

UNIVERSIDAD DE CONCEPCION  
DIRECCION DE POSTGRADO  
CONCEPCION-CHILE



**Modelamiento y Esquemas Numéricos para Problemas de Flujo Cinemático  
Multi-Especies  
(Modelling and Numerical Schemes for Multi-Species Kinematic Flow Problems)**

*Tesis para optar al grado de Doctor  
en Ciencias Aplicadas con mención en Ingeniería Matemática*

**Luis Miguel Villada Osorio**

**FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS  
DEPARTAMENTO DE INGENIERIA MATEMATICA**

2013



# Modelling and Numerical Schemes for Multi-species Kinematic Flow Problems

**Luis Miguel Villada Osorio**

**Profesor guía:** Dr. Raimund Bürger, Universidad de Concepción, Chile.

**Co-tutor:** Dr. Pep Mulet, Universitat de València, España.

**Co-tutor:** Dr. Christophe Chalons, Université Paris Diderot-Paris 7 & L.J.L.L., France.

**Director de Programa:** Dr. Raimund Bürger, Universidad de Concepción, Chile.

## COMISION EVALUADORA

Dr. Stefan Diehl, Centre for Mathematical Sciences, Lund Universit, Sweden.

Dr. Enrique Fernandez-Nieto, Universidad de Sevilla, Spain.

Dr. Frédéric Lagoutière, Université Paris-Sud, France.

Dr. Carlos E. Mejía, Universidad Nacional de Colombia, Medellín, Colombia.

Firma: \_\_\_\_\_  
Dr. Raimund Bürger  
Universidad de Concepción, Chile.

Firma: \_\_\_\_\_  
Dr. Pep Mulet  
Universitat de València, España.

Firma: \_\_\_\_\_  
Dr. Cristóbal Castro  
Universidad de Tarapacá, Chile.

Firma: \_\_\_\_\_  
Dr. Mauricio Sepúlveda  
Universidad de Concepción, Chile.

**Fecha Examen de Grado:** \_\_\_\_\_

**Calificación:** \_\_\_\_\_



## AGRADECIMIENTOS

Comienzo expresando un especial agradecimiento a mis directores de tesis, al profesor Raimund Bürger por aceptar trabajar conmigo, a Pep Mulet por su hospitalidad en mi estadía en la Universitat de València. A ambos les agradezco infinitamente su paciencia, su buena disposición, su amistad, su apoyo y que a la vez han compartido conmigo sus conocimientos durante todo este proceso. A Christophe Chalons siempre dispuesto a colaborar y por su hospitalidad en el Laboratoire Jacques-Louis Lions en Paris. Los considero unos científicos dignos de imitar.

A los miembros del Departamento de Ingeniería Matemática, en particular, a los profesores que orientaron los cursos y seminarios en el doctorado.

A mis padres y hermanas, que entendieron mi decisión y desde el primer momento y en la distancia todos estos años conté con su apoyo y optimismo.

A todos compañeros y personal administrativo con los que compartí en el CI<sup>2</sup>MA, por su excelente camaradería durante todos estos años. A su director Gabriel Gatica por permitirme tener un espacio en tan importante centro de investigación.

El desarrollo de esta tesis y las estadias en el Departament de Matemàtica Aplicada de la Universitat de València y el Laboratoire Jacques-Louis Lions de la Université Pierre et Marie Curie, fueron financiadas gracias al proyecto MECESUP UCO0713.



*A mi hijo Daniel Augusto  
y a mi familia.*



# Contents

<b>Resumen</b>	<b>xi</b>
<b>Summary</b>	<b>xiii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Spectral WENO schemes with Adaptive Mesh Refinement for models of polydisperse sedimentation</b>	<b>9</b>
1.1 Introduction . . . . .	9
1.1.1 Scope . . . . .	9
1.1.2 Related work . . . . .	10
1.2 Preliminaries . . . . .	11
1.2.1 Sedimentation of polydisperse suspensions . . . . .	11
1.2.2 Secular equation and hyperbolicity analysis . . . . .	11
1.3 Numerical schemes . . . . .	12
1.3.1 SPEC-INT and COMP-GLF schemes . . . . .	12
1.3.2 Adaptive Mesh Refinement (AMR) . . . . .	15
1.4 Numerical results . . . . .	17
1.4.1 Example 1.1 ( $N = 4$ ) . . . . .	19
1.4.2 Example 1.2 ( $N = 7$ ) . . . . .	22
1.5 Conclusions of Chapter 1 . . . . .	27
<b>2 Regularized nonlinear solvers for Implicit-Explicit methods</b>	<b>29</b>
2.1 Introduction . . . . .	29
2.1.1 Scope . . . . .	29
2.1.2 Related work . . . . .	30
2.2 Diffusively corrected multi-species kinematic flow models . . . . .	31
2.2.1 Polydisperse sedimentation . . . . .	31
2.2.2 Hyperbolicity and parabolicity analysis for the polydisperse sedimentation model . . . . .	32
2.2.3 A new diffusively corrected MCLWR model (DCMCLWR model) . . . . .	33

2.2.4	Hyperbolicity and parabolicity analysis for the diffusively corrected MCLWR model . . . . .	34
2.3	Numerical schemes . . . . .	36
2.3.1	Spatial discretization . . . . .	36
2.3.2	Boundary conditions . . . . .	37
2.3.3	Explicit schemes . . . . .	37
2.3.4	Implicit-explicit schemes . . . . .	38
2.3.5	IMEX Runge-Kutta schemes . . . . .	39
	Algorithm 2.1 Implicit-explicit Runge-Kutta scheme . . . . .	39
2.4	Nonlinear solvers . . . . .	41
	Algorithm 2.2 Newton-Raphson method with line search strategy . . . . .	43
2.5	Numerical results . . . . .	43
2.5.1	Example 2.1: settling of a tridisperse suspension . . . . .	45
2.5.2	Examples 2.2 and 2.3: diffusively corrected kinematic traffic model . . . . .	48
2.6	Derivation of the model of polydisperse sedimentation with compression . . . . .	54
2.7	Conclusions of Chapter 2 . . . . .	57
<b>3</b>	<b>A DCMCLWR traffic model with anticipation lengths and reaction times</b>	<b>59</b>
3.1	Introduction . . . . .	59
3.1.1	Scope . . . . .	59
3.1.2	Related work . . . . .	61
3.2	Diffusively Corrected MCLWR model DCMCLWR . . . . .	62
3.2.1	The model equations . . . . .	62
3.2.2	Stability analysis . . . . .	62
3.2.3	DCLWR model with drivers having the same maximum speed . . . . .	64
3.3	Parabolicity analysis . . . . .	65
3.4	Numerical results . . . . .	68
3.4.1	Example 3.1 (DG model, $N = 4$ , stable behaviour). . . . .	69
3.4.2	Examples 3.2–3.5 (DG model, $N = 2$ , unstable behaviour) . . . . .	72
3.4.3	Examples 3.6–3.8 (DG model, $N = 2$ , mildly unstable behaviour) . . . . .	72
3.4.4	Examples 3.9–3.11 (GS model, $N = 4$ and $N = 2$ ) . . . . .	78
3.4.5	Examples 3.12 and 3.13 (DG model, $N = 5$ , drivers having the same maximum speed) . . . . .	79
3.4.6	Example 3.14 (Daganzo’s test, $N = 4$ ) . . . . .	83
3.5	Conclusions of Chapter 3 . . . . .	86
<b>4</b>	<b>Lagrangian-Remap schemes for the MCLWR traffic models</b>	<b>89</b>
4.1	Introduction . . . . .	89
4.1.1	Scope . . . . .	89
4.1.2	Related work . . . . .	91
4.2	Preliminaries . . . . .	92

4.2.1	Weak solutions and entropy admissibility . . . . .	92
4.2.2	Interlacing property of the MCLWR model . . . . .	93
4.2.3	Some simple difference schemes for the MCLWR model . . . . .	93
4.3	Discretization of the Lagrangian step . . . . .	94
4.4	Anti-diffusive schemes for the remap step . . . . .	96
4.4.1	Anti-diffusive schemes . . . . .	96
4.4.2	Choice of the numerical fluxes . . . . .	98
4.4.3	Lagrangian-anti-diffusive remap (L-AR) schemes, scalar case ( $N = 1$ ) . . . .	100
4.4.4	The multi-species case ( $N \geq 1$ ) and CFL condition . . . . .	102
4.5	Statistically conservative schemes . . . . .	103
4.5.1	Integral remap . . . . .	103
4.5.2	Random sampling remap, scalar case ( $N = 1$ ) . . . . .	104
4.5.3	Random sampling remap, multiclass case ( $N > 1$ ) . . . . .	104
4.5.4	Lagrangian-random sampling (L-RS) scheme . . . . .	106
4.6	Numerical results . . . . .	106
4.6.1	CFL condition, errors, and entropy test . . . . .	106
4.6.2	Example 4.1: $N = 1$ , linear velocity . . . . .	108
4.6.3	Example 4.2: $N = 1$ exponential velocity . . . . .	110
4.6.4	Example 4.3: $N = 5$ , linear velocity . . . . .	112
4.6.5	Examples 4.4 and 4.5: $N = 9$ , exponential velocity . . . . .	114
4.7	Conclusions of Chapter 4 . . . . .	121
<b>Discusión</b>		<b>123</b>
4.8	Conclusiones . . . . .	123
4.9	Trabajo Futuro . . . . .	125
<b>Bibliography</b>		<b>127</b>



# Resumen

En este trabajo de tesis se desarrollan esquemas numéricos para aproximar la solución de problemas de flujo cinemático multi-especies con (posiblemente) término difusivo fuertemente degenerado y se deduce y analiza un nuevo modelo de tráfico vehicular multi-especies con corrección difusiva. Específicamente se plantean esquemas numéricos para modelos multi-especies de tráfico vehicular y para problemas de sedimentación polidispersa. La tesis tiene los siguientes objetivos.

El primer objetivo de esta tesis es demostrar que un algoritmo de refinamiento de malla adaptativo (AMR) es eficiente respecto al tiempo de ejecución y memoria en la simulación de un modelo de sedimentación polidispersa. Se aplica esta técnica adaptativa a dos esquemas shock-capturing de alto orden.

El segundo objetivo es demostrar que esquemas Implícitos-Explícitos Runge-Kutta permiten obtener una eficiente solución numérica para los problemas de flujo cinemático multi-especies con término difusivo fuertemente degenerado. Estos esquemas consisten en una combinación de esquemas explícitos Runge-Kutta para el término convectivo con un tratamiento implícito para el término difusivo. El sistema no lineal que resulta de la discretización implícita se resuelve mediante una novedosa técnica que consiste en regularizar los coeficientes de difusión y el uso del método de Newton-Raphson con algunas técnicas adecuadas. Finalmente se obtiene un esquema con una condición CFL menos restrictiva que para un esquema con un tratamiento explícito para el término difusivo.

El tercer objetivo de esta tesis es proponer un modelo multi-clase Lighthill-Whitham-Richards de tráfico vehicular con corrección difusiva que considera tiempos reacción y distancias de anticipación. El modelo puede ser parabólico bajo una cierta relación entre las velocidades máximas y los tiempos de reacción para cada especie. Se analiza la estabilidad del modelo, basado en la información característica de la matriz de difusión. Se muestra que la solución del problema con corrección difusiva puede desarrollar inestabilidades debido a altos tiempos de reacción o pequeñas distancias de anticipación, tales inestabilidades pueden estar controladas debido a la naturaleza no lineal del problema.

Finalmente, se proponen una nueva clase de esquemas numéricos de dos pasos para la solución de problemas de tráfico vehicular multi-clase que combina en un primer paso la solución de ecuaciones

en coordenadas Lagrangianas, y el segundo paso consistente en resolver una ecuación de transporte. Se busca obtener un esquema numérico anti-difusivo convergente. Para el segundo paso se consideran dos diferentes estrategias, una basada en recientes y novedosos esquemas anti-difusivos los cuales son sencillos de implementar y con la propiedad T.V.D., tales esquemas se denominaran “L-AR”. La segunda estrategia se basa en el clásico método aleatorio de Glimm, tales esquemas se denominaron “L-RS”. Para el caso  $N = 1$  se demuestra que los esquemas L-AR son conservativos, tienen la propiedad T.V.D. y satisfacen el principio del máximo, con lo cual se muestra que convergen a una solución débil de la ley de conservación. Ambos esquemas L-AR y L-RS se generalizaron para el caso  $N > 1$ . Se muestra mediante resultados numéricos que estos esquemas resultan competitivos con respecto a una variedad de otros esquemas, aparte de que no requieren del uso de la información característica del flujo.

# Summary

In this dissertation, numerical schemes for the approximate solution of problems of multi-species kinematic flow models with (possibly) strongly degenerate diffusion terms are introduced and an analysis of a new diffusively corrected multi-species traffic flow model is presented. In particular, new numerical schemes for multi-species traffic flow models and polydisperse sedimentation problems are proposed. The thesis has the following aims.

The first aim of this thesis is to demonstrate that an adaptive mesh refinement algorithm saves computational resources in simulations of polydisperse sedimentation model. The implementation of this adaptive technique is applied to two state-of-the-art high resolution shock capturing techniques.

The second goal of this thesis is to demonstrate that implicit-explicit Runge-Kutta schemes efficiently generate a numerical solution of multi-species kinematic flow models with strongly degenerate diffusive term. These schemes consist in combining an explicit Runge-Kutta scheme for the time integration of the convective part with an implicit one for the diffusive part. To solve the highly nonlinear and non-smooth system that arises in the implicit discretization, it is proposed to regularize the diffusion coefficients and to apply the Newton-Raphson method with suitable globalization techniques. The CFL condition for the numerical scheme obtained is less severe than for an explicit treatment of the diffusive term.

The third goal of this thesis is to propose a diffusively corrected multi-class Lighthill-Whitham-Richards traffic model with anticipation lengths and reaction times. We analyse the stability of this diffusively corrected model under varying reaction times and anticipation lengths. It is demonstrated that instabilities may develop for high reaction times and short anticipation lengths, and that these instabilities may have controlled frequencies and amplitudes due to their nonlinear nature.

Finally, we propose to introduce a new class of two-step numerical schemes for the multi-class Lighthill-Whitham-Richards traffic model. The new class of schemes combines in the first step the solutions of equations in Lagrangian coordinates and in a second step, a transport equation is solved to remap the solution to the original coordinates. The new schemes are referred to as “Lagrangian-remap” (LR) schemes. In the second step, two different strategies are considered. One strategy for LR schemes incorporates recent anti-diffusive techniques for transport equations.

The corresponding subclass of LR schemes are named “Lagrangian-anti-diffusive-remap” (L-AR) schemes. The second strategy consists in handling the remap step by Glimm-like random sampling, which gives rise to a statistically conservative “Lagrangian-random sampling” (L-RS) scheme. The LR schemes for the MCLWR model are supported by a partial analysis of the L-AR schemes for  $N = 1$ , which are total variation diminishing (TVD) under a suitable CFL condition and therefore produce numerical solutions that converge to a weak solution, and by numerical examples for both L-AR and L-RS subclasses of schemes.

# Introduction

## Multi-species kinematic flow models

Multi-species kinematic flow models arise in the mathematical description of a wide variety of phenomena that involve the flow of one disperse substance through a continuous phase, and where the disperse substance consists of particles belonging to a number  $N$  of species that can be distinguished by some characteristic property. To mention a few examples, problems arising in different contexts: wastewater treatment, mineral processing, chemical engineering, volcanology, traffic flows, petrology and medicine.

The term “kinematic” means that the velocity  $v_i$  of species  $i \in \{1, \dots, N\}$  is an explicit function of the vector  $\Phi := (\phi_1, \dots, \phi_N)^T \in \mathbb{R}^N$  of the unknown concentrations (volume fractions or densities)  $\phi_i$  of each species. Thus, standard multi-species kinematic flow models are given by systems of  $N$  scalar, in general nonlinear, first-order conservation laws

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = \mathbf{0}, \quad \mathbf{f}(\Phi) = (f_1(\Phi), \dots, f_N(\Phi))^T = (\phi_1 v_1(\Phi), \dots, \phi_N v_N(\Phi))^T, \quad (1)$$

where  $t$  is time and  $x$  is the spatial coordinate. This work focuses particularly on two applications of multi-species kinematic flow models. As we will see throughout this thesis, our study will be highly motivated by these applications.

*Polydisperse suspensions* consist of small solid particles dispersed in a viscous fluid, where the particles are assumed to belong to a number  $N$  of species where species  $i$  is associated with the volume fraction  $\phi_i$ , the phase velocity  $v_i$  that is assumed to be given as an explicit function of  $\Phi$ , and a size (diameter)  $d_i$ , where  $d_1 > d_2 > \dots > d_N$ . The continuity equations for all species lead to a system of conservation laws in the form (1). This model was studied intensively in recent years [12, 22, 23, 52, 54, 77, 79, 94, 115, 116] in where a number of algebraic expressions for  $\mathbf{f}(\Phi)$  have been proposed in the literature. We will consider the Masliyah-Lockett-Bassoon (MLB) model [77, 79].

*Multi-class traffic flow model.* The well-known Lighthill-Whitham-Richards (LWR) kinematic traffic model [75, 92] considers the density of cars  $\phi = \rho/\rho_{\max}$ , where  $\rho$  is the local number of cars per mile and  $\rho_{\max}$  is some maximum bumper-to-bumper density and the local velocity  $v = v(x, t)$  as function of the local density,  $v = v(\phi(x, t))$ , where  $t$  is time,  $x$  is the spatial coordinate along either an unbounded, one-directional highway or a closed circuit. Usually it is assumed that  $v(\phi) = v^{\max} V(\phi)$ , where  $v^{\max}$  is the preferential velocity of drivers on a free highway and  $V$  is a hindrance function describing the drivers’ behaviour of reducing speed in presence of other cars.

The function  $V$  satisfies the assumption

$$V(0) = 1, \quad V(\phi) \geq 0, \quad V'(\phi) \leq 0, \quad 0 \leq \phi \leq \phi_{\max}. \quad (2)$$

These assumptions lead to the one-dimensional scalar conservation law

$$\partial_t \phi + \partial_x f(\phi) = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (3)$$

where the flux density function  $f$  is given by

$$f(\phi) = \phi v(\phi) = v^{\max} \phi V(\phi). \quad (4)$$

Benzoni-Gavage and Colombo [8] and Wong and Wong [113] introduced the multi-class (MCLWR) traffic flow model where we distinguish  $N$  classes of drivers associated with preferential velocities  $v_1^{\max} > v_2^{\max} > \dots > v_N^{\max}$  and assume that the sought quantity is the vector  $\Phi := (\phi_1, \dots, \phi_N)^T$  of the densities  $\phi_i$  of the cars of the different driver classes. The local velocity  $v_i$  of vehicles of driver class  $i$  is given by  $v_i = v_i(\phi) = v_i^{\max} V(\phi)$  for  $i = 1, \dots, N$ , where  $V$  is a function of the total local density  $\phi := \phi_1 + \dots + \phi_N$ . Thus, the MCLWR model is given by a strongly coupled system of nonlinear first-order conservation laws in the form (1), where the components of the flux vector  $\mathbf{f}(\Phi)$  are given by

$$f_i(\Phi) = \phi_i v_i(\phi) = \phi_i v_i^{\max} V(\phi), \quad i = 1, \dots, N. \quad (5)$$

Solutions of (1) include kinematic shocks separating areas of different composition. The accurate numerical approximation of these solutions is a challenge since closed-form eigenvalues and eigenvectors of the flux Jacobian  $\mathcal{J}_{\mathbf{f}}(\Phi)$  are usually not available, and the characteristic fields are neither genuinely nonlinear nor linearly degenerate. For some polydisperse models [64, 77, 79] and for the MCLWR model [50], the eigenstructure can be analyzed through a convenient hyperbolicity criterion that has become known as the ‘‘secular equation’’ [3, 22]. Based on this idea, Bürger et al. [23] and Donat and Mulet [50] proposed a shock-capturing high order or spectral weighted essentially non-oscillatory (WENO) scheme for the numerical solution of these models calculating in a numerical form the eigenstructure of  $\mathcal{J}_{\mathbf{f}}(\Phi)$ . The computational cost of these schemes for obtaining simulations with fine resolution and near steady states can be quite high, in part due to the costly operations involving the computation of the eigenstructure and the nonlinear reconstructions.

It is one of the purposes in this thesis to discuss two different techniques for solving numerically (1) that perform better in terms of resolution accuracy and efficiency, using characteristic information in **Chapter 1** and a new class of schemes that do not rely on characteristic information in **Chapter 4**.

In **Chapter 1**, for models of polydisperse sedimentation, an adaptive technique, namely the Adaptive Mesh Refinement (AMR) algorithm [11], is applied to two different shock capturing WENO schemes proposed in [23], one implemented in a component-wise fashion combined with global Lax-Friedrichs flux vector splitting (denoted ‘‘COMP-GLF’’) and another one applied in

a characteristic-wise (spectral) fashion which uses the eigenstructure of  $\mathcal{J}_f(\Phi)$  (denoted ‘‘SPEC-INT’’). The two schemes equipped with AMR allow us to obtain speed-up of CPU time and saves in memory requirement when are compared with the corresponding non-adaptive version. The AMR technique consists in using a fine mesh near the shocks, heads and tails of rarefactions, while employing a coarse mesh near smooth regions.

The contents of **Chapter 1** corresponds to the article [32]:

- R. Bürger, P. Mulet, and L.M. Villada. Spectral WENO schemes with Adaptive Mesh Refinement for models of polydisperse sedimentation. *ZAMM Z. Angew. Math. Mech.*, **93** (2013), pp. 373–386.

## Diffusively Corrected Multi-species kinematic flow models

In some applications, the velocities also depend on the spatial variation of  $\Phi$  to account for additional effects such as sediment compressibility in sedimentation or drivers’ reaction time and anticipation length in traffic flow. In the sedimentation model, this property reflects that the solid particles possibly form a compressible sediment layer, while in the traffic flow models, it reflects that each driver adjusts his speed to the concentration slightly ahead and that this reaction is subject to a small delay.

For traffic flow models in the case  $N = 1$ , Nelson [81, 82] showed that introducing an anticipation length  $L$  and a reaction time  $\tau$ , replacing  $V(\phi(x, t))$  by  $V(\phi(x + L - v^{\max}V\tau, t - \tau))$  and neglecting  $\mathcal{O}(L^2 + \tau^2)$  terms when expanding the latter expression around  $(x, t)$ , one obtains a ‘‘diffusively corrected’’ version of (3), (4) of the form

$$\partial_t \phi + \partial_x f(\phi) = A(\phi)_{xx}, \quad (6)$$

where the function  $A$  is Lipschitz continuous and increasing so that the governing equation (6) of the diffusively corrected LWR model (‘‘DCLWR model’’) is a strongly degenerate parabolic PDE in the sense that  $A(\phi) = 0$  for  $\phi \leq \phi_c$ , where  $\phi_c$  is a critical density value (e.g., a perception threshold), and  $A'(\phi) > 0$  for  $\phi > \phi_c$ .

In **Chapter 2**, first we introduce a new model, called diffusively corrected MCLWR model (‘‘DCMCLWR model’’), by combining the assumptions of the DCLWR model with those of the MCLWR model. In particular, we associate class  $i$  of drivers with the triple  $(v_i^{\max}, L_i, \tau_i)$ ,  $i = 1, \dots, N$ , which means that drivers of different classes may have different preferential velocities, anticipation lengths, and reaction times. The resulting model can be written as a system of partial differential equations with an extra, possibly strongly degenerate diffusive term of the type

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = \partial_x (\mathbf{B}(\Phi) \partial_x \Phi), \quad (7)$$

where  $\mathbf{B}(\Phi)$  is a given  $N \times N$  matrix function expressing the diffusive correction. On the other hand, for models of polydisperse sedimentation, diffusive terms leading to the form (7) were first

proposed by Stamatakis and Tien [106], and a theory of sedimentation of polydisperse suspensions forming compressible sediments was advanced in [12].

Although the available mathematical theory does not allow us to be conclusive about the existence, uniqueness and well-posedness of the solutions of such strongly degenerate hyperbolic-parabolic systems, it is plausible to perform simulations with appropriate numerical methods.

With respect to the use of efficient numerical schemes it is important to remark that explicit numerical schemes to solve (7) are widely used in many applications for example the explicit high-resolution central difference scheme by Kurganov-Tadmor (KT) was used in [12], however, the diffusive term imposes a drastic time step size restriction to comply with the stability condition [71, Sect. 4.2]. On the other hand, an implicit treatment in the parabolic term in (7) requires the solution of highly nonlinear and non-smooth systems of algebraic equations and the efficient solution of these systems with nonlinear solver requires some degree of smoothness in the coefficients of the matrix function  $\mathbf{B}$ .

It is the purpose of **Chapter 2** to apply implicit-explicit (IMEX) Runge-Kutta schemes to solve the strongly degenerate parabolic partial differential equations (7). These schemes consist in combining a Runge-Kutta scheme with an implicit discretization, denoted  $D(\Phi)$ , of the diffusive term with an explicit one, denoted  $C(\Phi)$ , for the convective term. With this idea, the problem (7) is expressed in the form

$$\partial_t \Phi = C(\Phi) + D(\Phi). \quad (8)$$

The stability restriction on the time step  $\Delta t$  that IMEX Runge-Kutta schemes impose when applied to (8) is expressed as

$$\frac{\Delta t}{\Delta x} \max_{\Phi} \rho(\mathcal{J}_{\mathbf{f}}(\Phi)) \leq C_{\text{eff}} \leq 1,$$

where  $\Delta x$  is the size of the space discretization and  $\rho(\cdot)$  is the spectral radius. This  $\Delta t$  restriction depends on the convective term alone which is discretized by a spectral WENO scheme like SPEC-INT (see [23]). This IMEX strategy has been used in other contexts, such as convection-diffusion problems, convection problems with stiff reaction terms [4, 49], and stiff terms in hyperbolic systems with relaxation [17, 18, 19, 20, 89].

In the context of strongly degenerate parabolic partial differential equations, we propose to use different types of IMEX Runge-Kutta proposed in the literature [4, 17, 18, 68, 89, 122] with an adequate regularization of the diffusive coefficients and apply a suitable technique based on Newton-Raphson's method to solve the nonlinear systems in an efficient way. The proposed regularization is applied to the non-smooth diffusion coefficient, in a way that does not change the strong degeneracy of the diffusion term. The final schemes are much more efficient, in term of error reduction versus CPU time, than the explicit schemes.

The contents of **Chapter 2** corresponds to the article [31]:

- R. Bürger, P. Mulet, and L.M. Villada. Regularized nonlinear solvers for IMEX methods applied to diffusively corrected multi-species kinematic flow models. *SIAM J. Sci. Comput.*, **35** (2013), pp. B751–B777.

## Analysis of DCMCLWR model

In **Chapter 3**, an analysis of the conditions on the non-negative parameters  $v_i^{\max}$ ,  $\tau_i$  and  $L_i$  under which  $\mathbf{B}(\Phi)$  has eigenvalues with positive real part for all  $\Phi \in \mathcal{D}_{\phi_{\max}} \{\Phi \in \mathbb{R}^N \mid \phi_1 > 0, \dots, \phi_N > 0, \phi < 1\}$  is performed. This generates the following restriction:

$$L_k(\phi) + \tau_k v_k^{\max} \phi V'(\phi) \leq 0 \quad \text{for } 0 \leq \phi \leq \phi_{\max}, k = 1, \dots, N. \quad (9)$$

If (9) is violated, then the model is likely to exhibit anti-diffusive phenomena such as formation of clusters, steep density gradients, stop-and-go waves, and other instability phenomena. A similar conclusion (though based on a slightly different model) has been drawn, for example, in [87]. To analyse the eigenstructure of the matrix  $\mathbf{B}(\Phi)$  it is not sufficient to establish a stability criterion for the model (7). A linearized stability analysis consists in analysing the eigenvalues of the matrix

$$\mathbf{M}(\Phi, \xi) := \frac{i}{\xi} \mathcal{J}_f(\Phi) + \mathbf{B}(\Phi) \in \mathbb{C}^{N \times N}, \quad \xi \in \mathbb{R}_+,$$

where  $i = \sqrt{-1}$  and  $\xi$  is a frequency parameter.

Numerical experiments showed that stable behavior is obtained when the eigenvalues of  $\mathbf{M}$  have positive real parts and that instabilities may be triggered otherwise, although the nonlinear character of the equations stabilizes some initial traffic configurations that would explode under the linearized equations. The nonlinearities also help to control the amplitude (and in some cases the frequency) of instabilities in the simulations. While we associate oscillations in the numerical solution with unstable behaviour in general, we distinguish between situations where there is a blow-up of frequency, which means that violations of the stability condition lead to strongly oscillating solutions (akin to those studied in [13]), and situations of mildly unstable behaviour with finite frequencies of oscillation, and where numerical solutions can be interpreted as the formation of stop-and-go waves (although the latter phenomenon is usually associated with much larger amplitudes, cf., e.g., [86, 87]).

In a series of numerical examples we compare the performance of the IMEX-Runge-Kutta schemes introduced in **Chapter 2** with KT scheme and one scheme introduced in [24] adapted to the context of DCMCLWR models. This allows to show the efficiency of IMEX-Runge-Kutta schemes, and also allows to conclude that some oscillations are not numerical artefacts.

The contents of **Chapter 3** corresponds to the accepted article [30]:

- R. Bürger, P. Mulet, and L.M. Villada. A diffusively corrected multiclass Lighthill-Whitham-Richards traffic model with anticipation lengths and reaction times. *Adv. Appl. Math. Mech.*, to appear.

## Lagrangian-Remap schemes for MCLWR traffic model

In **Chapter 4** we propose to introduce a new class of schemes for multi-species traffic flow model (1). These new schemes are based on a Lagrangian-Projection decomposition for each density, i.e. the approach is to combine the solution of the governing equation in a Lagrangian reference frame with an algorithm to remap the original mesh. Here, we propose an idea of numerical strategy easy to implement that is independent of  $N$  (number of species), does not use the knowledge of the eigenstructure of the model and is expected to be anti-diffusive. These schemes are referred to as “Lagrangian-remap” (LR) schemes.

For this thesis, we introduce these LR schemes only for MCLWR models (1), (5) (we consider here the local number of cars per mile  $\rho$  instead of the density  $\phi = \rho/\rho_{\max}$ ). In the case  $N = 1$  the idea is to solve (3), (4) by exploiting the assumption on the hindrance function  $V(\rho)$  described in (2). The two steps are described below: we first apply a Lagrangian method [56] (Lagrangian step) to solve

$$\partial_t \rho + \rho \partial_x v(\rho) = 0, \quad (10)$$

and use this solution, evolved over the time interval of length  $\Delta t$ , as the initial condition for solving in a second step the transport equation

$$\partial_t \rho + v(\rho) \partial_x \rho = 0, \quad (11)$$

whose solution, again evolved over a time interval of length  $\Delta t$ , provides the sought approximate solution valid for  $t + \Delta t$  (remap step).

The idea behind the introduction of LR schemes is to solve (11) using anti-diffusive techniques that have been developed recently for transport equations and thereby to increase the overall efficiency of the proposed splitting strategy, while keeping its simplicity. More precisely, the remap step can be handled in two different ways. One alternative is to employ an anti-diffusive but stable numerical scheme for the transport equation (11), where care is taken to design the scheme for the remap step in such a way that the resulting scheme (first step followed by a second step) is conservative with respect to (3). Anti-diffusive techniques are based on the pioneering work of Després and Lagoutière [46], Bouchut [15], Bokanowski and Zidani [14]. This subclass of LR schemes will be addressed as “Lagrangian-anti-diffusive remap” (L-AR) schemes. Alternatively, the remap step can be handled by random sampling in a Glimm-like approach [55]. The resulting scheme, denoted here as “Lagrangian-random-sampling” (L-RS) scheme, is only statistically conservative.

We discuss a partial analysis of the L-AR scheme with the conclusion that under the CFL condition

$$-1 \leq \lambda \rho_{\max} v'(\rho) \leq 0 \quad \text{for } 0 \leq \rho \leq \rho_{\max}. \quad (12)$$

$$\lambda v(\rho) \leq 1 \quad \text{for } 0 \leq \rho \leq \rho_{\max}, \quad (13)$$

where  $\lambda = \Delta t/\Delta x$ , the L-AR schemes have the total variation diminishing (TVD) property and satisfy the maximum principle and therefore converge to a weak solution of (3). We remark that CFL condition (12) is obtained in the analysis of stability for the Lagrangian step, while (13) is the CFL condition for the transport step (11).

Both L-AR and L-RS subclasses of LR schemes can readily be extended to the multiple-species case ( $N > 1$ ). For that case, we propose to equip the L-RS scheme with random sampling among the fan of states of the simple Harten, Lax and van Leer (HLL) approximate Riemann solver [61, 111]. Numerical experiments show that the proposed schemes are competitive with respect to recent schemes introduced in [24].

The contents of **Chapter** 4 corresponds to the following preprint submitted [29]:

- R. Bürger, C. Chalons, and L.M. Villada. *Anti-diffusive and random-sampling Lagrangian-remap schemes for the multiclass Lighthill-Whitham-Richards traffic model*. Preprint 2013-14, Centro de Investigación en Ingeniería Matemática, Universidad de Concepción; (submitted).



# Chapter 1

## Spectral WENO schemes with Adaptive Mesh Refinement for models of polydisperse sedimentation

### 1.1 Introduction

#### 1.1.1 Scope

The sedimentation of a polydisperse suspension of small rigid equal-density spheres that belong to a finite number  $N$  of species differing in size can be described by a spatially one-dimensional system of first-order nonlinear conservation laws

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = 0, \quad \Phi = (\phi_1, \dots, \phi_N)^T, \quad \mathbf{f}(\Phi) = (f_1(\Phi), \dots, f_N(\Phi))^T; \quad 0 < x < K, \quad t > 0. \quad (1.1)$$

The unknowns are the volume fractions (concentrations)  $\phi_i$  of species  $i$ ,  $i = 1, \dots, N$ , as functions of depth  $x$  and time  $t$ . The flux density functions  $f_i$  are given by  $f_i(\Phi) = \phi_i v_i(\Phi)$ , where  $v_i = v_i(\Phi)$  is the phase velocity of particle species  $i$  that is assumed to be given as an explicit function of  $\Phi$ . The model (1.1) and its variants are widely used in applications including wastewater treatment, mineral processing, chemical engineering and volcanology. Moreover, a very similar model describes multi-class traffic flow. See [12, 23, 50] for references.

Typical solutions of (1.1), for instance for batch settling of an initially homogeneous suspension in a column, include moving and stationary discontinuities (kinematic shocks) separating areas of different composition. The accurate numerical approximation of these solutions is a challenge since closed-form eigenvalues and eigenvectors of the flux Jacobian  $\mathcal{J}_{\mathbf{f}}(\Phi)$  are usually not available, and the characteristic fields are neither genuinely nonlinear nor linearly degenerate. Some of these sedimentation models, including the widely used models by Masliyah, Lockett and Bassoon (MLB model) [77, 79] and Höfler and Schwarzer (HS model) [64], give rise to flux Jacobians whose eigenstructure can be analyzed through a convenient hyperbolicity criterion that has become known as the “secular equation” [3, 22]. When this approach applies, hyperbolicity can be ensured under easily verifiable conditions and the eigenstructure of the Jacobian can be computed numerically,

so that efficient shock capturing schemes may be applied for the numerical computation of the solutions of the models.

The computational cost of these schemes for obtaining simulations with fine resolution and near steady states can be quite high, in part due to the costly operations involving the computation of the eigenstructure and the nonlinear reconstructions. Adaptive techniques, as the Adaptive Mesh Refinement (AMR) algorithm [11], aim to reduce the computational cost of these schemes, by using a higher resolution near salient flow features (shocks, heads and tails of rarefactions, etc.), while employing a coarse mesh near smooth regions of the flow.

In the present chapter, we apply the AMR technique to two different WENO schemes introduced in [23], namely to a WENO scheme implemented in a component-wise fashion combined with global Lax-Friedrichs flux vector splitting (denoted by “COMP-GLF”), and alternatively, to a WENO scheme applied in a characteristic-wise (spectral) fashion, and which makes essential use of the interlacing property of the velocities  $v_1, \dots, v_N$  with the eigenvalues of  $\mathcal{J}_f(\Phi)$  (see Section 1.2.1). The second version is denoted by “SPEC-INT”. The scheme COMP-GLF does not rely on characteristic information, is much easier to implement than SPEC-INT, and on a fixed uniform grid is several times faster than SPEC-INT. However, SPEC-INT is substantially more accurate than COMP-GLF, and turns out to be even more efficient than COMP-GLF in terms of reduction of numerical error per CPU time [23]. It turns out that equipping both versions with AMR produces substantial gains in computational efficiency when compared with the corresponding non-adaptive version, and that the adaptive versions based on SPEC-INT are consistently more efficient than those relying on COMP-GLF.

### 1.1.2 Related work

Any kind of adaptativity that permits to restrict the use of a high-resolution scheme on a fine grid to a portion of the computational domain will produce a benefit in terms of computational efficiency. Common methods are multiresolution algorithms [40, 41], moving mesh methods [107] and approximations on unstructured meshes [65]. Adaptive Mesh Refinement (AMR) is a grid adaptation technique, introduced by Berger and Olinger [11] for hyperbolic conservation laws, which is based not so much on the reduction of the number of cells on the grid as on the reduction of the overall number of applications of the integration algorithm. This algorithm is very time-consuming especially for high-resolution shock capturing schemes. The AMR algorithm is a two-fold adaptive method. The goal of allowing arbitrary grid resolution is attained by the definition of a set of overlapping grids of different resolutions –a grid hierarchy– being the grid at each resolution level defined only on the part of the domain that is foreseen to require such a resolution. The way in which the grids are overlapped allows to refine also in time, in the sense that each grid is integrated with temporal steps adapted to its spatial grid size. This time refinement is another key feature for improving the overall performance of the algorithm [10, 11].

## 1.2 Preliminaries

### 1.2.1 Sedimentation of polydisperse suspensions

The MLB model arises from the continuity and linear momentum balance equations for the solid species and the fluid through suitable constitutive assumptions and simplifications. We refer to [12] for details and introduce the model here in its final form. For particles that have the same density, the velocities  $v_1, \dots, v_N$  are given by

$$v_i(\Phi) := \frac{(\varrho_s - \varrho_f)gd_1^2}{18\mu_f}(1 - \phi)V(\phi)(\delta_i - \boldsymbol{\delta}^T\Phi), \quad i = 1, \dots, N, \quad (1.2)$$

where  $d_1 > d_2 > \dots > d_N$  are the respective species diameters,  $\delta_i := d_i^2/d_1^2$ ,  $\boldsymbol{\delta} := (\delta_1 = 1, \delta_2, \dots, \delta_N)^T$ ,  $\varrho_s$  and  $\varrho_f$  are the solid and fluid densities,  $g$  is the acceleration of gravity,  $\mu_f$  is the fluid viscosity,  $\phi := \phi_1 + \dots + \phi_N$  is the total solids volume fraction, and  $V(\phi)$  is a hindered settling factor that is assumed to satisfy  $V(0) = 1$ ,  $V(\phi_{\max}) = 0$  and  $V'(\phi) \leq 0$  for  $\phi \in [0, \phi_{\max}]$ , where the constant  $\phi_{\max}$  denotes the maximum total solids concentration. A standard choice in [94] is

$$V(\phi) = (1 - \phi)^{n_{\text{RZ}}-2} \quad \text{if } \Phi \in \mathcal{D}_{\phi_{\max}}, n_{\text{RZ}} > 2; \quad V(\phi) = 0 \quad \text{otherwise}, \quad (1.3)$$

where  $\mathcal{D}_{\phi_{\max}} := \{\Phi \in \mathbb{R}^n \mid \phi_1 \geq 0, \dots, \phi_N \geq 0, \phi \leq \phi_{\max}\}$  is the set of physically relevant concentration vectors and  $n_{\text{RZ}}$  is the material-dependent Richardson-Zaki exponent [93].

The components  $f_1(\Phi), \dots, f_N(\Phi)$  of the flux vector  $\mathbf{f}(\Phi)$  of the MLB model are given by

$$f_i(\Phi) := v_i(\mathbf{0})\phi_i(1 - \phi)V(\phi)(\delta_i - \boldsymbol{\delta}^T\Phi), \quad i = 1, \dots, N. \quad (1.4)$$

### 1.2.2 Secular equation and hyperbolicity analysis

For general kinematic models with  $v_i = v_i(\phi_1, \dots, \phi_N)$ , the Jacobian  $\mathcal{J}_{\mathbf{f}}(\Phi)$  has no definite structure, hence its spectral information cannot be readily obtained. However, when  $v_1, \dots, v_N$  do not depend on each of the  $N$  components of  $\Phi$  in an individual way, but are functions of a small number  $m \ll N$  of scalar functions of  $\Phi$  (as is the case of the MLB model), i.e.,  $v_i = v_i(p_1, \dots, p_m)$  and  $p_l = p_l(\Phi)$  for  $i = 1, \dots, N$  and  $l = 1, \dots, m$ , then the entries of  $\mathcal{J}_{\mathbf{f}}(\Phi)$  are given by

$$f_{ij} = \frac{\partial(\phi_i v_i)}{\partial \phi_j} = v_i \delta_{ij} + \sum_{l=1}^m \phi_i \frac{\partial v_i}{\partial p_l} \frac{\partial p_l}{\partial \phi_j}, \quad i, j = 1, \dots, N,$$

i.e.,  $\mathcal{J}_{\mathbf{f}}(\Phi)$  is a rank- $m$  perturbation of the diagonal matrix  $\mathbf{D} := \text{diag}(v_1, \dots, v_N)$  of the form  $\mathcal{J}_{\mathbf{f}} = \mathbf{D} + \mathbf{B}\mathbf{A}^T$ , where

$$\mathbf{B} := (B_{il}) = (\phi_i \partial v_i / \partial p_l), \quad \mathbf{A} := (A_{jl}) = (\partial p_l / \partial \phi_j), \quad 1 \leq i, j \leq N, \quad 1 \leq l \leq m. \quad (1.5)$$

The hyperbolicity analysis is then based on the following theorem, which can be found in [3], but we give here the form in [51], which provides the explicit formulas to be used in the applications.

**Theorem 1.1 (The secular equation, [3, 51])** *Assume that  $v_i > v_j$  for  $i < j$ , and that  $\mathbf{A}$  and  $\mathbf{B}$  have the formats specified in (1.5). We denote by  $S_r^p$  the set of all (ordered) subsets of  $r$  elements taken from a set of  $p$  elements. If  $\mathbf{X}$  is an  $m \times N$  matrix,  $I := \{i_1 < \dots < i_k\} \in S_k^N$  and  $J := \{j_1 < \dots < j_l\} \in S_l^m$ , then we denote by  $\mathbf{X}^{I,J}$  the  $k \times l$  submatrix of  $\mathbf{X}$  given by  $(\mathbf{X}^{I,J})_{p,q} = X_{i_p, j_q}$ . Let  $\lambda \neq v_i$  for  $i = 1, \dots, N$ . Then  $\lambda$  is an eigenvalue of  $\mathbf{D} + \mathbf{B}\mathbf{A}^\top$  if and only if*

$$R(\lambda) := \det(\mathbf{I} + \mathbf{A}^\top(\mathbf{D} - \lambda\mathbf{I})^{-1}\mathbf{B}) = 1 + \sum_{i=1}^N \frac{\gamma_i}{v_i - \lambda} = 0. \quad (1.6)$$

The relation  $R(\lambda) = 0$  is known as the secular equation [3]. The coefficients  $\gamma_i$  are given by the following expression:

$$\gamma_i = \sum_{r=1}^{\min\{N,m\}} \sum_{i \in I \in S_r^N, J \in S_r^m} \frac{\det \mathbf{A}^{I,J} \det \mathbf{B}^{I,J}}{\prod_{l \in I, l \neq i} (v_l - v_i)}, \quad i = 1, \dots, N. \quad (1.7)$$

When  $m \leq 2$ , the quantities (1.7) can be easily computed and the hyperbolicity analysis via the secular equation is much less involved than discussing the zeros of  $\det(\mathcal{J}_{\mathbf{f}}(\Phi) - \lambda\mathbf{I})$ , as was done in [12, 94]. For  $m = 3$  or  $m = 4$ , the computations are more involved [22], but have turned out very useful in providing at least partial results concerning hyperbolicity, where the theoretical analysis of  $\det(\mathcal{J}_{\mathbf{f}}(\Phi) - \lambda\mathbf{I})$  is essentially out of reach. For the special case of the MLB model with equal-density spheres,  $v_i$  depends on the parameters  $p_1 := \phi$  and  $p_2 := \delta^\top \Phi$ . Therefore, we are in the case  $m = 2$  and we can compute explicitly the coefficients  $\gamma_i = -v_1(\mathbf{0})(n-1)(1-\phi)^{n-2}\phi_i\delta_i > 0$  if  $\phi_i > 0$  and  $\phi < 1$ .

The following Theorem illustrates the importance of the secular equation. Its proof (see [22]) follows from Theorem 1.1 by a discussion of the poles of  $R(\lambda)$  and its asymptotic behavior as  $\lambda \rightarrow \pm\infty$ .

**Theorem 1.2 (Interlacing property, [22])** *With the notation of Theorem 1.1, assume that  $\gamma_i \cdot \gamma_j > 0$  for  $i, j = 1, \dots, N$ . Then  $\mathbf{D} + \mathbf{B}\mathbf{A}^\top$  is diagonalizable with real eigenvalues  $\lambda_1, \dots, \lambda_N$  which are the roots of the secular equation (1.6). If  $\gamma_1, \dots, \gamma_N > 0$ , the following so-called interlacing property holds:*

$$v_N < \lambda_N < v_{N-1} < \lambda_{N-1} < \dots < v_1 < \lambda_1 < M_2 := v_1 + \gamma_1 + \dots + \gamma_N. \quad (1.8)$$

As a consequence, we see that the model (1.1) with the flux vector  $\mathbf{f}(\Phi)$  of the MLB model given by (1.4) is strictly hyperbolic in  $\mathcal{D}_{\phi_{\max}}$  if  $\phi_1 > 0, \dots, \phi_N > 0$  and  $\phi < \phi_{\max} < 1$ .

## 1.3 Numerical schemes

### 1.3.1 SPEC-INT and COMP-GLF schemes

It is well known that nonlinear hyperbolic systems of conservation laws of the type (1.1) can develop discontinuities (shocks), even for smooth initial data. By Lax-Wendroff's theorem [73],

conservative schemes can cope with this situation since their limits are weak solutions of the conservation law. For grid points  $x_j := j\Delta x$  for  $j = 1, \dots, M$ , where  $\Delta x := \mathcal{L}/M$ , and  $t_n := n\Delta t$  for  $n \in \mathbb{N}_0$ , a conservative scheme for  $\Phi_j^n \approx \Phi(x_j, t_n)$  is given by

$$\Phi_j^{n+1} = \Phi_j^n - \frac{\Delta t}{\Delta x} (\hat{\mathbf{f}}_{j+1/2} - \hat{\mathbf{f}}_{j-1/2}), \quad \hat{\mathbf{f}}_{j+1/2} = \hat{\mathbf{f}}(\Phi_{j-s+1}^n, \dots, \Phi_{j+s}^n), \quad j = 1, \dots, M,$$

where we impose the zero-flux boundary conditions

$$\hat{\mathbf{f}}_{1/2} = \hat{\mathbf{f}}_{M+1/2} = \mathbf{0}.$$

The key point is the design of the numerical flux  $\hat{\mathbf{f}}_{i+1/2}$  so that the resulting scheme is (at least formally second-order) accurate and stable. The most common approach for this task is to solve Riemann problems, either exactly (as in the original Godunov scheme, which is very costly), or approximately (e.g., as in the Roe scheme). For polydisperse sedimentation, exact Riemann solvers are out of reach, since the eigenstructure of the flux Jacobian is hard to compute.

In [12], the authors used central schemes [71] for the MLB model. In [23] we used Shu-Osher's technique [100] along with the information provided by the secular equation to get efficient schemes for polydisperse sedimentation based on MLB and HS models. We here briefly describe this scheme, which is based on the method of lines, that is, on applying an ODE solver (we use the third-order TVD Runge-Kutta method of [100]) to a spatially semi-discretized equations. For the discretization of the flux derivative we use local characteristic projections. Local characteristic information to compute  $\hat{\mathbf{f}}_{i+1/2}$  is provided by the eigenstructure of the flux Jacobian  $\mathcal{J}_{\mathbf{f}}(\Phi_{j+1/2})$ , where  $\Phi_{j+1/2} := \frac{1}{2}(\Phi_j + \Phi_{j+1})$ , given by the right and left eigenvectors,  $\mathbf{r}_{j+1/2,i}$  and  $\mathbf{l}_{j+1/2,i}$ , respectively for each species  $i$ , that form the respective matrices

$$\mathbf{R}_{j+1/2} = \begin{bmatrix} \mathbf{r}_{j+1/2,1} & \dots & \mathbf{r}_{j+1/2,N} \end{bmatrix}, \quad (\mathbf{R}_{j+1/2}^{-1})^T = \begin{bmatrix} \mathbf{l}_{j+1/2,1} & \dots & \mathbf{l}_{j+1/2,N} \end{bmatrix}.$$

From a local flux-splitting  $\mathbf{f}^{\pm,k}$  (we omit its dependency on  $j + 1/2$ ) given by  $\mathbf{f}^{-,k} + \mathbf{f}^{+,k} = \mathbf{f}$ , where  $\pm\lambda_k(\mathcal{J}_{\mathbf{f}^{\pm,k}}(\Phi)) \geq 0$ ,  $\Phi \approx \Phi_{j+1/2}$  and  $\lambda_k$  is the  $k$ -th eigenvalue,  $k = 1, \dots, N$ , we can define the  $k$ -th characteristic flux as

$$g_j^{\pm,k} = \mathbf{l}_{i+1/2,k}^T \cdot \mathbf{f}^{\pm,k}(\Phi_j).$$

If  $\mathcal{R}^+$  and  $\mathcal{R}^-$  denote upwind-biased reconstructions (in our experiments we use the fifth-order WENO method introduced in [66]), then

$$\begin{aligned} \hat{g}_{i+1/2,k} &= \mathcal{R}^+(g_{i-s+1}^{+,k}, \dots, g_{i+s-1}^{+,k}; x_{i+1/2}) + \mathcal{R}^-(g_{i-s+2}^{-,k}, \dots, g_{i+s}^{-,k}; x_{i+1/2}), \\ \hat{\mathbf{f}}_{i+1/2} &= \mathbf{R}_{i+1/2} \hat{\mathbf{g}}_{i+1/2} = \sum_{k=1}^n \hat{g}_{i+1/2,k} \mathbf{r}_{i+1/2,k}. \end{aligned}$$

If we do not want to use local characteristic information, we can use the previous formula with  $\mathbf{R}_{j+1/2} = \mathbf{I}_N$ , where  $\mathbf{I}_N$  denotes the  $N \times N$  identity matrix, and a global flux splitting  $\mathbf{f}^- + \mathbf{f}^+ = \mathbf{f}$ , where  $\pm\lambda_k(\mathcal{J}_{\mathbf{f}^{\pm}}(\Phi)) \geq 0$  for all  $k$ . With this choice, and denoting by  $\mathbf{e}_k$  the  $k$ th unit vector, we get  $g_j^{\pm,k} = \mathbf{e}_k^T \mathbf{f}^{\pm}(\Phi_j) = f_k^{\pm}(\Phi_j)$ , i.e.,  $g_j^{\pm,k}$  are the components of the split fluxes, and the numerical flux

is computed component by component by reconstructing the split fluxes component by component, i.e.,  $\hat{\mathbf{f}}_{j+1/2} = (\hat{f}_{j+1/2,1}, \dots, \hat{f}_{j+1/2,N})^\top$ , where

$$\hat{f}_{j+1/2,k} = \mathcal{R}^+(g_{j-s+1}^{+,k}, \dots, g_{j+s-1}^{+,k}; x_{j+1/2}) + \mathcal{R}^-(g_{j-s+2}^{-,k}, \dots, g_{j+s}^{-,k}; x_{j+1/2}), \quad k = 1, \dots, N.$$

This scheme will be referred to as COMP-GLF and it is a high-order extension of the Lax-Friedrichs scheme.

We now explain the SPEC-INT scheme. If  $\lambda_k(\mathcal{J}_{\mathbf{f}}(\Phi)) > 0$  (respectively,  $< 0$ ) for all  $\Phi \in [\Phi_j, \Phi_{j+1}]$ , where  $[\Phi_j, \Phi_{j+1}] \subset \mathbb{R}^N$  denotes the segment joining both states, then we upwind (since then there is no need for flux splitting):

$$\mathbf{f}^{+,k} = \mathbf{f}, \quad \mathbf{f}^{-,k} = 0 \quad \text{if } \lambda_k(\mathcal{J}_{\mathbf{f}}(\Phi)) > 0, \quad \mathbf{f}^{+,k} = 0, \quad \mathbf{f}^{-,k} = \mathbf{f} \quad \text{if } \lambda_k(\mathcal{J}_{\mathbf{f}}(\Phi)) < 0.$$

On the other hand, if  $\lambda_k(\mathcal{J}_{\mathbf{f}}(\Phi))$  changes sign on  $[\Phi_j, \Phi_{j+1}]$ , then we use a Local Lax-Friedrichs flux splitting given by  $\mathbf{f}^{\pm,k}(\Phi) = \mathbf{f}(\Phi) \pm \alpha_k \Phi$ , where the numerical viscosity parameter  $\alpha_k$  should satisfy

$$\alpha_k \geq \max_{\Phi \in [\Phi_j, \Phi_{j+1}]} |\lambda_k(\mathcal{J}_{\mathbf{f}}(\Phi))|. \quad (1.9)$$

The usual choice of the numerical viscosity

$$\alpha_k = \max\{|\lambda_k(\mathcal{J}_{\mathbf{f}}(\Phi_j))|, |\lambda_k(\mathcal{J}_{\mathbf{f}}(\Phi_{j+1}))|\}$$

produces oscillations in the numerical solution indicating that the amount of numerical viscosity is insufficient. The right-hand side of (1.9) can usually not be evaluated exactly since closed-form expressions for the eigenvalues are not available. However, for the present class of models, we may use the interlacing property (see Corollary 1.2) to generate a fairly sharp bound of that expression. In the case of the MLB model, we have  $\gamma_k < 0$  (see [22, 51]) and the interlacing property can be written as

$$v_{k+1}(\Phi) \leq \lambda_k(\Phi) \leq v_k(\Phi), \quad k = 1, \dots, N$$

so we have

$$|\lambda_k(\Phi)| \leq \max\{|v_k(\Phi)|, |v_{k+1}(\Phi)|\}$$

and therefore we can get efficiently computable bounds

$$\max_{\Phi \in [\Phi_j, \Phi_{j+1}]} |\lambda_k(\Phi)| \leq \alpha_k := \max\left\{\max_{\Phi \in [\Phi_j, \Phi_{j+1}]} |v_k(\Phi)|, \max_{\Phi \in [\Phi_j, \Phi_{j+1}]} |v_{k+1}(\Phi)|\right\}, \quad k = 1, \dots, N. \quad (1.10)$$

(The same property also holds for other models, under appropriate circumstances [22].) We denote by ‘‘SPEC-INT’’ the scheme for which  $\alpha_1, \dots, \alpha_N$  are defined by (1.10).

### 1.3.2 Adaptive Mesh Refinement (AMR)

We now outline the main building blocks of the AMR algorithm and refer to [6] for details. We denote by  $G_0, \dots, G_L$  a 1D grid hierarchy composed of  $L + 1$  grids, such that, except for the coarsest grid  $G_0$ , cells of a given grid are obtained by the subdivision of cells of the immediately coarser grid into  $r$  parts (we assume  $r = 2$ ). The unit interval is thus divided into  $N_0, \dots, N_L$  subintervals of length  $h_l = 1/N_l$ , with  $N_l = 2^l N_0$ ,  $l = 0, \dots, L$ , whose centers will be denoted by  $x_j^l = (j + 1/2)h_l$ ,  $j = 0, \dots, N_l - 1$ ,  $l = 0, \dots, L$ . A “mesh”  $G_l$  at resolution level  $l$  is just a subset of the index set  $\{0, \dots, N_l - 1\}$  whose “extent”, the union of the cells indexed by elements of  $G_l$ , is denoted by  $\Omega_l(G_l)$ . We consider only “nested” grid hierarchies, i.e.,  $\Omega_l(G_l) \subseteq \Omega_{l-1}(G_{l-1})$  for  $1 \leq l \leq L$  is assumed to hold along with  $\Omega_0(G_0) = \Omega$ .

The meshes will be dynamically updated so that they adapt to the features of the solution, and we denote by  $G_l^{t_l}$  the mesh that corresponds to the resolution level  $l$  and time  $t_l$ . Over each mesh we consider a numerical solution defined by a discrete function  $\Phi_l^{t_l} = (\Phi_{l,j}^{t_l})$ , with  $\Phi_{l,j}^{t_l} \approx \Phi(x_j^l, t_l)$  and  $j \in G_l^{t_l}$ . For a given time instant  $t$  we denote  $\Phi^t = \{\Phi_l^t\}$ ,  $0 \leq l \leq L$ . This set includes the solution values on all scales.

The algorithm can be described by the time evolution of the meshes and their associated numerical solutions, starting with  $t_l = 0$ ,  $l = 0, \dots, L$  and ending at  $t_l = T$ ,  $l = 0, \dots, L$ , for some  $T > 0$ . The main building blocks of the AMR algorithm — integration and adaptation of the grids and projection from fine to coarse grids — are described next.

We now describe how the integration of the grids corresponding to the various refinement levels is organized. The first step is to select a time step  $\Delta t_0$  for the coarsest grid, so that the following CFL condition relevant for the grid  $G_0^t$  is satisfied:

$$\Delta t_0 \leq \frac{C_{\text{cfl}} h_0}{K}, \quad 0 < C_{\text{cfl}} \leq 1, \quad (1.11)$$

where  $K$  is an upper bound of the spectral radius of all Jacobian matrices  $\mathcal{J}_{\mathbf{f}}(\Phi)$ ,  $\Phi \in \Phi^t$ : For the SPEC-INT scheme, the eigenvalues of  $\mathcal{J}_{\mathbf{f}}(\Phi)$  are readily available to compute the spectral radius, whereas for COMP-GLF one can use (1.8) to obtain an upper bound for it. The time steps for the rest of the grids are taken by  $\Delta t_l = \Delta t_{l-1}/2$  for  $l = 1, \dots, L$ , which implies that the equivalent CFL condition holds for each grid. A time step for  $G_0$  corresponds therefore to  $2^l$  time steps for  $G_l$ . The grids are integrated according to the order dictated by the following condition:  $t_{l'} \leq t_l \leq t_{l'} + \Delta t_l$  if  $l \leq l'$ .

At some step of this time evolution,  $(\Phi_l^{t_l+k\Delta t_l}, G_l^{t_l})$ ,  $k = 1, 2$ , are sequentially computed from  $(\Phi_l^{t_l}, G_l^{t_l})$ , supplemented by boundary conditions at a band surrounding  $\Omega_l(G_l^{t_l})$  obtained by (cubic in space and linear in time) interpolation from  $(\Phi_{l-1}^{t_l}, G_{l-1}^{t_l})$  and  $(\Phi_{l-1}^{t_l+2\Delta t_l}, G_{l-1}^{t_l+2\Delta t_l})$ , which must have been computed in previous steps. Once  $(\Phi_l^{t_l+2\Delta t_l}, G_l^{t_l+2\Delta t_l})$  is computed, there is data that overlay  $\Omega_l(G_l^{t_l})$  at different resolution levels. It is at this point that the projection of the data at the fine resolution level should be applied to modify the values  $\Phi_{l-1,j}^{t_l+2\Delta t_l}$  of the immediately coarser grid function that correspond to cells overlaid by cells at  $G_l^{t_l}$  and adjacent to them as well, i.e., such that  $\{2j, 2(j-1), 2(j+1)\} \cap G_l^{t_l} \neq \emptyset$ . In this case, the numerical fluxes corresponding to interfaces

of refined coarse cells are modified in such a way that

$$\Phi_{l-1,j}^{t_l+2\Delta t_l} = \frac{1}{2}(\Phi_{l,2j}^{t_l+2\Delta t_l} + \Phi_{l,2j+1}^{t_l+2\Delta t_l}) \quad \text{if} \quad \Phi_{l-1,j}^{t_l} = \frac{1}{2}(\Phi_{l,2j}^{t_l} + \Phi_{l,2j+1}^{t_l}),$$

i.e., discrete conservation is maintained.

The next issue is the update of the grids. The grids corresponding to the various levels  $G_l$ ,  $1 \leq l \leq L$  have to be modified according to the current characteristics of the flow. The adaptation of each refinement level is performed by discarding the current grid and creating a new one according to specified refinement criteria. In this way, coarsening is not directly performed on refined areas, but implicitly obtained by not refining.

For our cell-centered approach, if  $x_j^l = (j + 1/2)h_l$  is the center of a cell of a grid  $G_l^t$  and  $\mathcal{I}(\Phi_{l-1}^t, x)$  is an MUSCL interpolation operator defined on the data  $\Phi_{l-1}^t = \{\Phi_{l-1,i}^t\}_{i \in G_{l-1}^t}$ , then the cell defined by  $x_j^l$  will be selected for refinement if

$$\left| \Phi_{l,j}^t - \mathcal{I}(\Phi_{l-1}^t, x_j^l) \right| > \tau_p \cdot \max_{l,j} \left| \Phi_{l,j}^t - \mathcal{I}(\Phi_{l-1}^t, x_j^l) \right|,$$

where  $\tau_p$  is a given tolerance. Note that only the cells present in the current grid are considered for refinement. New cells are included only because of the addition of some extra cells around each marked cell. We also ensure that the refined grid is obtained by subdivision of coarse cells: if a cell  $x_{l,j}^t$  is selected for refinement, then every cell that overlaps the same coarse cell as  $x_{l,j}^t$  is also included in the refined grid. Further, we also include a cell in the refinement list if the modulus of the discrete gradient, computed in the coarser grid, exceeds some large threshold, so that shock formation can be detected from steepened data. For the discrete gradient we use the approximation

$$\frac{\partial \Phi}{\partial x}(x_j^{l-1}, t) \approx \frac{1}{h_{l-1}} \max\{|\Phi_{l-1,j+1}^t - \Phi_{l-1,j}^t|, |\Phi_{l-1,j}^t - \Phi_{l-1,j-1}^t|\}.$$

Once the cells that will compose the refined grid have been selected we add a certain number of extra cells forming a band around each marked cell to ensure that the cells adjacent to a singularity are refined. This device of creating “safety points” follows the spirit of [60, 74, 91]. These extra cells will avoid singularities to escape from the fine grid during one coarse time step. To this aim it would suffice to add a band of one coarse cells around each marked cell. Another criterion for adding cells is dictated by the need of interpolating ghost cell values from *relatively* smooth regions: the length of the stencil of the interpolation operator must be less than twice the length — measured in number of coarse cells — of the band of added cells. In our case we use third-order linear interpolation, and this imposes the addition of two coarse cells at each side of a marked cell. For analogous reasons, if the computation of the numerical flux depends on  $2\mathcal{N}$  values of the fine grid, then, in order to ensure that it is computed using non-interpolated data, the length of the band has to be greater than  $\mathcal{N}/2$ . In the case of the method used in this work, described in Section 1.3.1, we have  $\mathcal{N} = 3$ , and thus the number of coarse cells added should be at least 2. According to the criteria above, we add two coarse cells in our implementation.

The last observation for this refinement procedure is that it should be performed from fine to coarse resolution levels to ensure that at every moment of the update process it holds that

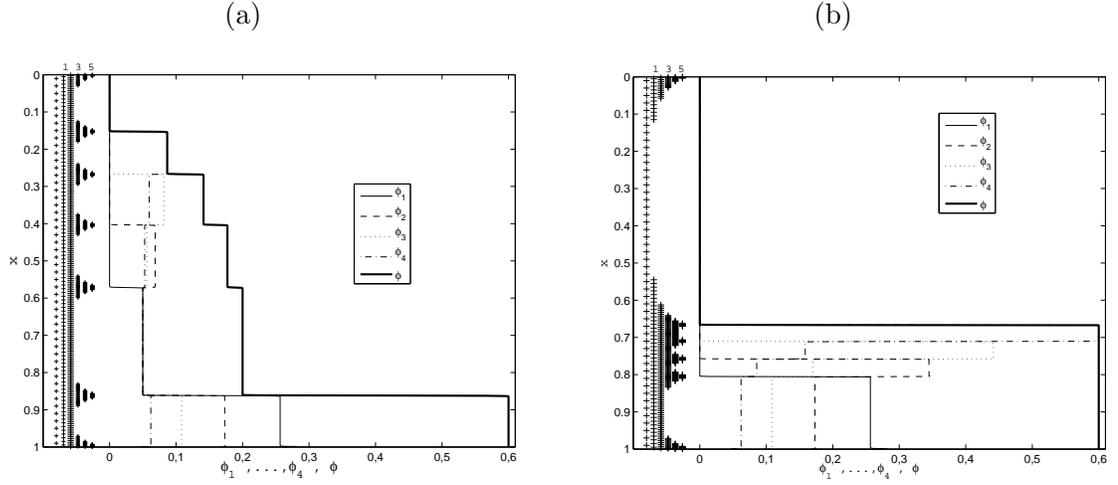


Figure 1.1: Example 1.1 ( $N = 4$ ): numerical solution obtained with SPEC-INT-AMR with  $L+1 = 6$  levels with coarsest grid of 50 subintervals at (a)  $t = 50$  s and (b)  $t = 300$  s. Here and in Figure 1.5, the numbers 1, 3 and 5 on the top of the frame bottom represent the level  $l$  of the grid hierarchy.

$\Omega_l(G_l^t) \subseteq \Omega_{l-1}(G_{l-1}^t)$ . We also enforce the inclusion  $\Omega_l(G_l^t) \supseteq \Omega_{l+1}(G_{l+1}^t)$ , so that the whole sequence of grids verifies the desired inclusions. Finally note that the process of computing data at the corresponding surrounding bands is possible because the grids are nested, and this implies that  $\tilde{\Omega}_l(G_l^t) \subseteq \tilde{\Omega}_{l-1}(G_{l-1}^t)$ .

Once the new grid  $\hat{G}_l$  is computed such that  $\Omega_l(\hat{G}_l^t) \subseteq \Omega_{l-1}(G_{l-1}^t)$  is satisfied, one sets

$$\hat{\Phi}_{l,j}^t = \begin{cases} \mathcal{I}(\Phi_{l-1}^t, x_j^l) & \text{if } j \in \hat{G}_l^t \setminus G_l^t, \\ \Phi_{l,j}^t & \text{if } j \in G_l^t, \end{cases}$$

i.e., the value at the  $j$ -th cell is interpolated from data at the next coarser level for cells not in  $G_l^t$ . The refined grid is therefore defined by  $(\hat{G}_l^t, \hat{\Phi}_l^t)$ . Discrete boundary conditions are also applied if the grid overlaps the domain boundary.

## 1.4 Numerical results

The threshold value  $\tau_p$  plays a special role for the performance of the SPEC-INT-AMR and COMP-GLF-AMR methods, since computational time and accuracy are related to their variations. Initially, we perform additional numerical experiments based on different examples in order to study the effect of different choices of the parameter  $\tau_p$ . In preliminary computations (not shown here) we tested the values  $\tau_p = 10^{-q}$ ,  $q = 2, 3, 4, 5$ . For  $q = 2$  the AMR schemes turned out to be most efficient in terms of reduction of error per CPU time. Moreover, in our experiments we use a CFL number of  $C_{\text{cfl}} = 0.5$ , cf. (1.11).

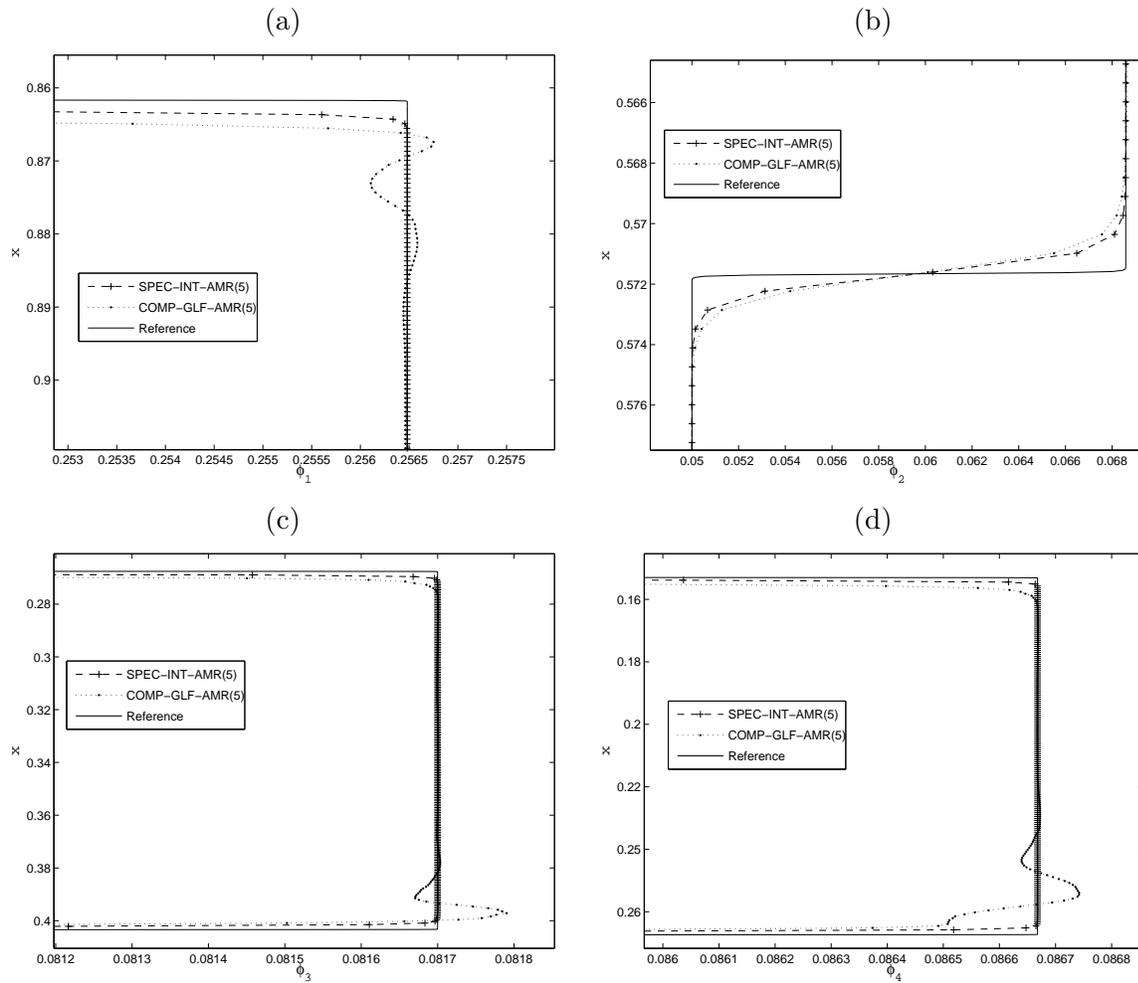


Figure 1.2: Example 1.1 ( $N = 4$ ): numerical solution at time  $t = 50$  s: details of using SPEC-INT-AMR and COMP-GLF-AMR methods with a grid hierarchy of  $L + 1 = 5$  levels on a coarsest grid with  $N_0 = 50$  subintervals. Here and in Figure 1.3, the reference solution is computed by SPEC-INT on a fixed grid with  $N_{\text{ref}} = 25600$  subintervals.

Fixed grid size	SPEC-INT-AMR(5)			COMP-GLF-AMR(5)		
	% Integrations	% Memory	% CPU time	% Integrations	% Memory	% CPU time
800	23.77	20.11	23.21	28.11	23.85	20.80
1600	12.34	11.31	12.60	14.16	13.30	13.91
3200	6.41	6.09	7.54	7.60	7.31	6.50
6400	3.31	3.30	4.95	4.17	4.02	4.92

Table 1.1: Example 1.1 ( $N = 4$ ): Percentage of storage space (memory), number of integrations and CPU time of the adaptive algorithm with respect to the fixed grid algorithm with  $\tau_p = 10^{-2}$  at simulated time  $t = 50$  s, for a hierarchy of  $L + 1 = 5$  levels and four different values of  $N_0$ .

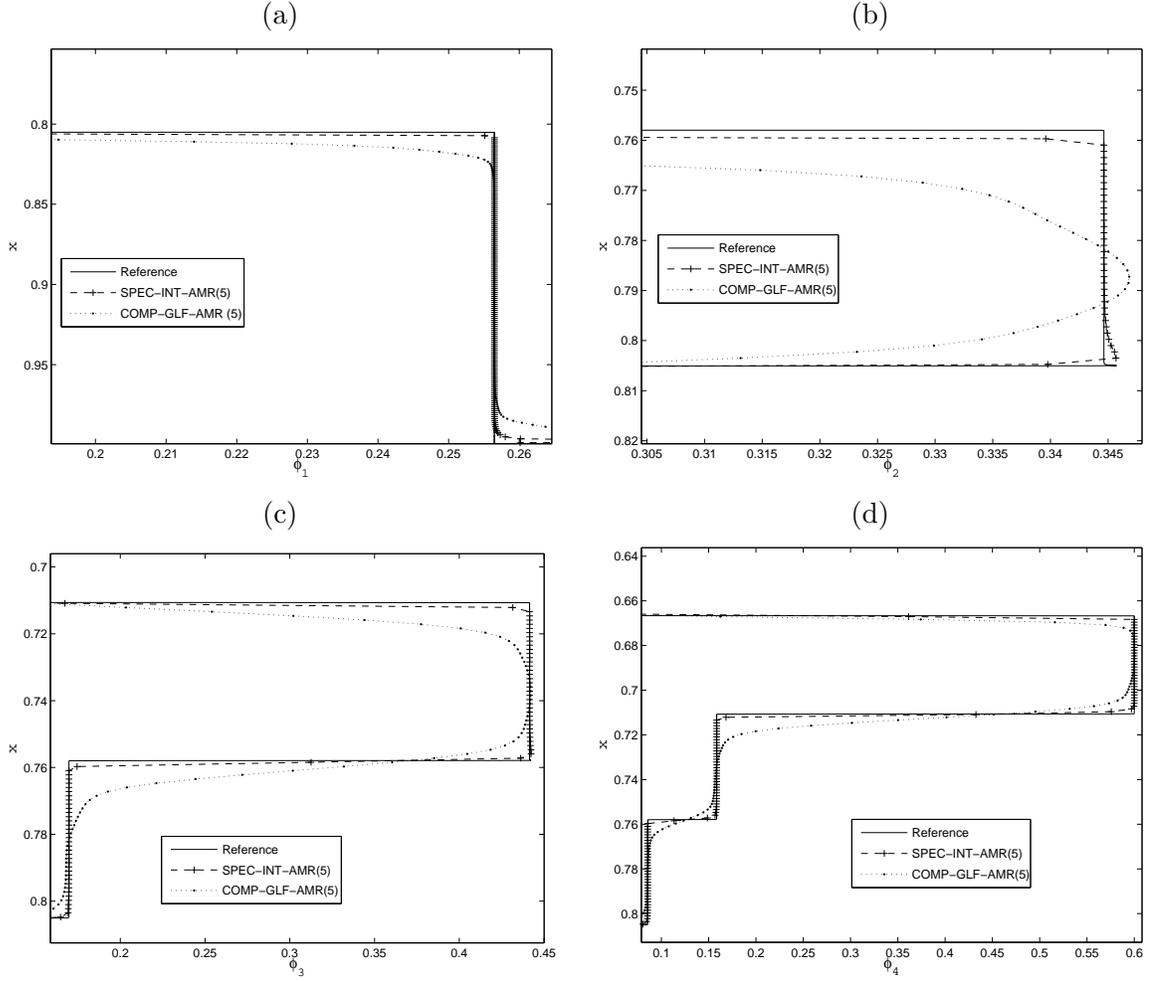


Figure 1.3: Example 1.1 ( $N = 4$ ): details of numerical solution at time  $t = 300$ s obtained by SPEC-INT-AMR and COMP-GLF-AMR methods with a grid hierarchy of  $L + 1 = 5$  levels on a coarsest grid with  $N_0 = 50$  subintervals.

#### 1.4.1 Example 1.1 ( $N = 4$ )

As in [22, 23], we consider the standard test case of batch settling of an initially homogeneous suspension in a column characterized by the velocity functions (1.2) with parameters  $N = 4$ ,  $\rho_s = 2790 \text{ kg/m}^3$ ,  $\rho_f = 1208 \text{ kg/m}^3$ ,  $\mu_f = 0.02416 \text{ Pa s}$ ,  $g = 9.8 \text{ m/s}^2$ ,  $\delta_1 = 1$ ,  $\delta_2 = 0.64$ ,  $\delta_3 = 0.36$  and  $\delta_4 = 0.16$ . We employ the hindered settling factor (1.3) with exponent  $n_{RZ} = 4.7$  and initial concentrations  $\phi_i^0 = 0.05$  for  $i = 1, \dots, 4$ .

We simulate the process until the sedimentation process attains a steady state, that is when all particles are captured in a sediment. This sediment usually consists of layers of different composition (which is part of the solution of the problem) separated by stationary kinematic shocks. Figure 1.1 shows the numerical solution obtained with SPEC-INT-AMR as concentration profiles at two different times together with the corresponding hierarchical grids. We have used a grid

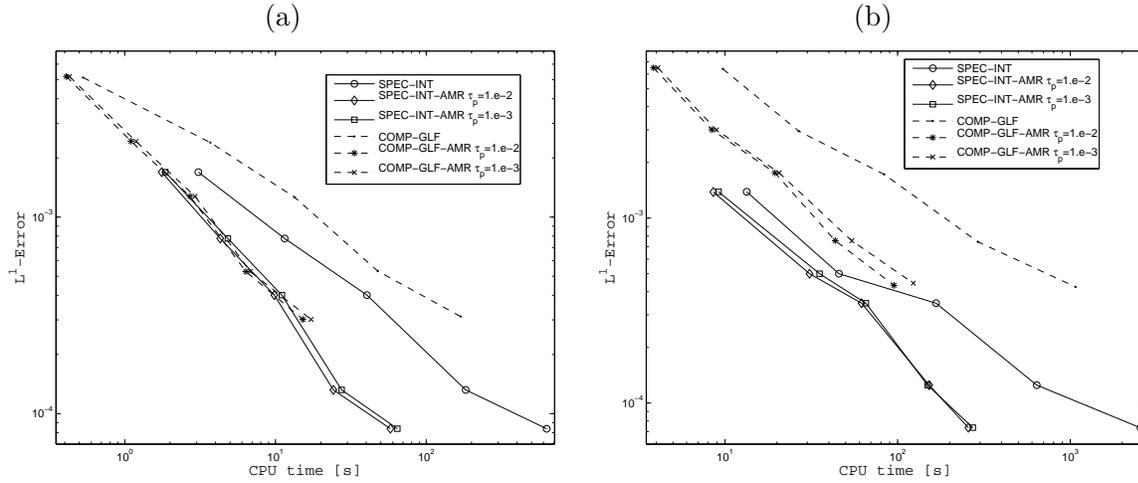


Figure 1.4: Example 1.1 ( $N = 4$ ): approximate  $L^1$  errors versus CPU time for SPEC-INT-AMR and COMP-GLF-AMR for different values of  $\tau_p$  at simulated times  $t = 50$  s (a) and  $t = 300$  s (b). Here and in Figure 1.7, for each scheme the interpolated symbols correspond to different values of  $N_0$ , and the AMR versions have been implemented with  $L + 1 = 4$  levels of refinement.

hierarchy of  $L + 1 = 6$  levels with a coarsest grid of  $N_0 = 50$  subintervals so that the finest grid has  $N_5 = N_L = 1600$  subintervals. The tolerance for the interpolation error sensor has been set to  $\tau_p = 10^{-2}$ . The plotted positions indicate that the adaptive mesh refinement technique works correctly, in the sense that the scheme correctly detects the formation of shocks and refines the corresponding areas.

In Figures 1.2 and 1.3 we compare the results obtained by SPEC-INT-AMR and COMP-GLF-AMR with a reference solution computed with SPEC-INT scheme on a fixed uniform grid with  $N_{\text{ref}} = 25600$  subintervals at times  $t = 50$  s and  $t = 300$  s, respectively. We have used a grid hierarchy of  $L + 1 = 5$  levels with a coarsest grid of  $N_0 = 50$  subintervals so that  $N_L = N_4 = 800$ . The tolerance for the interpolation error sensor has been set to  $\tau_p = 10^{-2}$ . The solution of the adaptive computation is mapped to the finest grid using the computed solution where available, and by interpolation from coarse to fine elsewhere. A few enlarged views of relevant parts of the numerical solutions shown in Figures 1.2 and 1.3 illustrate that the algorithm correctly identifies the regions corresponding to the shocks. The percentage of the number of integrations of the AMR algorithm with respect to a solution on the uniform finest grid with  $N_4$  subintervals is 25.20% and 29.70% for the SPEC-INT-AMR and COMP-GLF-AMR methods, respectively, for this experiment. We observe a slight “overshoot” in the  $\phi_1$  plot of Figure 1.3 close to  $x = 1$ . This phenomenon also appears in the reference solution, and is possibly associated with the fact that the function  $V(\phi)$  is cut off abruptly at  $\phi = \phi_{\text{max}}$ , see (1.3).

In Table 1.4, we present the percentages of storage space, number of integrations and CPU time required by AMR with respect to schemes on the uniform finest mesh for constant tolerance  $\tau_p = 10^{-2}$ . The indicated percentage represents the average memory load over all iterations. The values of the table correspond to coarsest grids of  $N_0 = 50, 100, 200$  and  $400$  subintervals and

		SPEC-INT			COMP-GLF			SPEC-INT			COMP-GLF		
		$t = 50\text{s}$			$t = 50\text{s}$			$t = 300\text{s}$			$t = 300\text{s}$		
$N_0$	error	cr	cpu	error	cr	cpu	error	cr	cpu	error	cr	cpu	
200	169.3	—	3.0	512.6	—	0.5	138.7	—	13.3	640.3	—	9.7	
400	77.8	1.12	11.5	239.5	1.09	3.7	50.0	1.47	45.7	295.1	1.11	26.7	
800	40.0	0.95	40.4	126.4	0.92	13.3	34.6	0.52	166.9	172.3	0.77	83.6	
1600	13.2	1.60	182.6	53.1	1.25	47.7	12.5	1.47	639.7	74.3	1.21	291.5	
3200	8.3	0.65	630.0	31.1	0.76	169.1	7.3	0.76	2551.2	42.4	0.81	1074.1	
$t =$	SPEC-INT-AMR			SPEC-INT-AMR			COMP-GLF-AMR			COMP-GLF-AMR			
50 s	$\tau_p = 10^{-2}$			$\tau_p = 10^{-3}$			$\tau_p = 10^{-2}$			$\tau_p = 10^{-3}$			
$N_0$	error	cr	cpu	error	cr	cpu	error	cr	cpu	error	cr	cpu	
25	169.3	—	1.7	169.3	—	1.8	519.1	—	0.4	518.9	—	0.43	
50	77.8	1.12	4.3	77.8	1.12	4.8	243.4	1.09	1.1	242.4	1.09	1.1	
100	40.0	0.95	9.8	40.0	0.95	11.0	127.6	0.93	2.7	127.4	0.92	2.9	
200	13.2	1.60	24.2	13.2	1.60	27.4	52.8	1.27	6.3	52.8	1.27	6.8	
400	8.3	0.65	57.8	8.3	0.65	64.3	30.1	0.88	15.2	30.1	0.80	17.2	
$t =$	SPEC-INT-AMR			SPEC-INT-AMR			COMP-GLF-AMR			COMP-GLF-AMR			
300 s	$\tau_p = 10^{-2}$			$\tau_p = 10^{-3}$			$\tau_p = 10^{-2}$			$\tau_p = 10^{-3}$			
$N_0$	error	cr	cpu	error	cr	cpu	error	cr	cpu	error	cr	cpu	
25	138.3	—	8.5	138.5	—	9.1	648.8	—	3.8	649.0	—	4.0	
50	50.0	1.47	30.8	50.0	1.47	35.3	300.2	1.11	8.3	299.6	1.11	8.9	
100	34.6	0.52	61.8	34.6	0.52	65.4	175.7	0.77	19.4	175.6	0.77	20.5	
200	12.5	1.47	152.3	12.5	1.47	149.0	75.4	1.21	43.5	75.3	1.22	54.2	
400	7.3	0.76	256.8	7.3	0.76	272.5	43.2	0.80	95.2	44.4	0.76	123.2	

Table 1.2: Example 1.1: approximate  $L^1$  errors (in short, “error”), multiplied by  $10^5$ ; convergence rates (cr), and CPU times (in seconds), at two different times for SPEC-INT and COMP-GLF on a fixed grid, and SPEC-INT-AMR and COMP-GLF-AMR with two different values of  $\tau_p$ .

$i$	1	2	3	4	5	6	7
$\phi_i^0 [10^{-2}]$	0.2365	1.1039	3.5668	3.8776	6.0436	10.8998	4.2718
$d_i [10^{-5} \text{ m}]$	290	250	210	170	130	90	50
$\delta_i$	1.000000	0.743163	0.524376	0.343639	0.200951	0.0963	0.029727

Table 1.3: Example 1.2: Initial conditions  $\phi_i^0$ , real  $d_i$  and normalized squared  $\delta_i$  particle sizes.

$L + 1 = 5$  levels of refinement. We observe that CPU time and the percentage of memory allocated by the SPEC-INT-AMR and COMP-GLF-AMR codes decreases as  $N_0$  increases, as expected. However, we observe in Figures 1.2 that COMP-GLF-AMR code produces spurious oscillations in the piecewise constant part of the solution. These oscillations also are present in the results obtained with COMP-GLF with uniform mesh and do not disappear upon mesh refinement, as was observed in [23]. In Table 1.2 and Figure 1.4, we display approximate  $L^1$  errors and CPU times at two different times for the methods SPEC-INT-AMR and COMP-GLF-AMR using a grid hierarchy for different levels, corresponding to values of  $N_0 = 25, 50, 100, 200$  and  $400$ , and for the methods

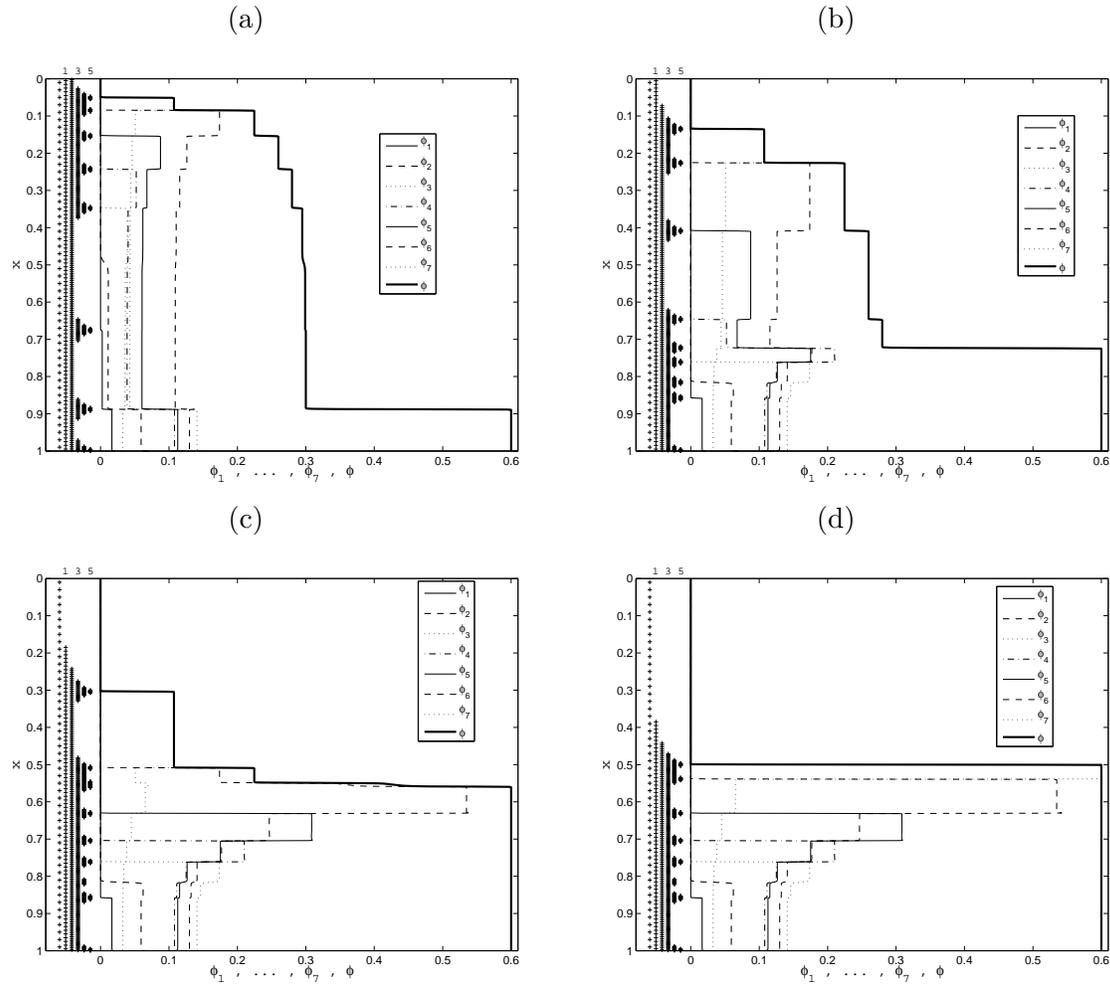


Figure 1.5: Example 1.2. ( $N = 7$ ): numerical solution with SPEC-INT-AMR with  $L + 1 = 6$  levels, coarsest grid of  $N_0 = 50$  cells at  $t = 228.05$  s (a),  $t = 400$  s (b),  $t = 1286.94$  s (c) and  $t = 2500$  s (d).

SPEC-INT and COMP-GLF using a fixed uniform grid corresponding to values of  $N_0 = 200, 400, 800, 1600$  and  $3200$ . For a given base resolution  $N_0$  and hierarchy depth  $L$ , we observe that COMP-GLF-AMR is faster than SPEC-INT-AMR (as should be expected, since the former scheme avoids the costly computation of the spectral decomposition of the flux Jacobian), but the latter provides results that are much closer to the reference solution. In addition, for a fixed  $L^1$  error, the CPU time is smaller for the AMR technique than for the equivalent fixed-grid computation. In many cases the AMR technique is around ten times faster. We plotted different choices for the threshold value  $\tau_p$  and observe that, for this case, the choice  $\tau_p = 10^{-2}$  gives the best efficiency.

### 1.4.2 Example 1.2 ( $N = 7$ )

This example is based on experimental data from [52], where the settling of suspension in a column of height  $h = 0.227$  m was considered. The initial concentrations  $\phi_i^0$ , diameters  $d_i$  and

normalized diameters  $\delta_i = d_i/d_1$  are given in Table 1.3. The maximum total concentration is  $\phi_{\max} = 0.6$  and the hindered settling factor is given by (1.3) with the exponent  $n_{\text{RZ}} = 5$ . These values were also used in [52]. The value of the exponent  $n_{\text{RZ}}$  is close to 4.65, which is the limiting value for this exponent for small particle Reynolds numbers and if particles are small compared with the width of the settling column, according to the well-known theoretical analysis of [93].

Other authors use similar values. For instance, the value  $n_{\text{RZ}} = 4.7$  of Example 1.1 has been adopted from [96], while  $n_{\text{RZ}} = 5$  is supported in [103]. On the other hand, the value of  $\phi_{\max}$  is supported by experimental results by Amy et al. [2], who observe that particle segregation is completely stopped beyond that value, which in turn is close to the packing density of around 0.65 for random arrangements of equal-sized spheres. (It is well known that imposing a limit on  $\phi$ , as we do here, is actually a strong simplification for a multi-size polydisperse suspension since smaller spheres may fill the interstices between larger ones, and thereby denser packings could be achieved; see the discussion e.g. in [96].)

We simulate the process until the phenomenon enters in a steady state. Figs 1.5 shows the numerical solution obtained with SPEC-INT-AMR as concentration profiles at four different times together with the corresponding grid hierarchy. We have used a grid hierarchy of  $L + 1 = 6$  levels with a coarsest grid of  $N_0 = 50$  points so that results are comparable with those for a fixed grid of  $N_5 = 1600$  points. The tolerance for the interpolation error sensor has been set to  $\tau_p = 10^{-2}$ .

The plotted positions indicate that the adaptive mesh refinement technique works correctly, in the sense that the scheme detects the shock formation and refines these areas.

In Figures 1.6 we show enlarged views of relevant parts of the numerical solutions for individual concentration and compare the results obtained at times  $t = 400$  s by SPEC-INT-AMR method with a grid hierarchy of  $L + 1 = 6$  levels on a coarsest grid with  $N_0 = 50$ , and SPEC-INT method on a fixed grid with  $N_0 = 1600$ , with a reference solution computed by SPEC-INT on a fixed grid with  $N_{\text{ref}} = 25600$  subintervals. The tolerance for the interpolation error sensor has been set to  $\tau_p = 10^{-2}$ . The percentage of integrations of the algorithm with respect to the solution on the finest grid is a 16.88% and the percentage of memory allocated by AMR technique is only 14.62% of the amount of memory allocated after each iteration of the coarse grid.

We comment that the original information in [52] is related to a suspension with a continuous initial particle size distribution determined from experiments in [2], which is then discretized into  $N = 28$  particle size ‘bins’. This particle size distribution also defines the initial data for 14 or 7 size classes (by suitable summation of the initial volume fractions for  $N = 28$  over two or four neighboring bins). The final (stationary) result for  $N = 28$  obtained by a first-order numerical method, is compared in [52] with original experimental data from [2]. Since the mathematical model of [52] is the same as ours, with the same parameters, the solutions are nearly the same, up to small errors due to the different numerical methods. We verified this by calculating  $\Phi$  for  $N = 28$  with the present method; results practically coincided with those of [52, Figure 2] (not shown here). In this work we chose  $N = 7$  to allow for a reasonable graphic representation while focusing on fine details of the transient profiles of the individual species.

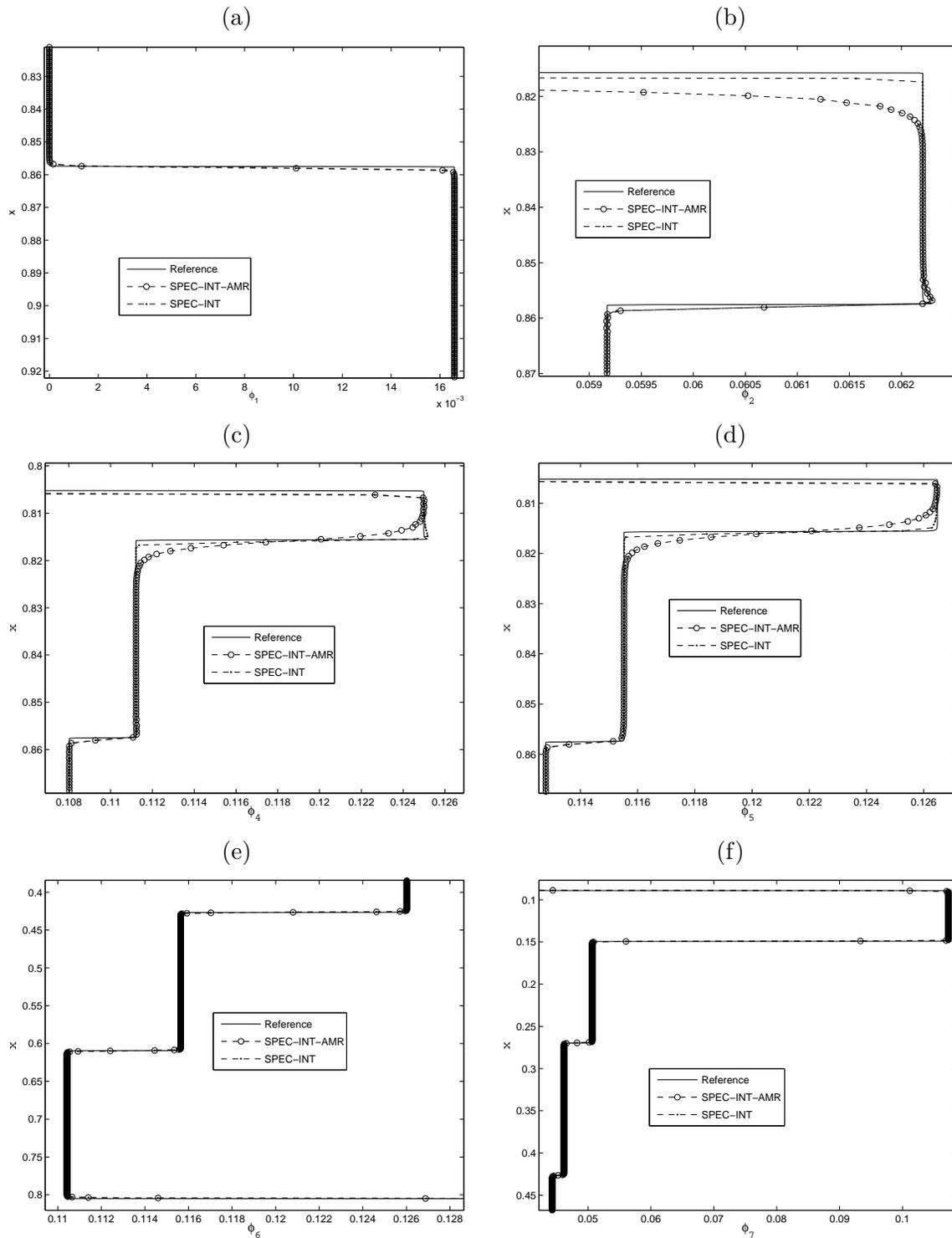


Figure 1.6: Example 1.2 ( $N = 7$ ): enlarged view in Figure 1.5(b) for  $\phi_1$  (a),  $\phi_2$  (b),  $\phi_4$  (c),  $\phi_5$  (d),  $\phi_6$  (e) and  $\phi_7$  (f). In addition, the solution produced by SPEC-INT on a fixed grid with 1600 subintervals and the reference solution computed by SPEC-INT on a fixed grid with  $N_{\text{ref}} = 12800$  is shown.

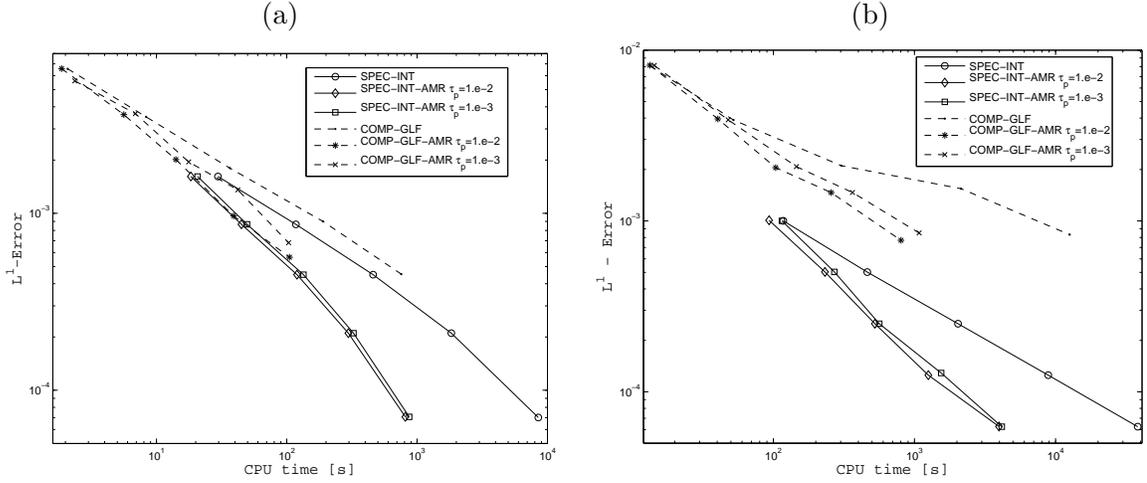


Figure 1.7: Example 1.2 ( $N = 7$ ): approximate  $L^1$  errors versus CPU time for SPEC-INT-AMR, COMP-GLF-AMR for different values of  $\tau_p$  at simulated times  $t = 400$  s (a) and  $t = 2500$  s (b).

		SPEC-INT			COMP-GLF			SPEC-INT			COMP-GLF		
		$t = 400$ s			$t = 400$ s			$t = 2500$ s			$t = 2500$ s		
$N_0$	error	cr	cpu	error	cr	cpu	error	cr	cpu	error	cr	cpu	
200	161.5	—	29.6	656.0	—	2.0	100.0	—	117.6	812.5	—	13.2	
400	86.6	0.89	117.3	350.8	0.90	8.3	50.2	0.99	462.7	389.7	1.06	52.1	
800	45.0	0.94	459.3	180.0	0.96	36.5	25.0	1.00	2036	210.6	0.88	303.6	
1600	21.0	1.09	1830.6	90.1	0.99	189.2	12.5	1.00	8876	154.5	0.44	2137	
3200	7.0	1.57	8520.5	45.3	0.99	756.2	6.2	1.00	38187	83.1	0.89	12579	
$t =$	SPEC-INT-AMR			SPEC-INT-AMR			COMP-GLF-AMR			COMP-GLF-AMR			
400 s	$\tau_p = 10^{-2}$			$\tau_p = 10^{-3}$			$\tau_p = 10^{-2}$			$\tau_p = 10^{-3}$			
$N_0$	error	cr	cpu	error	cr	cpu	error	cr	cpu	error	cr	cpu	
25	161.3	—	18.3	161.4	—	20.6	657.5	—	1.8	560.2	—	2.3	
50	86.6	0.89	44.8	86.6	0.89	49.5	361.0	0.86	5.6	365.8	0.61	6.8	
100	45.0	0.94	119.7	45.0	0.94	134.2	201.3	0.84	14.0	195.1	0.91	17.6	
200	21.0	1.09	296.8	21.0	1.09	324.9	96.7	1.05	39.0	135.6	0.52	42.0	
400	7.0	1.57	712.3	7.0	1.57	869.4	56.4	0.77	104.8	68.1	0.99	103.2	
$t =$	SPEC-INT-AMR			SPEC-INT-AMR			COMP-GLF-AMR			COMP-GLF-AMR			
2500 s	$\tau_p = 10^{-2}$			$\tau_p = 10^{-3}$			$\tau_p = 10^{-2}$			$\tau_p = 10^{-3}$			
$N_0$	error	cr	cpu	error	cr	cpu	error	cr	cpu	error	cr	cpu	
25	100.8	—	93.2	100.0	—	114.9	815.1	—	13.3	814.1	—	14.3	
50	50.3	1.00	232.0	50.3	0.99	270.7	395.5	1.04	40.4	390.7	1.05	47.8	
100	25.0	1.00	524.0	25.0	1.00	563.5	205.2	0.94	103.8	207.8	0.91	145.7	
200	12.5	1.00	1254.2	12.8	0.95	1546.0	146.5	0.48	256.6	146.7	0.50	363.7	
400	6.2	0.99	3971.6	6.2	1.04	4158.5	77.0	0.92	802.4	85.1	0.78	1078	

Table 1.4: Example 1.2: approximate  $L^1$  errors (in short, “error”), multiplied by  $10^5$ ; convergence rates (cr), and CPU times (in seconds), at two different times for SPEC-INT and COMP-GLF on a fixed grid, and SPEC-INT-AMR and COMP-GLF-AMR with two different values of  $\tau_p$ .

The approximate errors in the  $L^1$ -norm between the adaptive scheme for different levels and the corresponding scheme on an equivalent uniform fine grid are displayed in Table 1.4 and in Figure 1.7. From that figure we infer that the adaptive techniques SPEC-INT-AMR and COMP-GLF-AMR are always cheaper in CPU time than SPEC-INT and COMP-GLF methods on the finest grid. The reference solution was computed with SPEC-INT scheme on a fixed grid with 12800 subintervals at times  $t = 400$  s and  $t = 2500$  s. The CPU time used with adaptive technique is roughly a tenth of the CPU time required for the equivalent uniform grid. This CPU time can further be decreased when using more levels. We plotted different choices for the threshold value  $\tau_p$  and observe that  $\tau_p = 10^{-2}$  is the most efficient choice.

## 1.5 Conclusions of Chapter 1

We have applied an adaptive mesh refinement algorithm to save computational resources in simulations of polydisperse sedimentation by two state-of-the-art high resolution shock capturing techniques. In our experiments with 4 and 7 species and a given CPU time, the scheme that utilizes the characteristic information obtained through the use of the secular equation is more efficient than the scheme that makes no use of this information.

We are well aware that although one-dimensional kinematic models such as (1.1) are widely used in engineering and other fields where besides the vertical segregation the composition of the final deposit is of interest, many relevant applications involve flow fields that are lateral to the direction of gravity and can be described by spatially two- or three-dimensional models only. For instance, an interesting phenomenon of that kind is the settling of a suspension beneath inclined walls [44]. However, the extension of the present model to multiple space dimensions does not only give rise to a multi-dimensional analogue of (1.1); rather, additional equations of motion for the flow field of the mixture need to be solved (e.g., a variant of the Stokes system [12]). In one space dimension, this flow field is determined by boundary conditions and vanishes for batch settling in a column. The development of adaptive techniques to save computational effort for such multi-dimensional “coupled” sedimentation-flow models is topic of some of the authors’ current research (see e.g. [33] for first results for the monodisperse case  $N = 1$ ).

In view of the additional equations of motion, the computational difficulties for these models go beyond applying, say, AMR just to a multi-dimensional version of (1.1), so the treatment of two- or three-dimensional flow models is not within the scope of the paper. In fact, already in one space dimension AMR gives rise to considerable speedup. Finally, we mention that models of segregation of polydisperse dry granular materials (see e.g. [97, 108] and papers cited in these works) address similar physical phenomena, but are constructed in a different way than the MLB model. Although these models do not require the solution of additional flow equations in multiple space dimensions (in contrast to the sedimentation model), they do not give rise to strongly coupled systems of conservation laws (as does the sedimentation model), and therefore do not form a suitable case of example for spectral WENO schemes, and their AMR refinement.



## Chapter 2

# Regularized nonlinear solvers for Implicit-Explicit methods

### 2.1 Introduction

#### 2.1.1 Scope

In this chapter we focus on multi-species flow models defined by equation (1.1) in Chapter 1 in which the velocities also depend on the spatial variation of  $\Phi$  to account for additional effects such as sediment compressibility or drivers' reaction time and anticipation length in traffic flow. These corrections can be usually posed in such a way that the resulting system of partial differential equations (PDEs) has an extra, possibly strongly degenerate diffusive term. We therefore consider systems of PDEs of the type

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = \partial_x (\mathbf{B}(\Phi) \partial_x \Phi), \quad (2.1)$$

where  $\mathbf{B}(\Phi)$  is a given  $N \times N$  matrix function expressing the diffusive correction. The system (2.1) is supplied with an initial condition and depending on the application, zero-flux or periodic boundary conditions.

Although the available mathematical theory does not allow us to be conclusive about the existence, uniqueness and well-posedness of the solutions of such strongly degenerate hyperbolic-parabolic systems, it is plausible to perform simulations with appropriate numerical methods. Explicit schemes for hyperbolic systems of first-order conservation laws are widely used in many applications. Although they can be rather slow for some steady-state computations, due to CFL stability restrictions on the time step size, their use for unsteady computations is deemed as practical in many situations. This does not hold when diffusion terms are present. However, one can resort to an implicit treatment of these terms to overcome the drastic step size restrictions imposed by the stability condition for explicit schemes applied to parabolic equations.

The purpose of this chapter is to design nonlinear solvers for suitable regularizations of the systems that appear when using implicit-explicit (IMEX) schemes for the efficient solution of initial-boundary value problems for (2.1) under the specific assumptions of diffusively corrected kinematic

flow models. Moreover, we demonstrate the overall efficiency of these schemes when compared with schemes that are explicit or based on other nonlinear solvers. The specific assumptions, which are reflected in the design of the numerical schemes and in our analysis, include that the number  $N$  of species (and therefore of scalar equations) may be arbitrarily large; that the flux vector  $\mathbf{f}(\Phi)$  is constructed in a systematic way that makes characteristic-wise schemes applicable to the convective part of (2.1) (even though the eigenstructure of the flux Jacobian  $\mathcal{J}_{\mathbf{f}}(\Phi)$  is not available in closed algebraic form); and that (2.1) is often strongly degenerate, where the location of the type-change interface is unknown beforehand and  $\mathbf{B}$  may even be discontinuous as a function of  $\Phi$ . We focus on a model of sedimentation of polydisperse suspensions forming compressible sediment layers, and a diffusively corrected multi-class Lighthill-Whitham-Richards (LWR) model for vehicular traffic that includes anticipation length and reaction time.

The main novelty of this chapter is the particular method of solution of the nonlinear systems that appear with the implicit treatment of the degenerate diffusion term. This method is based on the Newton-Raphson method applied after regularizing the non-smooth diffusion coefficient. The final schemes are much more efficient, in term of error reduction versus CPU time, than the explicit schemes.

### 2.1.2 Related work

First-order models of the type (1.1) were widely studied in recent years, with an emphasis on polydisperse sedimentation [12, 22, 23, 52, 115] and multiclass vehicular traffic [8, 50, 85, 113, 117, 118, 120]. Other applications include the settling and creaming of emulsions and dispersions [54, 94]. Among the polydisperse sedimentation models, one of the most widely used velocity model is the Masliyah-Lockett-Bassoon (MLB) model [77, 79]. We refer to [22, 116] for alternate velocity models. On the other hand, the multi-class extension of the LWR model [75, 92], the MCLWR model, was introduced by Benzoni-Gavage and Colombo [8] and Wong and Wong [113]. All these models have in common that although the functions  $v_i$  are constructed in a systematic manner, the eigenvectors and eigenvalues of the Jacobian  $\mathcal{J}_{\mathbf{f}}(\Phi) = (\partial f_i(\Phi)/\partial \phi_j)_{1 \leq i, j \leq N}$  are usually not available in closed algebraic form. It is in general difficult to estimate the subregion of  $N$ -dimensional phase space where (1.1) is strictly hyperbolic, i.e.,  $\mathcal{J}_{\mathbf{f}}(\Phi)$  has pairwise distinct real eigenvalues, or to solve the Riemann problem for (1.1) exactly or approximately. However, for some of these models, the functions  $v_i$  depend on a small number of independent scalar functions of  $\Phi$  only, so that  $\mathcal{J}_{\mathbf{f}}(\Phi)$  is a low-rank perturbation of a diagonal matrix. In this case, the calculus of the so-called secular equation, advanced first by Anderson [3], allows one to establish that, under determined circumstances, the eigenvalues of  $\mathcal{J}_{\mathbf{f}}(\Phi)$  are real and interlace with the velocities  $v_i$  [22, 51]. This information provides starting values to determine the exact eigenvalues by a root finder, and eventually to determine the corresponding eigenvectors. This has led to the construction of involved but efficient characteristic-wise weighted essentially non-oscillatory (WENO) schemes [66, 76, 98, 99] for (1.1) [23, 50]. These schemes are employed herein to discretize the convective part of (2.1).

For models of polydisperse sedimentation, diffusive terms leading to the form (2.1) were first

proposed by Stamatakis and Tien [106]. A theory of sedimentation of polydisperse suspensions forming compressible sediments was advanced in [12], where the system (2.1) was solved by the Kurganov-Tadmor (KT) explicit high-resolution central difference scheme [71]. Its application to strongly degenerate convection-diffusion systems is explicitly proposed in [71, Sect. 4.2]. On the other hand, the multi-class version of the diffusively corrected LWR model proposed by Nelson [81], which can also be understood as a diffusively corrected version of the MCLWR traffic model [8, 113], is newly derived herein. We also mention that in a very recent paper, Abeynaïke et al. [1] propose a model for the sedimentation and creaming of size-distributed droplets in glycerol/biodiesel dispersions that is equivalent to (2.1).

An IMEX Runge-Kutta scheme consists in combining a Runge-Kutta scheme with an implicit discretization of the diffusive term with an explicit one for the convective term. To introduce the main idea, we consider the problem

$$\partial_t \Phi = C(\Phi) + D(\Phi), \quad (2.2)$$

where  $C(\Phi)$  and  $D(\Phi)$  are discretizations of the convective and diffusive terms, respectively. The stability restriction on the time step  $\Delta t$  that explicit schemes impose when applied to (2.2) is very severe ( $\Delta t$  must be proportional to the square  $\Delta x^2$  of the grid spacing), due to the presence of  $D(\Phi)$ . The implicit treatment of both  $C(\Phi)$  and  $D(\Phi)$  would remove any stability restriction on  $\Delta t$ , but the upwind nonlinear discretization of  $C(\Phi)$  that is needed for stability makes its implicit treatment extremely involved. In fact, after the pioneering work of Crouzeix [42], numerical integrators that deal implicitly with  $D(\Phi)$  and explicitly with  $C(\Phi)$  can be used with a time step restriction dictated by the convective term alone. These schemes, apart of having been profusely used in convection-diffusion problems and convection problems with stiff reaction terms (see [4, 49] and references therein), have been recently used to deal with stiff terms in hyperbolic systems with relaxation (see [17, 18, 19, 20, 89]). Finally, we mention that many authors have proposed IMEX Runge-Kutta schemes for the solution of semi-discretized partial differential equations [4, 68, 89, 122].

## 2.2 Diffusively corrected multi-species kinematic flow models

### 2.2.1 Polydisperse sedimentation

We consider a model of sedimentation of a suspension of equal-density particles belonging to  $N$  species with sizes  $d_1 > d_2 > \dots > d_N$ . We let  $\phi_i$  denote the local volume fraction of species  $i$  having size  $d_i$ , and define  $\phi := \phi_1 + \dots + \phi_N$ . The evolution of  $\Phi = \Phi(x, t)$  as a function of depth  $x$  and time  $t$  in a column is then governed by the combined effects of hindered settling and sediment compressibility. These effects determine the convective and diffusive parts, respectively, of the following system of convection-diffusion equations (see [12]):

$$\partial_t \phi_i + \partial_x f_i(\Phi) = \partial_x (a_i(\Phi, \partial_x \Phi)), \quad i = 1, \dots, N, \quad 0 < x < K, \quad t > 0, \quad (2.3)$$

this phenomena is described in Section 2.6, and is supplemented by the initial condition  $\Phi(x, 0) = \Phi_0(x)$  for  $0 \leq x \leq K$ , where  $\Phi_0$  is the given initial concentration distribution, and zero-flux

boundary conditions corresponding to settling in a closed column of height  $K$ , i.e.,

$$\phi_i v_i = f_i(\Phi) - a_i(\Phi, \partial_x \Phi) = 0 \quad \text{for } x = 0 \text{ and } x = K, \quad t > 0. \quad (2.4)$$

Here the flux density functions  $f_1, \dots, f_N$  are those of the MLB model described in Section 1.2.1 are given by (1.4) with hindered settling function  $V(\phi)$  (1.3).

The diffusion functions on the right-hand side of (2.3) are given by

$$a_i(\Phi, \partial_x \Phi) = \alpha_{i,1}(\Phi) \partial_x \phi_1 + \dots + \alpha_{N,1}(\Phi) \partial_x \phi_N, \quad i = 1, \dots, N,$$

where

$$\alpha_{ij} := \frac{\mu V(\phi)}{g\phi} \left\{ (1 - \phi) \phi_i (\delta_i - \delta^T \Phi) \sigma'_e(\phi) - \left[ \delta_i \delta_{ij} - \delta_j \phi_i - \frac{\phi_i}{\phi} (\delta_i - \delta^T \Phi) \right] \sigma_e(\phi) \right\}, \quad i, j = 1, \dots, N, \quad (2.5)$$

where  $\delta_{ij}$  is the standard Kronecker symbol. Here  $\sigma_e$  denotes the effective solid stress function, and  $\sigma'_e$  is its derivative. This function is assumed to satisfy

$$\sigma_e(\phi), \sigma'_e(\phi) \begin{cases} = 0 & \text{for } \phi \leq \phi_c, \\ > 0 & \text{for } \phi > \phi_c, \end{cases} \quad (2.6)$$

where  $\phi_c$  is a critical concentration at which the particles touch each other. A typical function  $\sigma_e$  having these properties is given by

$$\sigma_e(\phi) = \begin{cases} 0 & \text{for } \phi \leq \phi_c, \\ \sigma_0 \left( (\phi/\phi_c)^k - 1 \right) & \text{for } \phi > \phi_c, \end{cases} \quad \sigma_0, k > 0. \quad (2.7)$$

Defining the matrix  $\mathbf{B}(\Phi) := (\alpha_{ij})_{1 \leq i, j \leq N}$  and  $\mathbf{f}(\Phi) = (f_1(\Phi), \dots, f_N(\Phi))$  we can rewrite (2.3) in the form (2.1).

## 2.2.2 Hyperbolicity and parabolicity analysis for the polydisperse sedimentation model

For the flux  $\mathbf{f}(\Phi)$ ,  $\mathcal{J}_{\mathbf{f}}(\Phi)$  is a rank-two perturbation of a diagonal matrix. This property allows one to analyze hyperbolicity, to localize eigenvalues, and to eventually calculate the corresponding eigenvectors of  $\mathcal{J}_{\mathbf{f}}(\Phi)$ , by using the so-called secular equation [3], see [22, 51]. This hyperbolicity analysis was shown in Section 1.2.2. With respect to the diffusion matrix  $\mathbf{B}(\Phi)$ , in [12] it is proved that its eigenvalues are positive and pairwise distinct on  $\mathcal{D}_{\phi_{\max}}^0 \setminus \mathcal{D}_{\phi_c}$  by evaluating the characteristic polynomial in a fashion similar to that used for the eigenvalues of  $\mathcal{J}_{\mathbf{f}}(\Phi)$ .

**Theorem 2.1 ([12])** *Let  $G(\phi) := \phi(1 - \phi)^2 \sigma'_e(\phi) - \sigma_e(\phi)$ ,  $W(\phi) := \mu V(\phi)/(g\phi)$  and assume that  $V(\phi) \neq 0$  for  $\phi < \phi_{\max}$  and  $V(\phi) = 0$  otherwise. Then, for all  $\Phi \in \mathcal{D}_{\phi_{\max}}^0 \setminus \mathcal{D}_{\phi_c}$  the matrix  $\mathbf{B}(\Phi)$  has  $N$  distinct positive eigenvalues  $\Lambda_1, \dots, \Lambda_N$ ; i.e., the system (2.1) is strictly parabolic on  $\Phi \in \mathcal{D}_{\phi_{\max}}^0 \setminus \mathcal{D}_{\phi_c}$ . Moreover, we have the following interlacing properties, where for brevity we write  $W = W(\phi)$  and  $\sigma_e = \sigma_e(\phi)$ :*

1. If  $\Phi$  is chosen such that  $G(\phi) > 0$ , then these eigenvalues satisfy

$$0 < W\sigma_e\delta_N < \Lambda_N < W\sigma_e\delta_{N-1} < \cdots < W\sigma_e\delta_1 < \Lambda_N < W\delta_1\phi(1-\phi)^2\sigma'_e(\phi).$$

2. At those points  $\Phi$  where  $G(\phi) < 0$ , we have

$$0 < W\sigma_e\phi(1-\phi)^2\delta_N < \Lambda_N < W\sigma_e\delta_{N-1} < \cdots < W\sigma_e\delta_1 < \Lambda_N < W\delta_1\sigma'_e(\phi).$$

3. If  $G(\phi) = 0$ , then the eigenvalues are given by  $\Lambda_i = W\sigma_e\delta_i$  for  $i = 1, \dots, N$ .

### 2.2.3 A new diffusively corrected MCLWR model (DCMCLWR model)

In this subsection we consider a multi-class version of the diffusively corrected kinematic traffic flow model introduced in [81] (see also [25, 82]) for  $N = 1$ . This diffusively corrected multi-class LWR model (“DCMCLWR model”) will be derived and analyzed in Chapter 3. Assume now that  $\phi_i$ ,  $i = 1, \dots, N$ , is the density of vehicles of class  $i$  having the preferential velocity  $v_i^{\max}$ , where

$$v_1^{\max} > v_2^{\max} > \cdots > v_N^{\max} > 0. \quad (2.8)$$

According to the MCLWR model [8, 113], the local velocity  $v_i$  of vehicles of species  $i$  is given by  $v_i = v_i^{\max}V(\phi)$ , where  $v_i^{\max}$  is the preferential velocity of drivers of species  $i$  on a free highway, as usual,  $\phi = \phi_1 + \cdots + \phi_N$ , and  $V$  is a non-increasing function satisfying  $V(0) = 1$ ,  $V(\phi_{\max}) = 0$ , and  $V'(\phi) \leq 0$  for  $0 \leq \phi \leq \phi_{\max}$ . Thus, the standard MCLWR model (without diffusive correction) is given by the equation (1.1), where

$$f_i(\Phi) = \phi_i v_i^{\max} V(\phi), \quad i = 1, \dots, N. \quad (2.9)$$

Let us now assume that the behavior of drivers of species  $i$  is associated with an anticipation distance  $L_i$  and a reaction time  $\tau_i$ ,  $i = 1, \dots, N$ . Then, following the reasoning in [25], the reaction of the driver does not depend on the value of  $\phi$  seen at the point  $\phi(x, t)$ , but rather on  $p_i(x, t) := \phi(x + L_i - v_i^{\max}V\tau_i, t - \tau_i)$ . This formulation takes into account that  $v_i^{\max}V\tau_i$  is the distance travelled by a car of species  $i$  in a time interval of length  $\tau_i$ . (Note that notation is ambiguous here, since we are not specific about the argument of  $V$ , cf. [25].) To obtain a usable expression for the flux  $f_i$ , we expand  $V(p_i(x, t))$  around  $\phi(x, t)$ . Writing  $\phi = \phi(x, t)$  and denoting  $\tau := \max\{\tau_1, \dots, \tau_N\}$ ,  $L := \max\{L_1, \dots, L_N\}$ , we obtain

$$V(p_i(x, t)) = V(\phi) + V'(\phi)[\partial_x\phi(L_i - v_i^{\max}V(\phi)\tau_i) - \tau_i\partial_t\phi] + \mathcal{O}(\tau^2 + L^2). \quad (2.10)$$

On the other hand, summing the conservation laws  $\partial_t\phi_i + \partial_x(v_i^{\max}\phi_iV(\phi)) = 0$  over  $i = 1, \dots, N$  and defining  $\mathbf{v}^{\max} := (v_1^{\max}, \dots, v_N^{\max})^T$ , we get

$$\partial_t\phi = \partial_t\phi_1 + \cdots + \partial_t\phi_N = -\partial_x(V(\phi)(\mathbf{v}^{\max})^T\Phi).$$

Inserting this result into (2.10) we get

$$V(p_i(x, t)) = V(\phi) + V'(\phi)[(L_i - \tau_i v_i^{\max}V(\phi))\partial_x\phi + \tau_i\partial_x(V(\phi)(\mathbf{v}^{\max})^T\Phi)] + \mathcal{O}(\tau^2 + L^2).$$

Neglecting the  $\mathcal{O}(\tau^2 + L^2)$  term and inserting the result into the conservation equations

$$\partial_t \phi_i(x, t) + \partial_x (\phi_i(x, t) v_i(x, t)) = 0, \quad v_i(x, t) = v_i^{\max} V(p_i(x, t)), \quad i = 1, \dots, N,$$

we obtain a system of the form (2.1), where the components of the flux vector  $\mathbf{f}(\Phi)$  are given by (2.9) and the entries of the diffusion matrix  $\mathbf{B}(\Phi)$  are now given by

$$\alpha_{ij}(\Phi) = -V'(\phi) [L_i + \tau_i (V'(\phi) (\mathbf{v}^{\max})^T \Phi + (v_j^{\max} - v_i^{\max}) V(\phi))] \phi_i v_i^{\max}, \quad 1 \leq i, j \leq N. \quad (2.11)$$

For traffic flow models we will use periodic boundary conditions corresponding to a circular road of length  $K$ , namely

$$\Phi(0, t) = \Phi(K, t), \quad t > 0. \quad (2.12)$$

## 2.2.4 Hyperbolicity and parabolicity analysis for the diffusively corrected MCLWR model

According to [51], the MCLWR model (2.1), (2.9) with  $\mathbf{B} \equiv 0$  is strictly hyperbolic and the Jacobian  $\mathcal{J}_{\mathbf{f}}(\Phi)$  is a rank-one perturbation of a diagonal matrix. The eigenstructure of  $\mathcal{J}_{\mathbf{f}}(\Phi)$  can again be computed via the secular equation [3], as is explained in the following version of Theorem 1.1 Section 1.2.2.

**Theorem 2.2 ([51])** *Consider the first-order multiclass kinematic traffic flow model (1.1), (2.9) (i.e., without diffusive terms) and assume that the velocities  $v_i^{\max}$  are ordered according to (2.8). If  $\Phi \in \mathcal{D}_1^0$ , then the Jacobian  $\mathcal{J}_{\mathbf{f}}(\Phi)$  has  $N$  distinct real eigenvalues  $\lambda_1, \dots, \lambda_N$  which are the roots of the secular equation (1.6) with  $\gamma_i = v_i^{\max} \phi_i V'(\phi)$ , and the following interlacing property holds:*

$$v_N^{\max} + V'(\phi) (\mathbf{v}^{\max})^T \Phi < \lambda_N < v_N^{\max} < \lambda_{N-1} < v_{N-1}^{\max} < \dots < v_2^{\max} < \lambda_1 < v_1^{\max}.$$

We now wish to state sufficient conditions on the non-negative parameters  $v_i^{\max}$ ,  $\tau_i$  and  $L_i$  under which  $\mathbf{B}(\Phi)$  has eigenvalues with positive real part for all  $\Phi \in \mathcal{D}_{\phi_{\max}}$ . The latter property will only hold under restrictions on the parameters  $L_i$  and  $\tau_i$ . In fact, already in the case  $N = 1$ , where  $B(\phi) = -V'(\phi)(L + \tau v_{\max} \phi V'(\phi)) \phi v_{\max}$ , and considering that  $V'(\phi) \leq 0$ , we get that  $B(\phi) \geq 0$  for all  $0 \leq \phi \leq \phi_{\max}$  if and only if

$$\phi V'(\phi) \geq -\frac{L}{\tau v_{\max}} \quad \text{for all } 0 \leq \phi \leq \phi_{\max}. \quad (2.13)$$

Thus, we cannot expect  $\mathbf{B}(\Phi)$  to have non-negative eigenvalues only without further limitations and structural conditions between the parameters  $v_i^{\max}$ ,  $L_i$  and  $\tau_i$ .

**Theorem 2.3** *The eigenvalues  $\mu_1, \dots, \mu_N$  of  $\mathbf{B}(\Phi)$  are given by  $\mu_i = -V'(\phi) \lambda_i$ ,  $i = 1, \dots, N$ , where*

$$\lambda_1 = \frac{C_1}{2} - \left( \frac{C_1^2}{4} - C_2 \right)^{1/2}, \quad \lambda_2 = \frac{C_1}{2} + \left( \frac{C_1^2}{4} - C_2 \right)^{1/2}, \quad \lambda_3 = \dots = \lambda_N = 0,$$

where we have

$$\begin{aligned} C_1 &= \sum_{k=1}^N \phi_k v_k^{\max} (L_k + \tau_k V'(\phi) (\mathbf{v}^{\max})^T \Phi), \\ C_2 &= \sum_{\substack{i,j=1 \\ i < j}}^N \phi_i v_i^{\max} \phi_j v_j^{\max} \tau_i \tau_j \left( \frac{L_i}{\tau_i} - \frac{L_j}{\tau_j} + (v_j^{\max} - v_i^{\max}) V(\phi) \right) (v_j^{\max} - v_i^{\max}) V(\phi). \end{aligned} \quad (2.14)$$

**Proof.** We have  $\mathbf{B}(\Phi) = -V'(\phi) \tilde{\mathbf{B}}(\Phi)$ , where  $\tilde{\mathbf{B}}(\Phi) = (\tilde{\alpha}_{ij}(\Phi))_{1 \leq i, j \leq N}$  is defined in an obvious manner via (2.11). (Clearly, since  $V'(\phi) \leq 0$ ,  $\mathbf{B}(\Phi)$  has non-negative eigenvalues on  $\mathcal{D}_{\phi_{\max}}$  if  $\tilde{\mathbf{B}}(\Phi)$  has.) Since  $\tilde{\mathbf{B}}(\Phi)$  is a rank-2 matrix of size  $N \times N$ , we know that  $\det(\lambda \mathbf{I} - \tilde{\mathbf{B}}(\Phi)) = \lambda^{N-2} (\lambda^2 - C_1 \lambda + C_2)$ , where  $C_k = C_k(\Phi)$  is the sum of the  $k$ -rowed principal minors of  $\tilde{\mathbf{B}}(\Phi)$ , that is,

$$C_1 = \text{tr } \tilde{\mathbf{B}}(\Phi), \quad C_2 = \sum_{\substack{i,j=1 \\ i < j}}^N (\tilde{\alpha}_{ii} \tilde{\alpha}_{jj} - \tilde{\alpha}_{ji} \tilde{\alpha}_{ij}). \quad (2.15)$$

From (2.11) we get that  $\tilde{\alpha}_{ij} = (L_i + \tau_i V'(\phi) (\mathbf{v}^{\max})^T \Phi + \tau_i (v_j^{\max} - v_i^{\max}) V(\Phi)) \phi_i v_i^{\max}$  for  $i, j = 1, \dots, N$ . Evaluating (2.15) then yields (2.14).  $\blacksquare$

We recall that the system (2.1) is called *parabolic* at a state  $\Phi_0$  if the eigenvalues of  $\mathbf{B}(\Phi_0)$  have non-negative real parts. This is precisely the case if  $C_1(\Phi_0) \geq 0$  and  $C_2(\Phi_0) \geq 0$ . Thus, we can expect the system to be well-posed only if

$$C_1(\Phi) > 0, \quad C_2(\Phi) > 0 \quad \text{on } \mathcal{D}_{\phi_{\max}}^0. \quad (2.16)$$

In view of  $V'(\phi) (\mathbf{v}^{\max})^T \Phi \leq 0$ , a sufficient condition for  $C_1(\Phi) \geq 0$  to hold is that

$$L_k(\phi) + \tau_k v_1^{\max} \phi V'(\phi) \leq 0 \quad \text{for } 0 \leq \phi \leq \phi_{\max}, \quad k = 1, \dots, N. \quad (2.17)$$

Note that this condition is the multi-class ( $N \geq 1$ ) extension of (2.13). Furthermore, note that we can write (2.17) as

$$\tau_k \leq \min_{0 \leq \phi \leq \phi_{\max}} - \frac{L_k(\phi)}{\phi V'(\phi) v_1^{\max}}. \quad (2.18)$$

If (2.18) is violated, that is, when the reaction time of a driver is not sufficiently small, then the model is likely to exhibit anti-diffusive phenomena such as formation of clusters, steep density gradients, stop-and-go waves, and other instability phenomena. A similar conclusion (though based on a slightly different model) has been drawn, for example, in [87]. For the diffusively corrected MCLWR model we expand on this observation, and more closely analyze instability phenomena chapter 3 and in [30].

In the case  $N = 1$ , Nelson [81] (cf. [25, 82]) suggests to employ

$$L = L(\phi) = \max \left\{ \frac{(v_{\max} V(\phi))^2}{2a}, L_{\min} \right\}, \quad (2.19)$$

where the first argument is the distance required to decelerate to full stop from speed  $v_{\max}V(\phi)$  at deceleration  $a$ , and the second is a minimal anticipation distance  $L_{\min} > 0$  regardless of how small the velocity is. In the multi-class case we could define  $L_i$ , for instance, by (2.19) with  $v_{\max}$  replaced by  $v_i^{\max}$ . However, in our numerical experiments, we select  $L_i$  and  $\tau_i$  constant to ensure that (2.16) is always satisfied.

## 2.3 Numerical schemes

### 2.3.1 Spatial discretization

For grid points  $x_j := (j - \frac{1}{2})\Delta x$  for  $j = 1, \dots, M$ , where  $\Delta x := K/M$ , and  $t_n := n\Delta t$  for  $n \in \mathbb{N}_0$ , and using the notation  $\Delta^- \mathbf{g}_k = \mathbf{g}_k - \mathbf{g}_{k-1}$  we discretize (2.1) in space as follows:

$$\frac{d\Phi_j(t)}{dt} = \mathcal{L}_j(\Phi) := -\frac{1}{\Delta x} \Delta^- \mathbf{f}_{j+1/2} + \frac{1}{\Delta x} \Delta^- \mathbf{g}_{j+1/2}, \quad j = 1, \dots, M, \quad (2.20)$$

where  $\Phi_j(t) \approx \Phi(x_j, t)$ , the convective numerical flux  $\mathbf{f}_{j+1/2} := \mathbf{f}(\Phi_{j-2}, \dots, \Phi_{j+3})$  is calculated by the characteristic-wise fifth-order WENO approximation see Section 1.3.1 and [23], and the flux corresponding to the parabolic term is given by

$$\mathbf{B}(\Phi) \partial_x \Phi|_{x=x_{j+1/2}} \approx \mathbf{g}_{j+1/2} := \frac{1}{2\Delta x} (\mathbf{B}(\Phi_{j+1}) + \mathbf{B}(\Phi_j)) \Delta^- \Phi_{j+1},$$

which gives a second-order approximation for the diffusive term. (Higher order approximations of these terms could be used to match the order of approximation of the convective term, but we will not pursue this herein.)

The operator  $\mathcal{L} := (\mathcal{L}_1, \dots, \mathcal{L}_M)^\top$  appearing in (2.20) is now given by

$$\mathcal{L}_j(\Phi) = -\frac{1}{\Delta x} \Delta^- \mathbf{f}_{j+1/2} + \frac{1}{2\Delta x^2} \Delta^- \left( (\mathbf{B}(\Phi_{j+1}) + \mathbf{B}(\Phi_j)) \Delta^- \Phi_{j+1} \right)$$

for  $j = 2, \dots, M-1$ , along with appropriate modifications of this formula for  $j = 1$  and  $j = M$  to account for boundary conditions. This can be further written as follows:

$$\mathcal{L}(\Phi) = -\frac{1}{\Delta x} (\Delta^- \mathbf{f})(\Phi) + \frac{1}{\Delta x^2} \mathcal{B}(\Phi) \Phi, \quad (2.21)$$

where  $\mathcal{B}(\mathbf{v}) = \{\mathcal{B}_{ij}(\mathbf{v})\}_{i,j=1,\dots,M} \in \mathbb{R}^{(NM) \times (NM)}$  is a block tridiagonal matrix formed by blocks  $\mathcal{B}_{ij} \in \mathbb{R}^{N \times N}$  generally given by

$$\begin{aligned} \mathcal{B}_{ii}(\mathbf{v}) &= \frac{1}{2\Delta x^2} (\mathbf{B}(\mathbf{v}_{i+1}) + 2\mathbf{B}(\mathbf{v}_i) + \mathbf{B}(\mathbf{v}_{i-1})), \quad i = 1, \dots, M, \\ \mathcal{B}_{i,i-1}(\mathbf{v}) &= \mathcal{B}_{i-1,i}(\mathbf{v}) = -\frac{1}{2\Delta x^2} (\mathbf{B}(\mathbf{v}_{i-1}) + \mathbf{B}(\mathbf{v}_i)), \quad i = 2, \dots, M. \end{aligned} \quad (2.22)$$

### 2.3.2 Boundary conditions

For the polydisperse sedimentation model we discretize the zero-flux boundary conditions (2.4) by setting  $\mathbf{f}_{1/2} - \mathbf{g}_{1/2} = \mathbf{0}$  and  $\mathbf{f}_{M+1/2} - \mathbf{g}_{M+1/2} = \mathbf{0}$ . This affects  $\mathcal{L}_1$  and  $\mathcal{L}_M$ , which now read as

$$\begin{aligned}\mathcal{L}_1(\Phi) &= -\frac{1}{\Delta x} \mathbf{f}_{3/2} + \frac{1}{2\Delta x^2} (\mathbf{B}(\Phi_2) + \mathbf{B}(\Phi_1)) \Delta^- \Phi_2, \\ \mathcal{L}_M(\Phi) &= \frac{1}{\Delta x} \mathbf{f}_{M-1/2} - \frac{1}{2\Delta x^2} (\mathbf{B}(\Phi_{M-1}) + \mathbf{B}(\Phi_M)) \Delta^- \Phi_M.\end{aligned}$$

This can be written as (2.21) with

$$\mathbf{B}_{11}(\mathbf{v}) = \frac{1}{2\Delta x^2} (\mathbf{B}(\mathbf{v}_2) + \mathbf{B}(\mathbf{v}_1)), \quad \mathbf{B}_{MM}(\mathbf{v}) = \frac{1}{2\Delta x^2} (\mathbf{B}(\mathbf{v}_{M-1}) + \mathbf{B}(\mathbf{v}_M)).$$

When we discretize the periodic boundary conditions (2.12), for the discretization of the flux at  $x = 0$  we formally need values  $\Phi_{-j}$  for  $j = 0, 1, 2$ . By periodicity the value  $\Phi_{-j}$  should agree with the value  $\Phi_{M-j}$ . Similarly, at  $x = K$  the value  $\Phi_{M+j}$ ,  $j = 1, 2, 3$ , should agree with  $\Phi_j$ . Therefore, we have the following:

$$\begin{aligned}\mathcal{L}_1(\Phi) &= -\frac{1}{\Delta x} \Delta^- \mathbf{f}_{3/2} + \frac{1}{2\Delta x^2} \left( (\mathbf{B}(\Phi_2) + \mathbf{B}(\Phi_1)) \Delta^- \Phi_2 \right. \\ &\quad \left. - (\mathbf{B}(\Phi_1) + \mathbf{B}(\Phi_M)) (\Phi_1 - \Phi_M) \right), \\ \mathcal{L}_M(\Phi) &= -\frac{1}{\Delta x} \Delta^- \mathbf{f}_{M+1/2} + \frac{1}{2\Delta x^2} \left( (\mathbf{B}(\Phi_M) + \mathbf{B}(\Phi_1)) (\Phi_1 - \Phi_M) \right. \\ &\quad \left. - (\mathbf{B}(\Phi_{M-1}) + \mathbf{B}(\Phi_M)) \Delta^- \Phi_M \right).\end{aligned}$$

The blocks in the first and last rows of blocks of matrix  $\mathbf{B}$  in (2.22) that should be modified with respect to the general definition are

$$\begin{aligned}\mathbf{B}_{11}(\mathbf{v}) &= \frac{1}{2\Delta x^2} (\mathbf{B}(\mathbf{v}_2) + 2\mathbf{B}(\mathbf{v}_1) + \mathbf{B}(\mathbf{v}_M)), \\ \mathbf{B}_{MM}(\mathbf{v}) &= \frac{1}{2\Delta x^2} (\mathbf{B}(\mathbf{v}_1) + 2\mathbf{B}(\mathbf{v}_M) + \mathbf{B}(\mathbf{v}_{M-1})), \\ \mathbf{B}_{1,M}(\mathbf{v}) &= -\frac{1}{2\Delta x^2} (\mathbf{B}(\mathbf{v}_1) + \mathbf{B}(\mathbf{v}_M)), \quad \mathbf{B}_{M,1}(\mathbf{v}) = -\frac{1}{2\Delta x^2} (\mathbf{B}(\mathbf{v}_1) + \mathbf{B}(\mathbf{v}_M)).\end{aligned}$$

Therefore, the block structure of  $\mathbf{B}$  turns out to be *circulant* tridiagonal.

### 2.3.3 Explicit schemes

Given an approximation  $\Phi^n = (\Phi_1^n, \dots, \Phi_M^n)^\top$  for  $t = t_n$ , we can compute an approximation  $\Phi^{n+1} = (\Phi_1^{n+1}, \dots, \Phi_M^{n+1})^\top$  for  $t = t_{n+1}$  from (2.20) by using an ODE solver, such as Euler's method or third-order TVD Runge-Kutta method (see [57, 58, 100]). For instance, Euler's method can be written as

$$\Phi^{n+1} = \Phi^n - \frac{\Delta t}{\Delta x} (\Delta^- \mathbf{f})(\Phi^n) + \frac{\Delta t}{\Delta x^2} \mathbf{B}(\Phi^n) \Phi^n. \quad (2.23)$$

Other explicit schemes, such as the KT scheme [71] that we use in our numerical experiments, have a similar formulation.

A von Neumann analysis of the stability of these explicit schemes applied to suitable linearizations about constant states would suggest that

$$\frac{\Delta t}{\Delta x} \max_{\Phi} \rho(\mathcal{J}_f(\Phi)) + \frac{\Delta t}{2\Delta x^2} \max_{\Phi} \rho(\mathbf{B}(\Phi)) \leq C_{\text{cfl1}} \leq 1 \quad (2.24)$$

is an appropriate CFL stability condition, where  $\rho(\cdot)$  is the spectral radius. The constant  $C_{\text{cfl1}}$  depends on the method and should be empirically adjusted for nonlinear problems, since (2.24) is deduced for linearized problems and schemes.

### 2.3.4 Implicit-explicit schemes

The CFL stability condition (2.24) restricts the time step size dramatically when  $\mathbf{B}(\Phi) \neq 0$ . This restriction could be overcome by implicit schemes, but the implicit treatment of the convective term is complicated due to the highly nonlinear scheme that is used for its discretization. Therefore, implicit-explicit (IMEX) schemes, which treat the diffusive term implicitly and the convective term explicitly, could be an attractive alternative in this situation. The simplest IMEX scheme for the approximation of (2.20) is the following version of (2.23):

$$\Phi^{n+1} = \Phi^n - \frac{\Delta t}{\Delta x} (\Delta^- \mathbf{f})(\Phi^n) + \frac{\Delta t}{\Delta x^2} \mathbf{B}(\Phi^{n+1}) \Phi^{n+1}. \quad (2.25)$$

The CFL condition for IMEX schemes is

$$\frac{\Delta t}{\Delta x} \max_{\Phi} \rho(\mathcal{J}_f(\Phi)) \leq C_{\text{cfl2}} \leq 1,$$

which is much less restrictive than (2.24). Here, as mentioned above, the constant  $C_{\text{cfl2}}$  depends on the method used for the spatial and temporal discretizations.

For the case of zero-flux boundary conditions, the boundary condition at  $x = 0$  for the scheme (2.25) should be  $\mathbf{f}_{1/2}^k - \mathbf{g}_{1/2}^k = \mathbf{0}$  for  $k = n$  or  $k = n + 1$ . The use of different treatments for the convective and diffusive parts does not allow us to impose this condition in a natural manner as for explicit schemes. However, we impose this condition to avoid adding ghost cells and changing the structure of the matrices, but this treatment does generate a loss of precision at the boundary. In this manner we use for  $j = 1$  the equation

$$\Phi_1^{n+1} = \Phi_1^n - \frac{\Delta t}{\Delta x} \mathbf{f}_{3/2}^n + \frac{\Delta t}{\Delta x^2} (\mathbf{B}(\Phi_2^{n+1}) + \mathbf{B}(\Phi_1^{n+1})) \Delta^- \Phi_2^{n+1}.$$

The boundary condition at  $x = K$  is treated in a similar way. For periodic boundary conditions (cf. Section 2.3.2), the equations for  $j = 1$  and  $j = M$  in (2.25) are

$$\begin{aligned} \Phi_1^{n+1} &= \Phi_1^n - \frac{\Delta t}{\Delta x} \Delta^- \mathbf{f}_{3/2}^n + \frac{\Delta t}{2\Delta x^2} \left( (\mathbf{B}(\Phi_2^{n+1}) + \mathbf{B}(\Phi_1^{n+1})) \Delta^- \Phi_2^{n+1} \right. \\ &\quad \left. - (\mathbf{B}(\Phi_1^{n+1}) + \mathbf{B}(\Phi_M^{n+1})) (\Phi_1^{n+1} - \Phi_M^{n+1}) \right), \\ \Phi_M^{n+1} &= \Phi_M^n - \frac{\Delta t}{\Delta x} \Delta^- \mathbf{f}_{M+1/2}^n + \frac{\Delta t}{2\Delta x^2} \left( (\mathbf{B}(\Phi_M^{n+1}) + \mathbf{B}(\Phi_1^{n+1})) (\Phi_1^{n+1} - \Phi_M^{n+1}) \right. \\ &\quad \left. - (\mathbf{B}(\Phi_{M-1}^{n+1}) + \mathbf{B}(\Phi_M^{n+1})) \Delta^- \Phi_M^{n+1} \right). \end{aligned}$$

### 2.3.5 IMEX Runge-Kutta schemes

To introduce IMEX-RK methods for the initial-boundary value problems of (2.1) at hand, we basically follow the notation in [4, 89] and rewrite the semi-discrete formulation (2.20) in the form (2.2), where

$$C(\Phi) := -\frac{1}{\Delta x}(\Delta^- \mathbf{f})(\Phi), \quad D(\Phi) := \frac{1}{\Delta x^2} \mathcal{B}(\Phi)\Phi.$$

For the diffusive part we utilize an  $s$ -stage diagonally implicit (DIRK) scheme with coefficients  $\mathbf{A} \in \mathbb{R}^{s \times s}$ ,  $\mathbf{c}, \mathbf{b} \in \mathbb{R}^s$ , in the common Butcher notation, where  $\mathbf{A} = (a_{ij})$  with  $a_{ij} = 0$  for  $j > i$ . For the convective part we employ an  $s$ -stage explicit scheme with coefficients  $\hat{\mathbf{A}} \in \mathbb{R}^{s \times s}$ ,  $\hat{\mathbf{b}}, \hat{\mathbf{c}} \in \mathbb{R}^s$  with  $\hat{\mathbf{A}} = (\hat{a}_{ij})$  with  $\hat{a}_{ij} = 0$  for  $j \geq i$ . The idea is that the resulting scheme is explicit in  $C$  and implicit in  $D$ . The corresponding Butcher arrays are denoted by

$$D := \begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^\top \end{array}, \quad \hat{D} := \begin{array}{c|c} \hat{\mathbf{c}} & \hat{\mathbf{A}} \\ \hline & \hat{\mathbf{b}}^\top \end{array}.$$

The computations of an IMEX-RK scheme necessary to advance an approximate solution  $\Phi^n$  from time  $t^n$  to  $t^{n+1} = t^n + \Delta t$  are given in Algorithm 2.1 (see [89]; this is a version of the algorithm originally proposed in [4, p. 154]):

**Algorithm 2.1 (Implicit-explicit Runge-Kutta (IMEX-RK) scheme)**

*Input:* approximate solution vector  $\Phi^n$  for  $t = t_n$

**do**  $i = 1, \dots, s$

*solve for  $\Phi^{(i)}$  the nonlinear equation*

$$\Phi^{(i)} = \Phi^n + \Delta t \left( \sum_{j=1}^{i-1} a_{ij} K_j + a_{ii} D(\Phi^{(i)}) + \sum_{j=1}^{i-1} \hat{a}_{i,j} \hat{K}_j \right)$$

$$K_i \leftarrow D(\Phi^{(i)})$$

$$\hat{K}_i \leftarrow C(\Phi^{(i)})$$

**enddo**

$$\Phi^{n+1} \leftarrow \Phi^n + \Delta t \sum_{j=1}^s b_j K_j + \Delta t \sum_{j=1}^s \hat{b}_j \hat{K}_j$$

*Output:* approximate solution vector  $\Phi^{n+1}$  for  $t = t_{n+1} = t_n + \Delta t$ .

Algorithm 2.1 requires solving for the vector  $\mathbf{u} = \Phi^{(i)} \in \mathbb{R}^{MN}$  a nonlinear system of  $NM$  scalar equations of the form

$$\mathbf{F}_i(\mathbf{u}) := \mathbf{u} - a_{ii} \Delta t \mathcal{B}(\mathbf{u})\mathbf{u} - \mathbf{r}_i = \mathbf{0}, \quad i = 1, \dots, s, \quad (2.26)$$

where the vector  $\mathbf{r}_i \in \mathbb{R}^{MN}$  is given by

$$\mathbf{r}_i = \Phi^n + \Delta t \sum_{j=1}^{i-1} a_{ij} \mathcal{B}(\Phi^{(j)}) \Phi^{(j)} + \Delta t \sum_{j=1}^{i-1} \hat{a}_{i,j} (\Delta^- \mathbf{f})(\Phi^{(j)}). \quad (2.27)$$

The solution of systems (2.26), (2.27) will be discussed in Section 2.4.

There are different types of IMEX-RK schemes in the literature characterized by the structure of the matrix  $\mathbf{A} \in \mathbb{R}^{s \times s}$ . Following the classification by Boscarino [16], we mention IMEX-RK schemes of so-called type A, described in [89], for which the matrix  $\mathbf{A}$  is invertible. Moreover, IMEX-RK schemes of type CK, described in [68], are characterized by the property that  $\mathbf{A}$  can be written in the form

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ \mathbf{a} & \tilde{\mathbf{A}} \end{bmatrix},$$

where the submatrix  $\tilde{\mathbf{A}} \in \mathbb{R}^{(s-1) \times (s-1)}$  is invertible and  $\mathbf{a} \in \mathbb{R}^{(s-1) \times 1}$ . Finally, IMEX-RK schemes of type ARS, introduced in [4], can now be understood as a special case of schemes of type CK, namely with  $\mathbf{a} = \mathbf{0}$ .

We consider in this chapter the following three schemes which are identified with the notation IMEX-Name( $s, \sigma, p$ ), where this triplet characterizes the number  $s$  of stages of the implicit scheme, the number  $\sigma$  of effective stages of the explicit scheme ( $\sigma = s$  or  $\sigma = s + 1$ ) and the order  $p$  of the method: the scheme IMEX-ARS(1,1,1) (cf. [4]) defined by the pair of Butcher arrays

$$\mathbf{D} = \frac{0 \mid 0 \ 0}{1 \mid 0 \ 1} \quad \hat{\mathbf{D}} = \frac{0 \mid 0 \ 0}{1 \mid 1 \ 0} \quad \frac{\phantom{0} \mid \phantom{0} \ 0}{\phantom{1} \mid \phantom{1} \ 0}, \quad (2.28)$$

the scheme IMEX-ARS(2,2,2) (cf. [4]) defined by

$$\mathbf{D} = \frac{0 \mid 0 \ 0 \ 0}{\gamma \mid 0 \ \gamma \ 0}{1 \mid 0 \ 1 - \gamma \ \gamma} \quad \hat{\mathbf{D}} = \frac{0 \mid 0 \ 0 \ 0}{\gamma \mid \gamma \ 0 \ 0}{1 \mid \delta \ 1 - \delta \ 0} \quad \frac{\phantom{0} \mid \phantom{0} \ 0 \ 0}{\phantom{1} \mid \delta \ 1 - \delta \ 0}, \quad (2.29)$$

where  $\gamma = 1 - \frac{1}{\sqrt{2}}$ ,  $\delta = 1 - \frac{1}{2\gamma}$ ,

and the scheme IMEX-SSP2(3,3,2) introduced in [89], which is defined by

$$\mathbf{D} = \frac{1/4 \mid 1/4 \ 0 \ 0}{1/4 \mid 0 \ 1/4 \ 0}{1 \mid 1/3 \ 1/3 \ 1/3} \quad \hat{\mathbf{D}} = \frac{0 \mid 0 \ 0 \ 0}{1/2 \mid 1/2 \ 0 \ 0}{1 \mid 1/2 \ 1/2 \ 0} \quad \frac{\phantom{0} \mid \phantom{0} \ 0 \ 0}{\phantom{1} \mid 1/3 \ 1/3 \ 1/3} \quad (2.30)$$

Note that scheme IMEX-ARS(1,1,1) corresponds to a combination of the explicit and implicit Euler schemes (2.25).

## 2.4 Nonlinear solvers

Section 2.3.5 shows that IMEX-RK schemes are applicable to (2.1) as long as one can efficiently find the solution of the nonlinear system (2.26), whose existence and uniqueness is guaranteed if  $\Delta t$  is sufficiently small.

Due to the structure of the nonlinearity in (2.26), its solution could be obtained by a lagged diffusivity fixed point iteration that entails solving a convection-diffusion equation with a linear diffusion term at each iteration, and which consists in freezing the diffusion coefficient at the current iteration and solving the resulting linear equation for the next one. Nevertheless, the convergence of such a scheme could require a stringent restriction of  $\Delta t$ . This is just the kind of limitation that we want to avoid by the implicit treatment of the diffusion term. We therefore resort to an alternative nonlinear solver, based on the Newton-Raphson method, to handle the nonlinear systems (2.26).

Although some of the IMEX-RK schemes that we consider are singly-diagonally implicit Runge-Kutta (SDIRK) schemes, we do not exploit the fact that the Jacobians associated with the nonlinear function defined in (2.26) satisfy  $\mathcal{J}_{\mathbf{F}_i}(\mathbf{u}) = \mathcal{J}_{\mathbf{F}_j}(\mathbf{u})$  for the efficient solution of linear systems with these matrices, since this would entail “freezing” the Jacobians, a fact that would lead to a degradation of the convergence rate of the corresponding approximated NR method. Besides, since these Jacobians have a block tridiagonal structure, the triangularization step that could be saved by freezing Jacobians has a similar cost as the substitution steps that are needed to solve systems with distinct right hand sides. For the sake of simplicity, we denote  $\mathbf{F} = \mathbf{F}_i$  in the rest of this section.

To approximately solve the nonlinear system (2.26) by the NR iterative method it is necessary that the coefficients of the matrix function  $\mathbf{B}$ , and therefore those of  $\mathcal{B}$ , be at least continuously differentiable ([88, page 311]). However, the models of interest here, namely the diffusively corrected polydisperse sedimentation and MCLWR models, do not naturally satisfy this assumption. We therefore replace  $\mathbf{B}$  by a smooth approximation  $\mathbf{B}_\varepsilon$ , and denote the corresponding version of  $\mathcal{B}$  by  $\mathcal{B}_\varepsilon$ , where it is understood that  $\mathbf{B}_\varepsilon \rightarrow \mathbf{B}$  and  $\mathcal{B}_\varepsilon \rightarrow \mathcal{B}$  as  $\varepsilon \rightarrow 0$ . The precise algebraic form of this approximation is defined separately for each specific application in section 2.5. Note that the purpose of this approximation is to create smoothness, but not to convert the problem into a uniformly parabolic one.

We denote by  $\mathbf{F}_\varepsilon(\mathbf{u})$  the function (2.26), where  $\mathcal{B}(\mathbf{u})$  has been replaced by  $\mathcal{B}_\varepsilon(\mathbf{u})$ . The function  $\mathbf{F}_\varepsilon$  is highly nonlinear for small  $\varepsilon$ , in the sense that the second derivative of  $\mathbf{F}_\varepsilon$  is much larger than its first derivative. Therefore, by Kantorovich’s theorem (see [45]), the region of guaranteed convergence shrinks when  $\varepsilon \rightarrow 0$ . On the other hand, the linearity of  $\mathbf{F}_\varepsilon$  behaves in the opposite way when increasing  $\varepsilon$  (in fact, for the regularization used in Section 2.5,  $\mathbf{F}_\varepsilon(\mathbf{u}) \rightarrow \mathbf{u} - \mathbf{r}$  when  $\varepsilon \rightarrow \infty$ ), so the region of guaranteed convergence of the NR method increases. With these observations, we use a similar strategy as the one used in [39] to efficiently solve  $\mathbf{F}_\varepsilon(\mathbf{u}) = \mathbf{0}$  for a prescribed  $\varepsilon = \varepsilon_{\min}$  as follows:

If  $\mathbf{u}_\varepsilon$  is a solution of  $\mathbf{F}_\varepsilon(\mathbf{u}_\varepsilon) = \mathbf{0}$ , then  $\mathbf{u}_\varepsilon$  is used as an initial datum for approximating the solution of  $\mathbf{F}_{\varepsilon'}(\mathbf{u}) = \mathbf{0}$  for  $\varepsilon' < \varepsilon$  by the NR method with a line search strategy (see [45]). This process is started with a sufficiently large value  $\varepsilon_0$  and it is performed until a solution  $\mathbf{F}_{\varepsilon_{\min}}(\mathbf{u}) = \mathbf{0}$  is obtained. Based on the previous discussion on the linearity of  $\mathbf{F}_\varepsilon$ , we select  $\varepsilon_0$  as the smallest

power of 10 for which the NR method succeeds (i.e., converges within a given tolerance and a generously large maximum number of iterations) solving  $\mathbf{F}_{\varepsilon_0}(\mathbf{u}) = \mathbf{0}$  when given  $\mathbf{u} = \Phi^n$  as the initial guess.

The decrease of  $\varepsilon$  can be automated by using that  $\lim_{\varepsilon \rightarrow \varepsilon'} \mathbf{u}_\varepsilon = \mathbf{u}_{\varepsilon'}$  (this is ensured by compactness and uniqueness of solutions of the problems near  $\Phi^n$ ) enables us to choose  $\varepsilon' = \kappa\varepsilon$  for some  $\kappa \in (0, 1)$  and use  $\mathbf{u}_\varepsilon$  as the initial iteration to solve  $\mathbf{F}_{\varepsilon'}(\mathbf{u})$ , hoping that this initial guess is close enough to the solution for the NR method to converge within a given tolerance and a given maximum number of iterations. If the NR method does not succeed, then we take  $\kappa$  closer to 1 and try again; on the other hand, if the NR method takes a small number of iterations to converge (less than 3, say), then we diminish the factor  $\kappa$ .

It is easy to see that the direction obtained from the NR method for the solution of  $\mathbf{F}_\varepsilon(\mathbf{u}) = \mathbf{0}$ , namely the vector  $(-\mathcal{J}_{\mathbf{F}_\varepsilon}(\mathbf{u}))^{-1}\mathbf{F}_\varepsilon(\mathbf{u})$ , is a direction of descent for  $m(\mathbf{u}) = \|\mathbf{F}_\varepsilon(\mathbf{u})\|_2^2$ . Therefore we can use the following algorithm to ensure the convergence of the NR method.

**Algorithm 2.2 (Newton-Raphson (NR) method with line search strategy)**

*Input: approximate solution vector  $\Phi^n$  at  $t = t_n$  as a starting value*

$\mathbf{u}^{(0)} \leftarrow \Phi^n$ ,  $\varepsilon \leftarrow \varepsilon_0$ ,  $m(\mathbf{u}) \leftarrow \|\mathbf{F}_\varepsilon(\mathbf{u}^{(0)})\|_2^2$ ,  $\nu \leftarrow 0$

**while**  $\varepsilon \geq \varepsilon_{\min}$  **do**

**while**  $\nu \leq N_{\text{iter}}$  **and**  $m(\mathbf{u}^{(\nu)}) < \text{tol}$  **do**

*solve for  $\mathbf{p}_\nu$  the linear system  $\mathcal{J}_{\mathbf{F}_\varepsilon}(\mathbf{u}^{(\nu)})\mathbf{p}_\nu = -\mathbf{F}_\varepsilon(\mathbf{u}^{(\nu)})$*

$\alpha_0 \leftarrow 1$

$k \leftarrow 0$

**while**  $\alpha_k \geq \alpha_{\min}$  **do**

$\mathbf{u} \leftarrow \mathbf{u}^{(\nu)} + \alpha_k \mathbf{p}_\nu$ ,  $m(\mathbf{u}) \leftarrow \|\mathbf{F}_\varepsilon(\mathbf{u})\|_2^2$

**if**  $m(\mathbf{u}) < m(\mathbf{u}^{(\nu)})$  **then**

$\mathbf{u}^{(\nu+1)} \leftarrow \mathbf{u}$

**else**

$\alpha_{k+1} \leftarrow 0.8\alpha_k$

**endif**

$k \leftarrow k + 1$

**endwhile**

$\nu \leftarrow \nu + 1$

**endwhile**

*Decrease  $\varepsilon$*

**endwhile**

*Output: approximate solution  $\mathbf{u}$  of the nonlinear system  $\mathbf{F}_{\varepsilon_{\min}}(\mathbf{u}) = \mathbf{0}$ .*

Since the matrix  $\mathcal{J}_{\mathbf{F}_\varepsilon}(\mathbf{u})$  is block tridiagonal (and block circulant for periodic boundary conditions), an efficient block tridiagonal solver for the linear systems  $\mathcal{J}_{\mathbf{F}_\varepsilon}(\mathbf{u})\mathbf{z} = -\mathbf{F}_\varepsilon(\mathbf{u})$  can be used. The invertibility of these matrices (only for sufficiently small  $\Delta t$ ) is not ensured, but we have not experienced any invertibility failures in our tests.

## 2.5 Numerical results

For comparison purposes, we compute reference solutions for numerical tests by the KT scheme [71], which is employed in [12] for the numerical solution of (2.1) for the diffusively corrected polydisperse sedimentation model (sedimentation with compression). The reference solution is based on a fine discretization with  $M_{\text{ref}} = 12800$  cells and  $\Delta t$  is selected at each time step following the formula

$$\Delta t = C_{\text{cfl1}} \left( \frac{\max_{\Phi} \tilde{\rho}(\mathcal{J}_{\mathbf{f}}(\Phi))}{\Delta x} + \frac{\max_{\Phi} \tilde{\rho}(\mathbf{B}(\Phi))}{2\Delta x^2} \right)^{-1} \quad (2.31)$$

with  $C_{\text{cfl1}} = 0.25$  and with estimates  $\tilde{\rho}$  of the spectral radius of the corresponding matrices obtained from Theorems 1.2 and 2.1 for the polydisperse case and Theorems 2.2 and 2.3 for the traffic model. This CFL number has been adjusted empirically to be the largest multiple of 0.05 that yields an oscillation-free reference solution. The variable time step (2.31) with  $C_{\text{cfl1}} = 0.25$  has been used for the KT scheme in all tests. The following time steps for the IMEX-RK schemes are used:

$$\Delta t = C_{\text{cfl2}} \Delta x (\max_{\Phi} \rho(\mathcal{J}_{\mathbf{f}}(\Phi)))^{-1}, \quad (2.32)$$

where  $\rho(\mathcal{J}_{\mathbf{f}}(\Phi))$  is computed along with the characteristic information needed for the convective part and  $C_{\text{cfl2}}$  is empirically obtained as the largest multiple of 0.05 that yields oscillation-free simulations with  $M_{\text{ref}}$  cells. These numbers are  $C_{\text{cfl2}} = 0.25$  for the scheme IMEX-ARS(1,1,1) (2.28) in Example 2.1 and  $C_{\text{cfl2}} = 0.1$  for Examples 2.2 and 2.3, whereas  $C_{\text{cfl2}} = 0.7$  for the methods IMEX-ARS(2,2,2) (2.29) and IMEX-SSP2(3,3,2) (2.30) in all examples. We mention that the scheme IMEX-ARS(1,1,1) applied to fifth-order WENO (WENO5) spatial semi-discretizations should have a stability restriction related to that for the forward Euler method and WENO5. A modified von Neumann analysis carried out in [80] indicates that the CFL number of the schemes obtained by using Euler method to integrate semi-discretizations obtained by WENO5 have a stability restriction proportional to  $\Delta x^4$ . In our experiments we have not had to use such a small restriction, but we have had to reduce the Courant number considerably with respect to the other IMEX methods.

Total approximate  $L^1$  errors at different times for each scheme are computed as follows. Let us denote by  $(\phi_{j,i}^M(t))_{j=1}^M$  and  $(\phi_{l,i}^{\text{ref}}(t))_{l=1}^{M_{\text{ref}}}$  the numerical solution for the  $i$ -th component at time  $t$  calculated with  $M$  and  $M_{\text{ref}}$  cells, respectively. We use cubic interpolation from the reference grid to the  $M$  cells grid to compute  $\tilde{\phi}_{j,i}^{\text{ref}}(t)$  for  $j = 1, \dots, M$ . We then calculate the approximate  $L^1$  error in species  $i$  by

$$e_i(t) := \frac{1}{M} \sum_{j=1}^M |\tilde{\phi}_{j,i}^{\text{ref}}(t) - \phi_{j,i}^M(t)|, \quad i = 1, \dots, N.$$

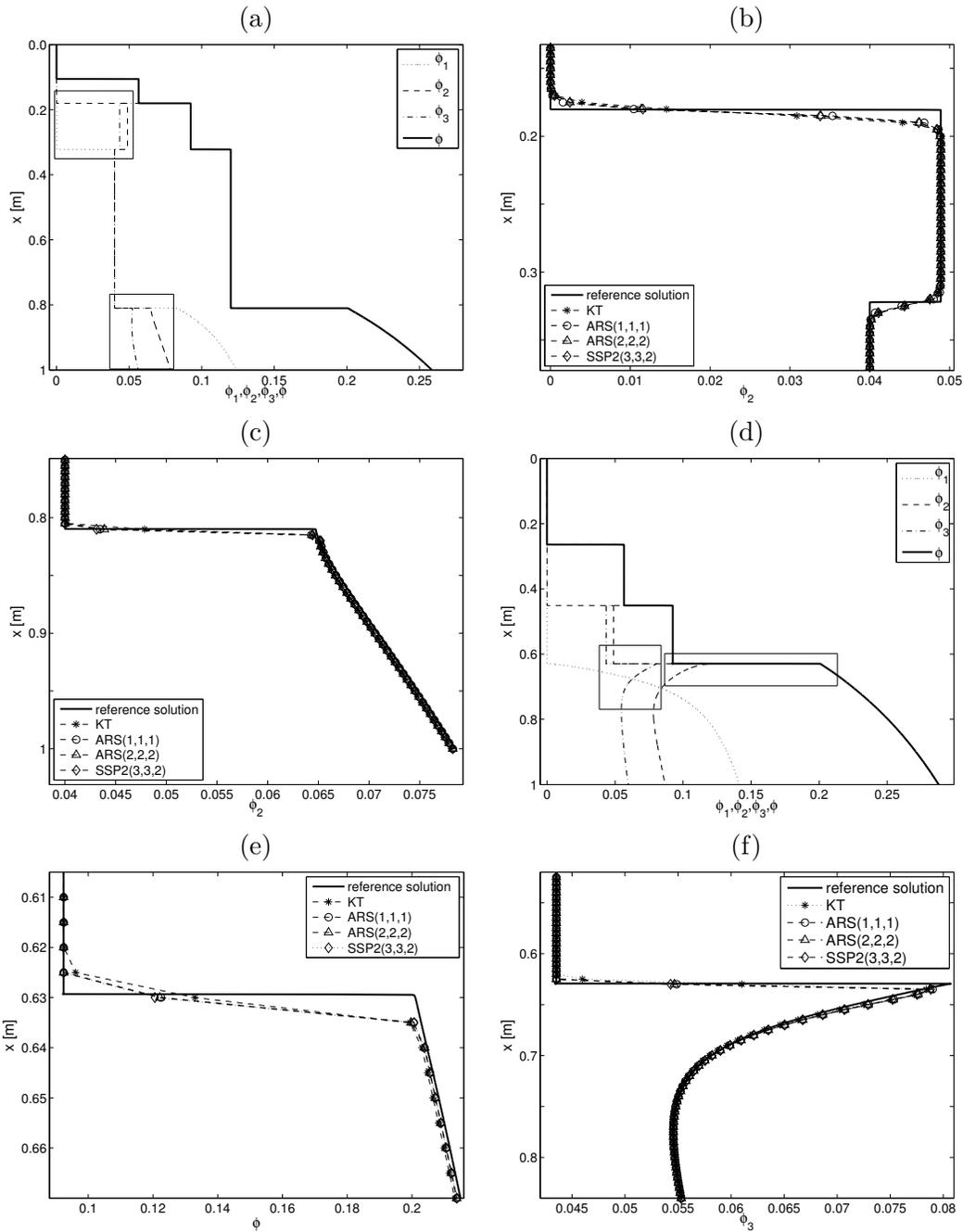


Figure 2.1: Example 2.1: (a) reference solution (KT scheme,  $M = 12800$ ) at  $T = 4000$  s, (b, c) enlarged views of numerical solutions with  $M = 200$  at  $T = 4000$  s, (d) reference solution at  $T = 10000$  s, (e, f) enlarged views of numerical solutions with  $M = 200$  at  $T = 10000$  s.

The total approximate  $L^1$  error at time  $t$  is defined as  $e_{\text{tot}}(t) := e_1(t) + \dots + e_N(t)$ .

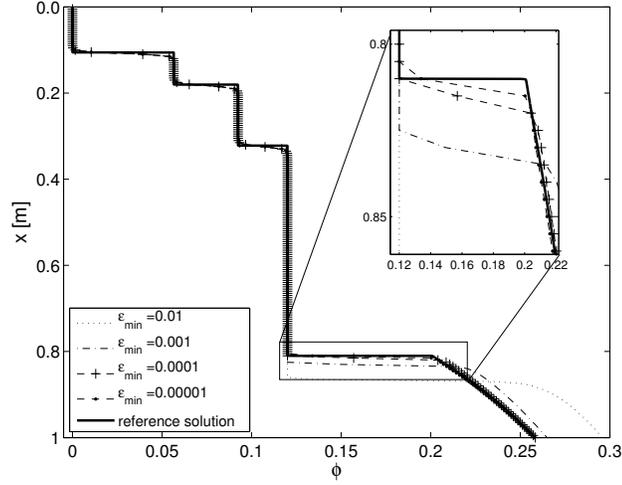


Figure 2.2: Example 2.1: numerical approximation by scheme IMEX-ARS(2,2,2) at simulated time  $T = 4000$  s using  $M = 200$  and a fixed value of  $\varepsilon = \varepsilon_{\min}$  in (2.33).

(a)					(b)				
IMEX scheme	$\nu_{\max}$	$k_{\max}$	$\bar{\nu}$	$\bar{k}$	IMEX scheme	$\nu_{\max}$	$k_{\max}$	$\bar{\nu}$	$\bar{k}$
ARS(1,1,1)	5	5	1.56	0.71	ARS(1,1,1)	3	1	1.12	0.001
ARS(2,2,2)	6	4	2.06	0.51	ARS(2,2,2)	3	1	1.10	0.01
SSP2(3,3,2)	6	3	2.21	0.23	SSP2(3,3,2)	3	1	1.07	0.02

Table 2.1: Examples 2.1 and 2.3: maximum number of iterations  $\nu_{\max}$  required by the NR method for solving (2.26), maximum number of reductions  $k_{\max}$  of  $\alpha$  in Algorithm 2.2, global average number of iterations  $\bar{\nu}$ , and global average number of reductions  $\bar{k}$  of  $\alpha$  required different IMEX schemes with  $M = 400$  for (a) Example 2.1 at  $T = 10000$  s and (b) Example 2.3 at  $T = 0.2$  h.

### 2.5.1 Example 2.1: settling of a tridisperse suspension

We simulate the settling of a tridisperse ( $N = 3$ ) suspension forming a compressible sediment. The mixture is described by the model functions (1.3), (1.4) and (2.7) with  $\phi_{\max} = 0.66$ ,  $n_{\text{RZ}} = 4.7$ ,  $\sigma_0 = 180$  Pa,  $\phi_c = 0.2$ ,  $k = 2$ ,  $\mu_f = 10^{-3}$  Pa s,  $d = 1.19 \times 10^{-5}$  m,  $\bar{\rho}_s = 1800$  kg/m<sup>3</sup>, and  $g = 9.81$  m/s<sup>2</sup> [12]. The initial concentration is  $\Phi_0 = (0.04, 0.04, 0.04)^T$  in a vessel of height  $K = 1$  m with normalized squared particle sizes  $\delta = (1, 0.5, 0.25)^T$ .

For this model, the coefficients are defined in terms of the function  $\sigma_e(\phi)$  and its derivative (cf. (2.5)). The regularization of Section 2.4 is achieved by replacing  $\sigma_e(\phi)$  by a regularized smooth function  $\sigma_e(\phi; \varepsilon)$  such that  $\sigma_e(\phi; \varepsilon) \rightarrow \sigma_e(\phi)$  for all  $\phi$  and  $\sigma'_e(\phi; \varepsilon) \rightarrow \sigma'_e(\phi)$  for all  $\phi \neq \phi_c$  as  $\varepsilon \rightarrow 0$ . Specifically, if  $\sigma_e$  satisfies (2.6), we choose

$$\sigma_e(\phi; \varepsilon) = \sigma_e(\phi) \exp(-\varepsilon/(\phi - \phi_c)^2), \quad \varepsilon > 0. \quad (2.33)$$

In Figure 2.1 we compare results obtained by schemes KT, IMEX-ARS(1,1,1), IMEX-ARS(2,2,2) and IMEX-SSP2(3,3,2). We observe good approximations for the IMEX schemes compared with

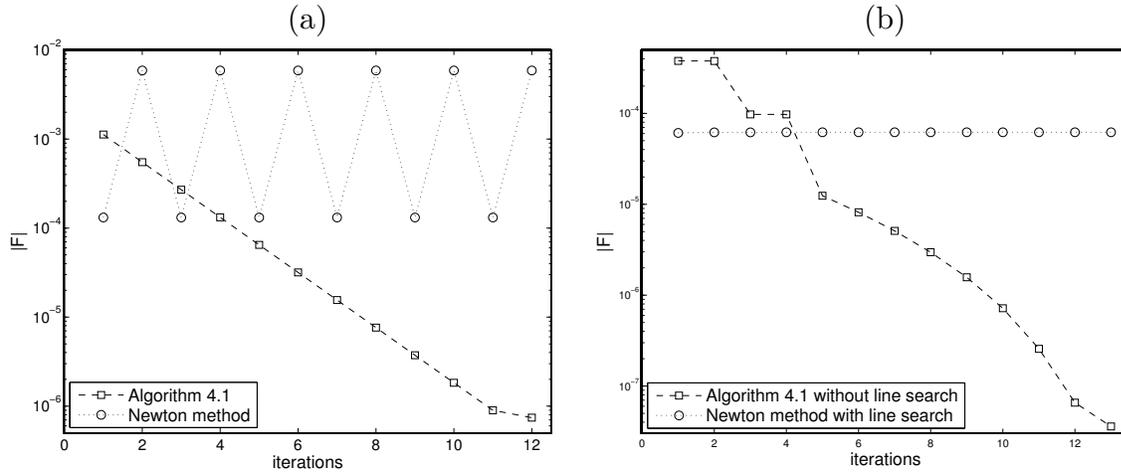


Figure 2.3: Example 2.1: comparison of the convergence history of nonlinear solvers obtained with scheme IMEX-ARS(1,1,1) with (a)  $M = 800$  at simulated time  $T = 19.2$ s, NR method versus Algorithm 2.2, (b)  $M = 1600$  at simulated time  $T = 76.4$ s, Algorithm 2.2 without line search versus NR method with line search.

		KT		IMEX-ARS(1,1,1)			IMEX-ARS(2,2,2)			IMEX-SSP2(3,3,2)			
$T$ [s]	$M$	error	cr	cpu	error	cr	cpu [s]	error	cr	cpu [s]	error	cr	cpu [s]
4000	100	150.4	—	0.6	130.3	—	0.2	145.1	—	0.1	146.1	—	0.2
	200	75.7	0.99	4.7	61.3	1.08	1.1	70.3	1.04	0.6	70.9	1.04	0.9
	400	38.4	0.98	31.4	31.1	0.96	4.3	36.2	0.95	2.4	36.8	0.94	6.0
	800	20.8	0.88	289.1	16.7	0.90	20.9	18.5	0.96	11.2	18.9	0.95	18.9
	1600	11.3	0.88	2349.0	9.0	0.89	108.7	9.9	0.89	65.5	10.1	0.90	97.8
10000	100	141.1	—	2.4	136.5	—	0.7	142.2	—	0.4	139.6	—	0.6
	200	75.6	0.89	18.1	64.9	1.07	3.1	68.7	1.04	1.7	68.8	1.02	2.7
	400	40.2	0.91	101.6	41.8	0.63	11.7	43.8	0.64	8.7	43.8	0.65	14.0
	800	21.6	0.89	1103.0	19.6	1.09	70.7	20.7	1.08	39.0	20.7	1.08	57.9
	1600	12.6	0.78	9058.6	12.0	0.70	318.3	12.3	0.75	173.5	12.2	0.76	280.1

Table 2.2: Example 2.1: total approximate  $L^1$  errors  $e_{\text{tot}}(T)$  (“error”, to be multiplied by  $10^{-5}$ ), convergence rates (cr), and CPU times (cpu), at two times  $T$  for the schemes KT and IMEX-ARS(1,1,1) with  $C_{\text{eff1}} = C_{\text{eff2}} = 0.25$ , and the schemes IMEX-ARS(2,2,2) and IMEX-SSP2(3,3,2) with  $C_{\text{eff2}} = 0.7$ .

the KT scheme near  $\phi_c$ . No oscillations appear at this scale. The nonlinear system (2.26) is solved by Algorithm 2.2, where  $\varepsilon$  varies from  $\varepsilon_0 = 10^{-4}$  to  $\varepsilon_{\min} = 10^{-7}$  and  $\text{tol} = 10^{-8}$ . Figure 2.2 displays numerical approximations obtained by IMEX-ARS(2,2,2) where the diffusion coefficient is regularized by (2.33) for different fixed values of  $\varepsilon = \varepsilon_{\min}$ . Note that for  $\varepsilon = 10^{-5}$ , the reference solution is approximated well.

Table 2.1 (a) informs some details of the convergence history of the NR method to obtain a numerical approximation within several IMEX–RK schemes. Observe that, for each fixed value of  $\varepsilon$ ,

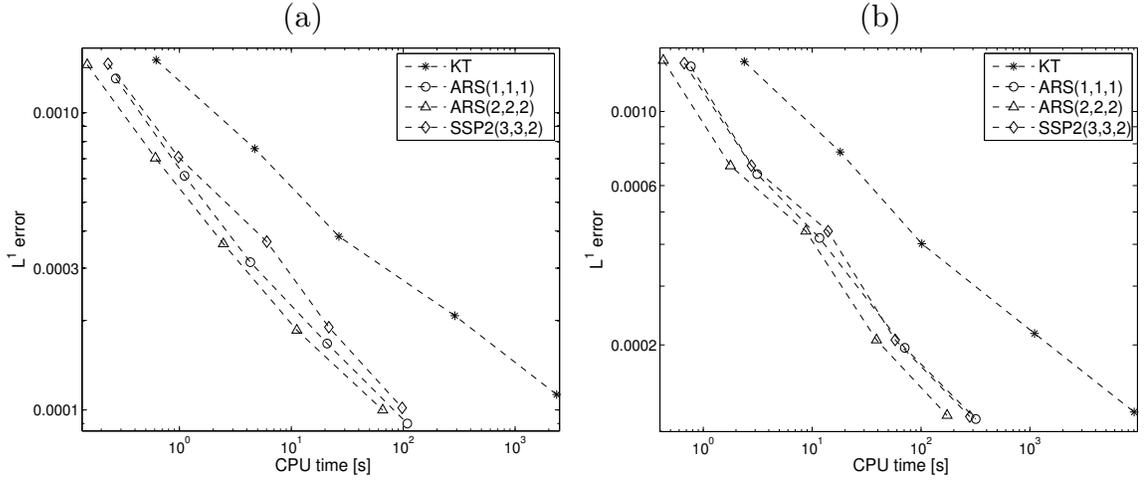


Figure 2.4: Example 2.1: total approximate  $L^1$  errors versus CPU time for KT and IMEX schemes at simulated times (a)  $T = 4000$  s, (b)  $T = 10000$  s. Here and in Figures 2.7 and 2.9, for each scheme the interpolated symbols correspond to different values of  $M$ .

within scheme IMEX-ARS(2,2,2) the NR method requires at most 6 iterations, but in general, only 2 or 3 iterations were necessary. With respect to  $\alpha$ , Algorithm 2.2 reduces this parameter at most five times, but less than once in average.

In Figure 2.3 we test different variants of Algorithm 2.2 to ascertain which of its internal mechanisms is the one that ensures convergence. In the first setup we compare the (undamped) NR method versus our proposed algorithm. The system to be solved corresponds to scheme IMEX-ARS(1,1,1). The simulated time is  $T = 19.2$  s, we set  $M = 800$ , and the parameters for Algorithm 2.2 are  $\varepsilon_0 = 10^{-4}$  and  $\varepsilon_{\min} = 10^{-8}$ . In the second setup we compare Algorithm 2.2 without line search with a gradual descent of  $\varepsilon$  from  $\varepsilon_0 = 10^{-4}$  to  $\varepsilon_{\min} = 10^{-6}$  with the NR method with line search applied directly to  $\mathbf{F}_{\varepsilon_{\min}}(\mathbf{u}) = \mathbf{0}$ . Again the system to be solved corresponds to scheme IMEX-ARS(1,1,1), and the simulated time is  $T = 76.4$  s with  $M = 1600$ . The simulated times  $T$  provided in these tests correspond to the time step when one of the nonlinear solvers in the comparison did not succeed converging within a generously prescribed maximum number of iterations. Figure 2.3 shows  $\|\mathbf{F}(\mathbf{u}^{(\nu)})\|_2$  (without regularization) for  $\nu = 0, 1, 2, \dots$  until Algorithm 2.2 reaches convergence. As can be deduced from Figure 2.3 (a), the NR method does not converge while Algorithm 2.2 does (although not at a quadratic rate for  $\|\mathbf{F}(\mathbf{u}^{(\nu)})\|_2$ , since NR is applied to  $\mathbf{F}_\varepsilon$  for decreasing  $\varepsilon > 0$ ). We deduce from Figure 2.3 (b) that the damped NR method applied to  $\mathbf{F}_{\varepsilon_{\min}}(\mathbf{u}) = \mathbf{0}$  stalls, due to the choice of an extremely small step-length to ensure decrease of the objective function, while Algorithm 2.2 without line search converges. Therefore, as shown in both experiments, the mechanism that ensures convergence to the solution is the gradual descent of  $\varepsilon$  towards the prescribed  $\varepsilon_{\min}$ ; the line search procedure may enhance efficiency, permitting a faster decrease of  $\varepsilon$ .

Table 2.2 and Figure 2.4 show approximate  $L^1$  total errors, convergence rates and CPU times for Example 2.1. According to Table 2.2, scheme IMEX-ARS(2,2,2) is the most efficient: for the same resolution  $M$ , scheme IMEX-ARS(1,1,1) has the smallest error, closely followed by schemes IMEX-ARS(2,2,2), IMEX-SSP2(3,3,2) and KT in increasing order with respect to error size; on the other hand, the CPU time of scheme IMEX-ARS(2,2,2) is the lowest, basically due to the increased CFL number compared with scheme IMEX-ARS(1,1,1) and the smaller number of implicit stages with respect to scheme IMEX-SSP2(3,3,2). One can also deduce from Table 2.2 that the CPU time scales as  $\mathcal{O}(M^2)$  for the IMEX-RK schemes, whereas it scales as  $\mathcal{O}(M^3)$  for the KT scheme. This implies a nearly fixed cost of the solution of nonlinear systems per time step and that the gap in CPU time increases with  $M$ . For instance, scheme IMEX-ARS(2,2,2) is about 60 times faster than the KT scheme for a resolution of  $M = 1600$  cells.

A careful observation of CPU times of the same scheme and resolution at different simulated times yields that the CPU time is not proportional to the simulated time. This is due to the variable time stepping in formulas (2.31) and (2.32) and to the fact that the spectral radius of the Jacobian of the fluxes and the diffusion matrix is smaller at the early stages of the simulation (in fact, the diffusion matrix may be null in a noticeable period of time) so the time steps may be larger at the beginning of the simulation.

### 2.5.2 Examples 2.2 and 2.3: diffusively corrected kinematic traffic model

We consider a circular road and a number  $N$  of driver classes associated with velocities  $v_i^{\max} > v_j^{\max}$  for  $i < j$ . If  $\rho_i$  denotes the number of cars of species  $i$  per mile, and  $\rho_{\max}$  is the maximal “bumper-to-bumper” number of cars per mile, we define  $\phi_i := \rho_i/\rho_{\max}$ . To make results comparable with those of [25], we employ the Dick-Greenberg model [48, 59]  $V(\phi) = V_{\text{DG}}(\phi) = \min\{1, -C \ln \phi\}$ , and choose (as in [25, 81, 82])  $C = e/7 \approx 0.38833$  so that

$$\begin{cases} V(\phi) = 1, V'(\phi) = 0 & \text{for } 0 \leq \phi \leq \phi_c = \exp(-1/C) \approx 0.076142, \\ V(\phi) = -C \ln \phi, V'(\phi) = -C/\phi & \text{for } \phi_c < \phi < 1. \end{cases}$$

Equipping class 1 with the same properties as the model of [25], we set  $\tau = 2 \text{ s} = 0.000\bar{5} \text{ h}$ . We choose all anticipation lengths  $L_i = L = 0.05 \text{ mi}$  and all reaction times  $\tau_i = \tau$ ,  $i = 1, \dots, N$  in such a way that (2.18) holds with  $\phi_{\max} = 1$ , i.e.,

$$\tau \leq \min_{0 \leq \phi \leq 1} \left( -\frac{L(\phi)}{\phi V'(\phi) v_1^{\max}} \right) = \frac{L}{C v_1^{\max}} = \frac{7 \cdot 0.05 \text{ mi}}{e \cdot 70 \text{ mi/h}} \approx 0.00184 \text{ h} = 6.622 \text{ s}.$$

We easily see that in this case

$$C_2 = \tau^2 (V(\phi))^2 \sum_{\substack{i,j=1 \\ i < j}}^N \phi_i v_i^{\max} \phi_j v_j^{\max} (v_j^{\max} - v_i^{\max})^2 \geq 0.$$

For this model, according to (2.11), the coefficients depend on  $V(\phi)$  and its derivative. The regularization mentioned in Section 2.4 is achieved by replacing  $V(\phi)$  by

$$V(\phi; \varepsilon) = 1 + (V(\phi) - 1) \exp(-\varepsilon/(\phi - \phi_c)^2), \quad \varepsilon > 0.$$

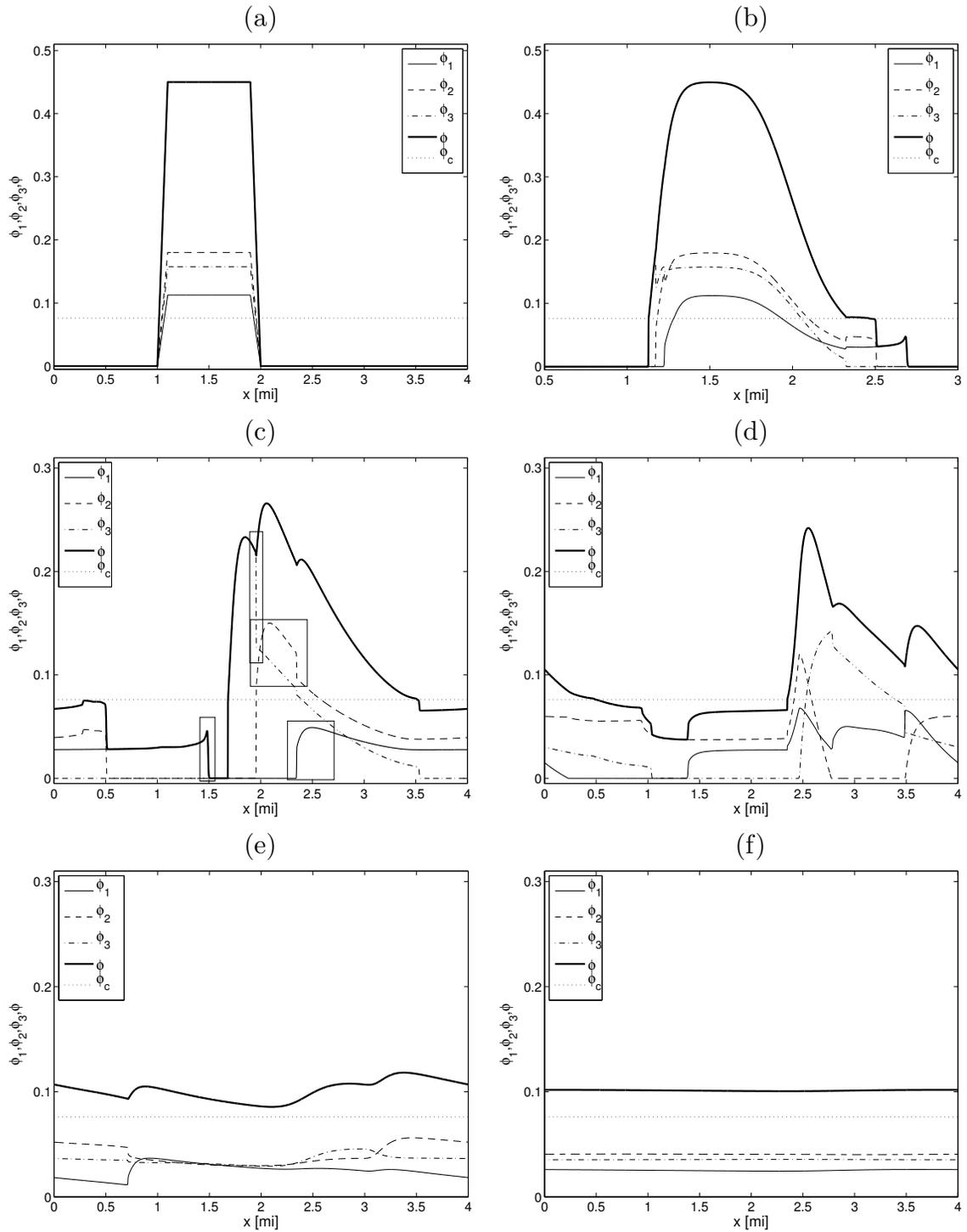


Figure 2.5: Example 2.2: numerical solution obtained with scheme IMEX-SSP2(3,3,2) with  $M = 1600$  at simulated times (a)  $T = 0.0$  h (initial datum), (b)  $T = 0.01$  h, (c)  $T = 0.05$  h (with marked areas of enlarged views shown in Figure 2.6), (d)  $T = 0.1$  h, (e)  $T = 0.5$  h and (f)  $T = 5.0$  h.

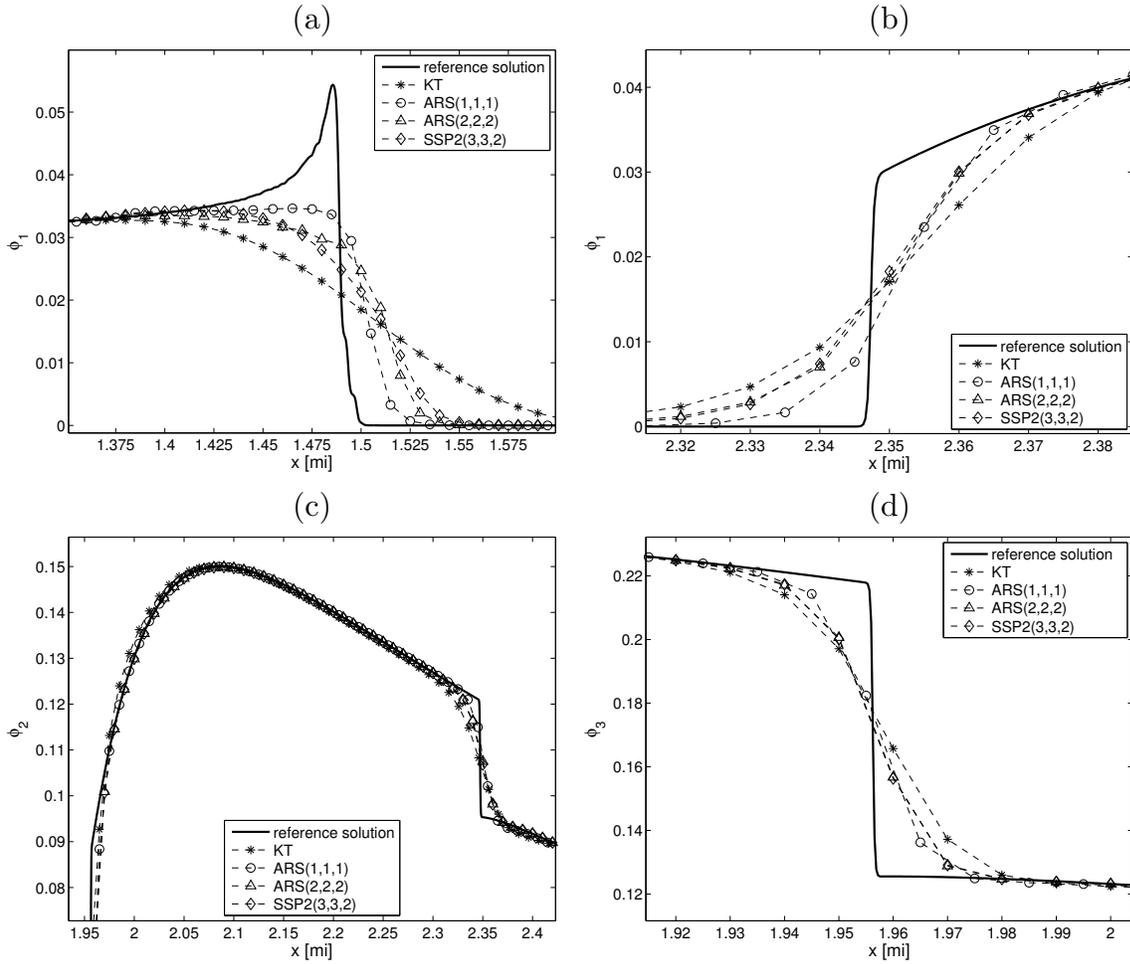


Figure 2.6: Example 2.2: (a, b, c, d) Enlarged views (cf. Figure 2.5 (c)) of reference solution (scheme KT,  $M_{\text{ref}} = 12800$ ) and numerical solutions for individual species for KT and IMEX schemes with  $M = 400$ , at simulated time  $T = 0.05$  h.

To be definite, we study  $N = 3$  vehicle classes with  $v_1^{\max} = 70$  mi/h,  $v_2^{\max} = 50$  mi/h and  $v_3^{\max} = 30$  mi/h on a circular roadway with length  $K = 4$  mi, i.e., we use periodic boundary conditions. The initial density distribution is given by an isolated platoon of maximum global density  $\rho_0$ ,  $\Phi_0(x, 0) = p(x - 1)\rho_0(0.25, 0.4, 0.35)^T$ , where

$$p(x) = \begin{cases} 10x & \text{for } 0 < x \leq 0.1, & -10(x - 1) & \text{for } 0.9 < x \leq 1, \\ 1 & \text{for } 0.1 < x \leq 0.9, & 0 & \text{otherwise,} \end{cases}$$

and we choose  $\rho_0 = 0.45$  and  $\rho_0 = 0.25$  in Examples 2.2 and 2.3, respectively. Algorithm 2.2 is used with  $\varepsilon$  varying from  $\varepsilon_0 = 10^{-4}$  to  $\varepsilon_{\min} = 10^{-6}$  and  $\text{tol} = 10^{-7}$ . Previous numerical tests indicate that  $\varepsilon_{\min} = 10^{-6}$  is sufficient to obtain good approximations.

Figure 2.5 shows the time evolution obtained with IMEX-SSP2(3,3,2) with  $M = 1600$  cells of the initial density platoon for Example 2.2.

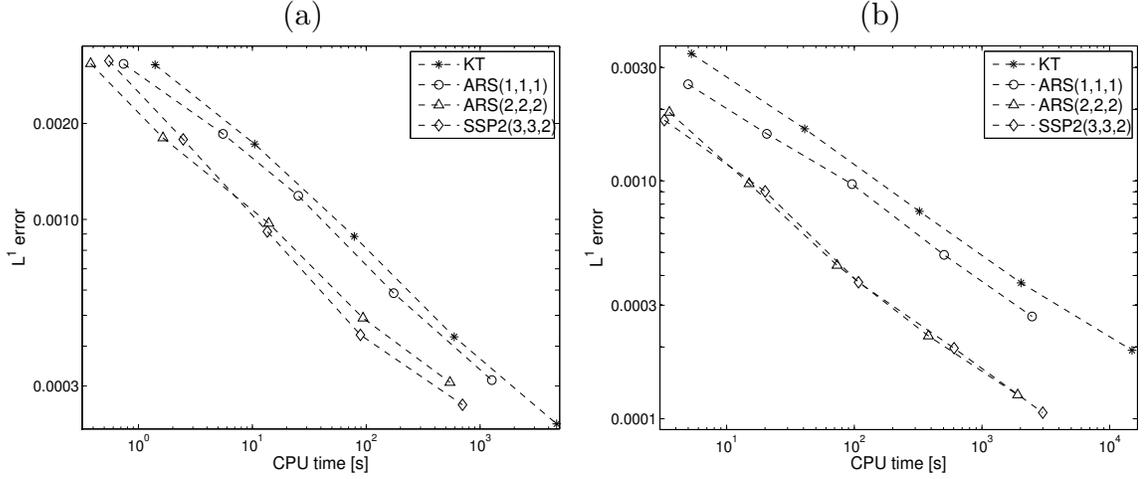


Figure 2.7: Example 2.2: total approximate  $L^1$  errors versus CPU time for KT and IMEX schemes at simulated times (a)  $T = 0.05$  h, (b)  $T = 0.25$  h.

$T$ [h]	$M$	KT			IMEX-ARS(1,1,1)			IMEX-ARS(2,2,2)			IMEX-SSP2(3,3,2)		
		error	cr	cpu [s]	error	cr	cpu [s]	error	cr	cpu [s]	error	cr	cpu [s]
0.05	200	305.7	—	1.4	307.8	—	0.7	308.3	—	0.3	314.4	—	0.5
	400	172.4	0.82	10.4	185.5	0.74	5.5	180.2	0.77	1.6	178.2	0.81	2.4
	800	88.4	0.96	78.6	118.6	0.64	25.3	97.2	0.89	13.9	91.6	0.96	13.5
	1600	42.8	1.04	592.5	58.7	1.01	174.3	49.9	0.98	93.7	43.4	1.07	89.0
	3200	22.8	0.90	4704.5	31.3	0.90	1268.4	30.8	0.76	543.9	26.2	0.72	701.6
0.25	200	343.2	—	5.3	255.2	—	4.9	194.4	—	3.5	180.2	—	3.2
	400	165.7	1.04	40.9	157.9	0.69	20.6	97.4	0.99	15.0	90.4	0.99	20.0
	800	74.5	1.15	324.1	97.0	0.70	95.8	44.2	1.13	73.3	37.5	1.26	107.8
	1600	37.2	0.99	2022.1	48.9	0.98	504.7	22.3	0.98	380.9	19.8	0.92	604.4
	3200	19.4	0.94	14965.5	26.9	0.86	2469.7	12.6	0.82	1904.3	10.6	0.90	2992.2

Table 2.3: Example 2.2: total approximate  $L^1$  errors  $e_{\text{tot}}(T)$  (“error”, to be multiplied by  $10^{-5}$ ), convergence rates and CPU times at time  $T$  for scheme KT with  $C_{\text{eff}1} = 0.25$ , scheme IMEX-ARS(1,1,1) with  $C_{\text{eff}2} = 0.1$ , and schemes IMEX-ARS(2,2,2) and IMEX-SSP2(3,3,2) with  $C_{\text{eff}2} = 0.7$ .

The average density exceeds  $\phi_c$ , i.e., the traffic is relatively dense. We observe that the numerical solution evolves to a stationary solution, which must lie in the parabolic region. In Figure 2.6 we compare the results obtained by the KT and IMEX-RK schemes. Plotted areas correspond to regions where the diffusive term acts. The IMEX-RK schemes approximate adequately the reference solution.

In Table 2.3 and Figure 2.7 we display the history of total approximate  $L^1$  errors and CPU times for Example 2.2. We infer that schemes IMEX-ARS(2,2,2) and IMEX-SSP2(3,3,2) are always more efficient than the KT scheme, with speedup factors above 10. However, for small resolutions, scheme IMEX-ARS(1,1,1) is penalized by the CFL reduction and, for instance, it is less efficient

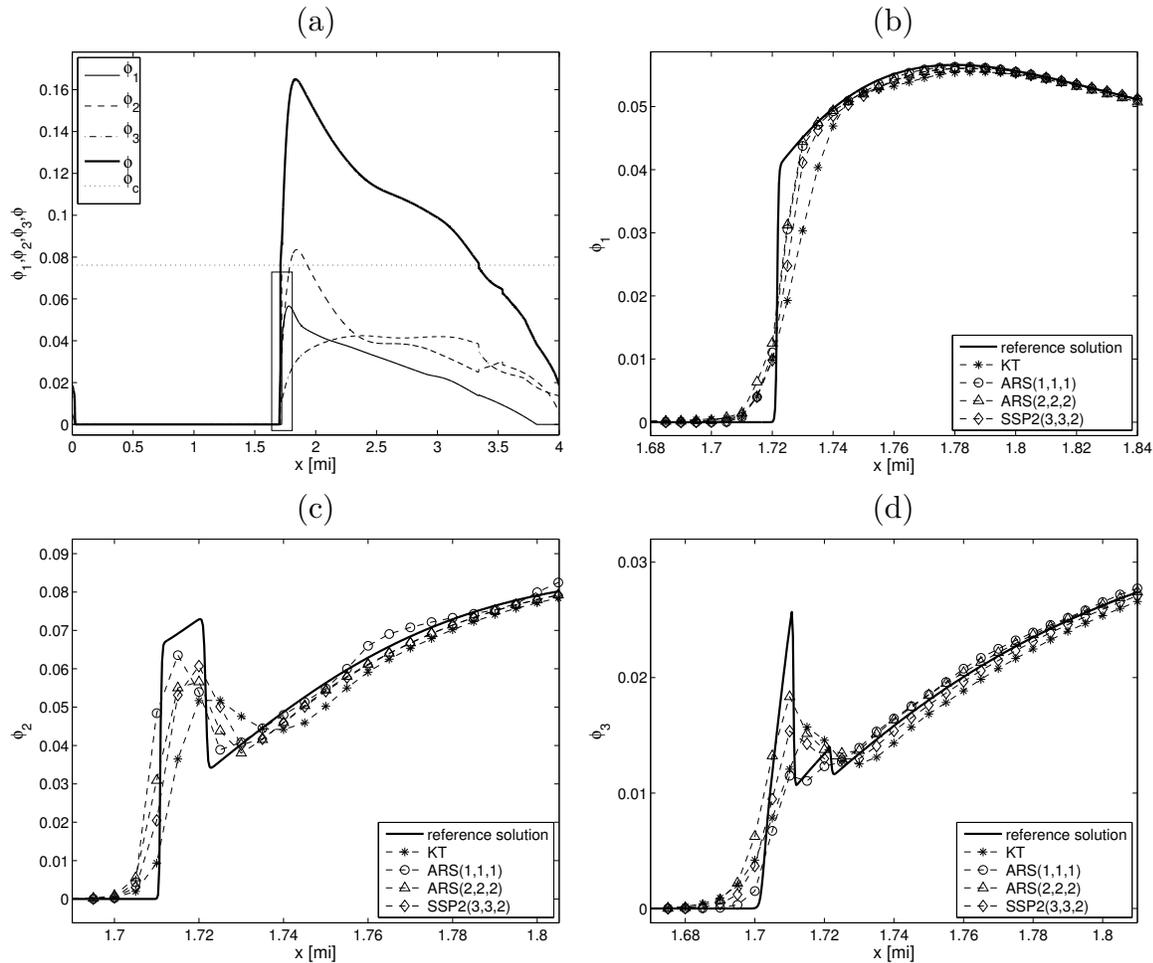


Figure 2.8: Example 2.3: (a) reference solution (scheme KT,  $M_{\text{ref}} = 12800$ ), (b, c, d) enlarged views of reference solution and numerical solutions for individual species for KT and IMEX schemes with  $M = 800$ , at simulated time  $T = 0.2$  h.

than the KT scheme for  $M = 200$ . Note that the numerical approximations with scheme IMEX-SSP2(3,3,2) have smaller error than those obtained with scheme IMEX-ARS(2,2,2), but the CPU time is larger.

In Example 2.3 we choose  $\rho_0 = 0.25$  so that the average density is below  $\phi_c$ . This case does not evolve into a stationary solution. In Figure 2.8 we compare results obtained by the KT, IMEX-RK schemes at simulated time  $T = 0.2$  h with respect to the reference solution. Numerical approximations are computed with  $M = 800$  cells. It is observed that the IMEX-RK schemes approximate adequately the reference solution. In Table 2.1 (b) some details of the convergence history of the NR method are described. Observe that for scheme IMEX-SSP2(3,3,2) and each fixed value of  $\varepsilon$ , the NR method requires at most 3 iterations, but in average, only one iteration is necessary. With respect to  $\alpha$ , at some point Algorithm 2.2 reduces it at maximum once but, in general, in most cases not even one reduction was necessary.

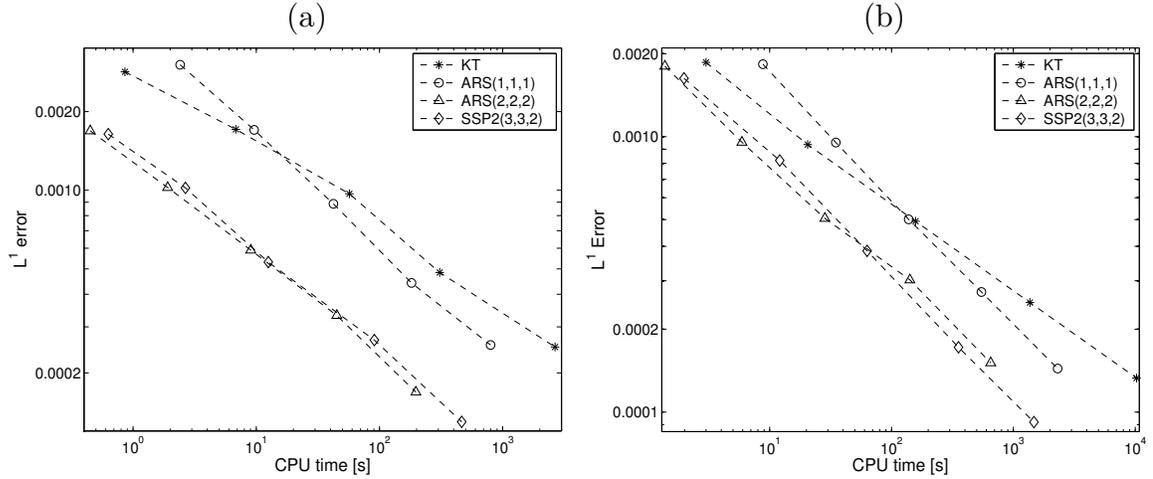


Figure 2.9: Example 2.3: total approximate  $L^1$  errors versus CPU time for KT and IMEX-RK schemes at simulated times (a)  $T = 0.06$  h, (b)  $T = 0.2$  h.

$T$ [h]	$M$	KT			IMEX-ARS(1,1,1)			IMEX-ARS(2,2,2)			IMEX-SSP2(3,3,2)		
		error	cr	cpu [s]	error	cr	cpu [s]	error	cr	cpu [s]	error	cr	cpu [s]
0.06	200	283.8	—	0.8	294.8	—	2.5	168.7	—	0.4	163.9	—	0.6
	400	170.9	0.71	6.8	150.4	0.97	10.1	102.4	0.72	1.9	102.1	0.68	2.6
	800	96.9	0.81	56.9	82.9	0.85	41.8	59.0	0.79	9.0	53.1	0.94	12.5
	1600	48.4	1.01	307.9	45.5	0.86	166.5	33.2	0.83	45.0	26.7	0.99	90.8
	3200	25.1	0.94	2660.1	26.2	0.79	874.4	16.9	0.97	197.4	13.0	1.03	464.3
0.20	200	186.0	—	3.0	183.3	—	8.8	180.2	—	1.3	162.8	—	1.9
	400	93.6	0.99	20.5	95.2	0.94	34.9	95.4	0.91	5.9	81.9	0.99	12.1
	800	49.3	0.92	157.4	50.1	0.92	138.3	50.5	0.91	28.3	38.5	1.08	63.0
	1600	25.0	0.97	1367.8	27.3	0.87	546.3	30.2	0.74	141.5	17.2	1.16	354.8
	3200	13.3	0.90	10261.1	14.4	0.92	2309.2	11.3	1.01	651.5	9.2	0.90	1474.4

Table 2.4: Example 2.3: total approximate  $L^1$  total errors  $e_{\text{tot}}(T)$  (“error”, to be multiplied by  $10^{-5}$ ), convergence rates and CPU times for KT scheme with  $C_{\text{eff1}} = 0.25$ , scheme IMEX-RK(1,1,1) with  $C_{\text{eff2}} = 0.1$ , and schemes IMEX-ARS(2,2,2) and IMEX-SSP2(3,3,2) with  $C_{\text{eff2}} = 0.7$ .

In Table 2.4 and Figure 2.9, we display total approximate  $L^1$  errors and CPU times for Example 2.3. Schemes IMEX-ARS(2,2,2) and IMEX-SSP2(3,3,2) appear to be always more efficient than the KT scheme. This reconfirms the same trends as those mentioned for the previous setup. As in Example 2.2, numerical approximations with scheme IMEX-SSP2(3,3,2) have smaller errors than those obtained with scheme IMEX-ARS(2,2,2), but the CPU time is larger.

## 2.6 Derivation of the model of polydisperse sedimentation with compression

In this subsection the derivation of the model of polydisperse sedimentation with compression described in Section 2 of [12] is summarized. A suspension may be represented as a superposition of continuous media, each following its own movement with the only restrictions imposed by the interaction between components. Each component obeys the laws of conservation of mass and momentum, incorporating terms to account for the interchange of momentum between components. Assuming that there is no mass transfer between species, the local mass balance equation of the solid species and the fluid are described as

$$\frac{\partial \phi_i}{\partial t} + \nabla \cdot (\phi_i \mathbf{v}_i) = 0, \quad i = 1, \dots, N, \quad \frac{\partial \phi}{\partial t} - \nabla \cdot ((1 - \phi) \mathbf{v}_f) = 0, \quad (2.34)$$

where  $\mathbf{v}_i$  is the phase velocity of solids species  $i$ ,  $i = 1, \dots, N$ , and  $\mathbf{v}_f$  is the fluid phase velocity. Defining the volume-average velocity of the mixture  $\mathbf{q} := (1 - \phi) \mathbf{v}_f + \phi_1 \mathbf{v}_1 + \dots + \phi_N \mathbf{v}_N$  and the relative velocities  $\mathbf{u}_i := \mathbf{v}_i - \mathbf{v}_f$  for  $i = 1, \dots, N$ , the solids mass balance equations can be rewritten in terms of  $\mathbf{q}$  and  $\mathbf{u}_1, \dots, \mathbf{u}_N$  as

$$\frac{\partial \phi_i}{\partial t} + \nabla \cdot (\phi_i \mathbf{u}_i + \phi_i \mathbf{q} - \phi_i (\phi_1 \mathbf{u}_1 + \dots + \phi_N \mathbf{u}_N)) = 0, \quad i = 1, \dots, N. \quad (2.35)$$

The sum of all equations in (2.34) produces the simple mass balance of the mixture  $\nabla \cdot \mathbf{q} = 0$ . The momentum balance equations for the  $N$  solid species and the fluid are

$$\varrho_i \phi_i \frac{D\mathbf{v}_i}{Dt} = \nabla \cdot \mathbf{T}_i + \varrho_i \phi_i \mathbf{b} + \mathbf{m}_i^f + \mathbf{m}_i^s, \quad i = 1, \dots, N, \quad (2.36)$$

$$\varrho_f (1 - \phi) \frac{D\mathbf{v}_f}{Dt} = \nabla \cdot \mathbf{T}_f + \varrho_f (1 - \phi) \mathbf{b} - (\mathbf{m}_1^f + \dots + \mathbf{m}_N^f). \quad (2.37)$$

Here  $\varrho_f$  is the mass density of the fluid,  $\mathbf{T}_i$  denotes the stress tensor of particles species  $i$ ,  $\mathbf{T}_f$  that of the fluid,  $\mathbf{b}$  is the body force,  $\mathbf{m}_i^f$  and  $\mathbf{m}_i^s$  is the particle-particle interaction term of species  $i$ , and  $D\mathbf{v}/Dt := \partial \mathbf{v}/\partial t + (\mathbf{v} \cdot \nabla) \mathbf{v}$ .

We assume that the stress tensors of the solid and fluid phase can be written as  $\mathbf{T}_i = -p_i \mathbf{I} + \mathbf{T}_i^E$  for each  $i$ , and  $\mathbf{T}_f = -p_f \mathbf{I} + \mathbf{T}_f^E$ , respectively, where  $p_i$  denotes the phase pressure of particles species  $i$ ,  $p_f$  that of the fluid,  $\mathbf{I}$  denotes the identity tensor and  $\mathbf{T}_i^E$ ,  $\mathbf{T}_f^E$  are the corresponding extra (or viscous) stress tensors. We assume that viscous effects due to the motion of the mixture are not dominant, all viscous effects are assigned to the fluid extra-stress tensor.

As the phase pressures  $p_1, \dots, p_N$  and  $p_f$  are theoretical variables which cannot be measured experimentally, they are replaced by the pore pressure  $p$  and the effective solid stress  $\sigma_e$ , which are measurable. We assume that  $\sigma_e$  is a function of the local composition of the sediment,  $\sigma_e = \sigma_e(\phi)$ . In stating the generic assumption on  $\sigma_e$ , we consider that during sedimentation, when  $\phi \leq \phi_c$ , there is no permanent contact between the particles and the momentum transfer between the particles occurs entirely through the fluid or through collisions. This means that the total stress of the mixture,  $p_t$ , which can be decomposed in two different ways as

$$p_t = p_f + p_1 + \dots + p_N = p + \sigma_e(\phi) \quad (2.38)$$

equals the pore pressure, and therefore  $\sigma_e(\phi) = 0$  for  $\phi \leq \phi_c$ . During consolidation, when  $\phi > \phi_c$ , permanent contact is established between the solid particles, and the contact forces are transmitted through solid-solid contacts. Moreover, it can be assumed that the part of the total stress supported by the skeleton of networked solid particles is an increasing function of their concentrations  $\phi$ , i.e.,  $\sigma'_e(\phi) > 0$  for  $\phi > \phi_c$ . These generic assumptions on  $\sigma_e(\phi)$  can be summarized as

$$\sigma_e(\phi) \begin{cases} = 0, & \text{for } \phi \leq \phi_c, \\ > 0, & \text{for } \phi > \phi_c, \end{cases} \quad \sigma'_e(\phi) \begin{cases} = 0, & \text{for } \phi \leq \phi_c, \\ > 0, & \text{for } \phi > \phi_c; \end{cases} \quad (2.39)$$

Now, it is possible to relate the fluid phase pressures  $p_f$  and the pore pressure as  $p_f = (1 - \phi)p$ , which (2.38) can be rewritten as

$$p_1 + \dots + p_N = \phi p + \frac{\phi_1 + \dots + \phi_N}{\phi} \sigma_e(\phi). \quad (2.40)$$

Thus, the phase pressure  $p_i$  is related to  $p$  and  $\sigma_e$  by  $p_i = (\phi_i/\phi)(\phi p + \sigma_e(\phi))$  for  $i = 1, \dots, N$ .

On the other hand, we assume that the only body force is gravity,  $\mathbf{b} = -g\mathbf{k}$ , where  $g$  is the acceleration of gravity and  $\mathbf{k}$  is the upwards-pointing unit vector. With respect to the interaction force, the solid-fluid interaction term related to species  $i$  is given by  $\mathbf{m}_i^f = \alpha_i(\Phi)\mathbf{u}_i + \beta_i(\Phi)\nabla\phi_i$ , where the  $\beta_i$  can be shown to coincide with the pore pressure  $p$ ,  $\alpha_i$  is the resistance coefficient for the transfer of momentum between the fluid and solid phase species  $i$ ,  $i = 1, \dots, N$ . Doing a dimensional analysis we obtain the following simplified linear momentum balances

$$\alpha_i(\Phi)\mathbf{u}_i = \varrho_i\phi_i g\mathbf{k} + \phi_i \nabla p + \nabla \left( \frac{\phi_i}{\phi} \sigma_e(\phi) \right), \quad i = 1, \dots, N, \quad (2.41)$$

$$\nabla p = -\varrho_f g\mathbf{k} - \frac{1}{1-\phi} (\alpha_1(\Phi)\mathbf{u}_1 + \dots + \alpha_N(\Phi)\mathbf{u}_N) + \frac{1}{1-\phi} \nabla \cdot \mathbf{T}_f^E. \quad (2.42)$$

The small viscous term  $\nabla \cdot \mathbf{T}_f^E$  is retained in (2.42) when this equation acts as an equation for the motion of mixture and is, however, deleted when (2.42) is inserted into (2.41), in order to produce a solvable linear system for the slip velocities  $\mathbf{u}_1, \dots, \mathbf{u}_N$ . Thus, the system can be written as

$$\frac{\alpha_i(\Phi)(1-\phi)}{\phi_i} \mathbf{u}_i + \sum_{j=1}^N \alpha_j(\Phi)\mathbf{u}_j = (1-\phi) \left[ (\varrho_i - \varrho_f)g\mathbf{k} + \frac{1}{\phi_i} \nabla \left( \frac{\phi_i}{\phi} \sigma_e(\phi) \right) \right], \quad i = 1, \dots, N. \quad (2.43)$$

As in [12], we obtain an explicit formula for the slip velocities  $\mathbf{u}_i$  solving the equations in (2.43) and by choosing  $\phi_i/\alpha(\Phi) = -d_i^2 V(\Phi)/(18\mu_f)$  which is consistent with the Masliyah-Lockett-Bassoon model

$$\mathbf{u}_i = -\frac{d_i^2}{18\mu_f} V(\phi) \left[ (\varrho_i - \varrho_f(\Phi))g\mathbf{k} + \frac{\sigma_e(\phi)}{\phi} \nabla \left( \frac{\phi_i}{\phi} \right) + \frac{1-\phi}{\phi} \nabla \sigma_e(\phi) \right], \quad i = 1, \dots, N, \quad (2.44)$$

where  $\mu_f$  is the viscosity of the pure fluid,  $d_i$  are the respective species diameters, and the hindered settling factor  $V(\phi)$  can be chosen as  $V(\phi) = (1-\phi)^{n-2}$  for  $n > 2$ .

The final model equations are the continuity equations of the solids species and of the mixture  $\nabla \cdot \mathbf{q} = 0$ , the linear momentum balance of the fluid (2.42), and the equations (2.44) for the slip velocities  $\mathbf{u}_i$  derived from the linear momentum balances of the solid species. To derive explicit expressions for the fluxes  $\phi_1 \mathbf{v}_1, \dots, \phi_N \mathbf{v}_N$  appearing in these equations, we introduce the reduced densities  $\bar{\varrho}_i := \varrho_i - \varrho_f$ ,  $i = 1, \dots, N$ , the vector  $\bar{\boldsymbol{\varrho}} = (\bar{\varrho}_1, \dots, \bar{\varrho}_N)^T$ , and the parameters  $\mu := -gd_1^2/(18\mu_f)$  and  $\delta_i := d_i^2/d_1^2$ ,  $i = 1, \dots, N$ , such that (2.44) reads

$$\mathbf{u}_i = \mu \delta_i V(\phi) \left[ (\bar{\varrho}_i - \bar{\boldsymbol{\varrho}}^T \Phi) \mathbf{k} + \frac{\sigma_e(\phi)}{g \phi_i} \nabla \left( \frac{\phi_i}{\phi} \right) \frac{1-\phi}{g \phi} \nabla \sigma_e(\phi) \right], \quad i = 1, \dots, N. \quad (2.45)$$

We get  $\phi_i \mathbf{v}_i = f_i^{\text{MLB}}(\Phi) \mathbf{k} + \phi_i \mathbf{q} - \mathbf{a}_i(\Phi, \nabla \Phi)$  for  $i = 1, \dots, N$ , where the components of  $f_i^{\text{MLB}}(\Phi)$  (corresponding to the MLB model for suspensions of rigid spheres) are given by

$$f_i(\Phi) = f_i^{\text{MLB}}(\Phi) = \mu \delta_i V(\phi) \left[ \delta_i (\bar{\varrho}_i - \bar{\boldsymbol{\varrho}}^T \Phi) - \sum_{k=1}^N \delta_k \phi_k (\bar{\varrho}_k - \bar{\boldsymbol{\varrho}}^T \Phi) \right], \quad i = 1, \dots, N. \quad (2.46)$$

If we let  $\boldsymbol{\delta} := (\delta_1, \dots, \delta_N)^T$ , then the vectors  $\mathbf{a}_i(\Phi, \nabla \Phi)$  are given by

$$\begin{aligned} \mathbf{a}_i(\Phi, \nabla \Phi) = & -\frac{\mu V(\phi)}{g} \left\{ \frac{(1-\phi)\phi_i}{\phi} (\delta_i - \boldsymbol{\delta}^T \Phi) \nabla \sigma_e(\phi) + \right. \\ & \left. \sigma_e(\phi) \left[ \delta_i \nabla \left( \frac{\phi_i}{\phi} \right) - \phi_i \left( \delta_1 \nabla \left( \frac{\phi_1}{\phi} \right) + \dots + \delta_N \nabla \left( \frac{\phi_N}{\phi} \right) \right) \right] \right\}, \quad i = 1, \dots, N. \end{aligned}$$

The continuity equations for the solids, i.e., for the  $N$  unknowns  $\phi_1$  to  $\phi_N$ , can then be rewritten as

$$\frac{\partial \phi_i}{\partial t} + \nabla \cdot (\phi_i \mathbf{q} + f_i^{\text{MLB}}(\Phi) \mathbf{k}) = \nabla \cdot \mathbf{a}_i(\Phi, \nabla \Phi), \quad i = 1, \dots, N. \quad (2.47)$$

Finally, in a closed one-dimensional vessel, the mixture velocity at the bottom vanishes, hence  $\mathbf{q} = 0$  and the remaining equations that actually have to be solved are the system of convection-diffusion equations

$$\frac{\partial \phi_i}{\partial t} + \frac{\partial f_i^{\text{MLB}}(\Phi)}{\partial z} = \frac{\partial}{\partial z} \left[ \mathbf{a}_i(\Phi, \frac{\partial \phi}{\partial z}) \right], \quad i = 1, \dots, N. \quad (2.48)$$

## 2.7 Conclusions of Chapter 2

Some kinematic models can be enriched with nonlinear, non-smooth and strongly degenerate diffusive terms to account for certain additional features. Their long-term fine simulations with explicit schemes is limited by the typical parabolic time step restriction, but implicit-explicit Runge-Kutta schemes can overcome this burden. We have shown that classical Newton-Raphson methods (with or without damping) for the nonlinear systems that appear in this IMEX formulation do not work for vanishing smoothness regularizations of the diffusion coefficients whereas a smart solving strategy based on the damped Newton-Raphson's method and a gradual decrease of the regularizing parameter can be efficiently applied for this purpose. The speedup of these IMEX methods with respect to the Kurganov-Tadmor (explicit) scheme is computed for some selected tests and shown to be at least an order of magnitude for moderate spatial resolutions.

The limitations of this approach stem from the fact that a regularization of the diffusion coefficients has to be selected and there is a tradeoff between the fidelity to the original coefficients and computational time. Nevertheless, we stress that the proposed regularizations do not change the strong degeneracy of the diffusion.



## Chapter 3

# A diffusively corrected multiclass Lighthill-Whitham-Richards traffic model with anticipation lengths and reaction times

### 3.1 Introduction

#### 3.1.1 Scope

The well-known Lighthill-Whitham-Richards (LWR) kinematic traffic model [75, 92] states that the density of cars  $\phi = \rho/\rho_{\max}$ , where  $\rho$  is the local number of cars per mile and  $\rho_{\max}$  is some maximum bumper-to-bumper density, can be described by the conservation law  $\partial_t \phi + \partial_x(\phi v(\phi))_x = 0$ , where  $t$  is time,  $x$  is the spatial coordinate along either an unbounded, one-directional highway or a closed circuit, and the local velocity  $v = v(x, t)$  is a given function of the local density,  $v = v(\phi(x, t))$ . It is usually assumed that  $v(\phi) = v^{\max}V(\phi)$ , where  $v^{\max}$  is the preferential velocity of drivers on a free highway and  $V$  is a hindrance function describing the drivers' behaviour of reducing speed in presence of other cars. The function  $V$  satisfies  $V(0) = 1$  and  $V'(\phi) \leq 0$ . These assumptions lead to the one-dimensional scalar conservation law

$$\partial_t \phi + \partial_x f(\phi) = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (3.1)$$

where the flux density function  $f$  is given by

$$f(\phi) = \phi v(\phi) = v^{\max} \phi V(\phi). \quad (3.2)$$

The model (3.1), (3.2) has been extended in several directions. On one hand, Nelson [81, 82] showed that introducing an anticipation length  $L$  and a reaction time  $\tau$ , replacing  $V(\phi(x, t))$  by  $V(\phi(x + L - v^{\max}V\tau, t - \tau))$  and neglecting  $\mathcal{O}(L^2 + \tau^2)$  terms when expanding the latter expression around  $(x, t)$ , one obtains a “diffusively corrected” version of (3.1), (3.2) of the following form:

$$\partial_t \phi + \partial_x f(\phi) = A(\phi)_{xx}. \quad (3.3)$$

Here,  $L$  may also depend on  $\phi$ , and under certain restrictions on  $L = L(\phi)$ ,  $\tau$  and  $v(\phi)$ , the function  $A$  is Lipschitz continuous and increasing so that the governing equation (3.3) of the diffusively corrected LWR model (“DCLWR model”) is a strongly degenerate parabolic PDE in the sense that  $A(\phi) = 0$  for  $\phi \leq \phi_c$ , where  $\phi_c$  is a critical density value (e.g., a perception threshold), and  $A'(\phi) > 0$  for  $\phi > \phi_c$ . Properties of (3.3), under the additional assumption of abruptly varying road surface conditions, were analyzed in [25]. On the other hand, Benzoni-Gavage and Colombo [8] and Wong and Wong [113] extended the LWR model (3.1), (3.2) to a multi-class model, the so-called “MCLWR model”, by distinguishing  $N$  classes of drivers associated with preferential velocities  $v_1^{\max} > v_2^{\max} > \dots > v_N^{\max}$ . For the MCLWR model, the sought quantity is the vector  $\Phi := (\phi_1, \dots, \phi_N)^T$  of the densities  $\phi_i$  of the cars of the different driver classes. The local velocity  $v_i$  of vehicles of driver class  $i$  is given by  $v_i = v_i(\phi) = v_i^{\max} V(\phi)$  for  $i = 1, \dots, N$ , where we define  $\phi := \phi_1 + \dots + \phi_N$ . Thus, the MCLWR model is given by a strongly coupled system of nonlinear first-order conservation laws of the type

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = \mathbf{0}, \quad x \in \mathbb{R}, \quad t > 0; \quad \mathbf{f}(\Phi) = (f_1(\Phi), \dots, f_N(\Phi))^T, \quad (3.4)$$

where the components of the flux vector  $\mathbf{f}(\Phi)$  are given by

$$f_i(\Phi) = \phi_i v_i(\phi) = \phi_i v_i^{\max} V(\phi), \quad i = 1, \dots, N. \quad (3.5)$$

It can be shown [50, 118] that the system (3.4), (3.5) is strictly hyperbolic for  $\Phi \in \mathcal{D}^0 := \{\Phi \in \mathbb{R}^N \mid \phi_1 > 0, \dots, \phi_N > 0, \phi < 1\}$ .

It is the purpose of this chapter to analyse the new model called diffusively corrected multi-class LWR model (“DCMCLWR model”) introduced in Section 2.2.3, by combining the assumptions of the DCLWR model with those of the MCLWR model. In particular, we associate class  $i$  of drivers with the triple  $(v_i^{\max}, L_i, \tau_i)$ ,  $i = 1, \dots, N$ , which means that drivers of different classes may have different preferential velocities, anticipation lengths, and reaction times. The resulting model, which reduces to (3.1) and (3.2) in the respective limit cases  $N = 1$  and  $\mathbf{L} = \mathbf{0}$ ,  $\boldsymbol{\tau} = \mathbf{0}$ , where  $\mathbf{L} := (L_1, \dots, L_N)^T$  and  $\boldsymbol{\tau} := (\tau_1, \dots, \tau_N)^T$ , can be cast as a quasi-linear system of second-order PDEs of the form

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = \partial_x (\mathbf{B}(\Phi) \partial_x \Phi). \quad (3.6)$$

Here the flux vector  $\mathbf{f} = \mathbf{f}(\Phi)$  is the same as in the MCLWR model, and  $\mathbf{B} = \mathbf{B}(\Phi)$  is an  $N \times N$  matrix expressing the diffusive correction. The precise functional form of  $\mathbf{B}(\Phi)$  depends on the choice of  $V(\phi)$  and the vectors of parameters  $\mathbf{v}^{\max} := (v_1^{\max}, \dots, v_N^{\max})^T$ ,  $\mathbf{L}$  and  $\boldsymbol{\tau}$ .

The system (3.6) is supplied with an initial condition and periodic boundary conditions,

$$\Phi(x, t) = \Phi(x + \mathcal{L}, t), \quad x \in \mathbb{R}$$

where  $\mathcal{L}$  is the length of a circular road. We formulate, and in part evaluate, a stability criterion for the model (3.6) based on an analysis of the eigenvalues of the matrices

$$\mathbf{M}(\Phi, \xi) := \frac{i}{\xi} \mathcal{J}_{\mathbf{f}}(\Phi) + \mathbf{B}(\Phi) \in \mathbb{C}^{N \times N}, \quad \xi \in \mathbb{R}_+, \quad (3.7)$$

where  $i = \sqrt{-1}$  and  $\mathcal{J}_f(\Phi)$  denotes the Jacobian matrix of  $f(\Phi)$ . Furthermore, by a series of numerical experiments we illustrate the behaviour of solutions to (3.6), and in particular the effect of different values of  $L_i$  and  $\tau_i$  for different classes of drivers.

### 3.1.2 Related work

To put the chapter into the proper perspective, we mention first that the MCLWR model has been analyzed in a number of papers including [9, 51, 78, 118]. In particular, its hyperbolicity has been established [51, 118] and the admissible waves of the Riemann problem have been investigated [118]. Moreover, the model (3.4), (3.5) admits a separable, strictly convex entropy since the corresponding Jacobian matrix  $\mathcal{J}_f(\Phi)$  is diagonally symmetrizable [8, 9]. Component-wise or characteristic high-resolution numerical schemes for (3.4), (3.5) involving weighted essentially non-oscillatory (WENO) flux reconstructions are advanced in [28, 50, 117, 119]. On the other hand, particularly simple first- and second-order difference schemes for the same problem that rely on the structure of the fluxes  $f_i$  (3.5) along with the definite sign of the velocities  $v_i$  are introduced in [24]. Variants of the original MCLWR model (in the sense of [8, 113]) have been proposed and in part analyzed for highways with varying road surface conditions [27, 120, 121], traffic flow on networks [63, 84], and stochastic fundamental diagrams (equivalent to the velocity functions  $v_i$ ) [85].

To provide justification for the presence of second-order terms in (3.6), we mention that Nelson and Sopasakis [83] show that for one driver class, the first-order Chapman-Enskog expansion of the classical Prigogine-Herman [90] kinetic equation of vehicular traffic leads to a traffic stream model that can be expressed by the DCLWR equation (3.3). Thus, as is argued in [81], a DCLWR model is presumably the proper traffic-theoretical analogue of the Navier-Stokes equations of fluid dynamics. Moreover, Lighthill and Whitham themselves, at the end of their paper [75], propose to include a diffusion effect “due to the fact that each driver’s gaze is concentrated on the road in front of him, so that he adjusts his speed to the concentration slightly ahead” [75, p. 344]. Since the model development in Section 3.2 closely follows the original calculus by Nelson [81] in the case  $N = 1$ , the same arguments are valid for the present systems case ( $N \geq 1$ ). That said, we are well aware of the well-known general criticism by Daganzo [43] of second-order traffic models, who argues that diffusion terms may cause “wrong way travel” in determined situations. However, we demonstrate in Example 3.14 of Section 3.4.6 that the nonlinearity of the diffusion terms in (3.6) in conjunction with the (mild) assumption  $V'(0) = V'(1) = 0$  prevents this phenomenon from occurring.

Several alternative approaches have been pursued to extend the LWR model to finite reaction times and anticipation lengths. The treatment by Sopasakis and Katsoulakis [105] (see also [70]) for one driver class leads to a scalar conservation law with a non-local flux involving a non-symmetric “anticipation kernel”. In [86] a linear stability analysis is applied to a second-order macroscopic local traffic model, and a corrected “effective density” sensor accounts for aggressive or timid drivers. A related analysis is presented in [101]. Ngoduy and Tampere [87] study the influence of different reaction times (of a single driver class) in terms of the same model. Their condition for traffic

stability reads [87, Eq. (39)]

$$\tau < \frac{1}{2\phi^2|V'(\phi)|_{v^{\max}}} \quad \text{for } 0 < \phi < 1 \quad (3.8)$$

(in our notation). If this condition is violated, then their model can develop instabilities that can be considered as stop-and-go waves. The relation between reaction times and anticipation lengths and traffic stability is also discussed in [110].

Finally, we mention that other kinematic flow models that give rise to systems of the type (3.6) include the sedimentation of polydisperse suspensions [12] and the settling and creaming of dispersions of droplets [1]. These models are typically posed with zero-flux boundary conditions on a bounded  $x$ -interval.

## 3.2 Diffusively Corrected MCLWR model DCMCLWR

### 3.2.1 The model equations

We consider the DCMCLWR model introduced in Section 2.2.3, equations (2.1),(2.9) and where the coefficients of the diffusive matrix  $B$  are given by (2.11). In this chapter we focus on the hindrance functions  $V$  according to Dick [48] and Greenberg [59], namely

$$V(\phi) = \min\{1, -C \ln \phi\} = \begin{cases} 1 & \text{for } \phi \leq \phi_{\text{DG}}, \\ -C \ln \phi & \text{for } \phi > \phi_{\text{DG}} \end{cases} \quad (3.9)$$

with a parameter  $C > 0$ , where we employ the common value  $C = e/7 \approx 0.38833$  such that  $\phi_{\text{DG}} = \exp(-1/C) \approx 0.076142$ . Alternatively we use the common linear Greenshields (GS) velocity function

$$V(\phi) = 1 - \phi. \quad (3.10)$$

Note that in view of (2.11), the particular form of the velocity function (3.9) implies that  $B(\Phi) = 0$  for  $\phi \leq \phi_c = \phi_{\text{DG}}$ , so (3.6) degenerates to the first-order system (3.4) for the corresponding vectors  $\Phi$ . Thus, the resulting model is strongly degenerate. In general, and following [95], we assume that  $\phi_c$  is an explicitly known perception threshold or critical density such that the drivers' reaction is instantaneous in relatively free flow, i.e. when  $\phi \leq \phi_c$ , and otherwise is modeled by the diffusion term. Thus, for a unified treatment we assume that

$$\mathbf{B}(\Phi) = (B_{ij}(\Phi))_{i,j=1,\dots,N}, \quad \text{where } B_{ij}(\Phi) = \begin{cases} 0 & \text{if } \phi \leq \phi_c, \\ \alpha_{ij}(\Phi) & \text{if } \phi > \phi_c. \end{cases} \quad (3.11)$$

### 3.2.2 Stability analysis

We perform a linearized stability analysis for the system (3.6) under the assumptions of the DCMCLWR model. The linearized equation for a small perturbation  $\mathbf{u}$  about a constant state  $\Phi^{(0)}$

is obtained by substituting  $\Phi = \Phi^{(0)} + \mathbf{u}$  into (3.6) and neglecting quadratic terms in  $\mathbf{u}$ . This yields the following linearized version of (3.6):

$$\partial_t \mathbf{u} + \mathbf{J} \partial_x \mathbf{u} = \mathbf{B} \partial_x^2 \mathbf{u}, \quad \text{where } \mathbf{J} := \mathcal{J}_f(\Phi^{(0)}), \mathbf{B} := \mathbf{B}(\Phi^{(0)}). \quad (3.12)$$

We now seek solutions of (3.12) of the form  $\mathbf{u}(x, t) = \mathbf{z}(t; \xi) \exp(i\xi x)$  for a fixed frequency  $\xi$ . The vector function  $\mathbf{z}$  satisfies the system of ordinary differential equations

$$\mathbf{z}' = -\xi^2 \mathbf{M} \mathbf{z}, \quad (3.13)$$

where  $' \equiv d/dt$  and  $\mathbf{M} = \mathbf{M}(\Phi^{(0)}, \xi)$  is the matrix defined in (3.7). The general solution of (3.13) is of the well-known form

$$\mathbf{z}(t; \xi) = \sum_{j=1}^r \exp(-\xi^2 \lambda_j t) \mathbf{q}_j(t, \xi), \quad (3.14)$$

where  $\lambda_1, \dots, \lambda_r$  are the eigenvalues of  $\mathbf{M}$  that appear in its associated Jordan blocks of corresponding sizes  $m_1, \dots, m_r$ , where  $m_1 + \dots + m_r = N$ , and  $\mathbf{q}_j$  are polynomials (with vectorial coefficients related to the Jordan decomposition basis) of degree less or equal  $m_j - 1$ . If  $\lim_{t \rightarrow +\infty} |z(t; \xi)| < \infty$ , then  $\text{Re}(\lambda_j) \geq 0$  if  $m_j = 1$  or  $\text{Re}(\lambda_j) > 0$  if  $m_j > 1$ .

With respect to the instability phenomena predicted by eigenvalue analysis of  $\mathbf{J}$  and  $\mathbf{B}$ , we mention first that if  $\mathbf{B}$  has an eigenvalue  $\lambda$  with  $\text{Re}(\lambda) < 0$ , then it turns out that  $\mathbf{M}$  will do so for  $|\xi| > \xi_0$ , for some  $\xi_0$ . It would then follow that this would trigger a growth of  $\mathbf{z}(t; \xi)$  in (3.14) when  $t \rightarrow \infty$  for  $|\xi| > \xi_0$ . This would completely ruin the solution of the nonlinear system, for these oscillations would appear in all frequencies above  $\xi_0$ . Milder instabilities would be expected if  $\mathbf{B}$  has eigenvalues with positive real parts but  $\mathbf{M}$  does not, since this should only hold for relatively small values of  $\xi$ . These phenomena are illustrated in the numerical examples.

Considering separately the two terms of the matrix  $\mathbf{M}$  (cf. (3.7)), namely  $(i/\xi)\mathbf{J}$  and  $\mathbf{B}$ , we obtain that the linearized stability condition for (3.6) when  $\mathbf{B} = 0$  is exactly the hyperbolicity condition for the resulting system ( $\xi$  can take any sign), whereas the linearized stability condition when  $\mathbf{f} = \mathbf{0}$  is directly inherited by the condition on  $\mathbf{M}$ , i.e., the eigenvalues of  $\mathbf{B} = \mathbf{B}(\Phi^{(0)})$  should have non-negative real parts if they are simple and strictly positive real parts if they have some corresponding Jordan block of non-trivial dimension. Unfortunately,  $\mathbf{J}$  having real eigenvalues and  $\mathbf{B}$  having eigenvalues with strictly positive real parts does not imply that eigenvalues of  $\mathbf{M}$  have strictly positive real parts, as the following simple counterexample shows: with  $\xi = 1$  and

$$\mathbf{J} = \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & -3 \\ 3 & 3 \end{bmatrix},$$

the eigenvalues of  $\mathbf{M} = i\mathbf{J} + \mathbf{B}$  are  $-0.2332 - 2.2436i$  and  $2.2332 + 4.2436i$ , whereas the eigenvalues of  $\mathbf{B}$  are  $1 \pm 2.2361i$ .

This discussion illustrates that the satisfaction of the stability criterion stipulated by (3.14), namely that the pairwise distinct eigenvalues  $\lambda_1, \dots, \lambda_r$  of the matrix  $\mathbf{M} = \mathbf{M}(\Phi^{(0)}, \xi)$  satisfy

$$\text{Re } \lambda_1 \leq 0, \dots, \text{Re } \lambda_r \leq 0, \quad (3.15)$$

can in general not be evaluated exactly by analyzing  $\mathbf{J}$  and  $\mathbf{B}$  separately. However, some special cases are tractable. These include the DCMCLWR with drivers having the same maximum speed so that classes of drivers are distinguished by their reaction times and anticipation lengths (see Sect. 3.2.3).

### 3.2.3 DCLWR model with drivers having the same maximum speed

Let us consider the model (3.5), (3.6), (3.11) under the assumption

$$v_1^{\max} = \dots = v_N^{\max} =: v^{\max}. \quad (3.16)$$

This means that the classes of vehicles are distinguished only by the drivers' reaction times  $\tau_i$  and anticipation lengths  $L_i$ . Under the assumption (3.16), and defining the constant vector  $\mathbf{e} := (1, \dots, 1)^T$ ,  $\mathbf{D}_\tau := \text{diag}(\tau_1, \dots, \tau_N)$  and  $\mathbf{D}_L := \text{diag}(L_1, \dots, L_N)$ , we obtain

$$\begin{aligned} \mathcal{J}_f(\Phi) &= v^{\max} (V(\phi)\mathbf{I} + V'(\phi)\Phi\mathbf{e}^T), \\ \mathbf{B}(\Phi) &= -V'(\phi)v^{\max} (v^{\max}\phi V'(\phi)\mathbf{D}_\tau\Phi + \mathbf{D}_L\Phi)\mathbf{e}^T. \end{aligned} \quad (3.17)$$

Under the present assumptions, and setting  $\Phi := \Phi^{(0)}$  and  $\phi := \phi^{(0)}$ , we obtain

$$\mathbf{M} = v^{\max} \left[ \frac{i}{\xi} V(\phi)\mathbf{I} + V'(\phi) \left( \frac{i}{\xi} \Phi - (v^{\max}\phi V'(\phi)\mathbf{D}_\tau + \mathbf{D}_L)\Phi \right) \mathbf{e}^T \right].$$

This matrix is a rank-one perturbation of a multiple of the identity matrix  $\mathbf{I}$ , and its eigenvalues are given by

$$\begin{aligned} \tilde{\lambda}_1 &= v^{\max} \left[ \frac{i}{\xi} V(\phi) + V'(\phi) \left( \frac{i}{\xi} \phi - (v^{\max}\phi V'(\phi)\boldsymbol{\tau}^T + \mathbf{L}^T)\Phi \right) \right], \\ \tilde{\lambda}_2 &= \dots = \tilde{\lambda}_N = v^{\max} \frac{i}{\xi} V(\phi), \end{aligned} \quad (3.18)$$

with the corresponding one- and  $(N-1)$ -dimensional eigenspaces

$$\begin{aligned} \mathcal{V}_1 &= \left\{ \mathbf{w} \in \mathbb{C}^N : \mathbf{w} = \alpha \left( \frac{i}{\xi} \Phi - (v^{\max}\phi V'(\phi)\mathbf{D}_\tau + \mathbf{D}_L)\Phi \right), \alpha \in \mathbb{R} \right\}, \\ \mathcal{V}_{2,\dots,N} &= \{ \mathbf{w} \in \mathbb{R}^N : \mathbf{e}^T \mathbf{w} = 0 \}, \end{aligned}$$

so that all Jordan blocks are trivial. On the other hand, the rank-one matrix  $\mathbf{B}(\Phi)$  defined by (3.17) has the eigenvalues

$$\beta_1 = -V'(\phi)v^{\max} (v^{\max}\phi V'(\phi)\boldsymbol{\tau}^T + \mathbf{L}^T)\Phi, \quad \beta_2 = \dots = \beta_N = 0,$$

which are the real parts of  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_N$  given by (3.18). We have proved the following lemma.

**Lemma 3.1** *Under the assumption (3.16), the stability criterion (3.15) is violated for a vector  $\Phi := \Phi^{(0)}$ , i.e. the matrix  $\mathbf{M}$  has an eigenvalue  $\mu$  with  $\text{Re } \mu > 0$ , if and only if  $V'(\phi) < 0$  and*

$$(v^{\max}\phi V'(\phi)\boldsymbol{\tau}^T + \mathbf{L}^T)\Phi < 0, \quad (3.19)$$

that is, if the matrix  $\mathbf{B}$  has a negative eigenvalue  $\beta_1$ .

### 3.3 Parabolicity analysis

In Theorem 2.2 and Theorem 2.3 in Section 2.2.4 was established the conditions of stability for convective and diffusive terms respectively that appearing in (3.6). Now we remark that in the stability criterion for the diffusive term, it is observed in Theorem 2.3 that for quantities in (2.14),  $C_1^2 \neq 4C_2$  generically, we henceforth use  $C_1, C_2 > 0$  as a sufficient condition for  $\mathbf{B}(\Phi)$  to have eigenvalues with non-negative real parts only. From the next result we can obtain reaction times that ensure that the matrix  $\mathbf{B}(\Phi)$  has eigenvalues with non-negative real parts only with velocity functions  $V(\phi)$  that satisfy

$$V(\phi) = \begin{cases} 1 & \text{for } \phi \leq \phi_c, \\ W(\phi) & \text{for } \phi > \phi_c, \end{cases} \quad (3.20)$$

where  $W$  is a function that satisfies  $W(\phi_c) = 1$ ,  $W(1) = 0$  and  $W'(\phi) < 0$  for  $\phi_c < \phi < 1$ . For instance,  $W(\phi) = -C \ln \phi$  with  $\phi_c = \phi_{\text{DG}}$  gives the Dick-Greenberg model (3.9) and  $W(\phi) = (1 - \phi)/(1 - \phi_v)$  produces a variant of the Greenshields model (3.10).

The following Theorem, a relationship between reaction time and free velocity is established.

**Theorem 3.2** *Let  $v_1^{\max}, \dots, v_N^{\max}$  be free velocities such that  $v_1^{\max} > v_2^{\max} > \dots > v_N^{\max}$ ,  $V(\phi)$  a velocity function that satisfies (3.20) and*

$$L_i = L_i(\phi) = \max\{L_{\min}, \beta(v_i^{\max} V(\phi))^2\}, \quad (3.21)$$

where the parameters  $L_{\min}, \beta > 0$  are chosen such that

$$L_{\min} \leq \beta(v_i^{\max})^2. \quad (3.22)$$

Then there exist reaction times  $\tau_i = \tau(v_i^{\max})$  for  $i = 1, \dots, N$ , where  $\tau(v)$  is a monotone increasing function and  $\tau_1 \leq L_{\min}/(v_1^{\max} \kappa)$ , with  $\kappa = \max_{\phi} |\phi V'(\phi)|$ , such that the matrix  $\mathbf{B}(\Phi)$  has eigenvalues with positive real part.

Note that (3.21) is a multiclass version of the equation

$$L(\phi) = \max\left\{L_{\min}, \frac{(v(\phi))^2}{2a}\right\}$$

proposed in [82] for  $N = 1$ , where  $a$  is the deceleration and  $L_{\min}$  is a minimum anticipation distance (regardless of how small the velocity is).

**Proof.**[Proof of Theorem 3.2] We consider  $\phi > \phi_c$ , since  $\mathbf{B}(\Phi) = 0$  otherwise. By Theorem 2.3 it is ensured that  $\mathbf{B}(\Phi)$  has eigenvalues with non-negative real parts only when  $C_1, C_2 > 0$ , and this is in turn guaranteed when

$$L_k + \tau_k S(\Phi) \geq 0 \quad \text{for all } \Phi \text{ and } k = 1, \dots, N, \quad S(\Phi) := V'(\phi) \Phi^T \mathbf{v}^{\max}, \quad (3.23)$$

$$\Delta_{ij} := \frac{L_i}{\tau_i} - \frac{L_j}{\tau_j} + (v_j^{\max} - v_i^{\max}) V(\phi) \leq 0 \quad \text{for all } \Phi \text{ and } 1 \leq i < j \leq N. \quad (3.24)$$

Let  $\phi = \phi_i^*$  be the unique solution of  $L_{\min} = \beta(v_i^{\max})^2 V(\phi)^2$ , then

$$L_i(\phi) = \begin{cases} \beta(v_i^{\max})^2 V(\phi)^2 & \text{for } \phi \leq \phi_i^*, \\ L_{\min} & \text{for } \phi \geq \phi_i^*. \end{cases} \quad (3.25)$$

From the assumption (3.22) we deduce that  $\phi_c \leq \phi_i^*$  for  $i = 1, \dots, N$ . Furthermore,  $\phi_i^* \geq \phi_j^*$  for  $i < j$ . Moreover,  $S(\Phi) = V'(\phi)\Phi^T \mathbf{v}^{\max} \geq V'(\phi)\phi v_1^{\max}$  implies that a sufficient condition for (3.23) to hold is given by

$$\tau_i \leq \frac{L_{\min}}{\kappa v_1^{\max}} \quad \text{for } i = 1, \dots, N. \quad (3.26)$$

We consider now condition (3.24). From (3.20) and (3.25) we get

$$\frac{L_i}{\tau_i} - v_i^{\max} V(\phi) = \begin{cases} \frac{\beta(v_i^{\max})^2}{\tau_i} - v_i^{\max} & \text{if } \phi \leq \phi_c, \\ \frac{\beta(v_i^{\max})^2 W(\phi)^2}{\tau_i} - v_i^{\max} W(\phi) & \text{if } \phi_c \leq \phi \leq \phi_i^*, \quad i = 1, \dots, N, \\ \frac{L_{\min}}{\tau_i} - v_i^{\max} W(\phi) & \text{if } \phi_i^* \leq \phi \leq 1, \end{cases}$$

We consider a pair of indices  $i < j$  and discuss the cases determined by  $\phi$  belonging to  $[\phi_c, \phi_j^*]$  (Case 1),  $[\phi_j^*, \phi_i^*]$  (Case 2) or  $[\phi_i^*, 1]$  (Case 3).

In Case 1, if we use the functional form  $\tau_i = \tau(v_i^{\max})$ , for some  $\tau$  to be determined, and denote

$$\psi(v, \phi) := \frac{\beta(vW(\phi))^2}{\tau} - vW(\phi),$$

then  $\Delta_{ij} = \psi(v_i^{\max}, \phi) - \psi(v_j^{\max}, \phi)$ . If  $\psi_v \leq 0$  then  $\Delta_{ij} \leq 0$  for  $i < j$ , and this is equivalent to

$$\tau' \geq \frac{2\tau}{v} - \frac{\tau^2}{v^2 \tilde{\beta}}, \quad \tilde{\beta} := \tilde{\beta}(\phi) := \beta W(\phi).$$

We consider  $\mu(v) = v^n \tau(v)$ , with  $n$  to be determined so to simplify the latter expression:

$$\mu' = nv^{n-1}\tau + v^n \tau' \geq nv^{n-1}\tau + v^n \left( \frac{2\tau}{v} - \frac{\tau^2}{v^2 \tilde{\beta}} \right) = nv^{n-1}\tau + 2v^{n-1}\tau - \frac{v^{n-2}\tau^2}{\tilde{\beta}}.$$

We take  $n = -2$  so that this expression yields:

$$\mu' \geq -\frac{v^{-4}\tau^2}{\tilde{\beta}} = -\frac{\mu^2}{\tilde{\beta}} \implies \frac{\mu'}{\mu^2} = -\left(\frac{1}{\mu}\right)' \geq -\frac{1}{\tilde{\beta}} = -\left(\frac{v}{\tilde{\beta}}\right)',$$

and, upon integration and some algebra,

$$\tau = v^2 \mu \geq \frac{\tilde{\beta} v^2}{v + \mathcal{A}}, \quad (3.27)$$

for some positive  $\mathcal{A}$  to avoid null denominators.

In Case 2, and taking into account that  $L_{\min} \geq \beta(v_j^{\max})^2 W^2$  for  $\phi \geq \phi_j^*$ , we get

$$\begin{aligned} \Delta_{ij} &= \frac{L_i}{\tau_i} - v_i^{\max} W(\phi) - \left( \frac{L_j}{\tau_j} - v_j^{\max} W(\phi) \right) \\ &= \frac{\beta(v_i^{\max})^2 W(\phi)^2}{\tau_i} - v_i^{\max} W(\phi) - \left( \frac{L_{\min}}{\tau_j} - v_j^{\max} W(\phi) \right) \\ &\leq \frac{\beta(v_i^{\max})^2 W(\phi)^2}{\tau_i} - v_i^{\max} W(\phi) - \left( \frac{\beta(v_j^{\max})^2 W(\phi)^2}{\tau_j} - v_j^{\max} W(\phi) \right), \end{aligned}$$

As in Case 1, if  $\tau$  satisfies (3.27) then  $\Delta_{ij} \leq 0$ .

In Case 3 we get

$$\begin{aligned} 0 &\geq \frac{L_{\min}}{\tau_i} - v_i^{\max} W(\phi) - \left( \frac{L_{\min}}{\tau_j} - v_j^{\max} W(\phi) \right) \\ &= L_{\min} \left( \frac{1}{\tau_i} - \frac{1}{\tau_j} \right) - W(\phi) (v_i^{\max} - v_j^{\max}). \end{aligned} \tag{3.28}$$

Since  $W(1) = 0$  and  $v_i^{\max} > v_j^{\max}$ , then (3.28) holds if and only if  $\tau_i \geq \tau_j$ .

Recapitulating, we deduce that (3.24) holds if  $\tau_i = \tau(w_i)$  with  $\tau$  satisfying

$$\tau \geq \frac{\beta W(\phi) v^2}{v + \mathcal{A}}, \quad \text{for all } \phi, \quad \tau' \geq 0,$$

and, since  $W(\phi) \leq 1$ , this is equivalent to

$$\tau \geq \frac{\beta v^2}{v + \mathcal{A}}, \quad \tau' \geq 0. \tag{3.29}$$

We consider the increasing function

$$\tilde{\tau}(v) := \frac{\beta v^2}{v + \mathcal{A}}$$

for  $\mathcal{A} > 0$  that certainly satisfies (3.29). Therefore, to ensure that  $\mathbf{B}(\Phi)$  with  $\tau_i = \tilde{\tau}(w_i)$  has eigenvalues with non-negative real parts only, we use (3.26), so we need to find conditions on  $\mathcal{A}$  so that

$$\frac{\beta v^2}{v + \mathcal{A}} \leq \frac{L_{\min}}{\kappa v_1^{\max}} \iff \beta v^2 - \frac{L_{\min}}{\kappa v_1^{\max}} (v + \mathcal{A}) \leq 0 \quad \text{for all } 0 \leq v \leq v_1^{\max}. \tag{3.30}$$

The roots of

$$\beta v^2 - \frac{L_{\min}}{\kappa v_1^{\max}} (v + \mathcal{A}) = 0$$

are

$$w_{\pm} = \frac{1}{2\beta} \left[ \frac{L_{\min}}{\kappa v_1^{\max}} \pm \left( \left( \frac{L_{\min}}{\kappa v_1^{\max}} \right)^2 + 4\beta \frac{L_{\min}}{\kappa v_1^{\max}} \mathcal{A} \right)^{1/2} \right]$$

and (3.30) will hold if  $w_- \leq 0$  (which is true) and  $v_1^{\max} \leq w_+$ , which yields after some algebraic manipulations

$$\beta(v_1^{\max})^2 \leq \frac{L_{\min}}{\kappa} + \frac{L_{\min}}{\kappa v_1^{\max}} \mathcal{A} \implies \mathcal{A} \geq \left( \frac{\kappa \beta (v_1^{\max})^2}{L_{\min}} - 1 \right) v_1^{\max} > 0.$$

This concludes the proof. ■

### 3.4 Numerical results

In the subsequent series of examples, we solve the system (3.6) numerically for  $0 \leq t \leq T$  and  $0 \leq x < K$  along with the initial and periodic boundary conditions

$$\Phi(x, 0) = \Phi_0(x), \quad 0 \leq x < K; \quad \Phi(0, t) = \Phi(K, t) \quad \text{for } 0 \leq t \leq T,$$

corresponding to a circular one-directional road of length  $K$ . Numerical approximations are obtained by the Kurganov-Tadmor (KT) scheme [71] applied to the DCMCLWR model. In fact, Kurganov and Tadmor [71] explicitly propose a version of their scheme for convection-diffusion problems of the type (3.6), even though a well-posedness analysis for systems of PDEs of this type is not available in the strongly degenerate case. In [12] the same method was applied to (3.6) in the context of a model of polydisperse sedimentation. To further support the use of the KT scheme, we mention that numerical experiments conducted in [31] indicate that the KT scheme, a class of schemes introduced in [24] and based on MUSCL-type spatial differencing and Runge-Kutta temporal differencing and an alternative implicit-explicit (IMEX) scheme designed for (3.6) that involves a spectral WENO scheme for the convective part converge to the same solution of (3.6) as  $\Delta t, \Delta x \rightarrow 0$  (under suitable CFL conditions). In some examples we will compare the performance of the KT scheme with Scheme 10 introduced in [24] which is originally defined for (3.4) and is adapted to (3.6) by adding the same discretization of  $\partial_x(\mathbf{B}(\Phi)\partial_x\Phi)$  as that of the KT scheme, also we will compare with one of the same class of schemes introduced in [31] and described in Section 2.3.5, namely the scheme IMEX-ARS(3,4,3), which is order of accuracy 3 and is defined by the Butcher array

$$\mathbf{D} = \begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \gamma & 0 & \gamma & 0 & 0 \\ \frac{1+\gamma}{2} & 0 & \frac{1-\gamma}{2} & \gamma & 0 \\ 1 & 0 & b_1(\gamma) & b_2(\gamma) & \gamma \\ \hline & 0 & b_1(\gamma) & b_2(\gamma) & \gamma \end{array}, \quad \hat{\mathbf{D}} = \begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 & 0 \\ \frac{1+\gamma}{2} & \hat{a}_{31}(\gamma) & \hat{a}_{32}(\gamma) & 0 & 0 \\ 1 & \hat{a}_{41}(\gamma) & \hat{a}_{42} & \hat{a}_{43} & 0 \\ \hline & 0 & b_1(\gamma) & b_2(\gamma) & \gamma \end{array}, \quad (3.31)$$

where  $\gamma$  is the middle root of  $6x^3 - 18x^2 + 9x - 1 = 0$ , and

$$\begin{aligned} b_1(\gamma) &= -\frac{3}{2}\gamma^2 + 4\gamma - \frac{1}{4}, & b_2(\gamma) &= \frac{3}{2}\gamma^2 - 5\gamma + \frac{5}{4}, \\ \hat{a}_{31}(\gamma) &= \left(1 - \frac{9}{2}\gamma + \frac{3}{2}\gamma^2\right)\hat{a}_{42} + \left(\frac{11}{4} - \frac{21}{2}\gamma + \frac{15}{4}\gamma^2\right)\hat{a}_{43} - \frac{7}{2} + 13\gamma - \frac{9}{2}\gamma^2, \\ \hat{a}_{32}(\gamma) &= \left(-1 + \frac{9}{2}\gamma - \frac{3}{2}\gamma^2\right)\hat{a}_{42} + \left(-\frac{11}{4} + \frac{21}{2}\gamma - \frac{15}{4}\gamma^2\right)\hat{a}_{43} + 4 - \frac{25}{2}\gamma + \frac{9}{2}\gamma^2, \\ \hat{a}_{41}(\gamma) &= 1 - \hat{a}_{42} - \hat{a}_{43}. \end{aligned}$$

In the following numerical examples, and unless otherwise stated, the  $x$ -interval  $[0, K]$  is subdivided into  $M = 3200$  subintervals of length  $\Delta x = K/M$ . We denote by  $\Delta t$  the time step used to advance the numerical solution from time  $t = t_n$  to  $t_{n+1} = t_n + \Delta t$  and by  $\Phi_j^n$  the vector of numerical solutions associated with cell  $[j\Delta x, (j+1)\Delta x)$ ,  $j = 0, \dots, M-1$ , at time  $t_n$ . For each iteration, the time step  $\Delta t$  is determined anew by using the following formula (derived from a linearized CFL condition):

$$\frac{\Delta t}{\Delta x} \max_{0 \leq j \leq M-1} \varrho(\mathcal{J}_f(\Phi_j^n)) + \frac{\Delta t}{2\Delta x^2} \max_{0 \leq j \leq M-1} \varrho(\mathbf{B}(\Phi_j^n)) = C_{\text{cfl1}}$$

for the KT scheme and Scheme 10 of [24], and

$$\frac{\Delta t}{\Delta x} \max_{0 \leq j \leq M-1} \varrho(\mathcal{J}_f(\Phi_j^n)) = C_{\text{cfl2}}$$

for IMEX-ARS(3,4,3), where  $\varrho(\cdot)$  is the spectral radius.

### 3.4.1 Example 3.1 (DG model, $N = 4$ , stable behaviour).

In Example 3.1 we consider the DG velocity function (3.9), a circular road of length  $K = 10$  mi,  $N = 4$  driver classes with the respective preferential velocities  $v_1^{\max} = 60$  mi/h,  $v_2^{\max} = 55$  mi/h,  $v_3^{\max} = 50$  mi/h and  $v_4^{\max} = 45$  mi/h, and a uniform minimum anticipation length  $L_{\min} = 0.03$  mi. The reaction times are chosen such that the eigenvalues of the diffusion matrix  $\mathbf{B}(\Phi)$  have non-negative real parts for  $\Phi \in \mathcal{D}^0 \subset \mathbb{R}^4$ . According to (2.14) this is ensured if the parameters  $\tau_1, \dots, \tau_N$  satisfy the following condition:

$$\tau_1 \leq \frac{L_{\min}}{Cv_1^{\max}}; \quad \tau_i \leq \left( \frac{v_i^{\max}}{v_{i-1}^{\max}} \right)^2 \tau_{i-1}, \quad i = 2, \dots, N. \quad (3.32)$$

To satisfy (3.32) here, we choose  $\tau_1 = 0.0013$  h,  $\tau_2 = 0.0011$  h,  $\tau_3 = 0.0008$  h and  $\tau_4 = 0.0006$  h. Figure 3.1 shows the evolution of the initial traffic ‘‘platoon’’ given by

$$\Phi_0(x, 0) = p(x) \begin{pmatrix} 0.2 \\ 0.3 \\ 0.2 \\ 0.3 \end{pmatrix}, \quad p(x) = \begin{cases} 10x & \text{for } 0 < x \leq 0.1, \\ 1 & \text{for } 0.1 < x \leq 0.9, \\ -10(x-1) & \text{for } 0.9 < x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (3.33)$$

We observe that the system tends to a stationary constant solution.

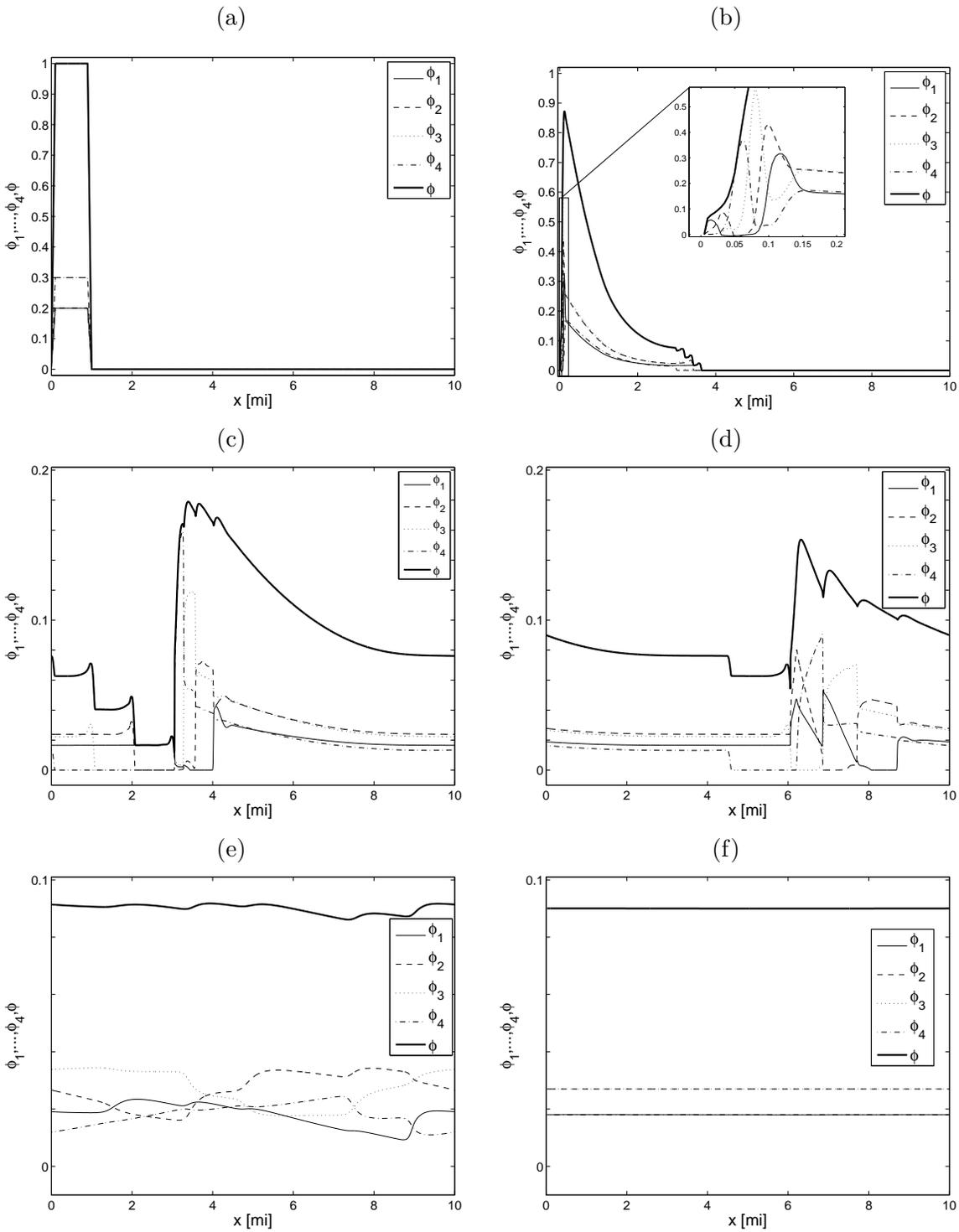


Figure 3.1: Example 3.1 (DG model,  $N = 4$ ): (a) initial datum (3.33) and (b–f) numerical solution at simulated times (b)  $t = 0.08$  h, (c)  $t = 0.2$  h, (d)  $t = 0.3$  h, (e)  $t = 9$  h and (f)  $T = 50.0$  h.

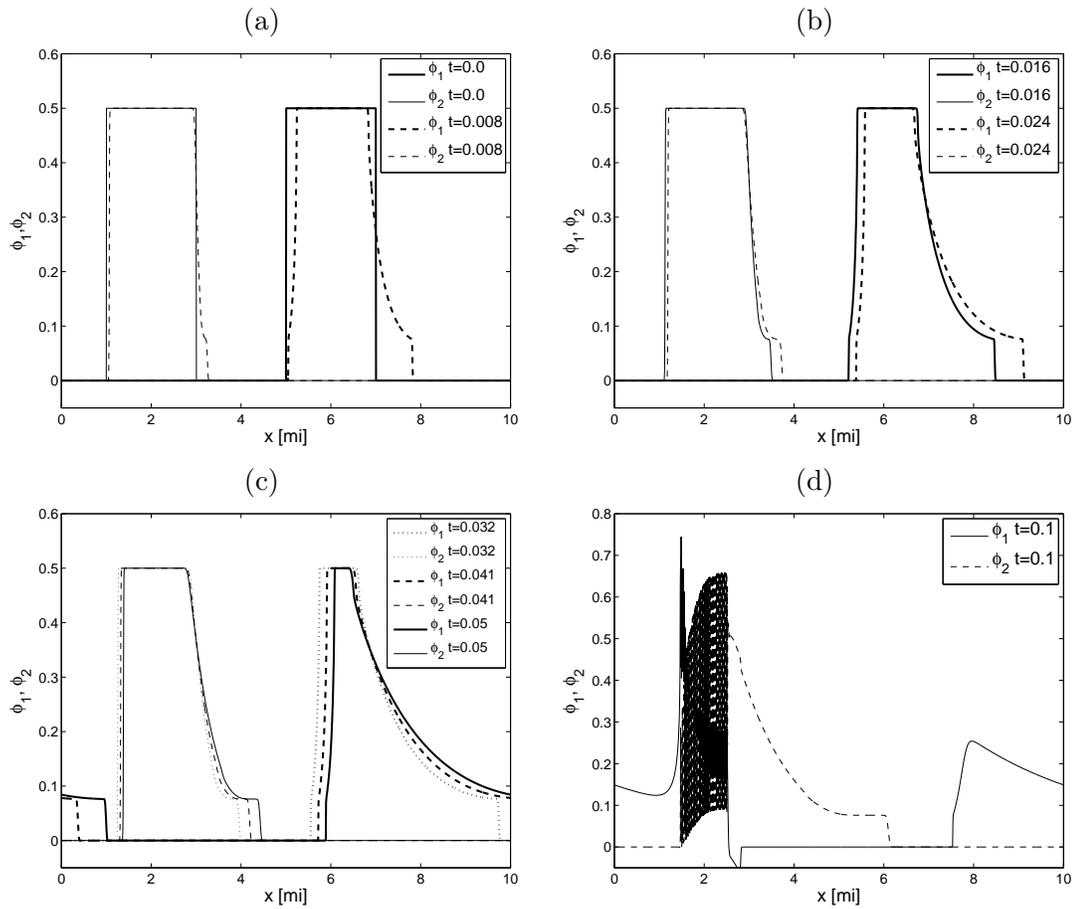


Figure 3.2: Example 3.2 (DG model,  $N = 2$ ): (a, b, c) stable behaviour for individual driver classes (spatially separated), followed by (d) unstable behavior for mixed driver classes.

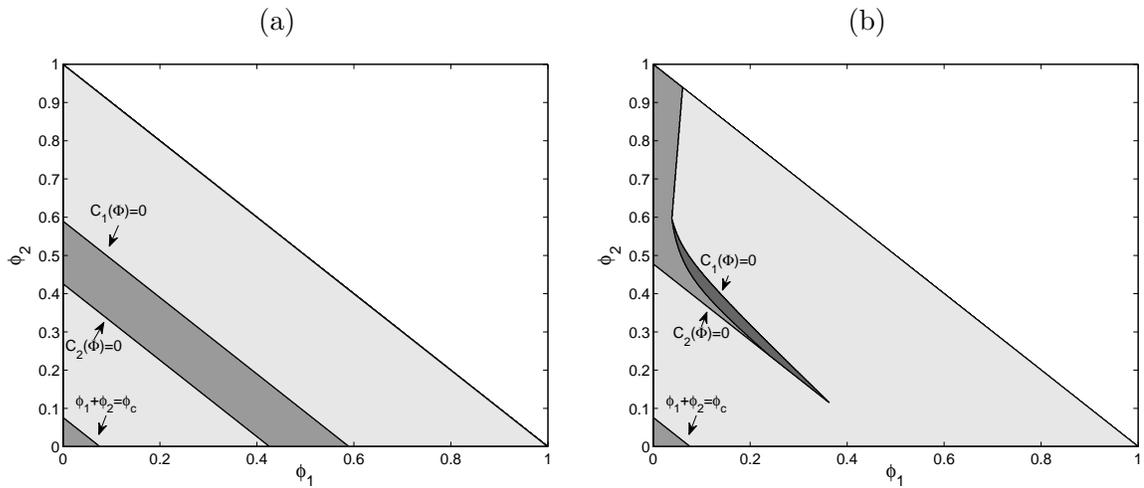


Figure 3.3: Examples 3.2–3.8 (DG model,  $N = 2$ ). Stability region for the diffusion matrix  $B$  and instability region for  $M$  for  $\xi \in [0, 100]$  for (a) Examples 3.2–3.5, (b) Examples 3.6–3.8.

### 3.4.2 Examples 3.2–3.5 (DG model, $N = 2$ , unstable behaviour)

In Examples 3.2 to 3.5 we consider the DG model (3.9), a circular road of length  $K = 2$  mi, and  $N = 2$  driver classes. The preferential velocities of the two classes are given by  $v_1^{\max} = 80$  mi/h and  $v_2^{\max} = 30$  mi/h, and a minimum anticipation distance  $L_{\min} = 0.03$  mi. For Example 3.2, the parameters  $\tau_1 = 0.00096$  h and  $\tau_2 = 0.0025$  h have been chosen in such a way that the condition for the PDE (3.3) (for  $N = 1$ ) to be parabolic for  $\phi > \phi_c$ , namely

$$\tau \leq \frac{L}{|\phi V'(\phi)|v^{\max}} \quad \text{for } \phi_c \leq \phi < 1, \quad (3.34)$$

is satisfied by both triples  $(v^{\max}, L, \tau) = (v_i^{\max}, L_i, \tau_i)$ ,  $i = 1, 2$ , but that at the same time  $C_2 < 0$  in a subregion of  $\mathcal{D}^0$ . In Figure 3.2 we show a numerical example obtained for these values of parameters in which  $\phi_1$  and  $\phi_2$  initially have disjoint support, i.e. drivers of both classes are well separated. The “convoys” of both species initially evolve according to the scalar model studied in [25, 81, 82], see Figures 3.2 (a–c). As soon as both classes enter in contact, unstable solution behaviour emerges, as can be seen in the oscillatory part of the solution visible in Figure 3.2 (d).

To ensure the parabolicity condition, we choose reaction times according to (3.32) by setting  $\tau_1 = 0.0008$  h and  $\tau_2 = 0.0011$  h in Examples 3.3–3.5. As in Example 3.2, we observe that with these reaction times each driver class is associated with stable behaviour when the respective other class is absent. In Figure 3.3 (a) we describe a stability region in the  $(\phi_1, \phi_2)$ -plane (phase space) corresponding to points at which the real parts of the eigenvalues of  $\mathbf{B}(\Phi)$  are positive. This is a subregion of  $\mathbb{R}_+^2$  bounded by curves  $C_1(\Phi) = 0$  and  $C_2(\Phi) = 0$ . Next, we choose the initial condition

$$\phi_i(x, 0) = \phi_i^0 + \delta\phi^0 \left[ \cosh^{-2} \left( \frac{320}{K} \left( x - \frac{5K}{16} \right) \right) - 0.25 \cosh^{-2} \left( \frac{40}{K} \left( x - \frac{11K}{32} \right) \right) \right] \quad (3.35)$$

for  $i = 1, 2$  (similar to the one proposed in [87]), where  $\delta\phi^0$  is the amplitude of perturbation; we here choose  $\delta\phi^0 = 0.08$ . We select the initial density  $\Phi^0$  in the different regions and compute the solution until a finite time. For the initial conditions  $\phi_1^0 = \phi_2^0 = 0.15$  (Example 3.3) or  $\phi_1^0 = \phi_2^0 = 0.4$  (Example 3.4) (Figures 3.4 (a–d)), which lie both in the instability region, we observe that amplitudes present in the initial datum are expanded but remain bounded in the instability region, while the frequencies are extended to maximum frequency. When  $\phi_1^0 = 0.25$  and  $\phi_2^0 = 0.25$  in the stability region (Example 3.5), simulations (Figures 3.4 (e, f)) show that amplitudes of the disturbance decrease with time, and that the corresponding frequency of oscillation does not increase. In both cases, initial perturbations generate waves traveling downstream and upstream.

### 3.4.3 Examples 3.6–3.8 (DG model, $N = 2$ , mildly unstable behaviour)

We continue using the DG model (3.9), consider a circular road of length  $K = 4$  mi and employ  $v_1^{\max} = 80$  mi/h,  $v_2^{\max} = 30$  mi/h,  $\tau_1 = 0.00095$  h,  $\tau_2 = 0.00075$  h and  $L_{\min} = 0.01$  mi. For this choice of parameters, we observe in Figure 3.3 (b) that the instability region of  $\mathbf{M} = \mathbf{M}(\Phi, \xi)$  is a subset of the stability region of  $\mathbf{B}$ , which indicates that bounded or unbounded instabilities could be generated even when the parabolicity conditions (2.14) are satisfied. As in the last example, we choose two initial conditions close to the instability region.

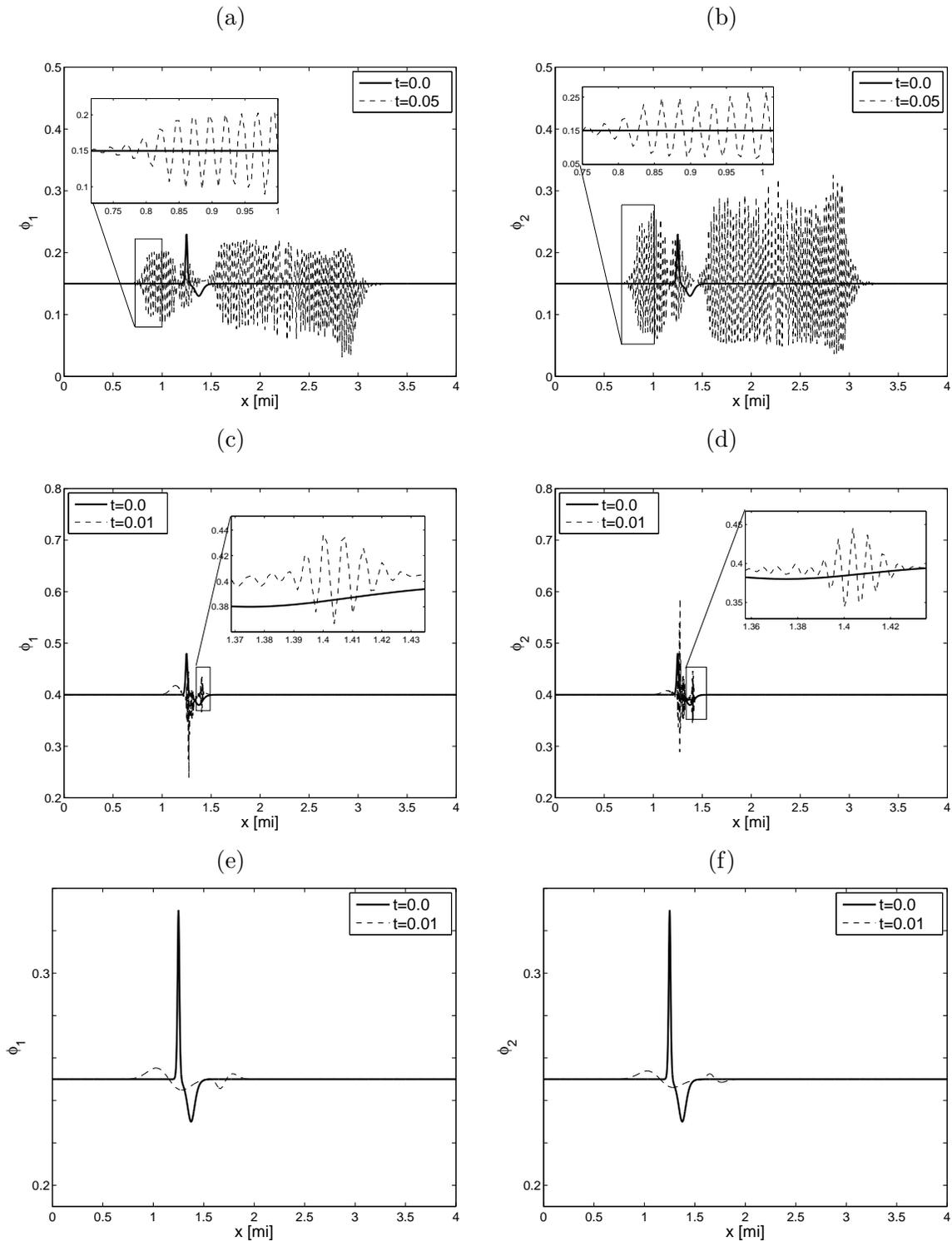


Figure 3.4: Examples 3.3–3.5 (DG model,  $N = 2$ ): Simulations for different initial conditions which lie in the stability or instability region (cf. Figure 3.3): (a, b)  $\phi_1^0 = \phi_2^0 = 0.15$  (Example 3.3), (c, d)  $\phi_1^0 = \phi_2^0 = 0.4$  (Example 3.4), (e, f)  $\phi_1^0 = \phi_2^0 = 0.25$  (Example 3.5).

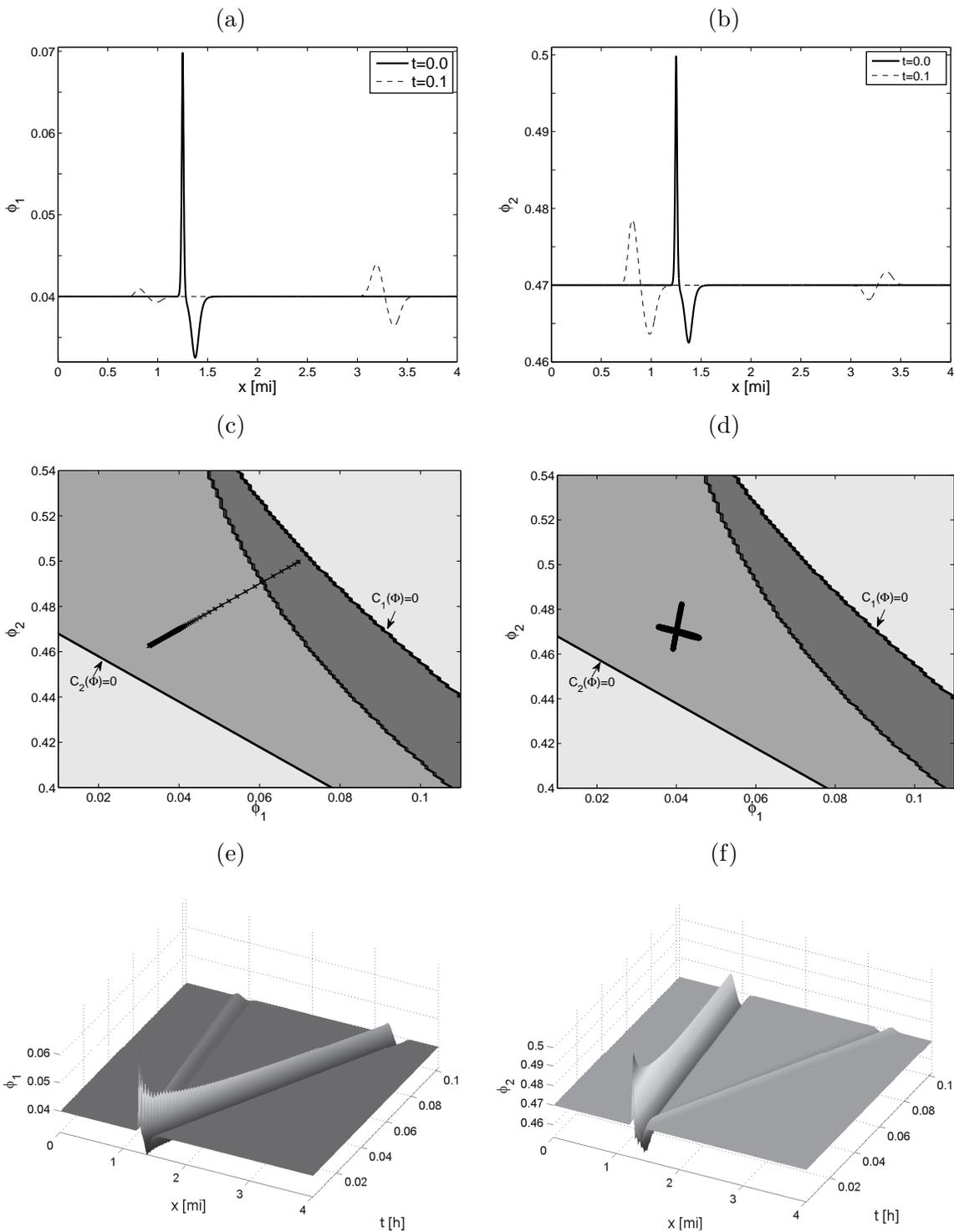


Figure 3.5: Example 3.6 (DG model,  $N = 2$ ). Unperturbed initial state in stability region of  $\mathbf{B}$  and initial perturbation in stability region of  $\mathbf{B}$  and partially in instability region of  $\mathbf{M}$ : (a, b) initial datum and numerical solution at  $t = 0.1$  h; (c, d) phase plane plots of (c) the initial datum and (d) the numerical solution for  $t = 0.1$  h; (e, f) numerical solution for  $0 \leq t \leq 0.1$  h.

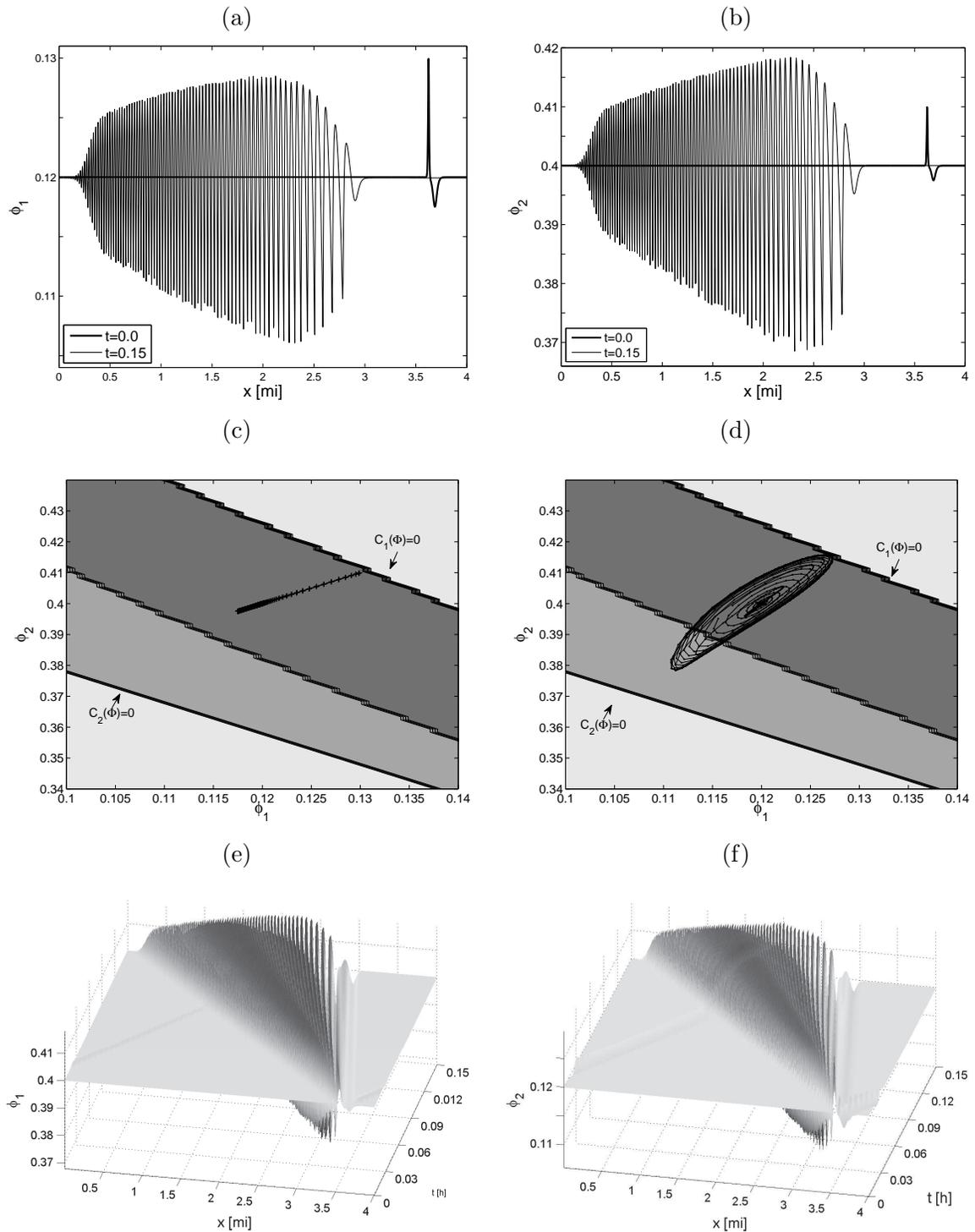


Figure 3.6: Example 3.7 (DG model,  $N = 2$ ). Unperturbed initial state and initial perturbation in stability region of  $\mathbf{B}$  and in instability region of  $\mathbf{M}$ : (a, b) initial datum and numerical solution at  $t = 0.1$  h; (c, d) phase plane plots of (c) the initial datum and (d) the numerical solution for  $t = 0.1$  h; (e, f) numerical solution for  $0 \leq t \leq 0.1$  h.

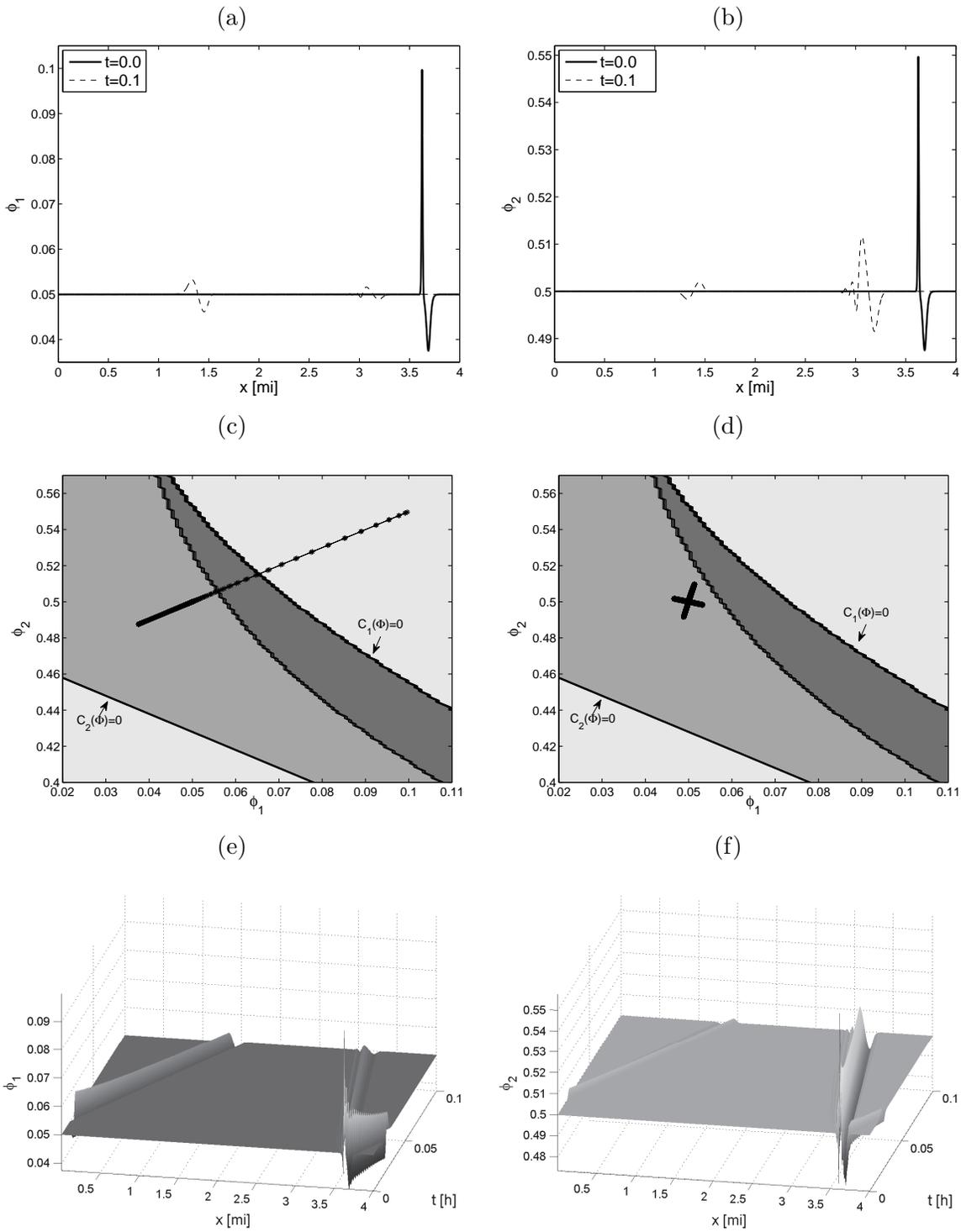


Figure 3.7: Example 3.8 (DG model,  $N = 2$ ). Unperturbed initial state in stability region of  $\mathbf{B}$  and large-amplitude initial perturbation partially in instability region of  $\mathbf{B}$ : (a, b) initial datum and numerical solution at  $t = 0.1$  h; (c, d) phase plane plots of (c) the initial datum and (d) numerical solution for  $t = 0.1$  h; (e, f) numerical solution for  $0 \leq t \leq 0.1$  h.

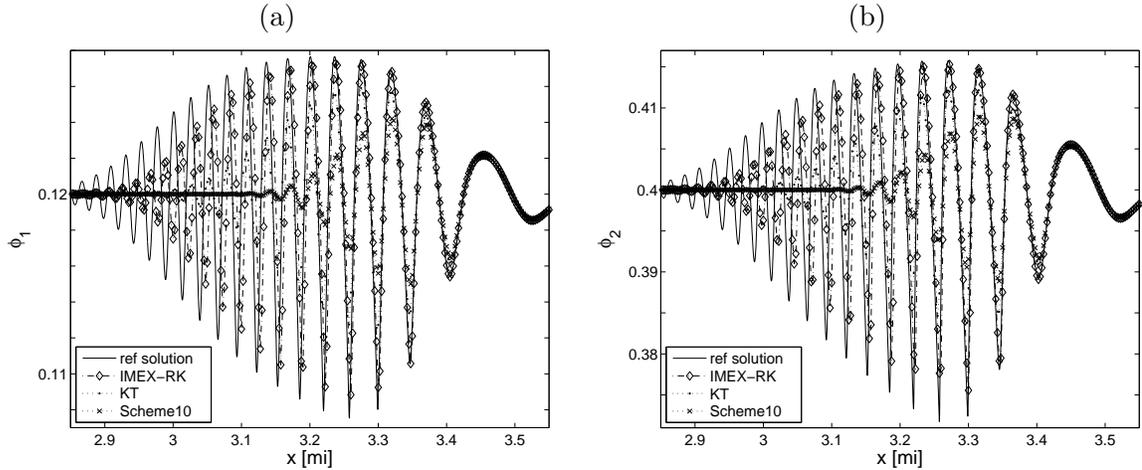


Figure 3.8: Example 3.7 (DG model,  $N = 2$ ): comparison of reference solution ( $M_{\text{ref}} = 12800$ ) with approximate solutions computed by schemes KT and IMEX-ARS(3,4,3) (3.31) and Scheme 10 of [24] with  $M = 1600$ .

	IMEX-ARS(3,4,3)		KT		Scheme 10	
$M$	error	CPU time [s]	error	CPU time [s]	error	CPU time [s]
400	1.7e-3	1.92	1.8e-3	2.29	2.3e-3	0.34
800	1.4e-3	6.60	1.7e-3	9.92	1.9e-3	1.23
1600	9.3e-4	25.74	1.2e-3	55.94	1.4e-3	5.02
3200	2.9e-4	105.91	8.9e-4	308.26	9.1e-4	26.41

Table 3.1: Example 3.7 (DG model,  $N = 2$ ): approximate total  $L^1$  errors and CPU times at time  $t = 0.03$  h for the KT scheme with  $C_{\text{eff1}} = 0.1$ , scheme IMEX-ARS(3,4,3) (3.31) with  $C_{\text{eff2}} = 0.6$  and Scheme 10 of [24] with  $C_{\text{eff1}} = 0.25$ .

We display numerical solutions for different initial conditions. We observe in Figure 3.5 that an initial perturbation is split into two waves, a wave traveling downstream which decreases rapidly in amplitude, and another wave traveling upstream which can cause traffic instability depending on the initial condition. For  $\phi_1^0 = 0.04$  and  $\phi_2^0 = 0.47$  (Example 3.6) and an initial perturbation with amplitude  $\delta\phi^0 = 0.03$  which does not lie in the instability region for  $\mathbf{M}$ , waves traveling upstream and downstream decrease in amplitude until a steady state is nearly reached. Numerical solutions are displayed in Figure 3.5.

For  $\phi_1^0 = 0.12$  and  $\phi_2^0 = 0.4$  (Example 3.7) and an initial perturbation with amplitude  $\delta\phi^0 = 0.01$ , which lie in the instability region for  $\mathbf{M}$ , waves traveling downstream decrease in amplitude, while waves traveling upstream grow in amplitude and until some frequency, which cause traffic instabilities. Numerical solutions at different times are displayed in Figure 3.6. Phase space diagrams are also shown in order to display how instabilities may be triggered. In Table 3.1 we calculate total approximate  $L^1$  errors and CPU times at time  $t = 0.03$  h for two different numerical schemes. The reference solution was calculated using the KT scheme with  $M = 12800$  subintervals.

$M$	IMEX-ARS(3,4,3)		KT		Scheme 10	
	error	CPU time [s]	error	CPU time [s]	error	CPU time [s]
400	3.1e-4	2.52	4.9e-4	4.05	6.7e-4	0.36
800	1.2e-4	10.60	2.4e-4	16.82	2.5e-4	1.31
1600	5.3e-5	39.43	8.9e-5	87.65	9.3e-5	6.33
3200	2.3e-5	145.04	2.8e-5	490.67	4.5e-5	30.24

Table 3.2: Example 3.8 (DG model,  $N = 2$ ): approximate total  $L^1$  errors and CPU times at time  $t = 0.03$  h for the KT scheme KT with  $C_{\text{eff1}} = 0.1$ , scheme IMEX-ARS(3,4,3) (3.31) with  $C_{\text{eff2}} = 0.6$  and Scheme 10 of [24] with  $C_{\text{eff1}} = 0.25$ .

This information indicates that the numerical solutions produced by all three schemes converge to the same solution as  $\Delta t, \Delta x \rightarrow 0$ . In particular, the oscillations visible in the numerical solution (see Figure 3.6 (a, b)) are not artifacts produced by the numerical scheme. That the oscillations are not a numerical artifact is further supported by Figure 3.8, where we compare the numerical solutions obtained for  $M = 1600$  with all three schemes with the reference solution, obtained by the KT scheme with  $M_{\text{ref}} = 12800$ .

In Example 3.8 we choose the constants  $\phi_1^0 = 0.05$  and  $\phi_2^0 = 0.5$ , which lie in the stability region, and add an initial perturbation with amplitude  $\delta\phi^0 = 0.05$  so that the initial function  $(\phi_1(x, 0), \phi_2(x, 0))^T$  defined by (3.35) assumes values that are in the instability region for  $B(\Phi)$ . In the numerical solution (see Figure 3.7) we observe that waves traveling upstream generate a wave that decreases in amplitude, and we also observe instabilities that remain controlled. In Table 3.2 we calculate total approximate  $L^1$  errors and CPU times at time  $t = 0.03$  h for three different numerical schemes. This table indicates that oscillations present in the numerical solution (cf. Figures 3.7 (a) and (b)) are not produced by the numerical scheme.

### 3.4.4 Examples 3.9–3.11 (GS model, $N = 4$ and $N = 2$ )

Now, we consider a DCMCLWR model with the Greenshields (GS) velocity function (3.10) and assume that  $\mathbf{B}(\Phi)$  is given by (3.11) with the perception threshold  $\phi_c = 0.05$ . In Example 3.9 we choose  $N = 4$  and  $v_1^{\text{max}} = 60$  mi/h,  $v_2^{\text{max}} = 55$  mi/h,  $v_3^{\text{max}} = 50$  mi/h,  $v_4^{\text{max}} = 45$  mi/h and  $L_{\text{min}} = 0.03$  mi. The reaction times are chosen such that the eigenvalues of  $\mathbf{B}(\Phi)$  have positive sign. In fact, we ensure that (2.14) holds by choosing  $\tau_1 = 0.0005$  h,  $\tau_2 = 0.0004$  h,  $\tau_3 = 0.0003$  h and  $\tau_4 = 0.0002$  h. Figure 3.10 shows a time evolution of the initial concentration platoon to a final (nearly) constant steady state reached at  $t = 50$  h.

To analyse chaotic behaviour, we consider in Example 3.10 a circular road of length 4 mi and interaction between  $N = 2$  classes with the respective free velocities  $v_1^{\text{max}} = 60$  mi/h and  $v_2^{\text{max}} = 30$  mi/h with an anticipation distance  $L_{\text{min}} = 0.01$  mi. Instabilities occur when we choose reaction times as  $\tau_1 = 0.0024$  h and  $\tau_2 = 0.0008$  h. In Figure 3.9 we display the stability region for the diffusion matrix  $\mathbf{B}(\Phi)$  and the instability region for the matrix  $\mathbf{M}$ . As for the Dick-Greenberg model, we choose two different initial conditions and show that traffic instabilities can occur. In Figure 3.11 we display a time evolution of an initial condition with  $\phi_1 = 0.2$  and  $\phi_2 = 0.23$  and a perturbation with amplitude  $\delta\phi^0 = 0.02$  in the instability region (Example 3.10).

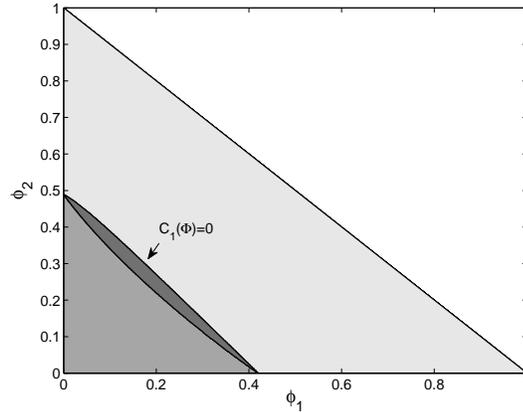


Figure 3.9: Examples 3.10 and 3.11 (GS model,  $N = 2$ ): stability region for diffusion matrix  $B$  and instability region for  $M$ .

We observe that the solution is a wave travelling upstream which grows in amplitude and frequency which are instabilities remain controlled, both in amplitude and frequency. We provide in Table 3.4 and in Figure 3.13 information similar to that of Table 3.1 and Figure 3.8 for Example 3.7, illustrating that also for this case, the oscillations observed are not a numerical artifice and that all three numerical schemes apparently approximate the same solution. In Figure 3.12 we display a time evolution of an initial condition with  $\phi_1 = \phi_2 = 0.18$  and a perturbation with amplitude  $\delta\phi^0 = 0.05$  in the stability region of  $B$  but with some values in the instability region of  $M$  (Example 3.11). We observe that the solution consists of two waves traveling downstream and decreasing in amplitude.

$i$	1	2	3	4	5
$L_i$ [mi]	0.006	0.012	0.03	0.008	0.028
$\tau_i$ [h]	0.00028	0.00052	0.00132	0.00036	0.00122

Table 3.3: Example 3.12 (DG model,  $N = 5$ , drivers having the same maximum speed): reaction times and anticipation distances.

### 3.4.5 Examples 3.12 and 3.13 (DG model, $N = 5$ , drivers having the same maximum speed)

In Example 3.12 we consider a circular road of 10 mi and choose  $N = 5$  classes of drivers with the same free velocities  $v^{\max} = 50$  mi/h. To satisfy the parabolicity condition (3.15), it is sufficient to choose reaction times  $\tau_i$  and anticipation distances  $L_i$  such that  $\tau_i \leq L_i(|\phi V'(\phi)|v^{\max})^{-1}$  for  $i = 1, \dots, N$ . We employ the DG velocity function (3.9) and choose the reaction times and anticipation distances given in Table 3.3. Figure 3.14 shows a time evolution of the initial concentration platoon  $\Phi_0(x, 0) = p(x)(0.2, 0.2, 0.2, 0.2, 0.2)^T$ , where  $p(x)$  is given in (3.33), for which stable behavior is observed. When condition (3.15) is not satisfied, unstable behavior with non-controlled oscillations appears. As Example 3.13 we consider the same initial platoon (3.33) but we choose  $\tau_2 = 0.00104$  h. Simulations are displayed in Figure 3.15(a).

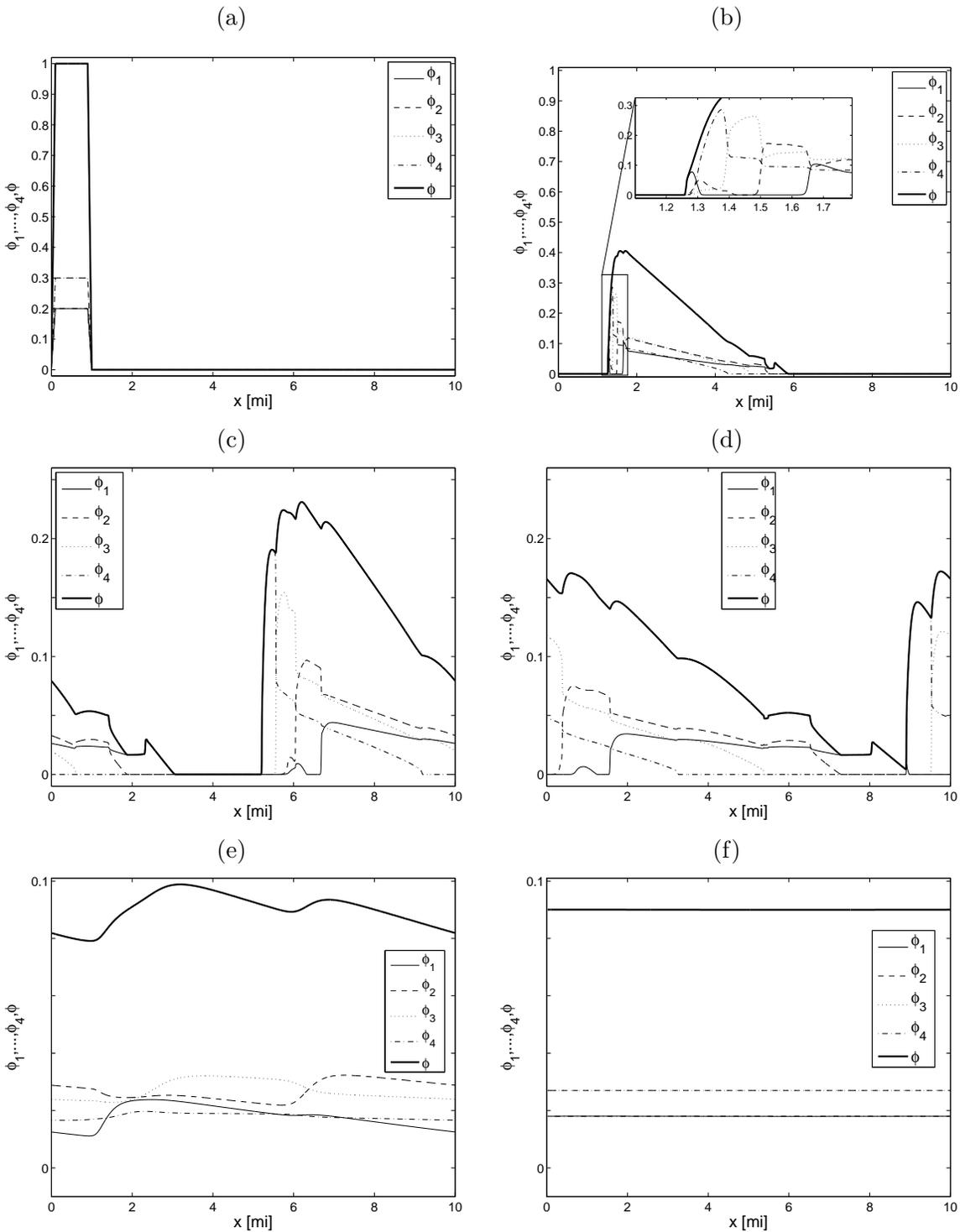


Figure 3.10: Example 3.9 (GS model,  $N = 4$ ): (a) initial datum and solution at simulated times (b)  $t = 0.08$  h, (c)  $t = 0.2$  h, (d)  $t = 0.3$  h, (e)  $t = 9$  h and (f)  $t = T = 50.0$  h.

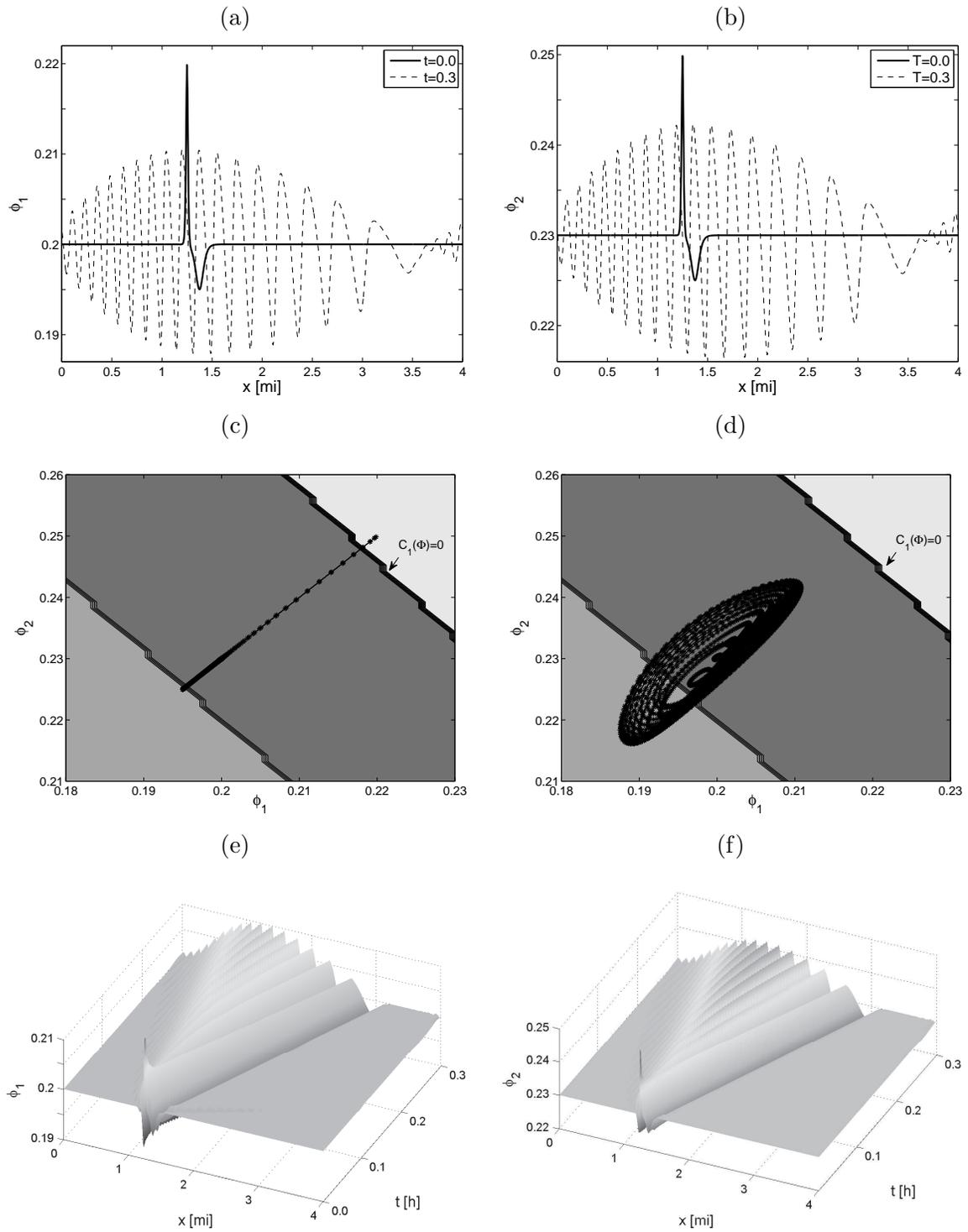


Figure 3.11: Example 3.10 (GS model,  $N = 2$ ). Unperturbed initial state and small-amplitude initial perturbation in stability region of  $\mathbf{B}$  and in instability region of  $\mathbf{M}$ : (a, b) initial datum and numerical solution at  $t = 0.1$  h; (c, d) phase plane plots of (c) the initial datum and (d) the numerical solution for  $t = 0.1$  h; (e, f) numerical solution for  $0 \leq t \leq 0.1$  h.

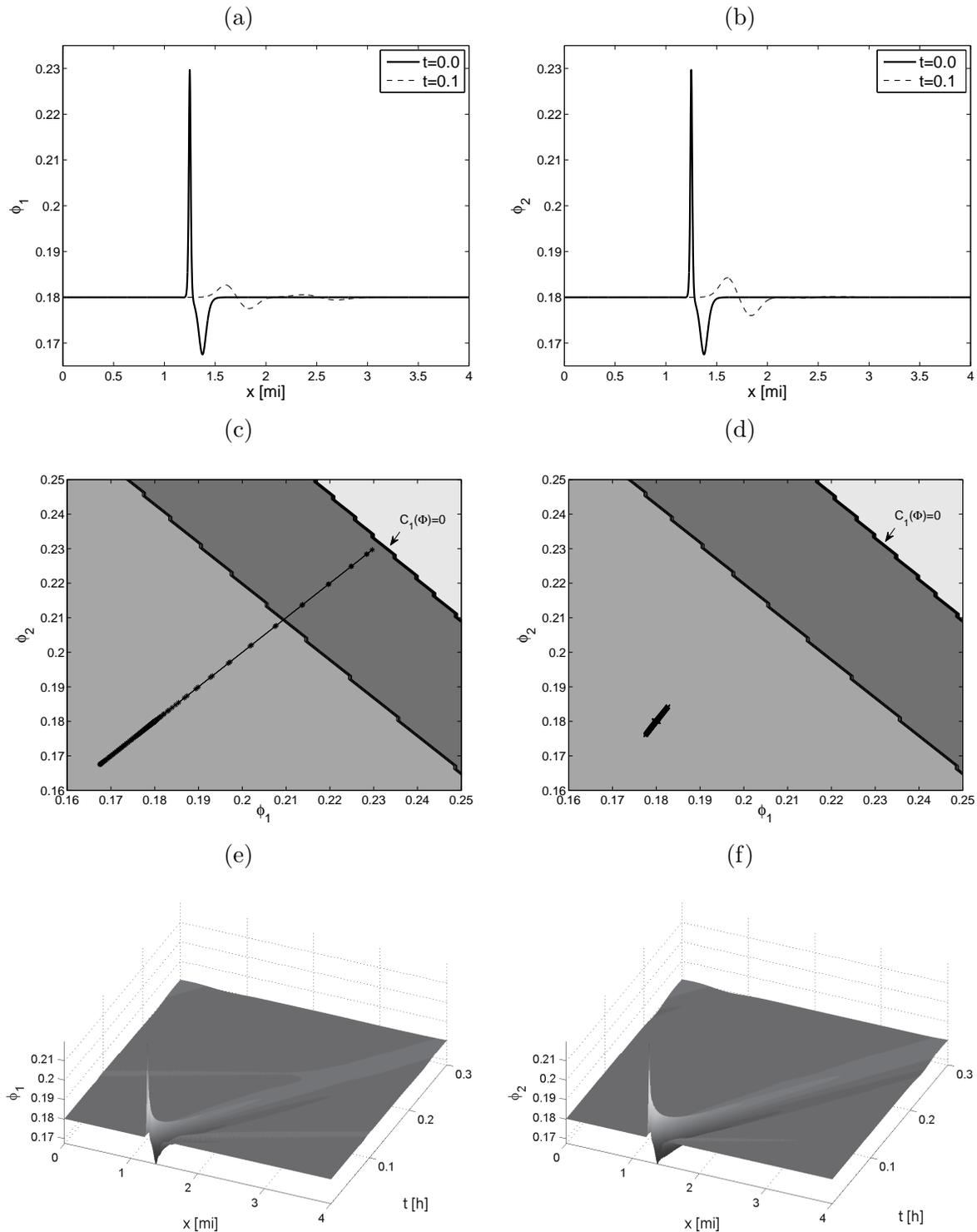


Figure 3.12: Example 3.11 (GS model,  $N = 2$ ): Unperturbed initial state in stability region of  $\mathbf{B}$  and initial perturbation in stability region of  $\mathbf{B}$  and in instability region of  $\mathbf{M}$ : (a, b) initial datum and numerical solution at  $t = 0.1$  h; (c, d) phase plane plots of (c) the initial datum and (d) the numerical solution for  $t = 0.1$  h; (e, f) numerical solution for  $0 \leq t \leq 0.1$  h.

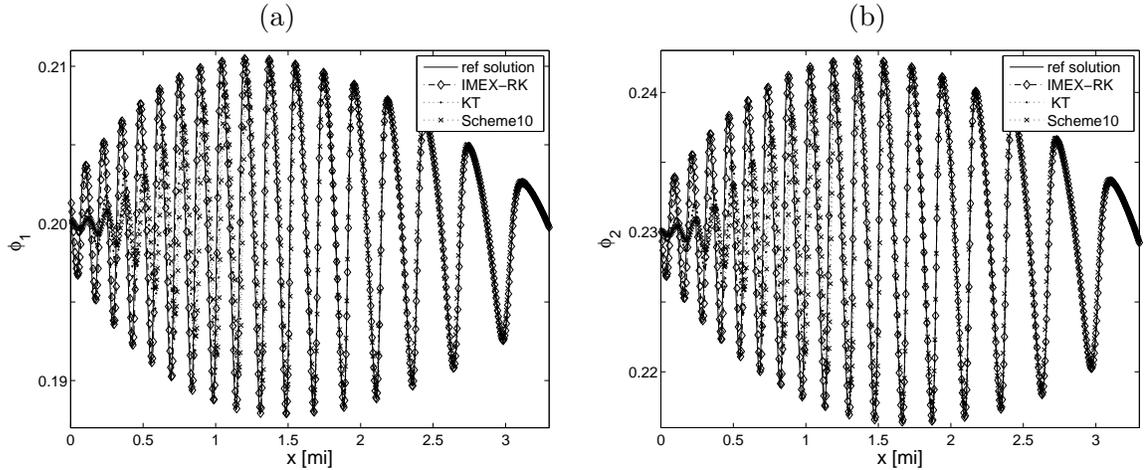


Figure 3.13: Example 3.10 (GS model,  $N = 2$ ): comparison of reference solution ( $M_{\text{ref}} = 12800$ ) with approximate solutions computed by schemes KT and IMEX-ARS(3,4,3) (3.31) and Scheme 10 of [24] with  $M = 800$ .

$M$	IMEX-ARS(3,4,3)		KT		Scheme10	
	error	CPU time [s]	error	CPU time [s]	error	CPU time [s]
400	2.1e-3	2.12	2.5e-3	4.05	8.1e-3	2.54
800	1.8e-3	9.12	2.0e-3	67.80	4.4e-3	12.08
1600	1.4e-3	47.38	1.9e-3	274.75	2.1e-3	50.45
3200	7.2e-5	205.44	1.0e-3	1059.48	1.0e-3	253.94

Table 3.4: Example 3.10 (GS model,  $N = 2$ ): approximate total  $L^1$  errors and CPU times at time  $t = 0.03$  h for the KT scheme with  $C_{\text{cf}1} = 0.05$ , scheme IMEX-ARS(3,4,3) (3.31) with  $C_{\text{cf}2} = 0.6$  and Scheme 10 of [24] with  $C_{\text{cf}2} = 0.1$ .

### 3.4.6 Example 3.14 (Daganzo's test, $N = 4$ )

We finish this numerical section with a test proposed by Daganzo in [43]. In that paper, the author argues that second-order (scalar) partial differential equations modelling traffic flow will give negative fluxes, i.e., cars moving backwards, at traffic jams. To perform a multiclass test that corresponds to this situation, we specify zero-flux boundary conditions and set up the following initial condition:

$$\Phi(x, 0) = \begin{cases} \Phi_L & \text{for } x < 0, \\ \Phi_R & \text{for } x > 0, \end{cases} \quad (3.36)$$

where  $\Phi_L = \mathbf{0}$  and  $\Phi_R$  is such that  $\phi_R = 1$ . This initial condition prescribes an initially stopped traffic after some point ( $x = 0$ ) and no cars behind it. This density distribution should be a stationary solution for the model, but, as Daganzo predicts in [43], the numerical scheme applied

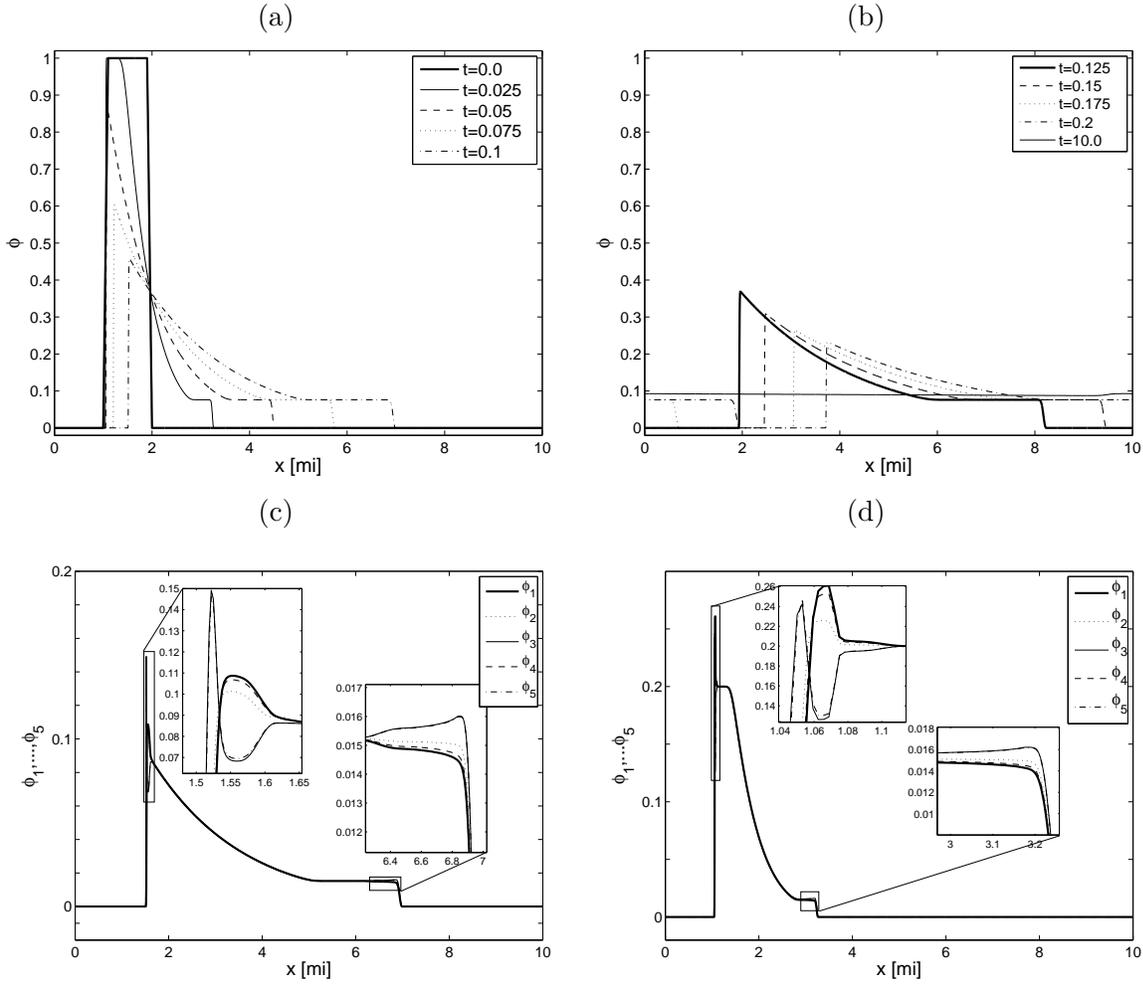


Figure 3.14: Example 3.12 (DG model,  $N = 5$ , drivers having the same maximum speed): (a, b) total concentration at different times, (c, d) individual concentrations and enlarged views at times (c)  $t = 0.025$  h and (d)  $t = 0.1$  h.

to our model and this initial condition produce a smearing of the profile around  $x = 0$ , i.e., cars move backwards, see the dashed numerical solution in Figure 3.15(b).

Nevertheless, the analysis in [43] is performed for linear diffusion, but our model yields strongly degenerate, nonlinear diffusion terms. Therefore, we can modify the numerical scheme for preventing this smearing as the following reasoning indicates: A stationary solution for the KT scheme satisfies

$$\hat{\mathbf{f}}_{j+1/2} = (\mathbf{B}(\Phi)\Delta\Phi)_{j+1/2},$$

where  $\hat{\mathbf{f}}_{j+1/2}$  and  $(\mathbf{B}(\Phi)\Delta\Phi)_{j+1/2}$  are the numerical convective and diffusive fluxes, respectively, of the KT scheme defined at the boundary between cells centered at  $x_j$  and  $x_{j+1}$ , where  $x_j = j\Delta x$  and  $j \in \mathbb{Z}$ . The function (3.36) would be a stationary solution if

$$\hat{\mathbf{f}}_{j+1/2} = \mathbf{0} \tag{3.37}$$

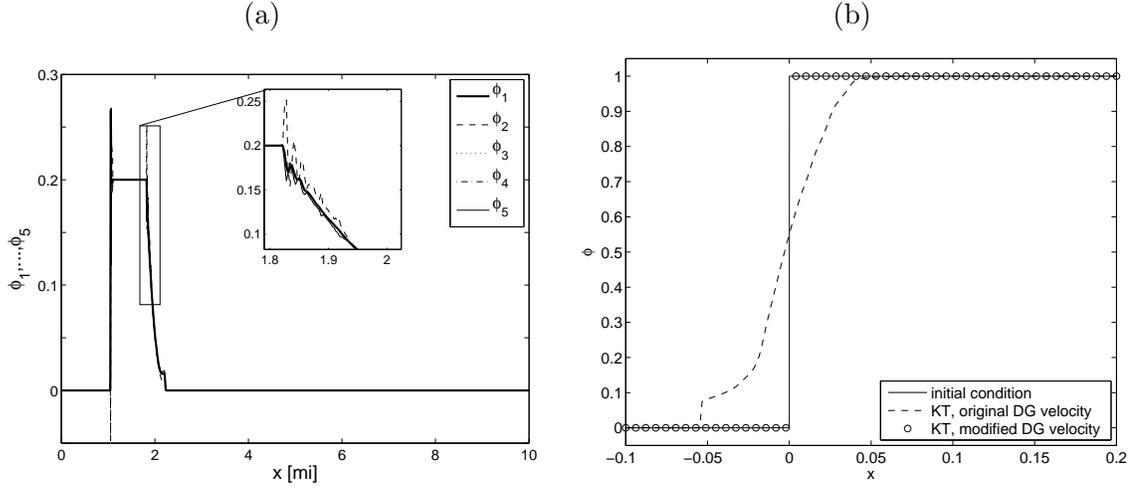


Figure 3.15: (a) Example 3.13 (DG model,  $N = 5$ , drivers having the same maximum speed): instabilities may occur when the parabolicity condition (3.15) is not satisfied. (b) Example 3.14 (Daganzo's test): solutions with the original DG velocity and modified velocity.

and

$$(\mathbf{B}(\Phi)\Delta\Phi)_{j+1/2} = \mathbf{0}. \quad (3.38)$$

Since the total concentration of (3.36) is zero for  $x < 0$  and one for  $x > 0$ , the physical convective fluxes are  $\mathbf{0}$  for  $\Phi$  in (3.36). If no artificial viscosity is added to numerical fluxes, then (3.37) holds for (3.36). One therefore deduces that condition (3.38) is sufficient for (3.36) to be a stationary solution for the numerical scheme. Since the discretization of the latter is

$$\frac{\Delta x}{2}(\mathbf{B}(\Phi_j) + \mathbf{B}(\Phi_{j+1}))(\Phi_{j+1} - \Phi_j),$$

it is then sufficient that  $\mathbf{B}(\Phi_j) = \mathbf{0}$ . Since the latter is proportional to  $V'(\phi_j)$  and either  $\phi_j = 0$  or  $\phi_j = 1$ , this would be ensured if  $V'(0) = V'(1) = 0$ . The first requirement is satisfied for the DG velocity (3.9) and the second one could be obtained by modifying the velocity function. We have performed the modification to the KT scheme in this manner and obtained that (3.36) is indeed a stationary solution for the numerical scheme (see Figure 3.15).

### 3.5 Conclusions of Chapter 3

We have analyzed the stability of a diffusively corrected multiclass Lighthill-Whitham-Richards (DCMCLWR) traffic model that takes into account anticipation lengths and reaction times. The basic result is that to achieve stability, defined in terms of the linearized version (3.12) of (3.6), it is not sufficient to ensure that the diffusion matrix  $\mathbf{B}$  has eigenvalues with positive real parts; rather, one also has to consider a contribution from the convective part defined by the Jacobian  $\mathbf{J}$ , multiplied by  $i/\xi$ . Thus, it is not possible in general to identify stable or unstable solution behaviour with a particular type of (3.6), unless we consider the special cases  $\mathbf{J} = \mathbf{0}$ ,  $\mathbf{J}$  being a rank-one perturbation of a multiple of  $\mathbf{I}$  (as for the case of equal free velocities discussed in Section 3.2.3) or  $\mathbf{B} = \mathbf{0}$  (as for the standard MCLWR model). This contrasts, for example, with the stability analysis of a model of polydisperse sedimentation [26], whose governing equations can be written as (3.4), and for which a stability analysis similar to the one conducted in Section 3.2.2 shows that a criterion for stable segregation (formation of horizontal concentration interfaces that move vertically), introduced in [7] for  $N = 2$  and supported by experimental results, is equivalent to hyperbolicity of (3.4). The experiments show that stable behavior is obtained when the eigenvalues of  $\mathbf{M}$  have positive real parts and that instabilities may be triggered otherwise, although the nonlinear character of the equations stabilizes some initial traffic configurations that would explode under the linearized equations. The nonlinearities also help to control the amplitude (and in some cases the frequency) of instabilities in the simulations. While we associate oscillations in the numerical solution with unstable behaviour in general, we distinguish between situations where there is a blow-up of frequency (such as in Examples 3.2, 3.3, 3.4 and 3.13), which means that violations of the stability condition lead to strongly oscillating solutions (akin to those studied in [13]), and situations of mildly unstable behaviour (such as the ones observed in Examples 3.5 to 3.8 and 3.10) with finite frequencies of oscillation, and where numerical solutions can be interpreted as the formation of stop-and-go waves (although the latter phenomenon is usually associated with much larger amplitudes, cf., e.g., [86, 87]).

The present model, the stability analysis and its numerical simulations allow us to draw some conclusions of stable and unstable traffic flow caused by heterogeneous drivers' behaviour. To elucidate this issue, let us first point out that the condition (3.34), which is precisely the condition for the scalar equation (3.3) to be (degenerate) parabolic, is very similar to the condition (3.8) derived in [87]. While it is plausible that traffic flow is stable, and for instance free of marked stop-and-go waves, if reaction times of drivers are sufficiently small, our analysis leads to a further conclusion for  $N \geq 2$ . Namely, for that case it turns out that to ensure stable traffic flow it is not sufficient so require that (3.34) be satisfied with  $L$ ,  $\tau$  and  $v^{\max}$  replaced by  $L_i$ ,  $\tau_i$  and  $v_i^{\max}$  for  $i = 1, \dots, N$ . This is vividly illustrated in Example 3.2: two populations of drivers may produce stable traffic flow when separated spatially, however, when they start to "mix", then instabilities occur. This behaviour is essentially produced by the fact that the larger reaction time  $\tau_2$  of species 2 (in Example 3.2) is not sufficiently small in presence of the significantly faster drivers of class 1. In fact, in view of the assumption (2.8) the criterion (3.32) states that the eigenvalues of  $\mathbf{B}$  have non-negative real parts (a condition necessary, but in general not sufficient, to ensure stability of

traffic flow) if the reaction time  $\tau_i$  of drivers of a given class  $i$  is adapted to the velocities of drivers of the faster classes 1 to  $i - 1$ . In particular, (3.32) means that  $\tau_1 \geq \tau_2 \geq \dots \geq \tau_N$ .

Finally, we remark that in the numerical examples the reaction times have been chosen relatively high (up to the value of  $\tau_2 = 0.0025 \text{ h} = 9 \text{ s}$  in Example 3.2). We are well aware that these values are not representative for real-world traffic flow, where reaction times of about  $\tau = 1 \text{ s}$  are realistic (cf., e.g., [109, 110]), but for non-attentive drivers the reaction time may be substantially larger. Moreover, Nelson [81] employs  $\tau = 2 \text{ s}$  for his order-of-magnitude study, and Ngoduy and Tampere [87] study the effect of reaction time on flow stability for several values of  $\tau$  up to  $\tau = 2 \text{ s}$ . However, we have not made the effort in this work to maximally adjust parameters to real systems, but rather chose the parameters in such a way that the effects of instability become well visible. That said, it should be pointed out that the choices of reaction times alone do not determine stable or unstable behaviour; the key information is expressed by the spectrum of the matrix  $\mathbf{M}(\Phi, \xi)$  defined in (3.7), and this spectrum decisively depends not only on the parameter vectors  $\mathbf{L}$ ,  $\boldsymbol{\tau}$  and  $\mathbf{v}^{\max}$ , but also on the choice of the function  $V(\phi)$  via its derivative  $V'(\phi)$ , which is present in the parabolicity or stability bounds (3.19) and (3.34) and in Theorem 3.2. For the latter function we have used the simple given functional forms (3.9) and (3.10); other choices of  $V(\phi)$ , especially those with large values of  $V'(\phi)$ , possibly give rise to appreciable instability phenomena at more realistic reaction times. In light of the comments made in Section 3.2.2, evaluation of the instability criterion (3.15) unfortunately requires analyzing the spectrum of  $\mathbf{M}(\Phi, \xi)$ , unless we are in the special case of equal preferential velocities (Section 3.2.3), where Lemma 3.1 provides an easily applicable criterion to decide whether traffic instabilities are excluded (which occurs if (3.19) cannot be satisfied).



## Chapter 4

# Lagrangian-Remap schemes for the MCLWR traffic model

### 4.1 Introduction

#### 4.1.1 Scope

In this chapter we consider exclusively the multiclass Lighthill-Whitham-Richards (MCLWR) traffic model described in the previous chapters, where the components of the unknowns vector  $\boldsymbol{\rho} = (\rho_1, \dots, \rho_N)^\top$  corresponds to  $\rho_i = \rho_i(x, t)$  the local number of cars per mile of class  $i$  which the MCLWR model is expressed as the following system of strongly coupled nonlinear first-order conservation laws

$$\partial_t \boldsymbol{\rho} + \partial_x \mathbf{f}(\boldsymbol{\rho}) = \mathbf{0}, \quad x \in \mathbb{R}, \quad t > 0, \quad (4.1)$$

where  $x$  is horizontal distance and either  $I = \mathbb{R}$  for an unbounded highway or  $I = (0, L)$  for a traffic circle of length  $L > 0$ ,  $t$  is time,  $\mathbf{f}(\boldsymbol{\rho}) = (f_1(\boldsymbol{\rho}), \dots, f_N(\boldsymbol{\rho}))^\top$ , where

$$f_i(\boldsymbol{\rho}) = \rho_i v_i(\rho), \quad i = 1, \dots, N, \quad (4.2)$$

and  $v_i(\rho)$  is the velocity of cars of class  $i$ , which is assumed to be an explicit function of the local total density  $\rho := \rho_1 + \dots + \rho_N$ . We assume that for all  $i$ ,  $0 \leq \rho_i \leq \rho \leq \rho_{\max}$ , where  $\rho_{\max}$  is a maximum density corresponding to a bumper-to-bumper situation. It is usually assumed that

$$v_i(\rho) = v_i^{\max} V(\rho), \quad i = 1, \dots, N, \quad (4.3)$$

where  $v_i^{\max}$  is the preferential velocity of class  $i$  corresponding to a free highway, and  $V(\rho)$  is a hindrance factor that takes into account drivers' attitude to reduce speed in presence of other cars. The function  $V$  is usually assumed that

$$V(0) = 1, \quad V'(\rho) \leq 0 \quad \text{for } 0 \leq \rho \leq \rho_{\max}, \quad V(\rho_{\max}) = 0. \quad (4.4)$$

The numerical solution of (4.1), (4.2) is a challenge since the eigenvalues and eigenvectors of the Jacobian matrix  $\mathcal{J}_{\mathbf{f}}(\boldsymbol{\rho}) = (\partial f_i(\boldsymbol{\rho}) / \partial \rho_j)_{1 \leq i, j \leq N}$  are not available in closed algebraic form.

Alternatively, one can construct easy-to-implement numerical schemes for (4.1), (4.2) by exploiting the concentration-times-velocity form (4.2) of the fluxes, and utilizing that by (4.3) and (4.4), the functions  $v_i$  are non-negative, bounded, and strictly decreasing. This idea was first used in [24] to design a family of relatively simple difference schemes for (4.1), (4.2).

It is the purpose of this chapter to introduce a new class of schemes for (4.1), (4.2) that do not rely on spectral (characteristic) information and are as easy to implement as the schemes introduced in [24], but perform better in terms of resolution, accuracy and efficiency. To explain the main idea, let us consider the continuity equation for the case  $N = 1$  of a single driver class

$$\partial_t \rho + \partial_x(\rho v(\rho)) = 0, \quad x \in I, \quad t > 0, \quad (4.5)$$

corresponding to the original LWR model [75, 92], where

$$v(\rho) = v^{\max} V(\rho), \quad (4.6)$$

and for ease of the argument, in the scalar case we assume that time is scaled such that  $v^{\max}$  equals unity. We formally rewrite (4.5) as

$$\partial_t \rho + \rho \partial_x(v(\rho)) + v(\rho) \partial_x \rho = 0, \quad x \in I, \quad t > 0. \quad (4.7)$$

The new class of schemes for (4.5) is based on splitting (4.7) into two different equations, which are solved successively for each time iteration. To advance the solution from time  $t$  to  $t + \Delta t$ , we first apply a Lagrangian method [56] to solve

$$\partial_t \rho + \rho \partial_x v(\rho) = 0, \quad (4.8)$$

and use this solution, evolved over the time interval of length  $\Delta t$ , as the initial condition for solving in a second step the transport equation

$$\partial_t \rho + v(\rho) \partial_x \rho = 0, \quad (4.9)$$

whose solution, again evolved over a time interval of length  $\Delta t$ , provides the sought approximate solution valid for  $t + \Delta t$ . These steps will be identified as ‘‘Lagrangian’’ and ‘‘remap’’ steps, respectively, so the class of schemes introduced herein is addressed as ‘‘Lagrangian-remap’’ (LR) schemes. (The names ‘‘Lagrangian’’ and ‘‘remap’’ will be given more insight in Sections 4.3 and 4.4 below.)

The idea behind the introduction of LR schemes is to solve (4.9) using anti-diffusive techniques that have been developed recently for transport equations and thereby to increase the overall efficiency of the proposed splitting strategy, while keeping its simplicity. More precisely, the remap step can be handled in two different ways. One alternative is to employ an anti-diffusive but stable numerical scheme [46, 15, 14] for the transport equation (4.9) (remap step), where care is taken to design the scheme for the remap step in such a way that the resulting scheme (first step followed by second step) is conservative. This subclass of LR schemes will be addressed as ‘‘Lagrangian-anti-diffusive remap’’ (L-AR) schemes. The L-AR schemes are discussed in several variants defined by different choices of a particular numerical flux. Alternatively, the remap step

can be handled by random sampling in a Glimm-like approach [55]. The resulting scheme, denoted here as “Lagrangian-random-sampling” (L-RS) scheme, is only statistically conservative, but turns out to be less diffusive than, for example, a (deterministic) integral remap step. Note that the loss of strictly conservativity property does not prevent the convergence to a weak solution in this context.

Both L-AR and L-RS subclasses of LR schemes can readily be extended to the multiple-species case ( $N > 1$ ). For that case, we propose to equip the L-RS scheme with random sampling among the fan of states of the simple Harten, Lax and van Leer (HLL) approximate Riemann solver [61, 111].

Our proposal of the class of LR schemes for the MCLWR model is supported by a partial analysis of the L-AR schemes for  $N = 1$ , with the conclusion that under suitable CFL conditions, the L-AR schemes have the total variation diminishing (TVD) property and therefore converge to a weak solution and by a number of numerical experiments that show that the proposed schemes are competitive with respect to recent schemes introduced in [24], see Section 4.2.3.

#### 4.1.2 Related work

The MCLWR model has been analyzed by several groups of authors, cf. e.g. [9, 50, 78, 118]. In particular, its hyperbolicity has been established [50, 118] and the admissible waves of the Riemann problem have been investigated [118]. Moreover, the model (4.1), (4.2) admits a separable, strictly convex entropy since its Jacobian matrix  $\mathcal{J}_f(\boldsymbol{\rho})$  is diagonally symmetrizable [8, 9]. Component-wise or characteristic high-resolution numerical schemes for (4.1), (4.2) involving weighted essentially non-oscillatory (WENO) flux reconstructions are advanced in [23, 50, 117, 119]. On the other hand, as mentioned above, particularly simple first- and second-order difference schemes for the same problem that rely on the structure of the fluxes  $f_i$  (4.2) along with the definite sign of the velocities  $v_i$  are introduced in [24]. Variants of the original MCLWR model (in the sense of [8, 113]) have been proposed, and in part analyzed, for highways with varying road surface conditions [27, 120, 121], traffic flow on networks [63, 84], stochastic fundamental diagrams (equivalent to the velocity functions  $v_i$ ) [85], and diffusive corrections modeling anticipation lengths and reaction times [30, 32].

Anti-diffusive numerical schemes used in this paper have been advanced in the pioneering work by Després and Lagoutière [46] for the linear transport equation with application to gas dynamics, and then extended to monotone scalar conservation laws by Bouchut [15] and applied to Hamilton-Jacobi-Bellman equations by Bokanowski and Zidani in [14]. We also refer to [47, 67, 69, 72] for further extensions. Furthermore, variants of the Glimm-like and mixed approach of the L-R and especially L-RS schemes have turned out successful in a number of contexts, ranging from the computation of classical and nonclassical shock waves [34, 35, 36], contacts discontinuities in two-phase flow and traffic flow models [5, 36, 62], phase transitions in traffic flow models set on a non-convex state space [37], to Vlasov equations [104].

## 4.2 Preliminaries

### 4.2.1 Weak solutions and entropy admissibility

We briefly recall the concepts of a weak solution and of entropy for (4.1) under the specific assumptions (4.2) and (4.3), and considering the initial condition

$$\boldsymbol{\rho}(x, 0) = \boldsymbol{\rho}_0(x), \quad x \in I. \quad (4.10)$$

We closely follow the preliminary remarks of [27]. First of all, it is well known that even if  $\boldsymbol{\rho}_0$  is smooth, solutions of (4.1), (4.10) develop discontinuities, and so we seek a weak solution, which is a bounded measurable function  $\boldsymbol{\rho} = \boldsymbol{\rho}(x, t)$  satisfying

$$\int_{\mathbb{R}^+} \int_I (\phi_t \boldsymbol{\rho} + \phi_x \mathbf{f}(\boldsymbol{\rho})) \, dx \, dt + \int_I \phi(x, 0) \boldsymbol{\rho}_0(x) \, dx = \mathbf{0} \quad (4.11)$$

for any smooth test function  $\phi = \phi(x, t)$  with compact support contained in  $I \times \mathbb{R}^+$ . If a weak solution  $\boldsymbol{\rho}$  has a discontinuity along a smooth curve  $x = x(t)$  and  $\boldsymbol{\rho}$  is continuous on either side of  $x(t)$  with limits  $\boldsymbol{\rho}_-$  and  $\boldsymbol{\rho}_+$  to the left and right of the jump, respectively, then the weak formulation (4.11) implies the following Rankine-Hugoniot (RH) jump condition, where  $s = dx/dt$  is the shock speed:

$$\mathbf{f}(\boldsymbol{\rho}_+) - \mathbf{f}(\boldsymbol{\rho}_-) = s(\boldsymbol{\rho}_+ - \boldsymbol{\rho}_-). \quad (4.12)$$

As is well known, weak solutions of (4.1) are not unique, so an additional admissibility criterion (usually motivated by the “physics” of the problem) needs to be imposed. Suppose that (4.1) admits a strictly convex entropy, meaning that there exists a strictly convex function  $\mathcal{E} = \mathcal{E}(\boldsymbol{\rho})$  and an entropy flux  $\mathcal{F} = \mathcal{F}(\boldsymbol{\rho})$  such that  $\nabla \mathcal{F}(\boldsymbol{\rho}) = \nabla \mathcal{E}(\boldsymbol{\rho}) \mathcal{J}_{\mathbf{f}}(\boldsymbol{\rho})$ . Then a weak solution  $\boldsymbol{\rho}$  of (4.1) is said to be *entropy-admissible* [21] if for every smooth nonnegative test function  $\varphi$  with compact support in  $I \times (0, \infty)$ , the inequality

$$\int_{\mathbb{R}^+} \int_I (\varphi_t \mathcal{E}(\boldsymbol{\rho}) + \varphi_x \mathcal{F}(\boldsymbol{\rho})) \, dx \, dt \geq 0 \quad (4.13)$$

is valid. This inequality follows from a parabolic regularization of (4.1) if one lets the regularization parameter tend to zero, assuming that the corresponding solutions converge boundedly a.e. to a limit  $\boldsymbol{\rho}$  (see [27, 102] for details). Note that (4.13) can also be expressed as  $\partial_t \mathcal{E}(\boldsymbol{\rho}) + \partial_x \mathcal{F}(\boldsymbol{\rho}) \leq 0$  in the sense of distributions. Moreover, (4.13) implies that all discontinuities satisfy the following entropy jump condition (in addition to (4.12)):

$$\mathcal{F}(\boldsymbol{\rho}_+) - \mathcal{F}(\boldsymbol{\rho}_-) \leq s(\mathcal{E}(\boldsymbol{\rho}_+) - \mathcal{E}(\boldsymbol{\rho}_-)). \quad (4.14)$$

For general  $N$  and systems of the type (4.1), the existence of an entropy pair  $(\mathcal{E}(\boldsymbol{\rho}), \mathcal{F}(\boldsymbol{\rho}))$  that satisfies  $\nabla \mathcal{F}(\boldsymbol{\rho}) = \nabla \mathcal{E}(\boldsymbol{\rho}) \mathcal{J}_{\mathbf{f}}(\boldsymbol{\rho})$  is an exceptional property. However, for the MCLWR model with  $\mathbf{f}(\boldsymbol{\rho})$  defined by (4.2), (4.3), Benzoni-Gavage and Colombo [8] showed that a convex entropy function  $\mathcal{E}(\boldsymbol{\rho})$  and corresponding entropy flux  $\mathcal{F}(\boldsymbol{\rho})$  are given by

$$\mathcal{E}(\boldsymbol{\rho}) = \sum_{i=1}^N \frac{\rho_i (\ln \rho_i - 1)}{v_i^{\max}}, \quad \mathcal{F}(\boldsymbol{\rho}) = V(\boldsymbol{\rho}) \sum_{i=1}^N \rho_i \ln \rho_i - \mathcal{V}(\boldsymbol{\rho}), \quad (4.15)$$

where  $\mathcal{V}$  is any primitive of  $V$ , i.e.,  $\mathcal{V}'(\rho) = V(\rho)$ . Finally, for later use we mention that in [8] it is also shown that for the special choice

$$V(\rho) = 1 - \frac{\rho}{\rho_{\max}}, \quad (4.16)$$

the entropy jump condition (4.14) is equivalent to

$$\rho_- \leq \rho_+. \quad (4.17)$$

In the present chapter, we will not attempt to *prove* that any of the newly introduced schemes converges to an entropy solution, i.e., a weak solution that satisfies (4.13). However, further support of the new schemes is provided by a heuristic argument based on evaluating a discrete analogue of  $\mathcal{E}(\boldsymbol{\rho})$  for given numerical solutions, and for some examples involving (4.16), we will verify whether the numerical solution approximates discontinuities that are consistent with (4.17).

### 4.2.2 Interlacing property of the MCLWR model

We assume that for  $N > 1$ , the velocities  $v_i^{\max}$  are ordered as

$$0 < v_1^{\max} \leq v_2^{\max} \leq \dots \leq v_N^{\max}. \quad (4.18)$$

Then a version of Theorem 2.2 in Section 2.2.4 indicate that the eigenvalues  $\lambda_i = \lambda_i(\boldsymbol{\rho})$  interlace with the velocities  $v_i = v_i(\rho)$  as follows:

$$v_1 - \sum_{i=1}^N \rho_i v_i^{\max} V'(\rho) \leq \lambda_1 \leq v_1 \leq \dots \leq v_{j-1} \leq \lambda_j \leq v_j \leq \dots \leq v_N. \quad (4.19)$$

Note that (4.19) implies that  $\lambda_1$  may be negative (corresponding to backwards-propagating characteristic information), while always  $\lambda_2, \dots, \lambda_N \geq 0$ .

### 4.2.3 Some simple difference schemes for the MCLWR model

The decisive advantage of our treatment is the simplicity of the new schemes. In that respect these schemes are comparable with a class of schemes introduced in [24]. If  $\Delta x = 1/M$  denotes a spatial meshsize,  $x_j = j\Delta x$  for  $j \in \mathbb{Z}$ ,  $\Delta t > 0$  is a time step,  $t^n := n\Delta t$ ,  $\lambda := \Delta t/\Delta x$ , and  $\rho_{i,j}^n$  denotes the approximate cell average of  $\rho_i$  on the cell  $[x_{j-1/2}, x_{j+1/2}] \times [t^n, t^{n+1}]$ , then Scheme 4 of that paper is defined by

$$\begin{aligned} \rho_{i,j}^{n+1} &= \rho_{i,j}^n - \lambda(h_{i,j+1/2}^n - h_{i,j-1/2}^n), \\ h_{i,j+1/2}^n &:= h_i(\boldsymbol{\rho}_j^n, \boldsymbol{\rho}_{j+1}^n) := \rho_{i,j}^n v_i(\rho_{j+1}^n), \quad i = 1, \dots, N. \end{aligned} \quad (4.20)$$

Scheme 10 of [24] is a version of (4.20) that is second-order accurate both in space and time. It is based on MUSCL-type spatial differencing and Runge-Kutta (RK) time stepping. The MUSCL version of  $h_i(\cdot, \cdot)$  is given by

$$h_{i,j+1/2}^{\text{MUSCL}} := h_i^{\text{MUSCL}}(\boldsymbol{\rho}_{j-1}^n, \dots, \boldsymbol{\rho}_{j+2}^n) = h_i\left(\boldsymbol{\rho}_{j+1}^n - \frac{1}{2}\boldsymbol{\sigma}_{j+1}^n, \boldsymbol{\rho}_j^n + \frac{1}{2}\boldsymbol{\sigma}_j^n\right), \quad i = 1, \dots, N, \quad (4.21)$$

where the so-called slope vector  $\sigma_j^n := (\sigma_{1,j}^n, \dots, \sigma_{N,j}^n)^\top$  is defined by the van Leer limiter [112], namely

$$\sigma_{i,j}^n = \frac{|\phi_{i,j}^n - \phi_{i,j-1}^n|(\phi_{i,j+1}^n - \phi_{i,j}^n) + |\phi_{i,j+1}^n - \phi_{i,j}^n|(\phi_{i,j}^n - \phi_{i,j-1}^n)}{|\phi_{i,j}^n - \phi_{i,j-1}^n| + |\phi_{i,j+1}^n - \phi_{i,j}^n|}. \quad (4.22)$$

Furthermore, if we write the scheme resulting from using the flux (4.21) as

$$\begin{aligned} \rho_j^{n+1} &= \rho_j^n - \Gamma_j(\rho_{j-2}^n, \dots, \rho_{j+2}^n) \\ &= \rho_j^n - \lambda[\mathbf{h}^{\text{MUSCL}}(\rho_{j-1}^n, \dots, \rho_{j+2}^n) - \mathbf{h}^{\text{MUSCL}}(\rho_{j-2}^n, \dots, \rho_{j+1}^n)], \end{aligned} \quad (4.23)$$

then the resulting MUSCL-RK version takes the following two-step form:

$$\begin{aligned} \tilde{\rho}_j^{n+1} &= \tilde{\rho}_j^n - \Gamma_j(\rho_{j-2}^n, \dots, \rho_{j+2}^n), \\ \rho_j^{n+1} &= \frac{1}{2}(\rho_j^n + \tilde{\rho}_j^{n+1} - \Gamma_j(\tilde{\rho}_{j-2}^{n+1}, \dots, \tilde{\rho}_{j+2}^{n+1})). \end{aligned} \quad (4.24)$$

Scheme 10 of [24] is defined by (4.24), with the ingredients (4.20)–(4.23). For the ease of presentation, in the remainder of the chapter we will address Schemes 4 and 10 of [24] simply as “Scheme 4” and “Scheme 10”, respectively.

### 4.3 Discretization of the Lagrangian step

Before introducing a Lagrangian method we observe that defining  $\tau := 1/\rho$ , we obtain from (4.8) the conservation of mass equation in Lagrangian coordinates

$$\rho \partial_t \tau - \partial_x v = 0. \quad (4.25)$$

In other words, solving (4.8), or equivalently (4.25), means solving the original equation (4.5) on a moving referential mesh with velocity  $v$ . Let us then denote  $v_{j+1/2}^n$  an approximate value of  $v(\rho)$  at the interface point  $x = x_{j+1/2}$  at time  $t^n$  and assume now that  $\{\rho_j^n\}_{j \in \mathbb{Z}}$  is an approximate solution (4.5) at time  $t = t^n$  and used as initial condition for (4.25). Then a numerical solution  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  of (4.25) at time  $\Delta t$  can be naturally computed by

$$\rho_j^{n+1,-} [\Delta x + (v_{j+1/2}^n - v_{j-1/2}^n) \Delta t] = \rho_j^n \Delta x, \quad j \in \mathbb{Z}, \quad (4.26)$$

since (4.26) expresses that the initial mass in the cell  $[x_{j-1/2}, x_{j+1/2}]$  at time  $t^n$  (the right-hand side) equals the mass on the modified cell  $[\bar{x}_{j-1/2}, \bar{x}_{j+1/2}]$  at time  $\Delta t$  (the left-hand side), where the new interface positions are

$$\bar{x}_{j+1/2} = x_{j+1/2} + v_{j+1/2}^n \Delta t \quad \text{for all } j.$$

This is illustrated in Figure 4.1. In particular, with this discretization and using the transformation  $\tau_j^n = 1/\rho_j^n$  in (4.26), we obtain the following discretization of (4.25):

$$\rho_j^n (\tau_j^{n+1,-} - \tau_j^n) = \lambda (v_{j+1/2}^n - v_{j-1/2}^n).$$

A natural choice for the velocity values in the interface point is  $v_{j+1/2} := v(\rho_{j+1}^n)$  for all  $j$ . A general theory about Lagrangian schemes can be found in [56]. Now, we indicate some properties of the numerical solution of the Lagrangian scheme (4.26) under certain CFL conditions.

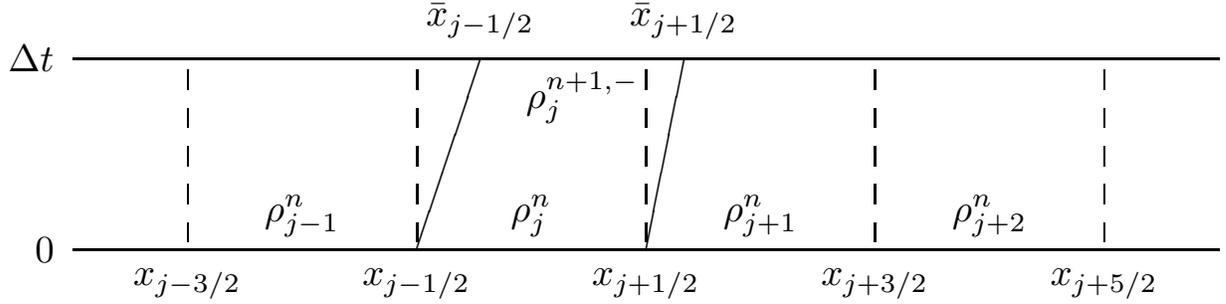


Figure 4.1: Illustration of the Lagrangian step.

**Lemma 4.1** *Assume that the following pair of CFL conditions hold:*

$$\lambda v(\rho) \leq 1 \quad \text{for } 0 \leq \rho \leq \rho_{\max}, \quad (4.27)$$

$$-1 \leq \lambda \rho_{\max} v'(\rho) \leq 0 \quad \text{for } 0 \leq \rho \leq \rho_{\max}. \quad (4.28)$$

If  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  denotes the numerical solution produced by the scheme (4.26), then the following maximum property holds:

$$\min\{\rho_j^n, \rho_{j+1}^n\} \leq \rho_j^{n+1,-} \leq \max\{\rho_j^n, \rho_{j+1}^n\} \quad \text{for all } j \in \mathbb{Z}. \quad (4.29)$$

**Proof.** We have  $\Delta x + (v_{j+1}^n - v_j^n)\Delta t \geq \Delta x$  if  $v_{j+1}^n \geq v_j^n$ , that is, if  $\rho_j^n \geq \rho_{j+1}^n$ . In this case we obtain from (4.26)

$$\rho_j^{n+1,-} \leq \rho_j^n = \max\{\rho_j^n, \rho_{j+1}^n\},$$

and the lower bound

$$\rho_j^{n+1,-} = \frac{\rho_j^n \Delta x}{\Delta x + \Delta t(v_{j+1}^n - v_j^n)} = \frac{\rho_j^n}{1 + \lambda(v_{j+1}^n - v_j^n)} \geq \min\{\rho_j^n, \rho_{j+1}^n\} = \rho_{j+1}^n$$

is valid if and only if

$$(\rho_{j+1}^n - \rho_j^n) \left( 1 + \lambda \frac{v_{j+1}^n - v_j^n}{\rho_{j+1}^n - \rho_j^n} \rho_{j+1}^n \right) = (\rho_{j+1}^n - \rho_j^n) (1 + \lambda v'(\zeta_{j+1/2}^n) \rho_{j+1}^n) \leq 0,$$

where  $\zeta_{j+1/2}^n \in [\rho_{j+1}^n, \rho_j^n]$  is an intermediate value; that is, if

$$1 + \lambda v'(\zeta_{j+1/2}^n) \rho_{j+1}^n \geq 0, \quad (4.30)$$

However, (4.30) holds if (4.28) is satisfied. Now, if  $v_{j+1}^n \leq v_j^n$ , by (4.27) we get

$$0 \leq \Delta t v_{j+1}^n \leq \Delta t v_j^n \leq \Delta x,$$

which is equivalent to

$$0 \leq \Delta t(v_{j+1}^n - v_j^n) + \Delta x \leq \Delta x,$$

which in turn implies that

$$\rho_j^n = \min\{\rho_j^n, \rho_{j+1}^n\} \leq \rho_j^{n+1,-}.$$

To establish the upper bound, we note that under the present conditions,

$$1 + \lambda(v_{j+1}^n - v_j^n) = 1 + \lambda v'(\zeta_{j+1/2}^n)(\rho_{j+1}^n - \rho_j^n) \geq 1 + \lambda v'(\zeta_{j+1/2}^n)\rho_{j+1}^n,$$

from which it is easy to deduce that

$$\rho_j^{n+1,-} = \frac{\rho_j^n}{1 + \lambda(v_{j+1}^n - v_j^n)} \leq \max\{\rho_j^n, \rho_{j+1}^n\} = \rho_{j+1}^n$$

holds if (4.30) is valid, where  $\rho_j^n \leq \zeta_{j+1/2}^n \leq \rho_{j+1}^n$  is an intermediate value, which again is ensured if (4.28) is valid.  $\blacksquare$

**Remark 4.2** *The CFL conditions (4.27) and (4.28) are equivalent to classical CFL condition for one-dimensional conservation laws,  $\lambda|f'(\rho)| \leq 1$ .*

## 4.4 Anti-diffusive schemes for the remap step

After the Lagrangian step, the new values  $\rho_j^{n+1,-}$  represent approximate values of the density on a moved mesh with new cells  $[\bar{x}_{j-1/2}, \bar{x}_{j+1/2}]$  for all  $j$ . In order to avoid dealing with moving meshes, a so-called remap step is necessary to define the new approximations  $\rho_j^{n+1}$  on the uniform mesh with cells  $[x_{j-1/2}, x_{j+1/2}]$ . Figure 4.1 illustrates that this step amounts to “averaging” the density values at time  $\Delta t$  on the cells  $[x_{j-1/2}, x_{j+1/2}]$ . It is clear that this average step can equivalently be reformulated by the solution of the transport equation (4.9) with initial data defined by  $\rho_j^{n+1,-}$  on each cell  $[x_{j-1/2}, x_{j+1/2}]$ . This is illustrated in Figure 4.2. The aim of this section is to propose and investigate several discretizations of (4.9), whereby we seek to introduce as little numerical diffusion as possible.

### 4.4.1 Anti-diffusive schemes

Here, we describe the conditions analyzed in [46, 15, 14] for solving (4.9) with initial condition  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  by using an anti-diffusive numerical scheme in the form

$$\rho_j^{n+1} = \rho_j^{n+1,-} - \bar{V}_j^n \lambda (\rho_{j+1/2}^{n+1,-} - \rho_{j-1/2}^{n+1,-}). \quad (4.31)$$

Here  $\bar{V}_j^n$  is a velocity value,

$$\bar{V}_j^n = \bar{V}(\rho_{j+1/2}^{n+1,-}, \rho_{j-1/2}^{n+1,-}, \rho_j^n, \rho_{j+1}^n, \rho_j^{n+1,-}),$$

which will be chosen in such a way that the whole scheme (4.26), (4.31) is conservative, and the quantities  $\rho_{j+1/2}^{n+1,-}$  are numerical fluxes associated with the cell interfaces  $x_{j+1/2}$ ,  $j \in \mathbb{Z}$ , and they

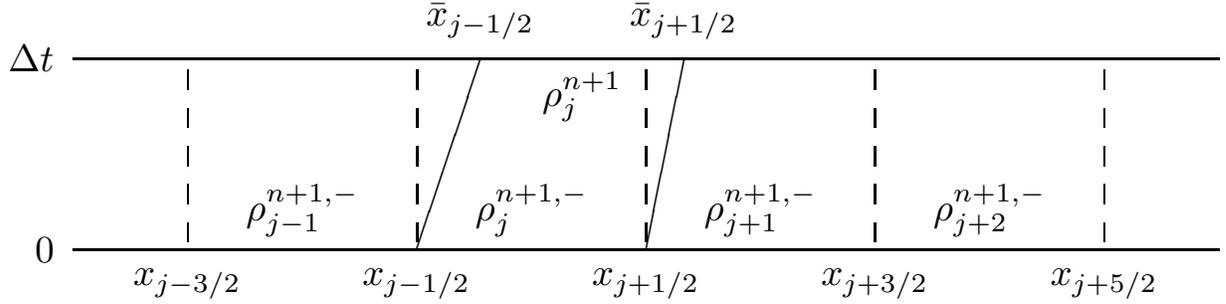


Figure 4.2: Illustration of the remap step.

will be chosen so that the scheme (4.31) satisfies certain stability and consistency requirements. To discuss this issue, we define for  $\bar{V}_j^n > 0$

$$m_{j-1/2} := \min\{\rho_j^{n+1,-}, \rho_{j-1}^{n+1,-}\}, \quad M_{j-1/2} := \max\{\rho_j^{n+1,-}, \rho_{j-1}^{n+1,-}\},$$

$$b_j^+ := M_{j-1/2} + \frac{\rho_j^{n+1,-} - M_{j-1/2}}{\bar{V}_j^n \lambda}, \quad B_j^+ := m_{j-1/2} + \frac{\rho_j^{n+1,-} - m_{j-1/2}}{\bar{V}_j^n \lambda}.$$

According to [46], to ensure the consistency property

$$\rho_{j+1/2}^{n+1,-} \rightarrow \rho \quad \text{as } \rho_j^{n+1,-}, \rho_{j+1}^{n+1,-} \rightarrow \rho, \quad 0 \leq \rho \leq \rho_{\max}, \quad \text{for all } j \in \mathbb{Z}$$

it is sufficient that

$$m_{j+1/2} \leq \rho_{j+1/2}^{n+1,-} \leq M_{j+1/2} \quad \text{for all } j \in \mathbb{Z}, \quad (4.32)$$

while for the  $L^\infty$  and TVD stability conditions it is necessary to have

$$b_j^+ \leq \rho_{j+1/2}^{n+1,-} \leq B_j^+ \quad \text{for all } j \in \mathbb{Z}. \quad (4.33)$$

For the definition of the flux  $\rho_{j+1/2}^{n+1,-}$ , note that the choice

$$\rho_{j+1/2}^{n+1,-} = \rho_j^{n+1,-} \quad \text{for all } j \in \mathbb{Z} \quad (4.34)$$

produces a diffusive and stable scheme, while for

$$\rho_{j+1/2}^{n+1,-} = \rho_{j+1}^{n+1,-} \quad \text{for all } j \in \mathbb{Z}$$

we obtain an anti-diffusive but unstable scheme. For this reason, Després and Lagoutière [46] proposed to choose  $\rho_{j+1/2}^{n+1,-}$  as close to the value  $\rho_{j+1}^{n+1,-}$  as possible, subject to the constraints (4.32) and (4.33). In the following section we discuss how to choose this numerical flux. In the following lemma, the first part of which is proved in [46] and extended to the case of nonlinear conservation laws by Bouchut [15], we resume the existence and properties of the schemes defined by (4.31). For this purpose, conditions (4.32) and (4.33) can be reduced by using the notation

$$a_{j+1/2} := \max\{b_j^+, m_{j+1/2}\}, \quad A_{j+1/2} := \min\{B_j^+, M_{j+1/2}\}.$$

**Lemma 4.3** *Assume that condition (4.27) is in effect. Then*

$$a_{j+1/2} \leq \rho_j^{n+1,-} \leq A_{j+1/2} \quad \text{for all } j \in \mathbb{Z},$$

and for any flux that satisfies

$$\rho_{j+1/2}^{n+1,-} \in [a_{j+1/2}, A_{j+1/2}] \quad \text{for all } j \in \mathbb{Z}, \quad (4.35)$$

the scheme (4.31) is  $L^\infty$ -stable and Total Variation Diminishing (TVD), i.e.

$$\min\{\rho_{j-1}^{n+1,-}, \rho_j^{n+1,-}\} \leq \rho_j^{n+1} \leq \max\{\rho_{j-1}^{n+1,-}, \rho_j^{n+1,-}\} \quad \text{for all } j \in \mathbb{Z}, \quad (4.36)$$

$$\sum_{j \in \mathbb{Z}} |\rho_{j+1}^{n+1} - \rho_j^{n+1}| \leq \sum_{j \in \mathbb{Z}} |\rho_{j+1}^{n+1,-} - \rho_j^{n+1,-}| \quad \text{for } n \in \mathbb{N}_0 := \{0, 1, 2, \dots\}. \quad (4.37)$$

Moreover, for each  $n$  there exist numbers  $\alpha_j \in [0, 1]$  such that

$$\rho_j^{n+1,-} = \alpha_j \rho_{j-1/2}^{n+1,-} + (1 - \alpha_j) \rho_{j+1/2}^{n+1,-}. \quad (4.38)$$

**Proof.** In [46] the properties (4.36) and (4.37) are proved. We only prove property (4.38). To this end, assume that (4.27) is in effect and that  $\rho_{j+1/2}^{n+1,-}$  satisfies (4.32) and (4.33). If  $\rho_j^{n+1,-} = M_{j-1/2}$ , then (4.33) implies that

$$\rho_j^{n+1,-} \leq \rho_{j+1/2}^{n+1,-} \leq \max\{\rho_{j-1/2}^{n+1,-}, \rho_{j+1/2}^{n+1,-}\}. \quad (4.39)$$

On the other hand, from (4.32) we have  $\rho_{j-1/2}^{n+1,-} \leq M_{j-1/2} = \rho_j^{n+1,-}$ , and thus

$$\min\{\rho_{j-1/2}^{n+1,-}, \rho_{j+1/2}^{n+1,-}\} \leq \rho_j^{n+1,-}. \quad (4.40)$$

Combining (4.39) and (4.40) we obtain

$$\min\{\rho_{j-1/2}^{n+1,-}, \rho_{j+1/2}^{n+1,-}\} \leq \rho^{n+1,-} \leq \max\{\rho_{j-1/2}^{n+1,-}, \rho_{j+1/2}^{n+1,-}\} \quad (4.41)$$

The proof is similar if  $\rho_j^{n+1,-} = m_{j-1/2}$ . From (4.41) we deduce (4.38).  $\blacksquare$

#### 4.4.2 Choice of the numerical fluxes

We have specified in Section 4.4.1 the stability bounds (4.35) for the numerical fluxes that guarantee that the whole L-AR scheme (4.44) converges to a weak solution of (4.5). In this Subsection, we describe numerical techniques following the methodology outlined in [46, 15, 14, 47, 114] for solving (4.9) by an anti-diffusive scheme in the form (4.31). As we mentioned before, the choice (4.34) leads to a stable but diffusive scheme, so the idea is to choose the numerical flux  $\rho_{j+1/2}^{n+1,-}$  as close as possible to the value the downwind value of the numerical solution  $\rho_{j+1}^{n+1,-}$ , when the CFL condition (4.27) is satisfied.

##### Limited downwind anti-diffusive flux

This numerical flux was formulated by Desprès and Lagoutière [46] and is defined by

$$\rho_{j+1/2}^{n+1,-} := \operatorname{argmin}_{\rho \in [a_{j+1/2}, A_{j+1/2}]} |\rho - \rho_{j+1}^{n+1,-}| = \min\left\{\max\{\rho_{j+1}^{n+1,-}, a_{j+1/2}\}, A_{j+1/2}\right\}. \quad (4.42)$$

In Lemma 4.3 it is proved that this scheme satisfies (4.35). Following [14] we denote this scheme for (4.9) by *UBee*, and refer to the corresponding complete L-AR scheme based on (4.42) as *L-UBee* scheme.

### Relaxed anti-diffusive flux

An equivalent form of scheme (4.42) was formulated by Bouchut [15] for the advection equation as follows:

$$\rho_{j+1/2}^{n+1,-} := \rho_{j+1/2}^{\text{diss}} + \text{minmod}(\rho_{j+1/2}^{\text{L}} - \rho_{j+1/2}^{\text{diss}}, \rho_{j+1/2}^{\text{R}} - \rho_{j+1/2}^{\text{diss}}),$$

where the standard minmod function is defined by

$$\text{minmod}(a, b) := \begin{cases} \text{sgn}(a) \min\{|a|, |b|\} & \text{if } \text{sgn } a = \text{sgn } b, \\ 0 & \text{otherwise,} \end{cases}$$

the dissipative flux  $\rho_{j+1/2}^{\text{diss}}$  is the classical upwind flux, and  $\rho_{j+1/2}^{\text{L}}$  and  $\rho_{j+1/2}^{\text{R}}$  are the extremal left-wind and right-wind fluxes. These quantities are defined as follows:

$$\rho_{j+1/2}^{\text{diss}} = \rho_j^{n+1,-}, \quad \rho_{j+1/2}^{\text{L}} = \frac{\rho_j^{n+1,-} - \rho_{j-1}^{n+1,-}}{\lambda \bar{V}_j^n} + \rho_{j-1}^{n+1,-}, \quad \rho_{j+1/2}^{\text{R}} = \rho_{j+1}^{n+1,-}.$$

A modification described in [114] for advection equation consists in applying a relaxed anti-diffusive flux as follows:

$$\rho_{j+1/2}^{n+1,-} := \rho_{j+1/2}^{\text{diss}} + \varphi_j \text{minmod}(\rho_{j+1/2}^{\text{L}} - \rho_{j+1/2}^{\text{diss}}, \rho_{j+1/2}^{\text{R}} - \rho_{j+1/2}^{\text{diss}}).$$

Here  $\varphi_j \in [0, 1]$  is a discontinuity indicator with  $\varphi_j \approx 0$  in smooth regions and  $\varphi_j \approx 1$  near a discontinuity. This choice of  $\varphi_j$  guarantees that (4.35) is satisfied. We denote this scheme for (4.9) by *rUBee*, and the corresponding L-AR scheme by *L-rUBee* scheme. The discontinuity indicator is chosen as

$$\varphi_j = \frac{\beta_j}{\beta_j + \gamma_j},$$

where

$$\alpha_j = |\rho_{j-1}^{n+1,-} - \rho_j^{n+1,-}|^2 + \varepsilon, \quad \beta_j = \left| \frac{\alpha_j}{\alpha_{j-1}} + \frac{\alpha_{j+1}}{\alpha_{j+2}} \right|^2, \quad \gamma_j = \frac{|\rho_{\max}^{n+1,-} - \rho_{\min}^{n+1,-}|^2}{\alpha_j},$$

where  $\rho_{\max}^{n+1,-}$  and  $\rho_{\min}^{n+1,-}$  are the maximum and minimum values of  $\rho_j^{n+1,-}$  for all grid points,  $\varepsilon$  is a small positive number taken as  $\varepsilon = 10^{-6}$ . Clearly,  $0 \leq \alpha_j \leq 1$ . Near a discontinuity,  $\gamma_j \ll \beta_j$ , so  $\varphi_j \approx 1$ , and  $\varphi_j = O(\Delta x^2)$  in smooth regions.

### NBee scheme

This scheme, which was proposed by Bokanowski and Zidani in [14] for linear transport equation, corresponds to a second-order scheme in space which is more diffusive than the U-Bee scheme, and which is defined by

$$\rho_{j+1/2}^{n+1,-} := \rho_j^{n+1,-} + \frac{1 - \bar{\lambda}}{2} \varphi_j (\rho_{j+1}^{n+1,-} - \rho_j^{n+1,-}), \quad (4.43)$$

where  $\bar{\lambda} = \lambda \bar{V}_j^n$  and

$$\varphi_j = \varphi^{\text{NB}}(r_j, \bar{\lambda}), \quad r_j = \frac{\rho_j^{n+1,-} - \rho_{j-1}^{n+1,-}}{\rho_{j+1}^{n+1,-} - \rho_j^{n+1,-}},$$

$$\varphi^{\text{NB}}(r, \bar{\lambda}) = \max \left\{ 0, \min \left\{ 1, \frac{2r}{\bar{\lambda}} \right\}, \min \left\{ r, \frac{2}{1 - \bar{\lambda}} \right\} \right\}.$$

It is proved in [14] that the numerical flux (4.43) satisfies the assumptions of Lemma 4.3. This so-called NBee scheme (for (4.9)) is written here in a limiter version. We refer to the corresponding L-AR scheme as *L-NBee* scheme.

#### 4.4.3 Lagrangian-anti-diffusive remap (L-AR) schemes, scalar case ( $N = 1$ )

We are now able to describe the subclass of Lagrangian-anti-diffusive remap (L-AR) schemes of LR schemes. Assume we have a numerical solution  $\{\rho_j^n\}_{j \in \mathbb{Z}}$  that approximate the solution of (4.5) at time  $t = t^n$  and wish to advance this solution to  $t = t^{n+1} = t^n + \Delta t$ , where  $\Delta t$  is subject to certain CFL-type restrictions. To this end, two steps are performed successively:

1. *Lagrangian step.* Consider that  $\{\rho_j^n\}_{j \in \mathbb{Z}}$  is an initial solution for (4.8). Then we can obtain a numerical solution  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  after an evolution over a time interval of length  $\Delta t$ , by using scheme (4.26).
2. *Anti-diffusive remap step.* Solve equation (4.9) with initial condition  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  using an anti-diffusive scheme (4.31) for a specific choice of  $\bar{V}_j^n$ , obtaining a numerical solution  $\{\rho_j^{n+1}\}_{j \in \mathbb{Z}}$  which approximate the solution of (4.5) at time  $t = t^{n+1} = t^n + \Delta t$ .

In the next lemma, the choice of  $\bar{V}_j^n$  is motivated by the existence of a classical conservative update formula for the whole LR scheme (4.26), (4.31).

**Lemma 4.4** *Assume that the CFL conditions (4.27) and (4.28) are satisfied. Then there exists a definition of  $\bar{V}_j^n \in [\min\{v_j^n, v_{j+1}^n\}, \max\{v_j^n, v_{j+1}^n\}]$  such that the complete L-AR scheme can be written in the form*

$$\rho_j^{n+1} = \rho_j^n - \lambda(\rho_{j+1/2}^{n+1,-} v_{j+1}^n - \rho_{j-1/2}^{n+1,-} v_j^n), \quad j \in \mathbb{Z}, \quad n \in \mathbb{N}_0. \quad (4.44)$$

**Proof.** Let  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  be a solution of (4.8) obtained by numerical scheme (4.26). Using this solution we solve equation (4.9) by the scheme (4.31), where the value  $\bar{V}_j^n$  still needs to be determined in such a way that the resulting scheme is conservative. Replacing  $\rho_j^{n+1,-}$  in (4.31) by

$$\rho_j^{n+1,-} = \rho_j^n - \lambda(v_{j+1}^n - v_j^n) \rho_j^{n+1,-},$$

we obtain

$$\rho_j^{n+1} = \rho_j^n - \lambda \bar{V}_j^n (\rho_{j+1/2}^{n+1,-} - \rho_{j-1/2}^{n+1,-}) - \lambda (v_{j+1}^n - v_j^n) \rho_j^{n+1,-}. \quad (4.45)$$

Since  $\rho_{j+1/2}^{n+1,-}$  satisfies the assumptions of Lemma 4.3, we may conclude from (4.38) that there exist numbers  $\alpha_j^n \in [0, 1]$  such that

$$\rho_j^{n+1,-} = \alpha_j^n \rho_{j-1/2}^{n+1,-} + (1 - \alpha_j^n) \rho_{j+1/2}^{n+1,-}, \quad j \in \mathbb{Z},$$

i.e. we may define

$$\alpha_j^n = \alpha(\rho_{j+1/2}^{n+1,-}, \rho_{j-1/2}^{n+1,-}, \rho_j^{n+1,-}) = \begin{cases} \frac{\rho_j^{n+1,-} - \rho_{j+1/2}^{n+1,-}}{\rho_{j-1/2}^{n+1,-} - \rho_{j+1/2}^{n+1,-}} & \text{if } \rho_{j-1/2}^{n+1,-} - \rho_{j+1/2}^{n+1,-} \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

With  $\alpha_j^n$  defined in this way, we set

$$\bar{V}_j^n := (1 - \alpha_j^n)v_j^n + \alpha_j^n v_{j+1}^n$$

and replacing in equation (4.45) we obtain (4.44). ■

**Remark 4.5** *From a practical point of view, the L-AR schemes are implemented by simply using the equivalent form (4.44). From a theoretical point of view, the Lagrangian-remap decomposition of (4.44) is used to prove the stability properties in Lemma 4.7 below, using Lemmas 4.1 and 4.3 for the Lagrangian and remap steps, respectively.*

**Remark 4.6** *Note that (4.31) define a fully implicit numerical scheme according with definition of  $\bar{V}_j^n$ . A way to overcome this difficulty is replacing  $\bar{V}$  by  $\max\{v_j^n, v_{j+1}^n\}$  in (4.31) and in the definitions of  $b_j^+$  and  $B_j^+$ , with this we can calculate the numerical fluxes  $\rho_{j+1}^{n+1,-}$  satisfying the constraints (4.33).*

Note that the numerical scheme (4.44) is written conservative form

$$\rho_j^{n+1} = \rho_j^n - \lambda(F_{j+1/2}^n - F_{j-1/2}^n),$$

where we define the numerical flux as

$$F_{j+1/2}^n := F(\rho_{j-1}^n, \rho_j^n, \rho_{j+1}^n, \rho_{j+2}^n) := \rho_{j+1/2}^{n+1,-} v_{j+1}^n.$$

This four-point numerical flux is consistent with the flux  $f(\rho) = \rho v(\rho)$  since by (4.26) and (4.32), we have

$$\rho_{j-1}^{n+1,-}, \rho_j^{n+1,-}, \rho_{j+1}^{n+1,-} \rightarrow \rho \quad \text{as } \rho_{j-1}^n, \dots, \rho_{j+2}^n \rightarrow \rho.$$

This eventually means that  $F(\rho, \dots, \rho) = \rho v(\rho)$ .

Next, we prove some properties for the numerical scheme (4.44).

**Lemma 4.7** *Assume that the CFL conditions (4.27) and (4.28) are satisfied. Then the numerical scheme (4.44) has the TVD property, is  $L^\infty$ -stable, and as consequence of (4.29) and (4.36) it satisfies the maximum property*

$$\min\{\rho_{j-1}^n, \rho_j^n, \rho_{j+1}^n\} \leq \rho_j^{n+1} \leq \max\{\rho_{j-1}^n, \rho_j^n, \rho_{j+1}^n\} \quad \text{for all } j \in \mathbb{Z}, n \in \mathbb{N}_0.$$

**Proof.** We recall from Lemma 4.3 that if (4.27) is satisfied and the scheme associated with the remap step, (4.31), satisfies (4.35), then (4.31) has the TVD property (4.37). For the Lagrangian step, we obtain from (4.26)

$$\begin{aligned} \rho_{j+1}^{n+1,-} - \rho_j^{n+1,-} &= [1 + \lambda \rho_j^{n+1,-} v'(\zeta_{j+1/2}^n)] (\rho_{j+1}^n - \rho_j^n) \\ &\quad - \lambda \rho_{j+1}^{n+1,-} v'(\zeta_{j+3/2}^n) (\rho_{j+3/2}^n - \rho_{j+1}^n). \end{aligned} \tag{4.46}$$

Since  $\rho_j^{n+1,-} \geq 0$ ,  $v'(\rho) \leq 0$  and

$$1 + \lambda \rho_j^{n+1,-} v'(\zeta_{j+1/2}^n) \geq 1 + \lambda \rho_{\max} v'(\zeta_{j+1/2}^n) \geq 0$$

due to the CFL condition (4.28), (4.46) implies that

$$\begin{aligned} |\rho_{j+1}^{n+1,-} - \rho_j^{n+1,-}| &\leq [1 + \lambda \rho_j^{n+1,-} v'(\zeta_{j+1/2}^n)] |\rho_{j+1}^n - \rho_j^n| \\ &\quad - \lambda \rho_{j+1}^{n+1,-} v'(\zeta_{j+3/2}^n) |\rho_{j+2}^n - \rho_{j+1}^n| \quad \text{for all } j \in \mathbb{Z}. \end{aligned}$$

Summing over  $j \in \mathbb{Z}$ , we get

$$\begin{aligned} \sum_{j \in \mathbb{Z}} |\rho_{j+1}^{n+1,-} - \rho_j^{n+1,-}| &\leq \sum_{j \in \mathbb{Z}} |\rho_{j+1}^n - \rho_j^n| + \lambda \sum_{j \in \mathbb{Z}} \rho_j^{n+1,-} v'(\zeta_{j+1/2}^n) |\rho_{j+1}^n - \rho_j^n| \\ &\quad - \lambda \sum_{j \in \mathbb{Z}} \rho_{j+1}^{n+1,-} v'(\zeta_{j+3/2}^n) |\rho_{j+2}^n - \rho_{j+1}^n| \\ &= \sum_{j \in \mathbb{Z}} |\rho_{j+1}^n - \rho_j^n|. \end{aligned} \tag{4.47}$$

Then, from (4.37) and (4.47) we obtain the TVD property for the numerical scheme (4.44) under the natural CFL conditions (4.27) and (4.28). The  $L^\infty$  bound is a consequence of the TVD property, i.e.,

$$|\rho_j^{n+1}| \leq |\rho_j^{n+1} - \rho_{j-1}^{n+1}| + |\rho_{j-1}^{n+1} - \rho_{j-2}^{n+1}| + \cdots \leq \sum_{j \in \mathbb{Z}} |\rho_{j+1}^{n+1} - \rho_j^{n+1}| \leq \sum_{j \in \mathbb{Z}} |\rho_{j+1}^0 - \rho_j^0|.$$

■

**Remark 4.8** A consequence of Lemma 4.7 is that under CFL conditions (4.27) and (4.28) if  $\rho_0 \in L^1(\mathbb{R})$ , the numerical solution of scheme (4.44) converges in  $L^\infty([0, T], L^1_{\text{loc}})$  to a weak solution of (4.5), see [56].

#### 4.4.4 The multi-species case ( $N \geq 1$ ) and CFL condition

For  $N$  species we can apply formula (4.44) in a component-wise manner for each species as follows:

$$\rho_{i,j}^{n+1} = \rho_{i,j}^n - \lambda (\rho_{i,j+1/2}^{n+1,-} v_{i,j+1}^n - \rho_{i,j-1/2}^{n+1,-} v_{i,j}^n), \quad i = 1, \dots, N, \quad j \in \mathbb{Z}, \quad n \in \mathbb{N}_0. \tag{4.48}$$

With respect to the CFL condition, inequalities (4.27) and (4.28) are conditions to guarantee positivity, TVD property and maximum principle for numerical solution of scalar conservation laws by is not generally hold in the system case. Below we derive the form of a CFL condition by requiring that a certain invariant region be preserved. This invariant region is defined as

$$\mathcal{D}_{\rho_{\max}} := \{(\rho_1, \dots, \rho_N)^T \in \mathbb{R}^N : \rho_1 \geq 0, \dots, \rho_N \geq 0, \rho = \rho_1 + \dots + \rho_N \leq \rho_{\max}\}.$$

**Theorem 4.9** Consider the numerical scheme (4.48) where the velocities function are defined as (4.3) with  $v_1^{\max} < \dots < v_N^{\max}$  and the hindrance factor  $V(\rho)$  satisfies (4.4). If  $\rho_j^n \in \mathcal{D}_{\rho_{\max}}$  for  $j \in \mathbb{Z}$  and the CFL conditions

$$\lambda v_N^{\max} \leq 1, \quad \lambda \rho_{\max} v_N^{\max} \max_{0 \leq \rho \leq \rho_{\max}} |V'(\rho)| \leq 1 \quad (4.49)$$

are satisfied at level time  $n$ , then  $\rho_j^{n+1} \in \mathcal{D}_{\rho_{\max}}$  for  $j \in \mathbb{Z}$ .

**Proof.** Suppose that  $\rho_j^n \in \mathcal{D}_{\rho_{\max}}$ . Then, as a consequence of Lemmas 4.1 and 4.7, we obtain that  $\rho_{i,j}^{n+1} \geq 0$  for  $i = 1, \dots, N$  and  $j \in \mathbb{Z}$ . On the other hand,

$$\begin{aligned} \rho_{i,j}^{n+1} &\leq \rho_{i,j}^n + \lambda \rho_{i,j-1/2}^{n+1,-} v_i(\rho_j^n) \\ &\leq \rho_{i,j}^n + \lambda \rho_{i,j-1/2}^{n+1,-} v_N^{\max} V(\rho_j^n), \quad i = 1, \dots, N, \quad j \in \mathbb{Z}, \end{aligned}$$

where summing over  $i$  gives

$$\rho_j^{n+1} \leq \rho_j^n + \lambda \sum_{i=1}^N \rho_{i,j-1/2}^{n+1,-} v_N^{\max} V(\rho_j^n) \leq \rho_j^n + \lambda \rho_{\max} v_N^{\max} V(\rho_j^n) =: G(\rho_j^n),$$

where we have use the hypothesis that under CFL conditions (4.49) and by Lemmas 4.1 and 4.7, we have  $\rho_{1,j-1/2}^{n+1,-} + \dots + \rho_{N,j-1/2}^{n+1,-} \leq \rho_{\max}$ . Assumption (4.4) implies that  $G(\rho_{\max}) = \rho_{\max}$ , and as  $G'(\rho_j^n) = 1 + \lambda \rho_{\max} v_N^{\max} V'(\rho_j^n)$ , the second CFL condition in (4.49) implies that  $G$  is a decreasing function of  $\rho_j^n$ . Thus,  $\max_{0 \leq \rho_j^n \leq \rho_{\max}} G(\rho_j^n) = \rho_{\max}$ , implying that  $\rho_j^{n+1} \leq \rho_{\max}$ . ■

The CFL conditions (4.49) can be simplified, as in the case of linear velocity (4.16), to

$$\lambda v_N^{\max} \leq C_{\text{CFL}}, \quad C_{\text{CFL}} = 1. \quad (4.50)$$

## 4.5 Statistically conservative schemes

We now introduce an alternative for solving the remap step in the one-species or multi-species cases to recover updated values of the unknown on the initial mesh.

### 4.5.1 Integral remap

For the scalar case ( $N = 1$ ), if  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  is the numerical solution given by (4.26), we set

$$\rho^{n+1,-}(x) := \sum_{j \in \mathbb{Z}} \rho_j^{n+1,-} \chi_{[\bar{x}_{j-1/2}, \bar{x}_{j+1/2}]}(x).$$

To define the new approximation  $\rho_j^{n+1}$  on the uniform mesh with cells  $[x_{j-1/2}, x_{j+1/2}]$  at time  $t^{n+1}$ , we then may apply an integral remap

$$\rho_j^{n+1} := \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \rho^{n+1,-}(x) dx, \quad j \in \mathbb{Z}.$$

After some calculations one obtains

$$\Delta x \rho_j^{n+1} = (\bar{x}_{j-1/2} - x_{j-1/2}) \rho_{j-1}^{n+1,-} + (x_{j+1/2} - \bar{x}_{j-1/2}) \rho_j^{n+1,-},$$

which yields

$$\rho_j^{n+1} = \lambda v_j^n \rho_{j-1}^{n+1,-} + (1 - \lambda v_j^n) \rho_j^{n+1,-}. \quad (4.51)$$

According to the definition (4.26), a completed scheme for Lagrangian step plus remap step can be written in the form

$$\rho_j^{n+1} = \rho_j^n - \lambda (v_{j+1}^n \rho_j^{n+1,-} - v_j^n \rho_{j-1}^{n+1,-}). \quad (4.52)$$

As consequence of (4.51) and Lemma 4.1, under a CFL condition (4.27) and (4.28), the numerical scheme (4.52) is conservative, TVD,  $L^\infty$ -stable and satisfies the maximum principle; however, (4.51) produces a diffusive numerical solution.

#### 4.5.2 Random sampling remap, scalar case ( $N = 1$ )

In order to define  $\rho_j^{n+1}$  without introducing numerical diffusion, we follow a Glimm-type random sampling strategy [55]. More precisely, for given well-distributed random sequence  $\{a_n\}_{n \in \mathbb{N}}$  taking values in  $(0, 1)$  (e.g. the Van der Corput sequence), we simply set

$$\rho_j^{n+1} = \begin{cases} \rho_{j-1}^{n+1,-} & \text{if } a_{n+1} \in (0, \lambda v_j^n), \\ \rho_j^{n+1,-} & \text{if } a_{n+1} \in (\lambda v_j^n, 1). \end{cases} \quad (4.53)$$

A CFL condition obtained from (4.53) is  $\lambda v(\rho) \leq 1$ ,  $0 \leq \rho \leq \rho_{\max}$ .

#### 4.5.3 Random sampling remap, multiclass case ( $N > 1$ )

For the multi-species cases, first, we use a multi-Lagrangian approach to calculate from  $\{\rho_{i,j}^n\}_{i,j}$  the numerical solution of equation (4.8) after an evolution over a time interval of length  $\Delta t$  for each species

$$\rho_{i,j}^{n+1,-} [\Delta x + (v_{i,j+1}^n - v_{i,j}^n) \Delta t] = \rho_{i,j}^n \Delta x, \quad i = 1, \dots, N. \quad (4.54)$$

Let us set  $\boldsymbol{\rho}_j^{n+1,-} = (\rho_{1,j}^{n+1,-}, \dots, \rho_{N,j}^{n+1,-})^T$ . One first idea to define the new approximation  $\boldsymbol{\rho}_j^{n+1}$  is using a multi-species version of formula (4.53), but numerical experiments show that this strategy generates spurious oscillations and instabilities that do not disappear with the refinement. The strategy to avoid this undesirable behaviour is to use the approximate HLL Riemann solver to locally obtain an intermediate value limited by the curves generated for the extremal maximum velocities. It consists in redefining the numerical solution  $\tilde{\boldsymbol{\rho}}(x, t)$  of the Lagrangian step in the region

$[x_{j-1/2}, x_{j+1/2}] \times [0, \Delta t]$  by

$$\tilde{\rho}(x, t) = \begin{cases} \rho_{j-1}^{n+1,-} & \text{if } \frac{x - x_{j-1/2}}{t} < \sigma_L, \\ \rho_j^* & \text{if } \sigma_L < \frac{x - x_{j-1/2}}{t} < \sigma_R, \\ \rho_j^{n+1,-} & \text{if } \frac{x - x_{j-1/2}}{t} > \sigma_R, \end{cases} \quad (4.55)$$

where  $(x, t) \in [x_{j-1/2}, x_{j+1/2}] \times [0, \Delta t]$ .

Here  $\sigma_L = v_1^{\max} V(\rho_j^n)$  and  $\sigma_R = v_N^{\max} V(\rho_j^n)$  are the extremal maximum velocities and the state  $\rho_j^*$  is calculated according to the following consistency relation:

$$\begin{aligned} v_i(\rho_j^n)(\rho_{i,j}^{n+1,-} - \rho_{i,j-1}^{n+1,-}) &= \sigma_L(\rho_{i,j}^* - \rho_{i,j-1}^{n+1,-}) \\ &+ \sigma_R(\rho_{i,j}^{n+1,-} - \rho_{i,j}^*), \quad i = 1, \dots, N. \end{aligned} \quad (4.56)$$

This relation is consistent with the integral form of the system of equations

$$\partial_t \rho_i + v_i(\rho) \partial_x \rho_i = 0, \quad i = 1, \dots, N \quad (4.57)$$

over the control volume  $[x_L, x_R] \times [0, \Delta t]$  where  $x_L \leq \Delta t \sigma_L$  and  $x_R \geq \Delta t \sigma_R$ , see [111].

**Remark 4.10** We could have defined the new approximation  $\rho_j^{n+1}$  as

$$\rho_j^{n+1} = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{\rho}(x, t^{n+1,-}) dx.$$

As in the scalar case, after some calculations we obtain

$$\Delta x \rho_j^{n+1} = \Delta t \sigma_L \rho_{j-1}^{n+1,-} + \Delta t (\sigma_R - \sigma_L) \rho_j^* + (\Delta x - \sigma_R \Delta t) \rho_j^{n+1,-}. \quad (4.58)$$

Now, replacing  $\rho_j^*$  according to (4.56) we obtain

$$\rho_{i,j}^{n+1} = \lambda \rho_{i,j}^{n+1,-} - \lambda v_i(\rho_j^n)(\rho_{i,j}^{n+1,-} - \rho_{i,j-1}^{n+1,-}), \quad i = 1, \dots, N. \quad (4.59)$$

According to the definition (4.54), the system (4.59) is equivalent to

$$\rho_{i,j}^{n+1} = \rho_{i,j}^n - \lambda (v_{i,j+1}^n \rho_{i,j}^{n+1,-} - v_{i,j}^n \rho_{i,j-1}^{n+1,-}), \quad i = 1, \dots, N,$$

which is a multi-species version of scheme (4.52).

Here, we use relation (4.55) to define an anti-diffusive scheme to update the numerical value  $\rho_j^{n+1}$ , namely we set

$$\rho_j^{n+1} = \begin{cases} \rho_{j-1}^{n+1,-} & \text{if } a_{n+1} \in (0, \lambda \sigma_L), \\ \rho_j^* & \text{if } a_{n+1} \in (\lambda \sigma_L, \lambda \sigma_R), \\ \rho_j^{n+1,-} & \text{if } a_{n+1} \in (\lambda \sigma_R, 1). \end{cases} \quad (4.60)$$

A CFL condition obtained from (4.60) is  $\lambda v_N^{\max} V(\rho) \leq 1$ ,  $0 \leq \rho \leq \rho_{\max}$ ,  $i = 1, \dots, N$ .

#### 4.5.4 Lagrangian-random sampling (L-RS) scheme

We now summarize the Lagrangian-Random Sampling (L-RS) scheme for (4.1). Assume we have a numerical solution  $\{\rho_j^n\}_{j \in \mathbb{Z}}$  for time  $t = t^n$  and wish to advance the solution to  $t = t^{n+1} = t^n + \Delta t$ , where  $\Delta t$  is subject to the CFL-type restriction (4.50). To this end the following two steps are performed successively:

1. *Lagrangian step.* Suppose that  $\{\rho_j^n\}_{j \in \mathbb{Z}}$  is an initial solution for

$$\partial_t \rho_i + \rho_i \partial_x v_i(\rho) = 0, \quad i = 1, \dots, N.$$

Then we obtain a numerical solution  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  by the Lagrangian scheme (4.54).

2. *Random sampling remap step.* For  $N = 1$ , we use the two-state-per-cell random sampling step (4.53) to advance the solution to  $t^{n+1}$ . For  $N > 1$ , we approximately solve equation (4.57) with initial data  $\{\rho_j^{n+1,-}\}_{j \in \mathbb{Z}}$  by reconstructing a solution via (4.55) using the HLL approximate Riemann solver. After that, we perform the three-state-per-cell random sampling step (4.60) to obtain a numerical solution  $\{\rho_j^{n+1}\}_{j \in \mathbb{Z}}$  for  $t = t^{n+1}$ .

## 4.6 Numerical results

### 4.6.1 CFL condition, errors, and entropy test

For  $N = 1$  and a given value of  $\Delta x$ , we choose  $\Delta t$  so that (4.50) is satisfied with  $C_{\text{CFL}} = 0.95$  for the LR schemes and  $C_{\text{CFL}} = 0.8$  for Schemes 4 and 10, in agreement with the CFL condition stated in [24]. For  $N > 1$ , we use  $C_{\text{CFL}} = 0.9$  for all schemes. Furthermore, to measure the efficiency of the new schemes, we present plots of total (approximate)  $L^1$  error versus CPU time. For  $N = 1$ , numerical solutions at moderately fine discretizations are compared with the exact entropy solution, while for  $N > 1$ , where no exact solution is at hand, we employ a reference solution obtained by a high-resolution spectral WENO scheme (WENO-SPEC; see [50]) with  $\Delta x = 1/M_{\text{ref}} = 1/25600$ .

We emphasize that for  $N > 1$  we only include numerical tests that turned out non-oscillatory results. In fact, numerical tests with multiclass versions of the L-UBee and L-rUBee schemes produced bounded strongly oscillatory numerical solutions. For this reason, for the multiclass case  $N > 1$ , only the L-NBee version of L-AR schemes has been selected for further study and numerical tests.

For the total  $L^1$  error, let us denote by  $\{\rho_{i,j}^M(t)\}_{j=1}^{\mathcal{M}}$  and  $\{\rho_{i,l}^{\text{ref}}(t)\}_{l=1}^{\mathcal{M}_{\text{ref}}}$  the numerical and reference solution for the  $i$ -th component at time  $t$  calculated with  $\Delta x = 1/M$  and  $\Delta x = 1/M_{\text{ref}}$  using  $\mathcal{M} = ML$  and  $\mathcal{M}_{\text{ref}} = M_{\text{ref}}L$  cells, respectively. We use cubic interpolation from the reference grid to the  $\mathcal{M}$  cells grid to compute  $\tilde{\rho}_{i,j}^{\text{ref}}(t)$  for  $j = 1, \dots, \mathcal{M}$ . We then calculate the approximate  $L^1$  error in species  $i$  by

$$e_i(t) := \frac{1}{M} \sum_{j=1}^{\mathcal{M}} |\tilde{\rho}_{i,j}^{\text{ref}}(t) - \rho_{i,j}^M(t)|, \quad i = 1, \dots, N.$$

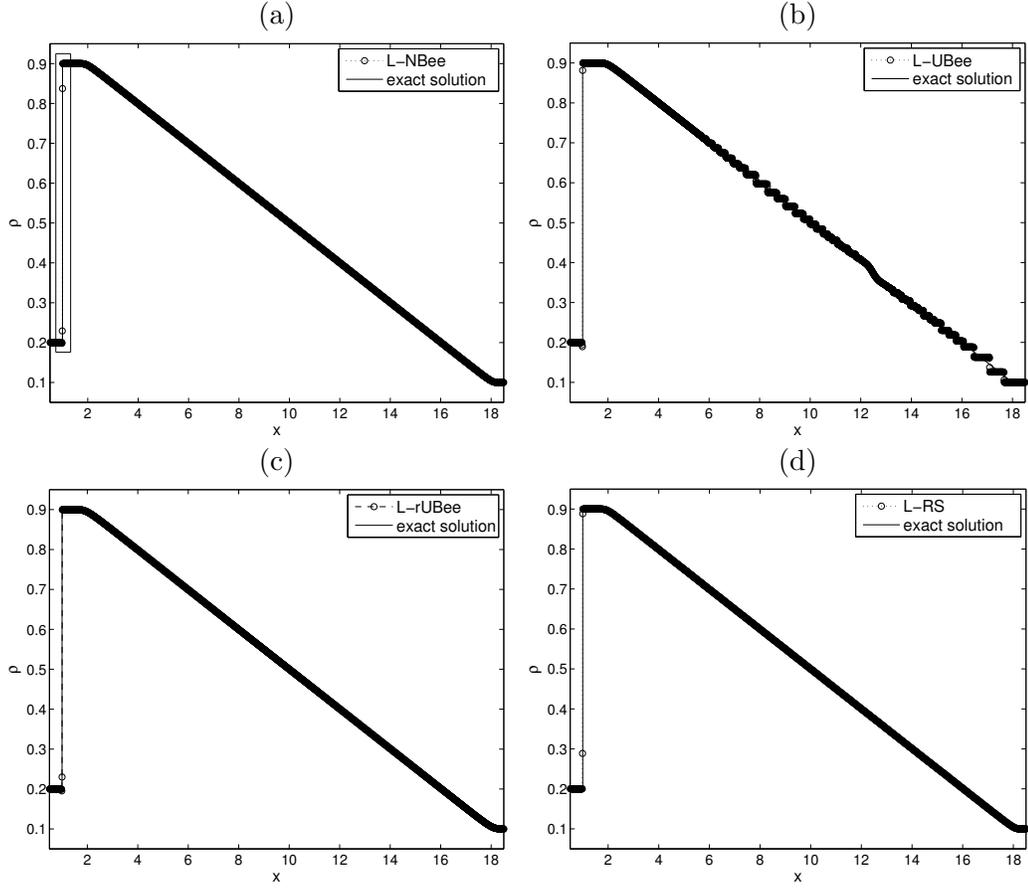


Figure 4.3: Example 4.1: numerical solution at  $t = 10$  for  $\Delta x = 0.01$  and schemes (a) L-NBee, (b) L-UBee, (c) L-rUBee, (d) L-RS.

The L-RS scheme described in Section 4.5 it is clearly non-conservative, since the remap step is based on random sampling. To measure the relative mass error as a function of time and  $\Delta x$  for the numerical solution for each species  $i$ , we evaluate

$$E_i(\Delta x; t) := \frac{1}{m_i(0)} \left| m_i(t) - \Delta x \sum_{j=1}^{\mathcal{M}} \rho_{i,j}^n \right|, \quad m_i(t) := \int_0^L \rho_i(x, t) dx,$$

where  $n = \lfloor t/\Delta t \rfloor$  and  $E_i(\Delta x; t)$  is the relative conservation of mass error of species  $i$  at time  $t$  in the interval  $[0, L]$ . In this way we define the total relative mass error by  $E(\Delta x; t) = E_1(\Delta x; t) + \dots + E_N(\Delta x; t)$ . Note that  $E(\Delta x; t)$  overestimates the relative mass error in  $\rho = \rho_1 + \dots + \rho_N$ .

Finally, in some cases we wish to test numerically whether the scheme under consideration possibly approximates an entropy solution. To this end, we define

$$E^\Delta(\Delta x, t) := \Delta x \sum_{j=1}^{\mathcal{M}} U(\rho_j^n), \quad \text{where } n = \lfloor t/\Delta t \rfloor. \quad (4.61)$$

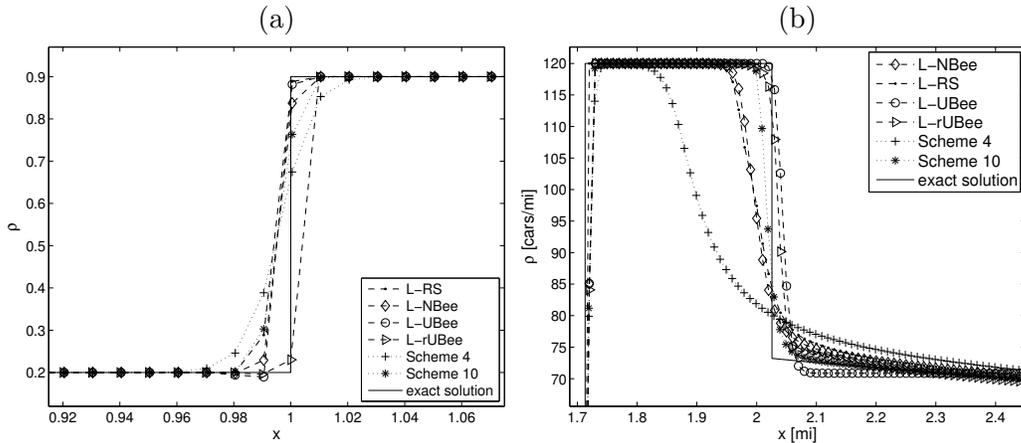


Figure 4.4: Enlarged views of (a) Example 4.1 at  $t = 10$  (cf. Figure 4.3), (b) Example 4.2 at  $t = 12.7$  (cf. Figure 4.7), calculated with  $\Delta x = 0.01$  and including solutions by Schemes 4 and 10.

Here  $U(\rho)$  is a convex entropy function. For  $N = 1$  we choose  $U(\rho) = \rho^2/2$ , and for  $N > 1$  we choose  $U(\rho) = \mathcal{E}(\rho)$ , where  $\mathcal{E}(\rho)$  is defined in (4.15). In this form, for a given value of  $\Delta x$  the function  $t \mapsto E^\Delta(\Delta x, t)$  must be non-increasing.

#### 4.6.2 Example 4.1: $N = 1$ , linear velocity

We consider the conservation law (4.5) with  $v(\rho)$  given by (4.6) with  $v^{\max} = 1$  and  $V(\rho)$  defined by (4.16) with  $\rho_{\max} = 1$ , along with the initial condition

$$\rho(x, 0) = \rho_0(x) := \begin{cases} 0.2 & \text{for } x < 2, \\ 0.9 & \text{for } 2 \leq x \leq 9, \\ 0.1 & \text{for } x > 9. \end{cases}$$

For this test, the flux  $f(\rho) = \rho v(\rho)$  is concave ( $f'' \equiv -2 < 0$ ), so according to the Lax entropy condition, the discontinuity in  $\rho_0$  sitting at  $x = 2$  evolves as a shock propagating at speed  $(f(0.9) - f(0.2))/(0.9 - 0.2) = -0.1$ , while the jump in  $\rho_0$  at  $x = 9$  gives rise to a rarefaction wave centered at that position. Figure 4.3 shows numerical results at simulated time  $t = 10$ , at which the backward-propagating shock of the exact solution has reached the position  $x = 1$  and the shock and the rarefaction wave do not yet interact, i.e., on a short interval the solution value  $\rho = 0.9$  is still present.

We observe that the shock and the rarefaction wave are adequately approximated by the L-RS, L-rUBee and L-NBee schemes, while Figure 4.3 (b) indicates that the scheme L-UBee generates “stairs” in the rarefaction wave. An enlarged view around  $x = 1$  is shown in Figure 4.4 (a), where numerical solutions are compared with those produced by Schemes 4 and 10. It appears that results by LR schemes are less diffusive those produced by other schemes.

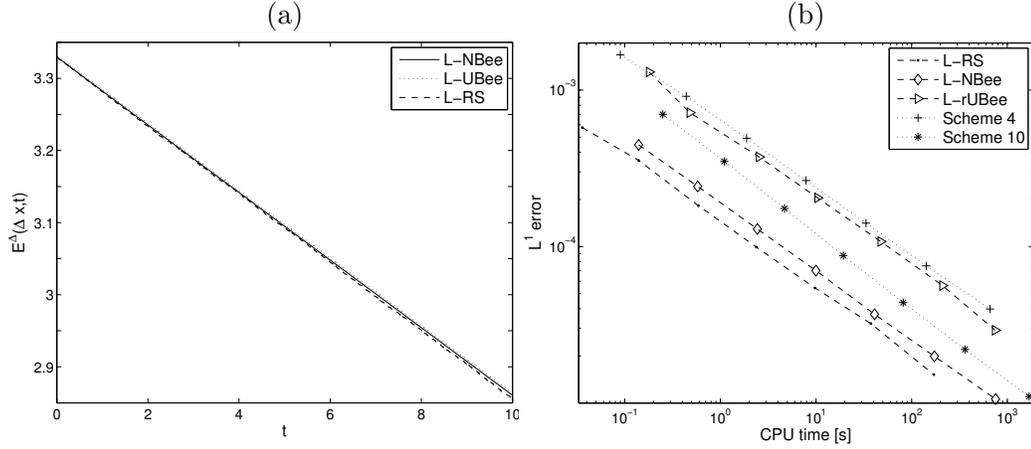


Figure 4.5: Example 4.1: (a) numerical entropy test for  $\Delta x = 1/200$ , (b)  $L^1$  error versus CPU time at  $t = 10$ . Here and in Figure 4.8,  $C_{\text{CFL}} = 0.95$  for LR schemes and  $C_{\text{CFL}} = 0.8$  for Schemes 4 and 10.

$M$	error	cr	CPU [s]	error	cr	CPU [s]	error	cr	CPU [s]	
			L-rUBee			L-UBee			L-NBee	
100	130.5	—	0.1	197.9	—	0.13	44.6	—	0.1	
200	71.7	0.86	0.4	186.5	0.08	0.56	24.2	0.88	0.5	
400	37.3	0.94	2.5	179.2	0.05	2.32	12.9	0.90	2.4	
800	20.4	0.86	10.3	163.7	0.13	9.43	7.0	0.88	9.9	
1600	10.7	0.92	47.5	159.2	0.04	38.78	3.6	0.92	40.9	
3200	5.6	0.93	210.3	146.4	0.12	166.41	1.9	0.89	173.0	
6400	2.9	0.94	747.6	134.8	0.11	709.00	1.0	0.90	753.9	
			L-RS			Scheme 4			Scheme 10	
100	57.6	—	0.03	168.1	—	0.1	70.0	—	0.2	
200	35.5	0.61	0.14	91.2	0.88	0.4	35.0	1.00	1.1	
400	18.2	0.96	0.59	49.2	0.89	1.8	17.5	0.99	4.6	
800	9.9	0.87	2.38	26.4	0.89	7.8	8.7	0.99	19.3	
1600	5.4	0.86	9.77	14.1	0.90	33.3	4.3	1.00	81.6	
3200	3.2	0.85	37.49	7.5	0.90	142.2	2.2	0.99	361.3	
6400	1.5	1.08	170.43	3.9	0.91	660.8	1.1	1.00	1691.3	

Table 4.1: Example 4.1:  $L^1$ -errors (“error”, to be multiplied by  $10^{-5}$ ), convergence rates (cr), and CPU times (CPU) at  $t = 10$  for three LR schemes and Schemes 4 and 10.

Table 4.1 and Figure 4.5 (b) show the error history, namely the approximate total  $L^1$  error, convergence rates cr and CPU time CPU for Example 4.1 for different schemes. According to Table 4.1, for the L-NBee and L-RS schemes, the error goes to zero when the mesh is refined. Furthermore, for a fixed relative error, say  $9.9 \times 10^{-5}$ , LR schemes require less CPU time than Schemes 4 and 10, but L-RS require less levels of refinement than Scheme 4, L-NBee require less than L-RS and Scheme 10 requires less than L-NBee, L-rUBee require as CPU time as Scheme 4. In this example the values of the convergence rates cr for LR schemes lie between those of Schemes 4 and 10.

As consequence of Lemma 4.7, each scheme converges to a weak solution of (4.5). To determine

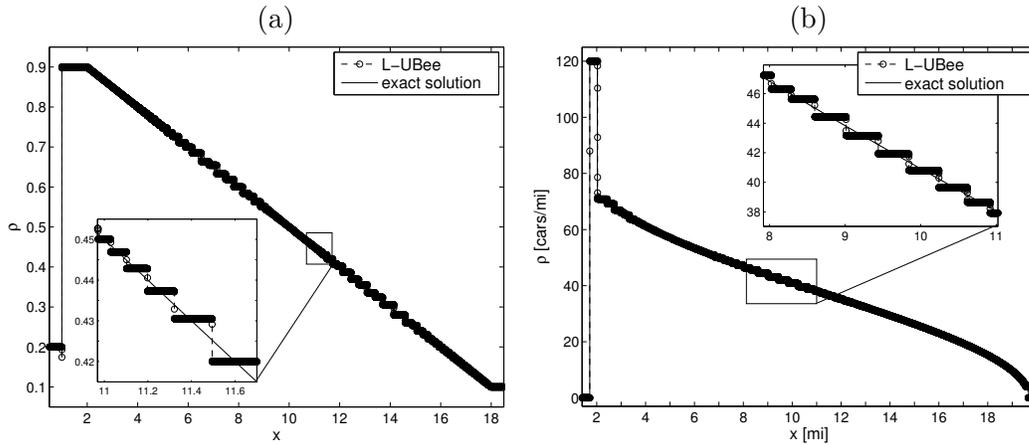


Figure 4.6: Numerical solution by scheme L-UBee with  $\Delta x = 1/12800$  and  $C_{\text{CFL}} = 0.95$  for (a) Example 4.1, (b) Example 4.2.

	Example 4.1		Example 4.2		Example 4.4		Example 4.5	
$M$	$t = 10$	$t = 12.7$	$t = 0.03$	$t = 0.11$	$t = 0.03$	$t = 0.14$		
100	3.45E-4	9.27E-3	2.82E-3	3.16E-3	6.65E-3	7.83E-3		
200	8.72E-5	3.46E-3	2.34E-3	2.61E-3	2.32E-3	2.82E-3		
400	8.60E-5	1.89E-3	1.81E-3	2.19E-3	6.07E-4	7.26E-4		
800	3.37E-5	1.63E-3	7.60E-4	9.00E-4	3.48E-4	3.87E-4		
1600	6.41E-6	1.59E-3	6.70E-4	8.50E-4	8.73E-5	1.04E-4		
3200	4.22E-6	1.47E-3	5.40E-4	6.30E-4	2.11E-5	8.82E-5		

Table 4.2: Relative mass error  $E(\Delta x; t)$  in dependence of  $\Delta x = 1/M$  for the L-RS scheme at the indicated simulated times  $t$ .

whether this weak solution is an entropy solution, in Figure 4.5 (a), we plot the total entropy (4.61) as a function of  $t$  for each LR scheme, using  $\Delta x = 1/200$ . We observe that  $E^\Delta(\Delta x, t)$  is non-increasing in  $t$ . This behaviour is maintained with the mesh is refined.

For the statistically conservative L-RS scheme, we display in Table 4.2 the relative mass error at simulated time  $t = 10$  for different levels of discretization. For Example 1 we observe that the conservation error is small already for a coarse grid and decreases with the refinement of the mesh.

In Figure 4.6 (a) we observe that when the mesh is refined, the numerical solution obtained with the L-UBee scheme produces “staircaising” that does not disappear upon refinement and the  $L^1$ -error does not appreciably go to zero when  $\Delta x \rightarrow 0$ . This phenomenon is due to the particular choice of the anti-diffusive scheme (UBee scheme) and has also been reported elsewhere for the linear advection, transport, and other equations [46, 15, 14, 47, 67, 69, 72].

### 4.6.3 Example 4.2: $N = 1$ exponential velocity

In this numerical test, we use a velocity function  $v(\rho)$  given by (4.6) with  $v^{\max} = 1$  and

$$V(\rho) = \exp(-(\rho/50)^2/2), \quad (4.62)$$

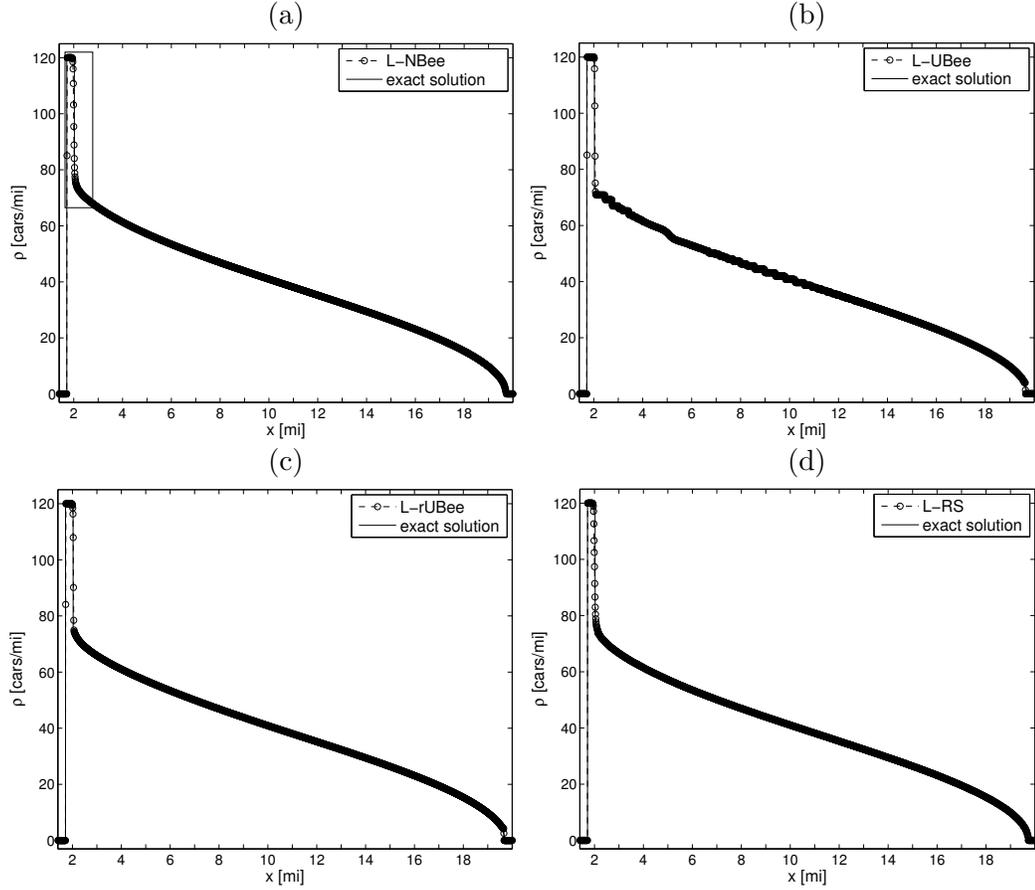


Figure 4.7: Example 4.2: schemes (a) L-NBee, (b) L-UBee, (c) L-rUBee, (d) L-RS at  $t = 12.7$  with  $\Delta x = 0.01$ .

which is the hindrance function proposed by Drake [53], and an initial condition

$$\rho(x, 0) = \rho_0(x) := \begin{cases} 120 \text{ cars/mi} & \text{for } 1 \leq x \leq 7, \\ 0 & \text{otherwise} \end{cases}$$

for  $x \in [0, 10]$ , where we employ the boundary conditions  $\rho(0, t) = \rho(20, t) = 0$  for  $t > 0$ . For this test, the flux  $f(\rho) = \rho v(\rho)$  is not concave, so the entropy solution for this initial condition contains a shock and a rarefaction wave followed by a shock. Results at  $t = 12.7$  are displayed in Figure 4.7. In Figures 4.7 (a) and (b) we observe that the L-NBee and L-RS schemes adequately approximate the shock and the rarefaction wave, while Figure 4.7 (c) indicates that the L-UBee scheme produces staircasing in the rarefaction wave (as in the linear velocity case). The “stairs” do not disappear under refinement, as is shown in Figure 4.6 (b). For the L-RS scheme in Table 4.2 we observe that the relative mass error at simulated time is small with a coarse grid and decreases with refinement of the mesh. Table 4.3 and Fig 4.8 (b) show approximate total  $L^1$  error, convergence rates  $cr$  and the CPU time  $cpu$  for Example 4.2 for different schemes. According to Table 4.3, the LR schemes are almost as efficient as Scheme 10. In Figure 4.8(a) we display the total entropy as a function

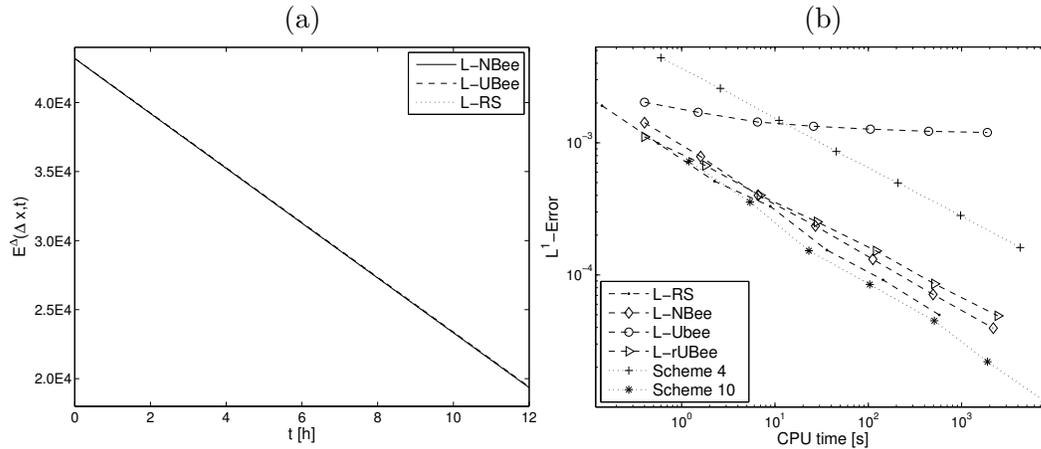


Figure 4.8: Example 4.2: (a) numerical entropy test for  $\Delta x = 1/200$ , (b)  $L^1$  error versus CPU time at  $t = 12.7$ .

$M$	L-rUBee			L-UBee			L-NBee		
	error	cr	CPU [s]	error	cr	CPU [s]	error	cr	CPU [s]
100	110.9	—	0.4	202.7	—	0.4	141.9	—	0.4
200	67.2	0.72	1.8	170.1	0.25	1.5	78.5	0.85	1.6
400	40.0	0.74	6.9	143.6	0.24	6.5	39.9	0.97	6.6
800	25.2	0.66	28.2	133.3	0.10	25.9	23.4	0.76	27.1
1600	15.2	0.72	121.3	126.7	0.07	105.2	13.1	0.83	111.8
3200	8.5	0.83	510.5	122.1	0.05	443.2	7.1	0.88	494.4
6400	4.9	0.80	2441.6	119.6	0.02	1883.5	3.9	0.84	2181.2
L-RS			Scheme 4			Scheme 10			
100	191.0	—	0.1	440.0	—	0.6	72.2	—	1.2
200	98.9	0.94	0.5	258.0	0.76	2.6	35.6	1.02	5.4
400	51.2	0.95	2.2	147.6	0.80	11.0	15.2	1.22	23.1
800	33.0	0.63	8.9	85.9	0.78	45.4	8.4	0.84	103.6
1600	15.4	1.10	35.9	49.7	0.79	208.0	4.5	0.91	510.0
3200	9.1	0.75	144.0	28.1	0.81	973.9	2.2	1.03	1892.0
6400	4.9	0.87	575.9	16.0	0.80	4230.3	1.1	1.00	7831.7

Table 4.3: Example 4.2: approximate  $L^1$ -errors (“error”, to be multiplied by  $10^{-5}$ ), cr, and CPU times for three LR schemes and Schemes 4 and 10 of [24] at simulated time  $t = 12.7$ .

of  $t$ . We observe that the function (4.61) is non-increasing in the time.

#### 4.6.4 Example 4.3: $N = 5$ , linear velocity

We consider the model (4.1) with the hindrance function (4.16),  $v_i^{\max} = i/N$ , and the initial datum

$$\rho(x, 0) = \begin{cases} (0.2, \dots, 0.2)^T & \text{for } 0 \leq x \leq 1, \\ \mathbf{0} & \text{for } x < 0 \text{ and } x > 1. \end{cases}$$

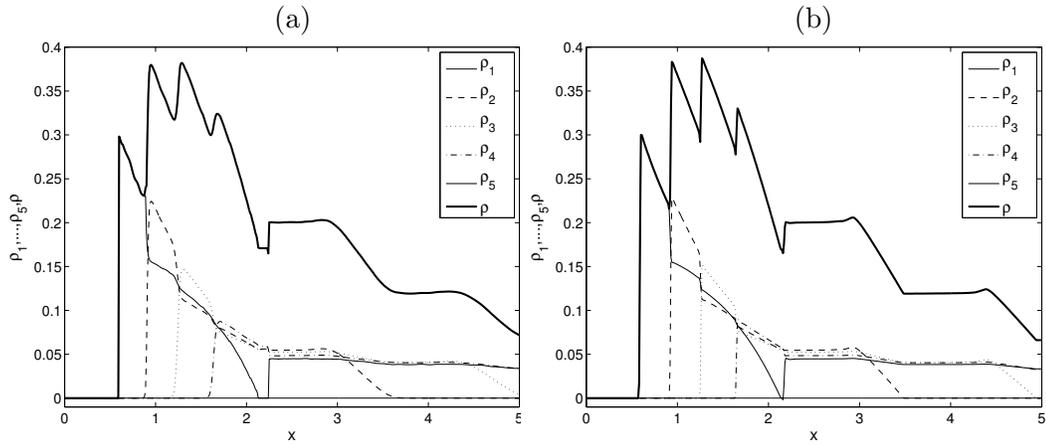


Figure 4.9: Example 4.3: numerical solution at simulated time  $t = 7$ , calculated with  $\Delta x = 0.01$  and schemes (a) L-RS, (b) L-NBee.

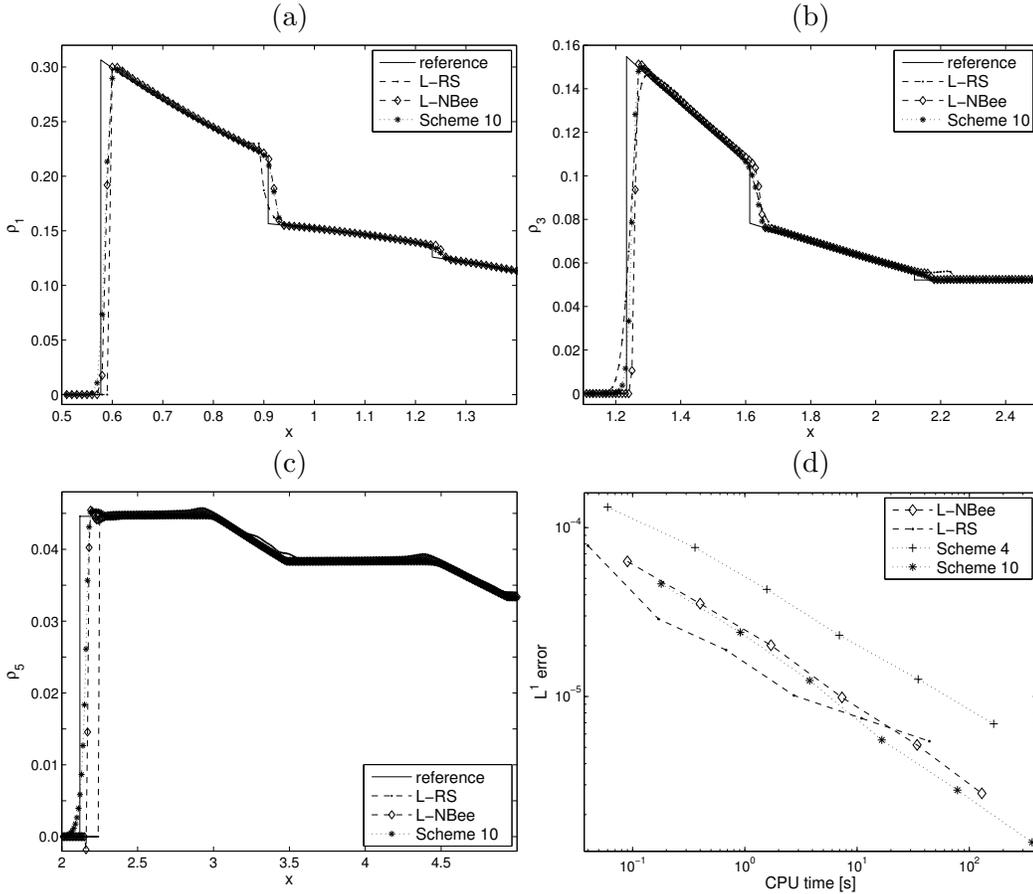


Figure 4.10: Example 4.3: (a–c) enlarged views of parts of Figure 4.9 for selected species, including numerical solutions by Schemes 4 and 10; (d) approximate total  $L^1$  errors versus CPU time.

In Figures 4.9 (a) and (b) we display the numerical solution obtained with the L-RS and L-NBee scheme, respectively, at simulated time  $T = 7$  with  $\Delta x = 1/100$ . The solution produced by L-NBee

$M$	L-NBee			L-RS			Scheme 4			Scheme 10		
	error	cr	CPU [s]	error	cr	CPU [s]	error	cr	CPU [s]	error	cr	CPU [s]
100	63.1	—	0.09	78.1	—	0.04	132.2	—	0.06	46.4	—	0.18
200	35.4	0.83	0.40	28.7	1.44	0.17	76.2	0.79	0.36	23.9	0.95	0.91
400	20.1	0.81	1.71	18.8	0.61	0.68	43.0	0.82	1.57	12.4	0.94	3.78
800	9.8	1.02	7.34	10.1	0.89	2.74	23.0	0.90	6.95	5.5	1.16	16.6
1600	5.1	0.93	34.1	7.4	0.55	10.9	12.6	0.86	34.9	2.7	0.98	78.3
3200	2.6	0.95	129.2	5.4	0.44	42.8	6.8	0.87	164.9	1.3	1.02	362.1

Table 4.4: Example 4.3: total approximate  $L^1$  errors  $e_{\text{tot}}(t)$  (“error”, to be multiplied by  $10^{-6}$ ), cr, and CPU times for two LR schemes and Schemes 4 and 10 at simulated time  $t = 7$ .

$T$	$M$	L-NBee			L-RS			Scheme 4			Scheme 10		
		error	cr	CPU [s]	error	cr	CPU [s]	error	cr	CPU [s]	error	cr	CPU [s]
0.03	100	684.9	—	0.05	1124.4	—	0.03	1197.7	—	0.04	670.7	—	0.09
	200	414.3	0.72	0.2	582.8	0.94	0.09	938.4	0.55	0.16	449.5	0.60	0.36
	400	192.1	1.10	1.0	567.1	0.59	0.39	708.2	0.70	0.74	224.2	1.00	1.58
	800	83.7	1.19	4.7	360.7	0.75	1.58	507.9	0.77	3.50	110.4	1.02	6.96
	1600	47.0	0.83	20.4	254.9	0.60	6.85	321.3	0.86	16.91	51.4	1.10	34.23
	3200	31.4	0.68	89.3	170.8	0.77	30.72	185.1	0.89	74.70	29.8	0.78	152.58
0.11	100	217.2	—	0.1	817.8	—	0.04	1477.1	—	0.02	444.9	—	0.2
	200	100.5	1.11	0.7	755.4	1.14	0.17	926.3	0.67	0.37	208.2	1.09	0.9
	400	53.4	0.91	3.1	414.6	0.86	0.70	555.6	0.73	1.57	103.6	1.00	4.0
	800	28.3	0.91	12.7	220.4	0.91	2.94	313.1	0.82	6.58	50.8	1.02	17.6
	1600	15.3	0.88	53.5	102.1	1.10	13.31	168.2	0.89	32.20	24.6	1.04	90.7
	3200	9.4	0.70	230.1	50.1	1.02	63.42	89.7	0.90	181.60	12.8	0.93	445.2

Table 4.5: Example 4.4: total approximate  $L^1$  errors  $e_{\text{tot}}(t)$  (“error”, to be multiplied by  $10^{-5}$ ), cr, and CPU times for two LR schemes and Schemes 4 and 10.

scheme appears to be less affected by numerical diffusion than the one corresponding to the L-RS scheme. In Figures 4.10 (a–c) enlarged views of the relevant parts for individual species are shown. Observe that L-NBee produces fairly sharp solutions in each individual species. Furthermore, note carefully that the numerical results of Figure 4.9 show that jumps in the total density  $\rho$  only occur “upwards”, i.e. from smaller to higher values. This is in agreement with the entropy jump condition (4.17) valid for the present case.

#### 4.6.5 Examples 4.4 and 4.5: $N = 9$ , exponential velocity

We consider the MCLWR model (4.1), where we utilize the hindrance function (4.62) proposed by Drake [53]. We present numerical test proposed in [113], where the initial density distribution is given by an isolated platoon in the congested regime for the Drake model with optimal car density  $\rho^* = 50$  cars/mi and maximum velocities  $v_i^{\max} = (52.5 + 7.5i)$  mi/h,  $i = 1, \dots, 9$ . We consider a circular highway of length  $L = 10$  mi, i.e. we set  $I := [0, 10]$  with periodic conditions in the boundary. As initial condition, a platoon is considered. Specifically we set

$$\rho(x, 0) = p(x)\rho_0(0.04, 0.08, 0.12, 0.16, 0.2, 0.16, 0.12, 0.08, 0.04)^T,$$

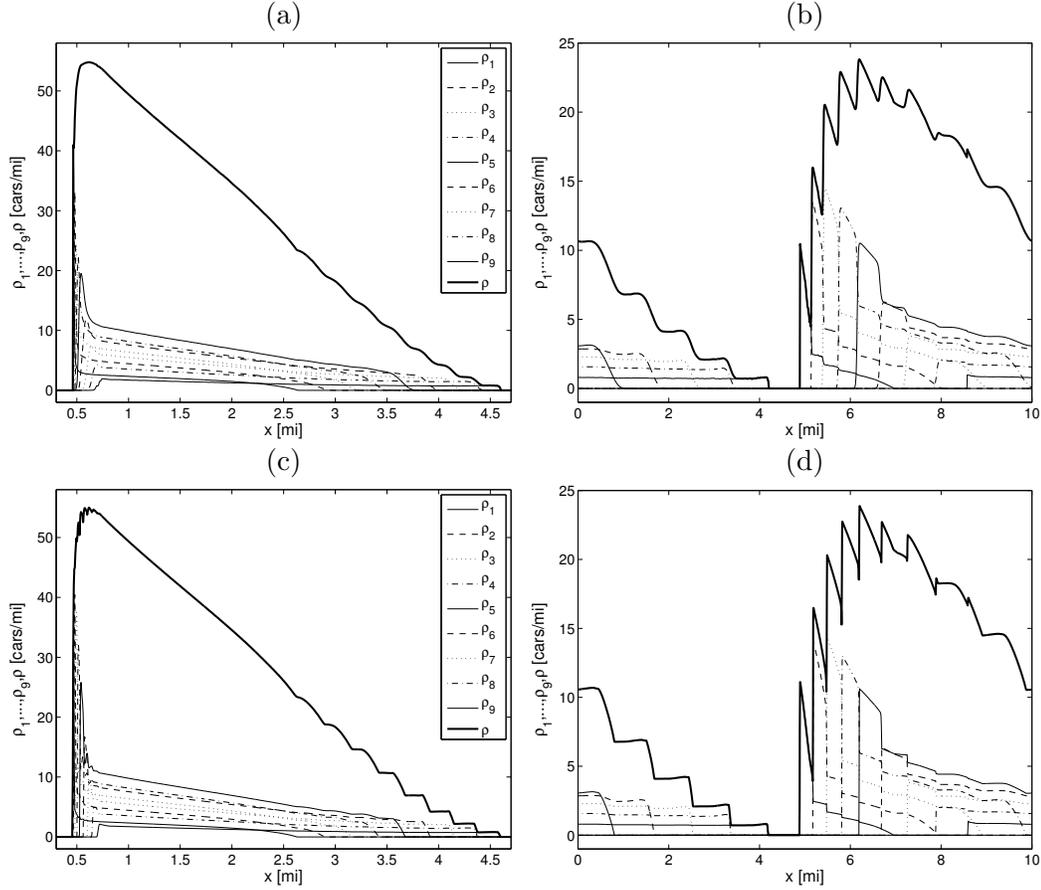


Figure 4.11: Example 4.4: schemes (a, b) L-RS, (c, d) L-NBee at (a, c)  $t = 0.03$  h, (b, d)  $t = 0.11$  h, and  $\Delta x = 1/200$ .

where the “shape function”  $p(x)$  is given by

$$p(x) = \begin{cases} 10x & \text{for } 0 < x \leq 0.1, & 1 & \text{for } 0.1 < x \leq 0.9, \\ -10(x-1) & \text{for } 0.9 < x \leq 1, & 0 & \text{otherwise.} \end{cases}$$

In Example 4.4 we set  $\rho_0 = 120$  cars/mi, which is well over the optimal density  $\rho^*$  of the model, and leads to a congested regime. In Example 4.5 we set  $\rho_0 = 40$  cars/mi, which leads to a non-congested regime.

For Example 4.4, in Figures 4.11 (a) and (b) we display the numerical solution obtained with the L-RS scheme and in Figures 4.11 (c) and (d) those obtained with the L-NBee scheme at simulated times  $t = 0.03$  h and  $t = 0.11$  h with  $\Delta x = 1/200$  mi. The traffic phenomenon is represented adequately by each scheme. Observe that L-NBee is more anti-diffusive than the L-RS scheme. This behaviour is maintained for long simulated times. Enlarged views of relevant parts of the numerical solutions of Figure 4.11 for some selected species are shown in Figures 4.12 and 4.13. We compare the numerical solution for each species with a reference solution obtained by a fifth-order WENO-SPEC scheme [50].

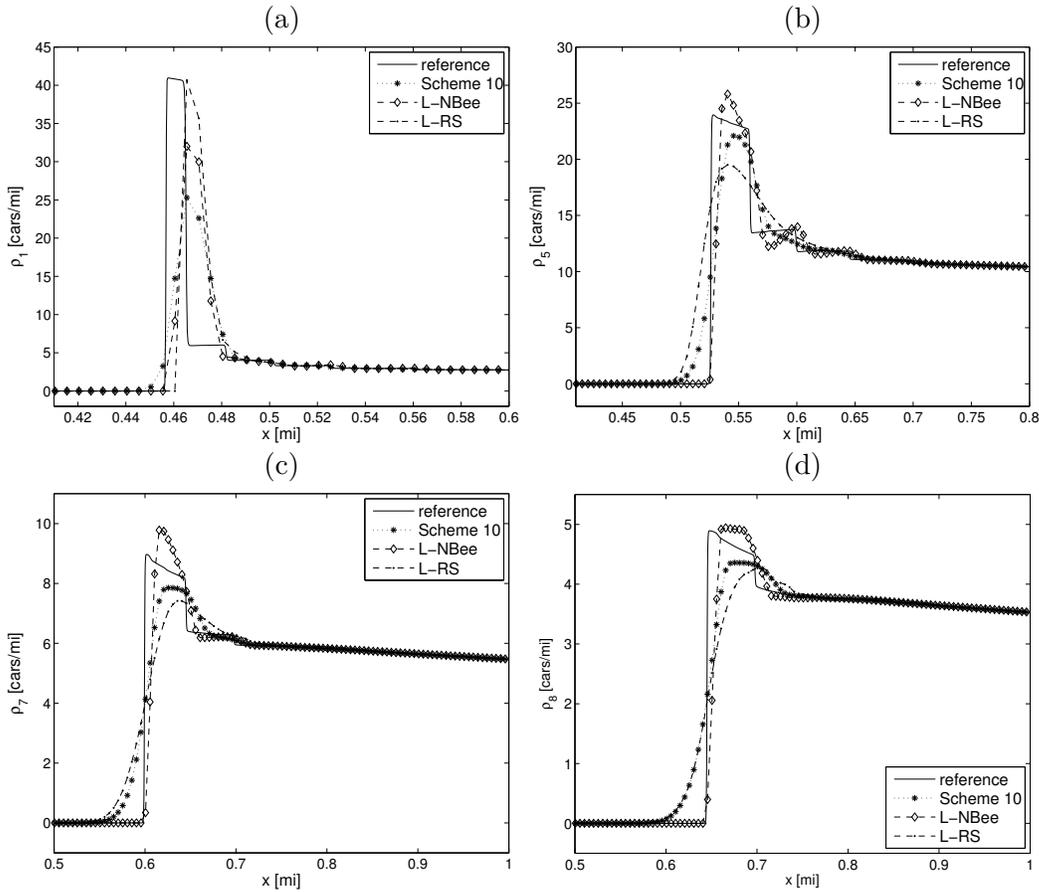


Figure 4.12: Example 4.4: enlarged views of relevant parts of the numerical solutions of Figure 4.11 for  $t = 0.03$  h and selected species, including numerical solutions obtained by Scheme 10.

The numerical tests indicate that for several species, the numerical solution obtained by the L-NBee scheme are anti-diffusive in each species and this behaviour is maintained during a long percentage of the simulated time, see Figures 4.12, 4.13 and 4.16 for Example 4.4 and Figures 4.17 and 4.18 for Example 5. However, the L-RS scheme is more diffusive than Scheme 10, and in Figure 4.13 we observe a delay in the approximation of the shock wave. This behaviour is observed for several individual species. As in another Examples, in Table 4.2 we observe that the percentage relative mass loss at simulate times decreasing with the refinement of the mesh.

In Figures 4.14 and 4.19 we display the efficiency of the numerical schemes in comparison with that of Schemes 4 and 10. We observe that the efficiency of the L-RS scheme is comparable with that of the first-order accurate Scheme 4, while the L-NBee scheme is even more efficient than the second-order accurate Scheme 10. In Figure 4.15 we display the total entropy as a function of  $t$ , we observe that the function (4.61) is non-increasing in the time.

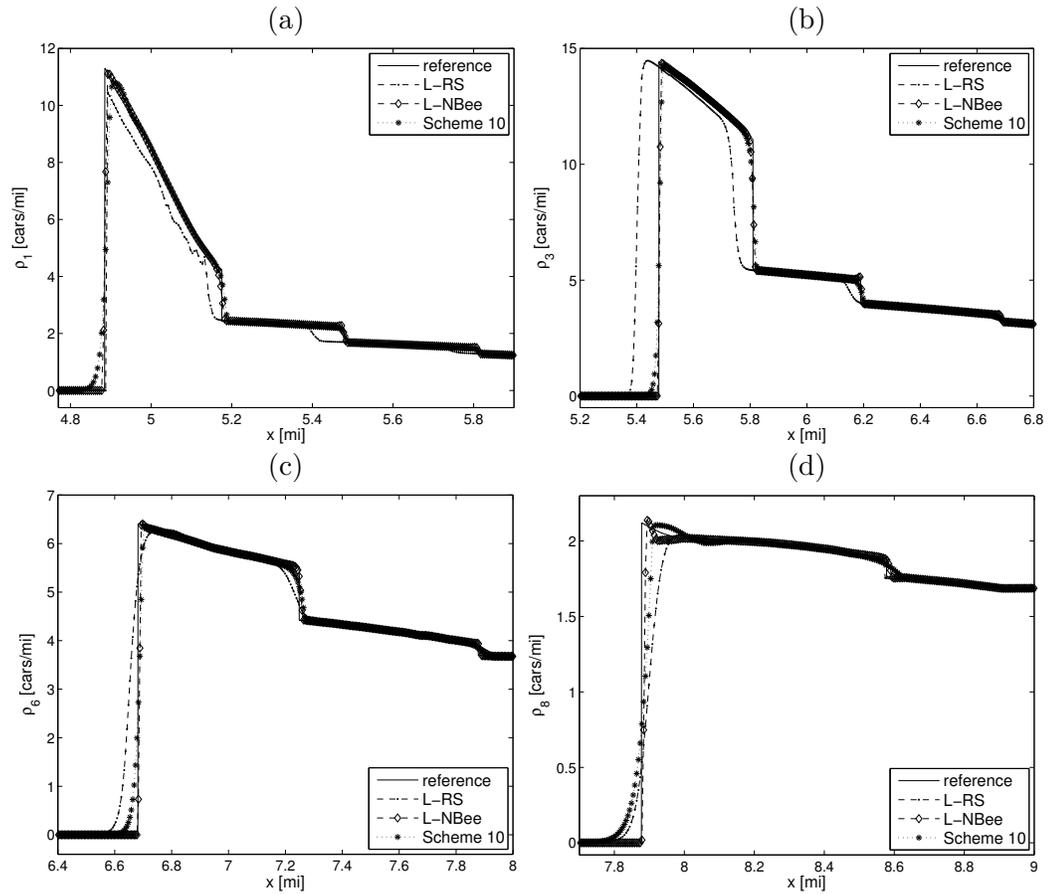


Figure 4.13: Example 4.4: enlarged views of relevant parts of the numerical solutions of Figure 4.11 for  $t = 0.11$  h and selected species, including numerical solutions obtained by Scheme 10.

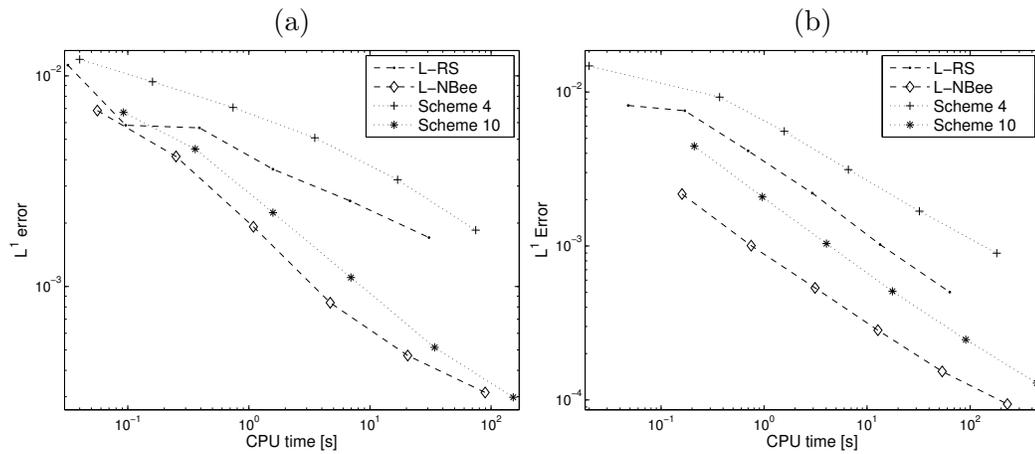


Figure 4.14: Example 4.4: approximate total  $L^1$  versus CPU time at simulated time (a)  $t = 0.03$  h, (b)  $t = 0.11$  h.

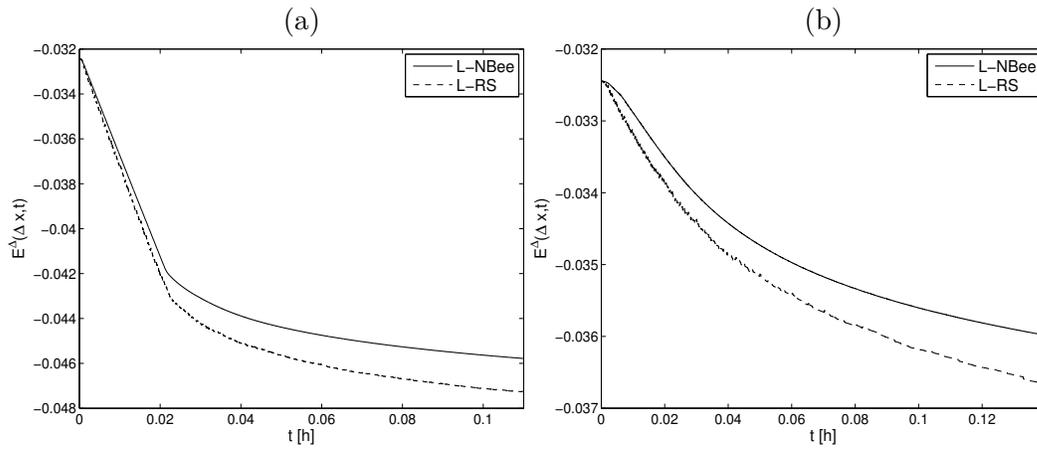


Figure 4.15: Numerical entropy with  $\Delta x = 1/200$ : (a) Example 4.4, (b) Example 4.5.

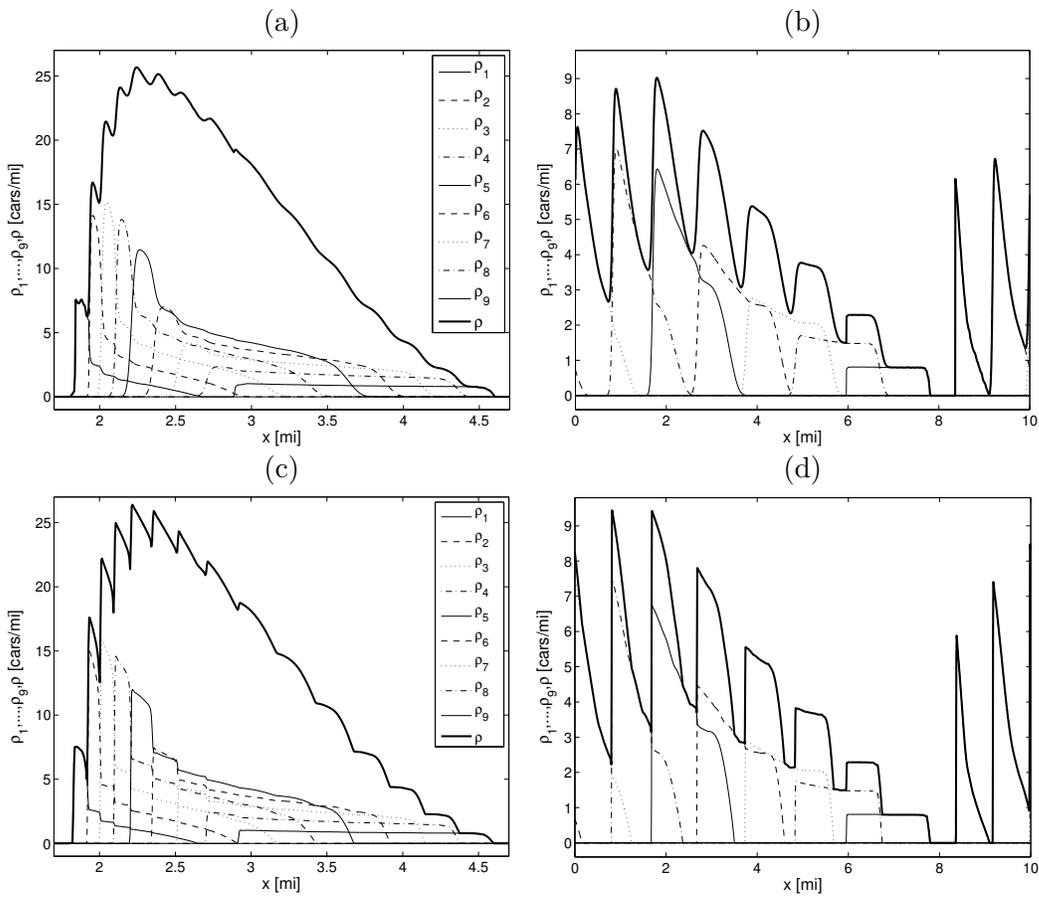


Figure 4.16: Example 4.5: schemes (a, b) L-RS, (c, d) L-NBee, (a, c)  $t = 0.03$  h, (b, d)  $t = 0.14$  h, and  $\Delta x = 1/200$ .

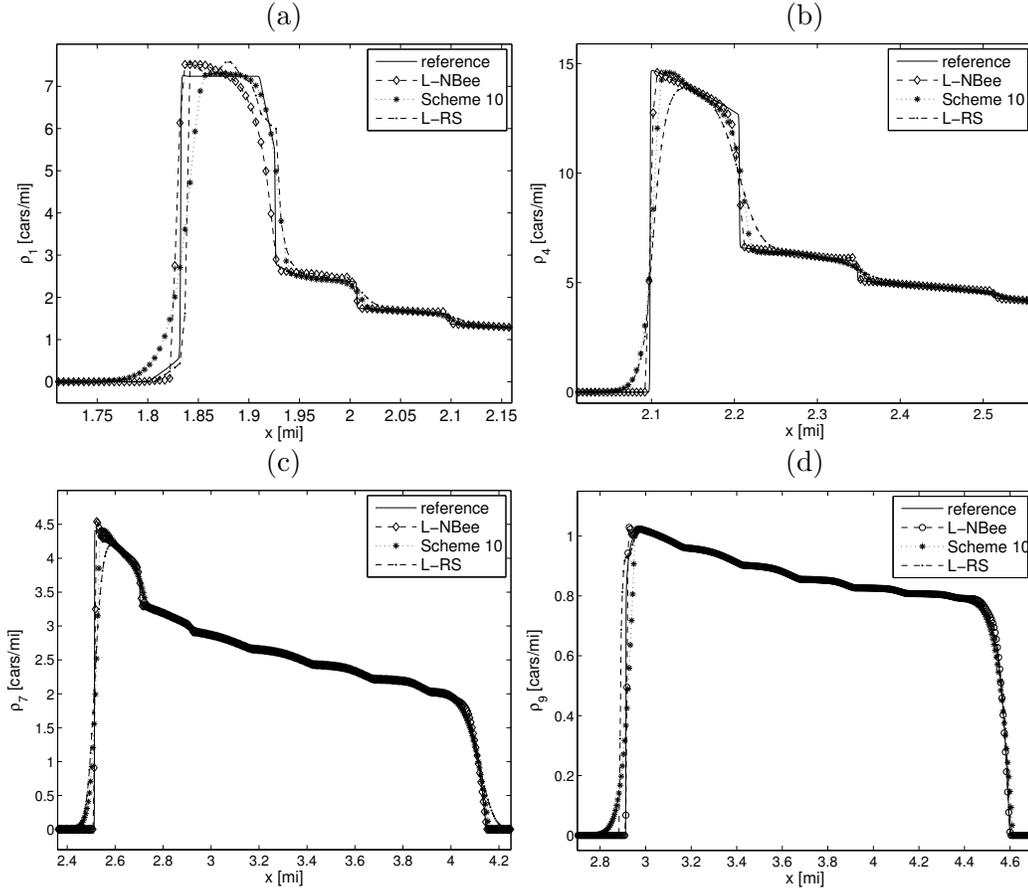


Figure 4.17: Example 4.5: enlarged views of relevant parts of the numerical solutions of Figure 4.16 for  $t = 0.03$  h and selected species, including numerical solutions obtained by Scheme 10.

$T$	$M$	L-NBee			L-RS			Scheme 4			Scheme10		
		error	cr	CPU [s]	error	cr	CPU [s]	error	cr	CPU [s]	error	cr	CPU [s]
0.03	100	178.9	—	0.05	567.2	—	0.02	1037.8	—	0.04	386.3	—	0.09
	200	89.4	0.99	0.2	329.6	0.78	0.08	714.3	0.53	0.1	194.8	0.98	0.3
	400	45.98	0.96	1.0	190.6	0.79	0.3	440.3	0.69	0.6	96.5	1.01	1.5
	800	20.7	1.14	4.1	92.3	1.04	1.3	243.3	0.85	2.9	45.6	1.08	6.9
	1600	8.9	1.21	18.5	47.3	0.96	5.5	126.1	0.94	13.5	20.2	1.17	33.9
	3200	4.4	1.01	80.7	28.3	0.74	23.8	65.3	0.94	60.5	10.1	1.00	150.4
0.14	100	115.4	—	0.2	499.8	—	0.07	931.9	—	0.1	340.2	—	0.2
	200	56.9	1.01	0.9	302.4	0.72	0.29	635.7	0.55	0.4	168.6	1.01	1.2
	400	25.7	1.14	4.0	175.0	0.78	1.15	406.6	0.64	1.9	81.1	1.05	5.0
	800	12.2	1.06	16.5	101.6	0.78	4.63	242.8	0.74	8.2	38.9	1.05	23.8
	1600	5.7	1.08	71.2	58.7	0.79	18.51	138.6	0.80	38.4	18.6	1.06	114.1
	3200	2.3	1.27	307.3	31.6	0.89	77.33	76.6	0.85	194.5	8.8	1.07	503.4

Table 4.6: Example 4.5: total approximate  $L^1$  errors  $e_{\text{tot}}(t)$  (“error”, to be multiplied by  $10^{-5}$ ), cr, and CPU times for two LR schemes and Schemes 4 and 10.

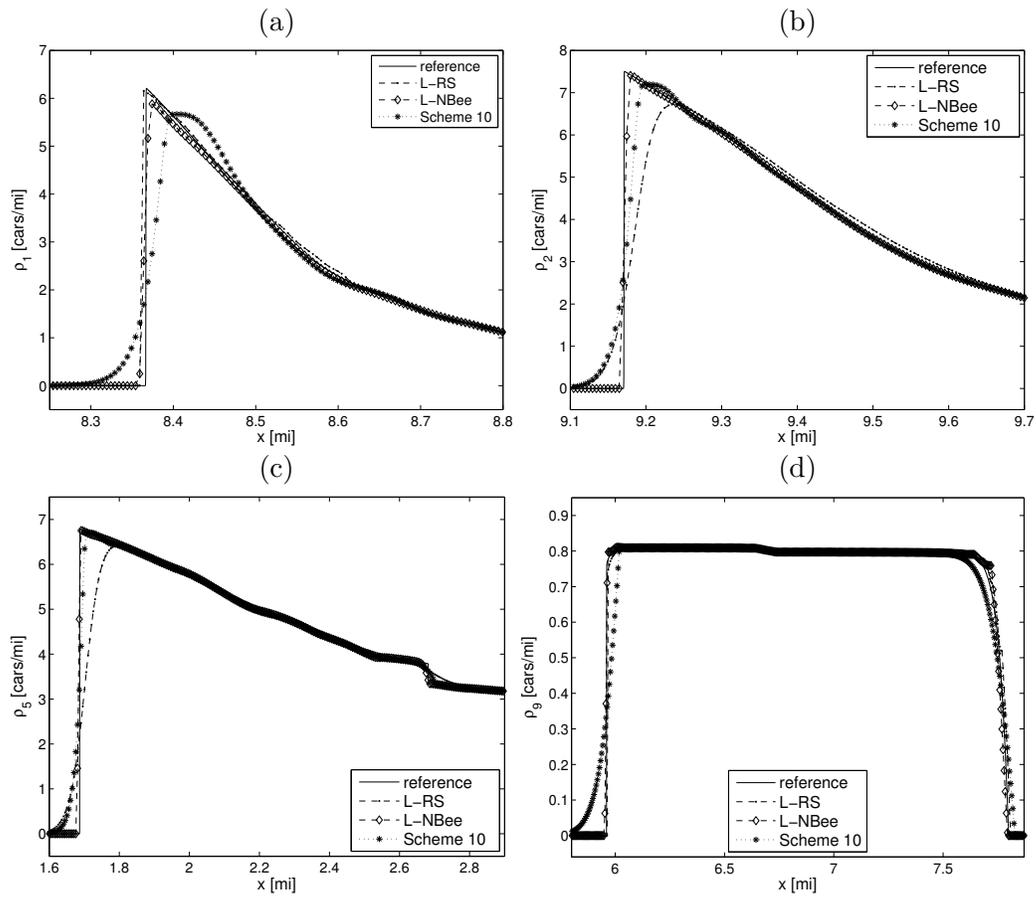


Figure 4.18: Example 4.5: enlarged views of relevant parts of the numerical solutions of Figure 4.16 for  $t = 0.14$  h and selected species, including numerical solutions obtained by Scheme 10.

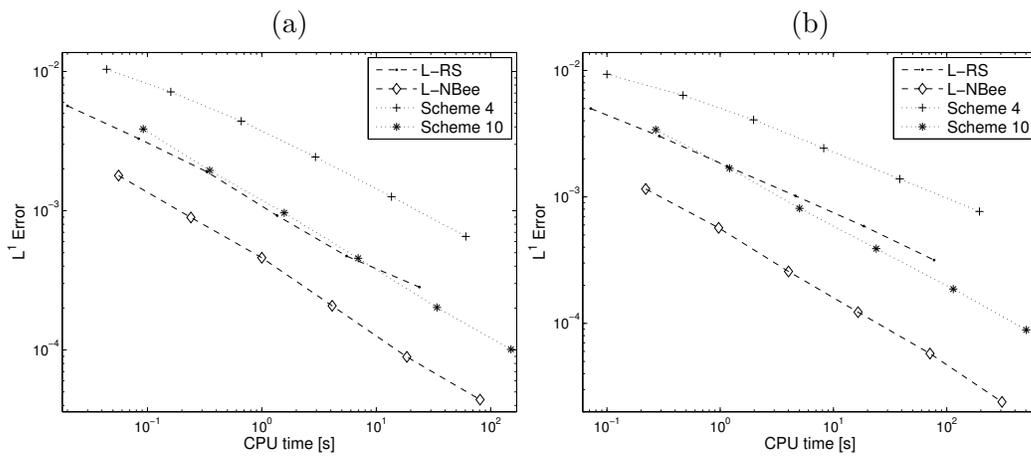


Figure 4.19: Example 4.5: approximate total  $L^1$  error versus CPU time at simulated time (a)  $t = 0.03$  h, (b)  $t = 0.14$  h.

## 4.7 Conclusions of Chapter 4

We have investigated in this paper a new class of numerical schemes for the challenging issue of approximating the solutions of the strongly coupled MCLWR models. These schemes are based on a Lagrangian-Projection decomposition for each car density, and the use of anti-diffusive techniques for solving the projection step. The proposed strategies turn out to be very easy and competitive with respect to existent schemes, especially for large values of  $N$  (the number of densities or equivalently the size of the system), where not characteristic decomposition of the eigenstructure are needed and are stable up to a CFL number 1.0. In the case  $N = 1$ , the strategy is supported by a partial numerical analysis since an  $L^\infty$  bound and a T.V.D. property are established. Proving the validity of an entropy inequality is still an open problem at this stage but numerical experiments show the convergence to entropy solution.

In this first investigation we focus on traffic flow models for which velocities  $v_i(\rho) = v_i^{\max} V(\rho)$  are assumed to be positive and such that  $v'(\rho) \leq 0$ . An interesting extension of this work could be envisaged to the polydisperse sedimentation models with single variable velocity. Extension to second or high order accuracy is a much more involved issue to be considered in the near future (even if the LR scheme are shown to be already competitive with second-order scheme “Scheme 10”).



# Discusión

En este capítulo presentamos una discusión de los principales resultados de esta tesis y una descripción del trabajo futuro a desarrollar.

## 4.8 Conclusiones

La tesis aporta diferentes esquemas numéricos para modelos de flujo cinemático multi-especies incluyendo aquellos modelos que contienen un término fuente difusivo, además se generaliza y se analiza a múltiples especies un modelo para problemas de tráfico vehicular en los que se tienen en cuenta tiempos de reacción y distancias mínimas para cada una de las especies.

La propuesta de nuevos esquemas numéricos para resolver modelos de flujo cinemático multi-especies se motiva a partir del hecho que actualmente los esquemas numéricos que generan mejor resolución en la solución de dichos problemas, esquemas shock-capturing de alto orden, requieren del uso de toda la información característica del Jacobiano del flujo, la cual en la mayoría de los modelos no se encuentra disponible debido a su complejidad, además los hace más costosos desde el punto de vista computacional. Por ello, con el fin de motivar el uso de esquemas numéricos requieran menor esfuerzo computacional y/o esquemas sencillos de implementar y que no requieran del uso de la información característica del modelo, en la tesis se han estudiado diferentes alternativas de esquemas numéricos para la resolución de este tipo de modelos.

A continuación se enumeran los resultados más importantes divididos en tres grandes tópicos.

### 1. Esquemas numéricos para modelos de flujo cinemático multi-especies.

Se aplicó un algoritmo de refinamiento de malla adaptativo (AMR) para ahorrar el esfuerzo computacional en memoria y tiempo de ejecución, en la simulación de modelo de sedimentación polidispersa a dos esquemas shock-capturing de alto orden. Las simulaciones para  $N = 4$  y  $N = 7$  mostraron que los esquemas numéricos con AMR requieren menor esfuerzo computacional que los esquemas sin refinamiento adaptativo. Además, los esquemas que utilizan toda la información característica son mucho más eficientes que los esquemas que no hacen uso de esta información.

Se propusieron esquemas numéricos de dos pasos para la resolución de problemas de tráfico vehicular multi-especies que combina en un primer paso la solución de ecuaciones en coordenadas Lagrangianas, y el segundo paso consistente en resolver una ecuación de transporte.

Para el segundo paso se consideraron dos diferentes estrategias, una basada en recientes y novedosos esquemas anti-difusivos los cuales son sencillos de implementar y con la propiedad T.V.D., tales esquemas se denominaron “L-AR”. La segunda estrategia se basa en el clásico método aleatorio de Glimm, tales esquemas se denominaron “L-RS”. Para el caso  $N = 1$  se demostró que los esquemas L-AR son conservativos, tienen la propiedad T.V.D. y satisfacen el principio del máximo, con lo cual se muestra que convergen a una solución débil de la ley de conservación. Ambas esquemas L-AR y L-RS se generalizaron para el caso  $N > 1$ . Los resultados numéricos indicaron que estos esquemas resultan competitivos con respecto a otras variedad de esquemas, aparte de que no requieren del uso de la información característica del flujo.

## 2. Esquemas implícitos-explicitos para modelos de flujo cinemático multi-especies con corrección difusiva.

Se propuso resolver un problema de flujo cinemático multi-especie con termino fuente difusivo y de carácter (posiblemente) fuertemente degenerado mediante una discretización temporal explícita para el término convectivo y una discretización temporal implícita para el término difusivo el avance temporal se realiza mediante un método Runge-Kutta. Si bien la formulación IMEX-Runge-Kutta requiere de resolver problemas altamente no lineal, tienen la principal ventaja que la condición de estabilidad CFL depende exclusivamente de la estabilidad requerida por el término convectivo, lo cual lo hace mucho más eficiente respecto a usar un método totalmente explícito.

Se propuso una estrategia para resolver el problema altamente no lineal que resulta después de discretizar de manera implícita el término difusivo la cual está basada en la aplicación del método de Newton-Raphson junto con una regularización del coeficiente discontinuo de la matriz de difusión, de manera que el carácter fuertemente degenerado del término difusivo se vea afectado.

La discretización espacial se hace mediante un esquema shock-capturing de alto orden y los resultados numéricos para diferentes modelos muestran que comparando con respecto a esquemas explícitos, los esquemas IMEX-Runge-Kutta tienen un rendimiento de hasta un orden de magnitud en niveles de discretización moderada.

## 3. Modelo de tráfico vehicular multi-especies con corrección difusiva.

Se propuso una generalización del modelo LWR de tráfico vehicular difusivamente corregido que considera un tiempo de reacción y distancia mínima de anticipación a un modelo multi-clase LWR difusivamente corregido en el que individuos de cada clase se caracterizan por tener una velocidad máxima preferencial y un tiempo de reacción en el que ajustan su velocidad. Se analizó la estabilidad del modelo desde el punto de vista de la información característica solo de la matriz de difusión, lo cual requiere que tenga valores propios positivos. Esto conduce a una relación entre las velocidades máximas y los tiempos de reacción que permite al sistema sea estable.

Numéricamente se analizó para  $N = 2$  la estabilidad del sistema completo que incluye los

términos convectivo y difusivo, esto permitió identificar configuraciones (velocidad máxima y tiempos de reacción) en las cuales el sistema completo es inestable aun cuando la matriz difusión tenga valores propios positivos. Esto genera que pequeñas perturbaciones en la densidad total del tráfico generen soluciones oscilatorias, algunas de ellas que permanecen controladas en relación a amplitud y frecuencia. Estas están relacionadas a tiempos de reacción grandes o distancias de anticipación muy pequeñas.

## 4.9 Trabajo Futuro

1. Aplicar un algoritmo de refinamiento de malla adaptativo (AMR) a los esquemas IMEX-Runge-Kutta para resolver problemas de flujo cinemático multi-especies con término fuente difusivo.
2. Se demostró para  $N = 1$  que los esquemas L-AR convergen a una solución débil, adicionalmente los resultados numéricos muestran que tales esquemas convergen a la solución de entropía. Con lo cual, demostrar que los esquemas LR convergen a la única solución de entropía es aun un problema abierto.
3. Extensión de los esquemas LR para problemas de sedimentación polidispersa en la cual la velocidad puede variar en signo. Esto puede ser logrado aplicando esquemas anti-difusivos para resolver la ecuación de transporte con velocidad variable en signo.
4. Obtener versiones de segundo o alto orden de exactitud para los esquemas LR. Aquí también se considera una extensión a problemas cinemático multi-especies donde la función de flujo varia espacialmente.
5. Para el caso del modelo de tráfico vehicular multi-especies con corrección difusiva, el análisis de estabilidad aplicado al sistema completo se realizó desde el punto de vista numérico. Sin embargo, se requiere de un análisis teórico que permita determinar relaciones entre las velocidades máximas y los tiempos de reacción con las inestabilidades del modelo.



# Bibliography

- [1] A. Abeynaïke, A.J. Sederman, Y. Khan, M.L. Johns, J.F. Davidson, and M.R. Mackley. The experimental measurement and modelling of sedimentation and creaming for glycerol/biodiesel droplet dispersions. *Chem. Eng. Sci.*, **79** (2012), pp. 125–137.
- [2] L.A. Amy, P.J. Talling, V.O. Edmonds, E.J. Sumner, and A. Lesueur. An experimental investigation of sand-mud suspension settling behaviour: implications for bimodal mud contents of submarine flow deposits. *Sedimentology*, **53** (2006), pp. 1411–1434.
- [3] J. Anderson. A secular equation for the eigenvalues of a diagonal matrix perturbation. *Lin. Alg. Appl.*, **246** (1996), pp. 49–70.
- [4] U. Ascher, S. Ruuth, and J. Spiteri. Implicit-explicit Runge-Kutta methods for time dependent partial differential equations. *Appl. Numer. Math.*, **25** (1997), pp. 151–167.
- [5] M. Bachmann, P. Helluy, H. Mathis, and S. Müller. Random sampling remap for compressible two-phase flows. Preprint (2010), see <http://hal.archives-ouvertes.fr/hal-00546919/fr/>
- [6] A. Baeza, A. Martínez-Gavara, and P. Mulet. Adaptation based on interpolation errors for high order mesh refinement methods applied to conservation laws. *Appl. Numer. Math.*, **62** (2012), pp. 278–296.
- [7] G.K. Batchelor and R.W. Janse van Rensburg. Structure formation in bidisperse sedimentation. *J. Fluid Mech.*, **166** (1986), pp. 379–407.
- [8] S. Benzoni-Gavage and R.M. Colombo. An  $n$ -populations model for traffic flow. *Eur. J. Appl. Math.*, **14** (2003), pp. 587–612.
- [9] S. Benzoni-Gavage, R.M. Colombo, and P. Gwiazda. Measure valued solutions to conservation laws motivated by traffic modelling. *Proc. Royal Soc. A*, **462** (2006), pp. 1791–1803.
- [10] M.J. Berger and P. Colella. Local adaptive mesh refinement for shock hydrodynamics. *J. Comput. Phys.*, **82** (1989), pp. 64–84.
- [11] M.J. Berger and J. Olinger. Adaptive mesh refinement for hyperbolic partial differential equations, *J. Comput. Phys.*, **53** (1984), pp. 484–512.

- [12] S. Berres, R. Bürger, K.H. Karlsen, and E.M. Tory. Strongly degenerate parabolic-hyperbolic systems modeling polydisperse sedimentation with compression. *SIAM J. Appl. Math.*, **64** (2003), pp. 41–80.
- [13] S. Berres, R. Bürger, and A. Kozakevicius. Numerical approximation of oscillatory solutions of hyperbolic-elliptic systems of conservation laws by multiresolution schemes. *Adv. Appl. Math. Mech.*, **1** (2009), pp. 581–614.
- [14] O. Bokanowski and H. Zidani. Anti-dissipative schemes for advection and application to Hamilton-Jacobi-Bellman equations. *J. Sci. Comput.*, **30** (2007), pp. 1–33.
- [15] F. Bouchut. An anti-diffusive entropy scheme for monotone scalar conservation laws. *J. Sci. Comput.*, **21** (2004), pp. 1–30.
- [16] S. Boscarino. Error analysis of IMEX Runge-Kutta methods derived from differential-algebraic systems. *SIAM J. Numer. Anal.*, **45** (2007), pp. 1600–1621.
- [17] S. Boscarino, P.G. LeFloch, and G. Russo. *High order asymptotic-preserving methods for fully nonlinear relaxation problems*. [arxiv.org/pdf/1210.4761](https://arxiv.org/pdf/1210.4761).
- [18] S. Boscarino, L. Pareschi, and G. Russo. Implicit-explicit Runge-Kutta schemes for hyperbolic systems and kinetic equations in the diffusion limit. *SIAM J. Sci. Comput.*, **35** (2013), pp. A22–A51.
- [19] S. Boscarino and G. Russo. Flux-Explicit IMEX RungeKutta Schemes for Hyperbolic to Parabolic Relaxation Problems. *SIAM J. Numer. Anal.*, **51** (2013), pp. 163–190.
- [20] S. Boscarino and G. Russo. On a class of uniformly accurate IMEX Runge-Kutta schemes and applications to hyperbolic systems with relaxation. *SIAM J. Sci. Comput.*, **31** (2009), pp. 1926–1945.
- [21] A. Bressan. *Hyperbolic Systems of Conservation Laws*. Oxford University Press, Oxford, UK, (2000).
- [22] R. Bürger, R. Donat, P. Mulet, and C.A. Vega. Hyperbolicity analysis of polydisperse sedimentation models via a secular equation for the flux Jacobian. *SIAM J. Appl. Math.*, **70** (2010), pp. 2186–2213.
- [23] R. Bürger, R. Donat, P. Mulet, and C.A. Vega. On the implementation of WENO schemes for a class of polydisperse sedimentation models. *J. Comput. Phys.*, **230** (2011), pp. 2322–2344.
- [24] R. Bürger, A. García, K.H. Karlsen, and J.D. Towers. A family of numerical schemes for kinematic flows with discontinuous flux. *J. Eng. Math.*, **60** (2008), pp. 387–425.
- [25] R. Bürger, and K.H. Karlsen. On a diffusively corrected kinematic-wave traffic flow model with changing road surface conditions. *Math. Models Methods Appl. Sci.*, **13** (2003), pp. 1767–1799.

- [26] R. Bürger, K.H. Karlsen, E.M. Tory, and W.L. Wendland. Model equations and instability regions for the sedimentation of polydisperse suspensions of spheres. *ZAMM Z. Angew. Math. Mech.*, **82** (2002), pp. 699–722.
- [27] R. Bürger, K.H. Karlsen, and J.D. Towers. On some difference schemes and entropy conditions for a class of multi-species kinematic flow models with discontinuous flux. *Netw. Heterog. Media*, **5** (2010), pp. 461–485.
- [28] R. Bürger and A. Kozakevicius. Adaptive multiresolution WENO schemes for multi-species kinematic flow models. *J. Comput. Phys.*, **224** (2007), pp. 1190–1222.
- [29] R. Bürger, C. Chalons, and L.M. Villada. *Anti-diffusive and random-sampling Lagrangian-remap schemes for the multiclass Lighthill-Whitham-Richards traffic model*. Preprint 2013-14, Centro de Investigación en Ingeniería Matemática, Universidad de Concepción; (submitted).
- [30] R. Bürger, P. Mulet, and L.M. Villada. A diffusively corrected multiclass Lighthill-Whitham-Richards traffic model with anticipation lengths and reaction times. *Adv. Appl. Math. Mech.*, to appear.
- [31] R. Bürger, P. Mulet, and L.M. Villada. Regularized nonlinear solvers for IMEX methods applied to diffusively corrected multi-species kinematic flow models. *SIAM J. Sci. Comput.*, **35** (2013), pp. B751–B777.
- [32] R. Bürger, P. Mulet, and L.M. Villada. Spectral WENO schemes with Adaptive Mesh Refinement for models of polydisperse sedimentation. *ZAMM Z. Angew. Math. Mech.*, **93** (2013), pp. 373–386.
- [33] R. Bürger, R. Ruiz-Baier, K. Schneider, and H. Torres. A multiresolution method for the simulation of sedimentation in inclined channels. *Int. J. Numer. Anal. Model.*, **9** (2012), pp. 479–504.
- [34] C. Chalons. Numerical approximation of a macroscopic model of pedestrian flows. *SIAM J. Sci. Comput.*, **29** (2007), pp. 539–555.
- [35] C. Chalons. Transport-equilibrium schemes for computing nonclassical shocks. Scalar conservation laws. *Numer. Meth. Partial Diff. Eqns.*, **24** (2008), pp. 1127–1147.
- [36] C. Chalons and F. Coquel. Computing material fronts with a Lagrange-projection approach. In: T. Li and S. Jiang (eds.), *Hyperbolic Problems: Theory, Numerics and Applications*. Series in Contemporary Applied Mathematics CAM 17/18, vol. 1. Higher Education Press, Beijing, China, 346–356, 2012.
- [37] C. Chalons and P. Goatin. Transport-equilibrium schemes for computing contact discontinuities in traffic flow modeling. *Commun. Math. Sci.*, **3** (2007), pp. 533–551.
- [38] C. Chalons and P. Goatin. Godunov scheme and sampling technique for computing phase transitions in traffic flow modeling. *Interfaces Free Boundaries*, **10** (2008), pp. 197–221.

- [39] T. F. Chan, G. H. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comput.*, **20** (1999), pp. 1964–1977.
- [40] G. Chiavassa and R. Donat. Point-value multiscale algorithms for 2D compressible flows. *SIAM J. Sci. Comput.*, **20** (2001), pp. 805–823.
- [41] A. Cohen, S. Kaber, S. Müller, and M. Postel. Fully adaptive finite volume schemes for conservation laws. *Math. Comp.*, **72** (2003), pp. 183–225.
- [42] M. Crouzeix. Une méthode multipas implicite-explicite pour l’approximation des équations d’évolution paraboliques. *Numer. Math.*, **35** (1980), pp. 257–276.
- [43] C. Daganzo. Requiem for second-order fluid approximations of traffic flow. *Transp. Res. B*, **29** (1995), pp. 277–286.
- [44] R.H. Davis, E. Herbolzheimer, and A. Acrivos. The sedimentation of polydisperse suspensions in vessels having inclined walls. *Int. J. Multiphase Flow* **8** (1982), pp. 571–585.
- [45] J.E. Dennis Jr. and R.B. Schnabel. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. SIAM, 1996.
- [46] B. Després and F. Lagoutière. Contact discontinuity capturing schemes for linear advection and compressible gas dynamics. *J. Sci. Comput.*, **16** (2001), pp. 479–524.
- [47] B. Després, E. Labourasse, F. Lagoutière, and I. Marmajou. An antidissipative transport scheme on unstructured meshes for multicomponent flows. *Int. J. Finite Volumes* **7**, no. 1, 2010.
- [48] A.C. Dick. Speed/flow relationships within an urban area. *Traffic Engrg. Control*, **8** (1966), pp. 393–396.
- [49] R. Donat and I. Higuera. On stability issues for IMEX schemes applied to 1D scalar hyperbolic equations with stiff reaction terms. *Math. Comp.*, **80** (2011), pp. 2097–2126.
- [50] R. Donat and P. Mulet. Characteristic-based schemes for multi-class Lighthill-Whitham-Richards traffic models. *J. Sci. Comput.*, **37** (2008), pp. 233–250 .
- [51] R. Donat and P. Mulet. A secular equation for the Jacobian matrix of certain multi-species kinematic flow models. *Numer. Methods Partial Differential Equations*, **26** (2010), pp. 159–175.
- [52] R.M. Dorrell, A.J. Hogg, E.J. Sumner, and P.J. Talling. The structure of the deposit produced by sedimentation of polydisperse suspensions. *J. Geophys. Res.* **116** (2011), paper F01024 .
- [53] J.S. Drake, J.L. Schofer, and A.D. May. A statistical analysis of speed density hypothesis. *Highway Research Record*, **154** (1967), pp. 53–87.

- [54] T. Frising, C. Noïk, and C. Dalmazzone. The liquid/liquid sedimentation process: from droplet coalescence to technologically enhanced water/oil emulsion gravity separators: a review. *J. Disp. Sci. Technol.*, **27** (2006), pp. 1035–1057.
- [55] J. Glimm. Solutions in the large for nonlinear hyperbolic systems of equations. *Comm. Pure Appl. Math.*, **18** (1965), pp. 697–715.
- [56] E. Godlewski and P.-A. Raviart. Numerical approximation of hyperbolic systems of conservation laws. Springer Verlag, New York, 1996.
- [57] S. Gottlieb and C.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Math. Comp.*, **67** (1998), pp. 73–85.
- [58] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Rev.*, **43** (2001), pp. 89–112.
- [59] H. Greenberg. An analysis of traffic flow. *Oper. Res.*, **7** (1959), pp. 79–85.
- [60] A. Harten. Multiresolution algorithms for the numerical solution of hyperbolic conservation laws, *Comm. Pure Appl. Math.* **48**, 1305–1342 (1995).
- [61] A. Harten, P.D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.*, **25** (1983), pp. 35–61.
- [62] P. Helluy and J. Jung. A coupled well-balanced and random sampling scheme for computing bubble oscillations. *ESAIM Proc.*, **35** (2012), pp. 245–250.
- [63] M. Herty, C. Kirchner, and S. Moutari. Multi-class traffic models on road networks. *Commun. Math. Sci.*, **4** (2006), pp. 591–608.
- [64] K. Höfler and S. Schwarzer, *The structure of bidisperse suspensions at low Reynolds numbers*, in A.-M. Sändig, W. Schiehlen, and W.L. Wendland (eds.), *Multifield Problems: State of the Art*, Springer Verlag, Berlin (2000), pp. 42–49.
- [65] C. Hu. Numerical Methods for Hyperbolic Equations on Unstructured Meshes. PhD Thesis, Brown University, 1999.
- [66] G.S. Jiang and C.-W. Shu. Efficient implementation of weighted ENO schemes. *J. Comput. Phys.*, **126** (1996), pp. 202–228.
- [67] S. Jaouen and F. Lagoutière. Numerical transport of an arbitrary number of components. *Comput. Meth. Appl. Mech. Engrg.*, **196** (2007), pp. 3127–3140.
- [68] C.A. Kennedy and M.H. Carpenter. Additive Runge-Kutta schemes for convection-diffusion-reaction equations. *Appl. Numer. Math.*, **44** (2003), pp. 139–181.

- [69] S. Kokh and F. Lagoutière. An anti-diffusive numerical scheme for the simulation of interfaces between compressible fluids by means of a five-equation model. *J. Comput. Phys.*, **229** (2010), pp. 2773–2809.
- [70] A. Kurganov and A. Polizzi. Non-oscillatory central schemes for traffic flow models with Arrhenius look-ahead dynamics. *Netw. Heterog. Media*, **4** (2009), pp. 431–451.
- [71] A. Kurganov and E. Tadmor. New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations. *J. Comput. Phys.*, **160** (2000), pp. 241–282.
- [72] F. Lagoutière. Stability of reconstruction schemes for hyperbolic PDEs. *Commun. Math. Sci.*, **6** (2008), pp. 57–70.
- [73] P.D. Lax and B. Wendroff. Systems of conservation laws. *Comm. Pure Appl. Math.*, **13** (1960), 217–237.
- [74] J. Liandrat. The wavelet transform; some applications to fluid dynamics and turbulence. *Eur. J. Mech. B/Fluids.*, **9** (1990), 1–19.
- [75] M.J. Lighthill and G.B. Whitham. On kinematic waves: II. A theory of traffic flow on long crowded roads. *Proc. Royal Soc. A*, **229** (1955), pp. 317–345.
- [76] X.-D. Liu, S. Osher, and T. Chan. Weighted essentially non-oscillatory schemes. *J. Comput. Phys.*, **115** (1994), pp. 200–212.
- [77] M.J. Lockett and K.S. Bassoon. Sedimentation of binary particle mixtures. *Powder Technol.*, **24** (1979), pp. 1–7.
- [78] S. Logghe and L.H. Immers. Multi-class kinematic wave theory of traffic flow. *Transp. Res. B*, **42** (2008), pp. 523–541.
- [79] J.H. Masliyah. Hindered settling in a multiple-species particle system. *Chem. Engrg. Sci.*, **34** (1979), pp. 1166–1168.
- [80] M. Motamed, C.B. Macdonald, and S.J. Ruuth. On the linear stability of the fifth-order WENO discretization. *J. Sci. Comput.*, **47** (2011), pp. 127–149.
- [81] P. Nelson. Synchronized traffic flow from a modified Lighthill-Whitman model. *Phys. Rev. E*, **61** (2000), pp. R6062–R6055.
- [82] P. Nelson. Traveling-wave solution of the diffusively corrected kinematic-wave model. *Math. Comp. Modelling*, **35** (2002), pp. 561–579.
- [83] P. Nelson and A. Sopasakis. The Chapman-Enskog expansion: a novel approach to hierarchical extension of Lighthill-Whitham models. In: A. Ceder (ed.), *Transportation and Traffic Theory: Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, Israel, 20–23 July 1999. Elsevier, Amsterdam 1999, pp. 51–79.

- [84] D. Ngoduy. Multiclass first-order modelling of traffic networks using discontinuous flow-density relationships. *Transportmetrica*, **6** (2010), pp. 121–141.
- [85] D. Ngoduy. Multiclass first-order traffic model using stochastic fundamental diagrams. *Transportmetrica*, **7** (2011), pp. 111–125.
- [86] D. Ngoduy. Effect of driver behaviours on the formation and dissipation of traffic flow instabilities. *Nonlin. Dynamics*, **69** (2012), pp. 969–975.
- [87] D. Ngoduy and C. Tampere. Macroscopic effect of reaction time on traffic flow characteristics. *Physica Scripta*, **80** (2009), paper 025802.
- [88] J.M. Ortega and W.C. Rheinboldt. Iterative Solution of Nonlinear Equations in Several Variables. *Classics in Applied Mathematics*. SIAM, 1987.
- [89] L. Pareschi, G. Russo. Implicit-Explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation. *J. Sci. Comp.*, **25** (2005), pp. 129–155.
- [90] I. Prigogine and R. Herman. Kinetic Theory of Vehicular Traffic. *American Elsevier*, New York, 1971.
- [91] J.J. Quirk. An Adaptive Grid Algorithm for Computational Shock Hydrodynamics. PhD Thesis, Cranfield Institute of Technology, 1991.
- [92] P.I. Richards. Shock waves on the highway. *Oper. Res.*, **4** (1956), pp. 42–51.
- [93] J.F. Richardson and W.N. Zaki. Sedimentation and fluidization: Part I, *Trans. Instn. Chem. Engrs. (London)*, **32** (1954), pp. 35–53.
- [94] F. Rosso and G. Sona. Gravity-driven separation of oil-water dispersions. *Adv. Math. Sci. Appl.*, **11** (2001), pp. 127–151.
- [95] E. Rouvre and G. Gagneux. Solution forte entropique de lois scalaires hyperboliques-paraboliques dégénérées. *C. R. Acad. Sci. Paris Sér. I*, **329** (1999), pp. 599–602.
- [96] W. Schneider, G. Anestis, and U. Schaffinger. Sediment composition due to settling of particles of different sizes. *Int. J. Multiphase Flow*, **11** (1985), pp. 419–423.
- [97] M. Shearer, J.M.N.T. Gray, and A.R. Thornton. Stable solutions of a scalar conservation law in dense granular avalanches. *Eur. J. Appl. Math.*, **9** (2008), pp. 61–86.
- [98] C.-W. Shu. *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*. In: B. Cockburn, C. Johnson, C.-W. Shu, and E. Tadmor, *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations* (A. Quarteroni, Ed.), Lecture Notes in Mathematics vol. 1697, Springer-Verlag, Berlin (1998), pp. 325–432.
- [99] C.-W. Shu. High order weighted essentially nonoscillatory schemes for convection dominated problems. *SIAM Rev.*, **51** (2009), pp. 82–126.

- [100] C.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock capturing schemes. *J. Comput. Phys.*, **77** (1988), pp. 439–471.
- [101] F. Siebel and W. Mauser. On the fundamental diagram of traffic flow. *SIAM J. Appl. Math.*, **66** (2006), pp. 1150–1162.
- [102] J. Smoller. *Shock Waves and Reaction-Diffusion Equations*. Springer-Verlag, New York, 1983.
- [103] P. Snabre, B. Pouligny, C. Metayer, and F. Nadal. Size segregation and particle velocity fluctuations in settling concentrated suspensions. *Rheol.*, **48** (2009), pp. 855–870.
- [104] E. Sonnendrücker, J. Roche, P. Bertrand, and A. Ghizzo. The semi-Lagrangian method for the numerical resolution of the Vlasov equation. *J. Comput. Phys.*, **149** (1999), pp. 201–220.
- [105] A. Sopasakis and M.A. Katsoulakis. Stochastic modeling and simulation of traffic flow: asymmetric single exclusion process with Arrhenius look-ahead dynamics. *SIAM J. Appl. Math.*, **66** (2006), pp. 921–944.
- [106] K. Stamatakis and C. Tien. Batch sedimentation calculations—the effect of compressible sediment. *Powder Technol.*, **72** (1992), pp. 227–240.
- [107] H. Tang and T. Tang. Adaptive mesh methods for one- and two-dimensional hyperbolic conservation laws. *SIAM J. Numer. Anal.*, **41** (2004), pp. 487–515.
- [108] A.R. Thornton and J.M.N.T. Gray. Breaking size segregation waves and particle recirculation in granular avalanches. *J. Fluid Mech.*, **596** (2008), 261–284.
- [109] M. Treiber and A. Kesting. *Verkehrsdynamik und -simulation*. Springer-Verlag, Berlin, 2010.
- [110] M. Treiber, A. Kesting, and D. Helbing. Influence of reaction times and anticipation on stability of vehicular traffic flow. *Transp. Res. Record*, **1999** (2007), pp. 23–29.
- [111] E.F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Third Edition, Springer-Verlag, Berlin, 2009.
- [112] B. van Leer. Towards the ultimate conservative difference scheme II. Monotonicity and conservation combined in a second order scheme. *J. Comput. Phys.*, **14** (1974), pp. 361–370.
- [113] G.C.K. Wong and S.C. Wong. A multi-class traffic flow model—an extension of LWR model with heterogeneous drivers. *Transp. Res. A*, **36** (2002), pp. 827–841.
- [114] Z. Xu and C. W. Shu. Anti-diffusive flux corrections for high order finite difference WENO schemes. *J. Comput. Phys.*, **205** (2005), pp. 458–485.
- [115] B. Xue and Y. Sun. Modeling of sedimentation of polydisperse spherical beads with a broad size distribution. *Chem. Engrg. Sci.*, **58** (2003), pp. 1531–1543.

- 
- [116] A. Zeidan, S. Rohani, A. Bassi, and P. Whiting. Review and comparison of solids settling velocity models. *Rev. Chem. Eng.*, **19** (2003), pp. 473–530.
- [117] M. Zhang, C.-W. Shu, G.C.K. Wong, and S.C. Wong. A weighted essentially non-oscillatory numerical scheme for a multi-class Lighthill-Whitham-Richards traffic flow model. *J. Comput. Phys.*, **191** (2003), pp. 639–659.
- [118] P. Zhang, R.-X. Liu, S.C. Wong, and S.Q. Dai. Hyperbolicity and kinematic waves of a class of multi-population partial differential equations. *Eur. J. Appl. Math.*, **17** (2006), pp. 171–200.
- [119] P. Zhang, S.C. Wong, and S.-Q. Dai. A note on the weighted essentially non-oscillatory numerical scheme for a multi-class Lighthill-Whitham-Richards traffic flow model. *Commun. Numer. Meth. Engrg.*, **25** (2009), pp. 1120–1126.
- [120] P. Zhang, S.C. Wong, and C.-W. Shu. A weighted essentially non-oscillatory numerical scheme for a multi-class traffic flow model on an inhomogeneous highway. *J. Comput. Phys.*, **212** (2006), pp. 739–756.
- [121] P. Zhang, S.C. Wong, and Z. Xu. A hybrid scheme for solving a multi-class traffic flow model with complex wave breaking. *Comput. Meth. Appl. Mech. Engrg.*, **197** (2008), pp. 3816–3827.
- [122] X. Zhong. Additive semi-implicit Runge-Kutta methods for computing high-speed nonequilibrium reactive flows. *J. Comput. Phys.*, **128** (1996), pp. 19–31.