SAPIENZA
UNIVERSITÀ DI ROMA

DEPARTMENT OF BASIC AND APPLIED SCIENCES FOR ENGINEERING

PHD THESIS IN MATHEMATICAL MODELS FOR ENGINEERING,
ELECTROMAGNETICS AND NANOSCIENCES
CURRICULUM: MATHEMATICS FOR ENGINEERING

Mathematical Models and Methods for Traffic Flow and Stop & Go Waves

Supervisors:
Prof. Fabio Camilli
Dott. Emiliano Cristiani

Candidate:
Elisa Iacomini

Cycle XXXII

Acknowledgments

Before focusing on the thesis, special thanks are due to the people who made it possible, directly or indirectly.

First of all, I wish to gratefully thank my supervisors Fabio Camilli and Emiliano Cristiani for guided and supported me during my PhD studies and research with their knowledge, expert advice and their immense patience. They always helped me in overcoming the numerous obstacles I have been facing during these three years.

I am really grateful to Simone Göttlich and all the SCICOM team for their kindness and help during my visiting period in Mannheim, and, lastly, for giving me the possibility to come back. Vielen dank!

A special thank is due to Maya Briani for her great collaboration and useful suggestions.

I wish to thank Michael Herty and Andrea Tosin for accepting to be the reviewers of my thesis, for the time devoted to me and their insightful comments for improving this work.

Moreover, I will never forget my first teaching experience. I would heartful thank Tommaso Leonori for his support, encouragements and useful suggestions.

I was also very lucky since I had the opportunity of sharing these three years at SBAI with colleagues that easily became my friends. A special thank to Raul, Francescantonio, Riccardo, Simone, Alessandro, Pierluigi, Davide, Lorenzo, Michela, Caterina, Laura and Gabriele for the support, the help and the patience they showed me. Furthermore, I would thank Luca and Giulio for the time we spent together after the Tuesday seminars.

An heartful thank to my friend and room mate Rossella, for making me fell at home, supporting and helping me anytime I needed.

I am also grateful to the biker family for the Sunday trips and the fun we had together. ("V")

A special thank to all my friends for being part of my life.

Lastly, I think that all the possible thanks I can do to my parents will be not enough for being always beside me. This thesis is for you.

"We shall not cease from exploration;
And the end of all our exploring
will be to arrive where we started,
and know the place for the first time."
T.S Eliot

Abstract

In this thesis we are concerned with mathematical methods and models for traffic flow, with special emphasis to second-order effects like Stop & Go waves.

To begin with, we investigate the sensitivity of the celebrated Lighthill-Whitham-Richards model on network to its parameters and to the network itself. The quantification of sensitivity is obtained by measuring the Wasserstein distance between two LWR solutions corresponding to different inputs. To this end, we propose a numerical method to approximate the Wasserstein distance between two density distributions defined on a network.

After that, we present a new multi-scale method for reproducing traffic flow, which couples a first-order macroscopic model with a second-order microscopic model, avoiding any interface or boundary conditions between them. The new multi-scale model is characterized by the fact that microscopic and macroscopic descriptions are not spatially or temporally separated.

Furthermore, a delayed LWR model on networks is proposed in order to allow simple first-order models to describe complex second-order effects caused by bounded accelerations. A time delay term is introduced in the flux term and its impact is studied from the numerical point of view.

Lastly, we focus on Stop & Go waves, a typical phenomenon of congested traffic flow. Real data are used to point out the main features of this phenomenon, then we investigate the possibility to reproduce it using new traffic models specifically conceived for this purpose.

Contents

Introduction

Since many years traffic models have become an indispensable tool in the urban management of vehicular traffic and in the last decades an intensive research activity in the field of traffic flow modelling flourished. The presence of hard congestions in urban networks may have dramatic implications in modern society, affecting, e.g. productivity, pollution, fuel consumption and life-style. Nowadays, many projects about urban intelligence and sustainable mobility require in some way the ability to forecast traffic flow. Therefore, new solutions to these issues will be of great socio-economical impact. Moreover, the modern technologies are an incessant source of new challenges for mathematicians, physicists and engineers. The big amount of data available gives opportunities to improve models calibrating and validating the existing ones, but also to catch the main features of traffic evolution in order to build new accurate models. To do that, researchers can analyse and process data coming from different sources as GPS devices that track the exact position of a vehicle, or fixed sensors that record the flux of vehicles passing through a certain point.

The idea of modelling vehicular traffic has a long history dating back to the beginning of the 20th century. In the last seventy years, many methods have been developed resorting to different approaches. Starting from the natural idea of tracking every single vehicle, several Follow-the-Leader models grew-up for computing positions, velocities and accelerations of each car by means of systems of ordinary differential equations (ODEs) [3, 25, 29]. Other ways go from kinetic [72] to macroscopic fluid-dynamic and measures approaches [4, 13, 14, 40, 60, 63], focusing on averaged quantities by means of systems of partial differential equations (PDEs), in particular conservation laws. The choice of the scale of observation mainly depends on the number of the involved vehicles, the size of the network and so on. Moreover the connections among different scales are well established [21, 29, 30, 37, 38, 45]. These nice links are the basis for developing multi-scale models [18, 42, 43, 50, 53].

The work presented in this thesis fits in the framework mentioned above since we deal with microscopic, macroscopic and multi-scale traffic flow models on networks. We focus also on comparing different solutions using Wasserstein distance, in particular we present a numerical approximation of it. Furthermore, we investigate specifically the phenomenon of Stop & Go waves starting from a set of real data coming from Autovie Venete S.p.A. (AV), an Italian company operating on the highways of the North-East of Italy.

In Chapter 1, we will focus on the mathematical background necessary to deal with traffic flow models. Conservation laws theory will be recalled together with the numerical counterpart [32, 54, 55]. Then we will focus on traffic flow modelling and in particular we will describe some approaches existing in literature. Furthermore, the extension to the network is required in order to describe real situations

where the vehicles move on several interconnected roads. On this topic, a large literature exists [8, 14, 40] and different techniques have been proposed so far. One useful tool in the theory of conservation laws is the Wasserstein distance and its link with the well known Monge Kantorovich mass transfer problem, or the optimal transport problem [65, 70, 71].

Focusing on the most famous macroscopic model introduced by Lighthill, Whitham [56] and Richards [61], in Chapter 2 we will propose a new approach to extend it on large networks, following the idea presented in [9, 10]. Vehicles are divided on the basis of their path assuming that drivers know the way to get their destination. Since the number of equations will grow as the network size increases, this approach will become unfeasible on large networks, then the idea is to divide the drivers on the basis of their path only next to the junctions. Then, the sensitivity of the model to its parameters will be investigated. To this end, we need a suitable tool to compare and quantify the difference (or the closeness) between various scenarios. Since the L^p- distances are not sufficient to catch the natural concept of distance among densities, the natural notion of distance we need is the Wasserstein one. The bottleneck for using this distance on networks comes from the computational side. Indeed, the classical definition of Wasserstein distance is not suitable for numerical approximation. Recent characterizations also seem to be unfit for this goal. Let us mention in this regard the variational approach proposed by Mazón et al. [57], which generalizes to networks the results by Evans and Gangbo [32]. It can be also shown [57] that the Wasserstein distance has a nice link with the p-Laplacian operator. This leads naturally to a PDE-based approach to solve the problem but both these approaches are highly ill-conditioned and then computationally infeasible. Exploiting the link with optimal transport problem, we will propose and implement a pure discrete, reasonably fast algorithm to approximate the Wasserstein distance on a network, based on a linear programming method. The sensitivity analysis will be done with respect to the initial data, the network size, the capacity of roads and traffic distribution at junctions.

It is well known that first order models represent a simplification of the reality since they are built on the assumption that accelerations are instantaneous, and the traffic conditions are always at equilibrium. Second order models instead are closer to the real dynamics of drivers since they consider bounded accelerations [3, 33]. On the other hand, second order macroscopic models are difficult to handle. In order to overcome issues coming from single scale models, many kinds of multi-scale models were proposed so far: first-order FtL and LWR [18], second-order FtL and LWR [43], second-order FtL and phase-transition model [42], second-order FtL and second order macroscopic model [53]. They all are characterized by the interface which separates micro- and macro-model, which can be either fixed or solution-dependent. In Chapter 3 we will propose a new multi-scale approach, where the interface is not needed. We will couple a first order macroscopic model with a second order microscopic one in order to inherit the advantages of both the approaches and get an easy-to-implement macroscopic model with second-order features, avoiding high computational costs.

In order to catch second order effects, mixing different scales is not the only way. Taking into account that the velocity can not change instantaneously, instead of switching to second order models, one could introduce a time delay in first order models. Since many phenomena need some transient to become visible or effective, several applications of delayed-systems are presented in literature: the study of the evolution of the HIV in medicine [27, 67], the feedback control loops in control engineering [52], and many applications in mechanics and economics [6]. In traffic flow models the delay represents the reaction time of both drivers and vehicles. Delayed first order models on networks are the core of Chapter 4. The impact of the delay on the solutions will be investigated, and the comparison with classical models will

be provided.

Since our aim is to develop new strategies to reproduce second order acceleration based effects, in Chapter 5 we will focus on a typical phenomenon of traffic flow, the so-called Stop-and-Go (S&G) waves: vehicles stop and restart without any apparent reason, and this perturbation travels backward with respect to the cars' trajectories. S&G waves often appear in traffic flow literature [62, 68] and one of the main open challenges is modelling and simulating them. To this aim we will use real traffic data coming from Autovie Venete, which will help us to point out the main features of this phenomenon. We will develop a new microscopic second order model specifically conceived to reproduce S&G waves, then mixing scales or introducing a time-delay in the model, we will be able to recover S&G waves at a macroscopic level too.

Publications

- C. Balzotti and E. Iacomini,
  Stop & Go waves: a microscopic and a macroscopic description,
  submitted.

- E. Cristiani and E. Iacomini,
  An interface-free multi-scale multi-order model for traffic flow,
  Discrete and Continuous Dynamical Systems - Series B, 24 (2019), pp. 6189-6207.

- F. Camilli, R. De Maio and E. Iacomini,
  A Hopf-Lax formula for Hamilton-Jacobi equations with Caputo time derivative,
  Journal of Mathematical Analysis and Applications, 477 (2019), pp. 1019-1032.

- M. Briani, E. Cristiani and E. Iacomini,
  Sensitivity analysis of the LWR model for traffic forecast on large networks using Wasserstein distance,
  Communications in Mathematical Science, 16 (2018), pp. 123-144.

Resumé of Activities

During these three years of PhD Program, I dealt with the study of models and numerical methods for vehicular traffic forecast. My research has been supervised by Prof. Fabio Camilli (SBAI, Sapienza) and Dott. Emiliano Cristiani (IAC-CNR).

During the first PhD year, I studied the sensitivity analysis for the LWR model on large networks and optimal transport problems.

After having investigated second order macroscopic models for traffic flow at the beginning of my second year, I focused my attention on multi-scale algorithms, which are able to reproduce real traffic phenomena with easy implementation.

During the last PhD year, I considered a new microscopic model specifically conceived to reproduce Stop and Go phenomena, which are typical of congested traffic.

Moreover, I spent two months at Mannheim University, hosted by Prof. Simone Göttlich with a scholarship founded by the German Government (IPID4all), where I worked on delayed models for traffic flow and their extension to networks.

Furthermore, I have studied fractional differential equations in collaboration with Fabio Camilli and Raul De Maio.

Last year, I won a grant for Young Researcher funded by Sapienza with the project PREVENT: accurate PREdictions for VEhicular Traffic on Networks in urban society.

In these 3 years I also passed 5 exams:

- Stochastic processes
  Prof. Emilio De Santis (Sapienza University of Rome)
- Uncertainty quantification for hyperbolic equations
  Prof. Pierangelo Marcati (GSSI, L'Aquila)
- Introduction to parallel computing with MPI and OpenMP
  Prof. Ruggiero Vittorio (CINECA, Rome)
- An introduction to the functions of bounded variation and existence results for elliptic equations with p-laplacian principal part (p greater or equal to 1) and singular lower order terms
  Prof. Virginia De Cicco (Sapienza University of Rome) and Prof. Daniela Giachetti (Sapienza University of Rome)
- Nonlinear diffusion in inhomogeneous environments
  Prof. Anatoli Tedeev (South Mathematical Inst. of VSC Russian Acad. Sci. Vladikavkaz Russia) and Prof. Daniele Andreucci (Sapienza University of Rome)

I have also followed these courses:

- Advanced graphics for scientific data
  Prof. Simone Cacace (Roma Tre, Rome)
- Fractional calculus and singular equations
  Prof. Mirko D'Ovidio (Sapienza University of Rome), Prof. Tommaso Leonori (Sapienza University of Rome) and Prof. Francesco Petitta (Sapienza University of Rome)

Workshop and conferences

I have attended the following schools, workshops and conferences:

- INdAM Workshop: PDE Models for Multi-Agent Phenomena
  INdAM, Rome, November 7-11, 2016
- INdAM Workshop: Transport Modeling and Management
  INdAM, Rome, March 6-10, 2017
- School on Uncertainty Quantification for Hyperbolic Equations and Related Topics
  GSSI (*with financial support*), L'Aquila, April 24-28, 2017.
- International School on Frontiers in Partial Differential Equations Analysis and Solvers
  Palazzo Vistarino (*with financial support*), Pavia (MI), May 22-25, 2017.
- Workshop on Data-Driven Methods for Multi-Scale Physics and Complex Systems
  Rome, 31 July-4 August 2017.
- The XVII Italian Meeting on Hyperbolic Equations IperPV2017
  Pavia (MI), September 6-8, 2017.
- School on Optimal Transport: Numerical Methods and Applications
  Lake Como School of Advanced Studies (*with financial support*), Como, May 7-11, 2018.
- Summer School on Mean Field Games
  CIME Fundation (*with partial financial support*), Cetraro, June 10-14, 2019.

Scientific contributions

I had the opportunity to communicate my research and results with talks and posters:

- **Poster presentation at RICAM Workshop on Feedback Control**
  Johann Radon Institute for Computational and Applied Mathematics,
  Linz (Austria), November 28, 2019.
- **Invited speaker** at departmental seminars by Dott. Giacomo Albi
  University of Verona, September 24, 2019.
- **Minisymposium's Talk at UMI19**
  Models and Applications,
  Pavia (MI), September 4, 2019.
- **Minisymposium's Talk at ICIAM19**
  Mathematical descriptions of traffic flow: micro, macro and kinetic models,
  Valencia (Spain), July 18, 2019.
- **Talk** at Seminario di Modellistica Differenziale Numerica
  Dept. of Mathematics, Sapienza University of Rome, December 11, 2018.

- **Invited speaker at KWIM: Women in Matematics**
  University of Konstanz (Germany), November 20, 2018.
- **Scientific Computing Research Seminar**
  University of Mannheim (Germany), October 18, 2018.
- **Contributed talk at SIMAI18**
  Dept. of Engineering, Rome, July 6, 2018.
- **Poster presentation at Numasp18**
  Numerical Aspects of Hyperbolic Balance Laws and Related Problems,
  Auditorium Santa Lucia, University of Ferrara, April 17, 2018.
- **Talk** at Seminari di Modellistica Differenziale Numerica
  Dept. of Mathematics, Sapienza University of Rome, November 14, 2017.

Chapter 1

Background

In this chapter we present the theoretical tools and results we will refer to throughout the thesis. We recall the basic notion about the theory of conservation laws, which are first order nonlinear hyperbolic partial differential equations (PDEs) with a particular structure, described in the first section.

The second part of the chapter is devoted to the applications. In particular we focus on traffic flow, giving an overview of the main approaches and the existing models. After investigating the main features of each model we deal with their extension to the road network, going deeper in the description of different approaches.

Lastly we introduce the notion of Wasserstein distance in order to compare different traffic scenarios and quantify how similar (or not) they are.

1.1 Conservation laws

Conservation laws are essential for understanding the physical world around us, indeed they describe the conservation in time of a quantity in an isolated system.

The typical form of a scalar conservation law is the following:

$$\partial_t u(x,t) + \partial_x f(u(x,t)) = 0, \tag{1.1.1}$$

where $u\colon\ \mathbb{R}\times\mathbb{R}^+ \to \mathbb{R}$ is the "conserved quantity", or state variable, and $f\colon\ \mathbb{R}^m \to \mathbb{R}^m$ is the flux.

They are also known as continuity equations, partial differential equations which give a relation between the amount of the quantity and the "transport" of that quantity, indeed (1.1.1) can be seen as a generalization of the rigid transport equation.

The formulation of equation (1.1.1) relies on the continuum assumption, in which u is assumed to be indefinitely divisible without changing its physical nature, for example a fluid has to be treated as continuous even if is composed of molecules. The *continuum assumption* is ensured by considering a very large mass compared to the size of the domain.

In order to see how conservation laws arise from physical principles, let us focus on a one-dimensional gas dynamics problem. For example consider a gas flowing in a tube where the properties of the gas such as density and velocity are assumed to be constant across each cross section of the tube.

Let u be the gas density in the section from a to b of the tube and $f(u)$ the flux through the boundaries.

Our aim is understanding how the gas quantity changes in the considered domain. The volume of the gas is fixed since it is in the tube, so the variation of the gas quantity depends on the variation in time of the density. On the other hand the variation of the gas is given by the difference between the incoming and outgoing gas quantities. Mathematically we have:

$$\frac{d}{dt}\int_a^b u(x,t)\,dx = f(u(a,t)) - f(u(t,b)) \tag{1.1.2}$$

that is called *integral form* of the conservation law. Under regularity assumptions for u, we can rewrite the equation:

$$\frac{d}{dt}\int_a^b u(x,t)\,dx = -\frac{d}{dt}\int_a^b f(u(x,t))_x\,dx \tag{1.1.3}$$

and switching the integral and derivative signs we obtain:

$$\int_a^b \partial_t u(x,t)\,dx + \int_a^b \partial_x f(u(x,t)) = 0, \tag{1.1.4}$$

and lastly:

$$\int_a^b (\partial_t u(x,t) + \partial_x f(u(x,t)))\,dx = 0. \tag{1.1.5}$$

Since (1.1.5) has to hold for any a and b, we easily recover (1.1.1).

Let us now focus on the main properties of the solutions. Consider the following Cauchy problem:

$$\begin{cases} \partial_t u(x,t) + \partial_x f(u(x,t)) = 0, & x \in \mathbb{R},\ t > 0, \\ u(x,0) = u_0(x). \end{cases} \tag{1.1.6}$$

It is well known that due to the nonlinearity of the problem, we can not define classical solutions, that means $u(x,t) \in C^1(\mathbb{R}\times\mathbb{R}^+)$, since the solution may develop discontinuities in a finite time also for regular initial data. For this reason we are looking for weak solutions to (1.1.6).

Definition 1.1.1 (*Weak solution*)**.** *A function $u(x,t) \in L^\infty\left(L^1 \cap L^\infty(\mathbb{R}), \mathbb{R}^+\right)$ that satisfy* (1.1.6) *in the sense of distributions:*

$$\int_{\mathbb{R}}\int_{\mathbb{R}^+} \Big[u(x,t)\partial_t\varphi(x,t) + f(u(x,t))\partial_x\varphi(x,t)\Big]dtdx + \int_{\mathbb{R}} u_0(x)\varphi(x,0)dx = 0 \tag{1.1.7}$$

$\forall \varphi \in C_c^\infty(\mathbb{R};\mathbb{R}^+\times\mathbb{R})$, is called weak solution to (1.1.6)*.*

On the other hand, the notion of weak solutions just introduced is not strong enough to provide uniqueness of solutions to (1.1.6), indeed we persist having infinite solutions.

Consider for example the Burger equation $u_t + \left(\frac{u^2}{2}\right)_x = 0$ with initial data $u_0 = 0$ for $x < 0$ and $u_0 = 1$ for $x \geq 0$:

$$u_\alpha(x,t) = \begin{cases} 0 & x < \frac{\alpha t}{2} \\ \alpha & \frac{\alpha t}{2} \leq x < \frac{(1+\alpha)t}{2} \\ 1 & x \geq \frac{(1+\alpha)t}{2} \end{cases} \tag{1.1.8}$$

are a family of weak solutions for every $0 < \alpha < 1$.

So we need a more accurate definition able to pick out the most physically relevant solution to (1.1.6).

Definition 1.1.2 (*Entropy solution*). *A function $u \in L^\infty(\mathbb{R}^+; L^1 \cap L^\infty(\mathbb{R}))$ is an entropy solution to* (1.1.6) *if it satisfies the entropy inequality:*

$$\int_{\mathbb{R}} \int_{\mathbb{R}^+} \Big[|u(x,t)-k| \partial_t \varphi(x,t) + sgn(u(x,t)-k)[f(u(x,t))-f(k)] \partial_x \varphi(x,t) \Big] dt dx + \int_{\mathbb{R}} |u(x)-k| \varphi(x,0) dx \geq 0 \tag{1.1.9}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^+ \times \mathbb{R}; \mathbb{R})$ with $\varphi \geq 0$, and for all constants $k \in \mathbb{R}$.

Clearly, any entropy solution is a weak solution too. The uniqueness of the entropy solution is given by the theorem below, proved by Kružkov in the '70s, see [51].

Theorem 1.1.3 (Kružkov). *Assume that the flux is locally Lipschitz. Then, for any given initial condition $\bar{u}$ in L^∞ with compact support, there exists a unique entropy solution to* (1.1.6) *in the sense of Definition 1.1.2.*

1.1.1 Riemann problem

For linear hyperbolic equations characteristics lines play a major role. For nonlinear problems, the generalization of this theory, which is most frequently used in understanding how the solution behaves and in developing numerical methods, is the solution of a Riemann problem.

A Riemann problem is nothing less than a Cauchy problem with a particular initial data consisting in two constant states separated by a single discontinuity:

$$\begin{cases} \partial_t u(x,t) + \partial_x f(u(x,t)) = 0 \\ u_0(x) = \begin{cases} u^- & x < 0 \\ u^+ & x \geq 0, \end{cases} \end{cases} \tag{1.1.10}$$

where $u^-, u^+ \in \mathbb{R}$. For example let us compute explicitly the entropy solution in the case of a concave flux.

- If $u^- < u^+$ (*shock wave*) the entropic solution is given by:

$$u(x,t) = \begin{cases} u^- & x < \lambda t \\ u^+ & x \geq \lambda t \end{cases} \qquad \lambda = \frac{f(u^+) - f(u^-)}{u^+ - u^-}, \tag{1.1.11}$$

  The shock wave will have a positive speed if $f(u^+) \geq f(u^-)$, negative on the contrary. Equation (1.1.11) is also known in the literature as the Rankine-Hugoniot condition.

- If $u^- \geq u^+$ (*rarefaction wave*) the entropic solution is given by:

$$u(t,x) = \begin{cases} u^- & \frac{x}{t} \leq f'(u^-) \\ \psi\left(\frac{x}{t}\right) & f'(u^-) \leq \frac{x}{t} \leq f'(u^+) \\ u^+ & \frac{x}{t} \geq f'(u^+), \end{cases} \tag{1.1.12}$$

  where the function $\psi(\xi)$ is defined by the solution if $f'(\psi(\xi)) = \xi$. The speed of the rarefaction wave will be positive if $f'(u^+), f'(u^-) > 0$, it will be negative if $f'(u^+), f'(u^-) < 0$.

1.1.2 System of conservation laws

In order to describe more complicated physical problems such as the motion of particles in a fluid with no viscosity, a scalar conservation law is not longer sufficient. The main example in this framework is represented by the Euler equations of fluid dynamics, in which we have three unknowns: the density, the velocity of particles and the internal energy that have to satisfy respectively three conservation laws, i. e. for the conservation of mass, momentum and global energy.

A system of conservation laws in one space variable is a partial differential equation in the form:

$$\partial_t \mathbf{u}(x,t) + \partial_x[\mathbf{F}(\mathbf{u}(x,t))] = 0 \tag{1.1.13}$$

where $x \in \mathbb{R}$, $t \in \mathbb{R}^+$, $\mathbf{u}: \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^m$ is the vector of unknowns and $\mathbf{F}: \mathbb{R}^m \to \mathbb{R}^m$ describes the fluxes.

Remark 1.1.4. *There exists also more complicated systems where the space variable is a vector $\boldsymbol{x} \in \mathbb{R}^n$. The problem can be rewritten taken $\boldsymbol{u}: \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^n$ and $\boldsymbol{F}: \mathbb{R}^m \to \mathcal{M}^{m \times n}$. The theory for this kind of systems is not so developed due to the difficulties we can easily imagine.*

Let us consider the Cauchy problem associated to (1.1.13):

$$\begin{cases} \partial_t \mathbf{u}(x,t) + \partial_x[\mathbf{F}(\mathbf{u}(x,t))] = 0, & x \in \mathbb{R}, t > 0, \\ \mathbf{u}(x,0) = \mathbf{u}_0(x). \end{cases} \tag{1.1.14}$$

Assume that $\mathbf{F}$ has a sufficient regularity in order to admit its Jacobian matrix $A(\mathbf{u})$.

Under hypothesis of regularity on $\mathbf{u}$, we can rewrite (1.1.13) in the quasi linear form:

$$\partial_t \mathbf{u} + A(\mathbf{u}) \partial_x \mathbf{u} = 0. \tag{1.1.15}$$

Definition 1.1.5 (Hyperbolic system)**.** *The system* (1.1.13) *is hyperbolic if for every $\boldsymbol{u} \in \mathbb{R}^m$, the matrix $A(\boldsymbol{u})$ admits m real eigenvalues $\lambda_1(\boldsymbol{u}), \dots, \lambda_m(\boldsymbol{u})$.*
Moreover if all the eigenvalues are distinct the system is strictly hyperbolic.

Let us denote by $r_i(\mathbf{u})$ a right eigenvector such that $A(\mathbf{u}) r_i(\mathbf{u}) = \lambda_i(\mathbf{u}) r_i(\mathbf{u})$.

Definition 1.1.6 (Characteristic field)**.** *$\lambda_i(\boldsymbol{u})$ is called characteristic field, in particular it is the λ_i-field.*

Definition 1.1.7. *Let $\lambda_1 < \lambda_2 < \cdots < \lambda_m$ be m distinct real eigenvalues of the Jacobian matrix A. Since A depends in the solution $\boldsymbol{u}$, so do the eigenvalues λ_i, $i = 1, \dots, m$ and the corresponding eigenvectors r_i. Let $\nabla_{\boldsymbol{u}}$ denote the gradient with respect to $\boldsymbol{u}$. Then the λ_i-field is called genuinely nonlinear if the function:*

$$\nabla_{\boldsymbol{u}} \lambda_i(\boldsymbol{u}) \cdot r_i(\boldsymbol{u}) \neq 0,$$

in other words if $\nabla_{\boldsymbol{u}} \lambda_i(\boldsymbol{u})$ is not orthogonal to the corresponding eigenvector $r_i(\boldsymbol{u})$.

On the other hand, if

$$\nabla_{\boldsymbol{u}} \lambda_i(\boldsymbol{u}) \cdot r_i(\boldsymbol{u}) = 0$$

$\lambda_i(\boldsymbol{u})$ is called linearly degenerate.

Note that for a linear system the eigenvalues are constant so we have $\nabla_{\mathbf{u}} \lambda_i(\mathbf{u}) = 0$ for every $i = 1, \dots, m$, thus they are linearly degenerate.

A crucial role is played by some quantities that are invariant along the characteristic fields called Riemann invariants.

Definition 1.1.8 (Riemann invariant)**.** *A Riemann invariant associated to the eigenvector $r_i(\boldsymbol{u})$ is a scalar function z that depends on $\boldsymbol{u}$ such that:*

$$\nabla z(\boldsymbol{u}) \cdot r_i(\boldsymbol{u}) = 0.$$

If the $i-$th characteristic field is genuinely nonlinear, we can assume without loosing of generality that:

$$\nabla_{\mathbf{u}} \lambda_i(\mathbf{u}) \cdot r_i(\mathbf{u}) > 0 \qquad \forall u \in \Omega. \tag{1.1.16}$$

This means that $\lambda_i(\mathbf{u})$ is strictly decreasing or increasing along the direction of r_i, otherwise there would exist a point $\mathbf{u}_0$ such that:

$$\nabla_{\mathbf{u}} \lambda_i(\mathbf{u}_0) \cdot r_i(\mathbf{u}_0) = 0 \tag{1.1.17}$$

that contradicts the hypothesis on the characteristic field.

1.2 Numerical aspects

After some investigations on the analytical properties on the theory of conservation laws, let us focus on their numerical counterpart.

Since the solution of conservation laws may develop jumps in finite time, we expect that a difference discretization by finite differences of the PDE is inappropriate near discontinuities.

One possible approach is to combine a standard finite difference method in smooth regions with some explicit procedure for tracking the location of discontinuities. These methods are usually called "*shock tracking*". They are quite easy in one dimensional problem, but become harder in more space dimensions in which the discontinuity lies on a curve or a surface that interacts each other.

On the other hand we would like to have a numerical method that will produce sharp approximations to discontinuous solutions automatically, without ad hoc techniques or using jump conditions. These methods are called "*shock capturing*". A great variety of these schemes is now available due to their wide applications and this is our framework.

Consider a typical Cauchy problem for a scalar conservation law:

$$\begin{cases} \partial_t u(x,t) + \partial_x f(u(x,t)) = 0, & x \in \mathbb{R}, t > 0, \\ u(x,0) = u_0(x). \end{cases} \tag{1.2.18}$$

The first step in any numerical approximation is to discretize both the spatial and temporal domain. Let $\Delta t > 0$ and $\Delta x > 0$ be respectively the time and the space steps. It will be useful define $u_j^n = u(x_j, t^n)$ as the solution at time $t^n = n\Delta t$ at point $x_j = j\Delta x$ with $j \in \mathbb{Z}$ and $n \in \mathbb{N}$. Since the solution we are looking for is not supposed to be continuous we have to be careful about considering pointwise values. To overcome this issue we will consider the cell average:

$$u_j^n = u(x_j, t^n) = \frac{1}{\Delta x} \int_{x_j - \frac{1}{2}\Delta x}^{x_j + \frac{1}{2}\Delta x} u(x, t^n)\, dx. \tag{1.2.19}$$

Moreover, the initial datum too has to be projected on a $\bar{u}_0(x)$ piecewise constant function. This is why the Riemann problem (1.1.10) plays such an important role here.

Before introducing the algorithms, let us briefly recall general notions about numerical schemes for conservation laws. To this aim let us introduce $\hat{u}(x,t)$ as the solution of the approximated problem. Since we can not handle numerically infinite domains, we assume $L>0$ as the length of the space interval, i.e. the length of a road, and $N_x=\frac{L}{\Delta x}$ the number of the cells in which the space is divided.

First of all, in order to guarantee that the numerical methods do not converge to non-solutions let us require that the schemes we will consider admit the *conservation form.*

Definition 1.2.1 (Conservative form)**.** *Given a uniform grid with time step Δt and spacial mesh size Δx, a numerical method is said to be conservative if the corresponding scheme can be written as:*

$$u_j^{n+1}=u_j^n-\frac{\Delta t}{\Delta x}\left(g_{j+\frac{1}{2}}^n-g_{j-\frac{1}{2}}^n\right), \qquad j=1,\dots,N_x, \quad n\in\mathbb{N} \tag{1.2.20}$$

where $g:\mathbb{R}^{2k}\to\mathbb{R}$ is a continuous function, called the numerical flux function, that defines a $(2k+1)$-point scheme: $g_{j+\frac{1}{2}}^n=g(u_{j-k+1}^n,\dots,u_{j+k}^n)$.

Essentially, the conservative form ensures that the discretization technique actually represents a discrete approximation to the integral form of the conservation laws.

Another important property of our numerical methods is the consistency with (1.1.1), that is the numerical flux function g reduces to the true flux f for the case of constant flow, i. e. $g(u,\dots,u)=f(u) \quad \forall u\in\mathbb{R}$. In order to guarantee the consistency it is sufficient assuming g as a Lipschitz continuous function in each variable.

The main advantage of conservative and consistent schemes is that, when they converge, they converge to solutions whose shocks or discontinuity satisfy automatically the jump conditions, that is, the discontinuities always travel at the correct velocity as the following theorem ensures.

Theorem 1.2.2 (Lax-Wendroff)**.** *Assume that $\hat{u}(x,t)$ is the solution of a conservative and consistent scheme. Let $\hat{u}(x,t)$ be uniformly bounded in L^∞ and convergent to $u(x,t)$ almost surely, as the grid size Δt and Δx go to 0. Then $u(x,t)$ is a weak solution of* (1.1.1).

For the proof see [55].

Note that the Lax-Wendroff theorem guarantees that the approximated solution converges only to a weak solution. In order to prove the convergence to the entropic solution we have to verify another condition.

Definition 1.2.3. *A function $U(u)$ and a function $F(u)$ are called respectively the entropy and the entropy flux of u if:*

i) $U(u)$ is absolutely continuous and convex;

ii) $F'(u)=U'(u)f'(u)$ where f is the flux of the conservation law and it is derived with respect to u.

Definition 1.2.4. *A numerical scheme written in the conservative form is consistent with the entropy inequality if $\forall U$ defined above, there exists a numeric entropic flux $\hat{F}=\hat{F}(u,\dots,u)$ such that:*

i) $U(u_j^{n+1})-U(_j^n)+\frac{\Delta t}{\Delta x}\left(\hat{F}_{j+\frac{1}{2}}^n-\hat{F}_{j-\frac{1}{2}}^n\right)\le 0$;

ii) $\hat{F}(u,\dots,u)=F(u)$.

Theorem 1.2.5. *Suppose that a conservative scheme is consistent with the entropic inequality, the approximated solution* $\hat{u}(x,t)$*is uniformly bounded in* L^∞ *and converges a.s. to* $u(x,t)$*. Then* $u(x,t)$ *is the entropic solution of* (1.1.1)*.*

1.2.1 Lax Friedrichs scheme

The Lax Friedrichs method presented by Lax in 1954 [54] is the simplest conservative and consistent first order scheme. It is a centred scheme based on finite differences. The typical form is given by:

$$u_j^{n+1}=\frac{1}{2}\left(u_{j+1}^n+u_{j-1}^n\right)-\frac{\Delta t}{2\Delta x}\Big(f(u_{j+1}^n)-f(u_{j-1}^n)\Big) \tag{1.2.21}$$

where f is the flux of the conservation law we have to solve. In order to guarantee the stability of the scheme we have to satisfy the Courant-Friedrichs-Levy condition (CFL):

$$\frac{\Delta t}{\Delta x}\max_u|f'(u)|\leq 1, \tag{1.2.22}$$

This scheme can be written also in the conservative form (1.2.20) choosing:

$$g_{j+\frac{1}{2}}^n=\frac{1}{2}\Big(f(u_{j+1}^n)+f(u_j^n)\Big)-\frac{\Delta x}{2\Delta t}\Big(u_{j+1}^n-u_j^n\Big). \tag{1.2.23}$$

Due to the presence of the average on the neighbour cells it smears out the solutions, hence it is quite dissipative.

1.2.2 Staggered Lax Friedrichs scheme

In order to reduce the dissipative behaviour of the Lax Friedrichs scheme, let us introduce a staggered grid as an intermediate step. The staggered density values are obtained by averaging over the neighbour densities. Then, centered differences are used with respect to the original grid points that are located a distance half of a step size to the considered point. Finally, the values are projected back to the original grid. For the detailed derivation of this scheme we refer to [44]. Note that the CFL condition has to be halved in order to preserve stability. The numerical scheme is the following:

$$\begin{aligned}
&\text{Left side:} && u_0^{n+1}=\tfrac{1}{4}(3u_0^n+u_1^n)-\tfrac{\Delta t}{2\Delta x}\left(f(u_1^n)+f(u_0^n)-2\bar{\gamma}^n\right)\\
&\text{Central points:} && u_j^{n+1}=\tfrac{1}{4}(u_{j-1}^n+2u_j^n+u_{j+1}^n)-\tfrac{\Delta t}{2\Delta x}\left(f(u_{j+1}^n)-f(u_{j-1}^n)\right)\\
&\text{Right side:} && u_{Nx}^{n+1}=\tfrac{1}{4}(3u_{Nx}^n+u_{Nx-1}^n)-\tfrac{\Delta t}{2\Delta x}\left(2\tilde{\gamma}^n-f(u_{Nx}^n)-f(u_{Nx-1}^n)\right),
\end{aligned} \tag{1.2.24}$$

where $\bar{\gamma}^n$ and $\tilde{\gamma}^n$ are known, recovered by the boundary conditions, and $f(u_i^n)$ is the flux of the conservation law.

1.2.3 Godunov scheme

One of the most important first order scheme for conservation laws is the Godunov scheme, introduced by Godunov in 1959. The main idea consists of solving several Riemann problems (1.1.10), one at each

cell boundary, to obtain the solution, since u has been projected on a piecewise constant function. Let us assume that the CFL condition is satisfied, so (3.2.4) holds. As said above, let us consider:

$$\bar{u}_j^n = \frac{1}{\Delta x}\int_{x_j-\frac{1}{2}\Delta x}^{x_j+\frac{1}{2}\Delta x} u(x,t^n)\,dx, \tag{1.2.25}$$

and each cell $C_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ for all $j \in \mathbb{Z}$.

The initial datum $u_0(x,t)$ has to be projected on a piecewise constant function. Recalling (1.1.2) and assuming $a = x_{j-\frac{1}{2}}$ and $b = j+\frac{1}{2}$, we obtain:

$$\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u\left(t^{n+1},x\right)dx - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u\left(t^n,x\right)dx = -\int_{t^n}^{t^{n+1}} f\left(u\left(t,x_{j+\frac{1}{2}}\right)\right)dt + \int_{t^n}^{t^{n+1}} f\left(u\left(t,x_{j-\frac{1}{2}}\right)\right)dt. \tag{1.2.26}$$

just applying the fundamental theorem of calculus. Note that on the left we have spacial integrals on the cell C_j, and, for construction, the values in the cells are constant:

$$\triangle x\left(\bar{u}_j^{n+1} - \bar{u}_j^n\right) = -\int_{t^n}^{t^{n+1}} f\left(u\left(t,x_{j+\frac{1}{2}}\right)\right)dt + \int_{t^n}^{t^{n+1}} f\left(u\left(t,x_{j-\frac{1}{2}}\right)\right)dt. \tag{1.2.27}$$

Now, focusing on the right side of (1.2.26), we observe that $f(u(t,x_{j-\frac{1}{2}}))$ and $f(u(t,x_{j-\frac{1}{2}}))$ are constant for $t \in [t^n, t^{n+1}]$. It happens because u at time t^n is piecewise constant and the Riemann problems generate constant solutions at the cell interfaces. Therefore we can write the scheme as follows:

$$\triangle x\left(\bar{u}_j^{n+1} - \bar{u}_j^n\right) = -\triangle t\; G\left(\bar{u}_j^n, \bar{u}_{j+1}^n\right) + \triangle t\; G\left(\bar{u}_{j-1}^n, \bar{u}_j^n\right) \tag{1.2.28}$$

where G is the numerical Godunov flux: $G(\bar{u}_j^n, \bar{u}_{j+1}^n) = f(u(x_{j+\frac{1}{2}}^n))$ and $G(\bar{u}_{j-1}^n, \bar{u}_j^n) = f(u(x_{j-\frac{1}{2}}^n))$.

From the construction of the scheme, it is clear that we can write it in the conservative form:

$$u_j^{n+1} = u_j^n - \frac{\triangle t}{\triangle x}\left[G\left(u_j^n, u_{j+1}^n\right) - G\left(u_{j-1}^n, u_j^n\right)\right]. \tag{1.2.29}$$

To our purpose, we are interested in studying the non convex flux case, so the Godunov flux is:

$$G\left(u^-,u^+\right) = \begin{cases} \min\limits_{u\in[u^-,u^+]} f(u) & u^- \le u^+ \\ \max\limits_{u\in[u^-,u^+]} f(u) & u^- > u^+ \end{cases}. \tag{1.2.30}$$

1.3 Application to traffic flow

Conservation laws are considered to be fundamental laws of nature, with broad application in physics, as well as in other fields such as chemistry, biology, geology, and engineering.

For instance let us consider a single lane road with two arbitrary points a and b, $a < b$, along the road. Let us call $\rho(x,t)$ the density of cars at the point $x \in [a,b]$ at time $t \ge 0$, that is the number of cars in a kilometre, and $v(x,t)$ the mean velocity. The variation of the number of cars is due only to the incoming flux in $x = a$ and to the outgoing flux in $x = b$, where the flux is given by the vehicles times their mean

velocity, $f=\rho v$. Mathematically we have:

$$\frac{d}{dt}\int_a^b \rho(x,t)dx = v(\rho(a,t))\rho(a,t) - v(\rho(b,t))\rho(b,t) = -\int_a^b \Big[f(\rho(x,t))\Big]_x dx. \tag{1.3.31}$$

Since it holds for all $a<b$, we recover the scalar conservation law:

$$\partial_t \rho(x,t) + \partial_x f(\rho(x,t), v(x,t)) = 0. \tag{1.3.32}$$

Therefore we are able to describe the evolution of traffic flow thanks to the theory of conservation laws. Moreover we can think of traffic as a fluid and use fluid dynamics equations to describe its behaviour in terms of macroscopic variables as we will see in the next section.

1.3.1 Macroscopic approach

In macroscopic models [40], traffic is described in terms of macroscopic variables such as the average density $\rho=\rho(x,t)$, that is the number of vehicles in a kilometre, and their mean velocity $v=v(x,t)$ at the point $x\in\mathbb{R}$ at time $t>0$. So we loose the detailed level of vehicles' description, indeed they become indistinguishable from each other.

Let us assume that $\rho\in[0,\rho^{max}]$, where $\rho=0$ means that there are no cars, i. e. the road is empty, while $\rho=\rho^{max}$ stands for the fully congested scenario, i. e. the cars are bump-to-bump. As well as density, $v\in[0,v^{max}]$, and in analogy with the real observations, a reasonable behaviour for the velocity is a decreasing function as the density increases.

The typical form of a first order macroscopic model is:

$$\begin{cases} \partial_t \rho(x,t) + \partial_x(\rho(x,t)\ v(x,t)) = 0, \\ \rho(x,0) = \rho^0(x) \end{cases}, \tag{1.3.33}$$

for $(x,t)\in\mathbb{R}\times[0,T_f]$, a final time $T_f>0$, and initial distribution ρ^0.

The law that describes the flux in terms of the density is the so-called *fundamental diagram* that plays a crucial role in this framework. Physically it represents the number of vehicles that pass through the point x in a given length of time. It is clear that the flux has to be 0 both in the case of empty road, since the density is null, and in the case of of fully congested road, since the velocity is null in order to avoid collisions. But for intermediate values of density, it is more complicated as real data show [7], see Fig. 1.1. Indeed drivers act differently in response of the same traffic conditions.

A huge literature exists about the shape of this given law has to show, since it is one of the main ingredient of macroscopic models.

In the first-order models like the most famous introduced in the early '50 by Lighthill, Witham and Richards (LWR), the velocity $v=v(\rho)$ is given as a function of the density, so the flux is depending only on the density, that is the only unknown of the PDE.

The following hypotheses are made on the flux:

1) $f\in C^0([0,\rho^{max}])\cap C^1((0,\rho^{max}))$;

2) is strictly concave;

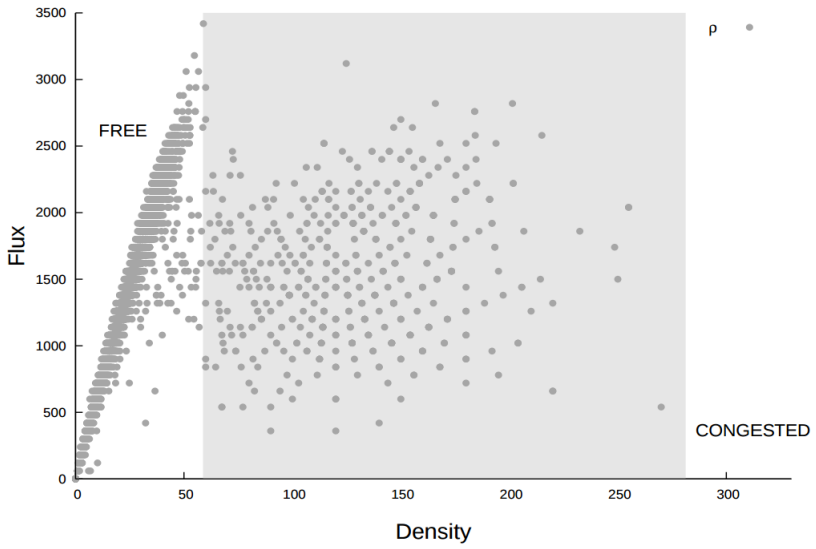

Figure 1.1: Fundamental diagram from real data [7].

3) $f(0) = f(\rho^{max}) = 0$;

4) $f(\sigma) = \max_{s \in (0,\rho^{max})} f(s)$.

There are different fundamental diagrams that satisfy the assumptions. The simplest one is the following:

$$f(\rho) = \rho \; v^{max}\left(1 - \frac{\rho}{\rho^{max}}\right), \tag{1.3.34}$$

mostly used by the mathematical community.

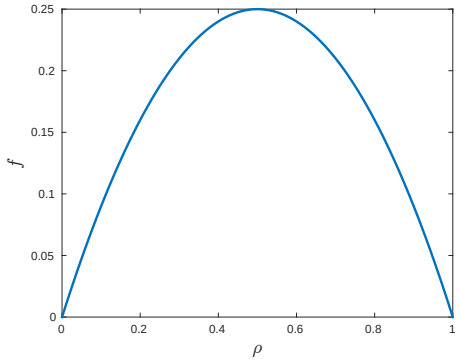

Figure 1.2: Typical parabolic fundamental diagram.

In order to fit the real data shown in Fig. 1.1, another fundamental diagram is also popular:

$$f(\rho) := \begin{cases} \frac{f_{\max}}{\sigma}\rho, & \text{if } \rho \leq \sigma \\ \frac{f_{\max}}{\sigma - 1}(\rho - 1), & \text{if } \rho > \sigma \end{cases} \tag{1.3.35}$$

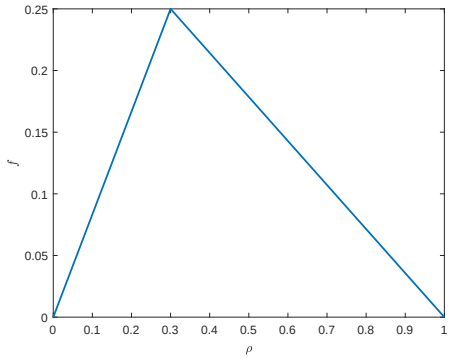

Figure 1.3: Typical triangular-shape fundamental diagram.

By the way, there exists also:

$$f(\rho) = \rho\ v_0 \log\left(\frac{\rho^{max}}{\rho}\right) \tag{1.3.36}$$

that comes from the experimental data from the Lincoln tunnel in New York and proposed for the first time by Greenberg. We have to mention the one studied by Greenshield:

$$f(\rho) = \rho\ v^{max}\left(1 - \left(\frac{\rho}{\rho^{max}}\right)^n\right) \tag{1.3.37}$$

and the California model too:

$$f(\rho) = \rho\ v_0\left(\frac{1}{\rho} - \frac{1}{\rho^{max}}\right). \tag{1.3.38}$$

Once chosen the fundamental diagram, there are no difficulties in solving (1.3.33). On the other hand the lacks of the LWR model, or first order models in general, are well known. For example, it fails to generate capacity drop, hysteresis, relaxation, platoon diffusion, or spontaneous congestions like stop and go waves, that are typical features of traffic dynamics.

In order to overcome this problem, second order models have been proposed. They take into account the non-equilibria states, assuming that accelerations are not instantaneous. To do this, the equation that describes the variation of the velocity in time has to be added to the system, replacing the typical given law $f(\rho)$ of first order models.

The first works on this direction were proposed by Payne and Whitham in the '70s, but they had some drawbacks, that were pointed out by Daganzo in 1995 [28]. The 'resurrection' of second order models is due to Aw and Rascle in the 2000 [4], and Zhang in 2002 [73], that highlighted for the first time the differences between traffic and fluid flows. Indeed the model is designed to simulate the anisotropic traffic behaviour, taking into account that the speed of each car does not change instantaneously and drivers look only ahead, that is the condition of the road in front of them, and do not care about what happens behind.

In formulas we have:

$$\begin{cases} \partial_t \rho(x,t) + \partial_x(\rho(x,t)\ v(x,t)) = 0, \\ \partial_t v(x,t) + v(x,t)\partial_x v(x,t) = \frac{v_e(\rho(x,t)) - v(x,t)}{\tau} - \frac{P'(\rho(x,t))\partial_x \rho(x,t)}{\rho}, \end{cases} \tag{1.3.39}$$

where $\tau > 0$ stands for the reaction/relaxation time, $v_e(\rho) : [0, \rho_{max}] \to [0, v_{max}]$ is a C^1 decreasing function

which gives the equilibrium (i.e. desired) velocity, that means the speed drivers would to achieve, and $P(\rho)$ is an increasing function representing a pressure law:

$$P(\rho) := \begin{cases} \frac{V_{\text{ref}}}{\gamma}\left(\frac{\rho}{\rho_{\max}}\right)^{\gamma}, & \gamma > 0 \\ V_{\text{ref}} \ln\left(\frac{\rho}{\rho_{\max}}\right), & \gamma = 0. \end{cases} \tag{1.3.40}$$

From the mathematical point of view the second order model (1.3.39) is a purely nonlinear hyperbolic system of two conservation laws, indeed we can rewrite (1.3.39) as follows:

$$\begin{cases} \partial_t \rho(x,t) + \partial_x(\rho(x,t)\ v(x,t)) = 0, \\ \partial_t[\rho(x,t)w(x,t)] + \partial_x[v(x,t)\rho(x,t)w(x,t)] = \rho(x,t)\frac{v_e(\rho(x,t))-v(x,t)}{\tau}, \end{cases} \tag{1.3.41}$$

with $w(x,t) = v(x,t) + P(\rho(x,t))$, also called the Lagrangian marker.

Let us forget for a while the relaxation term and focus our attention on the Aw Rascle model:

$$\begin{cases} \partial_t \rho(x,t) + \partial_x(\rho(x,t)\ v(x,t)) = 0, \\ \partial_t[\rho(x,t)w(x,t)] + \partial_x[v(x,t)\rho(x,t)w(x,t)] = 0, \end{cases} \tag{1.3.42}$$

In order to point out the main features of (1.3.42), let us compute the eigenvalues of the Jacobian matrix associated to the system:

$$\lambda_1 = v - \rho P'(\rho) \qquad \lambda_2 = v. \tag{1.3.43}$$

Therefore the system (1.3.42) is hyperbolic except for $\rho = 0$ when the two eigenvalues coalesce and the jacobian matrix is no longer diagonalizable.

Computing the corresponding right eigenvectors, we find out that $r^1 = (1, -P'(\rho))^T$ and $r^2 = (1,0)^T$. Reminding Sec. 1.1.2, we can say that λ_2 is always linearly degenerate, and λ_1 would be linearly degenerate if and only if $P(\rho) = A - \frac{B}{\rho}$, for $A, B > 0$ real constants. Under our assumption for $P(\rho)$, (1.3.40), λ_1 is always genuinely non linear, and therefore will admit either shock waves or rarefactions, while the second one admits contact discontinuities that correspond to waves whose propagation speed is always equal to the corresponding eigenvalue.
Note that $\lambda_1 \leq \lambda_2$ so all the waves propagate at a speed at most equal to the velocity v of the corresponding state.

Numerical approximation

The conservation law (1.3.33) will be numerically solve by means of the classical Godunov scheme (1.2.29) for the LWR model as described in [40].
Moreover we can solve the system (1.3.33) applying the Staggered Lax-Friedrichs method presented above (1.2.24).

1.3.2 Microscopic approach

On the other hand, the natural way of describing traffic evolution is tracking each single vehicle and following its trajectory. In order to do this, assume that N vehicles are moving along a single-lane

infinite road where overtaking is not possible. This means that cars are ordered and have to follow the first one that is called the *leader*. Cars that are not the leader are termed *followers*.

Let $X_k(t)$ be the position of the k-th car at time $t>0$ and $V_k(t)$ its instantaneous velocity, for $k=0,\dots,N$. The dynamics is governed by the distance between adjacent vehicles and described by a system of ordinary differential equations (ODEs) in the following way:

$$\begin{cases} \dot{X}_k(t) = V(X_k(t), X_{k+1}(t)), & k<N \\ \dot{X}_N(t) = \bar{V} \end{cases} \tag{1.3.44}$$

where $\dot{X}_k(t)$ stands for the derivative in time of the k-th vehicle's position.
Note that the leader needs a special dynamics since he has no one in front of him, so he/she moves at a constant velocity $\bar{V} \geq 0$.

Moreover if an accurate description is required, we have to consider also the acceleration term that will depend on the positions and the velocities of the nearby vehicles:

$$\begin{cases} \dot{X}_k(t) = V_k(t), & k \leq N \\ \dot{V}_k(t) = A(X_k(t), X_{k+1}(t), V_k(t), V_{k+1}(t)), & k<N \\ \dot{V}_N(t) = 0. \end{cases} \tag{1.3.45}$$

The systems (1.3.44) and (1.3.45) are the typical form respectively of a first order and a second order *Follow-the-Leader* type models, also known as *Car-Following* models. In this framework the difference between first and second order models lies in the acceleration term: the former kind of models represents a simplification of the reality since assume that accelerations are instantaneous, and the traffic conditions are always at equilibrium. On the other hand second order models are closer to the real dynamics since consider bounded accelerations.

There are many possible choices for the velocity and the acceleration terms, that identify which model we are considering, [3]. Let us focus first on the velocity term. Starting from real observations, the velocity function has to be 0 if the cars are nearer than a safe distance or the length of a vehicle called ℓ_N, so nobody can move in order to avoid collisions. On the other hand if cars are too far for influencing themselves they can travel at the maximum velocity allowed, $V_{\max}$. Thus we need an increasing function depending on the distance between nearby vehicles that we will denote with V^{des}. One of the most used velocity function is the following:

$$V^{des}(X_{k+1}, X_k) = \left(1 - \frac{\ell_N}{X_{k+1} - X_k}\right) \tag{1.3.46}$$

assuming that the distance between two vehicles is always greater than the length of a vehicle itself, i. e. $|X_{k+1} - X_k| \geq \ell_N, \forall k$.
The acceleration term is not so intuitive instead. There are a lot of parameters that we have to take into account such as the time-reaction of the drivers', the response of the vehicles' and so on. This is why there exist a lot of second order FtL models, depending on the choice for the acceleration term.

The most famous was proposed by Aw et al. in the early 2000s in [3], where the acceleration term is

defined as follows:

$$A\big(X_{k+1},X_k,V_{k+1},V_k\big)=C\frac{V_{k+1}-V_k}{(X_{k+1}-X_k)^{\gamma+1}}+\frac{1}{\tau}\left[v^*\left(\frac{\ell_N}{X_{k+1}-X_k}\right)-V_k\right] \tag{1.3.47}$$

where $\tau>0$ is modelling the reaction time, $\ell_N>0$ stands for the length of vehicles, $C,\gamma\geq 0$ are additional model parameters, and v^* is a C^1 decreasing function related to the macroscopic approach, which gives the equilibrium (i.e. desired) velocity of drivers as a function of the degree of congestion.

Numerical approximation

Systems (1.3.44) and (1.3.45) will be numerically integrated by means of the explicit Euler scheme.

1.3.3 From micro to macro

A natural question is now arising about the relationships between the two scales. Connections between microscopic agent-based and macroscopic fluid-dynamics traffic flow models are already well established. There are several works in the literature investigating the limit for first and second order models as [19, 30, 37, 38, 64] or applying the theory of mean field games to recover a new macroscopic model [15]. On the other hand links between the scales on networks are rarer, see [25].

The interest in pointing out and establishing connections between microscopic and macroscopic models comes from the willing of justify and validate the macroscopic model on the basis of the microscopic assumptions, which are easier to understand. For example the continuum hypothesis, see Sec. 1.3.1, is not immediately justifiable in this context, since the number of vehicles is typically lower then the number of molecules in a fluid, but it is accepted as a technical approximation thanks to the link with macroscopic quantities as measures for traffic dynamics.

Moreover the limit from the discrete to the continuous approaches could be seen as a theoretical analogue of the reconstruction of traffic conditions through high sampling data from GPS devices. Our aim is using this technique to recover a new macroscopic model able to reproduce the traffic dynamics. Let us now showing how the solution of the Follow-the-Leader model (1.3.44) tends to the entropic solution of LWR model (1.3.33).

After recalling the assumptions made in Section 1.3.2, i. e. N indistinguishable vehicles are moving on a single-lane infinite road where overtaking is not allowed, we denote by $\ell_N>0$ the vehicle's length (or mass) and by $\mathcal{M}$ the total length/mass of cars. It holds:

$$\ell_N=\frac{\mathcal{M}}{N} \tag{1.3.48}$$

that means $\mathcal{M}$ does not change for increasing number of cars, indeed their own length/mass is decreasing accordingly. In order to avoid collisions, we assume that $\delta_k(t):=X_{k+1}(t)-X_k(t)\geq\ell_N$, for $k=0,\dots,N-1$. Recalling the system (1.3.44), we are able to solve it by a recursion procedure. Starting from the leader, i. e. the Nth-vehicle, for which we obtain $X_N(t)=X_N+\int_0^t V(s)ds$. We recover the solution for $X_k(t)$, once knowing $X_{k+1}(t)$. The following result presented in [30], provides the non-degeneration of the quantities $\delta_k(t)$ for $k=0,\dots,N-1$ and $t\geq 0$, that ensures the existence and the uniqueness of the global solution once assuming that the microscopic velocity $V(\delta)$ is analogous to the macroscopic one $v(\rho)$, i. e. $V(\delta)=v(\frac{\ell_N}{\delta})$ and $v(\rho)=V(\frac{\ell_N}{\rho})$.

Lemma 1.3.1 (Discrete maximum principle). *For $k=1,\dots,N-1$, we have:*

$$\frac{\ell_N}{R} \le X_{k+1}(t) - X_k(t) \le X_N(0) - X_0(0) + (V_{\max} - v(R))t \qquad t \ge 0 \tag{1.3.49}$$

with $R = \max_{k=0,\dots,N-1}\left(\frac{\ell_N}{\delta_k}\right)$ the discrete density.

In order to formalize the connections between macro and microscopic descriptions we have to introduce the natural spaces for the density and the positions of cars:

$$\begin{aligned} D_{\mathcal{M}} &= \left\{\rho \in L^1(\mathbb{R};[0,1]) : \textstyle\int_{\mathbb{R}} \rho(x)\,dx = \mathcal{M}, \quad \text{with compact support}\right\} \\ C_N &= \left\{X \in \mathbb{R}^{n+1} : X_{k+1} - X_k \ge \ell_N, \quad \forall k = 0,\dots,N-1\right\} \end{aligned} \tag{1.3.50}$$

and the operators $E_N : C_N \to D_{\mathcal{M}}$ and $F_N : D_{\mathcal{M}} \to C_N$ as:

$$E_N[\rho] = X = \begin{cases} X_N = \max(supp(\rho)), \\ X_k = \max\{z \in \mathbb{R} : \int_z^{X_{k+1}} \rho(x)\,dx = \ell_N\}, \quad k = N-1,\dots,1 \end{cases} \tag{1.3.51}$$

$$F_N[X] = \rho_N = \sum_{k=0}^{N-1} \frac{\ell_N}{\delta_k} \chi_{[X_1, X_{k+1})}, \tag{1.3.52}$$

where χ_I stands for the characteristic function of the interval I.
The operator E_N gives the microscopic interpretation of the macroscopic density thanks to the vehicles' positions. On the other hand, the operator F_N plays the opposite role: given the discrete positions of cars it provides the corresponding piecewise constant density.

Proposition 1.3.2. *For every $N \in \mathbb{N}$ with $N \ge 2$ and for every $\mathcal{M} > 0$, $E_N \circ F_N = Id_{C_N}$. Viceversa, $F_N \circ E_N = Id_{D_{\mathcal{M}}}$ pointwise in $D_{\mathcal{M}} \cap BV(\mathbb{R};[0,1])$ with respect to the L^1-norm, so that*

- $\lim_{N\to\infty} ||(F_N \circ E_N)(\rho) - \rho||_{L^1} = 0$ *for all $\rho \in D_{\mathcal{M}} \cap BV(\mathbb{R};[0,1])$ and*
- $\lim_{N\to\infty}((F_N \circ E_N)(\rho))(x) = \rho(x)$ *for all $\rho \in D_{\mathcal{M}} \cap BV(\mathbb{R};[0,1])$, for a.e. $x \in \mathbb{R}$.*

We refer to [64] for the proof.

Moreover, the following two results ensure the well-posedness of the Cauchy problems (1.3.44) and (1.3.33).

Proposition 1.3.3. *After the assumptions made on v for the LWR model in Sec.1.3.1, for any $\mathcal{M} > 0$ and for any initial datum $\bar{\rho} \in D_{\mathcal{M}} \cap BV(\mathbb{R};[0,1])$, the Cauchy problem* (1.3.33) *admits a unique solution $\rho \in C^{0,1}(\mathbb{R}^+; D_{\mathcal{M}})$.*

Proposition 1.3.4. *After the assumptions made on v for the LWR model in Sec.1.3.1 and $V(\delta) = v(\frac{\ell_N}{\delta})$, for any $N \in \mathbb{N}$, with $N > 2$ and for any initial datum $\bar{X} \in C_N$, the Cauchy problem* (1.3.44) *admits a unique solution $X \in C^{0,1}(\mathbb{R}^+; C_N)$*

For the proofs we refer to [30].

1.3.4 Extension to the network

In order to describe more realistic situations where the vehicles move on several interconnected roads, the models presented above are not sufficient.

Let us deal with a connected and directed graph $\mathcal{G}=(\mathcal{E},\mathcal{J})$ consisting of a finite set of vertexes $\mathcal{J}$ and a set of oriented edges $\mathcal{E}$ connecting the vertexes. Starting from $\mathcal{G}$, we build the network $\mathcal{N}$ by assigning to each edge $\mathrm{E}\in\mathcal{E}$ a positive length $L_{\mathrm{E}}\in(0,+\infty)$. Moreover, a coordinate is assigned to each point of the edge. The coordinate will be denoted by x_{E} and it increases according to the direction of the edge, i.e. we have $x_{\mathrm{E}}=0$ at the initial vertex and $x_{\mathrm{E}}=L_{\mathrm{E}}$ at the terminal vertex. The mathematical definition is the following.

Definition 1.3.5. *A network $\mathcal{N}$ is a couple $(\mathcal{E},\mathcal{J})$ where:*

- *$\mathcal{E}$ is a finite collection of edges, which are intervals on $\mathbb{R}$, $I_i=[a_i,b_i]\subseteq\mathbb{R}$, with $i=1,..,N_e$;*
- *each edge has a system of local coordinates;*
- *$\mathcal{J}$ is a finite collection of vertices. Each vertex J is union of two non empty subsets $Inc(J)$ and $Out(J)$ of $\{1;\ldots;N_e\}$.*

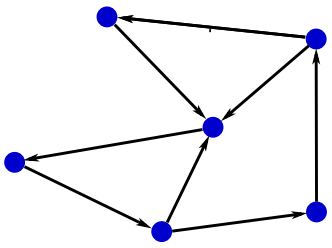

Figure 1.4: Network example.

In this context the edges represent the roads, the vertexes the junctions, $Inc(J)$ and $Out(J)$ are the sets of the indices of the incoming and outgoing roads, respectively, of a junction J. In order to determine in a unique way a junction we assume that:

1) For every $J\neq J'\in\mathcal{J}$ we have $Inc(J)\cap Inc(J')=\emptyset$ and $Out(J)\cap Out(J')=\emptyset$.
2) If $i\notin\cup_{J\in\mathcal{J}}Inc(J)$, then $b_i=+\infty$ and if $i\notin\cup_{J\in\mathcal{J}}Out(J)$ then $a_i=-\infty$.
   Moreover, the two cases are mutually exclusive.

The natural way to extend (1.3.33) to the network is to assume that the conservation law is satisfied on each arc for all times $t>0$ and consider separately the problem at the vertexes.

The first idea one could have is summing up the boundary conditions at the end/beginning of roads and going on solving (1.1.6), but it is not the right way to proceed. For example let us consider a junction with two incoming roads, e_1 and e_2, and one outgoing road, e_3. If we sum up the densities at the end of e_1 and e_2 it is easy that we may exceed ρ^{max} that would be the density value at the beginning of e_3, and it makes no sense. Thus the relevant physical condition is the conservation of fluxes at the junction, that means the number of the incoming vehicles in a unit of time has to be the same of the outgoing vehicles. So cars can not be created or deleted at junctions.

Moreover, additional conditions have to be imposed at junctions, because in general the conservation of the mass alone is not sufficient to characterize a unique solution.

Multiple approaches have been suggested in the literature for overcoming such ill-posedness: (i) maximization of the flux across junctions and introduction of priorities among the incoming roads [16, 40, 49]; (ii) introduction of a *buffer* to model the junctions by means of additional ODE coupled with (1.3.33) [8, 39, 41, 46]; (iii) reformulation of the problem on all possible paths on the network rather than on roads and junctions [9, 10, 11, 47]. In general, each of them allows to determine a unique solution for the traffic evolution on the network, but the solutions might be different.
Let us analyse in detail these approaches, pointing out their main features.

1.3.4.1 Classical approach

The classical approach introduced in [16, 40, 49] deals with the maximization of the flux at the junctions and the introduction of priorities on the incoming roads in order to select a unique solution and solve the ill-posedness of the problem.

Let us consider a generic junction J and denote by $\mathcal{E}_i=[a_i,b_i]$, for $i=1,\dots,n$ the incoming roads and by $\mathcal{E}_j=[a_j,b_j]$, for $j=n+1,\dots,n+m$ the outgoing roads.

Definition 1.3.6. *A traffic distribution matrix is given by*

$$A=\begin{pmatrix} \alpha_{n+1,1} & \dots & \alpha_{n+1,n} \\ \dots & \dots & \dots \\ \alpha_{n+m,1} & \dots & \alpha_{n+m,n} \end{pmatrix} \tag{1.3.53}$$

where $0\le\alpha_{j,i}\le 1$ *for every* $i\in\{1,\dots,n\}$ *and for every* $j\in\{n+1,\dots,n+m\}$ *and*

$$\sum_{j=n+1}^{n+m}\alpha_{j,i}=1, \tag{1.3.54}$$

for every $i=1,\dots,n$.

Roughly speaking, the $i-$th column of A describes how the cars coming from an incoming road i distributes in percentages to the outgoing roads. Note that each vertex has its own distribution matrix. In this framework we assume that the coefficient of A are constant but they could also depend on time, since the preferences of the drivers may change during the day.

A natural request for the densities on the incoming/outgoing roads at a generic junction, i.e. $(\rho^1,\dots,\rho^{n+m})$, in order to have an admissible solution in J is:

$$\sum_{i=1}^{n} f\left(\rho^i(b_i-,t)\right)=\sum_{j=n+1}^{n+m} f\left(\rho^j(a_j+,t)\right) \tag{1.3.55}$$

that means that the vehicles at the junction have to be conserved. Note that this assumption can be considered as a generalization of the Rankine-Hugoniot conditions at the junctions.

Let us define the maximum incoming flux and the maximum outgoing flux on each road, respectively, as follows:

$$\gamma^i_{max}(\rho(b_i,t))=\begin{cases} f(\rho(b_i,t)) & if\ \rho(b_i,t)\in[0,\sigma], \\ f(\sigma) & if\ \rho(b_i,t)\in(\sigma,\rho_{max}] \end{cases} \qquad i=1,\dots,n \tag{1.3.56}$$

and

$$\gamma^j_{max}(\rho(a_j,t)) = \begin{cases} f(\rho(a_j,t)) & if\ \rho(a_j,t) \in (\sigma,\rho_{max}], \\ f(\sigma) & if\ \rho(a_j,t) \in [0,\sigma] \end{cases} \qquad j = n+1,\dots,n+m. \tag{1.3.57}$$

Equations (1.3.56) and (1.3.57) are also known as *supply* and *demand* functions. Indeed the supply function measures the maximum flux that could outgoing from the incoming road, instead the demand function measures the maximum flux that an outgoing road could receive.

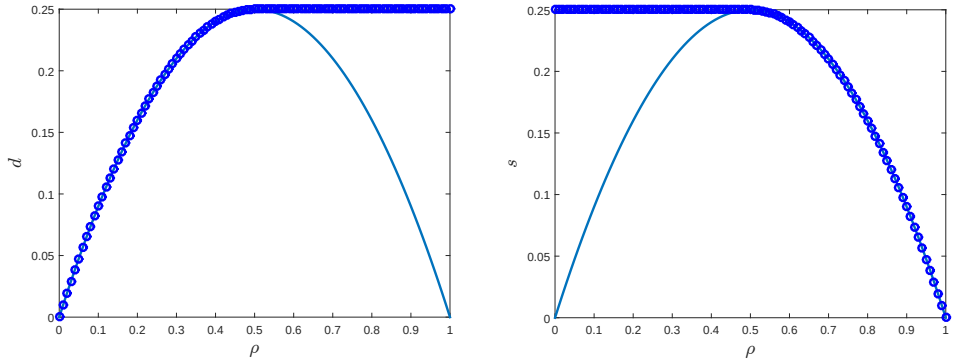

Figure 1.5: Example of demand (left) and supply (right) functions.

Let us denote:

$$\Omega_i = \left[0, \gamma^i_{max}(\rho(t,b_i))\right], \qquad i = 1,\dots,n,$$

$$\Omega_j = \left[0, \gamma^j_{max}(\rho(t,a_j))\right], \qquad j = n+1,\dots,n+m,$$

$$\Omega = \left\{ \left(\gamma^1,\dots,\gamma^n\right) \in \Omega_1 \times \dots \times \Omega_n \,|\, A\left(\gamma^1,\dots\gamma^n\right)^T \in \Omega_{n+1} \times \dots \times \Omega_{n+m} \right\}$$

where Ω_i and Ω_j are the sets of all the possible fluxes for the solution at the junctions, and Ω stands for the set of all the admissible fluxes at the end of the incoming roads, taking into account the preferences matrix A.

In order to maximize the flux at the junctions, we have to solve an optimization problem with linear constraints:

$$\max_{(\gamma^1,\dots,\gamma^n)\in\Omega} \sum_{i=1}^{n} \gamma^i. \tag{1.3.58}$$

On the other hand, (1.3.58) has not a unique solution in general. For example it can admit infinite solutions as shown in Fig. 1.6 where the admissible solutions are all the points belonging to the bold line, that represents the constraint $\gamma^1 + \gamma^2 \leq \gamma^3_{\max}$. This happens in the case of two incoming roads and one outgoing where the sum of the incoming densities have not to exceed the capacity of the outgoing road. So we have to impose another constraint that takes into account the priority, or the right of way of drivers denoted by the vector $\mathbf{q} = (q^1,\dots,q^n)$:

$$(\gamma^1,\dots,\gamma^n) \in \{\mathbf{q}s,\ s \in \mathbb{R}^+\}. \tag{1.3.59}$$

It is required that $q^i \geq 0\ \forall i = 1,\dots,n$ and $\sum_{i=1}^n q^i = 1$.

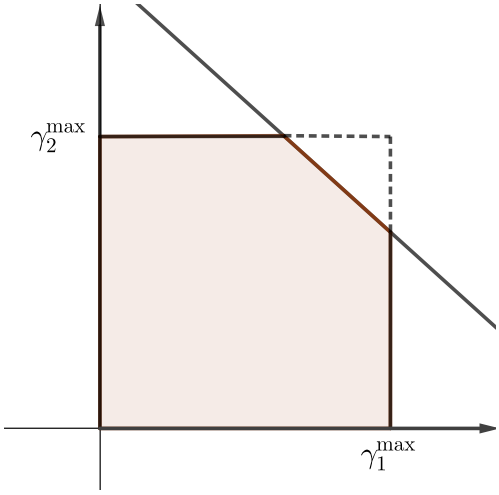

Figure 1.6: Example of infinite admissible solutions.

We are now able to guarantee the uniqueness of the solution.

Numerical approximation

We numerically solve the conservation law on each arc by means of the Godunov scheme (1.2.29) or applying the Staggered Lax Friedrichs scheme (1.2.24). On the other hand, at the junction we employ the supply and demand functions (1.3.56) (1.3.57) at the last cell of the outgoing road and the first cell of the incoming road.

1.3.4.2 Junction with buffer

An alternative way to describe the dynamics of a junction is to represent it with a buffer, see [39]. Since sometimes the geometry of a junction has no negligible effects on traffic conditions, we are not allowed to assume that a junction is a single point with no dynamics. The main idea behind modelling a junction with a buffer is that we would recover the microscopic description of vehicles' dynamics', by considering the fact that cars do not pass immediately from a road to another but spend some time crossing the junction, but, on the other hand, we would also maintain the macroscopic description of the problem. Even if the junction remains a 0-dimensional object, it has now its own dynamics.

Indeed the buffer allows some storage capacity, and mathematically, its evolution is described by an ODE that takes into account how, the total number of cars at the junction, change in time.

Let us call $r(t)$ the number of vehicles we can find in the buffer at time t. It can goes from 0 to r^{max}, that means the maximum capacity of the junction we are considering. Its dynamics is described by:

$$\dot{r}(t) = \sum_{i \in Inc(J)} f(\rho_i(b_i,t)) - \sum_{j \in Out(J)} f(\rho_j(a_j,t)) \tag{1.3.60}$$

and it is nothing else that the sum of the incoming vehicles less the sum of the outgoing ones at the same time t.

The last step we have to focus on is how the incoming and the outgoing roads interface with the

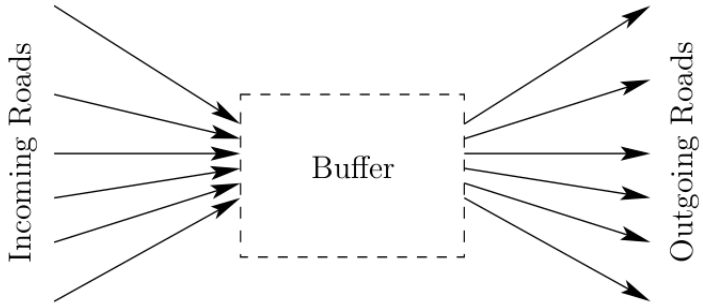

Figure 1.7: Buffer approach.

buffer. Let us introduce the supply and demand functions for the buffer, denoted respectively by $d_{\mathcal{B}}$ and $s_{\mathcal{B}}$. They are the functions representing the maximization of the flux for the incoming and the outgoing roads, as (1.3.56) and (1.3.57), see Fig 1.5. In other words we have to maximize the number of the incoming and outgoing vehicles at the buffer. Mathematically we have:

$$d_{\mathcal{B}} = \begin{cases} \beta & 0 < r \leq r^{max} \\ \min\{\gamma^i_{max}(\rho(b^i)), \beta\} & r(t) = 0 \end{cases} \tag{1.3.61}$$

$$s_{\mathcal{B}} = \begin{cases} \beta & 0 \leq r < r^{max} \\ \min\{\beta, \gamma^j_{max}(\rho(a^j))\} & r(t) = r^{max}, \end{cases} \tag{1.3.62}$$

where β is a constant parameter of the model. We can recover the incoming and the outgoing fluxes as follows:

$$f^{inc} = \min\{\gamma^i_{max}(\rho(b^i)), s_\beta\} \qquad f^{out} = \min\{d_\beta, \gamma^j_{max}(\rho(a^j))\}. \tag{1.3.63}$$

1.3.4.3 Multi-path approach

The main idea underlying this approach is considering the network not like a set of vertices and edges but as a set of all the possible overlapping paths joining all possible sources with all possible destinations. Indeed drivers are divided on the basis of their path. The dynamics has to be reformulated on the paths so the multi-path approach is based on a system of nonlinear conservation law with discontinuous flux.

Let us assume that the number of all the possible paths on the network is N_p and let us denote them by $P^1, P^2, \ldots, P^{N_p}$.

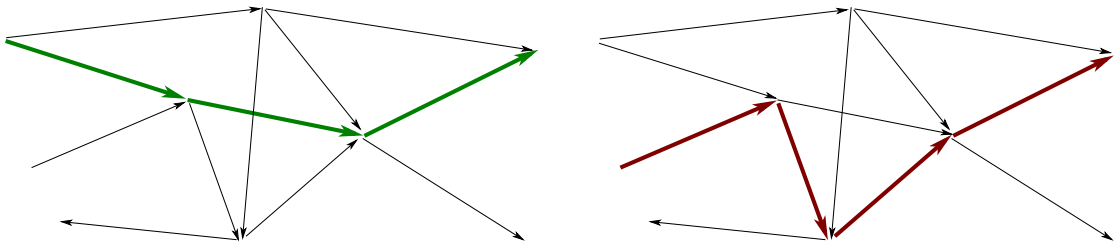

Figure 1.8: A generic network where two possible paths are highlighted.

Thus a point $x^{(p)}$ on the network is characterized by both the path it belongs to and the distance x from

the origin of that path.

Rather than tracking the total density $\rho(x,t)$, in this framework is more convenient to study the evolution of the so called sub-densities $\mu^p(x^{(q)},t)$, that is the density of the vehicles following the p-th path at point $x^{(q)}$ at time t. Obviously $\mu^p(x^{(q)},t)\in[0,\rho_{\max}]$ as well as ρ. Note that from the definition we have $\mu^p(x^{(q)},t)=0$ if $x^{(q)}\notin P^p$.

The total density $\rho(x,t)$ is recovered by summing up all the sub-densities at any point x:

$$\rho(x^{(p)},t)=\sum_{q=1}^{N_p}\mu^q(x^{(p)},t) \tag{1.3.64}$$

Clearly the dynamics of vehicles that follow the pth-path is influenced by the vehicles that are following the other paths because the same arc could belong to different paths. So we can not forget to compute the total density too, since the velocity depends on ρ.
Moreover $\rho(x,t)$ takes into account the whole network topology, as suggested in (1.3.64), indeed adding or deleting a road or modify a junction become easier [9, 10].

The LWR model (1.3.33) can be rewritten in this framework as a system of N_P conservation laws with discontinuous flux:

$$\frac{\partial}{\partial t}\mu^p(x^{(p)},t)+\frac{\partial}{\partial x^{(p)}}\left(\mu^p(x^{(p)},t)v(\rho(x^{(p)},t))\right)=0 \qquad x^{(p)}\in P^p,\ t>0,\ p=1,\dots,N_P, \tag{1.3.65}$$

where, for avoiding any confusion with the indices, we change the notation for the partial derivative: $\frac{\partial}{\partial t}=\partial_t$. Multiplying and dividing by the total density the second term inside the brackets, we obtain the typical form of LWR:

$$\frac{\partial}{\partial t}\mu^p(x^{(p)},t)+\frac{\partial}{\partial x^{(p)}}\left(\frac{\mu^p(x^{(p)},t)}{\rho(x^{(p)},t)}f(\rho(x^{(p)},t))\right)=0 \qquad x^{(p)}\in P^p,\ t>0,\ p=1,\dots,N_P. \tag{1.3.66}$$

Remark 1.3.7. *If $\rho(x,t)=0$, it implies that $\mu(x,t)=0$ too, so it is convenient assume $\frac{\mu^p(x,t)}{\rho(x,t)}=0$ to avoid any unphysical singularity.*

Note that the conservation laws in the system (1.3.66) are coupled by means of the total density, but it may happen that some paths do not share any arc so we can find decoupled equations.

Numerical approximation

Let us denote with $\mu^{n,p}_{k^{(q)}}$ the discrete subdensity $\mu^p\left(x^{(q)}_k,t^n\right)$, where $k^{(q)}$ is the kth-node belonging to the path P^q. In order to avoid any confusions, we drop the brackets $k^{(q)}=k^q$.

The numerical counterpart of the Multipath approach [9] is an "hybrid" scheme obtained by applying first the Upwind scheme [55] and then the Godunov scheme (1.2.29) to (1.3.66). In particular, employing directly the Godunov scheme on (1.3.66), leads to consider the derivative of the coefficient $\frac{\mu_p}{\rho}$ too, since it depends on x. Because it is a little bit hard to compute, we apply first an Upwind backward approximation of (1.3.66) in order to obtain:

$$\frac{\mu^{n+1,p}_{k^p}-\mu^{n,p}_{k^p}}{\Delta t}+\frac{1}{\Delta x}\left[\frac{\mu^{n,p}_{k^p}}{\rho^{n,p}_{k^p}}f(\rho^n_{k^p})\frac{\mu^{n,p}_{k^p-1}}{\rho^{n,p}_{k^p-1}}f(\rho_{k^p-1})\right]=0. \tag{1.3.67}$$

After that, we are ready to employ the Godunov scheme:

$$\mu_{k^p}^{n+1,p} = \mu_{k^p}^{n,p} - \frac{\Delta t}{\Delta x}\left[\frac{\mu_{k^p}^{n,p}}{\rho_{k^p}^{n,p}} G(\rho_{k^p+1}^n, \rho_{k^p}^n) \frac{\mu_{k^p-1}^{n,p}}{\rho_{k^p-1}^{n,p}} G(\rho_{k^p}^n, \rho_{k^p-1}^n)\right], \tag{1.3.68}$$

for $n \geq 0$, $p = 1, \ldots, N^P$ and G stands for the Godunov flux, that is:

$$G(\rho^r, \rho^l) = \begin{cases} \min\{f(\rho^r), f(\rho^l)\} & \rho^l \leq \rho^r \\ f(\rho^l) & \rho^l > \rho^r, \rho^l < \sigma \\ f(\rho^r) & \rho^l > \rho^r, \rho^r > \sigma \\ f(\sigma) & \rho^l > \rho^r, \rho^r < \sigma < \rho^l. \end{cases} \tag{1.3.69}$$

Note that this scheme is intrinsically asymmetric since the coefficient in front of the fluxes involve only the node k^p and $k^p - 1$, and not $k^p + 1$.

1.4 Wasserstein distance

It is well known that several problems in traffic modelling require the comparison of two density functions representing traffic conditions. For example it is necessary for the theoretical study of the property of the solution of scalar conservation laws or the study of the convergence of numerical schemes, indeed we have to verify (and quantify) that the numerical solution is close to the exact solution. Moreover the comparison between two densities plays a crucial role also in more application fields such as the calibration, that is to find the values of the parameters for the predicted outputs to be as close as possible to the observed ones, and the validation of models, to check if the outputs are close to the observed ones, as well as the sensitivity analysis of the model itself, in which we have to quantify how the uncertainty in the outputs can be apportioned to different sources of uncertainty in the inputs and/or model's parameters.

Our aim is indeed to quantify the difference (or the closeness) between different scenarios and the main ingredient we will need is the Wasserstein distance.
In order to understand why this notion of distance is more natural than L^p-distances in this framework, let us begin this section with a meaningful example.

Let us consider three different density functions ρ_i for $i = 1,2,3$, respectively the blue, the red and the green line in Fig. 1.9, corresponding to the same total mass $\mathcal{M}$. The L^1-distance between ρ_1 and ρ_2 is equal to:

$$\int_{\mathbb{R}} |\rho_1(x) - \rho_2(x)| = \int_{supp(\rho_1) \cup supp(\rho_2)} |\rho_1(x) - \rho_2(x)| = 2\mathcal{M}.$$

We recover the same result by computing the distance between ρ_1 and ρ_3 that is equal to $2\mathcal{M}$ as well. Thus we can conclude that all the L^p-distances are blind with respect to variation of the densities once the supports of them are disjoint. On the other hand our perception of distance suggests that $||\rho_3 - \rho_1|| > ||\rho_2 - \rho_1||$ and this is exactly what Wasserstein distance guarantees.

Let us now introduce the Wasserstein distance formally.
Let us denote by $(X, \mathfrak{D})$ a complete and separable metric space with distance $\mathfrak{D}$, and by $\mathcal{B}(X)$ a Borel σ-algebra of $(X, \mathfrak{D})$. Let us also denote by $\mathcal{M}^+(X)$ the set of non-negative finite Radon measures on $(X, \mathcal{B}(X))$. Let ν^{S} (S standing for *supply*) and ν^{D} (D standing for *demand*) be two Radon measures in $\mathcal{M}^+(X)$ such that $\nu^{\mathrm{S}}(X) = \nu^{\mathrm{D}}(X)$.

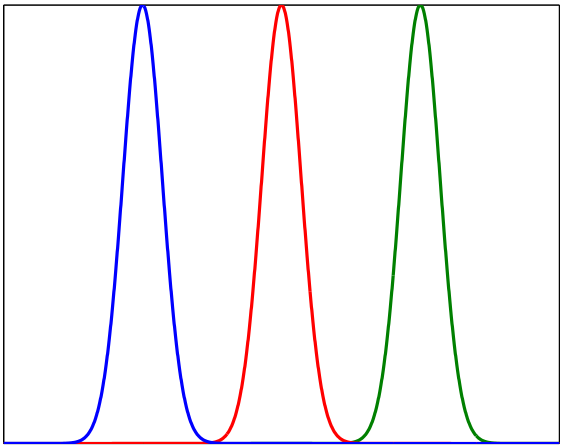

Figure 1.9: Example of different density functions with disjoint supports.

Definition 1.4.1 (Wasserstein distance)**.** *For any $q \in [1,+\infty)$, the L^q-Wasserstein distance between ν^{S} and ν^{D} is*

$$W_q(\nu^{\mathrm{S}},\nu^{\mathrm{D}}) := \left(\inf_{\gamma \in \Gamma(\nu^{\mathrm{S}},\nu^{\mathrm{D}})} \int_{X\times X} \mathfrak{D}(x,y)^q \, d\gamma(x,y) \right)^{1/q} \tag{1.4.70}$$

where

$$\Gamma(\nu^{\mathrm{S}},\nu^{\mathrm{D}}) := \left\{ \gamma \in \mathcal{M}^+(X\times X) \ s.t. \ \gamma(A\times X) = \nu^{\mathrm{S}}(A),\ \gamma(X\times B) = \nu^{\mathrm{D}}(B),\ \forall\ A,B \subset X \right\}.$$

Assuming that the measures $\nu^{\{\mathrm{S},\mathrm{D}\}}$ are absolutely continuous with respect to the Lebesgue measure, i.e. there are two density functions $\rho^{\{\mathrm{S},\mathrm{D}\}}$ such that $d\nu^{\{\mathrm{S},\mathrm{D}\}} = \rho^{\{\mathrm{S},\mathrm{D}\}} dx$, and considering the particular case $X = \mathbb{R}^n$, $\mathfrak{D}(x,y) = \|x-y\|_{\mathbb{R}^n}$, we have

$$W_q(\nu^{\mathrm{S}},\nu^{\mathrm{D}}) = W_q(\rho^{\mathrm{S}},\rho^{\mathrm{D}}) = \left(\inf_{T\in\mathcal{T}} \int_{\mathbb{R}^n} \|T(x)-x\|^q_{\mathbb{R}^n} \ \rho^{\mathrm{S}}(x)dx \right)^{1/q} \tag{1.4.71}$$

where

$$\mathcal{T} := \left\{ T : \mathbb{R}^n \to \mathbb{R}^n \text{ s.t. } \int_A \rho^{\mathrm{D}}(x)dx = \int_{\{x:T(x)\in A\}} \rho^{\mathrm{S}}(x)dx, \quad \forall A \subset \mathbb{R}^n \text{ bounded} \right\}.$$

To better understand the abstract definition of Wasserstein distance, we highlight the strict relation with the famous optimal mass transportation problem pointed out by Santambrogio in [65] and Villani [70, 71]. The idea is that a sand of soil or a sandpile with mass distribution ρ^{S}, has to be moved to an excavation with the same total volume and mass. The cost for moving mass depends on both the distance from the point of origin to the destination point, and the amount of mass that has to be moved along that path. We are looking for minimizing the total cost of the mass rearrangement, i. e. finding the optimal path to transport the mass from the initial to the final configuration.

In other words in (1.4.70) ν^{S} plays the role of the density of the mass we have to move, ν^{D} is the density that ν^{S} has to reach, i. e. the excavator, $d\gamma(x,y)$ denotes the amount of mass moving from two

points and $\mathfrak{D}(x,y)$ is the cost we have to pay, i. e. the distance between x and y.

More sophisticated characterizations based on a fluid-dynamic approach [5] , p-Laplacian [57], or a variational approach [57] are also available.

Remark 1.4.2. *In our framework, the mass to be moved corresponds to that of vehicles. We therefore measure the distance between two LWR solutions by computing the minimal cost to move vehicles from the scenario corresponding to one density distribution to the scenario corresponding to the other density distribution. Moreover, we assume that the mass transfer* is constrained to happen along the network $\mathcal{N}$ *(i.e. $X=\mathcal{N}$), but the transfer does not need to respect usual road laws (road direction, traffic distribution at junctions, etc.). This is reasonable since the measure of the distance between densities is conceptually different from the physical motion of vehicles.*

1.4.1 Optimal transportation problem

Let us briefly recall the optimal transport problem, also known as Monge-Kantorovich problem, and its main properties.

Definition 1.4.3 (Monge-Kantorovich problem)**.** *Fix two positive Radon measures ν^1 and ν^2, $\nu^1,\nu^2\in\mathcal{M}^+(X)$ satisfying the mass balance condition:*

$$\nu^1(X)=\nu^2(X). \tag{1.4.72}$$

The Monge-Kantorovich problem is the minimization problem:

$$\min_{\gamma\in\Gamma(\nu^1,\nu^2)}\int_{X\times X}\mathfrak{D}(x,y)^q\ d\gamma(x,y) \tag{1.4.73}$$

where $\Gamma(\nu^1,\nu^2)$ was defined in Def.1.4.1. The elements $\gamma\in\Gamma(\nu^1,\nu^2)$ are called transport plans between ν^1 and ν^2, and a minimizer γ^* is an optimal transport plan.

Let us introduce the main tools used for proving the existence of the solution to (1.4.73).

Definition 1.4.4 (Lower semi-continuity)**.** *On a metric space X, a function $g:\mathbb{R}\to\mathbb{R}^+$ is said to be lower semi continuous if for every sequence $x_n\to x$ we have $g(x)\leq\liminf_n f(x_n)$.*

Theorem 1.4.5 (Weierstrass)**.** *If $g:X\to\mathbb{R}^+$ is lower semi-continuous and X is compact, then there exists $\bar{x}\in X$ such that $g(\bar{x})=\min_{x\in X}g(x)$.*

Now we are ready to present the existence result proved in [65].

Theorem 1.4.6. *Let X be a compact metric space, $\nu^1,\nu^2\in\mathcal{M}^+(X)$ and $\mathfrak{D}(x,y):X\times X\to\mathbb{R}$ be lower semi-continuous and bounded from below. Then the Monge-Kantorovich problem admits a solution.*

Moreover, the Monge-Kantorovich problem has a dual formulation that is stated in the following theorem.

Theorem 1.4.7 (Duality formulation)**.** *Let $\nu^1,\nu^2\in\mathcal{M}^+(X)$ be two measures satisfying the mass balance condition. Then,*

$$\min_{\gamma\in\Gamma(\nu^1,\nu^2)}\int_{X\times X}\mathfrak{D}(x,y)^q d\gamma(x,y)=\sup_{z\in K_{\mathfrak{D}}(X)}\int_X z d(\nu^1-\nu^2) \tag{1.4.74}$$

where $K_{\mathfrak{D}}(X) = \{z : X \to \mathbb{R} : |z(x) - z(y)| \leq \mathfrak{D}(x,y)^q, \forall x, y \in X\}$.
Moreover, there exists $w \in K_{\mathfrak{D}}(X)$ *such that:*

$$\int_X w d(\nu^1 - \nu^2) = \sup_{z \in K_{\mathfrak{D}}(X)} \int_X z d(\nu^1 - \nu^2). \tag{1.4.75}$$

Such maximizers are called Kantorovich potentials.

The duality formulation and the link with the optimal transport problem help us to numerically compute the Wasserstein distance on networks. The successful result will be presented in the next chapter.

Chapter 2

Sensitivity analysis for the LWR model on networks

This chapter is devoted to quantifying the sensitivity of the LWR model on network to its parameters and to the network itself, in order to point out what influences most the solution of the LWR model. To do that, a numerical approximation of Wasserstein distance is required, since we need a suitable tool to quantify the difference between various scenarios. We will start from the network discretization rewriting the problem on the corresponding graph, and then we will develop the algorithm to compute Wasserstein distance using the link with optimal transport problems, see Sec. 1.4.1.

Moreover, the management of the LWR model on large networks is not so trivial also with the multi-path approach presented in Sec. 1.3.4.3. Since the number of paths increases exponentially as the network grows, the LWR solutions will be computed with a local version of the multi-path approach: indeed the sub-densities appear only in the cells next to the junctions, so they are not defined for all the possible paths on the whole network. In this way, the system of PDEs we have to solve remains feasible, even if each junction has its own system.

2.1 Approximation of Wasserstein distance on graph.

The aim of this section is the numerical approximation of Wasserstein distance on networks. To do that we have firstly to discretize the problem.

A network can be always approximated by a discrete graph at the cost of a loss of resolution.
For numerical purposes, the network has to be discretized by means of a grid. To avoid technicalities, let us assume that the length of all the edges is a multiple of the space step Δx, so we can use the same grid size everywhere in $\mathcal{N}$. In this way, each edge is divided in $N_{x_{\mathrm{E}}}$ cells, while the total number of cells on the network is given by N_x, in order to be consistent with the notations introduced in Sec. 1.2. The center of each cell on edge E will be denoted by $x_{\mathrm{E},j}$, $j=1,\dots,N_{x_{\mathrm{E}}}$ and the cell itself by $C_{\mathrm{E},j}=\left[x_{\mathrm{E},j-\frac{1}{2}},x_{\mathrm{E},j+\frac{1}{2}}\right)$. Now we can create a new, undirected graph whose vertexes coincide with the centres of the cells, and they are connected in agreement with the network, see Fig. 2.1. The graph created by this procedure will be denoted hereafter by $\mathcal{G}_{\mathcal{N}}^{\Delta}$. The number of vertices of $\mathcal{G}_{\mathcal{N}}^{\Delta}$ equals the total number of cells in $\mathcal{N}$, therefore will be denoted by J. In order to complete the discretization procedure, the mass distributed on each cell

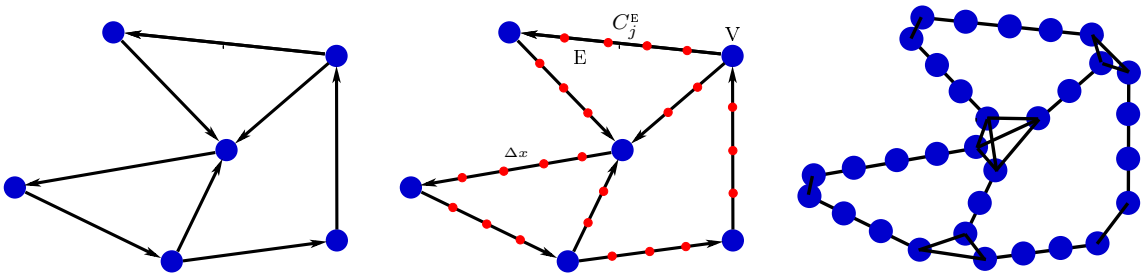

Figure 2.1: Original directed graph $\mathcal{G}$ (left), network $\mathcal{N}$ built on $\mathcal{G}$ discretized with space step Δx (center), and undirected graph $\mathcal{G}_{\mathcal{N}}^{\Delta}$ built on the discretized network $\mathcal{N}$ (right).

is accumulated to the vertex located at the center of the cell. Doing this, the problem is reformulated as an optimal mass transportation problem on a graph. The new problem clearly approximates the original one on the network $\mathcal{N}$ and the approximation error is controlled by Δx. Focusing on our particular case, after the numerical approximation of the LWR model, we are left with a single value $\rho_{\mathrm{E},j}$ for each cell $C_{\mathrm{E},j}$ of the network, which represents the average density in that cell, see (1.2.25). This means that the numerical procedure returns a constant density $\rho(x) \equiv \rho_{\mathrm{E},j}$ for all $x \in C_{\mathrm{E},j}$, which must be accumulated in the centre $x_{\mathrm{E},j}$ of the cell.

At this point one can resort to classical problems (see Hitchcock's paper [48]) and methods (see e.g., [65, Sec. 6.4.1] and [66, Chap. 19]), recasting the problem in the framework of linear programming (LP).

Let us enumerate the vertexes of $\mathcal{G}_{\mathcal{N}}^{\Delta}$ by $j = 1, \dots, J$, and denote by ρ_j^{s}, ρ_j^{d}, the supply and demand densities concentrated in vertex j, respectively. Following the mass transport interpretation, the supply mass at vertex j is $\mathrm{S}_j := \rho_j^{\mathrm{s}} \Delta x$, and the demand mass is $\mathrm{D}_j := \rho_j^{\mathrm{d}} \Delta x$. Let c_{jk} be the cost of shipping a unit quantity of mass from the origin $j \in \{1, \dots, J\}$ to the destination $k \in \{1, \dots, J\}$. Here we define c_{jk} as the length of the shortest path joining j and k on $\mathcal{G}_{\mathcal{N}}^{\Delta}$, which can be easily found by, e.g., the Dijkstra algorithm [31]. Let x_{jk} be the (unknown) quantity shipped from the origin j to the destination k. The problem is then formulated as

$$\begin{aligned} \text{minimize} \quad & \mathcal{H} := \sum_{j=1}^{J} \sum_{k=1}^{J} c_{jk} x_{jk} \\ \text{subject to} \quad & \sum_{k} x_{jk} = \mathrm{S}_j, \quad \forall j \\ & \sum_{j} x_{jk} = \mathrm{D}_k, \quad \forall k \\ & x_{jk} \geq 0. \end{aligned} \tag{2.1.1}$$

Note that the solution satisfies $x_{jk} \leq \min\{\mathrm{S}_j, \mathrm{D}_k\}$ since one cannot move more than S_j from any source vertex j and it is useless to bring more than D_k to any sink vertex k. From (2.1.1) it is easy to recover a standard LP problem

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^{\mathsf{T}} \mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0, \end{aligned} \tag{2.1.2}$$

simply defining

$$\mathbf{x} := (x_{11}, x_{12}, \ldots, x_{1J}, x_{21}, x_{22}, \ldots, x_{2J}, \ldots, x_{J1}, \ldots, x_{JJ})^{\intercal}$$
$$\mathbf{c} := (c_{11}, c_{12}, \ldots, c_{1J}, c_{21}, c_{22}, \ldots, c_{2J}, \ldots, c_{J1}, \ldots, c_{JJ})^{\intercal}$$
$$\mathbf{b} := (\mathrm{S}_1, \ldots, \mathrm{S}_J, \mathrm{D}_1, \ldots, \mathrm{D}_J)^{\intercal}$$

and $\mathbf{A}$ as the $2J \times J^2$ sparse matrix

$$\mathbf{A} := \begin{bmatrix} \mathbb{1}_J & 0 & 0 & \cdots & 0 \\ 0 & \mathbb{1}_J & 0 & \cdots & 0 \\ 0 & 0 & \mathbb{1}_J & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbb{1}_J \\ I_J & I_J & I_J & I_J & I_J \end{bmatrix} \tag{2.1.3}$$

where I_J is the $J \times J$ identity matrix and $\mathbb{1}_J := (\underbrace{1\ 1\ \cdots\ 1}_{J \text{ times}})$.

2.2 Error analysis of Wasserstein approximation

Let us now focus on quantifying the error introduced by the LP-based method presented above in computing the exact Wasserstein distance. We do that in the general case, without restricting ourselves to piecewise constant density functions.

Proposition 2.2.1. *Let $\rho^{\mathrm{S}}, \rho^{\mathrm{D}} : \mathcal{N} \to \mathbb{R}$ two densities defined on a network $\mathcal{N}$ such that*

$$M = \int_{\mathcal{N}} \rho^{\mathrm{S}} dx = \int_{\mathcal{N}} \rho^{\mathrm{D}} dx. \tag{2.2.4}$$

Then,

$$|W(\rho^{\mathrm{S}}, \rho^{\mathrm{D}}) - \mathcal{H}(\rho^{\mathrm{S}}, \rho^{\mathrm{D}})| \leq M \Delta x \tag{2.2.5}$$

where hereafter W denotes the Wasserstein distance W_1 and $\mathcal{H}$ is the solution of the problem (2.1.1).

Proof. To begin with, let us focus on a generic cell C_j of the network. In accordance with the optimal flow (found *a posteriori* as the solution of the optimal mass problem), the mass in C_j is transferred in one or more cells of the network. Let us denote by m_{jk} the mass which is moved from cell C_j to C_k for some $k = 1, \ldots, J$ (including $k = j$). Let us also denote by $\omega^{\mathrm{S}}_{jk}(\cdot)$ the density profile (with $\mathrm{supp}(\omega^{\mathrm{S}}_{jk}) \subseteq C_j$) associated to the leaving mass m_{jk} in C_j and by $\omega^{\mathrm{D}}_{jk}(\cdot)$ the density profile (with $\mathrm{supp}(\omega^{\mathrm{D}}_{jk}) \subseteq C_k$) associated to the arriving mass in C_k, see Fig. 2.2. By definition we have

$$m_{jk} = \int_{C_j} \omega^{\mathrm{S}}_{jk}(x) dx = \int_{C_k} \omega^{\mathrm{D}}_{jk}(x) dx.$$

Let us denote by $\underline{b}_j := x_{j-\frac{1}{2}}$, $\bar{b}_j := x_{j+\frac{1}{2}}$, and similarly by $\underline{b}_k$, $\bar{b}_k$, the two border points of the cell j and k, respectively.

By suitably accumulating the masses at the borders of the cells, and recalling that the discrete approach requires instead to accumulate the masses at the centers of the cells, we have:

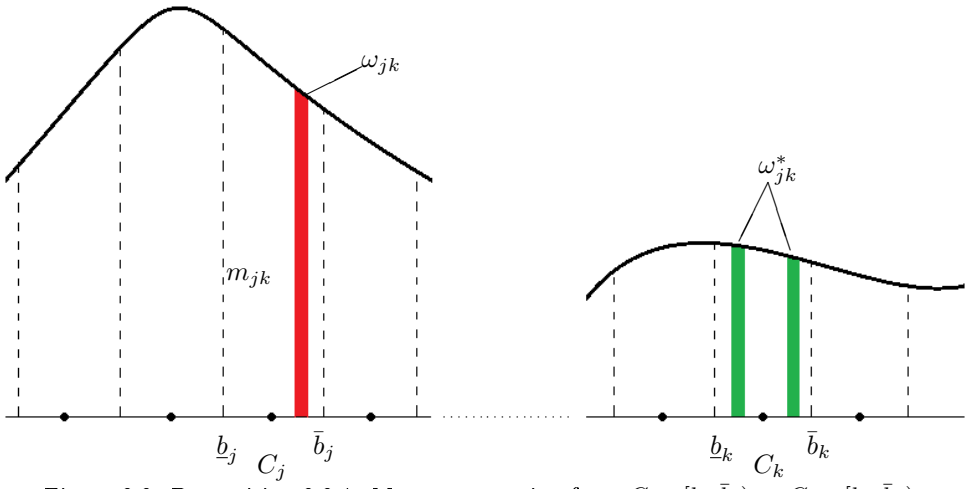

Figure 2.2: Proposition 2.2.1. Mass m_{jk} moving from $C_j=[\underline{b}_j,\bar{b}_j)$ to $C_k=[\underline{b}_k,\bar{b}_k)$.

Case A: $j\neq k$.

$$W(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})\leq m_{jk}\max_{\substack{b_j\in\{\bar{b}_j,\underline{b}_j\}\\ b_k\in\{\bar{b}_k,\underline{b}_k\}}} W(\delta_{b_j},\delta_{b_k})=\mathcal{H}(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})+2\frac{m_{jk}\Delta x}{2}, \tag{2.2.6}$$

and, equivalently,

$$W(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})\geq m_{jk}\min_{\substack{b_j\in\{\bar{b}_j,\underline{b}_j\}\\ b_k\in\{\bar{b}_k,\underline{b}_k\}}} W(\delta_{b_j},\delta_{b_k})=\mathcal{H}(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})-2\frac{m_{jk}\Delta x}{2} \tag{2.2.7}$$

(where the additional distance $\pm 2\frac{m_{jk}\Delta x}{2}$ comes from moving the mass from the borders to the centers of the cells in C_j and C_k).

Case B: $j=k$.

$$\mathcal{H}(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})=0 \qquad \text{and} \qquad 0\leq W(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})\leq m_{jk}\Delta x, \tag{2.2.8}$$

then we still have

$$W(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})\leq \mathcal{H}(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})+m_{jk}\Delta x \qquad \text{and} \qquad W(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})\geq \mathcal{H}(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})-m_{jk}\Delta x.$$

as in (2.2.6)-(2.2.7).

Summing up we obtain, by (2.1.1),

$$W(\rho^{\mathrm{S}},\rho^{\mathrm{D}})=\sum_j\sum_k W(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})\leq \sum_j\sum_k[\mathcal{H}(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})+m_{jk}\Delta x]=\mathcal{H}(\rho^{\mathrm{S}},\rho^{\mathrm{D}})+M\Delta x$$

and

$$W(\rho^{\mathrm{S}},\rho^{\mathrm{D}})=\sum_j\sum_k W(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})\geq \sum_j\sum_k[\mathcal{H}(\omega^{\mathrm{S}}_{jk},\omega^{\mathrm{D}}_{jk})-m_{jk}\Delta x]=\mathcal{H}(\rho^{\mathrm{S}},\rho^{\mathrm{D}})-M\Delta x.$$

Finally we have

$$|W(\rho^{\mathrm{S}},\rho^{\mathrm{D}})-\mathcal{H}(\rho^{\mathrm{S}},\rho^{\mathrm{D}})|\leq M\Delta x.$$

It is also easy to prove that this estimate is actually sharp. To see this it is sufficient to consider the one-dimensional case $\mathcal{N}=\mathbb{R}$ and choose $\rho^{\mathrm{S}}=M\delta_{\bar{b}_j}$ and $\rho^{\mathrm{D}}=M\delta_{\underline{b}_{j+1}}$. □

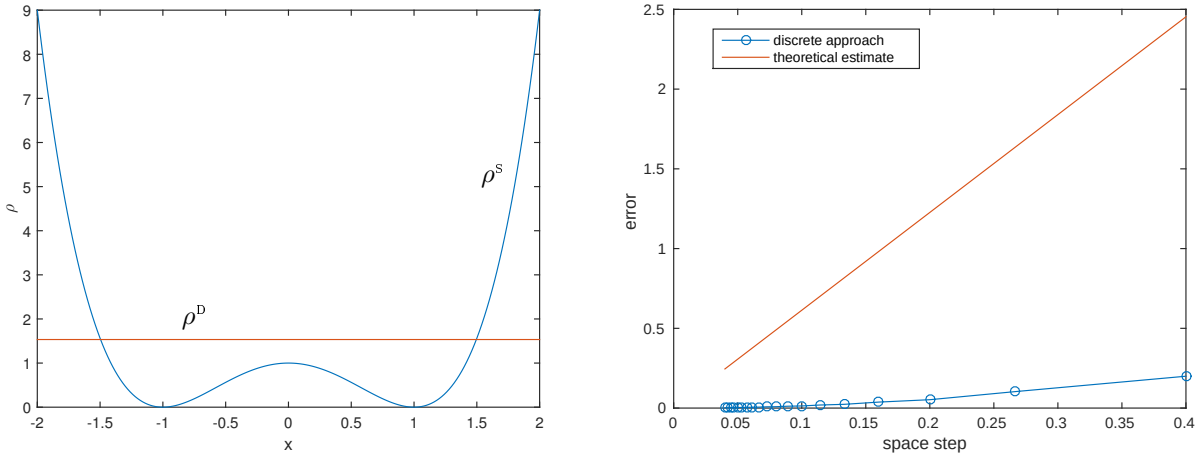

Figure 2.3: Exact vs. approximate Wasserstein distance. Functions ρ^{S}, ρ^{D} (left) and convergence of $|W-\mathcal{H}|$ as $\Delta x \to 0$ (right).

In the following we test the discrete approach described above against a one-dimensional problem where the Wasserstein distance can be analytically computed. We define

$$\rho^{\mathrm{S}}(x) = \begin{cases} x^4 - 2x^2 + 1, & x \in [-2,2] \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad \rho^{\mathrm{D}}(x) \equiv \frac{23}{15},$$

see Fig. 2.3(left). Note that the total mass is equal, i.e. $M = \int_{\mathbb{R}} \rho^{\mathrm{S}} = \int_{\mathbb{R}} \rho^{\mathrm{D}} = \frac{92}{15}$. The exact Wasserstein distance between the two densities can be easily computed by using this formula if $p=1$:

$$W_1(\rho^{\mathrm{S}}, \rho^{\mathrm{D}}) = \int_{\mathbb{R}} |F^{\mathrm{S}}(x) - F^{\mathrm{D}}(x)| dx, \qquad F^{\{\mathrm{S},\mathrm{D}\}}(x) := \int_{-\infty}^{x} \rho^{\{\mathrm{S},\mathrm{D}\}}(x) dx, \tag{2.2.9}$$

obtaining $W(\rho^{\mathrm{S}}, \rho^{\mathrm{D}}) = 3.2$. In Fig. 2.3(right) we report the value of the error $|W-\mathcal{H}|$ as a function of the space step Δx used to discretize the interval $[-2,2]$, and we compare it with the theoretical estimate given by Prop. 2.2.1. We note that in this special case the measured convergence rate is superlinear and the error is much lower than the theoretical estimate.

2.3 Local multi-path approach

In order to compute the solution of the LWR model on large networks following the multi-path approach presented in Sec. 1.3.4.3, the main issue we have to overcome is its unfeasibility given by the number of paths that grows exponentially with the size of the network, so the system becomes rapidly unmanageable.

The idea is then to divide vehicles on the basis of their path only next to the junctions. This means that in the cells surrounding the junctions we define "local" sub-densities that vanish right after. Far from the junctions, we consider only the global density so we avoid the increasing number of the possible routes.

We loose the global vision of the all paths on the whole network, in other words the total source-destination vision of the network, but we gain the feasibility of the problem since each junction has its own system of PDEs referring only to that junction itself (i.e. no more than 16 equations if we have 4 incoming and 4 outgoing roads).

Define as usual the unknown density at grid nodes as

$$\rho^n_{\mathrm{E},j} := \frac{1}{\Delta x}\int_{x_{\mathrm{E},j-\frac{1}{2}}}^{x_{\mathrm{E},j+\frac{1}{2}}} \rho(y, n\Delta t)dy, \qquad \mathrm{E}\in\mathcal{E}, \quad j=1,\dots,N_{x_\mathrm{E}}, \quad n=1,\dots,N_T. \tag{2.3.10}$$

Starting from the initial condition $\rho^0_{\mathrm{E},j} := \frac{1}{\Delta x}\int_{x_{\mathrm{E},j-\frac{1}{2}}}^{x_{\mathrm{E},j+\frac{1}{2}}} \rho^0(y)dy$, the approximate solution at any internal cell $j=2,\dots,N_{x_\mathrm{E}}-1$ of any edge $\mathrm{E}\in\mathcal{E}$ is easily found by the standard Godunov scheme:

$$\rho^{n+1}_{\mathrm{E},j} = \rho^n_{\mathrm{E},j} - \frac{\Delta t}{\Delta x}\Big(G(\rho^n_{\mathrm{E},j}, \rho^n_{\mathrm{E},j+1}) - G(\rho^n_{\mathrm{E},j-1}, \rho^n_{\mathrm{E},j})\Big), \quad n=0,\dots,N_T-1, \tag{2.3.11}$$

where the numerical flux G is defined as in (1.3.69).

Around vertexes we proceed as follows: let us focus on a generic vertex $\mathrm{V}\in\mathcal{J}$ and denote by $n^\mathrm{V}_{\mathrm{inc}}$ and $n^\mathrm{V}_{\mathrm{out}}$ the number of incoming and outgoing edges at V, respectively. As in the classical approach 1.3.4.1, we assume that it is given a traffic distribution matrix $A^\mathrm{V} = (\alpha^\mathrm{V}_{\mathrm{RR}'})$, $\mathrm{R}=1,\dots,n^\mathrm{V}_{\mathrm{inc}}$, $\mathrm{R}'=1,\dots,n^\mathrm{V}_{\mathrm{out}}$ which prescribes how the traffic distributes in percentage from any incoming edge R to any outgoing edge R′. Clearly $0\le\alpha^\mathrm{V}_{\mathrm{RR}'}\le 1\ \forall \mathrm{R},\mathrm{R}'$ and $\sum_{\mathrm{R}'}\alpha^\mathrm{V}_{\mathrm{RR}'}=1\ \forall\mathrm{R}$.

Now, following the multipath approach presented in 1.3.4.3 and keeping the same notations as above, we look at all possible paths across the vertex V. Having $n^\mathrm{V}_{\mathrm{inc}}$ incoming edges and $n^\mathrm{V}_{\mathrm{out}}$ outgoing edges, we have $n^\mathrm{V}_p := n^\mathrm{V}_{\mathrm{inc}}\times n^\mathrm{V}_{\mathrm{out}}$ possible paths.
Let us focus on a generic path $p=(\mathrm{E},\mathrm{E}')$ which joins edge E with E′, see Fig. 2.4. The problem is ready to

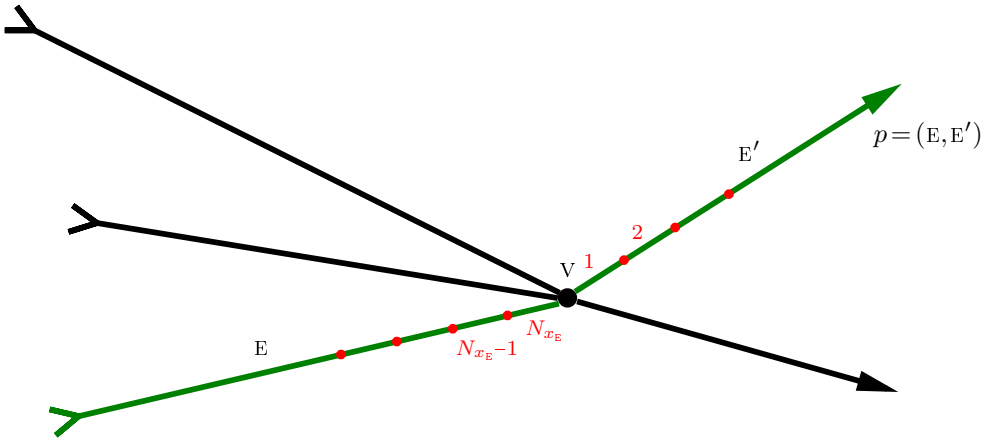

Figure 2.4: Zoom around vertex V. We show the path $p=(\mathrm{E},\mathrm{E}')$ and cells' labels on that path.

be solved by the Godunov-based multi-path scheme with minor modifications for all paths of all vertexes. Dropping the indexes p and V for readability, we have in the last cell of the incoming edge:

$$\mu^{n+1}_{\mathrm{E},J_\mathrm{E}} = \mu^n_{\mathrm{E},J_\mathrm{E}} - \frac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E},J_\mathrm{E}}}{\rho^n_{\mathrm{E},J_\mathrm{E}}} G\big(\rho^n_{\mathrm{E},J_\mathrm{E}}, \rho^n_{\mathrm{E}',1}\big) - \alpha^\mathrm{V}_{\mathrm{EE}'} G\big(\rho^n_{\mathrm{E},J_\mathrm{E}-1}, \rho^n_{\mathrm{E},J_\mathrm{E}}\big)\right), \tag{2.3.12}$$

for $n=0,\dots,N_T-1$. Note the presence of the parameter $\alpha^\mathrm{V}_{\mathrm{EE}'}$ in front of the incoming flux which tells that only a percentage of the total mass is following path p. In the first cell of the outgoing edge we have

instead:

$$\mu^{n+1}_{\mathrm{E}',1} = \mu^n_{\mathrm{E}',1} - \frac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}',1}}{\rho^n_{\mathrm{E}',1}} G\big(\rho^n_{\mathrm{E}',1}, \rho^n_{\mathrm{E}',2}\big) - \frac{\mu^n_{\mathrm{E},J_{\mathrm{E}}}}{\rho^n_{\mathrm{E},J_{\mathrm{E}}}} G\big(\rho^n_{\mathrm{E},J_{\mathrm{E}}}, \rho^n_{\mathrm{E}',1}\big)\right), \tag{2.3.13}$$

for $n = 0, \dots, N_T - 1$. The algorithm is completed by summing, at any time step, the sub-densities μ's (where defined) to compute the total density ρ, to be used at the next time step in (2.3.11), (2.3.12), and (2.3.13). More precisely, we have

$$\rho^{n+1}_{\mathrm{E},J_{\mathrm{E}}} = \sum_{q=1}^{n_p^{\mathrm{V}}} \mu^{n+1}_{\mathrm{E},N_{x_{\mathrm{E}}}}(q,\mathrm{V}) \quad \text{and} \quad \rho^{n+1}_{\mathrm{E}',1} = \sum_{q=1}^{n_p^{\mathrm{V}}} \mu^{n+1}_{\mathrm{E}',1}(q,\mathrm{V}). \tag{2.3.14}$$

Note that (2.3.12) and (2.3.13) are systems of n_p^{V} equations, coupled via (2.3.14), that takes into account vehicles moving along paths other than $p = (\mathrm{E}, \mathrm{E}')$. The other total densities ρ appearing in (2.3.12) and (2.3.13) are instead given by (2.3.11).

For the sake of clarity and dissipate any doubt, we write explicitly the scheme in the case of simple junctions.

Example 2.3.1 (Diverging case)**.** *First of all, we consider the diverging case, where we have one incoming* (E_1) *and two outgoing roads* (E_1', E_2')*. The total densities on the edges are defined by*

$$\begin{aligned} \rho_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}} &:= \mu_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}((\mathrm{E}_1,\mathrm{E}_1')),\mathrm{V}) + \mu_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}((\mathrm{E}_1,\mathrm{E}_2')),\mathrm{V}), \\ \rho_{\mathrm{E}_1',1} &:= \mu_{\mathrm{E}_1',1}((\mathrm{E}_1,\mathrm{E}_1')),\mathrm{V}), \\ \rho_{\mathrm{E}_2',1} &:= \mu_{\mathrm{E}_2',1}((\mathrm{E}_2,\mathrm{E}_2')),\mathrm{V}). \end{aligned} \tag{2.3.15}$$

Slightly simplifying the notation, system (2.3.12) *is explicitly written as*

$$\begin{aligned} \mu^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_1') &= \mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_1') - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_1')}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}, \rho^n_{\mathrm{E}_1',1}\big) - \alpha^{\mathrm{V}}_{\mathrm{E}_1\mathrm{E}_1'} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}-1}, \rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}\big)\right), \\ \mu^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_2') &= \mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_2') - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_2')}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}, \rho^n_{\mathrm{E}_2',1}\big) - \alpha^{\mathrm{V}}_{\mathrm{E}_1\mathrm{E}_2'} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}-1}, \rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}\big)\right), \end{aligned}$$

while system (2.3.13) *is explicitly written as*

$$\begin{aligned} \mu^{n+1}_{\mathrm{E}_1',1}(\mathrm{E}_1,\mathrm{E}_1') &= \mu^n_{\mathrm{E}_1',1}(\mathrm{E}_1,\mathrm{E}_1') - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}_1',1}(\mathrm{E}_1,\mathrm{E}_1')}{\rho^n_{\mathrm{E}_1',1}} G\big(\rho^n_{\mathrm{E}_1',1}, \rho^n_{\mathrm{E}_1',2}\big) - \frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_1')}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}, \rho^n_{\mathrm{E}_1',1}\big)\right), \\ \mu^{n+1}_{\mathrm{E}_2',1}(\mathrm{E}_1,\mathrm{E}_2') &= \mu^n_{\mathrm{E}_2',1}(\mathrm{E}_1,\mathrm{E}_2') - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}_2',1}(\mathrm{E}_1,\mathrm{E}_2')}{\rho^n_{\mathrm{E}_2',1}} G\big(\rho^n_{\mathrm{E}_2',1}, \rho^n_{\mathrm{E}_2',2}\big) - \frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_2')}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}, \rho^n_{\mathrm{E}_2',1}\big)\right). \end{aligned}$$

To complete the computation, we sum the sub-densities. Recalling that we have $\alpha_{\mathrm{E}_1\mathrm{E}_1'} + \alpha_{\mathrm{E}_1\mathrm{E}_2'} = 1$*, we get*

$$\begin{aligned} \rho^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}} = \rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}} - \frac{\Delta t}{\Delta x}\Bigg(&\frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_1') G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}, \rho^n_{\mathrm{E}_1',1}\big) + \mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}_2') G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}, \rho^n_{\mathrm{E}_2',1}\big)}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} \\ &- G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}-1}, \rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}\big)\Bigg), \end{aligned}$$

$$\rho^{n+1}_{\mathrm{E}'_1,1}=\rho^{n}_{\mathrm{E}'_1,1}-\frac{\Delta t}{\Delta x}\left(G\big(\rho^{n}_{\mathrm{E}'_1,1},\rho^{n}_{\mathrm{E}'_1,2}\big)-\frac{\mu^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1)G\big(\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^{n}_{\mathrm{E}'_1,1}\big)}{\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}}\right).$$

Example 2.3.2 (Merging case)**.** *Let us consider the opposite case: a single vertex* V *with two incoming edges* $\mathrm{E}_1,\mathrm{E}_2$ *and one outgoing edge* E'_1*.*

The total densities on E_i*,* $i=1,2$*, and* E'_1*, are defined by*

$$\begin{aligned}\rho_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}&:=\mu_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}((\mathrm{E}_1,\mathrm{E}'_1)),\mathrm{V}),\\ \rho_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}&:=\mu_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}((\mathrm{E}_2,\mathrm{E}'_1)),\mathrm{V})\\ \rho_{\mathrm{E}'_1,1}&:=\mu_{\mathrm{E}'_1,1}((\mathrm{E}_1,\mathrm{E}'_1)),\mathrm{V})+\mu_{\mathrm{E}'_1,1}((\mathrm{E}_2,\mathrm{E}'_1)),\mathrm{V}).\end{aligned}\tag{2.3.16}$$

In this case, since $\rho_{\mathrm{E}_i,N_{x_{\mathrm{E}_1}}}$*,* $i=1,2$*, coincide with the relative subdensities, system* (2.3.12) *is explicitly written as*

$$\begin{aligned}\mu^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1)&=\mu^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1)-\tfrac{\Delta t}{\Delta x}\left(G\big(\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^{n}_{\mathrm{E}'_1,1}\big)-G\big(\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}-1},\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}\big)\right),\\ \mu^{n+1}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_1)&=\mu^{n}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_1)-\tfrac{\Delta t}{\Delta x}\left(G\big(\rho^{n}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}},\rho^{n}_{\mathrm{E}'_1,1}\big)-G\big(\rho^{n}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}-1},\rho^{n}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}\big)\right);\end{aligned}\tag{2.3.17}$$

while system (2.3.13) *is explicitly written as*

$$\begin{aligned}\mu^{n+1}_{\mathrm{E}'_1,1}(\mathrm{E}_1,\mathrm{E}'_1)&=\mu^{n}_{\mathrm{E}'_1,1}(\mathrm{E}_1,\mathrm{E}'_1)-\tfrac{\Delta t}{\Delta x}\left(\tfrac{\mu^{n}_{\mathrm{E}'_1,1}(\mathrm{E}_1,\mathrm{E}'_1)}{\rho^{n}_{\mathrm{E}'_1,1}}G\big(\rho^{n}_{\mathrm{E}'_1,1},\rho^{n}_{\mathrm{E}'_1,2}\big)-G\big(\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^{n}_{\mathrm{E}'_1,1}\big)\right),\\ \mu^{n+1}_{\mathrm{E}'_1,1}(\mathrm{E}_2,\mathrm{E}'_1)&=\mu^{n}_{\mathrm{E}'_1,1}(\mathrm{E}_2,\mathrm{E}'_1)-\tfrac{\Delta t}{\Delta x}\left(\tfrac{\mu^{n}_{\mathrm{E}'_1,1}(\mathrm{E}_2,\mathrm{E}'_1)}{\rho^{n}_{\mathrm{E}'_1,1}}G\big(\rho^{n}_{\mathrm{E}'_1,1},\rho^{n}_{\mathrm{E}'_1,2}\big)-G\big(\rho^{n}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}},\rho^{n}_{\mathrm{E}'_1,1}\big)\right),\end{aligned}\tag{2.3.18}$$

To complete the computation, we sum the sub-densities and we get

$$\rho^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}=\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}-\frac{\Delta t}{\Delta x}\left(G\big(\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^{n}_{\mathrm{E}'_1,1}\big)\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}-G\big(\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}-1},\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}\big)\right),$$

the analogous holds for $\rho^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_2}}}$*, and:*

$$\rho^{n+1}_{\mathrm{E}'_1,1}=\rho^{n}_{\mathrm{E}'_1,1}-\frac{\Delta t}{\Delta x}\left(G\big(\rho^{n}_{\mathrm{E}'_1,1},\rho^{n}_{\mathrm{E}'_1,2}\big)-\frac{\mu^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1)G\big(\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^{n}_{\mathrm{E}'_1,1}\big)}{\rho^{n}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}}-\frac{\mu^{n}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_1)G\big(\rho^{n}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}},\rho^{n}_{\mathrm{E}'_1,1}\big)}{\rho^{n}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}}\right).$$

Example 2.3.3. *Now, we consider the case of a single vertex* V *with two incoming edges* $\mathrm{E}_1,\mathrm{E}_2$ *and two outgoing edges* $\mathrm{E}'_1,\mathrm{E}'_2$*.*

The total densities on E_i *and* E'_i, $i=1,2$, *are defined by*

$$\begin{aligned}
\rho_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}} &:= \mu_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}((\mathrm{E}_1,\mathrm{E}'_1),\mathrm{V}) + \mu_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}((\mathrm{E}_1,\mathrm{E}'_2),\mathrm{V}),\\
\rho_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}} &:= \mu_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}((\mathrm{E}_2,\mathrm{E}'_1),\mathrm{V}) + \mu_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}((\mathrm{E}_2,\mathrm{E}'_2),\mathrm{V}),\\
\rho_{\mathrm{E}'_1,1} &:= \mu_{\mathrm{E}'_1,1}((\mathrm{E}_1,\mathrm{E}'_1),\mathrm{V}) + \mu_{\mathrm{E}'_1,1}((\mathrm{E}_2,\mathrm{E}'_1),\mathrm{V}),\\
\rho_{\mathrm{E}'_2,1} &:= \mu_{\mathrm{E}'_2,1}((\mathrm{E}_1,\mathrm{E}'_2),\mathrm{V}) + \mu_{\mathrm{E}'_2,1}((\mathrm{E}_2,\mathrm{E}'_2),\mathrm{V}).
\end{aligned} \tag{2.3.19}$$

Slightly simplifying the notation, system (2.3.12) *is explicitly written as*

$$\begin{aligned}
\mu^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1) &= \mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1) - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1)}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^n_{\mathrm{E}'_1,1}\big) - \alpha^{\mathrm{V}}_{\mathrm{E}_1\mathrm{E}'_1} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}-1},\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}\big)\right),\\
\mu^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_2) &= \mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_2) - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_2)}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^n_{\mathrm{E}'_2,1}\big) - \alpha^{\mathrm{V}}_{\mathrm{E}_1\mathrm{E}'_2} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}-1},\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}\big)\right),\\
\mu^{n+1}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_1) &= \mu^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_1) - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_1)}{\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}} G\big(\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}},\rho^n_{\mathrm{E}'_1,1}\big) - \alpha^{\mathrm{V}}_{\mathrm{E}_2\mathrm{E}'_1} G\big(\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}-1},\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}\big)\right),\\
\mu^{n+1}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_2) &= \mu^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_2) - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_2)}{\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}} G\big(\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}},\rho^n_{\mathrm{E}'_2,1}\big) - \alpha^{\mathrm{V}}_{\mathrm{E}_2\mathrm{E}'_2} G\big(\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}-1},\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}\big)\right);
\end{aligned}$$

while system (2.3.13) *is explicitly written as*

$$\begin{aligned}
\mu^{n+1}_{\mathrm{E}'_1,1}(\mathrm{E}_1,\mathrm{E}'_1) &= \mu^n_{\mathrm{E}'_1,1}(\mathrm{E}_1,\mathrm{E}'_1) - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}'_1,1}(\mathrm{E}_1,\mathrm{E}'_1)}{\rho^n_{\mathrm{E}'_1,1}} G\big(\rho^n_{\mathrm{E}'_1,1},\rho^n_{\mathrm{E}'_1,2}\big) - \frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1)}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^n_{\mathrm{E}'_1,1}\big)\right),\\
\mu^{n+1}_{\mathrm{E}'_1,1}(\mathrm{E}_2,\mathrm{E}'_1) &= \mu^n_{\mathrm{E}'_1,1}(\mathrm{E}_2,\mathrm{E}'_1) - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}'_1,1}(\mathrm{E}_2,\mathrm{E}'_1)}{\rho^n_{\mathrm{E}'_1,1}} G\big(\rho^n_{\mathrm{E}'_1,1},\rho^n_{\mathrm{E}'_1,2}\big) - \frac{\mu^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_1)}{\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}} G\big(\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}},\rho^n_{\mathrm{E}'_1,1}\big)\right),\\
\mu^{n+1}_{\mathrm{E}'_2,1}(\mathrm{E}_1,\mathrm{E}'_2) &= \mu^n_{\mathrm{E}'_2,1}(\mathrm{E}_1,\mathrm{E}'_2) - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}'_2,1}(\mathrm{E}_1,\mathrm{E}'_2)}{\rho^n_{\mathrm{E}'_2,1}} G\big(\rho^n_{\mathrm{E}'_2,1},\rho^n_{\mathrm{E}'_2,2}\big) - \frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_2)}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}} G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^n_{\mathrm{E}'_2,1}\big)\right),\\
\mu^{n+1}_{\mathrm{E}'_2,1}(\mathrm{E}_2,\mathrm{E}'_2) &= \mu^n_{\mathrm{E}'_2,1}(\mathrm{E}_2,\mathrm{E}'_2) - \tfrac{\Delta t}{\Delta x}\left(\frac{\mu^n_{\mathrm{E}'_2,1}(\mathrm{E}_2,\mathrm{E}'_2)}{\rho^n_{\mathrm{E}'_2,1}} G\big(\rho^n_{\mathrm{E}'_2,1},\rho^n_{\mathrm{E}'_2,2}\big) - \frac{\mu^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_2)}{\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}} G\big(\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}},\rho^n_{\mathrm{E}'_2,1}\big)\right).
\end{aligned}$$

To complete the computation, we sum the sub-densities. Recalling that we have $\alpha_{\mathrm{E}_1\mathrm{E}'_1}+\alpha_{\mathrm{E}_1\mathrm{E}'_2}=1$ *and* $\alpha_{\mathrm{E}_2\mathrm{E}'_1}+\alpha_{\mathrm{E}_2\mathrm{E}'_2}=1$, *we get*

$$\begin{aligned}
\rho^{n+1}_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}} = \rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}} - \frac{\Delta t}{\Delta x}\Bigg(&\frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1) G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^n_{\mathrm{E}'_1,1}\big) + \mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_2) G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^n_{\mathrm{E}'_2,1}\big)}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}}\\
&- G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}-1},\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}\big)\Bigg),
\end{aligned}$$

$$\begin{aligned}
\rho^{n+1}_{\mathrm{E}'_1,1} = \rho^n_{\mathrm{E}'_1,1} - \frac{\Delta t}{\Delta x}\Bigg(&G\big(\rho^n_{\mathrm{E}'_1,1},\rho^n_{\mathrm{E}'_1,2}\big) - \frac{\mu^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}(\mathrm{E}_1,\mathrm{E}'_1) G\big(\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}},\rho^n_{\mathrm{E}'_1,1}\big)}{\rho^n_{\mathrm{E}_1,N_{x_{\mathrm{E}_1}}}}\\
&- \frac{\mu^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}(\mathrm{E}_2,\mathrm{E}'_1) G\big(\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}},\rho^n_{\mathrm{E}'_1,1}\big)}{\rho^n_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}}\Bigg),
\end{aligned}$$

and analogous expressions for $\rho^{n+1}_{\mathrm{E}_2,N_{x_{\mathrm{E}_2}}}$ *and* $\rho^{n+1}_{\mathrm{E}'_2,1}$.

2.4 Sensitivity analysis.

In this section we employ the discrete approach described in Sec. 2.1 to perform a sensitivity analysis of the LWR model. To solve the LP problem we used the GLPK [1] free C library.

For numerical tests we consider the "Manhattan"-like two-way road network depicted in Fig. 2.5. This choice is motivated by the fact that it allows one to easily compare networks of different size. Given the number ℓ of junctions per side, we get $4\ell(\ell-1)$ roads and ℓ^2 junctions. Roads are numbered starting from those going rightward, then leftward, upward, and finally downward. The length of each road is $L_{\mathrm{E}}=1$ and, if not otherwise stated, $\Delta x=0.1$ ($N_{x_{\mathrm{E}}}=10$, $N_x=40\ell(\ell-1)$).

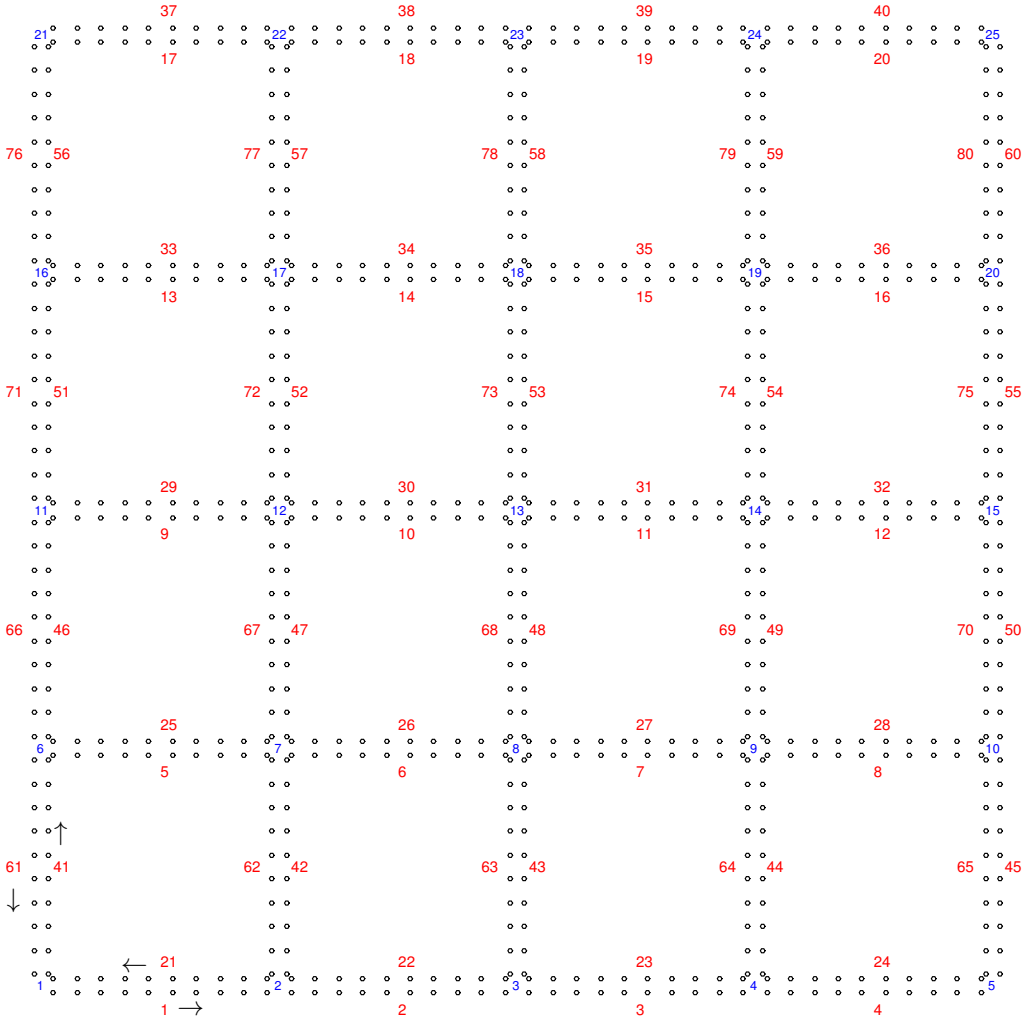

Figure 2.5: Manhattan-like road network with $\ell=5$ and $J_{\mathrm{E}}=10$. We draw the centers of the cells and report the numbering of roads and junctions. Roads are actually two-way, the small gap between lanes going in opposite directions is left for visualization purpose only. Road directions are indicated by the arrows at the bottom-left corner.

In order to fairly compare simulations with different number of vehicles, we report the normalized

[1] https://www.gnu.org/software/glpk/

approximate Wasserstein distance

$$\hat{\mathcal{H}} := \frac{\sum_j \sum_k c_{jk} x^*_{jk}}{M}, \tag{2.4.20}$$

where M as in (2.2.4), and x^*_{jk} is the solution of the LP problem (2.1.2).

2.4.1 Sensitivity to initial data.

In this test we measure the sensitivity to the initial position of vehicles. The goal is to quantify the impact of a possible error in locating vehicles at initial time (but still catching the correct amount of vehicles). In addition, this preliminary test aims at investigating some conceptual and numerical aspects of the proposed procedure. In particular we show the difference between Wasserstein and L^1 distance (see Sec. 1.4) and we study the convergence $\hat{\mathcal{H}} \to W$ as $\Delta x \to 0$ (see Sec. 2.2).

The parameters which remain fixed in this test are

- *Fundamental diagram*: $\sigma = 0.3$ and $f_{\max} = 0.25$ (see (1.3.35)).
- *Distribution matrix*:

$$\alpha^{\mathrm{V}}_{\mathrm{R}\mathrm{R}'} = \frac{1}{n^{\mathrm{V}}_{\mathrm{out}}}, \qquad \forall \mathrm{V} \in \mathcal{V}, \quad \mathrm{R} = 1, \dots, n^{\mathrm{V}}_{\mathrm{inc}}, \quad \mathrm{R}' = 1, \dots, n^{\mathrm{V}}_{\mathrm{out}}.$$

We consider the following two initial conditions, see Fig. 2.6: for all $\mathrm{E} \in \mathcal{E}$ and $j = 1, \dots, N_{x_{\mathrm{E}}}/2$,

$$\rho^{\mathrm{S},0}_{\mathrm{E},j} = \begin{cases} 0.5, & \text{on rightward roads} \\ 0, & \text{elsewhere,} \end{cases} \qquad \rho^{\mathrm{D},0}_{\mathrm{E},j} = \begin{cases} 0.5, & \text{on leftward roads} \\ 0, & \text{elsewhere.} \end{cases} \tag{2.4.21}$$

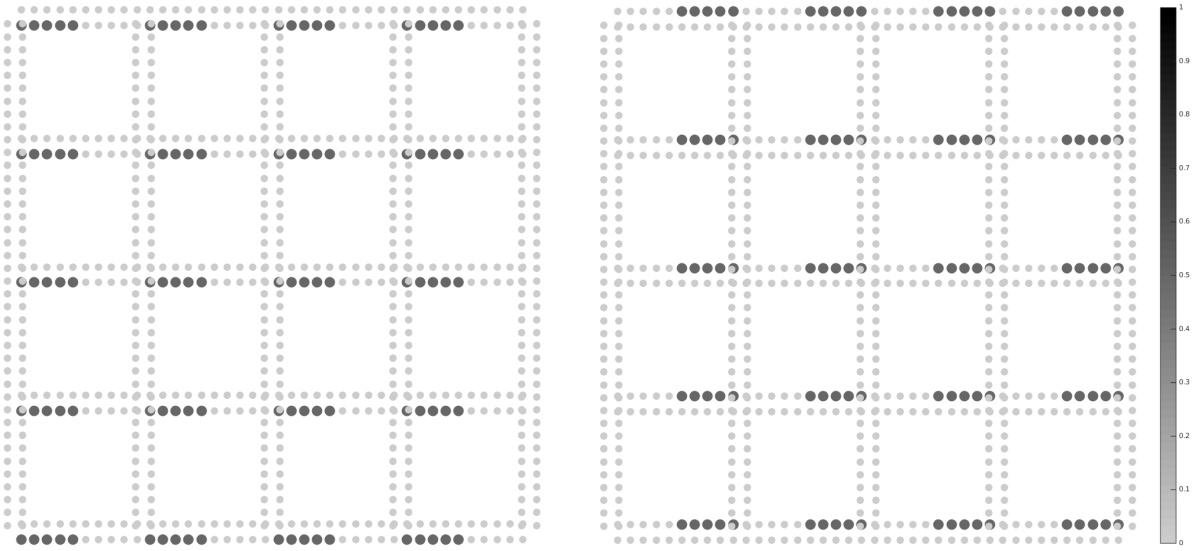

Figure 2.6: Sensitivity to initial data. $\rho^{\mathrm{S},0}$ (left) and $\rho^{\mathrm{D},0}$ (right).

Remark 2.4.1. *Due to the uniform traffic distribution at junctions, the density tends to become constant on the whole network as $t \to +\infty$, regardless of the initial datum. As a consequence, we expect that the distance between $\rho^{\mathrm{S}}(t)$ and $\rho^{\mathrm{D}}(t)$ (no matter how defined) tends to 0 as $t \to +\infty$.*

2.4.1.1 Comparison with L^1 distance.

In this test we compare the approximate Wasserstein distance with the discrete L^1 distance (normalized with respect to the mass as well), here denoted by $\hat{\mathcal{L}}^1$ and defined by

$$\hat{\mathcal{L}}^1(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t)) := \frac{\Delta x}{M}\sum_{\mathrm{E}\in\mathcal{E}}\sum_{j=1}^{N_{x_{\mathrm{E}}}}|\rho^{\mathrm{S}}_{\mathrm{E},j}(t)-\rho^{\mathrm{D}}_{\mathrm{E},j}(t)|. \tag{2.4.22}$$

Functions $t\to\hat{\mathcal{L}}^1(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t))$ and $t\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t))$ are shown in Fig. 2.7 for two different network size. Initially, the $\hat{\mathcal{L}}^1$ distance shows a plateau, which lasts until the supports of the densities ρ^{S} and

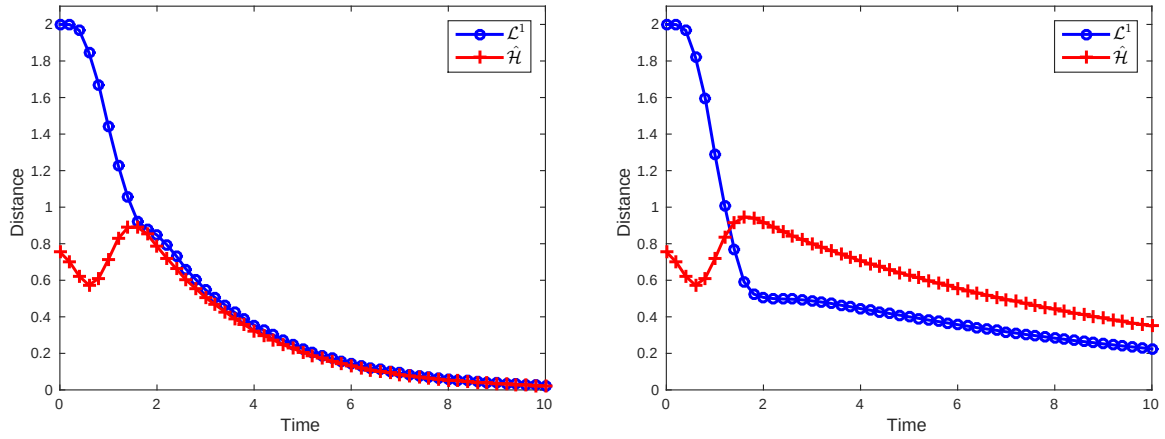

Figure 2.7: Sensitivity to initial data ($\hat{\mathcal{L}}^1$ vs. $\hat{\mathcal{H}}$). Comparison between functions $t\to\hat{\mathcal{L}}^1(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t))$ and $t\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t))$ for $\ell=3$ (left) and $\ell=5$ (right).

ρ^{D} are disjoint. This is not the case of the Wasserstein distance which instead immediately decreases. After that, the supports of the two densities start to overlap but the regions with maximal density move away from each other, see Fig. 2.8. When this process ends, we get the maximal value of the Wasserstein distance and the change of slope of the $\hat{\mathcal{L}}^1$ distance. Later on, the two densities uniformly distribute along the network and the two distances go smoothly to 0.

2.4.1.2 Numerical convergence as $\Delta x\to 0$.

In this test we consider a small network ($\ell=3$) and we compute the Wasserstein distance $\hat{\mathcal{H}}(\rho^{\mathrm{S}},\rho^{\mathrm{D}})$ for different values of $N_{x_{\mathrm{E}}}$. Fig. 2.9 shows the functions $t\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t))$ for $N_{x_{\mathrm{E}}}=10,20,40,80$ and $N_{x_{\mathrm{E}}}\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,T),\rho^{\mathrm{D}}(\cdot,T))$ at fixed time $T=1.4$. Fig. 2.9(left) suggests a relatively small sensitivity to the space step. We can safely assume that the difference between the values of $\hat{\mathcal{H}}$ obtained with $N_{x_{\mathrm{E}}}=10$ and $N_{x_{\mathrm{E}}}=160$ is lower than 10% with respect to the largest of the two values. We get similar results also for larger networks. The numerical convergence of $\hat{\mathcal{H}}=\hat{\mathcal{H}}(J_{\mathrm{E}})$ as $N_{x_{\mathrm{E}}}\to+\infty$ is also clearly visible in Fig. 2.9(right).

In the next sections, the sensitivity analysis will be obtained with $J_{\mathrm{E}}=10$ which seems to be a good compromise between accuracy of the results and computational costs.

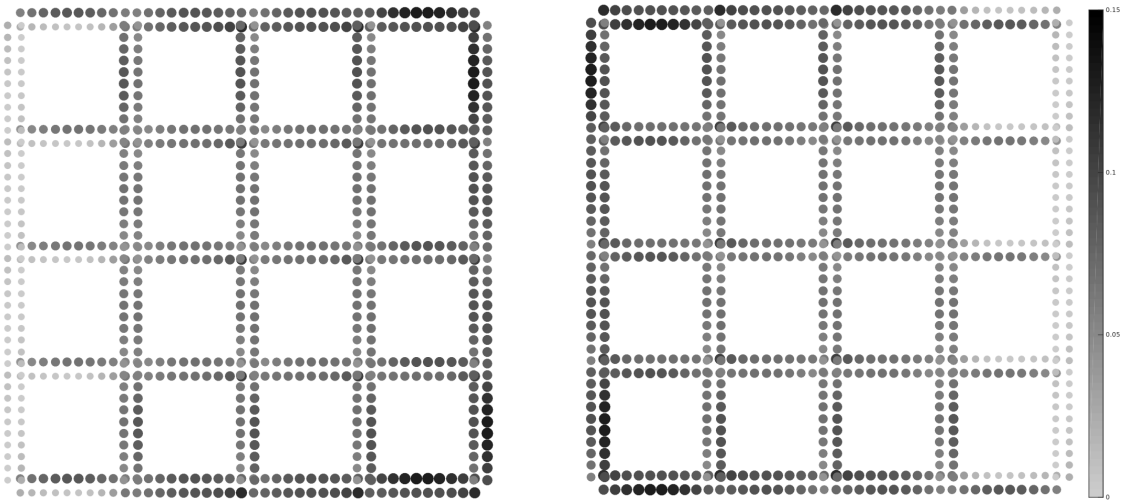

Figure 2.8: Sensitivity to initial data ($\hat{\mathcal{L}}^1$ vs. $\hat{\mathcal{H}}$). ρ^{S} (left) and ρ^{D} (right) at time $T=1.8$.

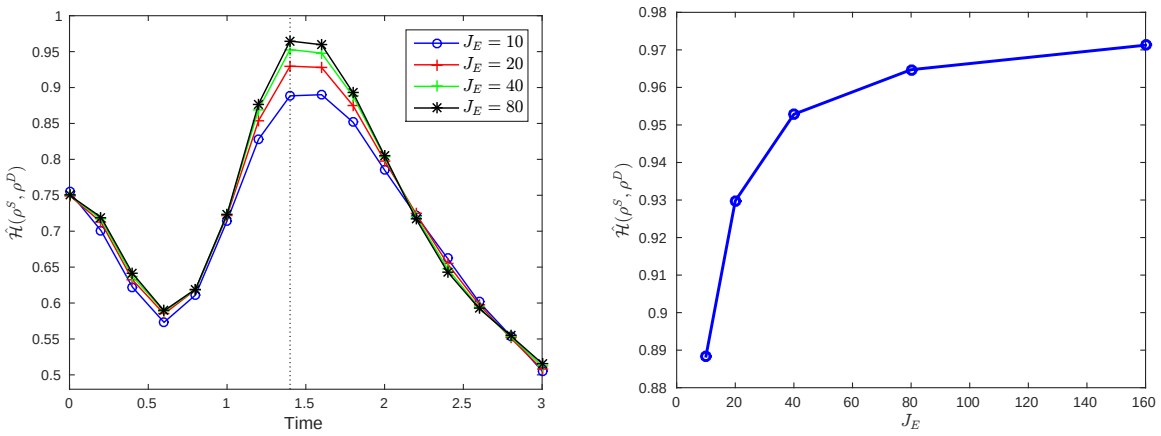

Figure 2.9: Sensitivity to initial data (convergence). Function $t \to \hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,t), \rho^{\mathrm{D}}(\cdot,t))$ for different values of $N_{x_{\mathrm{E}}}$(J_{E} for semplicity) (left) and function $N_{x_{\mathrm{E}}} \to \hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,T), \rho^{\mathrm{D}}(\cdot,T))$ with $T=1.4$ (right).

2.4.2 Sensitivity to fundamental diagram.

In this test we measure the sensitivity to the two parameters of the fundamental diagram, namely σ and $f_{\max}$. The goal is to quantify the impact of a possible error in measuring the capacity of the roads or in describing the drivers behavior. Note that the linear structure of the fundamental diagram used here (see Fig. 1.3) does not play any special role and any other fundamental diagram could be considered, as long as it is duly parametrized.

The parameters which remain fixed in this test are

- *Initial density*:

$$\rho^0_{\mathrm{E},j} = \begin{cases} 0.5 & \text{on rightward roads,} \\ 0 & \text{elsewhere,} \end{cases} \qquad \mathrm{E} \in \mathcal{E}, \quad j = 1\ldots, J_{\mathrm{E}}.$$

- *Distribution matrix*:

$$\alpha^{\mathrm{V}}_{\mathrm{RR}'} = \frac{1}{n^{\mathrm{V}}_{\mathrm{out}}}, \qquad \forall \mathrm{V} \in \mathcal{V}, \quad \mathrm{R} = 1,\ldots,n^{\mathrm{V}}_{\mathrm{inc}}, \quad \mathrm{R}' = 1,\ldots,n^{\mathrm{V}}_{\mathrm{out}}.$$

In Fig. 2.10(left) we report the distance between the solutions ρ^{S} and ρ^{D} at time $T=20$ obtained with $f^{\mathrm{S}}_{\max}=0.25$, $\sigma^{\mathrm{S}}=0.3$ and $f^{\mathrm{D}}_{\max}=0.25$, $\sigma^{\mathrm{D}}\in[0.15,0.5]$, respectively. In Fig. 2.10(right) we report the distance between the two solutions at time $T=20$ obtained with $\sigma^{\mathrm{S}}=0.3$, $f^{\mathrm{S}}_{\max}=0.25$, and $\sigma^{\mathrm{D}}=0.3$, $f^{\mathrm{D}}_{\max}\in[0.15,0.4]$, respectively. Errors in the calibration of σ or $f_{\max}$ lead to similar discrepancies, which

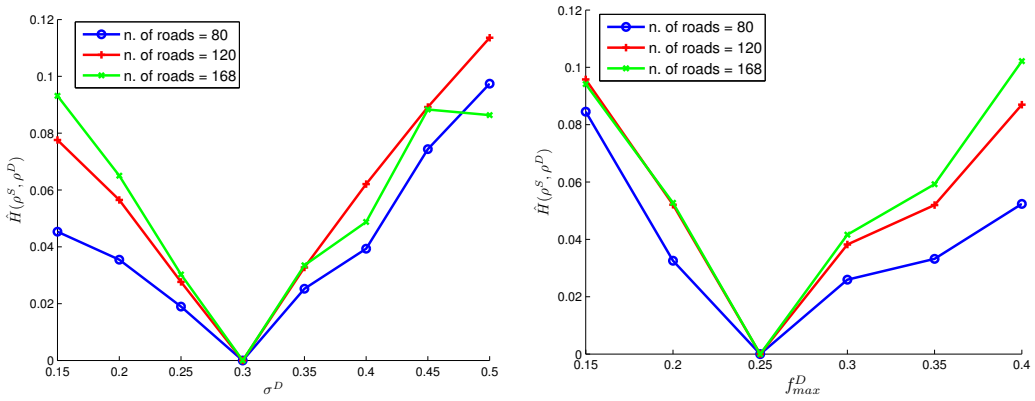

Figure 2.10: Sensitivity to fundamental diagram. Function $\sigma^{\mathrm{D}}\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,T),\rho^{\mathrm{D}}(\cdot,T))$ (left) and, $f^{\mathrm{D}}_{\max}\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,T),\rho^{\mathrm{D}}(\cdot,T))$ (right), for $\ell=5,6,7$.

are again amplified by the network size. Discrepancies grow approximately linearly with respect to both $|\sigma^{\mathrm{D}}-\sigma^{\mathrm{S}}|$ and $|f^{\mathrm{D}}_{\max}-f^{\mathrm{S}}_{\max}|$.

In Fig. 2.11 we report the distance between the solutions ρ^{S} and ρ^{D} obtained with $f^{\mathrm{S}}_{\max}=f^{\mathrm{D}}_{\max}=0.25$, $\sigma^{\mathrm{S}}=0.3$, $\sigma^{\mathrm{D}}=0.2$ (left), and $\sigma^{\mathrm{S}}=\sigma^{\mathrm{D}}=0.3$, $f^{\mathrm{S}}_{\max}=0.25$, $f^{\mathrm{D}}_{\max}=0.3$ (right), as a function of time. Again, we

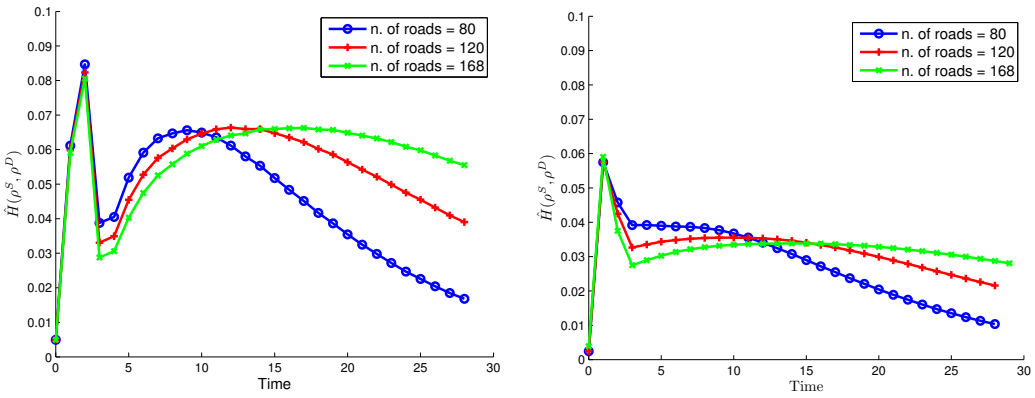

Figure 2.11: Sensitivity to fundamental diagram. With respect to σ (left) and $f_{\max}$ (right). Function $t\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t))$ for $\ell=5,6,7$.

see that the distances tend to 0 as $t\to\infty$ because vehicles spread across the networks toward a constant stationary density distribution (see Remark 2.4.1), and the size of the network affects the time scale only.

2.4.3 Sensitivity to the distribution matrix.

In this test we measure the sensitivity to the distribution coefficients at junctions, see Sec. 2.3. The goal is to quantify the impact of a possible error in the knowledge of the path choice at junctions.

The parameters which remain fixed in this test are

- *Initial density*: $\rho^0_{\mathrm{E},j} = 0.5, \quad \mathrm{E} \in \mathcal{E},\ j = 1,\dots,N_{x_\mathrm{E}}$.
- *Fundamental diagram*: $\sigma = 0.3$ and $f_{\max} = 0.25$.

Supply distribution ρ^{S} is obtained by means of equidistributed coefficients

$$\alpha^{\mathrm{S},\mathrm{V}}_{\mathrm{R}\mathrm{R}'} = \frac{1}{n^{\mathrm{V}}_{\mathrm{out}}}, \quad \forall \mathrm{V} \in \mathcal{J}, \quad \mathrm{R} = 1,\dots,n^{\mathrm{V}}_{\mathrm{inc}}, \quad \mathrm{R}' = 1,\dots,n^{\mathrm{V}}_{\mathrm{out}}.$$

Note that, due to the symmetry of the network and the initial datum, $\rho^{\mathrm{S}} \equiv 0.5$ for all x and t.

2.4.3.1 Single junction.

Here demand distribution ρ^{D} is obtained by varying the distribution coefficients at the junction $\bar{\mathrm{V}}$ located at the very center of the network (see, e.g., vertex 13 in Fig. 2.5). Variation is performed by means of a scalar parameter $\varepsilon > 0$. We have, for all incoming roads $\mathrm{R} = 1,2,3,4$,

$$\alpha^{\mathrm{D},\bar{\mathrm{V}}}_{\mathrm{R}1} = \frac{1}{n^{\mathrm{V}}_{\mathrm{out}}} + \varepsilon, \quad \alpha^{\mathrm{D},\bar{\mathrm{V}}}_{\mathrm{R}2} = \frac{1}{n^{\mathrm{V}}_{\mathrm{out}}} - \varepsilon, \quad \alpha^{\mathrm{D},\bar{\mathrm{V}}}_{\mathrm{R}3} = \frac{1}{n^{\mathrm{V}}_{\mathrm{out}}} + \varepsilon, \quad \alpha^{\mathrm{D},\bar{\mathrm{V}}}_{\mathrm{R}4} = \frac{1}{n^{\mathrm{V}}_{\mathrm{out}}} - \varepsilon.$$

In Fig. 2.12 we report the distribution ρ^{D} at time $t = 5$ and $t = 45$ obtained with $\varepsilon = 0.1$, to be compared with the constant distribution $\rho^{\mathrm{S}} \equiv 0.5$. Remarkably, a minor local modification of the traffic distribution

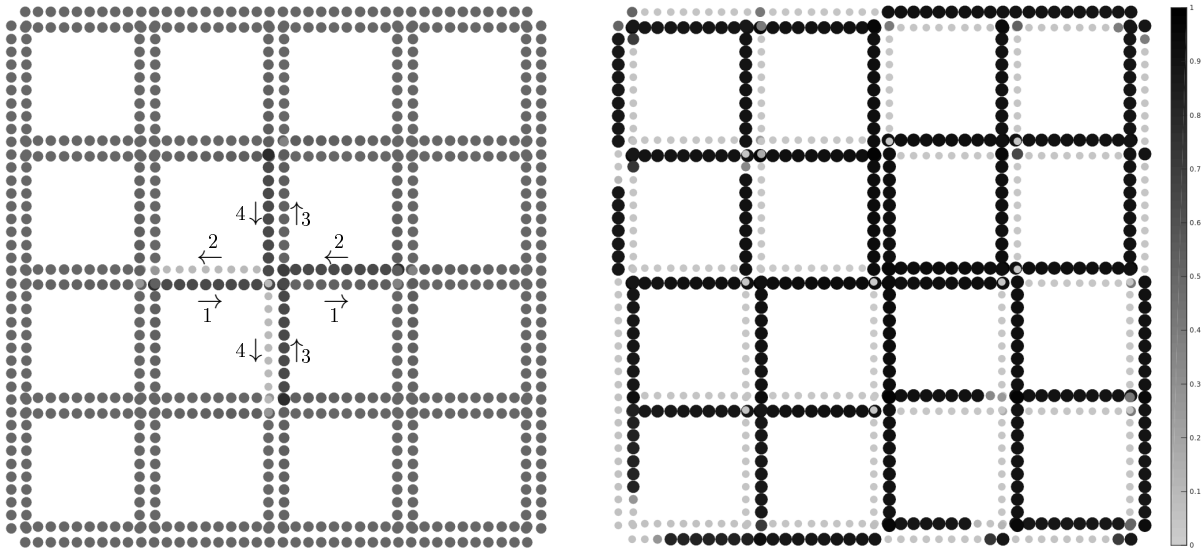

Figure 2.12: Sensitivity to distribution matrix (single junction). Density ρ^{D} at time $t = 5$ (left) and $t = 45$ (right).

in a single junction breaks the symmetry and has a great impact on the solution. This time the density does not tend to distribute uniformly across the network and then we expect the distance $W(\rho^{\mathrm{S}}, \rho^{\mathrm{D}})$ to increase in time, although the growth cannot continue indefinitely since the distance between two distributions on a finite network is finite.

In Fig. 2.13 we show the distance between the two densities as a function of time for $\varepsilon = 0.1,\ 0.2$. The distance is indeed increasing and bounded as expected. Moreover a larger ε accelerates the growth of the distance. Further comments will be given in the following Sec. 2.4.3.3.

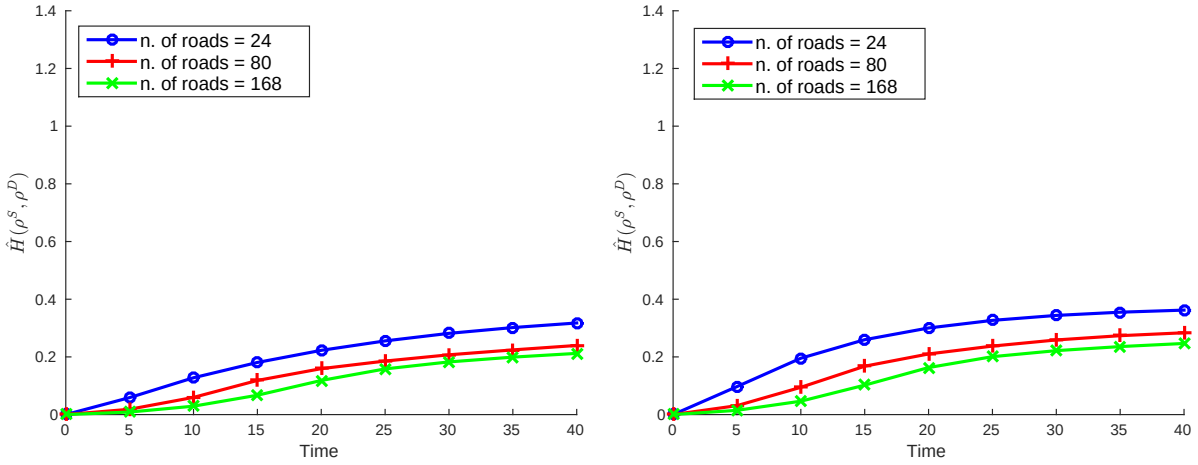

Figure 2.13: Sensitivity to traffic distribution at junctions (single junction). Function $t \to \hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,t), \rho^{\mathrm{D}}(\cdot,t))$ for $\ell = 3,5,7$. $\varepsilon = 0.1$ (left), $\varepsilon = 0.2$ (right).

2.4.3.2 All junctions.

Let us now modify *all* the distribution coefficients, and not only those at one junction. In the following test we set, for any V and $\mathrm{R} = 1,2,3,4$,

$$\alpha_{\mathrm{R}1}^{\mathrm{D},\mathrm{V}} = \frac{1}{n_{\mathrm{out}}^{\mathrm{V}}} + \varepsilon, \quad \alpha_{\mathrm{R}2}^{\mathrm{D},\mathrm{V}} = \frac{1}{n_{\mathrm{out}}^{\mathrm{V}}} - \varepsilon, \quad \alpha_{\mathrm{R}3}^{\mathrm{D},\mathrm{V}} = \frac{1}{n_{\mathrm{out}}^{\mathrm{V}}} + \varepsilon, \quad \alpha_{\mathrm{R}4}^{\mathrm{D},\mathrm{V}} = \frac{1}{n_{\mathrm{out}}^{\mathrm{V}}} - \varepsilon,$$

if V is labeled by an odd number and

$$\alpha_{\mathrm{R}1}^{\mathrm{D},\mathrm{V}} = \frac{1}{n_{\mathrm{out}}^{\mathrm{V}}} - \varepsilon, \quad \alpha_{\mathrm{R}2}^{\mathrm{D},\mathrm{V}} = \frac{1}{n_{\mathrm{out}}^{\mathrm{V}}} + \varepsilon, \quad \alpha_{\mathrm{R}3}^{\mathrm{D},\mathrm{V}} = \frac{1}{n_{\mathrm{out}}^{\mathrm{V}}} - \varepsilon, \quad \alpha_{\mathrm{R}4}^{\mathrm{D},\mathrm{V}} = \frac{1}{n_{\mathrm{out}}^{\mathrm{V}}} + \varepsilon,$$

otherwise (at border junctions only the first two incoming roads $\mathrm{R} = 1,2$ are considered). Results with $\varepsilon = 0.1$ are shown in Fig. 2.14.

2.4.3.3 Comparison

We observe a great difference between the density distributions ρ^{D}'s reported in Secs. 2.4.3.1 and 2.4.3.2. In Fig. 2.12(right) we see that free and congested roads segregate but remain close to each other. On the contrary, in Fig. 2.14(left) free and congested roads segregate *and separate spatially* from each other. The Wasserstein distance is able to catch this difference. Indeed, in the former test the Wasserstein distance is almost independent of the network size (Fig. 2.13(left)), while in the latter test (Fig. 2.14(right)) it is proportional to the network size (at large times).

2.4.4 Sensitivity to road network.

In this test we measure the sensitivity to the road network. The goal is to quantify the impact of a possible change in the network, specifically a road closure.

The parameters which remain fixed in this test are

- *Initial density*: $\rho_{\mathrm{E},j}^0 = 0.3, \quad \mathrm{E} \in \mathcal{E},\ j = 1,\dots,J_{\mathrm{E}}$.

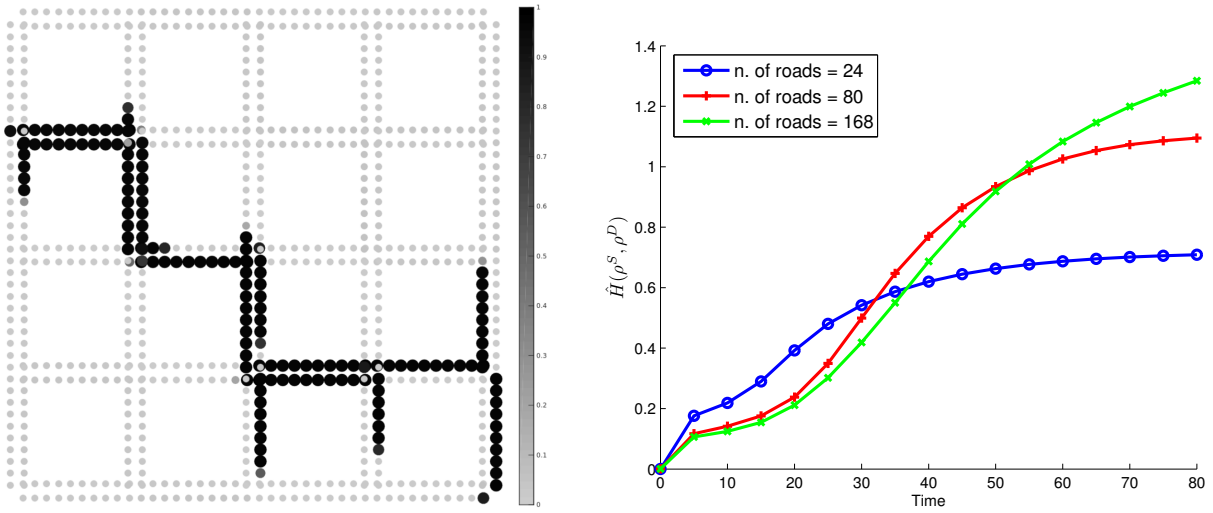

Figure 2.14: Sensitivity to traffic distribution at junctions (all junctions). Density ρ^{D} at time $t=55$ (left) and function $t\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t))$ for $\ell=3,5,7$ (right).

- *Fundamental diagram*: $\sigma=0.3$ and $f_{\max}=0.25$.
- *Distribution matrix:* equidistributed along outgoing roads.

Supply distribution ρ^{S} is obtained solving the equations on the complete network, while demand distribution ρ^{D} is obtained by closing the central rightward road $\bar{\mathrm{E}}$ (see, e.g., edge 11 in Fig. 2.5) just after the initial time, i.e. vehicles can come out of the road but none of them can enter. Note that, due to the symmetry of the network and the initial datum, $\rho^{\mathrm{S}}\equiv 0.3$ for all x and t.

In Fig. 2.15(left) we report the distribution ρ^{D} at time $t=55$, to be compared with the constant distribution. We see that the closure of a single road has a great impact on the solution. In Fig.

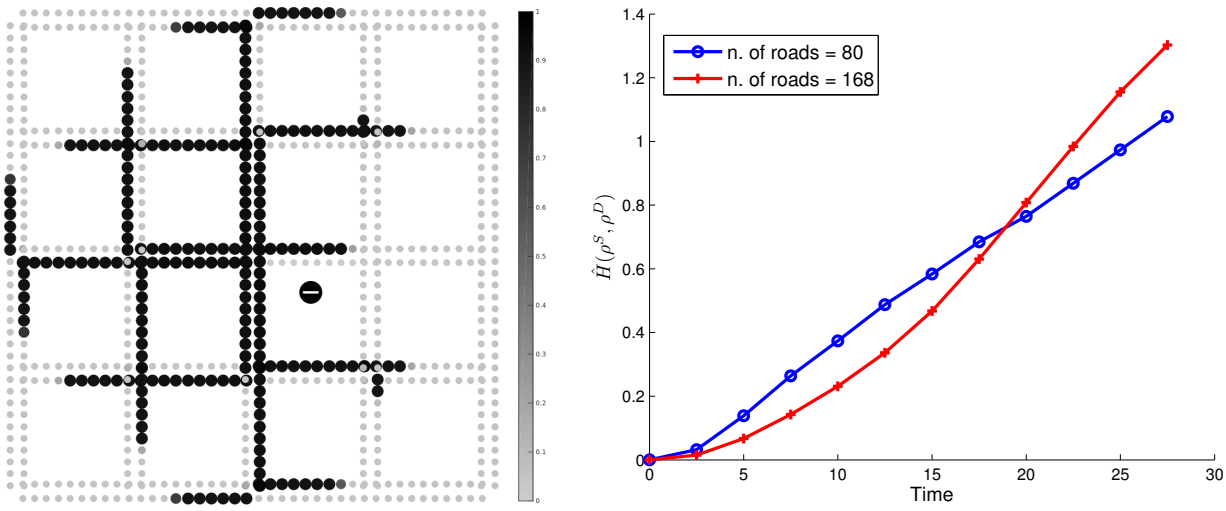

Figure 2.15: Sensitivity to road network. Density ρ^{D} at time $t=55$ (left) and function $t\to\hat{\mathcal{H}}(\rho^{\mathrm{S}}(\cdot,t),\rho^{\mathrm{D}}(\cdot,t))$ for $\ell=5,7$ (right).

2.15(right) we show the distance between the two densities as a function of time. The long-time behavior of the sensitivity is proportional to the network size. Even if the road closure is a *local* modification of

the network dynamics, the behavior with respect to the network size is more similar to that shown in Fig. 2.14(right) (all junctions perturbation) than that shown in Fig. 2.13(left) (single junction perturbation). Again the reason can be found by observing the densities: Fig. 2.15(left) shows that the free and congested roads segregate and separate spatially from each other, as in Fig. 2.14(left).

Chapter 3

A new multi-scale model

The analysis done in the previous chapter suggests us to improve first order macroscopic models, for example mixing different scales and define a new multi-scale multi-order model to reproduce accurately the density evolution.

In this chapter we will introduce a new multi-scale model conceived to inherit the advantages of a first-order macroscopic model, see Sec. 1.3.1, and a second-order microscopic model, see Sec. 1.3.2, avoiding any interface or boundary conditions between them.

Moreover the numerical counterpart will be developed and the complete algorithm will be explained in detail. We will prove also that the Euler-Godunov associated scheme is conservative, so the total mass is preserved.

3.1 Motivations

Single-scale models are often unsatisfactory for a number of reasons which involve both modeling and numerical considerations. Generally speaking, we can say that:

- Second-order models are more realistic and often perform better than first-order ones [33, 34];
- The numerical approximation of second-order macroscopic models is more difficult than that of first-order macroscopic models, especially if high-order numerical schemes are pursued;
- Microscopic models require a rather large CPU time if the number of vehicles involved in the simulation is large.

Ideally, one would have an easy-to-implement macroscopic model with second-order features. This is the goal which motivated the multi-scale model proposed. Multi-scale models are typically based on the spatial separation of the microscopic and macroscopic parts. On the contrary, the model we propose here is characterized by the fact that no interface (either fixed or mobile) is explicitly defined. More precisely, the macroscopic model is always and everywhere alive, while the microscopic model is activated only where and when it is needed, and it is able to correct (in full or in part) the macroscopic one. This procedure is expected to be advantageous if one couples an easy-to-use first-order macroscopic model with a more realistic but still easy-to-use second-order microscopic model. Advantages are complemented by the low computational cost, which comes from the fact that the microscopic model is used only in small

parts of the road, therefore the number of vehicles individually tracked is kept low. The question arises how the need of the second-order model can be detected. In principle, the second-order model should be activated when and where the traffic is not at equilibrium, i.e. the velocity of vehicles is far from the desired one (as for the current traffic conditions). This usually happens when nearby vehicles have very different speeds, since this implies the need of strong accelerations or decelerations which first-order models cannot handle. On the contrary, the second-order model can be safely deactivated when vehicles are moving at desired velocity, i.e. their acceleration is close to zero.

3.2 Multi-scale model with complete information

We are now ready to couple a microscopic model in the form (1.3.45) and the macroscopic model (1.3.33). In this section we assume, for illustrative purposes, that *both the macroscopic and the microscopic models are alive in the whole space-time domain*, i.e. we have a complete information coming from the two models.

It is given an initial condition at macroscopic level $\rho_0 := \rho(\cdot,0) \in L^1(\mathbb{R})$, such that $\mathcal{M} = \int_{\mathbb{R}} \rho_0(x)dx$ and N is the corresponding number of microscopic vehicles. One can set the initial position of vehicles $\{X_1(0),\dots,X_N(0)\}$ simply distributing them according to the probability density distribution $\rho_0/\mathcal{M}$. Initial velocities are set as $V_k(0) = v^*(\rho_0(X_k(0)))$, for $k < N$, and $V_N(0) = V_{\max}$.

Let us also define, as usual, $f : [0,\rho_{\max}] \to \mathbb{R}^+$, $\rho \mapsto f(\rho) := \rho v^*(\rho)$, the flux of vehicles as a function of their density.

The multi-scale model with complete information reads, in integral form, as

$$\begin{cases} \partial_t \displaystyle\int_a^b \rho(x,t)dx = \theta\big(f(a,t) - f(b,t)\big) \\ \qquad\qquad + (1-\theta)\left(\displaystyle\sum_{k=1}^{N} \ell_N \delta(X_k(t)-a) - \sum_{k=1}^{N} \ell_N \delta(X_k(t)-b)\right), \quad \forall a,b \in \mathbb{R} \\ \dot{X}_k(t) = V_k(t), \quad k \le N, \\ \dot{V}_k(t) = A(X_k(t), X_{k+1}(t), V_k(t), V_{k+1}(t); \mathbf{p}), \quad k < N, \\ \dot{V}_N(t) = 0, \end{cases} \tag{3.2.1}$$

where $\theta \in [0,1]$ is an additional parameter, $x \mapsto \delta(x - x_0)$ is the Dirac delta function centered at x_0, and the time derivative ∂_t is intended in the distributional sense. Moreover, $\mathbf{p}$ represents the vector of model parameters.

The idea underlying the model (3.2.1) is the following: The gain or loss of mass in any space interval $[a,b]$ between time t and $t+dt$ is only given by the flow of vehicles through the boundaries a and b, like in classical conservation laws. In this case, at the macroscopic level the flow is given by the classical LWR flux function f, while at the microscopic level the mass instantaneously (dis)appears at the passage of vehicles through the boundaries. The parameter θ is intended for tuning the contribution of the micro- and macro-scale, in the same spirit of [20, 22, 23, 24, 26]. Here we expect that large values of θ reduce the oscillations due to abrupt passage of microscopic vehicles across boundaries, while small values of θ increase the effectiveness of second-order dynamics.

3.2.1 Numerical approximation

The system (3.2.1) can be approximate by a suitable combination of existing numerical schemes. In the following, we employ the classical Godunov scheme for the PDE and the explicit Euler scheme for the system of ODEs. To do that, we recall the notation introduced above: for space and time steps $\Delta x, \Delta t > 0$ (with $\lambda := \frac{\Delta t}{\Delta x}$) and a grid in space $\{x_j := j\Delta x,\ j \in \mathbb{Z}\}$ and time $\{t^n := n\Delta t,\ n \in \mathbb{N}\}$. We denote by $C_j := [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$ the cell centered in x_j. As usual, we define

$$\rho_j^n := \frac{1}{\Delta x} \int_{C_j} \rho(x, t^n) dx$$

and we denote by (X_k^n, V_k^n) the approximation of $(X_k(t^n), V_k(t^n))$, for $k = 1, \dots, N$.

In order to define the correspondence between the micro- and the macro-scale we also introduce the scaling parameter $\Gamma_{\max}$, defined as the maximum number of vehicles which can fall in one cell of length Δx. $\Gamma_{\max}$ is the microscopic counterpart of $\rho_{\max}$ and it is naturally related to the scaling parameter ℓ_N by

$$\ell_N = \frac{\Delta x}{\Gamma_{\max}}.$$

By means of $\Gamma_{\max}$ we can define the number of vehicles to put in any cell C_j, which is equal to $\left\lfloor \frac{\rho_j^0}{\rho_{\max}} \Gamma_{\max} \right\rfloor$. In this way we have

$$\sum_{k=1}^{N} \ell_N = N\ell_N = \sum_{j \in \mathbb{Z}} \rho_j^0 \Delta x = \int_{\mathbb{R}} \rho_0(x) dx = \mathcal{M}.$$

We also assume that vehicles are initially equispaced in the cell and we assign to each of them the same velocity $v^*(\rho_j^0)$. This procedure defines the initial positions $\{X_1^0, \dots, X_N^0\}$ and velocities $\{V_1^0, \dots, V_N^0\}$ of the microscopic vehicles.

The numerical scheme reads as follows.

$$\begin{cases} \rho_j^{n+1} = \rho_j^n + \theta\lambda\Big(\mathcal{G}(\rho_{j-1}^n, \rho_j^n) - \mathcal{G}(\rho_j^n, \rho_{j+1}^n)\Big) + (1-\theta)\lambda\Big(\mathcal{F}_{j-\frac{1}{2}}^n - \mathcal{F}_{j+\frac{1}{2}}^n\Big), \\ X_k^{n+1} = X_k^n + \Delta t V_k^n, \quad k \le N, \\ V_k^{n+1} = V_k^n + \Delta t A(X_k^n, X_{k+1}^n, V_k^n, V_{k+1}^n; \mathbf{p}), \quad k < N, \\ V_N^{n+1} = V_N^n, \end{cases} \tag{3.2.2}$$

with the classical Godunov's numerical flux (1.3.69) and with the microscopic numerical flux:

$$\mathcal{F}_{j\pm\frac{1}{2}}^n := \frac{\ell_N}{\Delta t} \mathrm{Card}\left\{k \in \{1, \dots, N\} : x_k^n < x_{j\pm\frac{1}{2}} \le x_k^{n+1}\right\} \tag{3.2.3}$$

(where 'Card' denote the cardinality of a set). Note that the microscopic flux is simply computed by counting the number of vehicles passing through the boundaries of the cell C_j in the time interval $[t^n, t^{n+1})$.

Such a numerical scheme comes along with a natural CFL condition for the microscopic and the macroscopic part, in order to guarantee that in one time step both characteristic curves and vehicles

themselves do not move more than one cell apart. To get this, we need to impose

$$\lambda < \min\left\{ \frac{1}{\max\limits_{\rho\in[0,\rho_{\max}]} |f'(\rho)|}, \frac{1}{V_{\max}} \right\}. \tag{3.2.4}$$

3.3 The multi-scale algorithm

One of the main feature of the proposed multi-scale model is that the microscopic part of the model is activated only where and when the macroscopic model is expected to fail. This forces us to introduce two important modifications with respect to the model with complete information.

First, the number N of microscopic (singularly tracked) vehicles can change in time. To deal with that, we denote by Γ^n the total number of vehicles at time t^n, and by Γ_j^n the number of vehicles which, at time t^n, fall in the cell C_j.

Second, vehicle $k+1$ is no longer, in general, in front of vehicle k. To overcome this issue, we will denote by $\text{NEXT}(k)$ the vehicle in front of the k-th one, with the convection that leaders have $\text{NEXT}=0$.

Finally, we need some additional positive parameters which rule the activation and deactivation of the second-order microscopic model, which will be denoted by δt, δv and δV. Let us briefly describe their meaning, which will be even clearer after the description of the algorithm.

- δt is the minimal period of time that one microscopic vehicle is active. In other words, if a vehicle is activated at time t^n, it cannot be deactivated before time $t^n+\delta t$.
- δv is the minimal variation of the macroscopic velocity function which activates the second-order model.
- δV is the maximal difference between the current velocity and the equilibrium velocity of microscopic vehicles which allows the deactivation of the second-order model.

We are now ready to present the main steps of the algorithm which updates the density and the state of the microscopic vehicles

$$(\rho^n, X^n, V^n) \to (\rho^{n+1}, X^{n+1}, V^{n+1}).$$

1. Compute $\Gamma_j^n \quad \forall j$.
2. *Activation of new vehicles (Fig. 3.1).* For all j, if $|v^*(\rho_{j+1}^n) - v^*(\rho_j^n)| > \delta v$, for all $i \in \{j-1, j, j+1, j+2\}$ put new vehicles in cell C_i (unless the cell is already occupied, i.e. unless $\Gamma_i^n > 0$) with velocity $v^*(\rho_i^n)$. The number of vehicles to put in the cell C_i is $\left\lfloor \frac{\rho_i^n}{\rho_{\max}} \Gamma_{\max} \right\rfloor$. They are initially equispaced in the cell and they have the same velocity.

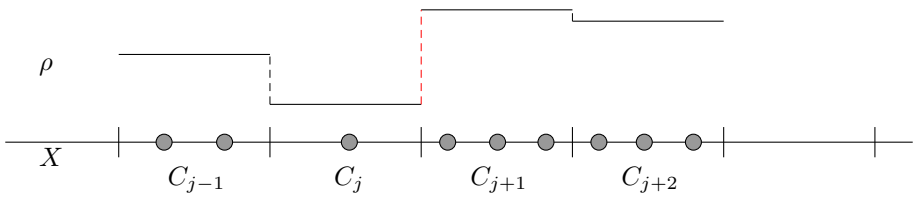

Figure 3.1: Step 1: Vehicles appear around large jumps of the macroscopic velocity (corresponding to large jumps of the macroscopic density).

3. *Labeling (Fig. 3.2).* Find $\text{NEXT}(k)$ for all k. The rightmost vehicle is labeled as leader ($\text{NEXT}=0$). Also, all vehicles h such that $|X^n_{\text{NEXT}(h)} - X^n_h| > \Delta x$ are also labeled as leader. This choice comes from the assumption that every time a vehicle has a free space of length $\geq \Delta x$ in front of it, its dynamics ceases to be dependent on the vehicle in front (if any). We will see, in Step 7, that the dynamics will depend instead on the macroscopic density.

Figure 3.2: Step 3: Green vehicles are leaders.

4. *Deactivation of followers (Fig. 3.3).* Remove all followers k which are active since more than δt units of time and such that $\left|V_k - v^*\left(\frac{\rho_{\max}\ell_N}{X^n_{\text{NEXT}(k)} - X^n_k}\right)\right| < \delta V$. Note that, without the first condition new vehicles would immediately deactivated since their velocity is initially at equilibrium (see Step 2). In this way, instead, vehicles have enough time to fully exploit their second-order dynamics. After that, if and when they get close to the equilibrium velocity again, they are deactivated.

Figure 3.3: Step 4: Red vehicles are going to be deactivated.

5. *Deactivation of leaders.* Remove all leaders which are not followed by anyone (lonely leaders).

6. Repeat steps 1 and 3 if needed.

7. *Update vehicles' positions and velocities.* We run the microscopic second-order model

$$\begin{cases} X^{n+1}_k = X^n_k + \Delta t V^n_k, \quad \forall k, \\ V^{n+1}_k = V^n_k + \Delta t A(X^n_k, X^n_{\text{NEXT}(k)}, V^n_k, V^n_{\text{NEXT}(k)}; \mathbf{p}), \quad \text{if } \text{NEXT}(k) > 0, \\ V^{n+1}_k = v^*(\rho^n_{j^*(k,n)+1}), \quad \text{if } \text{NEXT}(k) = 0, \end{cases} \tag{3.3.5}$$

where $j^*(k,n)$ is the cell occupied by the vehicle k at time t^n. Note that leaders' dynamics only depend on the density ρ. Indeed, the velocity of a leader is that of macroscopic vehicles located in the cell in front of it (cf. Colombo and Marcellini [18]).

8. Compute $\mathcal{F}^n_{j-\frac{1}{2}}$ and $\mathcal{F}^n_{j+\frac{1}{2}}$ as defined in (3.2.3).

9. *Update vehicles' density.* We run the multi-scale model, which reads as follows. For $\theta = 0$,

$$\rho^{n+1}_j = \rho^n_j + \lambda \cdot \begin{cases} \mathcal{F}^n_{j-\frac{1}{2}} - \mathcal{F}^n_{j+\frac{1}{2}} & \text{if } \Gamma^n_{j-1},\ \Gamma^n_j,\ \Gamma^n_{j+1} > 0 \\ \mathcal{F}^n_{j-\frac{1}{2}} - \mathcal{G}(\rho^n_j, \rho^n_{j+1}) & \text{if } \Gamma^n_{j-1},\ \Gamma^n_j > 0 \ \&\ \Gamma^n_{j+1} = 0 \\ \mathcal{G}(\rho^n_{j-1}, \rho^n_j) - \mathcal{F}^n_{j+\frac{1}{2}} & \text{if } \Gamma^n_{j-1} = 0 \ \&\ \Gamma^n_j,\ \Gamma^n_{j+1} > 0 \\ \mathcal{G}(\rho^n_{j-1}, \rho^n_j) - \mathcal{G}(\rho^n_j, \rho^n_{j+1}) & \text{otherwise.} \end{cases} \tag{3.3.6}$$

For $\theta \in [0,1]$,

$$\rho_j^{n+1} = \rho_j^n + \lambda \cdot \begin{cases} \left[\theta \mathcal{G}(\rho_{j-1}^n, \rho_j^n) + (1-\theta)\mathcal{F}_{j-\frac{1}{2}}^n\right] - \left[\theta \mathcal{G}(\rho_j^n, \rho_{j+1}^n) + (1-\theta)\mathcal{F}_{j+\frac{1}{2}}^n\right] \\ \qquad\qquad \text{if } \Gamma_{j-1}^n,\ \Gamma_j^n,\ \Gamma_{j+1}^n > 0 \\ \left[\theta \mathcal{G}(\rho_{j-1}^n, \rho_j^n) + (1-\theta)\mathcal{F}_{j-\frac{1}{2}}^n\right] - \mathcal{G}(\rho_j^n, \rho_{j+1}^n), \\ \qquad\qquad \text{if } \Gamma_{j-1}^n,\ \Gamma_j^n > 0 \ \&\ \Gamma_{j+1}^n = 0 \\ \mathcal{G}(\rho_{j-1}^n, \rho_j^n) - \left[\theta \mathcal{G}(\rho_j^n, \rho_{j+1}^n) + (1-\theta)\mathcal{F}_{j+\frac{1}{2}}^n\right], \\ \qquad\qquad \text{if } \Gamma_{j-1}^n = 0 \ \&\ \Gamma_j^n,\ \Gamma_{j+1}^n > 0 \\ \mathcal{G}(\rho_{j-1}^n, \rho_j^n) - \mathcal{G}(\rho_j^n, \rho_{j+1}^n), \qquad \text{otherwise.} \end{cases} \tag{3.3.7}$$

The scheme is different from the one with complete information (3.2.2) because here microscopic information is not always available. Where microscopic vehicles are present, one can choose a suitable combination of macroscopic and microscopic flux to update the macroscopic density. Where microscopic vehicles are missing, only the macroscopic flux is used, see Fig. 3.4.

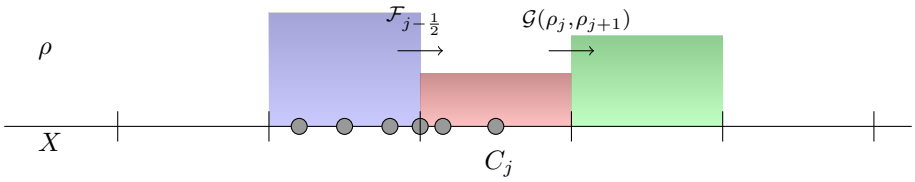

Figure 3.4: Step 9: Update of density ρ_j using microscopic flux on the left boundary and macroscopic flux on the right boundary of the cell j (case $\Gamma_{j-1},\ \Gamma_j > 0$ & $\Gamma_{j+1} = 0$, $\theta = 0$).

Remark 3.3.1. *We stress once again that the macroscopic density ρ is always updated (Step 9), regardless of the value of θ and regardless of the presence of microscopic vehicles (conversely we should manage the interface between the micro- and the macro-scale). Moreover, the total mass of the system must be evaluated by means on ρ only, as $\int_{\mathbb{R}} \rho dx$. The appearance of new microscopic vehicles should be seen as a temporary correction procedure which does not imply an additional injection of mass in the system. Furthermore this "correction" leads the macroscopic model showing a bounded acceleration behavior since it relies on a second order microscopic model. Note also that even if $\theta = 0$ (dynamics fully driven by the microscopic model), the macroscopic dynamics are still used, precisely where microscopic vehicles are not present. Moreover the density is used to define the velocity of microscopic leaders, so micro and macro models are no independent. Instead if $\theta = 1$ the microscopic model has no effect on the dynamics.* □

3.4 Numerical tests

In this section we present some numerical results for the multi-scale model described in Section 3.3, with scheme (3.3.6) ($\theta = 0$). In the first three tests A is defined as in (1.3.47), while in the last test A is defined as in (5.2.1).

We denote by L the length of the road, by T the final time of the simulation, by N_x the number of

space steps, by N_t the number of time steps. We set $\theta=\gamma=0$, $V_{\text{ref}}=V_{\max}=\rho_{\max}=1$ and $v^*(\rho)=1-\rho$. The others parameters used in the simulations are summarized in Table 3.1.

	T	L	N_x	N_t	τ	$\Gamma_{\max}$	δv	δt	δV	α	$\Delta_{\min}$
T1	3	20	100	300	0.01	20	0.08	$15\Delta t$	0.3	–	–
T2	3	20	100	600	0.01–3	30	0.1	$15\Delta t$	0.5	–	–
T3	12	20	100	1200	0.1	30	0.1	$30\Delta t$	0.2	–	–

Table 3.1: Model and algorithm parameters used for the numerical tests

3.4.1 Test 1: activation and deactivation of microscopic model

In this preliminary test we check the correctness of the steps 2, 3, 4 of the algorithm and we try to quantify the computational advantage of the new method. To do this, we consider a step function as initial condition ρ_0 and we plot the result of the classical LWR model (the density coming from the multi-scale is not plotted for better clarity) and the microscopic vehicles along the x-axis. We can see that at the initial time vehicles are correctly activated only around the three discontinuities (Fig. 3.5a). After some time, the first discontinuity is smoothed enough for allowing the deactivation of vehicles (Fig. 3.5b).

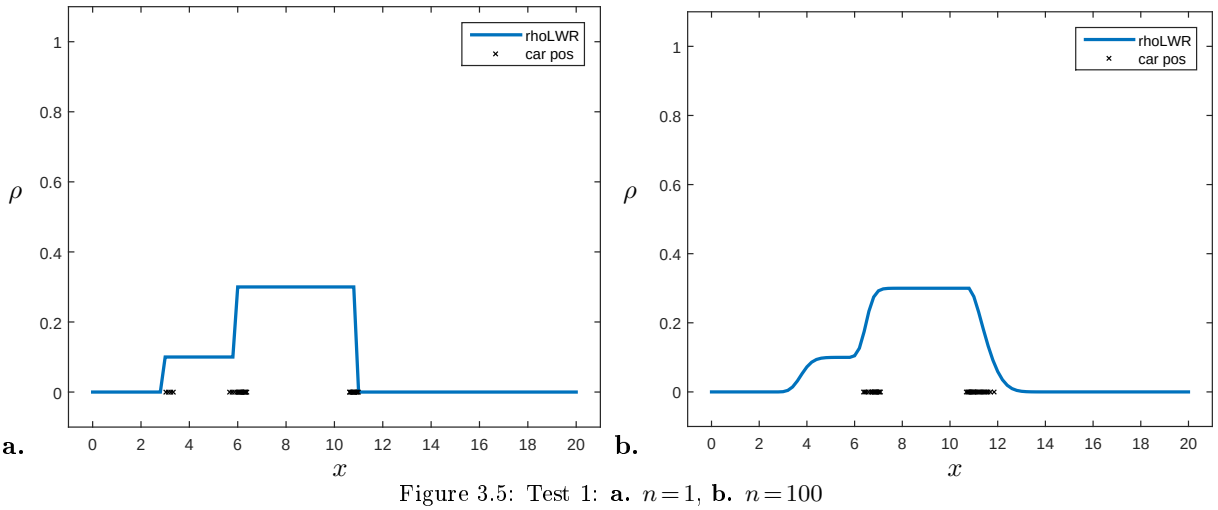

Figure 3.5: Test 1: **a.** $n=1$, **b.** $n=100$

In order to quantify the computational advantage, we compare the CPU time consumed by the fully microscopic model and the multi-scale model (the fully macroscopic model is not considered since we know that the LWR model is not able alone to provide a satisfactory traffic description). We run the two models for different values of the road length L, keeping fixed both Δx and $\Gamma_{\max}$. Positions of the discontinuities of ρ_0 (at $x=3,6,11$ if $L=20$) are scaled linearly with L. Note that the increase of L (and then of N_x accordingly), causes the increase of the total number of vehicles in the fully microscopic model, while in the multi-scale model the number of vehicles remains constant, being vehicles confined around discontinuities. In Fig. 3.6 we show the CPU time for the two models as a function of L. We see that the CPU time scales almost linearly with respect to L in both cases. This is a bit surprising for the multi-scale model but it can be explained by the fact that the computational effort of Steps 1, 2, 9 of the algorithm scales linearly with the number of cells N_x. Apart from that, we see that the CPU time for

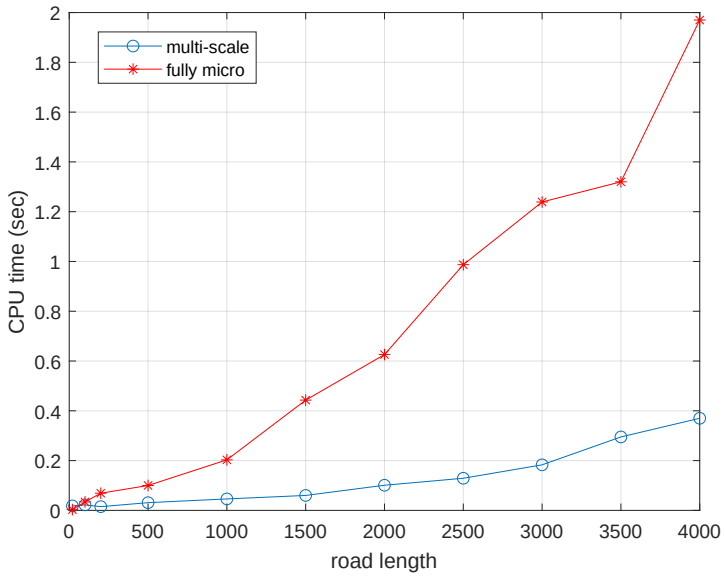

Figure 3.6: Test 1: CPU time for the fully microscopic model and the multi-scale model.

the multi-scale model is much lower than that of the fully microscopic model, and this is mainly due to the fact that the number of microscopic vehicles is kept low.

3.4.2 Test 2: effect of the relaxation term τ

In this test we investigate the effect of the relaxation parameter τ in the second-order model (1.3.47). This parameter is related to the *reactivity* of drivers. More precisely, the smaller τ, the more reactive the drivers are and the more the vehicles are able to accelerate and reach the equilibrium velocity rapidly. We consider the case of a road congested in the first part and totally free in the second part (Fig. 3.7a). In this case the LWR model shows immediately the classical rarefaction fan around the original discontinuity (Fig. 3.7bcd). Microscopic vehicles are activated only around the discontinuity (Fig. 3.7abcd) and are able to take into account the bounded acceleration of vehicles, see especially Fig. 3.7b where the velocity of microscopic vehicles is also plotted. At time step $n=400$ the difference between $\tau=0.01$ (Fig. 3.7c) and $\tau=3$ (Fig. 3.7d) is quite visible. In the former case (highly reactive drivers) the dynamics of the multi-scale model are very similar to that of the LWR model, while in the latest case (poorly reactive drivers) the multi-scale model differs from LWR and correctly take into account a delay in moving forward.

As discussed in Section 1.3.1, it is interesting to recover the fundamental diagram *a posteriori*, i.e. by means of the simulated traffic conditions. To do that, we plot the set of 2D points

$$\{(\rho^n_{j^*(k,n)}, \rho^n_{j^*(k,n)} V^n_k) \ \forall k \text{ and } \forall n, \text{ provided vehicle } k \text{ is active at } t=t^n\} \tag{3.4.8}$$

where $j^*(k,n)$ is the cell occupied by the vehicle k at time t^n, see Fig. 3.8. It can be seen that the multi-scale model is indeed able to recover a scattered (multivalued) fundamental diagram. Moreover, the scattering increases with τ, as expected, and decreases smoothly as $\rho \to 0$.

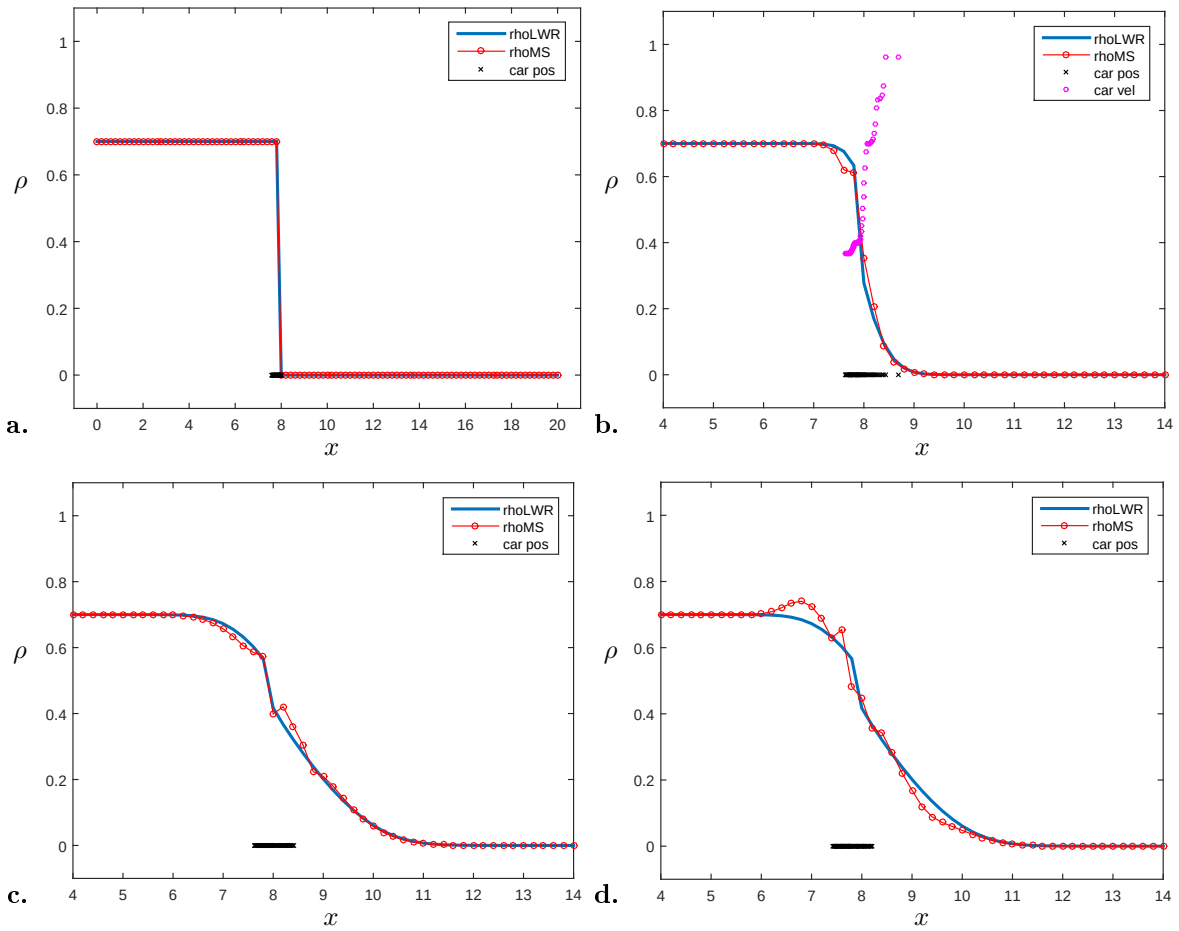

Figure 3.7: Test 2: **a.** $n=1$, **b.** $n=100$, **c.** $n=400$, $\tau=0.01$, **d.** $n=400$, $\tau=3$.

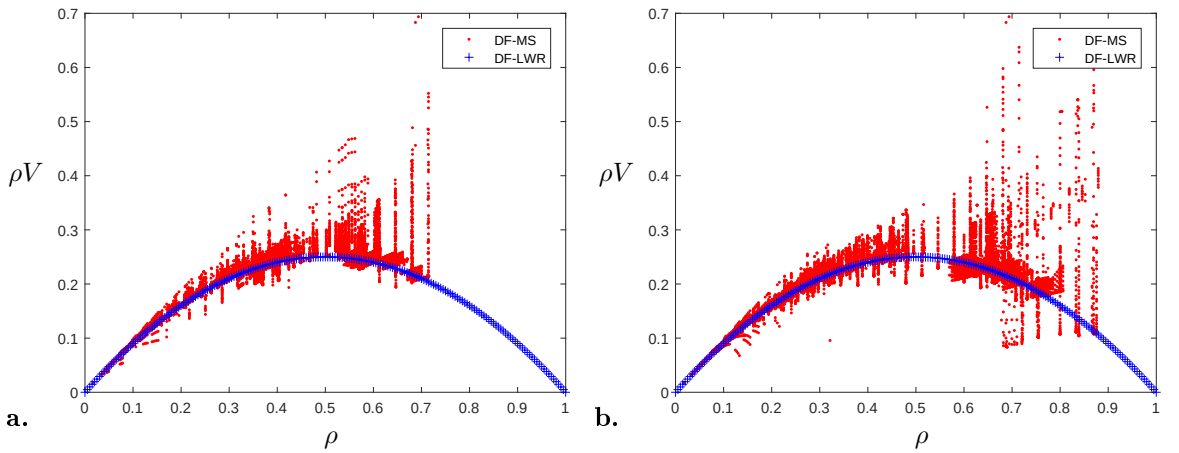

Figure 3.8: Test 2: Fundamental diagram of the multi-scale model compared with that of the LWR model. **a.** $\tau=0.01$, **b.** $\tau=3$.

3.4.3 Test 3: self-sustained perturbation

In this test we show the behavior of the model in presence of a perturbation. The perturbation is represented, at time $t=0$, by a small region where vehicles are moving slower than elsewhere (and

therefore their density is higher). A typical example is given by *sags*, which are road sections along which gradient changes significantly from downwards to upwards [62].

Microscopic vehicles are immediately activated in the region of perturbation (Fig. 3.9a), and, for a short time the multi-scale model and the LWR model behave similarly (Fig. 3.9b). At the microscopic level, it is clear that vehicles decelerates when get closer to the perturbation and then accelerate again (Fig. 3.9b). After that, the LWR model tends to smear out the perturbation as usual. The multi-scale model, instead, *self-sustains* the perturbation, which does not disappear, at least for a certain time (Fig. 3.9c). Around time step $n=900$ microscopic vehicles disappear since the perturbation is no longer strong enough to destroy the equilibrium, and the multi-scale model turns to be the LWR one (Fig. 3.9d).

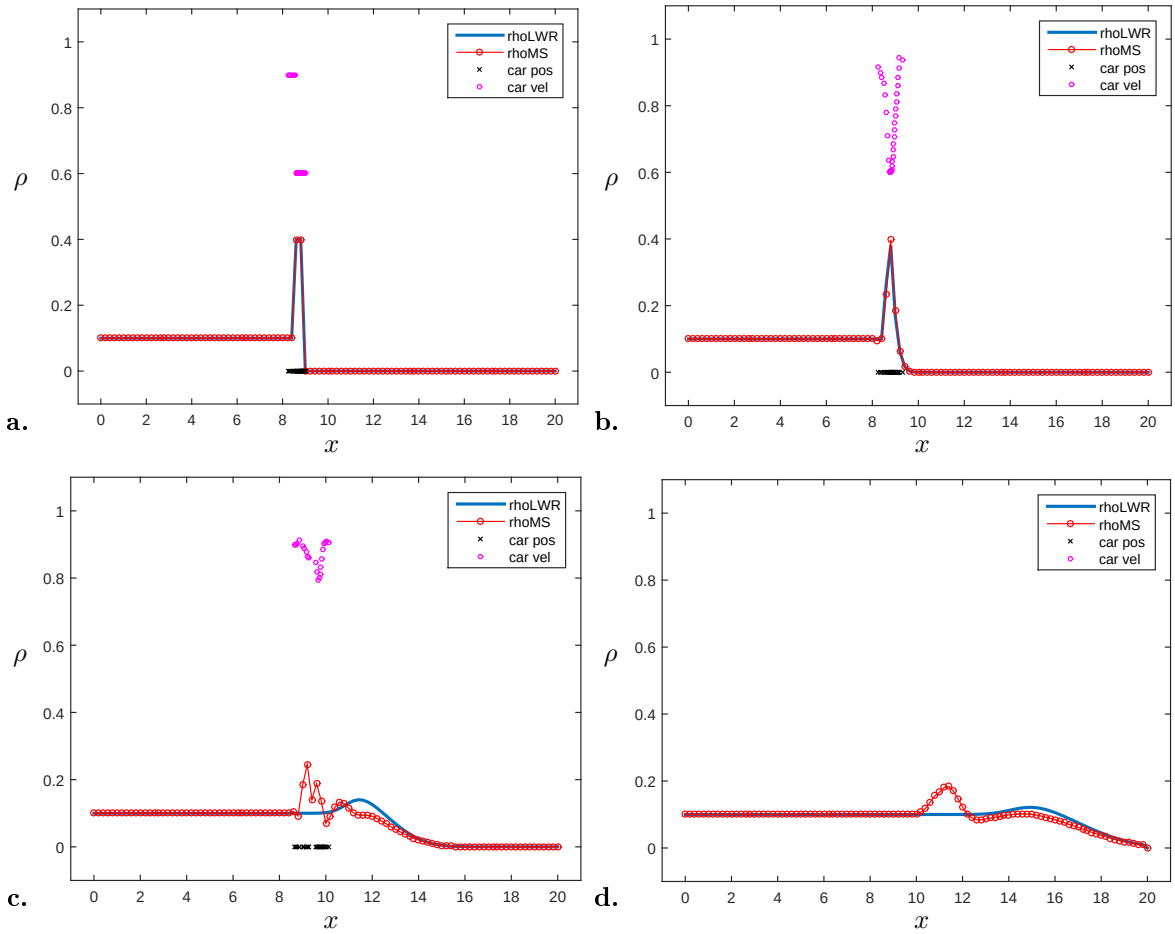

Figure 3.9: Test 3: **a.** $n=1$, **b.** $n=22$, **c.** $n=441$, **d.** $n=926$.

Chapter 4

Delayed LWR model on networks

In this chapter we deal with delayed models, which describe explicitly how past events influence the current evolution of a system by means of a time delay term. This kind of models encloses a wide area of applications, from epidemiology to traffic flow, as seen in the Introduction.

We will introduce a time delay in first order traffic flow model, which will allow us to introduce bounded accelerations as in the previous chapter, avoiding upgrading to second order models. The delay will be involved in the flux term, following the approach presented in [12]. The analytical results are still under investigations.

4.1 Derivation of the model

As shown in Sec. 1.3.3, we are able to derive macroscopic models starting from microscopic scale. Indeed let us consider the delayed microscopic model introduced by Newell in 1961 [59]:

$$\dot{X}_k = V\left(\frac{X_{k+1}(t-T)-X_k(t-T)}{\ell_N}\right) \qquad k=1,\dots,N-1, \tag{4.1.1}$$

where T is the reaction time, i. e. the delay. This means that now we face a delayed differential equation (DDE) system. Consequently, we have to provide an initial history function as initial data defined on $[-T,0]$, in order to guarantee the well-posedness of the problem in $[0,T_f]$.

Our aim is to explicitly include the delay in the flux function at the macroscopic scale avoiding the diffusion approximation as in [69], where the macroscopic model was obtained using a Taylor approximation.

Let us briefly recall that the density can be recovered by the microscopic quantities:

$$\rho_j(t) = \frac{\ell_N}{X_{k+1}(t)-X_k(t)} \tag{4.1.2}$$

as shown in Sec.1.3.3. Moreover, inserting this formula into (4.1.1), we get

$$\dot{X}_k = V\left(\frac{1}{\rho_j(t-T)}\right), \tag{4.1.3}$$

which is equivalent to $\dot{X}_k = v(\rho_j(t-T))$, since v is the macroscopic counterpart of V. Passing to the limit on the number of vehicles, i.e. $N \to \infty$, and keeping in mind that

$$\partial_t \frac{1}{\rho_j(t)} = \partial_t \frac{X_{k+1}(t) - X_k(t)}{\ell_N} = \frac{v(\rho_{j+1}(t-T) - \rho_j(t-T))}{\ell_N}, \tag{4.1.4}$$

we end up with

$$\partial_t \frac{1}{\rho(y,t)} - \partial_y v(\rho(y,t-T)) = 0. \tag{4.1.5}$$

Changing the coordinates from the Lagrangian to the Eulerian ones, i.e. $\rho(x,t) = \frac{1}{\rho(y,t)}$, we obtain

$$\partial_t \rho(x,t) + \partial_x \big(\rho(x,t) v(\rho(x,t-T))\big) = 0 \tag{4.1.6}$$

that is the Delayed LWR model (DLWR). To avoid any confusion, we will assume $v(\rho(x,t)) = 1 - \rho(x,t)$. Note that if the delay tends to 0, we recover the classical LWR model, see Fig. 5.12. On the other hand, if the delay is large enough, cars can overtake or crash each other, for example when a vehicle suddenly brakes and the car following is not reacting in time to slow down, see Fig. 5.11.

4.1.1 Numerical approximation

Equation (4.1.6) is numerically solved with the Staggered Lax Friedrichs scheme (1.2.24).

In order to observe the delay effect on the solution, let us compare the solutions coming from the standard LWR model and the delayed LWR one with $T = 15\Delta t$, with Dirichlet boundary conditions.

Starting from the same initial data, see Fig. 4.1 (left), we note how the DLWR density is sharper than

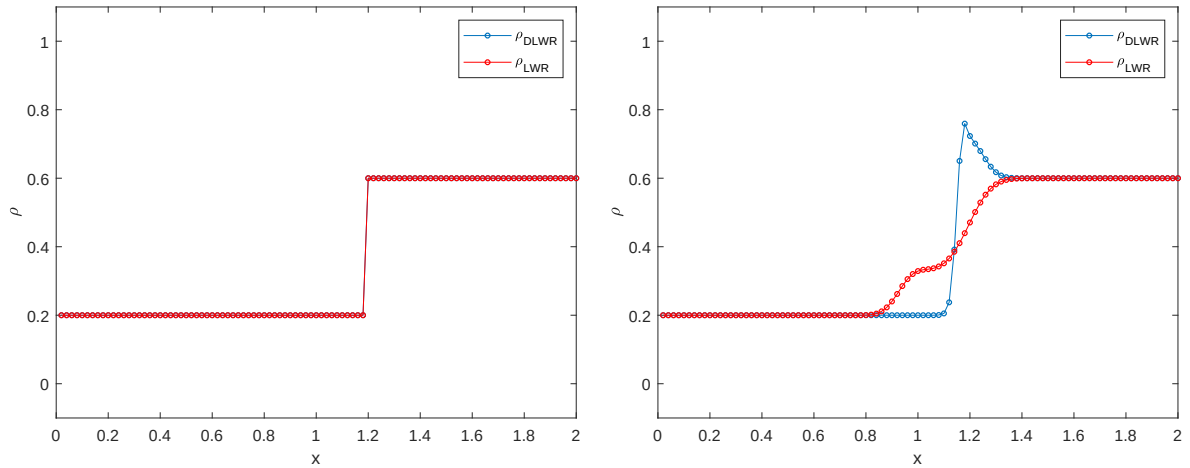

Figure 4.1: Density on network: without delay (red line), with delay (blue line) at time $n = 0$ and $n = 25$.

the other one. A little queue is formed immediately since vehicles' decelerations are not instantaneous, see Fig. 4.1(right). This behaviour is more and more evident as time increases, Fig. 4.2.

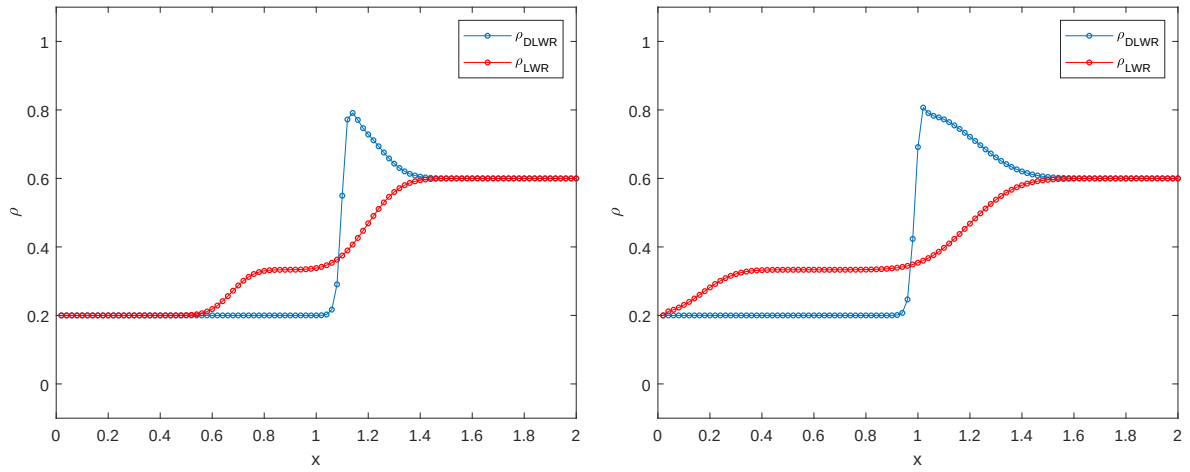

Figure 4.2: Density on network: without delay (red line), with delay (blue line) at time $n=50$ and $n=100$.

4.2 Extension to the network

Let us now focusing on the possibility to extend the delayed model on the road network. As shown in Sec.1.3.4, we can solve separately (4.1.6) along the roads and then analyse what happens at the junctions, following the classical approach. To do that, the "classical" supply and demand functions [40] (introduced in Sec.1.3.4.1)

$$D_e(\rho)=\begin{cases} f(\rho) & \rho\leq\sigma \\ f^{max} & \rho>\sigma, \end{cases} \qquad S_e(\rho)=\begin{cases} f^{max} & \rho\leq\sigma \\ f(\rho) & \rho>\sigma, \end{cases} \tag{4.2.7}$$

do not seem suitable for solving the problem at the junctions. Indeed we have to take into account the delay in the flux term since it depends on the velocity. We extend the definition of supply and demand functions in the case of delay as follows:

$$D_e(\rho,\rho^D)=\begin{cases} \rho(1-\rho^D) & \rho\leq\sigma,\rho^D\leq\sigma \\ \rho(1-\sigma) & \rho\leq\sigma,\rho^D>\sigma \\ \sigma(1-\rho^D) & \rho>\sigma,\rho^D\leq\sigma \\ \sigma(1-\sigma) & elsewhere, \end{cases} \qquad S_e(\rho,\rho^D)=\begin{cases} \rho(1-\rho^D) & \rho>\sigma,\rho^D>\sigma \\ \rho(1-\sigma) & \rho>\sigma,\rho^D\leq\sigma \\ \sigma(1-\rho^D) & \rho\leq\sigma,\rho^D>\sigma \\ \sigma(1-\sigma) & elsewhere, \end{cases} \tag{4.2.8}$$

where $\rho^D=\rho(x,t-T)$ stands for the delayed density.

Remark 4.2.1. *In order to avoid complications with supply and demand functions, one could in principle follow the multipath approach, see Sec. 1.3.4.3. Unfortunately, a Riemann solver is needed to apply the Godunov scheme, but we are not able to solve explicitly a Riemann problem with the delay in the flux term.*

Actually the multipath approach does not work with the Lax Friedrichs algorithm 1.2.21 as it is proved by the following counterexample.

COUNTEREXAMPLE: *Let us recall the merging scenario, see Example 2.3.2, where we have two incoming and one outgoing roads. From the multipath point of view, we have two possible paths, thus*

two sub-densities $\mu_1(\mathrm{E}_1,\mathrm{E}_1')$ *and* $\mu_2(\mathrm{E}_2,\mathrm{E}_1')$. *Assume that* $\mu_1(\mathrm{E}_1,\mathrm{E}_1')=\mu_2(\mathrm{E}_2,\mathrm{E}_1')=\sigma=\frac{1}{2}\rho_{\max}$. *The Lax-Friedrichs algorithm for the multipath approach is obtained from* (2.3.17) *and* (2.3.18) *by replacing* G *with the Lax Friedrichs flux defined in 1.2.21. Simplifying the notations and assuming that* J *is the merging cell, we can write the scheme as follows:*

$$\mu_{1,J}^{n+1}=\frac{1}{2}\left(\mu_{1,J+1}^n+\mu_{1,J-1}^n\right)-\frac{\Delta t}{2\Delta x}\left(\frac{\mu_{1,J+1}^n}{\rho_{J+1}^n}f(\rho_{J+1}^n)-f(\rho_{J-1}^n)\right)$$
$$\mu_{2,J}^{n+1}=\frac{1}{2}\left(\mu_{2,J+1}^n+\mu_{2,J-1}^n\right)-\frac{\Delta t}{2\Delta x}\left(\frac{\mu_{2,J+1}^n}{\rho_{J+1}^n}f(\rho_{J+1}^n)-f(\rho_{J-1}^n)\right).$$

We are now ready to compute the total density $\rho_{1,J}^{n+1}$, *reminding that the assumption is* $\mu_{1,J-1}^n=\mu_{2,J-1}^n=\mu_{1,J+1}^n=\mu_{2,J+1}^n=\frac{1}{2}\rho_{\max}$.

$$\begin{aligned}
\rho_J^{n+1}=&\mu_{1,J}^{n+1}+\mu_{2,J}^{n+1}=\\
&=\frac{1}{2}\left(\mu_{1,J+1}^n+\mu_{1,J-1}^n\right)-\frac{\Delta t}{2\Delta x}\left(\frac{\mu_{1,J+1}^n}{\rho_{J+1}^n}f(\rho_{J+1}^n)-f(\rho_{J-1}^n)\right)\\
&\quad+\frac{1}{2}\left(\mu_{2,J+1}^n+\mu_{2,J-1}^n\right)-\frac{\Delta t}{2\Delta x}\left(\frac{\mu_{2,J+1}^n}{\rho_{J+1}^n}f(\rho_{J+1}^n)-f(\rho_{J-1}^n)\right)\\
&=\frac{1}{2}\left(\mu_{1,J+1}^n+\mu_{1,J-1}^n+\mu_{2,J+1}^n+\mu_{2,J-1}^n\right)\\
&\quad-\frac{\Delta t}{2\Delta x}\left(\frac{\mu_{1,J+1}^n}{\rho_{J+1}^n}f(\rho_{J+1}^n)-f(\rho_{J-1}^n)+\frac{\mu_{2,J+1}^n}{\rho_{J+1}^n}f(\rho_{J+1}^n)-f(\rho_{J-1}^n)\right)\\
&=\frac{1}{2}(\sigma+\sigma+\sigma+\sigma)-\frac{\Delta t}{2\Delta x}(-f(\sigma))\\
&=\rho_{\max}+\frac{\Delta t}{2\Delta x}f(\sigma).
\end{aligned}$$

Then we have $\rho_J^{n+1}>\rho_{\max}$ *since* $f(\sigma)>0$, *which is physically unacceptable.*

4.3 Numerical tests

Let us focus on the numerical simulations on network. First of all, we will describe what happens on small networks (i.e. three arcs and one junction) in order to understand the density behaviour in presence of single junctions. After that, we will be ready to deal with a more complicated network.

We will show the density evolution in time. Moreover, we will compute the exit time, that is the time needed to empty the whole network, in order to compare different solutions coming from different delay values. To do that, we assume null Dirichlet boundary condition on the network incoming nodes and null Neumann condition on the network outgoing nodes, i.e. cars can only exit from the network.

4.3.1 Diverging case

Let us consider a network with three arcs and one junction, with one incoming road, called $\mathrm{E}_1=1$ and two outgoing roads, $\mathrm{E}_1'=2$ $\mathrm{E}_2'=3$.

The corresponding fluxes at the junction are:

$$\begin{aligned}\widehat{\gamma}_2 &= \min\{\alpha_{2,1} D_1(\rho_1,t), S_2(\rho_2,t)\}\\ \widehat{\gamma}_3 &= \min\{\alpha_{3,1} D_1(\rho_1,t), S_3(\rho_3,t)\}\\ \widehat{\gamma}_1 &= \widehat{\gamma}_2 + \widehat{\gamma}_3,\end{aligned} \tag{4.3.9}$$

In the following test we assume $\alpha_{3,1} = \alpha_{2,1} = \frac{1}{2}$ and the initial data: $\rho_1^0 = 0.9$, $\rho_2^0 = 0.5$, $\rho_3^0 = 0.3$.

Our aim is pointing out the impact of the delay on the density evolution and to do that we first overlap two solutions corresponding to different delays, and then we compute the exit time at different delay values.

Looking at Fig. 4.3, we note that the density profile corresponding to the higher value of the delay shows a longer queue on the incoming road.

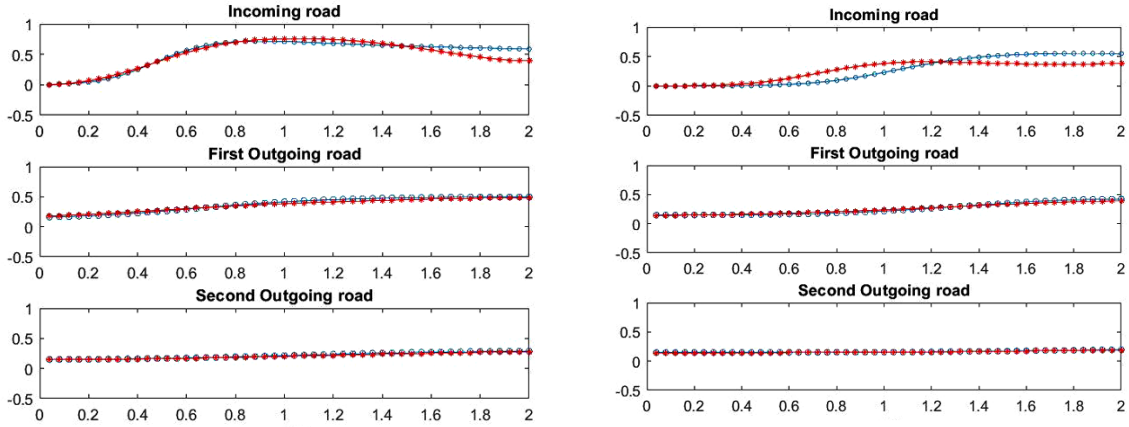

Figure 4.3: Comparison between solutions computed with a delay of 50 units of time (blue line) and 250 units of time (red line) at $T = 500\Delta t$ (left) and $T = 1000\Delta t$ (right).

Moreover, from Fig. 4.4, it is evident that the higher is the delay, the higher is the corresponding exit time, as intuition suggest, and the growth seems to be slightly super-linear.

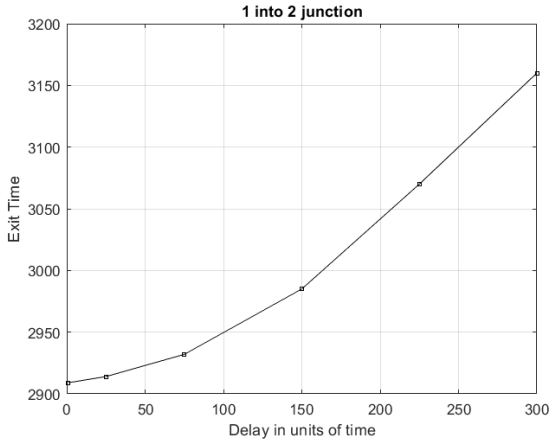

Figure 4.4: Exit Time computed at different delay values in case of diverging junction.

4.3.2 Merging case

Consider now the opposite case: a network with three arcs and one junction, with two incoming roads, called $\mathrm{E}_1=1$, $\mathrm{E}_2=2$, and one outgoing road, $\mathrm{E}_1'=3$. This is a more complicated example because we merge two arcs, so we have to be careful not to exceed $\rho_{\max}$, see Ex. 4.2.1.

The corresponding fluxes at the junction are given by:

$$\begin{aligned}
&\widehat{\gamma}_1 = \min\{D_1(\rho_1,t), \max\{qS_3(\rho_3,t), S_3(\rho_3,t) - D_2(\rho_2,t)\}\},\\
&\widehat{\gamma}_2 = \min\{D_2(\rho_2,t), \max\{(1-q)S_3(\rho_3,t), S_3(\rho_3,t) - D_1(\rho_1,t)\}\}\\
&\widehat{\gamma}_3 = \widehat{\gamma}_1 + \widehat{\gamma}_2.
\end{aligned} \tag{4.3.10}$$

where $q\in(0,1)$ is the priority coefficient, see Sec.1.3.4.1. In this framework $\bar{\gamma}^n$ for ρ_3 is given by $\widehat{\gamma}_3$, while $\tilde{\gamma}^n$ for ρ_1 and ρ_2 is respectively given by $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$.

Assume that on the first incoming road the density is $\rho_1^0=0.4$ and on the second one is $\rho_2^0=0.2$, while on the the outgoing road is initially empty. Moreover assume that $q=\frac{1}{2}$.

Looking at the density profile, Fig 4.5, we observe the formation of a queue on the first incoming road, which is higher if the delay is higher, in particular the density reaches its maximum value.

Note that nothing assures that the density does not exceed $\rho_{\max}$, indeed if the delay is too high vehicles may crash, as it happens in real life. In this case the model is not reliable any more and we have to change the dynamics.

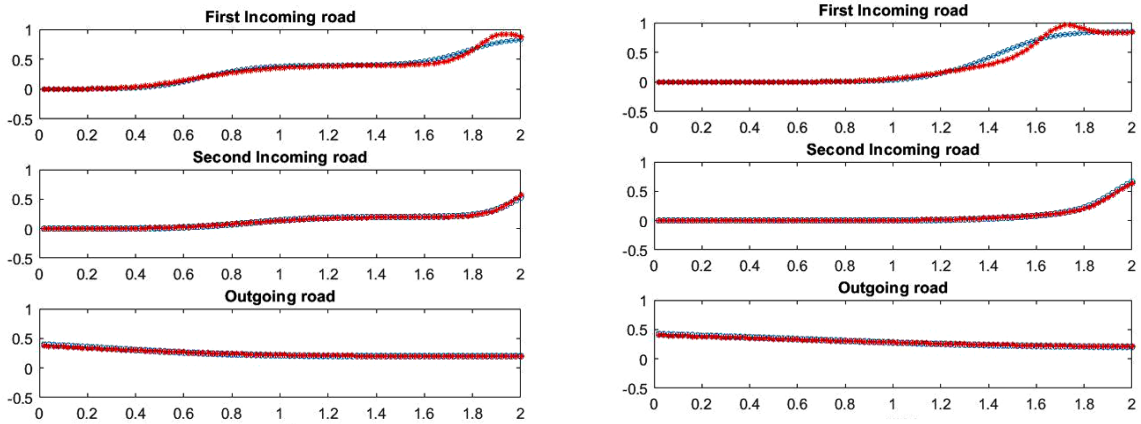

Figure 4.5: Comparison between solutions computed with a delay of 50 units of time (blue line) and 250 units of time (red line) at $t=500\Delta t$ (left) and $t=1000\Delta t$ (right).

As regards the exit time, it grows as the delay grows and the behaviour still seems to be slightly super-linear, Fig. 4.6, like the previous case.

4.3.3 Diamond network

We are now ready to join together different types of junctions and consider a more complicated network as we can see in Fig. 4.7, where we have 7 edges and 4 junctions,
we have ρ_i, $i=1,\dots,7$ densities, linked at the junctions thanks to (4.3.9) and (4.3.10) in case of diverging or merging junctions.

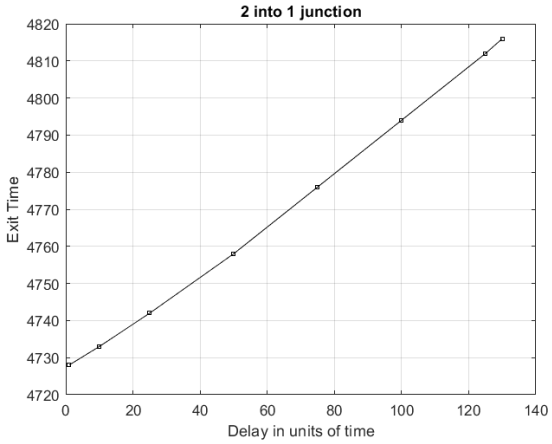

Figure 4.6: Exit Time computed at different delay values in case of merging junction.

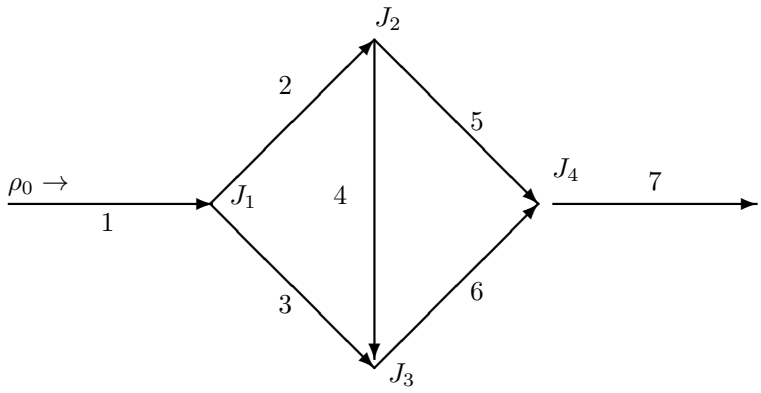

Figure 4.7: Example of "diamond" network.

Let $\rho_1^0 = 0.6$, $\rho_2^0 = 0.2$, $\rho_3^0 = 0.3$, $\rho_4^0 = 0.2$, $\rho_5^0 = 0.3$, $\rho_6^0 = 0.1$ and $\rho_7^0 = 0.2$ be the initial data, Fig. 4.8. Moreover, let the vehicles be equally distributed at junctions, i.e. $q = \frac{1}{2}$ and $\alpha_{j,i} = \frac{1}{2}$.
In the following figures the density is represented by colours on network as the colorbar suggests.

As shown in Fig. 4.9, the delay leads to higher density values next to the merging junctions, while in the classical LWR scenario the density does not turn to red, i.e. at the end of edge 3.

This behaviour is stressed for increasing time steps, Fig. 4.10, until the density reaches its maximum value only in the delayed scenario at the end of edge 6, Figs. 4.11-4.12. One of the main consequence is that the exit time is increasing with the delay.

Note that the exit time, on network, seems to have an exponential behaviour, see Fig. 4.13, so the slightly super-linear profile of the exit time shown on the single junction scenarios is emphasized.

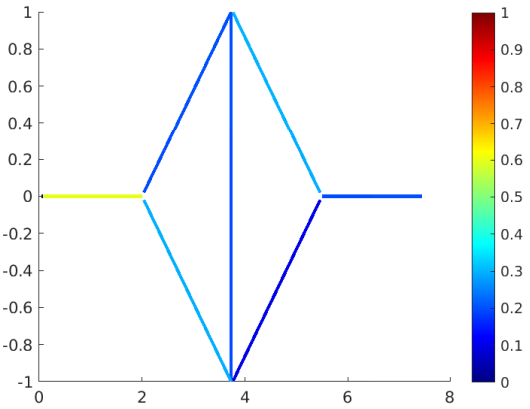

Figure 4.8: Initial data on the diamond network.

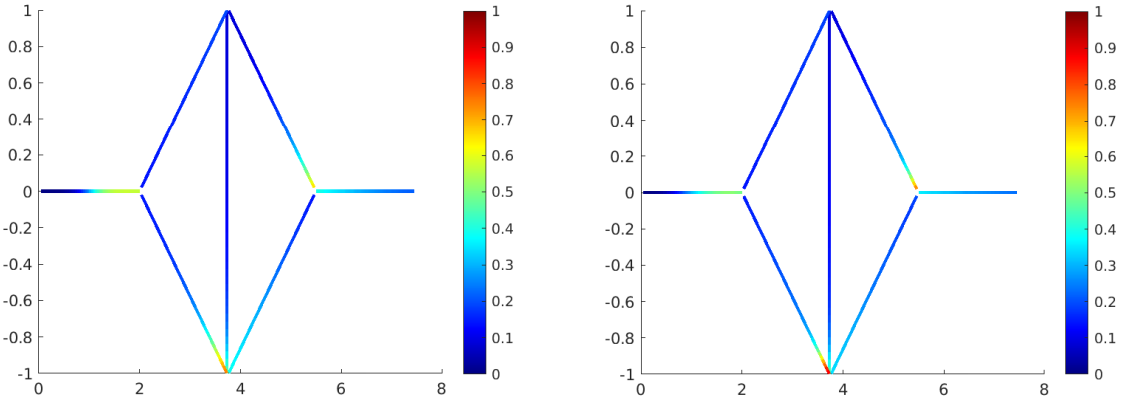

Figure 4.9: Density on network: without delay (left), with delay (right), for $n = 500$.

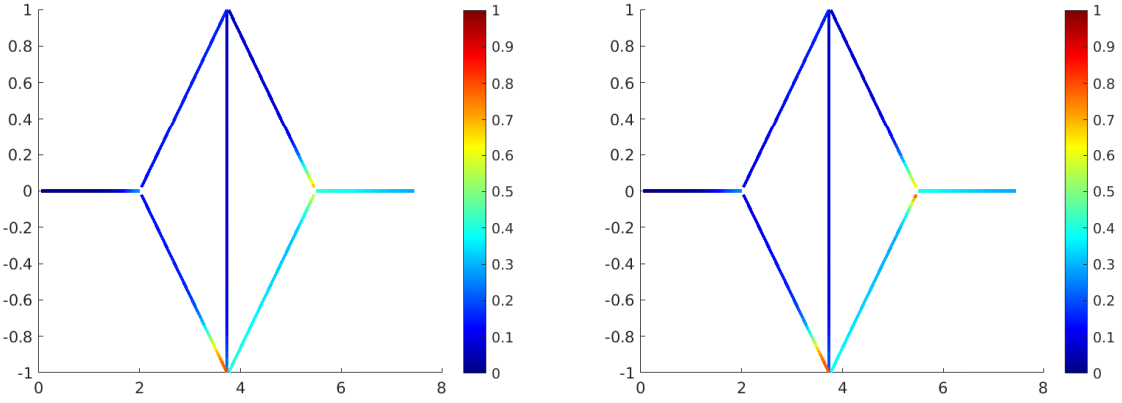

Figure 4.10: Density on network: without delay (left), with delay (right), for $n = 1000$.

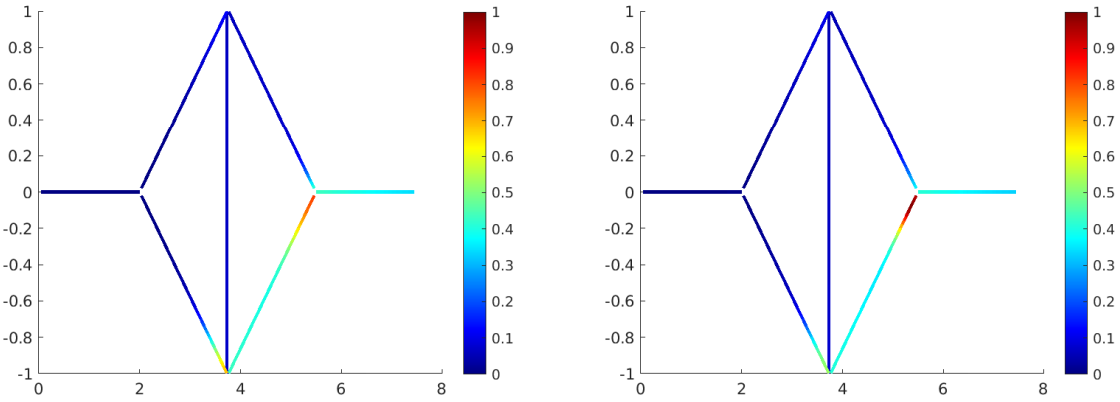

Figure 4.11: Density on network: without delay (left), with delay (right), for $n = 1500$.

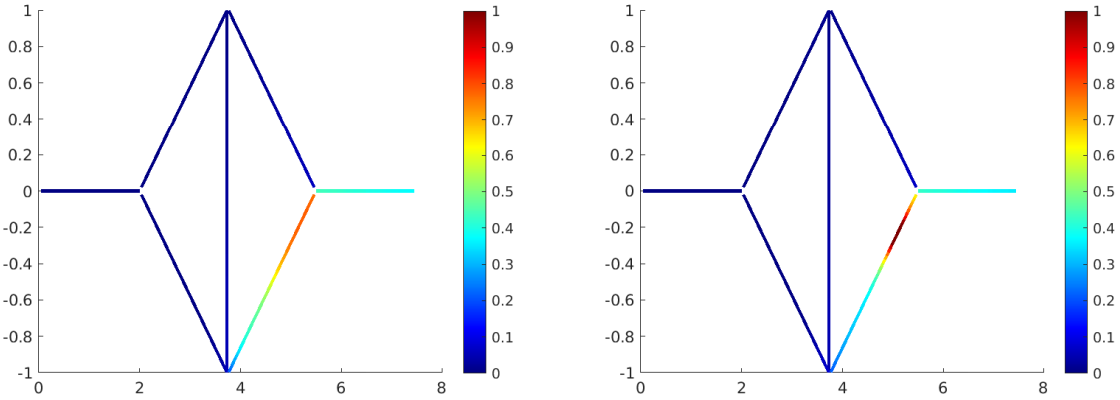

Figure 4.12: Density on network: without delay (left), with delay (right), for $n = 2000$.

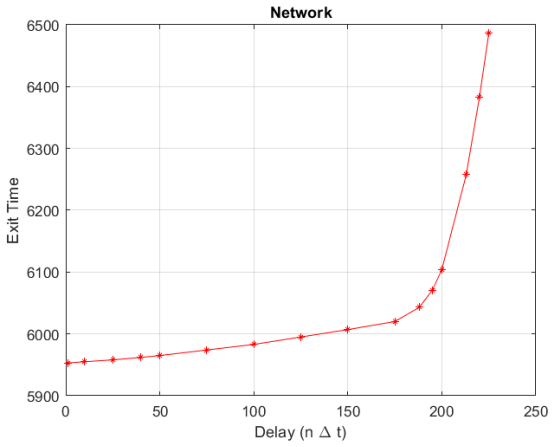

Figure 4.13: Exit Time computed at different delay values in case of the "diamond" network.

Chapter 5

Modelling Stop & Go waves

Stop & Go (S&G) waves are a typical feature of congested traffic and represent a real danger for drivers. A S&G wave is detected when vehicles stop and restart without any apparent reason, generating a wave that travels backward with respect to the cars' trajectories [36, 62].

Nowadays there are a few models able to reproduce S&G waves in literature, at least theoretically. In practice, it is very hard to find in the literature simulations which reproduce well recognizable S&G waves. In this chapter we will try to fill this gap, since modelling properly this phenomenon is crucial for developing techniques aimed at reducing it [17, 68]. For example, autonomous vehicles could be trained to avoid triggering S&G waves.

Real traffic data play a fundamental role in this framework, indeed detecting and observing real S&G waves point out the main features of this phenomenon.
In this chapter we will present a new microscopic model specifically conceived to recreate S&G waves and we will compare the numerical solutions with the real measurements.

Moreover, the new microscopic model will be coupled with the LWR model following the multi-scale approach presented in Sec. 3.

Lastly, we will try to recover the S&G waves from the delayed model, introduced in Sec. 4.

5.1 Real data and Stop & Go waves

Let us focus on real data coming from Autovie Venete S.p.A. (AV), an Italian company operating on the highways of the North-East of Italy, which has several sensors and radars to supervise and control traffic evolution all over its road network, which is about 234 km long and includes the highway Venezia-Trieste (A4), Palmanova-Udine (A23), Portogruaro-Pordenone-Conegliano (A28) and Villesse-Gorizia (A34), see Fig. 5.1.

Most roads have two lanes and only one is for overtaking. Moreover the geographical position is strategic for the trading with East-Europe, in particular with Slovenia, and roads often hold a lot of trucks, see Tab. 5.1. In order to monitor the traffic evolution on the road network, AV installed several fixed sensors, which count passing vehicles, their velocity and their road occupancy rate at each minute in each lane. Moreover they are also able to divide vehicles in five classes: cars, cars with trailer or caravan, trucks, trucks with trailer and buses. One more class has been added in order to detect all the vehicles

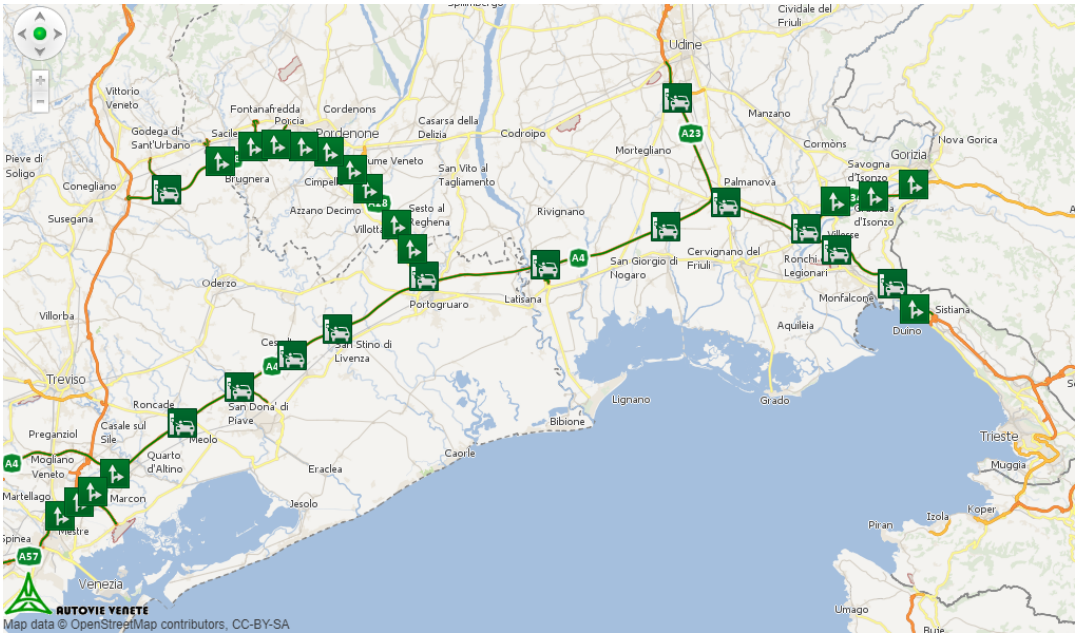

Figure 5.1: Autovie Venete road map.

that do not appear in the five classes, the "unknow". Furthermore we find in the *class 0* the total number of vehicles, i.e. the sum of all the previous classes.

The amount of available data is significant: we collected about 8,000,000 records per month. This is why we need a data analysis first.

We will focus on the period from 26.07.2016 to 25.10.2016.

First of all, our aim is understanding which class is mostly represented, i.e. which influences most the traffic evolution. To do that, we merge the two lanes corresponding to the same point of the road, i.e. which correspond to the same kilometre of the highway. For example, if we consider three different road points on A4, East direction, we obtain the results of Tab. 5.1.

	Detected vehicles	**Percentage**
All vehicles	6,615,488	100%
Cars	4,855,778	73.4%
Cars with trailer/Caravan	120,905	1.8%
Lorry/Truck	299,273	4.5%
Truck with trailer	1,152,817	17.4%
Bus	72,583	1.1%
Unknowns	114,132	1.8%

Table 5.1: Sensor data from 26.07.2016 until 25.10.2016.

Cars are the majority and represent the 70% of total vehicles, see Fig. 5.2, but the total impact on traffic flow must be computed taking into account also the length of vehicles. For example, let us assume that we have 100 vehicles on the road: 73 cars and 27 trucks, following Tab. 5.1. Assume also that a car is 4m long and a truck is 12m long. Therefore cars occupy $73 \cdot 4 = 292$m while trucks cover $27 \cdot 12 = 324$m. This means that even if trucks are less than cars, their impact on traffic flow is high due to the occupancy

rate.

Moreover we have to take into account that heavy vehicles' maximum speed is lower, i.e. 70-90km/h,

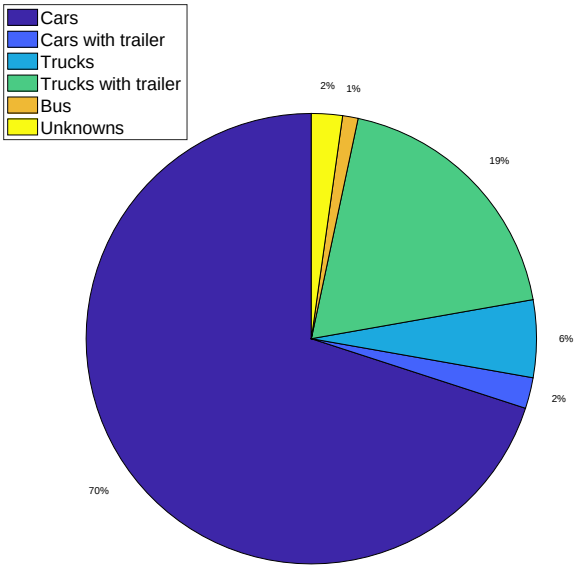

Figure 5.2: Percentage of different vehicles' classes from 26.07.2016 to 25.10.2016.

than cars' one, i.e. 130km/h, and we have to keep in mind also that the reaction time of trucks is greater than cars' reaction time.

On AV road network the *creeping* phenomenon is often observed: queues of heavy vehicles often appear on the rightmost lane, while light vehicles travel on the overtaking lane even if with reduced speed [35]. Heavy vehicles queues show the typical behaviour of S&G phenomenon, which we will focus on.

For example, let us analyse data collected on 21.09.2018 along the east direction of A4.

Starting from a small slowdown located at km 489 at time 12:10, a little queue is formed ten minutes later between km 485-486. From AV reports, we know that the queue grows and moves backward: at time 12:30 the queue is between km 482-483 and one hour later is between km 463-468, as shown in Fig. 5.3.

This is the typical S&G wave behaviour: the initial perturbation increases and propagates backward, so vehicles begin to stop and restart even far from the origin of perturbation itself.

Looking at the data from fixed sensors let us focus on the flux of heavy vehicles. We can easily observe that flux data are oscillating a lot, as the red line in Fig. 5.4 suggests. Moreover, there are minutes in which only a few vehicles are detected and this either means that a few vehicles travel along the highway or that vehicles are slow (or are stopped), i.e. the road is congested. In order to distinguish the two cases, we consider the Gaussian mean, the black line in Fig. 5.4. Considering the sensors on the East direction of A4 at km 466.7 from 12 p.m. to 16 p.m. on 21.09.2018 as before, we note that the flux detected, in particular the Gaussian mean, is almost constant until 13:30 p.m.. From 13:30 p.m. to 14 p.m. the flux is almost 0, as we expected seeing Fig. 5.3. Moreover from 15 p.m. to 15:30 p.m. another queue is detected.

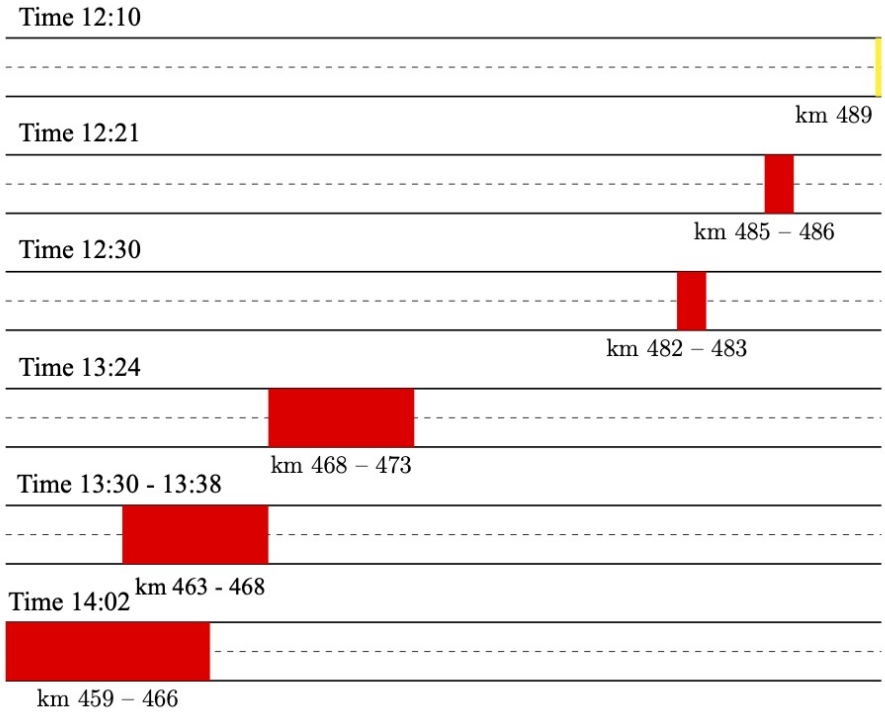

Figure 5.3: Reported S&G waves.

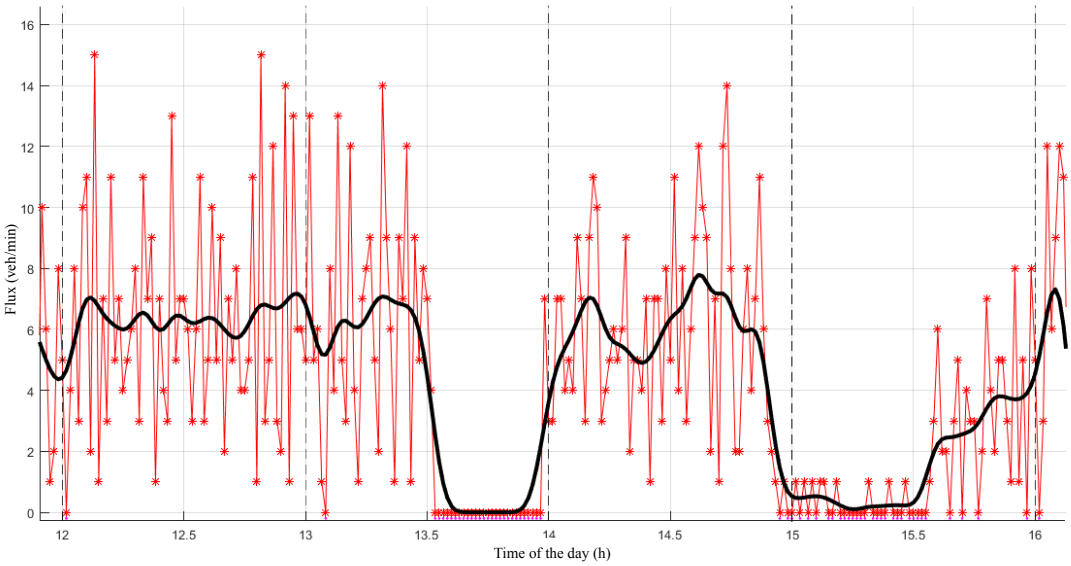

Figure 5.4: Flux of heavy vehicles on 21.09.2018.

5.2 A new microscopic model

In this section we describe a new microscopic second-order model specifically conceived to reproduce S&G waves. This model is nothing but a minimal version of the model recently introduced by Zhao and Zhang [74] to describe the dynamics of vehicles, bicycles and pedestrians in a unified framework. Our model

is 'minimal' in the sense that it is obtained from the Zhao and Zhang's model dropping all the terms which are not strictly necessary to reproduce realistic S&G waves. Recalling the second order microscopic models introduced in Sec. 1.3.2, it has the form (1.3.45) with

$$A\big(X,X',V,V';(\tau,\zeta,\Delta_{\min},V_{\max})\big)=\frac{1}{\tau}\Big(v^{ZZ}\big(X'-X\big)-V\Big) \tag{5.2.1}$$

and

$$v^{ZZ}(\Delta):=\begin{cases}0, & \Delta\le\Delta_{\min},\\ \zeta(\Delta-\Delta_{\min}), & \Delta_{\min}\le\Delta\le\Delta_{\min}+V_{\max}/\zeta,\\ V_{\max}, & \Delta\ge\Delta_{\min}+V_{\max}/\zeta.\end{cases} \tag{5.2.2}$$

Here $\zeta>0$ is a parameter and $\Delta_{\min}>\ell_N$ is the minimum critical spacing distance between the centers

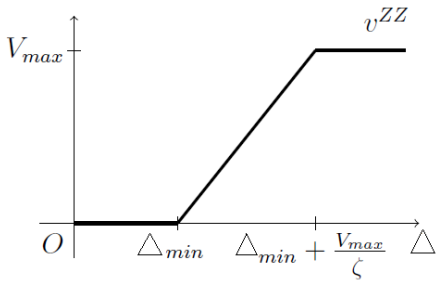

Figure 5.5: v^{ZZ} function.

of mass of a vehicle and the preceding one. Note that this minimal model, unlike the original one [74], is deterministic. Moreover, one should note that the condition $X_{k+1}(0)-X_k(0)\ge\ell_N\Rightarrow X_{k+1}(t)-X_k(t)\ge\ell_N\ \forall t$ is not *a priori* guaranteed.[1]

In Figs. 5.6-5.7 we show a typical solution to the system (1.3.45)-(5.2.1)-(5.2.2) in the case of a *circular road* of length L. Initial conditions are $X_k(0)=\frac{kL}{N+1}$ and $V_k(0)=0$, for $k=1,\dots,N$. Numerical integration is obtained by the explicit Euler scheme on a road segment $[0,L]$ with periodic boundary conditions. It can be seen that backward S&G waves are immediately generated by the small perturbation in the initial positions of the vehicles. Indeed, vehicles are initially equispaced

$$X_k(0)-X_{k-1}(0)=\frac{L}{N+1},\quad k>1$$

with the exception of the couple $(N,1)$ (first vehicle in X_1 is just in front of the N-th vehicle in X_N because of the periodic boundary conditions), for which we have

$$X_1(0)+L-X_N(0)=\frac{L}{N+1}+L-\frac{NL}{N+1}=\frac{2L}{N+1}.$$

Small perturbations in the initial velocity lead to similar effects as well.

Now let us consider a straight road in order to simulate the trajectories reported in AV real data. Starting with empty road, we use the flux data coming from sensors in order to recover left boundary

[1] The question arises why this condition should hold true in the context of traffic modeling, considering the fact that rear-end collisions are actually possible in real life. We have to be careful that in this case the model is not reliable any more and we have to change the dynamics.

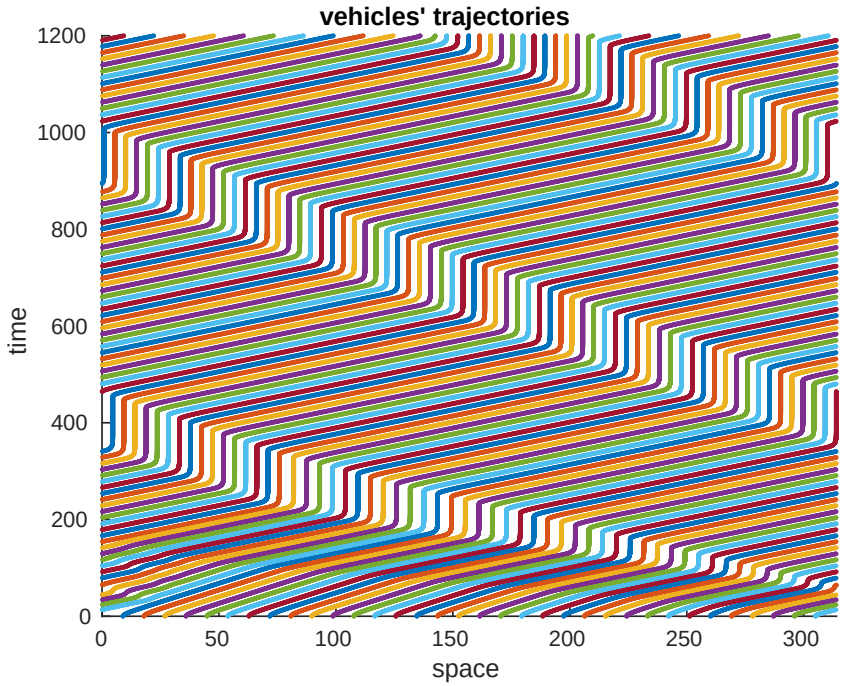

Figure 5.6: Space-time trajectories of vehicles obeying to the system (1.3.45)-(5.2.1)-(5.2.2) with $N=34$, $\alpha=0.6$, $\Delta_{\min}=7.89$, $V_{\max}=1$, $\tau=4.86$, $L=314$.

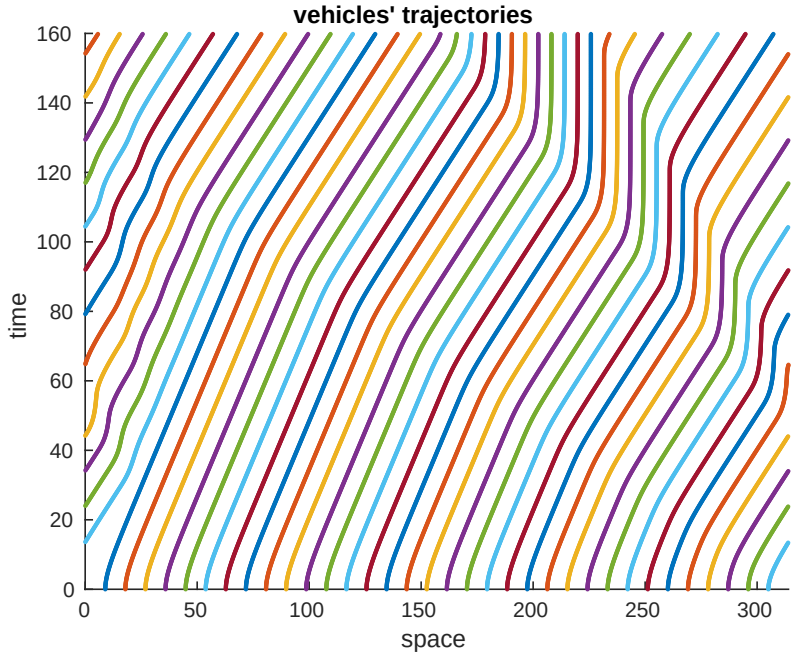

Figure 5.7: Zoom of the trajectories shown in Fig. 5.6 around initial time. It is well visible the emergence of the S&G wave from the interaction between the first and the last vehicle.

condition. These measurements are referring to the S&G wave detected on 21.09.2018 on the East direction of A4, as above, see Fig. 5.3-5.4.

Let us modify the model (5.2.1)-(5.2.2) considering two different reaction times: one for accelerations

and the other for decelerations. The system becomes as follows:

$$\begin{cases} \dot{X}_k(t) = V_k(t) \\ \begin{cases} \dot{V}_k(t) = \frac{1}{\tau_A}\left(v^{ZZ}\left(X_{k+1}(t) - X_k(t)\right) - V_k(t)\right) & \text{if } v^{ZZ} - V_k > 0 \\ \dot{V}_k(t) = \frac{1}{\tau_F}\left(v^{ZZ}\left(X_{k+1}(t) - X_k(t)\right) - V_k(t)\right) & \text{if } v^{ZZ} - V_k < 0 \end{cases} \end{cases} \tag{5.2.3}$$

where τ_A and τ_F are respectively the acceleration/deceleration reaction time. This choice is motivated by the fact that heavy vehicles reaction to accelerations and decelerations is different in time, i.e. braking must be more efficient (also for safety) than acceleration.

We modify also the function v^{ZZ}, introducing a discontinuity in its profile as in Fig. 5.8. After duly

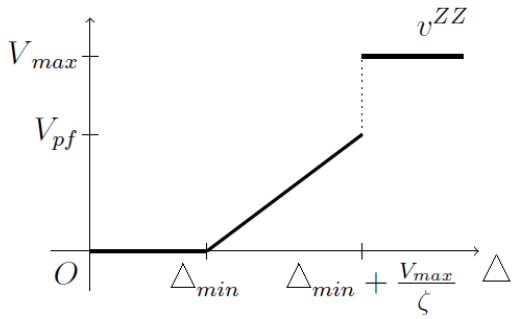

Figure 5.8: v^{ZZ} modified function.

calibration, the solution obtained with this microscopic model is really close to real data as shown in Fig. 5.9, where the light green lines stand for the length of the queue recorded in AV reports.

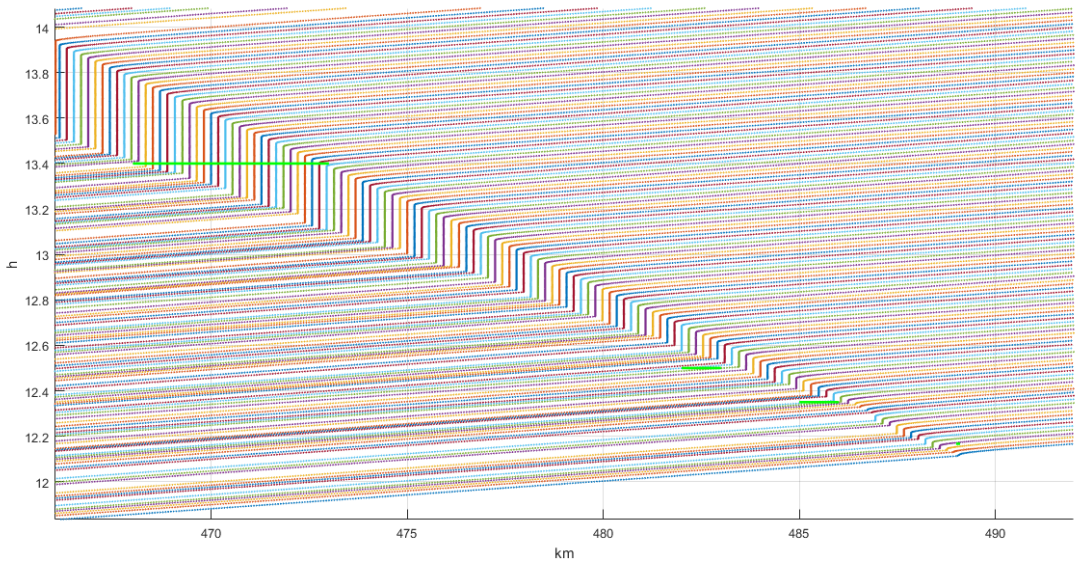

Figure 5.9: Heavy vehicles trajectories simulated with AV report measurements overlapping (green lines).

5.3 Multi-scale model for S&G waves

Let us now couple the new microscopic model described by the system (1.3.45)-(5.2.1)-(5.2.2), with the classical LWR model, using the multi-scale approach, described in Sec. 3.

A perturbation is created at initial time and microscopic vehicles are forced to be activated everywhere in the domain (Fig. 5.10a). After time δt, vehicles located in regions at equilibrium are correctly deactivated, while they stay alive around the initial perturbation (Fig. 5.10b). After that, perturbation increases until vehicles almost stop completely (Fig. 5.10c), and finally a large S&G wave is formed (Fig. 5.10d).

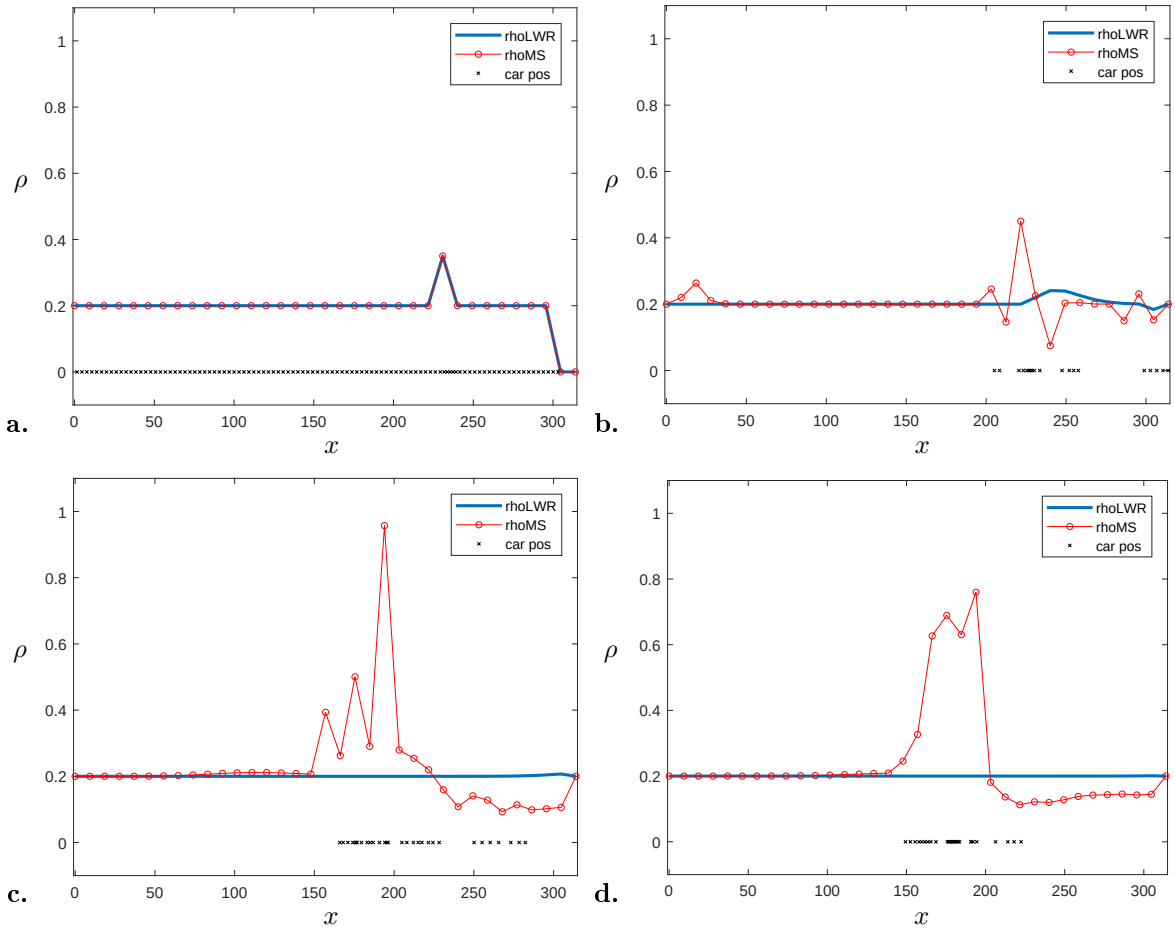

Figure 5.10: Multiscale approach for S&G waves at different time step **a.** $n=1$, **b.** $n=277$, **c.** $n=1666$, **d.** $n=2191$.

5.4 DLWR for S&G waves

In this section we see if the DLWR model is able to reproduce S&G wave on a single road when the velocity function is chosen as the macroscopic counterpart of (5.2.3). In order to avoid useless technicalities, we will choose the delay T as a multiple of the time step Δt.

We have to be careful in choosing the delay parameter: starting from the initial data Fig. 5.13a, if the

delay is too large the density grows more than one since an accident happened, i.e. cars are not reactive and brake too late, see Fig. 5.11. On the other hand, if the delay is too small the typical LWR density profile is shown, see Fig. 5.12.

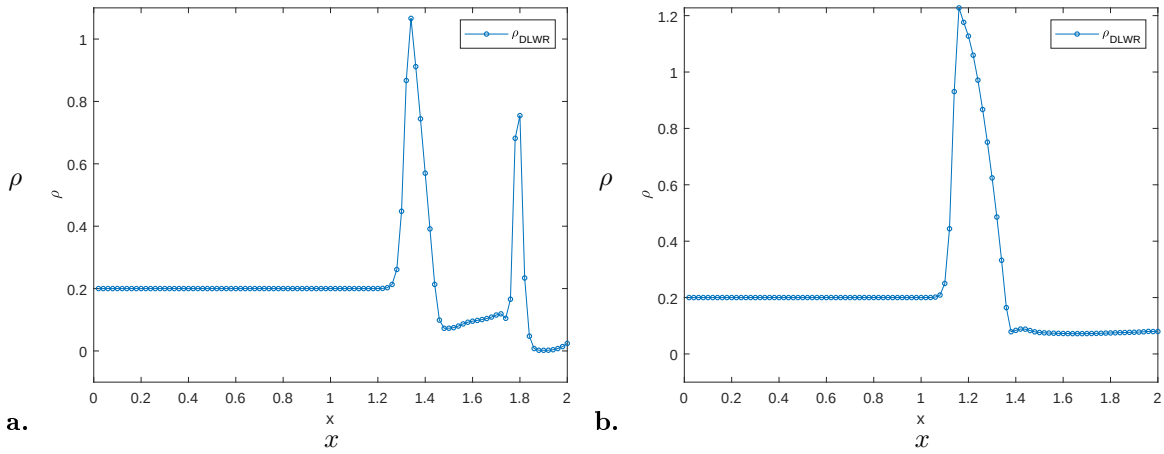

Figure 5.11: Effect of high delay value on the density profile.

Scaling the parameters of Sec. 5.3 properly, we are able to obtain qualitatively the same results of Sec. 5.3.

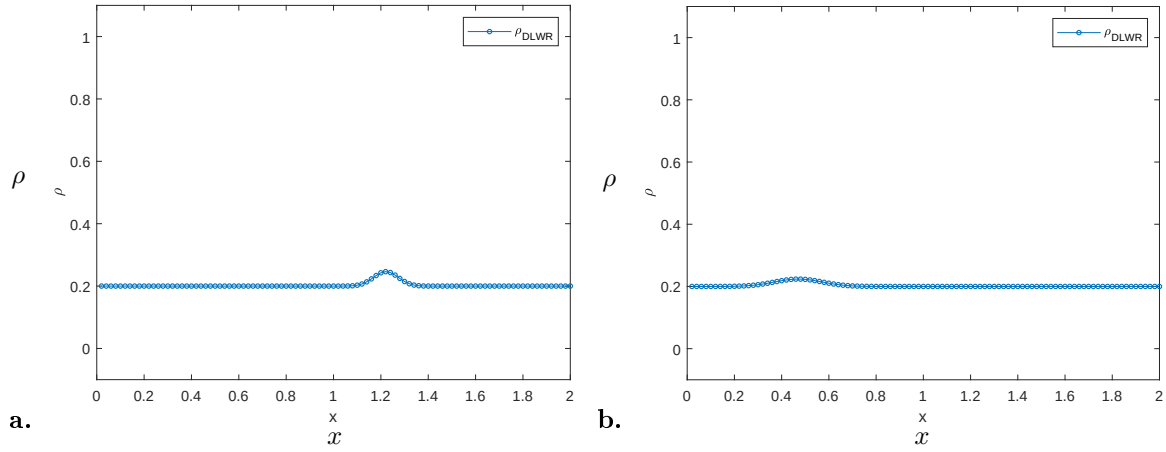

Figure 5.12: Effect of small delay value on the density profile.

5.4.1 S&G waves on networks

The DLWR model is able to reproduce S&G waves also on networks. Let us consider the diamond network Fig. 4.7 introduced in Sec. 4.3.3. Assume $\rho_1^0 = 0.9$, $\rho_2^0 = 0.4$, $\rho_3^0 = 0.3$, $\rho_4^0 = 0.2$, $\rho_5^0 = 0.4$, $\rho_6^0 = 0.1$ and $\rho_7^0 = 0.2$ as initial data, Fig. 5.14, and, moreover, that vehicles are equally distributed at junctions, i.e. $q = \frac{1}{2}$ and $\alpha_{j,i} = \frac{1}{2}$. The velocity function is $v(\rho) = 1 - \rho$.
In the following figures the density is represented by colors on network as the colorbar suggests. Looking at Fig. 5.15, we note that a congestion is formed for $n = 1000$ at the end of roads 5 and 6. The queue

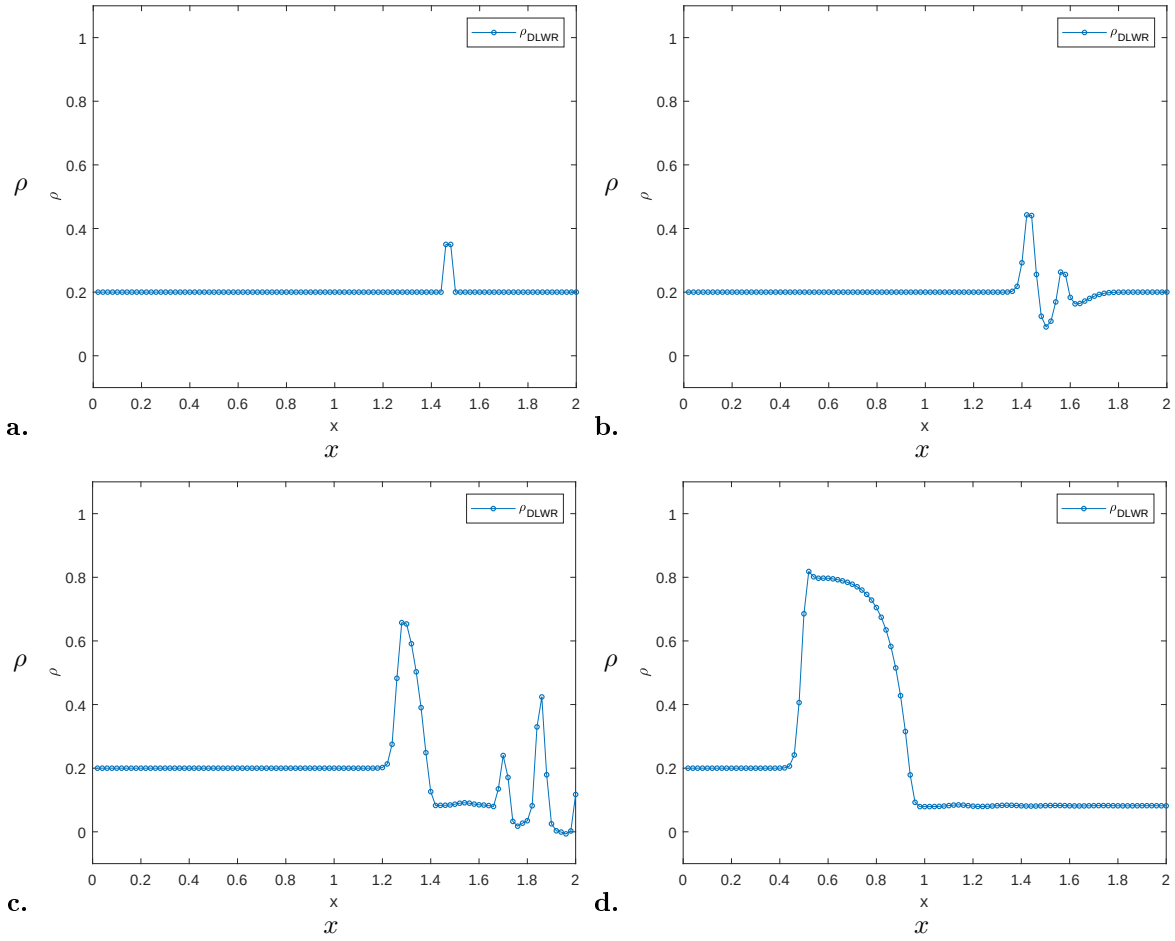

Figure 5.13: DLWR density at different time steps **a.** $n=1$, **b.** $n=75$, **c.** $n=200$, **d.** $n=500$.

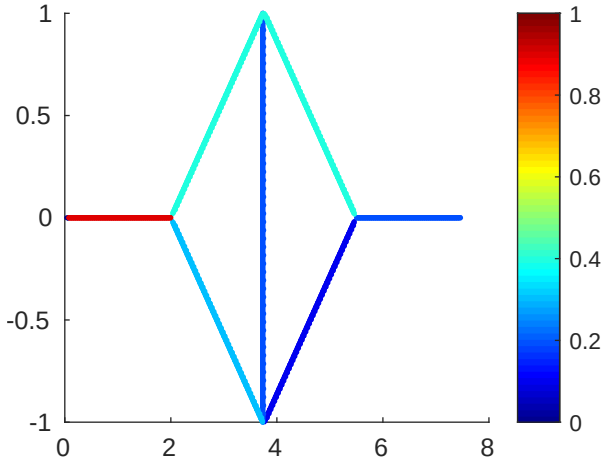

Figure 5.14: Initial data on the diamond network.

is increasing and density reaches its maximum value as we can see in Fig. 5.15 for $n=2000$ at the end of road 6. After a while, the congestion starts to propagate backward and vehicles begin to stop and restart, even if the queue at the junction is vanished. This is evident by following the evolution of the dark red spot on the road 6, Fig. 5.16.

Figs. 5.15-5.16 describe perfectly the evolution of a S&G wave.

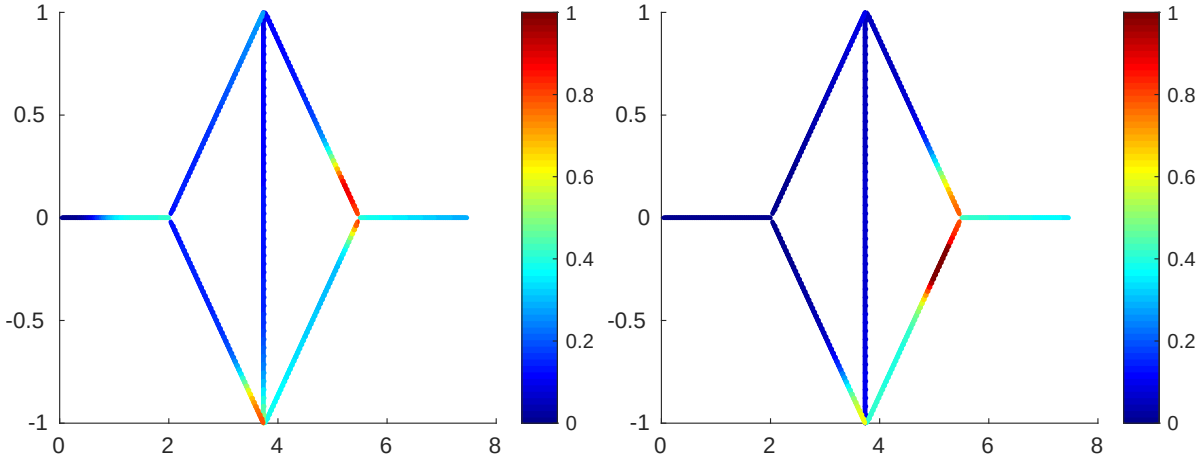

Figure 5.15: DLWR density on network at different time steps: $n = 1000$(left), $n = 2000$(right).

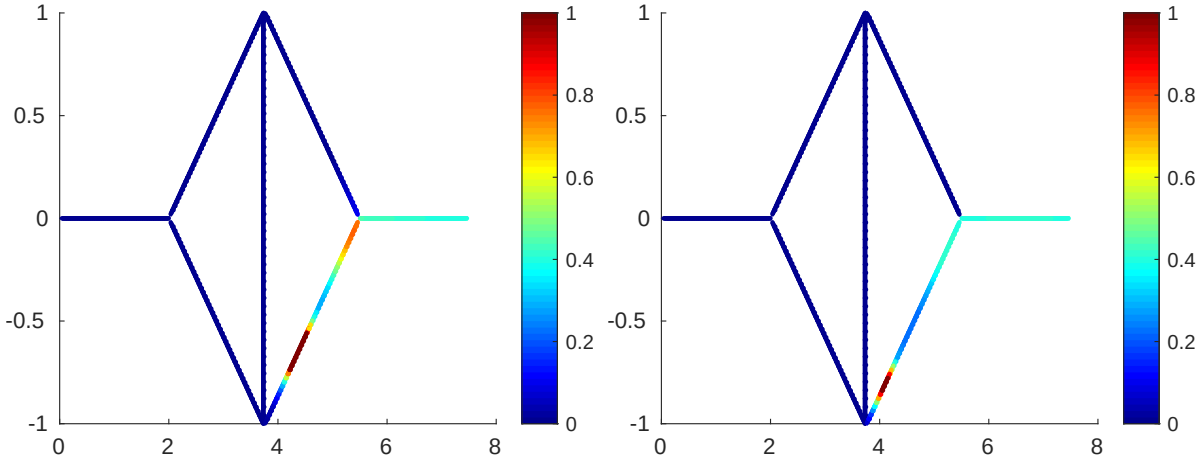

Figure 5.16: DLWR density on network at different time steps: $n = 3000$(left), $n = 3500$(right).

Conclusions

In this thesis, we proposed to employ Wasserstein distance for comparing traffic density distributions on networks. The numerical approximation of the Wasserstein distance on spaces other than the real line is not trivial. The linear-programming-based method proposed seems to be appropriate although the computational cost and memory requirements increase nonlinearly with the number of grid nodes.

As far as it concerns the sensitivity analysis of LWR model, we can say that wrong estimation of the position of vehicles at initial time does not seem to have a major impact on the final solution, at least for large times and small networks. Similar conclusions apply to wrong estimation of the fundamental diagram: errors on the capacity of the roads estimates have approximately the same impact on the final solution, and the discrepancy grows approximately linearly with respect to the parameters involved.

Conversely, wrong estimation of traffic distribution at junctions and road closures seem to have a far greater impact. Vehicles are redirected in the wrong direction at every passage across the junction, therefore the error grows in time.

Numerical investigation also shows that, in general, the sensitivity grows with the network size. Therefore we expect that LWR previsions based on real data become rapidly unusable on large networks.

Moreover, we introduced different approaches in order to reproduce second order traffic phenomena avoiding purely second order macroscopic models. The goal was to avoid to deal with systems of conservation laws, which are hard to solve from the numerical point of view. On the other hand second order effects are crucial in traffic evolution, as we have seen also from real data.

We proposed two different ways to overcome this issue: a new multiscale approach and a delayed macroscopic model. The former requires a smart mixing of two single scale models, i.e. a first order macroscopic model coupled with a second order microscopic one, in order to avoid high computational costs. The latter is based on the idea that the reaction time, which exists in braking and accelerating, has to appear in the macroscopic equation of the model. To do that, we put a time delay term in the flux of the first order macroscopic model. Our aim was to extend this approach on networks and recover S&G waves on networks.

Both approaches have shown to be perfectly able to reproduce second order effects caused by bounded accelerations, especially S&G waves, see Figs. 5.10-5.13-5.16. This is one of the main results of the thesis since simulations of S&G waves are largely missing in the literature about traffic flow modelling.

We conclude proposing some open questions strictly related to the work presented in this thesis:

- Extensions of the proposed multiscale approach: non-constant modelling parameters can be considered, i.e. the relaxation time or the tuning parameter can be assumed as functions of the density.

Allowing θ to be dependent on ρ itself, one can obtain a smooth passage from the macroscopic to the microscopic description as the traffic conditions become more and more congested. Also the relaxation parameter τ can depend on ρ, in such a way that drivers are described as more reactive if traffic conditions are more dangerous (e.g. in case of large flux characterized by a large number of vehicle moving at high velocity).

- Analysis and Control of S&G waves: techniques aimed at reducing traffic congestions and in particular S&G waves are highly desirable [17, 68] for improving drivers' safety and reducing air pollution. One can understand under which conditions S&G waves arise and disappear and the role of the stochasticity to simulate different drivers' behaviours. Starting from models able to reproduce this phenomenon, we would like to control the vehicles' dynamics in order to avoid the triggering of the S&G wave, for example with the help of trained autonomous vehicles [68].

- Many particles' limit: starting from the new microscopic model introduced in Sec. 5.2, one could recover the macroscopic counterpart passing through the many particles' limit. The idea is getting a new macroscopic model able to describe second order phenomena easier than the classical ARZ. From the numerical point of view we are investigating the relaxation schemes introduced by Natalini [1, 2, 58]. Since we have no pressure term, we consider $\varepsilon P(\rho)$ as $\varepsilon \to 0$.

- Theoretical results for delayed models: in this thesis we consider only the numerical solution to the delayed model. However it remains to investigate the numerical scheme properties, such as the mass preserving property and the local maximum principle depending on the delay. Furthermore, an estimate of the sensitivity of the solution with respect to the delay has to be studied.

Bibliography

[1] D. Aregba-Driollet and R. Natalini, *Convergence of relaxation schemes for conservation laws*, Applicable Analysis, 61 (1996), pp. 163–193.

[2] ——, *Discrete kinetic schemes for multidimensional systems of conservation laws*, SIAM Journal on Numerical Analysis, 37 (2000), pp. 1973–2004.

[3] A. Aw, A. Klar, M. Rascle, and T. Materne, *Derivation of continuum flow traffic models from microscopic follow-the-leader models*, SIAM J. Appl. Math., 63 (2002), pp. 259–278.

[4] A. Aw and M. Rascle, *Resurrection of "second order" models of traffic flow?*, SIAM J. Appl. Math., 60 (2000), pp. 916–938.

[5] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numerische Mathematik, 84 (2000), pp. 375–393.

[6] C. Bianca, M. Ferrara, and L. Guerrini, *The time delays' effects on the qualitative behavior of an economic growth model*, in Abstract and Applied Analysis, vol. 2013, Hindawi, 2013.

[7] S. Blandin, G. Bretti, A. Cutolo, and B. Piccoli, *Numerical simulations of traffic data via fluid dynamic approach*, Applied Mathematics and Computation, 210 (2009), pp. 441–454.

[8] A. Bressan and K. T. Nguyen, *Conservation law models for traffic flow on a network of roads*, Netw. Heterog. Media, 10 (2015), pp. 255–293.

[9] G. Bretti, M. Briani, and E. Cristiani, *An easy-to-use algorithm for simulating traffic flow on networks: Numerical experiments*, Discrete Contin. Dyn. Syst. Ser. S, 7 (2014), pp. 379–394.

[10] M. Briani and E. Cristiani, *An easy-to-use algorithm for simulating traffic flow on networks: Theoretical study*, Netw. Heterog. Media, 9 (2014), pp. 519–552.

[11] M. Briani, E. Cristiani, and E. Iacomini, *Sensitivity analysis of the LWR model for traffic forecast on large networks using Wasserstein distance*, Commun. Math. Sci.,. in press.

[12] M. Burger, S. Göttlich, and T. Jung, *Derivation of a first order traffic flow model of Lighthill-Whitham-Richards type*, IFAC-PapersOnLine, 51 (2018), pp. 49–54.

[13] S. Cacace, F. Camilli, R. De Maio, and A. Tosin, *A measure theoretic approach to traffic flow optimisation on networks*, European Journal of Applied Mathematics, (2018), pp. 1–23.

[14] F. Camilli, R. De Maio, and A. Tosin, *Measure-valued solutions to nonlocal transport equations on networks*, Journal of Differential Equations, 264 (2018), pp. 7213–7241.

[15] G. Chevalier, J. Le Ny, and R. Malhamé, *A micro-macro traffic model based on mean-field games*, American Control Conference, IEEE, (2015), pp. 1983–1988.

[16] G. M. Coclite, M. Garavello, and B. Piccoli, *Traffic flow on a road network*, SIAM J. Math. Anal., 36 (2005), pp. 1862–1886.

[17] R. M. Colombo and A. Groli, *Minimising stop and go waves to optimise traffic flow*, Appl. Math. Lett., 17 (2004), pp. 697–701.

[18] R. M. Colombo and F. Marcellini, *A mixed ODE-PDE model for vehicular traffic*, Math. Meth. Appl. Sci., 38 (2015), pp. 1292–1302.

[19] R. M. Colombo and E. Rossi, *On the micro-macro limit in traffic flow*, Rend. Sem. Mat. Univ. Padova, 131 (2014), pp. 217–235.

[20] E. Cristiani, *Blending Brownian motion and heat equation*, J. Coupled Syst. Multiscale Dyn., 3 (2015), pp. 351–356.

[21] E. Cristiani, B. Piccoli, and A. Tosin, *Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints*, in Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences, G. Naldi, L. Pareschi, and G. Toscani, eds., Modeling and Simulation in Science, Engineering and Technology, Birkhäuser, Boston, 2010, pp. 337–364.

[22] ——, *Multiscale modeling of granular flows with application to crowd dynamics*, Multiscale Model. Simul., 9 (2011), pp. 155–182.

[23] ——, *How can macroscopic models reveal self-organization in traffic flow?*, in 51st IEEE Conference on Decision and Control, 2012. Maui, Hawaii, December 10-13, 2012.

[24] ——, *Multiscale modeling of pedestrian dynamics*, Modeling, Simulation & Applications, Springer, 2014.

[25] E. Cristiani and S. Sahu, *On the micro-to-macro limit for first-order traffic flow models on networks*, Netw. Heterog. Media, 11 (2016), pp. 395–413.

[26] E. Cristiani and A. Tosin, *Reducing complexity of multiagent systems with symmetry breaking: an application to opinion dynamics with polls*, Multiscale Model. Simul., 16 (2018), pp. 528–549.

[27] R. V. Culshaw and S. Ruan, *A delay-differential equation model of hiv infection of cd4+ t-cells*, Mathematical Biosciences, 165 (2000), pp. 27–39.

[28] C. F. Daganzo, *Requiem for second-order fluid approximations of traffic flow*, Transportation Res. Part B, 29 (1995), pp. 277–286.

[29] M. Di Francesco, S. Fagioli, and M. D. Rosini, *Many particle approximation of the Aw-Rascle-Zhang second order model for vehicular traffic*, Math. Biosci. Eng., 14 (2017), pp. 127–141.

[30] M. Di Francesco and M. D. Rosini, *Rigorous derivation of nonlinear scalar conservation laws from follow-the-leader type models via many particle limit*, Arch. Rational Mech. Anal., 217 (2015), pp. 831–871.

[31] E. W. Dijkstra, *A note on two problems in connexion with graphs*, Numerische mathematik, 1 (1959), pp. 269–271.

[32] L. C. Evans and W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem*, no. 653, American Mathematical Soc., 1999.

[33] S. Fan and B. Seibold, *Data-fitted first-order traffic models and their second-order generalizations. Comparison by trajectory and sensor data*, Transportation Research Record, 2391 (2013), pp. 32–43.

[34] S. Fan, Y. Sun, B. Piccoli, B. Seibold, and D. B. Work, *A collapsed generalized Aw-Rascle-Zhang model and its model accuracy.* arXiv:1702.03624.

[35] S. Fan and D. B. Work, *A heterogeneous multiclass traffic flow model with creeping*, SIAM Journal on Applied Mathematics, 75 (2015), pp. 813–835.

[36] M. R. Flynn, A. R. Kasimov, J.-C. Nave, R. R. Rosales, and B. Seibold, *Self-sustained nonlinear waves in traffic flow*, Phys. Rev. E, 79 (2009), p. 056113.

[37] N. Forcadel and W. Salazar, *Homogenization of second order discrete model and application to traffic flow*, Differential Integral Equations, 28 (2015), pp. 1039–1068.

[38] N. Forcadel, W. Salazar, and M. Zaydan, *Homogenization of second order discrete model with local perturbation and application to traffic*, Discrete Contin. Dyn. Syst. Ser. A, 37 (2017), pp. 1437–1487.

[39] M. Garavello and P. Goatin, *The cauchy problem at a node with buffer*, Discrete Contin. Dyn. Syst. Ser. A, 32 (2012), pp. 1915–1938.

[40] M. Garavello and B. Piccoli, *Traffic Flow on Networks*, AIMS Series on Applied Mathematics, 2006.

[41] ——, *Advances in Dynamic Network Modeling in Complex Transportation Systems*, vol. 2 of Complex Networks and Dynamic Systems, Springer, 2013, ch. A multibuffer model for LWR road networks, pp. 143–161.

[42] ——, *Coupling of microscopic and phase transition models at boundary*, Netw. Heterog. Media, 8 (2013), pp. 649–661.

[43] ——, *Boundary coupling of microscopic and first order macroscopic traffic model*, Nonlinear Differ. Equ. Appl., 24 (2017), p. 43.

[44] S. Göttlich, M. Herty, and U. Ziegler, *Modeling and optimizing traffic light settings in road networks*, Computers & Operations research, 55 (2015), pp. 36–51.

[45] J. M. Greenberg, *Extensions and amplifications of a traffic model of Aw and Rascle*, SIAM J. Appl. Math., 62 (2001), pp. 729–745.

[46] M. Herty, J.-P. Lebacque, and S. Moutari, *A novel model for intersections of vehicular traffic flow*, Netw. Heterog. Media, 4 (2009), pp. 813–826.

[47] M. Hilliges and W. Weidlich, *A phenomenological model for dynamic traffic flow in networks*, Transportation Research Part B, 29 (1995), pp. 407–431.

[48] F. L. Hitchcock, *The distribution of a product from several sources to numerous localities*, J. Math. Phys., 20 (1941), pp. 224–230.

[49] H. Holden and N. H. Risebro, *A mathematical model of traffic flow on a network of unidirectional roads*, SIAM J. Math. Anal., 26 (1995), pp. 999–1017.

[50] M. Joueiai, L. Leclercq, H. van Lint, and S. P. Hoogendoorn, *Multiscale traffic flow model based on the mesoscopic Lighthill–Whitham and Richards models*, Transportation Research Record, 2491 (2015), pp. 98–106.

[51] S. N. Kružkov, *First order quasilinear equations in several independent variables*, Mathematics of the USSR-Sbornik, 10 (1970), p. 217.

[52] W. Kwon and A. Pearson, *Feedback stabilization of linear systems with delayed control*, IEEE Transactions on Automatic control, 25 (1980), pp. 266–269.

[53] C. Lattanzio and B. Piccoli, *Coupling of microscopic and macroscopic traffic models at boundaries*, Math. Models Methods Appl. Sci., 20 (2010), pp. 2349–2370.

[54] P. D. Lax, *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, vol. 11, SIAM, 1973.

[55] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Springer Basel AG, 1992.

[56] M. J. Lighthill and G. B. Whitham, *On kinematic waves. ii. a theory of traffic flow on long crowded roads*, Proc. R. Soc. Lond. Ser. A, 229 (1955), pp. 317–345.

[57] J. M. Mazón, J. D. Rossi, and J. Toledo, *Optimal mass transport on metric graphs*, SIAM J. Optim., 25 (2015), pp. 1609–1632.

[58] R. Natalini, *A discrete kinetic approximation of entropy solutions to multidimensional scalar conservation laws*, Journal of differential equations, 148 (1998), pp. 292–317.

[59] G. F. Newell, *Nonlinear effects in the dynamics of car following*, Operations Research, 9 (1961), pp. 209–229.

[60] B. Piccoli and A. Tosin, *Vehicular traffic: a review of continuum mathematical models*, in Encyclopedia of Complexity and Systems Science, R. Meyers, ed., Springer, New York, NY, 2009, pp. 9727–9749.

[61] P. I. Richards, *Shock waves on the highway*, Oper. Res., 4 (1956), pp. 42–51.

[62] B. G. Ros, V. L. Knoop, B. van Arem, and S. P. Hoogendoorn, *Empirical analysis of the causes of stop-and-go waves at sags*, IET Intell. Transp. Syst., 8 (2014), pp. 499–506.

[63] M. D. Rosini, *Macroscopic Models for Vehicular Flows and Crowd Dynamics: Theory and Applications*, Understanding Complex Systems, Springer International Publishing, 2013.

[64] E. Rossi, *A justification of a LWR model based on a follow the leader description*, Discrete Contin. Dyn. Syst. Ser. S, 7 (2014), pp. 579–591.

[65] F. Santambrogio, *Optimal Transport for Applied Mathematicians. Calculus of Variations, PDEs, and Modeling*, Birkhäuser, 2015.

[66] S. M. Sinha, *Mathematical Programming: Theory and Methods*, Elsevier Science & Technology Books, 2006.

[67] H. L. Smith, *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, vol. 57, Springer New York, 2011.

[68] R. E. Stern, S. Cui, M. L. Delle Monache, R. Bhadani, M. Bunting, M. Churchill, N. Hamilton, R. Haulcy, H. Pohlmann, F. Wu, B. Piccoli, B. Seibold, J. Sprinkle, and D. B. Work, *Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments*, Transportation Res. Part C, 89 (2018), pp. 205–221.

[69] A. Tordeux, G. Costeseque, M. Herty, and A. Seyfried, *From traffic and pedestrian follow-the-leader models with reaction time to first order convection-diffusion flow models*, SIAM Journal on Applied Mathematics, 78 (2018), pp. 63–79.

[70] C. Villani, *Topics in Optimal Transportation*, vol. 58 of Graduate Studies in Mathematics Series, American Mathematical Society, 2003.

[71] ——, *Optimal Transport: Old and New*, vol. 338 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2009.

[72] G. Visconti, M. Herty, G. Puppo, and A. Tosin, *Multivalued fundamental diagrams of traffic flow in the kinetic Fokker–Planck limit*, Multiscale Model. Simul., 15 (2017), pp. 1267–1293.

[73] H. M. Zhang, *A non-equilibrium traffic model devoid of gas-like behavior*, Transportation Res. Part B, 36 (2002), pp. 275–290.

[74] Y. Zhao and H. M. Zhang, *A unified follow-the-leader model for vehicle, bicycle and pedestrian traffic*, Transportation Res. Part B, 105 (2017), pp. 315–327.