Sapienza University of Rome
Department of Basic Sciences applied to Engineering,

Multiscale Methods for Traffic Flow on Networks

Raul De Maio

Supervisors: Prof. Fabio Camilli (Università di Roma, Sapienza), Prof. Andrea Tosin (Politecnico di Torino)

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Mathematical models in Engineering, Electromagnetism and Nanosciences of the University of Rome La Sapienza,
October 2018

Abstract

In this thesis we propose a model to describe traffic flows on network by the theory of measure-based equations. We first apply our approach to the initial/boundary-value problem for the measure-valued linear transport equation on a bounded interval, which is the prototype of an arc of the network. This simple case is the first step to build the solution of the respective linear problem on networks: we construct the global solution by gluing all the measure-valued solutions on the arcs by means of appropriate distribution rules at the vertices.

The linear case is adopted to show the well-posedness for the transport equation on networks in case of nonlocal velocity fields, i.e. which depends not only on the state variable, but also on the solution itself. It is also studied a representation formula in terms of the push-forward of the initial and boundary data along the network along the admissible trajectories, weighted by a properly defined measure on curves space. Moreover, we discuss an example of nonlocal velocity field fitting our framework and show the related model features with numerical simulations.

In the last part, we focus on a class of optimal control problems for measure-valued nonlinear transport equations describing traffic flow problems on networks. The objective is to optimize macroscopic quantities, such as traffic volume, average speed, pollution or average time in a fixed area, by controlling only few agents, for example smart traffic lights or automated cars. The measure-based approach allows to study in the same setting local and nonlocal drivers interactions and to consider the control variables as additional measures interacting with the drivers distribution. To complete our analysis, we propose a gradient descent adjoint-based optimization method and some numerical experiments in the case of smart traffic lights for a 2-1 junction.

Acknowledgements

I would like to express my gratitude to:

- My supervisors Prof. F. Camilli and Prof. A. Tosin for their support and mentoring during the last three years. A good Ph.D. period strongly depends on the supervisor and I'm well conscious that I've been extremely lucky since I have found not just one but two good supervisors who have always listened and supported my ideas and opinions. I wish them to have more students to work with.

- Paola Goatin and all ACUMES team for their kindness and help during my stay at INRIA near Nice for the OptiCoNet program. I have spent three wonderful and useful months thanks to you all.

- Simone Cacace for its help and suggestions. Working with him has been extremely enjoyable and challenging. I wish him, his wife Anna Chiara and their son all the possible luck for their life.

- All my colleagues Elisa, Riccardo, Simone, Alessandro, Francesco Antonio, Salvo and Davide for their friendship and the fun moments. I hope to still have the opportunity to meet and to joke with them and wish them to find their path to happiness.

- All my families, friends and researchers who have supported until now without which all these years wouldn't have been so fruitful and stimulating.

- My beloved Rossella who has always been next to me with her love and immense patience for which I will never find enough words to express my love and admiration.

“Was einmal gedacht wurde, kann nicht mehr zurckgenommen werden.”
Die Physiker - Friedrich Dúrrenmatt

Contents

List of Figures

Chapter 1

Introduction

Transportations and related issues have always been a key issues in every society. The technological progress and the increase in world population only made more evident its relevance and impact nowadays. Even if the improving of transportation facilities has progressively led to many benefit, such as speed and efficiency, it has simultaneously led to many issues: pollution, economical costs, incidents, inadequate infrastructure, maintenance, etc.

All these difficulties have attracted in the last century the attention of the scientific community, such as mathematicians, physicians and engineers; however thanks to the modern technologies and big amount of data nowadays, researchers have the opportunity to study and manage information from multiple sources (sensors on the highways, traffic lights, GPS data, smartphones, tech companies, etc.) to build new models in dependance on the kind of data and problem and then apply them to solve one or more issues in real situations.

The mathematical community has mainly focused on the definition of several models with the ambition to describe and predict vehicular traffic on roads. For these purpose, there have been proposed different approaches: microscopic models [50, 53, 42, 55], macroscopic ones [4, 38, 10, 26], kinetic models [34, 28, 27], etc. Since the literature on vehicular/pedestrian traffic is wide, we suggest [6, 49] for a detailed review on this topic.

All these approaches have successfully highlighted some features related to vehicular traffic such as congestions and "stop 'n go" waves; however they do not sufficiently deal two issues: scalability and wide variety of data from different sources.

The choice of the scale would depend on the phenomena we need to describe, based on the average number of involved drivers, number of junctions, network complexity, etc. However, the microscopic interactions at a junction have a big impact on the macroscopic dynamic in a large network. For this reason is important to tackle with multiscale modeling and to take advantage of all available data.

A first answer to these issues has been provided by the theory in [2, 54], then applied in [24] to the problem of pedestrian motion in euclidean spaces. With the above cited approach, the dynamic of pedestrian is described by a transport equation:

$$\begin{cases} \partial_t \mu_t + \nabla_x \cdot (v[\mu_t]\mu_t) = 0, & (x,t) \in \mathbb{R}^2 \times [0,T], \\ \mu_{t=0} = \mu_0, & \mu_0 \in \mathcal{P}_2(\mathbb{R}^2), \end{cases} \tag{1.1}$$

with the nonlocal velocity field defined as

$$v[\mu] = v_d(x) + \int_{D(x)} K(|x-y|)d\mu(y),$$

where v_d is the free flow speed, $D(x) \subset \mathbb{R}^2$ is the visual field for an agent in $x \in \mathbb{R}^2$ and $K(|x-y|)$ the interaction strength between two agents in x and y whose distance is $|x-y|$.

The main idea behind this thesis is to apply this approach, used for pedestrians, to drivers and vehicular traffic. In this way, we would be able to work with different scales and degrees of certainty. Our aim is to build a model able to describe vehicular traffic over any kind of networks and still able to show classical features such as congestions, "stop 'n go" and drivers' interaction.
If our network is equivalent to the real line, the model would immediately derive from the results in [24]. However, due to its geometric structure, the transport of a measure of a network deserve a deeper analysis to be defined properly. For this purpose, we decided to adopt a constructive approach: starting from the definition of measure transports on a single bounded road, we define the problem on networks by gluing its arcs thanks to a fixed a-priori transmission rule.

A first attempt with the measure-based approach on a single road can be found in [32, 33] in which Evers et al. describe traffic dynamic on a single (bounded) road and the behavior of drivers crossing the junction by a decrement of mass,

$$\partial_t \mu_t + \partial_x (v[\mu_t]\mu_t) = f \cdot \mu_t, (x,t) \in [0,1] \times [0,T] \tag{1.2}$$

where $f : [0,1] \to \mathbb{R}$ is a piecewise bounded lipschitz function which act as a sink/source term for the mass distribution μ_t.
In particular, the authors assume that drivers always stop at the junction and the term $f \cdot \mu_t$ is used to model the outflowing/inflowing of mass; this approach has as a consequence an exponential decay of the mass at the vertex $x = 1$ which implies that, for every $t > 0$, there is positive probability to be blocked at the junction until time t. for example, chosen $f = -a\chi_{\{1\}}$, where $a > 0$ is a decrement

rate, and denoted with ν_t the mass in $x = 1$, from the previous equation follows

$$\dot{\nu}_t = \mu_t(\{1\})v(1) - a\nu_t,$$

which implies an exponential decay for ν_t.

Even if this is a valid hypothesis and realistic in some scenarios, we believe it too restrictive and difficult to deal with on networks and in numerical schemes.

To allow an instantaneous flow of mass at the junctions, we decided to introduce measures to describe the outflowing mass along time. The dynamic of drivers on a single road is still described, as in [2, 24], by a transport equation

$$\begin{cases} \partial_t \mu + \nabla_x \cdot (v_t \mu_t) = 0, & (x,t) \in [0,1] \times [0,T], \\ \mu_{t=0} = \mu_0, & \mu_0 \in \mathcal{P}([0,1]), \\ \mu_{x=0} = \sigma_0, & \sigma_0 \in \mathcal{P}([0,T]) \end{cases} \tag{1.3}$$

where μ_0, σ_0 are, respectively, the initial distribution of cars along the road and the distribution of inflowing car along time through $x = 0$.

Here μ is not a continuous map in $C([0,T]; \mathcal{P}_2(\mathbb{R}))$ but a positive measure, with bounded mass, in $\mathcal{M}^+([0,1] \times [0,T])$ such that the measure μ can be "sliced" horizontally, i.e.

$$\mu(dxdt) = \mu_t(dx)dt,$$

with $\mu_{t=0} = \mu_0$, and vertically, i.e.

$$m(dxdt) = \mu_x(dt)dx,$$

with $\mu_{x=0} = \sigma_0$.

Since the outflow/inflow of mass is allowed, the measures μ_t does not have fixed mass along time. Hence it is necessary to adopt a different distance respect to the usual Wasserstein's one.

The existence and uniqueness of such measure, in particular of the family $\{\mu_t\}_{t\in[0,T]}$ derives from the theory in [2, 54] and it is possible to build explicitly, from the initial/boundary data μ_0, σ_0, the distribution $\mu_T \in \mathcal{M}^+([0,1])$ of cars along the road at time T and the distribution $\mu_{x=1} \in \mathcal{M}^+([0,T])$ of outflowing cars at the junction along time.

The transport equation on networks is then defined

$$\begin{cases} \partial_t \mu + \nabla \cdot (v[\mu_t]\mu) = 0, & (x,t) \in \Gamma \times [0,T], \\ \mu_{t=0} = \mu_0, & \mu_0 \in \mathcal{P}_2(\Gamma), \\ \mu_{x \in \mathcal{S}} = \sigma_0, & \sigma_0 \in \mathcal{P}(\mathcal{S} \times [0,T]), \\ \mu^j_{x=y} = \sum_{k \in \mathrm{Inc}(y)} p_{kj} \cdot \mu^k_{x=x_i}, & \forall e_j \in \mathrm{Out}(y), \forall y \in \mathcal{V} \setminus \mathcal{S}, \end{cases} \tag{1.4}$$

where μ_0 is the initial distribution of drivers on Γ and σ_0 is the data of drivers inflowing in our network along the time interval $[0,T]$ from a finite number of sources $x \in \mathcal{S}$. We will see that to define properly this problem on networks it is necessary, as expected from [38, 46], to add a condition.
In particular, we have chosen the transmission condition in the third line which is the key ingredient in the gluing procedure to obtain the solution on networks. Indeed, fixed a topological order of the network, the outflowing data introduce in the single road case can be used as inflow data for the next roads. Since we assume a conservation of mass at the junction, the transmission has to be weighted by distribution weights represented by a stochastic matrix $P = (p_{kj})$ which describes the percentage of drivers flowing across roads.

We remark that this is not the only way to define and build a transport of measure on networks to model vehicular traffic; indeed, we could adopt two other choices: direct definition and a multi-population model. In the first case, we could directly directly proposed a definition of the problem constrained on a network; however, it would have been more difficult to motivate and choose a transmission condition at the vertices, while our approach allow us to justify and introduce new ones with buffers. Otherwise, we could choose to describe driver dynamics by a multi-population based model constrained on the real line $\mathbb{R}$ where every population of drivers is characterized by paths "departure/destination". This approach is simpler and immediate from a theoretical point of view, however it adds several practical difficulties. Indeed, the geometrical complexity of a network would be transferred into the topological interactions between cars and the full knowledge of departure/destination of every driver, which is the assumption at the basis of this model, is not available in practice.

Once the model is defined, we study some applications to traffic management. Indeed, many important problems, such as pollution, car incidents' frequency, congestions, travel cost, etc., can be mathematically describe as optimal control problems where our problem (1.4) is the main constraint, i.e.

$$\min_{\theta \in \Theta} \left\{ \int_0^T L[\mu_t] dt + g(\mu_T) \right\},$$

where μ is solution of (1.4) which depends on controls in Θ a class of controls, $\int_0^T L[\mu_t]dt$ a time average cost w.r.t. μ and g a target function.
In particular, we used this framework to study the problem of optimal traffic light setting and the impact of few moving agents, such as autonomous cars, police, etc., to influence other drivers.

This thesis is organized as follows: in Chapter 2, we recall the background theory on metric space of measures, transport equation in $\mathbb{R}^d$ and multiscale modeling; in Chapter 3, it is defined the transport equation on networks in case of linear velocity field; while in Chapter 4 and 5, we investigate analytically and numerically some problem in case of (anisotropic) nonlocal velocity fields. Lastly, Chapter 6 deals with optimal control problems on networks and show applications of the proposed models to mobility optimization.

Statement of Originality

I declare that this thesis is my own original work. I confirm that I have composed this work without assistance, I have clearly referenced in accordance with departmental requirements all the sources used in this thesis.
I appreciate that any false claim in respect of this work will result in disciplinary action in accordance with university or departmental regulations and I confirm that I understand that this thesis may be electronically checked for plagiarism by the use of plagiarism detection software and stored on a third party's server for eventual future comparison.

Publications

The theory exposed in this thesis has been partially exposed in the following papers:

- S. Cacace, F. Camilli, R. De Maio and A. Tosin, *A measure theoretic approach to traffic flow optimization on networks*, accepted by European Journal of Applied Mathematics;
- F. Camilli, R. De Maio and A. Tosin, *Measure-valued solutions to transport equations on networks with nonlocal velocity*, J. Differential Equations, Vol. 254, Iss. 12, pg. 7213 - 7241, June 2018
- F. Camilli, R. De Maio and A. Tosin, *Transport of measures on networks* Networks and Heterogeneous Media, Vol. 12, No. 2, June 2017

Resumé of Activities

During the three years of the PhD Program, I've researched and developed mathematical models for vehicular traffic on network. This research has been supervised and co-authored by Prof. A. Tosin and Prof. F. Camilli. Here we propose innovative methods to describe vehicular traffic and congestion on complex network based on the theory of transport of measures. The first year I focused on studying the literature and defining rigorously the transport of measures on networks with linear velocity fields. During the second year, we have studied a more realistic scenario introducing non-local terms in the velocity field and proved the well-posedness of the respective problem. Lastly, in the last year we have studied applications and optimal control models based on the transport equation to model the impact of smart traffic lights on traffic jam and pollution. The activities of thes years have been partially funded with a grant (ca. 2000 Euros) for young researchers offered by Sapienza University. Moreover, at the beginning of the same academic year, I've spent three months at INRIA in Sophia Antipols in team ACUMES, as guest of Prof. P. Goatin, with a scholarship offered by the French Government. This experience has been fundamental to develop and understand the numerical aspects of my research.

In parallel, from the second year, in collaboration with Prof. F. Camilli and the PhD student E. Iacomini, I've studied fractional differential equations. In particular, our group focused on time-fractional mean field game, hamilton-jacobi equation and transport equations with memory terms. These activities have also been correlated to participation to workshops and congress.

I've attended the following events to improve knowledge and technical skills:

- AnCoNet, Analysis and Controls on Networks, March 2016, Padua;
- PDE Models for Multi-agent phenomena, November 2016, Rome;
- Symposium: PDEs in Socio-Economic Sciences, May 2017, Warwick, UK
- MFG2017, workshop on Mean Field Games, June 2017, Rome;

while in the following I had the opportunity to communicate my research and results with contributed talks or posters:

- contributed **talk at mathtech-wmft2017**, VIII Workshop on the Mathematical Foundations of Traffic, March 2017, Rome;
- contributed **talk at IperPV17**, September, 2017, Pavia (MI);

- departmental **talk at Inria-Sophia Antipolis**, 3rd October 2017;
- **local organizer and contributed talk, Workshop Day: Meeting on Fractional Derivatives,** 15 December, 2017, Rome (RM);
- departmental **talk at Politecnico di Torino**, 17th January 2018;
- **poster presentation at Numasp18**, University of Ferrara, April 2018;
- **Talk for PhD's students at University of Pisa**, 8th May 2018, invited by PhD. student V. Pagliari
- **Talk** on transport of measure **at SIMAI18** in Rome, July 2018;
- **Talk** on time-fractional mean field games **at SIMAI18** in Rome, July 2018;
- **Minisymposium's Talk at IFIP18** in Essen, July 2018;
- **Talk at Interactive workshop on hyperbolic equations**, September 2018, invited by Prof. A. Corli and M. Rosini.

Besides the papers which are the basis of this thesis, I've contributed to the following:

- F. Camilli and R. De Maio, *A time-fractional mean field game*, January 2018, accepted by Adv. in Diff. Eq.
- F. Camilli, R. De Maio and E. Iacomini, *A Hopf-Lax formula for Hamilton-Jacobi equations with Caputo time derivative*, submitted
- F. Camilli, R. De Maio, *Memory effects in measure transport equations*, 2018, submitted

Chapter 2

Background Theory

The aim of this chapter is to introduce the theoretical tools and results, introduced in [2, 29, 43, 54], that we will use in this thesis. We recall the basic notion about spaces of measures with a particular focus on probability measures. We highlight the connection between a fundamental operator, such as the pushforward of measures, and the transport equation. This is reinforced by the superposition principle which we will find again in the next chapters. Lastly, we focus on the possibility to build multiscale models thanks to the transport equations for measures.

2.1 Spaces of measures

Let (X, d) be a Polish space, i.e. a complete separable metric space, and $\mathcal{B} = \mathcal{B}(X)$ the associated Borel $\sigma-$algebra; we denote with $\mathcal{M}(X)$ is the set of all finite, real-valued, countably additive (signed) measures on $\mathcal{B}$. By the well-known Jordan's decomposition theorem, a measure $\mu \in \mathcal{M}(X)$ can be decomposed into two non-negative measures μ^+, μ^- such that it can be uniquely written as $\mu = \mu^+ - \mu^-$.

In particular, we focus on the set of Borel positive measures with finite mass $\mathcal{M}^+(X)$ and its subset $\mathcal{P}(X)$ of probability measures, i.e. the positive measures $\mu \in \mathcal{M}^+(X)$ with $\mu(X) = 1$.

Since $\mathcal{M}(X)$ is a subspace of the dual of bounded continuous function space $C_b(X)$, we endow it with the weakest topology, called weak* topology, over $C_b(X)^*$ which makes continuous all the linear functional $P \to P(f)$, with $f \in C_b(X)$.

Once we have fixed this topology, we provide a notion of convergence for measures: we say that a sequence $\{\mu_n\} \subset \mathcal{M}(X)$ narrowly converge to $\mu \in \mathcal{M}(X)$ as $n \to \infty$, shortly $\mu_n \rightharpoonup \mu$, if

$$\lim_{n\to\infty} \int_X f d\mu_n = \int_X f d\mu,$$

for every $f \in C_b(X)$.

In many applications, a relevant set of measure is given by positive measures with compact support, i.e $\mu \in \mathcal{M}^+(X)$ such that the set

$$\mathrm{Supp}(\mu) := \overline{\{x \in X : x \in N_x \in \mathcal{B}(X) \Rightarrow \mu^+(N_x) > 0\}}$$

is compact. This property is extended by the following:

Definition 2.1. *A finite positive measure μ is* tight *if for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset X$ such that $\mu(S \setminus K_\epsilon) < \epsilon$.*
A family of measures $M \subset \mathcal{M}^+(X)$ is uniformly tight *if for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset X$ such that $\mu(X \setminus K_\epsilon) < \epsilon$ for all $\mu \in M$.*

The tightness property is essential for our purposes. Since it is not always easy to verify the condition in the previous definition, we observe that it is equivalent to the existence of a function $\phi : X \to [0, +\infty)$, whose sublevels are compact in X, such that

$$\sup_{\mu\in M} \int_X \phi(x) d\mu(x) < +\infty. \tag{2.1}$$

Indeed, let $\{\epsilon_n\}$ be a sequence such that $\sum_n \epsilon_n < +\infty$ and $K(\epsilon_n)$ is an increasing sequence of compact sets such that $\mu(X \setminus K(\epsilon_n)) \leq \epsilon_n$ for every $\mu \in M$. Define the function

$$\phi(x) := \inf\{n \geq 0 : x \in K(\epsilon_n)\} = \sum_{n=0}^{\infty} \chi_{X\setminus K(\epsilon_n)}(x).$$

Then, ϕ satisfies (2.1). Conversely, if M satisfies the integral condition for a function ϕ, then its sublevels satisfies the properties cited in the tightness definition.
More information about a measure can be obtained by representation/comparison with another one.

Definition 2.2. *Given $\mu, \eta \in \mathcal{M}^+(X)$, we say that μ is absolutely continuous respect to η, $\mu \ll \eta$, if $\mu(E) = 0$, for $E \in \mathcal{B}(X)$, whenever $\eta(E) = 0$.*
We say that μ, η are mutually singular, $\mu \perp \eta$, if they are concentrated on two disjoint measurable sets.

These concepts are necessary to give a result on the representation of a measure μ with respect to a fixed one η.

Theorem 2.1. *Let $\mu, \eta \in \mathcal{M}^+(X)$ be $\sigma-$finite measures. Then*

- *there exists a unique pair of measure $\mu_a, \mu_s \in \mathcal{M}^+(X)$ such that $\mu = \mu_a + \mu_s$, with $\mu_a \ll \eta$ and $\mu_s \perp \eta$;*
- *there exists a unique non negative function ρ, integrable on X with respect to μ, such that $\mu_a(E) = \int_E \rho(x) d\eta(x)$, for any $E \in \mathcal{B}(X)$. The function ρ, also denoted by $\frac{d\mu_a}{d\eta}$, is called the density of μ_a respect to η;*
- *lastly, there exists measures $\mu_p, \mu_c \in \mathcal{M}^+(X)$ such that $\mu_s = \mu_p + \mu_c$ where μ_p is a discrete measure concentrated on a countable set and μ_c is the Cantor part.*

The theory previously exposed is still not complete because it does not give us any information or estimates about many useful properties such as the barycenter, the variation of the distribution or the cost functional since they can not be described by a pairing between a measure μ and a function in $C_b(X)$. We need to extend the pairing to function which are unbounded or semicontinuous. In this way we can define the p^{th} momentum or functionals relevant in modeling. Let $\mu_n, \mu \in \mathcal{M}(X)$ with $\mu_n \rightharpoonup \mu$, then

$$\liminf_{n\to\infty} \int_X g(x) d\mu_n(x) \geq \int_X g(x) d\mu(x) \tag{2.2}$$

for every lower semicontinuous function g over X.
In particular, choosing g as a characteristic function we obtain,

$$\liminf_{n\to\infty} \mu_n(A) \geq \mu(A),$$

for any open set A in X, and

$$\limsup_{n\to\infty} \mu_n(C) \geq \mu(C),$$

for any closed set C in X.

Definition 2.3. *A borel function $g : X \to [0, +\infty]$ is uniformly integrable w.r.t. a set $M \subset \mathcal{M}(X)$ if*

$$\lim_{k\to\infty} \int_{\{x: g(x)\geq k\}} g(x) d\mu(x) = 0, \quad \textit{uniformly for } \mu \in M.$$

Definition 2.4. *For $p > 0$, a measure $\mu \in \mathcal{M}(X)$ has finite p^{th} momentum if for one, hence any, $x_0 \in X$*

$$\langle \mu, d(\cdot, x_0)^p \rangle := \int_X d(x, x_0)^p d\mu(x) < +\infty.$$

Then, we denote with $\mathcal{M}_p^+(X)$ the set of positive measure with finite p^{th} momentum and with $\mathcal{P}_p(X)$ the set of probability measures with finite p^{th} momentum.

For $p > 0$ we define $\mathcal{M}_p^+(X)$ and $\mathcal{P}_p^+(X)$ as the sets of positive measures and probabilistic measures with finite p^{th} moment, i.e. for some (hence any) $x_0 \in X$,

$$\int_X d(x, x_0)^p d\mu(x) < +\infty.$$

Lemma 2.1. *Let $\{\mu_n\} \subset \mathcal{M}(X)$ be a sequence which narrowly converges to $\mu \in \mathcal{M}(X)$. If $f \in C(X)$,g l.s.c. and $|f|, g^-$ are uniformly integrable w.r.t. to the sequence $\{\mu_n\}$, then*

$$\liminf_{n\to\infty} \int_X g d\mu_n \geq \int_X g d\mu$$

$$\lim_{n\to\infty} \int_X f d\mu_n = \int_X f d\mu.$$

2.1.1 Pushforward over $\mathcal{M}(X)$

In this subsection we introduce the pushforward of measure, one of the main object in measure theory. It has a relevant role in optimal transport theory and a deep connection with the transport equations. Let X, Y be separable metric spaces, $\mu \in \mathcal{M}(X)$, and $\phi : X \to Y$ a $\mu-$measurable map.

Definition 2.5. *The pushforward $\phi\#\mu \in \mathcal{M}(Y)$ of a measure μ through ϕ is defined by*

$$\phi\#\mu(E) = \mu(\phi^{-1}(E)), \quad E \in \mathcal{B}(Y),$$

or, equivalentely,

$$\langle \phi\#\mu, f\rangle = \int_X f d(\phi\#\mu) = \int_X f \circ \phi d\mu = \langle \mu, f \circ \phi\rangle,$$

for every bounded Borel function $f : Y \to \mathbb{R}$.

It is easy to verify that the pushforward operator satisfies the following hypothesis:

- if $\mu \ll \eta \to \phi\#\mu \ll \phi\#\eta$, for any $\mu, \eta \in \mathcal{P}(X)$;
- (Composition rule) $(\phi \circ \psi)\#\mu = \phi\#(\psi\#\mu)$, for $\psi : X \to Y, \phi : Y \to Z$, and $\mu \in \mathcal{M}(X)$;
- if $\phi : X \to Y$ is continuous, then $\phi\#$ is continuous w.r.t. the narrow convergence; in particular

$$\phi(\mathrm{Supp}(\mu)) \subset \mathrm{Supp}(\phi\#\mu) = \overline{\phi(\mathrm{Supp}(\mu))}.$$

The following result provides an interesting link between sequence of maps ϕ_n and $\{\mu_n\} \subset \mathcal{M}(X)$.

Lemma 2.2. *Let $\phi_n : X \to Y$ be a sequence of Borel maps uniformly converging to ϕ on compact subsets of X and let $\{\mu_n\} \subset \mathcal{M}(X)$ a tight sequence narrowly converging to μ.*
If ϕ is continuous, then $\phi_n \# \mu_n$ narrowly converge to $\phi \# \mu$.

Proof. We restrict w.l.o.g. to sequences $\{\mu_n\} \subset \mathcal{M}^+(X)$. Taken $f \in C_b(Y)$ positive, for any compact set $K \subset X$, by the uniform convergence of ϕ_n to ϕ follows the same result for $f \circ \phi_n$ and $f \circ \phi$. Then, by (2.2) it follows

$$\begin{aligned}
\liminf_{n\to\infty} \int_X f \circ \phi_n d\mu_n &\geq \liminf_{n\to\infty} \int_K f \circ \phi_n d\mu_n = \liminf_{n\to\infty} \int_K f \circ \phi d\mu_n \\
&\geq (-\sup f) \sup_n \mu_n(X \setminus K) + \liminf_{n\to\infty} \int_X f \circ \phi d\mu_n \\
&\geq (-\sup f) \sup_n \mu_n(X \setminus K) + \int_X f \circ \phi d\mu.
\end{aligned}$$

By the tightness of $\{\mu_n\}$, the first term in the last inequality can be controlled by an arbitrary $\epsilon > 0$. Then,

$$\liminf_{n\to\infty} \int_X f \circ \phi_n d\mu_n \geq \int_X f \circ \phi d\mu.$$

The thesis follows replacing f in the previous argument with $-f$. $\square$

2.1.2 Disintegration of measures

Let $X \ni x \to \mu_x \in \mathcal{M}^+(Y)$ be a measure valued map. We are now interested in studying the map

$$x \in X \to \int_Y f(x, y) d\mu_x(y),$$

for a bounded and nonnegative Borel function $f : X \times Y \to \mathbb{R}$.
In particular, for $\nu \in \mathcal{M}^+(X)$, it is uniquely defined the measure $\mu = \nu \otimes \mu_x$, i.e.

$$\langle \mu, f \rangle := \int_X \int_Y f(x, y) d\mu_x(y) d\nu(x).$$

This representation is justified by the following theorem.

Theorem 2.2. *Let X_1, X_2 be Radon separable metric spaces, i.e. separable metric space on which every probability measure is a Radon measure. Given $\mu \in \mathcal{M}^+(X_1)$ and a Borel-measurable map*

$\phi : X_1 \to X_2$, let $\nu = \phi\#\mu \in \mathcal{M}^+(X_2)$. Then there exists a $\nu - a.e.$ uniquely determined Borel family of probability $\{\mu_x\}_{x\in X_2} \subset \mathcal{P}(X_1)$ such that

$$\mu_x(X_1 \setminus \pi^{-1}(x)) = 0, \quad \textit{for } \nu - a.e. \ x \in X_2,$$

and

$$\int_{X_1} f(z)d\mu(z) = \int_{X_2} \left(\int_{\pi^{-1}(x)} f(z)d\mu_x(z) \right) d\nu(x),$$

for every Borel-measurable map $f : X_1 \to [0, +\infty)$.

A particular case of the previous theorem is the following: let $X_1 = X \times Y$, $X_2 = X$, $\mu \in \mathcal{P}(X \times Y)$, $\nu = \pi^1\#\mu$, where $\pi^1 : X \times Y \to X$ is the projection map on the first component.
We can identify each fiber $(\pi^1)^{-1}(x)$ with Y and find a family of probability measures $\{\mu_x\}_{x\in X} \subset \mathcal{P}(Y)$ $\nu - a.e.$ uniquely determined and $\mu := \nu \otimes \mu_x$.

2.1.3 Metrics on space of measures

In recent years, the research of appropriate definition of a metric over a space of measures has been extremely vivid. In literature (see for example [2, 9, 29, 30, 43, 54, 57]) there exist a wide range of metrics which could be studied and adopted to build mathematical models. The choice strongly depends on the problem itself. In this subsection, we will discuss a small range of metrics and stress the main properties of the related metric space.

The most used metric on space of measure is the Wasserstein distance which has recently found many successful applications such as image classification, pattern recognition, deep learning, optimal transport or parameter calibration.
Given a metric space (X, d), we start with the following:

Definition 2.6. *The p^{th} Wasserstein distance over $\mathcal{P}_p(X)$ between two probability measure $\mu, \eta \in \mathcal{P}_p(X)$ defined as*

$$W_p(\mu, \eta) := \left(\inf_{\pi\in\Pi(\mu,\eta)} \int_{X\times X} d(x, y)^p dP(x, y) \right)^{1/p},$$

where $\Pi(\mu, \eta) = \{P \in \mathcal{P}(X \times X) : \pi_1\#P = \mu, \pi_2\#P = \eta\}$.

Any measure $P \in \Pi(\mu, \eta)$ is called *transference plan* between μ and η. This is due to the interpretation of W_p as global cost for transporting a mass represented by μ into a mass distribution η. Moreover the infimum is a minimum; indeed it is well known that there exist plans $P_* \in \Pi(\mu, \eta)$, called *optimal* *transference plans* such that

$$W_p(\mu,\eta) = \left(\int_{X\times X} d(x,y)^p dP_*(x,y)\right)^{1/p}.$$

The following result shows the importance of this distance.

Theorem 2.3. *For any $p \geq 1$, the Wasserstein distance W_p is a metric on $\mathcal{P}_p(X)$. Moreover, if (X,d) is complete and separable then $(\mathcal{P}_p(X), W_p)$ also is a complete separable space.*

In this thesis, it has a central role the 1^{st} Wasserstein distance over $\mathcal{P}_1(X)$

$$W_1(\mu,\eta) = \inf_{P\in\Pi(\mu,\eta)} \int_{X\times X} d(x,y)dP(x,y).$$

This choice is due to the *Kantorovich-Rubinstein's duality* which states that it can be also written as

$$W_1(\eta,\mu) = \sup_{f\in \mathrm{Lip}_1(X)} \int_X f(x)d(\mu-\nu)(x). \tag{2.3}$$

where $\mathrm{Lip}_1(X) := \{f : X \to \mathbb{R} : f \text{ is Lipschitz continuous with } \mathrm{Lip}(f) \leq 1\}$. This formulation will be largely preferred since it relies on the definition of measures as distributions. This is extremely useful to obtain continuity inequalities about our problems.
Another choice is the 2-Wasserstein distance

$$W_2(\mu,\eta) = \inf_{P\in\Pi(\mu,\eta)} \left(\int_{X\times X} d(x,y)^2 dP(x,y)\right)^{\frac{1}{2}}.$$

This choice is preferred for a variational approach for the transport equation studied in [2, 18, 51]. The relationship between W_1 and W_2 is generalized by the following result.

Proposition 2.1. *For $1 \leq p \leq q \geq +\infty$, it holds $W_p(\mu,\eta) \leq W_q(\mu,\eta)$ for every $\mu,\eta \in \mathcal{P}_p(X)$.*

The Wasserstein distance is not the unique choice. Indeed, since we want to work on $\mathcal{M}^+(X)$ where measures have different total mass, the Wasserstein distance is not anymore a suitable choice. Here we propose an alternative metric introduced by Dudley [29, 30] and chosen in [32, 33, 43] for the models therein.

Let $BL(X)$ be the subset of bounded functions in $Lip(X)$. Then, defined the bounded Lipschitz or Dudley norm $\|f\|_{BL} := \|f\|_\infty + |f|_L$, it follows that $(BL(X), \|\cdot\|_{BL})$ is a Banach space.

Observe also that $\mathcal{M}(X)$ embeds naturally in the dual space $BL(X)^*$, thanks to the linear map $\mathcal{M}(X) \ni \mu \to I_\mu \in BL(X)^*$, where

$$I_\mu(f) = \langle \mu, f \rangle := \int_X f(x) d\mu(x).$$

It is necessary to observe that the topological space, denoted with $\mathcal{M}(X)_{BL}$, which derives from $\mathcal{M}(X)$ equipped with the norm topology induced by $\|\cdot\|_{BL}$ is not generally complete.
On the other side, the space $\mathcal{M}^+(X)$ is clearly a convex cone in $\mathcal{M}(X)$ and it is closed and complete with respect to the dual norm $\|\cdot\|^*_{BL}$, defined as

$$\|\mu\|^*_{BL} := \sup_{f \in BL(X): \|f\|_{BL}=1} \left| \int_X f(x) d\mu(x) \right|.$$

It is easy to observe that, for $\mu \in \mathcal{M}^+(X)$, it holds $\|\mu\|^*_{BL} = \|\mu\|_{TV} = \mu(X)$.
From the Prokhorov's Theorem, if (X, d) is a complete separable metric space, a set of Borel measures $M \subset \mathcal{P}(X)$ is tight if and only if it is a precompact in $\mathcal{P}(X)_{BL}$.

Theorem 2.4 (see [29, 43])**.** *Let (X,d) be a complete separable metric space. Let $\{\mu_n\} \subset \mathcal{M}(X)$ be a sequence such that $\sup_n \mu_n(X) < \infty$ and for every $f \in BL(X)$, $\langle \mu_n, f \rangle$ converges. Then,*

- *$\langle \mu_n, f \rangle$ converges for every $f \in C_b(X)$;*
- *there exists $\mu \in \mathcal{M}(X)$ such that $\|\mu_n - \mu\|^*_{BL} \to 0$.*

Theorem 2.5. *Let (X, d) be a metric space and $\{\mu_n\} \subset \mathcal{M}(X)$ a tight sequence. If μ_n weak* converges to μ, then $\|\mu_n - \mu\|^*_{BL} \to 0$.*

The previous theorems imply the following results.

Corollary 2.5.1. *Let (X, d) be a complete separable metric space and $M \subset \mathcal{M}(X)$ such that $\sup_{\mu \in M} |\mu|(X) < \infty$ and M is uniformly tight. Then the weak* and the $\|\cdot\|^*_{BL}$-norm topology coincide.*

Proposition 2.2. *Let (X, d) be a complete separable metric space and $M := \{\mu \in \mathcal{M}(X) : |\mu|(X) = 1\}$. Then, the weak topology and the $\|\cdot\|^*_{BL}$ coincide.*

Thanks to these results, we can recover all the benefits obtained in the dual representation formula (2.3). On the other side, if we restrict on $\mathcal{P}_1(X)$, or equivalentely on a set of positive measures with constant mass, we have already observe that generally $\|\mu - \eta\|^*_{BL} \le W_1(\mu, \eta)$. The equality is clearly taken for any compact space X , however it is not generally true. Indeed, the following result clarifies this difference.

Theorem 2.6. *Let $\mu_n, \mu \subset P_p(X)$, with $p \geq 1$; then $W_p(\mu_n, \mu) \to 0$ iff $\mu_n \rightharpoonup \mu$ and $\int_X d(x, x_0)^p d\mu_n(x) \to \int_X d(x, x_0)^p d\mu(x)$.*

Proof. Since $W_1 \leq W_p$ for every $p \geq 1$, the W_p-convergence implies the $\| \cdot \|_{BL}^*$-convergence and the weak-*. The convergence of p^{th} momentum follows noticing, fixed $x_0 \in X$

$$\int_X d(x, x_0)^p d\mu_n(x) = W_p^p(\mu_n, \delta_{x_0}) \to W_p^p(\mu, \delta_{x_0}) = \int_X d(x, x_0)^p d\mu(x).$$

We prove the opposite. Let μ_n be a sequence of probability measures weakly converging to μ satisfying also $\int_X d(x, x_0)^p d\mu_n(x) \to \int_X d(x, x_0)^p d\mu(x)$. Fix $R > 0$ and consider the truncated function $\phi(x) = (\max\{d(x, x_0), R\})^p \in BL(X)$. Then,

$$\begin{aligned} \int_X (d(x, x_0)^p - \phi) d\mu_n &= \int d(x, x_0)^p d\mu_n - \int_X \phi d\mu_n \to \int_X d(x, x_0)^p d\mu(x) - \int_X \phi(x) d\mu(x) \\ &= \int_X (d(x, x_0)^p - \phi) d\mu \leq \int_{B_R(x_0)^c} d(x, x_0)^p d\mu. \end{aligned}$$

Due to the last inequality, for $\epsilon > 0$ we can choose R so that $\int_{B_R(x_0)^c} d(x, x_0)^p d\mu < \frac{\epsilon}{2}$, hence $\int_X (d(x, x_0)^p - \phi) d\mu_n < \frac{\epsilon}{2}$ for $n \gg 1$.
Since $(d(x, x_0) - R)^p \leq (d(x, x_0)^p - \phi(x))$, it follows

$$\int_X (d(x, x_0)^p - \phi(x)) d\mu_n, \int_X (d(x, x_0)^p - \phi(x)) d\mu < \epsilon, \quad n \gg 1.$$

Let $\pi_R : X \to \overline{B_R(x_0)}$ be the projection over $B_R(x_0)$ and observe that $d(x, \pi_R(x)) = d(x, x_0) - R$. Then

$$\begin{aligned} W_p(\mu, \pi_R \# \mu) &\leq \int_X (d(x, x_0) - R)^p d\mu \leq \epsilon \\ W_p(\mu_n, \pi_R \# \mu_n) &\leq \int_X (d(x, x_0) - R)^p d\mu_n \leq \epsilon. \end{aligned}$$

Due to π_R, these measures are concentrated on a compact set, then the weak convergence over $B_R(x_0)$ implies the W_p convergence. We can conclude:

$$\begin{aligned} \limsup_n W_p(\mu_n, \mu) &\leq \limsup_n (W_p(\mu_n, \pi_R \# \mu_n) + W_p(\pi_R \# \mu_n, \pi_R \# \mu) + W_p(\mu, \pi_R \# \mu)) \\ &\leq 2\epsilon^{1/p} + \lim_n W_p(\pi_R \# \mu_n, \pi_R \# \mu) = 2\epsilon^{1/p}. \end{aligned}$$

Since this is valid for any $\epsilon > 0$, we have the $\limsup_n W_p(\mu_n, \mu) = 0$. $\square$

2.2 Transport equation in $\mathbb{R}^d$

This section is devoted to the transport equation in case of $X = \mathbb{R}^d$ in measure-valued sense, which has been largely studied, for example in [2, 7, 18, 24, 48, 51, 54]. This equation has a central role in many applications such as optimal transport and mean field game. The problem is given by

$$\begin{cases} \partial_t \mu_t + \nabla \cdot (v_t \mu_t) = 0, & (x,t) \in \mathbb{R}^d \times [0,T], \\ \mu_{t=0} = \mu_0 & \mu_0 \in \mathcal{M}^+(\mathbb{R}^d), \end{cases} \tag{2.4}$$

where $v : (t,x) \to v_t(x) \in \mathbb{R}^d$ is a Borel-measurable velocity field which satisfies

$$\int_0^T \int_{\mathbb{R}^d} |v_t(x)| d\mu_t(x) dt < \infty. \tag{2.5}$$

Our aim in this section is to show the existence of solutions for such equation in the following sense:

Definition 2.7. *A measure-valued solution to* (2.4) *is an absolutely continuous map* $\mu : [0,T] \to \mathcal{M}^+(\mathbb{R}^d)$ *which satisfies* (2.4) *in sense of distribution,i.e.*

$$\frac{d}{dt} \int_{\mathbb{R}^d} f(t,x) d\mu_t(x) = \int_{\mathbb{R}^d} (\partial_t f(t,x) + v_t(x) \cdot \nabla f(t,x)) d\mu_t(x) dt, \tag{2.6}$$

or, equivalentely,

$$\int_{\mathbb{R}^d} f(T,x) d\mu_T(x) - \int_{\mathbb{R}^d} f(0,x) d\mu_0(x) = \int_0^T \int_{\mathbb{R}^d} (\partial_t f(t,x) + v_t(x) \cdot \nabla f(t,x)) d\mu_t(x) dt, \tag{2.7}$$

$\forall\ f \in C^1([0,T]; C_c^\infty(\mathbb{R}^d))$.

Observe that in (2.7) appears the first order partial differential equation

$$\partial_t f + v_t \cdot \nabla f = \psi,$$

whose solution depends on the characteristics associated to v_t. A key feature of this problem is related to the characteristics associated to the velocity field v_t:

Lemma 2.3 (see [2] - Characteristics of v_t)**.** *Let v_t be a Borel-measurable vector field such that* $v \in L^1([0,T]; L^\infty_{loc}(\mathbb{R}^d) \cap Lip_{loc}(\mathbb{R}^d))$*. Then, for every $x \in \mathbb{R}^d$ and $s \in [0,T]$ there exists a unique maximal solution to*

$$\begin{cases} \frac{d}{dt} \Phi_t(x,s) = v_t(\Phi_t(x,s)), & t \in [s,T] \\ \Phi_s(x,s) = x, \end{cases} \tag{2.8}$$

defined over an interval, relatively open in $[0,T]$ *and with* $s \in I$ *as internal point. Moreover, if* $|\Phi(x,s)|$ *is bounded over this interval, then it is defined over all* $[0,T]$.
Finally, if $v \in L^1([0,T]; L^\infty(\mathbb{R}^d) \cap Lip(\mathbb{R}^d))$ *then there exists a constant* $C > 0$ *such that*

$$\sup_{s,t \in [0,T]} Lip(\Phi_t(\cdot, s), \mathbb{R}^d) \leq e^C.$$

Under the hypotheses of the previous lemma, we can provide a fundamental representation formula for (2.4). In (2.7), we have the weak form of the adjoint backward equation to (2.4):

$$\begin{cases} \partial_t f + v_t \cdot \nabla f = \psi, & (t,x) \in [0,T] \times \mathbb{R}^d, \\ f(T,x) = f_T(x), & x \in \mathbb{R}^d, \end{cases}$$

where $f_T \in C^1_c(\mathbb{R}^d)$, v satisfies (2.5) and $\psi \in C_c(\mathbb{R}^d)$.
By the semigroup property of the characteristics, the solution of the adjoint backward problem is given by

$$f(\Phi_t(x,0),t) = f_T(\Phi_T(x,0)) - \int_t^T \psi(\Phi_s(x,0))ds.$$

Proposition 2.3 (see [2])**.** *Let* v_t *be a Borel velocity field satisfying all the hypotheses in Lemma 2.3 and the global bound* (2.5)*. Denoted with* Φ_t *the associated characteristic to* v_t*, solution of* (2.8)*, the map* $t \to \mu_t := \Phi_t \# \mu_0$ *is a continuous solution of (2.4) in* $[0,T]$ *in sense of Definition 2.7.*

Proof. By the continuity in $t-$variable of $\Phi_t(x,0)$ for μ_0-a.e. $x \in \mathbb{R}^d$ and the dominated convergence theorem, for every continuous and bounded function $f : \mathbb{R}^d \to \mathbb{R}$ we have

$$\lim_{s\to t} \int_{\mathbb{R}^d} f(x) d\mu_s(x) = \lim_{s \to t} \int_{\mathbb{R}^d} f(\Phi_s(x,0)) d\mu_0(x) = \int_{\mathbb{R}^d} f(\Phi_t(x,0)) d\mu_0(x) = \int_{\mathbb{R}^d} f(x) d\mu_t(x).$$

For any $f \in C^1([0,T]; C_c^\infty(\mathbb{R}^d))$ and for μ_0-a.e. $x \in \mathbb{R}^d$ the maps $t \to f(\Phi_t(x,0),t)$ are absolutely continuous and

$$\frac{d}{dt} f(\Phi_t(x,0),t) = \partial_t f(\Phi_t(x,0),t) + v_t(\Phi_t(x,0)) \cdot \nabla f(\Phi_t(x,0),t) =: A(\Phi_t(x,0),t).$$

Then,

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \left| \frac{d}{dt} f(\Phi_t(x,0),t) \right| d\mu_0(x) dt &= \int_0^T \int_{\mathbb{R}^d} |A(\Phi_t(x,0),t)| d\mu_0(x) dt \\ &= \int_0^T \int_{\mathbb{R}^d} |A(x,t)| d\mu_t(x) dt \\ &\leq \|f\|_{C^1} \left(T + \int_0^T \int_{\mathbb{R}^d} |v_t(x)| d\mu_t(x) dt \right) < \infty.\end{aligned}$$

Therefore,

$$\begin{aligned}0 = \int_{\mathbb{R}^d} f(x,T) d\mu_T(x) - \int_{\mathbb{R}^d} f(x,0) d\mu_0(x) &= \int_{\mathbb{R}^d} \left(f(\Phi_T(x,0),T) - f(x,0) \right) d\mu_0(x) \\ &= \int_{\mathbb{R}^d} \int_0^T \frac{d}{dt} f(\Phi_t(x,0),t) dt d\mu_0 \\ &= \int_0^T \int_{\mathbb{R}^d} (\partial_t f(x,t) + v_t(x) \nabla f(x,t)) d\mu_t(x) dt.\end{aligned}$$

□

The previous result states that the pushforward of the initial distribution via the characteristics of v_t is a solution of a transport equation. The next one that the sign of the solution is preserved by the transport equation. A corollary of this result is the uniqueness of the solution which implies that *every* solution is characterized by a pushforward.

Proposition 2.4. *Let σ_t be a (narrowly) continuous family in $\mathcal{M}(\mathbb{R}^d)$ solving $\partial_t \sigma_t + \nabla \cdot (v_t \sigma_t) = 0$ in $\mathbb{R}^d \times (0,T)$, with $\sigma_0 \leq 0$, such that*

$$\int_0^T \int_{\mathbb{R}^d} |v_t| d|\sigma_t| dt < \infty,$$

and the local boundness property

$$\int_0^T \left(|\sigma_t|(K) + \sup_K |v_t| + Lip(v_t, K) \right) dt < \infty,$$

for any compact $K \subset \mathbb{R}^d$. Then $\sigma_t \leq 0$ for every $t \in [0,T]$.

Proof. Fix $\phi \in C_c^\infty(\mathbb{R}^d \times [0,T])$ with $0 \geq \phi \geq 1$, and, for $R > 0$, χ_R a smooth cut-off function, such that

$$0 \geq \chi_R \geq 1, \quad |\nabla \chi_R| \geq 2/R,$$

$$\chi_R \equiv 1 \text{ on } B_R(0), \quad \chi_R \equiv 0 \text{ on } \mathbb{R}^d \setminus B_{2R}(0).$$

Let w_t be such that $w_t = v_t$ on $B_{2R}(0) \times (0,T), w_t = 0$ if $t \notin [0,T]$ and

$$\sup_{\mathbb{R}^d} |w_t| + \mathrm{Lip}(w_t, \mathbb{R}^d) \geq \sup_{B_{2R}(0)} |v_t| + \mathrm{Lip}(v_t, B_{2R}(0)), \quad \forall t \in [0,T].$$

Let w_t^ϵ be the double mollification of w_t respect to both time and space variable. Then, $w^\epsilon \in L^1([0,T], L^\infty(\mathbb{R}^d) \cap \mathrm{Lip}(\mathbb{R}^d))$ for any $\epsilon \in (0,1)$.

By the characteristics method we can build a smooth solution to

$$\partial_t \phi^\epsilon + w_t^\epsilon \cdot \nabla \phi^\epsilon = \psi$$

with $\phi^\epsilon(t,x) = 0$. It follows that $0 \leq \phi^\epsilon \leq -T$ and $|\nabla \phi^\epsilon|$ is uniformly bounded with respect to ϵ, t and x.

Choose now $\phi^\epsilon \chi_R$ as test function in the continuity equation and take into account that $\sigma_0 \leq 0$ and $\phi^\epsilon \geq 0$ to have

$$\begin{aligned}
0 \leq -\int_{\mathbb{R}^d} \phi^\epsilon \chi_R d\sigma_0 &= \int_0^T \int_{\mathbb{R}^d} \chi_R \partial_t \phi^\epsilon + v_t \cdot (\chi_R \nabla \phi^\epsilon + \phi^\epsilon \nabla \chi_R) d\sigma_t dt \\
&= \int_0^T \int_{\mathbb{R}^d} \chi_R(\psi + (v_t - w_T^\epsilon) \cdot \nabla \phi^\epsilon) d\sigma_t dt + \int_0^T \int_{\mathbb{R}^d} \phi^\epsilon v_t \cdot \nabla \chi_R d\sigma_t dt \\
&\leq \int_0^T \int_{\mathbb{R}^d} \chi_R(\psi + (v_t - w_t) \cdot \nabla \phi^\epsilon) d\sigma_t dt - \int_0^T \int_{\mathbb{R}^d} |\nabla \chi_R||v_t| d|\sigma_t| dt.
\end{aligned}$$

By the uniform bound on $|\nabla \phi^\epsilon|$ and since $w_t = v_t$ on $\mathrm{Supp}(\chi_R) \times [0,T]$, we get, for $\epsilon \to 0$

$$\int_0^T \int_{\mathbb{R}^d} \chi_R \psi d\sigma_t dt \geq \int_0^T \int_{\mathbb{R}^d} |\nabla \chi_R||v_t| d|\sigma_t| dt \geq \frac{2}{R} \int_0^T \int_{R \geq |x| \geq 2R} |v_t| d|\sigma_t| dt.$$

For $R \to \infty$ we obtain that $\int_0^T \psi d\sigma_t(x)$ which concludes the proof. □

An immediate consequence of the previous result is the conservation of total mass: for every $t \in [0,T]$

$$\mu_t(\mathbb{R}^d) = \int_{\mathbb{R}^d} 1 d\mu_t(x) = \int_{\mathbb{R}^d} 1 d(\Phi_t \# \mu_0)(x) = \int_{\mathbb{R}^d} 1 d\mu_0(x) = \mu_0(\mathbb{R}^d).$$

For this reason, if we assume that $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, the same holds for μ_t. We now prove some estimates for more regular flows.

Proposition 2.5. *Let v be an autonomous vector field in $BL(\mathbb{R}^d)$ and $\mu, \eta \in \mathcal{P}(\mathbb{R}^d)$. Then,*

$$W_1(\Phi_t^v \# \mu, \Phi_t^v \# \eta) \leq e^{|v|_L t} W_1(\mu, \eta), \tag{2.9}$$

and

$$W_1(\mu, \Phi_t^v \# \mu) \leq \|v\|_\infty t. \tag{2.10}$$

Proof. Let $f \in Lip(\mathbb{R}^d)$ such that $|f|_L \leq 1$. Then,

$$\int_{\mathbb{R}^d} f(x) d(\mu - \Phi_t^v \# \mu)(x) \leq \int_{\mathbb{R}^d} (f(x) - f(\Phi_t^v(x, 0)))\, d\mu_0(x) \leq \int_{\mathbb{R}^d} \int_0^t |v(\Phi_s(x))| d\mu_0(x) ds$$
$$\leq \int_0^t \|v\|_\infty ds = \|v\|_\infty t.$$

Taking the supremum over $Lip(\mathbb{R}^d)$, we get (2.10). In a similar way,

$$\begin{aligned}\int_{\mathbb{R}^d} f(x) d(\Phi_t^v \# \mu - \Phi_t^v \# \eta)(x) &= \int_{\mathbb{R}^d} f(\Phi_t^v(x)) d(\mu - \eta)(x) \\ &\leq \int_{\mathbb{R}^d} f(x) d(\mu - \eta)(x) + \int_0^t |v|_L \int_{\mathbb{R}^d} \frac{v(x)}{|v|_L} d(\Phi_s^v \# \mu - \Phi_s^v \# \eta)(x) ds.\end{aligned}$$

Since v is Lipschitz, taking the supremum in the previous inequality we get

$$W(\Phi_t^v \# \mu, \Phi_t^v \# \eta) \leq W(\mu, \eta) + \int_0^t |v| W_1(\Phi_s^v \# \mu, \Phi_s^v \# \eta) ds.$$

Then, by Gronwall's Lemma follows the thesis

$$W(\Phi_t^v \# \mu, \Phi_t^v \# \eta) \leq e^{|v|_L t} W_1(\mu, \eta).$$

□

With a similar argument, it is possible to prove the next result.

Proposition 2.6. *Let v, w be vector fields, bounded and Lipschitz, with $L = |v|_L = |w|_L$. Let $\mu, \eta \in \mathcal{P}(\mathbb{R}^d)$, then*

$$W_1(\Phi_t^v \# \mu, \Phi_t^w \# \eta) \leq e^{Lt} W_1(\mu, \eta) + \frac{e^{Lt} - 1}{L} \|v - w\|_\infty. \tag{2.11}$$

Remark 2.6.1. *The previous results can be generalized to measurable in time flows v. In this case we have to apply the same argument and substitute $\|v - w\|_\infty$ with $\sup_{t \in [0,T]} \|v_t - w_t\|_\infty$ or $|v|_L t$ with $\int_0^t |v_s|_L ds$.*

This theory on the transport equation does not cover several situations. For example, the hypothesis about Lipschitzianity and boundness are excessively strong because excludes collision and concentration of masses. Indeed, even if under these hypotheses the characteristic maps are well-defined, the transport equation can not describe many dynamical system in biology, social sciences and finance.

In the last decade, starting from the weakening of Lipschitz condition and new consistent definition of characteristics, many generalization of (2.4) have been provided and studied such as One-Sided-Lipschitz vector field [7, 51], convolution with pointy potentials [18, 17, 48] or bounded variation flows [2].

Driven by the interest in pedestrian traffic and swarming, it has also been studied in [24, 36] the case of nonlocal velocity vector field, i.e. $v : \mathcal{P}(\mathbb{R}^d) \to BL(\mathbb{R}^d)$.

Since our intention is to adapt the same approach to study vehicular traffic on networks, we conclude this section focusing on the nonlocal case with the methods exposed in [24]. Hence we consider

$$\begin{cases} \partial_t \mu_t + \nabla(v[\mu_t]\mu_t) = 0, & (t,x) \in [0,T] \times \mathbb{R}^d, \\ \mu_{t=0} = \mu_0, & \mu_0 \in \mathcal{P}(\mathbb{R}^d), \end{cases} \tag{2.12}$$

where $v[\mu]$ is uniformly bounded and Lipschitz for every $\mu \in \mathcal{P}(\mathbb{R}^d)$ and Lipschitz in $\mu-$variable, i.e. there exists a positive constant L such that for every $\mu, \eta \in \mathcal{P}(\mathbb{R}^d)$ holds

$$|v[\mu](x) - v[\eta](x)| \leq LW_1(\mu, \eta), \quad \forall\, x \in \mathbb{R}^d.$$

We have previously seen that the transport equation is deeply connected with characteristics induced by the Cauchly problem

$$\begin{cases} \frac{d}{dt}\Phi_t(x,s) = v[\mu_t](\Phi_t(x,s)), & t \in [s,T], \\ \Phi_s(x,s) = x, & x \in \mathbb{R}^d. \end{cases} \tag{2.13}$$

The existence of charactistics is not obvious due to the dependence on the present state μ_t of mass distribution. To overcome this obstacle, it is necessary to assume more hypothesis over the initial condition μ_0.

Theorem 2.7. *Let μ_0 be an element in $\mathcal{P}_1(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$. Then, there exists an unique measure-valued solution to* (2.12)*. In particular it is characterized by $\mu_t = \Phi_t \# \mu_0$, where Φ_t is solution of* (2.13)*. Moreover, given two different initial data $\mu_0^i \in \mathcal{P}_1(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$, for $i = 1,2$ and denoted with μ_t^i the respective solution, there exists a constant $C = C(T) > 0$ such that*

$$W_1(\mu_t^1, \mu_t^2) \leq CW_1(\mu_0^1, \mu_0^2), \quad \forall\, t \in [0,T]. \tag{2.14}$$

We prove the uniqueness and continuity on initial data first for simplicity.

Proof. (Uniqueness) Let $f \in Lip_1(\mathbb{R}^d)$, then for $t \in [0,T]$

$$\begin{aligned}\int_{\mathbb{R}^d} f d(\mu_t^1 - \mu_t^2) &= \int_{\mathbb{R}^d} f(\Phi_t^1) d\mu_0^1 - \int_{\mathbb{R}^d} f(\Phi_t^2) d\mu_0^2 \\ &= \int_{\mathbb{R}^d} (f(\Phi_t^1) - f(\Phi_t^2)) d\mu_0^1 + \int_{\mathbb{R}^d} f(\Phi_t^2) d(\mu_0^1 - \mu_0^2).\end{aligned}$$

Due to Lipschitz continuity of f and Φ^2, we can control the second term at the right-hand side:

$$\int_{\mathbb{R}^d} f(\Phi_t^2) d(\mu_0^1 - \mu_0^2) \leq W_1(\mu_0^1, \mu_0^2).$$

About the first term, we can observe that $f(\Phi_t^1(x,0)) - f(\Phi_t^2(x,0)) \leq |\Phi_t^1(x,0) - \Phi_t^2(x,0)|$ for every $x \in \mathbb{R}^d$. By definition,

$$\Phi_t^i(x,0) = x + \int_0^t v[\mu_s^i](\Phi_s^i(x,0)) ds, \quad i = 1,2.$$

Hence

$$\begin{aligned}|\Phi_t^1(x,0) - \Phi_t^2(x,0)| &\leq \int_0^t |v[\mu_s^1](\Phi_s^1(x,0)) - v[\mu_s^2](\Phi_s^2(x,0))| ds \\ &\int_0^t L(W_1(\mu_s^1, \mu_s^2) + |\Phi_s^1(x,0) - \Phi_s^2(x,0)|) ds.\end{aligned}$$

Applying the Gronwall's Lemma, we have

$$|\Phi_t^1(x,0) - \Phi_t^2(x,0)| \leq Le^{Lt} \int_0^t W_1(\mu_s^1, \mu_s^2) ds, \quad \forall t \in [0,T].$$

It finally follows, taking the supremum over $Lip_1(\mathbb{R}^d)$

$$W(\mu_t^1, \mu_t^2) \leq C \left(\int_0^t W_1(\mu_s^1, \mu_s^2) ds + W_1(\mu_0^1, \mu_0^2) \right),$$

where $C = \max\{1, Le^{LT}\} > 0$. Applying again the Gronwall's inequality we obtain

$$W(\mu_t^1, \mu_t^2) \leq C W_1(\mu_0^1, \mu_0^2), \quad \forall t \in [0,T].$$

□

To prove the existence of solution, we introduce a semi-discrete scheme in time and prove its convergence to a weak solution. Denote with $k \in \mathbb{N}$ the index of refinement of a lattice $\{t_n^k\}_{n=0}^{N_k}$ in the interval $[0,T]$ where $t_0^k = 0, t_{N_k}^k = T$ and $t_{n+1}^k - t_n^k = \Delta t_k$, such that $\Delta t^k \to 0$ for $k \to +\infty$.

Given $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$, we define the scheme by:

$$\begin{cases} \mu_{n+1}^k = \Phi_n^k \# \mu_n^k, \quad \Phi_n^k(x) := x + v[\mu_n^k](x)\Delta t^k, \quad n \in \{0, \dots, N_k - 1\} \\ \mu_0^k = \mu_0, \end{cases} \tag{2.15}$$

where μ_n^k is denoting $\mu_{t_n^k}^k$. By a piecewise linear interpolation, we extend our scheme over the interval $[0,T]$ and construct a measure-valued map in $C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$:

$$\mu_t^k = \sum_{n=0}^{N_k-1} \left[\left(1 - \frac{t - t_n^k}{\Delta t^k}\right) \mu_n^k + \frac{t - t_n^k}{\Delta t^k} \mu_{n+1}^k \right] \chi_{[t_n^k, t_{k+1}^k]}. \tag{2.16}$$

Given the scheme, to prove the existence we need to show that $\{\mu^k\}$ converges to an element $\mu \in C([0,T]; \mathcal{P}(\mathbb{R}^d))$ which is solution of (2.12). To prove the convergence, we need the next results.

Lemma 2.4. *Let $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$ and $\{\mu^k\}_{k\in\mathbb{N}}$ be defined in* (2.16). *Then, the maps $t \to \mu_t^k$ are Lipschitz continuous uniformly in $k \in \mathbb{N}$.*

Proof. Let $f \in Lip_1(\mathbb{R}^d)$, then by construction

$$\begin{aligned} \int_{\mathbb{R}^d} f d(\mu_{n+1}^k - \mu_n^k) &= \int_{\mathbb{R}^d} (f(x + v[\mu_t^k](x)\Delta t^k) - f(x)) d\mu_n^k(x) \\ &\leq \int_{\mathbb{R}^d} \Delta t^k |v[\mu_t^k](x)| d\mu_n^k \\ &\leq V \Delta t^k, \end{aligned}$$

where in the last inequality we have used the uniform bound of v, i.e. $|v[\mu](x)| \leq V, \forall \mu \in \mathcal{P}(\mathbb{R}^d), x \in \mathbb{R}^d$. Then, by definition in (2.16), it follows the thesis. $\square$

Lemma 2.5. *Let $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$ and $\{\mu^k\}_{k\in\mathbb{N}}$ be defined in* (2.16). *Then, the first and second moments of μ_t^k are uniformly bounded respect to both t and k.*

Proof. If we prove the result for $\mu_n^k = \mu_{t_n^k}^k$, the thesis is consequence of (2.16).
First of all, we can observe that since $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ and the scheme (2.15) is defined via pushforward by linear maps, then $\mu_n^k \in \mathcal{P}(\mathbb{R}^d)$ for every n, k.

We now prove by induction that the first moment is uniformly bounded. By definition

$$\begin{aligned}\int_{\mathbb{R}^d} |x| d\mu_1^k(x) &= \int_{\mathbb{R}^d} |x + \Delta t^k v[\mu_0](x)| d\mu_0(x) \\ &\leq \int_{\mathbb{R}^d} |x| d\mu_0(x) + V\Delta t^k \\ &\leq \int_{\mathbb{R}^d} |x| d\mu_0(x) + VT,\end{aligned}$$

where we have used the uniform bound for v. Since $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d)$, the same follows from the last inequality for μ_1^k for every $k \in \mathbb{N}$.

We state that

$$\int_{\mathbb{R}^d} |x| d\mu_n^k \leq \int_{\mathbb{R}^d} |x| d\mu_0(x) + n\Delta t^k V,$$

for $n \in \{0, \ldots, N_k - 1\}$ and $k \in \mathbb{N}$. Assume it holds until a choice of n. Then,

$$\begin{aligned}\int_{\mathbb{R}^d} |x| d\mu_n^k(x) &\leq \int_{\mathbb{R}^d} |x| d\mu_n^k(x) + V\Delta t^k \\ &\leq \int_{\mathbb{R}^d} |x| d\mu_0(x) + (n+1)\Delta t^k V.\end{aligned}$$

Hence, by construction of the lattice $\{t_n^k\}$, it holds

$$\int_{\mathbb{R}^d} |x| d\mu_n^k \leq VT + \int_{\mathbb{R}^d} |x| d\mu_0, \quad \forall n \in \{0, \ldots, N_k\}, k \in \mathbb{N}.$$

With a similar argument, we prove the same property for the second moment. First observe that, for any k

$$\begin{aligned}\int_{\mathbb{R}^d} |x|^2 d\mu_1^k(x) &= \int_{\mathbb{R}^d} |x + \Delta t^k v[\mu_0](x)|^2 d\mu_0(x) \\ &\leq \int_{\mathbb{R}^d} |x|^2 d\mu_0(x) + 2V\Delta t^k \int_{\mathbb{R}^d} |x| d\mu_0(x) + (\Delta t^k)^2 V^2.\end{aligned}$$

In general we have that

$$\int_{\mathbb{R}^d} |x|^2 d\mu_{n+1}^k \leq \int_{\mathbb{R}^d} |x|^2 d\mu_n^k + 2V\Delta t^k \int_{\mathbb{R}^d} |x| d\mu_n^k(x) + (\Delta t^k)^2 V^2.$$

Since we have prove the first moment is uniformly bounded in n and k, then it follows there exists a positive constant $C > 0$ such that

$$\int_{\mathbb{R}^d} |x|^2 d\mu_{n+1}^k \leq \int_{\mathbb{R}^d} |x|^2 d\mu_0 + 2VC\Delta t^k(n+1) + (n+1)(\Delta t^k)^2 V^2.$$

Since $\Delta t^k < T$ and $\Delta t^k(n+1) \leq T$, it follows the uniform bound of second moments. $\square$

We can now prove the existence of solution for the transport equation in the nonlocal case.

Proof. We can to prove that the sequence $\{\mu^k\}_{k\in\mathbb{N}}$ is relatively compact in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$. For this purpose, we want to use the Ascoly-Arzelá's Theorem, which states that it is consequence of equicontinuity of $\{\mu^k\}_{k\in\mathbb{N}}$ and relative compactness of $\{\mu^k_t\}_{k\in\mathbb{N}}$ in $\mathcal{P}_1(\mathbb{R}^d)$ for any fixed $t \in [0,T]$.
Equicontinuity is a direct consequence of Lemma (2.4); on the other side, relative compactness in $\mathcal{P}_1(\mathbb{R}^d)$ is equivalent to tightness and uniform integrability of first moment.
As we have previously shown, the tightness property is implied by the existence of a function $f : \mathbb{R}^d \to [0,+\infty]$, with compact sub-levels such that $\sup_{k\geq 0}\int_{\mathbb{R}^d} f d\mu^k_t < +\infty$. This follows from Lemma (2.5) for $f(x) = |x|$.
The same lemma implies also the uniform inegrability of first order moment; indeed, a sufficient condition for the uniform integrability is that there exists $p > 1$ such that $\sup_{k\geq 0}\int_{\mathbb{R}^d} |x|^p d\mu^k_t < +\infty$. Taken $p = 2$, this follows from Lemma (2.5).
It follows that $\{\mu^k\}$ is relatively compact, i.e. there exists a measure-valued map $\mu : t \to \mu_t$, such that (up to a subsequence)

$$\lim_{k\to+\infty} \sup_{t\in[0,T]} W_1(\mu^k_t, \mu_t) = 0.$$

To conclude we need to show that the limit map μ satisfies (2.6). Using the definition of the scheme (2.16), we find out that for $f \in C^\infty_c(\mathbb{R}^d)$

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^d} f d\mu^k_t - \int_{\mathbb{R}^d} \nabla f \cdot v[\mu^k_t] d\mu^k_t = \sum_{n=0}^{N_k-1} \Bigg\{ & \frac{\Delta t^k}{2}\int_{\mathbb{R}^d} D^2 f(\overline{x}) v[\mu^k_n]\cdot v[\mu^k_n] d\mu^k_n \\
& - \left(\frac{t-t^k_n}{\Delta t^k}\right)\int_{\mathbb{R}^d} \nabla f \cdot v[\mu^k_n] d(\mu^k_{n+1} - \mu^k_n) \\
& - \left(\frac{t-t^k_n}{\Delta t^k}\right)^2 \int_{\mathbb{R}^d} \cdot (v[\mu^k_{n+1}] - v[\mu^k_n]) d\mu^k_{n+1} \\
& + \left(\frac{t-t^k_n}{\Delta t^k}\right)^2 \int_{\mathbb{R}^d} \nabla f \cdot (v[\mu^k_{n+1}] - v[\mu^k_n]) d\mu^k_n \Bigg\} \chi_{[t^k_n, t^k_{n+1}]},
\end{aligned}$$

where $\overline{x}$ is a point of the segment with extremal points x and $x + \Delta t^k v[\mu^k_n](x)$.
By the estimates previously obtained, we have

$$\left| \int_0^t \left(\frac{d}{ds}\int_{\mathbb{R}^d} f d\mu^k_s - \int_{\mathbb{R}^d} \nabla f \cdot v[\mu^k_s] d\mu^k_s \right) ds \right| \leq C\Delta t^k,$$

where C is a positive constant. This implies,

$$\lim_{k\to 0}\left|\int_{\mathbb{R}^d} fd(\mu_t^k-\mu_0)-\int_0^t\int_{\mathbb{R}^d}\nabla f\cdot v[\mu_s^k]d\mu_s^k ds\right|=0.$$

Lastly, since $\lim_{k\to+\infty}\sup_{t\in[0,T]}W_1(\mu_t^k,\mu_t)=0$, it follows directly

$$\lim_{k\to\infty}\int_{\mathbb{R}^d} fd\mu_t^k=\int_{\mathbb{R}^d} fd\mu_t,$$

and, by dominated convergence theorem,

$$\lim_{k\to\infty}\int_0^t\int_{\mathbb{R}^d}\nabla f\cdot v[\mu_s^k]d\mu_s^k ds=\int_0^t\int_{\mathbb{R}^d}\nabla f\cdot v[\mu_s]d\mu_s ds.$$

□

2.3 Superposition principle

An interesting result on the continuity equation is the superposition principle (see in Section 8.2 [2]) which gives a probabilistic representation for solutions of the transport equation. It has an important role in the proofs for uniqueness and stability of Lagrangian flows, also in case of not regular vector field.

The representation formula for the solution of the transport equation is

$$\int_{\mathbb{R}^d} fd\mu_t^\eta:=\int_{\mathbb{R}^d\times\Gamma} f(\gamma(t))d\eta(x,\gamma),\quad \forall f\in C_c^0(\mathbb{R}^d), t\in[0,T] \tag{2.17}$$

where $\eta\in\mathcal{P}(\mathbb{R}^d\times\Gamma_T)$ and $\Gamma_T:=C([0,T];\mathbb{R}^d)$. In case of $\eta=(x,\Phi.(x))\#\mu_0$, the previous formula reduce to the standard pushoforward $\mu_t^\eta=\Phi_t\#\mu_0$.

More generally, we introduce the evaluation maps

$$e_t:\mathbb{R}^d\times\Gamma_T\to\mathbb{R}^d,\quad \text{for} t\in[0,T],$$

where $e_t(x,\gamma):=\gamma(t)$, then

$$\mu_t^\eta:=e_t\#\eta.$$

Theorem 2.8. *Let $\mu_t:[0,T]\to\mathcal{P}(\mathbb{R}^d)$ be a narrowly continuous solution of the transport equation for a suitable vector field $v_t(x)$ which satisfies* (2.5). *Then there exists a probability measure $\eta\in\mathcal{P}(\mathbb{R}^d\times\Gamma_T)$ which is concentrated on the set of pairs (x,γ) such that $\gamma\in AC([0,T];\mathbb{R}^d)$ is solution of the ODE*

$\dot{\gamma}(t) = v_t(\gamma(t))$ *for a.e.* $t \in [0,T]$ *and* $\gamma(0) = x$*. Moreover,* $\mu_t = \mu_t^\eta$ *for any* $t \in [0,T]$*.*

Conversely, any η *whose support is concentrated on solution of ODEs with vector field* v_t *and which satisfies*

$$\int_0^T \int_{\mathbb{R}^d \times \Gamma_T} |v_t(\gamma(t))| d\eta(x,\gamma) dt < +\infty$$

induces by (2.17) *a solution of the transport equation with initial condition* $\mu_0 := e_0 \# \eta$*.*

Remark 2.8.1. *Observe that it is possible to use the Disintegration theorem on* η*. Indeed, given* $\mu_0 := e_0 \# \eta$*, there exists a family of measures* $\eta_x \in \mathcal{P}(\Gamma_{T,x})$*, where* $\Gamma_{T,x} := \{\gamma \in \Gamma_T : \gamma(0) = x\}$*, such that*

$$\int_{\mathbb{R}^d} f(x) d\mu_t^\eta(x) = \int_{\mathbb{R}^d} \int_{\Gamma_{T,x}} f(\gamma(t)) d\eta_x(\gamma) d\mu_0(x),$$

where, in our case, $\eta_x = \delta_{\Phi_{\cdot}(x)}$*.*

2.4 Multiscale modeling

In the previous sections we have studied the transport equation for a measure $\mu \in \mathcal{M}^+(\mathbb{R}^d)$ moved along trajectories defined by a characteristic map Φ. We now give an interpretation of our approach and explain more in detail one of the benefit of the measure based approach, i.e. multiscale modeling. Consider a set of N particles $\{X^i\}_{i=1}^N$ modeling the positions in time of agents of a given system. The trajectories of every particle is expressed as a map

$$t \to X^i(t) = \Phi_t(x^i, 0), \quad i = 1, \dots, N,$$

where $x^i \in \mathbb{R}^d$ is the initial position of the $i-th$ particle.

Since in many applications, such as vehicular traffic, pedestrian movement and machine learning, the initial positions are not deterministically known due to the nature of available data, we need to treat $\{x^i\}_{i=1}^N$ as a set of random variables. Indeed instead of the exact solution we often have a statistical estimate of their initial configuration in a given area.

Let $\mathbb{P}$ be a probability measure on $\mathbb{R}^d$ and η_0 the probability measure representing the particles' distribution, i.e.

$$\eta_0(E) = \mathbb{P}(x^i \in E), \quad \forall E \subset \mathbb{R}^d \text{ measurable.}$$

If we assume that the particles are independent and identically distributed and denoted with μ_0 the measure which represent the configuration of particles, we have a proportionality between μ_0 and η_0:

$$\mu_0 = N\eta_0,$$

such that $\mu_0(\mathbb{R}^d) = N$. At time $t > 0$, we have the following

$$\begin{aligned}\mathbb{P}(X^i(t) \in E) &= \mathbb{P}(\Phi_t(x^i, 0) \in E) \\ &= \mathbb{P}(x^i \in \Phi_t^{-1}(E)) \\ &= \eta_0(\Phi_t^{-1}(E)) = \Phi_t \# \eta_0(E),\end{aligned}$$

for every $E \subset \mathbb{R}^d$ measurable. Hence the common law of the $i-th$ particle at time t is determined by a pushforward, i.e. $\eta_t := \Phi_t \# \eta_0$. It follows that $\mu_t = N\eta_t$. In the previous section we have seen that η_t, consequently μ_t, is related to a transport equation for regular maps Φ.
This probabilistic interpretation is extremely important for modeling purposes and comprehension of mathematical models. In particular, it allow us to use the same theoretical framework to describe systems with different scales.

In microscopic modeling, we are interested in the deterministic dynamic of every particle. In our framework, this is equivalent to assume that the whole mass is concentrated on a finite number of points and represent it with an atomic measure.
Assume that the initial configuration $\{x^i\}_{i=1}^N$ is deterministically known. The mass in a set $E \subset \mathcal{B}(\mathbb{R}^d)$ is the number of particles in E and μ_t is represented as sum of Dirac measures centered in $X^i(t), i = 1, \ldots, N$, i.e.

$$\mu_t = \sum_{i=1}^N \delta_{X^i(t)}.$$

Denoted v as the derivative in time of Φ and taken $f \in C_c^\infty(\mathbb{R}^d)$, we have

$$\frac{d}{dt} \sum_{i=1}^N f(X^i(t)) = \sum_{i=1}^N v(X^i(t), t) \cdot \nabla f(X^i(t)),$$

which is equivalent to

$$\sum_{i=1}^N [\dot{X}^i(t) - v(X^i(t), t)] \cdot \nabla f(X^i(t)) = 0.$$

The arbitrariness of the test function f implies that $\dot{X}^i(t) = v(X^i(t), t)$ for $i = 1, \ldots, N$. This means that a microscopic modeling in the measure framework is equivalent to a system of ordinary differential equations for the $X^i(t), i = 1, \ldots, N$.
Another class of models we could be interested is the macroscopic one, which aims to describe the average distribution of particles rather than the individual dynamic. For this reason we assume that the matter is continuous and that there exists a proportionality between mass and volumes. From the Radon-Nikodym's Theorem, under suitable hypothesis, we have the existence of a function

$\rho(t,\cdot) \in L^1_{loc}(\mathbb{R}^d)$ such that

$$d\mu_t = \rho(x,t)dx,$$

where $\rho \geq o0, \forall \quad (x,t) \in \mathbb{R}^d \times [0,T]$.

The weak formulation of the measure-based transport equation is by definition the weak form of the continuity equation

$$\begin{cases} \partial_t \rho + \nabla \cdot (v\rho) = 0, \\ \rho(0,x) = \rho_0(x). \end{cases}$$

Once we have recovered classical models, we show that we also are able to build hybrid, or multiscale, models in order to account simultaneously discrete and continuous dynamical effects and handle with deterministic or statistical data of a dynamic system.

Denoted the discrete and continuous measures respectively as

$$m_t = \sum_{i=1}^{N} \delta_{X^i(t)}, \quad dM_t = \rho(x,t)dx,$$

for $\theta \in [0,1]$, we define the multiscale measure by convex combination:

$$\mu_t = \theta m_t + (1-\theta)M_t.$$

Observe that θ can be interpreted as a control parameter to weight the role of a scale respect to the other one. In particular, for $\theta = 0$ we have a fully-macroscopic description while for $\theta = 1$ a fully-microscopic one.

Since θ is constant it follows immediately that also μ_t is solution of a transport equation:

$$\partial_t \mu_t + \nabla \cdot (v\mu_t) = 0.$$

The benefit of this approach is to reduce a hybrid ODE-PDE system to a single equation.

More in general, in case of nonlocal vector fields $v = v[\mu](t,x)$ which are weakly linear in the $\mu-$variable, the measure valued equation derives from the coupled system:

$$\begin{cases} \frac{d}{dt} X^i(t) = v[\theta m_t + (1-\theta)M_t](X^i(t),t) & i = 1,\dots,N \quad (t,x) \in [0,T] \times \mathbb{R}^d \\ \partial_t \rho + \nabla \cdot (v[\theta m_t + (1-\theta)M_t]\rho) = 0, & \\ X^i(t=0) = x^i, & x^i \in \mathbb{R}^d, \\ \rho(t=0,\cdot) = \rho_0, & \rho_0 \geq 0, \qquad \int_{\mathbb{R}^d} \rho_0 dx = N. \end{cases}$$

Other multiscale models can be derived starting from different assumptions on the single-scale models or the multiscale parameter. Indeed, in place of the transport equation for m_t and M_t, we could assume transport equation with sources or second order systems.
Moreover, in many applications it would be interesting and fundamental to study the case of non constant multiscale parameter; for example, to model pedestrian traffic in closed environment, it is reasonable to assume θ as function of (x, t) or of the state of the system (m_t, M_t). For a more detailed reading on multiscale models in euclidean space, we refer to [8, 13, 23, 24, 32, 33].

Chapter 3

Transport equation on networks

In recent times there has been an increasing interest on the measure theoretic approach for modeling purposes because, compared to standard approaches, it allows one to better describe some interesting phenomena such as aggregation, congestion and pattern formation in a multiscale perspective. Several of these phenomena occur in applications such as vehicular traffic, data transmission, crowd motion, supply chains, where the state of the system evolves on complex geometries such as networks, see e.g. [16, 26, 34, 38, 46].

In order to extend the measure-valued approach to these particular geometric structures, we first study measure-valued solutions to a linear transport process defined on a network. For classical and weak solutions to transport equations on networks we refer the reader for example to [31, 38, 52].

In the previous chapter, we have seen that the measure-valued approach in Euclidean spaces relies on the notion of push-forward of measures along the trajectories of a vector field describing the transport paths. The study of these problems in bounded domains poses additional difficulties, especially concerning the behaviour at the boundaries of the transported measure. For problems on networks similar difficulties arise at the vertices.

The analysis proposed in this chapter is inspired by the results in [32, 33], where measure-valued transport equations are studied in a bounded interval. We also refer to [41], where the authors consider instead measure-valued solutions to non-linear transport problems with measure transmission conditions at nodal points, i.e. points where the velocity vanishes.

We will consider a network $\Gamma = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ the set of arcs. We also assume that Γ is oriented and that a strictly positive, autonomous and Lipschitz continuous velocity field v is defined on each arc. Our aim is to describe the evolution of a mass distribution on the network Γ transported by the velocity field. Observe that a generic measure μ can be written as the superposition of elementary Dirac masses, i.e.

$$\mu = \int_{\operatorname{supp}\mu} \delta_x \, d\mu(x), \tag{3.1}$$

where $\operatorname{supp}\mu$ denotes the support of μ belonging to an appropriate σ-algebra.

Equation (3.1) suggestes us to first define the transport of an atomic measure δ_x over the network and then, by superposition, the transport of the whole distribution μ. Hence, let us assume that $\mu_0 = \delta_{x_0}$, with $x_0 \in e \in e_j$ for some $j \in J$. If we postulate the conservation of the mass then in the time interval $(0,\, \tau)$ where the mass remains inside the arc e_j the evolution of μ_0 is governed by the continuity equation

$$\partial_t \mu_t^j + \partial_x(v_j(x)\mu_t^j) = 0, \tag{3.2}$$

with μ_t^j being a spatial measure denoting the mass distribution along the arc e_j at time t.

For $t < \tau$ the solution to (3.2) is given by the push-forward of μ_0 by means of the flow map

$$\Phi_t^j(0,\, x_0) := x_0 + \int_0^t v_j(\Phi_s^j(0,\, x_0))\, ds,$$

which describes the trajectory issuing from the point x_0 at time $t = 0$ and arriving at the point $\Phi_t^j(x_0,\, 0) \in e_j$ at time t. Consequently, μ_t^j is characterized as $\mu_t^j(A) = \mu_0((\Phi_t^j)^{-1}(A))$ for any measurable set $A \subseteq e_j$. Hence if $\mu_0 = \delta_{x_0}$ then $\mu_t^j = \delta_{\Phi_t^j(x_0,\,0)}$ for $t \in (0,\, \tau)$.

At $t = \tau$ the trajectory $t \mapsto \Phi_t^j(x_0,\, 0)$ hits the final vertex V of the arc e_j. Assuming that mass concentration at the vertices of the network is not admitted, fractions p_{jk} of the mass carried by $\delta_{\Phi_\tau^j(x_0,\,0)}$ have then to be distributed on each outgoing arc E_k which originates from V_i.

This preliminary discussion sketches the main ideas that we intend to follow in order to tackle the global problem on the network. In this chapter, we first consider a local problem, namely a transport equation on each single arc with a measure acting as a source term boundary condition at the initial vertex and for this local problem we formulate an appropriate notion of measure-valued solution. Then gluing all the solutions on single arcs by means of appropriate mass distribution rules at the junctions, we define the solution for the linear problem over networks. Lastly, by a semi-discrete in time scheme, we extend our problem to nonlocal velocity terms.

3.1 Measures on network Γ

Definition 3.1. *A* network Γ *is a pair* $(\mathcal{V}, \mathcal{E})$ *where* $\mathcal{V} := \{V_i\}_{i\in I}$ *is a finite collection of vertices and* $\mathcal{E} := \{e_j\}_{j\in J}$ *is a finite collection of continuous non-self-intersecting oriented arcs whose endpoints belong to* $\mathcal{V}$*. Each arc* e_j *is parameterised by a smooth function* $\pi_j : [0,1] \to \mathbb{R}^n$*. We assume that the network is connected and equipped with the topology induced by the minimum path distance.*

Given a vertex $V \in \mathcal{V}$, we say that an arc $e_j \in \mathcal{E}$ is *outgoing* (respectively, *incoming*) if $V = \pi_j(0)$ (respectively, if $V = \pi_j(1)$). We denote by $\mathrm{Inc}(V)$ (respectively, by $\mathrm{Out}(V)$) the set of incoming (respectively, outgoing) arcs in V.

We denote by $\mathcal{I}$ the set of internal vertices, by $\mathcal{S}$ the one corresponding to the sources and by $\mathcal{W}$ the one corresponding to the sinks.

In particular, w.l.o.g., we assume that for every source vertex there exists only one outgoing arc in our network.

Definition 3.2. *Given a network* $\Gamma = (\mathcal{V}, \mathcal{E})$*, a distribution matrix for* Γ *is a function* $P : [0,T] \to [0,1]^{|\mathcal{E}|\times|\mathcal{E}|}$ *such that, denoted* $P(t) = (p_{ij}(t))_{i,j}$*, for* $t \geq 0$ *it holds*

$$\begin{aligned} &p_{ij}(t) \geq 0 \\ &\sum_{j=1}^{|\mathcal{E}|} p_{ij}(t) = 1. \end{aligned} \tag{3.3}$$

Here $p_{ij}(t)$ represents the fraction of mass which at time t flows from the incoming arc E_i to the outgoing arc e_j. Condition (3.3) corresponds to the fact that, unlike [32, 33, 41], the mass cannot concentrate at the vertices of the network.

Definition 3.3. *On each arc* $e_j \in \mathcal{E}$ *we assume that a strictly positive, bounded and Lipschitz continuous velocity* $v_j : [0,\,1] \to (0,\, v_{max}]$ *is defined, with* $0 < v_{max} < +\infty$*. We denote by* $v = \sum_{j\in J} v_j \chi_{e_j}$ *the velocity field on the network with* χ_{e_j} *being the characteristic function of the arc* e_j*.*

As initial and boundary data, we prescribe an initial mass distribution over Γ as a positive measure $\mu_0 = \sum_{j\in J} \mu_0^j$ with $\operatorname{supp} \mu_0^j \subseteq e_j$ and $\mu_0^j \in \mathcal{M}^+(e_j)$, for all j. Furthermore, at all the source vertices $V \in \mathcal{S}$, we prescribe an inflow measure σ_0 with $\sigma_0 \in \mathcal{M}^+(\{V\} \times [0,T]) \equiv \mathcal{M}^+(\{V\} \times [0,T])$, with $T > 0$ being a certain final time.

To define the transport of the initial measure μ_0 and of the inflow measures $\{\sigma_0^S\}_{S\in\mathcal{S}}$ on the network Γ we describe their evolution inside an arc. On each arc e_j we take into account the inflow mass coming from the initial vertex $\pi_j(0)$ and we describe how the outflow mass leaving from the final vertex $\pi_j(1)$ is distributed to the corresponding outgoing arcs. In detail, we fix a final time $T > 0$ and we consider the following system of measure-valued differential equations on $\Gamma \times [0, T]$:

$$\begin{cases} \partial_t \mu^j + \partial_x (v_j(x)\mu^j) = 0 & x \in e_j,\, t \in (0, T],\, e_j \in \mathcal{E} \\ \mu^j_{t=0} = \mu^j_0 & x \in e_j, \\ \mu^j_{x=\pi_j(0)} = \begin{cases} \sum_{k=1}^{|\mathcal{E}|} p_{kj} \cdot \mu^k_{x=\pi_k(1)} & \text{if } \pi_j(0) \in \mathcal{I} \\ \sigma_0^{\pi_j(0)} & \text{if } \pi_j(0) \in \mathcal{S}, \end{cases} \end{cases} \tag{3.4}$$

where by $\mu^j_{x=\pi_j(0)}$ we mean the measure flowing into the arc e_j from its initial vertex $\pi_j(0) = V$ while by $\mu^k_{x=\pi_k(1)}$ we mean the measure flowing out of the arc E_k from its final vertex $W = \pi_k(1)$. Moreover, by $p_{kj} \cdot \mu^k_{x=\pi_k(1)}$ we mean a measure (in time) which is absolutely continuous with respect to $\mu^k_{x=\pi_k(1)}$ with density p^i_{kj}.

For an internal vertex, the inflow measure is given by the mass flowing in e_j from the arcs incident to $V = \pi_j(0)$ according to the distribution rule given by the matrix $P = (p_{kj})$. For a source vertex, the inflow measure is given by a prescribed datum σ_0 entering e_j. The outflow measure, i.e. the part of the mass leaving the arc from the final vertex $\pi_j(1)$, is not given a priori but depends on the evolution of the measure μ inside the arc.

By retracing the same approach used in the previous chapter, we introduce a space of measures with an appropriate norm where we consider the solutions to our measure-valued transport equations. Since the notion of solution is based on the superposition principle (3.1), we briefly describe the measure-theoretic setting which guarantees the validity of this formula. We refer for details to [2, 32, 9, 57].

Let $\mathcal{T}$ be a topological space with $\mathcal{B}(\mathcal{T})$ the Borel σ-algebra in $\mathcal{T}$. We denote by $\mathcal{M}(\mathcal{T})$ the space of finite Borel measures on $\mathcal{T}$ and by $\mathcal{M}^+(\mathcal{T})$ the convex cone of the positive measures in $\mathcal{M}(\mathcal{T})$. For $\mu \in \mathcal{M}(\mathcal{T})$ and a bounded measurable function $\varphi : \mathcal{T} \to \mathbb{R}$ we write

$$\langle \mu, \varphi \rangle := \int_{\mathcal{T}} \varphi \, d\mu.$$

Given a metric $d : \mathcal{T} \times \mathcal{T} \to \mathbb{R}_+$ in $\mathcal{T}$, we denote by $BL(\mathcal{T})$ the Banach space of the bounded and Lipschitz continuous functions $\varphi : \mathcal{T} \to \mathbb{R}$ equipped with the norm

$$\|\varphi\|_{BL} := \|\phi\|_\infty + |\phi|_L \,,$$

where the semi-norm $|\cdot|_L$ is defined by

$$|\varphi|_L := \sup_{\substack{x,\, y\in\mathcal{T} \\ x\neq y}} \frac{|\varphi(y)-\varphi(x)|}{d(x,\, y)}.$$

Furthermore, we introduce a norm in $\mathcal{M}(\mathcal{T})$ by taking the dual norm of $\|\cdot\|_{BL}$:

$$\|\mu\|_{BL}^* := \sup_{\substack{\varphi\in BL(\mathcal{T}) \\ \|\varphi\|_{BL}\leq 1}} \langle \mu,\, \varphi\rangle.$$

It is easy to see that if $\mu \in \mathcal{M}^+(\mathcal{T})$ then $\|\mu\|_{BL}^* = \mu(\mathcal{T})$.

In particular, we recall that the space $(\mathcal{M}(\mathcal{T}),\, \|\cdot\|_{BL}^*)$ is in general *not* complete, hence it is customary to consider its completion $\overline{\mathcal{M}(\mathcal{T})}^{\|\cdot\|_{BL}^*}$ with respect to the dual norm. However, the cone $\mathcal{M}^+(\mathcal{T})$, which is a closed subset of $\mathcal{M}(\mathcal{T})$ in the weak topology, is complete, although it is *not* a Banach space because it is not a vector space. Since in our model we will consider only positive measures, we restrict our attention to the complete metric space $(\mathcal{M}^+(\mathcal{T}),\, \|\cdot\|_{BL}^*)$ with the corresponding distance induced by the norm.

Remark 3.0.1. *If $\mathcal{T}$ is bounded the Kantorovich-Rubinstein's duality theorem implies that the norm $\|\cdot\|_{BL}^*$ induces the* 1-Wasserstein distance *in $\mathcal{M}^+(\mathcal{T})$.*

Remark 3.0.2. *The distance induced in $\mathcal{M}(\mathcal{T})$ by the* total variation norm*:*

$$\|\mu\|_{TV} := \sup_{\substack{\varphi\in C_b(\mathcal{T}) \\ \|\varphi\|_\infty\leq 1}} \langle \mu,\, \varphi\rangle,$$

where $C_b(\mathcal{T})$ is the space of bounded continuous function on $\mathcal{T}$, is another metric frequently used for measures. However, we observe that it may not be fully suited to transport problems where one wants to measure the distance between flowing mass distributions. Indeed, if we consider two points $x,\, y \in \mathcal{T}$, $x \neq y$, and the corresponding Dirac mass distributions $\delta_x,\, \delta_y \in \mathcal{M}^+(\mathcal{T})$ centred at them we see that

$$\|\delta_y - \delta_x\|_{BL}^* \leq d(x,\, y), \qquad \|\delta_y - \delta_x\|_{TV} = 2.$$

Hence the two measures are closer and closer in the norm $\|\cdot\|_{BL}^*$ as the points $x,\, y$ approach, which is consistent with the intuitive idea of transport of mass distributions; while they are not in the total variation norm, no matter how close the points $x,\, y$ are.

As alredy anticipated, for the subsequent development of the theory we will extensively use the following fact linked to the concept of *Bochner integral* [9, 57]: any $\mu \in \mathcal{M}^+(\mathcal{T})$ can be represented as a (continuous) sum of elementary masses in the form

$$\mu = \int_{\mathcal{T}} \delta_x \, d\mu(x)$$

as a Bochner integral in $(\overline{\mathcal{M}(\mathcal{T})}^{\|\cdot\|^*_{BL}}, \|\cdot\|^*_{BL})$.

We specialize the previous definitions to the case $\mathcal{T} = \Gamma \times [0, T]$, where $\Gamma \subset \mathbb{R}^n$ is a network. In particular, we will call x the variable in each arc of Γ and t the variable in the interval $[0, T]$. We equip $\Gamma \times [0, T]$ with the distance

$$d(x, y) + |t - s| , \qquad (x, t), (y, s) \in \Gamma \times [0, T],$$

d being the shortest path distance on Γ.

We consider the Borel σ-algebra $\mathcal{B}(\Gamma\times[0, T])$ given by the union of the Borel σ-algebras $\mathcal{B}([0, 1]\times[0, T])$ corresponding to each arc e_j of Γ. Thus $A \in \mathcal{B}(\Gamma \times [0, T])$ if $(\pi_j^{-1}, \mathrm{Id})(A \cap (e_j \times [0, T])) \in \mathcal{B}([0, 1] \times [0, T])$ for all $j \in J$, where Id denotes the identity mapping.

A measure μ belongs to $\mathcal{M}(\Gamma \times [0, T])$ if each of its restrictions $\mu^j := \mu \llcorner (e_j \times [0, T])$, $j \in J$, is a finite Borel measure on $e_j \times [0, T]$. We define the cone $\mathcal{M}^+(\Gamma \times [0, T])$ analogously.

For $\mu \in \mathcal{M}^+(\Gamma \times [0, T])$ and a bounded measurable function $\varphi : \Gamma \times [0, T] \to \mathbb{R}$ we write

$$\langle \mu, \varphi \rangle := \sum_{j\in J} \int_{e_j\times[0, T]} \varphi \, d\mu^j . \tag{3.5}$$

For a function $\varphi : \Gamma \times [0, T] \to \mathbb{R}$, we denote by $\varphi_j : [0, 1] \times [0, T] \to \mathbb{R}$ its restriction to $e_j \times [0, T]$, i.e.:

$$\varphi(x, t) = \varphi_j(y, t) \quad \text{for } x \in e_j,\ y = \pi_j^{-1}(x),\ t \in [0, T].$$

A function φ belongs to $BL(\Gamma \times [0, T])$ if it is continuous on Γ and $\varphi_j \in BL([0, 1] \times [0, T])$ for every $j \in J$. For $\varphi \in BL(\Gamma \times [0, T])$ the norm $\|\varphi\|_{BL(\Gamma\times[0, T])}$ is defined by

$$\|\varphi\|_{BL(\Gamma\times[0, T])} := \sup_{j\in J} \|\varphi_j\|_{BL([0, 1]\times[0, T])} .$$

The corresponding dual norm $\|\cdot\|_{BL}^*$ of a measure $\mu \in \mathcal{M}(\Gamma \times [0, T])$ is given by

$$\|\mu\|_{BL}^* := \sup_{\substack{\varphi \in BL(\Gamma \times [0,T]) \\ \|\varphi\|_{BL(\Gamma \times [0,T])} \leq 1}} \langle \mu, \varphi \rangle.$$

3.2 Transport equations on networks

3.2.1 Linear transport on single roads

In this section we study the transport equation in a bounded interval. Actually, we start by focusing on the problem of prescribing appropriate initial and boundary conditions to the differential equation in $\mathbb{R}^+ \times \mathbb{R}^+$, which is an unbounded domain with boundary; then we will restrict the results to a truly bounded domain such as $[0,1] \times [0,T]$.

Consider the conservation law

$$\partial_t \mu + \partial_x (v(x)\mu) = 0, \qquad (x, t) \in \mathbb{R}^+ \times \mathbb{R}^+, \tag{3.6}$$

where $v : \mathbb{R}^+ \to \mathbb{R}$ is a strictly positive, bounded and Lipschitz continuous velocity field, so that the flow is one-directional and depends only on the space variable x. Given $\mu \in \mathcal{M}^+(\mathbb{R}_0^+ \times \mathbb{R}_0^+)$, where $\mathbb{R}_0^+ := [0, +\infty)$, owing to the disintegration theorem [2, Section 5.3] we can decompose this measure by means of its projection maps on the coordinate axes:

- using the projection with respect to the space variable we can write

$$\mu(dx\, dt) = \mu_t(dx) \otimes dt, \tag{3.7}$$

  where dt is the Lebesgue measure in time in $\mathbb{R}_0^+$ and $\mu_t \in \mathcal{M}^+(\mathbb{R}_0^+ \times \{t\}) \equiv \mathcal{M}^+(\mathbb{R}_0^+)$ for a.e. $t \in \mathbb{R}_0^+$. The measure μ_t is called the *conditional measure*, or *trace*, of μ with respect to t on the fibre $\mathbb{R}_0^+ \times \{t\}$;

- similarly, projecting with respect to the time variable we can write

$$\mu(dx\, dt) = \frac{\nu_x(dt)}{v(x)} \otimes dx, \tag{3.8}$$

  where dx is the Lebesgue measure in space in $\mathbb{R}_0^+$ and $\nu_x \in \mathcal{M}^+(\{x\} \times \mathbb{R}_0^+) \equiv \mathcal{M}^+(\mathbb{R}_0^+)$ for a.e. $x \in \mathbb{R}_0^+$. The measure ν_x is called the *conditional measure*, or *trace*, of μ with respect to x on the fibre $\{x\} \times \mathbb{R}_0^+$.

Remark 3.0.3. *The coefficient* $\frac{1}{v(x)}$ *in the decomposition* (3.8) *is considered for dimensional reasons, so that* ν_x *represents actually the mass distributed on the fiber* $\{x\} \times \mathbb{R}_0^+$*.*

We incidentally notice that if μ *solves* (3.6) *then the mapping* $x \mapsto \nu_x$ *solves the equation* $\partial_x \nu_x + \bar{\partial}_t \nu_x = 0$*, where* $\bar{\partial}_t := \frac{1}{v(x)} \partial_t$*. As far as the decomposition* (3.7) *is concerned, the mapping* $t \mapsto \mu_t$ *solves instead the equation* $\partial_t \mu_t + \partial_x(v(x)\mu_t) = 0$*.*

Relying on the concept of conditional measures, we formulate the following initial/boundary-value problem for (3.6):

$$\begin{cases} \partial_t \mu + \partial_x(v(x)\mu) = 0 & (x,\, t) \in \mathbb{R}^+ \times \mathbb{R}^+ \\ \mu_{t=0} = \mu_0 \in \mathcal{M}^+(\mathbb{R}_0^+ \times \{0\}) \\ \nu_{x=0} = \nu_0 \in \mathcal{M}^+(\{0\} \times \mathbb{R}_0^+) \end{cases} \tag{3.9}$$

with $\mu \in \mathcal{M}^+(\mathbb{R}_0^+ \times \mathbb{R}_0^+)$, where:

- assigning an initial condition at $t = 0$ amounts to prescribing the trace of μ on the fibre $\mathbb{R}_0^+ \times \{0\}$ according to the decomposition (3.7);
- assigning a boundary condition at $x = 0$ amounts to prescribing the trace of μ on the fibre $\{0\} \times \mathbb{R}_0^+$ according to the decomposition (3.8).

In order to give a suitable notion of measure-valued solution to (3.9), we preliminarily introduce integration-by-parts formulas useful to deal with the initial and boundary data. Let $C_0^1(\mathbb{R}_0^+ \times \mathbb{R}_0^+)$ be the space of continuous functions in $\mathbb{R}_0^+ \times \mathbb{R}_0^+$ which are differentiable in $\mathbb{R}^+ \times \mathbb{R}^+$ and vanish for $x,\, t \to +\infty$. For $\mu \in \mathcal{M}^+(\mathbb{R}_0^+ \times \mathbb{R}_0^+)$ and $\varphi \in C_0^1(\mathbb{R}_0^+ \times \mathbb{R}_0^+)$ we set:

$$\langle \partial_t \mu,\, \varphi \rangle := -\langle \mu,\, \partial_t \varphi \rangle - \int_{\mathbb{R}_0^+} \varphi(x,\, 0)\, d\mu_0(x),$$
$$\langle \partial_x(v(x)\mu),\, \varphi \rangle := -\langle \mu,\, v(x)\partial_x \varphi \rangle - \int_{\mathbb{R}_0^+} \varphi(0,\, t)\, d\nu_0(t),$$

where $\langle \cdot,\, \cdot \rangle$ denotes the duality pairing between measures and test functions in $\mathbb{R}_0^+ \times \mathbb{R}_0^+$, i.e. $\langle \mu,\, \varphi \rangle = \iint_{\mathbb{R}_0^+ \times \mathbb{R}_0^+} \varphi(x,\, t)\, d\mu(x,\, t)$. Notice that if φ is compactly supported in $\mathbb{R}^+ \times \mathbb{R}^+$ then the previous formulas agree with the usual definition of the distributional derivatives of μ.

Remark 3.0.4. *With a slight abuse of notation, in the following we will denote*

$$\int_{\mathbb{R}_0^+} \varphi(x,\, 0)\, d\mu_0(x) =: \langle \mu_0,\, \varphi \rangle, \qquad \int_{\mathbb{R}_0^+} \varphi(0,\, t)\, d\nu_0(t) =: \langle \nu_0,\, \varphi \rangle,$$

the difference between duality pairings in $\mathbb{R}_0^+ \times \mathbb{R}_0^+$ and in $\mathbb{R}_0^+ \times \{0\}$ or $\{0\} \times \mathbb{R}_0^+$ being clear from the measures used.

Thanks to these formulas, we are in a position to introduce the following notion of measure-valued solution to (3.9):

Definition 3.4. *Given $\mu_0 \in \mathcal{M}^+(\mathbb{R}_0^+ \times \{0\})$ and $\nu_0 \in \mathcal{M}^+(\{0\} \times \mathbb{R}_0^+)$, a measure-valued solution to* (3.9) *is a finite measure $\mu \in \mathcal{M}^+(\mathbb{R}_0^+ \times \mathbb{R}_0^+)$ such that*

$$\langle \mu,\, \partial_t \varphi + v(x)\partial_x \varphi \rangle = -\langle \mu_0,\, \varphi \rangle - \langle \nu_0,\, \varphi \rangle, \quad \forall\, \varphi \in C_0^1(\mathbb{R}_0^+ \times \mathbb{R}_0^+). \tag{3.10}$$

Since (3.9) is a linear problem, its solution can be obtained from the superposition of two measures $\mu^1,\, \mu^2 \in \mathcal{M}^+(\mathbb{R}_0^+ \times \mathbb{R}_0^+)$, where μ^1 is the solution to (3.9) with data $\mu_{t=0} = \mu_0$ and $\nu_{x=0} = 0$ while μ^2 is the solution to (3.9) with data $\mu_{t=0} = 0$ and $\nu_{x=0} = \nu_0$. This is doable in a standard way in terms of the push-forward of the initial and boundary data by means of appropriate vector fields in $\mathbb{R}_0^+ \times \mathbb{R}_0^+$, cf. [2]. With this approach time and space play the same role, the former being understood in particular as an additional state variable of the system.

However, for the next purposes it is convenient to characterise the solution μ to (3.9) by means of the traces of μ^1 and μ^2 over the fibres $\mathbb{R}_0^+ \times \{t\}$, $t > 0$; i.e.

$$\mu(dx\, dt) = (\mu_t^1(dx) + \mu_t^2(dx)) \otimes dt,$$

where μ_t^1, μ_t^2 are given by the transport of μ_0, ν_0, respectively, along the characteristics generated in $\mathbb{R}^+ \times \mathbb{R}^+$ by the velocity field v.

In order to obtain a formula for μ_t^1, let $\Phi_t = \Phi_t(x,\, 0)$ be the position at time $t > 0$ of the particle which is in $x \in \mathbb{R}_0^+$ at time $t = 0$ and which moves following the velocity field $v = v(x)$:

$$\begin{cases} \dfrac{d}{dt}\Phi_t(x,\, 0) = v(\Phi_t(x,\, 0)), \quad t > 0 \\[2ex] \Phi_0(x,\, 0) = x. \end{cases} \tag{3.11}$$

By standard results, it is well known that

$$\mu_t^1 = \Phi_t \# \mu_0 = \int_{\mathbb{R}_0^+} \delta_{\Phi_t(x,0)}\, d\mu_0(x) \in \mathcal{M}^+(\mathbb{R}_0^+ \times \{t\}),$$

where $\#$ is the push-forward operator, δ is the Dirac delta measure, and the integral at the right-hand side is understood in the sense of Bochner.

Likewise, to obtain a formula for μ_t^2 we consider the characteristic lines issuing from the t axis. In particular, we denote now by $\Phi_t(0,\, s)$ the position at time $t > 0$ of the particle which is in $x = 0$ at time $s \in \mathbb{R}_0^+$ and which moves following the velocity field $v = v(x)$:

$$\begin{cases} \dfrac{d}{dt}\Phi_t(0,\, s) = v(\Phi_t(0,\, s)), \quad t > s \\ \\ \Phi_s(0,\, s) = 0. \end{cases} \tag{3.12}$$

By transporting the mass ν_0 along these characteristics we can write

$$\mu_t^2 = \int_{[0,\, t]} \delta_{\Phi_t(0,\, s)}\, d\nu_0(s) \in \mathcal{M}^+(\mathbb{R}_0^+ \times \{t\}),$$

where the integral is again meant in the sense of Bochner.

Summing up, we consider the following representation formula for μ:

$$\mu(dx\, dt) = \left(\int_{\mathbb{R}_0^+} \delta_{\Phi_t(\xi,\, 0)}(dx)\, d\mu_0(\xi) + \int_{[0,\, t]} \delta_{\Phi_t(0,\, s)}(dx)\, d\nu_0(s) \right) \otimes dt \tag{3.13}$$

and we check that it actually defines a solution to (3.9) in the sense of Definition 3.4. To this purpose we preliminarily observe that, since $\mu_t^1 = \Phi_t \# \mu_0$, for every (bounded and measurable) function $f : \mathbb{R}_0^+ \to \mathbb{R}$ it results

$$\int_{\mathbb{R}_0^+} f(x)\, d\mu_t^1(x) = \int_{\mathbb{R}_0^+} f(\Phi_t(x,\, 0))\, d\mu_0(x). \tag{3.14}$$

We can obtain a similar formula for μ_t^2 by observing that, given a simple function $f : \mathbb{R}_0^+ \to \mathbb{R}$, $f(x) = \sum_{k=1}^N \alpha_k \chi_{A_k}(x)$, where $\{A_k\}_{k=1}^N$ is a measurable finite disjoint partition of $\mathbb{R}_0^+$, it results

$$\begin{aligned} \int_{\mathbb{R}_0^+} f(x)\, d\mu_t^2(x) &= \sum_{k=1}^N \alpha_k \mu_t^2(A_k) = \sum_{k=1}^N \alpha_k \int_{[0,\, t]} \delta_{\Phi_t(0,\, s)}(A_k)\, d\nu_0(s) \\ &= \sum_{k=1}^N \alpha_k \int_{[0,\, t]} \chi_{A_k}(\Phi_t(0,\, s))\, d\nu_0(s) \\ &= \int_{[0,\, t]} \sum_{k=1}^N \alpha_k \chi_{A_k}(\Phi_t(0,\, s))\, d\nu_0(s) \\ &= \int_{[0,\, t]} f(\Phi_t(0,\, s))\, d\nu_0(s). \end{aligned}$$

Approximating a measurable function f with a sequence of simple functions we get in general

$$\int_{\mathbb{R}_0^+} f(x)\, d\mu_t^2(x) = \int_{[0,\, t]} f(\Phi_t(0,\, s))\, d\nu_0(s). \tag{3.15}$$

Interestingly, an integral with respect to the x variable is converted into one with respect to the t

variable.

Plugging (3.13) into the left-hand side of (3.10) and using (3.14), (3.15) we discover:

$$
\begin{aligned}
&\langle \mu,\, \partial_t\varphi + v(x)\partial_x\varphi \rangle \\
&\quad = \int_{\mathbb{R}_0^+}\int_{\mathbb{R}_0^+} \Big(\partial_t\varphi(\Phi_t(x,\,0),\,t) + v(\Phi_t(x,\,0))\partial_x\varphi(\Phi_t(x,\,0),\,t)\Big)\,d\mu_0(x)\,dt \\
&\qquad + \int_{\mathbb{R}_0^+}\int_{[0,\,t]} (\partial_t\varphi(\Phi_t(0,\,s),\,t) + v(\Phi_t(0,\,s))\partial_x\varphi(\Phi_t(0,\,s),\,t))\,d\nu_0(s)\,dt \\
&\quad = \int_{\mathbb{R}_0^+}\int_{\mathbb{R}_0^+} \frac{d}{dt}\varphi(\Phi_t(x,\,0),\,t)\,d\mu_0(x)\,dt + \int_{\mathbb{R}_0^+}\int_{[0,\,t]} \frac{d}{dt}\varphi(\Phi_t(0,\,s),\,t)\,d\nu_0(s)\,dt,
\end{aligned}
$$

where in the last passage we have invoked (3.11), (3.12). By switching the order of integration in view of Fubini-Tonelli's Theorem we further obtain

$$
\begin{aligned}
&= \int_{\mathbb{R}_0^+}\int_{\mathbb{R}_0^+} \frac{d}{dt}\varphi(\Phi_t(x,\,0),\,t)\,dt\,d\mu_0(x) + \int_{\mathbb{R}_0^+}\int_{[s,\,+\infty)} \frac{d}{dt}\varphi(\Phi_t(0,\,s),\,t)\,dt\,d\nu_0(s) \\
&= \int_{\mathbb{R}_0^+} \Big[\varphi(\Phi_t(x,\,0),\,t)\Big]_{t=0}^{t=+\infty} d\mu_0(x) + \int_{\mathbb{R}_0^+} \Big[\varphi(\Phi_t(0,\,s),\,t)\Big]_{t=s}^{t=+\infty} d\nu_0(s) \\
&= -\int_{\mathbb{R}_0^+} \varphi(x,\,0)\,d\mu_0(x) - \int_{\mathbb{R}_0^+} \varphi(0,\,s)\,d\nu_0(s) \\
&= -\langle \mu_0,\,\varphi\rangle - \langle \nu_0,\,\varphi\rangle,
\end{aligned}
$$

which confirms that (3.13) is indeed a measure-valued solution to (3.9). Uniqueness of such a solution is a consequence of continuous dependence estimates on the initial and boundary data, which can be proved by standard arguments in literature, cf. [2]. In conclusion, for the transport problem in $\mathbb{R}^+ \times \mathbb{R}^+$ we have the following well-posedness result:

Theorem 3.1. *For $\mu_0 \in \mathcal{M}^+(\mathbb{R}_0^+ \times \{0\})$, $\nu_0 \in \mathcal{M}^+(\{0\} \times \mathbb{R}_0^+)$ there exists a unique measure-valued solution to* (3.9) *in the sense of Definition 3.4, which can be represented by* (3.13).

We now pass to consider the transport problem on the bounded domain $Q := (0,\,1) \times (0,\,T)$, $T > 0$, i.e.

$$
\begin{cases}
\partial_t\mu + \partial_x(v(x)\mu) = 0, & (x,\,t) \in Q \\
\mu_{t=0} = \mu_0 \in \mathcal{M}^+([0,\,1] \times \{0\}) \\
\nu_{x=0} = \nu_0 \in \mathcal{M}^+(\{0\} \times [0,\,T])
\end{cases}
\tag{3.16}
$$

for a given bounded, strictly positive and Lipschitz continuous velocity field $v : [0,\,1] \to (0,\,v_{\max}]$. The solution to this problem can be obtained by restricting to Q the measure μ solving (3.9) (with the velocity field v possibly extended to the whole $\mathbb{R}_0^+$ as, e.g. $v(x) = v(1)$ for $x \geq 1$). Therefore we are going to consider the restriction of μ to Q defined as the measure $\mu \llcorner Q$ such that

$$\mu \llcorner Q(E) := \mu(E \cap Q)$$

for every measurable set $E \subseteq \mathbb{R}_0^+ \times \mathbb{R}_0^+$.

In particular, in view of the application of this problem to a network, it is important to characterise the traces of $\mu \llcorner Q$ on the fibres $[0,\,1] \times \{T\}$ and $\{1\} \times [0,\,T]$, which depend on the transport of μ_0 and ν_0 inside Q.

Let us introduce the exit time:

$$\theta(x, s) := \inf\{t \geq s\,:\,\Phi_t(x,\,s) = 1\}, \quad x \in [0,\,1], s \in [0,\,T] \tag{3.17}$$

corresponding to the time needed to the characteristic line issuing from either $(x,\,s)$ to hit the boundary $x = 1$. Since the considered transport problem is linear, in particular the velocity field v does not depend on the measure μ itself, both $\tau := \theta(\cdot, 0)$ and $\varsigma := \theta(0, \cdot)$ are one-to-one, thus invertible.

Recalling (3.13) and using τ, σ we write the trace of $\mu \llcorner Q$ on the fibre $[0,\,1] \times \{T\}$ as (cf. Figure 3.1)

$$\mu_T := \int_{[0,\,\max\{0,\,\tau^{-1}(T)\}]} \delta_{\Phi_T(x,\,0)}\,d\mu_0(x) + \int_{[\max\{0,\,\sigma^{-1}(T)\},\,T]} \delta_{\Phi_T(0,\,s)}\,d\nu_0(s) \tag{3.18}$$

whereas, following the characteristics, we construct the trace on the fibre $\{1\} \times [0,\,T]$ as

$$\nu_1 := \int_{(\max\{0,\,\tau^{-1}(T)\},\,1]} \delta_{\tau(x)}\,d\mu_0(x) + \int_{[0,\,\max\{0,\,\sigma^{-1}(T)\})} \delta_{\sigma(s)}\,d\nu_0(s). \tag{3.19}$$

We incidentally notice that the first term at the right-hand side of (3.18) is the push-forward of μ_0 by the flow map Φ_T then restricted to $x \in [0,\,1]$ while the second term at the right-hand side of (3.19) is the push-forward of ν_0 by the mapping σ then restricted to $t \in [0,\,T]$.

The relationship between these traces and the transport of μ_0, ν_0 inside Q is rigorously stated by the following theorem, which represents our main result on problem (3.16):

Theorem 3.2. *Given $\mu_0 \in \mathcal{M}^+([0,\,1] \times \{0\})$, $\nu_0 \in \mathcal{M}^+(\{0\} \times [0,\,T])$, the measure $\mu \llcorner Q \in \mathcal{M}^+(\bar{Q})$, $\mu \in \mathcal{M}^+(\mathbb{R}_0^+ \times \mathbb{R}_0^+)$ being the solution to* (3.9)*, is the unique measure which satisfies the balance*

$$\langle \mu \llcorner Q,\,\partial_t\varphi + v(x)\partial_x\varphi \rangle = \langle \mu_T - \mu_0,\,\varphi \rangle + \langle \nu_1 - \nu_0,\,\varphi \rangle, \quad \forall\,\varphi \in C^1(\bar{Q}), \tag{3.20}$$

where $\mu_T \in \mathcal{M}^+([0,\,1] \times \{T\})$, $\nu_1 \in \mathcal{M}^+(\{1\} \times [0,\,T])$ are the traces defined in (3.18), (3.19)*, respec-*

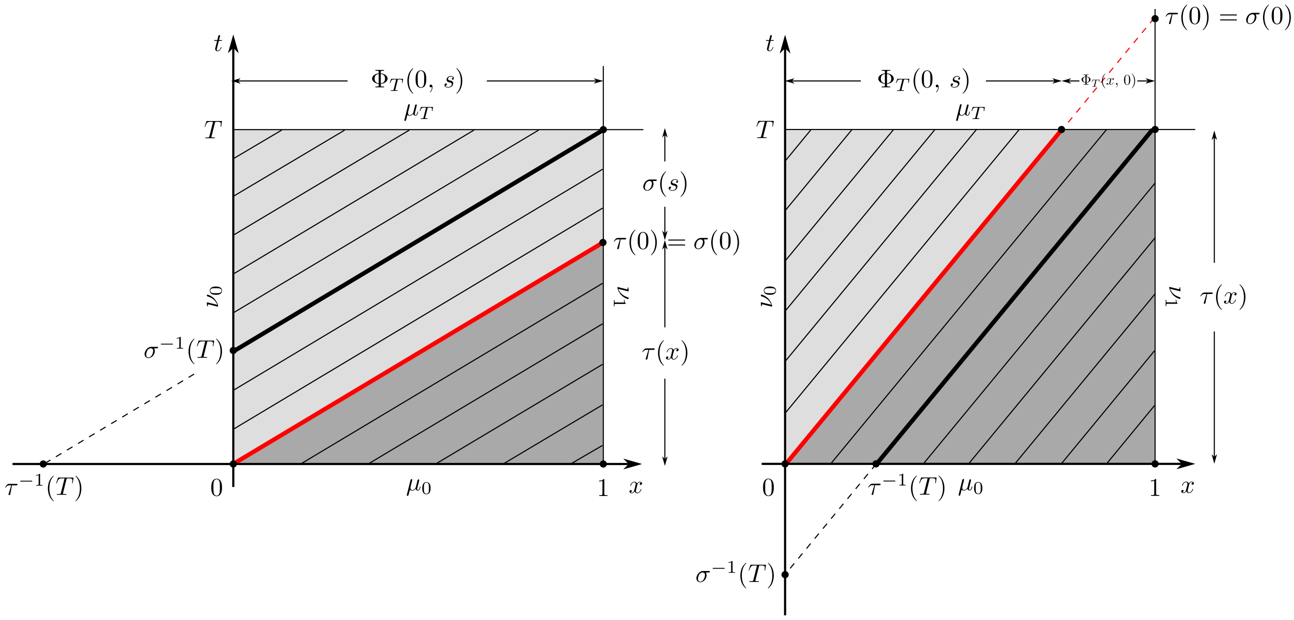


Figure 3.1: Sketch of the characteristics of problem (3.16) in the two cases $\tau(0) = \sigma(0) < T$ (left) and $\tau(0) = \sigma(0) > T$ (right). For pictorial purposes we imagine a constant velocity field, so that the characteristics are straight lines in the space-time.

tively.

Moreover, for $\mu_0^k \in \mathcal{M}^+([0,\, 1] \times \{0\})$*,* $\nu_0^k \in \mathcal{M}^+(\{0\} \times [0,\, T])$*,* $k = 1,\, 2$*, there exists a constant* $C = C(T) > 0$ *such that*

$$\left\|\mu_T^2 - \mu_T^1\right\|_{BL}^* + \left\|\nu_1^2 - \nu_1^1\right\|_{BL}^* \leq C\left(\left\|\mu_0^2 - \mu_0^1\right\|_{BL}^* + \left\|\nu_0^2 - \nu_0^1\right\|_{BL}^*\right). \tag{3.21}$$

We also give a result about the dependence on time.

Theorem 3.3. *Given* $\mu_0 \in \mathcal{M}^+([0,\, 1] \times \{0\})$*,* $\nu_0 \in \mathcal{M}^+(\{0\} \times [0,\, T])$*, there exists a constant* $C = C(T) > 0$ *such that*

$$\|\mu_t - \mu_{t'}\|_{BL}^* + \left\|\nu_1 \llcorner [0,\, t] - \nu_1 \llcorner [0,\, t']\right\|_{BL}^* \leq C\left|t - t'\right| + \nu_0([t',\, t]) \tag{3.22}$$

for all $t',\, t \in [0,\, T]$ *with* $t' < t$*.*

Remark 3.3.1. *Theorem 3.3 states virtually that the traces* μ_t *and* $\nu_1 \llcorner [0,\, t]$ *of* $\mu \llcorner Q$ *are Lipschitz continuous in time, a part from the presence of the term* $\nu_0([t',\, t])$ *in the estimate* (3.22)*.*

If the boundary datum ν_0 *is absolutely continuous with respect to the Lebesgue measure in the interval* $[t',\, t]$ *then for* $t \to t'$ *we get actually* $\|\mu_t - \mu_{t'}\|_{BL}^* + \|\nu_1 \llcorner [0,\, t] - \nu_1 \llcorner [0,\, t']\|_{BL}^* \to 0$*. If instead* ν_0 *contains singularities in* $[t',\, t]$ *then the distances* $\|\mu_t - \mu_{t'}\|_{BL}^*$*,* $\|\nu_1 \llcorner [0,\, t] - \nu_1 \llcorner [0,\, t']\|_{BL}^*$ *between two traces on horizontal and vertical fibres are in general not proportional to the time gap* $|t - t'|$*.*

In the applications, a Lebesgue-absolutely continuous ν_0 *corresponds to a* macroscopic *inflow mass provided with density. A Lebesgue-singular* ν_0 *corresponds instead to* microscopic *point masses flowing from the boundary* $x = 0$ *during the time interval* $[t',\, t]$ *and then propagating as singularities across* Q.

Proof of Theorem 3.3. We begin with the estimate of $\|\mu_t - \mu_{t'}\|^*_{BL}$. Let $\varphi \in BL(Q)$ be such that $\|\varphi\|_{BL} \leq 1$.
By (3.18), since

$$(-\infty,\, \tau^{-1}(t')) = (-\infty,\, \tau^{-1}(t)) \cup [\tau^{-1}(t),\, \tau^{-1}(t')],$$

we can write:

$$\begin{aligned}
&\int_{(-\infty,\, \tau^{-1}(t))\cap[0,\,1)} \varphi(\Phi_t(x,\,0),\,t)\,d\mu_0(x) - \int_{(-\infty,\, \tau^{-1}(t'))\cap[0,\,1)} \varphi(\Phi_{t'}(x,\,0),\,t')\,d\mu_0(x) \\
&= \int_{(-\infty,\, \tau^{-1}(t))\cap[0,\,1]} \Big(\varphi(\Phi_t(x,\,0),\,t) - \varphi(\Phi_{t'}(x,\,0),\,t')\Big)\,d\mu_0(x) \\
&\quad - \int_{[\tau^{-1}(t),\, \tau^{-1}(t'))\cap[0,\,1]} \varphi(\Phi_{t'}(x,\,0),\,t')\,d\mu_0(x) \\
&\leq \mu_0((-\infty,\, \tau^{-1}(t))\cap[0,\,1))\,\|v\|_\infty\,\left|t - t'\right| - \int_{[\tau^{-1}(t),\, \tau^{-1}(t'))\cap[0,\,1]} \varphi(\Phi_{t'}(x,\,0),\,t')\,d\mu_0(x).
\end{aligned}$$

Likewise, assuming for simplicity that $\sigma^{-1}(t) \leq t'$,

$$\begin{aligned}
&\int_{(\sigma^{-1}(t),\, t]\cap(0,\,T]} \varphi(\Phi_t(0,\,s),\,t)\,d\nu_0(s) - \int_{(\sigma^{-1}(t'),\, t']\cap(0,\,T]} \varphi(\Phi_t(0,\,s),\,t)\,d\nu_0(s) \\
&= -\int_{(\sigma^{-1}(t'),\, \sigma^{-1}(t)]} \varphi(\Phi_{t'}(0,\,s),\,t')\,d\nu_0(s) + \int_{(\sigma^{-1}(t),\, t']} \Big(\varphi(\Phi_t(0,\,s),\,t) - \varphi(\Phi_{t'}(0,\,s),\,t')\Big)\,d\nu_0(s) \\
&\quad + \int_{(t',\,t]} \varphi(\Phi_t(0,\,s),\,t)\,d\nu_0(s) \\
&\leq \nu_0((t',\,t]) + \nu_0((t - \tau(0),\, t'])\,\|v\|_\infty\,\left|t - t'\right| - \int_{(\sigma^{-1}(t'),\, \sigma^{-1}(t)]} \varphi(\Phi_{t'}(0,\,s),\,t')\,d\nu_0(s).
\end{aligned}$$

Hence

$$
\begin{aligned}
|\langle \mu_t - \mu_{t'},\, \varphi \rangle| &\le \left| \int_{(\tau^{-1}(t),\, \tau^{-1}(t')] \cap [0,\,1]} \Big(\varphi(\Phi_{t'}(x,\,0),\,t') - \varphi(1,\,\tau(x)) \Big)\, d\mu_0(x) \right| \\
&\quad + \left| \int_{(\sigma^{-1}(t'),\, \sigma^{-1}(t)]} \Big(\varphi(1,\,\sigma(s)) - \varphi(\Phi_{t'}(0,\,s),\,t') \Big)\, d\nu_0(s) \right| \\
&\le \mu_0((\tau^{-1}(t),\, \tau^{-1}(t')] \cap [0,\,1])\, \|v\|_\infty \left| t - t' \right| + \nu_0((\sigma^{-1}(t'),\, \sigma^{-1}(t)])\, \|v\|_\infty \left| t - t' \right| \\
&\le \|v\|_\infty \Big(\mu_0([0,\,1]) + \nu_0([0,\,t]) \Big) \left| t - t' \right| + \nu_0((t',\,t]) \\
&\le C \left| t - t' \right| + \nu_0([t',\,t])
\end{aligned}
$$

and finally, taking the supremum over φ at both sides,

$$
\|\mu_t - \mu_{t'}\|_{BL}^{*} \le C \left| t - t' \right| + \nu_0([t',\,t]).
$$

We now consider the estimate on the outflow measures. Taking again $\varphi \in BL(Q)$ with $\|\varphi\|_{BL} \le 1$, we compute:

$$
\begin{aligned}
\langle \nu_1 \llcorner [0,\,t] - \nu_1 \llcorner [0,\,t'],\, \varphi \rangle = &\int_{[0,\,1) \cap [\tau^{-1}(t),\,1)} \varphi(1,\,\tau(x))\, d\mu_0(x) + \int_{(0,\,t] \cap (0,\,\sigma^{-1}(t)]} \varphi(1,\,\sigma(s))\, d\nu_0(s) \\
&- \int_{[0,\,1) \cap [\tau^{-1}(t'),\,1)} \varphi(1,\,\tau(x))\, d\mu_0(x) - \int_{(0,\,t'] \cap (0,\,\sigma^{-1}(t')]} \varphi(1,\,\sigma(s))\, d\nu_0(s).
\end{aligned}
$$

We point out that if $\sigma^{-1}(t) < 0$ then the interval $(0,\, \sigma^{-1}(t)]$ is actually understood as $[\sigma^{-1}(t),\, 0)$ and, in this case, $(0,\,t] \cap (0,\,\sigma^{-1}(t)] = \emptyset$. Moreover, since $t > t'$ we have $\tau^{-1}(t') > \tau^{-1}(t)$, which implies $[\tau^{-1}(t'),\,1) = [\tau^{-1}(t'),\,\tau^{-1}(t)) \cup [\tau^{-1}(t),\,1)$. Then

$$
\begin{aligned}
&\int_{[0,\,1) \cap [\tau^{-1}(t),\,1)} \varphi(1,\,\tau(x))\, d\mu_0(x) - \int_{[0,\,1) \cap [\tau^{-1}(t'),\,1)} \varphi(1,\,\tau(x))\, d\mu_0(x) \\
&\qquad\qquad = \int_{[0,\,1) \cap [\tau^{-1}(t),\,\tau^{-1}(t'))} \varphi(1,\,\tau(x))\, d\mu_0(x).
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
&\int_{(0,\,\sigma^{-1}(t)] \cap (0,\,t]} \varphi(1,\,\sigma(s))\, d\nu_0(s) \\
&\qquad = \int_{(0,\,\sigma^{-1}(t')] \cap (0,\,t]} \varphi(1,\,\sigma(s))\, d\nu_0(s) + \int_{(\sigma^{-1}(t'),\,\sigma^{-1}(t)] \cap (0,\,t]} \varphi(1,\,\sigma(s))\, d\nu_0(s),
\end{aligned}
$$

which gives

$$\int_{(0,\sigma^{-1}(t)]\cap(0,t]} \varphi(1,\sigma(s))\,d\nu_0(s) - \int_{(0,\sigma^{-1}(t')]\cap(0,t']} \varphi(1,\sigma(s))\,d\nu_0(s)$$
$$= \int_{(\sigma^{-1}(t'),\sigma^{-1}(t)]\cap(0,t]} \varphi(1,\sigma(s))\,d\nu_0(s).$$

Therefore

$$\begin{aligned}\langle \nu_1\llcorner[0,t] - \nu_1\llcorner[0,t'],\,\varphi\rangle &= \int_{[0,1)\cap[\tau^{-1}(t),\tau^{-1}(t'))} \varphi(1,\tau(x))\,d\mu_0(x) + \int_{(\sigma^{-1}(t'),\sigma^{-1}(t)]\cap(0,t]} \varphi(1,\sigma(s))\,d\nu_0(s)\\ &\le \nu_0((t',t]) + \nu_0((\sigma^{-1}(t),t'])\,\|v\|_\infty\,|t-t'|\\ &\quad + \mu_0((-\infty,\tau^{-1}(t))\cap[0,1))\,\|v\|_\infty\,|t-t'|,\end{aligned}$$

whence, taking the supremum over φ at both sides,

$$\left\|\nu_1\llcorner[0,t] - \nu_1\llcorner[0,t']\right\|^*_{BL} \le C\,|t-t'| + \nu_0([t',t]).$$

Summing the estimates obtained so far for $\|\mu_t - \mu_{t'}\|^*_{BL}$, $\|\nu_1\llcorner[0,t] - \nu_1\llcorner[0,t']\|^*_{BL}$ we finally get (3.22).

□

Proof of Theorem 4.2. We observe that μ can be obtained, by linearity, as the sum of the solutions of two transport problems with $\nu_0 = 0$ and $\mu_0 = 0$, respectively. We begin by considering the case $\nu_0 = 0$ and assume, without loss of generality, that $T \le \tau(0)$. Then $\tau^{-1}(T) \ge 0$ whence, recalling (3.18), (3.19), we obtain

$$\mu_T = \int_{[0,\tau^{-1}(T)]} \delta_{\Phi_T(x,0)}\,d\mu_0(x), \qquad \nu_1 = \int_{(\tau^{-1}(T),1]} \delta_{\tau(x)}\,d\mu_0(x) \tag{3.23}$$

and we have to show that

$$\langle \mu\llcorner Q,\,\partial_t\varphi + v(x)\partial_x\varphi\rangle = \langle \mu_T - \mu_0,\,\varphi\rangle + \langle \nu_1,\,\varphi\rangle, \qquad \forall\,\varphi \in C^1(\bar{Q}), \tag{3.24}$$

where μ is the measure (3.13). Following the characteristics, its restriction to Q writes as

$$\mu\llcorner Q(dx\,dt) = \underbrace{\int_{[0,\tau^{-1}(t)]} \delta_{\Phi_t(\xi,0)}(dx)\,d\mu_0(\xi)}_{:=\mu_t\llcorner Q(dx)} \otimes dt,$$

thus for $\varphi \in C^1(\bar{Q})$ we discover:

$$\begin{aligned}
\langle \mu \llcorner Q,\, \partial_t\varphi + v(x)\partial_x\varphi\rangle &= \int_0^T \int_{[0,\,1]} (\partial_t\varphi + v(x)\partial_x\varphi)\; d\mu_t \llcorner Q(x)\, dt \\
&= \int_0^T \int_{[0,\,\tau^{-1}(t)]} \Big(\partial_t\varphi(\Phi_t(x,\,0),\,t) + v(\Phi_t(x,\,0))\partial_x\varphi(\Phi_t(x,\,0),\,t)\Big)\, d\mu_0(x)\, dt \\
&= \int_0^T \int_{[0,\,\tau^{-1}(t)]} \frac{d}{dt}\varphi(\Phi_t(x,\,0),\,t)\, d\mu_0(x)\, dt,
\end{aligned}$$

where in the last passage we have used (3.11). Switching the order of integration, we continue the calculation as:

$$\begin{aligned}
&= \int_{[0,\,1]} \int_0^{\min\{\tau(x),\,T\}} \frac{d}{dt}\varphi(\Phi_t(x,\,0),\,t)\, dt\, d\mu_0(x) \\
&= \int_{[0,\,\tau^{-1}(T)]} \int_0^T \frac{d}{dt}\varphi(\Phi_t(x,\,0),\,t)\, dt\, d\mu_0(x) \\
&\quad + \int_{(\tau^{-1}(T),\,1]} \int_0^{\tau(x)} \frac{d}{dt}\varphi(\Phi_t(x,\,0),\,t)\, dt\, d\mu_0(x) \\
&= \int_{[0,\,\tau^{-1}(T)]} \Big(\varphi(\Phi_T(x,\,0),\,T) - \varphi(\Phi_0(x,\,0),\,0)\Big)\, d\mu_0(x) \\
&\quad + \int_{(\tau^{-1}(T),\,1]} \Big(\varphi(\Phi_{\tau(x)}(x,\,0),\,\tau(x)) - \varphi(\Phi_0(x,\,0),\,0)\Big)\, d\mu_0(x) \\
&= \underbrace{\int_{[0,\,\tau^{-1}(T)]} \varphi(\Phi_T(x,\,0),\,T)\, d\mu_0(x)}_{\text{(i)}} + \underbrace{\int_{(\tau^{-1}(T),\,1]} \varphi(1,\,\tau(x))\, d\mu_0(x)}_{\text{(ii)}} \\
&\quad - \underbrace{\int_{[0,\,1]} \varphi(x,\,0)\, d\mu_0(x)}_{\text{(iii)}}.
\end{aligned}$$

From (3.23) we recognise that the term (i) is indeed $\int_{[0,\,1]} \varphi(x,\,T)\, d\mu_T(x) = \langle \mu_T,\, \varphi\rangle$ and that the term (ii) is $\int_{[0,\,T]} \varphi(1,\,t)\, d\nu_1(t) = \langle \nu_1,\, \varphi\rangle$, while the term (iii) is clearly $\langle \mu_0,\, \varphi\rangle$. Consequently (3.24) follows. We consider now the case $\mu_0 = 0$ and assume, without loss of generality, that $T \geq \sigma(0)$. Then $\sigma^{-1}(T) \geq 0$ whence, recalling again (3.18), (3.19), we find

$$\mu_T = \int_{[\sigma^{-1}(T),\,T]} \delta_{\Phi_T(0,\,s)}\, d\nu_0(s), \qquad \nu_1 = \int_{[0,\,\sigma^{-1}(T))} \delta_{\sigma(s)}\, d\nu_0(s) \tag{3.25}$$

and we have to show that

$$\langle \mu \llcorner Q,\, \partial_t\varphi + v(x)\partial_x\varphi\rangle = \langle \mu_T,\, \varphi\rangle + \langle \nu_1 - \nu_0,\, \varphi\rangle, \qquad \forall\, \varphi \in C^1(\bar{Q}), \tag{3.26}$$

where μ is again the measure (3.13). Following the characteristics we see that $\mu \llcorner Q$ is now expressed as

$$\mu\llcorner Q(dx\,dt) = \underbrace{\int_{[\max\{0,\sigma^{-1}(t)\},\,t]} \delta_{\Phi_t(0,\,s)}(dx)\,d\nu_0(s)}_{:=\mu_t\llcorner Q(dx)} \otimes dt,$$

hence for $\varphi \in C^1(\bar{Q})$ we obtain:

$$\begin{aligned}
&\langle \mu\llcorner Q, \partial_t\varphi + v(x)\partial_x\varphi\rangle \\
&\quad = \int_0^T \int_{[0,\,1]} \left(\partial_t\varphi + v(x)\partial_x\varphi\right)\,d\mu_t\llcorner Q(x)\,dt \\
&\quad = \int_0^T \int\limits_{[\max\{0,\sigma^{-1}(t)\},\,t]} \Big(\partial_t\varphi(\Phi_t(0,\,s),\,t) + v(\Phi_t(0,\,s))\partial_x\varphi(\Phi_t(0,\,s)\,t)\Big)\,d\nu_0(s)\,dt \\
&\quad = \int_0^T \int_{[\max\{0,\sigma^{-1}(t)\},\,t]} \frac{d}{dt}\varphi(\Phi_t(0,\,s),\,t)\,d\nu_0(s)\,dt,
\end{aligned}$$

where in the last passage we have used (3.12). We now switch the order of integration to discover:

$$\begin{aligned}
&= \int_{[0,\,T]} \int_s^{\min\{\sigma(s),\,T\}} \frac{d}{dt}\varphi(\Phi_t(0,\,s),\,t)\,dt\,d\nu_0(s) \\
&= \int_{[0,\,\sigma^{-1}(T)]} \int_s^{\sigma(s)} \frac{d}{dt}\varphi(\Phi_t(0,\,s),\,t)\,dt\,d\nu_0(s) \\
&\quad + \int_{(\sigma^{-1}(T),\,T]} \int_s^{T} \frac{d}{dt}\varphi(\Phi_t(0,\,s),\,t)\,dt\,d\nu_0(s) \\
&= \int_{[0,\,\sigma^{-1}(T)]} \Big(\varphi(\Phi_{\sigma(s)}(0,\,s),\,\sigma(s)) - \varphi(\Phi_s(0,\,s),\,s)\Big)\,d\nu_0(s) \\
&\quad + \int_{(\sigma^{-1}(T),\,T]} \Big(\varphi(\Phi_T(0,\,s)\,T) - \varphi(\Phi_s(0,\,s),\,s)\Big)\,d\nu_0(s) \\
&= \underbrace{\int_{[0,\,\sigma^{-1}(T)]} \varphi(1,\,\sigma(s))\,d\nu_0(s)}_{\text{(i)}} + \underbrace{\int_{(\sigma^{-1}(T),\,T]} \varphi(\Phi_T(0,\,s),\,T)\,d\nu_0(s)}_{\text{(ii)}} \\
&\quad - \underbrace{\int_{[0,\,T]} \varphi(0,\,s)\,d\nu_0(s)}_{\text{(iii)}}.
\end{aligned}$$

Thanks to (3.25) we recognise that the term (i) is $\int_{[0,\,T]} \varphi(1,\,t)\,d\nu_1(t) = \langle \nu_1,\,\varphi\rangle$ and that the term (ii) is $\int_{[0,\,1]} \varphi(x,\,T)\,d\mu_T(x) = \langle \mu_T,\,\varphi\rangle$, while the term (iii) is clearly $\langle \nu_0,\,\varphi\rangle$. Hence (3.26) follows.

To conclude the proof, we show the continuous dependence estimate (3.21). We consider two problems of the type (3.16) with respective initial data μ_0^1, μ_0^2 and source data ν_0^1, ν_0^2.

We begin by estimating the term $\left\|\mu_T^2 - \mu_T^1\right\|_{BL}^*$. Let $\varphi \in BL(Q)$ with $\|\varphi\|_{BL} \le 1$. Recalling (3.18) we have:

$$\begin{aligned}
\langle \mu_T^2 - \mu_T^1, \varphi \rangle &= \int_{[0,\,1]} \varphi(x,\,T)\, d(\mu_T^2 - \mu_T^1)(x) \\
&= \int_{[0,\,\max\{0,\,\tau^{-1}(T)\}]} \varphi(\Phi_T(x,\,0),\,T)\, d(\mu_0^2 - \mu_0^1)(x) \\
&\quad + \int_{[\max\{0,\,\sigma^{-1}(T)\},\,T]} \varphi(\Phi_T(0,\,s),\,T)\, d(\nu_0^2 - \nu_0^1)(s) \\
&\le \left|\mu_0^2 - \mu_0^1\right| ([0,\,\max\{0,\,\tau^{-1}(T)\}]) \\
&\quad + \left|\nu_0^2 - \nu_0^1\right| ([\max\{0,\,\sigma^{-1}(T)\},\,T])
\end{aligned}$$

where here $|\cdot|$ stands for the total variation of a measure. Thus

$$\le C \left(\left\|\mu_0^2 - \mu_0^1\right\|_{BL}^* + \left\|\nu_0^2 - \nu_0^1\right\|_{BL}^* \right)$$

and consequently, taking the supremum over φ at both sides,

$$\left\|\mu_T^2 - \mu_T^1\right\|_{BL}^* \le C \left(\left\|\mu_0^2 - \mu_0^1\right\|_{BL}^* + \left\|\nu_0^2 - \nu_0^1\right\|_{BL}^* \right).$$

Proceeding in a similar way for $\left\|\nu_1^2 - \nu_1^1\right\|_{BL}^*$, from (3.19) we have:

$$\begin{aligned}
\langle \nu_1^2 - \nu_1^1, \varphi \rangle &= \int_{[0,\,T]} \varphi(1,\,t)\, d(\nu_1^2 - \nu_1^1)(t) \\
&= \int_{(\max\{0,\,\tau^{-1}(T)\},\,1]} \varphi(1,\,\tau(x))\, d(\mu_0^2 - \mu_0^1)(x) \\
&\quad + \int_{[0,\,\max\{0,\,\sigma^{-1}(T)\})} \varphi(1,\,\sigma(s))\, d(\nu_0^2 - \nu_0^1)(s) \\
&\le \left|\mu_0^2 - \mu_0^1\right| ((\max\{0,\,\tau^{-1}(T)\},\,1]) \\
&\quad + \left|\nu_0^2 - \nu_0^1\right| ([0,\,\max\{0,\,\sigma^{-1}(T)\})) \\
&\le C \left(\left\|\mu_0^2 - \mu_0^1\right\|_{BL}^* + \left\|\nu_0^2 - \nu_0^1\right\|_{BL}^* \right),
\end{aligned}$$

hence, taking the supremum over φ at both sides,

$$\left\|\nu_1^2 - \nu_1^1\right\|_{BL}^* \le C \left(\left\|\mu_0^2 - \mu_0^1\right\|_{BL}^* + \left\|\nu_0^2 - \nu_0^1\right\|_{BL}^* \right).$$

Summing the two estimates just obtained yields finally (3.21).
Moreover, for $\mu_0^1 = \mu_0^2$, $\nu_0^1 = \nu_0^2$ the estimate (3.21) implies $\mu_T^1 = \mu_T^2$, $\nu_1^1 = \nu_1^2$, hence the uniqueness of (3.18) and (3.19). $\square$

3.2.2 Linear transport on networks

We now return to the study of problem (3.4). In order to make the notation consistent with the one introduced, we set

$$\nu_0^j := \mu^j_{x=\pi_j(0)}, \qquad \nu_1^j := \mu^j_{x=\pi_j(1)}$$

and we rewrite (3.4) as

$$\begin{cases} \partial_t \mu^j + \partial_x (v_j(x)\mu^j) = 0 & x \in e_j,\, t \in (0,\, T],\, e_j \in \mathcal{E} \\ \mu^j_{t=0} = \mu^j_0 & x \in e_j, \\ \nu^j_0 = \begin{cases} \displaystyle\sum_{k\,:\, V=\pi_k(1)} p_{kj} \cdot \nu^k_1 & \text{if } V = \pi_j(0) \in \mathcal{I} \\ \sigma^S_0 & \text{if } S \in \mathcal{S}. \end{cases} \end{cases} \tag{3.27}$$

Let $\varphi \in C^1(\Gamma \times [0,\, T])$. Given $\mu^j_0 \in \mathcal{M}^+(e_j \times \{0\})$, $\nu^j_0 \in \mathcal{M}^+(\{0\} \times [0,\, T])$, owing to Theorem 4.2 there exists $\mu^j \in \mathcal{M}^+(e_j \times [0,\, T])$ such that

$$\langle \mu^j,\, \partial_t \varphi + v_j(x)\partial_x \varphi \rangle = \langle \mu^j_T - \mu^j_0,\, \varphi \rangle + \langle \nu^j_1 - \nu^j_0,\, \varphi \rangle \tag{3.28}$$

for every $e_j \in \mathcal{E}$. Similarly to (3.18), (3.19), the traces μ^j_T, ν^j_1 are

$$\mu^j_T = \int_{[0,\, \max\{0,\, \tau_j^{-1}(T)\}]} \delta_{\Phi^j_T(x,0)}\, d\mu^j_0(x) + \int_{[\max\{0,\, \sigma_j^{-1}(T)\},\, T]} \delta_{\Phi^j_T(0,\, s)}\, d\nu^j_0(s) \tag{3.29}$$

$$\nu^j_1 = \int_{(\max\{0,\, \tau_j^{-1}(T)\},\, 1]} \delta_{\tau_j(x)}\, d\mu^j_0(x) + \int_{[0,\, \max\{0,\, \sigma_j^{-1}(T)\})} \delta_{\sigma_j(s)}\, d\nu^j_0(s), \tag{3.30}$$

where the flow maps $\Phi^j_t(x,\, 0)$ and $\Phi^j_t(0,\, s)$ are defined like in (3.11), (3.12), respectively, using the velocity field $v_j(x)$ on the arc e_j. Likewise, τ_j and σ_j are defined like in (3.17).

Summing (3.28) over j and recalling (3.5) we deduce

$$\langle \mu,\, \partial_t \varphi + v(x)\partial_x \varphi \rangle = \langle \mu_T - \mu_0,\, \varphi \rangle + \sum_{j\in J} \langle \nu^j_1 - \nu^j_0,\, \varphi \rangle, \tag{3.31}$$

where

$$\mu_0 = \sum_j \mu^j_0, \qquad \mu_T = \sum_j \mu^j_T. \tag{3.32}$$

In particular, the last term at the right-hand side in (3.31) can be rewritten in more detail by summing on the vertices of the network:

$$\begin{aligned}\sum_j \langle \nu_1^j - \nu_0^j,\, \varphi\rangle &= \sum_{V\in I}\left(\sum_{j\,:\,V=\pi_j(1)} \langle \nu_1^j,\, \varphi\rangle - \sum_{j\,:\,V=\pi_j(0)} \langle \nu_0^j,\, \varphi\rangle\right) \\ &= \sum_{V\in\mathcal{I}}\left(\sum_{j\,:\,V=\pi_j(1)} \langle \nu_1^j,\, \varphi\rangle - \sum_{j\,:\,V=\pi_j(0)} \langle \nu_0^j,\, \varphi\rangle\right) \\ &\quad + \sum_{V\in\mathcal{W}}\sum_{j\,:\,V=\pi_j(1)} \langle \nu_1^j,\, \varphi\rangle - \sum_{V\in\mathcal{S}}\sum_{j\,:\,V=\pi_j(0)} \langle \nu_0^j,\, \varphi\rangle.\end{aligned}$$

For an internal vertex $V \in \mathcal{I}$, using the corresponding boundary condition prescribed in (3.27) we obtain:

$$\begin{aligned}\sum_{j\,:\,V=\pi_j(1)} \langle \nu_1^j,\, \varphi\rangle - \sum_{j\,:\,V=\pi_j(0)} \langle \nu_0^j,\, \varphi\rangle &= \sum_{j\,:\,V=\pi_j(1)} \langle \nu_1^j,\, \varphi\rangle - \sum_{j\,:\,V=\pi_j(0)} \langle \sum_{k\,:\,V=\pi_k(1)} p_{kj}^i(t)\nu_1^k,\, \varphi\rangle \\ &= \sum_{j\,:\,V=\pi_j(1)} \langle \nu_1^j,\, \varphi\rangle - \sum_{k\,:\,V=\pi_k(1)} \langle \sum_{j\,:\,V=\pi_j(0)} p_{kj}^i(t)\nu_1^k,\, \varphi\rangle\end{aligned}$$

whence, taking (3.3) into account in the second term at the right-hand side,

$$\begin{aligned}\sum_{j\,:\,V=\pi_j(1)} \langle \nu_1^j,\, \varphi\rangle - \sum_{j\,:\,V=\pi_j(0)} \langle \nu_0^j,\, \varphi\rangle &= \sum_{j\,:\,V=\pi_j(1)} \langle \nu_1^j,\, \varphi\rangle - \sum_{k\,:\,V=\pi_k(1)} \langle \nu_1^k,\, \varphi\rangle \\ &= 0.\end{aligned}$$

This is the conservation of the mass through the internal vertices of the network.

For a source vertex $V \in \mathcal{S}$, we use the corresponding boundary condition prescribed in (3.27) to find:

$$\sum_{V\in\mathcal{S}}\sum_{j\,:\,V=\pi_j(0)} \langle \nu_0^j,\, \varphi\rangle = \sum_{V\in\mathcal{S}} \langle \sigma_0^V,\, \varphi\rangle = \langle \sigma_0,\, \varphi\rangle$$

where we have defined the measure $\sigma_0 := \sum_{S\in\mathcal{S}} \sigma_0^V \in \mathcal{M}^+(\cup_{\mathcal{S}}\{V\} \times [0,\, T])$. This is the total mass flowing into the network from the source vertices up to the time T.

Finally, for a sink $V \in \mathcal{W}$, we define

$$\begin{aligned}\omega^V &:= \sum_{j\,:\,V=\pi_j(1)} \nu_1^j \in \mathcal{M}^+(\{V\} \times [0,\, T]), \\ \omega &:= \sum_{V\in\mathcal{W}} \omega^V \in \mathcal{M}^+(\cup_{\mathcal{W}}\{V\} \times [0,\, T]),\end{aligned} \tag{3.33}$$

which represents the total mass flowing out of the network up to the time T. Equation (3.31) takes then the form

$$\langle \mu,\, \partial_t \varphi + v(x)\partial_x \varphi \rangle = \langle \mu_T - \mu_0,\, \varphi \rangle + \langle \omega - \sigma_0,\, \varphi \rangle, \qquad \forall\, \varphi \in C^1(\Gamma \times [0,\, T]), \tag{3.34}$$

thereby expressing the counterpart of (3.20) on the network.

Using the formulation just obtained, we are in a position to establish the well-posedness of the transport problem over networks.

Theorem 3.4. *Given $\mu_0 \in \mathcal{M}^+(\Gamma \times \{0\})$ and $\sigma_0 \in \mathcal{M}^+(\cup_{\mathcal{S}}\{V\} \times [0,\, T])$, there exists a unique measure $\mu \in \mathcal{M}^+(\Gamma \times [0,\, T])$ which satisfies the balance* (3.34) *with $\mu_T \in \mathcal{M}^+(\Gamma \times \{T\})$ defined in* (3.29)-(3.32) *and $\omega \in \mathcal{M}^+(\cup_{\mathcal{W}}\{V\} \times [0,\, T])$ defined in* (3.30)-(3.33).

Moreover, for $\mu_{0,k} \in \mathcal{M}^+(\Gamma \times \{0\})$, $\sigma_{0,k} \in \mathcal{M}^+(\cup_{\mathcal{S}}\{V\} \times [0,\, T])$, $k = 1,\, 2$, there exists a constant $C = C(T) > 0$ such that

$$\|\mu_{T,2} - \mu_{T,1}\|^*_{BL} + \|\omega_2 - \omega_1\|^*_{BL} \leq C \left(\|\mu_{0,2} - \mu_{0,1}\|^*_{BL} + \|\sigma_{0,2} - \sigma_{0,1}\|^*_{BL} \right). \tag{3.35}$$

Proof. We treat separately the cases in which the set of the source vertices is or is not empty.

(i) Assume $\mathcal{S} \neq \emptyset$. We introduce a partition of the set $\mathcal{E} = \{e_j\}_{j \in J}$ based on the distance from the source set:

$$\begin{aligned} \mathcal{E}_0 &= \{e_j \,:\, V_i = \pi_j(0) \text{ is a source}\} \\ \mathcal{E}_m &= \{e_j \,:\, \exists\, e_k \in \mathcal{E}_{m-1} \text{ s.t. } V = \pi_j(0) = \pi_k(1)\}, \qquad m = 1,\, 2,\, \ldots \end{aligned}$$

We first apply Theorem 4.2 to the problem defined on each arc in $\mathcal{E}_0$, i.e for each $e_j \in \mathcal{E}_0$ such that $V = \pi_j(0) \in \mathcal{S}$, we consider

$$\begin{cases} \partial_t \mu^j + \partial_x (v_j(x)\mu^j) = 0 & \text{in } e_j \times (0,\, T] \\ \mu^j_{t=0} = \mu^j_0 \in \mathcal{M}^+(e_j \times \{0\}) & \\ \nu^j_0 = \sigma^i \in \mathcal{M}^+(\{V\} \times [0,\, T]). & \end{cases}$$

Since ν^j_0 is prescribed, we obtain the existence of $\mu^j \in \mathcal{M}^+(e_j \times [0,\, T])$, $\mu^j_T \in \mathcal{M}^+(e_j \times \{T\})$ and $\nu^j_1 \in \mathcal{M}^+(\{\pi_j(1)\} \times [0,\, T])$ satisfying the balance (3.20). Next we proceed by induction on

$m = 1, 2, \ldots$ considering the problem on $e_j \in \mathcal{E}_m$ with $V = \pi_j(0)$:

$$\begin{cases} \partial_t \mu^j + \partial_x (v_j(x)\mu^j) = 0 & \text{in } e_j \times (0, T] \\ \mu^j_{t=0} = \mu^j_0 \in \mathcal{M}^+(e_j \times \{0\}) \\ \nu^j_0 = \sum_k p_{kj} \cdot \nu^k_1 \in \mathcal{M}^+(\{V\} \times [0, T]). \end{cases}$$

Since the arcs e_k belong to $\mathcal{E}_{m-1}$, the solution to the transport equation on them is known by the inductive step (using the case $m = 0$ as basis), hence the boundary measure ν^j_0 is well defined because so are the outflow measures ν^k_1. Therefore we can apply again Theorem 4.2 to fulfil the balance (3.20) on $e_j \in \mathcal{E}_m$.

In this way, after a finite number of steps we build arc by arc the measures $\mu \in \mathcal{M}^+(\Gamma \times [0, T])$, $\mu_T \in \mathcal{M}^+(\Gamma \times \{T\})$ and $\omega \in \mathcal{M}^+(\cup_{\mathcal{W}}\{V\} \times [0, T])$ which globally satisfy the balance (3.34).

(ii) Assume now $\mathcal{S} = \emptyset$. Fix an arbitrary internal vertex $V \in \mathcal{I}$, and choose

$$t_0 < \min_{j\,:\,V=\pi_j(1)} \tau_j(0).$$

From (3.30) we see that, up to the time t_0, on all the arcs e_j such that $V = \pi_j(1)$ the outflow measure ν^j_1 is given by

$$\nu^j_1 = \int_{(\tau_j^{-1}(t_0),\, 1]} \delta_{\tau_j(x)}\, d\mu^j_0(x),$$

because $\tau_j^{-1}(t_0) > 0$ while $\sigma_j^{-1}(t_0) < 0$ (cf. Figure 3.1, left). Hence ν^j_1 depends only on the initial datum μ^j_0 and not on the inflow measure ν^j_0. Let us consider the initial/boundary-value problem (3.27) for $t \in (0, t_0]$ with V as source vertex and corresponding source measure

$$\nu_0 = \sum_{j\,:\,V_i=\pi_j(1)} \nu^j_1 = \sum_{j\,:\,V_i=\pi_j(1)} \int_{(\tau_j^{-1}(t_0),\, 1]} \delta_{\tau_j(x)}\, d\mu^j_0(x).$$

From the case $\mathcal{S} \neq \emptyset$ we know that we can construct $\mu \in \mathcal{M}^+(e_j \times [0, t_0])$, $\mu_{t_0} \in \mathcal{M}(e_j \times \{t_0\})$ and $\omega \in \mathcal{M}^+(\cup_{\mathcal{W}}\{V\} \times [0, t_0])$ which satisfy the balance (3.20). Moreover, the inflow measures ν^j_0 of all the arcs e_j such that $V = \pi_j(0)$ coincide with those of the original problem without sources, because they are actually determined only by the initial datum. Hence μ is also a solution of the original problem in $[0, t_0]$. By repeating this argument on the intervals $(t_0, 2t_0]$, $(2t_0, 3t_0]$, $\ldots$, with initial data μ_{t_0}, μ_{2t_0}, $\ldots$, after a finite number of steps we obtain the solution of the problem without source in any interval $[0, T]$, $T > 0$.

Finally, the estimate (3.35) is in both cases an immediate consequence of the corresponding esti-

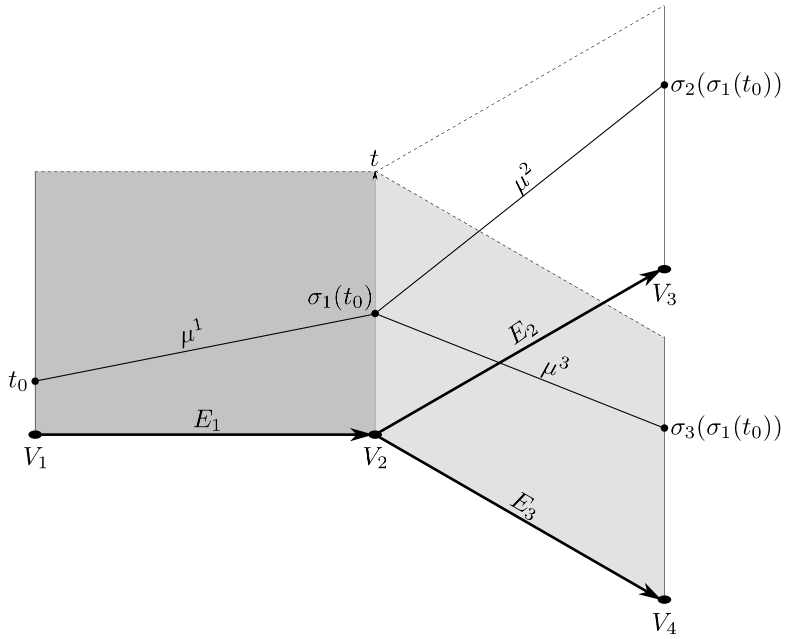


Figure 3.2: The 1-2 junction with a sketch of the characteristics along which the solution propagates.

mate (3.21) holding on each arc. □

3.2.3 Examples on junctions

In this section we write explicitly the solution to problem (3.27) for two typical junctions which occur frequently for instance in traffic flow on road networks. It is worth pointing out that, since in our linear equation the velocity depends only on the space variable but not on the measure μ itself, the transport model that we are considering may provide an acceptable description of the flow of vehicles at most in the so-called *free flow regime*. In fact, in such a case the number of vehicles is sufficiently small that their speed is almost independent of the presence of other vehicles on the road.

The 1-2 junction – Atomic inflow distribution. Let Γ be the road network shown in Figure 3.2 formed by 3 arcs, viz. roads, E_1, E_2, E_3 and 4 vertices $V_1, \dots, V_4$ such that E_1 connects the source vertex V_1 to the internal vertex V_2 while E_2 and E_3 connect the internal vertex V_2 to the well vertices V_3 and V_4. This gives also the orientation of the arcs. In practice, beyond the junction V_2 the road E_1 splits in the two roads E_2, E_3. We assume that the network is initially empty. At some time $t_0 > 0$ a *microscopic* vehicle enters the network from the vertex V_1 and then travels across it. At the junction V_2 we prescribe a flux distribution rule stating that a time-dependent fraction $p = p(t) : [0, T] \to [0, 1]$ of the incoming mass flows to the road E_2 while the complementary fraction $1 - p(t)$ flows to the road

E_3. Taking $T = +\infty$, the problem can be formalised as:

$$\begin{cases} \partial_t \mu^j + \partial_x (v_j(x)\mu^j) = 0 & x \in e_j,\, t \in \mathbb{R}^+,\, j = 1,\, 2,\, 3 \\ \mu_0 = 0 & x \in \Gamma \\ \nu_0^1 = \delta_{t_0} & t \in \mathbb{R}_0^+ \\ \nu_0^2 = p(t) \cdot \nu_1^1 & t \in \mathbb{R}_0^+ \\ \nu_0^3 = (1 - p(t)) \cdot \nu_1^1 & t \in \mathbb{R}_0^+, \end{cases}$$

where the velocity fields $v_j : e_j \to (0,\, v_{\max}^j]$, $0 < v_{\max}^j < +\infty$, are given Lipschitz continuous functions of x.

The solution on each road has the form $\mu^j(dx\, dt) = \mu_t^j(dx) \otimes dt$, where μ_t^j is the trace of μ^j on the fibre $e_j \times \{t\}$. Using (3.29), (3.30) we determine explicitly the expression of μ_t^j for all $t > 0$ and that of the outflow masses ν_1^j on the fibres $\{\pi_j(1)\} \times \mathbb{R}_0^+$ (notice that $\pi_1(1) = V_2$, $\pi_2(1) = V_3$, $\pi_3(1) = V_4$). We find (cf. Figure 3.2):

$$\begin{aligned} &\mu_t^1 = \delta_{\Phi_t^1(0,\, t_0)} \chi_{[t_0,\, \sigma_1(t_0)]}(t) \\ &\nu_1^1 = \delta_{\sigma_1(t_0)} \\ &\mu_t^2 = p(\sigma_1(t_0)) \delta_{\Phi_t^2(0,\, \sigma_1(t_0))} \chi_{[\sigma_1(t_0),\, \sigma_2(\sigma_1(t_0))]}(t) \\ &\nu_1^2 = \omega^3 = p(\sigma_1(t_0)) \delta_{\sigma_2(\sigma_1(t_0))} \\ &\mu_t^3 = [1 - p(\sigma_1(t_0))] \delta_{\Phi_t^3(0,\, \sigma_1(t_0))} \chi_{[\sigma_1(t_0),\, \sigma_3(\sigma_1(t_0))]}(t) \\ &\nu_1^3 = \omega^4 = [1 - p(\sigma_1(t_0))] \delta_{\sigma_3(\sigma_1(t_0))}. \end{aligned}$$

Furthermore, using Bochner integrals in the product space $e_j \times \mathbb{R}_0^+$ we can possibly write the solution μ^j on each road as

$$\begin{aligned} &\mu^1 = \int_{t_0}^{\sigma_1(t_0)} \delta_{(\Phi_t^1(0,\, t_0),\, t)}\, dt \\ &\mu^2 = p(\sigma_1(t_0)) \int_{\sigma_1(t_0)}^{\sigma_2(\sigma_1(t_0))} \delta_{(\Phi_t^2(0,\, \sigma_1(t_0)),\, t)}\, dt \\ &\mu^3 = [1 - p(\sigma_1(t_0))] \int_{\sigma_1(t_0)}^{\sigma_3(\sigma_1(t_0))} \delta_{(\Phi_t^3(0,\, \sigma_1(t_0)),\, t)}\, dt. \end{aligned}$$

Remark 3.4.1. *By carefully inspecting the expressions of μ_t^j, $j = 1,\, 2,\, 3$, we see that the unit-mass Dirac delta prescribed at the source vertex V_1 splits in two Dirac deltas beyond the junction V_2, cf. also Figure 3.2, each of which carries a fraction, $p(\sigma_1(t_0))$ and $1 - p(\sigma_1(t_0))$, respectively, of the initial* *mass.*

Unlike the Dirac delta entering the road E_1 from V_1, the two Dirac deltas propagating in the roads E_2, E_3 do not *represent physical microscopic vehicles. Rather, each of them is the same microscopic vehicle coming from the road E_1 and the coefficients $p(\sigma_1(t_0))$, $1-p(\sigma_1(t_0))$ have to be understood as the* probabilities *that such a vehicle takes either outgoing road beyond the junction V_2.*

This approach differs from the one proposed in [25], which instead assigns a path to each microscopic vehicle through the network in the spirit of the multipath traffic model introduced in [11, 12].

The 1-2 junction – Continuous inflow distribution. We now consider the same network as in the previous example but we prescribe an inflow measure ν_0^1 which is absolutely continuous with respect to the Lebesgue measure:

$$\nu_0^1(dt) := \rho(t)\,dt,$$

where $\rho \in L^1(\mathbb{R}_0^+)$ with $\operatorname{supp}\rho \subseteq \mathbb{R}_0^+$ is the density of the vehicles entering the network from the vertex V_1.

Recalling that the network is initially empty and using (3.29), we obtain that for each $t > 0$ the trace μ_t^1 of the solution μ^1 in the road E_1 is

$$\mu_t^1 = \int_{\max\{0,\,\sigma_1^{-1}(t)\}}^{t} \delta_{\Phi_t^1(0,\,s)}\rho(s)\,ds = \int_0^{t-\max\{0,\,\sigma_1^{-1}(t)\}} \delta_{\Phi_r^1(0,\,0)}\rho(t-r)\,dr,$$

where in the last passage we have set $r := t - s$ after observing from (3.12) that $\Phi_t^1(0,\,s) = \Phi_{t-s}^1(0,\,0)$ for all $0 \leq s \leq t$. Likewise, recalling (3.30) we find that the outflow mass ν_1^1 at the vertex V_2 is

$$\nu_1^1 = \int_0^{+\infty} \delta_{\sigma_1(s)}\rho(s)\,ds = \int_{\sigma_1(0)}^{+\infty} \delta_r\rho(r-\sigma_1(0))\,dr,$$

where in the second passage we have set $r := \sigma_1(s) = s + \sigma_1(0)$. In view of the Bochner representation (3.1) and considering that $\operatorname{supp}\rho(\cdot - \sigma_1(0)) \subseteq [\sigma_1(0),\,+\infty)$, we deduce in particular

$$\nu_1^1(dt) = \rho(t-\sigma_1(0))\,dt.$$

According to our transmission conditions, this mass is distributed to the outgoing roads E_2, E_3 as

$$\nu_0^2 = p(t)\nu_1^1, \qquad \nu_0^3 = (1-p(t))\nu_1^1,$$

which, owing to (3.29), implies that the traces μ_t^2, μ_t^3 of the solutions μ^2, μ^3 in the outgoing roads are respectively given by

$$
\begin{aligned}
\mu_t^2 &= \int_{\max\{0,\,\sigma_2^{-1}(t)\}}^{t} \delta_{\Phi_t^2(0,\,s)} p(s)\rho(s-\sigma_1(0))\,ds \\
&= \int_0^{t-\max\{0,\,\sigma_2^{-1}(t)\}} \delta_{\Phi_r^2(0,\,0)} p(t-r)\rho(t-r-\sigma_1(0))\,dr
\end{aligned}
$$

and by

$$
\begin{aligned}
\mu_t^3 &= \int_{\max\{0,\,\sigma_3^{-1}(t)\}}^{t} \delta_{\Phi_t^3(0,\,s)} (1-p(s))\rho(s-\sigma_1(0))\,ds \\
&= \int_0^{t-\max\{0,\,\sigma_3^{-1}(t)\}} \delta_{\Phi_r^3(0,\,0)} (1-p(t-r))\rho(t-r-\sigma_1(0))\,dr.
\end{aligned}
$$

It is interesting to note that, since in general the density ρ is split asymmetrically in the roads E_2 and E_3 (unless $p(t) = \frac{1}{2}$), the corresponding measure solution, even if possibly continuous inside the arcs of the network, is discontinuous across the vertex V_2.

Finally, the outflow masses $\nu_1^2 = \omega^3$ and $\nu_1^3 = \omega^4$ are recovered from (3.30) as

$$
\begin{aligned}
\nu_1^2 = \omega^3 &= \int_0^{+\infty} \delta_{\sigma_2(s)} p(s)\rho(s-\sigma_1(0))\,ds \\
&= \int_{\sigma_2(0)}^{+\infty} \delta_r p(r-\sigma_2(0))\rho(r-\sigma_1(0)-\sigma_2(0))\,dr
\end{aligned}
$$

and

$$
\begin{aligned}
\nu_1^3 = \omega^4 &= \int_0^{+\infty} \delta_{\sigma_3(s)} (1-p(s))\rho(s-\sigma_1(0))\,ds \\
&= \int_{\sigma_3(0)}^{+\infty} \delta_r (1-p(r-\sigma_3(0)))\rho(r-\sigma_1(0)-\sigma_3(0))\,dr.
\end{aligned}
$$

Observing that $\operatorname{supp}\rho(\cdot - \sigma_1(0) - \sigma_j(0)) \subseteq [\sigma_1(0) + \sigma_j(0), +\infty)$ for $j = 2, 3$, from the Bochner representation (3.1) of a measure we further deduce

$$
\begin{aligned}
\nu_1^2(dt) &= \omega^3(dt) = p(t-\sigma_2(0))\rho(t-\sigma_1(0)-\sigma_2(0))\,dt \\
\nu_1^3(dt) &= \omega^4(dt) = (1-p(t-\sigma_3(0))\rho(t-\sigma_1(0)-\sigma_3(0))\,dt.
\end{aligned}
$$

Remark 3.4.2. *The transport problem being linear, the case of an inflow measure ν_0^1 carrying both an atomic and a Lebesgue-absolutely continuous part can be addressed by simply superimposing the solutions obtained in the previous examples.*

The 2-1 junction. We consider now the road network Γ illustrated in Figure 3.3 with again 3 arcs,

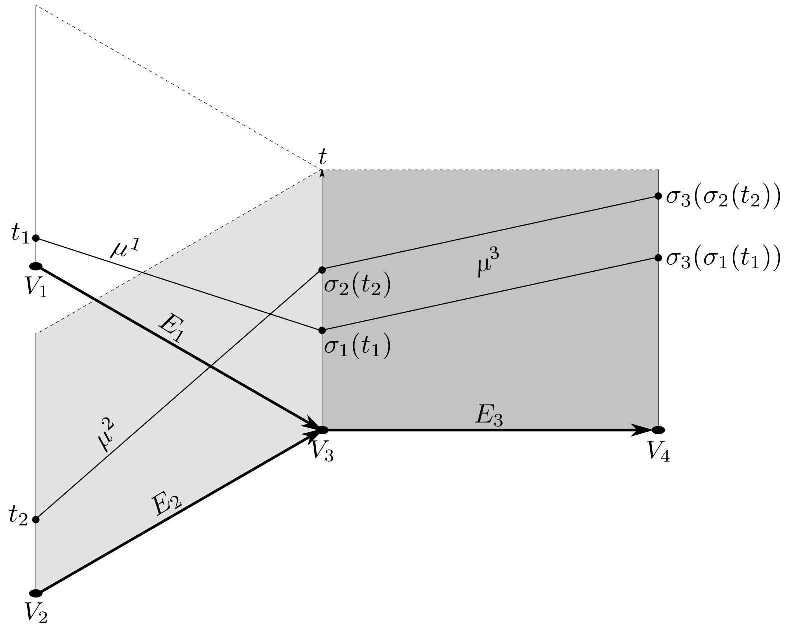


Figure 3.3: The 2-1 junction with a sketch of the characteristics along which the solution propagates in the space-time of the network.

viz. roads, E_1, E_2, E_3 and 4 vertices $V_1, \ldots, V_4$. However, in this case both vertices V_1, V_2 are sources and are connected by roads E_1, E_2 to the internal vertex V_3. The latter is finally connected to the well vertex V_4 by road E_3. In practice, beyond the junction V_3 the incoming roads E_1, E_2 merge into the outgoing road E_3.

Like in the previous examples, we assume that the network is initially empty. At two successive time instants $0 \leq t_1 \leq t_2$ two *microscopic* vehicles enter the network from the sources V_1, V_2, respectively. Their propagation across the network for $t > 0$ is then described by the problem:

$$\begin{cases} \partial_t \mu^j + \partial_x(v_j(x)\mu^j) = 0 & x \in e_j,\, t \in \mathbb{R}^+,\, j = 1, 2, 3 \\ \mu_0 = 0 & x \in \Gamma \\ \nu_0^1 = \delta_{t_1} & t \in \mathbb{R}_0^+ \\ \nu_0^2 = \delta_{t_2} & t \in \mathbb{R}_0^+ \\ \nu_0^3 = \nu_1^1 + \nu_1^2 & t \in \mathbb{R}_0^+, \end{cases}$$

where the velocity fields $v_j : e_j \to (0, v_{\max}^j]$, $0 < v_{\max}^j < +\infty$, are as usual given Lipschitz continuous functions of x. Notice that, for mass conservation purposes, the flux distribution coefficients at the junction V_3 are necessarily $p_{13}^3(t) = p_{23}^3(t) = 1$ for all $t > 0$.

Relying again on (3.29), (3.30) we write explicitly the solution $\mu^j \in \mathcal{M}^+(e_j \times \mathbb{R}_0^+)$ on each road as well as the outflow measures $\nu_1^j \in \mathcal{M}^+(\{\pi_j(1)\} \times \mathbb{R}_0^+)$, with $\pi_1(1) = \pi_2(1) = V_3$ and $\pi_3(1) = V_4$. We find (cf. Figure 3.3):

$$
\begin{aligned}
\mu_t^1 &= \delta_{\Phi_t^1(0,\,t_1)}\chi_{[t_1,\,\sigma_1(t_1)]}(t), & \nu_1^1 &= \delta_{\sigma_1(t_1)}\\
\mu_t^2 &= \delta_{\Phi_t^2(0,\,t_2)}\chi_{[t_2,\,\sigma_2(t_2)]}(t), & \nu_1^2 &= \delta_{\sigma_2(t_2)}\\
\mu_t^3 &= \delta_{\Phi_t^3(0,\,\sigma_1(t_1))}\chi_{[\sigma_1(t_1),\,\sigma_3(\sigma_1(t_1))]}(t) & \nu_1^3 &= \omega^4 = \delta_{\sigma_3(\sigma_1(t_1))} + \delta_{\sigma_3(\sigma_2(t_2))},\\
&\quad + \delta_{\Phi_t^3(0,\,\sigma_2(t_2))}\chi_{[\sigma_2(t_2),\,\sigma_3(\sigma_2(t_2))]}(t),
\end{aligned}
$$

whence, using Bochner integrals in the product spaces $e_j \times \mathbb{R}_0^+$, $j = 1, 2, 3$,

$$
\begin{aligned}
\mu^1 &= \int_{t_1}^{\sigma_1(t_1)} \delta_{(\Phi_t^1(0,\,t_1),\,t)}\,dt\\
\mu^2 &= \int_{t_2}^{\sigma_2(t_2)} \delta_{(\Phi_t^2(0,\,t_2),\,t)}\,dt\\
\mu^3 &= \int_{\sigma_1(t_1)}^{\sigma_3(\sigma_1(t_1))} \delta_{(\Phi_t^3(0,\,\sigma_1(t_1)),\,t)}\,dt + \int_{\sigma_2(t_2)}^{\sigma_3(\sigma_2(t_2))} \delta_{(\Phi_t^3(0,\,\sigma_2(t_2)),\,t)}\,dt.
\end{aligned}
$$

Chapter 4

Nonlocal interactions on Networks

Aim of this chapter is to study a nonlinear transport equation on an oriented network where the velocity field depends not only on the state variable, but also on the solution itself. We also provide a representation formula in terms of the push-forward of the initial and boundary data along the network and discuss an example of nonlocal velocity field fitting our framework.
We also refer to [5, 17, 24, 36, 48] for applications of the measure-theoretic approach to the study of various complex phenomena. Short and long range interaction mechanisms are efficiently taken into account by a velocity field depending on local terms, determined by the geometry of the space, and nonlocal terms, depending on the whole support of the measure solution or on a part of it; aggregation phenomena, leading in the classical setting to blow-up of the solution, are plainly taken into account by the measure setting.

We consider the nonlinear transport equation

$$\partial_t \mu + \partial_x (v[\mu_t]\mu) = 0, \qquad \text{in } \Gamma \times [0, T], \tag{4.1}$$

where the velocity v still depends on the state variable, but also on the distribution μ_t at time t. In this case, the evolution equation does not only depends on the portion of mass inflowing/outflowing from each arc, but also on the global distribution μ_t at time t.
This assumption adds several difficulties which have to be studied: to show the well posedness of (4.1), we approximate the nonlinear transport equation by a sequence of linear problems obtained via semi-discrete in time approximation of (4.1). We define a partition of the time interval $[0, T]$ in a family of subinterval $[t_k, t_k + \Delta t]$ and on each of these intervals we solve the linear problem (3.4) with the nonlinear velocity $v[\mu_t]$ replaced by the linear one $v[\mu_{t_k}]$. In such a way we obtain a sequence of measure $\{\mu^{\Delta t}\}$ defined on $[0, T]$. Using the results on the linear problem, we prove that for $\Delta t \to 0^+$,

the sequence $\{\mu^{\Delta t}\}$ converges (upon to subsequence) to a measure in $\mu \in \mathcal{M}^+(\Gamma \times [0,T])$ which is a solution of (4.1). A continuous dependence result and a representation formula in terms of the push-forward of the initial and boundary data along the admissible paths on the network complete the study of (4.1).

In the model discussed in the previous chapter, the sources represent the vertices where agents enter in the network, while the sinks the vertices where they exit from the network. Since the velocity term may depend on the distribution of the agents on all the network, in order to simplify the notations we prefer to consider a network without sinks, i.e. the terminal arcs always have infinite length. In any case, at the expense of an heavier notation, is not difficult to include in the model also the contribution of the sinks.

4.1 The nonlinear transport problem

This section is devoted to the study of the nonlinear transport problem

$$\begin{cases} \partial_t \mu + \partial_x (v[\mu_t]\mu) = 0, & \text{on } \Gamma \times [0,T], \\ \mu_{t=0} = \mu_0, & \\ \mu_{x=x_i} = \sigma_0^i, & \forall x_i \in \mathcal{S}, \\ \mu^j_{x=x_i} = \sum_{k \in \mathrm{Inc}(x_i)} p_{kj} \cdot \mu^k_{x=x_i}, & \forall e_j \in \mathrm{Out}(x_i), \, \forall x_i \in \mathcal{V} \setminus \mathcal{S}, \end{cases} \tag{4.2}$$

with $v[\mu]$, μ_0, σ_0 satisfying the assumptions in the previous chapter. In particular, thanks to the integration by part formulas, with $v(x)$ replaced by $v[\mu](x)$, we can state

Definition 4.1. *A measure-valued solution to* (4.2) *is a finite measure $m \in \mathcal{M}^+(\Gamma \times [0,T])$ such that for every $f \in C^1_0(\Gamma \times [0,T])$,*

$$\langle \mu_{t=T} - \mu_0,\, f \rangle - \langle \sigma_0,\, f \rangle = \langle \mu,\, \partial_t f + v[m_t]\partial_x f \rangle, \tag{4.3}$$

and $\forall\ x_i \in \mathcal{V} \setminus \mathcal{S}$, $\forall\ e_j \in Out(x_i)$,

$$\langle \mu^j_{x=x_i},\, f \rangle = \sum_{e_k \in Inc(x_i)} \langle \mu^k_{x=x_i},\, p_{kj} f \rangle. \tag{4.4}$$

We assume a *nonlinear velocity field* $v : \mathcal{M}^+(\Gamma) \times \Gamma \to \mathbb{R}$ with the following properties

$(H1)$ v is nonnegative and bounded by a positive constant V_{max};

$(H2)$ v is Lipschitz continuous with respect to the state variable, i.e. on each arc $e_j \in \mathcal{E}$

$$|v[\mu](x) - v[\mu](y)| \leq L|x - y|, \quad \forall \mu \in \mathcal{M}^+(\Gamma),\, x, y \in e_j;$$

$(H3)$ v is Lipschitz continuous with respect to the measure variable, i.e.

$$|v[\mu_1](x) - v[\mu_2](x)| \leq L\|\mu_1 - \mu_2\|^*_{BL} \quad \forall x \in \Gamma,\, \mu_1, \mu_2 \in \mathcal{M}^+(\Gamma).$$

When considered on a single arc isomorphic to $\mathbb{R}$, the previous assumptions coincide with the ones for the corresponding nonlinear transport model in [24], while, for a fixed $\mu \in \mathcal{M}^+(\Gamma)$, the velocity field $v[\mu]$ satisfies the hypotheses of linear transport problem considered in the previous chapter.
We conclude this section with a notion of p-moment for finite measures on networks. Even if it is a straightforward generalization of the corresponding concept in the Euclidean space, we give some details for reader's convenience.

Definition 4.2. *Fixed $p \in \mathbb{N}$ and $x \in \Gamma$, the $p-$moment centered at x of a finite measure $\mu \in \mathcal{M}^+(\Gamma)$ is defined by*

$$\langle \mu,\, d_\Gamma(\cdot, x)^p \rangle = \int_\Gamma d_\Gamma(y, x)^p d\mu(y). \tag{4.5}$$

Lemma 4.1. *A finite measure $\mu \in \mathcal{M}^+(\Gamma)$ has finite $p-$moment iff it has finite $p-$moment on every arc $e_j \in \mathcal{E}$ such that the length $\ell(e_j)$ is infinite.*

Proof. Assume w.l.o.g. that $x = x_i \in \mathcal{V}$ and set $d(\cdot) = d_\Gamma(\cdot, x_i)$. Given a measure $\mu \in \mathcal{M}^+(\Gamma)$, $\mu = \sum_{j\in J} \mu^j$ with $\text{supp}\{\mu^j\} \subseteq e_j$, we can write

$$\langle \mu; d^p \rangle = \sum_{\substack{j\in J \\ \ell(e_j)<+\infty}} \langle \mu^j; d^p \rangle + \sum_{\substack{j\in J \\ \ell(e_j)=+\infty}} \langle \mu^j; d^p \rangle.$$

If $e_j \in \mathcal{E}$ has finite length, then $d(\cdot)$ has its maximum value $\overline{d}_j$ on e_j . Then defined

$$\overline{d} := \max_{\substack{j\in J \\ \ell(e_j)<+\infty}} \overline{d}_j,$$

we have

$$\sum_{\substack{j\in J \\ \ell(e_j)<+\infty}} \langle \mu^j; d^p \rangle \leq \sum_{\substack{j\in J \\ \ell(e)<+\infty}} \overline{d}_j \cdot \mu^j(e_j) \leq \overline{d} \cdot \mu(\Gamma).$$

On the other side, if $\ell(e_j) = +\infty$ and $e_j = \pi_j([0, +\infty))$ with $x_k = \pi_j(0) \in \mathcal{V}$, by Jensen's inequality we have

$$\langle \mu^j; d^p\rangle = \int_{[0,+\infty)} (|y| + d(x_k))^p d\mu^j(y) \le 2^{p-1} \int_{[0,+\infty)} |y|^p d\mu^j(y) + 2^{p-1} d(x_k)^p \mu(e_j).$$

By the last inequality, the statement easily follows. □

The finite p-moment property for a measure μ is clearly independent of the point $x \in \Gamma$ chosen in the Definition 4.2.

To prove the core result of this chapter, i.e. existence of a measure-valued solution to (4.2), we introduce a semi-discretization in time procedure which allows to approximate the nonlinear problem by a family of linear problem. Fixed $N \in \mathbb{N}$, set $\Delta t^N := T/2^N$ and define a partition of $[0,T]$ by the intervals $I_n^N := [t_n^N; t_{n+1}^N]$ where $t_n^N := n\Delta t^N, n = 0, \dots, 2^N$ (in the following we write t_n in place of t_n^N when it is clear by the context). We consider the 2^N problems iteratively defined, $\forall n = 0, \dots, 2^N - 1$, by

$$\begin{cases} \partial_t \mu + \partial_x (v[\mu_{t_n}]\mu) = 0, & \text{on } \Gamma \times I_n^N \\ \mu_{t=t_n} = \mu_{t_n}, \\ \mu_{x_i \in \mathcal{S}} = \sigma_0 \llcorner_{I_n^N}, \\ \mu^j_{x=x_i} = \sum_{k \in \mathrm{Inc}(x_i)} p_{kj} \cdot \mu^k_{x=x_i}, & \forall e_j \in \mathrm{Out}(x_i), \forall x_i \in \mathcal{V} \setminus \mathcal{S}, \end{cases}$$

where $\sigma_0 \llcorner_{I_n^N}$ is the restriction of σ_0 to the interval I_n^N. We remark that on $\Gamma \times I_n^N$ the velocity term $v[\mu_{t_n}]$ is linear. Therefore, thanks to Theorem 3.4, there exists a unique solution $\mu^{N,n} \in \mathcal{M}^+(\Gamma \times I_n^N)$ which satisfies the balance equation

$$\langle \mu^{N,n}_{t=t^N_{n+1}} - \mu^{N,n}_{t=t^N_n}, f\rangle = \langle \sigma_0 \llcorner_{I_n^N}, f\rangle + \int_{t_n^N}^{t_{n+1}^N} \langle \mu_t^{N,n}, (\partial_t + v[\mu^{N,n}_{t_n^N}]\partial_x) \cdot f(\cdot, t)\rangle dt, \tag{4.6}$$

and the transition condition

$$\langle (\mu^{N,n})^j_{x=x_i} \llcorner_{I_n^N}, f\rangle = \sum_{k \in \mathrm{Inc}(x_i)} \langle p_{kj} (\mu^{n,N})^k_{x=x_i} \llcorner_{I_n^N}, f\rangle, \quad \forall e_j \in \mathrm{Out}(x_i), \forall x_i \in \mathcal{V} \setminus \mathcal{S}. \tag{4.7}$$

for every $f \in C_0^\infty(\Gamma \times \overline{I}_n^N)$. We denote by $\mu^N : [0,T] \to \mathcal{M}^+(\Gamma)$ the map defined by

$$\mu_t^N = \mu_t^{N,n} \qquad \text{for } t \in I_n^N, \ n = 0, \dots, 2^N - 1. \tag{4.8}$$

We first give some regularity properties of the map μ^N.

Proposition 4.1. *For any $t \in [0,T]$, the measure μ_t^N is bounded in $(\mathcal{M}^+(\Gamma), \|\cdot\|^*_{BL})$, uniformly in*

N, *i.e. there exists a positive constant $C = C(T)$ such that*

$$\|\mu_t^N\|_{BL}^* \leq C(\|\mu_0\|_{BL}^* + \|\sigma_0\|_{BL}^*), \qquad \forall t \in [0,T],\ N \in \mathbb{N}. \tag{4.9}$$

Moreover, there exists a positive constant $C = C(T)$ such that

$$\|\mu_t^N - \mu_s^N\|_{BL}^* \leq \sigma_0([s,t]) + C|t-s|, \qquad \forall\, 0 \leq s < t \leq T,\ N \in \mathbb{N}. \tag{4.10}$$

Proposition 4.2. *Assume that μ_0 has finite $p-$moment over Γ, $p = 1, 2$. Then, for any $t \in [0,T]$, the measure μ_t^N has finite $p-$moment, $p = 1, 2$, over Γ, uniformly in N, i.e. there exists a positive constant $C = C(T)$ such that*

$$\langle \mu_t^N,\, d^p \rangle \leq C, \qquad \forall t \in [0,T],\ N \in \mathbb{N}.$$

Proof of Proposition 4.1. Let $t \in [0,T]$ and $n \in \{0, \ldots, 2^N - 1\}$ such that $t \in I_n^N$. Then, by the representation formula (4.31), we write

$$\langle \mu_t^N,\, f \rangle = \int_\Gamma \sum_{\gamma \in \mathcal{A}(x)} f(\Phi_t^\gamma(x, t_n), t) p_\gamma(x, 0) d\mu_{t_n}^N(x) + \sum_{x_i \in \mathcal{S}} \int_{[0,T]} \sum_{\gamma \in \mathcal{A}(x_i)} f(\Phi_t^\gamma(x_i, s), t) p_\gamma(x_i, 0) d\sigma_0^i(s).$$

Hence, for every $f \in BL(\Gamma \times [0,T])$ such that $\|f\|_{BL} \leq 1$, it follows

$$\begin{aligned} |\langle \mu_t^N,\, f \rangle| &\leq \int_\Gamma \|f\|_{BL} \left(\sum_{\gamma \in \mathcal{A}(x)} p_\gamma(x, 0) \right) d\mu_{t_n}^N(x) + \sum_{x_i \in \mathcal{S}} \int_{[0,T]} \|f\|_{BL} \left(\sum_{\gamma \in \mathcal{A}(x_i)} p_\gamma(x_i, 0) \right) d\sigma_0^i(s) \\ &\leq \|\mu_{t_n}^N\|_{BL}^* + \sum_{x_i \in \mathcal{S}} \|(\sigma_0^i)\llcorner_{[t_n,t]}\|_{BL}^* = \|\mu_{t_n}^N\|_{BL}^* + \|(\sigma_0)\llcorner_{[t_n,t]}\|_{BL}^*, \end{aligned}$$

where we have used the property $\sum_{\gamma \in \mathcal{A}(x)} p_\gamma(x, 0) = 1$, for all $x \in \Gamma$. Taking the supremum over $f \in BL(\Gamma \times [0,T])$, we get

$$\|\mu_t^N\|_{BL}^* \leq \|\mu_{t_n}\|_{BL}^* + \|(\sigma_0)\llcorner_{(t_n,t]}\|_{BL}^*;$$

Applying the previous inequality recursively for $\mu \in \{0, \ldots, n\}$, we get (4.9).

We now prove (4.10). Fixed $N \in \mathbb{N}$, let $s, t \in [0,T]$ such that $s < t$ with $s \in I_n^N, t \in I_k^N$ for i $n, k \in \{0, \ldots, 2^N - 1\}$, $n \neq k$. This means that

$$t_n \leq s < t_{n+1} < \ldots < t_k \leq t \leq t_{k+1}.$$

We can clearly write

$$\mu_t^N - \mu_s^N = \left(\mu_t^N - \mu_{t_k}^N\right) + \left(\mu_{t_{n+1}}^N - \mu_s^N\right) + \sum_{l=n+1}^{k} \left(\mu_{t_{l+1}}^N - \mu_{t_l}^N\right).$$

This implies

$$\|\mu_t^N - \mu_s^N\|_{BL}^* \le \|\mu_t^N - \mu_{t_k}^N\|_{BL}^* + \|\mu_{t_{n+1}}^N - \mu_s^N\|_{BL}^* + \sum_{l=n+1}^{k} \|\mu_{t_{l+1}}^N - \mu_{t_l}^N\|_{BL}^*. \tag{4.11}$$

We need to estimate $\|\mu_{t_{l+1}}^N - \mu_{t_l}^N\|_{BL}^*$. Let $f \in BL(\Gamma \times [0,T])$ such that $\|f\|_{BL}^* \le 1$. Then, for every $t \in I_n^N$,

$$\begin{aligned}
&|\langle \mu_t^N - \mu_{t_n}^N, f\rangle| \\
&\quad \le \int_\Gamma \sum_{\gamma \in \mathcal{A}(x)} |f(\Phi_t^\gamma(x,t_n),t) - f(x,t_n)| d\mu_{t_n}^N(x) + \Big|\sum_{x_i \in \mathcal{S}} \int_{(t_n,t]} f(\Phi_t(x_i,s),t) d\sigma_0^i(s)\Big| \\
&\quad \le \int_\Gamma \left(\sum_{\gamma \in \mathcal{A}(x)} p_\gamma(x,0) \right) \left(d(\Phi_t^\gamma(x,t_n),x) + |t-t_n|\right) d\mu_{t_n}^N(x) + \|(\sigma_0)\llcorner_{(t_n,t]}\|_{BL}^*.
\end{aligned} \tag{4.12}$$

By definition of Φ^γ, it follows $d(\Phi_t^\gamma(x,t_n),x) \le \int_{t_n}^t v_n^N(\Phi_s^\gamma(x,t_n))ds \le |t-t_n|V_{max}$. Then, applying (4.9) and taking the supremum over $f \in BL(\Gamma \times [0,T])$ such that $\|f\|_{BL} \le 1$, we can write

$$\|\mu_t^N - \mu_{t_n}^N\|_{BL}^* \le C|t-t_n| + \sigma_0([t_n,t]), \tag{4.13}$$

where $C = (1+V_{max})(\|m_0\|_{BL}^* + \|\sigma_0\|_{BL}^*) > 0$. Using (4.13) and (4.12) in (4.11), we get (4.10).

□

Proof of Proposition 4.2. For $x_i \in \mathcal{V}$ fixed, we set $d(\cdot) := d_\Gamma(\cdot, x_i)$. Fixed $N \in \mathbb{N}$, we denote $v[\mu_{t_n^N}^N]$ with v_n^N and we consider $t \in [0,T]$ and $n \in \{0,\dots,2^N-1\}$ such that $t \in I_{n-1}^N$. By Lemma 4.1, $\mu_t^N \in \mathcal{M}^+(\Gamma)$ has finite $p-$moment over Γ iff it has finite $p-$moment on every arc $e_j \in \mathcal{E}$ such that $\ell(e_j) = +\infty$.

First consider the case $p=1$. If $e_j \in \mathcal{E}$ is such that $\ell(e_j) = +\infty$, there are two possibilities

i) $\exists x_i \in \mathcal{V}$ such that $e_j \in Inc(x_i)$;

ii) $\exists x_i \in \mathcal{V}$ such $e_j \in Out(x_i)$.

If (i) occurs, we parametrise $e_j \in \mathcal{E}$ as $(-\infty;0]$. For every $t \in I_{n-1}^N$, we denote with $\Phi_t^{e_j}$ the flow over

e_j with respect the velocity v_{n-1}^N. By the definition in (3.17), we have

$$\tau_{n,j}(x) = \inf\{t \geq t_{n-1}^N : \Phi_t^{e_j}(x, t_{n-1}^N) = \pi_j(0)\}.$$

Then, the first moment over e_j of μ_t^N can be estimated by

$$\begin{aligned}\int_{(-\infty;0]} |x| d\mu_t^{N,j}(x) &= \int_{(-\infty;\tau_{n,j}^{-1}(x)]} |\Phi_t^{e_j}(x, t_{n-1})| d\mu_{t_{n-1}}^{N,j}(x)\\ &\leq \int_{(-\infty;\tau_{n,j}^{-1}(x)]} |x| d\mu_{t_{n-1}}^{N,j}(x) + \mu_{t_{n-1}}^N((-\infty;\tau_{n,j}^{-1}(x)])V_{max}\Delta t^N\\ &\leq \int_{(-\infty;0]} |x| d\mu_{t_{n-1}}^{N,j} + \mu_{t_{n-1}}^{N,j}(e_j)\Delta t^N V_{max}.\end{aligned}$$

Applying iteratively the previous argument for $k \in \{0, 1, \ldots, n-1\}$, we get

$$\begin{aligned}\int_e |x| d\mu_t^N(x) &\leq \int_e |x| d\mu_0^j(x) + V_{max}\Delta t^N \sum_{k=0}^{n-1} \mu_{t_k}^N(e_j)\\ &\leq \int_e |x| d\mu_0^j(x) + V_{max}\frac{T}{2^N} \sum_{k=0}^{2^N-1} \mu_{t_k}^N(e_j).\end{aligned}$$

By Lemma 4.1 we have

$$\int_e |x| d\mu_t^N(x) \leq \int_e |x| d\mu_0^j(x) + V_{max}TC. \tag{4.14}$$

For the measure $\mu_{x=x_i}^{N,j} \in \mathcal{M}^+([0,T])$, projection of $\mu^{N,j}$ at x_i, by (3.19) we estimate

$$\|\mu_{x=x_i}^{N,j}\|_{BL}^* = \mu_{x=x_i}^{N,j}([0,T]) \leq \mu_0^j((-V_{max}T, 0]) \leq \mu_0^j(e_j). \tag{4.15}$$

If (ii) occurs, we have a similar proof. Indeed, thanks to the characterization (3.19), we can write

$$\int_{[0,+\infty)} |x| d\mu_t^{N,j}(x) = \int_{[0,+\infty)} |\Phi_t^{e_j}(x, t_{j-1})| d\mu_{t_{n-1}}^{N,j}(x) + \int_{[\varsigma_j^{-1}(t),t]} |\Phi_t^{e_j}(0, s)| d\mu_{x=x_i}^{N,j}(s).$$

The first integral on the right side can be estimated as in (4.14), while for the second one we have

$$\begin{aligned}\int_{[\sigma_e^{-1}(t),t]} |\Phi_t^e(0, s)| d\mu_{x=x_i}^j(s) &\leq V_{max} \int_{[\sigma_e^{-1}(t),t]} |t-s| d\mu_{x=x_i}^j(s)\\ &\leq V_{max}\Delta t^N \mu_{x=x_i}^j(I_{n-1}^N),\end{aligned}$$

which is finite and bounded by a constant which only depends on T, thanks to (4.15) and Theorem 3.4.

To conclude the proof, we need to show an analogous statement for $p = 2$. However, we can observe that

$$\begin{aligned}|\Phi_t^{e_j}(x, t_{n-1})|^2 &= |x|^2 + |\int_{t_{j-1}}^{t} v_{n-1}^N(\Phi_s^{e}(x, t_{n-1}))|^2 + 2|x| \left|\int_{t_{n-1}}^{t} v_{n-1}^N(\Phi_s^{e_j}(x, t_{n-1}))\right| \\ &\leq |x|^2 + (V_{max}\Delta t^N)^2 + 2V_{max}\Delta t^N|x|,\end{aligned}$$

and, for $s \in [\varsigma_j^{-1}(t), t]$,

$$|\Phi_t^{e_j}(0, s)|^2 = |\int_s^t v_{n-1}^N(\Phi_u^{e_j}(0, s))du|^2 \leq (V_{max}\Delta t^N)^2.$$

Then, we can repeat the argument used to estimate the first moment to obtain a similar uniform bound for the second moment. □

Inequality (4.10) shows that the map $t \mapsto \mu_t^N$ is not Lipschitz continuous in t if the source measure σ_0 is not absolutely continuous with respect to the Lebesgue measure. To prove the convergence of μ^N, we need to assume that $\sigma_0 \in \mathcal{M}^+(\mathcal{S} \times [0, T))$ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}(dt)$ on $[0, T)$ for every source $x_i \in \mathcal{S}$, i.e.

$$\sigma_0^i \ll \mathcal{L}(dt) \qquad \forall x_i \in \mathcal{S}. \tag{4.16}$$

Theorem 4.1. *Assume μ_0 has finite p-momentum for $p = 1, 2$ and* (4.16)*, then the sequence $\{\mu^N\}_{N\in\mathbb{N}}$ defined in* (4.8) *converges (up to a subsequence) to a map $\mu : [0, T] \to \mathcal{M}^+(\Gamma)$ in $C([0, T]; \mathcal{M}^+(\Gamma))$, i.e.*

$$\lim_{N\to+\infty} \sup_{t\in[0,T]} \|\mu_t^N - \mu_t\|_{BL}^* \to 0. \tag{4.17}$$

In addition, the measure $\mu := \int_0^T \int_\Gamma \delta_{(x,t)} d\mu_t(x)dt$ is a solution of (4.2) *in sense of Definition 4.1.*

Proof. Step (i): Convergence.
To show that $\{\mu^N\}_{N\in\mathbb{N}}$ is relatively compact in $C([0, T], \mathcal{M}^+(\Gamma))$, it is sufficient to verify that the sequence satisfies the conditions of the Ascoli-Arzelà criterion in the space of measures (see [2]): equicontinuity, tightness and uniform integrability.
Equicontinuity is consequence of Proposition 4.1, taking into account that by (4.10), (4.16) $\{\mu^N\}_{N\in\mathbb{N}}$ is uniformly Lipschitz continuous in t. The other two properties, tightness and uniform integrability, are implied by the uniform estimates on the first and second moments of the measure μ_t^N in Proposition 4.2. Hence we conclude that that, up to a subsequence, there exists $\mu \in C([0, T], \mathcal{M}^+(\Gamma))$ such that (4.17) holds.

Step (ii): μ satisfies the balance equation (4.3)

We now show that the measure $\mu \in \mathcal{M}^+(\Gamma \times [0,T])$ defined by

$$\mu(dxdt) := \mu_t(dx) \otimes dt = \int_0^T \int_\Gamma \delta_{(x,t)} d\mu_t(x) dt, \tag{4.18}$$

where μ_t is as in(4.17), satisfies (4.3). Set $v_n^N = v[\mu_{t_n^N}^N]$. Summing over n the identities (4.6) and (4.7), we get that the measure $\mu^N = \mu_t^N(dx) \otimes dt \in \mathcal{M}^+(\Gamma \times [0,T])$ satisfies

$$\langle \mu_T^N - \mu_0,\, f\rangle - \langle f,\, \sigma_0\rangle = \sum_{n=0}^{2^N-1} \int_{I_n^N} \langle \mu_t^N,\, (\partial_t + v_n^N \partial_x) \cdot f(\cdot, t)\rangle dt, \tag{4.19}$$

for every $f \in C_0^\infty(\Gamma \times [0,T])$. Passing to the limit for $N \to +\infty$ in (4.19), we first observe that by (4.17) we have the convergence of the left hand side of (4.19) to the one of (4.3), i.e.

$$\langle \mu_T^N - \mu_0,\, f\rangle \to \langle \mu_T - \mu_0,\, f\rangle \quad \text{for } N \to \infty.$$

We show the convergence of the right hand side of (4.19) to the one of (4.3) by estimating

$$\begin{aligned} &\sum_{n=0}^{2^N-1} \int_{I_n^N} \left(\langle \mu_t^N,\, (\partial_t + v_n^N \partial_x) \cdot f(\cdot, t)\rangle dt - \int_0^T \langle \mu_t,\, (\partial_t + v[\mu_t]\partial_x) \cdot f(\cdot, t)\rangle \right) dt \\ = &\int_{I_n^N} \langle \mu_t^N - \mu_t,\, (\partial_t + v[\mu_t]\partial_x) \cdot f(\cdot, t)\rangle dt + \sum_{n=0}^{2^N-1} \int_{I_n^N} \langle \mu_t^N,\, (v_n^N - v[\mu_t])\partial_x \cdot f(\cdot, t)\rangle dt. \end{aligned} \tag{4.20}$$

For $f \in C_0^\infty(\Gamma \times [0,T])$ by (4.17)

$$\begin{aligned} &\sum_{n=0}^{2^N-1} \int_{I_n^N} \langle \mu_t - \mu_t^N,\, (\partial_t + v[\mu_t]\partial_x) \cdot f(\cdot, t)\rangle dt \\ &\leq T \sup_{t\in[0,T]} |\langle \mu_t - \mu_t^N,\, (\partial_t + v[\mu_t]\partial_x) \cdot f(\cdot, t)\rangle| \leq CT \sup_{t\in[0,T]} \|\mu_t^N - \mu_t\|_{BL}^*. \end{aligned} \tag{4.21}$$

Moreover, fixed $n = 0, \ldots, 2^N - 1$, for $t \in I_n^N$ and $x \in \Gamma$, we have

$$|v_n^N(x) - v[\mu_t](x)| \leq L\|\mu_{t_n^N}^N - \mu_t\|_{BL}^* \leq L\|\mu_{t_n^N}^N - \mu_t^N\|_{BL}^* + L\|\mu_t^N - \mu_t\|_{BL}^*. \tag{4.22}$$

By (4.10), we estimate

$$\|\mu_{t_n^N}^N - \mu_t^N\|_{BL}^* \leq \sigma_0([t_n^N, t]) + C|t - t_n^N|,$$

and therefore

$$\int_{t_n^N}^{t_n^N+\Delta t^N} \langle \mu_t^N, (v_n^N - v[\mu_t])\partial_x f(\cdot,t)\rangle dt \le$$
$$\le \int_{t_n^N}^{t_n^N+\Delta t^N} \left[C_1(\sigma_0([t_n^N,t]) + |t-t_n^N| + \|\mu_t^N - \mu_t\|_{BL}^*)\langle \mu_t^N, \partial_x f(\cdot,t)\rangle\right] dt$$
$$\le C_1 \left(\sigma_0(I_n^N)\Delta t^N + \frac{1}{2}(\Delta t^N)^2 + \Delta t^N \sup_{t\in I_n^N} \|\mu_t^N - \mu_t\|_{BL}^*\right) \sup_{t\in I_n^N} \|\mu_t^N\|_{BL}^*,$$

where $C_1 = \max\{L, CL\}$. By (4.9), we have the estimate

$$\sup_{t\in[0,T]} \|\mu_t^N\|_{BL}^* < D$$

for a positive constant D independent of N. Hence

$$\int_{I_n^N} \langle \mu_t^N, (v_n^N - v[\mu_t])\partial_x f(\cdot,t)\rangle dt \le DC_1\Delta t^N \left(\sigma_0(I_n^N) + \frac{1}{2}\Delta t^N + \sup_{t\in I_n^N} \|\mu_t^N - \mu_t\|_{BL}^*\right),$$

and therefore

$$\begin{aligned} &|\sum_{n=0}^{2^N-1} \int_{I_n^N} \langle \mu_t^N, (\partial_t + (v_n^N - v[\mu_t])\partial_x)\cdot f(\cdot,t)\rangle dt| \\ &\qquad \le DC_1 \left(\Delta t^N \sigma_0([0,T]) + \frac{T}{2}\Delta t^N + T \sup_{t\in[0,T]} \|\mu_t^N - \mu_t\|_{BL}^*\right). \end{aligned} \tag{4.23}$$

Substituting (4.21) and (4.23) in (4.20) and passing to the limit for $N \to \infty$, we finally get that the measure μ satisfies the balance equation (4.3).

Step (iii): μ satisfies the vertex condition (4.4).

We have to show that the restrictions of μ to the vertices, defined by the identities

$$\langle \mu^j_{x=x_i}, f\rangle = \int_0^T \langle \mu_t^j, (\partial_t + v[\mu_t]\partial_x) f\rangle dt - \langle \mu_T^j - \mu_0^j, f\rangle, \tag{4.24}$$

if $e_j \in Inc(x_i)$, or

$$-\langle \mu^j_{x=x_i}, f\rangle = \int_0^T \langle \mu_t^j, (\partial_t + v[\mu_t]\partial_x) f\rangle dt - \langle \mu_T^j - \mu_0^j, f\rangle, \tag{4.25}$$

if $e_j \in Out(x_i)$, satisfy that the vertex condition (4.4). By (4.19), we have that

$$\langle \mu_T^N - \mu_0^N, f\rangle = \langle \sigma_0, f\rangle + \sum_{n=0}^{2^N-1} \int_{t_n}^{t_{n+1}} \langle \mu_t^N, (\partial_t + v[\mu_{t_n}^N]\partial_x)\cdot f(\cdot,t)\rangle dt; \tag{4.26}$$

Let $f \in C_0^\infty(\Gamma \times [0,T])$ be such that there exists a unique vertex $x_i \in \mathcal{V}$ which belongs to the support of $f(\cdot,t)$, for every $t \in [0,T]$. Then, taking into account that the support of f does not contain source vertices, we have

$$\langle \mu_T^N - \mu_0,\, f\rangle = \sum_{n=0}^{2^N-1} \int_{I_n^N} \langle \mu_t^N,\, (\partial_t + v_n^N \partial_x) f(\cdot,t)\rangle dt;$$

Moreover, by (4.6) and (4.7), if $supp(f) \subseteq e_j \times [0,T]$ and $e_j \in Inc(x_i)$, then

$$\langle \mu_{x=x_i}^{N,j},\, f\rangle = \sum_{n=0}^{2^N-1} \int_{I_n^N} \langle \mu_t^{N,j},\, (\partial_t + v_n^N \partial_x) f\rangle dt - \langle \mu_T^{N,j} - \mu_0^j,\, f\rangle; \tag{4.27}$$

otherwise, if $e_j \in Out(x_i)$, then

$$-\langle \mu_{x=x_i}^{N,j},\, f\rangle = \sum_{n=0}^{2^N-1} \int_{I_n^N} \langle \mu_t^{N,j},\, (\partial_t + v_n^N \partial_x) f\rangle dt - \langle \mu_T^{N,j} - \mu_0^j,\, f\rangle. \tag{4.28}$$

Passing to the limit for $N \to +\infty$ in either (4.27) or (4.28), by (4.17) we get that there exists measures $\mu_{x=x_i}^j \in \mathcal{M}(\{x_i\} \times [0,T])$ which satisfy (4.24) or (4.25), and such that $\|\mu_{x=x_i}^{N,j} - \mu_{x=x_i}^j\|_{BL}^* \to 0$ for $N \to +\infty$. Since by construction $\mu_{x=x_i}^{N,j} = \sum_{k \in Inc(x_i)} p_{ke} \cdot \mu_{x=x_i}^{N,k}$, we get that the same transmission condition (4.4) is satisfied by the limit measure μ. □

Remark 4.1.1. *For traffic flow problems on road networks, the assumption* (4.16) *excludes the presence of atomic terms in the source measure σ_0. Recall that* (4.16) *gives the uniform continuity with respect to t of the maps μ_t^N, $t \in [0,T]$, necessary to apply the Ascoli-Arzela criterion. We now explain how to partially overcome this difficulty if the source measure is of the type*

$$\sigma_0 = \sum_{x_i \in \mathcal{S}} (\sigma_{ac,0}^{x_i} + \sigma_{d,0}^{x_i}) \tag{4.29}$$

where $\sigma_{ac,0}^{x_i} \ll \mathcal{L}(dt)$ and $\sigma_{d,0}^{x_i}$ is an atomic measure in $\mathcal{M}^+(\mathcal{S} \times [0,T])$ with a finite number of atoms. Consider first the case of a source measure $\sigma_0 = \delta_{(x_i,\tau)}$, for $x_i \in \mathcal{S}$ and $\tau \in (0,T)$. We can apply Theorem 4.1 in $[0,\tau]$ where $\sigma_0 \equiv 0$ is absolutely continuous with respect to $\mathcal{L}(dt)$ to obtain the existence of a solution μ to (4.2) *in $[0,\tau]$. Then we consider* (4.2) *in $[\tau,T]$ with initial condition $\mu_\tau + \delta_{(x_i,\tau)}$ and boundary measure $(\sigma_0)\llcorner_{(\tau,T]} \equiv 0$. Again, since $\sigma_0 \equiv 0$ is absolutely continuous with respect to $\mathcal{L}(dt)$ in $[\tau,T]$, we obtain a solution of the problem in $[\tau,T]$. Gluing together the solutions previously obtained in $[0,\tau]$ and $[\tau,T]$, we obtain a piecewise continuous solution of* (4.2) *on $[0,T]$. Clearly this procedure can be repeated if the source measure σ_0 contains a finite number of atoms. The resulting solution of* (4.2) *is piecewise Lipschitz continuous on a finite number of disjoint intervals in $[0,T]$.*

4.2 Superposition principle on networks

The next result is a representation formula which characterizes the solution μ for the linear transport equation on networks in terms of the distribution matrix $P(t)$ and of the push-forward of the initial and boundary data on the paths over Γ.

Definition 4.3. *Given $x \in \Gamma$, a path γ starting from x is a sequence of edges $(e_{j_0}, e_{j_1}, \dots, e_{j_n}, \dots)$ where, for $i \in \mathbb{N}$, $e_{j_i} \cap e_{j_{i+1}} = x_{j_i} \in \mathcal{V}$, $e_{j_i} \to e_{j_{i+1}}$, e_{j_0} is the sub-edge with endpoints x and $x_{j_0} \in \mathcal{V}$ and the length $\ell(\gamma)$ of γ is infinite. We denote with $\mathcal{A}(x)$ the set of paths γ starting from x.*

Since the network Γ is oriented and $\mathcal{E}$ finite, a path γ is necessarily of one of the following types

- γ is composed by a finite number of arcs and the last one e_{j_n} has infinite length.
- γ is composed by an infinite number of arcs and there exists $n_0, k_0 \in \mathbb{N}$ such that for $n \geq n_0$, γ is given by a cycle $(e_{j_{n_0}}, \dots, e_{j_{n_0+k_0}})$ with $e_{j_{n_0+k_0}} = e_{j_{n_0}}$.

We denote by Φ^γ the flow map associated to the velocity field v restricted to γ, i.e. $\Phi^\gamma_s(x,s) = x$ and there are $t_0 := s < t_1 < \dots < t_n < \dots$ such that for any $\mu \in \mathbb{N}$, we have $\Phi^\gamma([t_m, t_{m+1}]) \subset e_{j_m}$ and

$$\frac{d}{dt}\Phi^\gamma_t(x,s) = v(\Phi^\gamma_t(x,s)), \qquad t \in [t_m, t_{m+1}).$$

We define the exit times from the arc $e_{j_k} = \pi_{j_k}([0, L_{j_k}])$ of γ as

$$\begin{aligned}
\theta^\gamma_0(x,s) &= \inf\{t \geq s : \Phi^\gamma(x,s) = \pi_{j_0}(L_{j_0})\}, \\
\theta^\gamma_k(x,s) &= \inf\{t \geq \theta^\gamma_{k-1}(x,s) : \Phi^\gamma(x,s) = \pi_{j_k}(L_{j_k})\} \quad k \in \mathbb{N},\, k > 0,
\end{aligned}$$

and we associate to each $(x,s) \in \Gamma \times [0,T]$ and to each $\gamma \in \mathcal{A}(x)$ a coefficient $p_\gamma(x,s) \in [0,1]$ defined by

$$p_\gamma(x,s) := \prod_k p_{j_k j_{k+1}}(\theta^\gamma_k(x,s)). \tag{4.30}$$

where $p_{j_k j_{k+1}}$ are the entries of the distribution matrix P defined in (3.3). The coefficient $p_\gamma(x,s)$ can be interpreted as the fraction of the total mass transported along the path γ. Due to the properties of P, it follows that

$$0 \leq p_\gamma(x,s) \leq 1, \qquad \sum_{\gamma \in \mathcal{A}(x)} p_\gamma(x,s) = 1, \quad \forall x \in \Gamma,\ s \in [0,T].$$

Theorem 4.2. *If $\mu \in \mathcal{M}^+(\Gamma \times [0,T])$ is a solution of* (3.4)*, then for any $t \in [0,T]$, μ_t is given by*

$$\mu_t = \int_\Gamma \sum_{\gamma \in \mathcal{A}(x)} \delta_{(\Phi_t^\gamma(x,0),t)} p_\gamma(x,0) dm_0(x) + \sum_{x_i \in \mathcal{S}} \int_{[0,t]} \sum_{\gamma \in \mathcal{A}(x_i)} \delta_{(\Phi_t^\gamma(0,s),t)} p_\gamma(0,s) d\sigma_0^i(s). \tag{4.31}$$

We preliminarily recall a characterization of the traces of the solution μ of (3.27) on the fibers $e_j \times \{t\}$ and $\{x\} \times [0,\, t]$, where $x = \pi_j(L_j)$ is the terminal point of e_j, in dependence of the initial and boundary data inside e_j (see [14]).

Proof of Theorem 4.2. At first, we consider a simple network with $\mathcal{V}$ given by two vertices $\{S, V\}$, where S is a source and V an internal vertex, and $\mathcal{E}$ given by an arc e_1 connecting S to V and by $n-1$ unbounded arcs $e_k \in \mathrm{Out}(V)$.
Due to this choice, we can observe that the structure of paths is simplified. Indeed, a path $\gamma \in \mathcal{A}(x)$ is either a subset of (e_1, e_k) if $x \in e_1$ or a subset of e_k if $x \in e_k$.
The solution can be written as $m = \int_0^T \delta_{(x,t)} dm_t(x) dt$, where $\mu_t = \sum_{k=1}^n \mu_t^k$. If $k = 1$, by (3.18) with $S = \pi_1(0)$ and $V = \pi_1(L_1)$, the solution restricted on e_1 is given by

$$\mu_t^1 = \int_{[0,\, \max\{0,\, (\tau_1)^{-1}(t)\}]} \delta_{\Phi_t^{e_1}(x,\,0)}\, d\mu_0^1(x) + \int_{(\max\{0,\, (\varsigma_1)^{-1}(t)\},\, t]} \delta_{\Phi_t^{e_1}(0,\, s)}\, d\sigma_0(s);$$

otherwise for $k \in \{2, \ldots, n\}$, by (3.18) with $V = \pi_k(0)$, on e_k it is given by

$$\mu_t^k = \int_{[0, \max\{0, (\tau_k)^{-1}(t)\}]} \delta_{\Phi_t^{e_k}(y,0)} d\mu_0^k(x) + \int_{(\max\{0, (\varsigma_k)^{-1}(t)\}, t]} \delta_{\Phi_t^{e_k}(0,s)} d\mu_{x=V}^k(s).$$

Observe that the first term on the right hand-side of the previous equation is the pushforward of the mass μ_0^k which is at time $t = 0$ on e_k; the second term is the fraction of the mass which flows from e_1 in e_k. Using the transmission condition $\mu_{x=V}^k = p_{1k} \cdot \mu_{x=V}^1$ and recalling that by (3.19) we have

$$\mu_{x=V}^1 = \int_{(\max\{0,\, (\tau_1)^{-1}(t)\},\, L_1]} \delta_{\tau_1(x)}\, d\mu_0^1(x) + \int_{[0,\, \max\{0,\, (\varsigma_1)^{-1}(t)\}]} \delta_{\varsigma_1(s)}\, d\sigma_0(s)$$

we get for $f \in C_0^\infty(\Gamma \times [0,T])$

$$\begin{aligned}
&\int_{(\max\{0,(\varsigma_k)^{-1}(t)\},t]} f(\Phi_t^{e_k}(0,s),t)d\mu^k_{x=V}(s)\\
&\qquad\qquad = \int_{(\max\{0,(\tau_1)^{-1}(t)\},L_1]} f(\Phi_t^{e_k}(0,\tau_1(x)),t)p_{1k}(\tau_1(x))d\mu_0^1(x)\\
&\qquad\qquad\quad + \int_{[0,\max\{0,(\varsigma_1)^{-1}(t)\}]} f(\Phi_t^{e_k}(0,\varsigma_1(s)),t)p_{1k}(\varsigma_1(s))d\sigma_0(s)\\
&\qquad\qquad = \int_{(\max\{0,(\tau_1)^{-1}(t)\},L_1]} f\big(\Phi_t^{e_k}(\Phi^{e_1}_{\tau_1(x)}(x,0),\tau_1(x)),t\big)p_{1k}(\tau_1(x))d\mu_0^1(x)\\
&\qquad\qquad\quad + \int_{[0,\max\{0,(\varsigma_1)^{-1}(t)\}]} f(\Phi_t^{e_k}(\Phi_{\varsigma_1(s)}(0,s),\varsigma_1(s)),t)p_{1k}(\varsigma_1(s))d\sigma_0(s)\\
&\qquad\qquad = \int_{(\max\{0,(\tau_1)^{-1}(t)\},L_1]} f(\Phi_t^{\gamma}(x,0),t)p_{1k}(\tau_1(x))d\mu_0^1(x)\\
&\qquad\qquad\quad + \int_{[0,\max\{0,(\varsigma_1)^{-1}(t)\}]} f(\Phi_t^{\gamma}(0,s),t)p_{1k}(\varsigma_1(s))d\sigma_0(s),
\end{aligned} \tag{4.32}$$

We observe that μ_t^1 can be split in $n-1$ parts, in dependence of the distribution terms p_{1k}. Indeed, if we write $\mu_t^1 = \sum_{k=2}^n (p_{1k}\circ\theta_0^\gamma)\cdot\mu_t^1$, then

$$\mu_t = \mu_t^1 + \sum_{k=2}^n \mu_t^k = \sum_{k=2}^n(((p_{1k}\circ\theta_0^\gamma)\cdot\mu_t^1) + \mu_t^k).$$

Concerning the first term, observing that $\tau_1(x) = \theta_0^\gamma(x,0)$ and $\varsigma_1(s) = \theta_0(0,s)$, we compute for any $f \in C_0^\infty(\Gamma \times [0,T])$

$$\begin{aligned}
\langle (p_{1k}\circ\theta_0^\gamma)\mu_t^1,\, f\rangle &= \int_{[0,\max\{0,(\tau_1)^{-1}(t)\}]} f(\Phi_t^{e_1}(x,0),t)p_{1k}(\theta_0^\gamma(x,0))d\mu_0^1(x)\\
&\quad + \int_{(\max\{0,(\varsigma_1)^{-1}(t)\},t]} f(\Phi_t^{e_1}(0,s),t)p_{1k}(\theta_0^\gamma(0,s))d\sigma_0(s)\\
&= \int_{[0,\max\{0,(\tau_1)^{-1}(t)\}]} f(\Phi_t^{e_1}(x,0),t)p_{1k}(\tau_1(x))d\mu_0^1(x)\\
&\quad + \int_{(\max\{0,(\varsigma_1)^{-1}(t)\},t]} f(\Phi_t^{e_1}(0,s),t)p_{1k}(\varsigma_1(s))d\sigma_0(x).
\end{aligned} \tag{4.33}$$

By (4.33),(4.32) it follows that, from the parametrization used for each arc,

$$\begin{aligned}
\langle (p_{1k}\circ\theta_0^\gamma)\cdot\mu_t^1,\, f\rangle + \langle \mu_t^k,\, f\rangle &= \int_{e_1} f(\Phi_t^\gamma(x,0),t)p_{1k}(\theta_0^\gamma(x,0))d\mu_0^1(x) + \int_{e_2} f(\Phi_t^\gamma(x,0),t)d\mu_0^k(x)\\
&\quad + \int_{[0,t]} f(\Phi^\gamma(0,s),t)p_{1k}(\theta_0^\gamma(0,s))d\sigma_0(s).
\end{aligned}$$

If we sum the previous formula over $e_k \in \mathrm{Out}(V)$ we have

$$\langle \mu_t,\, f\rangle = \int_{[0,t]} \sum_{\gamma\in\mathcal{A}(x)} f(\Phi_t^\gamma(0,s),t)p_\gamma(0,s)d\sigma_0(s) + \int_\Gamma \sum_{\gamma\in\mathcal{A}(x)} f(\Phi_t^\gamma(x,0),t)p_\gamma(x,0)d\mu_0(x), \tag{4.34}$$

Hence we have proved formula (4.34) for the special case of the simple network as above. If we consider a network with a similar structure but with multiple sources, it is sufficient to sum the contribution of each source $x_i \in \mathcal{S}$ to get the thesis.
Finally, the case of a general network can be studied in a similar way by taking a test function localized around a single vertex and repeating the argument used in the previous proof for the simple network so obtained. □

We extend to the nonlinear transport problem (4.2) the representation formula for the solution of the linear problem (3.27) proved in Theorem 4.2. Given $\mu \in \mathcal{M}^+(\Gamma\times[0,T])$, we denote by Φ^γ the flow map associated to the velocity field $v[\mu_t]$ restricted to γ, i.e. $\Phi_s^\gamma(x,s) = x$ and there are $t_0 := s < t_1 < \dots < t_n < \dots$ such that for any $m = 0,1,\dots$, we have $\Phi^\gamma([t_m,t_{m+1}]) \subset e_{j_m}$ and

$$\frac{d}{dt}\Phi_t^\gamma(x,s) = v[\mu_t](\Phi_t^\gamma(x,s)), \qquad t\in[t_m,t_{m+1}).$$

Proposition 4.3. *If $\mu \in \mathcal{M}^+(\Gamma\times[0,T])$ is given by (4.17), then for any $t\in[0,T]$, μ_t is given by*

$$\mu_t = \int_\Gamma \sum_{\gamma\in\mathcal{A}(x)} \delta_{(\Phi_t^\gamma(x,0),t)}p_\gamma(x,0)dm_0(x) + \sum_{x_i\in\mathcal{S}}\int_{[0,t]} \sum_{\gamma\in\mathcal{A}(x_i)} \delta_{(\Phi_t^\gamma(0,s),t)}p_\gamma(0,s)d\sigma_0^i(s). \tag{4.35}$$

where the coefficients p_γ are defined as in (4.30).

Proof. We observe that by (4.22), it follows that

$$\sup_{x\in e_j}|v_j^N(x) - v[\mu_t](x)| \to 0 \qquad \text{for } N\to+\infty,\ j\in J$$

The previous estimate implies the uniform convergence of the respective flow maps on a given path γ and the convergence of (4.31) to (4.35). □

Proposition 4.4. *Given initial data $\mu_0^1,\mu_0^2 \in \mathcal{M}^+(\Gamma\times\{0\})$, and boundary data $\sigma_0^1,\sigma_0^2 \in \mathcal{M}^+(\mathcal{S}\times[0,\,T])$ satisfying (4.16) and have bounded $p-$momentum, for $p=1,2$, and denoted by μ^1 and μ^2 the corresponding solutions, then*

$$\sup_{t\in[0,T]}\|\mu_t^1-\mu_t^2\|_{BL}^* \le C(\|\mu_0^1-\mu_0^2\|_{BL}^* + \|\sigma_0^1-\sigma_0^2\|_{BL}^*), \tag{4.36}$$

where $C = C(T)$ is a positive constant.

Proof. Fixed $x \in \Gamma$, we consider a path $\gamma \in \mathcal{A}(x)$ starting from x and the flow maps $\Phi^{1,\gamma}$ and $\Phi^{2,\gamma}$ associated, respectively, to $v[\mu_t^1]$ and $v[\mu_t^2]$. Let $f \in BL(\Gamma \times [0,T])$ with $\|f\|_{BL}^* \le 1$, then by formula (4.35) we have

$$\begin{aligned}
&\langle \mu_t^1 - \mu_t^2,\, f\rangle \\
&= \int_\Gamma \left(\sum_{\gamma\in\mathcal{A}(x)} f(\Phi_t^{1,\gamma}(x,0),t)p_\gamma(x,0)d\mu_0^1(x) - \sum_{\gamma\in\mathcal{A}(x)} f(\Phi_t^{2,\gamma}(x,0),t)p_\gamma(x,0)d\mu_0^2(x) \right) + \\
&+ \sum_{x_i\in\mathcal{S}} \left(\int_{[0,T]} \sum_{\gamma\in\mathcal{A}(x_i)} f(\Phi_t^{1,\gamma}(x_i,s),t)p_\gamma(x_i)d\sigma^1_{x_i}(s) - \int_{[0,T]} \sum_{\gamma\in\mathcal{A}(x_i)} f(\Phi_t^{2,\gamma}(x_i,s),t)p_\gamma(x_i)d\sigma^2_{x_i}(s) \right).
\end{aligned} \tag{4.37}$$

To estimate the right hand side in (4.37), we rewrite the first term as

$$\begin{aligned}
&\int_\Gamma \left(\sum_{\gamma\in\mathcal{A}(x)} f(\Phi_t^{1,\gamma}(x,0),t)p_\gamma(x,0)d\mu_0^1(x) - \int_\Gamma \sum_{\gamma\in\mathcal{A}(x)} f(\Phi_t^{2,\gamma}(x,0),t)p_\gamma(x,0)d\mu_0^2(x) \right) = \\
&+ \int_\Gamma \sum_{\gamma\in\mathcal{A}(x)} f(\Phi_t^{2,\gamma}(x,0),t)p_\gamma(x,0)d(\mu_0^1 - \mu_0^2)(x) + \\
&\int_\Gamma \sum_{\gamma\in\mathcal{A}(x)} \left(f(\Phi_t^{1,\gamma}(x,0),t) - f(\Phi_t^{2,\gamma}(x,0),t) \right) p_\gamma(x,0)d\mu_0^1(x).
\end{aligned}$$

Since $\|f\|_{BL}^* \le 1$ and $\sum_{\gamma\in\mathcal{A}(x)} p_\gamma(x,0) = 1$ for every $x \in \Gamma$, we have the estimate

$$\int_\Gamma \sum_{\gamma\in\mathcal{A}(x)} f(\Phi_t^{2,\gamma}(x,0),t)p_\gamma(x,0)d(\mu_0^1 - \mu_0^2)(x) \le \|\mu_0^1 - \mu_0^2\|_{BL}^*. \tag{4.38}$$

Moreover

$$|f(\Phi_t^{1,\gamma}(x,0),t) - f(\Phi_t^{2,\gamma}(x,0),t)| \le d(\Phi_t^{1,\gamma}(x,0), \Phi_t^{2,\gamma}(x,0)) \le d_\gamma((\Phi_t^{1,\gamma}(x,0), \Phi_t^{2,\gamma}(x,0))),$$

where d_γ is the path distance d restricted to γ. It follows that

$$\begin{aligned}
d_\gamma(\Phi_t^{1,\gamma}(x,0), \Phi_t^{2,\gamma}(x,0)) &\le \int_0^t \left| v[m_s^1](\Phi_s^{1,\gamma}(x,0)) - v[\mu_s^2](\Phi_s^{2,\gamma}(x,0)) \right| ds \\
&\le \int_0^t L\left(\|\mu_s^1 - \mu_s^2\|_{BL}^* + d_\gamma(\Phi_s^{1,\gamma}(x,0), \Phi_s^{2,\gamma}(x,0)) \right) ds.
\end{aligned}$$

By Gronwall's inequality, we get

$$d_\gamma(\Phi_t^{1,\gamma}(x,0), \Phi_t^{2,\gamma}(x,0)) \leq L\left(\int_0^t \|\mu_s^1 - \mu_s^2\|_{BL}^* ds\right) e^{Lt},$$

and consequently

$$|f(\Phi_t^{1,\gamma}(x,0),t) - f(\Phi_t^{2,\gamma}(x,0),t)| \leq L\left(\int_0^t \|\mu_s^1 - \mu_s^2\|_{BL}^* ds\right) e^{Lt}.$$

The previous inequality implies that

$$\begin{aligned}\int_\Gamma \sum_{\gamma\in\mathcal{A}(x)} \left(f(\Phi_t^{1,\gamma}(x,0),t) - f(\Phi_t^{2,\gamma}(x,0),t)\right) p_\gamma(x,0) d\mu_0^1(x)\\ \leq L\left(\int_0^t \|\mu_s^1 - \mu_s^2\|_{BL}^* ds\right) e^{Lt}\|\mu_0^1\|_{BL}^*.\end{aligned} \tag{4.39}$$

Proceeding in a similar way for the second term in (4.37), we obtain the inequality

$$\begin{aligned}&\sum_{x_i\in\mathcal{S}}\left(\int_{[0,T]} \sum_{\gamma\in\mathcal{A}(x_i)} f(\Phi_t^{1,\gamma}(x_i,s),t) p_\gamma(x_i) d\sigma_{x_i}^1(s) - \int_{[0,T]} \sum_{\gamma\in\mathcal{A}(x_i)} f(\Phi_t^{2,\gamma}(x_i,s),t) p_\gamma(x_i) d\sigma_{x_i}^2(s)\right)\\ &\leq \|\sigma_0^1 - \sigma_0^2\|_{BL}^* + L\left(\int_0^t \|\mu_s^1 - \mu_s^2\|_{BL}^* ds\right) e^{Lt}\|\sigma_0^1\|_{BL}^*.\end{aligned} \tag{4.40}$$

By using (4.38), (4.39) and (4.40) in (4.37), we get

$$\langle \mu_t^1 - \mu_t^2,\, f\rangle \leq \|\mu_0^1 - \mu_0^2\|_{BL}^* + \|\sigma_0^1 - \sigma_0^2\|_{BL}^* + C\int_0^t \|\mu_s^1 - \mu_s^2\|_{BL}^* ds,$$

where $C = Le^{LT}(\|\mu_0^1\|_{BL}^* + \|\sigma_0^1\|_{BL}^*)$. Taking the supremum with respect to f we get

$$\|\mu_t^1 - \mu_t^2\|_{BL}^* \leq (\|\mu_0^1 - \mu_0^2\|_{BL}^* + \|\sigma_0^1 - \sigma_0^2\|_{BL}^*) + C\int_0^t \|\mu_s^1 - \mu_s^2\|_{BL}^* ds$$

and applying again Gronwall's inequality, we finally obtain

$$\|\mu_t^1 - \mu_t^2\|_{BL}^* \leq (\|\mu_0^1 - \mu_0^2\|_{BL}^* + \|\sigma_0^1 - \sigma_0^2\|_{BL}^*)e^{Ct}.$$

□

As an immediate consequence of the continuous dependence result we have

Corollary 4.2.1. *The solution of the nonlinear transport problem* (4.2) *is unique.*

Up to this point, we have used a constructive approach to build the solution of the transport equation on networks. This approach has had the advantage to induce a natural transmission condition thanks to a stochastic matrix which distributes the mass at the junction.

Observe that even if the velocity field is regular on every arc we lose this property at the vertices of Γ. It is well known that the loss of regularity implies the lose of the uniqueness of solution.

The transmission condition we have adopted allows us to select a solution to our problem. In particular, we have selected the transport of mass which is instantaneous at the junctions, i.e. drivers in our model are not allowed to stop in the junction and have to move to the next roads. This seems a reasonable hypothesis since the behavior of drivers is influenced by the others or external controls, such as traffic lights, rather than their will.

However, in many applications, for example to model supply chains or computer systems, other transmission condition are reasonable which includes, for example, buffers or capacities. For this purpose, it is important to understand how to translate them into the distribution matrix P and which is the effect on superposition formula (4.35).

An interesting condition is suggested in [32, 33] where it is assumed that the particles stop at the junction and there's a positive probability that a driver is stuck at the junction until time t, for any $t \geq 0$. This assumption has not still been studied in case of networks and it is place in the research topic of the admissible conditions at the junction.

Chapter 5

Numerical aspects of multiscale modeling on networks

In this chapter we focus on modeling and numerical aspects of the previously defined traffic flow model; these are relevant for application and calibration based on real data. We will also investigate the impact of parameters on traffic flow dynamic.
In the first section, we propose an example of nonlocal velocity field, suitable to describe interactions among drivers; then, we discuss a macroscopic Godunov scheme introduced in [37] and some numerical tests on simple junctions. Lastly, we discuss perspectives and open problems related to applications and real traffic forecasting.

5.1 Nonlocal fields on networks

In this section we study an example of nonlocal velocity term $v[\mu]$ suitable to describe and predict the evolution of traffic flow on a road network. Taking inspiration by similar models for collective dynamics of crowds (see [24]), we consider a positive velocity fields given by

$$v[\mu](x) := \max\{0, v_{des}(x) - v_i[\mu](x)\}. \tag{5.1}$$

The function $v_{des} : \Gamma \to \mathbb{R}^+$ is the desired velocity, or free flow speed, representing the speed of a car over a free road, while $v_i : \mathcal{M}^+(\Gamma) \times \Gamma \to \mathbb{R}^+$ is the interaction among drivers due to the presence of a car distribution $\mu \in \mathcal{M}^+(\Gamma)$ over the network Γ. Our aim is to identify an appropriate expression for $v[\mu](x)$ consistent with the traffic flow model and satisfying hypothesis **(H1)-(H3)**.

Concerning the free flow speed v_{des}, which depends only on the state variable x, we assume that it is positive, bounded and Lipschitz continuous on each arc e_j of the network Γ.
Observing that for every $x \in \Gamma$, the interaction term is a map $v_I[\cdot](x) : \mathcal{M}^+(\Gamma) \to \mathbb{R}^+$, it seems natural to define it as the functional

$$v_I[\mu](x) := \int_\Gamma K(x,y)d\mu(y),$$

or, more in general,

$$v_I[\mu](x) := \phi\left(\int_\Gamma K(x,y)d\mu(y)\right),$$

where non-decreasing function $\phi \in Lip_{loc}(\mathbb{R}^+)$ such that $\phi(0) = 0$.
Observe that if K is nonnegative and bounded by a positive constant C, then, for every $x \in \Gamma$,

$$0 \leq v_I[\mu](x) \leq C\mu(\Gamma),$$

and therefore **(H1)** is satisfied.
As in the Euclidean case (see [24, Section 5]) the Lipschitz continuity with respect to x is the most delicate hypothesis and we need to focus on our application to traffic flow.
We assume the interactions among drivers depend on their position and relative distance, hence the interaction kernels is the form

$$K(x,y) = k(d_\Gamma(x,y))\chi_{\mathcal{D}(x)}(y), \tag{5.2}$$

where $k : \mathbb{R}^+ \to \mathbb{R}^+$ is a Lipschitz continuous non-increasing function representing the interaction among cars on road networks in dependence of their distance and $\chi_{\mathcal{D}(x)}$ is the characteristic function of the set $\mathcal{D}(x)$. The crucial point is to properly define the set $\mathcal{D}(x)$ which represents the visual field of the driver.
It is reasonable to assume that a driver has only the knowledge of the distribution of the cars on the roads adjacent to his/her current position and, on the basis of this information, he/she gives a certain priority to a possible route. Hence we define the visual field as

$$\mathcal{D}(x) = \{y \in \Gamma : \, x \to y, \, d_\Gamma(x,y) \leq R\},$$

where $R > 0$ is the visual radius; moreover, for simplicity we assume

$$R \leq \min_{e \in \mathcal{E}} \ell(e).$$

Observe that the interaction between drivers is different from the Follow the Leader model or the Zhao-Zang model [55] where driver's interactions are binary.

In this way, $\forall\ e \in \mathcal{E}$ and $\forall\ x \in e$, we have $\mathcal{D}(x) \subset e \cup (\bigcup_{e_j \in Out(V)} e_j)$ where $V = \pi_e(\ell(e))$. The previous assumption allows us to study the interaction between drivers as the average of interactions concentrated on (local) path. For any $e_k \in \mathcal{E}$, we prescribe weights α_{kj} satisfying

$$0 \leq \alpha_{kj} \leq 1, \qquad \sum_{j=1}^{J} \alpha_{kj} = 1,$$

$$\alpha_{kj} = 0 \quad \textit{if either } e_k \cap e_j = \emptyset \textit{ or } e_j \to e_k.$$

Then, we define the interaction term in $x \in e_k$ as

$$v_I[\mu](x) = \sum_{e_j \in \mathcal{E}} \alpha_{kj} \int_\Gamma k(d_\Gamma(x,y)) \chi_{\mathcal{D}_{kj}(x)}(y) d\mu(y), \tag{5.3}$$

where $\mathcal{D}_{kj}(x) = \mathcal{D}(x) \cap (e_k \cup e_j)$.

We remark that the difference among the coefficient $p_{kj}(t)$ in (3.3) and α_{kj} previously defined is that the former represents the capacity of junction $e_k \cup e_j$ to allocate traffic distribution, while the latter the priority of a given route in the choice of the driver depending on observed traffic distribution. In this thesis, the weights α_{kj} are constant and do not depend on time variable or mass distribution on networks. Even if these hypotheses are reasonable for several applications, a further analysis would be necessary.

To prove the Lipschitz continuity in the x variable, it is enough to prove this property for the term

$$\int_\Gamma k(d_\Gamma(x,y)) \chi_{\mathcal{D}_{kj}(x)}(y) d\mu(y).$$

Without loss of generality, we assume that $e_k = [0, L_k]$ and $e_j = [L_k, L_k + L_j]$; hence

$$\mathcal{D}_{kj}(x) = \{y \in [x, L_k + R] \subset [0, L_k + L_j] :\ x \leq y,\ |x - y| \leq R\} =: \mathcal{A}(x).$$

Taken $x_1, x_2 \in [0, L_k]$ with $x_1 \leq x_2$ and defined $h = |x_2 - x_1|$, we can observe that $\mathcal{A}(x_2) = \mathcal{A}(x_1) + h$; then,

$$\chi_{\mathcal{A}(x_2)}(y) = \begin{cases} 1, & \text{if} \quad (y - h) \in \mathcal{A}(x_1) \\ 0, & \text{otherwise} \end{cases}$$

and therefore

$$\chi_{\mathcal{A}(x_2)} = \chi_{\mathcal{A}(x_1)}(y-h) = (\chi_{\mathcal{A}(x_1)} \circ \tau_{-h})(y),$$

where τ_{-h} is the translation on $\mathbb{R}$ with step equal to $-h$.
It follows

$$\begin{aligned}\int k(|x_1-y|)\chi_{\mathcal{A}(x_1)}(y)d\mu(y) - \int k(|x_2-y|)\chi_{\mathcal{A}(x_2)}(y)d\mu(y) = \\ \int k(|x_1-y|)\chi_{\mathcal{A}(x_1)}(y)d\mu(y) - \int k(|x_1-(y-h)|)\chi_{\mathcal{A}(x_1)}(y-h)d\mu(y) = \\ \int \left[k(|x_1-y|)\chi_{\mathcal{A}(x_1)}(y) - (k(|x_1-\cdot|)\chi_{\mathcal{A}(x_1)}(\cdot))\circ\tau_{-h}(y)\right]d\mu(y) \leq \\ \int k(|x_1-y|)\chi_{\mathcal{A}(x_1)}(y)d(\mu(y)-\tau_{-h}\#\mu(y)) \leq K\|\mu-\tau_{-h}\#\mu\|^*_{BL} = Kh = K|x_2-x_1|.\end{aligned}$$

5.2 Numerical scheme and simulations

The proofs of Theorems 3.4 and 4.1 provide us a semi-discrete in time scheme which is the basis for the numerical one:

Step 1: define a time grid with step Δt;

Step 2: define the topologic order of network Γ and $k=0$;

Step 3: while $k\Delta t < T : \mu_{k+1} = Update(\mu_k)$ for $kin\mathbb{N}$.

In this section we describe the $Update(\cdot)$ function used for our simulations.

5.2.1 Macroscopic description

Given a network Γ, a time horizon $T>0$ and a velocity field v, we are interested in simulating the dynamic of distributions $(\mu_t)_{t\in[0,T]}$ resulting from transport equation, starting from $\mu_0 \in \mathcal{M}^+(\Gamma)$.[1]
Assume that μ_0 can be approximated by an absolutely continuous distribution, i.e. there exists $\epsilon > 0$ and a scalar and positive function $\rho_0 \in L^1(\Gamma)\cap L^\infty(\Gamma)$ such that

$$\|\mu_0 - \mathcal{L}(\rho_0^{(\epsilon)})\|^*_{BL} < \epsilon.$$

For absolutely continuous initial data, there are recent papers on hyperbolic equations with nonlocal flux and related numerical methods in the case $\Gamma = \mathbb{R}$ (see [22, 37] and reference therein). We adapt

[1] To simplify the discussion, we assume w.l.o.g. boundary data $\sigma_0 \in \mathcal{M}^+(\mathcal{S}\times[0,T])$ be null.

these schemes as Update for μ^k on networks: parametrized every arc $e \in \mathcal{E}$ as $[0, L_e]$, we take a space step $\Delta x > 0$ such that for every arc $e \in \mathcal{E}$ there exists $N_e \in \mathbb{N}$ such that $L_e = \Delta x N_e$ and the interaction radius $R = N_R \Delta x$, with $N_R \in \mathbb{N}$, we denote $\lambda = \Delta t / \Delta x$. We define the space grid on e by $x_{j+\frac{1}{2}} = (j + \frac{1}{2})\Delta x$ as the cell centers and $x_j = j\Delta x$ as the interfaces for $j \in \{0, \dots, N_e\}$.
The approximated solution on arc e is denoted by $\rho^{n,e}_{j+\frac{1}{2}}$ for $(x,t) \in (x_j, x_{j+1}] \times [t^n, t^{n+1})$, for $j = 0, \dots, N-1$.
In the previous section, we have seen that the nonlocal term is written as the average among one-dimensional integral.
Fixed a Lipschitz continuous non-increasing function $k : \mathbb{R}^+ \to \mathbb{R}^+$, we denote with $\gamma_j = \int_{x_{j+\frac{1}{2}}}^{x_{j+\frac{3}{2}}} k(r)dr, \ \forall\, j \in \{0, \dots, N-1\}$, the (one-dimensional) nonlocal term is written as

$$\sum_{k=0}^{N_R} \gamma_j \rho^{n,e}_{j+k+\frac{3}{2}}; \tag{5.4}$$

it follows the speed v is approximated as

$$V^{n,e}_{j+1} = \sum_{e':e\to e'} \alpha_{ee'} \sum_{k=0}^{N_R} \gamma_j \rho^{n,ee'}_{j+k+\frac{3}{2}}, \tag{5.5}$$

where $\rho^{ee'}$ denotes the density restricted to the local path (e, e').
Then, the numerical flux function of the Godunov scheme introduced in [37] is determined as

$$F^e(\rho^{n,e}_{j+\frac{1}{2}}, \rho^n) = V^{n,e}_{j+1} \rho^{n,e}_{j+\frac{1}{2}}. \tag{5.6}$$

Then, we can initialize the Godunov type scheme defining the initial data as

$$\rho^{0,e}_{j+\frac{1}{2}} = \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \rho_0 dx, \quad \forall\ j \in \{0, \dots, N_e - 1\}$$

and the finite volume scheme

$$\rho^{n+1,e}_{j+\frac{1}{2}} = \rho^{n,e}_{j+\frac{1}{2}} - \lambda(V^{n,e}_{j+1} \rho^{n,e}_{j+\frac{1}{2}} - V^{n,e}_{j} \rho^{n,e}_{j-\frac{1}{2}}), \quad j \in \{1, \dots, N_e - 1\}. \tag{5.7}$$

Observe that for $j = 0$ we need to apply the boundary condition: if there is a source and boundary data given by $\sigma_0 \in \mathcal{M}^+([0,T])$, then

$$F^e(\rho^{n,e}_{-\frac{1}{2}}, \rho^n) := \sigma_0([t_n, t_{n+1})) = 0;$$

otherwise the incoming flux it is determined towards the distribution matrix P

$$F^e(\rho^{n,e}_{-\frac{1}{2}}, \rho^n) = \sum_{e':e' \to e} p^n_{e'e} \left(V^{n,e}_{N_{e'}} \rho^{n,e}_{N_{e'}-\frac{1}{2}} \right).$$

5.2.2 Test 1: 2-to-1 junction.

With the Godunov-type scheme previously introduced, we can test our model in some specific scenarios. In the first test, we want to observe the dynamic in a 2-to-1 junction, i.e. a network with a central node, two incoming and one outgoing arc. For simplicity, we assume that each arc has the same length equal to 1 and no boundary data. We fix $v_{des} \equiv 2.5$ and the interaction kernel is given by

$$k(r) = \frac{2}{R}\left(1 - \frac{r}{R}\right),$$

where R is the interaction radius. In Fig. 5.1 we have plotted the dynamic along every arc (horizontal axis) during the time interval $[0, 1]$ (vertical axis). The density shows interesting features. The bright regions indicate congestions, where the density is close to 1 while a darker color means low density values. We can observe high concentration at time $t = 0$ due to the choice of the initial condition. Another high density region is at the beginning of the third arc due to the transmission conditions and the density coming from the previous arcs.
For $t > 0$, the density flows towards the third arc but, due to the density distribution, it is possible to observe the back-propagation of jam on the first arc, where it meets two front-propagation, and on the second one, which is fastly absorbed.

5.2.3 Test 2: "stop 'n go" waves.

In Fig. 5.1 and 5.2.3 we can observe the so-called "stop 'n go" waves, an interesting behavior which occurs frequently in real situations. The empirical data and previous simulations show that traffic jam is not necessarily restricted to a precise position but it propagates backward with a speed proportional to v_{des}.
In our model, these waves are a consequence of the non-local field; to properly understand, we show in Fig. 5.2.3 the dynamic determined by the nonlocal transport equation over a cycle, i.e. a single arc with periodic boundary condition. We can observe that smaller is the visual radius R, slower they dissipate. Indeed, for large R driver's speed has small variations and the car density reach an equilibrium value. Otherwise, for small R the driver "sees" high concentration region only when they

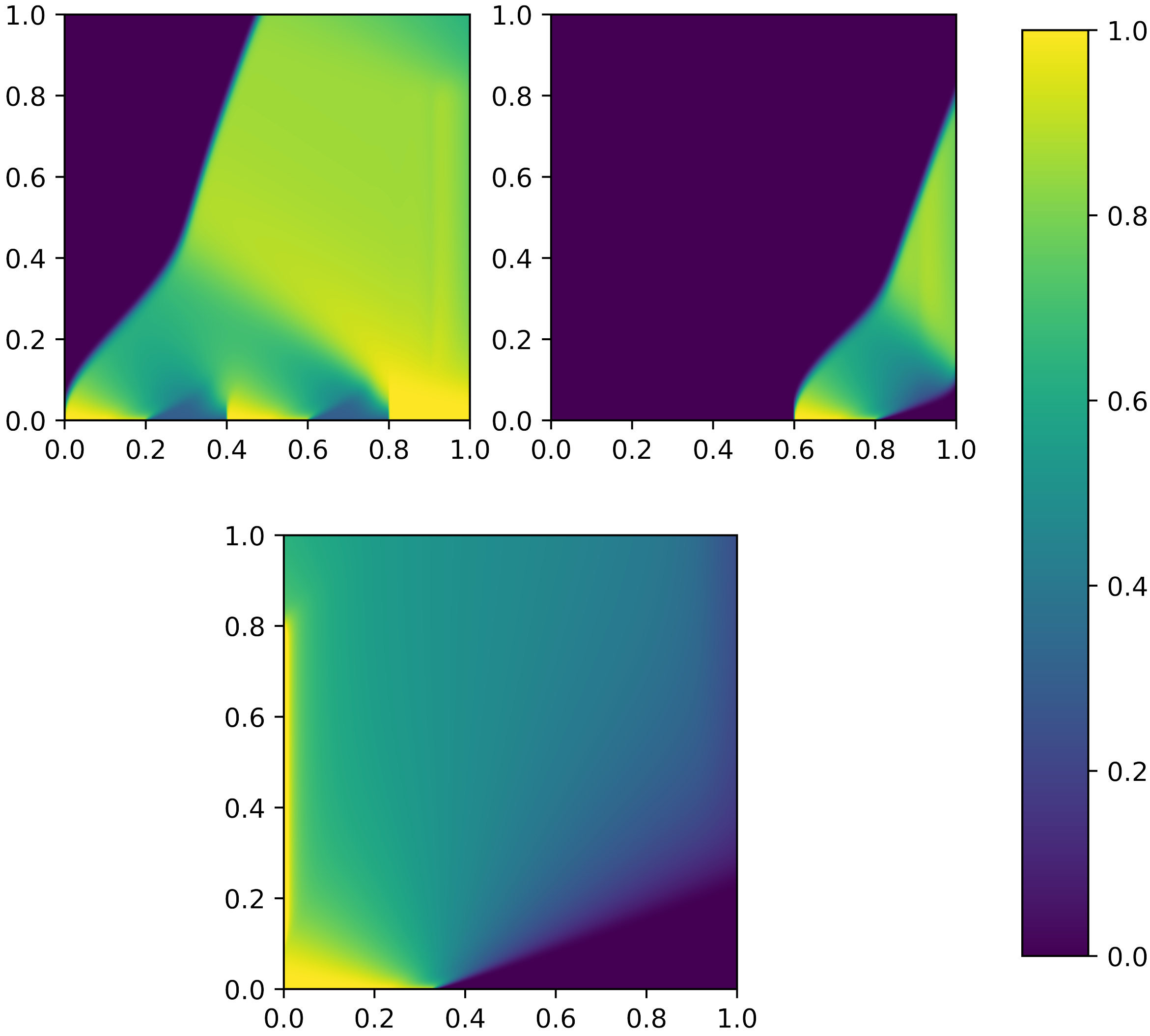


Figure 5.1: Traffic density dynamic on incoming arcs (top) and the outgoing arc (bottom). Parameters: $\Delta x = 5 \times 10^{-3}, \Delta t = 4 \times 10^{-4}, R = 25\Delta x$.

are too close, leading to sudden variation in their speed.

5.2.4 Test 3: 1-to-2 junction.

In this last test, we focus on a $1 - to - 2$ junction, a network with one incoming road and two outgoing ones. In this test, we will assume the same hypothesis of the first test. Our goal is to understand and observe the role of the matrix distribution and of the priority rules αs (see formula (5.3)). Indeed, for $2 - to - 1$ junctions (or $n - to - 1$, more in general) the transmission matrix is equivalent to a vector $(1 \quad 1)^T$ and all weights α are equal to 1. In the $1 - to - 2$ scenario (or $1 - to - n$ more generally) it is necessary to fix a distribution rule represented by a vector $(p_{12}, p_{13})^T$ and pririoty rules α_{12}, α_{13}. We propose three different scenarios following the next table:

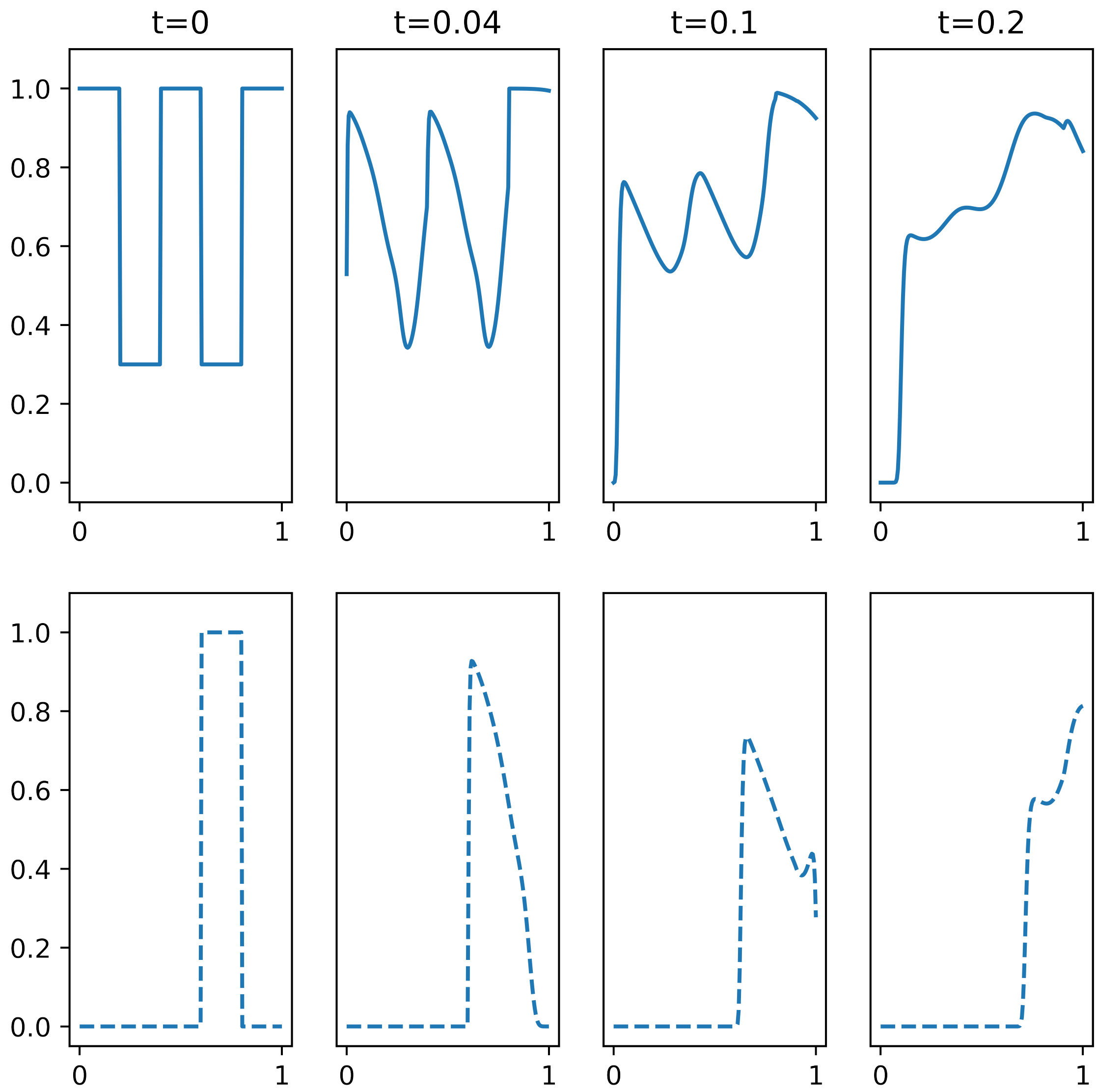


Figure 5.2: Traffic density dynamic on first (top, continuous line) and second (bottom, dashed line) incoming arcs at different times. From left to right: $t = 0, 0.04, 0.1, 0.2$.
Parameters: $\Delta x = 5 \times 10^{-3}, \Delta t = 4 \times 10^{-4}, R = 25\Delta x$, time horizon $T = 1$ and arcs' length $L = 1$.

Priority rules	*Test 3.1*	*Test 3.2*	*Test 3.3*
$\begin{pmatrix} \alpha_{12} \\ \alpha_{13} \end{pmatrix}$	$\begin{pmatrix} p_{12} \\ p_{13} \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

(5.8)

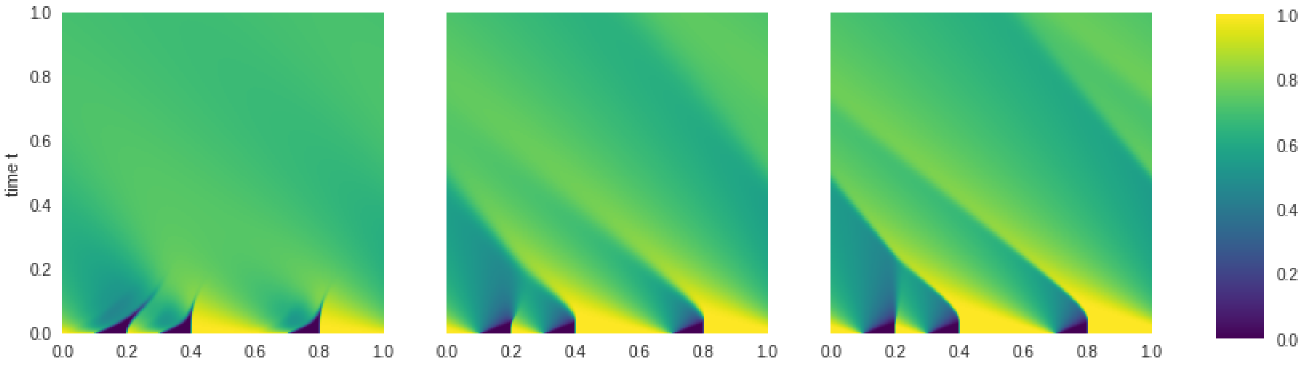


Figure 5.3: Backpropagation of congestion on a cycle. From left to right: $R = 100\Delta x, 25\Delta x, 10\Delta x$. Parameters: $\Delta x = 2 \times 10^{-3}, \Delta t = 10^{-4}, v_{des} = 5, T = L = 1$.

where

$$\begin{pmatrix} p_{12} \\ p_{13} \end{pmatrix} = \begin{pmatrix} 0.6\chi_{[0,\frac{1}{2}]} + 0.4\chi_{(\frac{1}{2},1]} \\ 0.4\chi_{[0,\frac{1}{2}]} + 0.6\chi_{(\frac{1}{2},1]} \end{pmatrix}$$

This test is useful to stress the role of parameters in the model like the distribution matrix P and the priority rules α. These parameters are important since they are able to catch many phenomena which appears in traffic flow systems. In all the simulation we observe the distribution matrix affects the distribution of mass crossing the junction. These effects are clearly observed on the outgoing arc e_2 where at time $t = 0.5$ we can observe the change in the inflowing mass.

The most interesting consequences follow from the parameters αs; indeed, under our assumption, the junction is a natural point of discontinuity of the velocity field. These affect the density as observed in Test 3.1 and 3.3 where a congestion borns at the beginning of e_3; after crossing the junction, the interaction changes rapidly due to the immediate change of perceived mass.

Test 3.2 does not show the creation of congestion on e_3 thanks to the choice of α. Indeed, the interaction is continuous on arcs e_1 and e_3. On the other hand, crossing into e_2, the interaction has a negative jump which means an high speed, hence any congestion is created.

This test suggests that the choice of $\alpha's$ as constant or, generally, functions in BV([0,T]) is a restrictive hypothesis. A possible solution, which is not covered by the theory developed in this thesis, is the dependence of α on the distribution μ_t. This choice would allow a richer description but the lipschitzianity is no more guaranteed, even if it would allow us to describe the possible strategies adopted by drivers depending on congestions and traffic flows.

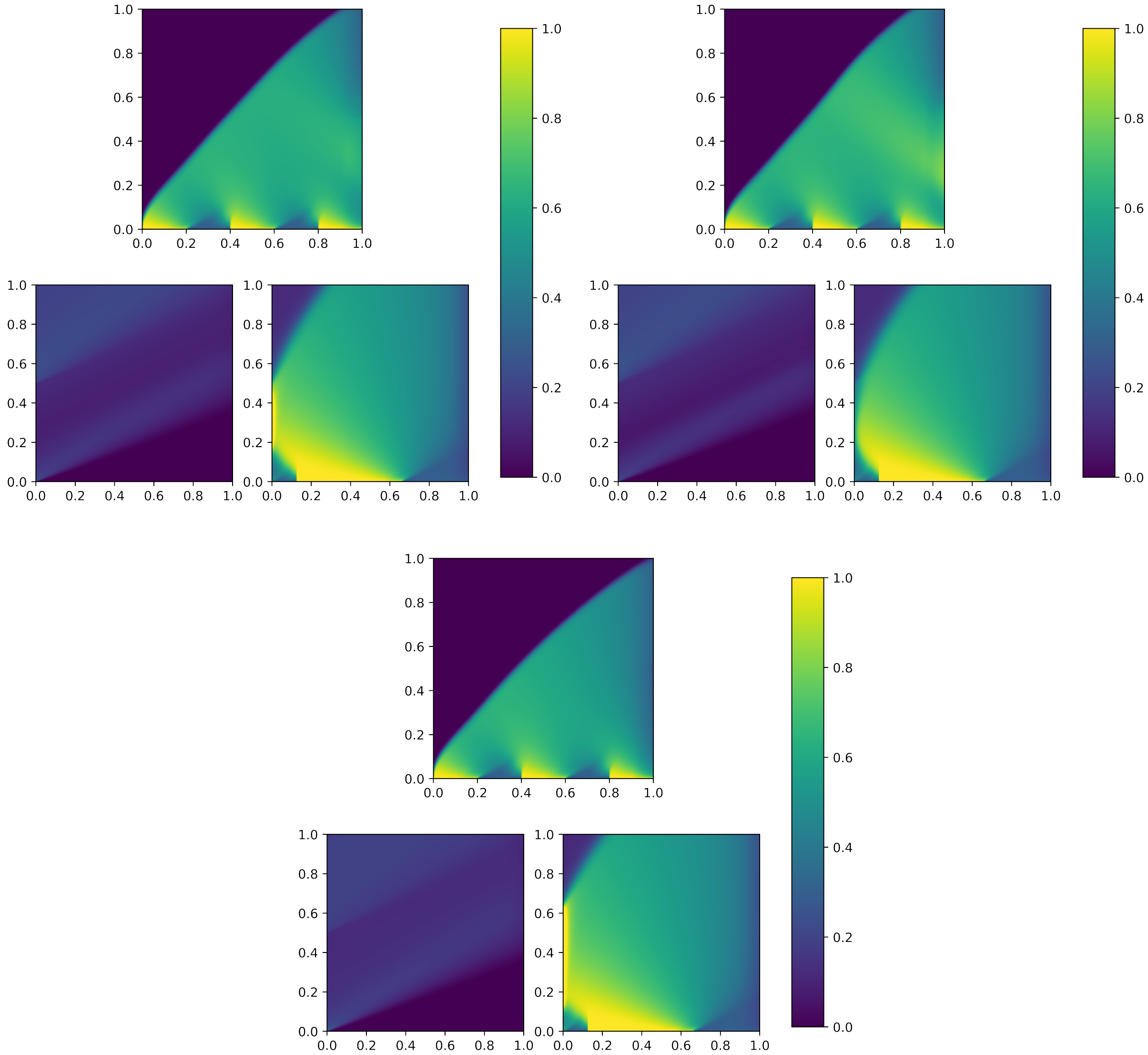


Figure 5.4: Test 3.1 (top-left), 3.2 (top-right) and 3.3 (bottom-center), based on 5.2.4, with an incoming arc and two outgoing ones.
Parameters: $\Delta x = 5 \times 10^{-3}, \Delta t = 4 \times 10^{-4}, R = 25\Delta x$, time horizon $T = 1$ and arcs' lenght $L = 1$.

5.3 Monte Carlo method on networks

Up to this point, we have based our analysis on pre-existing numerical methods for PDEs, assuming that the initial condition is approximated by an absolutely continuous function. These are not always the proper choice. In real applications, it would be based on the nature of available data. For example, if the traffic flow on an highway system and data on flux are provided by sensors we would be oriented towards numerical schemes for PDEs; otherwise, if data of the trajectories of a group of cars, such as GPS data, are available, then it is reasonable to use ODE-oriented schemes.

In this section, we propose methods useful to handle with deterministic data on the trajectories of vehicles. These data are necessary for calibration analysis even if their implementation is complex and

numerically expensive for problems with many drivers and large networks.
Assume that the initial distribution is an atomic measures, i.e.

$$\mu_0 = \sum_{i=1}^{N} \delta_{x_i},$$

where $x_i \in \Gamma$ are the initial positions of $N < +\infty$ vehicles. Then, the dynamic of vehicles is described by the superposition formula (4.35). Hence the distribution at time $t > 0$ is given by

$$\mu_t = \sum_i \sum_{\gamma \in \mathcal{A}(x_i)} \delta_{(\Phi_t^\gamma(x_i,0),t)} p_\gamma(x_i, 0). \tag{5.9}$$

The previous formula highlights how the main ingredient of a particles-based scheme for the transport equation on networks is a scheme for the trajectories and the path weights p_γ.
Denoted with x_i^n the position of $i-$th driver at time t_n and m_i^n its mass, its speed is determined as

$$V_i^n = V[\mu_{t_n}](x_i^n) = \max\left\{0, v_d(x_n^i) - \sum_{j: x_j^n \in \mathcal{D}(x_i^n)} \alpha_{ee'} K(d_\Gamma(x_i^n, x_j^n)) m_j^n\right\},$$

where e, e' are, respectively, the arcs which x_i^n, x_j^n belong to. The position at time $n+1$ can be determined by any kind of scheme for ODEs; for example, in case of first order Euler scheme we have:

$$\tilde{x}_i^{n+1} = x_i^n + \Delta t V_i^n; \tag{5.10}$$

if x_i^n does not cross the junction, i.e. $\tilde{x}_i^{n+1} \leq L_e$, we set $x_i^{n+1} = \tilde{x}_i^{n+1}$ and $m_i^{n+1} = m_i^n$; otherwise, we need to split the particle on each possible arc by the transmission matrix P. Denoted with e the arc which x_i^n belongs to and with $e_1, \dots, e_k$ the possible destination, we create new particles $x_{i_1}, \dots, x_{i_k}$ with mass $m_{i_1}, \dots, m_{i_k}$ such that

$$\begin{array}{lll} m_{i_j}^l = m_i^l, & x_{i_j}^l = x_i^l, & l = 0, \dots, n, \\ m_{i_j}^{n+1} = p_{ee_j} m_i^n, & x_{i_j}^{n+1} = \tilde{x}_i^{n+1} - L_e, & j = 1, \dots, k \end{array} \tag{5.11}$$

Even if this method is coherent with (4.35), it is not suitable for large networks due to the high computational costs for large time, proportional to the increasing number of particles.
For example if our network is given by a binary tree with $M + 1$ layers, with $M \in \mathbb{N}$, and the initial distribution is given by N particles over the first arc then for big times we would have a distribution with at most $N * 2^M$.
This obstacle could be solved by a "Monte Carlo" approach. We can observe that, chosen an arc

$e \in \mathcal{E}$, the respective row of the transmission matrix P is a probability distribution.

Let $C_e : \Omega \to \mathcal{E}$ be a discrete random variable with probability distribution given $P_e = (p_{ee'})_{e' \in \mathcal{E}}$ which defines the crossing from arc e to $e' = C_e(\omega)$; instead of splitting the mass of particles crossing the junction via the transmission matrix rows, we deterministically transport the particles to the roads determined by samples of C_e. A single simulation built with this scheme produce a distribution $\mu = \mu(\omega)$ which is not solution of the transport equation on Γ with transmission matrix P but, if $\mu_n(\omega)$ is properly defined, then the solution of the original problem would be given by the empirical average

$$\bar{\mu}_n = \sum_{i=1}^{N_{samples}} \mu_n(\omega_i), \tag{5.12}$$

where $N_{samples} > 0$. In this way the computational cost would be proportional to $N * N_{samples}$.

From a mathematical point of view, given the exact solution $\mu \in C([0,T]; \mathcal{P}_1(\Gamma))$, we define the samples $\mu(\omega) \in C([0,T]; \mathcal{P}_1(\Gamma))$ as solution of

$$\begin{cases} \partial_t \mu(\omega) + \nabla \cdot (v[\mu]\mu(\omega)) = 0, \\ \mu_{t=0}(\omega) = \mu_0 \\ \mu^e_{x=V}(\omega) = \sum_{e' \in Inc(V)} \delta_{C_{e'}(\omega)e} \mu^{e'}_{x=V}(\omega), \end{cases} \tag{5.13}$$

where δ_{ij} denotes the Kronecker symbol.

The samples $\mu(\omega)$ are solution of the transport equation with velocity field determined by the exact solution μ and by the transmission matrix induced by the random variable $(C_e)_{e \in \mathcal{E}}$.

Hence the empirical average $\bar{\mu} = \mathbb{E}[\mu(\omega)]$ satisfies the conservation equation

$$\partial_t \bar{\mu}_t + \nabla \cdot (v[\mu_t]\bar{\mu}_t) = 0,$$

and the initial condition, since

$$\bar{\mu}_{t=0} = \sum_i \frac{1}{N_{samples}} \mu_{t=0}(\omega_i) = \sum_i \frac{1}{N_{samples}} \mu_0 = \mu_0.$$

Moreover, for any vertex V and arc $e \in \mathrm{Out}(V)$ we have

$$\bar{\mu}^e_{x=V} = \sum_i^{N_{samples}} \frac{1}{N_{samples}} \mu^e_{x=V}(\omega_i) = \sum_{e' \in Inc(V)} \sum_i^{N_{samples}} \frac{1}{N_{samples}} \delta_{C_{e'}(\omega_i)e} \mu^{e'}_{x=V}(\omega_i).$$

Thanks to the law of Large Numbers, it follows that the matrix defined by C_e converges to P as $N_{samples}$ grows. Thus, by the uniqueness of the solution, $\bar{\mu}$ converges to the exact solution.

We can use this formal argument to build a Monte Carlo type scheme even if we do not know the exact solution which is necessary to compute the velocity field. This obstacle can be easily solved adopting an explicit discrete in time scheme and approximating the exact solution with the empirical average. Since $\bar{\mu}$ converges to μ for $N_{samples} \to +\infty$ and $\bar{\mu}_{t=0} = \mu_0$, we can define our Monte Carlo scheme as follows:

Step 1: define a time grid with step Δt and a number of samples $N_{samples}$;

Step 2: define the topological order of network Γ and k=0;

Step 3: while $k\Delta t < T$:

$$\begin{cases} \mu_{k+1} = Update(\mu_k^i, \bar{\mu}_k), & \text{for} 1 \leq i \leq N_{samples}, \\ \bar{\mu}_{k+1} = \frac{1}{N_{samples}} \sum_i \mu_{k+1}^i, & k \in \mathbb{N}. \end{cases}$$

Chapter 6

Optimal control problems on networks

In this last chapter, we focus on a class of optimal control problems for measure-valued nonlinear transport equations describing traffic flow problems on networks. The objective is to minimise/maximise macroscopic quantities, such as traffic volume or average speed, controlling few agents, for example smart traffic lights and automated cars. The measure theoretic approach allows to study in a same setting local and nonlocal drivers interactions and to consider the control variables as additional measures interacting with the drivers distribution.
As in [1, 56], we show that a small number of external agents can improve the global behavior of the population and, indeed, the typical examples of control variables we consider are smart traffic lights and automated cars. Since the external distribution is described by a measure evolving according to an appropriate dynamics, other control variables, such as information about the behavior of the traffic on the global network, can be considered.

We also discuss a gradient descent adjoint-based optimization method, obtained by deriving first-order optimality conditions for the control problem, and we provide some numerical experiments in the case of smart traffic lights for a 2-1 junction.
The chapter is organized as follows: in Section 6.1 we introduce the control problem from a theoretical point of view: network structure, transport equation and cost functional; Section 6.2 is devoted to two examples of control problem: traffic lights and self-driving cars as controls for vehicular traffic, while Section 6.3 focuses on numerical analysis for these problems: description and properties of the chosen scheme and numerical tests on some case studies.

6.1 Problem Formulation and theoretical setting

In this section we recall the main components of the traffic flow model: the dynamics of drivers motion (velocity, interaction with other drivers, influence of the structural components) and the control problem which has to be solved in order optimize the traffic flow on the network.

6.1.1 Driver motion

As described in the previous chapters, we use a nonlinear transport equation to describe drivers' motion. Mathematically, it will act as a constraint for drivers' distribution μ. The components of the system are the differential equations governing the evolution of the traffic inside the arcs and the transition conditions at the vertices regulating the distribution of the traffic flow at the junctions. It is important to remark that the velocity term is nonlocal since drivers usually have a local knowledge of the traffic distribution in a visual area in front of them; moreover they may have a global knowledge of the traffic distribution on the entire network thanks to appropriate navigation equipments.

We prescribe the initial mass distribution over Γ

$$\mu_0 = \sum_{j \in J} \mu_0^j \in \mathcal{M}^+(\Gamma),$$

where μ_0^j is restriction of μ_0 to e_j, and the incoming traffic measure at the source nodes

$$\sigma_0 = \sum_{V_i \in \mathcal{S}} \sigma_0^i, \quad \sigma_0^i \in \mathcal{M}^+([0,T]),$$

where σ_0^i is the restriction of σ_0 to V_i, representing the flow of cars entering in the road network at the vertex V_i. Then, the constraint on μ is given by the following:

$$\begin{cases} \partial_t \mu + \partial_x (v^j[\mu_t, m_t]\mu) = 0 & t \in (0,\, T] \\ \mu_{t=0} = \mu_0 & \\ \mu^j_{V=\pi_j(0)} = \begin{cases} \sum\limits_{k: e_k \in (V_i)} p_{kj}(t) \mu^k_{V=\pi_k(1)} & \text{if } V \in \mathcal{I} \\ \sigma_0^V & \text{if } V \in \mathcal{S} \end{cases} & j = 1, \ldots, |\mathcal{E}|. \end{cases} \tag{6.1}$$

Observe that, for each arc e_j, if the initial vertex $V = \pi_j(0)$ is internal, then the boundary condition at V is given by a measure representing the mass flowing in e_j from the arcs incident to the vertex according to the distribution matrix $P(t)$; if the initial vertex $V = \pi_j(0)$ is incoming traffic vertex, the inflow measure is the prescribed datum σ_0^V. The outflow measure, i.e. the part of the mass leaving the arc from the final vertex $V = \pi_j(1)$, is not given a priori but depends on the evolution of the measure μ inside the arc.

The velocity v depends on the solution μ_t itself, as well as on another distribution $m_t \in \mathcal{M}^+(\Gamma)$, representing external forces acting on the drivers such as traffic lights and autonomous vehicles (more details will be given in the next section where we consider specific models). Analogously to the previous chapters, we assume:

(H1) v is non-negative and bounded by $V_{max} > 0$;

(H2) v is Lipschitz with respect to the state variable, i.e. there exists $L > 0$ such that $\forall x, y \in e_j$, m_i, $\mu_i \in \mathcal{M}^+(\Gamma)$, for $i = 1, 2$

$$|v^j[\mu_1, m_1](x) - v^j[\mu_2, m_2](y)| \leq L(|x - y| + \|m_1 - m_2\|_{BL}^* + \|\mu_1 - \mu_2\|_{BL}^*);$$

Under these hypotheses, the solution clearly exists for every given $m \in \mathcal{M}^+([0,T]\times\Gamma)$. In this chapter, we assume a velocity field of the form

$$v[\mu, m](x) := \max\{v_f(x) - v_I[\mu](x) - v_E[m], 0\} \tag{6.2}$$

where $v_f : \Gamma \to \mathbb{R}^+$ is the free-flow speed, $v_I[\mu](x)$ is the interaction term described in the previous chapter, while $v_E[m]$ is an interaction term with an external distribution m.

6.1.2 Mobility optimization

We introduce a class of optimization problems on networks involving the distribution μ, given by the solution of (6.1), the external distribution m and a control variable u which has to be designed in order to minimize/maximize a given objective functional, or loss function.

We assume that the set of the admissible controls is given by a Banach space $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$. We also denote by $\mathcal{M}_M^+(\Gamma_T)$ the set of the measures $\mu \in \mathcal{M}^+(\Gamma_T)$ such that $\|\mu\|_{BL}^* \leq M$. Then the state space of the control problem is given by the space $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ where

$$\mathcal{X} = \mathcal{M}_M^+(\Gamma_T) \times \mathcal{M}_M^+(\Gamma_T) \times \mathcal{U},$$
$$\|\cdot\|_{\mathcal{X}} = \|\cdot\|_{BL}^* + \|\cdot\|_{BL}^* + \|\cdot\|_{\mathcal{U}}.$$

For a given initial distribution $\mu_0 \in \mathcal{M}^+(\Gamma)$ and an incoming traffic distribution $\sigma_0 \in \mathcal{M}^+([0,T])$, we consider the optimization problem

$$\begin{cases} \min\{J(\mu, m, u) : (\mu, m, u) \in \mathcal{X}\}, \\ \text{ subject to the state equation (6.1).} \end{cases} \tag{6.3}$$

It is convenient to rewrite the previous minimization problem in the following equivalent form

$$\min\{J(\mu, m, u) + 1_A(\mu, m, u) : (m, \mu, u) \in \mathcal{X}\}, \tag{6.4}$$

where $A := \{(\mu, m, u) \in \mathcal{X};\, m \,\text{solves}\,(6.1)\}$ and 1_A is the indicator function of the set A defined as

$$1_A(x) := \begin{cases} 0, & x \in A, \\ +\infty & \text{otherwise.} \end{cases}$$

A straightforward application of the direct method in Calculus of Variations gives the following existence result for the minima of (6.4).

Theorem 6.1. *Assume that*

- $J : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *is bounded from below;*
- J *is lower semicontinuous in* $\mathcal{X}$*, i.e. for any* $(\mu_n, m_n, u_n) \subset \mathcal{X}$ *such that* $(m_n, \mu_n, u_n) \to (m, \mu, u)$*, it holds* $J(m, \mu, u) \leq \liminf_{n\to\infty} J(m_n, \mu_n, u_n)$*;*
- *the set* A *is closed under the topology induced by* $\|\cdot\|_{\mathcal{X}}$*.*

Then the minimization problem (6.4) *has a solution.*

Even if the first two hypotheses are tautological for the existence of minima, the closure of A, w.r.t. $\|\cdot\|_{\mathcal{X}}$, is extremely important in our framework and it will be stressed in the next section.
A typical example of functional to be minimized is of the form

$$J(\mu, m, u) := -\int_0^T \int_\Gamma v[\mu_t, m_t] d\mu_t(y) dt + \int_{\Gamma\times[0,T]} f(x, t, u) d\mu_t(x) dt, \tag{6.5}$$

where the first term in (6.5) represents the mean velocity on the network, while the second one is a feedback term which depends on the choice of f. For example, if $f(t, x, u) = \chi_B(x)$, where $B \subset \Gamma$ is closed, the functional minimizes the amount of mass μ_t in a closed region B during the time interval $[0, T]$. Another interesting class of control problems are minimum time control ones introduced, in a measure theoretic setting, in [20, 21].

6.2 Smart traffic controls

This section is devoted to applications of the abstract setting previously described with the discussion of two significative problems in traffic flow optimization which have an increasing interest in the last years [45, 39, 55, 56] such as the optimization of traffic lights setting in order to improve the circulation on the road networks or the impact of autonomous car over the global traffic flow.

For both these models we assume that the control variable u influences the traffic flow distribution μ only by means of an external distribution $m = m[u]$. Hence the functional to be minimized in (6.4) is of the form $J(\mu, u)$ with μ subject to (6.1) and m determined by another dynamical system for a given initial configuration m_0.

6.2.1 Smart traffic lights

An important element of a road network model is given by *traffic lights*: they influence the behavior of the drivers near the junction and can be used as an external control to regulate the traffic flow. To model a traffic light, we follow the approach in [40]. Relying on the measure-theoretic setting, we describe a traffic light as a measure $\theta \in \mathcal{M}^+(\Gamma_T)$, which is a Dirac measure in space and a density with bounded variation in time.

We assume that there is at most one traffic light for each road and that it is closed to the terminal vertex $V \in \mathcal{V}$ of the arc e_j. Since the position is fixed a priori while the activity changes in time, a traffic light can be represented, with an abuse of notation, as the measure

$$\sum_{j\in(V)} \int_0^T u_j(t)\delta_V(y)dt, \tag{6.6}$$

where $u_j \in BV([0,T], \{0,1\})$ is a function representing the state of the traffic light: $u_j(t) = 1$ if the light is red, $u_j(t) = 0$ if green (for simplicity, we do not consider a yellow phase since the corresponding driver reaction is strongly influenced by drivers' culture).

Concerning the light phases, in order to exclude unrealistic scattering phenomena, we fix two positive times $T^R, T^G > 0$ and we assume that the red phase cannot last more then T^R and, analogously, the green phase must last at least T^G to guarantee a proper traffic flow. Hence denoted by $\tau_1, \tau_2 \in [0,T]$ two consecutive switching times of the traffic light on the arc e_j (corresponding to jump discontinuities of u_j), we assume that

$$\begin{aligned} &\text{if } u_j(\tau_1^+) = 1, \text{ then } |\tau_1 - \tau_2| < T^R, \\ &\text{if } u_j(\tau_1^+) = 0, \text{ then } |\tau_1 - \tau_2| > T^G. \end{aligned} \tag{6.7}$$

Moreover we assume that a traffic light can be green only for one of the incoming roads in a junction, i.e.

$$\begin{aligned} &\textstyle\sum_{j\in(V)} u_j + 1 = N \\ &T^R \geq (N-1)T^G \end{aligned} \tag{6.8}$$

where $N = \#(\mathrm{Inc}(V))$.

Denote by $\mathcal{F} \subset \mathcal{E}$ the set of the arcs containing a traffic light. Recalling (6.6), we consider the measure $dm(x,t) = \sum_{j=1}^{|\mathcal{E}|} u_j(t) dm^j(x,t)$ on Γ_T where $dm^j(x,t) \equiv 0$ if $e_j \notin \mathcal{U}$ and $dm^j(x,t) = \delta_V(x)dt$ if $e_j \in \mathcal{U} \cap \mathrm{Inc}(V_i)$. The term u_j, the phase duration of the traffic light on the road e_j, can be interpreted as the control variable. The set of admissible controls is given by

$$\mathcal{U} = \{u = \{u_j\}_{j=1,\dots,|\mathcal{U}|} : u_j \in BV([0,T], \{0,1\}) \text{ and satisfies (6.7), (6.8) } \} \tag{6.9}$$

To describe the interaction of the drivers with the traffic lights, we define an external velocity term $v_E[m]$ in (6.2). Fixed an arc $e_j \in \mathcal{U} \cap \mathrm{Inc}(V)$, then the restriction of $v_E[\mu]$ to the arc e_j is given by

$$v_E^j[m](x) := \int_\Gamma H(x,y) dm_t(y) = u_j(t) H(x,V) \delta_{e_j}(x).$$

We assume that the interaction kernel H is given by

$$H(x,y) = \begin{cases} v_f \max\left\{\left(1 - \frac{d_\Gamma(x,y)}{R}\right), 0\right\}, & \text{if } x \to y,\ d_\Gamma(x,y) \leq R, \\ 0 & \text{otherwise,} \end{cases} \tag{6.10}$$

where v_f is the desired velocity and $R \leq L_0$ is the visibility radius and L_0 the minimal arc length. The driver interaction with the traffic light, tuned by the signal u_j, occurs only if the driver is sufficiently close to the junction and becomes stronger getting closer.

We need to show that the chosen set of control (6.9) satisfies the hypotheses for the existence of minima for $\mathcal{X} = \mathcal{M}_M^+(\Gamma_T) \times \mathcal{M}_M^+(\Gamma_T) \times \mathcal{U}$.

Lemma 6.1. *The set of positive measures with bounded mass $\mathcal{M}_M^+(\Gamma_T)$ is compact with respect to $\|\cdot\|_{BL}^*$.*

Proof. Assume without loss of generality that $M = 1$. It is well known that for $m \in \mathcal{M}_M^+(\Gamma_T)$, $|\mu|_{TV} =$

$m(\Gamma_T) \leq 1$.

By Banach-Alaoglu Theorem it follows the compactness with respect to the weak*-convergence, which implies the same property with respect to the $\|\cdot\|_{BL}^*$ convergence. $\square$

Lemma 6.2. *The set $\mathcal{U}$ defined in* (6.9) *is compact in* $(BV^{|\mathcal{E}|}([0,T]), \|\cdot\|_{L^1})$.

Proof. Since (6.8) is just a condition which defines the dependence among the components of $u \in \mathcal{U}$, we prove the compactness of

$$\mathcal{U} = \{u \in BV([0,T], \{0,1\}) \,\text{and}\, u \,\text{satisfies (6.7),}\}.$$

Let $(u_n)_{n\in} \subset \mathcal{U}$. Denote by τ_i^n the switching times of u_n. By (6.7), for every two consecutive switching times $\tau_k^n, \tau_{k+1}^n \in [0,T]$, if $u^n(\tau_k^n) = 1$, then

$$|\tau_k^n - \tau_{k+1}^n| < T^R,$$

otherwise,

$$|\tau_k^n - \tau_{k+1}^n| > T^G.$$

Since $u_n(t) \in \{0,1\}$, we can assume that there exists a subsequence, still denoted by u_n, such that either $u_n(0) = 1$ or $u_n(0) = 0$ for every $n \in$. Assume now that, w.l.o.g., $u_n(0) = 1$ for every $n \in$ and denote by I_n the set of switching times of u_n. It follows that

$$\frac{T}{T^R} \leq \#(I_n) \leq \frac{T}{T^G}.$$

As before, we can assume, w.l.o.g., that that there exists $N \in$ such that $\#(I_n) = N$ for all $n \in$. Since $I_n \subset [0,T]$, applying the Cantor diagonal procedure, it follows that there exists a subsequence $(I_{n_k})_{k\in}$ such that $\tau_i^{n_k} \to \tau_i$ for $i = 1, \ldots, N$. In this way, we define a candidate u as limit for the subsequence u_{n_k} from the switching times set $\{\tau_1, \ldots, \tau_N\}$ and $u(0) = 1$. To conclude, we only need to show that $u_{n_k} \to u$ in L^1. By construction,

$$\|u_{n_k} - u\|_{L^1} = \sum_{i=1}^{N} |\tau_i^{n_k} - \tau_i| \leq N \sup_{i=1,\ldots,N} |\tau_i^{n_k} - \tau_i| \to_{k\to\infty} 0$$

$\square$

Lemma 6.3. *Assume $\mathcal{X} = \mathcal{M}_M^+(\Gamma_T) \times \mathcal{M}_M^+(\Gamma_T) \times \mathcal{U}$, where $\mathcal{U}$ satisfies the hypothesis of Lemma 6.2. The set A is closed under the topology induced by $\|\cdot\|_{\mathcal{X}}$.*

Proof. In this case, the distirbution μ has no role since it depends exclusively on u. Hence, we reduce on $\mathcal{X} = \mathcal{M}^+(\Gamma_T) \times \mathcal{U}$, where $\mathcal{U}$ defined by (6.9).
Let $(m_n, u_n)_{n\in} \subset A$ such that $(m_n, u_n) \to (m, u)$ with respect the norm $\|\cdot\|_{BL}^* + \|\cdot\|_{L^1}$.
The closure on the first component derives from the proof of Lemma 4.1 in [8] and the results in Chapter 4.
Instead, the closure on the second component derives from the compactness of $\mathcal{U}$. Indeed, there exists a subsequence $(u_{n_k})_{k\in}$ which converges to $\tilde{u} \in \mathcal{U}$, but it also converges to u by assumption. Then, it follows that $u = \tilde{u} \in \mathcal{U}$. □

6.2.2 Regulating traffic flow by means of autonomous cars

In this second application, we aim to optimize the traffic flow by exploiting another distribution of cars, possibly given by autonomous vehicles, of which we can control the velocity. Indeed some experiments (see [56]) have shown that it is possible to avoid stop-and-go phenomena regulating the interactions among drivers by means of external agents (autonomous vehicles, traffic light, signaling panels,etc.). The approach in this section is inspired to [8] where the authors present an optimization problem for a transport equation in the euclidean space with the control represented by a second distribution μ evolving according to another transport equation.
The dynamics of the autonomous cars is similar to the ones of rest of the driver, with the difference that it can be controlled in order to minimize the objective functional. Hence for a given initial distribution m_0 (typically $m_0 = \sum_{V_i \in \Gamma_a} \delta_{V_i}$ for some finite set $\Gamma_a \subset \Gamma$), the measure $\mu \in \Gamma_T$ representing the distribution of the fleet of the autonomous car satisfies the nonlinear transport equation

$$\begin{cases} \partial_t m^j + \partial_x(u \cdot v^j[\mu_t, m_t] m^j) = 0 & x \in e_j,\, t \in (0,T],\, j = 1, \dots, |\mathcal{E}| \\ m^j_{t=0} = m^j_0 & x \in e_j, j = 1, \dots, |\mathcal{E}| \\ m^j_{V=\pi_j(0)} = \begin{cases} \sum\limits_{k: e_k \in (V)} q_{kj}(t) m^k_{V=\pi_k(1)} & \text{if } V \in \mathcal{I} \\ 0 & \text{if } V \in \mathcal{S}, \end{cases} & j = 1, \dots, |\mathcal{E}| \end{cases} \tag{6.11}$$

We assume that the velocity fields $v[\mu_t, m_t]$ in (6.11) is the same of problem (6.1) and it is defined as in (6.2). Moreover we assume that the drivers are not able to discern between not-autonomous and autonomous cars and therefore $v_I = v_E$. Hence we can rewrite the velocity field (6.2) as

$$v[\eta] = \max\{0, v_f - v_I[\eta]\},$$

where, in our setting, $\eta = m + \mu$.

On the other side, since we want to regulate the velocity of the distribution μ we add a control term u and we assume that the control set is given by

$$\mathcal{U} = \mathrm{Lip}_L(\Gamma_T, [0, 1]), \tag{6.12}$$

i.e. the set of Lipschitz functions from $\Gamma \times [0, T]$ to $[0, 1]$ with Lipschitz constant $L > 0$. In this way, if $v[\mu_t, m_t]$ satisfies the assumptions (H1-H2), then also $u \cdot v[\mu_t, m_t]$ satisfies the same assumptions and therefore system (6.11), given $(\mu_t)_{t\in[0,T]}$, admits a unique measure-valued solution. Moreover, since we require that $u(x, t) \in [0, 1]$, then the autonomous cars can only slow the traffic distribution. Observe that system (6.11) also differs from (6.1) for the distribution matrix $Q = (q_{kj}(t))_{k,j=1}^{|\mathcal{E}|}$ at the junctions. Actually it is reasonable to assume that Q does not coincide with the distribution matrix P since the autonomous cars can behave differently from the rest of the drivers at the junctions. We assume that the matrix Q satisfies the assumptions of P. Hence, the existence of solutions (μ, m) of the coupled transport system follows by a standard fixed point argument.

Existence of a solution (μ, m) to the coupled transport system (6.1)-(6.11) can be proved by a fixed point argument.

Given $m \in C([0, T], \mathcal{M}^+(\Gamma))$, consider the map

$$\Phi_1 : C([0, T], \mathcal{M}^+(\Gamma)) \to C([0, T], \mathcal{M}^+(\Gamma))$$

which associates with m the unique solution of (6.11). Similarly, given $\mu \in C([0, T], \mathcal{M}^+(\Gamma))$, define a map

$$\Phi_2 : C([0, T], \mathcal{M}^+(\Gamma)) \to C([0, T], \mathcal{M}^+(\Gamma))$$

which associates with μ the solution $\Phi_2(\mu)$ of (6.1). Hence, defined a map $\Phi := (\Phi_1, \Phi_2)$, the solution of the coupled system (6.1)-(6.11) is given by a fixed point of Φ. By an argument similar to the one already used in [23, 24] for analogous results, it is possible to prove that Φ is a contraction and therefore existence of a unique solution to the system (6.1)-(6.11) is obtained.

We conclude this section with the following Lemma:

Lemma 6.4. *Assume $\mathcal{X} = \mathcal{M}^+_M(\Gamma_T) \times \mathcal{M}^+_M(\Gamma_T) \times \mathcal{U}$, where $\mathcal{U}$ is defined by* (6.12)*. The set A is closed under the topology induced by $\|\cdot\|_{\mathcal{X}}$.*

This result can be proven as in the proof of Lemma 6.3, using the Ascoli-Arzelà Theorem instead of Lemma 6.2.

Proof. It follows adopting the argument in the previous proof, for $\mathcal{X} = \mathcal{M}^+_M(\Gamma_T) \times \mathcal{M}^+_M(\Gamma_T) \times \mathcal{U}$ endowed with the norm $\|\cdot\|^*_{BL} + \|\cdot\|^*_{BL} + \|\cdot\|_\infty$. □

6.3 Numerical solution via optimality conditions

In this section we formally derive first-order optimality conditions for the optimization problem (6.3) in the case of a traffic light for a 2-1 junction. Then we build a gradient descent adjoint-based method to approximate the solution of the discretized optimality system and present some numerical experiments.

6.3.1 Optimality conditions

We consider a network Γ composed of a junction with two roads converging in a single one, namely we have $\mathcal{E} = \{e_1, e_2, e_3\}$, $\mathcal{V} = \{V_0, V_1, V_2, V_3\}$ and $= \{V_0\}$, $\mathcal{S} = \{V_1, V_2\}$, $\mathcal{W} = \{V_3\}$, $\mathrm{Inc}(V_0) = \{e_1, e_2\}$ and $\mathrm{Out}(V_0) = \{e_3\}$, as shown in Figure 6.1.

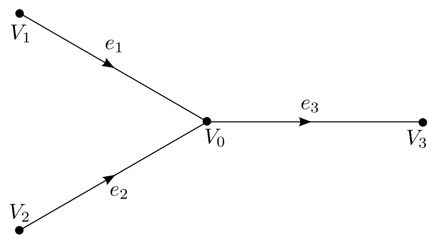


Figure 6.1: Example of 2-1 junction

To simplify the presentation, we neglect the drivers' interaction term, since the computation in the general case is similar but more involved. We place a traffic light at V_0 in order to maximize the average speed on the network. In this setting a single control $u \in BV([0,T],\{0,1\})$ is enough to describe the system, indeed we define edge-wise the velocity v by

$$v^1[u](x,t) = \max\{v^1_f(x) - u(t)H(x,V_0), 0\}\,,$$

$$v^2[u](x,t) = \max\{v^2_f(x) - (1-u(t))H(x,V_0), 0\}\,,$$

$$v^3(x,t) = v^3_f(x)\,,$$

where for $j = 1,2,3$, v^j_f is the free flow speed on e_j and H is defined as in (6.10). It is important to stress that, in general, our modeling choice about traffic lights, combined with our choice for nonlocal speed, does not guarantees that cars stop at the junction since it depends on the proper choice of parameters. Indeed, in our case, we need to require that, for $u = 1$, $v_f(V_0) - H(V_0, V_0) \leq 0$ which is satisfied since $H(V_0, V_0) = v_f$ by assumption.

Since the switching of the traffic light is intrinsically a discrete process, we translate the control problem into a finite dimensional setting. More precisely, we consider a vector $s = (s_1, ..., s_S) \in \mathbb{R}^S$, whose components represent the durations of $S-1$ successive switches, where the integer number $S > 1$ is fixed a priori. Then the control $u(t)$ is easily reconstructed from a given value $u(0) = u_0 \in \{0, 1\}$ at initial time and from the switching times $\tau_i = \sum_{k=1}^{i} s_i$ for $i = 1, ..., S$. Defining recursively $u_i = 1 - u_{i-1}$ for $i = 1, ..., S$ and $\tau_0 = 0$ we set (see Figure 6.2)

$$u(t) = u^s(t) = \sum_{i=0}^{S-1} u_i \chi_{[\tau_i, \tau_{i+1})}(t)$$

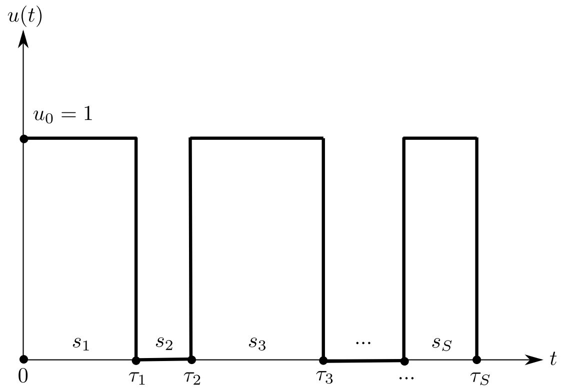


Figure 6.2: Reconstruction of control u from switching durations $s = (s_1, ..., s_S)$

Following this approach we avoid several difficulties. Indeed, $BV([0,T], \{0,1\})$ is not even a vector space and taking admissible variations of a given control or imposing constraints on the switching durations is in practice not easy at all. One could work instead with the convex subset $BV([0,T]; [0,1])$ of $L^2(0,T)$ and look for bang-bang controls. This can prevent unrealistic mixing of mass at the junction, due to the additional yellow phase for the traffic light (intermediate values in $(0,1)$), but chattering phenomena can occur. In our setting we just work in $\mathbb{R}^S$, chattering is not allowed by construction, and we can easily apply variations/constraints to the switching durations being sure that the control always remains in $BV([0,T], \{0,1\})$.

Assuming that the measure μ has a density, i.e. $d\mu = \mu(x,t)dx\,dt$ for some function $\mu : \Gamma \times [0,T] \to \mathbb{R}$, we want to minimize the cost functional

$$J(\mu, u^s) = -\int_0^T \int_\Gamma v[u^s](x,t)\mu(x,t)\,dxdt\,, \tag{6.13}$$

subject to

$$\begin{cases} \partial_t \mu^j + \partial_x(v^j \mu^j) = 0 & \text{in } e_j \times (0,T),\ j=1,2,3 \\ \mu^j(\cdot,0) = \mu_0^j & \text{in } e_j \end{cases} \tag{6.14}$$

We also assume null incoming traffic in the network during the whole evolution, imposing

$$\mu^1_{x=V_1} = 0, \quad \mu^2_{x=V_2} = 0, \quad t \in [0,T], \tag{6.15}$$

and the mass conservation condition at the internal vertex V_0

$$\mu^3_{x=V_0} = \mu^1_{x=V_0} + \mu^2_{x=V_0}. \tag{6.16}$$

We formally apply the method of Lagrange multipliers in order to derive first-order optimality conditions. We define the Lagrangian as

$$\begin{aligned} L(\mu,u^s,\lambda) &:= J(\mu,u^s) + \int_0^T \int_\Gamma (-\partial_t\lambda - v\partial_x\lambda)\mu \, dxdt \\ &+ \int_\Gamma (\lambda(x,T)\mu(x,T) - \lambda(x,0)\mu_0(x))\, dx \\ &+ \sum_{j=1,2,3} \int_0^T (\lambda^j(V_j^E,t)v^j(V_j^E,t)\mu^j(V_j^E,t) - \lambda^j(V_j^I,t)v^j(V_j^I,t)\mu^j(V_j^I,t))\, dt\,, \end{aligned}$$

where V_j^I and V_j^E denote the initial and, respectively, the final vertex of the arc e_j. Observe that the terms involving the Lagrange multiplier λ derive from the weak formulation of the transport equation on Γ.

We evaluate the derivates of the Lagrangian with respect to μ and s (recall that $u = u^s$). We first consider an admissible increment w for μ which preserves the boundary and transition conditions, i.e.

$$w^1(V_1,t) = 0\,, \qquad w^2(V_2,t) = 0\,, \qquad w^3(V_0,t) = w^1(V_0,t) + w^2(V_0,t) \qquad t \in [0,T]\,, \tag{6.17}$$

and we compute

$$\begin{aligned} \langle \partial_\mu L, w\rangle &= \int_0^T \int_\Gamma (-\partial_t\lambda - v\partial_x\lambda - v) w \, dxdt + \int_\Gamma \lambda(x,T)w(x,T)\, dx \\ &+ \int_0^T \sum_{j=1,2,3} (\lambda^j(V_j^E,t)v^j(V_j^E,t)w^j(V_j^E,t) - \lambda^j(V_j^I,t)v^j(V_j^I,t)w^j(V_j^I,t))\, dt\,. \end{aligned} \tag{6.18}$$

Imposing $\langle \partial_\mu L, w\rangle = 0$ for any admissible w, we get the following time-backward advection equation with a source term

$$-\partial_t\lambda^j - v^j\partial_x\lambda^j = v^j \qquad \text{in} \quad e_j \times (0,T),\ j=1,2,3, \tag{6.19}$$

and the final condition

$$\lambda^j(x,T) = 0 \qquad \text{in} \quad e_j,\ j = 1,2,3.$$

Note that for (6.19), V_3 is an inflow vertex where a boundary condition has to be prescribed, while V_1 and V_2 are outflow ones. Writing explicitly the remaining boundary terms in (6.18), we have

$$\int_0^T (\lambda^1 v^1 w^1(V_0,t) - \lambda^1 v^1 w^1(V_1,t) + \lambda^2 v^2 w^2(V_0,t)$$
$$-\lambda^2 v^2 w^2(V_2,t) + \lambda^3 v^3 w^3(V_3,t) - \lambda^3 v^3 w^3(V_0,t))\, dt = 0\,.$$

By taking w compactly supported in a neighborhood of V_3, we get the boundary condition

$$\lambda^3(V_3,t) = 0 \qquad \text{in} \quad [0,T]\,,$$

whereas for w compactly supported in a neighborhood of V_0, recalling (6.17), we get

$$\int_0^T \{(\lambda^1 v^1 - \lambda^3 v^3) w^1(V_0,t) + (\lambda^2 v^2 - \lambda^3 v^3) w^2(V_0,t)\}\, dt = 0\,. \tag{6.20}$$

The mass conservation condition (6.16) can be rewritten as

$$v^3(V_0,t)\mu^3(V_0,t) = v^1(V_0,t)\mu^1(V_0,t) + v^2(V_0,t)\mu^2(V_0,t) \qquad t \in [0,T]\,,$$

since the control law u models a traffic light which bring to halt the speed of the drivers at V_0 in e_1 and, alternatively, in e_2, in such a way that there is mass flow either from e_1 to e_3 or from e_2 to e_3. If $I_1 \subseteq [0,T]$ is an interval where $u(t) = 1$ (red light for e_1), then in this interval the speed $v^1(V_0,t)$ is null and therefore $\mu^1(V_0,t) = 0$ (recall that mass concentration at the vertices is not admitted). Similarly if $u(t) = 0$ for $t \in I_2$ (red light for e_2), we get $\mu^2(V_0,t) = 0$ for $t \in I_2$. An admissible increment, in order to preserve the transition condition for m, has to satisfy the same property and by (6.20) we get

$$\lambda^3(V_0,t)v^3(V_0,t) = \lambda^1(V_0,t)v^1(V_0,t) + \lambda^2(V_0,t)v^2(V_0,t),$$

or, more explicitly,

$$\lambda^1(V_0,t)v^1(V_0,t) = \lambda^3(V_0,t)v^3(V_0,t) \qquad \text{if} \quad t \in \{v^1(V_0,t) \neq 0\}\,,$$

$$\lambda^2(V_0,t)v^2(V_0,t) = \lambda^3(V_0,t)v^3(V_0,t) \qquad \text{if} \quad t \in \{v^2(V_0,t) \neq 0\}\,.$$

We now compute the derivative of L with respect to u^s for an increment $\varphi \in \mathbb{R}^S$

$$\langle \partial_s L, \varphi \rangle = -\int_0^T \int_\Gamma \partial_s v \cdot \varphi (\partial_x \lambda + 1) \mu \, dx dt + \int_0^T \{ \sum_{j=1,2,3} \lambda^j (V_j^E, t) \partial_s v^j (V_j^E, t) \cdot \varphi \, \mu^j (V_j^E, t)$$
$$- \lambda^j (V_j^I, t) \partial_s v^j (V_j^I, t) \cdot \varphi \, \mu^j (V_j^I, t) \} \, dt \, .$$

Recalling (6.15) and since v^3 is independent of u^s, we get

$$\langle \partial_s L, \varphi \rangle = \int_0^T \Big\{ - \int_{e_1} \partial_s v^1 \cdot \varphi (\partial_x \lambda^1 + 1) \mu^1 \, dx - \int_{e_2} \partial_s v^2 \cdot \varphi (\partial_x \lambda^2 + 1) \mu^2 \, dx$$
$$+ \lambda^1 (V_0, t) \partial_s v^1 (V_0, t) \cdot \varphi \, \mu^1 (V_0, t) + \lambda^2 (V_0, t) \partial_s v^2 (V_0, t) \cdot \varphi \, \mu^2 (V_0, t) \Big\} \, dt,$$

where

$$\partial_s v^1 (x, t) \cdot \varphi = -H(x, V_0) \nabla_s u^s (t) \cdot \varphi \, , \qquad \partial_s v^2 (x, t) \cdot \varphi = H(x, V_0) \nabla_s u^s (t) \cdot \varphi$$

and

$$\nabla_s u^s (t) \cdot \varphi = \sum_{i=1}^S (-1)^{u_{i-1}} \delta_{\tau_i} (t) \varphi_i \, .$$

We conclude

$$\langle \partial_s L, \varphi \rangle = \sum_{i=1}^S (-1)^{u_{i-1}} \Big\{ \int_{e_1} H(x, V_0) (\partial_x \lambda^1 (x, \tau_i) + 1) \mu^1 (x, \tau_i) \, dx - \lambda^1 (V_0, \tau_i) H(V_0, V_0) \mu^1 (V_0, \tau_i)$$
$$- \int_{e_2} H(x, V_0) (\partial_x \lambda^2 (x, \tau_i) + 1) \mu^2 (x, \tau_i) \, dx + \lambda^2 (V_0, \tau_i) H(V_0, V_0) \mu^2 (V_0, \tau_i) \Big\} \varphi_i \, .$$

Summarizing, the dual problem for (6.14)-(6.15)-(6.16) is

$$\begin{cases} -\partial_t \lambda^j - v^j \partial_x \lambda^j = v^j & \text{in } e_j \times (0, T), \; j = 1, 2, 3, \\ \lambda^j (\cdot, T) = 0 & \text{in } e_j, \end{cases}$$

with the boundary condition

$$\lambda^3 (V_3, t) = 0, \quad \text{in } [0, T],$$

and the transmission condition

$$\lambda^j (V_0, t) v^j (V_0, t) = \lambda^3 (V_0, t) v^3 (V_0, t) \quad \text{if } t \in \{ v^j \neq 0 \}, \; j = 1, 2 \, .$$

Finally, if we impose box constraints $T^G < s_i < T^R$ for $i = 1, ..., S$, the optimal solution (m, u^s, λ) should satisfy, for all $\bar{s} \in \mathbb{R}^S$ such that $T^G < \bar{s}_i < T^R$, the variational inequality

$$\langle \partial_s L(\mu, u^s, \lambda), \bar{s} - s \rangle \geq 0. \tag{6.21}$$

Remark 6.1.1. *If the velocity field contains the drivers interaction term, then the dual problem for* (6.14)-(6.15)-(6.16) *is given by*

$$\begin{cases} -\partial_t \lambda^j - v^j \partial_x \lambda^j - \nu * (\mu \partial_x \lambda) = v^j + \nu * \mu & \text{in } e_j \times (0,T),\ j = 1,2,3 \\ \lambda^j(\cdot, T) = 0 & \text{in } e_j \end{cases}$$

with the same boundary and transition conditions, where $(\nu * \phi)(x) = \int_\Gamma K(y,x)\phi(y)dy$. *The additional terms in the equation represent a time-backward counterpart of the nonlocal term in the forward equation. Indeed, note that the kernel* K *is not symmetric by definition and the integration is here performed with respect to the first variable, looking at* $y \to x$ *and not* $x \to y$ *as in* (5.2) .

6.3.2 Discretization

The above optimality system can be discretized using, for instance, finite difference schemes and solved by some root-finding algorithm. Here we do not solve the whole discrete system at once, we instead obtain an approximate solution splitting the problem in three simple steps. With a fixed control, we first solve the forward equation in μ, then we solve the backward equation in λ, and finally update the control using the expression we obtained for the gradient $\partial_s L$, iterating up to convergence. The resulting procedure is a gradient descent method, summarized in the following algorithm.

Algorithm [Forward-Backward system with Gradient Descent]

Step 0. Choose $\varepsilon > 0$, $\beta > 0$ and set $J^{(0)} = 0$;

Step 1. Fix an initial guess for $s^{(0)} \in \mathbb{R}^S$, $u_0 \in \{0, 1\}$ and set $k = 0$;

Step 2. Use $s^{(k)}$ to build the control $u^{(k)}$;

Step 3. Solve the forward problem for $\mu^{(k)}$ with control $u^{(k)}$;

Step 4. Solve the backward problem for $\lambda^{(k)}$ with control $u^{(k)}$;

Step 5. Compute $J^{(k+1)} = J(\mu^{(k)}, s^{(k)})$.
If $|J^{(k+1)} - J^{(k)}| < \varepsilon$ go to Step 8, otherwise update $J^{(k)} \leftarrow J^{(k+1)}$ and continue;

Step 6. Compute $\partial_s L$ at $(\mu^{(k)}, u^{(k)}, \lambda^{(k)})$;

Step 7. Update $s^{(k)} \leftarrow \Pi_{\{T^G, T^R\}} \left(s^{(k)} - \beta \partial_s L(\mu^{(k)}, u^{(k)}, \lambda^{(k)}) \right)$, $k \leftarrow k + 1$ and go to Step 2
($\Pi_{\{T^G, T^R\}}$ denotes the component-wise projection on the interval $[T^G, T^R]$);

Step 8. Accept $(\mu^{(k)}, u^{(k)}, \lambda^{(k)})$ as an approximate solution of the optimal control problem for (6.13).

In the actual implementation of the algorithm, we employ the scheme introduced in the previous chapter, to solve the forward equation in μ. On the other hand, the adjoint advection equation in λ is solved by means of a standard time-backward upwind scheme. We choose the numerical grid in space and time subject to a sharp CFL condition, in order to mitigate the numerical diffusion and better observe the nonlocal interactions. Moreover, we compute all the integrals appearing in the functional J, in the nonlocal terms and in the expression of the gradient $\partial_s L$, by means of a rectangular quadrature rule. We also employ a simple inexact line search technique to compute a suitable step β for the gradient update in Step 7. Finally, the application of control constraints is easily obtained by projection. More precisely, given compatible durations $0 < T^G < T^R$ and the updated $s^{(k)}$ in Step 7, we set $s_i^{(k)} \leftarrow \max\{T^G, \min\{s_i^{(k)}, T^R\}\}$ for $i = 1, ..., S$.

6.3.3 Numerical experiments

As a preliminary test we compare the local and the nonlocal case. We consider only the evolution of the density m along the edge e_1 and we set the control $u(t) \equiv 1$ to keep the traffic light at the end of the road activated (red) during the whole simulation. We choose the length $\ell(e_1) = 1$ and $R_1 = \frac{1}{8}$ for the visibility radius of the traffic light. On the other hand, we choose the nonlocal interaction kernel (5.2) with $k(r) = \frac{25}{1+r}$ and visibility radius $R = 15dx$, where dx is the step size of the space grid. Finally, we set the free flow speed $v_f^1 \equiv 1$ and the initial distribution $\mu_0(x) = \chi_{[0.1,0.15]}(x)$. Figure 6.3 shows the evolution of μ and v at different times. Top panels refer to the local case, bottom panels to the nonlocal one. We represent the density μ in black and the velocity v in red, decreasing from v_f^1 to zero with a linear ramp while approaching the traffic light, according to the definition (6.10) for H.

In the local case v does not depend on time, since u is constant. The density μ proceeds without changing profile (except some numerical diffusion at the boundary of its support), then starts concentrating close to the traffic light. At the final time, all the mass is concentrated at the point closest to the traffic light.
In the nonlocal case, drivers interactions are clearly visible both in μ and v. The initial density readily activates the nonlocal term in v, and μ starts assuming the well known triangle-shaped profile. Close to the traffic light we observe a slowing-down, that propagates backward up to the beginning of the queue, preventing mass concentration. At final time the profile becomes stationary, we observe that v is zero in the whole support of μ.

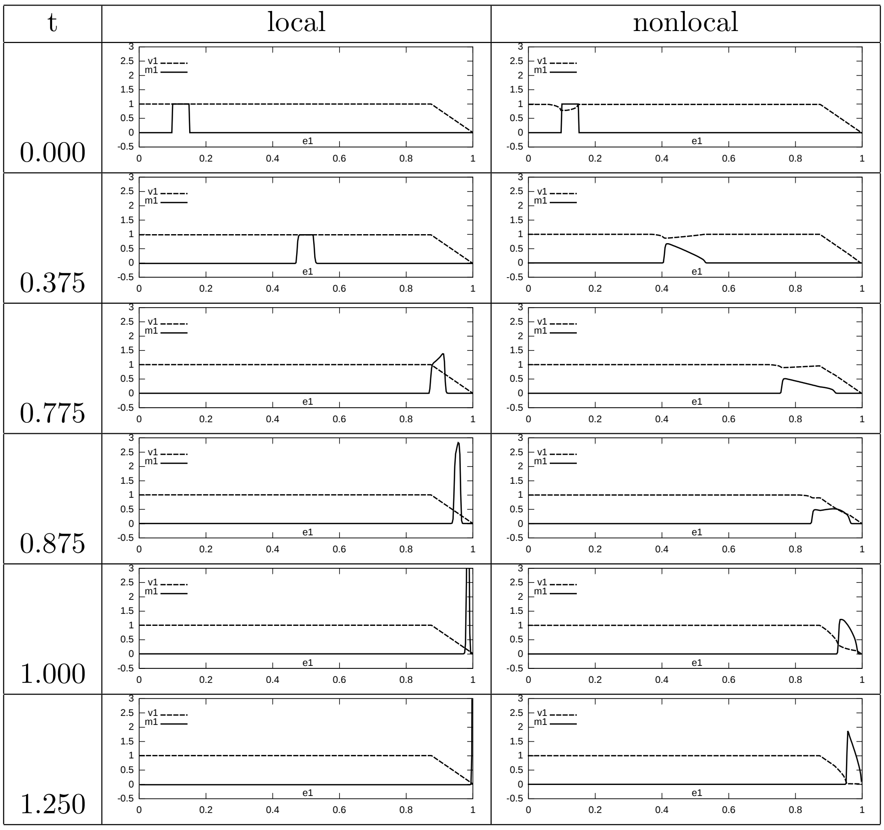
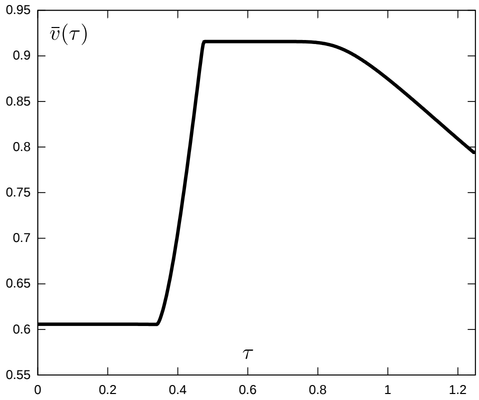


Figure 6.3: Red traffic light: local case vs nonlocal case

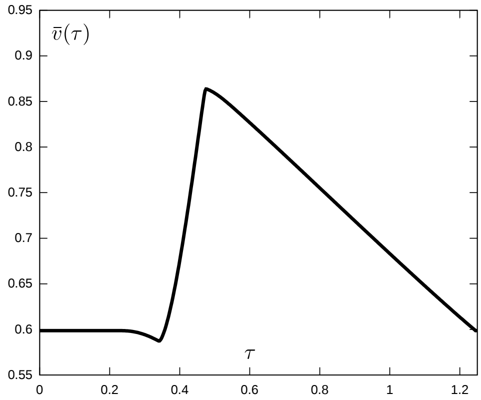
We proceed with a test for validating the proposed numerical method. We consider the case of a single switching time $\tau \in [0, T]$, namely we choose $s = (s_1, s_2) = (\tau, T - \tau)$ without constraints and $u_0 = 1$, so that the corresponding control is just $u^s(t) = \chi_{[0,\tau]}(t)$ (red light on e_1 for $t \leq \tau$). This reduces the optimization problem to a minimization in dimension one, that can be analyzed by an exhaustive search in τ and then compared with our adjoint-based algorithm. We set all the parameters as in the previous test, in particular we choose constant free flow speeds $v_1^f = v_2^f = v_3^f \equiv 1$ and set $T = 1.25$. We also assume that, apart from μ_0, no additional mass enters or leaves the network for all $t \in [0, T]$.

We start with $\mu_0 = (\mu_0^1, \mu_0^2, \mu_0^3) = (\chi_{[0.1,0.15]}(x), \chi_{[0.6,0.65]}(x), 0)$, i.e. two distributions of equal mass on e_1 and e_2 that arrive at the traffic light at different times (μ_2 first and then μ_1). In Figure 6.4(a) we plot the corresponding (normalized) mean velocity $\bar{v}(\tau) = -J(\mu, u^s)/M$ as a function of τ, where $M = \int_0^T \int_\Gamma m(x,t)dx\,dt$.

The scenario is pretty clear. If the switch occurs before μ_2 reaches the traffic light, then only μ_1 will move from e_1 to e_3 and the mean velocity cannot improve. For larger values of τ, also μ_2 will gradually move to e_3, and $\bar{v}(\tau)$ increases. If now the switch is placed just after μ_2 leaves e_2 and before μ_1 approaches the traffic light, we get the best performance, both distributions move as they are on a free road. Note that, due to the nonlocal interactions, the maximum of $\bar{v}$ is less than the free flow



(a)



(b)

Figure 6.4: Mean velocity for a single switch of the traffic light: well separated (a) vs overlapping (b) densities

speed. Finally, as τ keeps increasing up to T, μ_1 starts getting stuck at the traffic light, and $\bar{v}(\tau)$ decreases.

Now let us repeat the exaustive computation of the mean velocity $\bar{v}(\tau)$ with $\mu_0 = (\mu_0^1, \mu_0^2, \mu_0^3) = (\chi_{[0.6,0.65]}(x), \chi_{[0.6,0.65]}(x), 0)$, two distributions of equal mass on e_1 and e_2, starting at the same distance from the traffic light. Figure 6.4(b) shows the shape of the corresponding $\bar{v}$. We observe that the maximum of $\bar{v}$ is lower than in the previous test, and it is achieved at a single point instead of an interval. This clearly depends on the fact that the two densities are not well separated as before and it is not possible to place a switch without penalizing the overall traffic flow. Moreover, note that an absolute minimum appears just after the initial plateau. Interestingly, this means that if the switch occurs too early both densities slowdown, whereas the optimal choice corresponds to switch just after μ_2 leaves e_2 (see Figure 6.4).

These two simple examples show that, in general, the numerical optimization of the traffic light is a very challenging problem, since there is a wide number of local extrema where the gradient descent algorithm can stop. To overcome this issue, we perform several runs with random initial guesses for the controls, and we select the solution obtaining the best result.

Figure 6.5 shows the optimal solution at different times in the case of well separated. The solution is computed by the gradient descent method and achieves the absolute maximum of the corresponding mean velocity. Similarly, Figure 6.6 refers to the case of overlapping densities. We clearly observe that on e_1 the traffic is stopped until m_2 leaves e_2.

We conclude with a more complete example, also including control constraints. All the parameters are the same of the previous tests, but we fix to $S = 5$ the number of switching durations (corresponding

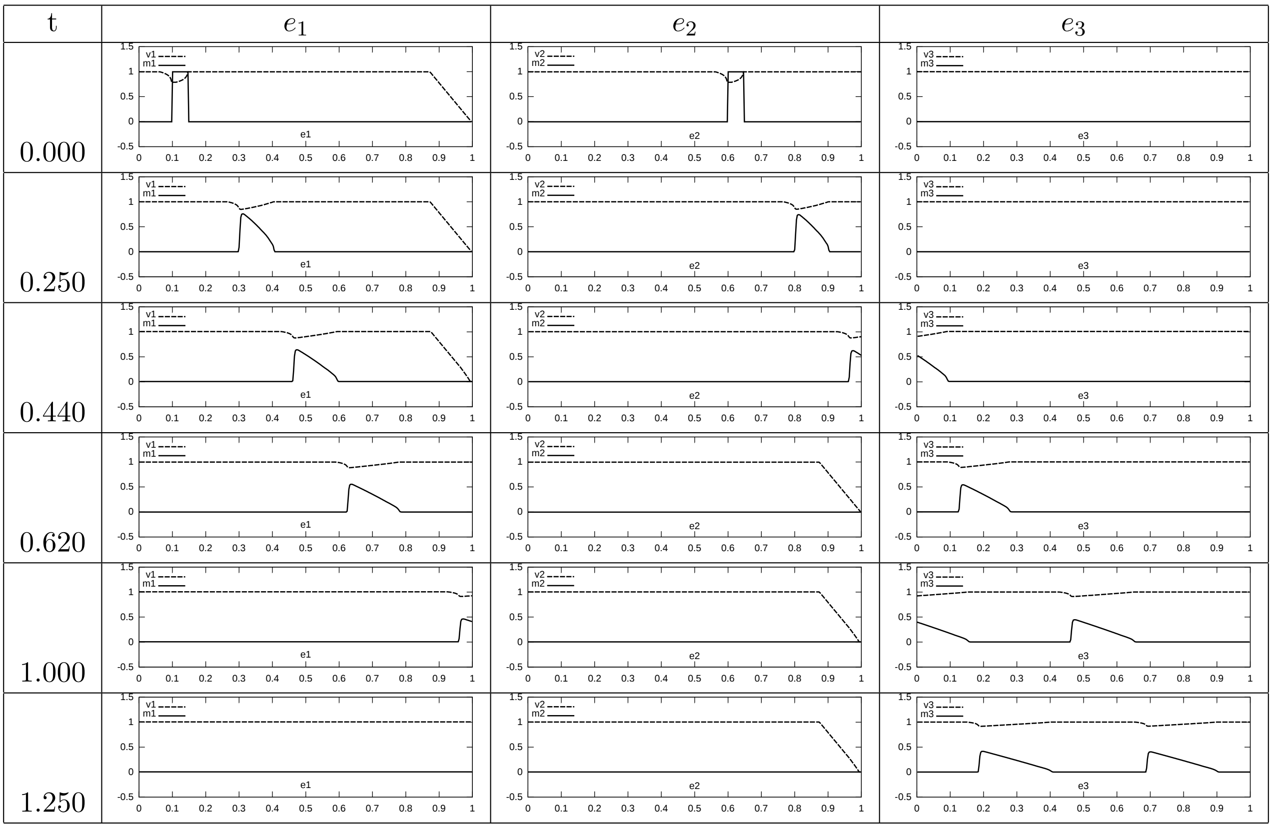


Figure 6.5: Optimal solution for well separated densities

to 4 switching times) and we start with $u_0 = 0$, i.e. green light on e_1. Moreover, we set the constraints $T^G = 0.15$, $T^R = 0.3$, and m_0 is given edge-wise by

$$\mu_0^1(x) = \chi_{[0.1,0.15]}(x) + \chi_{[0.4,0.45]}(x)\,, \quad \mu_0^2(x) = \chi_{[0.1,0.15]}(x) + \chi_{[0.6,0.65]}(x)\,, \quad \mu_0^3(x) = 0\,.$$

Note that, with this choice, we are mixing together the two cases analyzed before. Indeed, the initial density consists of four blocks which are, respectively, pairwise overlapped and well separated. The optimal solution produced by the gradient descent algorithm is $s^* = (0.227, 0.251, 0.259, 0.3, 0.21)$. Figure 6.7 shows the corresponding evolution at different times. We observe that the first switch occurs before μ_2 approaches the traffic light. This allows the first block of μ_2 to proceed without slowdowns from e_2 to e_3. The second switch occurs immediately after this block leaves e_2, so that also the first block of μ_1 can leave e_1 almost undisturbed before the traffic light switches again. Now, the remaining densities on e_1 and e_2 are in overlapping configuration, μ_2 goes first, while μ_1 stops. Finally, the last switch occurs just after μ_2 leaves e_2, so that also μ_1 can move to e_3 for the remaining time.

The numerical method proposed in this section has many interesting analogies with numerical methods related to the Pontryagin Principle or the numerical methods adopted for Mean Field Games. In our context, the complexity arises from the presence of one or more junctions. Moreover, we have focused

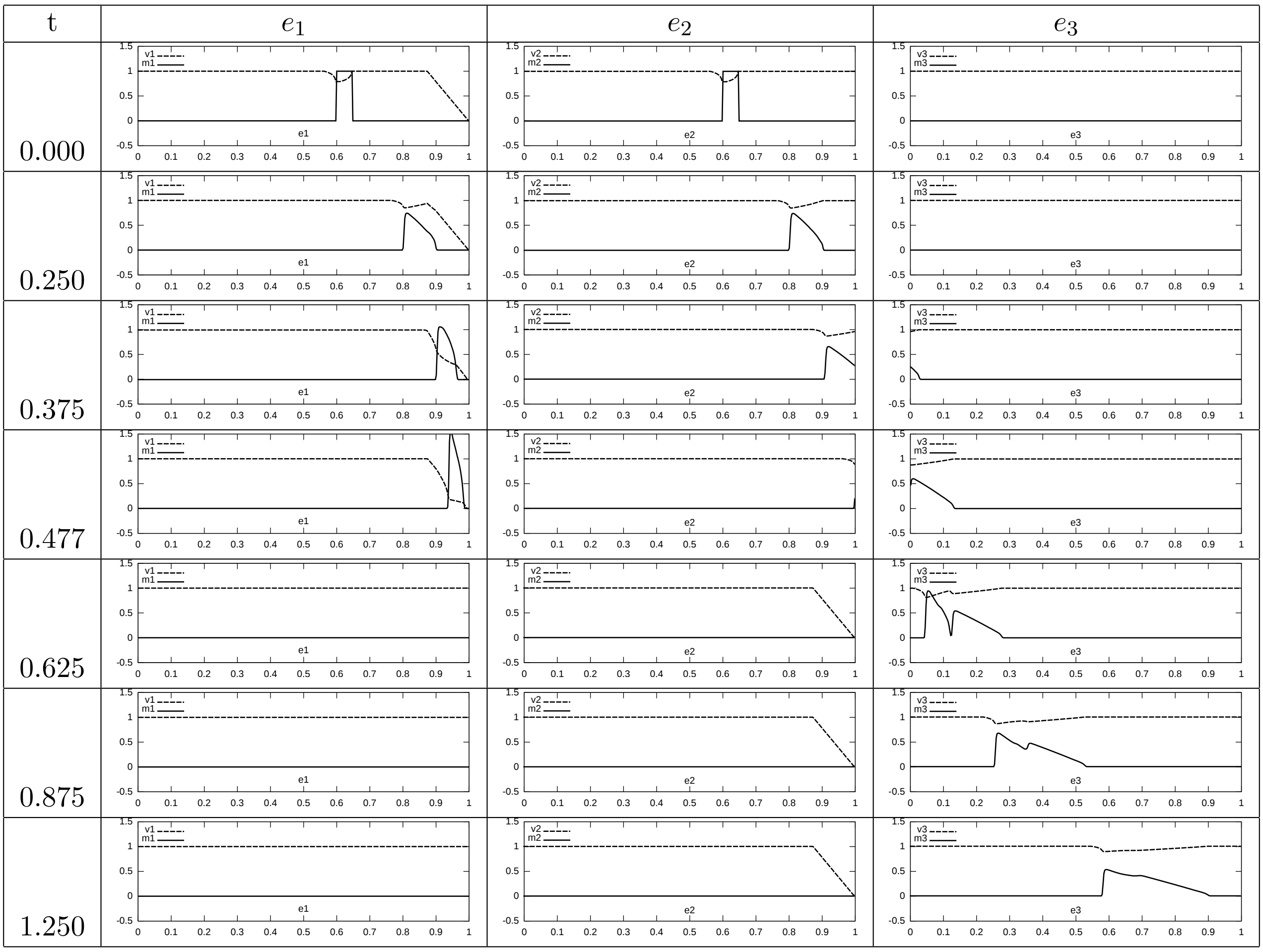


Figure 6.6: Optimal solution for overlapping densities

only on smart traffic lights for the simplest (nontrivial) junction and the generalization to more complex network is not difficult. The case of autonomous cars, i.e. atomic dynamic controls, is difficult and too complex to be addressed with the before mentioned methods. We think that this problem and the correlated numerical methods offer an interesting perspective for future research.

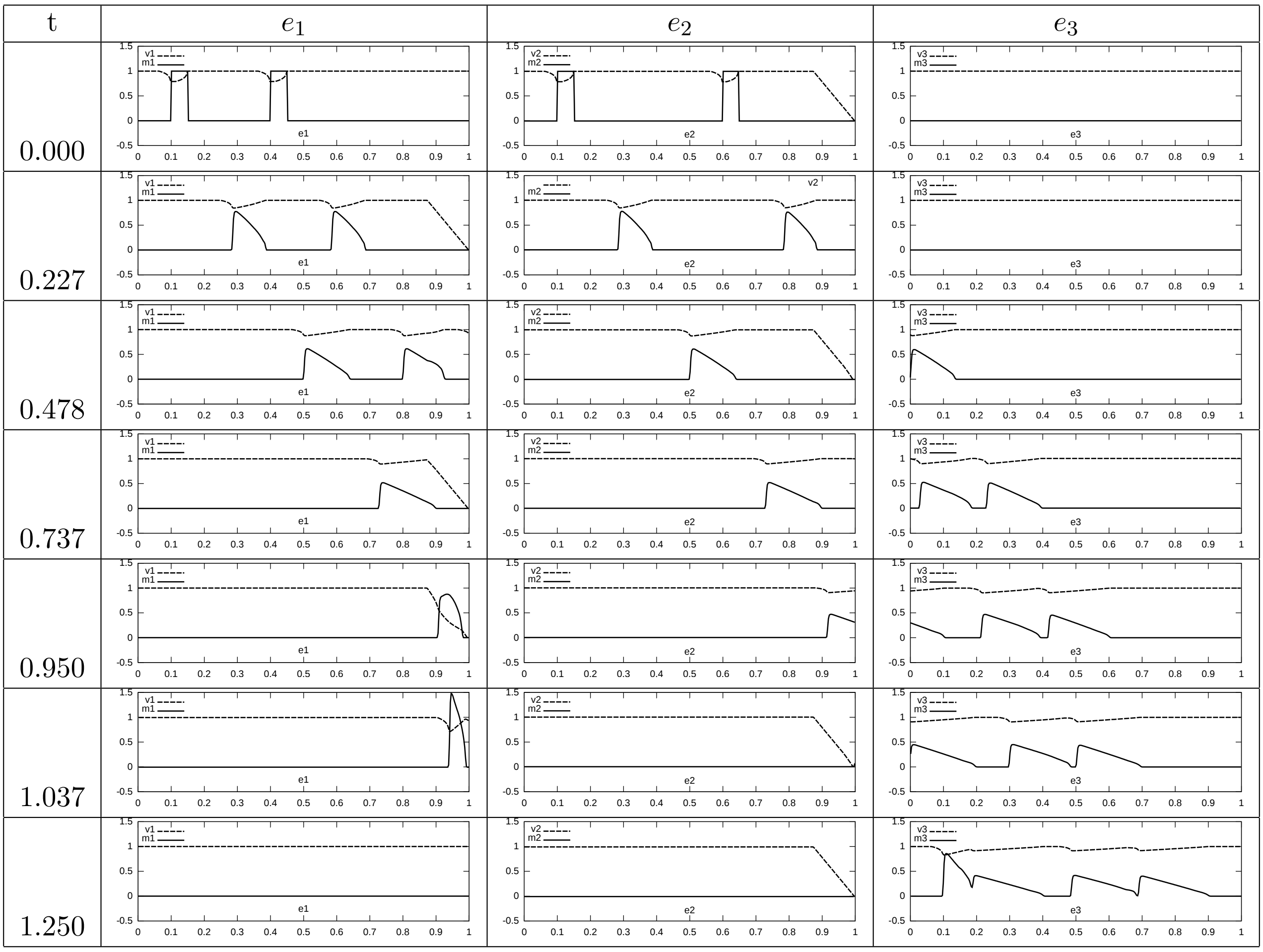


Figure 6.7: Optimal solution for a traffic light with 4 switches

Chapter 7

Conclusion

In this thesis we have introduced a class of measure valued transport equations on networks with non-local velocity fields to describe vehicular traffic and drivers' interactions. We have also investigated model's features and numerical aspects, useful for applications such as mobility optimization, traffic light settings and traffic control. The theory exposed in this manuscript covers only partially the complexity related to vehicular traffic but it can be a first step to investigate further on it. During the three years of the Ph.D. program we mainly focused on modeling aspects and partially on applications. However, a complete modeling of vehicular traffic would require a wider recursive process in four steps: data investigation, model, application and testing.
While we focused on the model phase, i.e. formulating hypothesis and building tools to be used, many opportunities and research areas are offered by the other three phases. For example, about the data it is important to understand their sources, characterization, relationships, noises and how to use them to understand the phenomena and build more efficient models. Once we have introduced a new one, it is necessary to understand and investigate its application, efficient numerical scheme based on the task, simulations and control problems; lastly, the testing phase proposes also many interesting challenges from a mathematical point of view, since it is important to calibrate and quantify parameters, errors and uncertainty.

Moreover after this last phase we can improve our model repeating the process starting from new data and informations. This four steps provide us four macro-area of research which could be investigated with different approaches. We conclude proposing some open questions and good opportunities, related to each area, for research on transport and traffic on networks:

- **Parameters' estimation**: our models depends on many parameters, such as the visual radius, priority rules and distribution matrixes, which we assumed given a priori; however, if we desire to apply such models, we need criteria to fix this parameters. For this purpose, a first opportunity would be some statistical knowledge from data; another opportunity, more interesting from a mathematical point of view, is assuming a relationship between this parameters and the state of the system. For example, we could assume that every driver in our population adopts strategy to choose the next road at the junction (i.e. $P = P[\mu]$) or interact with other drivers (i.e. $\alpha = \alpha[\mu]$).

- **Nonlocal in time velocity fields**: in this thesis we have assumed that every driver changes instantaneously its speed based on the present state of the system in front of him; this hypothesis is necessary to build the model but it does not describe correctly what really happens. Indeed, drivers' reactions are not instantaneous and it does not act directly on the car speed, but on its acceleration and these features are important since are the cause of car's incidents. For this purpose we could start from ordinary differential equations with delays or nonlocal acceleration. Both paths should lead to models with fully nonlocal, in time and space, velocity fields, i.e models where evolution of the state μ_t at time t depends on its history in a certain time range.

- **Relationship with classical models**: another area of research would be the investigation of these model with classical models, such as Follow-the-Leader, LWR or Aw-Rascle-Zhang. These investigation would lead to a deeper understanding of these models and respective parameters and, at the same time, lead to new models to be studied mathematically. From this point of view, there exists recent papers which show interesting connections between nonlocal velocity fields and the LWR model (see [22, 37]).

- **Transmission conditions**: until now we have proposed just one kind of condition at the internal vertices by assuming the instantaneous movement across the junctions. However, this is not the unique choice. For example, it would be possible to assume that every driver stop at the junction for a given amount of time, leading to shifted transmission conditions. Hence, it would be natural and important from a mathematical point of view to analyze which are the admissible conditions at the internal vertex and how they are characterized. A first step, as written in the Introduction, it is in [32, 33] and its possible extension to networks.

- **Algorithms and numerical methods**: for many applications it is fundamental to analyze both the accuracy and the computational speed. For this purpose it is important to study numerical methods and algorithms which can be parallelized in order to provide predictions in real time. A possible path is offered by the nowadays increasing interest in machine learning, in particular towards neural networks and GANs (see [45]).

Bibliography

[1] G. Albi, M. Bongini, E. Cristiani, D. Kalise, *Invisible control of self-organizing agents leaving unknown environments,* SIAM J. Appl. Math., 76 (2016), 1683-1710.

[2] L. Ambrosio, and N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures,* Birkhäuser Verlag, 2008, Lectures in Mathematics ETH Zürich

[3] J.P. Aubin, A. Cellina, *Differential Inclusions*, Springer-Verlag, 1984, A series of comprehensive studies in mathematics

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

[5] N. Bellomo, B. Piccoli, and A. Tosin, *Modeling Crowd Dynamics from a complex system viewpoint*, Math. Models Methods Appl. Sci., Vol. 02, no. 22, 2012

[6] N. Bellomo and C. Dogbé, *On the Modeling of Traffic and Crowds: A Survey of Models, Speculations, and Perspectives*, Vol. 53, no. 3, 2011, pg. 409-463

[7] A. Bertozzi, T. Laurent, and J. Rosado, *L^p Theory for the Multidimensional Aggregation Equation*, Comm. on Pure and Appl. Math., Vol. 64, no. 1, 2011, pg. 45-83

[8] M. Bongini, G. Buttazzo, *Optimal control problems in transport dynamics,* Math. Models Methods Appl. Sci. 27 (2017), 427-451.

[9] V. I. Bogachev,*Measure theory*, Springer-Verlag, 2007

[10] A. Bressan, S. Canic, M. Garavello, M. Herty, and B. Piccoli, *Flows on networks: recent results and perspectives*, EMS Surv. Math. Sci, Vol. 1, no. 1, 2014, pg. 47-111

[11] G. Bretti, and M. Briani, and E. Cristiani, *An easy-to-use algorithm for simulating traffic flow on networks: Numerical experiments*,Discrete Contin. Dyn. Syst. Ser. S, Vol. 7, no. 3, 2014, pg. 379-394

[12] M. Briani, and E. Cristiani, *An easy-to-use algorithm for simulating traffic flow on networks: Theoretical study*, Netw. Heterog. Media, Vol. 9, no. 3, 2014, pg. 519-552

[13] S. Cacace, F. Camilli, R. De Maio and A. Tosin, *A measure theoretic approach to traffic flow optimization on networks*, accepted by European J. of Applied Mathematics, 2018

[14] F. Camilli, R. De Maio and A. Tosin, *Transport of measures on networks,* Networks and Heterogeneous Media, vol. 12, no. 2, 2017, pg. 191-215

[15] F. Camilli, R. De Maio and A. Tosin, *Measure-valued solutions to nonlocal transport equations on networks*, Journal of Diff. Eq., Vol. 254, Iss. 12, 2018, pg. 7213 - 7241

[16] F. Camilli, and C. Marchi, *Stationary mean field games systems defined on networks*, SIAM J. Control. Optim., 54(2), 2016, pg. 1085-1103

[17] J. A. Cañizo, J. A. Carrillo, and J. Rosado, *A well-posedness theory in measures for some kinetic models of collective motion* Math. Models Methods Appl. Sci., Vol. 21, no. 3, 2011, pg. 515-539

[18] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepäv, *Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations*, Duke Math. J., Vol. 156, no. 2, 2011, pg. 229-271.

[19] E. Cascetta, *Transportation Systems Analysis*, Springer, Heidelberg (2009).

[20] G.Cavagnari, A. Marigonda, B. Piccoli, *Averaged time-optimal control problem in the space of positive Borel measures,* ESAIM: COCV, 24 2 (2018) 721-740.

[21] G. Cavagnari, A. Marigonda, K.T. Nguyen, F. S. Priuli, *Generalized control systems in the space of probability measures,* Set-Valued Var. Anal., 2017, pg. 1 - 29.

[22] F. A. Chiarello and P. Goatin, *Global entropy weak solutions for general non-local traffic flow models with anisotropic kernel*, ESAIM: M2AN, Vol. 52, no. 1, 2018, pg. 163-180

[23] G. Crippa and M. Lécureux-Mercier, *Existence and uniqueness of measure solutions for a system of continuity equations with non-local flow*, Nonlin. Diff. Eq. and App., Vol. 20, no. 3, 2013, pg. 523?537

[24] E. Cristiani, and B. Piccoli, and A. Tosin, *Multiscale Modeling of Pedestrian Dynamics*, Springer International Publishing, 2014, MS&A: Modeling, Simulation and Applications, Vol. 12

[25] E. Cristiani, and S. Sahu, *On the micro-to-macro limit for first-order traffic flow models on networks*, Netw. Heterog. Media no. 3, Vol. 11, 2016, pg. 395-413

[26] C. D'Apice, and R. Manzo, and B. Piccoli, *Modelling supply networks with partial differential equations*, Quart. Appl. Math., Vol. 67, no. 3, 2009, pg. 419-440

[27] P. Degond and S. Motsch, *Continuum limit of self-driven particles with orientation interaction*, Math. Models Methods Appl. Sci., Vol. 18, 2008,, pp. 1193-1217.

[28] M. Delitala, *Nonlinear models of vehicular traffic flow: New frameworks of the mathematical kinetic theory*, C. R. Mecanique, Vol. 331, 2003, pp. 817-822.

[29] R. Dudley, *Convergence of Baire measures*, Stud. Math. Vol. 27, no. 3, 1966

[30] R. Dudley, *Correction to "Convergence of Baire measures"*, Stud. Math. Vol. 55, no. 3, 1974

[31] K.-J. Engel, and M. K. Fijavž, and R. Nagel, and E. Sikolya, *Vertex control of flows in networks*, Netw. Heterog. Media, Vol. 3, no.4, 2008, pg. 709-722

[32] J.H.M. Evers, S.C. Hille, and A. Muntean, *Mild solutions to a measure-valued mass evolution problem with flux boundary conditions*, J. Differential Equations, vol. 259, no. 3, 2015, pg. 1068-1097,

[33] J.H.M. Evers, S.C. Hille, and A. Muntean, *Measure-valued mass evolution problems with flux boundary conditions and solution-dependent velocities*, SIAM J. Math. Anal., Vol. 48, no. 3, 2016, pg. 1929-1953

[34] L. Fermo, and A. Tosin, *A fully-discrete-state kinetic theory approach to traffic flow on road networks*, Math. Models Methods Appl. Sci., Vol. 25, no. 3, 2015, pg. 423 - 461

[35] W.H. Fleming, and H. Mete Soner, *Controlled Markov Processes and Viscosity solutions*, Springer-Verlag New York, Stochastic Modelling and Applied Probability, Vol. 25, 2006

[36] P. Frasca, and A. Tosin, *Existence and approximation of probability measure solutions to models of collective behaviors*, Netw. Heterog. Media, Vol. 3, no. 6, 2011, pg. 561-569

[37] J. Friedrich, S. Goettlich and O. Kolb, *A Godunov type scheme for a class of LWR traffic flow models with non-local flux*, Networks Heterogeneous Media, doi: 10.3934/nhm.2018024

[38] M. Garavello, and B. Piccoli, *Traffic Flow on Networks – Conservation Laws Models*, AIMS Series on Applied Mathematics, American Institute of Mathematical Sciences, 2006, pg. xvi+243

[39] S. Goettlich, and M. Herty, and U. Ziegler, *Modeling and optimizing traffic light settings in road networks* Computers and Operations Research, Vol. 55, 2015, pg. 36-51

[40] S.Gottlich, U.Ziegler, *Traffic light control: a case study*, Dis. and Con. Dyn. Sys. Vol. 7 (2014), 483–501.

[41] Gwiazda, P., Jamróz, G. and Marciniak-Czochra,A., *Models of discrete and continuous cell differentiation in the framework of transport equation*, SIAM J. Math. Anal., Vol. 44, no. 2, 2012, pg. 1103-1133

[42] D. Helbing, Traffic and related self-driven many-particle systems, Rev. Modern Phys., 73, 2001, pp. 1067-1141.

[43] S.C. Hille, T. Szarek, D. Worm, D. and M. Ziemlanska, *On a Schur-like property for spaces of measures*, 2017, submitted

[44] A. Marigonda, and M. Quincampoix, *Mayer control problem with probabilistic uncertainty on initial positions*, Journal of Diff. Eq., Vol. 264, no. 5, 2018, pg. 3212-3252

[45] S.S. Mousavi, M. Schukat, and E. Howley, *Traffic light control using deep policy-gradient and value-function-based reinforcement learning*, IET Intelligent Transport Systems, Vol. 11, no. 7, 2017, pg. 417-423

[46] D. Mugnolo, *Semigroup Methods for Evolution Equations on Networks*, Springer International Publishing, Understanding Complex Systems, 2014

[47] B. Piccoli, *Measure differential equations,* submitted, preprint on ArXiv

[48] B. Piccoli, and F. Rossi, *Transport Equation with Nonlocal Velocity in Wasserstein Spaces: Convergence of Numerical Schemes*, Acta Appl. Math. Vol. 124, no. 1, 2013, pg. 73-105

[49] B. Piccoli and A. Tosin, *Vehicular Traffic: A Review of Continuum Mathematical Models*, In: Meyers R. (eds) Mathematics of Complexity and Dynamical Systems. Springer, New York, NY, 2012

[50] L. A. Pipes, *An operational analysis of traffic dynamics,* J. Appl. Phys., 24 (1953), pp. 274-287

[51] F. Poupaud and M. Rascle, *Measure solutions to the linear multi-dimensional transport equation with non-smooth coefficients*, Comm. in Par. Diff. Equ., Vol. 22, no. 1-2, 1997

[52] Yu. V. Pokornyi, and A. V. Borovskikh, *Differential equations on networks*, J. Math. Sci. (N. Y.), 119 (6), 2004, pg. 691 - 718

[53] R. W. Rothery, *Car following models, in Traffic Flow Theory,* N. Gartner, C. J. Messer, and A. K. Rathi, eds., Transportation Research Board, Special Report 165, 1992.

[54] C. Villani, *Topics in Optimal Transportation*, American Mathematical Society, Graduate Studies in Mathematics, Vol. 58, 2003

[55] Y. Zhao and H.M.Zhang, *An unified follow-the-leader model for vehicle, bicycle and pedestrian traffic*, Transportation Research Part B, Vol. 105, 2017, pg. 315 - 327

[56] R.Wang, Y.Li, D. B. Work, *Comparing traffic state estimators for mixed human and automated traffic flows*, Transportation Research Part C: Emerging Technologies 78 (2017), 95–110.

[57] Worm, D. T. H., *Semigroups on spaces of measures*, PhD. Thesis, Leiden University, 2010

A Promise instead of a Dedication

It would be normal to conclude with a short dedication to someone of my past or present. However, these last years have taught me that we should focus a bit more on the future.
I started this Ph.D. because it was my intention to complete a path even if a little fuzzy and dark. I feel like I made the right choice and learned many things, reached my targets and found new goals.
I had the opportunity to understand more, learn methods, how to fail and how to look for innovation.
I believe that all the things I've acquired are precious and important not just for myself but for the whole society. Even if it seems that nowadays competence and study are no longer necessary and respected, I promise I will always study and try to understand the truth.