

A measure theoretic approach to traffic flow optimization on networks

*Original*

A measure theoretic approach to traffic flow optimization on networks / Cacace, Simone; Camilli, Fabio; De Maio, Raul; Tosin, Andrea. - In: EUROPEAN JOURNAL OF APPLIED MATHEMATICS. - ISSN 0956-7925. - ELETTRONICO. - 30:Spec. issue 6(2019), pp. 1187-1209. [10.1017/S0956792518000621]

*Availability:*

This version is available at: 11583/2742939 since: 2020-04-23T17:21:47Z

*Publisher:*

Cambridge University Press

*Published*

DOI:10.1017/S0956792518000621

*Terms of use:*

openAccess

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

# A MEASURE THEORETIC APPROACH TO TRAFFIC FLOW OPTIMIZATION ON NETWORKS

SIMONE CACACE, FABIO CAMILLI, RAUL DE MAIO, AND ANDREA TOSIN

**ABSTRACT.** We consider 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. We also propose 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.

## 1. INTRODUCTION

During the last years, the study of vehicular and pedestrian traffic flow problems has become a very active area and an opportunity of information exchange between mathematical investigation and applied research. From a mathematical point of view, these phenomena have been largely studied due to their high complexity and the literature offers a broad variety of models devoted to their description in a wide range of scenarios, see [4, 13, 15] for reviews. On the other side, from an engineering point of view, it is important to model, simulate, predict, control and optimize vehicular and pedestrian traffic in our society. These issues become more and more central with the fast technological progress and it is of particular interest to understand how the latest technologies, such as smart traffic lights, self-driving cars or big data, can be used to improve the quality of movement for drivers or pedestrians on road networks and urban roads, see [8, 20].

In this paper we propose a model to simulate and optimize traffic flow on networks based on the theory of measure-valued transport equations. In this approach, the population is represented by a probability distribution which evolves according to a velocity field depending on the position of the other individuals. In this way short and long range interaction mechanisms are readily taken into account into the dynamics of the problem. Moreover the measure approach easily catches the multi-scale nature of vehicular traffic, composed both by a continuous distribution of indistinguishable cars and by some special

---

*Date:* September 5, 2018.

*2010 Mathematics Subject Classification.* 49J20, 35B37, 35R02, 49M07, 65M99.

*Key words and phrases.* Network, transport equation, measure-valued solutions, transmission conditions, optimization.

individuals such as automated cars and traffic lights. With respect to other models considering transport equations with nonlocal interactions (see [1, 7, 11, 17]), the peculiarity of our model is to be defined on a network, posing additional difficulties for the interpretation in a measure-theoretic sense of the transition conditions at the vertices. Existence, uniqueness and continuous dependence results for the corresponding measure-valued transport equation were provided in [5, 6].

In [3, 9, 10], the authors consider optimal control problems for measure transport equations in the Euclidean space. Relying on a similar approach, we consider a model where, besides the driver distributions, the velocity field depends also on an external distribution which interacts with the original population in order to optimize, for example, traffic volume or average speed on the road network. As in [2, 20], our aim is to 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 and investigate 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.

The paper is organized as follows. In Section 2 we introduce the control problem from a theoretical point of view: network structure, transport equation and cost functional. Section 3 is devoted to two examples of control problem: traffic lights and self-driving cars as controls for vehicular traffic. Section 4 focuses on numerical analysis for these problems: description and properties of the chosen scheme and numerical tests on some case studies. In the Appendix we report the proofs of some theoretical results contained in the previous sections.

## 2. PROBLEM FORMULATION AND THEORETICAL SETTING

In this section we describe the main components of the traffic flow model, i.e. the structural components (roadway and priority rules at the junctions), 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 to optimize the traffic flow on the network.

**2.1. Structural components.** Traffic routes are mathematically described by a network  $\Gamma = (\mathcal{V}, \mathcal{E})$  where  $\mathcal{E} = \{e_1, e_2, \dots, e_{|\mathcal{E}|}\}$  is the set of arcs/roads while the crossroads are represented by the set of the vertexes  $\mathcal{V} = \{V_1, \dots, V_{|\mathcal{V}|}\}$ . The network is oriented and we write  $e_k \rightarrow e_j$  and, respectively,  $x \rightarrow y$  for  $x, y \in \Gamma$  to mean that  $e_k$  comes before  $e_j$  and, respectively,  $x$  before  $y$  in the orientation of the network. We assume that  $\Gamma$  is endowed with the minimum path distance  $d_\Gamma$  and each arc  $e_j \in \mathcal{E}$  is parametrised by a continuous bijective map  $\pi_j : [0, L_j] \rightarrow e_j$ ,  $L_j \in (0, +\infty]$ , which complies with the orientation of  $\Gamma$ , i.e. if  $V, W \in \mathcal{V}$  are the vertexes of the arc  $e_j$  oriented from  $V$  to  $W$ , then  $\pi_j(0) = V$  and  $\pi_j(L_j) = W$ .

For every  $V \in \mathcal{V}$ , we denote by  $\text{Inc}(V)$  the set of arcs in  $\mathcal{E}$  whose end point is  $V$  and by  $\text{Out}(V)$  the set of arcs in  $\mathcal{E}$  whose starting point is  $V$ . Then, we divide the set of the

vertexes respectively in the sets of sources, sinks and junctions

$$\begin{aligned}\mathcal{S} &= \{V \in \mathcal{V} \mid \text{Inc}(V) = \emptyset\}, \\ \mathcal{W} &= \{V \in \mathcal{V} \mid \text{Out}(V) = \emptyset\}, \\ \mathcal{J} &= \{V \in \mathcal{V} \mid \text{Out}(V) \neq \emptyset, \text{Inc}(V) \neq \emptyset\}.\end{aligned}$$

Since the velocity term depends on the distribution of the cars on all the network, in order to simplify the notations we prefer to consider a network without sinks, i.e. the set  $\mathcal{W}$  is empty and the terminal arcs always have infinite length. We also denote by  $L_0$  the minimal length of the edges in  $\mathcal{E}$ , i.e.

$$(2.1) \quad L_0 = \min_{j=1, \dots, |\mathcal{E}|} L_j.$$

A convenient framework to study transport problems is given by the measure theoretic one, since it allows to consider in a same setting macroscopic quantities such as a continuous distribution of drivers and microscopic ones such as traffic lights and other elements of the model. We set  $\Gamma_T = \Gamma \times [0, T]$  and we consider the metric space  $(\Gamma_T, d)$  where  $d((x, t), (y, s)) = d_\Gamma(x, y) + |t - s|$ . For a function  $\phi : \Gamma_T \rightarrow \mathbb{R}$  we define the norm

$$\|\phi\|_{BL} = \|\phi\|_\infty + \sup_{\substack{(x,t), (y,s) \in \Gamma_T \\ x \neq y \\ t \neq s}} \frac{|\phi(x, t) - \phi(y, s)|}{d((x, t), (y, s))},$$

and we consider the Banach space  $BL(\Gamma_T)$  of bounded and Lipschitz continuous functions equipped with norm  $\|\cdot\|_{BL}$ . Denoted by  $\mathcal{M}(\Gamma_T)$  the space of finite measure on  $\Gamma_T$ , we define a dual norm on this space by

$$\|\mu\|_{BL}^* = \sup_{\substack{\phi \in BL(\Gamma_T) \\ \|\phi\|_{BL} \leq 1}} |\langle \mu, \phi \rangle|.$$

Similar notations and definitions are employed for the Banach space  $\mathcal{M}(\Gamma)$  and  $\mathcal{M}([0, T])$ . In the following we will always consider measures in  $\mathcal{M}^+(\Gamma_T)$ , the cone of positive measures in  $\mathcal{M}(\Gamma_T)$ . By the Disintegration Theorem, we consider measures  $\mu \in \mathcal{M}^+(\Gamma_T)$  which can be decomposed as

$$\mu(dxdt) = d\mu_t(x)dt,$$

where  $\mu_t \in \mathcal{M}^+(\Gamma)$  represents the distribution at time  $t \in [0, T]$ . We remark that throughout the paper we only consider measures without Cantorian part, since this kind of measure does not have any significant interpretation for traffic flow problems. To model the behavior of drivers at junctions we assign a *distribution matrix*  $P(t) = (p_{kj}(t))_{k,j=1}^{|\mathcal{E}|}$ , for  $t \in [0, T]$ , satisfying the following properties

$$(2.2) \quad \begin{aligned}p_{kj} &\in BV([0, T]), \quad p_{kj}(t) \in [0, 1], \\ \sum_{j=1}^{|\mathcal{E}|} p_{kj}(t) &= 1, \quad \forall t \in [0, T], \forall k = 1, \dots, |\mathcal{E}|, \\ p_{kj}(t) &= 0 \quad \text{if either } e_k \cap e_j = \emptyset \text{ or } e_j \rightarrow e_k.\end{aligned}$$

Here  $p_{kj}(t)$  represents the fraction of drivers which at time  $t$  flows from an arc  $e_k$  to an arc  $e_j$ . Hence, for every arc  $e_k$ , we have a discrete probability distribution  $P_k(t) = \{p_{kj}(t)\}_j$

which describes the behaviour of drivers at the junction at time  $t$ . This quantity is defined on the basis of the knowledge of the statistical behavior of the traffic at a given day time (see [16, 19]). The assumptions in (2.2) imply the mass cannot concentrate at the vertexes and therefore the total mass is conserved at the internal junctions. Since we consider measures  $\mu \in \mathcal{M}^+(\Gamma_T)$  without Cantorian part, we assume that  $p_{kj} \in BV([0, T])$  so that for a measure  $\mu \in \mathcal{M}^+(\Gamma_T)$  the product  $p_{kj} \cdot \mu$  still has no Cantorian part.

**2.2. Driver motion.** We now describe the nonlinear transport system which models the evolution of the traffic on the network. 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  $\Gamma$

$$m_0 = \sum_{e_j \in \mathcal{E}} m_0^j \in \mathcal{M}^+(\Gamma),$$

where  $m_0^j$  is restriction of  $m_0$  to  $e_j$ , and the incoming traffic measure at the source nodes

$$\sigma_0 = \sum_{V_i \in \mathcal{S}} \sigma_0^i, \quad \sigma_0^i \in \mathcal{M}^+(\{V_i\} \times [0, T]),$$

where  $\sigma_0^i$  is the restriction of  $\sigma_0$  to  $V_i$ , representing the flow of cars entering in the road network at the vertex  $V_i$ . We consider the following system of measure-valued differential equations on  $\Gamma_T$  for the unknown measure  $m = \sum_{e_j \in \mathcal{E}} m^j \in \mathcal{M}^+(\Gamma_T)$

$$(2.3) \quad \begin{cases} \partial_t m^j + \partial_x(v^j[m_t, \mu_t]m^j) = 0 & x \in e_j, t \in (0, T], j = 1, \dots, |\mathcal{E}|, \\ m_{t=0}^j = m_0^j & x \in e_j, j = 1, \dots, |\mathcal{E}|, \\ m_{V_i=\pi_j(0)}^j = \begin{cases} \sum_{e_k \in \text{Inc}(V_i)} p_{kj}(t)m_{V_i=\pi_k(1)}^k & \text{if } V_i \in \mathcal{I} \\ \sigma_0^i & \text{if } V_i \in \mathcal{S}, \end{cases} & j = 1, \dots, |\mathcal{E}|. \end{cases}$$

Observe that, for each arc  $e_j$ , if the initial vertex  $V_i = \pi_j(0)$  is internal, then the boundary condition at  $V_i$  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  $V_i = \pi_j(0)$  is an incoming traffic vertex, the inflow measure is the prescribed datum  $\sigma_0^i$ . The outflow measure, i.e. the part of the mass leaving the arc from the final vertex  $V_k = \pi_j(1)$ , is not given a priori but depends on the evolution of the measure  $m$  inside the arc.

The velocity  $v = (v^j)_{j=1}^{|\mathcal{E}|}$  depends on the solution  $m_t$  itself, as well as on another distribution  $\mu_t \in \mathcal{M}^+(\Gamma)$ , 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). We assume that

- (H1)  $v$  is non-negative and bounded by  $v_{max} > 0$ ;  
(H2)  $v$  is Lipschitz with respect to the state variables, i.e. there exists  $L > 0$  such that for every  $e_j \in \mathcal{E}$ ,  $\forall x, y \in e_j$ ,  $m_i, \mu_i \in \mathcal{M}^+(\Gamma)$ , for  $i = 1, 2$

$$|v^j[m_1, \mu_1](x) - v^j[m_2, \mu_2](y)| \leq L(|x - y| + \|m_1 - m_2\|_{BL}^* + \|\mu_1 - \mu_2\|_{BL}^*).$$

For the definition of measure-valued solution to the system (2.3), we refer to [6]. The next theorem summarize the main results concerning existence, uniqueness and regularity of the measure-valued solution to (2.3) in case of a fixed  $\mu \in \mathcal{M}^+(\Gamma_T)$ .

**Theorem 2.1.** *There exists a unique  $m \in \mathcal{M}^+(\Gamma_T)$  which is a measure-valued solution to (2.3). Moreover,*

- i) *There exists a positive constant  $C = C(T)$  such that*

$$\|m_t - m_{t'}\|_{BL}^* \leq C |t - t'| + \sigma_0((t', t])$$

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

- ii) *Given initial data  $m_{0,1}, m_{0,2} \in \mathcal{M}^+(\Gamma)$  and boundary data  $\sigma_{0,1}, \sigma_{0,2} \in \mathcal{M}^+([0, T])$  and denoted by  $m_1, m_2 \in \mathcal{M}^+(\Gamma_T)$  the corresponding solutions, there exists a constant  $C = C(T) > 0$  such that*

$$\|m_{T,2} - m_{T,1}\|_{BL}^* \leq C (\|m_{0,2} - m_{0,1}\|_{BL}^* + \|\sigma_{0,2} - \sigma_{0,1}\|_{BL}^*).$$

We will consider a velocity field of the form

$$(2.4) \quad v[m, \mu](x) := \max\{v_f(x) - v_I[m](x) - v_E[\mu](x), 0\}$$

where  $v_f : \Gamma \rightarrow \mathbb{R}^+$  is the desired velocity representing the speed of a car over a free road,  $v_I[m](x)$  is the interaction term due to the presence of other cars on the roads and  $v_E[\mu]$  is the interaction term with an external distribution  $\mu$ . Here we describe the velocities  $v_f$  and  $v_I$ , while in the next section we will consider velocities  $v_E[\mu]$  corresponding to the specific models discussed.

Concerning the free flow speed  $v_f(x)$ , which depends only on the state variable  $x$ , we assume that this function is positive, bounded and Lipschitz continuous on each arc  $e_j$  of the network  $\Gamma$ . Hence (H1)-(H2) are easily verified for  $v_f$ .

We consider a interaction velocity  $v_i$  given by the functional

$$v_I[m](x) := \int_{\Gamma} K(x, y) dm(y).$$

The interaction kernel  $K$  is defined as

$$(2.5) \quad K(x, y) = k(d_{\Gamma}(x, y)) \chi_{\mathcal{D}(x)}(y)$$

where  $k$  is a Lipschitz continuous, non increasing, bounded function representing the strength of interaction among cars in dependence on their distance and  $\chi_{\mathcal{D}(x)}$  is the characteristic function of the set  $\mathcal{D}(x)$  representing the visual field of the driver. We assume that a driver has only the knowledge of the distribution of cars on the roads adjacent to the current position and therefore we define the visual field as

$$\mathcal{D}(x) = \{y \in \Gamma : x \rightarrow y, d_{\Gamma}(x, y) \leq R\}$$

with  $R < L_0$  and  $L_0$  defined in (2.1). Hence it follows that, given  $x \in e_k$ , if  $V = \pi_k(L_k) \in \mathcal{V}$  we have  $\mathcal{D}(x) \subset e_k \cup (\bigcup_{e_j \in \text{Out}(V_i)} e_j)$ . We prescribe for any  $e_j \in \text{Out}(V)$  a weight  $\alpha_{kj}$  satisfying

$$0 \leq \alpha_{kj} \leq 1, \quad \sum_{j=1}^{|\mathcal{E}|} \alpha_{kj} = 1,$$

$$\alpha_{kj} = 0 \quad \text{if either } e_k \cap e_j = \emptyset \text{ or } e_j \rightarrow e_k.$$

where the coefficients  $\alpha_{kj}$  represent the priority of a given route in the choice of the driver depending on the basis of the observed traffic distribution. In conclusion, the interaction velocity at  $x \in e_k$  is given

$$v_I[m](x) = \sum_{j=1}^{|\mathcal{E}|} \alpha_{kj} \int_{\Gamma} k(d_{\Gamma}(x, y)) \chi_{\mathcal{D}(x) \cap (e_k \cup e_j)}(y) dm(y).$$

Since the function  $K$  defined in (2.5) is nonnegative and bounded, then there exists a constant  $C > 0$  such that

$$0 \leq v_I[m](x) \leq Cm(\Gamma), \quad \forall x \in \Gamma,$$

$$|v_I[m_1](x) - v_I[m_2](x)| \leq C \|m_1 - m_2\|_{BL}^*, \quad \forall x \in \Gamma, \forall m_1, m_2 \in \mathcal{M}^+(\Gamma),$$

and therefore **(H1)** and **(H2)** are satisfied. The Lipschitz continuity with respect to  $x$  is more delicate and for its proof we refer to [6, Sect.5]. A specific example of function  $k$  is given by

$$k(x, y) = \frac{\rho_2}{(\rho_1 + d_{\Gamma}(x, y))^{\beta}}$$

which is inspired by a Cucker-Smale nonlocal interaction kernel (see [14]).

**2.3. Mobility optimization.** We introduce a class of optimization problems on networks involving the distribution  $m$ , given by the solution of (2.3), the external distribution  $\mu$  and a control variable  $u$  which has to be designed in order to minimize/maximize a given objective functional.

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  $m_0 \in \mathcal{M}^+(\Gamma)$  and an incoming traffic distribution  $\sigma_0 \in \mathcal{M}^+([0, T])$ , we consider the optimization problem

$$(2.6) \quad \begin{cases} \min\{J(m, \mu, u) : (m, \mu, u) \in \mathcal{X}\}, \\ \text{subject to the state equation (2.3)}. \end{cases}$$

It is convenient to rewrite the previous minimization problem in the following equivalent form

$$(2.7) \quad \min\{J(m, \mu, u) + \mathbb{1}_A(m, \mu, u) : (m, \mu, u) \in \mathcal{X}\},$$

where  $A := \{(m, \mu, u) \in \mathcal{X}; m \text{ solves (2.3)}\}$  and  $\mathbb{1}_A$  is the indicator function of the set  $A$  defined as

$$\mathbb{1}_A(x) := \begin{cases} 0, & x \in A, \\ +\infty & \text{otherwise.} \end{cases}$$

A straightforward application of the direct method in Calculus of Variations gives the following existence result for the minima of (2.7).

**Theorem 2.2.** *Assume that*

- $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  is bounded from below;
- $J$  is lower semicontinuous in  $\mathcal{X}$ , i.e. for any  $(m_n, \mu_n, u_n) \subset \mathcal{X}$  such that  $(m_n, \mu_n, u_n) \rightarrow (m, \mu, u)$ , it holds  $J(m, \mu, u) \leq \liminf_{n \rightarrow \infty} J(m_n, \mu_n, u_n)$ ;
- the set  $A$  is closed under the topology induced by  $\|\cdot\|_{\mathcal{X}}$ .

Then the minimization problem (2.6) has a solution.

A typical example of functional to be minimized is of the form

$$(2.8) \quad J(m, \mu, u) := - \int_0^T \int_{\Gamma} v[m_t, \mu_t] dm_t(y) dt + \int_{\Gamma \times [0, T]} f(x, t, u) dm(x, t),$$

where the first term in (2.8) 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  $m_t$  in a closed region  $B$  during the time interval  $[0, T]$ . Another interesting class of control problems is the Minimum Time control introduced, in a measure theoretic setting, in [9, 10].

### 3. MODEL EXAMPLES: TRAFFIC LIGHTS AND AUTONOMOUS CARS

This section is devoted to some applications of the abstract setting previously described with the discussion of two significative problems in traffic flow optimization. In the first example, we optimize the duration of traffic lights in order to improve the circulation on the road network; in the second example, we aim to regulate the traffic flow by a fleet of autonomous car.

For both these models we assume that the control variable  $u$  influences the traffic flow distribution  $m$  only by means of an external distribution  $\mu = \mu[u]$ . Hence the functional to be minimized in (2.7) is of the form  $J(m, u)$  with  $m$  subject to (2.3) and  $\mu$  determined by another dynamical system for a given initial configuration  $\mu_0$ .

**3.1. 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 [18]. 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 located 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

$$(3.1) \quad \sum_{j \in \text{Inc}(V)} \int_0^T u_j(t) \delta_V(y) dt,$$

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 than  $T_i^R$  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

$$(3.2) \quad \begin{aligned} &\text{if } u_j(\tau_1^+) = 1, \text{ then } |\tau_1 - \tau_2| < T^R, \\ &\text{if } u_j(\tau_1^+) = 0, \text{ then } |\tau_1 - \tau_2| > T^G. \end{aligned}$$

Moreover we assume that a traffic light can be green only for one of the incoming roads in a junction, i.e.

$$(3.3) \quad \begin{aligned} &\sum_{j \in \text{Inc}(V)} u_j + 1 = N \\ &T^R \geq (N - 1)T^G \end{aligned}$$

where  $N = \#(\text{Inc}(V))$ .

Denote by  $\mathcal{F} \subset \mathcal{E}$  the set of the arcs containing a traffic light. Recalling (3.1), we consider the measure  $d\mu(x, t) = \sum_{j=1}^{|\mathcal{E}|} u_j(t) d\mu^j(x, t)$  on  $\Gamma_T$  where  $d\mu^j(x, t) \equiv 0$  if  $e_j \notin \mathcal{F}$  and  $d\mu^j(x, t) = \delta_V(x) dt$  if  $e_j \in \mathcal{F} \cap \text{Inc}(V_i)$ . 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

$$(3.4) \quad \mathcal{U} = \{u = \{u_j\}_{j=1, \dots, |\mathcal{F}|} : u_j \in BV([0, T], \{0, 1\}) \text{ and satisfies (3.2), (3.3)}\}$$

To describe the interaction of the drivers with the traffic lights, we define an external velocity term  $v_E[\mu]$  in (2.4). Fixed an arc  $e_j \in \mathcal{F} \cap \text{Inc}(V)$ , then the restriction of  $v_E[\mu]$  to the arc  $e_j$  is given by

$$v_E^j[\mu](x) := \int_{\Gamma} H(x, y) d\mu_t(y) = u_j(t) H(x, V) \delta_{e_j}(x).$$

We assume that the interaction kernel  $H$  is given by

$$(3.5) \quad H(x, y) = \begin{cases} v_f \max \left\{ \left( 1 - \frac{d_{\Gamma}(x, y)}{R} \right), 0 \right\}, & \text{if } x \rightarrow y, d_{\Gamma}(x, y) \leq R, \\ 0 & \text{otherwise,} \end{cases}$$

where  $v_f$  is the desired velocity and  $R \leq L_0$ , for  $L_0$  as in (2.1), is the visibility radius. 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 (3.4) satisfies the hypotheses of Theorem 2.2 for  $\mathcal{X} = \mathcal{M}_M^+(\Gamma_T) \times \mathcal{M}_M^+(\Gamma_T) \times \mathcal{U}$ .

**Lemma 3.1.** *The set of positive measures with bounded mass  $\mathcal{M}_M^+(\Gamma_T)$  is compact with respect to  $\|\cdot\|_{BL}^*$ .*

**Lemma 3.2.** *The set  $\mathcal{U}$  defined in (3.4) is compact in  $(BV^{|\mathcal{E}|}([0, T]), \|\cdot\|_{L^1})$ .*

**Lemma 3.3.** *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 3.2. The set  $A$  is closed under the topology induced by  $\|\cdot\|_{\mathcal{X}}$ .*

The proofs of the previous results are given in Appendix.

**3.2. 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 [20]) 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 [3] where the authors present an optimization problem for a transport equation in the euclidean space with the control represented by a second distribution  $\mu$  evolving according to another transport equation.

The dynamics of the autonomous cars is similar to the one of rest of the drivers, with the difference that it can be controlled in order to minimize the objective functional. Hence for a given initial distribution  $\mu_0$  (typically  $\mu_0 = \sum_{x \in \Gamma_a} \delta_x$  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

$$(3.6) \quad \begin{cases} \partial_t \mu^j + \partial_x (u \cdot v^j [m_t, \mu_t] \mu^j) = 0 & x \in e_j, t \in (0, T], j = 1, \dots, |\mathcal{E}| \\ \mu_{t=0}^j = \mu_0^j & x \in e_j, j = 1, \dots, |\mathcal{E}| \\ \mu_{V=\pi_j(0)}^j = \begin{cases} \sum_{\epsilon_k \in \text{Inc}(V)} q_{kj}(t) \mu_{V=\pi_k(1)}^k & \text{if } V \in \mathcal{I} \\ 0 & \text{if } V \in \mathcal{S}, \end{cases} & j = 1, \dots, |\mathcal{E}| \end{cases}$$

We assume that the velocity fields  $v[m_t, \mu_t]$  in (3.6) is the same of problem (2.3) and it is defined as in (2.4). 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 (2.4) 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  $\mu$  we add a

control term  $u$  and we assume that the control set is given by

$$(3.7) \quad \mathcal{U} = \text{Lip}_L(\Gamma_T, [0, 1]),$$

i.e. the set of Lipschitz functions from  $\Gamma \times [0, T]$  to  $[0, 1]$  with Lipschitz constant  $L > 0$ . In this way, if  $v[m_t, \mu_t]$  satisfies the assumptions of Theorem 2.1, then also  $u \cdot v[m_t, \mu_t]$  satisfies the same assumptions and therefore system (3.6), given  $(m_t)_{t \in [0, T]}$ , 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 (3.6) also differs from (2.3) 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 and adopt different routes. We assume that the matrix  $Q$  satisfies the assumptions in (2.2). Existence of a solution  $(m, \mu)$  to the coupled transport system (2.3)-(3.6) can be proved by a fixed point argument.

Given  $m \in C([0, T], \mathcal{M}^+(\Gamma))$ , consider the map

$$\Phi_1 : C([0, T], \mathcal{M}^+(\Gamma)) \rightarrow C([0, T], \mathcal{M}^+(\Gamma))$$

which associates to  $m$  the unique solution of (3.6). Similarly, given  $\mu \in C([0, T], \mathcal{M}^+(\Gamma))$ , define a map

$$\Phi_2 : C([0, T], \mathcal{M}^+(\Gamma)) \rightarrow C([0, T], \mathcal{M}^+(\Gamma))$$

which associates to  $\mu$  the solution  $\Phi_2(\mu)$  of (2.3). Hence, defined a map  $\Phi := (\Phi_1, \Phi_2)$ , the solution of the coupled system (2.3)-(3.6) is given by a fixed point of  $\Phi$ . By an argument similar to the one already used in [12, 13] for analogous results, it is possible to prove that  $\Phi$  is a contraction and therefore existence of a unique solution to the system (2.3)-(3.6) is obtained.

We conclude this section with the following lemma, which allows to apply Theorem 2.2 to the present case .

**Lemma 3.4.** *Assume  $\mathcal{X} = \mathcal{M}_M^+(\Gamma_T) \times \mathcal{M}_M^+(\Gamma_T) \times \mathcal{U}$ , where  $\mathcal{U}$  is defined by (3.7). Then, the set  $A$  is closed under the topology induced by  $\|\cdot\|_{\mathcal{X}}$ .*

#### 4. NUMERICAL SOLUTION VIA OPTIMALITY CONDITIONS

In this section we formally derive first-order optimality conditions for the optimization problem (2.6) 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.

**4.1. Optimality conditions.** We consider a network  $\Gamma$  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  $\mathcal{J} = \{V_0\}$ ,  $\mathcal{S} = \{V_1, V_2\}$ ,  $\mathcal{W} = \{V_3\}$ ,  $\text{Inc}(V_0) = \{e_1, e_2\}$  and  $\text{Out}(V_0) = \{e_3\}$ , as shown in Figure 1.

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

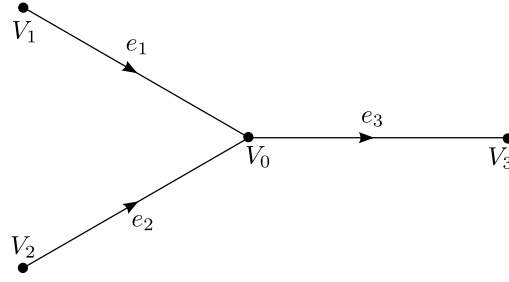


FIGURE 1. Example of 2-1 junction

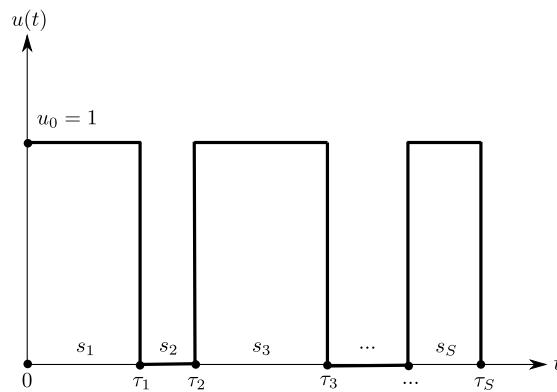
$u \in BV([0, T], \{0, 1\})$  is enough to describe the system, indeed we define edge-wise the velocity  $v$  by

$$\begin{aligned} v^1[u](x, t) &= \max\{v_f^1(x) - u(t)H(x, V_0), 0\}, \\ v^2[u](x, t) &= \max\{v_f^2(x) - (1 - u(t))H(x, V_0), 0\}, \\ v^3(x, t) &= v_f^3(x), \end{aligned}$$

where for  $j = 1, 2, 3$ ,  $v_f^j$  is the free flow speed on  $e_j$  and  $H$  is defined as in (3.5).

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  $s = (s_1, \dots, s_S) \in \mathbb{R}^S$ , 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  $\tau_i = \sum_{k=1}^i s_k$  for  $i = 1, \dots, S$ . Defining recursively  $u_i = 1 - u_{i-1}$  for  $i = 1, \dots, S$  and  $\tau_0 = 0$  we set (see Figure 2)

$$u(t) = u^s(t) = \sum_{i=0}^{S-1} u_i \chi_{[\tau_i, \tau_{i+1})}(t)$$

FIGURE 2. Reconstruction of control  $u$  from switching durations  $s = (s_1, \dots, s_S)$

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  $m$  has a density, i.e.  $dm = m(x, t)dx dt$  for some function  $m : \Gamma \times [0, T] \rightarrow \mathbb{R}$ , we want to minimize the cost functional

$$(4.1) \quad J(m, u^s) = - \int_0^T \int_{\Gamma} v[u^s](x, t)m(x, t) dx dt,$$

subject to

$$(4.2) \quad \begin{cases} \partial_t m^j + \partial_x(v^j m^j) = 0 & \text{in } e_j \times (0, T), j = 1, 2, 3 \\ m^j(\cdot, 0) = m_0^j & \text{in } e_j \end{cases}$$

We also assume null incoming traffic in the network during the whole evolution, imposing

$$(4.3) \quad m_{x=V_1}^1 = 0, \quad m_{x=V_2}^2 = 0, \quad t \in [0, T],$$

and the mass conservation condition at the internal vertex  $V_0$

$$(4.4) \quad m_{x=V_0}^3 = m_{x=V_0}^1 + m_{x=V_0}^2.$$

We formally apply the method of Lagrange multipliers in order to derive first-order optimality conditions. We define the Lagrangian as

$$\begin{aligned} L(m, u^s, \lambda) &:= J(m, u^s) + \int_0^T \int_{\Gamma} (-\partial_t \lambda - v \partial_x \lambda) m dx dt \\ &+ \int_{\Gamma} (\lambda(x, T)m(x, T) - \lambda(x, 0)m_0(x)) dx \\ &+ \sum_{j=1,2,3} \int_0^T (\lambda^j(V_j^E, t)v^j(V_j^E, t)m^j(V_j^E, t) - \lambda^j(V_j^I, t)v^j(V_j^I, t)m^j(V_j^I, t)) dt, \end{aligned}$$

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  $\lambda$  derive from the weak formulation of the transport equation on  $\Gamma$ .

We evaluate the derivatives of the Lagrangian with respect to  $m$  and  $s$  (recall that  $u = u^s$ ). We first consider an admissible increment  $w$  for  $m$  which preserves the boundary and transition conditions, i.e.

$$(4.5) \quad w^1(V_1, t) = 0, \quad w^2(V_2, t) = 0, \quad w^3(V_0, t) = w^1(V_0, t) + w^2(V_0, t) \quad t \in [0, T],$$

and we compute

$$(4.6) \quad \begin{aligned} \langle \partial_m L, w \rangle &= \int_0^T \int_{\Gamma} (-\partial_t \lambda - v \partial_x \lambda - v) w \, dx dt + \int_{\Gamma} \lambda(x, T) w(x, T) \, dx \\ &+ \int_0^T \sum_{j=1,2,3} (\lambda^j(V_j^E, t) v^j(V_j^E, t) w^j(V_j^E, t) - \lambda^j(V_j^I, t) v^j(V_j^I, t) w^j(V_j^I, t)) \, dt. \end{aligned}$$

Imposing  $\langle \partial_m L, w \rangle = 0$  for any admissible  $w$ , we get the following time-backward advection equation with a source term

$$(4.7) \quad -\partial_t \lambda^j - v^j \partial_x \lambda^j = v^j \quad \text{in } e_j \times (0, T), \, j = 1, 2, 3,$$

and the final condition

$$\lambda^j(x, T) = 0 \quad \text{in } e_j, \, j = 1, 2, 3.$$

Note that for (4.7),  $V_3$  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 (4.6), we have

$$\begin{aligned} &\int_0^T (\lambda^1 v^1 w^1(V_0, t) - \lambda^1 v^1 w^1(V_1, t) + \lambda^2 v^2 w^2(V_0, t) \\ &- \lambda^2 v^2 w^2(V_2, t) + \lambda^3 v^3 w^3(V_3, t) - \lambda^3 v^3 w^3(V_0, t)) \, dt = 0. \end{aligned}$$

By taking  $w$  compactly supported in a neighborhood of  $V_3$ , we get the boundary condition

$$\lambda^3(V_3, t) = 0 \quad \text{in } [0, T],$$

whereas for  $w$  compactly supported in a neighborhood of  $V_0$ , recalling (4.5), we get

$$(4.8) \quad \int_0^T \{(\lambda^1 v^1 - \lambda^3 v^3) w^1(V_0, t) + (\lambda^2 v^2 - \lambda^3 v^3) w^2(V_0, t)\} \, dt = 0.$$

The mass conservation condition (4.4) can be rewritten as

$$v^3(V_0, t) m^3(V_0, t) = v^1(V_0, t) m^1(V_0, t) + v^2(V_0, t) m^2(V_0, t) \quad t \in [0, T],$$

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  $m^1(V_0, t) = 0$  (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  $m^2(V_0, t) = 0$  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 (4.8) 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,

$$\begin{aligned} \lambda^1(V_0, t) v^1(V_0, t) &= \lambda^3(V_0, t) v^3(V_0, t) & \text{if } t \in \{v^1(V_0, t) \neq 0\}, \\ \lambda^2(V_0, t) v^2(V_0, t) &= \lambda^3(V_0, t) v^3(V_0, t) & \text{if } t \in \{v^2(V_0, t) \neq 0\}. \end{aligned}$$

We now compute the derivative of  $L$  with respect to  $u^s$  for an increment  $\varphi \in \mathbb{R}^S$

$$\begin{aligned} \langle \partial_s L, \varphi \rangle &= - \int_0^T \int_{\Gamma} \partial_s v \cdot \varphi (\partial_x \lambda + 1) m \, dx dt + \int_0^T \left\{ \sum_{j=1,2,3} \lambda^j(V_j^E, t) \partial_s v^j(V_j^E, t) \cdot \varphi m^j(V_j^E, t) \right. \\ &\quad \left. - \lambda^j(V_j^I, t) \partial_s v^j(V_j^I, t) \cdot \varphi m^j(V_j^I, t) \right\} dt. \end{aligned}$$

Recalling (4.3) and since  $v^3$  is independent of  $u^s$ , we get

$$\begin{aligned} \langle \partial_s L, \varphi \rangle &= \int_0^T \left\{ - \int_{e_1} \partial_s v^1 \cdot \varphi (\partial_x \lambda^1 + 1) m^1 \, dx - \int_{e_2} \partial_s v^2 \cdot \varphi (\partial_x \lambda^2 + 1) m^2 \, dx \right. \\ &\quad \left. + \lambda^1(V_0, t) \partial_s v^1(V_0, t) \cdot \varphi m^1(V_0, t) + \lambda^2(V_0, t) \partial_s v^2(V_0, t) \cdot \varphi m^2(V_0, t) \right\} dt, \end{aligned}$$

where

$$\partial_s v^1(x, t) \cdot \varphi = -H(x, V_0) \nabla_s u^s(t) \cdot \varphi, \quad \partial_s v^2(x, t) \cdot \varphi = H(x, V_0) \nabla_s u^s(t) \cdot \varphi$$

and

$$\nabla_s u^s(t) \cdot \varphi = \sum_{i=1}^S (-1)^{u_i-1} \delta_{\tau_i}(t) \varphi_i.$$

We conclude

$$\begin{aligned} \langle \partial_s L, \varphi \rangle &= \sum_{i=1}^S (-1)^{u_i-1} \left\{ \int_{e_1} H(x, V_0) (\partial_x \lambda^1(x, \tau_i) + 1) m^1(x, \tau_i) \, dx - \lambda^1(V_0, \tau_i) H(V_0, V_0) m^1(V_0, \tau_i) \right. \\ &\quad \left. - \int_{e_2} H(x, V_0) (\partial_x \lambda^2(x, \tau_i) + 1) m^2(x, \tau_i) \, dx + \lambda^2(V_0, \tau_i) H(V_0, V_0) m^2(V_0, \tau_i) \right\} \varphi_i. \end{aligned}$$

Summarizing, the dual problem for (4.2)-(4.3)-(4.4) 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  $i = 1, \dots, S$ , the optimal solution  $(m, u^s, \lambda)$  should satisfy, for all  $\bar{s} \in \mathbb{R}^S$  such that  $T^G < \bar{s}_i < T^R$ , the variational inequality

$$(4.9) \quad \langle \partial_s L(m, u^s, \lambda), \bar{s} - s \rangle \geq 0.$$

*Remark 4.1.* If the velocity field contains the drivers interaction term, then the dual problem for (4.2)-(4.3)-(4.4) is given by

$$\begin{cases} -\partial_t \lambda^j - v^j \partial_x \lambda^j - \nu * (m \partial_x \lambda) = v^j + \nu * m & \text{in } e_j \times (0, T), j = 1, 2, 3 \\ \lambda^j(\cdot, T) = 0 & \text{in } e_j \end{cases}$$

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  $y \rightarrow x$  and not  $x \rightarrow y$  as in (2.5).

**4.2. 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  $m$ , then we solve the backward equation in  $\lambda$ , 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  $m^{(k)}$  with control  $u^{(k)}$ ;
- Step 4. Solve the backward problem for  $\lambda^{(k)}$  with control  $u^{(k)}$ ;
- Step 5. Compute  $J^{(k+1)} = J(m^{(k)}, s^{(k)})$ .  
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  $(m^{(k)}, u^{(k)}, \lambda^{(k)})$ ;
- Step 7. Update  $s^{(k)} \leftarrow \Pi_{\{T^G, T^R\}}(s^{(k)} - \beta \partial_s L(m^{(k)}, u^{(k)}, \lambda^{(k)}))$ ,  $k \leftarrow k + 1$  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  $(m^{(k)}, u^{(k)}, \lambda^{(k)})$  as an approximate solution of the optimal control problem for (4.1).

In the actual implementation of the algorithm, we employ a standard scheme for conservation laws with a superbee flux limiter, to solve the forward equation in  $m$ . On the other hand, the adjoint advection equation in  $\lambda$  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  $\beta$  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  $i = 1, \dots, S$ .



**4.3. Numerical experiments.** As a preliminary test we compare the local and the non-local 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 (2.5) 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 final time  $T = 1.25$ , the free flow speed  $v_f^1 \equiv 1$  and the initial distribution  $m_0(x) = \chi_{[0.1,0.15]}(x)$ .

Figure 3 shows the evolution of  $m$  and  $v$  at different times for both local and nonlocal cases. Note that the velocity  $v$  decreases from  $v_f^1$  to zero with a linear ramp while approaching the traffic light, according to the definition (3.5) for  $H$ .

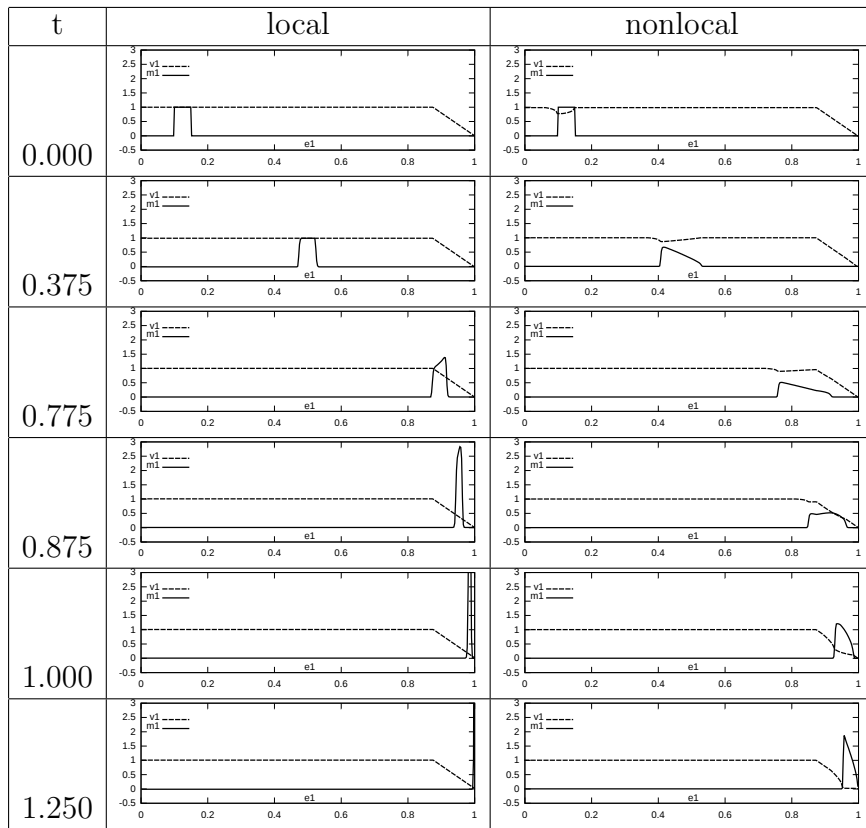


FIGURE 3. Red traffic light: local case vs nonlocal case

In the local case  $v$  does not depend on time, since  $u$  is constant. The density  $m$  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  $m$  and  $v$ . The initial density readily activates the nonlocal term in  $v$ , and  $m$  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  $m$ .

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  $\tau$  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  $v_1^f = v_2^f = v_3^f \equiv 1$ . We also assume that, apart from  $m_0$ , no additional mass enters or leaves the network for all  $t \in [0, T]$ .

We start with  $m_0 = (m_0^1, m_0^2, m_0^3) = (\chi_{[0.1, 0.15]}(x), \chi_{[0.6, 0.65]}(x), 0)$ , i.e. two distributions of equal mass on  $e_1$  and  $e_2$  that arrive at the traffic light at different times ( $m_2$  first and then  $m_1$ ). In Figure 4(a) we plot the corresponding (normalized) mean velocity  $\bar{v}(\tau) = -J(m, u^s)/M$  as a function of  $\tau$ , where  $M = \int_0^T \int_{\Gamma} m(x, t) dx dt$ .

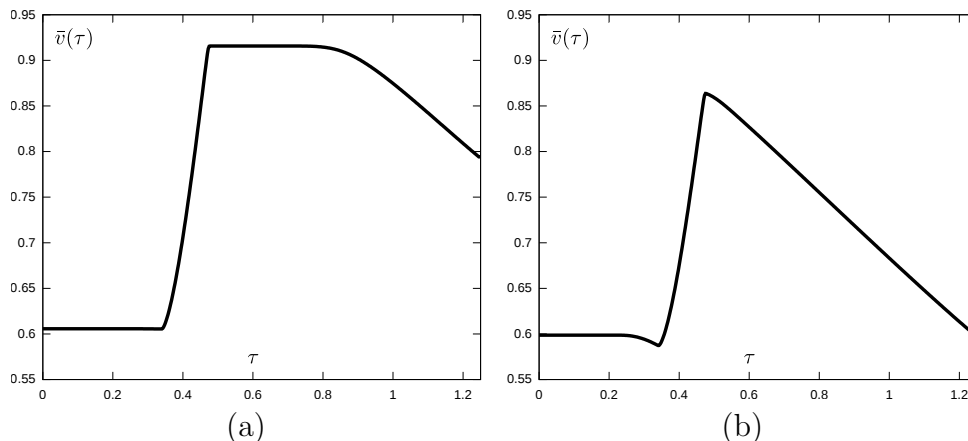


FIGURE 4. Mean velocity for a single switch of the traffic light: well separated (a) vs overlapping (b) densities

The scenario is pretty clear. If the switch occurs before  $m_2$  reaches the traffic light, then only  $m_1$  will move from  $e_1$  to  $e_3$  and the mean velocity cannot improve. For larger values of  $\tau$ , also  $m_2$  will gradually move to  $e_3$ , and  $\bar{v}(\tau)$  increases. If now the switch is placed just after  $m_2$  leaves  $e_2$  and before  $m_1$  approaches the traffic light,  $\bar{v}$  exhibits a plateau and 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 speed. Finally, as  $\tau$  keeps increasing up to  $T$ ,  $m_1$  starts getting stuck at the traffic light, and  $\bar{v}(\tau)$  decreases.

Now let us repeat the exhaustive computation of the mean velocity  $\bar{v}(\tau)$  with  $m_0 = (m_0^1, m_0^2, m_0^3) = (\chi_{[0.6, 0.65]}(x), \chi_{[0.6, 0.65]}(x), 0)$ , two distributions of equal mass on  $e_1$  and  $e_2$ , starting at the same distance from the traffic light. Figure 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  $m_2$  leaves  $e_2$  (see Figure 6 below).

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 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.

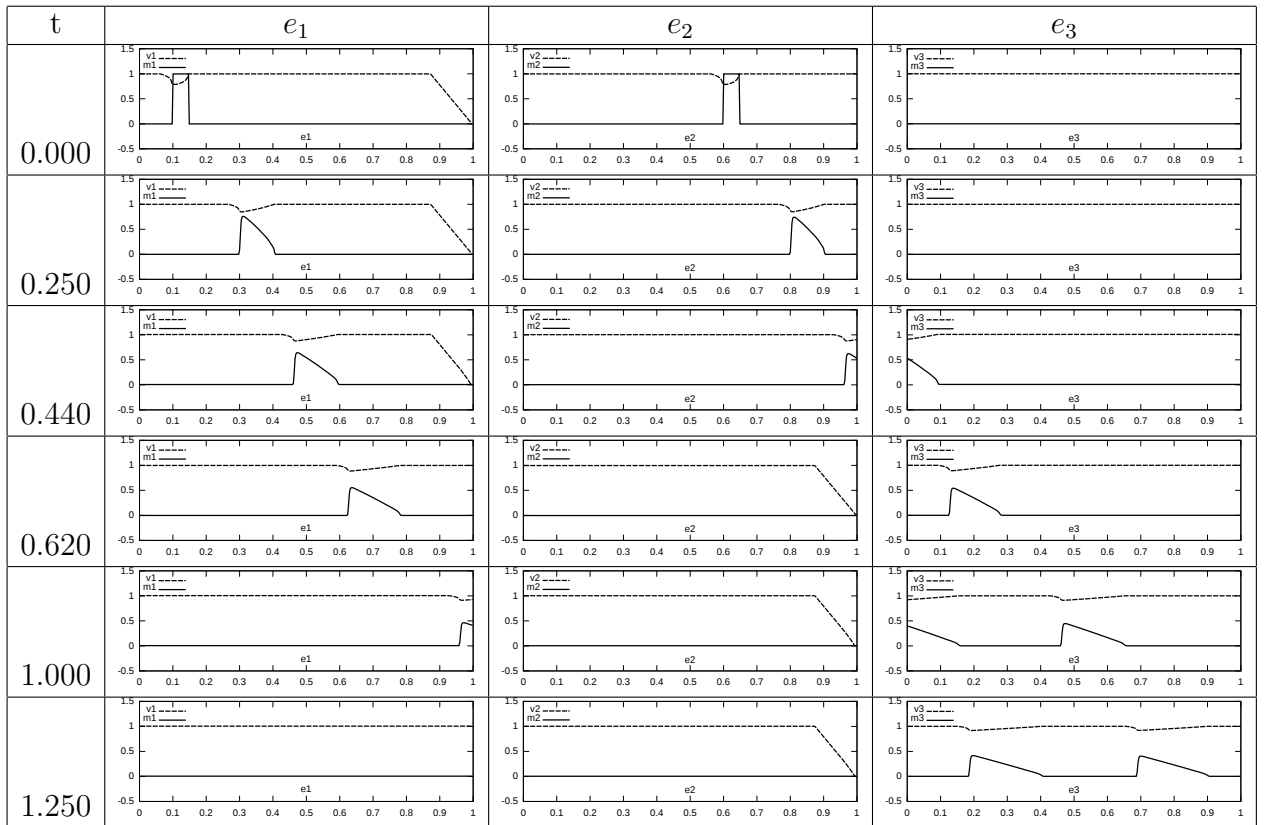


FIGURE 5. Optimal solution for well separated densities

Similarly, Figure 6 refers to the case of overlapping densities. We clearly observe that on  $e_1$  the traffic is stopped until  $m_2$  leaves  $e_2$ .

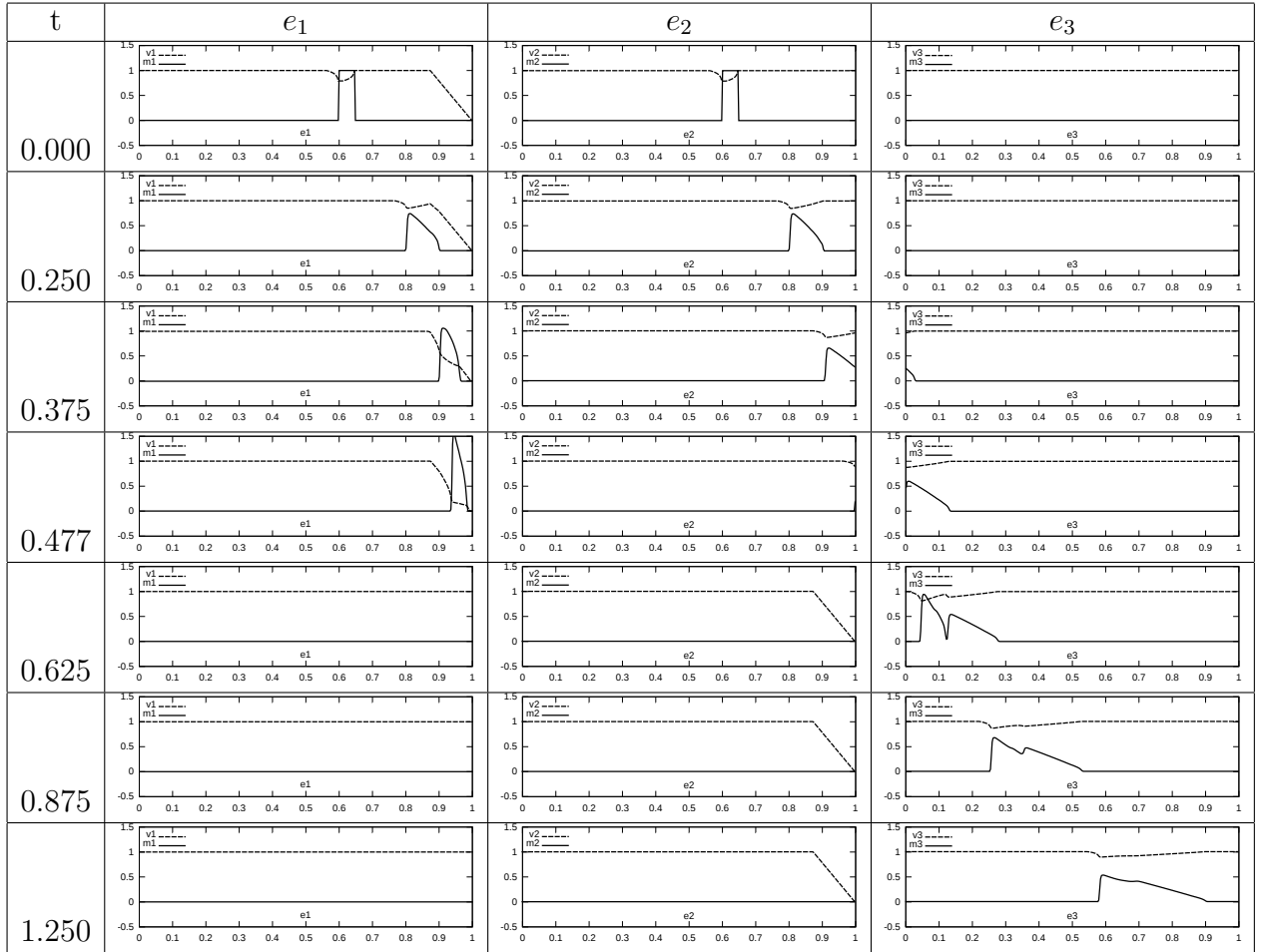


FIGURE 6. Optimal solution for overlapping densities

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 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

$$m_0^1(x) = \chi_{[0.1,0.15]}(x) + \chi_{[0.4,0.45]}(x), \quad m_0^2(x) = \chi_{[0.1,0.15]}(x) + \chi_{[0.6,0.65]}(x), \quad m_0^3(x) = 0.$$

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 7 shows the corresponding evolution at different times. We observe that the first switch occurs before  $m_2$  approaches the traffic light. This allows the first block of  $m_2$  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  $m_1$  can leave

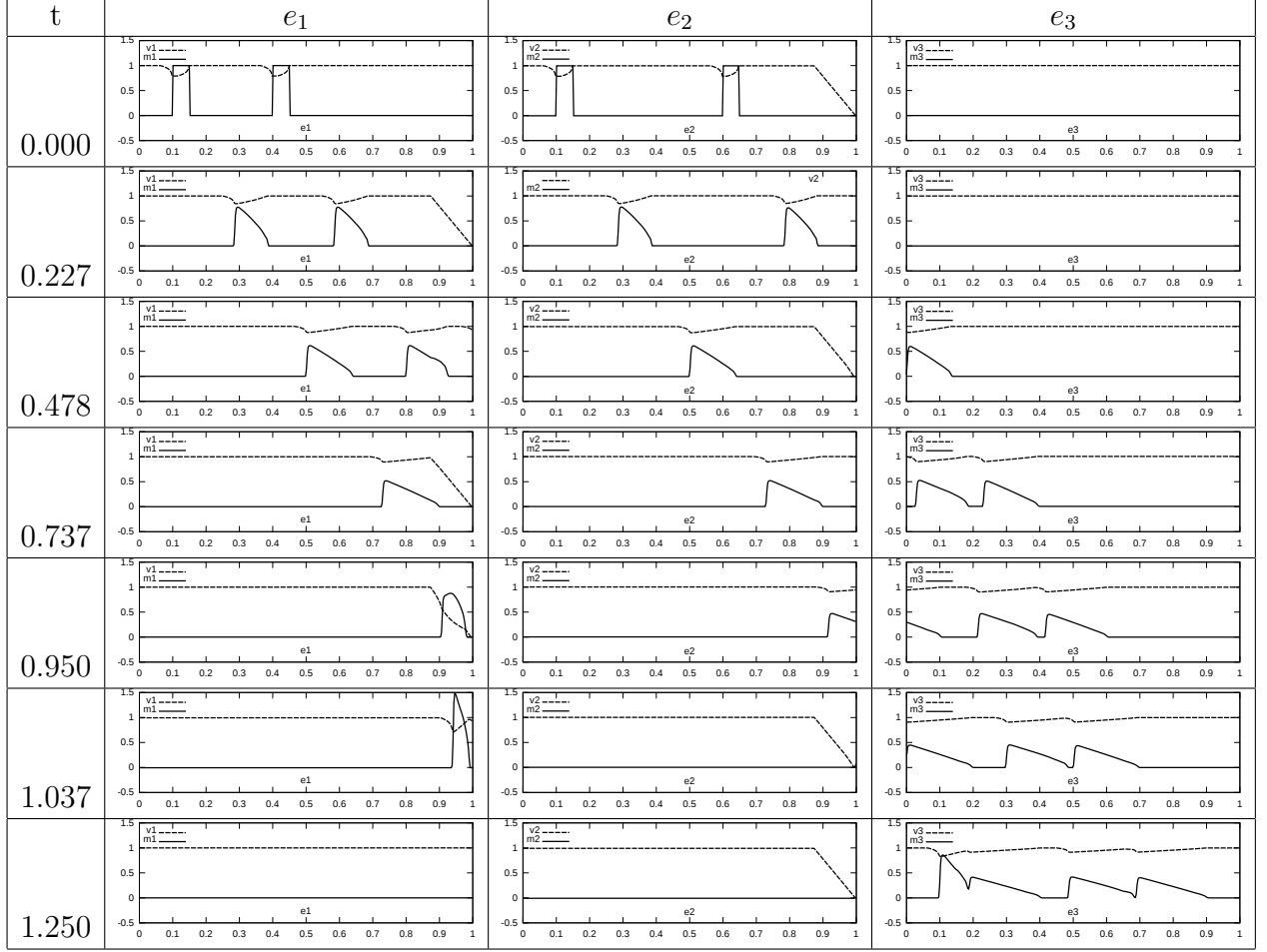


FIGURE 7. Optimal solution for a traffic light with 4 switches

$e_1$  almost undisturbed before the traffic light switches again. Now, the remaining densities on  $e_1$  and  $e_2$  are in overlapping configuration,  $m_2$  goes first, while  $m_1$  stops. Finally, the last switch occurs just after  $m_2$  leaves  $e_2$ , so that also  $m_1$  can move to  $e_3$  for the remaining time.

#### APPENDIX A. SOME COMPLEMENTARY RESULTS FOR THE VARIATIONAL PROBLEMS

*Proof of Lemma 3.1.* Assume without loss of generality that  $M = 1$ . It is well known that for  $\mu \in \mathcal{M}_M^+(\Gamma_T)$ ,  $|\mu|_{TV} = \mu(\Gamma_T) \leq 1$ .

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 3.2.* Since (3.3) is just a condition which defines the dependence among the components of  $u \in \mathcal{U}$ , we prove the compactness of

$$\mathcal{U} = \{u \in BV([0, T], \{0, 1\}) \text{ and } u \text{ satisfies (3.2)}\}.$$

Let  $\{u_n\}_{n \in \mathbb{N}} \subset \mathcal{U}$ . Denote by  $\tau_i^n$  the switching times of  $u_n$ . By (3.2), 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  $n \in \mathbb{N}$ . Assume now that, w.l.o.g.,  $u_n(0) = 1$  for every  $n \in \mathbb{N}$  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 there exists  $N \in \mathbb{N}$  such that  $\#(I_n) = N$  for all  $n \in \mathbb{N}$ . Since  $I_n \subset [0, T]$ , applying the Cantor diagonal procedure, it follows that there exists a subsequence  $(I_{n_k})_{k \in \mathbb{N}}$  such that  $\tau_i^{n_k} \rightarrow \tau_i$  for  $i = 1, \dots, N$ . In this way, we define a candidate  $u$  as limit for the subsequence  $u_{n_k}$  from the switching times set  $\{\tau_1, \dots, \tau_N\}$  and  $u(0) = 1$ . To conclude, we only need to show that  $u_{n_k} \rightarrow u$  in  $L^1$ . By construction,

$$\|u_{n_k} - u\|_{L^1} = \sum_{i=1}^N |\tau_i^{n_k} - \tau_i| \leq N \sup_{i=1, \dots, N} |\tau_i^{n_k} - \tau_i| \xrightarrow{k \rightarrow \infty} 0$$

□

*Proof of Lemma 3.3 (traffic lights).* In this case, the distribution  $\mu$  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 (3.4). Let  $(m_n, u_n)_{n \in \mathbb{N}} \subset A$  such that  $(m_n, u_n) \rightarrow (m, u)$  with respect the norm  $\|\cdot\|_{BL}^* + \|\cdot\|_{L^1}$ . The closure on the first component derives from the proof of Lemma 4.1 in [3] and the results in [6].

Instead, the closure on the second component derives from the compactness of  $\mathcal{U}$ . Indeed, there exists a subsequence  $(u_{n_k})_{k \in \mathbb{N}}$  which converges to  $\tilde{u} \in \mathcal{U}$ , but it also converges to  $u$  by assumption. Then, it follows that  $u = \tilde{u} \in \mathcal{U}$ .

□

*Proof of Lemma 3.4 (autonomous cars).* 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  $\|\cdot\|_{BL}^* + \|\cdot\|_{BL}^* + \|\cdot\|_{\infty}$ .

□

## REFERENCES

1. L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.
2. 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.

3. M. Bongini, G. Buttazzo, Optimal control problems in transport dynamics, *Math. Models Methods Appl. Sci.* 27 (2017), 427-451.
4. A. Bressan, S. Čanić, M. Garavello, M. Herty, B. Piccoli, Flows on networks: recent results and perspectives, *EMS Surv. Math. Sci.* 1 (2014), no. 1, 47-111.
5. F. Camilli, R. De Maio, A. Tosin, Transport of measures on networks, *Netw. Heterog. Media*, 12 (2017), 191 - 215.
6. F. Camilli, R. De Maio, A. Tosin, Measure-valued solutions to transport equations on networks with nonlocal velocity, *J. Differential Equations*, 254 (2018), 7213 - 7241.
7. J. A. Cañizo, J. A. Carrillo, J. Rosado, A well-posedness theory in measures for some kinetic models of collective motion, *Math. Models Methods Appl. Sci.* 21 (2011), no. 3, 515-539.
8. E. Cascetta, *Transportation Systems Analysis*, Springer, Heidelberg (2009).
9. G. Cavagnari, A. Marigonda, B. Piccoli, Averaged time-optimal control problem in the space of positive Borel measures, *Esaim COCV* (2018), in press.
10. G. Cavagnari, A. Marigonda, K.T. Nguyen, F.S. Priuli, Generalized control systems in the space of probability measures, *Set-Valued Var. Anal.* (2017), 1 - 29.
11. R. M. Colombo, M. Garavello, M. Lécureux-Mercier, A class of nonlocal models for pedestrian traffic, *Math. Models Methods Appl. Sci.* 22 (2012), no. 4, 1150023.
12. 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
13. E. Cristiani, B. Piccoli, and Tosin, A., *Multiscale Modeling of Pedestrian Dynamics*, Springer International Publishing, 2014, MS&A: Modeling, Simulation and Applications, Vol. 12
14. F. Cucker and S. Smale, Emergent behaviour in flocks, *IEEE Trans. on Auto. Con.* 52 (2007), 852–862.
15. M. Garavello, K. Han, B. Piccoli, *Models for vehicular traffic on networks*, AIMS Series on Applied Mathematics, American Institute of Mathematical Sciences, Springfield, MO, 2016.
16. G. Gentile, L. Meschini, Using dynamic assignment models for real-time traffic forecast on large urban networks, in *“2nd Int. Conf. on Mod. and Tech. for Int. Trans. Sys.”*, Leuven, Belgium (2011).
17. P. Goatin, S. Scialanga, Well-posedness and finite volume approximations of the LWR traffic flow model with non-local velocity, *Netw. Heterog. Media*, 11 (2016), no. 1, pg. 107-121
18. S. Göttlich, U. Ziegler, Traffic light control: a case study, *Dis. and Con. Dyn. Sys.* 7 (2014), 483–501.
19. V. Trozzi, I. Kaparias, M. G.H. Bell, G. Gentile, A dynamic route choice model for public transport networks with boarding queues, *Trans. Plan. and Tech.* 36 (2013), 44-61.
20. 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.

DIPARTIMENTO DI MATEMATICA E FISICA, UNIVERSITÀ DEGLI STUDI ROMA TRE, L.GO S.LEONARDO MURIALDO 1, 00146 ROMA, ITALY

*Email address:* cacace@mat.uniroma3.it

DIPARTIMENTO DI SCIENZE DI BASE E APPLICATE PER L'INGEGNERIA, "SAPIENZA" UNIVERSITÀ DI ROMA, VIA SCARPA 16, 00161 ROMA, ITALY

*Email address:* camilli@sbai.uniroma1.it

DIPARTIMENTO DI SCIENZE DI BASE E APPLICATE PER L'INGEGNERIA, "SAPIENZA" UNIVERSITÀ DI ROMA, VIA SCARPA 16, 00161 ROMA, ITALY

*Email address:* raul.demaio@sbai.uniroma1.it

DIPARTIMENTO DI SCIENZE MATEMATICHE "G. L. LAGRANGE", POLITECNICO DI TORINO, CORSO DUCA DEGLI ABRUZZI 24, 10129 TORINO, ITALY

*Email address:* andrea.tosin@polito.it