Integrated optimal control for multi-lane motorway networks*

Vasileios Markantonakis, Ioannis Papamichail

Abstract— **This paper presents a Quadratic Programming (QP) problem formulation, employing a modified multi-lane version of the well-known macroscopic Cell Transmission Model (CTM), to determine integrated optimal control actions for motorway networks. These include mainstream traffic flow control (MTFC), lane change control (LCC), ramp metering control (RM) and dynamic traffic assignment (DTA) actions to be applied by Connected and Automated Vehicles (CAVs). An algorithm based on the barrier method provides a fast solution of the convex QP problem. A case study, for a hypothetical motorway network with multiple destinations and routes, demonstrates the efficiency of the open-loop solutions.**

I. Introduction

Traffic flow modelling and control has been of prime interest since the mid-20th century. The first macroscopic traffic flow model was composed of a single conservation equation for the vehicles followed by a density-flow relationship, referred as the Fundamental Diagram (FD) [1]. A discrete approximation of the model in time and space, known as the CTM, was then presented for a single-origin destination highway stretch [2]. CTM was extended to simulate the propagation of traffic over complex networks of links and nodes using macroscopic variables (e.g., density, flow) [3]. An alternative modelling approach to calculate the evolution of density and mean speed on motorway networks, was proposed in [4], based on a second-order traffic flow model. The modelling tool METANET, in which the model is involved, allows for simulation of free flow, dense or congested conditions, capacity reducing events with prescribed characteristics and arbitrary network topology (bifurcations, on-ramps, off-ramps) [5]. The aforementioned macroscopic traffic flow models as well as many others found in the literature consider all traffic variables aggregated across all lanes and have been utilized by many researchers for developing model-based controllers. These are controllers where the known dynamics of the system and its response to actuation are considered in determining control actions [6]. Representative work using macroscopic traffic flow models, can be found in [7]-[12].

In contrast to the vast literature, where the main research interest lies in macroscopic traffic flow models aggregated over all lanes, there is a comparatively small number of works in which the proposed traffic flow model is defined on multiple lanes, employing traffic control measures that make use of CAVs. The era of CAVs has arrived, changing the role played by the traffic management centers (TMC). Vehicles enhanced with such technology may act as actuators, useful enough for the TMC to control the road traffic using innovative traffic control actions such as individual variable speed limit control, lane change control, dynamic traffic assignment control or ramp metering control. Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication protocols, ensure fast and accurate information services in a Vehicle-to-Everything (V2I) environment. In this section, such novelties are collected and presented from the work proposed so far from a small group of research institutions, that utilize connected and automated vehicle technology in real field and simulation applications.

A Model Predictive Control (MPC) architecture for platoons was proposed in [9], in which the authors regulate the lane change activity, speeds and on-ramp flows in a simple simulation example of a two-lane highway stretch. Specifically, the MCP approach leads to a non-convex optimization problem that requires proper optimization techniques in order to be solved fast in real-time, minimizing the TTS (Total Time Spent). In a similar manner, [14] studied the potential benefits of using V2I and I2V communications, to enable variable speed limit control per lane. Specifically, a cooperative variable speed limit system is proposed and evaluated with microscopic simulations. An optimal feedback control strategy, formulated as a Linear Quadratic Regulator (LQR), was proposed in [15] aiming at regulating the lane assignment of CAVs upstream of a bottleneck location so as to maximize the bottleneck throughput, targeting critical densities as set points. In [16], utilizing the same strategy, a new method was proposed that does not always aim at tracking the critical density, but through opportunely defined functions, it allows distribution of the total density at a bottleneck area over the lanes according to a given policy.

The combined deployment and evaluation of previously proposed control strategies, namely LCC [16] and MTFC via Variable Speed Limits (VSL) [17], using a microscopic simulation model for a lane-drop infrastructure was presented in [13]. The MTFC strategy is proven successful in avoiding the capacity drop, even at low penetration rates whereas the LCC strategy is able to achieve an appropriate lane assignment of vehicles upstream of the bottleneck and, as a result, to increase capacity. For low penetration rates of connected vehicles, the integrated use of the two strategies is demonstrated to be highly beneficial. The reported results were obtained assuming full compliance and no communication delays. A novel methodology for the integrated use of lane changing and ramp metering feedback control in the presence of connected vehicles, is proposed in [18] and is extended in [19] to include an extremum seeking algorithm to compute the optimal set-points of the feedback strategy. The optimal feedback control strategy is formulated as a Linear Quadratic Integral (LQI) regulator, based on a simple linear time-invariant traffic flow model. The

*This work was supported in part by the European Commission in the framework of the project FRONTIER under Grant Agreement No: 955317.

V. Markantonakis and I. Papamichail are with the Dynamic Systems and Simulation Laboratory, Technical University of Crete, 73100 Chania, Greece (phone: +30 2821037421; email: vmarkantonakis@dssl.tuc.gr; ipapa@dssl.tuc.gr).

methodology is evaluated via simulation experiments on a first-order traffic flow model and is compared to the well-known feedback controller ALINEA [20] in order to further investigate the potential benefits of ramp metering. A series of simulation experiments demonstrate the effectiveness of the recently developed methodology.

This work assumes the presence of CAVs for the purpose of supporting the formulation of an integrated optimal control problem for motorway networks under a mixture of traditional and novel traffic control measures. Various control variables are used to incorporate different options for control, namely, mainstream traffic flow control, lane change control, ramp metering control and dynamic traffic assignment control. The optimal control problem formulation is based on a linear dynamic multi-lane macroscopic traffic flow model for motorway networks [21], extended to include system-optimal dynamic traffic assignment policies. The model has been derived with a view to adopt different network structures that include lane drop areas, merge areas as well as multiple origins and destinations for the DTA policy. Following that approach, this paper explores the efficiency of the open-loop control strategy for a challenging and sizable hypothetical motorway network with multiple destinations and routes, on a realistic traffic demand scenario. It is in our future plans to cast the optimal control solutions, derived from the network level QP optimization problem, in an MPC mode for testing via application of control actions on CAVs, in a microsimulation environment.

II. Traffic Control Measures

The traffic control measures that are considered in the proposed framework are briefly explained as follows:

- Mainstream Traffic Flow Control (MTFC): The basic idea of MTFC is to enable the mainstream traffic flow that is approaching areas with a particular infrastructure, e.g. lane-drops or bottlenecks, to take values ordered by an appropriate control strategy and establish optimal traffic conditions for any appearing demand [22].
- Lane Changing Control (LCC): This control measure aims to achieve a desired distribution of vehicles among the lanes, so as to exploit the capacity of each and every lane, thus increase the overall (cross-lane) capacity and improve traffic flow efficiency. The optimal lateral flows are delivered by the control strategy for each lane.
- Ramp Metering (RM): These control actions consist in regulating the inflow from the on-ramps to the motorway mainstream [23]. Since they are applied directly at the on-ramps using ordinary traffic signals, they do not necessarily require any particular in-vehicle equipment. The optimal ramp flows are delivered by the control strategy for a segment-lane where the on-ramp is located.
- Dynamic Traffic Assignment (DTA): This control measure aims to distribute the traffic demand with the same origin-destination (O-D) among alternative routes of the network, so that some optimality principles are satisfied [24]. In our work, we focus on system-optimal DTA, where the sum of the travel times of all O-D pairs should be the lowest possible. The optimal lateral flows for each segment-lane and destination, are the ones also used for the LCC control actions.

III. CTM-Based Optimal Control Modelling

A. CTM equations

Based on a linear multi-lane traffic flow model proposed in [21] and further utilized within an optimal control problem [25], we consider a multi-lane motorway network, which is divided in n segments while each segment comprises a number of lanes. We use the index i for the segment, index j for the lane, while l_i and L_i are the number of lanes and the length of segment i, respectively. In addition, we denote as $D_{i,j}$ the set of all reachable destinations from segment-lane (i, j) and d a reachable destination from segment-lane (i, j). For each segment, we assume internally homogeneous characteristics such as the number of lanes. Our model is formulated in discrete time, considering the discrete time step T for a simulation horizon K indexed by $k = 1, 2, ..., K$ where the simulation time is $t = KT$. In addition, we denote as T^Q the MTFC control time step, as T^D the DTA and LCC control time step and as T^R the RM control time step. The corresponding discrete time indices for each control action are defined by:

$$k^Q = \left\lfloor \frac{kT}{T^Q} \right\rfloor,\ k^D = \left\lfloor \frac{kT}{T^D} \right\rfloor,\ k^R = \left\lfloor \frac{kT}{T^R} \right\rfloor$$

with $k = 0, 1, 2, ..., K$ and $\lfloor . \rfloor$ denoting the integer part.

Consider a network discretized in space by defining segment-lane entities. These entities are characterized by non-negative variables and parameters. The following notation is presented for the corresponding variables, used for the optimal control problem formulation:

- $\rho^d_{i,j}(k)$ is the partial density [veh/km]: the number of vehicles in segment-lane (i, j), at time step k, for reachable destination d divided by the segment length L_i, $d \in D_{i,j}$.
- $q^{in}_{i,j,d}(k^Q)$ is the partial inflow [veh/h]: longitudinal and diagonal traffic volume entering segment-lane (i, j) for reachable destination d during the time interval $\left(k^Q, k^Q+1\right]$, $d \in D_{i,j}$.
- $q^{out}_{i,j,d}(k^Q)$ is the partial outflow [veh/h]: longitudinal and diagonal traffic volume exiting segment-lane (i, j) for reachable destination d during the time interval $\left(k^Q, k^Q+1\right]$, $d \in D_{i,j}$.
- $f^d_{i,j,\bar{j}}(k^D)$ is the lateral flow [veh/h]: lateral traffic volume for destination d moving from lane j to lane $\bar{j} = j \pm 1$, during time interval $\left(k^D, k^D+1\right]$, $d \in D_{i,j}$.
- $f^d_{i,\bar{j},j}(k^D)$ is the lateral flow [veh/h]: lateral traffic volume for destination d entering lane j from lane $\bar{j} = j \pm 1$, during the time interval $\left(k^D, k^D+1\right]$, $d \in D_{i,j}$.
- $r^d_{i,j}(k^R)$ is the on-ramp flow [veh/h]: traffic volume entering segment-lane (i, j) from the on-ramp located at the corresponding segment-lane for destination d during the time interval $\left(k^R, k^R+1\right]$, $d \in D_{i,j}$.

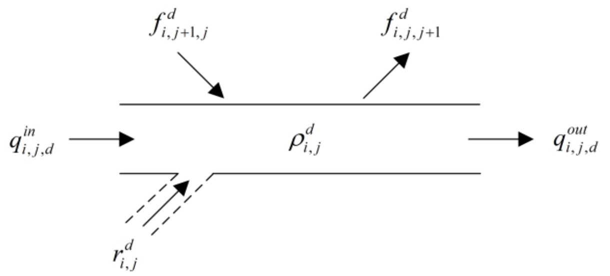

Fig. 1. Segment-lane entities used in the conservation equation

Based on the definitions given above, the conservation equation for the partial density for each segment-lane (i, j) and reachable destination d reads (see Fig. 1):

$$\rho_{i,j}^{d}(k+1)=\rho_{i,j}^{d}(k)+$$
$$\frac{T}{L_i}\left(q_{i,j,d}^{in}(k^Q)-q_{i,j,d}^{out}(k^Q)+f_{i,\bar{j},j}^{d}(k^D)-f_{i,j,\bar{j}}^{d}(k^D)+r_{i,j}^{d}(k^R)\right),$$
$$i=1,2,...,n,\ j=1,2,...,l_i,\ \bar{j}=j\pm 1, d\in D_{i,j} \quad (1)$$

In order to guarantee numerical stability, since the CTM model (and consequently the extension used in our formulation) is a discrete approximation to the LWR model [1], it is supposed that each control action is updated according to a specific control time step which may be specified according to human-factors and other operational requirements. The control time-steps T^Q, T^D, T^R are assumed to be integer multiples of the traffic flow model time step T, where T must be in agreement with the Courant-Friedrichs-Lewy (CFL) condition. Hence, the following inequality must hold:

$$T\le \min_{i,j}\frac{L_i}{v_{i,j}^{free}} \quad (2)$$

where $v_{i,j}^{free}$ is the maximum speed allowed in the corresponding segment-lane (i, j).

The partial density of segment-lane (i, j) is updated based on the conservation of inflows and outflows. In our model, we denote as $q_{i,j,d}^{in}(k)$ the partial inflow and as $q_{i,j,d}^{out}(k)$ the partial outflow of segment-lane (i, j). In normal conditions, where there is no change in the geometry of the network, the partial inflow $q_{i,j,d}^{in}(k)$ at time step k consists of the longitudinal partial flow sent from the upstream segment-lane $(\bar{i}, \bar{j})$ to the downstream segment-lane (i, j). In situations where there is a change in the geometry of the network (e.g., segments that feature a lane-drop, a lane addition or a bifurcation), strong lateral flows may appear close to the end of the segment. As a result, diagonal left/right lateral flows $q_{\bar{i},\bar{j}-1,d}^{left}(k)$, $q_{\bar{i},\bar{j}+1,d}^{right}(k)$ act as additional longitudinal flow to the downstream segment-lane (i, j) as well. In order to have a generic modelling development that takes into account any possible network layout, appropriate flags are used in the input files when and to the extend needed.

Based on the above, we may yield the following for the total partial inflow $q_{i,j,d}^{in}(k)$ at time-step k (see Fig. 2):

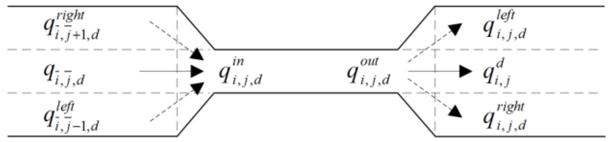

Fig. 2. Inflows and outflows of a segment-lane

$$q_{i,j,d}^{in}(k)=\begin{cases} q_{\bar{i},\bar{j},d}(k), \text{ if there is only longitudinal movement from the upstream segment-lane} \\ q_{\bar{i},\bar{j},d}(k)+q_{\bar{i},\bar{j}-1,d}^{left}(k), \text{ if there is longitudinal and left diagonal movement} \\ q_{\bar{i},\bar{j},d}(k)+q_{\bar{i},\bar{j}+1,d}^{right}(k), \text{ if there is longitudinal and right diagonal movement} \\ q_{\bar{i},\bar{j},d}(k)+q_{\bar{i},\bar{j}-1,d}^{left}(k)+q_{\bar{i},\bar{j}+1,d}^{right}(k), \text{ if there is longitudinal and left/right diagonal movement} \end{cases} \quad (3)$$

Following a similar manner, the total partial outflow $q_{i,j,d}^{out}(k)$ (see also Fig. 2) at time-step k yields as follows:

$$q_{i,j,d}^{out}(k)=\begin{cases} q_{i,j,d}(k), \text{ if there is only longitudinal movement to the downstream segment-lane} \\ q_{i,j,d}^{left}(k), \text{ if there is only left movement to the downstream segment-lane} \\ q_{i,j,d}^{right}(k), \text{ if there is only right movement to the downstream segment-lane} \\ q_{i,j,d}(k)+q_{i,j,d}^{left}(k), \text{ if there is longitudinal and left movement to the downstream segment-lane} \\ q_{i,j,d}(k)+q_{i,j,d}^{right}(k), \text{ if there is longitudinal and right movement to the downstream segment-lane} \end{cases} \quad (4)$$

Each segment-lane (i, j) is characterized by the total density $\rho_{i,j}$, the total inflow $q_{i,j}^{in}$, the total outflow $q_{i,j}^{out}$ and the total ramp-flow $r_{i,j}$ which are given by the sum of all the corresponding partial entities. The following equations express the aforementioned:

$$q_{i,j}^{out}(k)=\sum_{d\in D_{i,j}} q_{i,j,d}^{out}(k) \quad (5)$$

$$q_{i,j}^{in}(k)=\sum_{d\in D_{i,j}} q_{i,j,d}^{in}(k) \quad (6)$$

$$\rho_{i,j}(k)=\sum_{d\in D_{i,j}} \rho_{i,j}^{d}(k) \quad (7)$$

$$r_{i,j}(k)=\sum_{d\in D_{i,j}} r_{i,j}^{d}(k) \quad (8)$$

for $i=1,2,...,n,\ j=1,2,...,l_i,\ d\in D_{i,j}$.

If an on-ramp exists at segment-lane (i, j), it receives an uncontrollable (external) demand $D_{i,j}^{d}$ for each destination d that feeds the motorway network with the corresponding ramp flow $r_{i,j}^{d}$. The respective ramp outflows are controllable via corresponding ramp metering actions; hence, this may lead to the creation of ramp queues $w_{i,j}^{d}$ for each destination d. By construction of the traffic flow model, the motorway segments and the on-ramps have finite storage capacities. These values are imposed as hard constraints to the optimal control problem formulation and excessive external demand scenarios may lead to an infeasible optimization problem. In that case, no control action can accommodate the external demands without violaeting the storage capacity constraints. To address the need of avoiding such situations where the admissible control region is void, and enable the computation of optimal control for any arbitrary scenario, the following variables are introduced to the model:

- $W_{i,j}^{d}$ is the virtual extra queue [veh]: the number of vehicles waiting virtually at the queue for each destination d.
- $m_{i,j}^{d}$ is the internal demand [veh/h]: the demand flow that is capable to enter the real ramp queue $w_{i,j}^{d}$ for each destination d.

The dynamics at the on-ramps for each destination d are thus stated as follows:

$$w_{i,j}^{d}(k+1) = w_{i,j}^{d}(k) + T\left[m_{i,j}^{d}(k) - r_{i,j}^{d}\left(k^{R}\right)\right] \tag{9}$$

$$W_{i,j}^{d}(k+1) = W_{i,j}^{d}(k) + T\left[D_{i,j}^{d}(k) - m_{i,j}^{d}(k)\right] \tag{10}$$

and the equations for the total real ramp queues and total virtual extra queues read as follows:

$$w_{i,j}(k) = \sum_{d\in D_{i,j}} w_{i,j}^{d}(k) \tag{11}$$

$$W_{i,j}(k) = \sum_{d\in D_{i,j}} W_{i,j}^{d}(k) \tag{12}$$

for $i = 1,2,...,n,\ j = 1,2,...,l_i,\ d \in D_{i,j}$.

The external demand $D_{i,j}^{d}$ for each destination d is feeding the extra-queue $W_{i,j}^{d}$, while the internal demand $m_{i,j}^{d}$ connects the extra-queue with the real ramp queue $w_{i,j}^{d}$. A strong penalty factor for the total extra queues $W_{i,j}$, is used in the objective function, to keep the extra queues equal to zero. In special cases, where there is an excessive demand scenario, the problem will remain within its feasible region by charging the extra queues $W_{i,j}$. The per-destination lane inflows entering the first segment of the network are formally treated as uncontrollable on-ramps. In simple terms, this derives that for each lane $j = 1,2,...,l_1$ the total queues $w_{1,j}$ are set to zero and the total extra queues $W_{1,j}$ are also weighted in the objective function, to avoid infeasible solutions.

In the original CTM formulation, traffic flow is computed as the minimum between an upstream demand function and a downstream supply function (see Fig. 3). In our approach, we adopt the same logic for computing the longitudinal flows $q_{i,j}(k)$ enhanced with additional terms in the min-operator when required from the geometry of the network. Thus, the longitudinal flow $q_{i,j}(k)$ at time-step k is derived as follows:

$$q_{i,j}(k) = \min\Bigg[Q_{i,j}^{D}(k) - \sum_{d\in D_{i,j}} q_{i,j,d}^{left}\left(k^{Q}\right) - \sum_{d\in D_{i,j}} q_{i,j,d}^{right}\left(k^{Q}\right), \\ Q_{\tilde{i},\tilde{j}}^{S}(k) - \lambda_r r_{\tilde{i},\tilde{j}}\left(k^{R}\right) - \sum_{d\in D_{i,j}} q_{i,j-1,d}^{left}\left(k^{Q}\right) - \sum_{d\in D_{i,j}} q_{i,j+1,j}^{right}\left(k^{Q}\right)\Bigg] \tag{13}$$

for $i = 1,2,...,n,\ j = 1,2,...,l_i$ and $\tilde{i} = 2,3,...,n,\ \tilde{j} = 1,2,...,l_{\tilde{i}}$,

where (i, j) is the current segment-lane and $(\tilde{i},\tilde{j})$ is the downstream segment-lane, while the demand and supply functions are given by the following equations:

$$Q_{i,j}^{D}(k) = \min\left[v_{i,j}^{free}\rho_{i,j}(k),\ q_{i,j}^{cap} + \lambda_d q_{i,j}^{cap}\frac{\rho_{i,j}(k) - \rho_{i,j}^{cr}}{\rho_{i,j}^{cr} - \rho_{i,j}^{jam}}\right] \tag{14}$$

$$Q_{i,j}^{S}(k) = \min\left[q_{i,j}^{cap},\ w_{i,j}^{S}\left(\rho_{i,j}^{jam} - \rho_{i,j}(k)\right)\right] \tag{15}$$

where $\rho_{i,j}^{jam}$ and $w_{\tilde{i},\tilde{j}}^{S}$ are the total jam density and the back-wave speed, respectively.

It is well known that the original CTM does not reproduce capacity drop phenomena. To address the need and be able to investigate the impact of possible capacity drop, two additional terms are added in the model (appearing in (13), (14) as λ_r, λ_d). If these two parameters are set $\lambda_r = 1$ and $\lambda_d = 0$, by default no capacity drop is introduced as typical for CTM; if set between 0 and 1, a corresponding level of capacity drop is produced by the model [26].

For all first-order models described by one dynamic equation, speed (which is not a decision variable of the QP problem) is calculated only a posteriori. Hence, using both equations (5) and (7), speed for each segment-lane (i, j) is calculated from the non-linear relation of flow and density as:

$$v_{i,j}(k) = \frac{q_{i,j}^{out}(k)}{\rho_{i,j}(k)},\ i = 1,2,...,n,\ j = 1,2,...,l_i \tag{16}$$

B. *Linear inequality constraints*

The computation of the lateral flows for each destination d, used to determine the LCC and DTA control actions, is

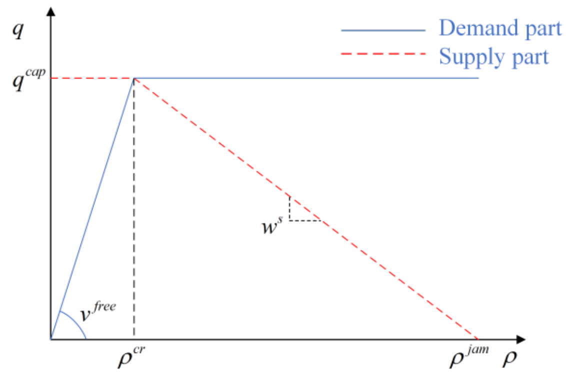

Fig. 3. Triangular fundamental diagram

fully delegated to the optimizer; only upper bounds are specified to the non-negative lateral flows as well to the on-ramp flows and real queues as follows:

$$\left[f_{i,j,j-1}^{d}\left(k^{D}\right)+f_{i,j,j+1}^{d}\left(k^{D}\right)\right] \le \frac{L_i}{T}\rho_{i,j}^{d}\left(k\right) \tag{17}$$

$$\sum_{d\in D_{i,j}} f_{i,j-1,j}^{d}\left(k^{D}\right)+f_{i,j+1,j}^{d}\left(k^{D}\right) \le \frac{L_i}{T}\left[\rho_{i,j}^{jam}-\rho_{i,j}\left(k\right)\right] \tag{18}$$

$$\sum_{d\in D_{i,j}} f_{i,j,j-1}^{d}\left(k^{D}\right) \le f_{i,j,j-1}^{\max} \tag{19}$$

$$\sum_{d\in D_{i,j}} f_{i,j,j+1}^{d}\left(k^{D}\right) \le f_{i,j,j+1}^{\max} \tag{20}$$

$$r_{i,j}\left(k^{R}\right) \le r_{i,j}^{\max} \tag{21}$$

$$w_{i,j}\left(k\right) \le w_{i,j}^{\max} \tag{22}$$

for $i=1,2,...,n,\ j=1,2,...,l_i,\ d\in D_{i,j}$.

Equation (17) represents the upper-bound as a restriction for the lateral movements determined by the number of vehicles in the current segment-lane; (18) is the upper-bound for the lateral flows entering segment-lane (i,j) from the adjacent lanes, considering the available space in the current segment-lane; whereas, equations (19), (20) serve as hard constraints on both directions, to limit any unrealistic lateral movements. In addition, equations (21), (22) represent the upper bounds for the total on-ramp flow and maximum queue allowed at the on-ramp, respectively.

The equations (1), (5), (6), (7), (8) are linear but due to the presence of the min-operator the CTM flow equations presented previously, (13), (14), (15) are non-linear. In this regard, such nonlinearities may be transformed to linear inequalities. This is achieved by requesting the left-hand side of the equation, where the min-operator appears, to be smaller than or equal to each of the terms included in the min-operator. Hence, the following transformations are obtained for (13), (14), (15):

$$q_{i,j}^{out}\left(k\right) \le v_{i,j}^{free}\rho_{i,j}\left(k\right) \tag{23}$$

$$q_{i,j}^{out}\left(k\right) \le q_{i,j}^{cap}+\lambda_d q_{i,j}^{cap}\frac{\rho_{i,j}\left(k\right)-\rho_{i,j}^{cr}}{\rho_{i,j}^{cr}-\rho_{i,j}^{jam}} \tag{24}$$

$$q_{i,j}^{in}\left(k\right) \le q_{i,j}^{cap}-\lambda_r r_{i,j}\left(k^{R}\right) \tag{25}$$

$$q_{i,j}^{in}\left(k\right) \le w_{i,j}^{S}\left(\rho_{i,j}^{jam}-\rho_{i,j}\left(k\right)\right)-\lambda_r r_{i,j}\left(k^{R}\right) \tag{26}$$

for $i=1,2,...,n,\ j=1,2,...,l_i$. Inequalities (23), (24) represent the demand part of the FD; while (25), (26) represent the supply part of the FD.

Each O-D pair in the network is described by use of the partial flow $q_{i,j,d}^{out}$ and the partial density $\rho_{i,j}^{d}$. For the optimizer both $q_{i,j,d}^{out}$ and $\rho_{i,j}^{d}$ are control variables; meaning that multiple pairs of $q_{i,j,d}^{out}$ and $\rho_{i,j}^{d}$ may satisfy the constraints of the optimal control problem. In order to limit any unrealistic pairs, the following inequality must hold:

$$q_{i,j,d}^{out} \le v_{i,j}^{free}\rho_{i,j}^{d} \tag{27}$$

for $i=1,2,...,n,\ j=1,2,...,l_i,\ d\in D_{i,j}$.

IV. QP Problem Formulation

The objective function (Z) to be minimized is defined in (28).

$$\begin{aligned}
Z ={}& T\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{l_i}\left[L_i\rho_{i,j}\left(k\right)+w_{i,j}\left(k\right)\right]+\beta_1\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{I_i}\left[W_{i,j}\left(k\right)\right]\\
&+\sum_{k^{D}=1}^{K^{D}}\sum_{i=1}^{n}\sum_{j=1}^{l_i}\sum_{d\in D_{i,j}}\left[a_{i,j,j-1}^{d}f_{i,j,j-1}^{d}\left(k^{D}\right)+a_{i,j,j+1}^{d}f_{i,j,j+1}^{d}\left(k^{D}\right)\right]\\
&+\beta_2\sum_{k^{D}=2}^{K^{D}}\sum_{i=1}^{n}\left\{\sum_{j=2}^{l_i}\sum_{d\in D_{i,j}}\left[f_{i,j,j-1}^{d}\left(k^{D}\right)-f_{i,j,j-1}^{d}\left(k^{D}-1\right)\right]^2\right.\\
&\left.+\sum_{j=1}^{l_i-1}\sum_{d\in D_{i,j}}\left[f_{i,j,j+1}^{d}\left(k^{D}\right)-f_{i,j,j+1}^{d}\left(k^{D}-1\right)\right]^2\right\}\\
&+\beta_3\sum_{k^{D}=2}^{K^{D}}\sum_{i=1}^{n}\left\{\sum_{j=1}^{l_i}\sum_{d\in D_{i,j}}\left[r_{i,j}^{d}\left(k^{R}\right)-r_{i,j}^{d}\left(k^{R}-1\right)\right]^2\right\}\\
&+\beta_4\sum_{k^{D}=2}^{K^{D}}\sum_{i=1}^{n}\left\{\sum_{j=1}^{l_i}\sum_{d\in D_{i,j}}\left[q_{i,j,d}^{out}\left(k^{Q}\right)-q_{i,j,d}^{out}\left(k^{Q}-1\right)\right]^2\right\}\\
&+\beta_5\sum_{k=1}^{K}\sum_{i=1}^{n}\left\{\sum_{j=1}^{l_i}\sum_{d\in D_{i,j}}\frac{\left[q_{i,j,d}^{out}\left(k^{Q}\right)-q_{i,j}^{out}\left(k^{Q}\right)-v_{i,j}^{*}\left(\rho_{i,j}^{d}\left(k\right)-\rho_{i,j}\left(k\right)\right)\right]^2}{\left(\rho_{i,j}^{*}\right)^2}\right\}\\
&+\beta_6\sum_{k=2}^{K-1}\sum_{i=1}^{n}\left\{\sum_{j=1}^{l_i}\sum_{d\in D_{i,j}}\frac{\left[q_{i,j,d}^{out}\left(k^{Q}\right)-q_{i,j,d}^{out}\left(k^{Q}-1\right)-v_{i,j}^{*}\left(\rho_{i,j}^{d}\left(k\right)-\rho_{i,j}^{d}\left(k-1\right)\right)\right]^2}{\left(\rho_{i,j}^{*}\right)^2}\right\}\\
&+\beta_7\sum_{k=1}^{K}\sum_{i=1}^{n}\left\{\sum_{j=1}^{l_i}\sum_{d\in D_{i,j}}\frac{\left[r_{i,j}^{d}\left(k^{R}\right)-r_{i,j}\left(k^{R}\right)-\delta_{i,j}^{*}\left(w_{i,j}^{d}\left(k\right)-w_{i,j}\left(k\right)\right)\right]^2}{\left(w_{i,j}^{*}\right)^2}\right\}\\
&+\beta_8\sum_{k=1}^{K}\sum_{i=1}^{n}\left\{\sum_{j=1}^{l_i}\sum_{d\in D_{i,j}}\frac{\left[m_{i,j}^{d}\left(k\right)-m_{i,j}\left(k\right)-\varepsilon_{i,j}^{*}\left(W_{i,j}^{d}\left(k\right)-W_{i,j}\left(k\right)\right)\right]^2}{\left(W_{i,j}^{*}\right)^2}\right\}
\end{aligned} \tag{28}$$

It is composed by the weighted sum of ten different terms, the first three linear and the following seven quadratic.

- The first linear term represents the Total Time Spent (TTS) [veh·h] it considers the time spent by vehicles to travel along the network and the waiting time at the queues. This is the most important term that is used to evaluate the goodness of the solution in terms of traffic flow efficiency.
- The second linear term weighted by β_1 is a penalty term for the extra queues. The value of the weight should be big enough to make sure that all variables $W_{i,j}$ are set equal to zero when $w_{i,j}$ remains within bounds.
- The third linear term weighted by coefficients $\alpha_{i,j,j-1}^{d},\alpha_{i,j,j+1}^{d}$ aims at penalizing the partial lateral flows; it has the purpose of reducing the unnecessary lateral movements. However, the use of lower values for the coefficients at specific locations across the network

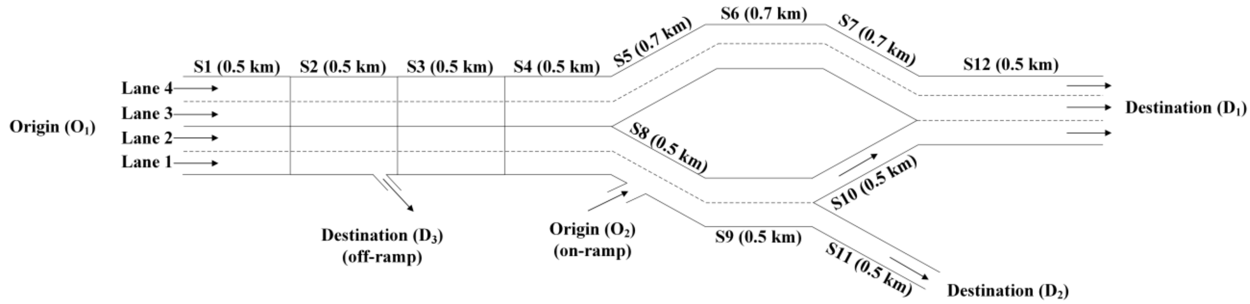

Fig. 4. Motorway network used for the case study

(e.g., upstream of a bottleneck location) may facilitate reasonable and beneficial lateral movements.

- The remaining quadratic terms are penalty terms aiming at penalizing the variation of control variables from a time step to the next one or/and within the current time step. These terms are introduced to reduce, or even suppress, time fluctuations of the control variables, that have a minor contribution to the resulting traffic flow efficiency. The first penalty term (weighted by β_2) is related to time-variations of the partial lateral flows; the second and third term (weighted by β_3, β_4) have the purpose of penalizing time-variations of the on-ramp flows and partial flows accordingly. The fourth term (weighted by β_5) penalizes variations of the partial flow in relation with the partial density. In fact, this term allows for having homogeneous flows that follow the relation of $q^{out}_{i,j,d} / q^{out}_{i,j} = \rho^{d}_{i,j} / \rho_{i,j}$. The fifth term (weighted by β_6) has the purpose of penalizing time variations of speed values per destination. The last two terms (weighted by β_7, β_8) penalize variations of the on-ramp flows in relation to the ramp-queues and time variations of the internal demand $m_{i,j}$ to the virtual extra queues.

The solution to our discrete time formulated optimal control problem is given through a convex QP problem in matrix form. A general quadratic programming formulation contains a quadratic objective function and linear equality and inequality constraints:

$$\min Z = \mathbf{c}^{\mathrm{T}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{H}\mathbf{x} \tag{29}$$

subject to

$$\mathbf{A}_i\mathbf{x} \le \mathbf{b}_i \tag{30}$$

$$\mathbf{A}_e\mathbf{x} = \mathbf{b}_e \tag{31}$$

$$\mathbf{0} \le \mathbf{x} \le \mathbf{b}_{ub} \tag{32}$$

where $\mathbf{x}$ is the decision vector including all the states and control variables. In (29) the vector $\mathbf{c}^{\mathrm{T}}$ and the matrix $\mathbf{H}$ contain all the coefficients for the linear terms and the specifics of the quadratic terms of (28) , respectively.

V. CASE STUDY

A description of the network used for the simulations, the configuration of the macroscopic traffic flow model, the configuration of the QP model as well as the results of the corresponding QP problem, are presented in this section. Useful conclusions are derived regarding the adequacy of the four available traffic control measures in an open loop control scenario. The goal is to perform a coordinated exploitation of the available measures so as to enable the computation of optimal control for any arbitrary demand scenario.

A. *Network description*

The considered motorway network is illustrated in Fig. 4. It is composed by a part of four lanes (segments 1-4), a part of three lanes (segment 12), parts with two lanes (segments 5-9), and parts with a single lane for segments 10, 11 and it presents one on-ramp at segment 8. It also features two origins denoted by O_1 and O_2 as well as three reachable destinations denoted by D_1, D_2 and D_3. Starting from origin O_1, two alternative routes of 4.0 km and 4.6 km length exist to reach D_1, while a unique route of 3.5 km length exists for destination D_2 and of 1.0 km for destination D_3 accordingly. For those entering from origin O_2, both destinations D_1 and D_2 can be reached via unique routes of 2.0 km and 1.5 km, respectively.

B. *Configuration of the macroscopic traffic flow model*

The utilized CTM parameters are $v^{free}_{i,j} = 100$ km/h, $\rho^{cr}_{i,j} = 22$ veh/km, $\rho^{jam}_{i,j} = 120$ veh/km and $w^{s}_{ij} = 22.4$ km/h. Also, the values used to enable capacity drop are $\lambda_d = 0.3$, $\lambda_r = 0.7$. The choice of the simulation step T is a crucial aspect that must be carefully taken into account. In fact, a too long simulation step could allow vehicles to travel in more than one cell during its duration, causing numerical instability of the mathematical model; on the other hand, the size of the optimization problem is also affected by this choice. In this case, a value $T = 10$ s is set, that satisfies the CFL condition (2) for model stability. Once the simulation step is chosen, the control steps are defined as a multiple of the simulation step. In our case, the control step for all the control variables is set equal to the simulation step. Hence, the control variables are updated synchronously with the state variables and have the highest degree of freedom. Based on the aforementioned, $T^Q = T^D = T^R = 10$ s.

C. *Configuration of the QP model*

The weighting parameters used in the objective function of the QP problem for the respective terms have been tuned and are $\beta_1 = 10^4$, $\beta_2 = 10^{-5}$, $\beta_3 = 10^{-5}$, $\beta_4 = 10^{-5}$, $\beta_5 = 10^{-2}$, $\beta_6 = 10^1$, $\beta_7 = 10^{-5}$, $\beta_8 = 10^{-5}$. For the linear penalty terms related to the partial lateral flows, the weight $\alpha^{d}_{i,j,\bar{j}}$ is set either equal to zero, to encourage vehicles to change lane if necessary, or equal to 10^2, when no lane

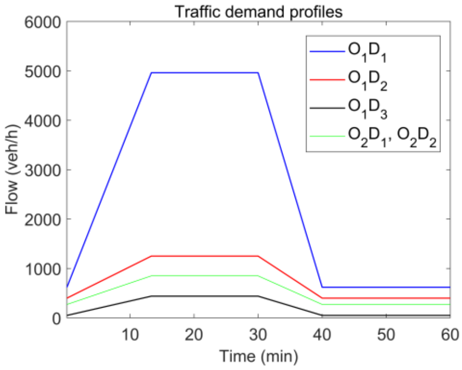

Fig. 5. Average traffic demand profiles for origin 1 destination 1 (blue line), origin 1 destination 2 (red line), origin 1 destination 3 (black line), on-ramp origin 2 destination 1, 2 (green line)

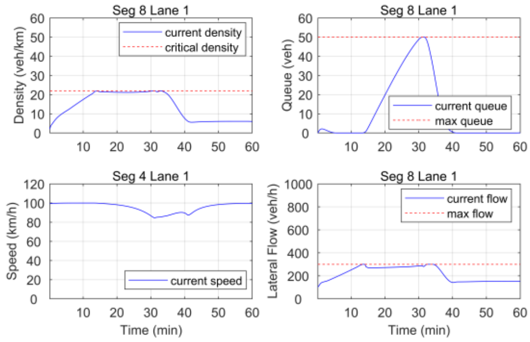

Fig. 6. Density values (in blue) of destination 1 (upper left) and destination 2 (upper right); flow values (in blue) of destination 1 (bottom left) and destination 2 (bottom right); and the corresponding critical, and capacity values (in dashed red) of segments 10 and 11

changing movement is allowed. As a result, the QP problem is a convex optimization problem and has a unique solution. In addition, upper bounds are provided for the lateral movements ($f^{\max} = 300$ veh/h), for the on-ramp queues at segment 8 ($w^{\max} = 50$ veh/h) and for the maximum flow that is allowed to enter from the on-ramp at segment 8 ($r^{\max} = 1800$ veh/h). Furthermore, the parameters for suppressing oscillations of the control variables have been chosen as $v^*_{i,j} = v^{free}_{i,j}$, $\rho^*_{i,j} = \rho^{cr}_{i,j}$, $\delta^*_{i,j} = 10^2$, $\varepsilon^*_{i,j} = 10^2$, $w^*_{i,j} = w^{\max}$, $W^*_{i,j} = W^{\max}$.

D. Open loop scenario and solution

A time-horizon equal to 60 min is considered in the scenario and trapezoidal traffic demand profiles for all O-D pairs are depicted in Fig. 5. An algorithm based on the barrier method [27] implemented in GUROBI [28] provides a fast solution of the QP problem in almost 48 sec. The Total Time Spent (TTS) as well as the Total Delay (TD) improvements compared to a no-control scenario (developed similarly to [21]), are 42% and 66%, respectively. Due to the increased demand from the on-ramp located at segment 8 lane 1, an integrated use of all the traffic control measures is performed. Specifically, strong lateral movements are performed from lane 1 to lane 2 at segment 8 for the vehicles traveling towards destination D_1, approaching the maximum value, so as to create the necessary space for those entering from the on-ramp (Fig. 6) (bottom right). In addition, effective RM actions are performed at segment 8 lane 1 throughout the simulation horizon for all vehicles entering, enabling the use of the ramp queues (Fig. 6) (upper right). While both control measures try to maintain the density at the merging area below its critical value, strong MTFC actions are performed in the upstream segments. In fact, as demonstrated in (Fig. 6) (bottom left), at segment 4 lane 1, this corresponds to VSL values ordered that are gradually decreasing close to the value of 80 km/h. Because of VSL actions, the flow entering the merging area is reduced; thus, capacity drop is avoided, obtaining the corresponding critical density value (Fig. 6) (upper left).

As mentioned earlier in section V.A., destination D_1 can be reached using three alternative routes. For those entering from origin O_1, two alternative routes are available: $R1: S1 \to S2 \to S3 \to S4 \to S5 \to S6 \to S7 \to S12$ and $R2: S1 \to S2 \to S3 \to S4 \to S8 \to S9 \to S10 \to S12$ while for those entering from origin O_2, the unique route $R3: S8 \to S9 \to S10 \to S12$ exists. For destination D_2, for those entering from origin O_1 the unique route $R4: S1 \to S2 \to S3 \to S4 \to S8 \to S9 \to S11$ exists while for those entering from origin O_2 the unique route $R5: S8 \to S9 \to S11$ exists.

It is reasonable that, when in free-flow conditions, all drivers travelling towards a destination want to use the shortest route as the best option to reduce their travel time. In our study, the optimality principle adopted for system optimal DTA implies that the travel times for each O-D pair along alternative routes are not equal, which means that some users may experience longer travel times on alternative routes compared to other users travelling for the same O-D. Starting from origin O_1, destination D_1 can be reached via the shortest route $R2$ and destination D_2 via the route $R4$. As expected, the displayed results in Fig. 7 confirm that densities are always equal or lower than the respective critical densities at both segments 10 and 11; hence, the outflows are always close to the nominal capacity ($\sim$2200 veh/h) for segment 10 and ($\sim$2000 veh/h) for segment 11. It is worth noting that a lower critical density value (lower capacity) for segment 11 is selected, under the assumption that destination D_2 with a single lane exits to an urban network. This implies that the optimizer tackles the DTA problem effectively utilizing the maximum capacity of the shortest routes of the network for all destinations, so as to reduce the total travel time of the system. For the remaining vehicles travelling towards destination D_1, the longest route $R1$ is utilized. Fig. 8 depicts that densities at segment 5 (immediate segment after the routing diversion) never reach the respective critical values; hence, the outflows are always lower than the nominal capacity per lane ($\sim$2200 veh/h).

VI. Conclusions

In this study, a first order multi-lane macroscopic traffic flow model supports the formulation of a Quadratic Programming (QP) problem to determine the coordinated use of MTFC, LCC, RM and DTA optimal control actions for motorway networks. Investigations were performed on a hypothetical motorway network featuring an on-ramp, an off-ramp and multiple destinations and routes. Considering the large size of the problem, due to the large number of state/control variables and equality/inequality constraints, a fast solution of the QP problem is observed. The results demonstrate the

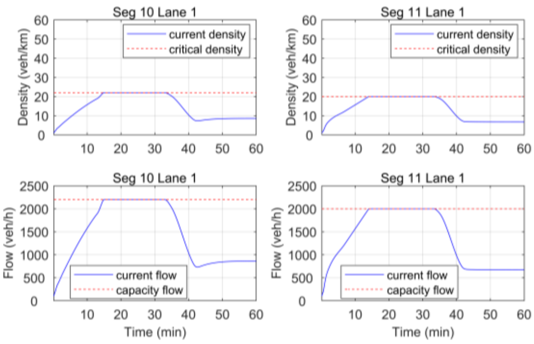

Fig. 7. Total density value (in blue) (upper left), ramp queue (in blue) (upper right), speed value (in blue) (bottom left), lateral flow of destination 1 (in blue) (bottom right); and the corresponding critical, maximum, and speed limit values (in dashed red)

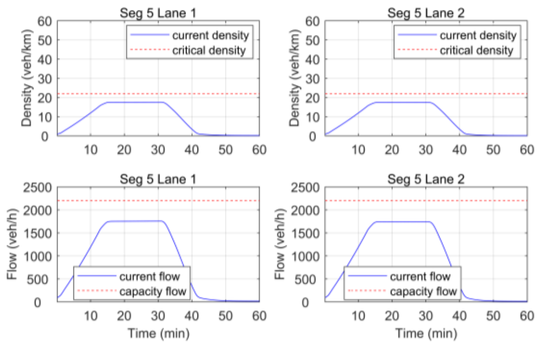

Fig. 8. Density values (in blue) of destination 1 (upper left, right); flow values (in blue) of destination 1 (bottom left, right); and the corresponding critical, and capacity values (in dashed red) of segment 5 lane 1 and lane 2

effective control actions, improving the motorway traffic flow efficiency significantly. It is in our future plans to cast the optimal control solutions, derived from the network level QP optimization problem, in an MPC framework for testing via application of control actions on CAVs, in a microsimulation environment.

References

[1] M. J. Lighthill and G. B. Whitham, "On kinematic waves II. A theory of traffic flow on long crowded roads," Proceedings of the Royal Society of London, vol. A229, pp. 317-345, 1955.

[2] C. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," Transportation Research Part B, vol. 28(4), pp. 269-287, 1994.

[3] C. Daganzo, "The cell transmission model, part II: Network traffic," Transportation Research Part B:, vol. 29(2), pp. 79-93, 1995.

[4] A. Messmer and M. Papageorgiou, "METANET: a macroscopic simulation program for motorway networks," Traffic Engineering and Control, vol. 31(8-9), pp. 466-470, 1990.

[5] A. Kotsialos, M. Papageorgiou, C. Diakaki, Y. Pavlis and F. Middelham, "Traffic flow modelling of large-scale motorway networks using the macroscopic modelling tool METANET," IEEE Trans. on Intelligent Transportation Systems, vol. 3(4), pp. 282-292, 2022.

[6] M. Berger, M. Van Den Berg, A. Hegyi, B. De Schutter and H. Hellendoorn, "Considerations for model-based traffic control," Transportation Research Part C, vol. 35, pp. 1-19, 2013.

[7] I. Papamichail, A. Kotsialos, I. Margonis and M. Papageorgiou, "Coordinated ramp metering for freeway-networks-a-model-predictive hierarchical control approach," Transportation Research Part C, vol. 18(3), pp. 311-331, 2010.

[8] R. Carlson, I. Papamichail, M. Papageorgiou and A. Messmer, "Optimal motorway traffic control involving variable speed limits and ramp metering," Transportation Science, vol. 44(2), pp. 238-253, 2010.

[9] L. D. Baskar, B. De Schutter and H. Hellendoorn, "Traffic management for automated highway systems using model-based predictive control," IEEE Trans. on Intelligent Transportation Systems, vol. 13(2), pp. 838-847, 2012.

[10] A. Muralidharan and R. Horowitz, "Computationally efficient model predictive control of freeway networks," Transportation Research Part C, vol. 58, pp. 532-553, 2015.

[11] Y. Han, Y. Yuan, A. Hegyi and S. P. Hoogendoorn, "Linear quadratic MPC for integrated route guidance and ramp metering," IEEE 18th Int. Conf. on Intelligent Transportation Systems, pp. 1150-1155, 2015.

[12] Y. Han, M. Ramezani, A. Hegyi, Y. Yuan and S. P. Hoogendoorn, "Hierarchical ramp metering in freeways: An aggregated modeling and control approach," Transportation Research Part C, vol. 110, pp. 1-19, 2020.

[13] V. Markantonakis, D. I. Skoufoulas, I. Papamichail and M. Papageorgiou, "Integrated traffic control for freeways using variable speed limits and lane change control actions," Transportation Research Record, vol. 2673(9), pp. 602-613, 2019.

[14] E. Grumert, X. Ma and A. Tapani, "Analysis of a cooperative variable speed limit system using microscopic traffic simulation," Transportation Research Part C, vol. 52, pp. 173-186, 2015.

[15] C. Roncoli, N. Bekiaris-Liberis and M. Papageorgiou, "Optimal lane-changing control at motorway bottlenecks," IEEE 19th Int. Conference on Intelligent Transportation Systems, pp. 1785-1791, 2016.

[16] C. Roncoli, N. Bekiaris-Liberis and M. Papageorgiou, "Lane-changing feedback control for efficient lane assignment at motorway bottlenecks," Transportation Research Record, vol. 2625(1), pp. 20-31, 2017.

[17] R. C. Carlson, I. Papamichail and M. Papageorgiou, "Comparison of local feedback controllers for the mainstream traffic flow on freeways using variable speed limits," Journal of Intelligent Transportation Systems, vol. 17(4), pp. 268-281, 2013.

[18] F. Tajdari, C. Roncoli, N. Bekiaris-Liberis and M. Papageorgiou, "Integrated ramp metering and lane-changing feedback control at motorway bottlenecks," 18th European Control Conference, pp. 3179-3184, IEEE, 2019.

[19] F. Tajdari, C. Roncoli and M. Papageorgiou, "Feedback-based ramp metering and lane-changing control with connected and automated vehicles," IEEE Trans. on Intelligent Transportation Systems, vol. 23(2), pp. 939-951, 2022.

[20] M. Papageorgiou, H. Hadj-Salem and J. M. Blosseville, "ALINEA: A local feedback control law for on-ramp metering," Transportation Research Record, vol. 1320(1), pp. 58-67, 1991.

[21] C. Roncoli, M. Papageorgiou and I. Papamichail, "Traffic flow optimisation in presence of vehicle automation and communication systems – Part I: A first-order multi-lane model for motorway traffic," Transportation Research Part C, vol. 57, pp. 241-259, 2015.

[22] M. Papageorgiou, E. Kosmatopoulos and I. Papamichail, "Effects of variable speed limits on motorway traffic flow," Transportation Research Record, vol. 2047(1), pp. 37-48, 2008.

[23] I. Papamichail, A. Kotsialos, I. Margonis and M. Papageorgiou, "Coordinated ramp metering for freeway networks–a model-predictive hierarchical control approach," Transportation Research Part C, vol. 18(3), pp. 311-331, 2010.

[24] Y. Wang, A. Messmer and M. Papageorgiou, "Freeway network simulation and dynamic traffic assignment with METANET tools," Transportation Research Record, vol. 1776(1), pp. 178-188, 2001.

[25] C. Roncoli, M. Papageorgiou and I. Papamichail, "Traffic flow optimisation in presence of vehicle automation and communication systems – Part II: Optimal control for multi-lane motorways," Transportation Research Part C, vol. 57, pp. 260-275, 2015.

[26] M. Kontorinaki, A. Spiliopoulou, C. Roncoli and M. Papageorgiou, "First-order traffic flow models incorporating capacity drop: Overview and real-data validation," Transportation Research Part B, vol. 106, pp. 52-75, 2017.

[27] I. Dikin, "Iterative solution of problems of linear and quadratic programming," Doklady Akademii Nauk, vol. 174, pp. 747-748, 1967.

[28] Gurobi Optimization L.L.C, "Gurobi optimizer reference manual, version 9.0," URL: https://www.gurobi.com, 2023.