Front-tracking transition system model for traffic state reconstruction, model learning, and control with application to stop-and-go wave dissipation

Mladen Čičić *, Karl Henrik Johansson

Division of Decision and Control Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Malvinas väg 10, 10044 Stockholm, Sweden

A R T I C L E   I N F O

A B S T R A C T

Connected and Autonomous Vehicles is a technology that will be disruptive for all layers of traffic control. The Lagrangian, in-the-flow nature of their operation offers untapped new potentials for sensing and actuation, but also presents new fundamental challenges. In order to use these vehicles for traffic state reconstruction and control, we need suitable traffic models, which should be computationally efficient and able to represent complex traffic phenomena. To this end, we propose the Front-tracking Transition System Model, a cell-free modelling approach that can incorporate Lagrangian measurements, and has a structure that yields itself to on-line model learning and control. The model is formulated as a transition system, and based on the front-tracking method for finding entropy solutions to the Lighthill–Whitham–Richards model. We characterize the solution of this model and show that it corresponds to the solution of the underlying PDE traffic model. Algorithms for traffic state reconstruction and model learning are proposed, exploiting the model structure. The model is then used to design a prediction-based control law for stop-and-go wave dissipation using randomly arriving Connected and Autonomous Vehicles. The proposed control framework is able to estimate the traffic state and model, adapt to changes in the traffic dynamics, and achieve a reduction in vehicles' Total Time Spent.

1. Introduction

The introduction of Connected and Autonomous Vehicles (CAVs) into highway traffic strike a delicate balance between the benefits and detriments that they bring. While long-term, we may expect significant throughput and capacity increases once CAVs reach a high market penetration rate (Shladover et al., 2012; Talebpour and Mahmassani, 2016), the low-to-mid-term effects are likely to be adverse, due to increased demand and overly conservative autonomous driving behaviour (Alonso Raposo et al., 2017). Luckily, the presence of even a small number of directly controllable CAVs on the road gives us new potentials for traffic sensing and control, with unprecedented level of detail, which we might use to improve the traffic flow and offset their negative effects.

Stop-and-go wave dissipation is one problem that yields itself to CAV-based traffic control. In this work, we define stop-and-go waves (also known as phantom jams, wide moving jams, traffic waves, etc.) as zones of congestion propagating upstream, with the traffic flow discharging from them lower than the capacity of the road (Yuan et al., 2017). While methods for dealing with this problem based on variable speed limits also exist (Hegyi et al., 2008), their implementation may suffer from the lack of

required signal equipment wherever a stop-and-go wave might arise. In contrast, CAV-based control relies on communicating with a small subset of vehicles, which may be done via the ubiquitous mobile phone network, and has been shown to work both in simulations (Simoni and Claudel, 2017; Piacentini et al., 2018; Čičić and Henrik Johansson, 2019), and experiment (Stern et al., 2018). In particular, using CAVs acting as moving bottlenecks to actuate the control actions, as the referenced works did, has some crucial advantages, i.e., it requires a relatively low number of directly controlled vehicles, and does not rely on complex behavioural models to capture the interaction between the CAVs and the rest of the traffic.

These new approaches to traffic control necessitate new approaches to traffic modelling, which accurately predict stop-and-go waves' evolution and the influence of individual CAVs on the rest of the traffic, while still being numerically tractable. In order to accurately model the discharging flow (which may be lower than the road capacity) and propagation of stop-and-go waves, we need to either use a microscopic traffic model (Laval and Leclercq, 2010), a second-order PDE model (Flynn et al., 2009), or an extended first-order model (Han et al., 2016). The first two options result in complex models that are hard to use for CAV-based traffic control design, and the third, while simpler, is ill-suited due to its cell-based nature. Namely, in order for cell-based models to represent moving features, such as stop-and-go waves and individual controlled CAVs, they need to have small cell lengths everywhere where these features are present. This greatly increases the number of states and necessitates use of lower time step, leading to numerical intractability when calculating predictions over a long horizon. Efficient Lax–Hopf based grid-free solution algorithms for the Lighthill–Whitham–Richards (LWR) model (Mazaré et al., 2011; Simoni and Claudel, 2017) are a good alternative, but while they can model the CAVs as moving bottlenecks (Monache and Goatin., 2014), they do not provide a straightforward way of implementing stop-and-go waves with discharging flow lower than the capacity of the road, or other similar phenomena. Furthermore, these approaches assume the traffic model and state to be known.

Another difficulty in applying CAV-based traffic control is the need for model calibration and traffic state reconstruction wherever the control is needed. This difficulty is even more prominent here than in case of classical traffic control methods (e.g., ramp metering and variable speed limits (Hou et al., 2008; Smaragdis et al., 2004; Hosein Ghods et al., 2009)), because the locations of the actuators and regions of interest might be unknown a-priori. Therefore adequate traffic sensor coverage might be lacking, preventing the use of some established traffic state estimation algorithms (e.g. Wang and Papageorgiou, 2005, see Seo et al., 2017 for a survey). Algorithms using Lagrangian sensing, where the measurements come from probe vehicles immersed in the traffic flow, to reconstruct the traffic density profile, have been gaining more traction in recent time (Herrera et al., 2010; Mehran et al., 2012; Seo and Kusakabe, 2015; Bekiaris-Liberis et al., 2016; Laura Delle Monache et al., 2019; Čičić et al., 2020a).

In addition to providing information about the traffic state, these local measurements may be used to learn the models governing the behaviour of the background traffic and the influence of the CAVs. Traffic model learning is most commonly done by choosing a model structure, and then calibrating the model parameters using stationary sensor measurements (Spiliopoulou et al., 2017) or probe vehicle trajectories (Seo et al., 2019). In order to be able to adapt to model variations in time, e.g., due to changing weather conditions (Billot et al., 2009), online model calibration methods have also been proposed (Antoniou et al., 2005), as well as extensions to traffic state estimation approaches with model adaptation (Wang et al., 2009).

The main contribution of this work is a cell-free event-based modelling approach that can handle stop-and-go waves, captures the influence of CAVs, can incorporate Lagrangian measurements, and has a structure that yields itself to on-line model learning, traffic state reconstruction and control. This model is the Front-tracking Transition System Model (FTSM), formulated as a transition system, which mirrors the procedure of finding the front-tracking solution to an extended LWR model with a space-dependent flux function and wave-speed bounds. In this work, we assume that the considered road segment is homogeneous, that stop-and-go waves have constant propagation speeds (as is empirically observed by Schönhof and Helbing, 2007), and that the CAVs can be modelled as moving bottlenecks. The FTSM allows regions with potentially moving boundaries to be described by different flux functions, by including their piecewise linear definitions as one of the model states. This makes it particularly convenient for use in complex traffic environments, e.g., where some individual vehicles greatly affect the overall behaviour of the traffic in their vicinity acting as moving bottlenecks or similar phenomena. By assuming constant stop-and-go wave propagation speeds, we are able to model their discharging flow as lower than the capacity of the road, by defining a new type of weak solution, the wave-speed-bounded solution, which violates the entropy condition at the stop-and-go wave boundary. A preliminary version of the model was presented in Čičić et al. (2020b).

Front-tracking has long been used as a method for finding entropy solutions to PDE traffic models (Lu et al., 2009; Monache and Goatin., 2014; Holden and Henrik Risebro, 2015), and is known to yield exact solutions given a piecewise linear flux function and piecewise constant initial conditions. A comparison of front-tracking and other prominent methods for solving the LWR equations (Lax–Hopf algorithm, variational method, and Godunov scheme) is summarized in Mazaré et al. (2011). The hybrid nature of the solution obtained by front-tracking inspired us to formulate the FTSM as a transition system (see Tabuada, 2009 for an introduction to the concept). While the well-posedness of the LWR model with modifications such as the proposed ones, e.g. discontinuous flux function (Adimurthi et al., 2011), is still an open question for the general case, in the specific case that this paper deals with, the solutions exist and are unique, allowing us to define FTSM as a nonblocking deterministic transition system without Zeno behaviour. This model formulation enables straightforward extensions which can be made by redefining some model components (e.g. transitions), without substantial changes to the remainder of the model. For example, in this work, stop-and-go waves were implemented by modifying one of the transitions. As a model based on transition systems, the FTSM bears a resemblance in spirit to the SHIFT framework (Deshpande et al., 1996) and other hybrid system approaches explored within the California PATH program.

Next, we leverage the specific form of the model to reconstruct the traffic state and learn different aspects of the model. Since the FTSM provides a cell-free description of the traffic situation, we are able to accurately represent the local traffic data,coupling the traffic state measurements with the position of the CAVs that acquired them. Local traffic density measurements are used to

Table 1
Overview of most important notation.

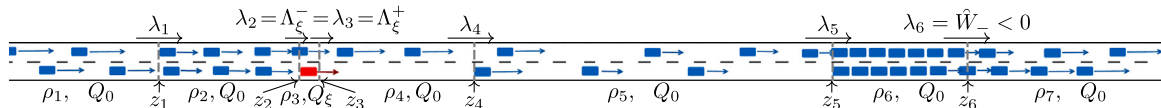| Symbol	Meaning	Symbol	Meaning
x	Position	n	Number of active fronts in the FTSM
t, τ	Time (absolute t, duration τ)	z	Front position
ρ	Traffic density	ξ	Flux function and CAV identifier
q	Traffic flow	F	Forced density in the FTSM
Q	Flux function, fundamental diagram	u_ξ	Reference speed for CAV ξ
σ, Σ	Flux function breakpoint densities	ω_ξ	Overtaking flow of CAV ξ
V	Flux function slopes	ψ	Stop-and-go wave identifier
λ	Front propagation speed	$\boldsymbol{f}$	Set of f
Λ	Flux function boundary propagation speed	$\check{f}$	Measurement of quantity f
W	Front speed bounds	$\hat{f}$	Estimate or prediction of quantity f
X	State of the FTSM	B_f	Bound on f in the algorithms

Fig. 1. An illustration of the motivating example. Flux function Q_0 describes the behaviour of the background traffic, while flux function Q_ξ describes the traffic flowing past a slow-moving CAV (from $x_1^Q(t) = z_2(t)$ to $x_2^Q(t) = z_3(t)$, $\dot{z}_2(t) = \dot{z}_3(t) = \Lambda_\xi^\pm$). The downstream end of a stop-and-go wave $z_6(t)$ propagates upstream at constant speed $\hat{W}_-$. See Fig. 2 for a more detailed explanation of the example, showing its evolution in time.

reconstruct the traffic state, as well as to learn the flux functions and other model components, along with measurements of traffic speed or flow. As opposed to e.g. Seo et al. (2019), where a triangular flux function was estimated from similar CAV data, we only adopt the broad assumption that the flux functions can be approximated as a piecewise linear function, as in Trafikverket (2014). We propose simple stream processing event-triggered methods for on-line learning of the flux function of the background traffic, moving bottleneck flux function, and front speed bounds which model stop-and-go waves.

Finally, we use the proposed model to design and implement an adaptive prediction-based traffic control law. We utilize CAVs as moving bottlenecks to dissipate stop-and-go waves, simultaneously with traffic state reconstruction and model learning based on the measurements provided by the same CAVs. This control problem is characterized by a significant delay from the time we apply some control action using a CAV to the time when this control action has an effect on the stop-and-go wave. Therefore, in order to efficiently dissipate stop-and-go waves, we need to be able to predict the evolution of the traffic over a long horizon, until the time all of the control actions have an effect on them. Since the FTSM is characterized by a diminishing increases in computational burden as the prediction horizon increases, it is a good candidate for use as a prediction model for the proposed control law.

The remainder of this article is organized as follows. First, in Section 2 we motivate the introduction of the FTSM through an illustrative example, describe the model, including how the influence of CAVs is captured, and give some of model properties. Discussing the front-tracking solutions of the LWR model, including the wave-speed-bounded solution and the solution to Riemann problems at the interface between zones with two different flux functions, is deferred to the Appendix, together with theorem proofs. Next, in Section 3, we present the traffic control architecture and control laws for stop-and-go wave dissipation, using CAVs as the only sensors and actuators, based on the reconstructed traffic state, and using a traffic model learned from the local measurements. Simple algorithms for traffic state reconstruction and learning different parts of the model (background traffic flux function, moving bottleneck diagram, and front speed bounds) are proposed. This control law is then tested in simulations in Section 4 and compared with control laws that have access to the full traffic state and know the model. Finally, in Section 5, we conclude.

2. Front-tracking transition system model

In this section we introduce the FTSM. The use of the model is first motivated through an illustrative example, followed by a description of the FTSM and its properties. An overview of the most important notation is given in Table 1. For succinctness, the background about the front-tracking method, upon which the model is based, is given in Appendix A, and the full definition and implementation of the FTSM is given in Appendix B, together with proofs of the theorems given in this section in Appendix C.

2.1. Motivating example

Consider an example consisting of a homogeneous stretch of road, with no on- and off-ramps, and with a single CAV upstream of a single stop-and-go wave, as illustrated in Fig. 1. We are interested in modelling the evolution of the traffic state, assuming the CAV moves at piecewise constant speed (i.e., acceleration and deceleration happens instantly) not faster than the rest of the traffic around it, and the stop-and-go wave propagates upstream at a constant speed. This situation can be described using the LWR model with a space-dependent flux function and a bound on the upstream propagation speed of the rarefaction wave.

The LWR model (James Lighthill and Beresford Whitham, 1955; Richards, 1956) is a first-order scalar hyperbolic conservation law given by the partial differential equation

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0, \tag{1}$$

where t is the time, x the position, the conserved quantity $\rho(x,t)$ is the traffic density, and $q(x,t)$ is the traffic flow. The traffic flow $q(x,t)$ is given according to the flux function, or fundamental diagram,

$$q(x,t) = Q(\rho(x,t), x, t),$$

depending on the traffic density, but also possibly on x and t, differently from the standard LWR formulation. We assume that the initial condition $\rho(x,0)$ is piecewise constant,

$$\rho(x,0) = \begin{cases} \rho_1, & x < x_1^\rho \\ \vdots \\ \rho_i, & x_{i-1}^\rho < x < x_i^\rho \\ \vdots \\ \rho_{N^\rho+1}, & x > x_{N^\rho}^\rho \end{cases} \tag{2}$$

and the space-dependent flux function $Q(\rho, x, t)$ is given piecewise, by

$$Q(\rho, x, t) = \begin{cases} Q_1(\rho), & x < x_1^Q(t), \\ \vdots \\ Q_j(\rho), & x_{j-1}^Q(t) < x < x_j^Q(t), \\ \vdots \\ Q_{N^Q+1}(\rho), & x > x_{N^Q}^Q(t). \end{cases} \tag{3}$$

Here $x_j^Q(t)$ are piecewise linear and continuous functions, $\dot{x}_j^Q(t) = \Lambda_j^Q(t)$, with piecewise constant $\Lambda_j^Q(t)$, and $x_j^Q(t) \le x_{j+1}^Q(t)$. Effectively, the space is divided into zones $[x_{j-1}^Q(t), x_j^Q(t)]$ where the dynamics of the traffic are described by different flux functions $Q = Q_j(\rho(x,t))$, e.g., zones where the road has different numbers of lanes. We assume each flux function $Q_j(\rho)$, $j = 1, \ldots, N^Q$, is a piecewise linear continuous function of the form

$$Q_j(\rho) = \begin{cases} V_{j,1}\rho, & 0 \le \rho \le \sigma_{j,1}, \\ Q_j(\sigma_{j,1}) + V_{j,2}(\rho - \sigma_{j,1}), & \sigma_{j,1} < \rho \le \sigma_{j,2}, \\ \vdots \\ Q_j(\sigma_{j,i-1}) + V_{j,i}(\rho - \sigma_{j,i-1}), & \sigma_{j,i-1} < \rho \le \sigma_{j,i}, \\ \vdots \\ Q_j(\sigma_{j,m_j}) + V_{j,m_j}(\rho - \sigma_{j,m_j}), & \sigma_{j,m_j} < \rho \le \sigma_{j,m_j+1}, \\ 0, & \rho > \sigma_{j,m_j+1}, \end{cases} \tag{4}$$

with $Q_j(\sigma_{j,m_j}) + V_{j,m_j}(\sigma_{j,m_j+1} - \sigma_{j,m_j}) = 0$. We denote values of Q_j at the breakpoints $q_{j,i}^\sigma = Q_j(\sigma_{j,i})$, the set of breakpoints

$$\Sigma_{Q_j} = \left\{ \sigma_{j,1}, \ldots, \sigma_{j,m_j+1} \right\},$$

and the set of slopes between the breakpoints as

$$\boldsymbol{V}_{Q_j} = \left\{ V_{j,1}, \ldots, V_{j,m_j} \right\}.$$

The minimum and maximum slopes are $V_{Q_j}^{\min} = \min \left\{ \boldsymbol{V}_{Q_j} \right\}$ and $V_{Q_j}^{\max} = \max \left\{ \boldsymbol{V}_{Q_j} \right\}$, respectively. Note that the maximum traffic speed $v_{Q_j}^{\max}$ need not be the same as the maximum slope, $v_{Q_j}^{\max} = \max_i q_{j,i}^\sigma / \sigma_{j,i} \le V_{Q_j}^{\max}$, although in practice, most flux functions are concave for $\rho \in [0, \sigma_{Q_j}^{\max}]$, where $\sigma_{Q_j}^{\max} = \arg\max_\rho Q_j(\rho)$. The set of all functions Q_j that satisfy these requirements is denoted $\mathcal{Q}$. Assuming this form of flux functions allows us to handle simple flux functions, such as the popular triangular flux function, but also to approximately handle generic smooth flux functions, such as Greenshields' flux function. In the example depicted in Fig. 1, the background traffic would be described by some flux function $Q_0(\rho)$, and the traffic flowing past the slow-moving CAV with position $x_\xi(t)$ and length l_ξ can, e.g., be described by a scaled-down version of the flux function $Q_\xi(\rho) = \frac{1}{2}Q_0(2\rho)$, yielding $Q(\rho, x, t) = Q_\xi(\rho)$ for $x_\xi(t) - l_\xi < x < x_\xi(t)$, and $Q(\rho, x, t) = Q_0(\rho)$ elsewhere. The simplest (triangular) form of flux functions $Q_0(\rho)$ and $Q_\xi(\rho)$ are shown in Fig. 2 on the left. Note that the triangular flux function was used here for its simplicity, and a more generic continuous piecewise-linear flux function will be used later.

Depending on the shape of the flux function, the solution defined by this model may contain parts that do not reflect the real-world traffic behaviour and phenomena, e.g., stop-and-go waves, caused by the difference between deceleration when entering congestion, and acceleration when leaving it. Since the stop-and-go wave propagation speed is empirically observed to be
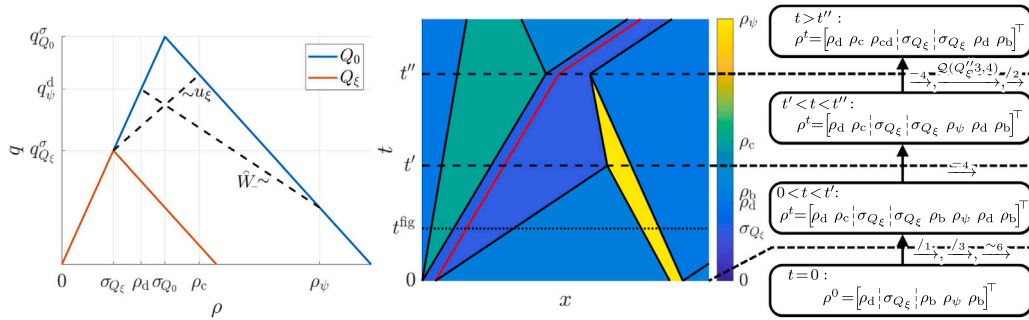
Fig. 2. An example of the evolution of $\rho(x,t)$ (centre), its hybrid representation (right) for the motivating example illustrated in Fig. 1, and the flux functions used therein (left). The evolution of the traffic density profile is colour-coded, with higher density corresponding to warmer colours. There is a slow-moving CAV restricting the traffic flow, represented by flux function Q_ξ and its downstream end is shown in red. At $t = 0$, the initial condition is given by ρ^0. Immediately at $t = 0+$, there is a jump, a zone of congestion ρ_c starts forming upstream of the platoon, a zone of lower traffic density σ_{Q_ξ} starts forming downstream of it, and the discharging flow of the stop-and-go wave causes a zone of density ρ_d to form. Note that the propagation speed of the stop-and-go wave $\hat{W}_-$ is lower than the slope of the congested part of flux function Q_0, as shown on the left, meaning that ρ_d will depend on its traffic density ρ_ψ, and $q_\psi^d < q_{Q_0}^\sigma$. Two more jumps occur at $t = t'$ and $t = t''$, when two fronts collide, and ρ^t is updated accordingly. After the stop-and-go dissipates at $t = t''$, CAV ξ accelerates and starts moving at the speed of the traffic around it. Between the jumps, the fronts move at constant speeds. The moment illustrated in Fig. 1 approximately corresponds to the traffic state at $t = t_{fig}$, with $\rho_1 = \rho_7 = \rho_d$, $\rho_2 = \rho_c$, $\rho_3 = \rho_4 = \sigma_1^\xi$, $\rho_5 = \rho_b$, and $\rho_6 = \rho_\psi$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

approximately constant (Schönhof and Helbing, 2007), we can model them by enforcing a bound on how quickly the downstream end of the stop-and-go wave can propagate. This is achieved by specifying the upstream propagation speed bound for rarefaction, denoted $\hat{W}_-$, which is set to the empirically observed propagation speed of stop-and-go waves.

The front-tracking method for numerically finding solutions to PDEs, upon which the proposed model is based, is described in Appendix A. This method corresponds to starting with piecewise constant initial conditions $\rho(x, 0)$, assuming piecewise linear flux functions, and then solving a sequence of Riemann problems which describe the behaviour of $\rho(x, t)$ around its discontinuities and discontinuities in $Q(\rho, x, t)$. These discontinuities in $\rho(x, t)$ and $Q(\rho, x, t)$ are known as fronts, and the resulting solution $\rho(x, t)$ will be piecewise constant for each t, consisting of periods when the fronts move at constant speeds. The number and speed of these fronts only changes when two of the fronts collide, or if there is an exogenous change in the system. An example of the solutions thus obtained is shown in Fig. 2 in the middle.

Note that since $\rho(x, t)$ is piecewise constant for all t, it is unequivocally defined by a vector of front positions z^t and a vector of traffic densities between them ρ^t. These vectors exhibit distinctly hybrid behaviour, with flow dynamics corresponding to fronts propagating on the road, i.e., changing all elements of z^t at constant rates, and jump dynamics corresponding to changes in ρ^t and z^t due to front collisions or exogenous changes to the system. An example of the jump dynamics of ρ^t is shown in Fig. 2 on the right.

2.2. Model description

Here we present the general structure and properties of the FTSM, with the details about the implementation of the model deferred to Appendix B. We follow the transition system formulation given in Tabuada (2009), and define the FTSM as the execution of the transition system given by the quadruple $(\mathcal{X}, \mathcal{X}_0, \mathcal{U}, \rightarrow)$, which describes the evolution of the front-tracking solution to the considered problem. The model described herein is geared towards the traffic scenario exemplified in the previous subsection. Namely, in the remainder of this work, we assume that the road segment under consideration is homogeneous and without on- and off-ramps, that the CAVs affect the rest of the traffic by acting as moving bottlenecks, and that the downstream front of stop-and-go waves propagates upstream at constant speed. Extensions of the FTSM that allow it to capture some more general traffic conditions will be discussed in future work.

The set of states $\mathcal{X} = (n, t, \bar{z}, \bar{\rho}, \bar{Q}, W)$ is composed of:

- Number of active fronts: $n \in \mathbb{N}$, $n \leq n_{max}$
- Time: $t \in \mathbb{R}_{\geq 0}$
- Front positions: $\bar{z} \in \mathbb{R}^{n_{max}}$, $z_i \leq z_{i+1}$ for $i = 1, \ldots, n$
- Traffic densities: $\bar{\rho} \in \mathbb{R}_{\geq 0}^{n_{max}+1}$
- Flux functions: $\bar{Q} \in \mathcal{Q}^{n_{max}+1}$, where $\mathcal{Q}$ is a set of flux functions
- Front speed bounds: $W \in \mathbb{R}^4$.

The four front speed bounds are written together as $W = (\check{W}_-, \check{W}_+, \hat{W}_-, \hat{W}_+)$, with $\check{W}_- < \check{W}_+$, and $\hat{W}_- < \hat{W}_+$. Each flux function can be described by a quintuple $q = (V, \Sigma, \Lambda^\pm, \xi, F)$, consisting of:

- Slopes and breakpoints: $(V, \Sigma) \in \mathcal{L}$

- Boundary speeds: $\Lambda^{\pm} \in \mathbb{R}^2$
- Identifier: $\xi \in \mathbb{Z}$
- Forced density: $F \in \mathbb{R} \cup \{\oslash\}$,

where the set of feasible slopes and breakpoints is

$$\mathcal{L} = \left\{ V \in \mathbb{R}^m, \Sigma \in \mathbb{R}_{>0}^{m+1} : \sigma_1 < \sigma_2 < \ldots < \sigma_{m+1}, \ldots, q_{Q,i}^\sigma \geq 0, i = 1, \ldots, m, q_{Q,m+1}^\sigma = 0 \right\},$$

$$q_{Q,1}^\sigma = V_1 \sigma_1, \quad q_{Q,i}^\sigma = q_{Q,i-1}^\sigma + V_i(\sigma_i - \sigma_{i-1}), i = 2, \ldots, m+1,$$

without any further constraints on the shape of the flux function. As the number of breakpoints of flux functions increases, we are able to more closely approximate any generic (possibly even non-concave) flux function, at the cost of increased computational complexity, but preserving the general form of the model. Note that in spite of its generality, the model does not sacrifice computational efficiency when dealing with simple flux functions, such as the triangular one. The states of the FTSM are exemplified in Fig. 1.

We denote the traffic flow, given traffic density ρ, as $q = Q_*(\rho)$ and calculate it as like $Q_j(\rho)$ in (4), with $V_{j,i}$ and $\sigma_{j,i}$ given in V and Σ, respectively. Boundary speeds $\Lambda^{\pm} = (\Lambda^-, \Lambda^+)$ represent the propagation speed of the upstream and downstream boundary of the region of Q_*. The identifier ξ differentiates flux functions, and is unique for each flux function. We also use it to define the precedence when determining the propagation speed of the boundary between regions with different flux functions. Namely, at a boundary between flux functions $Q_i \neq Q_{i+1}$, the propagation speed is given by

$$\lambda_i = \begin{cases} \Lambda_i^+, & \xi_i > \xi_{i+1}, \\ \Lambda_{i+1}^-, & \xi_i < \xi_{i+1}. \end{cases}$$

We define the active fronts and densities

$$z = \begin{bmatrix} z_1 & \cdots & z_n \end{bmatrix}^\top = \begin{bmatrix} I_n & \mathbb{0}_{n \times n^{\max}-n} \end{bmatrix} \bar{z}$$

$$\rho = \begin{bmatrix} \rho_1 & \cdots & \rho_{n+1} \end{bmatrix}^\top = \begin{bmatrix} I_{n+1} & \mathbb{0}_{n+1 \times n^{\max}-n} \end{bmatrix} \bar{\rho},$$

to be all the states required to reconstruct the full traffic density profile $\rho(x,t)$. The maximum number of fronts $n_{\max}$ can be taken large enough so that the number of active fronts never exceeds it, but effectively, the dimension of active states will vary as a part of the model dynamics. Furthermore, only the active fronts and densities, along with flux functions $Q = \begin{bmatrix} Q_1 & \cdots & Q_{n+1} \end{bmatrix}^\top = \begin{bmatrix} I_{n+1} & \mathbb{0}_{n+1 \times N-n} \end{bmatrix} \bar{Q}$, will influence the behaviour of the system, so when describing the transitions, we only define their updates, and the inactive states may take arbitrary values.

Given the current state $X \in \mathcal{X}$ of the transition system, the density function $\rho(x,t)$ describing the current state of the system can be reconstructed based on $z_1, \ldots, z_n$ and $\rho_1, \ldots, \rho_{n+1}$:

$$\rho(x,t) = \begin{cases} \rho_1, & x < z_1, \\ \vdots & \\ \rho_i, & z_{i-1} < x < z_i, \\ \vdots & \\ \rho_{n+1}, & x > z_n. \end{cases}$$

Note the abuse of notation as we use $\rho(x,t)$ for the reconstructed function, and $\rho = \begin{bmatrix} \rho_1 & \cdots & \rho_{n+1} \end{bmatrix}^\top$ for the vector of traffic densities.

The set of initial states $\mathcal{X}_0$ can be the same as the set of all states. Inputs to the system $\mathcal{U}$ describe the passage of time and exogenous transitions. The evolution of the system state is described by defining the various transitions that arise in the process of calculating the solution. Here we only give a short description of what each of the transitions does, and the full definition is deferred to Appendix B. The transitions are listed in order of increasing priority, i.e., we first present the transitions that can only be taken if no other transition can be taken, and end with transitions which do not depend on the state of the system, only on exogenous inputs. The required transitions are: passage of time $\tau(t_{\mathrm{end}})$, front interaction $-_i$, internal Riemann transition $\sim_i$, boundary Riemann transition $/_i$, state forcing $!_i$, state insertion $\vee(\rho_\vee, x_\vee)_i$, and flux function transition $\mathcal{Q}(q, i, j)$. All these transitions are illustrated in Fig. 3, as well as shown in the example in Fig. 2.

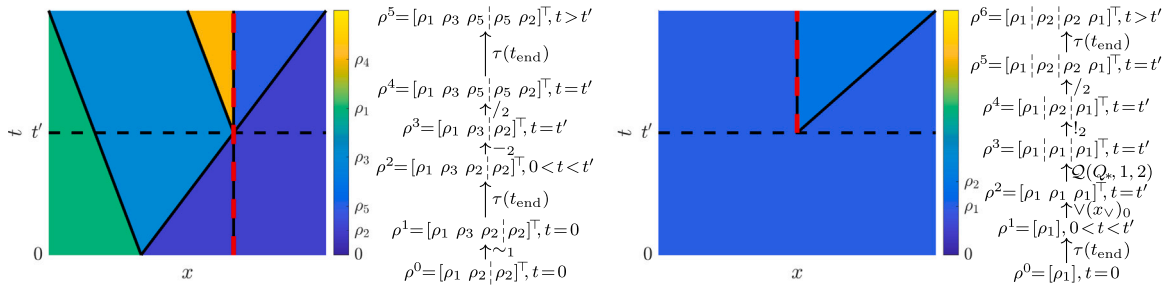The first transition, passage of time, describes the propagation of fronts,

$$(t, z) \xrightarrow{\tau(t_{\mathrm{end}})} (t', z'),$$

between their interactions, or until the externally provided goal time t_{end}. Traffic densities ρ, number of active states n and flux functions Q do not change in these transitions, and only the front positions of active states ($i = 1, \ldots, n$) are changed.

The front interaction transition describes what happens when two fronts collide, i.e., if the positions of these two fronts becomes equal, $z_i = z_{i+1}$ while their distance is decreasing, $\lambda_i > \lambda_{i+1}$,

$$(n, z, \rho, Q) \xrightarrow{-_i} (n', z', \rho', Q').$$

This transition corresponds to simply deactivating one state, with $n' = n - 1$.

(a) Transitions $\tau(t_{\text{end}})$, $-_i$, $\sim_i$, and $/_i$. Area downstream of the dashed red line is modelled by a different flux function (with lower capacity) from the area upstream (Q_+).

(b) Transitions $\tau(t_{\text{end}})$, $\vee(x_\vee)_i$, $\mathcal{Q}(Q_\mathcal{Q}, i, j)$, $!_i$, and $/_i$. The dashed red line indicates the infinitesimally narrow zone described by flux function Q_* with state forced to $F_* = \rho_2$.

Fig. 3. Illustration of the transitions used in the FTSM. Traffic density profiles are shown colour-coded, and the transition system path (only ρ and ι) and transitions are shown to the right of them.

Internal Riemann transition results from solutions to Riemann problem with the same flux function on both sides of the discontinuity, and can be described by

$$(n, z, \rho, Q) \xrightarrow{\sim_i} (n', z', \rho', Q')$$

Depending on ρ_i and ρ_{i+1}, the number of active states can decrease (if $\rho_i = \rho_{i+1}$), increase, or stay the same. Note that the solution to the basic LWR model can entirely be modelled by the three hitherto described transitions.

Next, the boundary Riemann transition can occur at interfaces between zones with different flux functions, and is given by

$$(n, z, \rho, Q) \xrightarrow{/_i} (n', z', \rho', Q').$$

Here, densities $\rho'_- = \rho'_-(\rho_i, \rho_{i+1}, Q_i, Q_{i+1})$ and $\rho'_+ = \rho'_+(\rho_i, \rho_{i+1}, Q_i, Q_{i+1})$ are obtained by maximizing the flow over the flux functions boundary, subject to entropy conditions in the interiors on both sides of the discontinuity.

We use the forced density F and the state forcing transition to potentially enforce internal boundary constraints, with $F_i \neq \varnothing$ signifying that the traffic density in the zones described by Q_i must be $\rho_i = F_i$, and $F_i = \varnothing$ signifying that the traffic density in said zone is not forced, and can evolve according to the other hitherto described transitions. The transition is given by

$$(\rho) \xrightarrow{!_i} (\rho').$$

The two exogenous transitions can be taken for any $X \in \mathcal{X}$ given the appropriate external input, are state insertion and the flux function transition. State insertion consists of adding two fronts at position $x_\vee$ downstream of front i, with $z_i \leq x_\vee \leq z_{i+1}$,

$$(n, z, \rho, Q) \xrightarrow{\vee(x_\vee)_i} (n', z', \rho', Q'),$$

whereas flux function transition covers changes in flux functions in specific areas,

$$(Q) \xrightarrow{\mathcal{Q}(Q_\mathcal{Q}, i, j)} (Q').$$

The potential for using CAVs as both traffic sensors and actuators makes them important for future traffic control systems. We describe how both these roles of CAVs can be modelled in the FTSM framework. We represent adding a CAV to the model, with identifier ξ at position x_ξ and moving at speed u_ξ, by first taking a transition $\vee(x_\xi)_{i_-}$, which creates two fronts downstream of front i_-, $z_{i_-} \leq x_\xi \leq z_{i_-+1}$, followed by taking a transition $\mathcal{Q}(Q_\xi, i_- + 1, i_- + 2)$, with $Q_\xi = (V_\xi, \Sigma_\xi, (u_\xi, u_\xi), \xi, F_\xi)$. The choice of V_ξ, Σ_ξ and F_ξ depends on the role that the CAV has. The identifier corresponding to each CAV is greater than the identifier of the flux function describing the background traffic.

A CAV ξ acting as a sensor communicates its local measurements of traffic density $\check{\rho}_\xi$, and traffic speed $\check{v}_\xi$, from which we get $\check{q}_\xi = \check{\rho}_\xi \check{v}_\xi$. In case CAV ξ is moving slower than the surrounding traffic, $u_\xi < v_{\xi+}$, we measure the overtaking flow $\check{\omega}_\xi$. In general, the values that are directly related to measurements are denoted by caron ($\check{\ }$). We model CAV ξ acting as an actuator by setting its reference speed u_ξ, which determines their actual speed $\Lambda_\xi^\pm$ together with the speed of the traffic immediately downstream of it $v_{\xi+}$,

$$\Lambda_\xi^\pm = \min\{u_\xi, v_{\xi+}\}, \tag{5}$$

$$v_{\xi+} = Q_{i_{\xi+}}(\rho_{i_{\xi+}})/\rho_{i_{\xi+}}, \tag{6}$$

$$i_{\xi+} = \min i, \text{ s.t. } z_i \geq x_\xi, Q_i \neq Q_\xi. \tag{7}$$

Note that both its reference speed u_ξ and the speed of traffic downstream of it $v_{\xi+}$ are piecewise constant, i.e., CAV ξ instantly adjusts its speed in response to a change in reference speed or traffic conditions according to (5). A vehicle moving slower than the
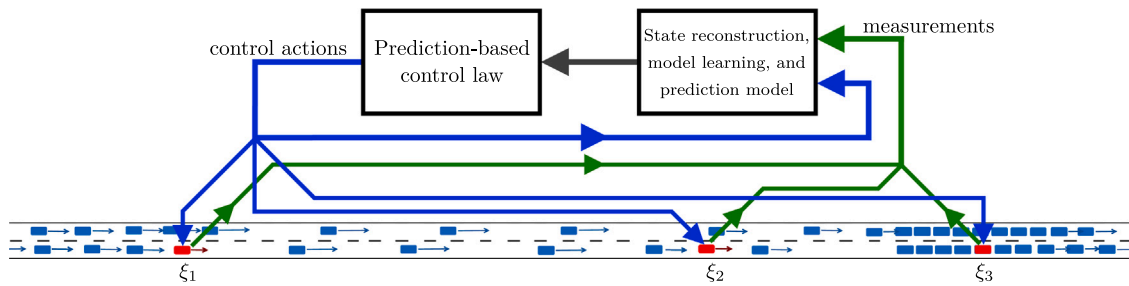
Fig. 4. Closed-loop system structure. Local traffic measurements are collected by the CAVs and sent to the infrastructure, where traffic state reconstruction and model learning is implemented. Based on the current state estimate, control actions for the CAVs are calculated using the learned predictions model.

rest of the traffic acts as a moving bottleneck, limiting the overtaking flow, which we may use for traffic control. We model the effect a CAV acting as a moving bottleneck has on the rest of the traffic by using V_ξ and Σ_ξ in the shape of a bottleneck diagram.

Finally, we study the properties of the solutions to the FTSM, and how they correspond to the wave-speed-bounded solutions of the LWR with piecewise linear flux functions and piecewise constant initial conditions. The results are stated in the following theorems, with proofs given in Appendix C.

Theorem 1. *Let X_0 be the initial state of the FTSM with $t = 0$. Consider the zone between two flux function boundaries $[z_-^*(t), z_+^*(t)]$, described by flux function Q^*, and assume that for $0 < t < T$, $T > 0$: $z_-^*(t) < z_+^*(t)$, boundary conditions $\rho_{i_-^*(t)}$ and $\rho_{i_+^*(t)+1}$ are constant, and fronts from outside of $[z_-^*(t), z_+^*(t)]$ do not collide with $z_-^*(t)$ or $z_+^*(t)$. Then $\rho(x, t')$, $x \in [z_-^*(t'), z_+^*(t')]$, given by X' with $t = t'$ is the unique wave-speed-bounded solution of the corresponding LWR model with initial conditions $\rho(x, 0)$, $x \in [z_-^*(0), z_+^*(0)]$, given by X_0, for $0 < t' < T$, $T > 0$, and its Total Variation T.V.$(\rho(\cdot, t'))$ and number of active fronts is nonincreasing in t'.*

This theorem establishes correspondence between the solutions of the FTSM and the appropriate LWR model within zones where the flux functions are homogeneous. It also holds if all flux functions Q_i are identical, in which case $[z_-^*(t), z_+^*(t)]$ corresponds to the entire space. In case the wave-speed bounds are such that the entropy solution is identical to the wave-speed-bounded solution, this theorem corresponds to Corollary 2.8 in Holden and Henrik Risebro (2015), for constant boundary conditions and inside a zone where the flux function is homogeneous. Furthermore, since the number of active fronts is nonincreasing in time, the numerical complexity of evaluating the FTSM is also nonincreasing with time, which results in diminishing increases in computational burden as the prediction horizon increases in case the model is used to predict the evolution of the traffic.

Finally, the solution of the FTSM is shown to exist and be well-defined, in particular, to be devoid of Zeno behaviour, ensuring that the execution of the FTSM indeed proceeds until the desired time. We define Zeno behaviour as the existence of a solution with an infinite series of transitions that are not passage of time $X \xrightarrow{\tau(t_{\text{end}})} X'$, with $t < t_{\text{end}}$, $t < t'$.

Theorem 2. *There exists a unique, non-Zeno solution to the FTSM, with initial state X_0 and a finite number of exogenous transitions, that defines a unique wave-speed-bounded solution of the LWR (1), $\rho(x, t)$, with corresponding: piecewise constant initial conditions (2), piecewise defined spacetime-dependent flux function (3) with each spatiotemporal zone defined by a continuous piecewise linear flux functions (4) and boundaries between zones described by different flux functions (4) propagating at constant speeds, and exogenous transitions.*

Note that here we only state the quantities of the FTSM, whereas the rigorous discussion of the nature of wave-speed-bounded solutions of the considered form of the LWR model is left to future work. As discussed in Appendix A, in the case studied in this paper, the wave-speed-bounded solutions and flux function boundary solutions exist and are unique. These solutions may cause an increase in the total variation of $\rho(x, t)$, but assuming a finite number of exogenous transitions, this increase is bounded.

In summary, the FTSM consists of a transition system which describes the process of finding solutions to the LWR model with space-dependent flux functions and front speed bounds. It can handle any piecewise linear flux function, and yields exact wave-speed-bounded solutions, as defined in Appendix A. As outlined in Mazaré et al. (2011), due to the event-based nature of the front-tracking method, the computation time will strongly depend on the particular traffic situation. In case there is a large number of fronts, and flux functions have a large number of breakpoints, the computation time for the exact solution can be high. Since we typically rely on noisy measurements, uncertain flux functions, and simplifications of more complex underlying models, calculating an exact solution is fruitless in practice. Therefore it is advantageous to introduce some level of approximation in order to lower the computational burden, without sacrificing too much performance.

3. Reconstruction-based stop-and-go wave dissipation control

With increasing numbers of CAVs entering the roads, traffic control with CAVs used as both sensors and actuators, as outlined in Fig. 4, is becoming feasible. We assume that CAVs can collect and communicate local traffic measurements, and that their reference speed can be used as control inputs and set from the infrastructure. These measurements of traffic density, speed, and overtaking
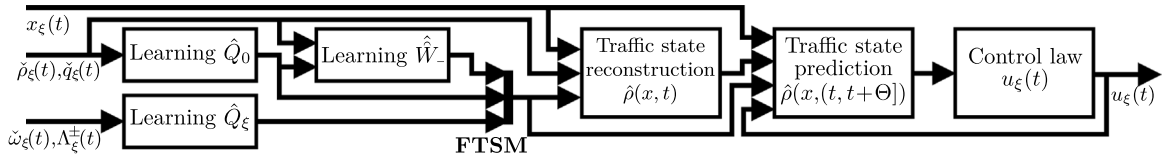
Fig. 5. Data flows for reconstruction-based control for stop-and-go wave dissipation, using the FTSM with learned model components for traffic state reconstruction and prediction. The control actions are calculated based on the predicted evolution of the traffic.

flow could be acquired using on-board sensors required for CAV situation awareness, or possibly estimated based on the speed of the CAV in case it is acting as a floating car, and we use them to reconstruct the traffic state, as well as to identify the flux function, moving bottleneck diagram, and stop-and-go wave speed. The traffic model thus learned is used to predict the evolution of the traffic state based on the chosen control actions for all CAVs, and based on these predictions, we are able to calculate the control actions. We use the CAVs acting as controlled moving bottlenecks to dissipate stop-and-go waves in order to minimize the Total Time Spent of all vehicles.

We adopt the same assumptions as in Section 2, i.e., that the considered stretch of road is homogeneous and can be described by a single flux function Q_0. The FTSM used for traffic state reconstruction and prediction is initialized at some start time $t^0 = 0$ to

$$\mathcal{X}^0 = \left(n^0, t^0, z^0, \rho^0, Q^0, W^0\right) = \left(0, 0, \emptyset, \rho^{\text{avg}}, Q^0, W^0\right),$$
$$Q^0 = Q_0 = \left(V^0, \Sigma^0, \Lambda^{\pm,0}, \xi^0, F^0\right) = \left([0], \left[\rho^{\max}\right], (0,0), 0, \emptyset\right),$$

with constant initial traffic density ρ^0 using the overall average density known from historical data, wave-speed bounds W^0 such that all wave-speeds are allowed, and $\rho^{\max}$ chosen large enough so that the traffic density will never exceed it. The flux functions, including front speed bounds, and traffic state reconstruction are updated as new measurements become available, according to the algorithms described in what follows.

We first propose algorithms that can be used to identify each aspect of the traffic model (flux function of the road Q_0, flux function of the CAVs acting as moving bottlenecks Q_ξ, and wave speed bound $\hat{W}_-$ used to model stop-and-go waves), using only local measurements of traffic density $\check{\rho}_\xi$, traffic flow $\check{q}_\xi$, and overtaking flow $\check{\omega}_\xi$ provided by the CAVs ξ. The proposed algorithms exploit the continuous piecewise linear structure of the flux functions to simplify calculations and avoid making assumptions about their shape. Indeed the only assumption made is that the dynamics of the traffic can, at least approximately, be described as a first-order conservation law, ignoring the dynamics of the traffic speed. Then, assuming that all flux functions and other parameters of the traffic model are known, we describe how local traffic density measurements $\check{\rho}_\xi$ can be used to reconstruct the traffic state, which will later be used to calculate control actions. Finally, we present a control law that uses the FTSM as the prediction model, along with the proposed traffic state reconstruction and model learning algorithms, as well as some simplified control laws which use more information and are used as benchmarks. The general flow of data, from CAV trajectories $x_\xi(t)$, current CAV speed $\Lambda_\xi^\pm(t)$, and current measurements $\check{\rho}_\xi(t)$, $\check{q}_\xi(t)$, and $\check{\omega}_\xi(t)$, to control action updates $u_\xi(t+)$, is shown in Fig. 5, and the remainder of this section broadly reflects its structure.

3.1. Learning the traffic flux function Q_0

Consider a CAV ξ travelling through a road segment described by flux function Q_0, and let its reference speed be $u_\xi \geq v_{Q_0}^{\max}$. Then the actual speed of the CAV depends only on the speed of the traffic immediately downstream of it, and does not influence the surrounding traffic, so $\Lambda_\xi^\pm = \check{v}_\xi$, where $\check{v}_\xi$ denotes the local measurement of the traffic speed collected by CAV ξ. Since the CAV is following the flow of the traffic, we have the measured overtaking flow $\check{\omega}_\xi = 0$, and the measurement of the traffic flow $\check{q}_\xi$ should be consistent with the flux function, $Q_0(\check{\rho}_\xi) \approx \check{q}_\xi$.

We propose a stream processing event-triggered method for on-line flux function update, summarized in Algorithm 1, where operator avg denotes finding the average value, and operator # the number of elements of some set. Let Ξ_{Q_0} be the set of CAVs collecting measurements about the flux function Q_0. If, for any CAV ξ, the deviation of the measured traffic flow $\check{q}_\xi$ from the modelled traffic flow based on the measured traffic density $Q_0(\check{\rho}_\xi)$ is greater than some margin B_q, we update the flux function based on the measurement. Choosing higher B_q reduces the update frequency and sensitivity to measurement noise, at the cost of higher model error. Instead of using all the collected data, the set of measurements which we use for flux function fitting $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$ consists only of the measurements that trigger a flux function update.

First, we add the $(\check{\rho}_\xi, \check{q}_\xi)$ for which $\left|Q_0(\check{\rho}_\xi) - \check{q}_\xi\right| > B_q$ to $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$. Next, if the flux function already has any breakpoints close to the measurement, $\left|\sigma_i - \check{\rho}_\xi\right| < B_\sigma$, they will be replaced by a new breakpoint; otherwise a new breakpoint is added without moving any of the other breakpoints. Choosing higher B_σ decreases the number of breakpoints in flux function Q_0, at the cost of potentially increasing the frequency of updates and higher model error. The new breakpoint σ' and its flux $Q_0(\sigma')$ are taken as the average values of all measurements from $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$ that are close to the new measurement, $\left(\check{\rho}_\xi^{Q_0,B_\sigma}, \check{q}_\xi^{Q_0,B_\sigma}\right) = \left\{(\rho, q) \in (\check{\rho}^{Q_0}, \check{q}^{Q_0}) : \left|\rho - \check{\rho}_\xi\right| < B_\sigma\right\}$.

We therefore add the new breakpoint σ' to Σ_{Q_0} and recalculate V_{Q_0} so that we have $Q_0'(\sigma') = q'$, where σ' is the average of $\check{\rho}_\xi^{Q_0, B_\sigma}$ and q' the average of $\check{q}_\xi^{Q_0, B_\sigma}$. The updated $\Sigma_{Q_0'}$ and $V_{Q_0'}$ are given by

$$\Sigma_{Q_0'} = \begin{bmatrix} \sigma_1 & \dots & \sigma_{i_{\xi_-}} & \vline & \sigma' & \vline & \sigma_{i_{\xi_-}} & \dots & \sigma_m \end{bmatrix}^\top,$$

$$V_{Q_0'} = \begin{bmatrix} V_1 & \dots & V_{i_{\xi_-}} & \vline & V_-' & V_+' & \vline & V_{i_{\xi_-}+1} & \dots & V_m \end{bmatrix}^\top,$$

$$i_{\xi_-} = \begin{cases} \max i \text{ s.t. } \sigma_i \in \Sigma_{Q_0}, \sigma_i < \check{\rho}_\xi - B_\sigma, & \sigma_1 \geq \check{\rho}_\xi - B_\sigma, \\ 0, & \sigma_1 < \check{\rho}_\xi - B_\sigma, \end{cases}$$

$$i_{\xi_+} = \min i \text{ s.t. } \sigma_i \in \Sigma_{Q_0}, \sigma_i > \check{\rho}_\xi + B_\sigma,$$

$$V_-' = \frac{q' - Q_0(\sigma_{i_{\xi_-}})}{\sigma' - \sigma_{i_{\xi_-}}}, \quad V_+' = \frac{Q_0(\sigma_{i_{\xi_+}}) - q'}{\sigma_{i_{\xi_+}} - \sigma'},$$

where the first blocks of $\Sigma_{Q_0'}$ and $V_{Q_0'}$ are empty if $i_{\xi_-} = 0$, and the third block of $V_{Q_0'}$ is empty if $i_{\xi_+} = m$. We also need to ensure by suitable selection of initial Q_0 that $\sigma_m > \check{\rho}^{\max} - B_\sigma$, i.e., the final breakpoint of Q_0 is never moved.

Finally, in order to accelerate the adaptation of Q_0 to potentially new condition on the road, we may choose to always calculate the breakpoint densities and flows using at most the $B_{\tilde{N}}$ most relevant measurements. Then, if after adding the new measurement to $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$ we have $\#\check{\rho}_\xi^{Q_0, B_\sigma} > B_{\tilde{N}}$, i.e., the number of measurements close to $\check{\rho}_\xi$ is higher than $B_{\tilde{N}}$, we may remove the measurement that differs from the new measurement $(\check{\rho}_\xi, \check{q}_\xi)$ the most.

3.2. Learning the moving bottleneck flux function Q_ξ

Consider again a CAV ξ travelling through a road segment described by flux function Q_0, with reference speed that is lower than the measured speed of the surrounding traffic, $u_\xi < \check{v}_\xi$. Then the CAV follows its reference speed, $\Lambda_\xi^\pm = u_\xi$, the measured overtaking flow is $\check{\omega}_\xi > 0$, and the presence of the slow moving vehicle affects the behaviour of the surrounding traffic. We aim to model this influence by using a different flux function Q_ξ at the position of the CAV, with $\Lambda_\xi^\pm = u_\xi$. We need the measurement to satisfy

$$\begin{aligned} (\forall \rho \geq 0) \quad & Q_\xi(\rho) \leq \check{\omega}_\xi + \rho \Lambda_\xi^\pm, \\ (\exists \rho \geq 0) \quad & Q_\xi(\rho) = \check{\omega}_\xi + \rho \Lambda_\xi^\pm, \end{aligned} \tag{8}$$

i.e., the line $\check{\omega}_\xi + \Lambda_\xi^\pm \rho$ should be tangent to $Q_\xi(\rho)$.

If the conditions (8) are violated by more than some margin B_ω, we update flux function Q_ξ according to the proposed Algorithm 2 so that (8) holds. Choosing a higher B_ω decreases the frequency of updates of Q_ξ and sensitivity to measurement noise, but increases the model error. We first calculate the maximum overtaking flow $\omega_\xi^{\max}$ that Q_ξ allows in case the boundary speed is $\Lambda_\xi^\pm$, which is achieved for traffic density $\sigma_{i_\xi^\omega}$, $\omega_\xi^{\max} = Q_\xi(\sigma_{i_\xi^\omega}) - \Lambda_\xi^\pm \sigma_{i_\xi^\omega}$. Since Q_ξ is piecewise linear, we only need to search for σ_ξ^ω within the set of its breakpoints Σ_{Q_ξ}. Then, if $\omega_\xi^{\max} > \check{\omega}_\xi + B_\omega$, the current flux function Q_ξ admits a higher overtaking flow than the measured

---

Algorithm 1: Learning the traffic flux function Q_0 using measurements from CAVs travelling with the flow

input : Flux function Q_0 defined by V_{Q_0} and Σ_{Q_0}, set of used measurements $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$, new measurements $(\check{\rho}_\xi, \check{q}_\xi)$
output: Updated flux function Q_0 defined by updated V_{Q_0} and Σ_{Q_0}, updated set of used measurements $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$

for $\xi \in \Xi_{Q_0}$ **do**

  if $\left| Q_0(\check{\rho}_\xi) - \check{q}_\xi \right| > B_q$ **then**

    Add $(\check{\rho}_\xi, \check{q}_\xi)$ to $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$;

    Remove from Σ_{Q_0} all $\left\{ \sigma \in \Sigma_{Q_0} : \left| \sigma - \check{\rho}_\xi \right| < B_\sigma \right\}$;

    Find $\left(\check{\rho}_\xi^{Q_0, B_\sigma}, \check{q}_\xi^{Q_0, B_\sigma} \right) = \left\{ (\rho, q) \in \left(\check{\rho}^{Q_0}, \check{q}^{Q_0} \right) : \left| \rho - \check{\rho}_\xi \right| < B_\sigma \right\}$;

    Add $\sigma' = \text{avg} \, \check{\rho}_\xi^{Q_0, B_\sigma}$ to Σ_{Q_0};

    Recalculate V_{Q_0}, so that $Q_0'(\sigma') = \text{avg} \, \check{q}_\xi^{Q_0 B_\sigma}$;

    if $\#\check{\rho}_\xi^{Q_0, B_\sigma} > B_{\tilde{N}}$ **then**

      Remove from $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$ measurements $(\rho, q) = \arg \max\limits_{(\rho, q) \in \left(\check{\rho}_\xi^{Q_0, B_\sigma}, \check{q}_\xi^{Q_0, B_\sigma} \right)} \left(\frac{\rho}{\check{\rho}_\xi} - 1 \right)^2 + \left(\frac{q}{\check{q}_\xi} - 1 \right)^2$

    end

  end

end

one, so we proceed to update it with

$$\Sigma_{Q'_\xi} = \begin{bmatrix} \sigma_1 & \cdots & \sigma_{i_{\xi-}} & \sigma_- & \sigma_+ & \sigma_{i_{\xi-}} & \cdots & \sigma_m \end{bmatrix}^\top,$$

$$V_{Q'_\xi} = \begin{bmatrix} V_1 & \cdots & V_{i_{\xi-}+1} & \Lambda_\xi^\pm & V_{i_{\xi+}} & \cdots & V_m \end{bmatrix}^\top,$$

$$\sigma_- = \min \rho \geq 0 \text{ s.t. } Q_\xi(\rho) = \Lambda_\xi^\pm \rho + \breve{\omega}_\xi$$

$$\sigma_+ = \max \rho \geq 0 \text{ s.t. } Q_\xi(\rho) = \Lambda_\xi^\pm \rho + \breve{\omega}_\xi$$

$$i_{\xi-} = \begin{cases} \max i \text{ s.t. } \sigma_i \in \Sigma_{Q_\xi}, \sigma_i < \sigma_-, \ \sigma_1 \geq \sigma_-, \\ 0, & \sigma_1 < \sigma_-, \end{cases}$$

$$i_{\xi+} = \min i \text{ s.t. } \sigma_i \in \Sigma_{Q_\xi}, \sigma_i > \sigma_+,$$

where the first block of $\Sigma_{Q'_\xi}$ is empty if $i_{\xi-} = 0$. Otherwise, if $\omega_\xi^{\max} < \breve{\omega}_\xi - B_\omega$, the current flux function Q_ξ admits a lower overtaking flow than the measured one, so we proceed to update it with

$$V_{Q'_\xi} = \begin{bmatrix} V_1 & \cdots & V_{i_\xi^\omega-1} & V'_- & V'_+ & V_{i_\xi^\omega+2} & \cdots & V_m \end{bmatrix}^\top,$$

$$V'_- = \frac{\breve{\omega}_\xi + \Lambda_\xi^\pm \sigma_{i_\xi^\omega} - Q_\xi(\sigma_{i_\xi^\omega-1})}{\sigma_{i_\xi^\omega} - \sigma_{i_\xi^\omega-1}}, \ V'_+ = \frac{Q_\xi(\sigma_{i_\xi^\omega+1}) - \breve{\omega}_\xi - \Lambda_\xi^\pm \sigma_{i_\xi^\omega}}{\sigma_{i_\xi^\omega+1} - \sigma_{i_\xi^\omega}},$$

where if $i_\xi^\omega = 1$, we take $\sigma_{i_\xi^\omega-1} = 0$, and the first block of $V_{Q'_\xi}$ is empty.

Algorithm 2: Learning the moving bottleneck flux function Q_ξ using measurements of overtaking flow

---

input : Flux function Q_ξ defined by V_{Q_ξ} and Σ_{Q_ξ}, CAV speed $\Lambda_\xi^\pm$, overtaking flow measurement $\breve{\omega}_\xi$
output: Updated flux function Q_ξ defined by updated V_{Q_ξ} and Σ_{Q_ξ}

Find $\omega_\xi^{\max} = \max_{\sigma \in \Sigma_{Q_\xi}} Q_\xi(\sigma) - \Lambda_\xi^\pm \sigma = Q_\xi(\sigma_{i_\xi^\omega}) - \Lambda_\xi^\pm \sigma_{i_\xi^\omega}$;

if $\omega_\xi^{\max} > \breve{\omega}_\xi + B_\omega$ **then**

    Find $\sigma_- = \min \rho \geq 0$ s.t. $Q_\xi(\rho) = \Lambda_\xi^\pm \rho + \breve{\omega}_\xi$, and $\sigma_+ = \max \rho \geq 0$ s.t. $Q_\xi(\rho) = \Lambda_\xi^\pm \rho + \breve{\omega}_\xi$;

    Remove from Σ_{Q_ξ} all $\left\{ \sigma \in \Sigma_{Q_\xi} : \sigma_- \leq \sigma \leq \sigma_+ \right\}$;

    Add $\{\sigma_-, \sigma_+\}$ to Σ_{Q_ξ};

    Recalculate V_{Q_ξ}, so that $Q'_\xi(\rho) = \breve{\omega}_\xi + \Lambda_\xi^\pm \rho, \sigma_- \leq \rho \leq \sigma_+$;

else if $\omega_\xi^{\max} < \breve{\omega}_\xi - B_\omega$ **then**

    Recalculate V_{Q_ξ}, so that $Q'_\xi(\sigma_{i_\xi^\omega}) = \breve{\omega}_\xi + \Lambda_\xi^\pm \sigma_{i_\xi^\omega}$;

end

Algorithm 3: Learning the front speed bounds W using traffic density and flow measurements

---

input : Front speed bounds W, sets of used front speed measurements $\breve{w}^{\bar{W}_\pm}$, previous measurements $(\breve{\rho}_\xi, \breve{q}_\xi)$, new measurements $(\breve{\rho}'_\xi, \breve{q}'_\xi)$
output: Updated front speed bounds W, updated sets of used front speed measurements $\breve{w}^{\bar{W}_\pm}$

for $\xi \in \Xi$ **do**

    if $|\breve{\rho}'_\xi - \breve{\rho}_\xi| > B_\rho, u_\xi > \frac{\breve{q}_\xi}{\breve{\rho}_\xi}, u_\xi > \frac{\breve{q}'_\xi}{\breve{\rho}'_\xi}, {}^{\breve{\rho}_\xi}\tilde{\Sigma}_{Q_0}^{\breve{\rho}'_\xi} \neq \begin{bmatrix} \breve{\rho}_\xi & \breve{\rho}'_\xi \end{bmatrix}^\top$ **then**

        Find $\breve{w}_\xi = \frac{\breve{q}'_\xi - \breve{q}_\xi}{\breve{\rho}'_\xi - \breve{\rho}_\xi}$;

        Add $\breve{w}_\xi$ to $\begin{cases} \breve{w}^{\hat{W}_-}, & \text{if } \breve{\rho}_\xi > \breve{\rho}'_\xi, \breve{w}_\xi > {}^{\breve{\rho}_\xi}\tilde{V}_{Q_0,1}^{\breve{\rho}'_\xi} \\ \breve{w}^{\hat{W}_+}, & \text{if } \breve{\rho}_\xi > \breve{\rho}'_\xi, \breve{w}_\xi < {}^{\breve{\rho}_\xi}\tilde{V}_{Q_0,1}^{\breve{\rho}'_\xi} \\ \breve{w}^{\hat{W}_-}, & \text{if } \breve{\rho}_\xi < \breve{\rho}'_\xi, \breve{w}_\xi > {}^{\breve{\rho}_\xi}\tilde{V}_{Q_0,1}^{\breve{\rho}'_\xi} \\ \breve{w}^{\hat{W}_+}, & \text{if } \breve{\rho}_\xi < \breve{\rho}'_\xi, \breve{w}_\xi < {}^{\breve{\rho}_\xi}\tilde{V}_{Q_0,1}^{\breve{\rho}'_\xi} \end{cases}$;

        Recalculate all W for which $\breve{w}^{\bar{W}_\pm} \neq \emptyset$ so that $\bar{W}_\pm = \text{avg} \breve{w}^{\bar{W}_\pm}$;

    end

end

3.3. Learning the front speed bounds W

Finally, we use the traffic density and speed measurements of the CAVs to learn the front speed bounds W. We detect the influence of front speed bounds upon a change in the measurements of some CAV ξ as it enters a zone of different traffic density, when $|\check{\rho}'_\xi - \check{\rho}_\xi| > B_\rho$, where $\check{\rho}_\xi$ is the old traffic density measurement, and $\check{\rho}'_\xi$ the new one. If the CAV ξ is travelling with the traffic flow without affecting it ($u_\xi > \check{v}_\xi$ and $u_\xi > \check{v}'_\xi$), then according to the flux function that describes the traffic flow Q_0, the solution to the Riemann problem between these two traffic densities should include zones of traffic density $\check{\rho}_\xi \bar{\Sigma}^{\check{\rho}'_\xi}_{Q_0}$, i.e., the transition from $\check{\rho}_\xi$ to $\check{\rho}'_\xi$ should happen along an upper concave envelope of Q_0 if $\check{\rho}_\xi > \check{\rho}'_\xi$, or over a lower convex envelope of Q_0 if $\check{\rho}_\xi < \check{\rho}'_\xi$ (discussed in the Appendix in Appendix A.1). Therefore, if $\check{\rho}_\xi \bar{\Sigma}^{\check{\rho}'_\xi}_{Q_0} \neq [\check{\rho}_\xi \; \check{\rho}'_\xi]^\top$, we infer that the reason for this discrepancy is the limit on front speeds, which we use to estimate $W = (\check{W}_-, \check{W}_+, \hat{W}_-, \hat{W}_+)$, as outlined in Algorithm 3.

We denote the front speed based on these measurements $\check{w}_\xi = \frac{\check{q}'_\xi - \check{q}_\xi}{\check{\rho}'_\xi - \check{\rho}_\xi}$. If $\check{\rho}_\xi \bar{\Sigma}^{\check{\rho}'_\xi}_{Q_0} \neq [\check{\rho}_\xi \; \check{\rho}'_\xi]^\top$, this front speed offers information about one front speed bound: $\check{W}_\pm$ or $\hat{W}_\pm$ depending on whether $\check{\rho}_\xi < \check{\rho}'_\xi$ or $\check{\rho}_\xi > \check{\rho}'_\xi$, respectively, and $\check{W}_-$ or $\check{W}_+$ depending on whether $\check{w}_\xi > \check{\rho}_\xi \tilde{V}^{\check{\rho}'_\xi}_{Q_0,1}$ or $\check{w}_\xi < \check{\rho}_\xi \tilde{V}^{\check{\rho}'_\xi}_{Q_0,1}$, respectively. Finally, we calculate those front speed bounds for which we have at least one such measurement as the average over the whole set, $\tilde{W}_\pm = \mathrm{avg}\check{w}^{\check{W}_\pm}$.

3.4. Traffic state reconstruction

Finally, we use the local measurements of traffic density $\check{\rho}_\xi$, $\xi \in \Xi$, where ξ denotes the CAV that collected the measurement, to reconstruct the traffic state. We force the reconstructed traffic density in the FTSM to be equal to the measured one at the position of the CAV ξ, by setting $F_\xi = \check{\rho}_\xi$. If the flux functions of all road segments are known, the influence of these measurements spreads to the rest of the space over time.

We update the traffic state in an event-based manner at some time t' when

1. a new CAV enters the road,
2. the difference between the current traffic density measurement of some CAV $\check{\rho}'_{\xi*}$, and the previous forced traffic density $F_{\xi*}$, exceeds some bound, $|\check{\rho}'_{\xi*} - F_{\xi*}| > B_\rho$, or
3. before predicting the traffic state evolution.

If the state $\mathcal{X}$ was last updated at time t, we may calculate the current state $\mathcal{X}'$ at time t' by chaining transitions until the current time becomes t'. During this update, the CAVs are described by their forced densities F_ξ, and boundary speeds $\Lambda^\pm_\xi = \frac{x_\xi(t') - x_\xi(t)}{t' - t}$, which ensure that the positions of the CAVs in $\mathcal{X}'$ correspond to their real positions on the road $x_\xi(t')$. If the state update was caused by the entry of a new CAV, we assign a new unique identifier $\xi^* > \max \Xi$ to it, where Ξ is the set of identifiers of all CAVs that were on the road before the new arrival. The newly arrived CAV ξ^* is then added to set $\Xi' = \Xi \cup \{\xi^*\}$, and added to the model at position $x_{\xi^*}(t')$, with forced traffic density set to $F'_{\xi^*} = \check{\rho}'_{\xi^*}$. Otherwise, if the state update was caused by a large deviation of the currently measured traffic density $\check{\rho}'_{\xi*}$ from $F_{\xi*}$, we update its forced traffic density to $F'_{\xi*} = \check{\rho}'_{\xi*}$ once the state reaches $\mathcal{X}'$. The process then repeats at the next update time t''. We can influence the frequency of updates by choosing a higher B_ρ, which also decreases the sensitivity to measurement noise, but may yield a higher state reconstruction error.

3.5. Stop-and-go wave dissipation control

We use the proposed model and learning algorithms for traffic state reconstruction and model learning to tackle the control problem of stop-and-go wave dissipation. At every time instant t, we consider some number of present CAVs $\xi \in \Xi(t)$ and stop-and-go waves $\psi \in \Psi(t)$, and these numbers will change over time as new CAVs leave the road segment and new ones enter it, and as stop-and-go waves get dissipated and new ones appear. The control action of CAV ξ is its reference speed, $u_\xi(t) \in [u^{\min}, u^{\max}]$. In practice, the only restriction on the reference speed is the minimum speed $u_\xi(t) \geq u^{\min}$, as the maximum speed of the CAV will be dictated by the speed of the surrounding traffic, $\Lambda^\pm_\xi \leq v_{\xi+}$. The proposed control laws follow the idea used in Čičić et al. (2020a). Once the presence of stop-and-go waves is detected, we use the CAVs as moving bottlenecks to restrict the inflow to them, accelerating their dissipation. All control laws use a prediction model to calculate the maximum speed at which the CAVs can drive while still avoiding entering any stop-and-go waves. The predictions and reference speeds of all CAVs are calculated at every time instant and applied as control actions if they satisfy the speed restrictions. Four cases of reconstruction-based control are considered:

(1) (FI) control with full access to the information about both the traffic state and traffic model,
(2) (EM) control using the actual traffic state but without considering the variation in the traffic model,
(3) (RS) control using the actual traffic model and reconstructed traffic density, and
(4) (FT) prediction-based control using the FTSM for traffic state reconstruction, model learning, and prediction.

The first three cases are used as benchmarks for comparison, using additional information about the traffic dynamics, the current traffic state, or both.

3.5.1. Full-information control (FI)

The number of vehicles in the zone of interest, between the CAV ξ and the stop-and-go wave it is dissipating ψ_ξ, evolves according to

$$\dot{N}_\xi(t) = \omega_\xi(t) - \omega_{\psi_\xi}(t),$$
$$\omega_\xi(t) = Q_0(\underline{\rho}^{\mathrm{d}}_{-\xi}(t)) - u_\xi(t)\underline{\rho}^{\mathrm{d}}_{-\xi}(t),$$
$$\omega_{\psi_\xi}(t) = Q_0(\underline{\rho}^{\mathrm{d}}_{-\psi_\xi}) - \hat{W}_-\underline{\rho}^{\mathrm{d}}_{-\psi_\xi},$$

where $\omega_\xi(t)$ is the traffic flow overtaking the moving bottleneck, $\underline{\rho}^{\mathrm{d}}_{-\xi}(t)$ is the traffic density of the overtaking flow dependent on $u_\xi(t)$, and $\omega_{\psi_\xi}(t)$ is the discharging flow from the stop-and-go wave. Since $\hat{W}_- < 0$ is assumed to be constant, the discharging traffic density is also constant $\underline{\rho}^{\mathrm{d}}_{-\psi_\xi}$ depending only on the traffic density of the stop-and-go wave $\underline{\rho}^{\mathrm{c}}_{-\psi_\xi}$. Likewise, assuming $u_\xi(t) \leq v_{\xi+}(t)$, the distance between these two borders of the zone follows

$$\dot{d}_\xi(t) = \hat{W}_- - u_\xi(t).$$

Given the initial number of vehicles $N_\xi(t)$ and distance $d_\xi(t)$, we readily get the times after which the number of vehicles in the zone and the distance become zero,

$$\tau^N_\xi(t) = \frac{N_\xi(t)}{Q_0(\underline{\rho}^{\mathrm{d}}_{-\xi}(t)) - u_\xi(t)\underline{\rho}^{\mathrm{d}}_{-\xi}(t) - Q_0(\underline{\rho}^{\mathrm{d}}_{-\psi_\xi}) + \hat{W}_-\underline{\rho}^{\mathrm{d}}_{-\psi_\xi}}, \tag{9}$$

$$\tau^d_\xi(t) = \frac{d_\xi(t)}{\hat{W}_- - u_\xi(t)}. \tag{10}$$

In order for the stop-and-go wave ψ_ξ to be dissipated before CAV ξ reaches it, we need to ensure that $N_\xi(t_\xi) = 0$, $d_\xi(t_\xi) \geq 0$ at some time $t_\xi \geq t$, i.e., $\tau^n_\xi(t) \leq \tau^d_\xi(t)$, which is achieved with minimum disruption to the rest of the traffic if $\tau^N_\xi(t) = \tau^d_\xi(t)$. We simplify the calculation of $u_\xi(t)$ by approximating $\underline{\rho}^{\mathrm{d}}_{-\xi}(t) \approx \underline{\rho}^{\mathrm{d}}_{-\mathrm{max}}$, where $\underline{\rho}^{\mathrm{d}}_{-\mathrm{max}}$ is $\underline{\rho}^{\mathrm{d}}_{-\xi}(t)$ in case $u_\xi(t) = u^{\min}$, so substituting (9)–(10) into $\tau^N_\xi(t) = \tau^d_\xi(t)$ yields

$$u_\xi(t) = \frac{Q_0(\underline{\rho}^{\mathrm{d}}_{-\mathrm{max}}) - Q_0(\underline{\rho}^{\mathrm{d}}_{-\psi_\xi}) + \hat{W}_-\left(\underline{\rho}^{\mathrm{d}}_{-\psi_\xi} - \underline{\rho}^{\mathrm{avg}}_{-\xi}(t)\right)}{\underline{\rho}^{\mathrm{d}}_{-\mathrm{max}} - \underline{\rho}^{\mathrm{avg}}_{-\xi}(t)}, \tag{11}$$

where $\underline{\rho}^{\mathrm{avg}}_{-\xi}(t) = N_\xi(t)/d_\xi(t)$ is the average traffic density in the considered zone. If thus calculated $u_\xi(t) < u^{\min}$, we infer that CAV ξ will not be able to dissipate stop-and-go wave ψ_ξ, and instead apply $u_\xi(t) = u^{\min}$ and use the next CAV upstream to dissipate the same stop-and-go wave.

3.5.2. Estimated-model control (EM)

In the full-information case, we assumed that both the model and the current traffic density profile are fully known. In reality, the parameters of the traffic model can change due to varying weather or other conditions, and the traffic density profile has to be reconstructed using available measurements. If the true model is not known, we calculate the control action using its current estimates $\hat{Q}_0$, $\hat{Q}_\xi$, and $\hat{\hat{W}}_-$, based on which we calculate $\hat{\rho}^{\mathrm{d}}_{-\mathrm{max}}$ and $\hat{\rho}^{\mathrm{d}}_{-\psi_\xi}$,

$$u_\xi(t) = \frac{\hat{Q}_0(\hat{\rho}^{\mathrm{d}}_{-\mathrm{max}}) - \hat{Q}_0(\hat{\rho}^{\mathrm{d}}_{-\psi_\xi}) + \hat{\hat{W}}_-\left(\hat{\rho}^{\mathrm{d}}_{-\psi_\xi} - \underline{\rho}^{\mathrm{avg}}_{-\xi}(t)\right)}{\hat{\rho}^{\mathrm{d}}_{-\mathrm{max}} - \underline{\rho}^{\mathrm{avg}}_{-\xi}(t)}. \tag{12}$$

When the actual model changes, or if it deviates from the estimated one, the calculated control action can become overly conservative, leading to superfluous traffic disruption, or overly optimistic, leading to failure to dissipate stop-and-go waves.

3.5.3. Reconstructed-state control (RS)

Conversely, if the traffic density profile is not known, we use the estimated average density $\hat{\underline{\rho}}^{\mathrm{avg}}_{-\xi}(t)$ instead,

$$u_\xi(t) = \frac{Q_0(\underline{\rho}^{\mathrm{d}}_{-\mathrm{max}}) - Q_0(\underline{\rho}^{\mathrm{d}}_{-\psi_\xi}) + \hat{W}_-\left(\underline{\rho}^{\mathrm{d}}_{-\psi_\xi} - \hat{\underline{\rho}}^{\mathrm{avg}}_{-\xi}(t)\right)}{\underline{\rho}^{\mathrm{d}}_{-\mathrm{max}} - \hat{\underline{\rho}}^{\mathrm{avg}}_{-\xi}(t)}, \tag{13}$$

We use the simple traffic state reconstruction algorithm given in Čičić et al. (2020a),

$$\hat{\rho}_i(t) = \begin{cases} \rho_i(t), & i = i^h_\xi(t) \vee i = i^h_\xi(t) + 1, \xi \in \Xi(t), \\ \hat{\rho}_i(t-1) + \frac{T}{L}\left(\hat{q}_{i-1}(t-1) - \hat{q}_i(t-1)\right), & \text{elsewhere}, \end{cases}$$

where $\Xi(t)$ is the set of all CAVs that are currently present on the road, and calculate $\hat{\underline{\rho}}^{\mathrm{avg}}_{-\xi}(t)$ based on the reconstructed traffic density profile $\hat{\rho}_i(t)$. Note that since here we assume we can only measure the traffic density at the position of the CAV and immediately downstream, in this case, we have no information about the congestion that builds up upstream of the CAV, which can lead to overly optimistic predictions of the traffic state.

3.5.4. FTSM-based control (FT)

Finally, we employ the proposed FTSM as a prediction model to calculate control actions for all CAVs. Employing the traffic state reconstruction and model learning algorithms, we acquire the current FTSM state $\mathcal{X}(t)$, which includes both the traffic density and model information. The control actions $u_\xi(t)$, $\xi \in \Xi(t)$, are based on the predicted state $\hat{\mathcal{X}}(t+\theta)$ with θ ranging from 0 to the prediction horizon Θ, taken long enough that all CAVs are guaranteed to either dissipate or run into stop-and-go waves, $\Theta = N_{\text{cell}} L/(u^{\min} - \hat{W}_-)$, where $N_{\text{cell}} L$ is the length of the considered road segment. During the process of finding $\hat{\mathcal{X}}(t + \theta)$, all traffic densities are unforced $\hat{F}_\xi = \varnothing$, and the boundary speeds of all CAVs are set to

$$\hat{\Lambda}^\pm_\xi = \begin{cases} u^{\min}, & \hat{v}_{\xi+} > u_{\min} \wedge (\exists i \geq i_{\xi+})\hat{\rho}_i \geq \hat{\sigma}^* \\ \hat{v}_{\xi+}, & \text{otherwise,} \end{cases}$$

where $\hat{v}_{\xi+}$ is given by (6), $i_{\xi+}$ by (7), and $\hat{\sigma}^*$ is the minimum traffic density considered as a stop-and-go wave,

$$\hat{\sigma}^* = \begin{cases} \sigma^{\max}_{\hat{Q}_0}, & (\nexists i)\hat{V}_i < \hat{W}_-, \\ \max\left\{ \sigma^{\max}_{\hat{Q}_0}, \sigma^{\hat{W}_-}_{\hat{Q}_0} \right\}, & (\exists i)\hat{V}_i < \hat{W}_-, \end{cases}$$

$$\sigma^{\hat{W}_-}_{\hat{Q}_0} = \hat{\sigma}_i, i = \max j \text{ s.t. } \hat{V}_k > \hat{W}_-, k = 1, \dots, j.$$

Essentially, the CAVs are driven at minimum speed, applying maximum control action, until they run into congestion or there is no congestion downstream of their location.

Next, for each CAV ξ, we identify the position and prediction time $(\hat{x}^*_\xi, \hat{\tau}^*_\xi)$ at which the final zone i where $\hat{\rho}_i > \hat{\sigma}^*$ is dissipated, i.e., the congestion is dissipated and the CAV no longer needs to apply control action. The prediction time is given in relative coordinates, compared to the initial time t at which the prediction is calculated. This congestion includes the stop-and-go waves initially present in the system, as well as potential congestion created in the wake of CAVs downstream of the considered one. If the CAV ξ fails to dissipate all congestion, and instead enters congested traffic, $\hat{v}_{\xi+} < u^{\min}$, we instead set $(\hat{x}^*_\xi, \hat{\tau}^*_\xi)$ to the point at which this happened, and if there was no congestion downstream of CAV ξ at the beginning of the prediction, we set $(\hat{x}^*_\xi, \hat{\tau}^*_\xi) = (x_\xi(t), 0)$ Then, the reference speed $u_\xi(t)$ is set to

$$u_\xi(t) = \begin{cases} u^{\max}, & \hat{\tau}^*_\xi = 0, \\ \max\left\{ u^{\min}, \min\left\{ u^{\max}, \frac{\hat{x}^*_\xi + \hat{x}_{\vec{\xi}} - \hat{x}_{\vec{\xi}}(t) + u^{\min}\hat{\tau}^*_\xi}{\hat{\tau}^*_\xi} \right\} \right\}, & \hat{\tau}^*_\xi > 0, \end{cases} \tag{14}$$

where $\vec{\xi}$ denotes the first CAV downstream of CAV ξ. The control action is therefore calculated by first predicting the evolution of the system under maximum control effort, and then reducing the control effort where possible, based on the prediction.

4. Numerical evaluations

Finally, we put the control laws described in the previous section to the test in simulations. First we give the simulation model, followed by presenting the parameters of the simulation setup, and then discuss the simulation results.

4.1. Simulation model

The simulation model used to test the proposed algorithms and control is a variant of Godunov-discretized LWR (1),

$$\rho_i(t + 1) = \rho_i(t) + \frac{T}{L}\left(q_{i-1}(t) - q_i(t) \right),$$

$$q_i(t) = \min\left\{ D_i(t), S_{i+1}(t) \right\}$$

$$D_i(t) = \min\left\{ \rho_i(t), \sigma^{\max}_{Q_0} \right\} \min\left\{ U_i(t), \max\left\{ 0, \mathcal{V}\left(\min\left\{ \rho_i(t), \sigma^{\max}_{Q_0} \right\}\right) + \delta_i(t) \right\} \right\},$$

$$S_i(t) = \max\left\{ \rho_i(t), \sigma^{\max}_{Q_0} \right\} \min\left\{ \max\left\{ 0, \mathcal{V}\left(\max\left\{ \rho_i(t), \sigma^{\max}_{Q_0} \right\}\right) + \delta_{i-1}(t) \right\} \right\},$$

where $\rho_i(t)$ is the traffic density in cell $i = 1, \dots, N_{\text{cell}}$ at time t, $q_i(t)$ the traffic flow which depends on the demand of cell i, $D_i(t)$, and supply of cell $i + 1$, $S_{i+1}(t)$. The average speed of each cell is given by the speed–density relation

$$\mathcal{V}(\rho) = V^{\max}_{Q_0} \exp\left(-\frac{1}{\alpha}\left(\frac{\rho}{\sigma^{\max}_{Q_0}} \right)^\alpha \right), \tag{15}$$

with an added normally distributed noise term $\delta_i(t) \sim \mathcal{N}(0, \Delta)$ that models the stochastic nature of human driving. The flux function (fundamental diagram) is therefore given by $Q_0(\rho) = \rho\mathcal{V}(\rho)$. The inflow to the first cell $q_0(t)$ is given as an external input, and can be delayed if the conditions in cell 1 are such that they cannot admit such flow, and the outflow from the final cell $q_{N_{\text{cell}}}(t)$ will be

limited by extraneously setting $S_{N_{\text{cell}}+1}(t)$ in order to generate stop-and-go waves entering the road segment from downstream. We use reference speed $U_i(t)$, to model the influence of moving bottlenecks and stop-and-go waves, as in Čičić et al. (2020a),

$$U_i(t) = V_{Q_0}^{\max} \min \left\{ 1, \max \left\{ 0, \frac{\rho_{i+1}^*(t) - \frac{V_{Q_0}^{\max} - U_{i+1}(t)}{V_{Q_0}^{\max}} \rho_{i+1}(t)}{\rho_i(t)} \right\} \right\},$$

where $U_{N_{\text{cell}}+1} = V_{Q_0}^{\max}$ and $\rho_i^*(t)$ is the reference density.

The reference density $\rho_i^*(t)$ is set to an arbitrary value higher than the maximum traffic density in the simulation that ensures $U_i(t) = V_{Q_0}^{\max}$ everywhere except in the vicinity of moving bottlenecks and stop-and-go waves. There, we use the reference density of the form

$$\rho_i^*(t) = \begin{cases} \underline{\rho}_k^c, & i = i_k^t(t), \ldots, i_k^h(t), \\ \underline{\rho}_k^d + \left(\underline{\rho}_k^c - \underline{\rho}_k^d \right) \frac{z_k(t) - i_k^h(t)L}{L}, & i = i_k^h(t), \\ \underline{\rho}_k^d, & i = i_k^h(t) + 1, \end{cases}$$

to model their influence on the rest of the traffic.

We model the influence a moving bottleneck ξ, whose position is $z_\xi(t)$ and velocity $u_\xi(t)$, by setting $i_\xi^h(t) = \lfloor z_\xi(t)/L \rfloor + 1$ ($i_\xi^h(t)$ is the cell where the moving bottleneck is), $i_\xi^t(t) = i_\xi^h(t) - 1$. If the moving bottleneck is slower than the traffic surrounding it, $u_\xi(t) < q_{i_\xi^h}(t)/\rho_{i_\xi^h}(t)$, we model its influence by setting a reduced reference density in its vicinity. Densities $r_\xi^d(t)$ and $r_\xi^c(t)$ are set to the solutions of $Q_0(\rho) = Q_{\text{sc}}(\rho^\tau) + u_\xi(t)(\rho - \rho^\tau)$, $Q'_{\text{sc}}(\rho^\tau) = u_\xi(t)$, $r_\xi^d(t) < r_\xi^c(t)$, i.e., the intersections between the flux function $Q_0(\rho)$ and the tangent of the scaled flux function $Q_{\text{sc}}(\rho)$ with slope $u_\xi(t)$ (Monache and Goatin., 2014). The flux function $Q_{\text{sc}}(\rho)$ is scaled down by the ratio of lanes not occupied by the moving bottleneck and the total number of lanes. The position of the moving bottlenecks is updated to $z_\xi(t+1) = z_\xi(t) + T \min \left\{ q_{i_\xi^h}(t)/\rho_{i_\xi^h}(t), u_\xi(t) \right\}$.

A stop-and-go wave k whose downstream end is at $z_k(t)$, is modelled by setting $i_k^h(t) = \lfloor z_k(t)/L \rfloor + 1$, and $i_k^t(t)$ is the cell where upstream end of its congestion is. We set $\underline{\rho}_k^c$ to the maximum congestion level in the wave, and if $Q'_0(\underline{\rho}_k^c) < \hat{W}_-$, where $\hat{W}_-$ is the constant negative wave-speed propagation bound, we reduce the reference in its area. The discharging density $\underline{\rho}_k^d$ is then given as $Q_0(\underline{\rho}_k^d) = Q_0(\underline{\rho}_k^c) + \hat{W}_-(\underline{\rho}_k^d - \underline{\rho}_k^c)$, $\underline{\rho}_k^d < \underline{\rho}_k^c$, and the position of the downstream end of the stop-and-go propagates backwards to $z_k(t+1) = z_k(t) + T\hat{W}_-$. Once the level of congestion of the stop-and-go wave falls to the level where $Q'_0(\underline{\rho}_k^c) \geq \hat{W}_-$, the stop-and-go wave ceases to exist as a separate phenomenon, and if at some part of the road sufficiently dense congestion arises, a new stop-and-go wave is created.

4.2. Simulation setup

The simulations were executed on a 10 km stretch of a two-lane road, which consists of $N_{\text{cell}} = 100$ cells of length $L = 100$ m each. The simulation length is taken to be $t_{\text{end}} = 3$ h, and the simulation time step is $T = 3$ s. The inflow to the road segment randomly varies in time, changing every minute, and is uniformly distributed $q_0(t) \sim \mathcal{U}(1450, 4350)$ veh/h. The CAVs arrive with random time gaps between them, $g_\xi = \max\{\gamma_\xi, 30\}$ s, $\gamma_\xi \sim \text{Exp}(\frac{1}{30})$. We create the stop-and-go waves, arriving from downstream, by limiting the maximum outflow at the downstream end of the road segment to $\mathcal{U}(200, 400)$ veh/h for 30 s, causing congestion to build up. Once this restriction is removed, the congestion starts dissipating, with discharging flow lower than the capacity of the road, propagating upstream as a stop-and-go wave. The time gap between two generated stop-and-go waves is uniformly distributed $q_k \sim \mathcal{U}(360, 1080)$ s (6 to 18 minutes).

In order to demonstrate how control adapts to time-varying traffic model, e.g., due to a change in weather conditions when it starts raining (Billot et al., 2009), we change the flux function at time $t = 0.5$ h. Both flux function are given by $Q_0(\rho) = \rho \mathcal{V}(\rho)$, (15), with parameters $\alpha = 2.34$, $V_{Q_0}^{\max} = 120$ km/h, $\sigma_{Q_0}^{\max} = 51.1$ veh/km, yielding capacity $q^{\max} = 4000$ veh/h, for $t \leq 0.5$ h, and $\alpha = 3.4$, $V_{Q_0}^{\max} = 75$ km/h, $\sigma_{Q_0}^{\max} = 60.4$ veh/km, yielding capacity $q^{\max} = 3375$ veh/h, for $t > 0.5$ h. The variance of the additive process noise of the velocity is in both cases $\Delta = 16$ km^2/h^2.

The FTSM state reconstruction and model learning algorithms used traffic flow bound $B_q = 1000$, flux function breakpoint bound $B_\sigma = 20$, bound on the maximum number of used data points $B_{\tilde{N}} = 10$, overtaking flow bound $B_\omega = 500$, and traffic density bound $B_\rho = 30$. Furthermore, all of the measurements were perturbed with multiplicative noise $\varepsilon \sim \mathcal{N}(1, 0.0025)$.

4.3. Simulation results

We executed 50 simulation runs, comparing the performance of the four described control laws. For each simulation run, six simulations with the same realization of all random variables were executed, one using each of the control laws, one with no control, and a benchmark simulation with no stop-and-go waves. The performance metric used was the delay ratio, defined as the percentage increase of Total Time Spent (TTS) compared to the TTS of the benchmark simulation, $\text{DR}_i^{\text{case}} = \frac{\text{TTS}_i^{\text{case}} - \text{TTS}^{\min}}{\text{TTS}^{\min}} 100\%$. For
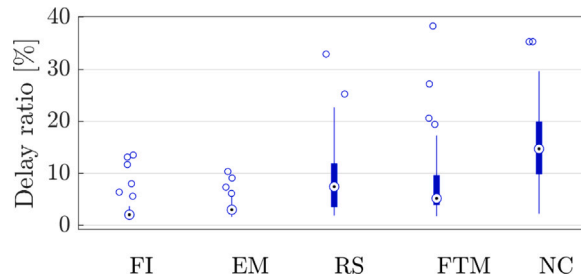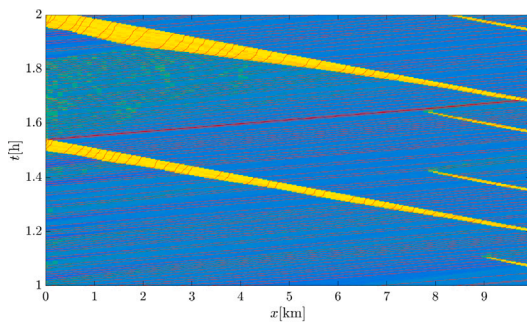
Fig. 6. Box plots of delay ratios with different control laws. FI: Full-information (11), EM: Estimated-model (12), RS: Reconstructed-state (13), FT: FTSM-based (14), NC: No control.
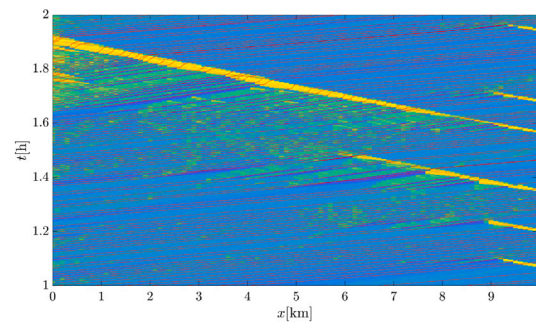
Table 2
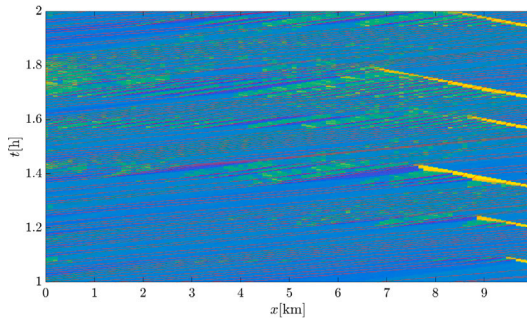Achieved mean and median delay ratios of the different control laws.

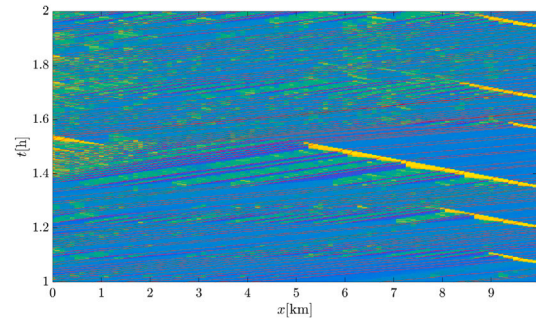| Delay ratio [%]	FI	EM	RS	FT	NC
Mean	2.92	3.44	9.44	8.02	15.69
Median	2.06	3.02	7.45	5.20	14.73

(a) No control

(b) Reconstructed-state control

(c) Full-information control

(d) FTSM-based control

Fig. 7. Detail from the simulations comparing different control cases.

example, $DR^{case} = 15\%$ indicates that the TTS of that case of simulation run i was $TTS_i^{case} = 1.15 \cdot TTS^{min}$, i.e., the stop-and-go waves caused the vehicles to spend 15% more time on the road in total. The results are shown in Fig. 6 and Table 2.

We can see that all control cases achieve significant reduction of delay compared to the uncontrolled case. In particular, control laws that use full information about the traffic density profile (FI and EM) perform significantly better than those based on state reconstruction (RS and FT). This is not surprising, since in this case, the control law is able to react as soon as a stop-and-go wave appears on the road, instead of waiting for one CAV to detect it first. In this case, not knowing the exact model only deteriorated the control performance slightly (in case of EM compared to FI), since the change in model parameters was not severe.

Reconstructed-state control performs worse than the other control laws, probably due to the fact that it does not correctly capture the buildup of congestion in the wake of CAVs. Conversely, since the traffic state reconstruction is done model-based in case of the FTSM-based control, this control law achieves better performance, in spite of using less information. Overall, in spite of being fully data-driven, this control law is able to dissipate stop-and-go waves and significantly reduce the overall Total Time Spent.
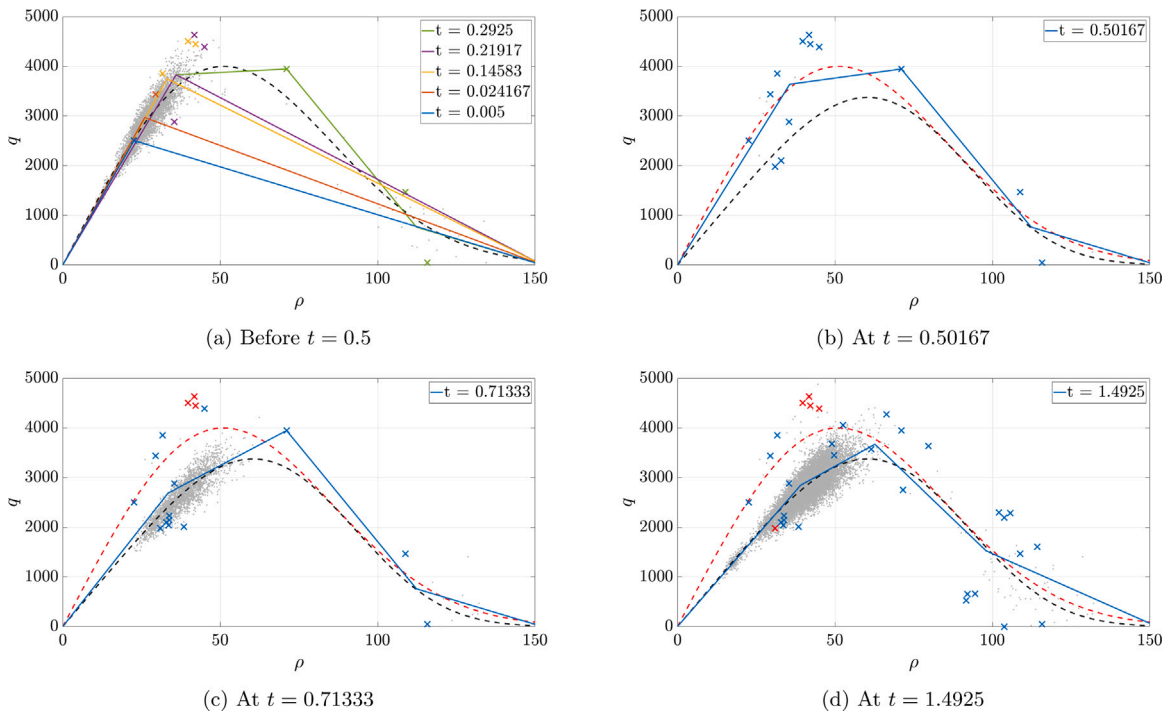
(a) Before $t = 0.5$

(b) At $t = 0.50167$

(c) At $t = 0.71333$

(d) At $t = 1.4925$

Fig. 8. Overview of the changes to $\hat{Q}_0$ through time in one simulation run. In 8(a), some updates before $t = 0.5$ when Q_0 is changed are shown. Density-flow measurement pairs for $t < 0.5$ are shown as grey dots and the current Q_0 is shown in dashed black. The estimation of $\hat{Q}_0$ is done based on measurements $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$, shown as ×-es, and newly added ×-es are shown in the same colour as the current $\hat{Q}_0$. In 8(b), 8(c), and 8(d), $\hat{Q}_0$ is shown at three different times $t > 0.5$. Here, only the measurements from $t = 0.5$ to the time when $\hat{Q}_0$ was estimated are shown. The current Q_0 is shown in dashed black and the old Q_0 in dashed red. Blue ×-es are the measurements used to estimate $\hat{Q}_0$, and red ×-es are the "forgotten" measurements, that were used at some previous time, but have since been removed from $(\check{\rho}^{Q_0}, \check{q}^{Q_0})$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In Fig. 7 we show a detail from one of the simulations. In case no control is applied, stop-and-go waves grow and propagate upstream. In this simulation run, the reconstructed-state control failed to dissipate the stop-and-go wave, and the other control laws were successful, albeit with the FTSM-based control causing some more congestion further upstream.

The traffic flux function Q_0 is the most impactful component of the overall traffic prediction model, so the process of estimating it using Algorithm 1 is shown in Fig. 8. It can be seen that the first estimates of Q_0 have a single breakpoint and are triangular, because all available measurements are still tightly clustered at that time. Later, as measurements of congested traffic become available, more breakpoints are added and the flux function takes a somewhat more complex shape. Once the flux function is changed at $t = 0.5$, the algorithm gradually adapts the estimate to the new flux functions and either offsets or eliminates the measurements that are the worst outliers. Even though the traffic speed is modelled with significant process noise, and the measurements used are also noisy, the learning algorithms are able to quickly estimate the flux function using a low number of memorized measurements. Moreover, even though the piecewise linear flux function deviates from the form of the underlying actual flux function, this deviation does not reflect particularly negatively on the control performance, while adopting such a simple form of the flux function greatly increases the numerical efficiency of the prediction model.

5. Conclusions

In this work we presented a cell-free traffic modelling approach based on front-tracking. We showed that the FTSM can capture discharging flows from stop-and-go waves lower than the capacity of the road, while also being able to model moving bottlenecks and other traffic phenomena. The solution of the FTSM was shown to be existing, unique, non-Zeno, and corresponding to the solution of the LWR model with space-dependent piecewise linear flux functions and piecewise constant initial state. In order to be able to model stop-and-go waves with discharging flow lower than the road capacity, a new type of weak solution for the LWR was introduced, the wave-speed-bounded solution, which constrains the propagation speed of fronts to some range, while only violating the entropy condition at fronts where this constraint is active. We also described how traffic state reconstruction using local traffic measurements from the CAVs can be implemented, as well as presented simple algorithms for model learning based on the same measurements, leveraging the structure of the model. The proposed algorithms were shown to be able to approximate a more general flux function with a piecewise linear one, only relying on a small dynamically changing subset of the collected traffic measurements. Using the FTSM as a prediction model, we designed a control law for stop-and-go wave dissipation using CAVs, and compared it

with other similar control laws. The control laws were tested in simulations, and were shown to significantly reduce the Total Time Spent by eliminating stop-and-go waves.

There are many possible directions for further extending this work. Firstly, a more rigorous validation of the FTSM needs to be performed, covering a wide variety of traffic conditions and infrastructure configurations. Next, while in this work we assumed that the considered road segment was homogeneous, the FTSM can easily be extended to the more general case. By introducing rule-based transitions, we are able to define exactly what happens when fronts at the boundary of different flux functions collide, which can be used to model the transition of specific vehicles through zones with different road geometry, and consequently also different flux functions. The general structure of the FTSM allows modelling complex scenarios with dynamically changing zones where the traffic is described by different flux functions. The modular nature of the model also allows for straightforward extensions, by specifying additional transitions or modifying the described transitions. For example, we may modify the Boundary Riemann transition to introduce on- and off-ramps and more complex road network topologies into the model, as well as model the capacity drop phenomenon at bottlenecks. Furthermore, these flux functions can be used to describe the aggregate effect of the CAVs on the traffic in some area, allowing us to capture more complex interactions and formulate different control laws. Finally, an in-depth study on the nature of solution that arise from introducing wave-speed bounds and zones described by different flux functions to the LWR model needs to be conducted. In particular, since these modifications can cause the total variation of the solution to increase, compared to that of the initial conditions, bounds on this increase need to be derived.

In this work we presented some basic model learning algorithms, designed to be simple and straightforward, but more efficient algorithms designed for specific purposes should also be developed. The proposed prediction-based traffic control based on the FTSM should also be applied to more general simulation setups, including in microscopic traffic simulations. Finally, for practical applications where the exactness of the calculated solution is not of critical importance, approximation schemes that will reduce the computation burden need to be implemented. For example, fronts that separate zones where the same flux function holds and the traffic densities do not differ too much could be removed, and flux functions could be simplified by reducing the number of breakpoints that define them. We caution that care needs to be taken so that these approximations do not compromise non-Zenoness of the transition system.

CRediT authorship contribution statement

Mladen Čičić: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Karl Henrik Johansson:** Supervision, Writing – review & editing.

Acknowledgments

Appendix A. Front-tracking with zones of different flux function and wave-speed bounds

The entropy solution to the LWR model (1), with flux functions (3), and initial conditions (2) is of the form

$$\rho(x,t) = \begin{cases} \rho'_1, & x < x_1^{\rho'} + \lambda_1 t, \\ \vdots & \\ \rho'_i, & x_{i-1}^{\rho'} + \lambda_{i-1}t < x < x_i^{\rho'} + \lambda_i t, \\ \vdots & \\ \rho'_{N'+1}, & x > x_{N'}^{\rho'} + \lambda_{N'}t, \end{cases} \tag{16}$$

with $\lambda_{i-1} \le \lambda_i$ wherever $x_{i-1}^{\rho'} = x_i^{\rho'}$. Here λ_i, $i = 1, \ldots, N'$ are the transition speeds, defined by the Rankine–Hugoniot condition

$$\lambda_i = \frac{Q(\rho'_{i+1}) - Q(\rho'_i)}{\rho'_{i+1} - \rho'_i}.$$

The solution consists of zones of constant density separated by fronts $x_i^{\rho'} + \lambda_i t$ where we have a discontinuity in the density. This solution holds for $t \in [0, \tau]$, where τ is the minimum time when two fronts collide, $x_{i-1}^{\rho'} + \lambda_{i-1}\tau = x_i^{\rho'} + \lambda_i\tau$, with $\lambda_{i-1} > \lambda_i$. To get the solution after that time, we solve a new composite Riemann problem for initial conditions $\rho(\tau, x)$, and by iterating this step, we can obtain an exact entropy solution $\rho(x,t)$ to the initial value problem (1), (2) for any t. Due to Corollary 2.8 from Holden and Henrik Risebro (2015), the front-tracking method yields *exact* entropy solutions in case when the flux function is continuous and piecewise linear and initial conditions piecewise constant, which is the case we consider here.

Since $\rho(x,0)$ is piecewise constant, we have $\partial_x Q(\rho(x,t),x,t) = 0$ everywhere except at discontinuities of $\rho(x,t)$, and at discontinuities of $Q(\rho,x,t)$ in x. Therefore, the process of finding (16) can be decomposed into finding the solutions to the Riemann problems at each discontinuity. A Riemann problem, is an initial value problem of (1) with initial conditions

$$\rho(x,0) = \begin{cases} \rho_-, & x < 0, \\ \rho_+, & x > 0, \end{cases} \tag{17}$$

and possibly different flux functions on either side of the discontinuity, with the boundary moving at a predefined constant speed Λ,

$$Q(\rho,x,t) = \begin{cases} Q_-(\rho), & x < \Lambda t, \\ Q_+(\rho), & x > \Lambda t, \end{cases} \tag{18}$$

where Λ is given as a parameter of the problem. Note that the space coordinate is shifted so that the discontinuity is at $x = 0$.

If the discontinuity is only in $\rho(x,t)$, and $Q_-(\rho) = Q_+(\rho)$, this Riemann problem corresponds to the basic form of finding the entropy solution to the LWR model, which is well-known in the literature. In the following, we first discuss this case, and then extend the results by defining a new type of weak solution, the wave-speed-bounded solution, for which we impose bounds on the speed of fronts that originate from the discontinuity. Note that this solution can violate the entropy condition, but it allows us to model some additional traffic phenomena. Finally, we discuss the solutions at the boundary between two different flux functions, when $Q_-(\rho) \neq Q_+(\rho)$.

A.1. Entropy solution

When finding the entropy solution to Riemann problems (1), (17), (18), we need to calculate *lower convex envelope* or *upper concave envelope* of the flux function if $\rho_- < \rho_+$ or $\rho_- > \rho_+$, respectively. We define these envelopes

$$^{\rho_-}\tilde{Q}_Q^{\rho_+}(\rho) = \begin{cases} ^{\rho_-}\breve{Q}_Q^{\rho_+}(\rho), & \rho_- < \rho_+, \\ ^{\rho_-}\hat{Q}_Q^{\rho_+}(\rho), & \rho_- > \rho_+, \end{cases}$$

$$^{\rho_-}\breve{Q}_Q^{\rho_+}(\rho) = \sup\left\{ Q_*(\rho) : q(\rho) \leq Q(\rho),\ Q_* \text{ convex}, \rho \in [\rho_-, \rho_+] \right\},$$

$$^{\rho_-}\hat{Q}_Q^{\rho_+}(\rho) = \inf\left\{ Q_*(\rho) : q(\rho) \geq Q(\rho),\ Q_* \text{ concave}, \rho \in [\rho_+, \rho_-] \right\},$$

on $[\rho_{\min}, \rho_{\max}]$, $\rho_{\min} = \min(\rho_-, \rho_+)$, $\rho_{\max} = \max(\rho_-, \rho_+)$. We follow this notation in further text, with $\sim$ signifying $\breve{}$ (lower convex envelope) if $\rho_- < \rho_+$, or $\hat{}$ (upper concave envelope) if $\rho_- > \rho_+$.

Note that $^{\rho_-}\tilde{Q}_Q^{\rho_+}(\rho)$ also is a continuous piecewise linear function on $[\rho_{\min}, \rho_{\max}]$ and it can be defined in a similar way as (4),

$$\tilde{Q}(\rho) = \begin{cases} Q(\tilde{\sigma}_1) + \tilde{V}_1(\rho - \tilde{\sigma}_1), & \tilde{\sigma}_1 \leq \rho \leq \tilde{\sigma}_2, \\ Q(\tilde{\sigma}_2) + \tilde{V}_2(\rho - \tilde{\sigma}_2), & \tilde{\sigma}_2 \leq \rho \leq \tilde{\sigma}_3, \\ \quad\vdots \\ Q(\tilde{\sigma}_{\tilde{m}-1}) + \tilde{V}_{\tilde{m}-1}(\rho - \tilde{\sigma}_{\tilde{m}-1}), & \tilde{\sigma}_{\tilde{m}-1} \leq \rho \leq \tilde{\sigma}_{\tilde{m}}, \end{cases}$$

omitting superscript ρ_- and ρ_+ and subscript Q for better readability, and defining $\tilde{\sigma}_1 = \rho_{\min}$, $\tilde{\sigma}_{\tilde{m}} = \rho_{\max}$. We write the column vector of slopes of such function $^{\rho_-}\tilde{V}_Q^{\rho_+}$, ordered from $\tilde{V}_1$ to $\tilde{V}_{\tilde{m}-1}$ for $\rho_- < \rho_+$ or from $\tilde{V}_{\tilde{m}-1}$ to $\tilde{V}_1$ for $\rho_- > \rho_+$. For all breakpoints of $\tilde{Q}$, we have $\tilde{Q}(\tilde{\sigma}_i) = Q(\tilde{\sigma}_i)$. Breakpoints $\tilde{\sigma}_i$, $i = 2,\ldots,\tilde{m}-1$ are also breakpoints of Q, σ_j for some j, on $[\rho_{\min}, \rho_{\max}]$, and can be determined using efficient convex hull algorithms. Finally, we denote the sorted (ascending if $\rho_- < \rho_+$ and descending if $\rho_- > \rho_+$) column vector of elements of $^{\rho_-}\tilde{\Sigma}_Q^{\rho_+}$, including ρ_- and ρ_+, as $^{\rho_-}\tilde{\Sigma}_Q^{\rho_+}$, and its length as $^{\rho_-}\tilde{m}_Q^{\rho_+}$. Same as with the envelopes $^{\rho_-}\tilde{Q}_Q^{\rho_+}(\rho)$, $^{\rho_-}\tilde{\Sigma}_Q^{\rho_+}$ will consist of breakpoints of the lower convex or upper concave envelope, depending on whether ρ_- or ρ_+ is larger.

The solution to the Riemann problem is then given by

$$\rho(x,t) = \begin{cases} \rho_-, & x < \tilde{V}_1 t, \\ \tilde{\sigma}_2, & \tilde{V}_1 t < x < \tilde{V}_2 t, \\ \quad\vdots \\ \tilde{\sigma}_{\tilde{m}-1}, & e_{\tilde{m}}^\top \tilde{V}_{\tilde{m}-2} t < x < \tilde{V}_{\tilde{m}-1} t, \\ \rho_+, & x > \tilde{V}_{\tilde{m}-1} t, \end{cases} \tag{19}$$

again omitting superscript symbols. Since $^{\rho_-}\tilde{\Sigma}_Q^{\rho_+}$ is monotonous, the Total Variation of the solution is equal to the Total Variation of the initial conditions, $\text{T.V.}(^{\rho_-}\tilde{\Sigma}_Q^{\rho_+}) = \text{T.V.}([\rho_- \ \rho_+])$.

A.2. Wave-speed-bounded solution

Let $[\breve{W}_-, \breve{W}_+]$ and $[\hat{W}_-, \hat{W}_+]$ be the admissible ranges of front speeds for compression ($\rho_- < \rho_+$) and rarefaction ($\rho_- > \rho_+$), respectively. In case $\breve{W}_- \leq \min\{\boldsymbol{V}_Q\} \geq \hat{W}_-$ and $\breve{W}_+ \geq \max\{\boldsymbol{V}_Q\} \leq \hat{W}_+$, the entropy solution never violates the wave-speed bounds,

and it thus coincides with the wave-speed-bounded solution. Otherwise, we define $\tilde{V}_- = \min {}^{\rho-}\tilde{V}_Q^{\rho+}$ and $\tilde{V}_+ = \max {}^{\rho-}\tilde{V}_Q^{\rho+}$, and if $\tilde{V}_- < \check{W}_-$ or $\tilde{V}_+ > \check{W}_+$, we need to look for a solution that fulfils the front speed constraints, while minimally violating the entropy condition. The wave-speed-bounded solution is again given by (19), with

$$
{}^{\rho-}\tilde{\Sigma}_{Q,W}^{\rho+} = \begin{cases}
\left[\rho_- \mid \bar{\rho}_-^{W}\, \tilde{\Sigma}_Q^{\bar{\rho}_+^{W\mathsf{T}}} \mid \rho_+\right]^{\mathsf{T}}, & \tilde{V}_- < \check{W}_-, \tilde{V}_+ > \check{W}_+ \\[4pt]
\left[\rho_- \mid \bar{\rho}_-^{W}\, \tilde{\Sigma}_Q^{\mathsf{T}}\right]^{\mathsf{T}}, & \tilde{V}_- < \check{W}_-, \tilde{V}_+ \le \check{W}_+ \\[4pt]
\left[{}^{\rho-}\tilde{\Sigma}_Q^{\bar{\rho}_+^{W\mathsf{T}}} \mid \rho_+\right]^{\mathsf{T}}, & \tilde{V}_- \ge \check{W}_-, \tilde{V}_+ > \check{W}_+ \\[4pt]
{}^{\rho-}\tilde{\Sigma}_Q^{\rho+}, & \tilde{V}_- \ge \check{W}_-, \tilde{V}_+ \le \check{W}_+
\end{cases}
\qquad
{}^{\rho-}\tilde{V}_{Q,W}^{\rho+} = \begin{cases}
\left[\check{W}_- \mid \bar{\rho}_-^{W}\, \tilde{V}_Q^{\bar{\rho}_+^{W\mathsf{T}}} \mid \check{W}_+\right]^{\mathsf{T}}, & \tilde{V}_- < \check{W}_-, \tilde{V}_+ > \check{W}_+ \\[4pt]
\left[\check{W}_- \mid \bar{\rho}_-^{W}\, \tilde{V}_Q^{\mathsf{T}}\right]^{\mathsf{T}}, & \tilde{V}_- < \check{W}_-, \tilde{V}_+ \le \check{W}_+ \\[4pt]
\left[{}^{\rho-}\tilde{V}_Q^{\bar{\rho}_+^{W\mathsf{T}}} \mid \check{W}_+\right]^{\mathsf{T}}, & \tilde{V}_- \ge \check{W}_-, \tilde{V}_+ > \check{W}_+ \\[4pt]
{}^{\rho-}\tilde{V}_Q^{\rho+}, & \tilde{V}_- \ge \check{W}_-, \tilde{V}_+ \le \check{W}_+
\end{cases}
$$

where we define the densities

$$
\hat{\rho}_-^{W} = \max \rho \ \text{s.t.} \ \rho < \rho_-, \frac{\mathrm{d}Q(\rho)}{\mathrm{d}\rho} \ge \hat{W}_-, \frac{Q(\rho)-Q(\rho_-)}{\rho-\rho_-} = \check{W}_-,
\qquad
\hat{\rho}_+^{W} = \min \rho \ \text{s.t.} \ \rho > \rho_+, \frac{\mathrm{d}Q(\rho)}{\mathrm{d}\rho} \le \hat{W}_+, \frac{Q(\rho_+)-Q(\rho)}{\rho_+-\rho} = \check{W}_+,
$$

$$
\check{\rho}_-^{W} = \min \rho \ \text{s.t.} \ \rho > \rho_-, \frac{\mathrm{d}Q(\rho)}{\mathrm{d}\rho} \ge \hat{W}_-, \frac{Q(\rho)-Q(\rho_-)}{\rho-\rho_-} = \check{W}_-,
\qquad
\check{\rho}_+^{W} = \max \rho \ \text{s.t.} \ \rho < \rho_+, \frac{\mathrm{d}Q(\rho)}{\mathrm{d}\rho} \le \hat{W}_+, \frac{Q(\rho_+)-Q(\rho)}{\rho_+-\rho} = \check{W}_+.
$$

Note that the resulting solution can be non-monotone in density, and therefore cause an increase of Total Variation from the initial conditions, since $\tilde{\rho}_-^{W}$ and $\tilde{\rho}_+^{W}$ might lie outside of $\left[\min\{\rho_-, \rho_+\}, \max\{\rho_-, \rho_+\}\right]$, or might be differently ordered than ρ_- and ρ_+. The entropy condition is only violated at the discontinuities between ρ_- and $\tilde{\rho}_-^{W}$, and ρ_+ and $\tilde{\rho}_+^{W}$, and the potential increase in Total Variation of the solution, compared to the initial condition, is

$$
\text{T.V.}({}^{\rho-}\tilde{\Sigma}_{Q,W}^{\rho+}) - \text{T.V.}([\rho_- \ \rho_+]^{\mathsf{T}}) = \begin{cases}
|\tilde{\rho}_-^{W} - \rho_-| + |\tilde{\rho}_+^{W} - \tilde{\rho}_-^{W}| + |\rho_+ - \tilde{\rho}_+^{W}| - |\rho_+ - \rho_-|, & \tilde{V}_- < \check{W}_-, \tilde{V}_+ > \check{W}_+, \\[4pt]
|\tilde{\rho}_-^{W} - \rho_-| + |\rho_+ - \tilde{\rho}_-^{W}| - |\rho_+ - \rho_-|, & \tilde{V}_- < \check{W}_-, \tilde{V}_+ \le \check{W}_+, \\[4pt]
|\tilde{\rho}_+^{W} - \rho_-| + |\rho_+ - \tilde{\rho}_+^{W}| - |\rho_+ - \rho_-|, & \tilde{V}_- \ge \check{W}_-, \tilde{V}_+ > \check{W}_+, \\[4pt]
0, & \tilde{V}_- \ge \check{W}_-, \tilde{V}_+ \le \check{W}_+.
\end{cases}
$$

In order to model stop-and-go waves with discharging flow lower than road capacity, it is enough to use $\hat{W}_- > \min\{V_Q\}$. If the considered flux function is concave, as is typically the case in practice, only $\hat{W}_-$ and $\check{W}_+$ constrain the solution.

A.3. Flux function boundary solution

Stationary and moving bottlenecks are modelled by using different flux functions in different regions, and we study the evolution of traffic conditions around a bottleneck by solving a Riemann problem (1), (17), (18), with $Q_-(\rho) \ne Q_+(\rho)$. In order to find the weak solution, the propagation speed of the boundary between two flux functions Λ needs to be defined first. Then, we find the solution that maximizes the flow over the boundary between two regions, while satisfying the Rankine–Hugoniot condition across the boundary, and yielding entropy solutions inside both regions. The solution with initial conditions (17) and flux function (18) can thus be split into two parts, consisting of Riemann problems between ρ_- and ρ'_-, and between ρ'_+ and ρ_+. The solutions to these problems both need to be constrained to only have discontinuities inside the zones of Q_- $(x < \Lambda t)$ and Q_+ $(x > \Lambda t)$, respectively. The Rankine–Hugoniot condition at the boundary between the two flux functions

$$
Q_+(\rho'_+) - Q_-(\rho'_-) = \Lambda\left(\rho'_+ - \rho'_-\right),
$$

can be rewritten in terms of the flow over the boundary ω,

$$
\omega = Q_-(\rho'_-) - \Lambda\rho'_- = Q_+(\rho'_+) - \Lambda\rho'_+.
$$

Boundary densities ρ'_- and ρ'_+ both depend on ρ_-, ρ_+, Q_- and Q_+, and are given as optimizers of the optimization problem

$$
\begin{aligned}
\underset{\rho'_-, \rho'_+}{\text{maximize}} \quad & \omega \\
\text{s.t.} \quad & Q_+(\rho'_+) - Q_-(\rho'_-) = \Lambda(\rho'_+ - \rho'_-), \\
& {}^{\rho-}\tilde{V}_{Q_-,W}^{\rho'_-} < \Lambda, \\
& {}^{\rho'_+}\tilde{V}_{Q_+,W}^{\rho_+} > \Lambda,
\end{aligned}
\tag{20}
$$

so that the flow over the boundary ω is maximized, under specified constraints.

For most simple flux functions used in practice, solving this maximization problem can be done explicitly. Since optimal ρ'_- and ρ'_+ will always be such that either $\rho'_- \in \Sigma_{Q_-} \cup \{\rho_-\}$ or $\rho'_+ \in \Sigma_{Q_+} \cup \{\rho_+\}$, the problem can be solved by forming a set of all possible pairings of (ρ'_-, ρ'_+) that satisfy the Rankine–Hugoniot condition, and then checking the second and third constraint for each of them, in order of descending boundary overtaking flow, so that the first pair to satisfy these constraints is the optimizer.

Another type of boundary that we consider is the one that arises when we force the density on one side to be equal to some externally defined value, $\rho(x,t) = F_-, \ x < \lambda t$ or $\rho(x,t) = F_+, \ x > \lambda t$. We write $F_\pm \ne \varnothing$ for those sides where the density is forced,

and $F_\pm = \varnothing$ where it is not forced. In this case, the dynamics of traffic on the forced side are ignored, and the forced traffic density instead acts as a boundary condition for the other side, and the solution is given as

$$
\begin{aligned}
&\underset{\rho'_+}{\text{minimize}} && \left|\rho'_+ - F_-\right| \\
&\text{s.t.} && Q_+(\rho'_+) - Q_+(F_-) = \Lambda(\rho'_+ - F_-), \\
& && \rho'_+ \tilde{V}^{\rho'_+}_{Q_+, W} > \Lambda,
\end{aligned}
$$

if $F_- \neq \varnothing$ and $F_+ = \varnothing$, i.e., the density upstream is forced, or

$$
\begin{aligned}
&\underset{\rho'_-}{\text{minimize}} && \left|\rho'_- - F_+\right| \\
&\text{s.t.} && Q_-(F_+) - Q_-(\rho'_-) = \Lambda(F_+ - \rho'_-), \\
& && \rho_- \tilde{V}^{\rho'_-}_{Q_-, W} < \Lambda,
\end{aligned}
$$

if $F_- = \varnothing$ and $F_+ \neq \varnothing$ and the density downstream is forced. Alternatively, we may define special rules for handling the behaviour at the boundary between two flux functions, e.g., when a moving bottleneck moves from a zone of one capacity to a zone of different capacity.

The potential increase in Total Variation of the solution, compared to the initial conditions, is given by

$$
\text{T.V.}([{}^{\rho_-}\tilde{\Sigma}^{\rho'_-\top}_{Q_-,W} {}^{\rho'_+}\tilde{\Sigma}^{\rho_+\top}_{Q_+,W}]) - \text{T.V.}([\rho_- \ \rho_+]^\top) = \begin{cases} |\rho'_- - \rho_-| + |\rho'_+ - \rho'_-| + |\rho_+ - \rho'_+| - |\rho_+ - \rho_-|, & F_- = \varnothing, F_+ = \varnothing, \\ |\rho'_+ - F_-| + |\rho_+ - \rho'_+| - |\rho_+ - F_-|, & F_- \neq \varnothing, F_+ = \varnothing, \\ |\rho'_- - \rho_-| + |F_+ - \rho'_-| - |F_+ - \rho_-|, & F_- = \varnothing, F_+ \neq \varnothing. \end{cases}
$$

Appendix B. FTSM implementation

In this section, we give the details of the implementation of the FTSM, defining how the particular transitions change the current state $\mathcal{X}$. We first define the notion of admissible state, which is used to define guard sets, and then specify the transitions.

B.1. Admissible states

Since the passage of time transition constitutes the flow dynamics of the system, we call the states for which this transition can be taken the *admissible states*. Depending on the initial state X_0, we can be forced to take some number of other transitions before the current state enters the set of admissible states. This set is defined by the following conditions:

$$
\left(z_{j+1} > z_j\right) \vee \left(\left(z_{j+1} = z_j\right) \wedge \left(\lambda_j \leq \lambda_{j+1}\right)\right), \ j = 1, \ldots, n-1, \tag{$\star-$}
$$

$$
\left(\rho_j \neq \rho_{j+1}, {}^{\rho_j}\tilde{\Sigma}^{\rho_{j+1}}_{Q_j,W} = [\rho_j \ \rho_{j+1}]^\top\right), \ \text{if } Q_j = Q_{j+1}, \ j = 1, \ldots, n, \tag{$\star\sim$}
$$

$$
\left(\rho_j = \rho'_-(\rho_j, \rho_{j+1}, Q_j, Q_{j+1}), \ \rho_{j+1} = \rho'_+(\rho_j, \rho_{j+1}, Q_j, Q_{j+1})\right), \ \text{if } Q_j \neq Q_{j+1}, \ j = 1, \ldots, n, \tag{$\star\ /$}
$$

$$
\rho_j = F_j, \ \text{if } F_j \neq \varnothing, \ j = 1, \ldots, n+1, \tag{$\star\ !$}
$$

where $\rho'_-(\rho_j, \rho_{j+1}, Q_j, Q_{j+1})$ and $\rho'_+(\rho_j, \rho_{j+1}, Q_j, Q_{j+1})$ are given as the optimizers from the solution described in Appendix A.3. These conditions are also used to define guards of the transitions, as will be described in the following subsection.

B.2. Transitions

Here we describe the various transitions that model the evolution of the FTSM. For each of the transitions, the states that do not change are omitted from the description. The transitions are listed in order of increasing priority, i.e., we first present the transitions that can only be taken if no other transition can be taken, and end with transitions which do not depend on the state of the system, only on exogenous inputs. We use notation $\circ$ to represent any of the transitions.

B.2.1. Passage of time transition $\tau(t_{\text{end}})$

The first transition we describe is the passage of time, which describes the propagation of fronts between their interactions, or until the externally provided goal time t_{end}. This transition is taken if the state X is in guard set

$$
X \in \mathcal{G}_\tau = \left\{X \in \mathcal{X} | (\star-), (\star\sim), (\star\ /), (\star\ !), \ \tau \in [0, \tau^*] \right\},
$$

i.e., the state is admissible (conditions $(\star-)$, $(\star\sim)$, $(\star\ /)$, and $(\star\ !)$ hold). The transition is defined by

$$
(t, z) \xrightarrow{\tau(t_{\text{end}})} (t', z')
$$
$$
t' = t + \tau^*, \quad z' = z + \lambda \tau^*
$$

where $\lambda = \begin{bmatrix} \lambda_1 & \dots & \lambda_n \end{bmatrix}^\top$, and the wave-speeds λ_i are given as

$$\lambda_i = \begin{cases} \frac{Q_{i+1}(\rho_{i+1}) - Q_i(\rho_i)}{\rho_{i+1} - \rho_i}, & \xi_i = \xi_{i+1}, \\ \lambda_i^+, & \xi_i > \xi_{i+1}, \\ \lambda_{i+1}^-, & \xi_i < \xi_{i+1}. \end{cases}$$

The maximum time shift τ^* is the minimum of the time for which condition $(\star-)$ is first violated,

$$\tau_z^* = \min \left\{ \frac{z_{i+1} - z_i}{\lambda_i - \lambda_{i+1}} \,\middle|\, z_{i+1} \ge z_i, \lambda_i > \lambda_{i+1}, i = 1, \dots, n-1 \right\}$$

and the time to specified goal time $\tau_{\text{end}}^* = t_{\text{end}} - t$, $\tau^* = \max\left\{0, \min\left\{\tau_z^*, \tau_{\text{end}}^*\right\}\right\}$. Note that if $t \ge t_{\text{end}}$, $X' = X$.

B.2.2. Front interaction transition $-_i$

A front interaction transition is taken when two fronts collide and the state is in guard set

$$X \in \mathcal{G}_{-_i} = \left\{ X \in \mathcal{X} \,\middle|\, \neg(\star-)_i, (\star-)_j, j > i, (\star\sim), (\star\,/), (\star\,!) \right\},$$

where by $\neg(\star-)_i$ we signify that the $j = i$-th condition in $(\star-)$ is violated. For all transitions $\circ \in \{-, \sim, /, !\}$, we write $X \in \mathcal{G}_\circ$ if $X \in \mathcal{G}_{\circ_i}$ for any i. In this case, the position of fronts becomes equal, $z_i = z_{i+1}$ while their distance is decreasing, $\lambda_i > \lambda_{i+1}$. The front interaction transition corresponds to deactivating one state,

$$(n, z, \rho, Q) \xrightarrow{-_i} (n', z', \rho', Q')$$
$$n' = n - 1,$$
$$z' = \begin{bmatrix} z_1 & \dots & z_i & \vdots & z_{i+2} & \dots & z_n \end{bmatrix}^\top,$$
$$\rho' = \begin{bmatrix} \rho_1 & \dots & \rho_i & \vdots & \rho_{i+2} & \dots & \rho_{n+1} \end{bmatrix}^\top,$$
$$Q' = \begin{bmatrix} Q_1 & \dots & Q_i & \vdots & Q_{i+2} & \dots & Q_{n+1} \end{bmatrix}^\top.$$

If $Q_i \ne Q_{i+2}$, this transition is likely to cause condition $(\star\,/)$ to be violated, and thus be followed by transition $/_i$. Note that the Total Variation of the traffic density is nonincreasing through the transition,

$$\text{T.V.}(\rho') - \text{T.V.}(\rho) = \begin{cases} 0, & \rho_{i+1} \in [\min(\rho_i, \rho_{i+2}), \max(\rho_i, \rho_{i+2})], \\ |\rho_{i+2} - \rho_i| - |\rho_{i+1} - \rho_i| - |\rho_{i+2} - \rho_{i+1}|, & \text{otherwise}. \end{cases}$$

B.2.3. Internal Riemann transition $\sim_i$

This transitions results from solutions to Riemann problem given in Appendix A.2, and it is taken when the state is in guard set

$$X \in \mathcal{G}_{\sim_i} = \left\{ X \in \mathcal{X} \,\middle|\, \neg(\star\sim)_i, (\star\sim)_j, j > i, (\star\,/), (\star\,!), Q_i = Q_{i+1} \right\},$$

The transition can be described by

$$(n, z, \rho, Q) \xrightarrow{\sim_i} (n', z', \rho', Q')$$
$$n' = n + m - 2, \quad m = {}^{\rho_i}\tilde{m}_{Q_i, W}^{\rho_{i+1}}$$
$$z' = \begin{bmatrix} z_1 & \dots & z_{i-1} & \vdots & z_i \mathbb{1}_{m-1}^\top & \vdots & z_{i+1} & \dots & z_n \end{bmatrix}^\top,$$
$$\rho' = \begin{bmatrix} \rho_1 & \dots & \rho_{i-1} & \vdots & {}^{\rho_i}\tilde{\Sigma}_{Q_i, W}^{\rho_{i+1}\top} & \vdots & \rho_{i+2} & \dots & \rho_{n+1} \end{bmatrix}^\top,$$
$$Q' = \begin{bmatrix} Q_1 & \dots & Q_{i-1} & \vdots & Q_i \mathbb{1}_m^\top & \vdots & Q_{i+2} & \dots & Q_{n+1} \end{bmatrix}^\top.$$

Depending on ρ_i and ρ_{i+1}, the number of active states can decrease (if $\rho_i = \rho_{i+1}$), increase, or stay the same.

B.2.4. Boundary Riemann transition $/_i$

This transition can occur at interfaces between zones with different flux functions and reflect the solution from Appendix A.3. It is taken when the state is in guard set

$$X \in \mathcal{G}_{/_i} = \left\{ X \in \mathcal{X} \,\middle|\, \neg(\star\,/)_i, (\star\,/)_j, j > i, (\star\,!), Q_i \ne Q_{i+1} \right\}.$$

The transition can be described by

$$(n, z, \rho, Q) \xrightarrow{/_i} (n', z', \rho', Q')$$
$$n' = n + m_- + m_+ - 2, \quad m_- = {}^{\rho_i}\tilde{m}_{Q_i, W}^{\rho'_-}, \quad m_+ = {}^{\rho'_+}\tilde{m}_{Q_{i+1}, W}^{\rho_{i+1}}$$
$$z' = \begin{bmatrix} z_1 & \dots & z_{i-1} & \vdots & z_i \mathbb{1}_{m_-+m_+}^\top & \vdots & z_{i+1} & \dots & z_n \end{bmatrix}^\top,$$
$$\rho' = \begin{bmatrix} \rho_1 & \dots & \rho_{i-1} & \vdots & {}^{\rho_i}\tilde{\Sigma}_{Q_i, W}^{\rho'_-\top} & \vdots & {}^{\rho'_+}\tilde{\Sigma}_{Q_{i+1}, W}^{\rho_{i+1}\top} & \vdots & \rho_{i+2} & \dots & \rho_{n+1} \end{bmatrix}^\top,$$
$$Q' = \begin{bmatrix} Q_1 & \dots & Q_{i-1} & \vdots & Q_i \mathbb{1}_{m_-}^\top & \vdots & Q_{i+1} \mathbb{1}_{m_+}^\top & \vdots & Q_{i+2} & \dots & Q_{n+1} \end{bmatrix}^\top,$$

where densities $\rho'_- = \rho'_-(\rho_i, \rho_{i+1}, Q_i, Q_{i+1})$ and $\rho'_+ = \rho'_+(\rho_i, \rho_{i+1}, Q_i, Q_{i+1})$ are obtained by solving the optimization problem (20), with $\rho_- = \rho_i$, $\rho_+ = \rho_{i+1}$, $Q_- = Q_i$, and $Q_+ = Q_{i+1}$.

B.2.5. State forcing $!_i$

This transition ensures that the density matches the forced traffic density where $F_i \neq \varnothing$; it is thus taken when the state is in guard set

$$X \in \mathcal{G}_{!_i} = \left\{ X \in \mathcal{X} \mid \neg(\star \; !)_i, (\star \; !)_j, j > i \right\},$$

and defined by

$$(\rho) \xrightarrow{!_i} (\rho')$$
$$\rho'_i = \begin{cases} \rho_i, & F_i = \varnothing, \\ F_i, & F_i \neq \varnothing. \end{cases}$$

B.2.6. State insertion $\vee(\rho_\vee, x_\vee)_i$

State insertion is an exogenous transition, i.e., it can be taken for any $X \in \mathcal{X}$ given the appropriate external input. It consists of adding two fronts at position $x_\vee$ downstream of front i, with $z_i \leq x_\vee \leq z_{i+1}$,

$$(n, z, \rho, Q) \xrightarrow{\vee(x_\vee)_i} (n', z', \rho', Q')$$
$$n' = n + 2,$$
$$z' = \begin{bmatrix} z_1 & \dots & z_i & x_\vee & x_\vee & z_{i+1} & \dots & z_n \end{bmatrix}^\top,$$
$$\rho' = \begin{bmatrix} \rho_1 & \dots & \rho_i & \rho_{i+1} & \rho_{i+1} & \rho_{i+1} & \dots & \rho_{n+1} \end{bmatrix}^\top,$$
$$Q' = \begin{bmatrix} Q_1 & \dots & Q_i & Q_{i+1} & Q_{i+1} & Q_{i+1} & \dots & Q_{n+1} \end{bmatrix}^\top.$$

It is only necessary to specify i if $z_i = x_\vee$ or $z_{i+1} = x_\vee$, in order to disambiguate the ordering of fronts.

B.2.7. Flux function transition $\mathcal{Q}(Q_Q, i, j)$

Finally, flux function transition is another exogenous transition, which covers changes in flux functions in specific areas. The transition is defined as

$$(Q) \xrightarrow{\mathcal{Q}(Q_Q, i, j)} (Q')$$
$$Q' = \begin{bmatrix} Q_1 & \dots & Q_i & Q_Q & \dots & Q_Q & Q_{j+1} & \dots & Q_n \end{bmatrix}^\top,$$

with $Q_Q \in \mathcal{Q}$ and $j > i$. Formally, this change has no immediate effect on any of the other states, but it is likely to cause $(\star \sim)$, $(\star \; /)$, or $(\star \; !)$ to be violated.

Appendix C. Proofs

We state the proofs of theorems from Section 2.2, along with some required lemmas.

Lemma 1. *The FTSM is nonblocking and deterministic.*

Proof. All $X \in \mathcal{X}$ are in exactly one guard set, since the guard sets $\mathcal{G}_\circ$ form a partition of $\mathcal{X}$. Furthermore, all transitions $\circ$ ensure that if $X \in \mathcal{X}$ and $X \xrightarrow{\circ} X'$ then $X' \in \mathcal{X}$. Therefore there exists a unique solution of the FTSM for every initial state $X_0 \in \mathcal{X}$. $\square$

This lemma establishes the basic properties of the FTSM based directly on the guard sets of the transitions. Next, we study the behaviour of the FTSM in zones where the flux function is homogeneous, and with constant boundary conditions.

Lemma 2. *If the state of FTSM $X \in \mathcal{G}_{\circ_i}$, where transition $\circ_i$ is $\sim_i$ if $X \xrightarrow{\sim_i} X'$ or $/_i$ if $X \xrightarrow{/_i} X'$, then $X' \notin \mathcal{G}_{\sim_j}$ and $X' \notin \mathcal{G}_{/_j}$ for $j = i, \dots, i + n' - n$.*

Proof. First, consider the case when $X \in \mathcal{G}_{\sim_i}$ and $(\star \; /)$. Based on the definition of the transition $\sim_i$, we have that $[\rho'_i \; \dots \; \rho'_{i+n'-n+1}]^\top = \rho_i \; \bar{\Sigma}^{\rho_{i+1}}_{Q_i, W}$. Therefore, for $k = 0, \dots, n' - n$, $[\rho'_{i+k} \; \rho'_{i+k+1}]^\top = \rho'_{i+k} \; \bar{\Sigma}^{\rho'_{i+k+1}}_{Q_i, W}$, so condition $(\star \sim)_{i+k}$ holds and $X' \notin \mathcal{G}_{\sim_j}$, $j = i + k$. Next, if $X \in \mathcal{G}_{/_i}$, based on the definition of the transition $/_i$, we have that $[\rho'_i \; \dots \; \rho'_{i+m_--1}]^\top = \rho_i \; \bar{\Sigma}^{\rho_{i+1}}_{Q_i, W}$ and $[\rho'_{i+m_-} \; \dots \; \rho'_{i+n'-n+1}]^\top = \rho'_+ \; \bar{\Sigma}^{\rho_{i+1}}_{Q_{i+1}, W}$, where $m_- = \rho_i \; \tilde{m}^{\rho'_-}_{Q_i, W}$ and $\rho'_\pm$ are the solutions to the optimization problem (20) for $\rho_- = \rho_i$, $\rho_+ = \rho_{i+1}$, $Q_- = Q_i$, $Q_+ = Q_{i+1}$. Therefore, for $k = 0, \dots, m_- - 2, m_-, \dots, n' - n$, $[\rho'_{i+k} \; \rho'_{i+k+1}]^\top = \rho'_{i+k} \; \bar{\Sigma}^{\rho'_{i+k+1}}_{Q'_{i+k}, W}$, so condition $(\star \sim)_{i+k}$ holds and $X' \notin \mathcal{G}_{\sim_j}$, $j = i + k$. For $i + m_- - 1$, we have $\rho'_{i+m_--1} = \rho'_-$, $\rho'_{i+m_-} = \rho'_+$, $Q'_{i+m_--1} = Q_i$, and $Q'_{i+m_-} = Q_{i+1}$, so since ρ'_- and ρ'_+ are the result of transition $/_i$, condition $(\star \; /)_{i+m_--1}$ holds and $X' \notin \mathcal{G}_{/_{i+m_--1}}$. $\square$

We can now state the proof of Theorem 1.

Proof of Theorem 1. First, if the density is forced in the zone of flux function Q^*, $F^* \neq \varnothing$, this holds trivially, with $\rho(x, t') = F^*$ on $[z_-^*(t'), z_+^*(t')]$ for $t' > 0$. Otherwise, after a finite number of transmissions $X_0 \xrightarrow{\circ_1} \dots \xrightarrow{\circ_K} X_0'$, the state X_0' satisfies $(\star -)$, $(\star \sim)$, $(\star /)$, and $(\star \,!)$ for $i = i_-^*(0), \dots, i_+^*(0)$. Here transitions $\circ_k$ are: $/_{i_-^*(0)}$, $/_{i_+^*(0)}$, and a finite number of transitions $\sim_i$ and $-_i$ for $i_-^*(0) < i < i_+^*(0)$, if the state enters their respective guard sets during the transitions from X_0 to X_0'. Afterwards, due to Lemma 2, while $t < T$, the evolution of the state can be described with only transitions $\xrightarrow{\tau(t_{\text{end}})}$ and $\xrightarrow{-_i}$, for which the Total Variation and number of active fronts is nonincreasing. $\square$

Next, we study an infinitesimally narrow zone with a different flux function, and show that it does not give rise to Zeno behaviour.

Lemma 3. *Let X be the state of the FTSM with $z_{i_-} = z_{i_-+1} = \dots = z_{i_+}$ and no other front positions equal, where $Q_{i_-} = Q_L$, $Q_i = Q_C$ for $i = i_-+1, \dots, i_+$, $Q_{i_++1} = Q_R$, $Q_L \neq Q_C$, $Q_C \neq Q_R$, and the density in the zones of Q_L and Q_R is not forced, $F_L = F_R = \varnothing$. Then $X' \in \mathcal{G}_\tau$, where $X \xrightarrow{\circ_1} \dots \xrightarrow{\circ_K} X'$, $\circ_k \in \{-_{i_k}, \sim_{i_k}, /_{i_k}, !_{i_k}\}$, $k = 1, \dots, K$ with finite K.*

Proof. Without loss of generality, assume $(\star \sim)_i$, $(\star /)_i$, and $(\star \,!)_i$ hold for $i \notin [i_-, i_+]$, which could always be achieved in a finite number of transitions. Firstly, if the density in the zone of Q_C is forced, after transitions $!_i$, $i = i_- + 1, \dots, i_+$ inside the zone, $/_{i_-}$ and $/_{i_+}$ at both borders of the zone, and $i_+ - i_- - 1$ instances of transition $-_{i_-+1}$, the resulting state $X' \in \mathcal{G}_\tau$. Otherwise, after up to two transitions $/_i$ and up to $i_+ - i_- - 1$ transitions $\sim_i$, the resulting state is X^0, and it holds that either $\lambda_{i_-^0} = \lambda_{i_-^0+1} = \dots = \lambda_{i_+^0}$ and $X^0 \in \mathcal{G}_\tau$, satisfying the requirements of the lemma, or $X^0 \in \mathcal{G}_{-_i}$ for some $i_-^0 \leq i \leq i_+^0$. We denote i_-^m and i_+^m based on state X^m so that $Q_{i_-^m}^m = Q_L$, $Q_i^m = Q_C$ for $i = i_-^m + 1, \dots, i_+^m$, $Q_{i_+^m+1}^m = Q_R$. If $X^0 \in \mathcal{G}_{-_i}$ for some i, after k^0 front interaction transitions, $1 \leq k^0 \leq i_+^0 - i_-^0 - 1$, the resulting state $X^{0'}$ will either satisfy the requirements of the lemma, or $X^{0'} \in \mathcal{G}_{/_{i_-^{0'}}} \cup \mathcal{G}_{/_{i_+^{0'}}}$. After transition $/_{i_-^{0'}}$ or $/_{i_+^{0'}}$ from $X^{0'}$ to X^1, the process repeats with k^m front interaction transitions, $1 \leq k^1 \leq i_+^1 - i_-^1 - 1$, and with $i_+^{m+1} < i_+^m$, $m = 1, \dots, M$, until state $X^{M'}$ satisfies the requirements of the lemma, with $\lambda_{i_-^{M'}} \leq \lambda_{i^{M'}} \leq \dots \leq \lambda_{i_+^{M'}}$ if $\lambda_{i_-^{M'}} \leq \lambda_{i_+^{M'}}$, and $Q_{i_-^{M'}}^{M'} = Q_L$, $Q_{i_-^{M'}+1}^{M'} = Q_R$ if $\lambda_{i_-^{M'}} > \lambda_{i_+^{M'}}$ (i.e., the zone described by flux function Q_C vanishes), or we have $Q_{i_-^{M'}}^{M'} = Q_L$, $Q_{i^{M'}}^{M'} = Q_C$, and $Q_{i^{M'}+1}^{M'} = Q_R$. Finally, if $Q_{i_-^{M'}} = Q_L$, $Q_{i^{M'}} = Q_C$, and $Q_{i^{M'}+1} = Q_R$, the state satisfies the requirements of the lemma after up to three series of transitions, where the first and third series consist of a finite numbers of transitions $/_{i_\pm^p}$ or $-_{i_\pm^p}$, and the second series consists of a finite number of transitions $/_{i_\mp^p}$ or $-_{i_\mp^p}$, i.e., if the first transition is $/_{i^{M'}}$, the second series of transitions will happen at the boundary between Q_C and Q_R, and if the first transition is $/_{i_+^{M'}}$, the second series of transitions will happen at the boundary between Q_L and Q_C. Therefore, no infinite internal behaviour of the FTSM, starting from X and with no states in $\mathcal{G}_\tau$ can exist, and after a finite number of transitions, the system reaches state $X' \in \mathcal{G}_\tau$. $\square$

Finally, the proof of Theorem 2 is stated.

Proof of Theorem 2. Consider first the case when there are no exogenous transitions and passage of time transition is taken with an arbitrarily large t_{end}. Firstly, due to Lemma 1, the solution to initial conditions X_0 from arbitrary $t = t_0$ until arbitrary $t_{\text{end}} > t_0$ exists and is unique. Due to Theorem 1, the only way Zeno behaviour can arise in FTSM is through interactions with flux function boundaries. Due to Lemma 2, the transitions at a flux function boundary can only happen once a front reaches the boundary, and the fronts created by the flux function boundary transition travel away from the boundary. Therefore any periodic behaviour would require interaction between two flux function boundaries. If the zone between two flux function boundaries is of non-zero length, the fronts originating from one boundary can only reach the other boundary after non-zero time, so the only case when Zeno behaviour could arise is if the length of the zone goes to zero. Therefore, due to Lemma 3, no Zeno behaviour can arise. If exogenous transitions are forced at times $T_1, \dots, T_{\text{exo}}$, the solution can be split into intervals $[t_0, T_1), [T_1, T_2), \dots, [T_{\text{exo}}, t_{\text{end}}]$. Since the exogenous transitions do not cause the output of the system to change, we may form $\rho(x, t)$ out of pieces between two exogenous transitions, with each exogenous transition changing some part of the LWR model. $\square$

Appendix D. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.trb.2022.10.008.

References

Adimurthi, Adi., Sundar Ghoshal, Shyam., Dutta, Rajib., Veerappa Gowda, G.D., 2011. Existence and nonexistence of tv bounds for scalar conservation laws with discontinuous flux. Comm. Pure Appl. Math. 64 (1), 84–115.

Alonso Raposo, Maria., Ciuffo, Biagio., Makridis, Michail., Thiel, Christian., 2017. The r-evolution of driving: from connected vehicles to coordinated automated road transport (C-ART). In: Part I: Framework for a Safe & Efficient Coordinated Automated Road Transport (C-ART) System. Publications Office of the European Union, EUR 28575.

Antoniou, Constantinos, Ben-Akiva, Moshe, Koutsopoulos, Haris N, 2005. Online calibration of traffic prediction models. Transp. Res. Rec. 1934 (1), 235–245.

Bekiaris-Liberis, Nikolaos., Roncoli, Claudio., Papageorgiou, Markos., 2016. Highway traffic state estimation with mixed connected and conventional vehicles. IEEE Trans. Intell. Transp. Syst. 17 (12), 3484–3497.

Billot, Romain., El Faouzi, Nour-Eddin., De Vuyst, Florian., 2009. Multilevel assessment of the impact of rain on drivers' behavior: standardized methodology and empirical analysis. Transp. Res. Rec. J. Transp. Res. Board 2107 (1), 134–142.

Čičić, Mladen., Barreau, Matthieu., Henrik Johansson, Karl., 2020a. Numerical investigation of traffic state reconstruction and control using connected automated vehicles. In: IEEE 23rd Intelligent Transportation Systems Conference. ITSC, Rhodes, Greece, pp. 1–6.

Čičić, Mladen., Henrik Johansson, Karl., 2019. Stop-and-go wave dissipation using accumulated controlled moving bottlenecks in multi-class CTM framework. In: IEEE 58th Conference on Decision and Control. CDC, Nice, France, pp. 3146–3151.

Čičić, Mladen., Mikolášek, Igor., Henrik Johansson, Karl., 2020b. Front tracking transition system model with controlled moving bottlenecks and probabilistic traffic breakdowns. In: The 21rst IFAC World Congress.

Deshpande, Akash., Göllü, Aleks., Varaiya, Pravin., 1996. SHIFT: A formalism and a programming language for dynamic networks of hybrid automata. In: International Hybrid Systems Workshop. Springer, pp. 113–133.

Flynn, Morris R, Kasimov, Aslan R, Nave, J-C, Rosales, Rodolfo R, Seibold, Benjamin, 2009. Self-sustained nonlinear waves in traffic flow. Phys. Rev. E 79 (5), 056113.

Han, Yu., Yuan, Yufei., Hegyi, Andreas., Hoogendoorn, Serge.P., 2016. New extended discrete first-order model to reproduce propagation of jam waves. Transp. Res. Rec. J. Transp. Res. Board (2560), 108–118.

Hegyi, Andreas., Paul Hoogendoorn, Serge., Schreuder, Marco., Stoelhorst, Henk., Viti, Francesco., 2008. SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory. In: 11th IEEE International Conference on Intelligent Transportation Systems. pp. 827–832.

Herrera, Juan C, Work, Daniel B, Herring, Ryan, Ban, Xuegang Jeff, Jacobson, Quinn, Bayen, Alexandre M, 2010. Evaluation of traffic data obtained via gps-enabled mobile phones: The mobile century field experiment. Transp. Res. C 18 (4), 568–583.

Holden, Helge., Henrik Risebro, Nils., 2015. Front Tracking for Hyperbolic Conservation Laws, Vol. 152. Springer.

Hosein Ghods, Amir., Rahimi Kian, Ashkan., Tabibi, Masoud., 2009. Adaptive freeway ramp metering and variable speed limit control: a genetic-fuzzy approach. IEEE Intell. Transp. Syst. Mag. 1 (1), 27–36.

Hou, Zhongsheng., Xu, Jian-Xin., Yan, Jingwen., 2008. An iterative learning approach for density control of freeway traffic flow via ramp metering. Transp. Res. C 16 (1), 71–97.

James Lighthill, Michael., Beresford Whitham, Gerald., 1955. On kinematic waves II. a theory of traffic flow on long crowded roads. Proc. R. Soc. London. Ser. A. Math. Phys. Sci. 229 (1178), 317–345.

Laura Delle Monache, Maria., Liard, Thibault., Piccoli, Benedetto., Stern, Raphael., Work, Dan., 2019. Traffic reconstruction using autonomous vehicles. SIAM J. Appl. Math. 79 (5), 1748–1767.

Laval, Jorge A, Leclercq, Ludovic, 2010. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. Phil. Trans. R. Soc. A 368 (1928), 4519–4541.

Lu, Yadong., Chun Wong, Sze., Zhang, Mengping., Shu, Chi-Wang., 2009. The entropy solutions for the Lighthill-Whitham-Richards traffic flow model with a discontinuous flow-density relationship. Transp. Sci. 43 (4), 511–530.

Mazaré, Pierre-Emmanuel, Dehwah, Ahmad H, Claudel, Christian G, Bayen, Alexandre M, 2011. Analytical and grid-free solutions to the Lighthill–Whitham–Richards traffic flow model. Transp. Res. B 45 (10), 1727–1748.

Mehran, Babak., Kuwahara, Masao., Naznin, Farhana., 2012. Implementing kinematic wave theory to reconstruct vehicle trajectories from fixed and probe sensor data. Transp. Res. C 20 (1), 144–163.

Monache, Maria Laura Delle, Goatin., Paola, 2014. A front tracking method for a strongly coupled PDE-ODE system with moving density constraints in traffic flow. Discrete Contin. Dyn. Syst.-Ser. S 7 (3), 435–447.

Piacentini, Giulia., Goatin, Paola., Ferrara, Antonella., 2018. Traffic control via moving bottleneck of coordinated vehicles. In: 15th IFAC symposium on control in transportation systems.

Richards, Paul I, 1956. Shock waves on the highway. Oper. Res. 4 (1), 42–51.

Schönhof, Martin., Helbing, Dirk., 2007. Empirical features of congested traffic states and their implications for traffic modeling. Transp. Sci. 41 (2), 135–166.

Seo, Toru, Bayen, Alexandre M., Kusakabe, Takahiko, Asakura, Yasuo, 2017. Traffic state estimation on highway: A comprehensive survey. Annu. Rev. Control 43, 128–151.

Seo, Toru., Kawasaki, Yutaka., Kusakabe, Takahiko., Asakura, Yasuo., 2019. Fundamental diagram estimation by using trajectories of probe vehicles. Transp. Res. B 122, 40–56.

Seo, Toru., Kusakabe, Takahiko., 2015. Probe vehicle-based traffic state estimation method with spacing information and conservation law. Transp. Res. C 59, 391–403.

Shladover, Steven., Su, Dongyan., Lu, Xiao-Yun., 2012. Impacts of cooperative adaptive cruise control on freeway traffic flow. Transp. Res. Rec. J. Transp. Res. Board (2324), 63–70.

Simoni, Michele D, Claudel, Christian G, 2017. A fast simulation algorithm for multiple moving bottlenecks and applications in urban freight traffic management. Transp. Res. B 104, 238–255.

Smaragdis, Emmanouil., Papageorgiou, Markos., Kosmatopoulos, Elias., 2004. A flow-maximizing adaptive local ramp metering strategy. Transp. Res. B 38 (3), 251–270.

Spiliopoulou, Anastasia., Papamichail, Ioannis., Papageorgiou, Markos., Tyrinopoulos, Yannis., Chrysoulakis, John., 2017. Macroscopic traffic flow model calibration using different optimization algorithms. Oper. Res. 17 (1), 145–164.

Stern, Raphael E., Cui, Shumo, Delle Monache, Maria Laura, Bhadani, Rahul, Bunting, Matt, Churchill, Miles, Hamilton, Nathaniel, Haulcy, R'mani, Pohlmann, Hannah, Wu, Fangyu, Piccoli, Benedetto, Seibold, Benjamin, Sprinkle, Jonathan, Work, Daniel B., 2018. Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. Transp. Res. C 89, 205–221.

Tabuada, Paulo., 2009. Verification and Control of Hybrid Systems: A Symbolic Approach. Springer Science & Business Media.

Talebpour, Alireza, Mahmassani, Hani S, 2016. Influence of connected and autonomous vehicles on traffic flow stability and throughput. Transp. Res. C 71, 143–163.

Trafikverket, 2014. TRVMB Kapacitet och framkomlighetseffekter: Trafikverkets metodbeskrivning för beräkning av kapacitet och framkomlighetseffekter i vägtrafikanläggningar. TRV 2013:64343.

Wang, Yibing., Papageorgiou, Markos., 2005. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. Transp. Res. B 39 (2), 141–167.

Wang, Yibing., Papageorgiou, Markos., Messmer, Albert., Coppola, Pierluigi., Tzimitsi, Athina., Nuzzolo, Agostino., 2009. An adaptive freeway traffic state estimator. Automatica 45 (1), 10–24.

Yuan, Kai, Knoop, Victor L, Leclercq, Ludovic, Hoogendoorn, Serge P, 2017. Capacity drop: a comparison between stop-and-go wave and standing queue at lane-drop bottleneck. Transp. B Transp. Dyn. 5 (2), 145–158.