Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Front tracking transition system model with controlled moving bottlenecks and probabilistic traffic breakdowns

Mladen Čičić * **Igor Mikolášek** ** **Karl Henrik Johansson** *

* *Division of Decision and Control Systems, KTH Royal Institute of Technology, Stockholm, Sweden (e-mail:* cicic@kth.se, kallej@kth.se*)*
** *CDV – Transport Research Centre, Brno, Czech Republic (e-mail:* igor.mikolasek@cdv.cz*)*

Abstract: Cell-based approximations of PDE traffic models are widely used for traffic prediction and control. However, in order to represent the traffic state with good resolution, cell-based models often require a short cell length, which results in a very large number of states. We propose a new transition system traffic model, based on the front tracking method for solving the LWR PDE model. Assuming piecewise-linear flux function and piecewise-constant initial conditions, this model gives an exact solution. Furthermore, it is easier to extend, has fewer states and, although its dynamics are intrinsically hybrid, is faster to simulate than an equivalent cell-based approximation. The model is extended to enable handling moving bottlenecks as well as probabilistic traffic breakdowns and capacity drops at static bottlenecks. A control strategy that utilizes controlled moving bottlenecks for bottleneck decongestion is described and tested in simulation. It is shown that we are able to keep the static bottleneck in free flow by creating controlled moving bottlenecks at specific instances along on the road, and using them to regulate the incoming traffic flow.

Keywords: Traffic Modelling, Front Tracking, Transition Systems, Moving Bottlenecks, Stochastic Capacity, Traffic Control

1. INTRODUCTION

Ever since the development of the Lighthill-Whitham-Richards (LWR) traffic model, Lighthill and Whitham (1955) and Richards (1956), predicting and tracking the evolution of traffic shock waves has been an important part of modelling traffic. This problem gave rise to a multitude of PDE traffic models (with LWR being the simplest of them), where the traffic flow is described using hyperbolic conservation laws, see Lax (1973). Drawing upon the substantial body of mathematical literature on this class of PDEs, these models have been used to capture a variety of complex phenomena that arise in traffic, such as moving bottlenecks, as in Delle Monache and Goatin (2014), phantom jams in Flynn et al. (2009), and traffic phase transitions in Blandin et al. (2011).

Although these models can be very rich, their complexity is not conducive to control design, leading to relatively few works explicitly considering them for control, e.g., Yu and Krstic (2019). For practical implementation, where issues like model calibration need to be resolved, most traffic control systems are using cell-based discretized models such as METANET, in Kotsialos et al. (2002), or the Cell Transmission Model (CTM), Daganzo (1994), see Baskar et al. (2011) for a survey. A notable exception are ad hoc control algorithms, like those based directly on shock wave theory in Hegyi et al. (2008), or controlled moving bottlenecks in Čičić and Johansson (2018), where the control actions are calculated by analysing and predicting the evolution of traffic waves, and then applied on a cell-based traffic model. In order to represent the traffic with good resolution, these cell-based models often require a short cell length, which results in a large number of states. Additionally, these models require various extensions to correctly model relevant traffic phenomena such as moving bottlenecks and stop-and-go waves, since diffusion, which is inherent in their formulation, destroys the representation of crisp wavefronts.

The main focus of this paper is on developing a simple traffic model that is not cell-based. Inspired by the fact that the CTM is equivalent to discretizing the LWR model, with appropriate flux function, using the Godunov scheme, see Lebacque (1996), we use another method for solving PDEs, known as front tracking, given in Glimm et al. (1981), Holden and Risebro (2015). The main contribution is in formulating a front tracking transition system traffic model, where the traffic situation is described by the

⋆ The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 674875, VINNOVA within the FFI program under contract 2014-06200, the Swedish Research Council, the Swedish Foundation for Strategic Research, Knut and Alice Wallenberg Foundation, and the Ministry of Education, Youth and Sports of Czech Republic within the National Sustainability Programme I, project of Transport R&D Centre (LO1610), on the research infrastructure acquired from the Operation Programme Research and Development for Innovations (CZ.1.05/2.1.00/03.0064). The first author is affiliated with the Wallenberg AI, Autonomous Systems and Software Program (WASP).

Copyright lies with the authors

positions of wavefronts and traffic densities between them, and its evolution is governed by appropriately defined transitions. This model eliminates diffusion, is easier to extend, has fewer states and, although its dynamics are intrinsically hybrid, is faster to simulate than an equivalent CTM. Furthermore, since the model directly corresponds to solving a composite Riemann problem, we are able to include many results derived for the LWR model, like handling moving bottlenecks.

To showcase the flexibility of this framework, we extend the basic model to capture the stochastic capacity at stationary bottlenecks and capacity drop. Traditionally, and in most practical applications, capacity has been considered as constant for given road geometry, but since in reality traffic breakdowns can occur at various traffic demand levels, capacity is often modelled as a probability distribution, as in Brilon et al. (2005). Furthermore, this approach allows us to model the reduction of capacity of the bottleneck following the onset of congestion, like in Srivastava and Geroliminis (2013).

The paper is structured as follows. First, in Section 2 we describe the front tracking solution of the LWR model. Next, using the said solution, in Section 3 we present the proposed front tracking transition system traffic model, and in Section 4 introduce moving and static bottlenecks with probabilistic traffic breakdown into the model. Then, in Section 5 we propose a control law for bottleneck decongestion using controlled moving bottlenecks and in Section 6 show its effectiveness in simulation. Finally, in Section 7 we conclude the paper and present some directions for future work.

2. SOLUTION OF THE LWR MODEL

The LWR model is a first-order scalar hyperbolic conservation law given by the partial differential equation

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

where the conserved quantity $\rho(t,x)$ is the traffic density, and $Q(\rho)$ is the flux function. The model assumes that the speed of the vehicles can be directly expressed as a function of traffic density as $v(\rho)$, and that the traffic flow is simply $Q(\rho) = v(\rho)\rho$. We are interested in finding the weak solution to the general initial value problem that satisfies the entropy conditions, i.e. the so called entropy solution, see Holden and Risebro (2015).

2.1 Initial conditions and flux function

The initial condition $\rho(0,x)$ is a function of position x, and we assume that it can be approximated by a piecewise-constant function

$$\rho(0,x) = \begin{cases} \rho_1, & x < X_1 \\ \quad\vdots & \\ \rho_i, & X_{i-1} < x < X_i \\ \quad\vdots & \\ \rho_{N+1}, & x > X_N \end{cases} \tag{2}$$

to an arbitrary degree of accuracy. Furthermore, $\rho(0,x)$ is non-negative and bounded by some value P, known as jam density, $0 \le \rho(0,x) \le P$. The flux function $Q(\rho)$ is Lipschitz continuous and has support $[0,p]$, $p \le P$. Traditionally, this function was taken to be strictly concave, with the earliest choice being the Greenshields fundamental diagram that assumes that speed decreases linearly with traffic density, $v_G(\rho) = V\left(1 - \frac{\rho}{P}\right)$, $\rho \in [0,P]$, yielding a quadratic flux function $Q_G(\rho) = v_G(\rho)\rho$, $\rho \in [0,P]$. We may approximate any such function to arbitrary degree of accuracy as a polygon, i.e., continuous piecewise-linear, function Q,

$$Q(\rho) = \begin{cases} V_0\rho, & 0 \le \rho \le \sigma_0, \\ Q(\sigma_0) + V_1(\rho - \sigma_0), & \sigma_0 < \rho \le \sigma_1, \\ \quad\vdots & \\ Q(\sigma_{i-1}) + V_i(\rho - \sigma_{i-1}), & \sigma_{i-1} < \rho \le \sigma_i, \\ \quad\vdots & \\ Q(\sigma_m) + V_m(\rho - \sigma_m), & \sigma_m < \rho \le P, \end{cases} \tag{3}$$

where $Q(\sigma_m) + V_m(P - \sigma_m) = 0$. We denote the set of nodes in the definition of Q as

$$\mathbf{\Sigma}_Q = \{\sigma_0, \ldots, \sigma_m\},$$

and the set of slopes between nodes as

$$\mathbf{V}_Q = \{V_0, \ldots, V_m\}.$$

The set of all functions $Q(\rho)$ that satisfy these requirements will be denoted $\mathcal{Q}$.

Greenshields fundamental diagram, and various modifications thereof, does not fit the actual traffic data particularly well. Another simple choice is the triangular (Newel-Daganzo) fundamental diagram,

$$Q_\Delta^\sigma(\rho) = \begin{cases} V\rho, & 0 \le \rho \le \sigma, \\ W(P - \rho), & \sigma < \rho \le P, \end{cases} \tag{4}$$

with $W = \frac{V\sigma}{P-\sigma}$, which distinguishes between two phases of traffic: free flow where $0 \le \rho \le \sigma$ and $v(\rho) = V$, and congestion where $\sigma < \rho \le P$ and $v(\rho)$ decreases as ρ increases. This flux function is piecewise-linear, and described by (3) with node $\sigma_0 = \sigma$ (critical density) and slopes $V_0 = V$ (free flow speed), $V_1 = -W$ (congestion wave speed). In this work, we will use (4) to model the default behaviour of the traffic.

The front tracking corresponds to solving a sequence of Riemann problems, to find the entropy solution for a piecewise-constant approximation of initial conditions $\rho(0,x)$, assuming a piecewise-linear flux function. Effectively, instead of solving the *exact* PDE problem *approximately*, as is done in cell-based discretization, this method solves the *approximate* problem *exactly*. If we assume that the flux function is continuous and piecewise-linear, the solution obtained is of the form

$$\rho(t,x) = \begin{cases} \rho_1, & x < X_1 + \Lambda_1 t, \\ \quad\vdots & \\ \rho_i, & X_{i-1} + \Lambda_{i-1}t < x < X_i + \Lambda_i t, \\ \quad\vdots & \\ \rho_{N'+1}, & x > X_{N'} + \Lambda_{N'}t, \end{cases}$$

with $\Lambda_{i-1} \le \Lambda_i$ wherever $X_{i-1} = X_i$. We denote by Λ_i the transition speeds, defined either by the Rankine-Hugoniot condition

$$\Lambda_i = \frac{Q_i(\rho_{i+1}) - Q_i(\rho_i)}{\rho_{i+1} - \rho_i}, \tag{5}$$

if $Q_i = Q_{i+1}$, or externally if $Q_i \ne Q_{i+1}$.

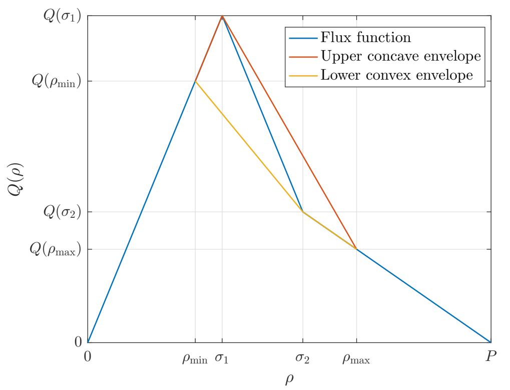


Fig. 1. The flux functions and its envelopes on $[\rho_{\min}, \rho_{\max}]$.

The solution consists of zones of constant density separated by fronts $X_i + \Lambda_i t$ where we have a discontinuity in density. This solution holds for $t \in [0, \tau]$, where τ is the minimum time when two fronts collide, $X_{i-1} + \Lambda_{i-1}\tau = X_i + \Lambda_i \tau$, 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 exact entropy solution to the initial value problem (1), (2), $\rho(t, x)$ for any t. Furthermore, the front tracking method yields *exact* solutions in case when the flux function is piecewise linear and initial conditions piecewise constant, which will be the case we consider here.

2.2 Lower convex and upper concave envelopes

When solving Riemann problems for an arbitrary $Q \in \mathcal{Q}$, with initial datum

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

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

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

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

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

on $[\rho_{\min}, \rho_{\max}]$, $\rho_{\min} = \min(\rho_-, \rho_+)$, $\rho_{\max} = \max(\rho_-, \rho_+)$. An illustration of upper concave and lower convex envelopes of a piecewise-linear flux function is given in Figure 1.

Note that $^{\rho_-}\tilde{Q}_Q^{\rho_+}(\rho)$ also is a polygon on $[\rho_{\min}, \rho_{\max}]$ and it can be defined in the same way as (3),

$$\tilde{Q}(\rho) = \begin{cases} Q(\rho_-) + \tilde{V}_0(\rho - \rho_{\min}), & \rho_{\min} \leq \rho \leq \tilde{\sigma}_1, \\ Q(\tilde{\sigma}_1) + \tilde{V}_1(\rho - \tilde{\sigma}_1), & \tilde{\sigma}_1 \leq \rho \leq \tilde{\sigma}_2, \\ \qquad \vdots & \\ Q(\tilde{\sigma}_{\tilde{m}}) + \tilde{V}_{\tilde{m}}(\rho - \tilde{\sigma}_{\tilde{m}}), & \tilde{\sigma}_{\tilde{m}} \leq \rho \leq \rho_{\max}, \end{cases}$$

omitting superscript ρ_- and ρ_+ and subscript Q, and not requiring that $\tilde{Q}(\rho_{\max}) = 0$. We write the vector of slopes of such polygon $^{\rho_-}\tilde{V}_Q^{\rho_+}$, ordered from $\tilde{V}_0$ to $\tilde{V}_{\tilde{m}}$ for $\rho_- < \rho_+$ or from $\tilde{V}_{\tilde{m}}$ to $\tilde{V}_0$ for $\rho_- > \rho_+$. All nodes of $\tilde{Q}$, $\tilde{\sigma}_i$, are also nodes of Q, σ_i, on $[\rho_{\min}, \rho_{\max}]$, and they can be determined by efficient convex hull algorithms.

Finally, we denote the sorted (ascending if $\rho_- < \rho_+$ and descending if $\rho_- > \rho_+$) column vector of elements of $\tilde{\boldsymbol{\Sigma}}$ as $\tilde{\Sigma}$, and its length as $\tilde{m}$. Same as with envelopes $^{\rho_-}\tilde{Q}_Q^{\rho_+}(\rho)$, $^{\rho_-}\tilde{\Sigma}_Q^{\rho_+}$ will consist of nodes of the lower convex or upper concave envelope, depending on whether ρ_- or ρ_+ is larger,

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

The flux function for which $\tilde{\Sigma}$ and $\tilde{m}$ are calculated is written in subscript.

3. FRONT TRACKING TRANSITION SYSTEM MODEL

The described procedure, with continuously changing solution between two composite Riemann problem solving instances and jumps resulting from them, lends itself to a transition system formulation. We follow the transition system formulation given in Tabuada (2009), and define the evolution of the front tracking solution to the scalar conservation as the execution of the transition system given by the quadruple $(\mathcal{X}, \mathcal{X}_0, U, \rightarrow)$. We will describe this formulation part by part in this section.

3.1 States and initial states

The set of states $\mathcal{X} = (n, t, Z, R, Q)$ is composed of:

- Number of active states: $n \in \mathbb{N}$, $n \leq N$
- Time: $t \in \mathbb{R}_{\geq 0}$
- Front positions: $Z \in \mathbb{R}^N$, $z_i \leq z_{i+1}$ for $i = 1, \ldots, n$
- Traffic densities: $R \in [0, P]^{N+1}$
- Flux functions: $Q \in \mathcal{Q}^{N+1}$, where $\mathcal{Q}$ is a set of piecewise linear continuous functions with support $[0, p]$, $p \leq P$

The maximum number of states N can be taken large enough so that the number of states never exceeds it. However, only the active states, which we will denote

$$z = [z_1 \ \ldots \ z_n]^\top = [\, I_n \,\vdots\, \mathbb{0}_{n \times N-n} \,]\, Z$$
$$\rho = [\rho_1 \ \ldots \ \rho_{n+1}]^\top = [\, I_{n+1} \,\vdots\, \mathbb{0}_{n+1 \times N-n} \,]\, R,$$

and flux functions $Q_1, \ldots, Q_{n+1}$, influence the behaviour of the system, so when describing the transitions, we only define their updates, and the inactive states may take arbitrary values. Effectively, the dimension of active states will vary as a part of the model dynamics.

The set of initial states X_0 can be the same as the set of all states, but in that case, we may be forced to take some number of Riemann transitions, described in the following section, at $t = 0$. This can be counteracted by imposing additional conditions on the set of initial states,

$$\rho_i \neq \rho_{i+1},\ ^{\rho_i}\tilde{\boldsymbol{\Sigma}}_{Q_i}^{\rho_{i+1}} = \emptyset, \text{ if } Q_i = Q_{i+1}, \tag{6a}$$
$$\rho_i = r_-,\ \rho_{i+1} = r_+, \text{ if } Q_i \neq Q_{i+1}, \tag{6b}$$

where r_- and r_+ are given as the optimizers of the optimization problem

$$
\begin{aligned}
\underset{r_-,r_+}{\text{maximize}}\ & Q_i(r_-) - \Lambda_i r_- \\
\text{s.t.}\quad & Q_{i+1}(r_+) - Q_i(r_-) = \Lambda_i(r_+ - r_-), \\
& {}^{\rho_i}\tilde{V}^{r_-}_{Q_i} < \Lambda_i, \\
& {}^{r_+}\tilde{V}^{\rho_{i+1}}_{Q_{i+1}} > \Lambda_i.
\end{aligned} \tag{7}
$$

These conditions define the *admissible set* of states. Solving the optimization problem (7) is equivalent to solving a Riemann problem at the boundary between zones described with flux functions Q_i and Q_{i+1}, assuming the transition speed Λ_i is imposed as an external input. For most simple flux functions used in practice, solving this maximization problem can be done explicitly. Furthermore, note that optimal r_- and r_+ will always be such that either $r_- \in \boldsymbol{\Sigma}_{Q_i} \cup \{\rho_i\}$ or $r_+ \in \boldsymbol{\Sigma}_{Q_{i+1}} \cup \{\rho_{i+1}\}$. Therefore, the problem can be solved by forming a set of all possible pairings of (r_-, r_+), and then checking the second and third constraint for each of them, in order of descending $Q_i(r_-) - \Lambda_i r_-$, so that the first pair to satisfy these constraints is the optimizer.

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

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

Note that we use notation $\rho(t,x)$ for the reconstructed function, and $\rho = [\rho_1 \ldots \rho_{n+1}]^\top$ for the vector of traffic densities.

3.2 Inputs and transitions

In this subsection, we will describe the various transitions that model the evolution of the transition system. For each of the transitions, the states that do not change will be omitted from the description.

Homogeneous Riemann transitions $\sim_i$: The first type of transitions results from solving a Riemann problem at the position of a wavefront that is not an interface between different flux functions. For this transition to be possible at z_i, we require that $Q_i = Q_{i+1}$ and that the condition (6a) is not satisfied. The transition can be described by

$$
\begin{gathered}
(n, z, \rho, Q) \xrightarrow{\sim_i} (n', z', \rho', Q') \\
n' = n + {}^{\rho_i}\tilde{m}^{\rho_{i+1}}_{Q_i} - 2, \\
z' = \left[z_1 \ldots z_{i-1} \;\vdots\; z_i \mathbb{1}^\top_{\rho_i \tilde{m}^{\rho_{i+1}}_{Q_i} - 1} \;\vdots\; z_{i+1} \ldots z_n \right]^\top, \\
\rho' = \left[\rho_1 \ldots \rho_{i-1} \;\vdots\; {}^{\rho_i}\tilde{\Sigma}^{\rho_i + 1\,\top}_{Q_i} \;\vdots\; \rho_{i+2} \ldots \rho_{n+1} \right]^\top, \\
Q' = \left[Q_1 \ldots Q_{i-1} \;\vdots\; Q_i \mathbb{1}^\top_{\rho_i \tilde{m}^{\rho_{i+1}}_{Q_i} - 1} \;\vdots\; Q_{i+2} \ldots Q_{n+1} \right]^\top.
\end{gathered}
$$

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

Heterogenous Riemann transitions q_i: These transitions can occur at interfaces between zones with different flux functions. They ensure that the condition (6b) is satisfied at the interface between two flux functions, $Q_i \neq Q_{i+1}$. The transition can be described by

$$
\begin{gathered}
(n, z, \rho, Q) \xrightarrow{q_i} (n', z', \rho', Q') \\
n' = n + {}^{\rho_i}\tilde{m}^{r_-}_{Q_i} + {}^{r_+}\tilde{m}^{\rho_{i+1}}_{Q_{i+1}} - 2, \\
z' = \left[z_1 \ldots z_{i-1} \;\vdots\; z_i \mathbb{1}^\top_{\rho_i \tilde{m}^{r_-}_{Q_i} + {}^{r_+}\tilde{m}^{\rho_{i+1}}_{Q_{i+1}} - 1} \;\vdots\; z_{i+1} \ldots z_n \right]^\top, \\
\rho' = \left[\rho_1 \ldots \rho_{i-1} \;\vdots\; {}^{\rho_i}\tilde{\Sigma}^{r_-}_{Q_i} \;\vdots\; {}^{r_+}\tilde{\Sigma}^{\rho_{i+1}}_{Q_{i+1}} \;\vdots\; \rho_{i+2} \ldots \rho_{n+1} \right]^\top, \\
Q' = \left[Q_1 \ldots Q_{i-1} \;\vdots\; Q_i \mathbb{1}^\top_{\rho_i \tilde{m}^{r_-}_{Q_i}} \;\vdots\; Q_{i+1} \mathbb{1}^\top_{r_+ \tilde{m}^{\rho_{i+1}}_{Q_{i+1}}} \;\vdots\; Q_{i+1} \ldots Q_{n+1} \right]^\top,
\end{gathered}
$$

where densities r_- and r_+ are again obtained by solving the optimization problem (7).

Passage of time τ ***and front interactions*** $-_i$: Passage of time transition describes the propagation of wave fronts between their interactions, and can only be taken if the state is admissible (i.e. the condition (6) holds), positions of active fronts z are monotonically non-decreasing, and for every $z_i = z_{i+1}$ we also have $\Lambda_i < \Lambda_{i+1}$. Traffic densities ρ, number of active states n and flux functions Q do not change in these transitions, so those will be omitted from the description. Only the wavefront positions of active states $(i = 1, \ldots, n)$ are changed. We define this transition by

$$
\begin{gathered}
(t, z) \xrightarrow{\tau} (t', z') \\
t' = t + \tau,\ z' = z + \Lambda\tau
\end{gathered}
$$

where $\Lambda = [\Lambda_1 \ \ldots \ \Lambda_n]^\top$, and the wave speeds Λ_i are given as the transition speeds of the Rankine-Hugoniot condition (5) if $Q_{i+1} = Q_i$, or are an external input in case $Q_{i+1} \neq Q_i$.

This transition can be taken for $\tau \leq \tau^*$, where τ^* is the minimum of the time to next front interaction

$$
\tau^*_z = \min \left\{ \frac{z_{i+1} - z_i}{\Lambda_i - \Lambda_{i+1}} \,\middle|\, z_{i+1} \geq z_i, \Lambda_i > \Lambda_{i+1} \right\}
$$

and the time to next externally generated event τ^*_e, $\tau^* = \min\{\tau^*_z, \tau^*_e\}$.

A front interaction transition is taken when two fronts interact, or collide, i.e., their position 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,

$$
\begin{gathered}
(n, z, \rho, Q) \xrightarrow{-_i} (n', z', \rho', Q') \\
n' = n - 1, \qquad z' = A_{n \setminus i+1} z, \\
\rho' = A_{n \setminus i+1} \rho,\quad Q' = A_{n \setminus i+1} Q,
\end{gathered}
$$

where $A_{n \setminus i} = [e_1 \ldots e_{i-1}\ e_{i+1} \ldots e_n]$ and e_i are the standard basis vectors. If $Q_i \neq Q_{i+1}$, this transition is likely to cause condition (6b) to be violated, thus it will be followed by transition q_i.

State insertion $+(\rho_+, x_+)_i$ ***and flux function transition*** $\mathcal{Q}(q, i, j, \Lambda_i, \Lambda_j)$: Here we describe two useful exogenous transitions. The state insertion transition consists of adding two fronts at position x_+ downstream of front i, with $z_i \leq x_+ \leq z_{i+1}$, with density ρ_+,

$$
\begin{aligned}
(n, z, \rho, Q) &\xrightarrow{+(\rho_+, x_+)_i} (n', z', \rho', Q') \\
n' &= n + 2, \\
z' &= \left[\, z_1 \ldots z_i \;\vdots\; x_+ \quad x_+ \;\vdots\; z_{i+1} \ldots z_n \right]^\top, \\
\rho' &= \left[\, \rho_1 \ldots \rho_i \;\vdots\; \rho_{i+1} \quad \rho_+ \;\vdots\; \rho_{i+1} \ldots \rho_{n+1} \right]^\top, \\
Q' &= \left[Q_1 \ldots Q_i \;\vdots\; Q_{i+1}\; Q_{i+1} \;\vdots\; Q_{i+1} \ldots Q_{n+1} \right]^\top.
\end{aligned}
$$

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

Finally, flux function transitions cover various changes done to flux functions in specific areas. The transition is defined as

$$
\begin{aligned}
(Q) &\xrightarrow{\mathcal{Q}(q, i, j, \Lambda_i, \Lambda_j)} (Q') \\
Q' &= \left[\, Q_1 \ldots Q_i \;\vdots\; q \ldots q \;\vdots\; Q_{j+1} \ldots Q_n \right]^\top,
\end{aligned}
$$

with $q \in \mathcal{Q}$ and $j > i$. It is required that wave speeds Λ_i and Λ_j are externally defined if $q \neq Q_{i-1}$ and $q \neq Q_{j+1}$, respectively. Formally, this change has no immediate effect on any of the other states, but it is likely to force a number of transitions at the boundaries of changed flux functions. Furthermore, since some wave speeds may be changed, the passage of time transition will now change the front positions in a different way.

4. MODELLING MOVING AND STATIC BOTTLENECKS

Using two exogenous transitions, it is easy to model the creation and influence of controlled moving bottlenecks through the new transition system traffic model. We denote the scaled triangular flux function

$$
Q_\Delta^{\sigma_s}(\rho) = \begin{cases} V\rho, & 0 \leq \rho \leq \sigma_s, \\ W\left(P\frac{\sigma_s}{\sigma} - \rho\right), & \sigma_s < \rho \leq P\frac{\sigma_s}{\sigma}, \end{cases}
$$

where σ_s is the new critical density. We may model the addition of a bottleneck at position x_b, moving at speed u_b and reducing the capacity of the road at its position to $V\sigma_b$ by:

(1) Taking a transition $+(\sigma_b, x_b)_{i_-}$. A zone of density σ_b is added donwstream of front i_-, $z_{i_-} \leq x_b \leq x_{i_-+1}$.
(2) Taking a transition $\mathcal{Q}(Q_\Delta^{\sigma_b}, i_- + 1, i_- + 2, u_b, u_b)$. The flux function at the position of the bottleneck is scaled down so that its capacity is $V\sigma_b$, and both the upstream and downstream ends of the bottleneck will move at speed u_b.

A similar procedure can be applied to create a moving bottleneck of nonzero length. Note that it is required to ensure that u_b is always taken such that

$$
u_b\left(\rho_+ - P\frac{\sigma_b}{\sigma}\right) \geq Q_+(\rho_+),
$$

where ρ_+ and Q_+ are the traffic density and flux function immediately downstream of the moving bottleneck. The speed or capacity of the moving bottleneck can be changed by taking a transition $\mathcal{Q}(Q_\Delta^{\sigma_b'}, i_b, i_b + 1, u_b', u_b')$, where σ_b' and u_b' are the new critical density and moving bottleneck speed, respectively.

Note that with $u_b = 0$, this approach also allows us to model stochastic traffic breakdown and capacity drop. A static bottleneck will be in free flow for low levels of demand. As the demand increases, the probability of traffic breakdown will start increasing. The stochastic capacity of a bottleneck is given by specifying the probability of traffic breakdown within some time interval T, given the demand level q, as in Brilon et al. (2005). The probability of traffic breakdown is taken to be Weibull-distributed,

$$
F_B(T, q) = 1 - e^{-\frac{T}{T_0}\left(\frac{q}{\beta}\right)^{-\alpha}},
$$

where parameters T_0, α_0 and β_0 are positive design parameters obtained from estimating the stochastic capacity of the bottleneck. Conversely, the time to breakdown is an exponentially distributed random variable parametrised by q, $\Theta_q \sim \exp(T_B(q)^{-1})$. We denote by $T_B(q)$ the mean time to breakdown, given as a function of the current traffic demand at the bottleneck,

$$
T_B(q) = T_0\left(\frac{q}{\beta_0}\right)^{-\alpha_0}.
$$

If the bottleneck at position x_b is in free flow, and demand at its position stays $q = V\rho(t, x_b)$ longer than the time to breakdown Θ_q, we say that there has been a traffic breakdown at the bottleneck, and drop its capacity to $V\sigma_d$. If the demand at the bottleneck changes to q' before the time to breakdown has elapsed, we generate a new time to breakdown parametrised by the new demand $\Theta_{q'}$ and repeat the process. We implement the bottleneck activation and capacity drop by simply adding a static bottleneck, $u_b = 0$, with critical density σ_d at the bottleneck position x_b. By choosing σ_d, we impose the discharge rate $V\sigma_d$ of the bottleneck with active capacity drop.

Once the bottleneck is active and congested, we consider it to have been resolved once all the congestion has been dissipated and the demand at its position drops below $q < V\sigma_d$. We implement its return to free flow by taking a transition $\mathcal{Q}(Q_\Delta^{\sigma}, i_b, i_b + 1, 0, 0)$. This way, the flux function in the zone of the bottleneck becomes equal to that of the road. Note that while technically a breakdown can happen for any q, it is only necessary to consider the case when $q > V\sigma_d$, since otherwise the breakdown would immediately be resolved.

5. BOTTLENECK DECONGESTION CONTROL

To showcase the use of the proposed transition system traffic model, we design a simple control law for bottleneck decongestion using controlled moving bottlenecks. We assume that we are able to create controlled moving bottlenecks at arbitrary positions on the road x_{b_i}, and control how many lanes they take; in real application, this will depend on availability of suitable infrastructure-controllable vehicles. Assuming a three lane road, we say that a moving bottleneck can either take two lanes, in which case it is described by flux function $Q_\Delta^{\sigma_{\min}}$, or one lane, corresponding to $Q_\Delta^{\sigma_{\max}}$, thus limiting the overtaking flow at its position to $V\sigma_{\min}$ and $V\sigma_{\max}$, respectively. This way, we are able to regulate the traffic flow and restrict it when and where it is required.

Once the flow at the bottleneck is high enough and a traffic breakdown occurs at time t_0, as described in Section 4, the controller can react by activating some vehicles on the road to act as controlled moving bottlenecks and help decongest the stationary bottleneck. We assume that the discharge rate of the static bottleneck $V\sigma_d$ lies between these two

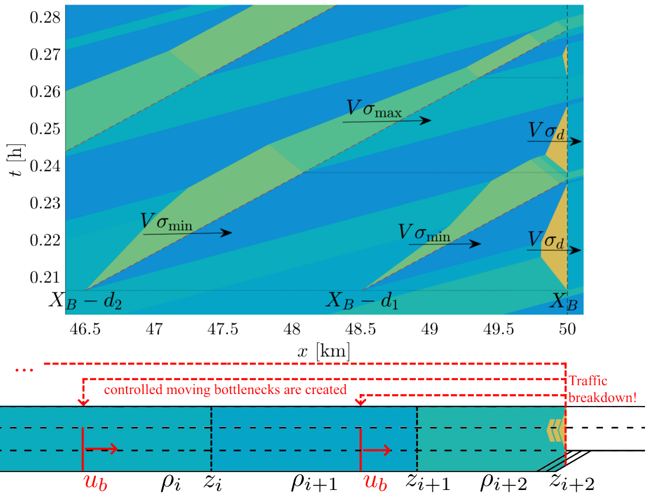


Fig. 2. An illustration of how bottleneck decongestion control is calculated. The trajectories of the bottlenecks are indicated by dashed lines and the text above the arrows indicates the intensity of traffic flow past moving and static bottleneck. Brighter colour indicates higher traffic density. Once a traffic breakdown is detected, controlled moving bottlenecks are created at desired positions.

overtaking flows, $V\sigma_{\min} < V\sigma_d < V\sigma_{\max}$. To simplify the calculation, we also assume that all moving bottlenecks move at same speed u_b.

Starting with $t = t_0$, the controlled moving bottlenecks will restrict the flow to the bottleneck by taking two lanes. Then, once we predict that the overtaking flow from a moving bottleneck will no longer feed into congestion, we change the flux function at that moving bottleneck from $Q_\Delta^{\sigma_{\min}}$ to $Q_\Delta^{\sigma_{\max}}$, allowing more vehicles to pass. As illustrated in Figure 2, the first controlled moving bottleneck is created at distance $d_1 \geq D_0$ from the bottleneck, where $D_0 > 0$ is some minimum distance and

$$d_1 \geq \tau_1(t_0) u_b = \frac{u_b}{V} \frac{\int_{X_B - d_1}^{X_B} \rho(t_0, x) \mathrm{d}x}{\sigma_d - \sigma_{\min}},$$

and $\tau_1(t_0)$ is how long the moving bottleneck is restricting the overtaking flow to $V\sigma_{\min}$. After $t = t_0 + \tau_1(t_0)$, the overtaking flow restriction is raised to $V\sigma_{\max}$, in order to start dissipating the congestion that built up behind the moving bottleneck. In order to handle the worst-case scenario, we assume that there will be an immediate traffic breakdown when this congestion reaches X_B.

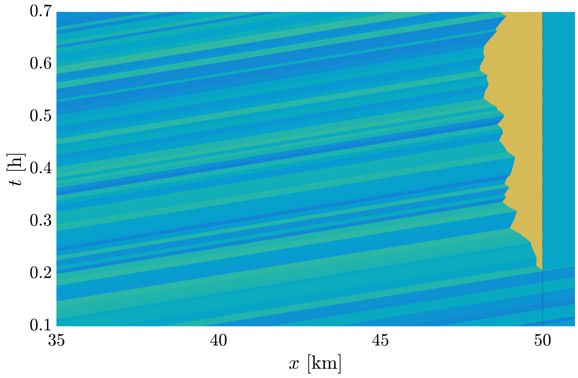
Each subsequent controlled moving bottleneck will dissipate the congestion that remains after the previous moving bottleneck reaches the X_B and the congestion built up behind it causes a new traffic breakdown. Subsequent moving bottlenecks are created with distance of at least $D_{\min}$ to the previous one, $d_i \geq d_{i-1} + D_{\min}$, at minimum distance such that

$$d_i \geq \tau_i(t_0) u_b = \frac{u_b}{V} \frac{n_i^0 + k_1 V \tau_{i-1} + k_2 d_{i-1} - \sigma_{\min} d_i}{\sigma_d - \sigma_{\max}},$$

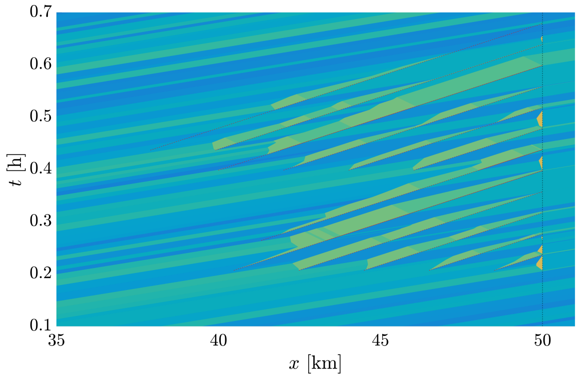
$$n_i^0(t_0) = \int_{X_B - d_i}^{X_B - d_{i-1}} \rho(t_0, x) \mathrm{d}x,$$

$$k_1 = \sigma_{\max} - \sigma_{\min},$$

$$k_2 = \sigma_{\min} + \frac{V}{u_b} (\sigma_d - \sigma_{\min}).$$



(a) No control.



(b) With control.

Fig. 3. A simulation example comparing the evolution of traffic without and with control. Brighter colour indicates higher traffic density.

The overtaking flow is limited to $V\sigma_{\min}$ for $t \in [t_0, t_0 + \tau_i(t_0)]$, and $V\sigma_{\max}$ for $t > t_0 + \tau_i(t_0)$ until the moving bottleneck i reaches X_B. Once we detect that moving bottleneck i will reach X_B with no congestion left to dissipate, i.e., when condition

$$\int_{X_B - \frac{V}{u_b} d_i}^{X_B - d_i} \rho(t_0, x) \mathrm{d}x > (V\tau_i - d_i)\,\sigma_{\min} + \left(\frac{V}{u_b} d_i - V\tau_i \right) \sigma_{\max}$$

is satisfied, no additional controlled moving bottlenecks will be created. Once the controlled moving bottlenecks are created, we update the timing of overtaking flow restriction changes every time the situation at the static bottleneck changes, e.g., if a new traffic breakdown is triggered, or if a predicted traffic breakdown is delayed.

6. SIMULATION RESULTS

The simulation results of an example are shown in Figure 3, comparing the case where we apply no control and let the traffic evolve freely, and the case where we apply bottleneck decongestion control. We consider a stretch of highway, with no on- and off-ramps and a bottleneck at position X_B. The initial density $\rho(0, x)$ is piecewise constant and randomly generated, with average value $\bar{\rho}$, resulting in a varying traffic flow at the position of the bottleneck. The time the first traffic breakdown happens (in this case at $t_0 \approx 0.21$) is taken to be the same in both cases, and the simulations run independently starting with that point.

As can be seen from Figure 3b, by delaying the arrival of a part of the traffic, we are able to maintain free flow at the

Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

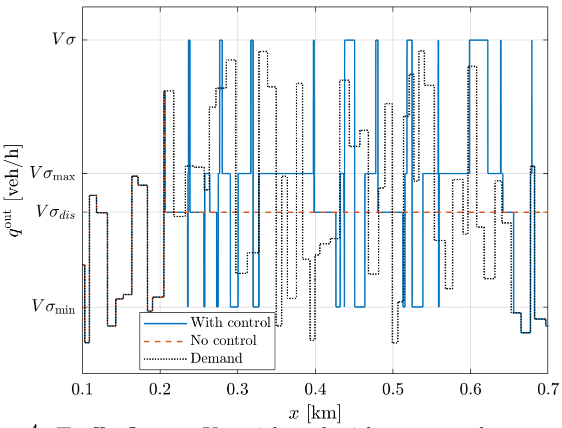


Fig. 4. Traffic flow at X_B with and without control, compared with the demand.

bottleneck. Control action is recalculated in order to react to changes at the bottleneck, and new controlled moving bottlenecks are added when needed. Since in this case the average initial traffic density is larger than the density at which the traffic flows out of the congested bottleneck, $\bar{\rho} > \sigma_d$, once a traffic breakdown happens, it is likely that congestion will persist and grow, since the average inflow to the queue will be larger than its outflow.

The flow at the position of the bottleneck is shown in Figure 4. We can see that the traffic flow follows the demand until t_0, when a traffic breakdown happens. In the controlled case, we manage to return to the unperturbed state around $t = 0.65$, whereas in the uncontrolled case the congestion at the bottleneck keeps accumulating. A total of $N_c = 2106$ vehicles was served from $t = 0.1$ to $t = 0.7$ in the controlled case, compared to $N_n = 1990$ vehicles in the uncontrolled case, corresponding to a queue of $N_q = 116$ vehicles at $t = 0.7$.

7. CONCLUSION

In this work, we propose a new transition system traffic model that is based on the front tracking solution of the LWR model. In case a piecewise linear fundamental diagram and piecewise constant initial conditions are used the presented model captures exactly the evolution of traffic, while also being computationally cheap to simulate. Additionally, the model can easily be extended, by including new types of transitions, which was demonstrated by capturing the dynamics of moving and static bottlenecks. A stochastic model of traffic breakdown at static bottlenecks is given, and an example control law was designed for bottleneck decongestion.

While effective, the described control law is derived for worst case and does not consider the stochastic nature of bottleneck dynamics. One direction for future work will be leveraging additional information to increase efficiency, e.g., by considering more complex traffic breakdown dynamics to avoid excessive traffic flow restriction after dissipating the queue at the bottleneck. This will also allow designing proactive control laws that act before the actual traffic breakdown and try to preempt it by spreading the delay more evenly even before the traffic breakdown, thus making it more likely that the traffic will stay in free flow.

REFERENCES

Baskar, L.D., De Schutter, B., Hellendoorn, J., and Papp, Z. (2011). Traffic control and intelligent vehicle highway systems: a survey. *IET Intelligent Transport Systems*, 5(1), 38–52.

Blandin, S., Work, D., Goatin, P., Piccoli, B., and Bayen, A. (2011). A general phase transition model for vehicular traffic. *SIAM journal on Applied Mathematics*, 71(1), 107–127.

Brilon, W., Geistefeldt, J., and Regler, M. (2005). Reliability of freeway traffic flow: a stochastic concept of capacity. In *Proceedings of the 16th International symposium on transportation and traffic theory*, volume 125143. College Park Maryland.

Daganzo, C.F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 28(4), 269–287.

Delle Monache, M.L. and Goatin, P. (2014). A front tracking method for a strongly coupled pde-ode system with moving density constraints in traffic flow. *Discrete and Continuous Dynamical Systems-Series S*, 7(3), 435–447.

Flynn, M.R., Kasimov, A.R., Nave, J.C., Rosales, R.R., and Seibold, B. (2009). Self-sustained nonlinear waves in traffic flow. *Physical Review E*, 79(5), 056113.

Glimm, J., Isaacson, E., Marchesin, D., and McBryan, O. (1981). Front tracking for hyperbolic systems. *Advances in Applied Mathematics*, 2(1), 91–119.

Hegyi, A., Hoogendoorn, S., Schreuder, M., Stoelhorst, H., and Viti, F. (2008). SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory. In *11th IEEE International Conference on Intelligent Transportation Systems*, 827–832.

Holden, H. and Risebro, N.H. (2015). *Front tracking for hyperbolic conservation laws*, volume 152. Springer.

Kotsialos, A., Papageorgiou, M., Diakaki, C., Pavlis, Y., and Middelham, F. (2002). Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool metanet. *IEEE Transactions on intelligent transportation systems*, 3(4), 282–292.

Lax, P.D. (1973). *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, volume 11. SIAM.

Lebacque, J.P. (1996). The Godunov scheme and what it means for first order traffic flow models. *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, 647–677.

Lighthill, M.J. and Whitham, G.B. (1955). On kinematic waves II. a theory of traffic flow on long crowded roads. *Proc. R. Soc. Lond. A*, 229(1178), 317–345.

Richards, P.I. (1956). Shock waves on the highway. *Operations research*, 4(1), 42–51.

Srivastava, A. and Geroliminis, N. (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C: Emerging Technologies*, 30, 161–177.

Tabuada, P. (2009). *Verification and control of hybrid systems: a symbolic approach*. Springer Science & Business Media.

Čičić, M. and Johansson, K.H. (2018). Traffic regulation via individually controlled automated vehicles: a cell transmission model approach. In *21st IEEE International Conference on Intelligent Transportation Systems*. Maui, US.

Yu, H. and Krstic, M. (2019). Traffic congestion control for Aw–Rascle–Zhang model. *Automatica*, 100, 38–51.