



**HAL**  
open science

## Optimal scenario for road evacuation in an urban environment

Mickael Bestard, Emmanuel Franck, Laurent Navoret, Yannick Privat

► **To cite this version:**

Mickael Bestard, Emmanuel Franck, Laurent Navoret, Yannick Privat. Optimal scenario for road evacuation in an urban environment. Zeitschrift für Angewandte Mathematik und Physik = Journal of Applied mathematics and physics = Journal de mathématiques et de physique appliquées, In press, 10.48550/arXiv.2310.15359 . hal-04253010v2

**HAL Id: hal-04253010**

**<https://hal.science/hal-04253010v2>**

Submitted on 3 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Optimal scenario for road evacuation in an urban environment

Mickael Bestard\* <sup>1</sup>, Emmanuel Franck <sup>† 2</sup>, Laurent Navoret <sup>‡ 1</sup>, and Yannick Privat <sup>§ 3,4</sup>

<sup>1</sup>IRMA, Université de Strasbourg, CNRS UMR 7501, Inria, 7 rue René Descartes, 67084 Strasbourg, France.

<sup>2</sup>Inria, IRMA, Université de Strasbourg, CNRS UMR 7501, 7 rue René Descartes, 67084 Strasbourg, France.

<sup>3</sup>Université de Lorraine, CNRS, Institut Elie Cartan de Lorraine, Inria, BP 70239 54506 Vandœuvre-lès-Nancy Cedex, France.

<sup>4</sup>Institut Universitaire de France (IUF)

June 3, 2024

## Abstract

How to free a road from vehicle traffic as efficiently as possible and in a given time, in order to allow for example the passage of emergency vehicles? We are interested in this question which we reformulate as an optimal control problem. We consider a macroscopic road traffic model on networks, semi-discretized in space and decide to give ourselves the possibility to control the flow at junctions. Our target is to smooth the traffic along a given path within a fixed time. A parsimony constraint is imposed on the controls, in order to ensure that the optimal strategies are feasible in practice.

We perform an analysis of the resulting optimal control problem, proving the existence of an optimal control and deriving optimality conditions, which we rewrite as a single functional equation. We then use this formulation to derive a new mixed algorithm interpreting it as a mix between two methods: a descent method combined with a fixed point method allowing global perturbations. We verify with numerical experiments the efficiency of this method on examples of graphs, first simple, then more complex. We highlight the efficiency of our approach by comparing it to standard methods. We propose an open source code implementing this approach in the Julia language.

**Keywords:** Traffic network, Optimal Control, Fluid model, Hyperbolic PDE, optimization methods

**AMS classification:** 35Q49, 65M08, 49K15, 49K30.

## 1 Introduction

**Road traffic modeling.** With the concentration of populations in cities where transport flows are constantly increasing, road traffic modeling has become a central issue in urban planning. Whether it is to configure traffic lights, to adapt public transport offers or to manage in an optimal way situations of high congestion. The knowledge of the behavior of a road network is an important issue for the

---

\*mickael.bestard@math.unistra.fr

†emmanuel.franck@inria.fr

‡laurent.navoret@math.unistra.fr

§yannick.privat@univ-lorraine.fr

management of a crisis in an urban environment. It is indeed necessary to quickly predict traffic, especially traffic jams, in order to organize rescue operations. Before creating decision support tools for crisis management involving road traffic, a first step is to develop a method for monitoring road traffic.

In this work, we will focus on the evacuation of a traffic road in finite time. It can be, for example, an axis that we wish to free in order to provide access to first aid. We approach this question in the form of an optimal control problem of a graph where each edge corresponds to a traffic road. The main objective of this work is the determination of a prototype algorithm answering this question.

**State of the art.** Depending on the situation of interest, several types of approaches co-exist to model this phenomenon, depending on the scale and the desired accuracy. At the microscopic scale, trajectories (speed and position) of each vehicle can be described through Lagrangian microscopic models such as car-following models [22, 43, 47] or cellular automaton models [7, 27, 37]. Kinetic traffic models were also introduced in the 60's [14, 16, 18, 36, 38, 40, 45] and deal with vehicles on a mesoscopic scale in the form of distributions in position-velocity space.

Finally, on a macroscopic scale, the vehicle flow is then considered as a continuous medium. This is the case of the famous model introduced by Lighthill, Whitham and Richards (denoted LWR in what follows) [33, 42]. We will use it in the rest of this article, since the continuous description of the flow is particularly well adapted to the writing of a dynamic programming problem. It is notable that this first order model introduced in the 50's was later extended to Euler-type systems with viscosity, of second order [3, 39, 46, 50].

In this paper, we introduce and analyze an optimal control problem modeling an evacuation scenario for an axis belonging to a road network. A close control model has already been introduced in [24], but under restrictive assumptions about the flow regime in order to avoid congestion phenomena, which is what we want to avoid. It has been in particular highlighted that the flow may not be invertible with respect to control parameters, meaning that density at the nodes may not be reconstructed from boundary conditions. A related model, also applicable to crisis management, is discussed in [35], where the objective function is the mean spatial velocity at final time computed on specific path between e.g. an accident location and an hospital. The control variables are the traffic distribution parameters at junctions. The exact solution is computed locally for each junction and a sub-optimal global solution is retrieved by applying the local optimal solution for each junction. In [23], a method to manage variable speed limits combined with coordinated ramp metering within the framework of the Lighthill-Whitham-Richards (LWR) network model has been introduced. They consider a "first-discretize-then-optimize" numerical approach to solve it numerically. Note that the optimal control problem we deal with in the following has some similarities with the one proposed by the authors of this article, but differs in several respects: our control variables aim to control not just the ramps metering, but an entire axis. Furthermore, our approach to solving the underlying optimal control problem is quite different, in particular our use of optimality conditions leading to an efficient solution algorithm.

In [41], a similar control problem is discussed, which is addressed using a piecewise-linear flux framework driving to a much simplified adjoint system involving piecewise constant Jacobian matrices in time and space. This approach enables faster gradient computation, allowing them to tackle real-world simulations related to ramp-metering configuration. In our study, we opted to use the model introduced in [13] and dealt with the non-linear flux using automatic differentiation techniques to compute the gradient. This kind of approach has also been used in [6] for air traffic flow with a modified LWR-based network model.

We also mention several contributions aiming at dealing with concrete or real-time cases. The paper [19] deals with this kind of control problem for simplified nonlinear (essentially without considering congestion) and linear formulations of the model, considering large-scale networks. In [28], an algorithm for real-time traffic state estimation and short-term prediction is introduced.

Several other close control problems have been investigated in the past: in [25], the authors use switching controls to deal with traffic lights at an  $8 \times 4$  junction in a piecewise linear flow framework.

The paper [2] deals with an optimal control problem of the same kind, formulated from the continuous LWR model, solved numerically using a heuristic random parameters search. Other works have looked at traffic regulation by comparing and discussing the use of lights and circles [12]. We refer to [1, 5] for a broad view on this topic.

Regarding now controlled microscopic models, let us mention [34] where an approach based on model predictive control (MPD) is developed and [4] using a reinforcement learning algorithm. In the book [44], the last chapter is devoted to the modeling of various optimization problems related to road traffic.

Concerning now the control of mesoscopic and macroscopic models, there are software contributions seeking in particular to bring real time answers and guidance tools in case of accident. Let us also mention [26, 31] dealing in part with traffic prediction and control problems.

This article uses a controlled model close to the one studied in [24] which is devoted to the control of the macroscopic LWR model on a directed graph using the framework introduced in [13]. However, we are interested in taking into account some properties of the model that were not addressed in [24], such as congestion phenomena.

Hence, we will not focus on individual paths since we are interested in the overall dynamics with a focus on congestion. Particular attention is paid to the modeling of junctions as they introduce a nonlinear coupling between roads.

**Main objectives.** Our goal is to design a numerical method to tackle the control of the flow of vehicles in all fluid regimes, including saturated regimes/congestions, and using dynamical barriers at each road in the graph.

We thus propose a road traffic model in the form of a controlled graph at each junction. This models an urban area structured by roads. For practical reasons, even if we wish to take into account possible congestion/shocks in the system, we position ourselves in a differentiable framework by adding weak diffusion terms in the system, through a semi-discretization in space. Hence, the resulting problem consists in minimizing the vehicle density at the final time on a given route under ODE constraints, obtained by a finite volume semi-discretization of the LWR system. We then enrich this problem with a constraint on the number of simultaneously active roadblocks in order to take into account staffing issues. We use this analysis to build a new algorithm adapted to this problem, mixing a classical gradient approach with a fixed-point method, which allows non-local perturbations. One of the advantages of this approach is that the constructed sequence can escape from a minimum well and survey a larger part of the admissible set than a standard gradient type algorithm.

**Plan of this article.** While the macroscopic model describing the flow is very standard (Lighthill-Whitham-Richards, 1955) the problem of flow distribution at junctions, inspired by more recent works [20, 21], realizes a nonlinear coupling between the different edges. The distribution of vehicles at junctions is thus modeled by an optimal allocation process that depends on the maximum possible flows at the junctions, through a linear programming (LP) problem targeting the maximization of incoming flows. In order to influence the traffic flow, we introduce control functions that are defined at each road entrance and acts as barriers by weighting the capacity of a road leaving a junction to admit new vehicles.

Section 2.1 is devoted to the description of the traffic flow model we consider, and to the optimal control problem modeling the best evacuation scenario for a road.

After investigating the existence of optimal control in section 3.1, we then derive necessary optimality conditions that we reformulate in an exploitable way in section 3.2. We hence introduce an optimization algorithm in section 4, based on a hybrid combination of a variable step gradient method and a well-chosen fixed point method translating the optimality conditions and allowing global perturbations. Finally, in section 5, we illustrate the efficiency of the introduced method using examples of various graphs modeling, in particular, a traffic circle or a main road surrounded by satellite roads.

Our contribution can be summarized in the form of a computational open source code allowing processing complex graphs, but nevertheless of rather small dimensions. This code, written in the Julia language, is available and usable at the following link:

<https://github.com/mickaelbestard/TRoN.jl>

**Notations.** Throughout this article, we will use the following notations:

- For  $x \in \mathbb{R}^n$ ,  $x_+$  will denote the positive part of  $x$ , namely  $x_+ = \max\{x, 0_{\mathbb{R}^n}\}$ , the max being understood component by component;
- $BV(0, T)$  denotes the space of all  $L^1(0, T)$  functions of bounded variation on  $(0, T)$ . For any  $u \in L^1(0, T)$ , the total variation of  $u$  is defined by

$$\text{TV}(f) = \sup \left\{ \int_0^T u(t)\phi'(t) dt, \phi \in C_c^1([0, T]), \|\phi\|_{L^\infty(0, T)} \leq 1 \right\};$$

- $W^{1, \infty}(0, T)$  denotes the space of all functions  $f$  in  $L^\infty(0, T)$  whose gradient in the sense of distributions also belongs to  $L^\infty(0, T)$ ;
- $\|\cdot\|_{\mathbb{R}^d}$  (resp.  $\langle \cdot, \cdot \rangle_{\mathbb{R}^d}$ ) stands for the standard Euclidean norm (resp. inner product) in  $\mathbb{R}^d$ ;
- Let  $F : \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}^m$  and  $(x_0, y_0) \in \mathbb{R}^p \times \mathbb{R}^q$ . We will denote by  $\partial_x F(x_0, y_0)$ , the Jacobian matrix of the mapping  $x \mapsto F(x, y_0)$ . When no ambiguity is possible, we will simply denote it by  $\partial_x F$ ;
- $N_r$ : number of roads in the model;
- $N_c$ : number of mesh cells per road for the first-order Finite Volume scheme;
- $N_T$ : number of time discretization steps in the numerical schemes;

## 2 A controlled model of traffic flow

In this section, we first present the traffic flow model on a road network and its semi-discretization. Then we introduce the precise control model: the control at junctions will translate regulation action using traffic signs or traffic lights.

### 2.1 Traffic dynamics on network without control

The road network is a directed graph of  $N_r$  roads, with  $N_r \in \mathbb{N}^*$ , where each edge corresponds to a road and each vertex to a junction. The roads will all be denoted as real intervals  $[a_i, b_i]$  for some  $i \in \llbracket 1, N_r \rrbracket$ . Of course, this writing does not determine the topology of the graph. It is necessary to define, for each junction, the indices of the incoming and outgoing roads in order to characterize the directed graph completely.

We are interested in the time evolution of the densities on each road. On the  $i$ -th road, the evolution of the density  $\rho_i(t, x)$  is provided by the so-called LWR model

$$\begin{cases} \partial_t \rho_i(t, x) + \partial_x f_i(\rho_i(t, x)) = 0, & (t, x) \in (0, T) \times [a_i, b_i], \\ \rho_i(0, x) = \rho_i^0(x), & x \in [a_i, b_i], \\ f_i(\rho_i(t, a_i)) = \gamma_i^L(t), & t \in (0, T), \\ f_i(\rho_i(t, b_i)) = \gamma_i^R(t), & t \in (0, T), \end{cases} \quad (\text{LWR})$$

where the flux  $f_i(\rho)$  is given by

$$f_i(\rho) = \rho v_i^{\max} \left( 1 - \frac{\rho}{\rho_i^{\max}} \right), \quad (1)$$

where  $v_i^{\max}$  and  $\rho_i^{\max}$  denotes respectively the maximal velocity and density allowed on the  $i$ -th road. The initial density is denoted  $\rho_i^0(x)$  and  $\gamma_i^L$  and  $\gamma_i^R$  are functions allowing prescribing the flux at the left and right boundaries of the interval.

The fluxes  $\gamma$  at junctions are determined by using the model introduced in [13]. Considering a junction  $J$ , the sets of indices corresponding to the in-going and outgoing roads are denoted  $J_{\text{in}}$  and  $J_{\text{out}}$ . We assume that there is a statistical traffic distribution matrix  $A_J$  defined by

$$A_J := (\alpha_{ji})_{(j,i) \in J_{\text{out}} \times J_{\text{in}}}, \quad 0 < \alpha_{ji} < 1, \quad \sum_{j \in J_{\text{out}}} \alpha_{ji} = 1,$$

with  $\alpha_{ji}$  the proportion of vehicles going to the  $j$ -th outgoing road among those coming from the  $i$ -th incoming road. Outgoing fluxes have to satisfy the relation

$$(\gamma_j^L)_{j \in J_{\text{out}}} = A_J (\gamma_i^R)_{i \in J_{\text{in}}}, \quad (2)$$

and we have the balance between in-going and outgoing fluxes thanks to the row-stochasticity of  $A_J$ :

$$\sum_{i \in J_{\text{in}}} \gamma_i^R = \sum_{j \in J_{\text{out}}} \gamma_j^L.$$

Let us note that we have natural constraints on the fluxes. Indeed the LWR flux function  $f_i$  reaches its maximum at  $\rho_i^{\max}/2$ . Therefore, incoming fluxes  $\gamma_i^R$  have to be smaller than the following upper bound:

$$\gamma_i^{R,\max} = f_i(\min\{\rho_i(t, b_i^+), \rho_i^{\max}/2\}),$$

which takes into account reduced demand when the density is lower than  $\rho_i^{\max}/2$  in the incoming road. Similarly outgoing fluxes  $\gamma_j^L$  have to be smaller than the upper-bound:

$$\gamma_j^{L,\max} = f_j(\max\{\rho_j(t, a_j^-), \rho_j^{\max}/2\}),$$

which takes into account reduced capacities when the density is larger than  $\rho_j^{\max}/2$  in the outgoing road. We refer to [13] for details. It is notable that, according to J.P. Lebacque,  $\gamma_i^{R,\max}$  and  $\gamma_j^{L,\max}$  can respectively be seen as the *upstream demand* and the *downstream supply* functions [29]. Consequently, the fluxes belong to the following set:

$$\Omega_J = \left\{ ((\gamma_i^R)_{i \in J_{\text{in}}}, (\gamma_j^L)_{j \in J_{\text{out}}}) \in \prod_{i \in J_{\text{in}}} [0, \gamma_i^{R,\max}] \prod_{j \in J_{\text{out}}} [0, \gamma_j^{L,\max}] \quad \text{with} \quad (\gamma_j^L)_{j \in J_{\text{out}}} = A_J (\gamma_i^R)_{i \in J_{\text{in}}} \right\}$$

Then we assume that drivers succeed in maximizing the total flow. Consequently the fluxes are solution to the following Linear Programming (LP) problem:

$$\max_{((\gamma_i^R)_{i \in J_{\text{in}}}, (\gamma_j^L)_{j \in J_{\text{out}}}) \in \Omega_J} \sum_{i \in J_{\text{in}}} \gamma_i^R, \quad (3)$$

This definition is valid at any junction of the network. Note that this problem does not have necessarily a unique solution in the case where there are more incoming roads than outgoing roads. In that case, a priority modeling assumption is added to select one solution. We refer to [8] for more details. In the following, we will only consider networks with at most two incoming roads and two outgoing roads, i.e. either  $1 \times 1$ ,  $1 \times 2$ ,  $2 \times 1$  or  $2 \times 2$  junctions. This gives existence and stability of the LWR Cauchy

problem [13], and has the advantage that the linear programming problem has a unique solution and can be explicitly provided [49], while it already enables to model a huge variety of road networks. We formally write:

$$\boldsymbol{\gamma}(t) = \phi^{LP}(\boldsymbol{\rho}(t)),$$

for the simultaneous resolution of the linear programming problems at all the junctions of the network. Note that the dependency on  $\boldsymbol{\rho}(t)$  results from the definition of the upper bounds involved in the definition of the sets  $\Omega_J$ . The explicit expression of  $\phi^{LP}$  for the 4 types of junctions is detailed in Appendix A.

## 2.2 Control at junctions

We introduce the vector of controls  $\mathbf{u} = (u_i(\cdot))_{1 \leq i \leq N_r} \in L^\infty([0, T], \mathbb{R}^{N_r})$  at every road entrance. Note that we implicitly assume that every road junction is controlled, but our model allows without any difficulty to neutralize some controls in order to model the fact that only some junctions are controlled.

We interpret each control  $u_j$  as a rate, assuming that at each time  $t \in (0, T)$ , the maximum flow out of a junction and coming into road  $j$  is weighted by a coefficient  $u_j(t) \in [0, 1]$ . At junction  $J$ , we thus define the following polytope of constraints

$$\Omega_J(\mathbf{u}) = \left\{ ((\gamma_i^R)_{i \in J_{\text{in}}}, (\gamma_j^L)_{j \in J_{\text{out}}}) \in \prod_{i \in J_{\text{in}}} [0, \gamma_i^{R, \text{max}}] \prod_{j \in J_{\text{out}}} [0, (1 - u_j) \gamma_j^{L, \text{max}}] \right. \\ \left. \text{with } (\gamma_j^L)_{j \in J_{\text{out}}} = A_J (\gamma_i^R)_{i \in J_{\text{in}}} \right\} \quad (4)$$

Thus the fluxes at junctions are determined by solving the LP problem (3) with this new set of constraints.

In the case where the junction has more than one outgoing road, this initial model is not completely satisfactory since full control on one outgoing road would result in zero incoming fluxes for all the outgoing roads: indeed relation (2) implies that outgoing fluxes are linear combination of the incoming fluxes with positive weights. Consequently, as soon as a control is fully activated, all the incoming fluxes vanish and then all the outgoing fluxes vanish, too. This is definitely not the desired behavior as we would expect that the traffic flow would be deviated to the uncontrolled outgoing roads. To solve this problem, we choose to make the traffic distribution matrix  $A_J$  also dependent on the control  $\mathbf{u}$ . More precisely, we ask that, as soon as a road entry is fully controlled and the other are not controlled, the proportion of vehicles entering the road is set to 0. To illustrate this point, let us describe the control of  $1 \times 2$  junction, i.e. with one incoming and 2 outgoing roads. Then the traffic distribution matrix  $A$  writes:

$$A(\mathbf{u}) = \begin{pmatrix} \alpha(\mathbf{u}) \\ 1 - \alpha(\mathbf{u}) \end{pmatrix},$$

where  $\alpha(\mathbf{u})$  denotes the proportion of vehicles going in the first outgoing road. Then, denoting  $\mathbf{u} = (u_1, u_2)$  the controls of the two outgoing roads, the function  $\alpha(u_1, u_2)$  has to be chosen such that:

- (a) if the first outgoing road is fully controlled and the second is not controlled, then  $\alpha(\mathbf{u})$  has to vanish, resulting in the condition:  $\alpha(1, 0) = 0$ , and then the entire vehicle flow goes in the second outgoing road.
- (b) inversely, if the second outgoing road is fully controlled and the first one is not controlled, then  $1 - \alpha(\mathbf{u})$  has to vanish, resulting in the condition:  $1 - \alpha(0, 1) = 0$ , and then the entire vehicle flow goes in the first outgoing road.

We also ask for the following additional property:

(c) the parameters are not modified if the outgoing roads are equally controlled, which results in the condition:

$$\forall u \in [0, 1], \quad \alpha(u, u) = \bar{\alpha},$$

where  $\bar{\alpha}$  is a given distribution coefficient.

In appendix A, we provide an explicit construction of such a function. Note that, in practice, to avoid the degeneracy of the junction problems, the distribution coefficient is set to a very small value  $\varepsilon^2$  (resp.  $1 - \varepsilon^2$ ) in situations (a) (resp. (b)). We also treat the case of  $2 \times 2$  junctions. Note that junctions with only one outgoing roads ( $1 \times 1$  and  $2 \times 1$ ) do not require such modification.

Like in the previous section, the simultaneous resolution of the Linear Programming problems at each junctions junction is now denoted:

$$\gamma = \phi^{LP}(\boldsymbol{\rho}, \mathbf{u})$$

The explicit expressions of  $\phi^{LP}$  in the set of cases we deal with are provided in appendix A. Beyond the explicit expression of  $\phi^{LP}(\boldsymbol{\rho}, \mathbf{u})$ , what matters is that  $\phi^{LP}$  is a Lipschitz function with respect to  $(\boldsymbol{\rho}, \mathbf{u})$ . In the following, in order to use tools of differentiable optimization, we will consider a  $C^1$  approximation of this function. This issue is commented at the end of appendix A. According to these comments, we will assume from now on:

$$\text{The function } \phi^{LP} \text{ is Lipschitz, and } C^1 \text{ with respect to its second variable } \mathbf{u}. \quad (\text{H}_{\phi^{LP}})$$

In this paper, we do not investigate the well-posedness of this continuous controlled dynamics but rather focus on a semi-discretization version.

### 2.3 Semi-discretized model

For algorithmic efficiency reasons, we prefer to be able to define the sensitivity of the different data of the problem with respect to the control. Since the model (LWR) is known to generate possible irregularities in the form of shocks, we have decided to introduce regularity through a semi-discretization of the model in space. The relevance of this choice will be discussed in the concluding section.

Hence, we discretize the model (LWR) with a first-order Finite Volume (FV) scheme. To simplify the notations, we consider the same number  $N_c \in \mathbb{N}^*$  of mesh cells per road. The discretization points on the  $i$ -th roads are denoted  $a_i = x_{i,1/2} < x_{i,3/2} < \dots < x_{i,N_c-1/2} = b_i$  and the space steps  $\Delta x_{i,j} = x_{i,j+1/2} - x_{i,j-1/2}$ . Then the discrete densities and the boundary fluxes are denoted:

$$\begin{aligned} \boldsymbol{\rho}(t) &= (\rho_{i,j}(t))_{\substack{1 \leq i \leq N_r \\ 1 \leq j \leq N_c}} = \left( \frac{1}{\Delta x_{i,j}} \int_{x_{i,j-\frac{1}{2}}}^{x_{i,j+\frac{1}{2}}} \rho(t, x) dx \right)_{\substack{1 \leq i \leq N_r \\ 1 \leq j \leq N_c}}, \\ \boldsymbol{\gamma}(t) &= (\gamma_1^L(t), \gamma_1^R(t), \dots, \gamma_{N_r}^L(t), \gamma_{N_r}^R(t)). \end{aligned}$$

With a slight abuse of notation, we have written the discrete variable similarly as the continuous one, since we will essentially work on the semi-discretized model. Then the semi-discretized dynamics are given by the differential equation

$$\begin{cases} \frac{d\boldsymbol{\rho}}{dt}(t) = f^{FV}(\boldsymbol{\rho}(t), \boldsymbol{\gamma}(t)), & t \in (0, T) \\ \boldsymbol{\gamma}(t) = \phi^{LP}(\boldsymbol{\rho}(t), \mathbf{u}(t)), & t \in (0, T) \\ \boldsymbol{\rho}(0) = \boldsymbol{\rho}_0, \end{cases} \quad (\text{LWR-sd})$$



where the finite volume flow at the  $j$ -th mesh cell of the  $i$ -th road reads

$$f^{FV}(\boldsymbol{\rho}, \boldsymbol{\gamma})_{i,j} = \begin{cases} \frac{-1}{\Delta x_{i,j}} (\mathcal{F}_i(\rho_{i,j}, \rho_{i,j+1}) - \gamma_i^L), & \text{if } j = 1, \\ \frac{-1}{\Delta x_{i,j}} (\mathcal{F}_i(\rho_{i,j}, \rho_{i,j+1}) - \mathcal{F}_i(\rho_{i,j-1}, \rho_{i,j})), & \text{if } 2 \leq j \leq N_c - 1, \\ \frac{-1}{\Delta x_{i,j}} (\gamma_i^R - \mathcal{F}_i(\rho_{i,j-1}, \rho_{i,j})), & \text{if } j = N_c, \end{cases} \quad (5)$$

with  $\mathcal{F}_i(u, v)$  the so-called local Lax-Friedrichs numerical flux given by

$$\mathcal{F}_i(u, v) = \frac{f_i(u) + f_i(v)}{2} - \max\{|f'_i(u)|, |f'_i(v)|\} \frac{(v - u)}{2}.$$

We refer to [32] for an introduction to finite volume approximations.

## 2.4 Conclusion: an optimal control problem

We are interested in an approximate controllability problem which consists in emptying a given route as much as possible for a fixed final time  $T > 0$ . Let us introduce  $\chi_{\text{path}} \subset \llbracket 1, N_r \rrbracket$ , a set of indices corresponding to the route we wish to empty in a time  $T$ .

We would like to minimize a functional with respect to the control  $\mathbf{u}$ , representing the sum of all the densities on this path, in other words

$$C_T(\mathbf{u}) = \sum_{i \in \chi_{\text{path}}} \sum_{j=1}^{N_c} \rho_{i,j}(T; \mathbf{u}) = \mathbf{c} \cdot \boldsymbol{\rho}(T; \mathbf{u}),$$

where  $\boldsymbol{\rho}(t; \mathbf{u})$  denotes the solution to the controlled system (LWR-sd) and  $\mathbf{c} = (c_i)_{1 \leq i \leq N_r}$  the vector defined by

$$c_i = 1 \text{ if } i \in \chi_{\text{path}} \quad \text{and} \quad c_i = 0 \text{ else.}$$

Of course, it is necessary to introduce a certain number of constraints on the sought controls, in accordance with the model under consideration, and to model that the obtained control is feasible in practice. We are thus driven to consider the following constraints:

- (i)  $0 \leq u_i(\cdot) \leq 1$ , meaning that at each time, the control is a vehicle acceptance rate on a road;
- (ii)  $\sum_{i=1}^{N_r} u_i(\cdot) \leq N_{\max}$ : this imposes a constraint on the total amount of control in order to take into account the staff required for roadblocking. In particular, there cannot be more than  $N_{\max}$  fully active controls ( $u_i = 1$ ) at the same time.
- (iii) Regularity: we will assume that  $\mathbf{u}$  is of bounded variation, and write  $\mathbf{u} \in \text{BV}([0, T]; \mathbb{R}^{N_r})$ . This constraint involves the *total variation* of the control and models that the roadblock is supposed not to "blink" over time. From a control point of view, we aim at avoiding the so-called *chattering phenomenon*.

The first constraint above will be included in the set of admissible controls. Concerning the other two constraints, we have chosen to include them as penalty/regularization terms in the functional. Of course, other choices would be quite possible and relevant.

Let  $N_{\max} \in \mathbb{N}^*$  be an integer standing for the maximal number of active controls for this problem and  $\theta = (\theta_S, \theta_B) \in \mathbb{R}_+^2$  be two non-negative parameters. According to all the considerations above, the optimal control problem we will investigate reads:

$$\boxed{\inf_{\mathbf{u} \in \mathcal{U}_{ad}} \mathcal{J}_\theta(\mathbf{u})}, \quad (\mathcal{P}_\theta)$$

where the admissible set of controls is defined by

$$\mathcal{U}_{ad} = L^\infty([0, T], [0, 1]^{N_r}), \quad (6)$$

and the regularized cost functional  $\mathcal{J}_\theta$  writes

$$\mathcal{J}_\theta(\mathbf{u}) = C_T(\mathbf{u}) + \frac{\theta_S}{2} S(\mathbf{u}) + \theta_B B(\mathbf{u}), \quad (7)$$

where  $S(\mathbf{u})$  denotes the regularizing term modeling the limitation on the number of active controls and  $B(\mathbf{u})$  is the total variation of  $\mathbf{u}$  in time, namely

$$S(\mathbf{u}) = \int_0^T \left( \sum_{i=1}^{N_r} u_i(t) - N_{\max} \right)_+^2 dt \quad \text{and} \quad B(\mathbf{u}) = \sum_{i=1}^{N_r} \text{TV}(u_i). \quad (8)$$

### 3 Analysis of the optimal control problem ( $\mathcal{P}_\theta$ )

In this section, we will investigate the well-posedness and derive the optimality conditions of problem ( $\mathcal{P}_\theta$ ). These conditions will form the basis of the algorithms used in the rest of this study. Without loss of generality, we will suppose that all roads have the same traffic parameters,  $v_i^{\max} = 1$ ,  $\rho_i^{\max} = 1$  and have same length and same number of mesh cells. In particular, the space steps  $\Delta x_i = \Delta x$  is the same for all roads.

#### 3.1 Well-posedness of problem ( $\mathcal{P}_\theta$ )

We follow the so-called direct method of calculus of variations. The key point of the following result is the establishment of uniform estimates of  $\boldsymbol{\rho}$  with respect to the control variable  $\mathbf{u}$ , in the  $W^{1,\infty}(0, T)$  norm.

**Theorem 1.** *Let us assume that  $\theta_B$  is positive. Then, problem ( $\mathcal{P}_\theta$ ) has a solution.*

*Proof.* Let  $(\mathbf{u}_n)_{n \in \mathbb{N}}$  be a minimizing sequence for problem ( $\mathcal{P}_\theta$ ). Observe first that, by minimality, the sequence  $(\mathcal{J}_\theta(\mathbf{u}_n))_{n \in \mathbb{N}}$  is bounded, and therefore, so is the total variation  $(B(\mathbf{u}_n))_{n \in \mathbb{N}}$ , as  $\theta_B$  is assumed to be positive. We infer that, up to a subsequence,  $(\mathbf{u}_n)_{n \in \mathbb{N}}$  converges in  $L^1(0, T)$  and almost everywhere toward an element  $\mathbf{u}^* \in \text{BV}(0, T)$ . Since the class  $\mathcal{U}_{ad}$  is closed for this convergence, we moreover get that  $\mathbf{u}^* \in L^\infty([0, T], [0, 1]^{N_r})$ .

For the sake of readability, we will denote similarly a sequence and any subsequence with a slight abuse of notation.

In what follows, it is convenient to introduce the function  $g : \mathbb{R}^{N_c \times N_r} \times [0, 1]^{N_r}$  defined by  $g(\boldsymbol{\rho}, \mathbf{u}) = f^{FV}(\boldsymbol{\rho}, \phi^{LP}(\boldsymbol{\rho}, \mathbf{u}))$ , so that  $\boldsymbol{\rho}$  solves the ODE system  $\boldsymbol{\rho}' = g(\boldsymbol{\rho}, \mathbf{u})$ . Let us set  $\boldsymbol{\rho}_n := \boldsymbol{\rho}(\cdot; \mathbf{u}_n)$ .

**Step 1: the function  $g$  is continuous.** Indeed, note that the function  $\phi^{LP}$  is continuous (and even Lipschitz), according to ( $\mathbf{H}_{\phi^{LP}}$ ). Explicit expressions of  $\phi^{LP}$  for the four considered type of junctions are detailed in Appendix A. According to (5),  $f^{FV}$  has therefore the same regularity as the numerical flow  $\boldsymbol{\rho} \mapsto \mathcal{F}(\boldsymbol{\rho})$ , defined by

$$\mathcal{F}_i^{j,j+1}(\boldsymbol{\rho}) = \frac{f_i^j + f_i^{j+1}}{2} - \frac{\max\{|df_i^j|, |df_i^{j+1}|\}}{2} (\rho_i^{j+1} - \rho_i^j),$$

with  $f_i^j = \rho_i^j(1 - \rho_i^j)$  and  $df_i^j = 1 - 2\rho_i^j$ , which is obviously continuous.

**Step 2: the sequence  $(\rho_n)_{n \in \mathbb{N}}$  is uniformly bounded in  $W^{1,\infty}(0, T)$ .** For  $\mathbf{u} \in \mathcal{U}_{ad}$  given, let us introduce the function  $\mathcal{L} : t \mapsto \frac{1}{2} \|\rho(t)\|_{\mathbb{R}^{N_c \times N_r}}^2$ . According to the chain rule,  $\mathcal{L}$  is differentiable, and its derivative reads<sup>1</sup>

$$\begin{aligned}
\mathcal{L}'(t) &= \langle \rho'(t), \rho(t) \rangle_{\mathbb{R}^{N_c \times N_r}} = \langle g(\rho(t), \mathbf{u}(t)), \rho(t) \rangle_{\mathbb{R}^{N_c \times N_r}} \\
&= -\frac{1}{\Delta x} \sum_{i=1}^{N_r} \sum_{j=1}^{N_c} \left( \mathcal{F}_i^{j,j+1}(\rho) - \mathcal{F}_i^{j-1,j}(\rho) \right) \rho_i^j \\
&= -\frac{1}{\Delta x} \sum_{i=1}^{N_r} \left( \gamma_i^R \rho_i^{N_c} + \sum_{j=1}^{N_c-1} \mathcal{F}_i^{j,j+1}(\rho) \rho_i^j - \gamma_i^L \rho_i^1 - \sum_{j=2}^{N_c} \mathcal{F}_i^{j-1,j}(\rho) \rho_i^j \right) \\
&= \frac{1}{\Delta x} \sum_{i=1}^{N_r} \left( \gamma_i^L \rho_i^1 - \gamma_i^R \rho_i^{N_c} + \sum_{j=1}^{N_c-1} \mathcal{F}_i^{j,j+1}(\rho) (\rho_i^{j+1} - \rho_i^j) \right).
\end{aligned}$$

For a given  $i \in \llbracket 1, N_r \rrbracket$ , and  $j \in \llbracket 1, N_c - 1 \rrbracket$ , we have

$$\begin{aligned}
\mathcal{F}_i^{j,j+1}(\rho) (\rho_i^{j+1} - \rho_i^j) &= \left( \frac{f_i^j + f_i^{j+1}}{2} - \frac{\max\{|df_i^j|, |df_i^{j+1}|\}}{2} (\rho_i^{j+1} - \rho_i^j) \right) (\rho_i^{j+1} - \rho_i^j) \\
&\leq (f_i(\sigma_i) + (\rho_i^{j+1} - \rho_i^j)) (\rho_i^{j+1} - \rho_i^j),
\end{aligned}$$

since  $f_i^j \leq f_i(\sigma_i) = 1/4$  according to the explicit expression (1) of  $f_i$ , and  $|df_i^j| \leq v_i^{\max} = 1$ . Since the local Lax-Friedrichs scheme is monotone, we infer that  $\rho_i^j$  is lower than  $\rho_i^{\max} = 1$ . We then have

$$\begin{aligned}
\mathcal{F}_i^{j,j+1}(\rho) (\rho_i^{j+1} - \rho_i^j) &\leq 2f_i(\sigma_i) + (\rho_i^{j+1} - \rho_i^j)^2 = \frac{1}{2} + (\rho_i^{j+1} - \rho_i^j)^2 \\
&\leq \frac{1}{2} + (\rho_i^j)^2 + (\rho_i^{j+1})^2 - 2\rho_i^j \rho_i^{j+1} \\
&\leq \frac{1}{2} + (\rho_i^j)^2 + (\rho_i^{j+1})^2,
\end{aligned}$$

by positivity of the  $\rho_i^j$ . Using the majoration  $\gamma_i^L \rho_i^1 - \gamma_i^R \rho_i^{N_c} \leq f_i(\sigma_i) \times 2 = 1/2$ , all the calculations above yield

$$\begin{aligned}
\mathcal{L}'(t) &\leq \frac{1}{\Delta x} \sum_{i=1}^{N_r} \left( \frac{1}{2} + \sum_{j=1}^{N_c-1} \left( \frac{1}{2} + (\rho_i^j)^2 + (\rho_i^{j+1})^2 \right) \right) \\
&\leq \frac{1}{\Delta x} \left( \frac{N_r N_c}{2} + \|\rho(t)\|^2 \right).
\end{aligned}$$

We infer the existence of two positive numbers  $\bar{\alpha}, \bar{\beta}$  that do not depend on  $\mathbf{u}$ , such that  $\mathcal{L}'(t) \leq \bar{\alpha} \mathcal{L}(t) + \bar{\beta}$  for a.e.  $t \in (0, T)$ . By using a Grönwall-type inequality, we get

$$\mathcal{L}(t) \leq \mathcal{L}(0) e^{\bar{\alpha} t} + \int_0^t \beta e^{\bar{\alpha}(t-s)} ds = \left( \mathcal{L}(0) + \frac{\bar{\beta}}{\bar{\alpha}} (1 - e^{-\bar{\alpha} t}) \right) e^{\bar{\alpha} t}, \quad (9)$$

**Step 3: conclusion.** According to (9), the sequence  $(\rho_n)_{n \in \mathbb{N}}$  is bounded in  $L^\infty(0, T)$ . Since  $\rho'_n = g(\rho_n, \mathbf{u}_n)$  a.e. in  $(0, T)$  and  $g$  is continuous, it follows that  $(\rho_n)_{n \in \mathbb{N}}$  is uniformly bounded in  $W^{1,\infty}(0, T)$ . Therefore, by using the Arzela-Ascoli theorem, this sequence converges up to a subsequence in  $C^0([0, T])$  toward an element  $\rho^* \in W^{1,\infty}(0, T)$ .

<sup>1</sup>we drop the time dependency notation for readability.

Let us recast System (LWR-sd) as a fixed point equation, as

$$\text{for every } t \in (0, T), \quad \boldsymbol{\rho}_n(t) - \boldsymbol{\rho}_0 = \int_0^t g(\boldsymbol{\rho}_n(s), \mathbf{u}_n(s)) ds,$$

we obtain by letting  $n$  go to  $+\infty$ ,

$$\boldsymbol{\rho}^*(t) - \boldsymbol{\rho}_0 = \int_0^t g(\boldsymbol{\rho}^*(s), \mathbf{u}^*(s)) ds,$$

meaning that  $\boldsymbol{\rho}^*$  solves System (LWR-sd) with  $\mathbf{u}^*$  as control. Finally, according to the aforementioned convergences and since the functionals  $S$  and  $B$  are convex, it is standard that one has

$$\lim_{n \rightarrow +\infty} C_T(\mathbf{u}_n) = C_T(\mathbf{u}^*), \quad \liminf_{n \rightarrow +\infty} S(\mathbf{u}_n) \geq S(\mathbf{u}^*), \quad \liminf_{n \rightarrow +\infty} B(\mathbf{u}_n) \geq B(\mathbf{u}^*).$$

We get that  $\mathcal{J}_\theta(\mathbf{u}^*) \leq \liminf_{n \rightarrow +\infty} \mathcal{J}_\theta(\mathbf{u}_n)$ , and we infer that  $\mathbf{u}^*$  solves problem  $(\mathcal{P}_\theta)$ .  $\square$

### 3.2 Optimality conditions

We have established the existence of an optimal control in Theorem 1. In order to derive a numerical solution algorithm, we will now state the necessary optimality conditions on which the algorithm we will build is based. One of the difficulties is that the functional we use involves non-differentiable quantities. We will therefore first use the notion of subdifferential to write the optimality conditions. In section 4 dedicated to numerical experiments, we will explain how we approximate these quantities.

Let us first compute the differential of the functional  $C_T$ . To this aim, we introduce the tangent cone to the set  $\mathcal{U}_{ad}$ .

**Definition 2.** Let  $\mathbf{u} \in \mathcal{U}_{ad}$ . A function  $\mathbf{h}$  in  $L^\infty(0, T)$  is said to be an ‘‘admissible perturbation’’ of  $\mathbf{u}$  in  $\mathcal{U}_{ad}$  if, for every sequence of positive real numbers  $(\varepsilon_n)_{n \in \mathbb{N}}$  decreasing to 0, there exists a sequence of functions  $\mathbf{h}^n$  converging to  $\mathbf{h}$  for the weak-star topology of  $L^\infty(0, T)$  as  $n \rightarrow +\infty$ , and such that  $\mathbf{u} + \varepsilon_n \mathbf{h}^n \in \mathcal{U}_{ad}$  for every  $n \in \mathbb{N}$ .

**Proposition 3.** Let  $\mathbf{u} \in \mathcal{U}_{ad}$  and  $(\boldsymbol{\rho}, \boldsymbol{\gamma})$  the associated solution to (LWR-sd). We introduce the two matrices  $M$  and  $N$  defined from the Jacobian matrices of  $f^{FV}$  and  $\phi^{LP}$  as

$$M(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u}) = (\partial_{\boldsymbol{\rho}} f^{FV})(\boldsymbol{\rho}, \boldsymbol{\gamma}) + (\partial_{\boldsymbol{\gamma}} f^{FV})(\boldsymbol{\rho}, \boldsymbol{\gamma}) (\partial_{\boldsymbol{\rho}} \phi^{LP})(\boldsymbol{\rho}, \mathbf{u}), \quad (10)$$

$$N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u}) = (\partial_{\boldsymbol{\gamma}} f^{FV})(\boldsymbol{\rho}, \boldsymbol{\gamma}) (\partial_{\mathbf{u}} \phi^{LP})(\boldsymbol{\rho}, \mathbf{u}), \quad (11)$$

where we use the notation conventions introduced in section 1.

The functional  $C_T$  is differentiable in the sense of Gâteaux and its differential reads

$$dC_T(\mathbf{u})\mathbf{h} = \int_0^T (N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})^\top \mathbf{p}) \cdot \mathbf{h} dt, \quad (12)$$

for every admissible perturbation  $\mathbf{h}$ , where  $\mathbf{p}$  is the so-called adjoint state, defined as the unique solution to the Cauchy system

$$\begin{cases} \mathbf{p}' + M(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})^\top \mathbf{p} = 0 & \text{in } (0, T), \\ \mathbf{p}(T) = \mathbf{c}. \end{cases} \quad (13)$$

**Remark 4.** In Proposition 3 above, the matrices  $M$  and  $N$  express respectively the way by which  $\boldsymbol{\rho}$  interacts with  $\boldsymbol{\gamma}$  and  $\boldsymbol{\gamma}$  interacts with  $\mathbf{u}$ .

*Proof of Proposition 3.* Let  $\mathbf{u} \in \mathcal{U}_{ad}$ . The Gâteaux differentiability of  $C_T$ ,  $\mathcal{U}_{ad} \ni \mathbf{u} \mapsto \boldsymbol{\rho}$  and  $\mathcal{U}_{ad} \ni \mathbf{u} \mapsto \boldsymbol{\gamma}$  is standard, and follows for instance directly of the proof of the Pontryagin Maximum Principle (PMP, see e.g. [30]). Although the expression of the differential of  $C_T$  could also be obtained by using the PMP, we provide a short proof hereafter to make this article self-contained.

Let  $\mathbf{h} \in L^\infty(0, T, \mathcal{U}_{ad})$  be an admissible perturbation of  $\mathbf{u}$  in  $\mathcal{U}_{ad}$ . One has

$$dC_T(\mathbf{u}) \mathbf{h} = \mathbf{c} \cdot \dot{\boldsymbol{\rho}}(T), \quad (14)$$

where  $\dot{\boldsymbol{\rho}} = d(\mathbf{u} \mapsto \boldsymbol{\rho})\mathbf{h}$  solves the system

$$\begin{cases} \dot{\boldsymbol{\rho}}' = (\partial_{\boldsymbol{\rho}} f^{FV})(\boldsymbol{\rho}, \boldsymbol{\gamma})\dot{\boldsymbol{\rho}} + (\partial_{\boldsymbol{\gamma}} f^{FV})(\boldsymbol{\rho}, \boldsymbol{\gamma})\dot{\boldsymbol{\gamma}}, \\ \dot{\boldsymbol{\gamma}} = (\partial_{\boldsymbol{\rho}} \phi^{LP})(\boldsymbol{\rho}, \mathbf{u})\dot{\boldsymbol{\rho}} + (\partial_{\mathbf{u}} \phi^{LP})(\boldsymbol{\rho}, \mathbf{u})\mathbf{h}, \\ \dot{\boldsymbol{\rho}}(0) = 0. \end{cases}$$

We infer that  $\dot{\boldsymbol{\rho}}$  solves

$$\begin{cases} \dot{\boldsymbol{\rho}}' = M(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})\dot{\boldsymbol{\rho}} + N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})\mathbf{h}, \\ \dot{\boldsymbol{\rho}}(0) = 0, \end{cases} \quad (15)$$

with  $M$  and  $N$  as defined in (10). Let us multiply the main equation of (13) by  $\dot{\boldsymbol{\rho}}$  in the sense of the inner product, and integrate over  $(0, T)$ . We obtain:

$$\int_0^T \dot{\boldsymbol{\rho}} \cdot \frac{d\mathbf{p}}{dt} dt + \int_0^T \dot{\boldsymbol{\rho}} \cdot (M(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})^\top \mathbf{p}) dt = 0.$$

Similarly, let us multiply the main equation of (15) by  $\mathbf{p}$  in the sense of the inner product, and integrate over  $(0, T)$ . We obtain:

$$\int_0^T \mathbf{p} \cdot \frac{d\dot{\boldsymbol{\rho}}}{dt} dt - \int_0^T M(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})\dot{\boldsymbol{\rho}} \cdot \mathbf{p} dt = \int_0^T N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})\mathbf{h} \cdot \mathbf{p} dt.$$

Summing the last two equalities above yields

$$\dot{\boldsymbol{\rho}}(T) \cdot \mathbf{p}(T) - \dot{\boldsymbol{\rho}}(0) \cdot \mathbf{p}(0) = \int_0^T \mathbf{h} \cdot (N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})^\top \mathbf{p}) dt.$$

Using this identity with  $\mathbf{p}(T) = c$  and  $\dot{\boldsymbol{\rho}}(0) = 0$  results in expression (12).  $\square$

From this result, we will now state the optimality conditions for problem  $(\mathcal{P}_\theta)$ . Let us first recall that, according to [17, Proposition I.5.1], the subdifferential of the total variation is well-known, given by

$$\partial \text{TV}(\mathbf{u}^*) = \{\boldsymbol{\eta} \in C^0([0, T]; \mathbb{R}^{N_r}) \mid \|\boldsymbol{\eta}\|_\infty \leq 1 \text{ and } \int \boldsymbol{\eta} d\mathbf{u}^* = \text{TV}(\mathbf{u}^*)\}.$$

Let us denote by  $e_i$  the  $i$ -th vector of the canonical basis of  $\mathbb{R}^{N_r}$ .

**Theorem 5.** *Let  $\mathbf{u}^* = (u_i^*)_{1 \leq i \leq N_r}$ , denote a solution to problem  $(\mathcal{P}_\theta)$ ,  $(\boldsymbol{\rho}, \boldsymbol{\gamma})$  the associated solution to (LWR-sd). and let  $i \in \llbracket 1, N_r \rrbracket$ . There exists  $\boldsymbol{\eta}^* \in \partial \text{TV}(\mathbf{u}^*)$  such that*

- on  $\{u_i^* = 0\}$ , one has  $\Psi \cdot e_i \geq 0$ ,
- on  $\{u_i^* = 1\}$ , one has  $\Psi \cdot e_i \leq 0$ ,
- on  $\{0 < u_i^* < 1\}$ , one has  $\Psi \cdot e_i = 0$ ,

where the function  $\Psi : [0, T] \rightarrow \mathbb{R}^{N_r}$  is given by

$$\Psi(t) = N(\boldsymbol{\rho}(t), \boldsymbol{\gamma}(t), \mathbf{u}^*(t))^\top \mathbf{p}^*(t) + \theta_S \left( \sum_{i=1}^{N_r} u_i^*(t) - N_{\max} \right)_+ + \theta_B \boldsymbol{\eta}^*(t)$$

and where  $\mathbf{p}^*$  denotes the adjoint state introduced in proposition 3, associated with the control choice  $\mathbf{u}^*$ .

**Remark 6.** Written in this way, the first order optimality conditions are difficult to use. In the next section, we will introduce an approximation of the total variation of  $\mathbf{u}^*$  leading to optimality conditions more easily usable within an algorithm.

*Proof of theorem 5.* To derive the first order optimality conditions for this problem, it is convenient to introduce the so-called indicator function  $\iota_{\mathcal{U}_{ad}}$  given by

$$\iota_{\mathcal{U}_{ad}}(\mathbf{u}) = \begin{cases} 0 & \text{if } \mathbf{u} \in \mathcal{U}_{ad} \\ +\infty & \text{else.} \end{cases}$$

Observing that the optimization problem we want to deal with can be recast as

$$\min_{\mathbf{u} \in L^\infty((0, T); \mathbb{R}^{N_r})} \mathcal{J}_\theta(\mathbf{u}) + \iota_{\mathcal{U}_{ad}}(\mathbf{u}),$$

it is standard in nonsmooth analysis to write the first order optimality conditions as:

$$0 \in \partial(\mathcal{J}_\theta(\mathbf{u}^*) + \iota_{\mathcal{U}_{ad}}(\mathbf{u}^*)),$$

or similarly, by using standard computational rules [17],

$$-\partial C_T(\mathbf{u}^*) - \frac{\theta_S}{2} \partial S(\mathbf{u}^*) \in \theta_B \partial B(\mathbf{u}^*) + \partial \iota_{\mathcal{U}_{ad}}(\mathbf{u}^*),$$

Let  $\mathbf{u} \in \mathcal{U}_{ad}$ . The condition above yields the existence of  $\boldsymbol{\eta}^* = (\eta_i^*)_{1 \leq i \leq N_r} \in \partial \text{TV}(\mathbf{u}^*)$  such that

$$dC_T(\mathbf{u}^*)(\mathbf{u} - \mathbf{u}^*) + \theta_S dS(\mathbf{u}^*)(\mathbf{u} - \mathbf{u}^*) + \theta_B \sum_{i=1}^{N_r} \langle \eta_i^*, u_i - u_i^* \rangle_{L^2(0, T)} \geq 0.$$

Since  $\mathbf{u}$  is arbitrary, we infer that for any admissible perturbation  $\mathbf{h}$  of  $\mathbf{u}^*$  (see definition 2), one has

$$dC_T(\mathbf{u}^*)\mathbf{h} + \theta_S dS(\mathbf{u}^*)\mathbf{h} + \theta_B \sum_{i=1}^{N_r} \langle \eta_i^*, h_i \rangle_{L^2(0, T)} \geq 0,$$

or similarly

$$\int_0^T \mathbf{h} \cdot \left( N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})^\top \mathbf{p}^* + \theta_S \left( \sum_{i=1}^{N_r} u_i^* - N_{\max} \right)_+ + \theta_B \boldsymbol{\eta}^* \right) dt \geq 0. \quad (16)$$

To analyze this optimality condition, let us introduce the function  $\Psi : [0, T] \rightarrow \mathbb{R}^{N_r}$  defined by

$$\Psi(t) = N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{u})^\top \mathbf{p}^*(t) + \theta_S \left( \sum_{i=1}^{N_r} u_i^* - N_{\max} \right)_+ + \theta_B \boldsymbol{\eta}^*(t). \quad (17)$$

In what follows, we will write the optimality conditions holding for the  $i$ -th component of  $\mathbf{u}^*$ , where  $i \in \llbracket 1, N_r \rrbracket$  is given.

Let us assume that the set  $\mathcal{I} = \{0 < u_i^* < 1\}$  has positive Lebesgue measure. Let  $x_0$  be a Lebesgue point of  $u_i^*$  in  $\mathcal{I}$  and let  $(G_n)_{n \in \mathbb{N}}$  be a sequence of measurable subsets with  $G_n$  included in  $\mathcal{I}$  and containing  $x_0$ . Let us consider  $\mathbf{h} = (h_j)_{1 \leq j \leq N_r}$  such that  $h_j = 0$  for all  $j \in \llbracket 1, N_r \rrbracket \setminus \{i\}$  and  $h_i = \mathbf{1}_{G_n}$ . Notice that  $\mathbf{u}^* \pm \eta \mathbf{h}$  belongs to  $\mathcal{U}_{ad}$  whenever  $\eta$  is small enough. According to (16), one has

$$\pm \int_{G_n} \Psi(t) \cdot e_i dt \geq 0.$$

Dividing this inequality by  $|G_n|$  and letting  $G_n$  shrink to  $\{x_0\}$  as  $n \rightarrow +\infty$  shows that one has

$$\Psi(t) \cdot e_i = 0, \quad \text{a.e. in } \mathcal{I}.$$

Let us now assume that the set  $\mathcal{I}_1 = \{u_i^* = 1\}$  has positive Lebesgue measure. Then, by mimicking the reasoning above, we consider  $x_1$ , a Lebesgue point of  $u_i^*$  in  $\mathcal{I}_1$ , and  $\mathbf{h} = (h_j)_{1 \leq j \leq N_r}$  such that  $h_j = 0$  for all  $j \in \llbracket 1, N_r \rrbracket \setminus \{i\}$  and  $h_i = -\mathbf{1}_{G_n}$ , where  $(G_n)_{n \in \mathbb{N}}$  is a sequence of measurable subsets with  $G_n$  included in  $\mathcal{I}_1$  and containing  $x_1$ . According to (16), one has

$$- \int_{G_n} \Psi(t) \cdot e_i dt \geq 0.$$

As above, we divide this inequality by  $|G_n|$  and let  $G_n$  shrink to  $\{x_1\}$  as  $n \rightarrow +\infty$ . We recover that  $\Psi(t) \cdot e_i \leq 0$ .

Regarding now the set  $\mathcal{I}_0 = \{u_i^* = 0\}$ , the reasoning is a direct adaptation of the case above, which concludes the proof.  $\square$

## 4 Towards a numerical algorithm

In this section, we introduce an exploitable approximation of the problem we are dealing with and describe the algorithm implemented in the numerical part.

### 4.1 An approximate version of problem $(\mathcal{P}_\theta)$

The fact that problem  $(\mathcal{P}_\theta)$  involves the total variation of control makes the problem non-differentiable. Of course, dedicated algorithms exist to take into account such a term in the solution, for instance Chambolle's projection algorithm [10]. Nevertheless, in order to avoid too costly numerical approaches, we have chosen to consider a simple differentiable approximation of the term  $B(\mathbf{u})$ , namely

$$B_\nu(\mathbf{u}) = \sum_{i=1}^{N_r} \text{TV}_\nu(u_i), \quad (18)$$

where  $\nu > 0$  is a small parameter and the total variation  $\text{TV}(u_i)$  is approximated by a differentiable functional in  $L^2(0, T)$ , denoted  $\text{TV}_\nu(u_i)$ , where  $\nu > 0$  stands for a regularization parameter. The concrete choice of the differentiable approximation of the TV standard will be discussed in the rest of this section. We will also give some elements on its practical implementation.

We are thus led to consider the following approximate version of problem  $(\mathcal{P}_\theta)$ :

$$\boxed{\inf_{\mathbf{u} \in \mathcal{U}_{ad}} \mathcal{J}_{\theta, \nu}(\mathbf{u})}. \quad (\mathcal{P}_{\theta, \nu})$$

where  $\mathcal{U}_{ad}$  is given by (6), and  $\mathcal{J}_{\theta, \nu}$  is given by

$$\mathcal{J}_{\theta, \nu}(\mathbf{u}) = C_T(\mathbf{u}) + \frac{\theta_S}{2} S(\mathbf{u}) + \theta_B B_\nu(\mathbf{u}), \quad (19)$$

We will see that this approximation is in fact well adapted to practical use. Indeed, the first order optimality conditions for this approximated problem can be rewritten in a very concise and workable way, unlike the result stated in theorem 5. This is the purpose of the following result.

**Theorem 7.** Let  $\mathbf{u}^* \in \mathcal{U}_{\text{ad}}$  denote a local minimizer for problem  $(\mathcal{P}_{\theta,\nu})$  and  $(\boldsymbol{\rho}, \boldsymbol{\gamma})$  the associated solution to (LWR-sd). Then,  $\mathbf{u}^*$  satisfies the first order necessary condition

$$\Lambda(\cdot) = 0 \quad \text{a.e. on } [0, T],$$

where  $\Lambda : [0, T] \rightarrow \mathbb{R}^{N_r}$  is defined by

$$\Lambda(t) = \min \{ \mathbf{u}^*(t), \max \{ \mathbf{u}^*(t) - 1, \nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*)(t) \} \},$$

and

$$\nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*) : [0, T] \ni t \mapsto N(\boldsymbol{\rho}(t), \boldsymbol{\gamma}(t), \mathbf{u}(t))^\top \mathbf{p}^*(t) + \theta_S \left( \sum_{i=1}^{N_r} u_i^*(t) - N_{\max} \right)_+ + \theta_B \sum_{i=1}^{N_r} \nabla_{\mathbf{u}} \text{TV}_\nu(u_i)(t),$$

where  $\mathbf{p}^*$  has been introduced in theorem 5, the min, max operations being understood componentwise, and the term  $\nabla_{\mathbf{u}}$  denoting the gradient with respect to  $\mathbf{u}$  in  $L^2(0, T)$ .

*Proof.* The proof is similar to the proof of theorem 5. Indeed, let  $\mathbf{u}^*$  be a local minimizer for problem  $(\mathcal{P}_{\theta,\nu})$ . The first order optimality conditions are given by the so-called Euler inequation and read  $d\mathcal{J}_{\theta,\nu}(\mathbf{u}^*)\mathbf{h} \geq 0$ , or similarly

$$\int_0^T \nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*) \cdot \mathbf{h} dt \geq 0.$$

for every admissible perturbation  $\mathbf{h}$ , as defined in definition 2. Let us fix  $i \in \llbracket 1, N_r \rrbracket$ . By mimicking the reasoning involving Lebesgue points in the proof of theorem 5, we get

- on  $\{u_i^* = 0\}$ , one has  $\nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*) \cdot e_i \geq 0$ ,
- on  $\{u_i^* = 1\}$ , one has  $\nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*) \cdot e_i \leq 0$ ,
- on  $\{0 < u_i^* < 1\}$ , one has  $\nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*) \cdot e_i = 0$ .

Note that, on  $\{u_i^* = 0\}$ , one has  $\nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*) \cdot e_i \geq 0$  and then  $\Lambda(t) \cdot e_i = \min\{0, \max\{-1, \nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*)(t) \cdot e_i\}\} = 0$ . On  $\{u_i^* = 1\}$ , one has  $\nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*) \cdot e_i \leq 0$  and therefore  $\Lambda(t) \cdot e_i = \min\{1, \max\{0, \nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*)(t) \cdot e_i\}\} = 0$ . On  $\{0 < u_i^* < 1\}$ , one has  $\nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*) \cdot e_i = 0$  and therefore  $\Lambda(t) \cdot e_i = \min\{u_i^*, \max\{u_i^*(t) - 1, 0\}\} = 0$ .

Conversely, let us assume that  $\Lambda(\cdot) = 0$ . On  $\{u_i^* = 0\}$ , one has

$$0 = \Lambda(t) \cdot e_i = \min\{0, \max\{-1, \nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*)(t) \cdot e_i\}\},$$

so that  $\max\{-1, \nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*)(t) \cdot e_i\} \geq 0$  and finally,  $\nabla_{\mathbf{u}} \mathcal{J}_{\theta,\nu}(\mathbf{u}^*)(t) \cdot e_i \geq 0$ . A similar reasoning yields the optimality conditions on  $\{u_i^* = 1\}$  and  $\{0 < u_i^* < 1\}$ . The conclusion follows.  $\square$

**Practical computation of the TV operator gradient in a discrete framework.** From a practical point of view, we discretize the space of controls, which leads us to consider a time discretization denoted  $(t_n)_{1 \leq n \leq N_T}$  with  $N_T \in \mathbb{N}^*$  fixed, as well as piecewise constant controls in time, denoted  $(u_i^n)_{\substack{1 \leq i \leq N_r \\ 1 \leq n \leq N_T}}$ .

Hence, the term  $u_i^n$  corresponds to the control of road  $i$  and time  $n$ . The simplest discrete version of the TV semi-norm is given by

$$\text{TV}(\mathbf{u}) \simeq \sum_{i=1}^{N_r} \sum_{n=2}^{N_T} |u_i^n - u_i^{n-1}|.$$



In order to manipulate differentiable expressions, we introduce for some  $\nu > 0$ , the smoothed discrete TV operator is given by

$$\text{TV}_\nu : \mathbb{R}^{N_r \times N_T} \ni (u_i^n)_{\substack{1 \leq i \leq N_r \\ 1 \leq n \leq N_T}} \mapsto \sum_{i=1}^{N_r} \sum_{n=2}^{N_T} a_\nu(u_i^n - u_i^{n-1}), \quad (20)$$

where  $a_\nu : \mathbb{R} \ni x \mapsto \sqrt{x^2 + \nu^2}$ .

Let  $(i_0, n_0) \in \llbracket 1, N_r \rrbracket \times \llbracket 1, N_T \rrbracket$ . In what follows, we will use a discrete equivalent of theorem 7, involving the gradient of  $\text{TV}_\nu$ , obtained from the expression

$$\partial_{u_{i_0}^{n_0}} \text{TV}_\nu(\mathbf{u}) = \begin{cases} a'_\nu(u_{i_0}^{n_0} - u_{i_0}^{n_0-1}) - a'_\nu(u_{i_0}^{n_0+1} - u_{i_0}^{n_0}) & \text{if } 2 \leq n_0 \leq N_T - 1 \\ a'_\nu(u_{i_0}^{N_T} - u_{i_0}^{N_T-1}) & \text{if } n_0 = N_T \\ -a'_\nu(u_{i_0}^2 - u_{i_0}^1) & \text{if } n_0 = 1, \end{cases}$$

where  $\mathbf{u} = (u_i^n)_{\substack{1 \leq i \leq N_r \\ 1 \leq n \leq N_T}}$  is given.

## 4.2 Numerical solving of the primal and dual problems

The models we use are already discretized in space. We now explain how we discretize them in time. We recall that, at the end of section 4.1, we have already considered that the controls are assimilated to piecewise constant functions on each cell of the considered mesh and on each time step.

**Finite volumes scheme for the primal problem.** The main transport equation on network is solved by integrating for each road the finite volume flow given by (LWR-sd) with an explicit Euler scheme:

$$\rho_{i,j}^{n+1} = \rho_{i,j}^n - \frac{\Delta t_n}{\Delta x_{i,j}} \left( \mathcal{F}_{i,j+\frac{1}{2}}^n - \mathcal{F}_{i,j-\frac{1}{2}}^n \right), \quad 1 \leq i \leq N_r, \quad 1 \leq j \leq N_c, \quad 1 \leq n \leq N_T - 1,$$

using the local Lax-Friedrichs numerical flux

$$\mathcal{F}_{i,j+\frac{1}{2}}^n = \mathcal{F}(\rho_{i,j}^n, \rho_{i,j+1}^n), \quad 2 \leq j \leq N_c - 1,$$

and the Neumann boundary conditions

$$\mathcal{F}_{i,\frac{1}{2}}^n := \gamma_{i,L}^n, \quad \mathcal{F}_{i,N_c+\frac{1}{2}}^n := \gamma_{i,R}^n,$$

where  $\gamma_{i,L}^n$  and  $\gamma_{i,R}^n$  are obtained as the solution to the linear programming system

$$\gamma^n = \phi^{\text{LP}}(\boldsymbol{\rho}^n, \mathbf{u}^n).$$

**Euler scheme for adjoint problem.** To solve the backward ODE (13), it is convenient to introduce  $Z(t) := (M(T-t))^T$ , and  $\mathbf{q}(t) := \mathbf{p}(T-t)$  such that we are now dealing with the Cauchy system:

$$\begin{cases} \mathbf{q}'(t) = Z(t) \mathbf{q}(t), & t \in (0, T), \\ \mathbf{q}(0) = \mathbf{c}, \end{cases}$$

that we integrate with a classical explicit Euler scheme:

$$\mathbf{q}^1 = \mathbf{c}, \quad \mathbf{q}^{n+1} = (I + \Delta t_n Z^n) \mathbf{q}^n, \quad n = 1, \dots, N_T - 1.$$

The solution is finally recovered using that  $\mathbf{p}^n = \mathbf{q}^{N_T-n+1}$ .

### 4.3 Optimization algorithms

The starting point of the algorithm we implement is based on a standard primal-dual approach, in which the state and the adjoint are computed in order to deduce the gradient of the considered functional. We combine it with a projection method in order to guarantee the respect of the  $L^\infty$  constraints on the control. This method has the advantage of being robust, as it generally allows a significant decrease in the cost functional. On the other hand, it often has the disadvantage of being very local, which results in an important dependence on the initialization. Moreover, one can expect that there are many local minima, since the targeted problem is intrinsically of infinite dimension, which makes the search difficult.

We will propose a modification of this well-known method, using a fixed point method inspired by the optimality condition stated in theorem 7.

#### 4.3.1 Projected gradient descent

A direct approach is to consider the gradient algorithm, in which we deal with the condition  $0 \leq u_i^k \leq 1$  by projection, according to algorithm 1. The specific difficulty of this approach is to find a suitable descent-step  $\delta_k$ , which must be small enough to ensure descent but large enough for the algorithm to converge in a reasonable number of iterations. We hereby combine this algorithm with a scheduler (21) inspired by classical learning rate scheduler in deep-learning [48] to select an acceptable descent step, where  $\delta_0$  and decay are given real numbers.

$$\delta_k := \frac{\delta_0}{1 + \text{decay} \times k}. \quad (21)$$

---

**Algorithm 1** Optimal control by projected gradient descent method (GD)

---

**Require:**  $\rho^0, \mathbf{u}^0, \text{tol} > 0, \text{itermax} > 0$

**Initialization:**  $k = 0$

**while**  $\|\mathbf{u}^{k+1} - \mathbf{u}^k\| > \text{tol}$  **and**  $k \leq \text{itermax}$  **do**

$$(\rho, \gamma) \leftarrow \text{solution to } \begin{cases} \rho' = f^{\text{FV}}(\rho, \gamma) \\ \gamma = \phi^{\text{LP}}(\rho, \mathbf{u}^k) \\ \rho(0) = \rho^0 \end{cases}$$

$$\mathbf{p} \leftarrow \text{solution to } \begin{cases} \mathbf{p}' + M(\rho, \gamma, \mathbf{u}^k)^T \mathbf{p} = \mathbf{0} \\ \mathbf{p}(T) = \mathbf{c} \end{cases}$$

$$\mathbf{u}^{k+1} \leftarrow \text{proj}_{[0,1]}(\mathbf{u}^k - \delta_k \nabla \mathcal{J}_{\theta, \nu}(\mathbf{u}^k)), \text{ with } \delta_k > 0 \text{ such that } \mathcal{J}_{\theta, \nu}(\mathbf{u}^{k+1}) \leq \mathcal{J}_{\theta, \nu}(\mathbf{u}^k)$$

$k \leftarrow k + 1$

**end while**

---

**Fixed point method** A fixed-point (FP) algorithm is derived using the first-order optimality conditions stated in theorems 5 and 7, by rewriting them in a fixed-point formulation. We have seen that they write under the form

$$u_i \in I_\mu \Rightarrow \partial_{u_i} \mathcal{J}_{\theta, \nu}(\mathbf{u}) \in E_\mu,$$

where  $\mu \in \{0, *, 1\}$ ,  $I_0 = \{0\}$ ,  $I_* = (0, 1)$ ,  $I_1 = \{1\}$ ,  $E_0 = \mathbb{R}_+$ ,  $E_* = \{0\}$ ,  $E_1 = \mathbb{R}_-$ . This rewrites as

$$u_i \in I_\mu \Rightarrow u_i \in F_\mu := \{u_i | \partial_{u_i} \mathcal{J}_{\theta, \nu}(\mathbf{u}) \in E_\mu\}.$$

This leads us to compute  $u_i$  by using the following fixed-point relationship

$$u_i^{k+1} = 1 \times \mathbb{1}_{\{\partial_{u_i} \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) < -\kappa\}} + u_i^k \mathbb{1}_{\{-\kappa \leq \partial_{u_i} \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) \leq \kappa\}} + 0 \times \mathbb{1}_{\{\partial_{u_i} \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) > \kappa\}}, \quad (22)$$

where  $\kappa > 0$  is a small threshold preventing too small gradients from modifying the control. This finally leads us to algorithm 2.

**Remark 8.** *One important advantage of the fixed-point formulation can be illustrated by the fact that some common control configurations can lead to a flat gradient while optimisation is still possible: gradient descent is too local to detect certain appropriate and existing descent direction.*

---

**Algorithm 2** Optimal control with Fixed Point method (FP)

---

**Require:**  $\rho^0, \mathbf{u}^0, \text{tol} > 0, \text{itermax} > 0$

**Initialization:**  $k = 0$

**while**  $\|\mathbf{u}^{k+1} - \mathbf{u}^k\| > \text{tol}$  **and**  $k \leq \text{itermax}$  **do**

$$(\rho, \gamma) \leftarrow \text{solution to } \begin{cases} \rho' = f^{\text{FV}}(\rho, \gamma) \\ \gamma = \phi^{\text{LP}}(\rho, \mathbf{u}^k) \\ \rho(0) = \rho^0 \end{cases}$$

$$\mathbf{p} \leftarrow \text{solution to } \begin{cases} \mathbf{p}' + M(\rho, \gamma, \mathbf{u}^k)^T \mathbf{p} = \mathbf{0} \\ \mathbf{p}(T) = \mathbf{c} \end{cases}$$

$$\mathbf{u}^{k+1} \leftarrow \mathbb{1}_{\{\nabla \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) < -\kappa\}} + \mathbf{u}^k \mathbb{1}_{\{-\kappa \leq \nabla \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) \leq \kappa\}} + 0 \times \mathbb{1}_{\{\partial_{u_i} \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) > \kappa\}}$$

$$k \leftarrow k + 1$$

**end while**

---

It is noteworthy that the convergence of a fixed-point algorithm applied to an extremal eigenvalue problem has been established in [11], for certain parameters regimes involved in the problem. To the best of our knowledge, there is no general convergence result for this type of optimal control problem.

### 4.3.2 Hybrid GDFP methods

Since gradient descent theoretically always guarantees a direction of descent (provided we choose a small enough step) but can be slow and/or remain trapped in a local extremum (see remark 8), and since the fixed point appears more exploratory but does not guarantee descent, it seems worthwhile to investigate the hybridization of both methods.

The chosen algorithm is implemented by computing most of the iterations with the gradient descent method and using the fixed point method every  $K \in \mathbb{N}$  iterations in the expectation of escaping from possible basins of attraction of local minimizers. Note that  $K$  is a hyper-parameter of the method. This is summarized in algorithm 3.

Instead of performing FP steps regularly, we can choose to space them in order to have more iterations for the gradient descent to converge. This second version is given in algorithm 4 where the time between two FP steps increases by a factor of  $\tau$ .

## 5 Numerical Results

This section presents some results obtained in various situations using the methods presented above. In all that follows, we assume that  $\rho^{\max} = 1, v^{\max} = 1$  and  $L = 1$ . We will first validate our approach on single junctions before studying more complex road networks. These are namely a traffic circle and a three routes network of intermediate size with a configuration unfavorable to our objective. In all the test cases, we impose a flow equal to  $f_i(\rho_{\text{in}}^b)$  with  $\rho_{\text{in}}^b = 0.66$  on the incoming boundaries of the network and a Neumann flow boundary condition on all outgoing boundaries.

In the following numerical tests, the regularization parameters  $\nu$  and  $\eta$  are set surpventively equal to 0, after testing the algorithm for very low values of these parameters. We found no unusual phenomena

---

**Algorithm 3** Optimal control with hybrid GD & FP algorithm (GDFP)

---

**Require:**  $\boldsymbol{\rho}^0, \mathbf{u}^0, \text{tol} > 0, \text{itermax} > 0$

**Initialization:**  $k = 0$

**while**  $\|\mathbf{u}^{k+1} - \mathbf{u}^k\| > \text{tol}$  **and**  $k \leq \text{itermax}$  **do**

$$(\boldsymbol{\rho}, \gamma) \leftarrow \text{solution to } \begin{cases} \boldsymbol{\rho}' = f^{\text{FV}}(\boldsymbol{\rho}, \gamma) \\ \gamma = \phi^{\text{LP}}(\boldsymbol{\rho}, \mathbf{u}^k) \\ \boldsymbol{\rho}(0) = \boldsymbol{\rho}^0 \end{cases}$$

$$\mathbf{p} \leftarrow \text{solution to } \begin{cases} \mathbf{p}' + M(\boldsymbol{\rho}, \gamma, \mathbf{u}^k)^T \mathbf{p} = \mathbf{0} \\ \mathbf{p}(T) = \mathbf{c} \end{cases}$$

**if**  $k$  is multiple of  $K$  **then**

$$\mathbf{u}^{k+1} \leftarrow \mathbb{1}_{\{\nabla \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) < -\kappa\}} + \mathbf{u}^k \mathbb{1}_{\{-\kappa \leq \nabla \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) \leq \kappa\}} + 0 \times \mathbb{1}_{\{\partial_{u_i} \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) > \kappa\}}$$

**else**

$$\mathbf{u}^{k+1} \leftarrow \text{proj}_{[0,1]}(\mathbf{u}^k - \delta_k \nabla \mathcal{J}_{\theta, \nu}(\mathbf{u}^k)), \text{ with } \delta_k > 0 \text{ such that } \mathcal{J}_{\theta, \nu}(\mathbf{u}^{k+1}) \leq \mathcal{J}_{\theta, \nu}(\mathbf{u}^k)$$

**end if**

$$k \leftarrow k + 1$$

**end while**

---

---

**Algorithm 4** Optimal control with GDFP with spaced FP steps

---

**Require:**  $\boldsymbol{\rho}^0, \mathbf{u}^0, \text{tol} > 0, \text{itermax} > 0, K_0, \tau$

**Initialization:**  $k = 0, K = K_0$

**while**  $\|\mathbf{u}^{k+1} - \mathbf{u}^k\| > \text{tol}$  **and**  $k \leq \text{itermax}$  **do**

$$(\boldsymbol{\rho}, \gamma) \leftarrow \text{solution of } \begin{cases} \boldsymbol{\rho}' = f^{\text{FV}}(\boldsymbol{\rho}, \gamma) \\ \gamma = \phi^{\text{LP}}(\boldsymbol{\rho}, \mathbf{u}^k) \\ \boldsymbol{\rho}(0) = \boldsymbol{\rho}^0 \end{cases}$$

$$\mathbf{p} \leftarrow \text{solution of } \begin{cases} \mathbf{p}'(t) + M^T \mathbf{p} = \mathbf{0} \\ \mathbf{p}(T) = \mathbf{c} \end{cases}$$

**if**  $k$  equals  $K$  **then**

$$\mathbf{u}^{k+1} \leftarrow \mathbb{1}_{\{\nabla \mathcal{J}_{\theta}(\mathbf{u}^k) < -\kappa\}} + \mathbf{u}^k \mathbb{1}_{\{-\kappa \leq \nabla \mathcal{J}_{\theta}(\mathbf{u}^k) \leq \kappa\}} + 0 \times \mathbb{1}_{\{\partial_{u_i} \mathcal{J}_{\theta, \nu}(\mathbf{u}^k) > \kappa\}}$$

$$K \leftarrow \tau K$$

**else**

$$\mathbf{u}^{k+1} \leftarrow \text{proj}_{[0,1]}(\mathbf{u}^k - \delta_k \nabla \mathcal{J}_{\theta}(\mathbf{u}^k)), \text{ with } \delta_k > 0 \text{ such that } \mathcal{J}_{\theta}(\mathbf{u}^{k+1}) \leq \mathcal{J}_{\theta}(\mathbf{u}^k)$$

**end if**

$$k \leftarrow k + 1$$

**end while**

---

in the limit, and the simulations suggest the continuity of the minimizers. The question of choosing  $\nu$  and  $\eta$  is therefore related to the study of the continuous problem mentioned in the concluding section 6.

### 5.1 Single junctions

We start by considering single junctions of type  $1 \times 1$ ,  $1 \times 2$ ,  $2 \times 1$  and  $2 \times 2$  as they will be the building blocks of the more complex networks.

We consider 50 mesh cells per road and the initial density equals 0.66. The route to empty is always composed of an incoming and an outgoing road. The time interval of the simulations is adjusted in order to allow the route to be completely emptied: the final time  $T$  thus equals respectively 6, 3.5, 10 and 5 for the four junctions. We start the optimization algorithms with initial controls equal to 0 and set the convergence threshold for having  $\|\Lambda\|$  less than  $10^{-1}$ , with a prescribed maximum of 100 iterations. Neither constraints on the number of controls ( $\theta_S = 0$ ) nor BV regularization ( $\theta_B = 0$ ) are considered for these test cases.

Fig. 1 shows us the comparison between the cost functional history when using the gradient descent (GD), the fixed-point (FP) and the hybrid (GDFP) methods. Specific numerical parameters of the algorithm are given by table 1. It should be noted that we consider given parameter values, but we have systematically verified that wide ranges of values guarantee the results mentioned, even if this means modifying the number of iterations to guarantee numerical convergence. For example, for the weights  $\theta_S$  and  $\theta_B$ , any initial value below  $10^{-5}$  works effectively. The acceptable range for the parameter  $\kappa$  is for instance between 0 and 1 for the  $1 \times 2$  junction. Of course, the choice of  $\kappa$  influences the number of iterations required to guarantee numerical convergence.

We observe fundamental differences in behavior between the cost functionals obtained by GD and the one obtained by FP. Indeed, GD allows a regular descent whereas FP generates jumps and oscillations, which allows a better exploration of the parameters and avoids certain unsatisfactory local minima.

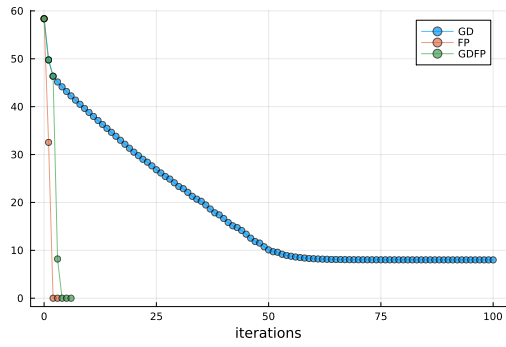
On fig. 2, we also plotted the evolution of the optimality conditions  $\Lambda$  for all methods: as expected, we observe that this quantity reaches values close to 0 at the optimal point.

The obtained controls are shown in fig. 3 and seem relevant. For instance, in the  $2 \times 1$  case, the control of the outgoing road 3 is fully activated only after time  $t \approx 8$  to enable the emptying of the incoming road 1 first. We further note that the controls are essentially sparse and non-oscillating. They are almost bang-bang, i.e. take only values 0 and 1.

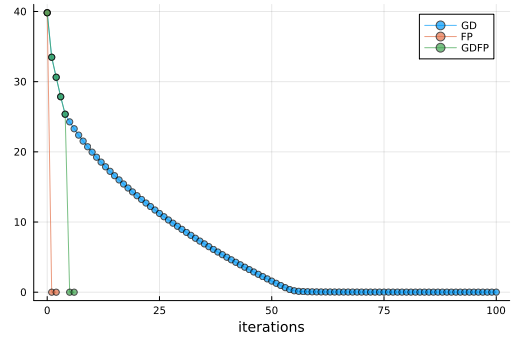
These results seem to suggest that the hybrid method is the most likely to generalize to larger graphs because of its ability to explore and find critical points while still being able to provide convergence. It is therefore the one we will use in the following.

Symbol	Name	1x1	1x2	2x1	2x2
K	FP trigger in GDFP	3	5	10	2
$\kappa$	vanishing gradient threshold in FP	0	0	0	$10^{-4}$
$\delta_0$	initial descent step	1	$5 \times 10^{-2}$	$2 \times 10^{-1}$	$10^{-1}$
decay	decay in scheduler	$10^{-2}$	$10^{-1}$	$10^{-1}$	$10^{-1}$

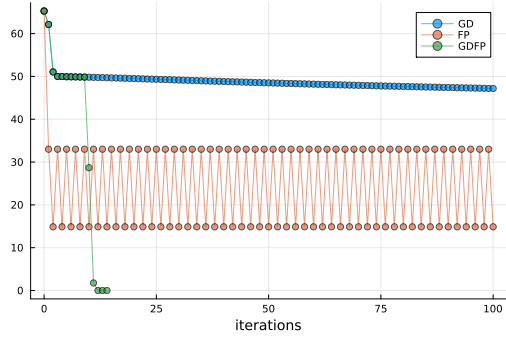
**Table 1:** Numerical parameters for single junctions.



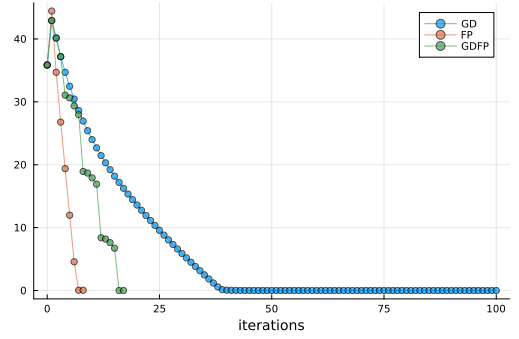
(a) 1x1



(b) 1x2

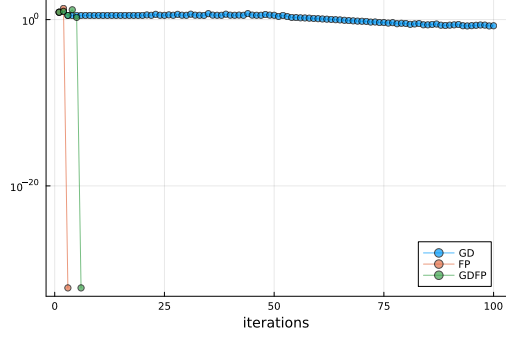


(c) 2x1

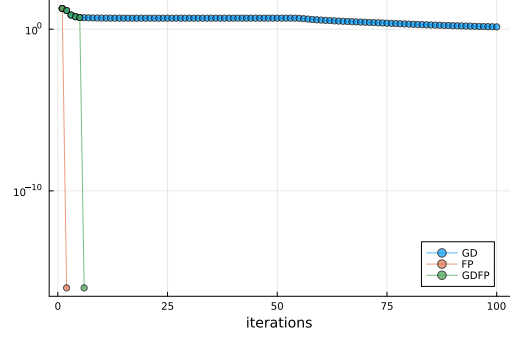


(d) 2x2

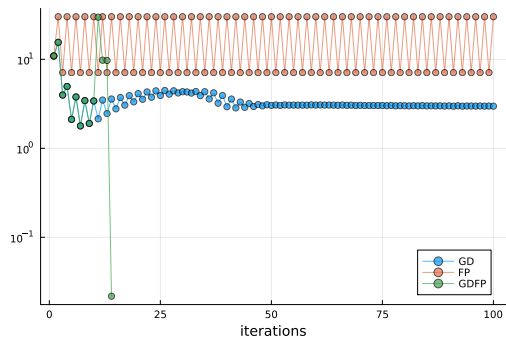
**Figure 1:** (Single junctions) Cost functional as a function of iterations for the gradient descent (GD), the fixed point (FP) and the hybrid (GDFP) methods.



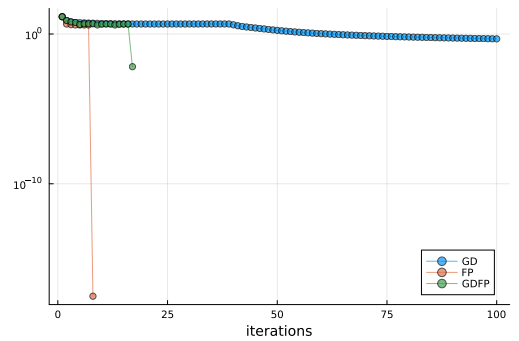
(a) 1x1



(b) 1x2

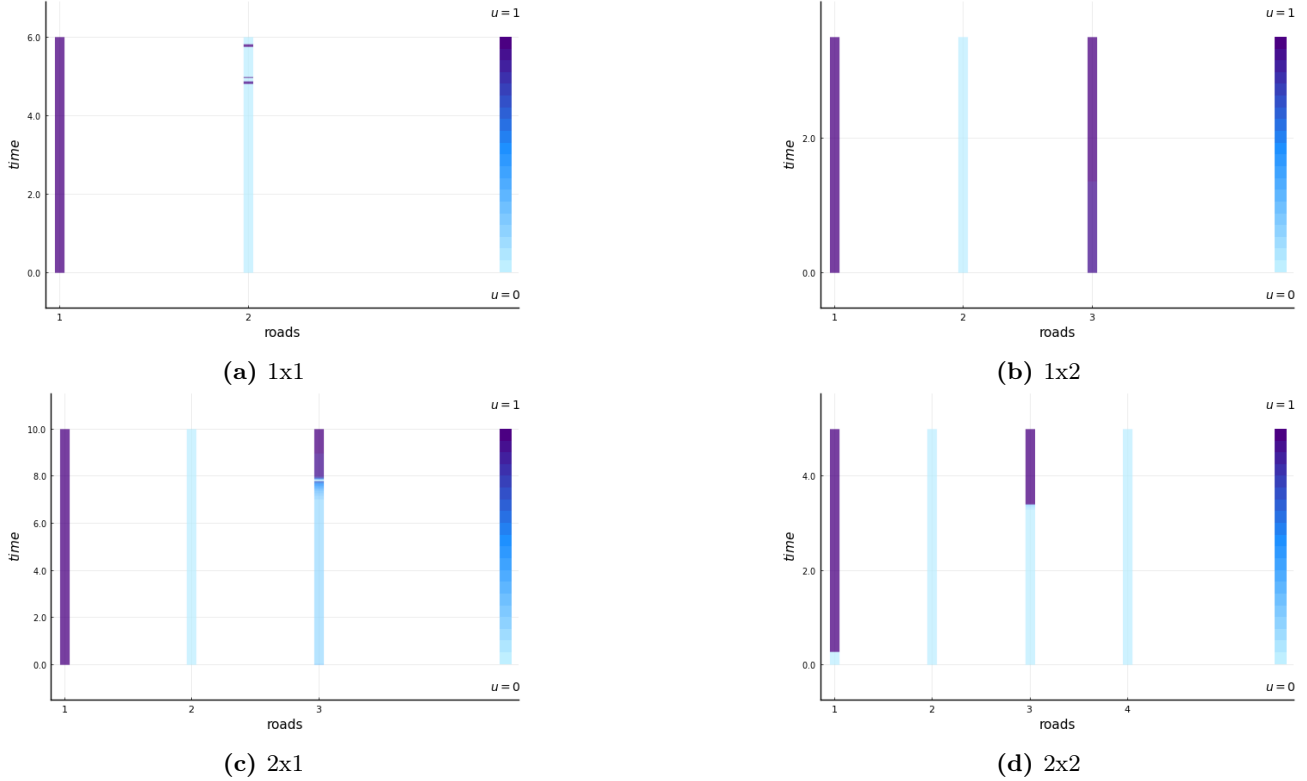


(c) 2x1



(d) 2x2

**Figure 2:** (Single junctions) Iterations of the convergence criterion  $\|\Lambda\|$ .



**Figure 3:** (Single junctions) Optimal controls obtained with the GDFP method. The routes to empty are respectively  $(1, 2)$ ,  $(1, 3)$ ,  $(1, 3)$ ,  $(1, 3)$ .

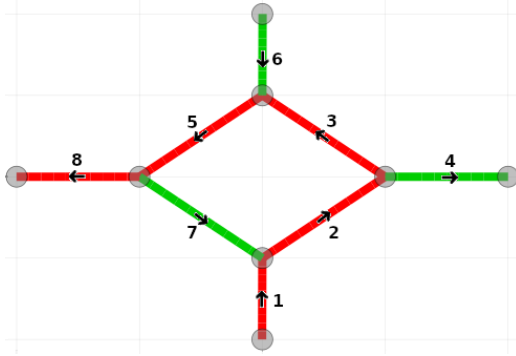
## 5.2 Traffic circle

We consider the traffic circle test case as proposed in [9]. It is composed of 8 roads as depicted in fig. 4a including 2 incoming roads, 2 outgoing roads and the 4 circle roads. The roads  $(2, 3, 5, 7)$  in the circle are counter-clockwise, roads  $(1, 6)$  are incoming and roads  $(4, 8)$  are outgoing. Initial density is taken constant equal to 0.66 on each road and, in fig. 4b, we observe that congestion appears if no control is applied, since boundary conditions make new vehicles constantly want to enter at the incoming boundaries.

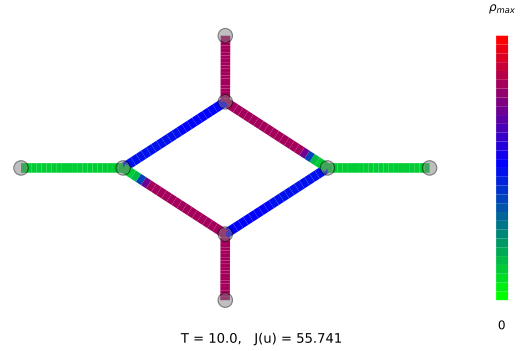
The route to evacuate is chosen to be  $(1, 2, 3, 5, 8)$  (see fig. 4a) and there are no constraints on the maximal number of controls ( $\theta_S = 0$ ) or BV regularization ( $\theta_B = 0$ ). As this test case are more difficult to handle than single junctions, we use Algorithm 4, which incorporates increasingly large spacing between fixed-point iterations: this gives the gradient iterations more time to converge. The parameters are given by table 2.

Results are gathered on fig. 5. On the upper left panel, the cost functional history is depicted. The first four iterations of the gradient descent make the cost functional decrease very slowly, except for the second iteration, and then a FP step triggers the jump seen at iteration 5. A few GD steps are then observed and then another FP step results in a second jump at iteration 10. Then the gradient descent is able to reach a satisfactory local minimum by the 12th iteration. We note that the optimality function  $\Lambda$  (fig. 5) follows essentially the same behavior and reaches a small value in the final iteration.

The control obtained by the algorithm is given in fig. 5c, bottom left. We observe first that roads 1 and 6 are always controlled since they are entry roads on the network. These controls acting on the incoming boundaries have an immediate impact on the total number of cars on the road network. That road 3 entrance is always controlled to drive the flow to the outgoing road 4. Furthermore, while road 5 entrance is mostly controlled after time  $t = 8$  to let cars from road 3 leave the route before, control on road 7 entrance is quite the opposite: up to time  $t = 8$ , it is activated to let the flow circulate as



(a) Road network and the route to empty in red.



(b) Density at the final time without control.

**Figure 4:** (Traffic circle) Left: Traffic circle configuration. The roads (2, 3, 5, 7) in the circle are counter-clockwise, roads (1, 6) are incoming and roads (4, 8) are outgoing. The route to empty (1, 2, 3, 5, 8) is in red. Right: Numerical simulations without control at time  $T = 10$ .

road 8 entrance is open, then from time  $t = 8$  to  $t = 10$ , it is deactivated as outgoing road 8 entrance is now closed. Finally, in fig. 5d, we can check that the route is effectively empty at the final time.

Parameters	$\rho_0$	$\mathbf{u}_0$	$N_c$	$\kappa$	tol	$K_0$	$\tau$	T	$\delta_0$	decay
Traffic circle	0.66	1	20	$10^{-6}$	$10^{-1}$	5	2	10	$5 \times 10^{-2}$	$10^{-1}$
Three routes network	0.66	1	5	$10^{-3}$	$10^{-2}$	5	3	30	1	$10^{-2}$

**Table 2:** Numerical parameters for the *traffic circle* and the *three routes* networks.

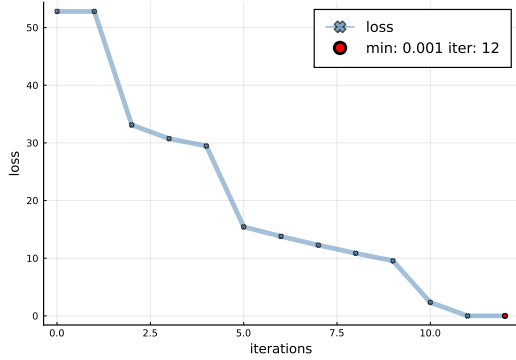
### 5.3 Three routes network

We consider the three routes network composed of 23 roads as shown in fig. 6a: all the roads are directed from left to right, with road 1 and road 11 being respectively the incoming and outgoing leaves. Initial density is still constant equal to 0.66 and a corresponding incoming flux is imposed at the entrance of the network on road 1. Neumann boundary conditions are used at the end of the outgoing road 11. The traffic distribution at each junction is uniform: there are no preferred trajectories. fig. 6b depicts the result of the simulation without control: the density in the central route tends to be saturated as several roads lead onto it.

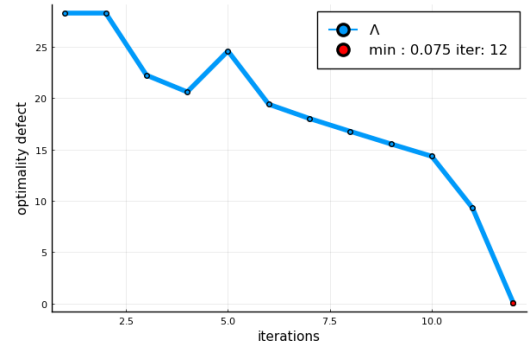
In the sequel, we would like to determine an optimal control in order to empty the central route, made of roads (4, 6, 7, 8, 9) (see fig. 6a), with  $N_{\max} = 5$ . Note that, contrary to the previous test cases, controlling this route may not prevent the flow from circulating from the incoming to the outgoing road.

We now give effect to the regularization by setting  $\theta_S = 10^{-4}$  and  $\theta_B = 10^{-6}$  for each iteration. We use the numerical parameters of table 2: in particular, FP step is used only at iteration 5.

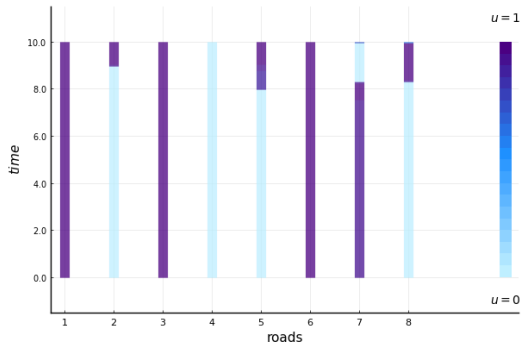




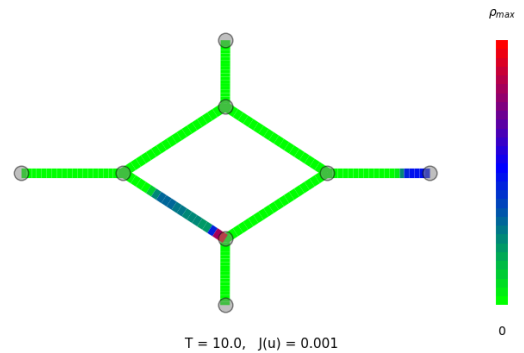
(a) Cost functional history



(b)  $\|\Lambda\|_2$  history

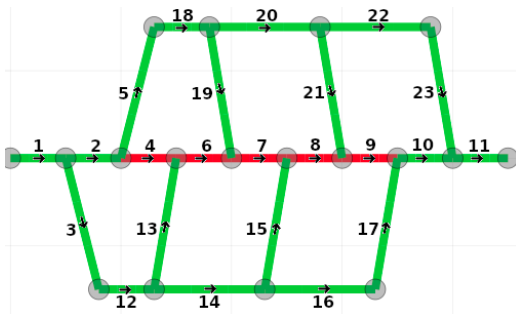


(c) Control

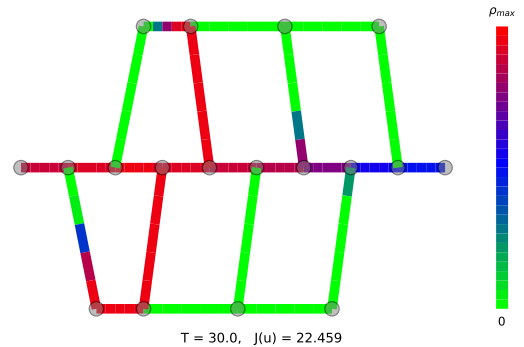


(d) Density at final time

**Figure 5:** (Traffic circle). Results obtained with GDFP algorithm with spaced FP steps.

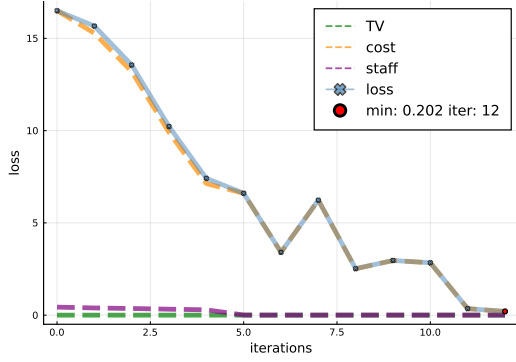


(a) Road network and the route to empty in red.

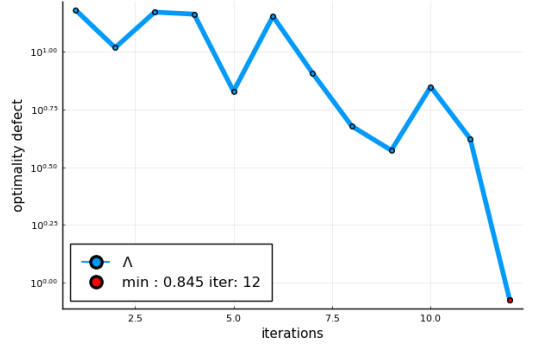


(b) Density at the final time without control.

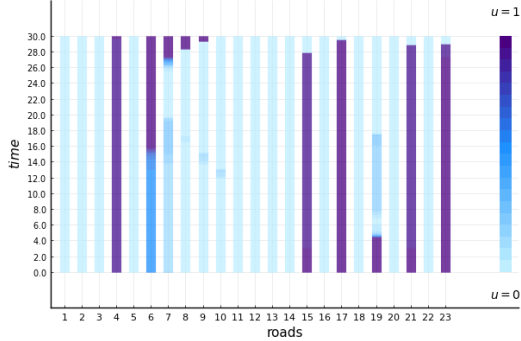
**Figure 6:** (Three routes network) Left: Three routes network configuration. Road 1 is incoming and road 11 is outgoing and all the others are directed from left to right. The route to empty (4, 6, 7, 8, 9) is in red. Right: Numerical simulations without control at time  $T = 30$ .



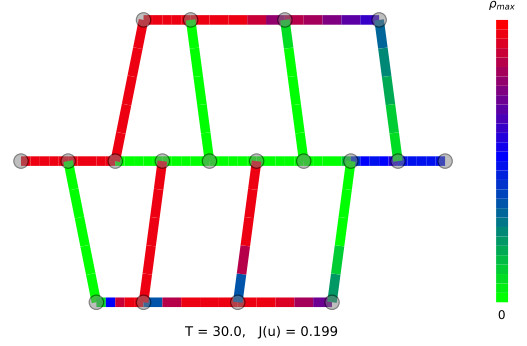
(a) Cost functional.



(b)  $\|\Lambda\|_2$  history.



(c) Control (at iteration 12).

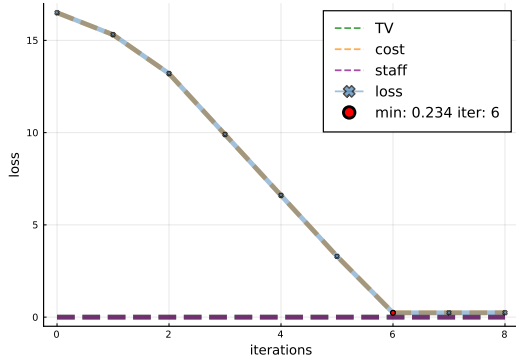


(d) Final time snapshot (at iteration 12).

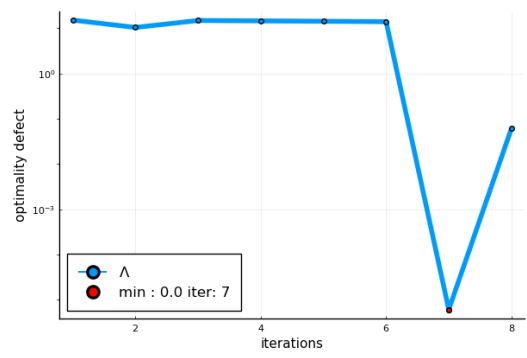
**Figure 7:** (Three routes network with activated constraints) Controls obtained with the GDFP algorithm.

The penalization algorithm has clearly solved the two issues: the number of active controls is now less than 7, which is still higher than 5 but far preferable to the earlier 16 in an approached control point of view, and there is no time oscillations of the controls even if it had been a mathematically correct way to reduce the staffing constraint.

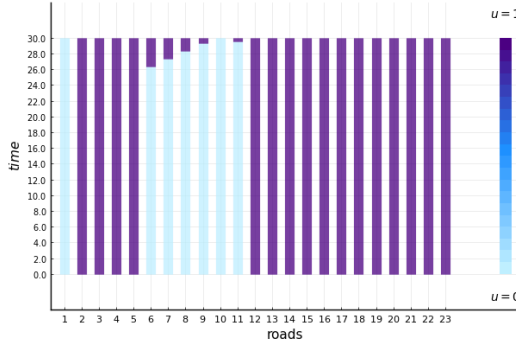
It is worth noticing that the fixed-point algorithm is also able to converge for this case, but only if we first consider the optimal control problem with essentially no limitation on the number of active controls ( $\theta_S = 10^{-8}$ ) nor BV regularization ( $\theta_B = 10^{-8}$ ) until  $\|\Lambda\|_2$  becomes lower than 1. We then consider that we are close to a basin of attraction and we increase the regularization weights by setting  $\theta_S = 10^{-4}$  and  $\theta_B = 10^{-6}$ . The associated plots are given in fig. 8 and fig. 9 where we used the parameters of table 2. Fig. 8a and Fig. 8b represents the cost functional history and the optimality defect history during the optimization process. We observe first a rapid decrease of the cost functional during the first six FP steps and  $\|\Lambda\|_2$  becomes lower than 1 at iteration 7. Fig. 8c shows the control provided by the algorithm at this iteration. The time evolution of the road densities with this control is represented in fig. 9c. The control performs rather well regarding the final cost functional and the final road densities on the central line. However, as it might be expected, this remains rather unsatisfactory given the relatively high number of active controls. When the regularization parameters  $\theta_S$  and  $\theta_B$  are changed at iteration 7, we observe a slight increase of the loss  $\mathcal{J}_\theta$ , mainly due to the increase of  $\theta_S$ . After the fixed-point step at iteration 8, the convergence criterion  $\|\Lambda\|_2 < 1$  is again satisfied and the algorithm stops.



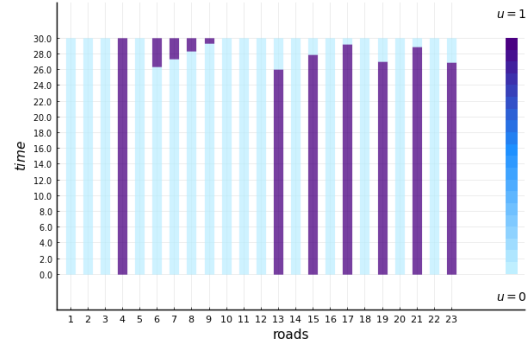
(a) Cost functional.



(b)  $\|\Lambda\|_2$  history.

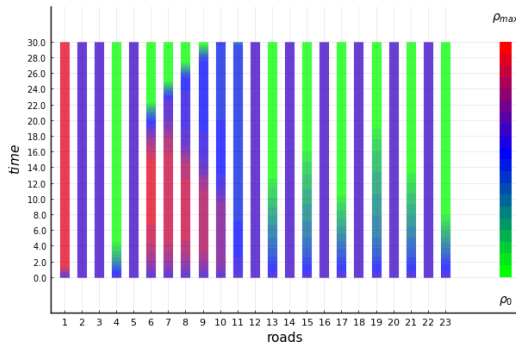


(c) Control (at iteration 6).

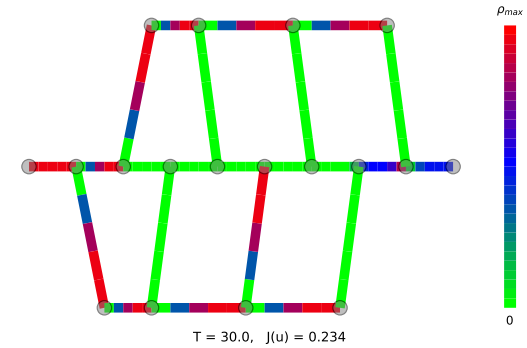


(d) Control (at iteration 8).

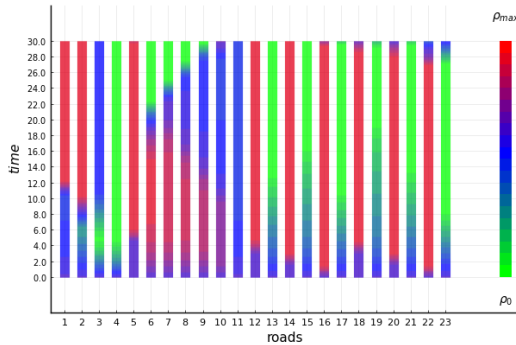
**Figure 8:** (Three routes network with activated constraints) Controls obtained with the FP algorithm.



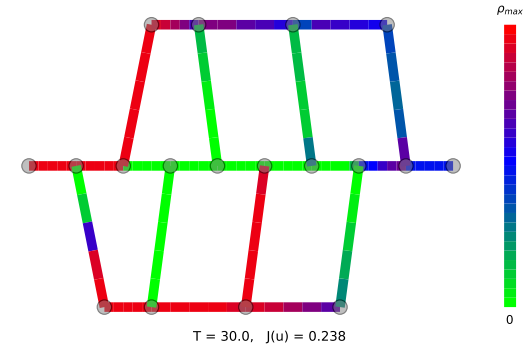
(a) Mean density (at iteration 6).



(b) Final time snapshot (at iteration 6).



(c) Mean density (at iteration 8).



(d) Final time snapshot (at iteration 8).

**Figure 9:** (Three routes network with scheduled constraints) Densities obtained with the FP algorithm.

## 6 Conclusion

This article has therefore introduced a new network traffic control problem where the control acts not only on the flux constraints at the entrance of the road but also on the traffic distribution matrices at the junctions. The theoretical analysis of the optimal control problem led to a hybrid numerical method, which combines the gradient descent method and a fixed-point strategy. The numerical results obtained on the traffic circle and on the three-routes network have shown that the method is relatively effective in providing relevant controls. It is also possible to impose constraints on the number of active controls.

This work will be pursued in several directions. Firstly, it would be interesting to show theoretically that optimal controls are indeed bang-bang, i.e. that they take only the two extreme values 0 and 1, as suggested by our numerical results. This could lead to further improvements in the optimization algorithm. For example, we plan to formulate a shape optimization problem whose unknown is the (temporal) domain over which control is equal to 1, and then apply efficient algorithms for such problems, such as the level set method.

Secondly, the indicator and projection functions make our algorithms sensitive to threshold or scaling effects. Further analysis could prove very useful in generically determining hyperparameters such as  $\theta$  or  $\kappa$ , given that they are chosen heuristically so far. Based on our testing, a wide range of values for  $\kappa$  have proven effective. However, in some test cases,  $\kappa$  should not be too small to ensure algorithm convergence. For instance in the  $2 \times 2$  test case, the FP algorithm is stuck in a local minimum for any  $\kappa$  whose order of magnitude is less than  $10^{-4}$ , and converges for every  $\kappa \in [10^{-4}, 1]$ . As for the value of  $K$ , it mainly influences the number of iterations required for convergence in single junctions. However, in more complex test cases, such as the traffic circle scenario, it can significantly alter the behavior of the algorithm. In these cases, solutions that are less easily reachable may not be found if the fixed-point iterations take steps that are too large, causing the control to move away from these solutions. Finally, the problem has been tackled here using a strategy of discretization (in space) followed by optimization. Another possibility would have been to optimize first and then discretize. However this would require to deal with a non-differentiable cost function since the density is non-differentiable anymore. Such a strategy may give rise to a different type of algorithm.

More generally, we consider it important to clarify the links between the solutions of fully discretized (in time and space), semi-discretized (in space) and continuous (without discretization) versions of the problem. These are fairly advanced  $\Gamma$ -convergence issues, which will be the subject of future reflection. In particular, we will seek to implement an argument to certify that the approach chosen in this article does not create "numerical pathology" when trying to approximate solutions to the continuous version of the optimization problem.

A longer-term objective we plan to tackle in the coming years is to generalize and test our algorithm both on larger road networks and on real data. To this end, we will seek to couple this method with relevant initialization methods for which initial control would be obtained by considering a macroscopic graph and implementing model reduction methods. Such an approach has, for example, been developed in a different context in epidemiology [15].

## Acknowledgments

The fourth author is supported by the Agence Nationale de la Recherche, Project TRECOS.

Appendix

## A Expression of the function $\phi^{LP}$

We provide hereafter the explicit expression of  $\phi^{LP}$  for junctions with at most 2 incoming and 2 outgoing roads obtained by solving the Linear Programming problem (3)-(4). We will write  $(\gamma^R, \gamma^L) = \phi^{LP}(\boldsymbol{\rho}, \mathbf{u})$ . We then introduce regularized approximations of these in order to deal with differential optimization techniques.

**One incoming and one outgoing roads ( $n = m = 1$ ).** In that case, the solution to the LP problem (3)-(4) is given by

$$\gamma_1^R = \gamma_2^L = \min\left(\gamma_1^{R,\max}, (1 - u_2)\gamma_2^{L,\max}\right).$$

**Two incoming and one outgoing roads ( $n = 2, m = 1$ ).** As proposed in [8], we introduce the priority parameter  $q \in [0, 1]$ : this sets the proportion of vehicles entering the outgoing road and coming from the first incoming road if the junction is saturated (meaning that not all vehicles can enter the outgoing road). Let us index the incoming roads by 1, 2 and the outgoing one by 3. Then the solution to the LP problem (3)-(4) leads to  $\gamma_3^L = \gamma_1^R + \gamma_2^R$ , and the outgoing fluxes  $\gamma_1^R, \gamma_2^R$  are computed as follows. Denoting

$$\gamma_{3,u}^{L,\max} = (1 - u_3)\gamma_3^{L,\max},$$

we have the alternative

- if  $\gamma_1^{R,\max} + \gamma_2^{R,\max} \leq \gamma_{3,u}^{L,\max}$ , then  $(\gamma_1^R, \gamma_2^R) = (\gamma_1^{R,\max}, \gamma_2^{R,\max})$ ,
- else
  - if  $\gamma_1^{R,\max} \geq q\gamma_{3,u}^{L,\max}$  and  $\gamma_2^{R,\max} \geq (1 - q)\gamma_{3,u}^{L,\max}$ , then  $(\gamma_1^R, \gamma_2^R) = (q\gamma_{3,u}^{L,\max}, (1 - q)\gamma_{3,u}^{L,\max})$ .
  - if  $\gamma_1^{R,\max} < q\gamma_{3,u}^{L,\max}$  and  $\gamma_2^{R,\max} \geq (1 - q)\gamma_{3,u}^{L,\max}$ , then  $(\gamma_1^R, \gamma_2^R) = (\gamma_1^{R,\max}, \gamma_{3,u}^{L,\max} - \gamma_1^{R,\max})$ .
  - if  $\gamma_1^{R,\max} \geq q\gamma_{3,u}^{L,\max}$  and  $\gamma_2^{R,\max} < (1 - q)\gamma_{3,u}^{L,\max}$ , then  $(\gamma_1^R, \gamma_2^R) = (\gamma_{3,u}^{L,\max} - \gamma_2^{R,\max}, \gamma_2^{R,\max})$ .

**One incoming and two outgoing roads ( $n = 1, m = 2$ ).** Let us index the incoming roads by 1 and the outgoing one by 2 and 3. Denoting  $\alpha = \alpha_{21}$ , the solution to the original LP problem (3)-(4) is given by:

$$\begin{aligned} \gamma_1^R &= \min\left(\gamma_1^{R,\max}, (1 - u_2)\frac{\gamma_2^{L,\max}}{\alpha}, (1 - u_3)\frac{\gamma_3^{L,\max}}{1 - \alpha}\right), \\ \gamma_2^L &= \alpha\gamma_1^R, \quad \gamma_3^L = (1 - \alpha)\gamma_1^R. \end{aligned}$$

However, this formulation raises a modeling problem. Indeed, this solution does not distinguish between a blocked road ( $\gamma_j^{L,\max} = 0$ ) and a controlled road ( $1 - u_j = 0$ ). Indeed, the first case corresponds to a congestion situation involving cars on both sides of the intersection, preventing vehicles from entering the second -potentially empty- road, while in the second case, drivers will not try to enter the controlled road and we revert to the  $1 \times 1$  configuration with the remaining roads.

To capture this difference in driver behavior, we will adjust the parameter  $\alpha$  with respect to the control  $\mathbf{u}$ . Therefore,  $\alpha$  will be considered as a function of  $(u_2, u_3)$ . More precisely, we introduce a function denoted  $P$  of the variable  $x = u_2 - u_3$ , taking the constant value  $\bar{\alpha} \in (0, 1)$  at  $x = 0$  and yielding the relevant  $1 \times 1$  case whenever one road is fully controlled. For the sake of simplicity, we look for a function  $(u_2, u_3) \mapsto P_{\bar{\alpha}}(u_2 - u_3)$ , where  $P_{\bar{\alpha}}$  is a polynomial of degree at most 2, that satisfies

$$P_{\bar{\alpha}}(-1) = \alpha(0, 1) = 1, \quad P_{\bar{\alpha}}(0) = \alpha(u, u) = \bar{\alpha}, \quad P_{\bar{\alpha}}(1) = \alpha(1, 0) = 0.$$

However, due to the division by  $\alpha$  in the expression of  $\gamma_1^R$ , the case  $\alpha(1,0) = 0$  is degenerate. To overcome this problem, we introduce a small perturbation parameter  $\varepsilon > 0$  and a modification of the polynomial  $P_{\bar{\alpha}}$  denoted  $P_{\bar{\alpha},\varepsilon}$ , which satisfies

$$P_{\bar{\alpha},\varepsilon}(-1) = 1 - \varepsilon^2, \quad P_{\bar{\alpha},\varepsilon}(0) = \alpha, \quad P_{\bar{\alpha},\varepsilon}(1) = \varepsilon^2.$$

The standard Lagrange interpolation formula yields

$$P_{\bar{\alpha},\varepsilon}(x) = \frac{x(x-1)}{2} + \bar{\alpha}(1-x^2) + \varepsilon^2 x. \quad (23)$$

Since the interpolation procedure does not guarantee that the range of the function is contained in  $[\varepsilon^2, 1 - \varepsilon^2]$ , we compose the obtained expression with a projection onto the set of admissible values, to get at the end

$$\alpha_\varepsilon(u_2, u_3) := \text{proj}_{[\varepsilon^2, 1-\varepsilon^2]}(P_{\bar{\alpha},\varepsilon}(u_2 - u_3)). \quad (24)$$

Similarly, we also replace the expression  $(1 - u_j)$  by  $(1 - u_j + \varepsilon)$  to avoid the case where  $(1 - u_2) = \alpha(u_2, u_3) = 0$ . We finally obtain the following regularized solution at  $1 \times 2$  junctions

$$\begin{aligned} \gamma_1^R &= \min \left( \gamma_1^{R,\max}, (1 - u_2 + \varepsilon) \frac{\gamma_2^{L,\max}}{\alpha_\varepsilon(u_2, u_3)}, (1 - u_3 + \varepsilon) \frac{\gamma_3^{L,\max}}{1 - \alpha_\varepsilon(u_2, u_3)} \right), \\ \gamma_2^L &= \alpha_\varepsilon(u_2, u_3) \gamma_1^R, \quad \gamma_3^L = (1 - \alpha_\varepsilon(u_2, u_3)) \gamma_1^R. \end{aligned}$$

**Two incoming and two outgoing roads ( $n = m = 2$ ).** The incoming roads are indexed by 1, 2 and the outgoing ones are indexed by 3, 4. As in the  $1 \times 2$  junction, the distribution traffic is chosen depending on the control  $u$ . We thus have the following relation

$$\begin{pmatrix} \gamma_3^L \\ \gamma_4^L \end{pmatrix} = A(\mathbf{u}) \begin{pmatrix} \gamma_1^R \\ \gamma_2^R \end{pmatrix} \quad \text{with} \quad A(\mathbf{u}) = \begin{pmatrix} \alpha_\varepsilon(\mathbf{u}) & \beta_\varepsilon(\mathbf{u}) \\ 1 - \alpha_\varepsilon(\mathbf{u}) & 1 - \beta_\varepsilon(\mathbf{u}) \end{pmatrix},$$

where the coefficients  $\alpha_\varepsilon(\mathbf{u})$  and  $\beta_\varepsilon(\mathbf{u})$  are defined by

$$\alpha_\varepsilon(\mathbf{u}) = \text{proj}_{[\varepsilon^2, 1-\varepsilon^2]}(P_{\bar{\alpha},\varepsilon}(u_3 - u_4)), \quad \beta_\varepsilon(\mathbf{u}) = \text{proj}_{[\varepsilon^2, 1-\varepsilon^2]}(P_{\bar{\beta},\varepsilon}(u_3 - u_4)).$$

with  $\bar{\alpha}, \bar{\beta}$  in  $(0, 1)$ .

To solve the resulting  $2 \times 2$  LP problem, we first analyze the simplex standing for the polytope of constraints, represented in Figure 10: the constraints related to the incoming roads form the rectangle  $\Omega_{\text{in}} := [0, \gamma_1^{R,\max}] \times [0, \gamma_2^{R,\max}]$ , while those related to the outgoing roads correspond to the regions below the lines  $\mathcal{C}_3$  and  $\mathcal{C}_4$ , whose equations are

$$\alpha_\varepsilon(\mathbf{u}) \gamma_1^R + \beta_\varepsilon(\mathbf{u}) \gamma_2^R = (1 - u_3) \gamma_3^{L,\max} \quad \text{and} \quad (1 - \alpha_\varepsilon(\mathbf{u})) \gamma_1^R + (1 - \beta_\varepsilon(\mathbf{u})) \gamma_2^R = (1 - u_4) \gamma_4^{L,\max}.$$

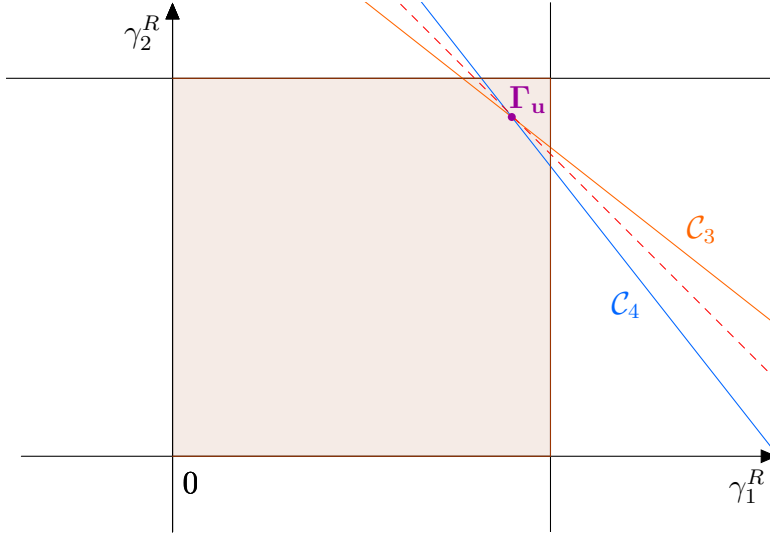
A convexity argument, standard in linear optimization, yields to the existence of at least one solution lying on a vertex of the set of constraints.

In the case where  $\alpha_\varepsilon(\mathbf{u}) = \beta_\varepsilon(\mathbf{u})$ , the two straight lines  $\mathcal{C}_3$  and  $\mathcal{C}_4$  are parallel. One of the constraint can be temporally removed and the problem actually boils down to the case of a  $2 \times 1$  junction. The outgoing fluxes  $\gamma_1^R, \gamma_2^R$  are then obtained using the same algorithm as described previously. Note that, in that case, a priority parameter  $q \in [0, 1]$  need to be introduced to select one condition.

We suppose now that  $\alpha_\varepsilon(\mathbf{u})$  and  $\beta_\varepsilon(\mathbf{u})$  take different values. To discuss the optimality of each vertex, let us introduce the intersection of lines  $\mathcal{C}_3$  and  $\mathcal{C}_4$  and denote it:

$$\Gamma_{\mathbf{u}} = (\Gamma_{\mathbf{u},1}, \Gamma_{\mathbf{u},2}) = ((1 - \beta_\varepsilon(\mathbf{u})) \gamma_{3,\mathbf{u}}^{\max} - \beta_\varepsilon(\mathbf{u}) \gamma_{4,\mathbf{u}}^{\max}) / \Delta, (-(1 - \alpha_\varepsilon(\mathbf{u})) \gamma_{3,\mathbf{u}}^{\max} - \alpha_\varepsilon(\mathbf{u}) \gamma_{4,\mathbf{u}}^{\max}) / \Delta),$$

with  $\Delta = \alpha_\varepsilon(\mathbf{u})(1 - \beta_\varepsilon(\mathbf{u})) - \beta_\varepsilon(\mathbf{u})(1 - \alpha_\varepsilon(\mathbf{u}))$  and  $\gamma_{i,\mathbf{u}}^{L,\max} = (1 - u_i) \gamma_i^{L,\max}$ , for  $i = 3, 4$ . A standard graphical reasoning shows that the solution is realized at  $\Gamma_{\mathbf{u}}$ , if  $\Gamma_{\mathbf{u}}$  belongs to  $\Omega_{\text{in}}$ , as in fig. 10. Otherwise, the solution is at intersection of the constraint with largest slope (in absolute value) with the boundary of  $\Omega_{\text{in}}$ . It is this disjunction that is performed in what follows.



**Figure 10:** Polytope of constraints for the 2x2 LP. The colored area represents the domain  $\Omega_{in}$ . The red dashed line represent the line  $\gamma_1^R + \gamma_2^R = \text{cst}$ , passing through  $\Gamma_u$ .

- **First case:**  $\Gamma_{u,1} \leq \gamma_1^{\max}$  and  $\Gamma_{u,2} \leq \gamma_2^{\max}$ ;

- if  $\Gamma_{u,1} \geq 0$  and  $\Gamma_{u,2} < 0$ , then  $(\gamma_1^R, \gamma_2^R) = \begin{cases} (\gamma_{3,\mathbf{u}}^{\max}/\alpha_\varepsilon(\mathbf{u}), 0) & \text{if } \alpha_\varepsilon(\mathbf{u}) < \beta_\varepsilon(\mathbf{u}) \\ (\gamma_{4,\mathbf{u}}^{\max}/(1 - \alpha_\varepsilon(\mathbf{u})), 0) & \text{else} \end{cases}$
- if  $\Gamma_{u,1} < 0$  and  $\Gamma_{u,2} \geq 0$ , then  $(\gamma_1^R, \gamma_2^R) = \begin{cases} (0, \gamma_{3,\mathbf{u}}^{\max}/\beta_\varepsilon(\mathbf{u})) & \text{if } \alpha_\varepsilon(\mathbf{u}) < \beta_\varepsilon(\mathbf{u}) \\ (0, \gamma_{4,\mathbf{u}}^{\max}/(1 - \beta_\varepsilon(\mathbf{u}))) & \text{else} \end{cases}$
- if  $\Gamma_{u,1} < 0$  and  $\Gamma_{u,2} < 0$ , then  $(\gamma_1^R, \gamma_2^R) = (0, 0)$ ;
- else  $(\gamma_1^R, \gamma_2^R) = (\Gamma_{u,1}, \Gamma_{u,2})$ .

- **Second case:**  $\Gamma_{u,1} > \gamma_1^{\max}$  and  $\Gamma_{u,2} > \gamma_2^{\max}$ . Then,  $(\gamma_1^R, \gamma_2^R) = (\gamma_1^{\max}, \gamma_2^{\max})$ .

- **Third case:**  $\Gamma_{u,1} > \gamma_1^{\max}$  and  $\Gamma_{u,2} \leq \gamma_2^{\max}$ ;

- if  $\alpha_\varepsilon(\mathbf{u}) < \beta_\varepsilon(\mathbf{u})$ , then

$$\gamma_2^R = \text{proj}_{[0, \gamma_2^{\max}]} \left( \frac{\gamma_{3,\mathbf{u}}^{\max} - \alpha_\varepsilon(\mathbf{u})\gamma_1^{\max}}{\beta_\varepsilon(\mathbf{u})} \right), \quad \gamma_1^R = \begin{cases} \gamma_3^{\max}/\alpha_\varepsilon(\mathbf{u}) & \text{if } \gamma_2^R = 0 \\ \gamma_1^{\max} & \text{if } \gamma_2^R \in (0, \gamma_2^{\max}] \end{cases};$$

- else,

$$\gamma_2^R = \text{proj}_{[0, \gamma_2^{\max}]} \left( \frac{\gamma_{4,\mathbf{u}}^{\max} - (1 - \alpha_\varepsilon(\mathbf{u}))\gamma_1^{\max}}{1 - \beta_\varepsilon(\mathbf{u})} \right), \quad \gamma_1^R = \begin{cases} \gamma_4^{\max}/(1 - \alpha_\varepsilon(\mathbf{u})) & \text{if } \gamma_2^R = 0 \\ \gamma_1^{\max} & \text{if } \gamma_2^R \in (0, \gamma_2^{\max}] \end{cases};$$

- **Fourth case:**  $\Gamma_{u,1} \leq \gamma_1^{\max}$  and  $\Gamma_{u,2} > \gamma_2^{\max}$ ;

- if  $\alpha_\varepsilon(\mathbf{u}) > \beta_\varepsilon(\mathbf{u})$ , then

$$\gamma_1^R = \text{proj}_{[0, \gamma_1^{\max}]} \left( \frac{\gamma_{3,\mathbf{u}}^{\max} - \beta_\varepsilon(\mathbf{u})\gamma_2^{\max}}{\alpha_\varepsilon(\mathbf{u})} \right), \quad \gamma_2^R = \begin{cases} \gamma_3^{\max}/\beta_\varepsilon(\mathbf{u}) & \text{if } \gamma_1^R = 0 \\ \gamma_2^{\max} & \text{if } \gamma_1^R \in (0, \gamma_1^{\max}] \end{cases};$$

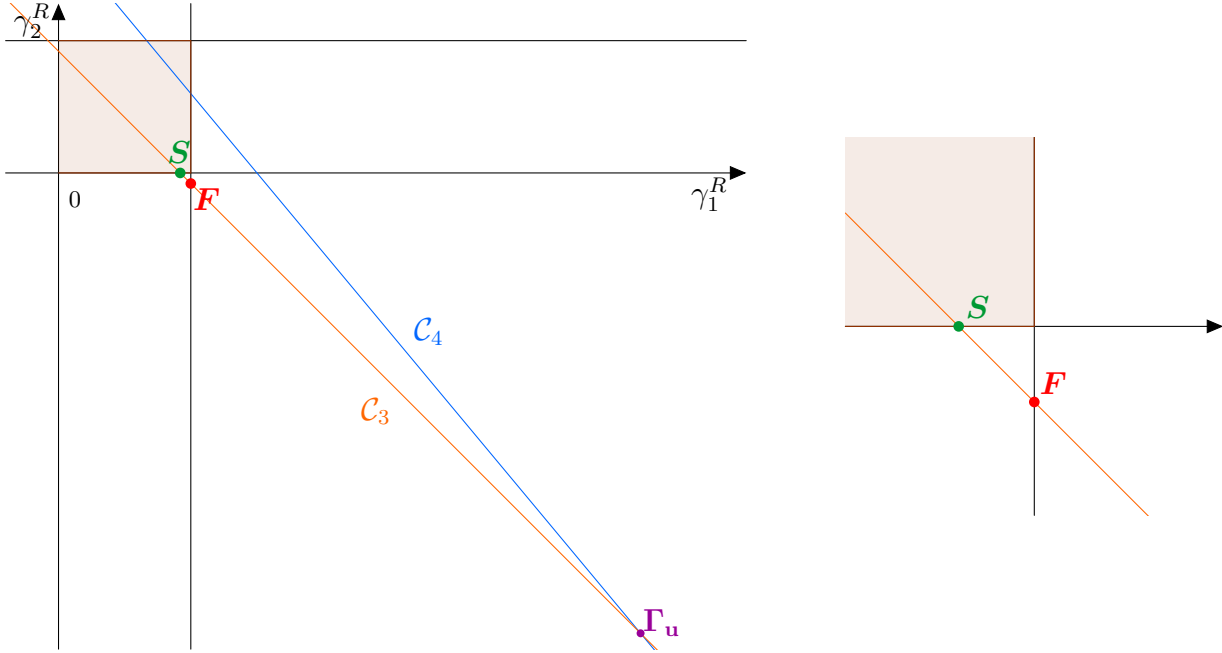
- else,

$$\gamma_1^R = \text{proj}_{[0, \gamma_1^{\max}]} \left( \frac{\gamma_{4,\mathbf{u}}^{\max} - (1 - \beta_\varepsilon(\mathbf{u}))\gamma_2^{\max}}{1 - \alpha_\varepsilon(\mathbf{u})} \right), \quad \gamma_2^R = \begin{cases} \gamma_4^{\max}/(1 - \beta_\varepsilon(\mathbf{u})) & \text{if } \gamma_1^R = 0 \\ \gamma_2^{\max} & \text{if } \gamma_1^R \in (0, \gamma_1^{\max}] \end{cases}.$$

**Remark 9.** It should be noted that a similar LP problem has been solved in [9]. However, we found that a case corresponding to a particular constraint configuration was forgotten in their analysis. This error could lead to erroneous vertices providing negative values to some fluxes. An example of such a configuration is illustrated on fig. 11, for the parameters given in table 3.

Parameter	$\gamma_1^{\max}$	$\gamma_2^{\max}$	$\gamma_3^{\max}$	$\gamma_4^{\max}$	$u_3$	$u_4$	$\bar{\alpha}$	$\bar{\beta}$	$\alpha_\varepsilon(\mathbf{u})$	$\beta_\varepsilon(\mathbf{u})$	$\tilde{\gamma}_1^R$	$\tilde{\gamma}_2^R$
Value	0.25	0.25	0.2244	0.2244	0	0.96	0.45	0.50	0.97608	0.98	0.25	-0.02002

**Table 3:** Parameters for the 2x2 LP problem in Figure 11.



**Figure 11:** Graphical resolution of the  $2 \times 2$  LP problem. Polytope of the constraints (left) and a zoom (right).  $F$  is the erroneous solution and  $S$  is the correct one.

**Regular approximation of  $\phi^{LP}$ .** In order to use differentiable optimization techniques, we will replace the function  $\phi^{LP}$ , which is only Lipschitz, by an approximation  $C^1$ , replacing all operations consisting in taking the minimum or the maximum of two quantities by regular approximations:

$$\min\{x, y\} \leftarrow \frac{x + y - \sqrt{(x + y)^2 + \eta^2}}{2} \quad \text{and} \quad \max\{x, y\} \leftarrow \frac{x + y + \sqrt{(x + y)^2 + \eta^2}}{2}$$

for a given  $\eta > 0$ .

## References

- [1] F. Ancona, L. Caravenna, A. Cesaroni, G. M. Coclite, C. Marchi, and A. Marson. Analysis and control on networks: Trends and perspectives. *Networks & Heterogeneous Media*, 12(3), 2017.
- [2] F. Ancona, A. Cesaroni, G. M. Coclite, and M. Garavello. On the optimization of conservation law models at a junction with inflow and flow distribution controls. *SIAM Journal on Control and Optimization*, 56(5):3370–3403, 2018.



- [3] A. Aw and M. Rascle. Resurrection of "second order" models of traffic flow. SIAM J. Appl. Math., 60:916–938, 2000.
- [4] U. Baumgart and M. Burger. Optimal control of traffic flow based on reinforcement learning. In C. Klein, M. Jarke, M. Helfert, K. Berns, and O. Gusikhin, editors, Smart Cities, Green Technologies, and Intelligent Transport Systems, pages 313–329, Cham, 2022. Springer International Publishing.
- [5] A. Bayen, M.-L. Delle Monache, M. Garavello, P. Goatin, and B. Piccoli. Control Problems for Conservation Laws with Traffic Applications: Modeling, Analysis, and Numerical Methods, volume 99 of Progress in Nonlinear Differential Equations and Their Applications. Springer International Publishing, Cham, 2022.
- [6] A. Bayen, R. Raffard, and C. Tomlin. Adjoint-based control of a new eulerian network model of air traffic flow. IEEE Transactions on Control Systems Technology, 14(5):804–818, Sept. 2006.
- [7] O. Biham, A. A. Middleton, and D. Levine. Self-organization and a dynamical transition in traffic-flow models. Phys. Rev. A, 46:R6124–R6127, Nov 1992.
- [8] G. Bretti, R. Natalini, and B. Piccoli. Numerical approximations of a traffic flow model on networks. Networks and Heterogeneous Media, 1(1):57, 2006.
- [9] S. Canic, B. Piccoli, J.-M. Qiu, and T. Ren. Runge-kutta discontinuous galerkin method for traffic flow model on networks. Journal of Scientific Computing, 2015.
- [10] A. Chambolle. An algorithm for total variation minimization and applications. Journal of Mathematical imaging and vision, 20:89–97, 2004.
- [11] A. Chambolle, I. Mazari-Fouquer, and Y. Privat. Stability of optimal shapes and convergence of thresholding algorithms in linear and spectral optimal control problems. arXiv preprint arXiv:2306.14577, 2023.
- [12] Y. Chitour and B. Piccoli. Traffic circles and timing of traffic lights for cars flow. Discrete Contin. Dyn. Syst. Ser. B, 5:599–630, Aug. 2005.
- [13] G. M. Coclite, M. Garavello, and B. Piccoli. Traffic flow on a road network. SIAM Journal on Mathematical Analysis, 36(6):1862–1886, 2005.
- [14] V. Coscia, M. Delitala, and P. Frasca. On the mathematical theory of vehicular traffic flow II: Discrete velocity kinetic models. International Journal of NonLinear Mechanics, 42(3):411–421, 2007.
- [15] C. Courtès, E. Franck, K. Lutz, L. Navoret, and Y. Privat. Reduced modelling and optimal control of epidemiological individual-based models with contact heterogeneity. Optimal Control Applications and Methods, 2023.
- [16] M. Delitala and A. Tosin. Mathematical modeling of vehicular traffic: a discrete kinetic theory approach. Mathematical Models and Methods in Applied Sciences, 17(06):901–932, 2007.
- [17] I. Ekeland and R. Témam. Convex analysis and variational problems, volume 28 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.
- [18] L. Fermo and A. Tosin. Fundamental diagrams for kinetic equations of traffic flow. Discrete and Continuous Dynamical Systems - Series S, 7:449–462, 2014.
- [19] A. Fügenschuh, M. Herty, A. Klar, and A. Martin. Combinatorial and Continuous Models for the Optimization of Traffic Flows on Networks. SIAM Journal on Optimization, 16(4):1155–1176, Jan. 2006.
- [20] B. P. Gabrielle Bretti, Roberto Natalini. A fluid-dynamic traffic model on road networks. Archives of Computational Methods in Engineering, 2007.
- [21] M. Garavello and B. Piccoli. Traffic flow on networks. American Institute of Mathematical Sciences, 2006.
- [22] P. Gipps. A behavioural car-following model for computer simulation. Transportation Research Part B: Methodological, 1981.
- [23] P. Goatin, S. Göttlich, and O. Kolb. Speed limit and ramp meter control for traffic flow networks. Engineering Optimization, 48(7):1121–1144, July 2016.

- [24] M. Gugat, M. Herty, A. Klar, and G. Leugering. Optimal Control for Traffic Flow Networks. Journal of Optimization Theory and Applications, 126(3):589–616, Sept. 2005.
- [25] S. Göttlich, M. Herty, and U. Ziegler. Modeling and optimizing traffic light settings in road networks. Computers & Operations Research, 55:36–51, Mar. 2015.
- [26] R. Jayakrishnan, H. S. Mahmassani, and T.-Y. Hu. An evaluation tool for advanced traffic information and management systems in urban networks. Transportation Research Part C: Emerging Technologies, 1994.
- [27] J. Krug and H. Spohn. Universality classes for deterministic surface growth. Physical Review A, 1988.
- [28] A. A. Kurzhanskiy and P. Varaiya. Active traffic management on road networks: a macroscopic approach. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 368(1928):4607–4626, Oct. 2010.
- [29] J. Lebacque. The godunov scheme and what it means for first order traffic flow models. In Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Lyon, France, July, volume 2426, 1996.
- [30] E. B. Lee and L. Markus. Foundations of optimal control theory, volume 24. Wiley New York, 1967.
- [31] D. Leonard, P. Gower, and N. Taylor. Contram: Structure of the model. Transport and Road Research Laboratory (TRRL) Research Report 178, Department of Transport, Crowthorne, 1989.
- [32] R. J. LeVeque. Numerical methods for conservation laws. Lectures in mathematics ETH Zürich. Birkhäuser Verlag, Basel ; Boston, 2nd ed edition, 1992.
- [33] M. Lighthill and G. Whitham. On kinematic waves II. A theory of traffic flow on long crowded roads. Proc. R. Soc. Lond., 1955.
- [34] K. Malena, C. Link, L. Bußemas, S. Gausemeier, and A. Trächtler. Traffic estimation and mpc-based traffic light system control in realistic real-time traffic environments. In C. Klein, M. Jarke, M. Helfert, K. Berns, and O. Gusikhin, editors, Smart Cities, Green Technologies, and Intelligent Transport Systems, pages 232–254, Cham, 2022. Springer International Publishing.
- [35] R. Manzo, B. Piccoli, and L. Rarita. Optimal distribution of traffic flows in emergency cases. European Journal of Applied Mathematics, 23(4):515–535, 2012.
- [36] T. Materne, A. Klar, M. Günther, and R. Wegener. An explicit kinetic model for traffic flow. In A. M. Anile, V. Capasso, and A. Greco, editors, Progress in Industrial Mathematics at ECMI 2000, pages 597–601, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.
- [37] K. Nagel and M. Schreckenberg. A cellular automaton model for freeway traffic. Journal de Physique I, 2(12):2221–2229, 1992.
- [38] S. L. Paveri-Fontana. On boltzmann-like treatments for traffic flow: A critical review of the basic model and an alternative proposal for dilute traffic analysis. Transportation Research, 9:225–235, 1975.
- [39] H. J. Payne. Model of freeway traffic and control. Mathematical Model of Public System, pages 51–61, 1971.
- [40] I. Prigogine, P. Resibois, R. Herman, and R. Anderson. On a generalized boltzmann-like approach for traffic flow. Bulletin de la Classe des sciences. Académie royale de Belgique, 48(9):805–814, 1962.
- [41] J. Reilly, W. Krichene, M. L. D. Monache, S. Samaranyake, P. Goatin, and A. Bayen. Adjoint-based optimization on a network of discretized scalar conservation law PDEs with applications to coordinated ramp metering. Journal of Optimization Theory and Applications, 167:733–760, 2015.
- [42] P. I. Richards. Shock waves on the highway. Operations Research, 4(1), 1956.
- [43] M. Treiber, A. Hennecke, and D. Helbing. Congested traffic states in empirical observations and microscopic simulations. Physical Review E, 62(2):1805, 2000.
- [44] M. Treiber and A. Kesting. Traffic Flow Dynamics. Springer Berlin Heidelberg, 2013.
- [45] R. Wegener and A. Klar. A kinetic model for vehicular traffic derived from a stochastic microscopic model. Transport Theory and Statistical Physics, 25(7):785–798, 1996.

- [46] G. B. Whitham. Linear and nonlinear waves. John Wiley & Sons, 2011.
- [47] R. Wiedemann. Simulation des strassenverkehrsflusses. Institute for Traffic Engineering, University of Karlsruhe, Tech. Rep, 1974.
- [48] Y. Wu, L. Liu, J. Bae, K.-H. Chow, A. Iyengar, C. Pu, W. Wei, L. Yu, and Q. Zhang. Demystifying Learning Rate Policies for High Accuracy Training of Deep Neural Networks. In 2019 IEEE International conference on big data (Big Data), pages 1971–1980. IEEE, 2019.
- [49] Y. G. YuFeng Shi. A maximum-principle-satisfying finite volume compact-weno scheme for traffic flow model on networks. Applied Numerical Mathematics, 108:21–36, 2016.
- [50] H. Zhang. A non-equilibrium traffic model devoid of gas-like behavior. Transportation Research Part B: Methodological, 36(3):275–290, March 2002.