u}{\partial t} + 0.5 \frac{\partial}{\partial x} (u^2) = 0$ ) with the smooth IC  $u(x, t = 0) = \sin(2\pi x) + \frac{\sin(\pi x)}{2}$  and BCs  $u(0, t) = u(1, t) = 0$ , is considered. This equation is a non-linear system and so discontinuous propagating fronts will develop even for smooth initial conditions. The numerical results are presented in Figure A.2. It is clear that spurious (non-physical) oscillations develop through time (and instability occurs around  $t = 0.225$ ): this is a direct consequence of continuous variation of the state

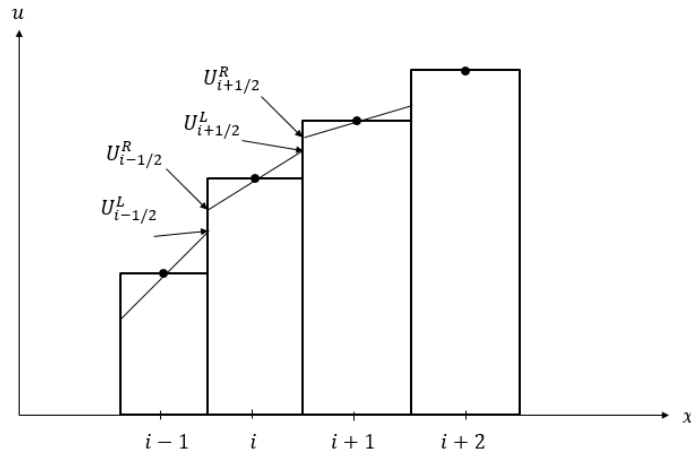


Figure A.3: Piecewise linear variation of the state variable and flux on cells.

variable and flux.

### A.1.2 Discontinuous variation of state variables and fluxes

To have a stable solution, the continuity constraint can be relaxed at cell edges, as shown in Figure A.3. Based on such variations, the Lax-Friedrich (LxF), NT [25] and KT [22] schemes have been developed; for the LxF and NT schemes both staggered and non-staggered formulations are provided [75]. they will be described in the following.

#### The LxF Method

The LxF method is the first scheme allowing existence of possible discontinuities at cell edges. In this method, a piecewise constant variation is assumed on cells with possible discontinuities at cell edges, see Figure A.4. Due to constant values of state variables, the LxF method leads to a first-order accuracy scheme: a dissipative method. In the following, formulations of the LxF method on staggered and non-staggered grids (cells) will be provided.

**The LxF method - staggered formulation** The following hyperbolic equation is considered:

$$u_t + f_x = 0. \quad (\text{A.8})$$

Variation of the state variable in LxF method is show in Figure A.4a. Based on this illustration, the hyperbolic equation is integrated on the spatio-temporal volume  $[x_i, x_{i+1}] \times$

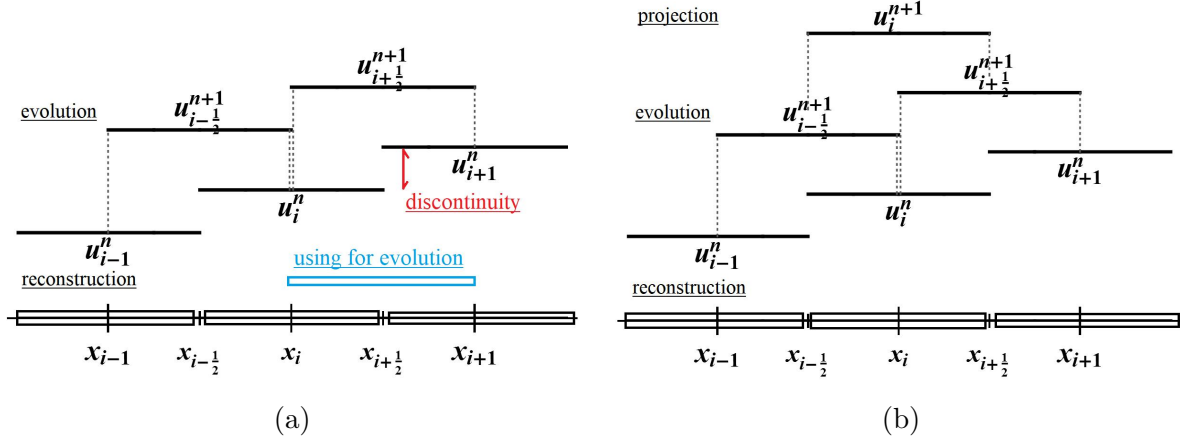


Figure A.4: Variation of solutions over cells (reconstructed and evolved ones) for the LxF scheme; (a) Staggered formulation, (b) Non-staggered one.  $\{x_i\}$  and  $\{x_{i\pm 1/2}\}$  denote cell-centers and cell-edges, respectively.

$[t_n, t_{n+1}]$ , as:

$$\int_{x_i}^{x_{i+1}} \int_{t^n}^{t^{n+1}} (u_{,t} + f_{,x}) dt dx = 0, \quad (\text{A.9})$$

or:

$$\int_{x_i}^{x_{i+1}} (u^{n+1} - u^n) dx + \int_{t^n}^{t^{n+1}} (f(x_{i+1}) - f(x_i)) dt = 0. \quad (\text{A.10})$$

Also, the integrals over the spatial domain can be approximated as:

$$\int_{x_i}^{x_{i+1}} u^{n+1} dx \approx \Delta x u_{i+1/2}^{n+1}, \quad (\text{A.11})$$

and

$$\int_{x_i}^{x_{i+1}} u^n dx = \int_{x_i}^{x_{i+1/2}} u^n dx + \int_{x_{i+1/2}}^{x_{i+1}} u^n dx \approx \frac{\Delta x}{2} (u_i^n + u_{i+1}^n), \quad (\text{A.12})$$

The integral of fluxes over time can also be approximated as:

$$\int_{t^n}^{t^{n+1}} f(x_i) dt \approx F^n(x_i) \Delta t, \quad (\text{A.13})$$

and

$$\int_{t^n}^{t^{n+1}} f(x_{i+1}) dt \approx F^n(x_{i+1}) \Delta t. \quad (\text{A.14})$$

By replacing Eqs. (A.11)-(A.14) in Eq. (A.10), the final form of the LxF method can be obtained on staggered cells, as:

$$u_{i+\frac{1}{2}}^{n+1} = \frac{1}{2}(u_i^n + u_{i+1}^n) - \frac{\Delta t}{\Delta x}(F^n(x_{i+1}) - F^n(x_i)). \quad (\text{A.15})$$

This formulation is called a staggered scheme, since at the next time step  $n + 1$ , the solutions are obtained at edge points  $x_{i+1/2}$ , Figure A.4a.

This form of formulation has more artificial dissipation than non-staggered one; so the latter is always preferable. In the non-staggered form, the spatial location of the solution in time step  $n + 1$  will be the same as that of time step  $n$ .

**The LxF scheme - non staggered** The non-staggered form of the LxF method can be obtained by the projection step. This can simply be done by evaluating an average of the staggered solutions (from two neighbor staggered cells, Figure A.4b), as: [75]:

$$u_i^{n+1} = \frac{u_{i-\frac{1}{2}}^{n+1} + u_{i+\frac{1}{2}}^{n+1}}{2}. \quad (\text{A.16})$$

This relationship can simply be obtained, since:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_i^{n+1} dx = \int_{x_{i-\frac{1}{2}}}^{x_i} u_{i-\frac{1}{2}}^{n+1} dx + \int_{x_i}^{x_{i+\frac{1}{2}}} u_{i+\frac{1}{2}}^{n+1} dx \approx \frac{\Delta x}{2}(u_{i-\frac{1}{2}}^{n+1} + u_{i+\frac{1}{2}}^{n+1}),$$

and also:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_i^{n+1} dx \approx \Delta x \cdot u_i^{n+1},$$

and so that Eq. (A.16) can be obtained. For staggered solution at point  $x_{i-1/2}$ , Eq. (A.15) yields:

$$u_{i-\frac{1}{2}}^{n+1} = \frac{1}{2}(u_i^n + u_{i-1}^n) - \frac{\Delta t}{\Delta x}(F^n(x_i) - F^n(x_{i-1})). \quad (\text{A.17})$$

By averaging Eqs. (A.15) and (A.17), the final non-staggered form of the LxF scheme becomes [75]:

$$u_i^{n+1} = \frac{1}{4}(u_{i-1}^n + 2u_i^n + u_{i+1}^n) - \frac{\Delta t}{2\Delta x}(F^n(x_{i+1}) - F^n(x_{i-1})). \quad (\text{A.18})$$

This formulation has a fully-discrete form.

**Example:** The previous Burgers' equation is considered again in order to study the performance of the LxF method. The numerical results are shown in Figure A.5 (with



Appendix A Central schemes on uniform grids

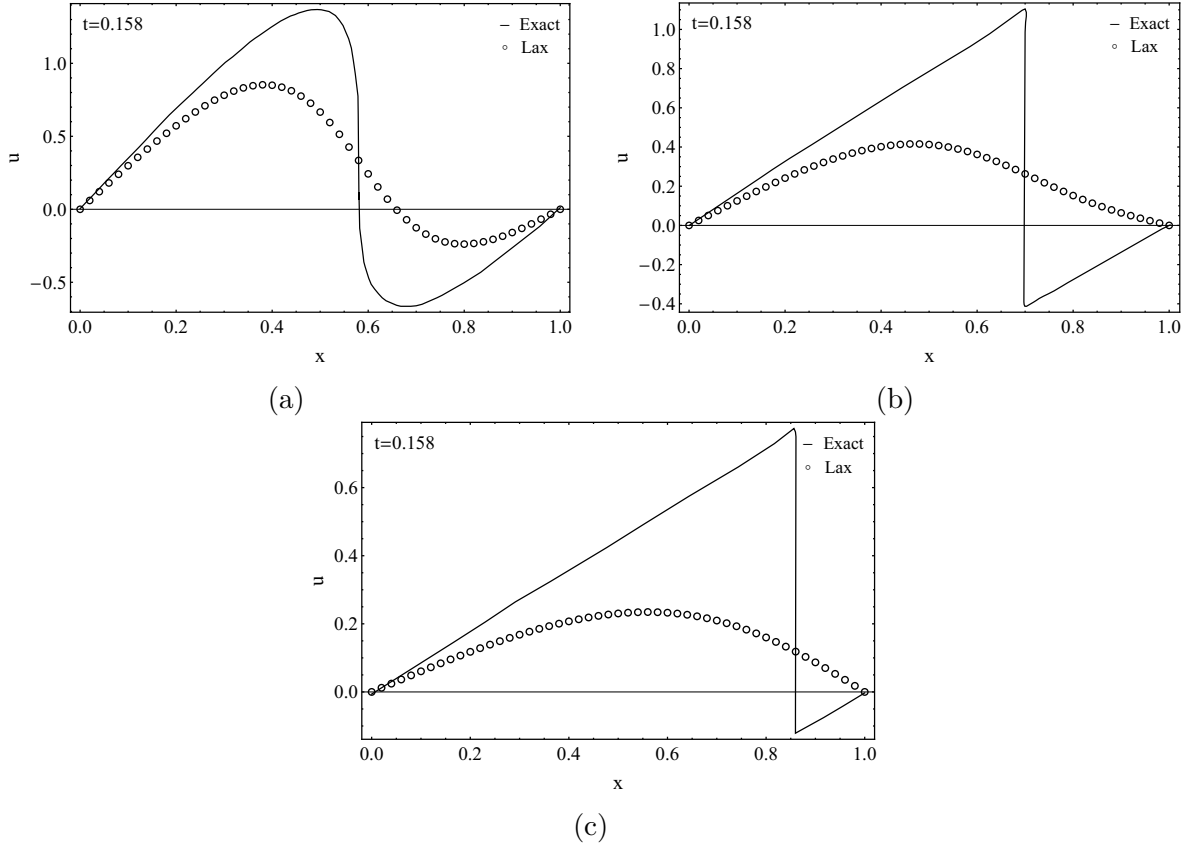


Figure A.5: Numerical results for the Burgers' equation obtained by the LxF scheme at  $t=0.158$ ,  $t=0.5$  and  $t=1$ ; solid lines and hollow circles denote the exact and numerical solutions, respectively.

the same CFL condition of Figure A.2). It is clear that the numerical dissipation is considerable in the LxF method, even-though the results are stable.

**The truncation error due to the LxF scheme** In this part, the truncation error will be provided for the LxF scheme. Let subtract  $u_{i+1/2}^n$  from both sides of Eq. (A.15) at  $x_{i+1/2}$ :

$$u_{i+1/2}^{n+1} - u_{i+1/2}^n = \frac{1}{2}(u_i^n - 2u_{i+1/2}^n + u_{i+1}^n) - \frac{\Delta t}{\Delta x}(F^n(x_{i+1}) - F^n(x_i)). \quad (\text{A.19})$$

Dividing Eq. (A.19) by  $\Delta t$  leads to:

$$\begin{aligned} \frac{u_{i+\frac{1}{2}}^{n+1} - u_{i+\frac{1}{2}}^n}{\Delta t} &= \frac{1}{2\Delta t}(u_i^n - 2u_{i+\frac{1}{2}}^n + u_{i+1}^n) - \frac{1}{\Delta x}(F^n(x_{i+1}) - F^n(x_i)) \\ &= \frac{1}{2\Delta t} \frac{(\frac{\Delta x}{2})^2[(u_{i+1}^n - u_{i+\frac{1}{2}}^n) - (u_{i+\frac{1}{2}}^n - u_i^n)]}{(\frac{\Delta x}{2})^2} - \frac{1}{\Delta x}(F^n(x_{i+1}) - F^n(x_i)). \end{aligned} \quad (\text{A.20})$$

According to approximation of the second derivative of  $u$  with the central difference definition  $((u_{xx})_i \approx \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2})$ <sup>12</sup>, Eq. (A.20) becomes:

$$\frac{u_{i+\frac{1}{2}}^{n+1} - u_{i+\frac{1}{2}}^n}{\Delta t} = \frac{\Delta x^2}{8\Delta t}(u_{xx})_{i+\frac{1}{2}} - \frac{1}{\Delta x}(F^n(x_{i+1}) - F^n(x_i)) + \Delta x^2 \mathcal{O}(\Delta x^2). \quad (\text{A.21})$$

Based on the Taylor expansion,  $(u_{xx})_{i+\frac{1}{2}}$  can be expanded as:

$$(u_{xx})_{i+\frac{1}{2}} = (u_{xx})_i + \frac{\Delta x}{2}(u_{xxx})_i + \mathcal{O}(\Delta x^2). \quad (\text{A.22})$$

Replacing Eq. (A.22) in Eq. (A.21) results as:

$$\frac{u_{i+\frac{1}{2}}^{n+1} - u_{i+\frac{1}{2}}^n}{\Delta t} = \frac{\Delta x^2}{2^3\Delta t}(u_{xx})_i + \frac{\Delta x^3}{2^4\Delta t}(u_{xxx})_i - \frac{1}{\Delta x}(F^n(x_{i+1}) - F^n(x_i)) + \mathcal{O}(\Delta x^4). \quad (\text{A.23})$$

The RHS of Eq. (A.23) contains two types of derivatives: odd and even. The smallest order of odd and even derivatives control the numerical dissipation and numerical dispersion, respectively. Also as  $\Delta t$  is in the denominator of the coefficients, it can not be chosen to be so small, since:  $\frac{1}{\Delta t} \rightarrow \infty$  as  $\Delta t \rightarrow 0$ .

### The NT scheme

The NT method is an improvement of the LxF method with the same assumption for the variation of state variable at edges: existence of possible discontinuities at cell-edges. In this method, a linear piecewise variation is considered through cells, Figure A.6. In the NT method, the solution has a second-order accuracy as a linear variation

---

<sup>12</sup>This central difference approximation of the second derivative has second order accuracy; this means  $(u_{xx})_i = \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} + \mathcal{O}(\Delta x^2)$ .

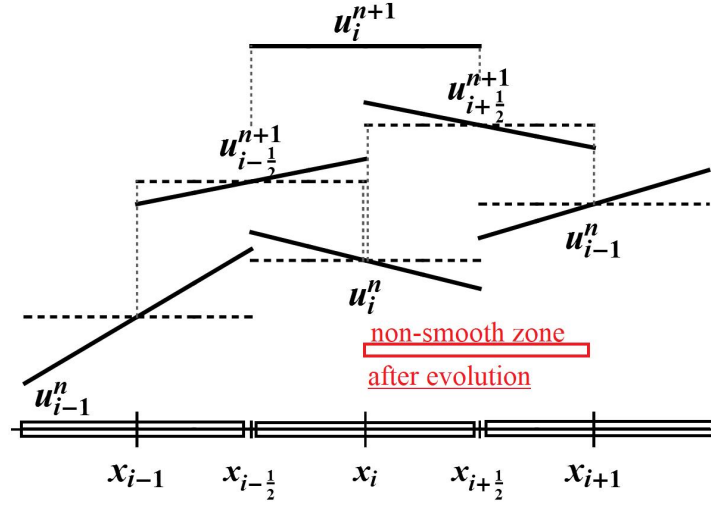


Figure A.6: Illustration of the NT scheme.

is assumed in each cell. The formulation for this scheme can be followed with the same procedure denoted in the LxF scheme. The NT formulation can also have staggered or non-staggered form. For studying the numerical dissipation phenomenon, the truncation error due to applying the NT method will be evaluated and compared with that of the LxF scheme.

**The NT scheme - Staggered** Again based on Figure A.6, the hyperbolic equation is integrated over the spatio-temporal volume  $[x_i, x_{i+1}] \times [t^n, t^{n+1}]$  as:

$$\int_{x_i}^{x_{i+1}} \int_{t^n}^{t^{n+1}} (u_t + f_x) dt dx = 0, \quad (\text{A.24})$$

$$\int_{x_i}^{x_{i+1}} (u^{n+1} - u^n) dx + \int_{t^n}^{t^{n+1}} (f(x_{i+1}) - f(x_i)) dt = 0.$$

Due to a possible discontinuity at  $x_{i+\frac{1}{2}}$ , the integration on the spatial domain is divided into two parts:

$$I = \int_{x_i}^{x_{i+1}} u^n dx = \int_{x_i}^{x_{i+\frac{1}{2}}} u^n dx + \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} u^n dx. \quad (\text{A.25})$$

Due to the linear variations in cells, the state variable  $u^n(x)$  can be reconstructed as:

i On the interval  $x_{i-\frac{1}{2}} \leq x \leq x_{i+\frac{1}{2}}$ :

$$u^n(x) = \bar{u}_i + (u_x)_i(x - x_i), \quad (\text{A.26})$$

ii On the interval  $x_{i+\frac{1}{2}} \leq x \leq x_{i+\frac{3}{2}}$ ,

$$u^n(x) = \bar{u}_{i+1} + (u_x)_{i+1}(x - x_{i+1}), \quad (\text{A.27})$$

where:  $(u_x)_i$  and  $(u_x)_{i+1}$  are proper approximations of the first derivative (approximated by slope limiters); and the cell averages  $\bar{u}_i^n$  and  $\bar{u}_{i+1}^n$  can be obtained as:

$$\bar{u}_i^n = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u^n dx, \quad (\text{A.28})$$

$$\bar{u}_{i+1}^n = \frac{1}{\Delta x} \int_{x_{i+\frac{1}{2}}}^{x_{i+\frac{3}{2}}} u^n dx. \quad (\text{A.29})$$

It should be mentioned that since variations are linear then:  $u_i^n = \bar{u}_i^n$ . By inserting the linear variations in Eq. (A.25), we have:

$$I = \bar{u}_i \frac{\Delta x}{2} + (u_x)_i \underbrace{\left(-x_i \frac{\Delta x}{2} + \frac{1}{2}(x_{i+\frac{1}{2}}^2 - x_i^2)\right)}_{J_1} + \bar{u}_{i+1} \frac{\Delta x}{2} + (u_x)_{i+1} \underbrace{\left(-x_{i+1} \frac{\Delta x}{2} + \frac{1}{2}(x_{i+1}^2 - x_{i+\frac{1}{2}}^2)\right)}_{J_2}. \quad (\text{A.30})$$

So, it is easy to see that:

$$\begin{aligned} J_1 &= \left(-x_i \frac{\Delta x}{2} + \frac{1}{2}(x_{i+\frac{1}{2}}^2 - x_i^2)\right) = -x_i \frac{\Delta x}{2} + \frac{1}{2}(x_{i+\frac{1}{2}} - x_i)(x_{i+\frac{1}{2}} + x_i) \\ &= -x_i \frac{\Delta x}{2} + \frac{1}{2} \frac{\Delta x}{2} (2x_i + \frac{\Delta x}{2}) = \frac{\Delta x^2}{8}, \end{aligned} \quad (\text{A.31})$$

and

$$\begin{aligned} J_2 &= -x_{i+1} \frac{\Delta x}{2} + \frac{1}{2}(x_{i+1}^2 - x_{i+\frac{1}{2}}^2) = -x_{i+1} \frac{\Delta x}{2} + \frac{1}{2}(x_{i+1} - x_{i+\frac{1}{2}})(x_{i+1} + x_{i+\frac{1}{2}}) \\ &= -x_{i+1} \frac{\Delta x}{2} + \frac{1}{2} \left(\frac{\Delta x}{2}\right) (2x_{i+1} - \frac{\Delta x}{2}) = -\frac{\Delta x^2}{8}. \end{aligned} \quad (\text{A.32})$$

Substituting Eqs. (A.31) and (A.32) in Eq. (A.30), yields:

$$I = \frac{\Delta x}{2}(\bar{u}_i + \bar{u}_{i+1}) + \frac{\Delta x^2}{8}((u_x)_i - (u_x)_{i+1}). \quad (\text{A.33})$$

Considering Eq. (A.33), Eq. (A.24) can be rewritten as:

$$\int_{x_i}^{x_{i+1}} u^{n+1} dx - I + \int_{t^n}^{t^{n+1}} (f(x_{i+1}) - f(x_i)) dt = 0. \quad (\text{A.34})$$

Using the mid-point approximation for the time integration (which has second order accuracy),

$$\int_{t^n}^{t^{n+1}} f(x_i) dt \approx F(x_i^{n+\frac{1}{2}}) \Delta t, \quad (\text{A.35})$$

and based on linear variations on the spatial domain, it is clear:

$$\int_{x_i}^{x_{i+1}} u^{n+1} dx \approx u_{i+\frac{1}{2}}^{n+1} \Delta x. \quad (\text{A.36})$$

So Eq. (A.34) can be rewritten as:

$$\Delta x u_{i+\frac{1}{2}}^{n+1} - \frac{\Delta x}{2}(\bar{u}_i^n + \bar{u}_{i+1}^n) - \frac{\Delta x^2}{8}((u_x)_i^n - (u_x)_{i+1}^n) + \Delta t(F(x_{i+\frac{1}{2}}^{n+\frac{1}{2}}) - F(x_i^{n+\frac{1}{2}})). \quad (\text{A.37})$$

Finally, the final form of the staggered NT scheme becomes:

$$u_{i+\frac{1}{2}}^{n+1} = \frac{1}{2}(\bar{u}_i^n + \bar{u}_{i+1}^n) + \frac{\Delta x}{8}((u_x)_i^n - (u_x)_{i+1}^n) - \frac{\Delta t}{\Delta x}(F(x_{i+\frac{1}{2}}^{n+\frac{1}{2}}) - F(x_i^{n+\frac{1}{2}})). \quad (\text{A.38})$$

This is a fully-discrete form. Based on the Taylor expansion, the mid-values can be approximated in time as:

$$u_i^{n+\frac{1}{2}} = \bar{u}_i^n - \frac{\Delta t}{2}(F_{,x})_i^n, \quad (\text{A.39})$$

and

$$u_{i+1}^{n+\frac{1}{2}} = \bar{u}_{i+1}^n - \frac{\Delta t}{2}(F_{,x})_{i+1}^n. \quad (\text{A.40})$$

It should be mentioned that for this formulation, the two first steps of the Gondunov-type central scheme (the reconstruction and evolution stages) are used.

**The NT method - non staggered** In order to project the staggered form of the NT method to the non-staggered form, the staggered solution  $u_{i+\frac{1}{2}}^{n+1}$  is averaged to obtain  $u_i^{n+1}$ . It is assume the evolved staggered solutions  $u^{n+1}$  have linear variation on

staggered cells (a linear piecewise interpolation, but this time on the staggered grid):

$$u_{i+\frac{1}{2}}^{n+1} = \bar{u}_{i+\frac{1}{2}}^{n+1} + (\bar{u}_x)_{i+\frac{1}{2}}(x - x_{i+\frac{1}{2}}), \quad (\text{A.41})$$

and

$$u_{i-\frac{1}{2}}^{n+1} = \bar{u}_{i-\frac{1}{2}}^{n+1} + (\bar{u}_x)_{i-\frac{1}{2}}(x - x_{i-\frac{1}{2}}). \quad (\text{A.42})$$

Based on Figure A.6, the integral of  $\bar{u}_i^{n+1}$  is divided into two parts due to possible discontinuity at  $x_i$  as [75]:

$$\bar{u}_i^{n+1} = \frac{1}{\Delta x} \left( \int_{x_{i-\frac{1}{2}}}^{x_i} u_{i-\frac{1}{2}}^{n+1} + \int_{x_i}^{x_{i+\frac{1}{2}}} u_{i+\frac{1}{2}}^{n+1} \right) = \frac{\bar{u}_{i-\frac{1}{2}}^{n+1} + \bar{u}_{i+\frac{1}{2}}^{n+1}}{2} - \frac{\Delta x}{8} ((u_x)_{i+\frac{1}{2}} - (u_x)_{i-\frac{1}{2}}). \quad (\text{A.43})$$

The final formulation of non-staggered scheme of the NT method can be represented as:

$$\begin{aligned} \bar{u}_i^{n+1} &= \frac{1}{4}(\bar{u}_{i-1}^n + 2\bar{u}_i^n + \bar{u}_{i+1}^n) - \frac{\Delta x}{16}((u_x)_{i+1} - (u_x)_{i-1}) \\ &\quad - \frac{\Delta t}{2\Delta x}(F(u_{i+\frac{1}{2}}^{n+\frac{1}{2}}) - F(u_{i-\frac{1}{2}}^{n+\frac{1}{2}})) - \frac{\Delta x}{8}(u_x)_{i+\frac{1}{2}} - (u_x)_{i-\frac{1}{2}}. \end{aligned} \quad (\text{A.44})$$

**The truncation error for the NT method** In this part, the truncation error for the NT scheme will be provided: the source of the numerical dissipation and dispersion. By subtracting  $u_{j+\frac{1}{2}}^n$  from both sides of the fully-discrete form of the NT scheme, Eq. (A.38), this equation can be rewritten as:

$$u_{j+\frac{1}{2}}^{n+1} - u_{j+\frac{1}{2}}^n = \frac{1}{2}(u_j^n - 2u_{j+\frac{1}{2}}^n + u_{j+1}^n) + \frac{\Delta x}{8}((u_x)_j^n + (u_x)_{j+1}^n) - \frac{\Delta t}{\Delta x}(F(x_{j+\frac{1}{2}}^{n+\frac{1}{2}}) - F(x_j^{n+\frac{1}{2}})). \quad (\text{A.45})$$

Dividing Eq. (A.45) by  $\Delta t$ , yields:

$$\begin{aligned} \frac{u_{j+\frac{1}{2}}^{n+1} - u_{j+\frac{1}{2}}^n}{\Delta t} &= \frac{1}{2\Delta t}(u_j^n - 2u_{j+\frac{1}{2}}^n + u_{j+1}^n) + \frac{\Delta x}{8\Delta t}((u_x)_j^n - (u_x)_{j+1}^n) \\ &\quad - \frac{1}{\Delta x}(F(x_{j+\frac{1}{2}}^{n+\frac{1}{2}}) - F(x_j^{n+\frac{1}{2}})) \\ &= \frac{1}{2\Delta t}[(u_{j+1}^n - u_{j+\frac{1}{2}}^n) - (u_{j+\frac{1}{2}}^n - u_j^n)] + \frac{\Delta x}{8\Delta t}((u_x)_j^n - (u_x)_{j+1}^n) \\ &\quad - \frac{1}{\Delta x}(F(x_{j+\frac{1}{2}}^{n+\frac{1}{2}}) - F(x_j^{n+\frac{1}{2}})). \end{aligned} \quad (\text{A.46})$$

According to the definition of the central second derivative of  $u$  with second order accuracy (i.e.:  $(u_{xx})_i = \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} + \mathcal{O}(\Delta x^2)$ ), Eq. (A.46) can be rewritten as:

$$\begin{aligned} \frac{u_{j+\frac{1}{2}}^{n+1} - u_{j+\frac{1}{2}}^n}{\Delta t} &= \frac{1}{2\Delta t} \frac{(\frac{\Delta x}{2})^2 [(u_{j+1}^n - u_{j+\frac{1}{2}}^n) - (u_{j+\frac{1}{2}}^n - u_j^n)]}{(\frac{\Delta x}{2})^2} + \frac{\Delta x \Delta x ((u_x)_j^n - (u_x)_{j+1}^n)}{8\Delta t \Delta x} \\ &\quad - \frac{1}{\Delta x} (F(x_{j+1}^{n+\frac{1}{2}}) - F(x_j^{n+\frac{1}{2}})) \\ &= \frac{(\Delta x)^2}{8\Delta t} (u_{xx})_{j+\frac{1}{2}} - \frac{(\Delta x)^2}{8\Delta t} (u_{xx})_j - \frac{1}{\Delta x} (F(x_{j+1}^{n+\frac{1}{2}}) - F(x_j^{n+\frac{1}{2}})) + \mathcal{O}(\Delta x^4). \end{aligned} \quad (\text{A.47})$$

Based on the Taylor expansion,  $(u_{xx})_{j+\frac{1}{2}}$  can be expanded as:

$$(u_{xx})_{j+\frac{1}{2}} = (u_{xx})_j + \frac{\Delta x}{2} (u_{xxx})_j + \mathcal{O}(\Delta x^2). \quad (\text{A.48})$$

Replacing Eq. (A.48) in Eq. (A.47), results:

$$\begin{aligned} \frac{u_{j+\frac{1}{2}}^{n+1} - u_{j+\frac{1}{2}}^n}{\Delta t} &= \frac{\Delta x^2}{8\Delta t} \left( (u_{xx})_j + \frac{\Delta x}{2} (u_{xxx})_j + \mathcal{O}(\Delta x^2) \right) \\ &\quad - \frac{\Delta x^2}{8\Delta t} (u_{xx})_j - \frac{1}{\Delta x} \left( F(x_{j+1}^{n+\frac{1}{2}}) - F(x_j^{n+\frac{1}{2}}) \right) + \mathcal{O}(\Delta x^4) \\ &= \frac{\Delta x^3}{2^4 \Delta t} (u_{xxx})_j - \frac{1}{\Delta x} \left( F(x_{j+1}^{n+\frac{1}{2}}) - F(x_j^{n+\frac{1}{2}}) \right) + \mathcal{O}(\Delta x^4). \end{aligned} \quad (\text{A.49})$$

From Eq. (A.49) it is clear that:

1. The smallest coefficients in the truncation error is for the odd derivative, showing the numerical dissipation effects,
2.  $\Delta t$  is in the denominator of the coefficients; so the truncation error approach infinity as  $\Delta t \rightarrow 0$  (since  $\frac{1}{\Delta t} \rightarrow \infty$ ). As a result, the fully-discrete form of the NT scheme can not be converted to the semi-discrete form,
3. The truncation error of the second order NT scheme is much less that those of the first-order LxF method.

**Example:** The previous Burgers' equation is studied for reveal numerical feature of the NT scheme with the previous CFL condition. The numerical results are shown in Figure A.7. It is clear that numerical dissipation is considerably less than the LxF scheme.

## Appendix A Central schemes on uniform grids

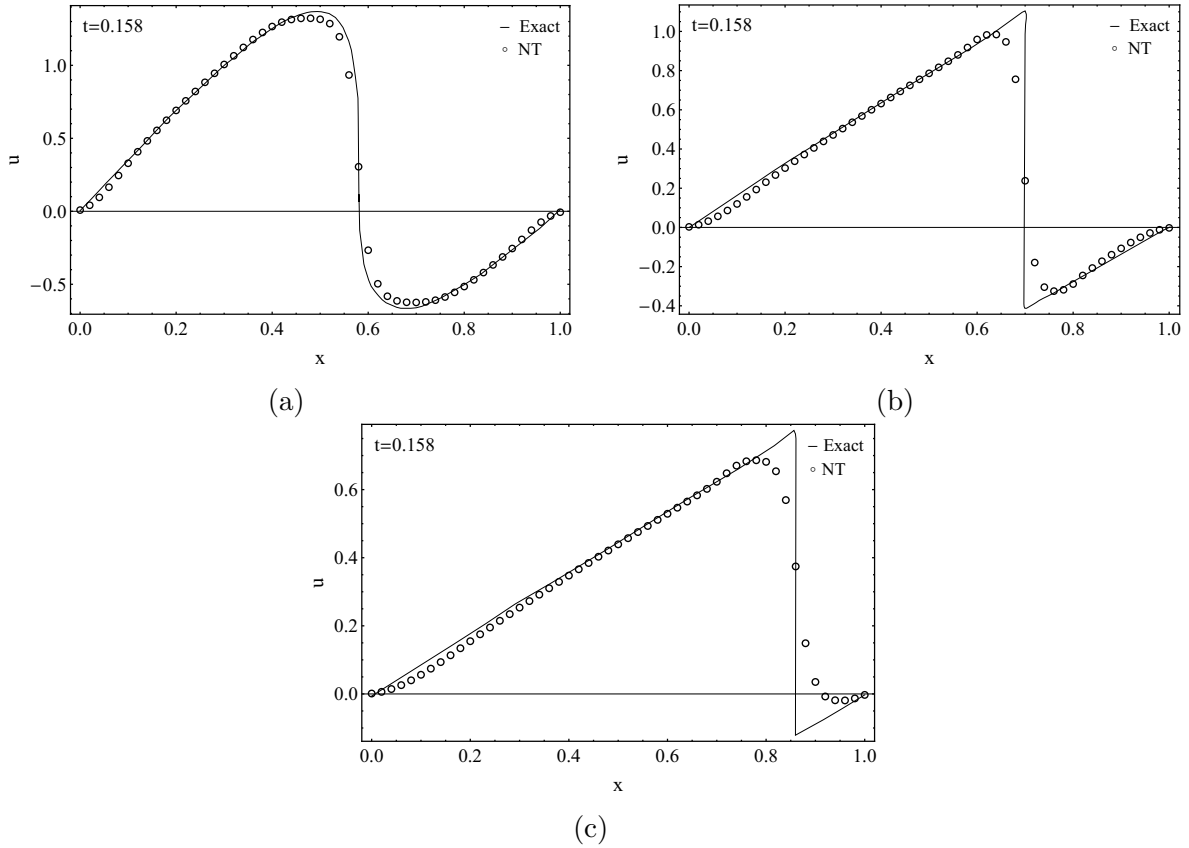


Figure A.7: Numerical results for the Burgers' equation by using the NT method at  $t=0.158$ ,  $t=0.5$  and  $t=1$ ; solid lines and hollow circles denote the exact and numerical solutions, respectively.

### The Kurganov and Tadmor (KT) method

The last method which will be derived in this section is the KT method. In this subsection, the KT formulation in fully-discrete and semi-discrete forms will be provided.

**The KT scheme - the fully discrete second-order central form** The main idea behind the introduction of the KT method is to use more precise information about the local speed of propagating waves for distinguishing smooth and non-smooth solutions from each other. Non-smooth solutions can develop due to possible discontinuities around cell edges  $x_{i+1/2}$ . By this distinguishing, the non-smooth solutions are averaged over narrower zones and this leads to the smaller numerical dissipation. In comparison to the NT scheme, in the KT method, the maximum local speeds around cell edges are needed as additional information [22].

In the KT scheme, in the reconstruction step, the linear piecewise interpolation is



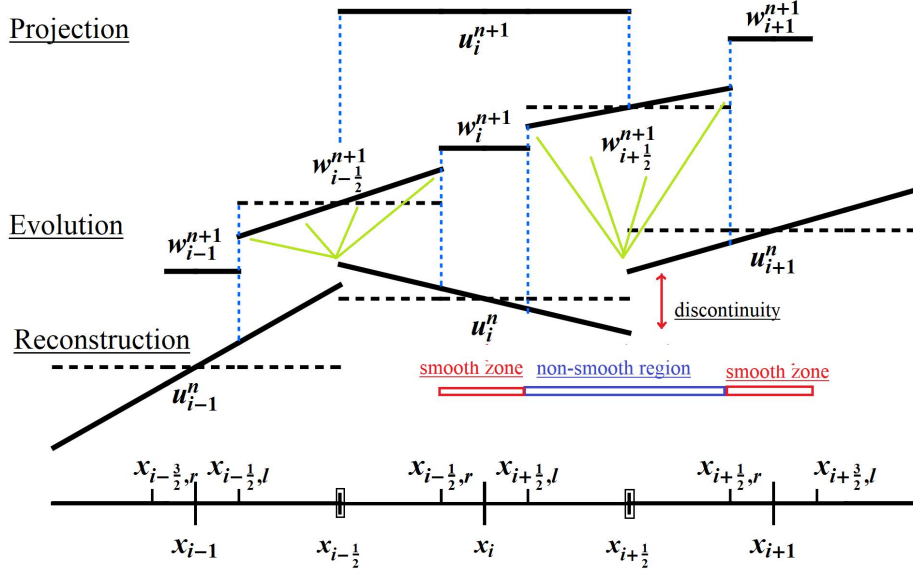


Figure A.8: Illustration of the KT scheme.

used over each cell; so, the state variable  $u(x)$  can be approximated as:

$$u(x) = u_j^n + (x - x_j)(u_x)_j, \quad x \in [x_{j-\frac{1}{2}}^n, x_{j+\frac{1}{2}}^n]. \quad (\text{A.50})$$

Since in each cell, the variation is linear, the method has second order accuracy in the spatial domain. As mentioned before due to possible discontinuities around cell-edges, solutions can be classified as smooth and non-smooth solutions. The left and right borders of the non-smooth zone around  $x_{j+1/2}$  are denoted by  $x_{j+1/2,l}$  and  $x_{j+1/2,r}$ , respectively. The locations can be obtained as:

$$\begin{aligned} x_{j+\frac{1}{2},r} &= x_{j+\frac{1}{2}} + a_{j+\frac{1}{2}} \cdot \Delta t, \\ x_{j+\frac{1}{2},l} &= x_{j+\frac{1}{2}} - a_{j+\frac{1}{2}} \cdot \Delta t, \end{aligned} \quad (\text{A.51})$$

where  $a_{j+\frac{1}{2}}$  is the maximum of absolute local propagation speed at  $x_{j+\frac{1}{2}}$ .

After the reconstruction step, the evaluation step will be done. For this purpose, according to Figure A.8, the evolution stage is done separately for smooth and non-smooth parts as:

- i Evolution on the non-smooth volume:  $[x_{j+\frac{1}{2},l}^n, x_{j+\frac{1}{2},r}^n] \times [t^n, t^{n+1}]$ ,
- ii Evolution on the smooth volume:  $[x_{j-\frac{1}{2},r}^n, x_{j+\frac{1}{2},l}^n] \times [t^n, t^{n+1}]$ .

i **The non-smooth interval**  $[x_{j+\frac{1}{2},l}^n, x_{j+\frac{1}{2},r}^n] \times [t^n, t^{n+1}]$ :

We begin by the spatio-temporal integration as:

$$\begin{aligned} \int_{t^n}^{t^{n+1}} \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} \left( \frac{\partial u}{\partial t} + \frac{\partial F}{\partial x} \right) dx dt &= 0, \\ \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} [u(x, t^{n+1}) - u(x, t^n)] dx + \int_{t^n}^{t^{n+1}} [F(x_{j+\frac{1}{2},r}, t) - F(x_{j+\frac{1}{2},l}, t)] dt &= 0, \\ \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} u(x, t^{n+1}) dx - \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} u(x, t^n) dx + \int_{t^n}^{t^{n+1}} [F(x_{j+\frac{1}{2},r}, t) - F(x_{j+\frac{1}{2},l}, t)] dt &= 0. \end{aligned} \tag{A.52}$$

Let  $\int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} u(x, t^n) dx$  be referred as  $I$ . Due to possible discontinuity at  $x_{i+1/2}$ , we have:

$$\begin{aligned} I &= \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} u(x, t^n) dx \\ &= \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2}}} [u_j + (x - x_j)(u_x)_j] dx + \int_{x_{j+\frac{1}{2}}}^{x_{j+\frac{1}{2},r}} [u_{j+1} + (x - x_{j+1})(u_x)_{j+1}] dx \\ &= u_j(x_{j+\frac{1}{2}} - x_{j+\frac{1}{2},l}) + (u_x)_j \left( \frac{x_{j+\frac{1}{2}}^2}{2} - \frac{x_{j+\frac{1}{2},l}^2}{2} - (x_j \cdot x_{j+\frac{1}{2}} - x_j \cdot x_{j+\frac{1}{2},l}) \right) \\ &\quad + u_{j+1}(x_{j+\frac{1}{2},r} - x_{j+\frac{1}{2}}) \\ &\quad + (u_x)_{j+1} \left( \frac{x_{j+\frac{1}{2},r}^2}{2} - \frac{x_{j+\frac{1}{2}}^2}{2} - (x_{j+1} \cdot x_{j+\frac{1}{2},r} - x_{j+1} \cdot x_{j+\frac{1}{2}}) \right). \end{aligned} \tag{A.53}$$

However:  $x_{j+\frac{1}{2},r} = x_{j+\frac{1}{2}} + a_{j+\frac{1}{2}} \cdot \Delta t$ ,  $x_{j+\frac{1}{2},l} = x_{j+\frac{1}{2}} - a_{j+\frac{1}{2}} \cdot \Delta t$  and  $x_{j+\frac{1}{2}} - x_{j+\frac{1}{2},l} = a_{j+\frac{1}{2}} \Delta t$ ; therefore  $x_{j+\frac{1}{2},r} - x_{j+\frac{1}{2}} = a_{j+\frac{1}{2}} \Delta t$ . Inserting these relationships in Eq. (A.53), we have:

$$\begin{aligned} I &= \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} u(x, t^n) dx = u_j(a_{j+\frac{1}{2}} \Delta t) + (u_x)_j (x_{j+\frac{1}{2}} \cdot a_{j+\frac{1}{2}} \Delta t - x_j \cdot a_{j+\frac{1}{2}} \Delta t - \frac{a_{j+\frac{1}{2}}^2 \cdot \Delta t^2}{2}) \\ &\quad + u_{j+1}(a_{j+\frac{1}{2}} \Delta t) + (u_x)_{j+1} (x_{j+\frac{1}{2}} \cdot a_{j+\frac{1}{2}} \Delta t - x_{j+1} \cdot a_{j+\frac{1}{2}} \Delta t + \frac{a_{j+\frac{1}{2}}^2 \cdot \Delta t^2}{2}). \end{aligned} \tag{A.54}$$

Appendix A Central schemes on uniform grids

Since  $x_{j+\frac{1}{2}} \cdot a_{j+\frac{1}{2}} \Delta t - x_j \cdot a_{j+\frac{1}{2}} \Delta t = \frac{\Delta x}{2} a_{j+\frac{1}{2}} \Delta t$  and  $x_{j+\frac{1}{2}} \cdot a_{j+\frac{1}{2}} \Delta t - x_{j+1} \cdot a_{j+\frac{1}{2}} \Delta t = -\frac{\Delta x}{2} a_{j+\frac{1}{2}} \Delta t$ , Eq. (A.54) will be simplified as:

$$\begin{aligned}
 I &= \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} u(x, t^n) dx \\
 &= (u_j + u_{j+1})(a_{j+\frac{1}{2}} \Delta t) + ((u_x)_j - (u_x)_{j+1}) \frac{\Delta x}{2} \cdot a_{j+\frac{1}{2}} \cdot \Delta t \\
 &\quad + ((u_x)_{j+1} - (u_x)_j) \frac{a_{j+\frac{1}{2}}^2 \Delta t^2}{2} \\
 &= (a_{j+\frac{1}{2}} \Delta t) [(u_j + u_{j+1}) + \frac{\Delta x - a_{j+\frac{1}{2}} \Delta t}{2} ((u_x)_j - (u_x)_{j+1})].
 \end{aligned} \tag{A.55}$$

Let  $\Delta x_{j+\frac{1}{2}} := x_{j+\frac{1}{2},r} - x_{j+\frac{1}{2},l} = 2(a_{j+\frac{1}{2}} \Delta t)$ , so  $a_{j+\frac{1}{2}} \Delta t = \frac{\Delta x_{j+\frac{1}{2}}}{2}$ . Then  $I$  can be written as:

$$I = \Delta x_{j+\frac{1}{2}} \left( \frac{u_j + u_{j+1}}{2} + \frac{\Delta x - a_{j+\frac{1}{2}} \Delta t}{4} ((u_x)_j - (u_x)_{j+1}) \right). \tag{A.56}$$

So the solution at  $t_{n+1}$  can be obtained by replacing  $I$  in Eq. (A.52), as:

$$\begin{aligned}
 \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} u(x, t^{n+1}) dx &= \Delta x_{j+\frac{1}{2}} \left( \frac{u_j + u_{j+1}}{2} + \frac{\Delta x - a_{j+\frac{1}{2}} \Delta t}{4} ((u_x)_j - (u_x)_{j+1}) \right) \\
 &\quad - \int_{t^n}^{t^{n+1}} [F(x_{j+\frac{1}{2},r}, t) - F(x_{j+\frac{1}{2},l}, t)] dt.
 \end{aligned} \tag{A.57}$$

The flux integrals can be approximated by the the midpoint rule. Let the average solution at  $t_{n+1}$  is denoted as:

$$w_{j+\frac{1}{2}}^{n+1} = \frac{1}{\Delta x_{j+\frac{1}{2}}} \int_{x_{j+\frac{1}{2},l}}^{x_{j+\frac{1}{2},r}} u(x, t^{n+1}) dx. \tag{A.58}$$

Then the evolved average solution on non-smooth zone is [22]:

$$w_{j+\frac{1}{2}}^{n+1} = \frac{u_j + u_{j+1}}{2} + \frac{\Delta x - a_{j+\frac{1}{2}}^n \Delta t}{4} ((u_x)_j^n - (u_x)_{j+1}^n) - \frac{1}{2a_{j+\frac{1}{2}}^n} [F(u_{j+\frac{1}{2},r}^{n+\frac{1}{2}}) - F(u_{j+\frac{1}{2},l}^{n+\frac{1}{2}})]. \tag{A.59}$$

ii **The smooth interval**  $[x_{j-\frac{1}{2},r}^n, x_{j+\frac{1}{2},l}^n] \times [t^n, t^{n+1}]$ : We follow the previous steps

again, beginning by integration on the smooth spatio-temporal volume.

$$\begin{aligned}
 \int_{t^n}^{t^{n+1}} \int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} \left( \frac{\partial u}{\partial t} + \frac{\partial F}{\partial x} \right) dx dt &= 0, \\
 \int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} [u(x, t^{n+1}) - u(x, t^n)] dx + \int_{t^n}^{t^{n+1}} [F(x_{j+\frac{1}{2},l}, t) - F(x_{j-\frac{1}{2},r}, t)] dt &= 0, \\
 \int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} u(x, t^{n+1}) dx - \int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} u(x, t^n) dx + \int_{t^n}^{t^{n+1}} [F(x_{j+\frac{1}{2},l}, t) - F(x_{j-\frac{1}{2},r}, t)] dt &= 0.
 \end{aligned} \tag{A.60}$$

Let  $\int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} u(x, t^n) dx$  is referred as  $J$ . The calculation of the cell averages at  $t^n$  on the spatial interval  $[x_{j-\frac{1}{2},r}^n, x_{j+\frac{1}{2},l}^n]$  is:

$$\begin{aligned}
 J &= \int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} u(x, t^n) dx = \int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} [u_j^n + (x - x_j)(u_x)_j] dx \\
 &= u_j^n (x_{j+\frac{1}{2},l} - x_{j-\frac{1}{2},r}) + (u_x)_j \left( \frac{x_{j+\frac{1}{2},l}^2}{2} - \frac{x_{j-\frac{1}{2},r}^2}{2} - (x_j x_{j+\frac{1}{2},l} - x_j x_{j-\frac{1}{2},r}) \right).
 \end{aligned} \tag{A.61}$$

For simplicity, the following relationships are considered:

$$\begin{aligned}
 x_{j+\frac{1}{2},l} &:= x_{j+\frac{1}{2}} - a_{j+\frac{1}{2}} \Delta t, & x_{j-\frac{1}{2},r} &:= x_{j-\frac{1}{2}} + a_{j-\frac{1}{2}} \Delta t, \\
 x_{j+\frac{1}{2}} &:= x_j + \frac{\Delta x}{2}, & x_{j-\frac{1}{2}} &:= x_j - \frac{\Delta x}{2}, \\
 x_j &= \frac{x_{j+\frac{1}{2}} + x_{j-\frac{1}{2}}}{2}, & \Delta x &:= x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}.
 \end{aligned}$$

Also the width of the smooth zone around  $x_j$  is  $\Delta x_j = x_{j+\frac{1}{2},l}^n - x_{j-\frac{1}{2},r}^n = \Delta x - \Delta t(a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n)$ . Considering the above mentioned equations, the simplified form of Eq. (A.61) is:

$$\int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} u(x, t^n) dx = \Delta x_j [u_j^n + (u_x)_j \frac{\Delta t}{2} (a_{j-\frac{1}{2}} - a_{j+\frac{1}{2}})]. \tag{A.62}$$

By substituting Eq. (A.62) in (A.60), the average of the evolved solution(at  $t_{n+1}$ ) is:

$$\int_{x_{j-\frac{1}{2},r}}^{x_{j+\frac{1}{2},l}} u(x, t^{n+1}) dx = \Delta x_j [u_j^n + (u_x)_j \frac{\Delta t}{2} (a_{j-\frac{1}{2}} - a_{j+\frac{1}{2}})] - \int_{t^n}^{t^{n+1}} [F(x_{j+\frac{1}{2},l}, t) - F(x_{j+\frac{1}{2},r}, t)] dt. \tag{A.63}$$

Appendix A Central schemes on uniform grids

The average evolved solution  $u(x, t^{n+1})$  on the strip  $[x_{j-\frac{1}{2},r}^n, x_{j+\frac{1}{2},l}^n]$  is denoted as:

$$w_j^{n+1} = \frac{1}{\Delta x_j} \int_{x_{j-\frac{1}{2},r}^n}^{x_{j+\frac{1}{2},l}^n} u(x, t^{n+1}) dx. \quad (\text{A.64})$$

Using the mid-point rule for the time integration, Eq. (A.63),  $w_j^{n+1}$  can be written as [22]:

$$w_j^{n+1} = u_j^n + (u_x)_j \frac{\Delta t}{2} (a_{j-\frac{1}{2}}^n - a_{j+\frac{1}{2}}^n) - \frac{\lambda}{1 - \lambda(a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n)} [F(u_{j+\frac{1}{2},l}^{n+\frac{1}{2}}) - F(u_{j-\frac{1}{2},r}^{n+\frac{1}{2}})], \quad (\text{A.65})$$

where  $\lambda := \frac{\Delta t}{\Delta x}$ .

In the above formulation, the midpoint values  $u^{n+\frac{1}{2}}$  can be obtained by the Taylor expansion as:

$$u_{j+\frac{1}{2},l}^{n+\frac{1}{2}} = u_{j+\frac{1}{2},l}^n - \frac{\Delta t}{2} F(u_{j+\frac{1}{2},l}^n)_x, \quad u_{j+\frac{1}{2},r}^{n+\frac{1}{2}} = u_{j+\frac{1}{2},r}^n - \frac{\Delta t}{2} F(u_{j+\frac{1}{2},r}^n)_x, \quad (\text{A.66})$$

$$u_{j+\frac{1}{2},l}^n = u_j^n + \Delta x (u_x)_j^n \left(\frac{1}{2} - \lambda a_{j+\frac{1}{2}}^n\right), \quad u_{j+\frac{1}{2},r}^n = u_{j+1}^n - \Delta x (u_x)_{j+1}^n \left(\frac{1}{2} - \lambda a_{j+\frac{1}{2}}^n\right).$$

To have a non-staggered formulation, the projection step is used (projection of solutions on the original grid). For this, at first, a piecewise interpolation is done on the results obtained in the evolution step. The piecewise interpolation contains both linear variations (in non-smooth zones) and constant values (in smooth areas). The piecewise interpolation is:

$$\tilde{w}(x, t^{n+1}) = \sum_j ([w_{j+\frac{1}{2}}^{n+1} + (u_x)_{j+\frac{1}{2}}^{n+1} (x - x_{j+\frac{1}{2}})] 1_{[x_{j+\frac{1}{2},l}^n, x_{j+\frac{1}{2},r}^n]} + w_j^{n+1} 1_{[x_{j-\frac{1}{2},r}^n, x_{j+\frac{1}{2},l}^n]}), \quad (\text{A.67})$$

where  $1_{[a,b]}(x) = 1$  for  $a \leq x \leq b$  and  $1_{[a,b]}(x) = 0$  elsewhere. Here, the exact spatial derivatives  $u_x(x_{j+\frac{1}{2}}, t^{n+1})$  are approximated by using the *MINMOD* limiter:

$$(u_x)_{j+\frac{1}{2}}^{n+1} = \text{MINMOD} \left( \frac{w_{j+1}^{n+1} - w_{j+\frac{1}{2}}^{n+1}}{\frac{\Delta x_{j+1}}{2} + a_{j+\frac{1}{2}}^n \Delta t}, \frac{w_{j+\frac{1}{2}}^{n+1} - w_j^{n+1}}{\frac{\Delta x_j}{2} + a_{j+\frac{1}{2}}^n \Delta t} \right), \quad (\text{A.68})$$

or:

$$(u_x)_{j+\frac{1}{2}}^{n+1} = \frac{2}{\Delta x} \cdot \text{MINMOD} \left( \frac{w_{j+1}^{n+1} - w_{j+\frac{1}{2}}^{n+1}}{1 + \lambda(a_{j+\frac{1}{2}}^n - a_{j+\frac{3}{2}}^n)}, \frac{w_{j+\frac{1}{2}}^{n+1} - w_j^{n+1}}{1 + \lambda(a_{j+\frac{1}{2}}^n - a_{j-\frac{1}{2}}^n)} \right). \quad (\text{A.69})$$

Let us define:

$$\Delta x_R = \frac{\Delta x_{j+1}}{2} + a_{j+1/2}^n \Delta t \quad \text{and} \quad \Delta x_L = \frac{\Delta x_j}{2} + a_{j+1/2}^n \Delta t.$$

Then it is easy to show that:

$$\Delta x_L = \frac{\Delta x}{2} - \frac{\Delta t}{2}(a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n) + a_{j+\frac{1}{2}}^n \Delta t = \frac{\Delta x}{2} + \frac{\Delta t}{2}(a_{j+\frac{1}{2}}^n - a_{j-\frac{1}{2}}^n) = \frac{2}{\Delta x}(1 + \lambda(a_{j+\frac{1}{2}}^n - a_{j-\frac{1}{2}}^n)),$$

$$\Delta x_R = \frac{\Delta x}{2} - \frac{\Delta t}{2}(a_{j+\frac{1}{2}}^n + a_{j+\frac{3}{2}}^n) + a_{j+\frac{1}{2}}^n \Delta t = \frac{\Delta x}{2} + \frac{\Delta t}{2}(a_{j+\frac{1}{2}}^n - a_{j+\frac{3}{2}}^n) = \frac{2}{\Delta x}(1 + \lambda(a_{j+\frac{1}{2}}^n - a_{j+\frac{3}{2}}^n)).$$

Finally, by averaging  $\tilde{w}(x, t^n)$  over  $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ , the fully-discrete form of the second-order central KT scheme can be obtained as [22]:

$$\begin{aligned} \bar{u}_j^{n+1} = u_j^{n+1} &= \frac{1}{\Delta x} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \tilde{w}(x, t^{n+1}) dx = \lambda a_{j-\frac{1}{2}}^n w_{j-\frac{1}{2}}^{n+1} + [1 - \lambda(a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n)] w_j^{n+1} \\ &\quad + \lambda a_{j+\frac{1}{2}}^n w_{j+\frac{1}{2}}^{n+1} + \frac{\Delta x}{2} [(\lambda a_{j-\frac{1}{2}}^n)^2 (u_x)_{j-\frac{1}{2}}^{n+1} - (\lambda a_{j+\frac{1}{2}}^n)^2 (u_x)_{j+\frac{1}{2}}^{n+1}]. \end{aligned} \quad (\text{A.70})$$

**The KT scheme - reduction to semi-discrete formulation** The KT scheme is the first central high-resolution scheme admitting the semi-discrete form. This is because its truncation error is independent of the term  $\mathcal{O}(1/\Delta t)$  [22]. A benefit of the semi-discrete form is that it lets one using of traditional time-stepping methods, developed for ODEs, such as the TVD Runge-Kutta method.

The fully-discrete central scheme shown in Eq. (A.70) is used to provide corresponding semi-discrete form. For this, Eq. (A.70) is rewritten as:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{a_{j-\frac{1}{2}}^n}{\Delta x} w_{j-\frac{1}{2}}^{n+1} + \left( \frac{1}{\Delta x} - \frac{a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n}{\Delta x} \right) w_j^{n+1} + \frac{a_{j+\frac{1}{2}}^n}{\Delta x} w_{j+\frac{1}{2}}^{n+1} - \frac{1}{\Delta t} u_j^n + \mathcal{O}(\lambda). \quad (\text{A.71})$$

By substituting Eq. (A.59) in Eq. (A.72), we have:

$$\begin{aligned}
 \frac{u_j^{n+1} - u_j^n}{\Delta t} &= \frac{a_{j-\frac{1}{2}}^n}{2\Delta x} (u_{j-1}^n - u_j^n) + \frac{1}{4} a_{j-\frac{1}{2}}^n ((u_x)_{j-1}^n - (u_x)_j^n) - \frac{1}{2\Delta x} [F(u_{j-\frac{1}{2},r}^{n+\frac{1}{2}}) - F(u_{j-\frac{1}{2},l}^{n+\frac{1}{2}})] \\
 &\quad - \frac{a_{j-\frac{1}{2}}^n + a_{j+\frac{1}{2}}^n}{\Delta x} u_j^n + \frac{1}{2} (a_{j-\frac{1}{2}}^n - a_{j+\frac{1}{2}}^n) (u_x)_j^n - \frac{1}{\Delta x} [F(u_{j+\frac{1}{2},l}^{n+\frac{1}{2}}) - F(u_{j-\frac{1}{2},r}^{n+\frac{1}{2}})] \\
 &\quad + \frac{a_{j+\frac{1}{2}}^n}{2\Delta x} (u_j^{n+1} + u_j^n) + \frac{1}{4} a_{j+\frac{1}{2}}^n ((u_x)_j^n - (u_x)_{j+1}^n) - \frac{1}{2\Delta x} [F(u_{j+\frac{1}{2},r}^{n+\frac{1}{2}}) - F(u_{j+\frac{1}{2},l}^{n+\frac{1}{2}})] + \mathcal{O}(\lambda) \\
 &= \frac{1}{2\Delta x} \left( -[(F(u_{j+\frac{1}{2},r}^{n+\frac{1}{2}}) + F(u_{j+\frac{1}{2},l}^{n+\frac{1}{2}})) - (F(u_{j-\frac{1}{2},r}^{n+\frac{1}{2}}) + F(u_{j-\frac{1}{2},l}^{n+\frac{1}{2}}))] \right. \\
 &\quad + \frac{a_{j+\frac{1}{2}}^n}{\Delta x} [(u_{j+1}^n - \frac{\Delta x}{2} (u_x)_{j+1}^n) - (u_j^n + \frac{\Delta x}{2} (u_x)_j^n)] \\
 &\quad \left. - \frac{a_{j-\frac{1}{2}}^n}{\Delta x} [(u_j^n - \frac{\Delta x}{2} (u_x)_j^n) - (u_{j-1}^n + \frac{\Delta x}{2} (u_x)_{j-1}^n)] \right) + \mathcal{O}(\lambda).
 \end{aligned} \tag{A.72}$$

By using the Taylor expansion in time and linear variation in the spatial domain, in the case  $\Delta t \rightarrow 0$ , the mid values can be approximated as (as  $\Delta t \rightarrow 0$ , then  $u^{n+1/2} \approx u^n$ ):

$$\begin{aligned}
 u_{j+\frac{1}{2},r}^{n+\frac{1}{2}} &\rightarrow u_{j+1}(t) - \frac{\Delta x}{2} (u_x)_{j+1}(t) = u_{j+\frac{1}{2}}^r(t), \\
 u_{j+\frac{1}{2},l}^{n+\frac{1}{2}} &\rightarrow u_j(t) + \frac{\Delta x}{2} (u_x)_j(t) = u_{j+\frac{1}{2}}^l(t).
 \end{aligned} \tag{A.73}$$

The semi-discrete formulation can be obtained if in Eq. (A.72),  $\Delta t \rightarrow 0$ . Regarding Eq. (A.72) and the above-mentioned simplifications, the semi-discrete form is:

$$\begin{aligned}
 \frac{d}{dt} u_j(t) &= - \frac{[F(u_{j+\frac{1}{2}}^r(t)) + F(u_{j+\frac{1}{2}}^l(t))] - [F(u_{j-\frac{1}{2}}^r(t)) + F(u_{j-\frac{1}{2}}^l(t))]}{2\Delta x} \\
 &\quad + \frac{1}{2\Delta x} \left( a_{j+\frac{1}{2}}^n(t) [u_{j+\frac{1}{2}}^r(t) - u_{j+\frac{1}{2}}^l(t)] - a_{j-\frac{1}{2}}^n(t) [u_{j-\frac{1}{2}}^r(t) - u_{j-\frac{1}{2}}^l(t)] \right),
 \end{aligned} \tag{A.74}$$

or:

$$\frac{d}{dt} u_j(t) = - \frac{F_{j+\frac{1}{2}}^*(t) - F_{j-\frac{1}{2}}^*(t)}{\Delta x}, \tag{A.75}$$

where:

$$F_{j+\frac{1}{2}}^*(t) = \frac{F(u_{j+\frac{1}{2}}^r(t)) + F(u_{j+\frac{1}{2}}^l(t))}{2} - \frac{a_{j+\frac{1}{2}}^n(t)}{2} [u_{j+\frac{1}{2}}^r(t) - u_{j+\frac{1}{2}}^l(t)], \tag{A.76}$$

$F_{j+\frac{1}{2}}^*$  denotes the corrected flux obtained by a proper combination of the estimated

solutions at the right ( $u_{j+\frac{1}{2}}^r(t)$ ) and left ( $u_{j+\frac{1}{2}}^l(t)$ ) sides of  $j$ th cell at the edge  $x_{j+1/2}$ . These reconstructed solutions can be obtained as:

$$\begin{aligned} u_{j+\frac{1}{2}}^r &= u_{j+1}(t) - \frac{\Delta x}{2}(u_x)_{j+1}(t), \\ u_{j+\frac{1}{2}}^l &= u_j(t) + \frac{\Delta x}{2}(u_x)_j(t). \end{aligned} \tag{A.77}$$

In order to approximate the derivatives  $(u_x)_j^n$ , a limited slope such as the generalized *MINMOD*-based one will be used; its definition is:

$$(u_x)_j^n = \text{MINMOD} \left( \theta \frac{u_j^n - u_{j-1}^n}{\Delta x}, \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x}, \theta \frac{u_{j+1}^n - u_j^n}{\Delta x} \right), \quad 1 \leq \theta \leq 2, \tag{A.78}$$

where  $\theta = 1$  and  $\theta = 2$  lead to the most and the least dissipative results, respectively.

**Example:** Numerical simulation of the Burgers' equation with the KT scheme is presented here, Figure A.9; the CFL condition is the same as the LxF and NT schemes. It is clear that numerical dissipation is reduced considerably.



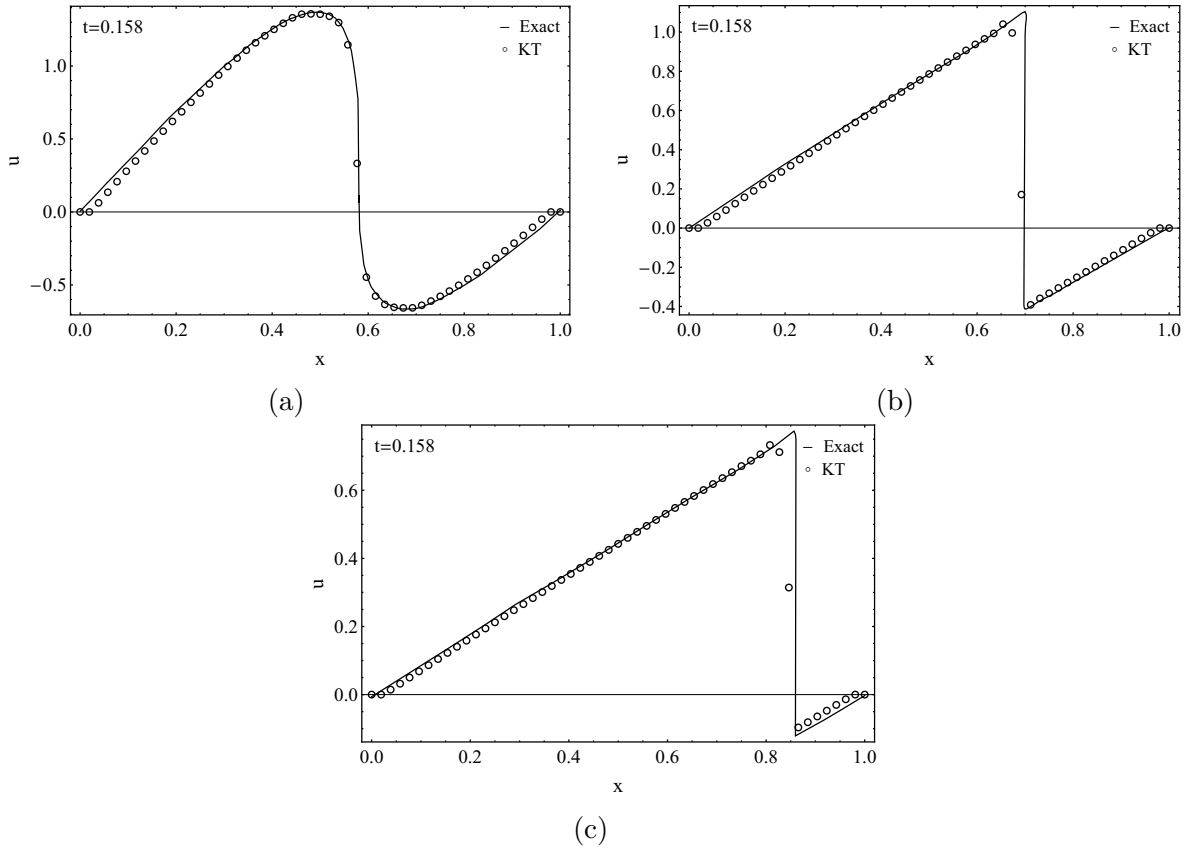


Figure A.9: Numerical results of the Burgers' equation by the KT method at  $t=0.158$ ,  $t=0.5$  and  $t=1$ ; solid lines and hollow circles denote the exact and numerical solutions, respectively.

# Appendix B

## The NVSF upwind scheme

In this appendix the upwind-based NVSF formulation is reviewed. Let us assume the hyperbolic equation  $u(x, t)_t + f(u(x, t))_x = 0$  with propagating velocity  $a(x) := \partial f / \partial u$ . To preserve the monotone reconstruction of cell-interface flux,  $f_{i+1/2}$ , this flux is obtained by the following interpolation function [20, 50]:

$$f_{i+1/2} = f_U + \hat{f}_{i+1/2} (f_D - f_U), \quad (\text{B.1})$$

where:  $f_{i+1/2}$  denotes the flux at the edge point  $x_{i+1/2}$ ;  $f_U$  and  $f_D$  denote upstream and downstream fluxes, respectively. These fluxes are determined by  $a_i$  as (see Figure B.1):

1. If  $a_{i+1/2} > 0$ :  $f_U = f_{i-1}$ ,  $f_D = f_{i+1}$  with central cell center  $x_P := x_i$  (for the cell  $I_i$ ),
2. If  $a_{i+1/2} < 0$ :  $f_U = f_{i+2}$  and  $f_D = f_i$  with central cell center  $x_P := x_{i+1}$  (for the cell  $I_{i+1}$ ),

where  $a_{i+1/2} := \frac{a_i + a_{i+1}}{2}$ . And, the parameter  $\hat{f}_{i+1/2}$  shows a limited flux (obtained by a nonlinear limiter).

For the MINMOD limiter,  $\hat{f}_{i+1/2}$  can be estimated as [50]:

$$\hat{f}_{i+1/2} = \text{Max} \left[ \hat{f}_P, \text{Min} \left( \frac{\hat{x}_{i+1/2}}{\hat{x}_P} \hat{f}_P, \frac{1 - \hat{x}_{i+1/2}}{1 - \hat{x}_P} \hat{f}_P + \frac{\hat{x}_{i+1/2} - \hat{x}_P}{1 - \hat{x}_P} \right) \right], \quad (\text{B.2})$$

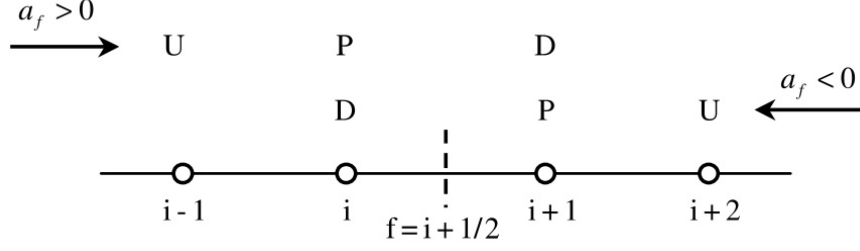


Figure B.1: Definition of the upstream and downstream cells based on the velocity  $a_{i+1/2} = a_f$ .

and for the SMART limiter<sup>13</sup>,  $\hat{f}_{i+1/2}$  is [50]:

$$\hat{f}_{i+1/2} = \text{Max} \left[ \hat{f}_P, \text{Min} \left( \frac{\hat{x}_{i+1/2} (1 - 3\hat{x}_P + 2\hat{x}_{i+1/2})}{\hat{x}_P (1 - \hat{x}_P)} \hat{f}_P, \frac{\hat{x}_{i+1/2} (1 - \hat{x}_{i+1/2})}{\hat{x}_P (1 - \hat{x}_P)} \hat{f}_P + \frac{\hat{x}_{i+1/2} (\hat{x}_{i+1/2} - \hat{x}_P)}{1 - \hat{x}_P}, 1 \right) \right], \quad (\text{B.3})$$

where the normalized variables  $\hat{f}_P$ ,  $\hat{x}_P$  and  $\hat{x}_{i+1/2}$  are defined as:

$$\begin{aligned} \hat{f}_P &:= \frac{f_P - f_U}{f_D - f_U}, \\ \hat{x}_P &:= \frac{x_P - x_U}{x_D - x_U}, \\ \hat{x}_{i+1/2} &:= \frac{x_{i+1/2} - x_U}{x_D - x_U}. \end{aligned} \quad (\text{B.4})$$

Using the upwind formulation (Eq. (4.6)), the evolved solution  $u^{n+1}$  can be obtained based on  $u^n$  and edge flux  $f_{i\pm 1/2}$ .

<sup>13</sup>The SMART limiter, itself, is defined as [231]:

$$\phi_{smart} := \text{Max} \left[ 0, \text{Min} \left[ 2r, \left( \frac{1+3r}{4} \right) \right] \right],$$

where  $r_i := r(x_i) = \frac{u_i - u_{i-1}}{u_{i+1} - u_i}$  denotes the ratio of successive gradients.

# Appendix C

## The TVD condition, Entropy functions and E-schemes

In this appendix, the following concepts are reviewed:

1. The TVD conditions and limiters,
2. Entropy functions and entropy conditions,
3. E-schemes and slope limiters.

### C.1 The TVD conditions and limiters

This subsection contains:

1. Deriving of the global TVD (positivity) conditions,
2. Some local TVD conditions based on the positivity relationships,
3. Essential constraints for designing of flux limiters satisfying the TVD conditions,
4. Relationship between flux and slope limiters,

#### C.1.1 Global TVD conditions

**Theorem 4.** *In order to the fully-discrete form (4.40) is TVD, the following conditions are sufficient:*

$$\begin{aligned}C_{i+1/2}^+ + C_{i+1/2}^- &\leq 1, \\C_{i+1/2}^- &\geq 0, \\C_{i+1/2}^+ &\geq 0.\end{aligned}\tag{C.1}$$

**Proof.** Using Eq. (4.40) for both  $u_i^{n+1}$  and  $u_i^n$ , the difference  $u_{i+1}^{n+1} - u_i^{n+1}$  can be written as:

$$\begin{aligned} u_{i+1}^{n+1} - u_i^{n+1} &= (u_{i+1}^n - u_i^n) + C_{i+3/2}^+(u_{i+2}^n - u_{i+1}^n) - C_{i+1/2}^+(u_{i+1}^n - u_i^n) \\ &\quad - C_{i+1/2}^-(u_{i+1}^n - u_i^n) + C_{i-1/2}^-(u_i^n - u_{i-1}^n) \\ &= (u_{i+1}^n - u_i^n) \left[ 1 - C_{i+1/2}^+ - C_{i+1/2}^- \right] + (u_i^n - u_{i-1}^n) C_{i-1/2}^- + (u_{i+2}^n - u_{i+1}^n) C_{i+3/2}^+. \end{aligned} \quad (\text{C.2})$$

By summation the both sides of the above equation over  $i$ , we have:

$$\begin{aligned} \sum_i (u_{i+1}^{n+1} - u_i^{n+1}) &= \\ &\sum_i \left( (u_{i+1}^n - u_i^n) \left[ 1 - C_{i+1/2}^+ - C_{i+1/2}^- \right] \right) + \sum_i \left( (u_i^n - u_{i-1}^n) C_{i-1/2}^- \right) + \sum_i \left( (u_{i+2}^n - u_{i+1}^n) C_{i+3/2}^+ \right) \\ &= \sum_i \left( (u_{i+1}^n - u_i^n) \left[ 1 - C_{i+1/2}^+ - C_{i+1/2}^- \right] \right) + \sum_i \left( (u_{i+1}^n - u_i^n) C_{i+1/2}^- \right) + \sum_i \left( (u_{i+1}^n - u_i^n) C_{i+1/2}^+ \right). \end{aligned} \quad (\text{C.3})$$

So it is clear, if the conditions (C.1) are met, then

$$\begin{aligned} \sum_i \left| u_{i+1}^{n+1} - u_i^{n+1} \right| &\leq \\ &\sum_i \left( \left| u_{i+1}^n - u_i^n \right| \left[ 1 - C_{i+1/2}^+ - C_{i+1/2}^- \right] \right) + \sum_i \left( \left| u_{i+1}^n - u_i^n \right| C_{i+1/2}^- \right) + \sum_i \left( \left| u_{i+1}^n - u_i^n \right| C_{i+1/2}^+ \right), \end{aligned} \quad (\text{C.4})$$

or:

$$\sum_i \left| u_{i+1}^{n+1} - u_i^{n+1} \right| \leq \sum_i \left| u_{i+1}^n - u_i^n \right|, \quad (\text{C.5})$$

and the solution is TVD. ■

**Theorem 5.** *In order to the semi-discrete form Eq. (4.42) is TVD, the following conditions are sufficient:*

$$\begin{aligned} C_{i+1/2}^+ + C_{i+1/2}^- &\leq 1, \\ C_{i+1/2}^- &\geq 0, \\ C_{i+1/2}^+ &\geq 0. \end{aligned} \quad (\text{C.6})$$

**Proof:** Let us rewrite the semi-discrete form Eq. (4.42) as:

$$\frac{d}{dt}u_i(t) = C_{i+1/2}^+ \Delta u_{i+1/2} - C_{i-1/2}^- \Delta u_{i-1/2}, \quad \Delta u_{i+1/2} := u_{i+1}(t) - u_i(t). \quad (\text{C.7})$$

Forward differencing of Eq. (C.7) leads:

$$\frac{d}{dt} \Delta u_{i+1/2} = \left( C_{i+3/2}^+ \Delta u_{i+3/2} - C_{i+1/2}^- \Delta u_{i+1/2} \right) - \left( C_{i+1/2}^+ \Delta u_{i+1/2} - C_{i-1/2}^- \Delta u_{i-1/2} \right). \quad (\text{C.8})$$

Multiplying Eq. (C.8) by the *Sign* function  $s_{i+1/2} := \text{Sign}(\Delta u_{i+1/2})$  and then by summing by parts, we have:

$$\begin{aligned} \frac{d}{dt} \sum_i \left| \Delta u_{i+1/2} \right| &= \sum_i \frac{d}{dt} s_{i+1/2} \Delta u_{i+1/2} \\ &= \sum_i \left( C_{i+3/2}^+ s_{i+1/2} \Delta u_{i+3/2} - C_{i+1/2}^- s_{i+1/2} \Delta u_{i+1/2} \right) \\ &\quad - \sum_i \left( C_{i+1/2}^+ s_{i+1/2} \Delta u_{i+1/2} - C_{i-1/2}^- s_{i+1/2} \Delta u_{i-1/2} \right) \\ &= \sum_i \left( C_{i+1/2}^+ s_{i-1/2} \Delta u_{i+1/2} - C_{i+1/2}^- s_{i+1/2} \Delta u_{i+1/2} \right) \\ &\quad - \sum_i \left( C_{i+1/2}^+ s_{i+1/2} \Delta u_{i+1/2} - C_{i+1/2}^- s_{i+3/2} \Delta u_{i+1/2} \right), \end{aligned} \quad (\text{C.9})$$

or:

$$\begin{aligned} \frac{d}{dt} \sum_i \left| \Delta u_{i+1/2} \right| &= \sum_i \frac{d}{dt} s_{i+1/2} \Delta u_{i+1/2} \\ &= - \sum_i \left[ \left( s_{i+1/2} - s_{i-1/2} \right) C_{i+1/2}^+ + \left( s_{i+1/2} - s_{i+3/2} \right) C_{i+1/2}^- \right] \Delta u_{i+1/2}, \end{aligned} \quad (\text{C.10})$$

and it is clear that:  $\Delta u_{i+1/2} = s_{i+1/2} \left| \Delta u_{i+1/2} \right|$  and  $s_{i+1/2}^2 = 1$ <sup>14</sup>. So, Eq. (C.10) becomes:

$$\frac{d}{dt} TV [u(t)] = - \sum_i \left[ \left( 1 - s_{i+1/2} s_{i-1/2} \right) C_{i+1/2}^+ + \left( 1 - s_{i+1/2} s_{i+3/2} \right) C_{i+1/2}^- \right] \left| \Delta u_{i+1/2} \right|. \quad (\text{C.11})$$

Coefficients  $\left( 1 - s_{i+1/2} s_{i-1/2} \right)$  and  $\left( 1 - s_{i+1/2} s_{i+3/2} \right)$  are either 0 or 2. So the summation in RHS is positive if  $C_{i+1/2}^+ \geq 0$  and  $C_{i+1/2}^- \geq 0$ . In this case the method remains TVD.

<sup>14</sup>The signum function at zero is defined to be  $\pm 1$ , so that its square equals 1.

■

### C.1.2 Local TVD conditions

In the following, for the following semi-discrete form

$$\frac{d}{dt}u_i(t) = -\frac{1}{\Delta x_i} [F_{i+1/2} - F_{i-1/2}], \quad \Delta x_i = x_{i+1/2} - x_{i-1/2}, \quad (\text{C.12})$$

some local TVD conditions are provided.

**Lemma 1.** *The semi-discrete scheme (C.12) is TVD if, the numerical fluxes  $F_{i+1/2}$  satisfy:*

$$\begin{aligned} F_{i+1/2} &\geq F_{i-1/2} & u_i \text{ is a maximum value,} \\ F_{i+1/2} &\leq F_{i-1/2} & u_i \text{ is a minimum value.} \end{aligned} \quad (\text{C.13})$$

**Proof:** At first, we forward-difference (C.12) as:

$$\frac{d}{dt}\Delta u_{i+1/2} = -\frac{1}{\Delta x_{i+1}} [F_{i+3/2} - F_{i+1/2}] + \frac{1}{\Delta x_i} [F_{i+1/2} - F_{i-1/2}], \quad (\text{C.14})$$

multiplying by  $s_{i+1/2} := \text{Sign}(\Delta u_{i+1/2})$  and then summing by parts, we have:

$$\begin{aligned} \sum_i \frac{d}{dt}\Delta u_{i+1/2} &= \sum_i \left( -\frac{1}{\Delta x_{i+1}} [F_{i+3/2} - F_{i+1/2}] s_{i+1/2} + \frac{1}{\Delta x_i} [F_{i+1/2} - F_{i-1/2}] s_{i+1/2} \right), \\ &= \sum_i \left( -\frac{1}{\Delta x_i} [F_{i+1/2} - F_{i-1/2}] s_{i-1/2} + \frac{1}{\Delta x_i} [F_{i+1/2} - F_{i-1/2}] s_{i+1/2} \right), \end{aligned} \quad (\text{C.15})$$

or:

$$\frac{d}{dt}TV[u(t)] = \sum_i \frac{1}{\Delta x_i} [s_{i+1/2} - s_{i-1/2}] [F_{i+1/2} - F_{i-1/2}]. \quad (\text{C.16})$$

To have a TVD solution, RHS of the above-equation should be zero or negative. Now, for different possible solutions, we check these possibilities.

1. *Monotone* solution: in this case we have:  $s_{i+1/2} = s_{i-1/2}$ , and so RHS is zero,
2.  $u_i$  is *maximum*: then  $s_{i+1/2} < 0$  and  $s_{i-1/2} > 0$ ; so it is necessary that:  
 $[F_{i+1/2} - F_{i-1/2}] \geq 0$ , or  $F_{i+1/2} \geq F_{i-1/2}$ ,

3.  $u_i$  is *minimum*: then  $s_{i+1/2} > 0$  and  $s_{i-1/2} < 0$ ; so it is necessary that:

$$\left[ F_{i+1/2} - F_{i-1/2} \right] \leq 0, \text{ or } F_{i+1/2} \leq F_{i-1/2}.$$

■

The above-mentioned proof shows that the TVD condition is satisfied automatically in monotone regions. And so we need to check the TVD condition around some isolated points (extrema points).

### C.1.3 Designing of flux/slope limiters based on the global TVD conditions

For satisfying the TVD condition, the concept of flux/slope limiters can be used. Here, fluxes (in the vicinity of high gradient solutions) are limited in a way that forming of spurious oscillations are prevented. In this regard, a high-order flux  $F^H$  can be decomposed to a low-order  $F^L$  and a correction (residual) term as [68]:

$$F_i^H = F_i^L + (F_i^H - F_i^L). \quad (\text{C.17})$$

The high-order flux is used in smooth regions while  $F^L$  is utilized around discontinuities. For such switching, a flux limiter function can be used, as:

$$F_i^H = F_i^L + \phi_i \left[ F_i^H - F_i^L \right], \quad (\text{C.18})$$

where  $\phi_i := \phi(u_j; x_i)$  denotes a limiter and is defined for all points  $x_j$  where  $|j - i| \leq c$  (for the three point stencil,  $c$  is 1). In smooth regions, the limiter  $\phi_i$  approaches 1, and around discontinuities, it is near to zero.

For deriving TVD conditions for the limiters  $\phi_i$ , a linear advection equation,  $u_t + (au)_x = 0$ , is assumed, where  $a$  is a constant velocity and here we assume:  $a > 0$ . Using the *Lax-Wendroff*<sup>15</sup> discretization with second-order-accuracy, the advection equation

---

<sup>15</sup>The Lax-Wendroff method: Let us assume the linear advection equation  $u_t + au_x = 0$ , where  $a$  is a constant. Using the Taylor expansion,  $u^{n+1} := u(x, t_n + \Delta t)$  can be written as:  $u^{n+1} = u^n + \Delta t \dot{u}^n + \frac{\Delta t^2}{2} \ddot{u}^n$ ; but:  $\dot{u}(t) = -au_x$  and  $\ddot{u} = -a u_{tx} = -a(u_t)_x = -a(-au_x)_x = a^2 u_{xx}$ . Using the central approximations for the spatial derivatives, we have:  $(u_j^n)_x = \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x}$  and  $(u_j^n)_{xx} = \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x^2}$ , where  $u_j^n := u(x_j, t^n)$ . By substituting  $\dot{u}$  and  $\ddot{u}$  in the Taylor expansion, we have:  
 $u_j^{n+1} = u_j^n - \frac{a\Delta t}{2\Delta x} (u_{j+1}^n - u_{j-1}^n) + \frac{a^2 \Delta t^2}{2\Delta x^2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n).$



becomes:

$$u_i^{n+1} = u_i^n - \frac{\nu}{2}(u_{i+1}^n - u_{i-1}^n) + \frac{\nu^2}{2}(u_{i+1}^n - 2u_i^n + u_{i-1}^n), \quad (\text{C.19})$$

where  $\nu := \frac{a\Delta t}{\Delta x}$ .

The next step is to rewrite Eq. (C.19) in the conservative<sup>16</sup> fully-discrete form. It is easy to show that on the spatio-temporal cell  $[x_{i-1/2}, x_{i+1/2}] \times [t^n, t^{n+1}]$ , the conservative fully-discrete form of the equation  $u_t + f(u)_x = 0$  is:  $u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} [F_{i+1/2}^n - F_{i-1/2}^n]$ , where  $u_i^n = \int_{x_{i-1/2}}^{x_{i+1/2}} u(x, t_n) dx / \Delta x$ , and  $F_{i+1/2}^n = \int_{t^n}^{t^{n+1}} f(x_{i+1/2}, t) dt / \Delta t$ ;  $F_{i+1/2}^n$  is known as the *numerical flux*.

By adding and subtracting the term  $(-\nu u_i^n + \nu u_{i-1}^n)$  to Eq. (C.19) and some simplifications, Eq. (C.19) can be written as:

$$u_i^{n+1} = u_i^n - \nu(u_i^n - u_{i-1}^n) - \frac{\nu}{2}(1 - \nu)(u_{i+1}^n - 2u_i^n + u_{i-1}^n). \quad (\text{C.20})$$

Regarding the conservative fully-discrete form, in Eq. (C.20), the numerical flux is:

$$F_{i+1/2}^n = au_i^n + \frac{a}{2}(1 - \nu)(u_{i+1}^n - u_i^n), \quad (\text{C.21})$$

where, the term  $au_i^n$  is the first order upwind approximation of the flux, or  $F^L = au_i^n$ . So, the concept of flux limiter can be inserted in Eq. (C.21) as:

$$F_{i+1/2}^n = au_i^n + \phi_i \left\{ \frac{a}{2}(1 - \nu)(u_{i+1}^n - u_i^n) \right\}. \quad (\text{C.22})$$

By substituting Eq. (C.22) in the conservative fully-discrete form and some simplifi-

---

<sup>16</sup>Regarding the conservation law  $u_t + f(u)_x = 0$ , the total quantity of a conserved variable in any region changes only due to flux through the boundaries. This leads to the integral form of the conservation law as:  $\int_a^b u(x, t_2) dx = \int_a^b u(x, t_1) dx - \left( \int_{t_1}^{t_2} f(u(b, t)) dt - \int_{t_1}^{t_2} f(u(a, t)) dt \right)$  for every  $a, b, t_1$  and  $t_2$  values. Let  $u$  has constant values outside the finite interval  $a \leq x \leq b$  during  $t_1 \leq t \leq t_2$ ; say  $u(x < a) \equiv u_{-\infty}$  and  $u(x > b) \equiv u_{+\infty}$ , then the conserved equation becomes:  $\int_a^b u(x, t_2) dx = \int_a^b u(x, t_1) dx - (t_2 - t_1) [f(u_{+\infty}) - f(u_{-\infty})]$ . If  $u_{+\infty} = u_{-\infty}$ , then the integral  $\int_a^b u(x, t) dx$  remains constant in time. Discretizing the conservation law for the case  $a = x_{i-1/2}$  and  $b = x_{i+1/2}$ , leads to  $u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (F_{i+1/2}^n - F_{i-1/2}^n)$ . This conservative method has also a global conservation form. By summation over cells  $I_i$ , we have:  $\Delta x \sum_i u_i^{n+1} = \Delta x \sum_i u_i^n - \Delta t \sum_i (F_{i+1/2}^n - F_{i-1/2}^n)$ . The fluxes in the summation drop out each other except those at the extreme edges, say  $x_{J-1/2}$  and  $x_{K+1/2}$ , then we have:  $\Delta x \sum_{i=J}^K u_i^{n+1} = \Delta x \sum_{i=J}^K u_i^n - \Delta t (F_{K+1/2}^n - F_{J-1/2}^n)$ . And if  $F_{K+1/2}^n = F_{J-1/2}^n$ , then  $\sum_{i=J}^K u_i^{n+1} = \sum_{i=J}^K u_i^n$ ; or equivalently  $\int_{x_{J-1/2}}^{x_{K+1/2}} u(x, t^{n+1}) dx = \int_{x_{J-1/2}}^{x_{K+1/2}} u(x, t^n) dx$ , and so the discrete method is *conservative*.

cations we have:

$$u_i^{n+1} = u_i^n - \left( \nu - \frac{\nu}{2}(1 - \nu)\phi_{i-1} \right) (u_i^n - u_{i-1}^n) + \left( -\frac{\nu}{2}(1 - \nu)\phi_i \right) (u_{i+1}^n - u_i^n). \quad (\text{C.23})$$

Regarding Eq. (4.40), one possible choice for  $C_{i+1/2}^+$  and  $C_{i-1/2}^-$  is:

$$\begin{aligned} C_{i-1/2}^- &= \left( \nu - \frac{\nu}{2}(1 - \nu)\phi_{i-1} \right), \\ C_{i+1/2}^+ &= -\frac{\nu}{2}(1 - \nu)\phi_i. \end{aligned} \quad (\text{C.24})$$

But for  $\phi \approx 1$ ,  $C_{i+1/2}^+ < 0$ , and this is not correct based on Theorem 1. Another possible choice is:

$$\begin{aligned} C_{i-1/2}^- &= \nu + \frac{\nu}{2}(1 - \nu) \left( \frac{\phi_i \times (u_{i+1}^n - u_i^n) - \phi_{i-1} \times (u_i^n - u_{i-1}^n)}{u_i^n - u_{i-1}^n} \right), \\ C_{i+1/2}^+ &= 0. \end{aligned} \quad (\text{C.25})$$

Since  $0 \leq \nu \leq 1$ , then  $0 \leq C_{i-1/2}^- \leq 1$ .

If we define the ratio of successive gradients as  $r_i = \Delta_- u_i / \Delta_+ u_i$ , where  $\Delta_- u_i := u_i - u_{i-1}$  and  $\Delta_+ u_i := u_{i+1} - u_i$ , and notice that  $\phi_i = \phi(r_i)$ , then  $C_{i-1/2}^-$  can be simplified as:

$$C_{i-1/2}^- = \nu + \frac{\nu}{2}(1 - \nu) \left( \frac{\phi(r_i)}{r_i} - \phi(r_{i-1}) \right). \quad (\text{C.26})$$

In the above equation,  $r_i$  acts as a *smoothness monitor* or *smoothness indicator*. Due to the positivity condition ( $0 \leq C_{i-1/2}^- \leq 1$ ), it is straitforward to show that:

$$-\frac{2}{1 - \nu} \leq \left( \frac{\phi(r_i)}{r_i} - \phi(r_{i-1}) \right) \leq \frac{2}{\nu} \quad (\text{C.27})$$

Since  $\nu \leq 1$  (due to the numerical stability), from Eq. (C.27), we must have:

$$\left| \frac{\phi(r_i)}{r_i} - \phi(r_{i-1}) \right| \leq 2. \quad (\text{C.28})$$

For  $r_i \leq 0$ , neighbor gradients have opposite signs, and this means near  $x_i$  an extremum point exists. To guarantee the TVD condition around such points, it is safer

to set  $\phi(r_i) = 0$  for  $r_i \leq 0$ . By considering this new setting, Eq. (C.28) leads to two conditions for TVD-limiters, as:

$$\begin{aligned} 0 \leq \frac{\phi(r_i)}{r_i} &\leq 2, \\ \phi(r_i) &\leq 2. \end{aligned} \tag{C.29}$$

### C.1.4 Relationship between slope and flux limiters

The linear convection equation,  $u_t + au_x = 0$  where  $a > 0$ , is assumed. For this equation the REP concept is used for deriving the relationship between slope and flux limiter concepts. For this equation, the REP procedure can be summarized as:

1. The *reconstruction* step: for each cell  $I_j$ , the reconstructed linear solution is:  $u(x, t^n) = u_j^n + (u_x)_j^n(x - x_j)$ , where  $x_{j-1/2} \leq x \leq x_{j+1/2}$ ,  $(u_x)_j^n$  denotes a limited slope, and  $u_j^n$  is the average solution,
2. The *evolution* step: here  $u(x, t^{n+1})$  can be obtained exactly, since  $u(x, t^{n+1}) = u(x - a\Delta t, t^n)$ ; then the linear function  $u(x, t^n)$  on  $x \in [x_{j-1/2}, x_{j+1/2}]$  becomes a linear function  $u(x, t^{n+1})$  on  $x \in [x_{j-1/2} + a\Delta t, x_{j+1/2} + a\Delta t]$ ,
3. The *projection* step:  $u_j^{n+1}$  can be obtained as:

$$\begin{aligned} u_j^{n+1} &= \frac{1}{\Delta x} \left[ \int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t^{n+1}) dx \right] \\ &= \frac{1}{\Delta x} \left[ \int_{x_{j-1/2}}^{x_{j-1/2} + a\Delta t} (u_{j-1}^n + (u_x)_{j-1}^n(x - x_{j-1} - a\Delta t)) dx \right. \\ &\quad \left. + \int_{x_{j-1/2} + a\Delta t}^{x_{j+1/2}} (u_j^n + (u_x)_j^n(x - x_j - a\Delta t)) dx \right] \\ &= u_j^n - \nu(u_j^n - u_{j-1}^n) - \frac{1}{2}\nu(1 - \nu)(\Delta x(u_x)_j^n - \Delta x(u_x)_{j-1}^n), \end{aligned} \tag{C.30}$$

where  $\nu = a\Delta t/\Delta x$ . So corresponding numerical flux for Eq. (C.30) is:

$$F_j^n = au_j^n + \frac{1}{2}a(1 - \nu)\Delta x(u_x)_j^n. \tag{C.31}$$

If  $(u_x)_j^n = (\frac{u_{j+1}^n - u_j^n}{\Delta x})\phi_j$ , the numerical flux from Eq. (C.31) is the same as Eq. (C.22). In this regard, flux limiters can act as slope limiters, as well.

## C.2 Entropy functions and entropy conditions

Let  $\eta(u)$  denotes an entropy function satisfying a conservation law for smooth solutions, as [68]:

$$\eta(u)_t + \psi(u)_x = 0, \quad (\text{C.32})$$

for some unknown flux  $\psi(u)$ . This equation can be rewritten as:

$$\eta'(u)u_t + \psi'(u)u_x = 0, \quad (\text{C.33})$$

where  $\eta'(u) = \frac{\partial \eta}{\partial u}$  and  $\psi'(u) = \frac{\partial \psi}{\partial u}$ .

A general conservation law  $u_t + f(u)_x = 0$  is considered; this equation can be rewritten as:  $u_t + f'(u)u_x = 0$ . By multiplying this equation with  $\eta'(u)$ , the following updated conservation law will be obtained:

$$\eta'(u)u_t + \eta'(u)f'(u)u_x = 0, \quad (\text{C.34})$$

and from this,  $\psi'(u)$  has the following relationship with  $f$  and  $\eta$ :

$$\psi'(u) = \eta'(u)f'(u). \quad (\text{C.35})$$

An additional condition is assumed for the general entropy function  $\eta(u)$ ; it is assumed that the flux function is *convex*:  $\eta''(u) > 0$  where  $\eta''(u) := \partial^2 \eta / \partial u^2$ . The reason of such assumption will be explained soon.

The conservation property of the entropy function is only valid on smooth solutions. The aim is to study the entropy behavior around discontinuities. For this reason, the following viscous problem is considered:

$$u_t + f(u)_x = \epsilon u_{xx}. \quad (\text{C.36})$$

For this equation, corresponding vanishing viscosity case is studied to capture behavior of corresponding conservation law, as  $\epsilon \rightarrow 0$ . It should be mentioned that, the viscous problem admits only the smooth solutions and its weak solution is a physical and unique solution.

Both sides of Eq. (C.36) are multiplied by  $\eta'(u)$  as:

$$\begin{aligned}\eta'(u)u_t + \eta'(u)f(u)_x &= \epsilon\eta'(u)u_{xx}, \quad \text{or} \\ \eta(u)_t + \psi(u)_x &= \epsilon\eta'(u)u_{xx}.\end{aligned}\tag{C.37}$$

The RHS of the above equation can be expanded as:

$$\eta(u)_t + \psi(u)_x = \epsilon(\eta'(u)u_x)_x - \epsilon\eta''(u)u_x^2.\tag{C.38}$$

Integrating this equation over the spatio-temporal rectangle  $[x_1, x_2] \times [t_1, t_2]$ , we have:

$$\begin{aligned}\int_{t_1}^{t_2} \int_{x_1}^{x_2} \eta(u)_t + \psi(u)_x dx dt &= \epsilon \int_{t_1}^{t_2} [\eta'(u(x_2, t))u_x(x_2, t) - \eta'(u(x_1, t))u_x(x_1, t)] dt \\ &\quad - \epsilon \int_{t_1}^{t_2} \int_{x_1}^{x_2} \eta''(u)u_x^2 dx dt.\end{aligned}\tag{C.39}$$

If the solution  $u$  is smooth at  $x_1$  and  $x_2$ , then the first term in RHS approaches zero as  $\epsilon \rightarrow 0$  (since it is bounded). However this discussion is not true for the second term in RHS, especially if  $u$  has a discontinuity between  $x_1$  and  $x_2$ ; this is clear since this term includes  $u_x^2$ . However, as  $u_x^2 > 0$ ,  $\epsilon > 0$  and  $\eta'' > 0$ , at the limit  $\epsilon \rightarrow 0$ , we have:

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} [\eta(u)_t + \psi(u)_x] dx dt \leq 0,\tag{C.40}$$

for all possible  $x_1$  &  $x_2$ , and  $t_1$  &  $t_2$ . And this shows that the solution  $u(x, t)$  is the *entropy solution* for all convex entropy functions and corresponding entropy fluxes, if:

$$\eta(u)_t + \psi(u)_x \leq 0.\tag{C.41}$$

### C.3 E-schemes and slope limiters

Considering a scalar conservation law  $u_t + f(u)_x = 0$  with the semi-discrete form  $\frac{d}{dt}u + \frac{F_{i+1/2}^* - F_{i-1/2}^*}{\Delta x_i} = 0$ . A simple class of numerical flux functions,  $F^*$  can be introduced which resulted solutions converge to unique entropy ones as  $\Delta x_i \rightarrow 0$ . Such schemes is known as the *E - schemes* [168].

**Definition 1.** A consistent scheme <sup>17</sup> whose numerical flux satisfies

$$\text{Sign}(u_i - u_{i-1}) \left( F_{i-1/2}^* - f(u) \right) \leq 0, \quad (\text{C.42})$$

is called a E-scheme, for all  $u$  between  $u_{i-1}$  and  $u_i$ .

It is easy to show that the three point stencil monotone schemes (where  $F_{i-1/2}^* = F^*(u_i, u_{i-1})$ ) are E-schemes.

### Deriving an extra condition for limiters

To guarantee Eq. (C.42), it is sufficient that:

$$\text{Sign}(u_i - u_{i-1}) = \text{Sign} \left( f(u) - F_{i-1/2}^* \right). \quad (\text{C.43})$$

Let us consider the linear advection equation  $u_t + a.u_x = 0$ , where the flux is:  $f(u) = u$ . We assume  $u(x_i, t^n) = u_i$  is an extremum point; at this point, the TVD condition leads to zero slope:  $(u_x)_i = 0$ . Considering this condition, for two set  $\{u_{i-1}, u_i\}$  and  $\{u_{i+1}, u_i\}$ , the E-condition, Eq. (C.43) leads to following two equations:

$$\text{Sign}(u_{i+1} - u_i) = \text{Sign} \left[ (f(u)) - (F_{i+1/2}^*) \right], \quad (\text{C.44a})$$

$$\text{Sign}(u_i - u_{i-1}) = \text{Sign} \left[ (f(u)) - (F_{i-1/2}^*) \right]. \quad (\text{C.44b})$$

Let us use the numerical flux from the KT formulation:

$F_{i+1/2}^* := \frac{1}{2} \left[ (f_{i+1/2}^r + f_{i+1/2}^l) - a_{i+1/2} (u_{i+1/2}^r - u_{i+1/2}^l) \right]$ , where  $a_{j+1/2} = a$ ,  $u_{i+1/2}^r := u_{i+1} - (u_x)_{i+1} \Delta x / 2$ ,  $u_{i+1/2}^l := u_i + (u_x)_i \Delta x / 2$ ,  $f_{i+1/2}^r = a u_{i+1/2}^r$  and  $f_{i+1/2}^l = a u_i$  (since  $(u_x)_i = 0$ ).

As mentioned, in our example, it is assumed that  $u_i$  is an extremum point (i.e.  $(u_x)_i = 0$ ); for the KT scheme <sup>18</sup> this means:  $F_{i+1/2}^* = u_i$  and  $F_{i-1/2}^* = (u_{i-1} + \frac{\Delta x}{2} (u_x)_{i-1})$ . In Eqs. (C.44a) and (C.44b), we set  $u = u_{i+1/2}^r$  and  $u = u_i$ , respectively. Hence, Eqs.

<sup>17</sup>A numerical method is consistent with the original conservation law ( $u_t + f_x = 0$ ), if the numerical flux  $F$  reduces to the true flux  $f$  for the case of constant state variable (flow); this means for  $u(x, t) \equiv \bar{u}$ , we have:  $F(\bar{u}, \bar{u}) = f(\bar{u})$ .

<sup>18</sup>For the KT scheme the numerical flux at  $x_{i+1/2}$  reads:  $F_{i+1/2}^* = \frac{1}{2} \left( f_{i+1/2}^r + f_{i+1/2}^l \right) - \frac{a_{i+1/2}}{2} \left( u_{i+1/2}^r - u_{i+1/2}^l \right)$ .

(C.44) become:

$$\text{Sign}(u_{i+1} - u_i) = \text{Sign} \left[ \left( u_{i+1} - \frac{\Delta x}{2} (u_x)_{i+1} \right) - (u_i) \right], \quad (\text{C.45a})$$

$$\text{Sign}(u_i - u_{i-1}) = \text{Sign} \left[ (u_i) - \left( u_{i-1} + \frac{\Delta x}{2} (u_x)_{i-1} \right) \right]. \quad (\text{C.45b})$$

RHS of these equations can be rewritten as:

$$\text{Sign}(u_{i+1} - u_i) = \text{Sign} \left[ (u_{i+1} - u_i) - \left( \frac{\Delta x}{2} (u_x)_{i+1} \right) \right] \quad (\text{C.46a})$$

$$= \text{Sign} \left[ \Delta u_{i+1/2} - \frac{\Delta x}{2} (u_x)_{i+1} \right] = \text{Sign} \left[ \Delta u_{i+1/2} \left( 1 - \frac{1}{2} \frac{\Delta x (u_x)_{i+1}}{\Delta u_{i+1/2}} \right) \right],$$

$$\text{Sign}(u_i - u_{i-1}) = \text{Sign} \left[ (u_i - u_{i-1}) - \left( \frac{\Delta x}{2} (u_x)_{i-1} \right) \right] \quad (\text{C.46b})$$

$$= \text{Sign} \left[ \Delta u_{i-1/2} - \frac{\Delta x}{2} (u_x)_{i-1} \right] = \text{Sign} \left[ \Delta u_{i-1/2} \left( 1 - \frac{1}{2} \frac{\Delta x (u_x)_{i-1}}{\Delta u_{i-1/2}} \right) \right].$$

So, it is necessary that:

$$\frac{1}{2} \left| \frac{\Delta x}{\Delta u_{i+1/2}} (u_x)_{i+1} \right| \leq 1, \quad (\text{C.47a})$$

$$\frac{1}{2} \left| \frac{\Delta x}{\Delta u_{i-1/2}} (u_x)_{i-1} \right| \leq 1. \quad (\text{C.47b})$$

# Appendix D

## Numerical methods for stress waves

### D.1 The Runge-Kutta 4<sup>th</sup> order for stress wave problems: second order systems

Let us assume that values of displacement ( $u(t)$ ), velocity ( $v(t) := du/dt$ ) and acceleration ( $a(t) := d^2u/dt^2$ ) of a particle in motion (with spatial position  $x_j$ ) are known at time  $t_n := n\Delta t$ ; the response at the next time step,  $t_{n+1} = (n+1)\Delta t$  can be estimated by the Runge-Kutta 4<sup>th</sup> order time integration as [232]:

$$\left\{ u(t + \Delta t) = u(t) + \frac{\Delta t}{6} \left[ v(t) + 2\tilde{v}_2\left(t + \frac{\Delta t}{2}\right) + 2\tilde{v}_3\left(t + \frac{\Delta t}{2}\right) + \tilde{v}_4(t + \Delta t) \right] \right\},$$
$$\left\{ v(t + \Delta t) = v(t) + \frac{\Delta t}{6} \left[ a(t) + 2\tilde{a}_2\left(t + \frac{\Delta t}{2}\right) + 2\tilde{a}_3\left(t + \frac{\Delta t}{2}\right) + \tilde{a}_4(t + \Delta t) \right] \right\},$$

where  $\tilde{v}_i$  and  $\tilde{a}_i$  denote guess (intermediate) velocity and acceleration at  $i^{\text{th}}$  stage, respectively; intermediate displacements,  $\tilde{u}_i$  can also be defined. These guess values can be obtained as:

$$\left\{ \tilde{u}_2\left(t + \frac{\Delta t}{2}\right) = u(t) + \frac{\Delta t}{2}v(t) \right\}, \left\{ \tilde{v}_2\left(t + \frac{\Delta t}{2}\right) = v(t) + \frac{\Delta t}{2}a(t) \right\},$$
$$\left\{ \tilde{u}_3\left(t + \frac{\Delta t}{2}\right) = u(t) + \frac{\Delta t}{2}\tilde{v}_2\left(t + \frac{\Delta t}{2}\right) \right\}, \left\{ \tilde{v}_3\left(t + \frac{\Delta t}{2}\right) = v(t) + \frac{\Delta t}{2}\tilde{a}_2\left(t + \frac{\Delta t}{2}\right) \right\},$$
$$\left\{ \tilde{u}_4(t + \Delta t) = u(t) + \Delta t\tilde{v}_3\left(t + \frac{\Delta t}{2}\right) \right\}, \left\{ \tilde{v}_4(t + \Delta t) = v(t) + \Delta t\tilde{a}_3\left(t + \frac{\Delta t}{2}\right) \right\}.$$

Regarding the scalar wave equation  $c^2\partial^2u/\partial x^2 = \partial^2u/\partial t^2$ , the equation is rewritten in a semi-discrete form: discrete in spatial domain and continuous in time. The resulted



system of ODEs can then be solved by the method of lines scheme; for this, the time integration can be done by the Runge-Kutta 4<sup>th</sup> order method. Let us assume the spatial discretization is done by a finite difference method at spatial location  $x_j$ ; the Runge-Kutta 4<sup>th</sup> order method can be implemented by considering:  $u_j(t) := u(x_j, t)$ ,  $v_j(t) := du_j/dt$  and  $a_j(t) := \partial^2 u_j / \partial t^2 = c^2 \partial^2 u_j / \partial x^2$ .

## D.2 The generalized $\alpha$ -dissipative time integration method

To use this scheme, a semi-discrete form of wave equations is considered: discrete in the spatial domain and continuous in time; the spatial discretization is done with the finite element method. The semi-discrete form can then be represent as:  $\mathbf{M}\ddot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{F}$ .

Having solution values at time step  $n$ , corresponding values at the time step  $n + 1$  can be obtained by the generalized  $\alpha$ -time integration algorithm as [130]:

$$\begin{aligned}\mathbf{F}_{n+1} &= \mathbf{M}\mathbf{a}_{n+1} - \alpha\mathbf{K}\mathbf{u}_n + (\alpha + 1)\mathbf{K}\mathbf{u}_{n+1}, \\ \mathbf{u}_{n+1} &= \mathbf{u}_n + \Delta t\mathbf{v}_n + \Delta t^2 [(0.5 - \beta)\mathbf{a}_n + \beta\mathbf{a}_{n+1}], \\ \mathbf{v}_{n+1} &= \mathbf{v}_n + \Delta t [(1 - \gamma)\mathbf{a}_n + \gamma\mathbf{a}_{n+1}],\end{aligned}\tag{D.1}$$

where  $\alpha$ ,  $\beta$  and  $\gamma$  are free parameters controlling the stability and numerical dissipation of the algorithm. At the initial step, we have  $\mathbf{u}_0 = \mathbf{u}$ ,  $\mathbf{v}_0 = \mathbf{v}$ , and  $\mathbf{a}_0 = \mathbf{M}^{-1}(\mathbf{F}_0 - \mathbf{K}\mathbf{d}_0)$ . For case  $\gamma > 0.5$  numerical dissipation exists and for  $\beta \geq 0.25(\gamma + 0.5)^2$  the mentioned algorithm is unconditionally stable [130].

## D.3 The time discontinuous Galerkin method

For second order dynamical systems, the time discontinuous Galerkin method formulation can be obtained by considering the following assumptions: 1) employing the concept of the finite element method in the time domain; 2) possible existence of a discontinuity (jump) at each time step; 3) rewriting the second-order equations as first-order ones.

To account possible discontinuities, the following notations are introduced:

$Z_n^\pm := \lim_{\epsilon \pm \rightarrow 0} Z(t_n \pm \epsilon)$  and  $t_n^\pm := \lim_{\epsilon \pm \rightarrow 0} (t_n \pm \epsilon)$ ; where  $Z \in \{\mathbf{u}, \mathbf{v}, \mathbf{w}_i\}$ , in which:  $\mathbf{w}_i := \mathbf{w}_i(t)$  denotes a weight function and  $\mathbf{v} := \dot{\mathbf{u}}$ . By considering the new variable  $\mathbf{v}$ , the second order (semi-discrete) equation  $\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{F}$  can be written as:  $\mathbf{M}\dot{\mathbf{v}} + \mathbf{C}\mathbf{v} + \mathbf{K}\mathbf{u} = \mathbf{F}$

&  $\mathbf{K}(\dot{\mathbf{u}} - \mathbf{v}) = \mathbf{0}$ . The weighted residual form of these first order equations on time interval  $t \in [t_n^-, t_{n+1}^-]$  is [233]:

$$R_n = \int_{t_n^-}^{t_{n+1}^-} \mathbf{w}_1^T (\mathbf{M}\dot{\mathbf{v}} + \mathbf{C}\mathbf{v} + \mathbf{K}\mathbf{u} - \mathbf{F}) dt + \int_{t_n^-}^{t_{n+1}^-} \mathbf{w}_2^T (\mathbf{K}(\dot{\mathbf{u}} - \mathbf{v})) dt = 0, \text{ for } n \in \{1, \dots, N\}, \quad (\text{D.2})$$

where:  $t_1 = 0$ ,  $t_{N+1} = T$  and  $t_n \leq t_{n+1}$ . It should be mentioned that solutions  $\mathbf{u}$  &  $\mathbf{v}$ , and weight functions  $\mathbf{w}_i$  can have discontinuities in time interval  $t \in [t_n^-, t_n^+]$ .

By considering linear shape functions and linear weight functions (in time domain), the following matrix form equation can be obtained from Eq. (D.2) for  $t \in [t_n^-, t_{n+1}^-]$  [217]:

$$\begin{pmatrix} \frac{\mathbf{K}}{2} & \frac{\mathbf{K}}{2} & -\frac{\Delta t_n \mathbf{K}}{3} & -\frac{\Delta t_n \mathbf{K}}{6} \\ -\frac{\mathbf{K}}{2} & \frac{\mathbf{K}}{2} & -\frac{\Delta t_n \mathbf{K}}{6} & -\frac{\Delta t_n \mathbf{K}}{3} \\ \frac{\Delta t_n \mathbf{K}}{3} & \frac{\Delta t_n \mathbf{K}}{6} & \frac{\Delta t_n \mathbf{C}}{3} + \frac{\mathbf{M}}{2} & \frac{\Delta t_n \mathbf{C}}{6} + \frac{\mathbf{M}}{2} \\ \frac{\Delta t_n \mathbf{K}}{6} & \frac{\Delta t_n \mathbf{K}}{3} & \frac{\Delta t_n \mathbf{C}}{6} - \frac{\mathbf{M}}{2} & \frac{\Delta t_n \mathbf{C}}{3} + \frac{\mathbf{M}}{2} \end{pmatrix} \cdot \begin{bmatrix} \mathbf{u}_n^+ \\ \mathbf{u}_{n+1}^- \\ \mathbf{v}_n^+ \\ \mathbf{v}_{n+1}^- \end{bmatrix} = \begin{bmatrix} \mathbf{K}\mathbf{u}_n^- \\ \mathbf{0} \\ \mathbf{F}_1 + \mathbf{M}\mathbf{v}_n^- \\ \mathbf{F}_2 \end{bmatrix}, \quad (\text{D.3})$$

where:  $\mathbf{F}_1 = \int_{t_n}^{t_{n+1}} \frac{t_{n+1}-t}{\Delta t_n} \mathbf{F} dt$  and  $\mathbf{F}_2 = \int_{t_n}^{t_{n+1}} \frac{t-t_n}{\Delta t_n} \mathbf{F} dt$ .

## D.4 Taylor-Galerkin discretizations

The second order equation of motion is again rewritten as a first order system, like:  $\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{E}}{\partial x} = \mathbf{H}$ ; in this equation we have:  $\mathbf{U} = [\mathbf{u}, \rho \mathbf{v}]^T$ ,  $\mathbf{E} = [\mathbf{0}, -\boldsymbol{\sigma}]^T$  and  $\mathbf{H} = [\mathbf{v}, -\rho \mathbf{F}]^T$ ; where:  $\boldsymbol{\sigma}$  and  $\mathbf{F}$  are the stress and load vectors, respectively; and  $\mathbf{v} := \dot{\mathbf{u}}$ . The momentum  $\mathbf{V} = \rho \mathbf{v}$  is expanded by the Taylor series in time  $t = t^n$ , as:  $\mathbf{V}^{n+1} = \mathbf{V}^n + \Delta t \dot{\mathbf{V}}^n + (\Delta t^2/2) \ddot{\mathbf{V}}^n + \mathcal{O}(\Delta t^3)$  (where  $\mathbf{V}^n := \mathbf{V}(t^n)$ ). Inserting this expansion in the first-order system, and then using the Galerkin discretization in space, the conventional (second-order) Taylor-Galerkin method can be obtained as [216]:

$$\mathbf{M}\mathbf{v}^{n+1} = \mathbf{M}\mathbf{v}^n + \Delta t \{\mathbf{H}^n - \mathbf{K}\mathbf{u}^n\} + \frac{\Delta t^2}{2} \{\dot{\mathbf{H}}^n - \mathbf{K}\dot{\mathbf{v}}^n\}, \quad (\text{D.4})$$

where  $\mathbf{M}$  and  $\mathbf{K}$  are the mass and stiffness matrices. Displacement  $\mathbf{u}^{n+1}$  can be obtained as:

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t \left\{ (1 - \gamma) \mathbf{v}^n + \gamma \mathbf{v}^{n+1} \right\}, \quad (\text{D.5})$$

where for  $\gamma = 0.5$ , the method is stable with maximum time step.

If a higher order Taylor expansion is used for momentum  $\mathbf{V}$ , a higher order Taylor Galerkin scheme can be obtained. For the third order expansion,  $\mathbf{v}^{n+1}$  can be obtained as [216]:

$$\mathbf{v}^{n+1} = \mathbf{v}^n + \left( \mathbf{M} + \alpha \frac{\Delta t^2}{6} \mathbf{K} \right)^{-1} \left[ \Delta t \mathbf{H}^n + \frac{\Delta t^2}{2} \dot{\mathbf{H}}^n + \alpha \frac{\Delta t^2}{6} (\dot{\mathbf{H}}^{n-1} - \dot{\mathbf{H}}^n) - \mathbf{K} \left( \frac{\Delta t^2}{2} \mathbf{v}^n + \Delta t \mathbf{u}^n \right) \right], \quad (\text{D.6})$$

where it is needed  $\alpha \geq 1/2$ ; for accurate solutions  $\alpha$  near to  $1/2$  should be used. For unconditional stability, it is necessary  $\alpha = 3/2$  and  $\gamma = 1/2$ . Displacement  $\mathbf{u}^{n+1}$  can be obtained from Eq. (D.5).

# Bibliography

- [1] A. Schadschneider, D. Chowdhury, and K. Nishinari. *Stochastic Transport in Complex Systems: from Molecules to Vehicles*. Elsevier, 2010.
- [2] M. A. J. Chaplain. Avascular growth, angiogenesis and vascular growth in solid tumours: the mathematical modelling of the stages of tumour development. *Mathematical and Computer Modelling*, 23(6):47–87, 1996.
- [3] X. Li and E. Weinan. Multiscale modeling of the dynamics of solids at finite temperature. *Journal of the Mechanics and Physics of Solids*, 53(7):1650–1685, 2005.
- [4] C. Cattani, IA. IA. Rushchitskiĭ, and J. Rushchitsky. *Wavelet and Wave Analysis as Applied to Materials with Micro or Nanostructure*. World Scientific Publishing Company Incorporated, 2007.
- [5] J. A. Trangenstein. *Numerical Solution of Hyperbolic Partial Differential Equations*. Cambridge University Press, 2009.
- [6] F. Montomoli, M. Carnevale, A. D’Ammaro, M. Massini, and S. Salvadori. *Uncertainty quantification in computational fluid dynamics and aircraft engines*. Springer, 2015.
- [7] E. B. Tadmor, M. Ortiz, and R. Phillips. Quasicontinuum analysis of defects in solids. *Philosophical magazine A*, 73(6):1529–1563, 1996.
- [8] F. F. Abraham, J. Q. Broughton, N. Bernstein, and E. Kaxiras. Spanning the continuum to quantum length scales in a dynamic simulation of brittle fracture. *EPL (Europhysics Letters)*, 44(6):783, 1998.
- [9] Y. Wang and D. L. McDowell. *Uncertainty Quantification in Multiscale Materials Modeling*. Woodhead Publishing Limited, 2020.
- [10] J. C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, 2008.
- [11] U. M. Ascher. *Numerical Methods for Evolutionary Differential Equations*, volume 5. SIAM, 2008.
- [12] T. Cebeci. *Computational Fluid Dynamics for Engineers: from Panel to Navier-Stokes Methods with Computer Programs*. Springer, 2005.

## Bibliography

- [13] D. Gottlieb and J. S. Hesthaven. Spectral methods for hyperbolic problems. *Journal of Computational and Applied Mathematics*, 128(1):83–131, 2001.
- [14] S. Bertoluzza and L. M. Castro. Adaptive wavelet collocation for elasticity: first results. *Pavia, Italy: IAN-CNR*, 2001.
- [15] O. V. Vasilyev and N. K.-R. Kevlahan. An adaptive multilevel wavelet collocation method for elliptic problems. *Journal of Computational Physics*, 206(2):412 – 431, 2005.
- [16] J. M. Alam, N. K.-R. Kevlahan, and O. V. Vasilyev. Simultaneous space time adaptive wavelet solution of nonlinear parabolic differential equations. *Journal of Computational Physics*, 214(2):829 – 857, 2006.
- [17] M. Mehra and N. K.-R. Kevlahan. An adaptive wavelet collocation method for the solution of partial differential equations on the sphere. *Journal of Computational Physics*, 227(11):5610 – 5632, 2008.
- [18] L. Jameson and T. Miyama. Wavelet analysis and ocean modeling: A dynamically adaptive numerical method 'WOFD-AHO'. *Monthly Weather Review*, 128(5):1536–1549, 2000.
- [19] H. Yousefi and T. Rabczuk. Multiresolution-based adaptive central high resolution schemes for modeling of nonlinear propagating fronts. *Engineering Analysis with Boundary Elements*, 103:172–195, 2019.
- [20] M. S. Darwish and F. H. Moukalled. Normalized variable and space formulation methodology for high-resolution schemes. *Numerical Heat Transfer, Part B*, 26(1):79–96, 1994.
- [21] J. C. Santos, P. Cruz, M. A. Alves, P. J. Oliveira, F. D. Magalhães, and A. Mendes. Adaptive multiresolution approach for two-dimensional PDEs. *Computer Methods in Applied Mechanics and Engineering*, 193(3):405–425, 2004.
- [22] A. Kurganov and E. Tadmor. New high-resolution central schemes for nonlinear conservation laws and convection–diffusion equations. *Journal of Computational Physics*, 160(1):241–282, 2000.
- [23] A. Kurganov and G. Petrova. Central schemes and contact discontinuities. *ESAIM: Mathematical Modelling and Numerical Analysis*, 34(06):1259–1275, 2000.
- [24] A. Kurganov and Ch.-T. Lin. On the reduction of numerical dissipation in central-upwind schemes. *Communications in Computational Physics*, 2(1):141–163, 2007.
- [25] H. Nessyahu and E. Tadmor. Non-oscillatory central differencing for hyperbolic conservation laws. *Journal of Computational Physics*, 87(2):408–463, 1990.

## Bibliography

- [26] S. Karni, A. Kurganov, and G. Petrova. A smoothness indicator for adaptive algorithms for hyperbolic systems. *Journal of Computational Physics*, 178(2):323–341, 2002.
- [27] S. Karni and A. Kurganov. Local error analysis for approximate solutions of hyperbolic conservation laws. *Advances in Computational Mathematics*, 22(1):79–99, 2005.
- [28] G. Puppo. Numerical entropy production on shocks and smooth transitions. *Journal of Scientific Computing*, 17(1-4):263–271, 2002.
- [29] G. Puppo. Numerical entropy production for central schemes. *SIAM Journal on Scientific Computing*, 25(4):1382–1415, 2004.
- [30] A. Kurganov, G. Petrova, and B. Popov. Adaptive semidiscrete central-upwind schemes for nonconvex hyperbolic conservation laws. *SIAM Journal on Scientific Computing*, 29(6):2381–2401, 2007.
- [31] G. Puppo and M. Semplice. Numerical entropy and adaptivity for finite volume schemes. *Communications in Computational Physics*, 10(5):1132, 2011.
- [32] M. Ersoy, F. Golay, and L. Yushchenko. Adaptive multiscale scheme based on numerical density of entropy production for conservation laws. *Central European Journal of Mathematics*, 11(8):1392–1415, 2013.
- [33] A. Cohen. *Numerical Analysis of Wavelet Methods*, volume 32. Elsevier, 2003.
- [34] W. Sweldens and P. Schröder. Building your own wavelets at home. In *Wavelets in the Geosciences*, pages 72–107. Springer, 2000.
- [35] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.
- [36] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. *Communications on pure and applied mathematics*, 44(2):141–183, 1991.
- [37] K. Urban. *Wavelet Methods for Elliptic Partial Differential Equations*. Oxford university press, 2009.
- [38] S. Müller. *Adaptive Multiscale Schemes for Conservation Laws*, volume 27. Springer, 2003.
- [39] B. G. Müller and S. Müller. Application of multiscale techniques to hyperbolic conservation laws. In *Computational Mathematics, Lecture Notes in Pure & Applied Mathematics*, pages 113–138, 1998.
- [40] S. Müller and Y. Stiriba. Fully adaptive multiscale schemes for conservation laws employing locally varying time stepping. *J. Sci. Comput.*, 30(3):493–531, 2007.

## Bibliography

- [41] M. Holmström and J. Waldén. Adaptive wavelet methods for hyperbolic PDEs. *J. Sci. Comput.*, 13(1):19–49, 1998.
- [42] W. Dahmen, B. Gottschlich-Müller, and S. Müller. Multiresolution schemes for conservation laws. *Numerische Mathematik*, 88(3):399–443, 2001.
- [43] M. Griebel and F. Koster. Adaptive wavelet solvers for the unsteady incompressible Navier-Stokes equations. In J. Málek, J. Nečas, and M. Rokyta, editors, *Advances in Mathematical Fluid Mechanics*, pages 67–118. Springer Berlin Heidelberg, 2000.
- [44] Wei Cai and Jianzhong Wang. Adaptive multiresolution collocation methods for initial-boundary value problems of nonlinear pdes. *SIAM Journal on Numerical Analysis*, 33(3):937–970, 1996.
- [45] P. Cruz, A. Mendes, and F. D. Magalhães. Using wavelets for solving PDEs: an adaptive collocation method. *Chemical Engineering Science*, 56(10):3305 – 3309, 2001.
- [46] P. Cruz, A. Mendes, and F. D. Magalhães. Wavelet-based adaptive grid method for the resolution of nonlinear PDEs. *AIChE Journal*, 48(4):774–785, 2002.
- [47] M. A. Alves, P. Cruz, A. Mendes, F. D. Magalhães, F. T. Pinho, and P. J. Oliveira. Adaptive multiresolution approach for solution of hyperbolic PDEs. *Computer Methods in Applied Mechanics and Engineering*, 191(36):3909 – 3928, 2002.
- [48] H. Yousefi, A. Noorzad, and J. Farjoodi. Simulating 2D waves propagation in elastic solid media using wavelet based adaptive method. *Journal of Scientific Computing*, 42(3):404–425, 2010.
- [49] H. Yousefi, A. Noorzad, and J. Farjoodi. Multiresolution based adaptive schemes for second order hyperbolic PDEs in elastodynamic problems. *Applied Mathematical Modelling*, 37(12):7095–7127, 2013.
- [50] P. Cruz, M. A. Alves, F. D. Magalhães, and A. Mendes. Solution of hyperbolic PDEs using a stable adaptive multiresolution method. *Chemical Engineering Science*, 58(9):1777 – 1792, 2003.
- [51] A. Harten. Multiresolution algorithms for the numerical solution of hyperbolic conservation laws. *Communications on Pure and Applied Mathematics*, 48(12):1305–1342, 1995.
- [52] J. Glimm, M. J. Graham, J. V. Grove, and B. J. Plohr. Hyperbolic problems: theory, numerics, applications. In *Proceedings of the 5th International Conference held at the University of New York*, 1994.
- [53] A. Cohen, S. Kaber, S. Müller, and M. Postel. Fully adaptive multiresolution finite volume schemes for conservation laws. *Mathematics of Computation - American Mathematical Society*, 72(241):183–225, 2003.

## Bibliography

- [54] M. Holmström. Solving hyperbolic PDEs using interpolating wavelets. *SIAM Journal on Scientific Computing*, 21(2):405–420, 1999.
- [55] H. Yousefi, A. Noorzad, J. Farjoodi, and M. Vahidi. Multiresolution-based adaptive simulation of wave equation. *Applied Mathematics & Information Sciences*, 6(1S):47S–58S, 2012.
- [56] S. Paolucci, Z. J. Zikoski, and D. Wirasaet. WAMR: An adaptive wavelet method for the simulation of compressible reacting flow. part I. accuracy and efficiency of algorithm. *J. Comput. Phys.*, 272:814–841, 2014.
- [57] S. Paolucci, Z. J. Zikoski, and T. Grenga. Wamr: An adaptive wavelet method for the simulation of compressible reacting flow. part II. the parallel algorithm. *J. Comput. Phys.*, 272:842–864, 2014.
- [58] L. Tang and S. Song. A multiresolution finite volume scheme for two-dimensional hyperbolic conservation laws. *J. Comput. Appl. Math.*, 214(2):583–595, 2008.
- [59] R. Bürger and A. Kozakevicius. Adaptive multiresolution WENO schemes for multi-species kinematic flow models. *Journal of Computational Physics*, 224(2):1190 – 1222, 2007.
- [60] M. O. Domingues, S. M. Gomes, O. Roussel, and K. Schneider. Space–time adaptive multiresolution methods for hyperbolic conservation laws: applications to compressible Euler equations. *Appl. Numer. Math.*, 59(9):2303–2321, 2009.
- [61] H. Yousefi, A. Taghavi, and I. Mahmoudzadeh. Multiscale RBF-based central high resolution schemes for simulation of generalized thermoelasticity problems. *Frontiers of Structural and Civil Engineering*, 3(2):429–455, 2019.
- [62] Jingyang Guo and Jae-Hun Jung. A rbf-weno finite volume method for hyperbolic conservation laws with the monotone polynomial interpolation method. *Applied Numerical Mathematics*, 112:27–50, 2017.
- [63] Daniel Caviedes-Voullième, Nils Gerhard, Aleksey Sikstel, and Siegfried Müller. Multiwavelet-based mesh adaptivity with discontinuous galerkin schemes: Exploring 2d shallow water problems. *Advances in Water Resources*, 138:103559, 2020.
- [64] S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.
- [65] M. Calhoun-Lopez and M. D. Gunzburger. A finite element, multiresolution viscosity method for hyperbolic conservation laws. *SIAM Journal on Numerical Analysis*, 43(5):1988–2011, 2005.



## Bibliography

- [66] A. Cohen, N. Dyn, S. M. Kaber, and M. Postel. Multiresolution schemes on triangles for scalar conservation laws. *Journal of Computational Physics*, 161(1):264–286, 2000.
- [67] E. Tadmor and K. Waagan. Adaptive spectral viscosity for hyperbolic conservation laws. *SIAM Journal on Scientific Computing*, 34(2):A993–A1009, 2012.
- [68] R. J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser, 1992.
- [69] J.-L. Guermond, R. Pasquetti, and B. Popov. Entropy viscosity method for nonlinear conservation laws. *Journal of Computational Physics*, 230(11):4248–4267, 2011.
- [70] B. Van Leer. Towards the ultimate conservative difference scheme. V. a second-order sequel to Godunov’s method. *Journal of Computational Physics*, 32(1):101–136, 1979.
- [71] Ch.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77(2):439–471, 1988.
- [72] X.-D. Liu, S. Osher, and T. Chan. Weighted essentially non-oscillatory schemes. *Journal of Computational Physics*, 115(1):200–212, 1994.
- [73] Ch.-W. Shu. Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. In *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations*, pages 325–432. Springer, 1998.
- [74] K. O. Friedrichs and P. D. Lax. Systems of conservation equations with a convex extension. *Proceedings of the National Academy of Sciences*, 68(8):1686–1688, 1971.
- [75] G.-Sh. Jiang, D. Levy, Ch.-T. Lin, S. Osher, and E. Tadmor. High-resolution nonoscillatory central schemes with nonstaggered grids for hyperbolic conservation laws. *SIAM Journal on Numerical Analysis*, 35(6):2147–2168, 1998.
- [76] X.-D. Liu and E. Tadmor. Third order nonoscillatory central scheme for hyperbolic conservation laws. *Numerische Mathematik*, 79(3):397–425, 1998.
- [77] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*, volume 153. Springer Science & Business Media, 2006.
- [78] D. Zhang, L. Cheng, H. An, and M. Zhao. Direct numerical simulation of flow around a surface-mounted finite square cylinder at low Reynolds numbers. *Physics of Fluids*, 29(4):045101, 2017.
- [79] J. G. Bartzis, D. Vlachogiannis, and A. Sfetsos. Thematic area 5: Best practice advice for environmental flows. *The QNET-CFD Network Newsletter*, 2(4):34–39, 2004.

## Bibliography

- [80] A. Scaperdas and S. Gilham. Thematic area 4: Best practice advice for civil construction and HVAC. *The QNET-CFD Network Newsletter*, 2(4):28–33, 2004.
- [81] J. Franke, A. Hellsten, H. Schlünzen, and B. Carissimo. Best practice guideline for the CFD simulation of flows in the urban environment. COST action 732: Quality assurance and improvement of microscale meteorological models. *Hamburg, Germany*, 2007.
- [82] Y. Tominaga, A. Mochida, R. Yoshie, H. Kataoka, T. Nozu, M. Yoshikawa, and T. Shirasawa. AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(10):1749–1761, 2008.
- [83] L. White and A. Adcroft. A high-order finite volume remapping scheme for nonuniform grids: The piecewise quartic method (PQM). *Journal of Computational Physics*, 227(15):7394–7422, 2008.
- [84] J. Velechovský, R. Liska, and M. Shashkov. High-order remapping with piece-wise parabolic reconstruction. *Computers & Fluids*, 83:164 – 169, 2013. Numerical methods for highly compressible multi-material flow problems.
- [85] L. Cueto-Felgueroso and I. Colominas. High-order finite volume methods and multiresolution reproducing kernels. *Archives of Computational Methods in Engineering*, 15(2):185–228, 2008.
- [86] E. Tadmor. Local error estimates for discontinuous solutions of nonlinear hyperbolic equations. *SIAM Journal on Numerical Analysis*, 28(4):891–906, 1991.
- [87] E. Tadmor and T. Tang. Pointwise error estimates for scalar conservation laws with piecewise smooth solutions. *SIAM Journal on Numerical Analysis*, 36(6):1739–1758, 1999.
- [88] G. Bartholomeeusen, H. De Sterck, and G. Sills. Non-convex flux functions and compound shock waves in sediment beds. In *Hyperbolic Problems: Theory, Numerics, Applications*, pages 347–356. Springer, 2003.
- [89] G. Bartholomeeusen. *Compound shock waves and creep behaviour in sediment beds*. PhD thesis, University of Oxford, 2003.
- [90] D. R. Reynolds. *A nonlinear thermodynamic model for phase transitions in shape memory alloy wires*. PhD thesis, Rice University, 2003.
- [91] J. Liu, Ch. M. Landis, H. Gomez, and Th. J. R. Hughes. Liquid–vapor phase transition: Thermomechanical theory, entropy stable numerical formulation, and boiling simulations. *Computer Methods in Applied Mechanics and Engineering*, 297:476–553, 2015.

## Bibliography

- [92] A. Kurganov and G. Petrova. A third-order semi-discrete genuinely multidimensional central scheme for hyperbolic conservation laws and related problems. *Numerische Mathematik*, 88(4):683–729, 2001.
- [93] X.-D. Liu and S. Osher. Nonoscillatory high order accurate self-similar maximum principle satisfying shock capturing schemes I. *SIAM Journal on Numerical Analysis*, 33(2):760–779, 1996.
- [94] X. Zhang and Ch.-W. Shu. A genuinely high order total variation diminishing scheme for one-dimensional scalar conservation laws. *SIAM Journal on Numerical Analysis*, 48(2):772–795, 2010.
- [95] X. Zhang and Ch.-W. Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *Journal of Computational Physics*, 229(9):3091–3120, 2010.
- [96] X. Zhang and Ch.-W. Shu. Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2134):2752–2776, 2011.
- [97] T. Barth and D. Jespersen. The design and application of upwind schemes on unstructured meshes. In *27th Aerospace sciences meeting*, page 366, 1989.
- [98] L. Cueto-Felgueroso and I. Colominas. High-order finite volume methods and multiresolution reproducing kernels. *Archives of Computational Methods in Engineering*, 15(2):185–228, 2008.
- [99] Ph. Colella and P. R. Woodward. The piecewise parabolic method (PPM) for gas-dynamical simulations. *Journal of Computational Physics*, 54(1):174–201, 1984.
- [100] M. V. Popov and S. D. Ustyugov. The piecewise parabolic method (PPM) for gas-dynamical simulations. *Computational Mathematics and Mathematical Physics*, 47(12):1970–1989, 2007.
- [101] Ph. Colella and M. D. Sekora. A limiter for PPM that preserves accuracy at smooth extrema. *Computational Mathematics and Mathematical Physics*, 227(15):7069–7076, 2008.
- [102] W. J. Rider and L. G. Margolin. Simple modifications of monotonicity-preserving limiter. *Journal of Computational Physics*, 174(1):473–488, 2001.
- [103] D. Engwirda and M. Kelley. A WENO-type slope-limiter for a family of piecewise polynomial methods. *arXiv preprint arXiv:1606.08188*, 2016.
- [104] M. R. Norman and R. D. Nair. Inherently conservative nonpolynomial-based remapping schemes: Application to semi-Lagrangian transport. *Monthly Weather Review*, 136(12):5044–5061, 2008.

## Bibliography

- [105] D. Levy, G. Puppo, and G. Russo. Central WENO schemes for hyperbolic systems of conservation laws. *ESAIM: Mathematical Modelling and Numerical Analysis*, 33(3):547–571, 1999.
- [106] D. Levy, G. Puppo, and G. Russo. Compact central WENO schemes for multidimensional conservation laws. *SIAM Journal on Scientific Computing*, 22(2):656–672, 2000.
- [107] L. White, A. Adcroft, and R. Hallberg. High-order regridding–remapping schemes for continuous isopycnal and generalized coordinates in ocean models. *Journal of Computational Physics*, 228(23):8665–8692, 2009.
- [108] I. Cravero and M. Semplice. On the accuracy of WENO and CWENO reconstructions of third order on nonuniform meshes. *Journal of Scientific Computing*, 67(3):1219–1246, 2016.
- [109] M. Dumbser, W. Boscheri, M. Semplice, and G. Russo. Central weighted ENO schemes for hyperbolic conservation laws on fixed and moving unstructured meshes. *SIAM Journal on Scientific Computing*, 39(6):A2564–A2591, 2017.
- [110] M. Semplice, A. Coco, and G. Russo. Adaptive mesh refinement for hyperbolic systems based on third-order compact WENO reconstruction. *Journal of Scientific Computing*, 66(2):692–724, 2016.
- [111] O. Kolb. On the full and global accuracy of a compact third order WENO scheme. *SIAM Journal on Numerical Analysis*, 52(5):2335–2355, 2014.
- [112] P. S. Verma and W.-Ch. Müller. Higher order finite volume central schemes for multi-dimensional hyperbolic problems. *Journal of Scientific Computing*, 75(2):941–969, 2018.
- [113] G. Capdeville. A central WENO scheme for solving hyperbolic conservation laws on non-uniform meshes. *Journal of Computational Physics*, 227(5):2977–3014, 2008.
- [114] R. Artebrant and H. J. Schroll. Conservative logarithmic reconstructions and finite volume methods. *SIAM J. Sci. Comput.*, 27(1):294–314, 2005.
- [115] R. Artebrant and H. J. Schroll. Limiter-free third order logarithmic reconstruction. *SIAM J. Sci. Comput.*, 28(1):359–381, 2006.
- [116] A. Marquina. Local piecewise hyperbolic reconstruction of numerical fluxes for nonlinear scalar conservation laws. *SIAM J. Sci. Comput.*, 15(4):892–915, 1994.
- [117] F Xiao, T Yabe, X Peng, and H Kobayashi. Conservative and oscillation-less atmospheric transport schemes based on rational functions. *J. Geophys. Res. Atmos.*, 107(D22), 2002.

## Bibliography

- [118] Y. Ha, Y. J. Lee, and J. Yoon. Modified essentially nonoscillatory schemes based on exponential polynomial interpolation for hyperbolic conservation laws. *SIAM Journal on Numerical Analysis*, 51(2):864–893, 2013.
- [119] Y. Ha, Ch. H. Kim, H. Yang, and J. Yoon. Sixth-order weighted essentially nonoscillatory schemes based on exponential polynomials. *SIAM Journal on Scientific Computing*, 38(4):A1987–A2017, 2016.
- [120] Y. B. Wang, X. Z. Jia, and J. Cheng. A numerical differentiation method and its application to reconstruction of discontinuity. *Inverse Problems*, 18(6):1461, 2002.
- [121] T. Aboiyar, E. H. Georgoulis, and A. Iske. High order WENO finite volume schemes using polyharmonic spline reconstruction. In *Proceedings of the Intl. Conference on Numerical Analysis and Approximation Theory, ClujNapoca, Romania*, pages 113–126, 2006.
- [122] A. Iske and T. Sonar. On the structure of function spaces in optimal recovery of point functionals for ENO-schemes by radial basis functions. *Numer. Math.*, 74(2):177–201, 1996.
- [123] A. Harten and S. Osher. Uniformly high-order accurate nonoscillatory schemes. I. In *Upwind and High-Resolution Schemes*, pages 187–217. Springer, 1997.
- [124] M. Dehghan and R. Jazlanian. A high-order non-oscillatory central scheme with non-staggered grids for hyperbolic conservation laws. *Computer Physics Communications*, 182(6):1284–1294, 2011.
- [125] S. Gottlieb and Ch.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Mathematics of Computation of the American Mathematical Society*, 67(221):73–85, 1998.
- [126] A. Kurganov and M. Pollack. Semi-discrete central-upwind schemes for elasticity in heterogeneous media. *SIAM Journal on Scientific Computing (Submitted for publication)*, 2011.
- [127] G. M. Hulbert and J. Chung. Explicit time integration algorithms for structural dynamics with optimal numerical dissipation. *Computer Methods in Applied Mechanics and Engineering*, 137(2):175–188, 1996.
- [128] C. Hoff and P. J. Pahl. Development of an implicit method with numerical dissipation from a generalized single-step algorithm for structural dynamics. *Computer Methods in Applied Mechanics and Engineering*, 67(3):367–385, 1988.
- [129] C. Hoff and P. J. Pahl. Practical performance of the  $\theta_1$ -method and comparison with other dissipative algorithms in structural dynamics. *Computer Methods in Applied Mechanics and Engineering*, 67(1):87–110, 1988.

## Bibliography

- [130] H. M. Hilber, Th. J. R. Hughes, and R. L. Taylor. Improved numerical dissipation for time integration algorithms in structural dynamics. *Earthquake Engineering & Structural Dynamics*, 5(3):283–292, 1977.
- [131] D. W. Zingg, H. Lomax, and H. Jurgens. High-accuracy finite-difference schemes for linear wave propagation. *SIAM Journal on Scientific Computing*, 17(2):328–346, 1996.
- [132] D. W. Zingg. Comparison of high-accuracy finite-difference methods for linear wave propagation. *SIAM Journal on Scientific Computing*, 22(2):476–502, 2000.
- [133] S. M. Day, L. A. Dalguer, N. Lapusta, and Y. Liu. Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture. *Journal of Geophysical Research: Solid Earth*, 110(B12), 2005.
- [134] S. M. Day and G. P. Ely. Effect of a shallow weak zone on fault rupture: Numerical simulation of scale-model experiments. *Bulletin of the Seismological Society of America*, 92(8):3022–3041, 2002.
- [135] Th. J. R. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Courier Corporation, 2012.
- [136] E. Tadmor. Convergence of spectral methods for nonlinear conservation laws. *SIAM Journal on Numerical Analysis*, 26(1):30–44, 1989.
- [137] M. Calhoun-Lopez and M. D. Gunzburger. The efficient implementation of a finite element, multi-resolution viscosity method for hyperbolic conservation laws. *Journal of Computational Physics*, 225(2):1288–1313, 2007.
- [138] S. K. Lele. Compact finite difference schemes with spectral-like resolution. *Journal of Computational Physics*, 103(1):16–42, 1992.
- [139] J. Waldén. Filter bank methods for hyperbolic PDEs. *SIAM Journal on Numerical Analysis*, 36(4):1183–1233, 1999.
- [140] G. Wei and Y. Gu. Conjugate filter approach for solving Burgers’ equation. *Journal of Computational and Applied mathematics*, 149(2):439–456, 2002.
- [141] Y. Gu and G. Wei. Conjugate filter approach for shock capturing. *Communications in Numerical Methods in Engineering*, 19(2):99–110, 2003.
- [142] C. Loader. Smoothing: Local regression techniques. In J. E. Gentle, W. K. Härdle, and Y. Mori, editors, *Handbook of Computational Statistics: Concepts and Methods*, pages 571–596. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [143] J. W. Banks and W. D. Henshaw. Upwind schemes for the wave equation in second-order form. *Journal of Computational Physics*, 231(17):5854–5889, 2012.

## Bibliography

- [144] D. Appelö, J. W. Banks, W. D. Henshaw, and D. W. Schwendeman. Numerical methods for solid mechanics on overlapping grids: Linear elasticity. *Journal of Computational Physics*, 231(18):6012–6050, 2012.
- [145] Y. P. Petrov and V. S. Sizikov. *Well-posed, Ill-posed, and Intermediate Problems with Applications*. Walter de Gruyter, 2005.
- [146] P. Ch. Hansen. *Rank-Deficient and Discrete Ill-posed Problems: Numerical Aspects of Linear Inversion*. SIAM, 1998.
- [147] I. R. Khan and R. Ohba. New finite difference formulas for numerical differentiation. *Journal of Computational and Applied Mathematics*, 126(1):269–276, 2000.
- [148] C. Lanczos. *Applied Analysis*, 1956.
- [149] C. W. Groetsch. Lanczo’s generalized derivative. *The American Mathematical Monthly*, 105(4):320–326, 1998.
- [150] S. K. Rangarajan and S. P. Purushothaman. Lanczos’ generalized derivative for higher orders. *Journal of Computational and Applied Mathematics*, 177(2):461–465, 2005.
- [151] Z. Wang and R. Wen. Numerical differentiation for high orders by an integration method. *Journal of Computational and Applied Mathematics*, 234(3):941–948, 2010.
- [152] T. Wei, Y. C. Hon, and Y. B. Wang. Reconstruction of numerical derivatives from scattered noisy data. *Inverse Problems*, 21(2):657, 2005.
- [153] G. Nakamura, Sh. Wang, and Y. Wang. Numerical differentiation for the second order derivatives of functions of two variables. *Journal of Computational and Applied Mathematics*, 212(2):341–358, 2008.
- [154] F. Jauberteau and J. L. Jauberteau. Numerical differentiation with noisy signal. *Applied Mathematics and Computation*, 215(6):2283–2297, 2009.
- [155] A. G. Ramm and A. Smirnova. On stable numerical differentiation. *Mathematics of Computation*, 70(235):1131–1153, 2001.
- [156] A. G. Ramm and A. Smirnova. Stable numerical differentiation: when is it possible? *Journal–Korean Society for Industrial and Applied Mathematics*, 7(1):47–62, 2003.
- [157] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

## Bibliography

- [158] J. P. Boyd and J. R. Ong. Exponentially-convergent strategies for defeating the Runge phenomenon for the approximation of non-periodic functions, part I: Single-interval schemes. *Communications in Computational Physics*, 5(2-4):484–497, 2009.
- [159] J. P. Boyd. Defeating the Runge phenomenon for equispaced polynomial interpolation via Tikhonov regularization. *Applied Mathematics Letters*, 5(6):57–59, 1992.
- [160] M. F. Hutchinson and F. R. De Hoog. Smoothing noisy data with spline functions. *Numerische Mathematik*, 47(1):99–106, 1985.
- [161] Th. C. M. Lee. Smoothing parameter selection for smoothing splines: a simulation study. *Computational Statistics & Data Analysis*, 42(1):139–148, 2003.
- [162] D. L. Ragozin. Error bounds for derivative estimates based on spline smoothing of exact or noisy data. *Journal of Approximation Theory*, 37(4):335–355, 1983.
- [163] R. Arcangéli and B. Ycart. Almost sure convergence of smoothing dm-splines for noisy data. *Numerische Mathematik*, 66(1):281–294, 1993.
- [164] D. L. Donoho. Interpolating wavelet transforms. *Preprint, Department of Statistics, Stanford University*, 2(3), 1992.
- [165] G.-Sh. Jiang and E. Tadmor. Nonoscillatory central schemes for multidimensional hyperbolic conservation laws. *SIAM Journal on Scientific Computing*, 19(6):1892–1917, 1998.
- [166] A. Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 49(3):357–393, 1983.
- [167] A. Jameson and P. D. Lax. Conditions for the construction of multi-point total variation diminishing difference schemes. *Applied Numerical Mathematics*, 2(3):335–345, 1986.
- [168] S. Osher. Riemann solvers, the entropy condition, and difference approximations. *SIAM Journal on Numerical Analysis*, 21(2):217–235, 1984.
- [169] R. Sanders. On convergence of monotone finite difference schemes with variable spatial differencing. *Mathematics of Computation - American Mathematical Society*, 40(161):91–106, 1983.
- [170] E. Tadmor. Convenient total variation diminishing conditions for nonlinear difference schemes. *SIAM Journal on Numerical Analysis*, 25(5):1002–1014, 1988.
- [171] S. Osher. Convergence of generalized MUSCL schemes. *SIAM Journal on Numerical Analysis*, 22(5):947–961, 1985.



## Bibliography

- [172] M. Berger, M. J. Aftosmis, and S. M. Murman. Analysis of slope limiters on irregular grids. In *43rd AIAA Aerospace Science Meeting*, January 2005.
- [173] S. Jain, P. Tsiotras, and H.-M. Zhou. A hierarchical multiresolution adaptive mesh refinement for the solution of evolution PDEs. *SIAM Journal on Scientific Computing*, 31(2):1221–1248, 2008.
- [174] Ch.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes, II. *Journal of Computational Physics*, 83:32–78, 1989.
- [175] K.-A. Lie and S. Noelle. On the artificial compression method for second-order nonoscillatory central difference schemes for systems of conservation laws. *SIAM Journal on Scientific Computing*, 24(4):1157–1174, 2003.
- [176] M. Brio and Ch. Ch. Wu. An upwind differencing scheme for the equations of ideal magnetohydrodynamics. *Journal of Computational Physics*, 75(2):400–422, 1988.
- [177] J. Goodman, A. Kurganov, and P. Rosenau. Breakdown in Burgers’-type equations with saturating dissipation fluxes. *Nonlinearity*, 12:247–268, 1999.
- [178] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*, volume 31. Cambridge University Press, 2002.
- [179] C.-W. Shu. TVB uniformly high-order schemes for conservation laws. *Mathematics of Computation*, 49(179):105–121, 1987.
- [180] S. D. Conte and C. De Boor. *Elementary Numerical Analysis: An Algorithmic Approach*. SIAM, 2017.
- [181] C. B. Laney. *Computational Gasdynamics*. Cambridge University Press, New York, 1998.
- [182] S. Serna. A class of extended limiters applied to piecewise hyperbolic methods. *SIAM Journal on Scientific Computing*, 28(1):123–140, 2006.
- [183] H. Yousefi, S. Mohammadi, and T. Rabczuk. Multiscale polynomial-based high-order central high resolution schemes. *Journal of Scientific Computing*, 80(1):555–613, 2019.
- [184] C. Bigoni and J. S. Hesthaven. Adaptive WENO methods based on radial basis function reconstruction. *Journal of Scientific Computing*, 72(3):986–1020, 2017.
- [185] U. S. Fjordholm and D. Ray. A sign preserving WENO reconstruction method. *Journal of Scientific Computing*, 68(1):42–63, 2016.
- [186] G. A. Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of Computational Physics*, 27(1):1–31, 1978.

## Bibliography

- [187] P. D. Lax. Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Communications on Pure and Applied Mathematics*, 7(1):159–193, 1954.
- [188] E. F. Toro and E. F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*, volume 16. Springer, 1999.
- [189] A. Kurganov and D. Levy. A third-order semidiscrete central scheme for conservation laws and convection-diffusion equations. *SIAM Journal on Scientific Computing*, 22(4):1461–1488, 2000.
- [190] G.-Sh. Jiang and Ch.-W. Shu. Efficient implementation of weighted ENO schemes. *Journal of Computational Physics*, 126(1):202–228, 1996.
- [191] A. J. Kriel. Error analysis of flux limiter schemes at extrema. *Journal of Computational Physics*, 328:371–386, 2017.
- [192] Ch. Hu and Ch.-W. Shu. Weighted essentially non-oscillatory schemes on triangular meshes. *Journal of Scientific Computing*, 150(1):97–127, 1999.
- [193] P. Woodward and Ph. Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *Journal of Computational Physics*, 54(1):115–173, 1984.
- [194] W. J. Rider, J. A. Greenough, and J. R. Kamm. Accurate monotonicity-and extrema-preserving methods through adaptive nonlinear hybridizations. *Journal of Computational Physics*, 225(2):1827–1848, 2007.
- [195] F. Kemm. On the proper setup of the double mach reflection as a test case for the resolution of gas dynamics codes. *Computers and Fluids*, 132:72–75, 2016.
- [196] J. Cullum. Numerical differentiation and regularization. *SIAM Journal on Numerical Analysis*, 8(2):254–265, 1971.
- [197] V. Barakat, B. Guilpart, R. Goutte, and R. Prost. Model-based Tikhonov-Miller image restoration. In *Image Processing, 1997. Proceedings., International Conference on*, volume 1, pages 310–313. IEEE, 1997.
- [198] D. Ruan, J. A. Fessler, M. Roberson, J. Balter, and M. Kessler. Nonrigid registration using regularization that accomodates local tissue rigidity. In *Medical Imaging*, pages 614412–614412. International Society for Optics and Photonics, 2006.
- [199] B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1-3):185–203, 1981.
- [200] S. N. Kersey. Mixed interpolating–smoothing splines and the  $\nu$ -spline. *Journal of Mathematical Analysis and Applications*, 322(1):28–40, 2006.

## Bibliography

- [201] P. Rentrop and U. Wever. Computational strategies for the tension parameters of the exponential spline. In *Optimal Control*, pages 122–134. Springer, 1987.
- [202] S. Pruess. An algorithm for computing smoothing splines in tension. *Computing*, 19(4):365–373, 1978.
- [203] S. Pruess. Properties of splines in tension. *Journal of Approximation Theory*, 17(1):86–96, 1976.
- [204] P. Rentrop. An algorithm for the computation of the exponential spline. *Numerische Mathematik*, 35(1):81–93, 1980.
- [205] T. Rohlfing, C. R. Maurer, D. A. Bluemke, and M. A. Jacobs. Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. *IEEE Transactions on Medical Imaging*, 22(6):730–741, 2003.
- [206] T. Le, R. Chartrand, and Th. J. Asaki. A variational approach to reconstructing images corrupted by Poisson noise. *Journal of Mathematical Imaging and Vision*, 27(3):257–263, 2007.
- [207] R. Chartrand and V. Staneva. Total variation regularisation of images corrupted by non-Gaussian noise using a quasi-Newton method. *IET Image Processing*, 2(6):295–303, 2008.
- [208] P. Rodríguez. Total variation regularization algorithms for images corrupted with different noise models: A review. *Journal of Electrical and Computer Engineering*, 2013:10, 2013.
- [209] B. Gautier, G. Prudon, and J. C. Dupuy. Toward a better reliability in the deconvolution of SIMS depth profiles. *Surface and Interface Analysis*, 26(13):974–983, 1998.
- [210] F. Boulsina, M. Berrabah, and J. C. Dupuy. Deconvolution of SIMS depth profiles: Towards simple and faster techniques. *Applied Surface Science*, 255(5):1946–1958, 2008.
- [211] D. Terzopoulos. Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (4):413–424, 1986.
- [212] M. L. de Silanes. Convergence and error estimates for  $(m, l, s)$ -splines. *Journal of Computational and Applied Mathematics*, 87(2):373–384, 1997.
- [213] B. Fornberg. Classroom note: calculation of weights in finite difference formulas. *SIAM Review*, 40(3):685–691, 1998.

## Bibliography

- [214] K. Yomogida and R. Benites. Scattering of seismic waves by cracks with the boundary integral method. In *Seismic Waves in Laterally Inhomogeneous Media*, pages 1771–1789. Springer, 2002.
- [215] J. Kawahara and T. Yamashita. Scattering of elastic waves by a fracture zone containing randomly distributed cracks. *Pure and Applied Geophysics*, 139(1):121–144, 1992.
- [216] S.-K. Youn and S.-H. Park. A new direct higher-order Taylor-Galerkin finite element method. *Computers & Structures*, 56(4):651–656, 1995.
- [217] X. D. Li and N.-E. Wiberg. Structural dynamic analysis by a time-discontinuous Galerkin finite-element method. *International Journal for Numerical Methods in Engineering*, 39(12):2131–2152, 1996.
- [218] J. Sochacki, R. Kubichek, J. George, W. R. Fletcher, and S. Smithson. Absorbing boundary conditions and surface waves. *Geophysics*, 52(1):60–71, 1987.
- [219] H. Liu and J. Qiu. Finite difference Hermite WENO schemes for hyperbolic conservation laws. *Journal of Scientific Computing*, 63(2):548–572, 2015.
- [220] J. Qiu and Ch.-W. Shu. Hermite WENO schemes for Hamilton–Jacobi equations. *Journal of Computational Physics*, 204(1):82–99, 2005.
- [221] Y. Ha, Y. J. Lee, and J. Yoon. Modified essentially nonoscillatory schemes based on exponential polynomial interpolation for hyperbolic conservation laws. *SIAM Journal on Numerical Analysis*, 51(2):864–893, 2013.
- [222] Zh. J. Wang. Spectral (finite) volume method for conservation laws on unstructured grids. Basic formulation: Basic formulation. *Journal of Computational Physics*, 178(1):210–251, 2002.
- [223] I. Christov and B. Popov. New non-oscillatory central schemes on unstructured triangulations for hyperbolic systems of conservation laws. *Journal of Computational Physics*, 227(11):5736–5757, 2008.
- [224] A. Kurganov and G. Petrova. Central-upwind schemes on triangular grids for hyperbolic systems of conservation laws. *Numerical Methods for Partial Differential Equations*, 21(3):536–552, 2005.
- [225] M. Jansen, R. Baraniuk, and S. Lavu. Multiscale approximation of piecewise smooth two-dimensional functions using normal triangulated meshes. *Applied and Computational Harmonic Analysis*, 19(1):92–130, 2005.
- [226] F. Arandiga, A. Cohen, R. Donat, N. Dyn, and B. Matei. Approximation of piecewise smooth functions and images by edge-adapted (ENO-EA) nonlinear multiresolution techniques. *Applied and Computational Harmonic Analysis*, 24(2):225–250, 2008.

## Bibliography

- [227] W. J. Rider and D. B. Kothe. Constrained minimization for monotonic reconstruction. In *Proceedings of the Thirteenth AIAA Computational Fluid Dynamics Conference*, pages 955–964, 1997.
- [228] G. Awanou, M.-J. Lai, and P. Wenston. The multivariate spline method for scattered data fitting and numerical solutions of partial differential equations. *Wavelets and Splines: Athens*, pages 24–74, 2005.
- [229] Gh. Micula. A variational approach to spline functions theory. *Universitae Politecnico di Torino*, page 209, 2003.
- [230] I. J. Schoenberg. On trigonometric spline interpolation. *Journal of Mathematics and Mechanics*, 13(5):795, 1964.
- [231] P. H. Gaskell and A. K. C. Lau. Curvature-compensated convective transport: SMART, a new boundedness- preserving transport algorithm. *International Journal for Numerical Methods in Fluids*, 8(6):617–641, 1988.
- [232] W. G. Hoover, C. G. Hoover, O. Kum, V. M. Castillo, H. A. Posch, and S. Hess. *Smooth Particle Applied Mechanics*. World Scientific Singapore, 2006.
- [233] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method, Volume 1, The Basis*. John Wiley & Sons Canada, Ltd., 2000.

# List of Figures

2.1	The projection of the operator $d/dx$ into wavelet spaces by the wavelet Db-12. . . . .	8
2.2	Non-uniform cells obtained by a GF (growth factor) for simulation of flow around a rectangle obstacle [78]. . . . .	23
3.1	Schematic illustration for calculation of the D-D wavelet transform. . . .	26
3.2	Grid adaptation for a discontinuous and a smooth but high-gradient data where $\epsilon = 10^{-5}$ . a) the considered data (solid line) and adapted points (solid points); b) distribution of adapted points in different resolution levels: solid points are distribution of multiresolution-based sampled points, and hollow circles denotes the adapted points. . . . .	28
3.3	Spatial locations of points belonging to subspaces $V_2$ and $V_3$ , where $\bullet \in V_2$ and $\bullet \cup \circ \in V_3$ , for $(x_i, y_j) \in [0, 1] \times [0, 1]$ . . . . .	29
3.4	Modification of an adapted grid; solid points and hollow circles correspond to adapted and extra added points, respectively. For grid modification stage, we assume: $N_s = N_c = 1$ . . . . .	31
3.5	Adding procedure in the same resolution for point types $s_1, s_2$ or $s_3$ , where $\bullet \in V_2, \bullet \cup \circ \in V_3, \circ \in W_2$ and bright gray solid points are the added points belonging to $W_2$ . . . . .	32
3.6	Adding procedure in the successive coarser resolution for point types $s_1, s_2$ or $s_3$ , where $\bullet \in V_2, \bullet \cup \circ \in V_3, \circ \in W_2$ and bright gray solid points are the added points belonging to $V_2$ . . . . .	32
4.1	The reconstruction/evolution/projection concept: the second order accuracy. In this figure $x_{i\pm k, m} \equiv x_{i\pm k, m}^n$ , for $m \in \{l, r\}$ . . . . .	40
4.2	The first local TVD condition. . . . .	53
4.3	The second local TVD condition. . . . .	54
4.4	The third local TVD condition. . . . .	54
4.5	The definition of the parameter $f$ . . . . .	58
4.6	The performance of GMINMOD limiters on non-uniform grids. Regarding the top row in each figure, the hollow circles and the solid points represent the center of cells and the cell edges, respectively. . . . .	62
4.7	Wavelet based adaptive distribution of cell centers $x_j$ , corresponding edge locations and TVD feature of the GMINMOD limiter on resulted non-uniform grids. . . . .	64

*List of Figures*

4.8	Effects of a weaker restriction on wavelet-based grid variation. For this figure, we assume $x_{j+1} - x_j = 0.5(x_j - x_{j-1})$ ; a) cell configurations; b) the GMINMOD limiter performance. . . . .	70
4.9	The MRA-based algorithm for the simulation of hyperbolic problems. . .	72
4.10	Numerical solutions and corresponding adapted grid points represented in different resolution levels; a,b) for $t = 0.4$ ; c,d) for $t = 0.8$ . . . . .	73
4.11	Number of grids points, $N_g$ during time; at times $t = 0.75$ and $t = 0.85$ discontinuous fronts reach the boundaries. In the finest resolution, number of grid points is 2049. . . . .	74
4.12	Numerical results of the Burgers' equation and corresponding adapted grid points in different resolution levels; in figure (a), exact solutions are presented by the solid line and numerical solutions are illustrated by points and hollow shapes. In the finest resolution, number of grid points is 2049. . . . .	75
4.13	Number of grid points $N_g$ used during the simulation; the uniform grid has 2049 points. . . . .	76
4.14	Parameter effect of the generalized MINMOD slope limiter; a) result for $\theta = 2$ ; b) solution for $\theta = 1$ . . . . .	76
4.15	Comparison of the proposed adaptive central method and NVSF-based upwind scheme. . . . .	77
4.16	The Comparison between the adaptive KT and ENO-Roe [173] schemes on MRA-based adapted grids at $t = 0.158, 0.5$ and $1$ . . . . .	79
4.17	The Comparison between the adaptive KT and SPR-FD [54] schemes on MRA-based adapted grids at $t = 0.25$ , where $J_{max} = 14$ and $J_{min} = 4$ . . .	80
4.18	$N_g$ values during simulations for different $\epsilon$ values; a) adapted grids, b) comparison of $N_g$ values between the uniform grid and corresponding adapted ones. . . . .	81
4.19	Adapted solutions for different threshold values at $t = 0.04$ ; a, b) $\epsilon = 10^{-3}$ , c, d) $\epsilon = 0.5 \times 10^{-3}$ , e, f) $\epsilon = 10^{-4}$ . Solid line and shapes are the exact solution and adapted ones, respectively. In all calculations it is assumed that: $N_c = 1$ and $N_s = 2$ . . . . .	81
4.20	Effect of adaptive simulation without grid modification stage. a) Adapted solution and corresponding adapted grid points in case $\{N_c = N_s = 0\}$ ; b) $N_g$ values of solutions in two cases: 1) with grid modification ( $\{N_c = 1 \ \& \ N_s = 2\}$ ), 2) without grid modification ( $\{N_c = N_s = 0\}$ ). . . . .	82
4.21	$N_g$ values during simulations for different threshold values; it is assumed $\epsilon_0 = 5 \times 10^{-4}$ , $\theta = 2$ ; $J_{max} = 11$ and $N_d = 6$ . . . . .	82
4.22	The solution at $t = 1.2$ and corresponding adapted points for $\epsilon = \epsilon_0$ ; the solid line and hollow circle illustrate exact and numerical solutions, respectively. . . . .	83
4.23	Numerical solutions and corresponding adapted grids of interaction of two blast waves with parameters $\epsilon = 10^{-2}$ , and $\theta = 2$ ; a-b) $t = 0.01$ ; c-d) $t = 0.016$ ; e-f) $t = 0.026$ ; g-h) $t = 0.028$ ; . . . . .	85
4.23	(continued) i-j) $t = 0.03$ ; k-l) $t = 0.032$ ; m-n) $t = 0.034$ ; o-p) $t = 0.038$ . .	86

*List of Figures*

4.24	Reference and numerical solutions of entropy and shock wave interaction at $t = 1.8$ . Solid lines and shapes are the reference and adaptive results, respectively. In all calculations it is assumed that: $N_c = 1$ and $N_s = 2$ . . . . .	87
4.25	Threshold effects on density approximations; it is assumed that $J_{max} = 13$ , $N_c = 1$ , and $N_s = 2$ . . . . .	88
4.26	Threshold effects on adaptation; it is assumed that $J_{max} = 13$ , $N_c = 1$ , and $N_s = 2$ . . . . .	88
4.27	$N_g$ values during simulations for different $\epsilon$ values; it is assumed $\epsilon_0 = 5 \times 10^{-4}$ and $J_{max} = 13$ ; a) adapted grids, b) comparison of $N_g$ values between the uniform and the adapted grids. . . . .	89
4.28	Threshold effects on solutions and adaptations; it is assumed that $J_{min} = 5$ , $\epsilon = \epsilon_0$ , $N_c = 1$ , and $N_s = 2$ . . . . .	90
4.29	Adaptive solutions and corresponding distribution of adapted points in different resolutions for various $J_{max}$ values at $t = 0.12$ . In all calculations, it is assumed that: $N_c = 1$ and $N_s = 2$ . . . . .	92
4.30	Effect of the $J_{max}$ values regarding dispersion and dissipation phenomena; the illustrations are from zoomed in solutions at $t = 0.12$ . . . . .	92
4.31	Number of grid points used in adaptive solutions for different $J_{max}$ values; a) adapted grids, b) comparison of $N_g$ values between the uniform grids with different finest resolutions and corresponding adapted ones. . . . .	93
4.32	Number of grid points ( $N_g$ ) used in the adaptation procedure for different threshold values. . . . .	94
4.33	Numerical solutions of the non-linear wave propagation problem and corresponding adapted grids where $\epsilon = 10^{-6}$ and $\theta = 2$ ; a-b) $t = 80$ ; c-d) $t = 160$ ; e-f) $t = 240$ . In the top row, solid lines are reference solutions, and shapes are numerical ones. In all calculations it is assumed that: $N_c = 1$ and $N_s = 2$ . . . . .	95
4.34	Adapted solution for the threshold value $\epsilon = \epsilon_0 = 6 \times 10^{-2}$ at $t = 1.5$ . Here, the solid line denotes the asymptotic solution. In all calculations it is assumed that: $N_c = 1$ and $N_s = 2$ . . . . .	96
4.35	Number of grid points ( $N_g$ ) used in the adaptation procedure for the threshold value $\epsilon = \epsilon_0 = 6 \times 10^{-2}$ ; a) adapted grids, b) comparison of $N_g$ values between the uniform and the corresponding adapted grids. . . . .	96
4.36	The solution at $t = 0.4$ and corresponding adapted points; a) the exact and numerical results, b) distribution of adapted points in different resolutions. In (a), the solid line and hollow circle illustrate exact and numerical solutions, respectively. . . . .	96
4.37	Adaptive simulation of the Buckley-Leverett equation with and without the gravitational effects; a) adaptive solutions; b) distribution of adapted points in different resolution levels for the case with the gravity effects; c) distribution of adapted points for the case without the gravity effects. In all calculations it is assumed that: $N_c = 1$ and $N_s = 2$ . . . . .	98



*List of Figures*

4.38	Number of grid points, $N_g$ , used for adaptive simulations; a) adapted grids, b) comparison of $N_g$ values between the uniform grid and corresponding adapted ones. . . . .	98
4.39	Right propagating ice shield simulated by the adaptive method and corresponding distribution of adapted points in different resolution levels; shapes are the adaptive solutions and the solid lines are the reference ones obtained by SD2 method with 800 uniform cells. . . . .	100
4.40	Number of grid points $N_g$ used for adaptive simulation; a) adapted grids, b) comparison of $N_g$ values between the uniform grid and corresponding adapted one. . . . .	100
4.41	Glacier growth; the shapes are numerical solutions and the solid lines are the reference ones; where $N_c = 1$ and $N_s = 2$ . . . . .	102
4.42	Threshold effects on the localized feature detection; hollow circles and solid line are the adaptive and reference solutions, respectively. . . . .	103
4.43	Number of grid points during adaptive solutions; a) adapted grids, b) comparison of $N_g$ values between the uniform grid and corresponding adapted ones. . . . .	103
4.44	Effect of threshold values on adaptive solution and the corresponding grid; a-b) $\epsilon = 10^{-4}$ ; c-d) $\epsilon = 2.5 \times 10^{-4}$ ; e-f) $\epsilon = 10^{-3}$ . . . . .	104
4.45	Number of grid points ( $N_g$ ) used in the adaptation procedure for different threshold values. . . . .	104
4.46	2-D adaptive solution of the Burgers' type equation at $t = 0.7$ ; a) adapted grid; b) numerical solution. In calculations it is assumed: $N_c = 2$ , $N_s = 2$ , $\epsilon = 0.25 \times 10^{-3}$ . . . . .	106
4.47	Comparison of 2-D adapted solution with 1-D one; the 1-D solution is obtained on a uniform grid of 200 cells. . . . .	106
4.48	Adapted grid points for different $\epsilon'$ values; a-c) $\epsilon' = 0.1$ ; d-f) $\epsilon' = 0$ . In the calculations it is assumed: $\epsilon = 0.25 \times 10^{-3}$ , $N_c = 2$ , $N_s = 2$ . . . . .	107
4.49	Solution of 2-D Burgers' equation for different $\epsilon'$ values; a-c) $\epsilon' = 0.1$ ; d-f) $\epsilon' = 0$ . . . . .	108
4.50	Number of grid points used in adaptive solutions for different $\epsilon'$ values. . . . .	108
4.51	Solutions and corresponding adapted grid points of the shallow water problem, where: $\epsilon = 0.25 \times 10^{-3}$ , $N_c = 1$ , $N_s = 2$ . . . . .	109
5.1	Estimation of the slope $(u_x)^{n+1}$ in different schemes; a) KT and M1, b) M2. . . . .	112
5.2	The variation of some commonly used limiters in terms of the parameter $f$ ; the black point represent the linearity preserving location. . . . .	121
5.3	The modified van Leer limiters on a transmitting cell. . . . .	123
5.4	The modified van Albada limiter on a transmitting cell, derived from a weighted harmonic average. . . . .	125
5.5	The modified van Albada limiters on a transmitting cell with different $n$ values. . . . .	126
5.6	The performance of the SuperBee limiter on non-uniform cells. . . . .	127

*List of Figures*

5.7	The performance of $(\phi_j D_j)_1$ and $(\phi_j D_j)_2$ limiters; a) The limiter $(\phi_j D_j)_1$ ; b) The limiter $(\phi_j D_j)_2$ . . . . .	128
5.8	The function $f_1(x)$ (the solid lines) and corresponding adapted grid points (the plot markers $\bullet$ ) in different resolution levels for different $J_{max}$ values; in all illustrations, $J_{min} = 5$ and $\epsilon = 0.5 \times 10^{-4}$ . . . . .	135
5.9	The function $f_2(x)$ (the solid lines) and corresponding adapted grid points (the plot markers $\bullet$ ) in different resolution levels for different $J_{max}$ values; in all illustrations, $J_{min} = 5$ and $\epsilon = 0.5 \times 10^{-4}$ . . . . .	136
5.10	Numerical solutions of the Burgers' equation with different TVD second-order limiters at $t = 0.159$ . . . . .	140
5.11	Distribution of adapted cells in different resolution levels for the Burgers' equation obtained with different TVD second-order limiters at $t = 0.159$ . . . . .	141
5.12	Numerical results, corresponding truncation errors, entropy productions, and adapted points in different resolutions for the Sod problem with the KT, $M1$ and $M2$ schemes at $t = 0.2$ ; a-c) the KT scheme; d-f) the $M1$ method; g-i) the $M2$ scheme. . . . .	143
5.13	Number of adapted grid points $N_g$ during simulations. In the finest resolution, number of grid points is $2^{11} + 1$ . . . . .	143
5.14	Post-processing effects on stability of solutions; a-b) with the post-processing stage; c-d) without the post-processing step. In figures (a) and (c), solid lines and hollow shapes are exact and numerical solutions, respectively. . . . .	144
5.15	Full and semi post-processing effects; a) full post processing, $N_s = N_c = 2$ ; b) semi-post processing, $N_s = 2$ and $N_c = 0$ . Solid lines and hollow shapes are exact and numerical solutions, respectively. . . . .	145
5.16	Numerical results, corresponding entropy production, truncation errors, and adapted points of different resolutions for the Lax problem at $t = 0.1$ ; a-c) the KT scheme; d-f) the $M1$ method; g-i) the $M2$ scheme. . . . .	146
5.17	Numerical results, and entropy productions for the right propagating front with the KT, $M1$ and $M2$ schemes at $t = 1.8$ . In these figures, solid black lines are the reference solutions, hollow shapes are numerical ones, and the gray solid lines are numerical entropy productions. . . . .	147
5.18	Distribution of adapted grid points in different levels of resolution for the KT, $M1$ and $M2$ schemes at $t = 1.8$ . . . . .	148
5.19	Local truncation errors for the right propagating front with the KT, $M1$ and $M2$ schemes at $t = 1.8$ . . . . .	148
5.20	Adaptive numerical solutions ( $\rho$ ) for the right propagating front obtained with KT scheme using the GMM and the MM-based UNO limiters at $t = 1.8$ . . . . .	149
5.21	The 2-D Euler gas dynamic adaptive solutions with corresponding adapted grids at $t = 0.4$ . . . . .	151
5.22	Comparison of numerical results obtained by the KT and $M1$ schemes along $y = 0$ at $t = 0.4$ . . . . .	151
5.23	Local truncation errors for the KT and $M1$ schemes at $t = 0.214$ . . . . .	152

*List of Figures*

5.24	Numerical solution of the non-convex conservation law- case 1 with corresponding local truncation error; a, b) constant parameter $\theta = 2$ and corresponding local truncation error; c,d) with adaptive $\theta$ and corresponding local truncation error; e) number of adapted grid points during simulations.	153
5.25	Numerical solution of the non-convex conservation law- case 2 with corresponding local truncation error; a, b) constant parameter $\theta = 2$ and corresponding local truncation error; c,d) with adaptive $\theta$ and corresponding local truncation error; e) number of adapted grid points during simulations.	154
5.26	Numerical solutions, corresponding adapted grids and local truncation errors for the polymer system; a, b) with $\theta = 2$ ; c,d) with adaptive $\theta$ ; e,f) adaptive $\theta$ .	155
5.27	Adaptive solutions of 2-D non-convex conservation law system at $t = 1$ ; a, b) obtained with $\theta$ -adaptive KT scheme; c, d) based on the KT scheme with $\theta = 2$ .	157
6.1	The REP procedure for the third-order central scheme.	160
6.2	The three-average solutions used for evaluation of $q_j(x)$ .	160
6.3	Different variation patterns over three successive cells $\{I_j\}$ .	167
6.4	The fourth reconstruction feature at max and min cells.	170
6.5	Possible forming of non-physical extrema at cell-edges.	170
6.6	Parameters used for scaling limiting at cell-edges for monotone increasing data.	172
6.7	CWENO scheme defined over three cells with three sub-stencils.	175
6.8	CWENO scheme defined over five cells with three sub-stencils.	177
6.9	Using local stencil for the first PPM method.	180
6.10	Modification of $u'_j(x) = v_j(x)$ by limiting $u''(j)$ , to prevent sign changing of $v_j(x)$ . The original data and the limited one are denoted by the red dashed the solid blue lines, respectively.	182
6.11	Using local stencil for the second PPM method.	183
6.12	Modification of $v_j(x)$ , the red dashed line, to prevent sign changing of $v_j(x)$ as the solid blue line.	184
6.13	The first test function $f_1(x)$ and distribution of adapted grids in different resolution levels for different $J_{max}$ values.	189
6.14	The second test function $f_2(x)$ and distribution of adapted grids in different resolution levels for different $J_{max}$ values.	191
6.15	Numerical results and corresponding adapted points in different resolutions for the Burgers' problem with different approaches.	196
6.16	Number of adapted grid points $N_g$ and corresponding $TV$ during simulations with different approaches. In the finest resolution, the number of grid points is $2^{10} + 1$ .	197
6.17	Numerical results of $\rho$ ( $u \equiv \rho$ ) and corresponding adapted points in different resolutions for the Sod problem with different approaches at $t = 0.2$ . Solid lines and shapes denote exact and numerical solutions, respectively.	199

List of Figures

6.18	Number of adapted grid points $N_g$ and corresponding total variations $TV$ of $\rho$ for the Sod problem during simulations with different approaches. In the finest resolution, number of grid points is $2^{11} + 1$ . . . . .	200
6.19	$L^1$ error as a function of run-time for different schemes for the Sod problem, evaluated for $0 \leq t \leq 0.2$ . . . . .	200
6.20	Numerical results of $\rho$ ( $u \equiv \rho$ ) and corresponding adapted points in different resolutions for the Lax problem with different approaches at $t = 0.16$ .	201
6.21	Number of adapted grid points $N_g$ and corresponding $TV$ of $\rho$ during simulations for the Lax problem with different approaches. In the finest resolution, the number of grid points is $2^{11} + 1$ . . . . .	202
6.22	$L^1$ error as a function of run-time for different schemes for the Lax problem, evaluated for $0 \leq t \leq 0.16$ . . . . .	202
6.23	Numerical results from adaptive solvers for the entropy and shock wave interaction with $M_1$ , $M_2$ and $CWENO$ methods at $t = 1.8$ . . . . .	205
6.24	Numerical results from adaptive solvers for the entropy and shock wave interaction with $PPM - 1$ and $PPM - 2$ methods at $t = 1.8$ ; the $KT$ , $C_1$ and $C_2$ second order central high resolution schemes with the GMM limiter (with $\theta = 2$ ) are used for the evaluations of velocities and accelerations.	206
6.25	$L^1$ error as a function of run-time for different schemes for the Shu-Osher problem, evaluated for $0 \leq t \leq 1.8$ . . . . .	207
6.26	Numerical results from adaptive different solvers for the interaction of two shock waves; in each row, solutions are for times 0.016, 0.028 and 0.032 respectively from left to right. . . . .	208
6.27	Numerical results for the 2-D Burgers' problem with different schemes for $\Omega \in [-1, 1] \times [-1, 1]$ at $t = 0.5$ . . . . .	210
6.28	The 3-D representations of numerical results for the 2-D Burgers' problem with different schemes for $\Omega \in [-1, 1] \times [-1, 1]$ at $t = 0.5$ . . . . .	211
6.29	Adapted grids for the 2-D Burgers' problem at different times for the $M_1$ scheme where $\Omega \in [-2, 2] \times [-2, 2]$ . For adaptation, we have: $J_{max} = 8$ , $J_{min} = 5$ and $\epsilon = 0.25 \times 10^{-3}$ . . . . .	211
6.30	Numerical results obtained by the $M_1$ , $M_2$ , $CWENO-3P$ and $CWENO-5P$ schemes at $t = 0.2$ ; the number of contours is 30. . . . .	214
6.31	Adapted grids for the numerical simulations of the $M_1$ , $M_2$ , $CWENO-3P$ and $CWENO - 5P$ schemes at $t = 0.2$ . . . . .	215
6.32	Numerical results obtained by the $PPM - 1$ and $PPM - 2$ schemes at $t = 0.2$ ; the number of contours is 30. . . . .	216
6.33	Adapted grids for the numerical simulations of the $PPM-1$ and $PPM-2$ schemes at $t = 0.2$ . . . . .	217
6.34	Numerical results obtained by the $M_1$ , $M_2$ , $CWENO-3P$ and $CWENO-5P$ schemes at $t = 3$ ; the number of contours is 30. . . . .	219
6.35	Adapted grids for the numerical simulations of the $M_1$ , $M_2$ , $CWENO-3P$ and $CWENO - 5P$ schemes at $t = 3$ . . . . .	220
6.36	Adapted solutions (with 30 contours) and their grids at times 0.04, 0.17, 0.35 and 0.60 obtained with the $CWENO - 3P$ scheme. . . . .	221

List of Figures

6.37	Adapted solutions obtained with the PPM schemes with different resolution levels at $t = 1.8$ with 30 contours. . . . .	222
7.1	Numerical solutions with different Tikhonov methods with parameters $p = 0.875$ and $\alpha = 1$ at $t = 0.2$ . The numerical dispersion appears in solutions obtained by the constraints $\Omega_1^2$ and $\Omega_3^2$ . . . . .	229
7.2	Numerical solutions of the wave propagation problem, denoised with the model-based Tikhonov scheme with constraints $\Omega_1^2$ , $\Omega_2^2$ and $\Omega_3^2$ at $t = 0.2$ ; for each figure, for pair $\{\mathbf{f}_{model}, \mathbf{f}_{min}\}$ , corresponding the constraint pairs are: a) $\{\Omega_1^2, \Omega_1^2\}$ ; b) $\{\Omega_2^2, \Omega_2^2\}$ , c) $\{\Omega_3^2, \Omega_3^2\}$ , d) $\{\Omega_{TV}, \Omega_1^2\}$ , e) $\{\Omega_{TV}, \Omega_2^2\}$ , f) $\{\Omega_{TV}, \Omega_3^2\}$ . . . . .	230
7.3	Numerical solutions of the wave propagation problem having discontinuous fronts, denoised with the model-based Tikhonov scheme with constraints $\Omega_1^2$ , $\Omega_2^2$ and $\Omega_3^2$ at $t = 0.2$ ; for each figure, for the pair $\{\mathbf{f}_{model}, \mathbf{f}_{min}\}$ , corresponding constraints are: a) $\{\Omega_1^2, \Omega_1^2\}$ ; b) $\{\Omega_2^2, \Omega_2^2\}$ , c) $\{\Omega_3^2, \Omega_3^2\}$ , d) $\{\Omega_{TV}, \Omega_1^2\}$ , e) $\{\Omega_{TV}, \Omega_2^2\}$ , f) $\{\Omega_{TV}, \Omega_3^2\}$ . . . . .	231
7.4	The L-curves for numerical solutions of the 1D wave propagation problem at $t = 0.2$ , denoised with different regularization approaches. . . . .	232
7.5	Convergence rates in the $l_2$ norm for different FD methods for smooth (top row) and discontinuous (bottom row) solutions; a) different compact FDs without filtering; b, e) different compact FDs with filtering; c,f) the inherent filtering approach; d,g) Tikhonov based smoothing by a post-processor stage. . . . .	238
7.6	Convergence rates in the $l_1$ norm for different FD methods for discontinuous solutions. . . . .	238
7.7	Discontinuous solutions $u(x, t = 0.2)$ for different FD schemes and different filtering approaches. . . . .	239
7.8	Filter coefficients for different constraints in the Tikhonov method, where $p = 0.8$ & $\alpha = 1$ . . . . .	241
7.9	Regularized solutions with different constraints, where $p = 0.8$ & $\alpha = 1$ . . . . .	241
7.10	Control of the discrete Picard condition for regularized solutions with different constraints, where $p = 0.8$ & $\alpha = 1$ . . . . .	242
7.11	The test function $y(x)$ . . . . .	248
7.12	Error in estimations, the numerical dispersion and Runge phenomena resulted from the Tikhonov method with different constraints; (a) & (b) regularization with constraint $\Omega_1^2$ ; (c) & (d) regularization with constraint $\Omega_2^2$ ; . . . . .	249
7.13	Convergence rates for Tikhonov regularization in $L_1$ and $L_2$ norms with the constraints $\Omega_1^2$ and $\Omega_2^2$ for different $p$ and $\alpha$ values. . . . .	250
7.14	Variable $\alpha$ parameter in Tikhonov-based regularization with the constraint $\Omega_2^2$ ; a) regularized results ( $f$ ) and $\alpha(x)$ ; b) estimation errors; c) convergence rates. . . . .	251
7.15	Local smoothing with $C^2$ continuity with surrounding data; value of smoothing parameter is $p = 0.7$ . . . . .	256

*List of Figures*

7.16	Numerical results for wave propagation in a bar due to the box-shaped imposed initial condition at $t = 0.2$ . . . . .	259
7.17	Direct simulation of the second order scalar wave propagation problem at $t = 0.2$ ; symbols $P - 4$ , $UW2$ , $UW4$ and $HR2$ denote the 4 <sup>th</sup> order Padé, second order upwind, fourth order upwind and second-order high resolution approximations, respectively. The results of $UW2$ , $UW4$ and $HR2$ are from [143]. . . . .	261
7.18	Numerical results with different assumptions of regularization constraints at $t = 0.24$ ; a) the IC condition ( $U(x, y)$ ); b) the solution based on using the constraint $\Omega_2^2$ ; c) the solution based on using the constraint $\Omega_1^2$ ; d) the solution without any regularization. . . . .	262
7.19	Comparison of the numerical method having the post-processing stage with constraint $\Omega_2^2$ , and the modal analysis based solution (with $60 \times 60$ first modes) at spatial position $(x, y) = (0.5, 0.5)$ . . . . .	263
7.20	Schematic shape of an infinite-periodic medium including a fluid-filled narrow crack. . . . .	264
7.21	Snapshots of solutions $u_x$ and $u_z$ . . . . .	265
7.22	Snapshot of solutions $u_z$ at $t = 0.067$ . . . . .	266
7.23	A homogeneous medium with stochastic fluid-filled small cavities. . . . .	266
7.24	Formation of stochastic-like solutions $u_x$ and $u_z$ from the fluid-filled cavities due to diffraction of the incident plane $P$ wave. . . . .	267
7.25	Snapshots of the P-SV wave propagation in a medium having fluid-filled cavities with stochastic locations; a) the solution with the regularization step with the constraint $\Omega_2^2$ ; b) the solution without the regularization step. . . . .	268
A.1	Continuous variation of the state variable and flux through cells. . . . .	275
A.2	Simulation of the Burgers' equation with numerical solution using continuous variable and flux. . . . .	276
A.3	Piecewise linear variation of the state variable and flux on cells. . . . .	277
A.4	Variation of solutions over cells (reconstructed and evolved ones) for the LxF scheme; (a) Staggered formulation, (b) Non-staggered one. $\{x_i\}$ and $\{x_{i\pm 1/2}\}$ denote cell-centers and cell-edges, respectively. . . . .	278
A.5	Numerical results for the Burgers' equation obtained by the LxF scheme at $t=0.158$ , $t=0.5$ and $t=1$ ; solid lines and hollow circles denote the exact and numerical solutions, respectively. . . . .	280
A.6	Illustration of the NT scheme. . . . .	282
A.7	Numerical results for the Burgers' equation by using the NT method at $t=0.158$ , $t=0.5$ and $t=1$ ; solid lines and hollow circles denote the exact and numerical solutions, respectively. . . . .	287
A.8	Illustration of the KT scheme. . . . .	288
A.9	Numerical results of the Burgers' equation by the KT method at $t=0.158$ , $t=0.5$ and $t=1$ ; solid lines and hollow circles denote the exact and numerical solutions, respectively. . . . .	296

*List of Figures*

B.1 Definition of the upstream and downstream cells based on the velocity  $a_{i+1/2} = a_f$ . . . . . 298

# List of Tables

5.4.1	Reconstruction error for $f_1(x)$ at cell-interfaces for the norms with $p_0 = 1$	134
5.4.2	Reconstruction error for $f_2(x)$ at cell-interfaces for the norms with $p_0 = 1$	136
6.4.1	The computation cost of different high-order central schemes on uniform cells; for all cases $J_{min} = 5$ .	186
6.4.2	The computation cost of different high-order central schemes on adaptive cells. The parameter $C_{MS}$ is measured in Minutes and $J_{min} = 5$ for all cases.	187
6.5.1	Reconstruction error at interfaces with $p = 1$ for the first test function.	189
6.5.2	Reconstruction error at interfaces with $p = 1$ for the first test function.	190
6.5.3	Reconstruction error at interfaces with $p = 1$ for the second test function.	192
6.5.4	Reconstruction error at interfaces with $p = 1$ for the second test function.	192
6.6.1	The accuracy of different schemes for solving the Sod problem at $t = 0.2$ .	198
6.6.2	Accuracy for different schemes for solving the Lax problem at $t = 0.16$ .	203
7.1.1	Different constraint definitions in the $L_2$ space for the Tikhonov method.	226
7.3.1	Coefficients of generalized Padé approximations for the second derivative with different accuracy and periodic boundary conditions [138].	235
7.3.2	Filter coefficient values used for filtering the generalized Padé approximations by a post-processing stage for the periodic boundary condition [138].	235



# Academic Curriculum Vitae

## Hassan Yousefi

Institute of Structural Mechanics  
Bauhaus-Universität Weimar  
Marienstrasse 15, 99423 Weimar, Germany  
Email: hyousefi@ut.ac.ir  
Date of Birth: Sep 15th, 1978  
Place of Birth: Tehran, Iran  
Citizenship: Iranian

## Education

1. Postdoctoral: Bauhaus University Weimar, Germany (research advisor: Prof. Dr. Timon Rabczuk), Oct. 2012- Sep. 2015,
2. First PhD in Civil Engineering-Earthquake Engineering, School of Civil Engineering, Faculty of Engineering, University of Tehran, Tehran, Iran, 2005-2010,
3. MSc in Civil Engineering-Earthquake Engineering, School of Civil Engineering, Faculty of Engineering, University of Tehran, Tehran, Iran, 2003-2005,
4. BSc in Civil Engineering, Department of Civil and Environmental Engineering, Polytechnic of Tehran (Amir-Kabir University), Tehran, Iran, 2000-2003.

## Publications: International peer-reviewed journals

*Derived from the thesis*

1. H. Yousefi, T. Rabczuk. Wave propagation in generalized thermo-poro-elastic media via wavelet-based cell-adaptive central high resolution schemes using UNO limiters, *Applied Numerical Mathematics* (2021) 173, 112-143,
2. H. Yousefi, S. Mohammadi, T. Rabczuk. Multiscale Polynomial-Based High-Order Central High Resolution Schemes, *Journal of Scientific Computing* (2019) 80(1), 555-613,
3. H. Yousefi, T. Rabczuk. Multiresolution-based adaptive central high resolution schemes for modeling of nonlinear propagating fronts. *Engineering Analysis with Boundary Elements* (2019) 103, 172-195,
4. H. Yousefi, Sh. Ghorashi, T. Rabczuk. Directly Simulation of Second Order Hyperbolic Systems in Second Order Form via the Regularization Concept, *Communication in Computational Physics* (2016) 20(1), 86-135,

5. H. Yousefi, T. Rabczuk. A comparative study of different types of slope limiters over wavelet-based adapted cells via central high resolution schemes (submitted).

*From other projects*

1. H. Yousefi, A. Taghavi, I. Mahmoudzadeh. Multiscale RBF-based central high resolution schemes for simulation of generalized thermoelasticity problems, *Frontiers of Structural and Civil Engineering* (2019) 13(2), 429-455,
2. H. Yousefi, J. Farjoodi, I. Mahmoudzadeh. Adaptive simulation of wave propagation problems including dislocation sources and random media, *Frontiers of Structural and Civil Engineering* (2019) 13(5), 1054-1081,
3. H. Yousefi, A. Taghavi, I. Mahmoudzadeh. Wavelet-based iterative data enhancement for implementation in purification of modal frequency for extremely noisy ambient vibration tests in Shiraz-Iran, *Frontiers of Structural and Civil Engineering* (2020) 14(2), 446-472,
4. H. Yousefi, A. Taghavi, I. Mahmoudzadeh. Response of a spherical cavity in a fully-coupled thermo-poro-elastodynamic medium by cell-adaptive second-order central high resolution schemes, *Underground Space* (2018) 3(3) 206-217
5. H. Yousefi. Uncertainty quantification of asymmetric spherical cavities with lining in coupled saturated thermo-poro-elastic media via adaptive second-order central high resolution schemes, *International Journal of Hydromechatronics* (2019) 2(4), 111-132,
6. H. Yousefi, A. Noorzad, J. Farjoodi. Simulating 2D Waves Propagation in Elastic Solid Media Using Wavelet Based Adaptive Method, *Journal of Scientific Computing* (2010) 42, 404-425,
7. H. Yousefi, A. Noorzad, J. Farjoodi. Multiresolution Based Adaptive Schemes for Second order Hyperbolic PDEs in Elastodynamic problems, *Applied Mathematical Modeling*, (2013) 37(12-13), 7095-7127,
8. H. Yousefi, A. Noorzad, J. Farjoodi, M. Vahidi. Multiresolution-Based Adaptive Simulation of Wave Equation, *Applied Mathematics and Information Sciences*, (2012) 6(1), 47-58,
9. A. M. Boroujerdian, M. Saffarzadeh, H. Yousefi, H. Ghassemian. A Model to Identify High Crash Road Segments with the Dynamic Segmentation Method, *Accident Analysis and Prevention* (2014) 73, 274-287.