# A multiclass link transmission model for dynamic network loading of mixed legacy and automated vehicle flow

Michael W. Levin, Di Kang

Department of Civil, Environmental, and Geo- Engineering
University of Minnesota

**Abstract**

As automated vehicles gradually become available to travelers, many cities will experience a mixed traffic flow consisting of both legacy and automated vehicles. Although the overall market penetration of automated vehicles may be known, the proportion of automated vehicles may vary in space and time due to spatial and temporal variations in automated vehicle demand. Since automated vehicles are expected to behave differently than legacy vehicles, this results in a flow-density relationship that varies in both time and space with the local proportion of automated vehicles. We model this scenario using a multiclass kinematic wave theory. Assuming a triangular flow-density relationship (with shape parameters varying with the automated vehicle proportion), we develop a multiclass Newell's method for finding exact solutions to the multiclass kinematic wave theory. The solution method takes the form of a linear program with postprocessing. We then extend this method to a multiclass link transmission model. We develop a faster solution method for the receiving flow which consists of iteratively solving a system of linear equations. Numerical results from dynamic traffic assignment on the downtown Austin city network demonstrate the computational tractability of this method and explore the effects of automated vehicles on traffic congestion.

# 1 Introduction

Many previous studies have predicted that automated vehicles will behave significantly differently in traffic than legacy (human-driven) vehicles. Of course, for platooning (cooperative adaptive cruise control), automated vehicles (AVs) have a significantly different flow-density relationship (Shladover et al., 2012; Melson et al., 2018; Ye and Yamamoto, 2018). Even when automated and legacy vehicles are mixed together, previous studies have suggested that differing reaction times and car-following behaviors could lead to greater capacity (Levin and

Boyles, 2016; Ghiasi et al., 2017) and stability (Schakel et al., 2010; Talebpour and Mahmassani, 2016) of traffic flow.

Since AVs will likely comprise only a fraction of all vehicles in the network for many years, predicting the future impacts of a mixed traffic flow is a relevant problem for practitioners and policymakers. This necessitates a model in which the traffic flow varies based on the dynamic proportion of AVs. Although the overall market penetration of AVs in a city may be known, spatial and temporal variations in demand and turning movements may result in space- and time-varying AV proportions that differ significantly from the network-wide market penetration. For instance, AVs may be substantially more expensive than legacy vehicles. Higher-income residences in the network may generate a significantly higher proportion of AV trips than lower-income residences, resulting in spatial variations across the network.

Perhaps the most obvious modeling approach is to construct microsimulation models in which individual vehicles use different car-following models based on their driver type (legacy or automated) (e.g. Schakel et al., 2010; Zhu and Zhang, 2018). Unfortunately, the detailed dynamics within microsimulation require high computation times for large city networks. Solving dynamic traffic assignment (Chiu et al., 2011) to predict the effects of route choice on traffic congestion further increases the computation time. Solving traffic assignment is important because Braess (1968) and Daganzo (1998) showed that improvements to the network capacity could result in greater congestion, which Melson et al. (2018) demonstrated for AVs specifically.

Although they are not as detailed as microsimulation, dynamic network loading models based on the kinematic wave theory of traffic flow (Lighthill and Whitham, 1955; Richards, 1956) are used in traffic flow modeling for their ability to model time-dependent traffic flow with manageable computation times. For instance, the dynamic traffic assignment model VISTA (Ziliaskopoulos and Waller, 2000) uses the cell transmission model (Daganzo, 1994, 1995) to propagate flow. We propose a multiclass kinematic wave theory in which the flow-density relationship varies in space and time with respect to the local proportion of AVs. Levin and Boyles (2016) developed a multiclass cell transmission model to solve the multiclass kinematic wave theory and used it for dynamic network loading. However, the cell transmission model is only an approximate solution to the kinematic wave theory and contains numerical errors, particularly in the propagation of congested states. Newell (1993) observed that the kinematic wave theory is easily solved along certain characteristic curves, which Yperman et al. (2005) converted into the link transmission model (LTM). We develop a multiclass LTM for the multiclass kinematic wave theory, which is made challenging by the fact that characteristic curves may have to pass through space- and time-dependent variations in the flow-density relationship.

The kinematic wave theory can be alternatively solved by transforming it to Lagrangian coordinates. For instance, van Wageningen-Kessels et al. (2009, 2010) used Lagrangian coordinates to solve a multiclass kinematic wave theory, although their work did not consider a mix of human and automated vehicles changing the flow-density relationship. Lagrangian coordinates are definitely worth consideration for modeling a mix of human and AV flows. However, our goal is to construct a model suitable for dynamic network loading in a network of links and nodes. To the best of the authors' knowledge, all kinematic wave theory-based models for dynamic network loading on realistic city networks are solved using Eulerian coordinates. This is because intersections would be "moving" in the Lagrangian perspective while in Eulerian coordinates they remain stationary. Therefore, we have chosen to address this problem with Eulerian coordinates, and we construct a link flow model which is compatible with the large body of existing work on dynamic network loading and dynamic traffic assignment based on the kinematic wave theory.

The contributions of this paper are as follows. We formulate a multiclass kinematic wave theory and derive some general results that assist with the structure of the solution method. We then assume that the flow-density relationship is triangular, with the congested wave speed varying with the proportion of automated vehicles. This assumption is used to derive a multiclass version of Newell (1993)'s method to find exact solutions to the multiclass kinematic wave theory. The solution method involves solving a linear program then post-processing the optimal solution with other constraints to find an exact value for the cumulative count at an arbitrary $(t, x)$ point. The primary motivation of this paper is to extend the multiclass Newell's method to a multiclass LTM. We present a faster solution algorithm for the multiclass LTM which involves iteratively solving a system of linear equations. Numerical results on a calibrated large city network demonstrate solving a dynamic traffic assignment model using this multiclass LTM, and explore the impacts of AV market penetration on city-wide congestion.

The remainder of this paper is organized as follows. Section 2 reviews literature relevant to multiclass modeling of traffic flow, in particular automated vehicles. In Section 3, we formally define the multiclass kinematic wave theory. Section 4 derives the multiclass Newell's method, first presenting a simplified example, then presenting the general solution. Section 5 adapts the multiclass Newell's method to the specific characteristics of LTM. Numerical results from dynamic traffic assignment are presented in Section 6, and we conclude in Section 7.

# 2 Background

For decades, traffic assignment models have modeled travel times as a continuous function of traffic volumes for their favorable analytical properties (Beckmann et al., 1956). Such models have recently been extended to mixed legacy and AV flow. Levin and Boyles (2015) modified the standard Bureau of Public Roads travel time function to include a capacity that varied with the proportion of AVs. However, static traffic assignment models cannot capture how time-varying proportions of automated vehicles affect traffic congestion. Although the overall proportion of AVs in a network may be known, local variations in the proportion affect the local capacity and propagation of congested waves. The effects of time-varying congestion (separate from AV flow differences) have long been recognized and modeled through microsimulation or mesoscopic flow models such as the kinematic wave theory (Lighthill and Whitham, 1955; Richards, 1956). Research on efficient solutions to the kinematic wave theory has been used for dynamic network loading, yielding dynamic traffic assignment models that can predict time-dependent congestion across city networks (Chiu et al., 2011). The multiclass flow of a mix of legacy and automated vehicles creates additional challenges for these network models. In addition to time-varying densities, the flow-density relationship itself can vary in space and time with the proportion of AVs present. These variations in the flow-density relationship move with the traffic flow, and are therefore interdependent on the solution to the kinematic wave theory.

Modeling multiclass flow is important because of the demonstrated behavioral differences of AVs. Results from microsimulation (Kesting et al., 2007; Davis, 2007; Shladover et al., 2012; Zhu and Zhang, 2018; Liu et al., 2018) and live demonstrations (Ploeg et al., 2011; Milanés et al., 2013) on homogeneous segments of road have shown increases in capacity from cooperative adaptive cruise control (Shladover et al., 2015) that reduces following headways. It is not clear whether these capacity improvements apply to networks as well, but these studies nevertheless demonstrate that multiclass flow may behave differently from legacy vehicle flow. Connected and cooperative vehicle control has also resulted in improvements in string stability (Schakel et al., 2010; Ploeg et al., 2011; Talebpour and Mahmassani, 2016), although these results may not be true of non-cooperative AVs. One major limitation demonstrated traffic flow improvements is driver comfort (Nowakowski et al., 2010). Improvements in capacity and stability from automated driving largely result from lower reaction times admitting shorter following headways (Kesting and Treiber, 2008; Chen et al., 2017; Makridis et al., 2019). Until vehicle manufacturers and drivers are comfortable with short headways, capacity increases may be limited (Vander Werf et al., 2002; Shladover et al., 2012; Makridis et al., 2019), or capacity may even decrease (Calvert et al., 2017; James et al., 2019). Nevertheless, most studies suggest significant (positive or negative) changes in traffic flow resulting from AVs. Some of these traffic flow differences are assumed to be limited to AV-exclusive lanes (Chen et al., 2016; Vander Laan and Sadabadi, 2017; Liu and Song, 2019; Melson et al., 2018) where connectivity and automation are minimally interrupted by human driving. Other studies have explored technologies for efficient automated driving within mixed traffic flow (Kesting et al., 2007; Gong and Du, 2018) hat vary with the AV market penetration (Talebpour and Mahmassani, 2016; Zhu and Zhang, 2018; Ye et al., 2018). This paper focuses on the latter category. Since the exact characteristics of future automated driving are difficult to determine, the focus of this study is on solving a multiclass kinematic wave theory where the flow-density relationship can be adjusted to model AV differences in traffic flow.

Although many previous studies have used car-following models of automated driving to model traffic flow differences, the computation times required for these models often discourage their use for solving dynamic traffic assignment on large city networks. Static multiclass traffic assignment models (Levin and Boyles, 2016; Wang et al., 2019) have been studied but fail to capture the time-varying changes in congestion, especially due to the varying flow-density relationship. The kinematic wave theory is often an useful middle ground with dynamic traffic flow but more computational efficiency than microsimulation. Thus far, multiclass variants of the kinematic wave theory have only been solved by approximation methods like the cell transmission model. This is partially because solving the standard kinematic wave theory is challenging by itself. Although the kinematic wave theory was proposed decades ago by Lighthill and Whitham (1955) and Richards (1956), efficient solutions were discovered much later. The cell transmission model (Daganzo, 1994, 1995), a Godunov (1959) approximation, is still widely used for dynamic network loading. Newell (1993)'s method for finding exact solutions with a triangular flow-density relationship was used in the relatively recent development of LTM (Yperman et al., 2005). A Lax-Hopf method for exact solutions to the general kinematic wave theory was developed by Claudel and Bayen (2010a,b).

Given the difficulty in solving the standard kinematic wave theory, it is perhaps not surprising that most previous work on multiclass kinematic wave theory has relied on numerical approximation methods. Levin and Boyles (2016) developed a multiclass cell transmission model which was used by Patel et al. (2016) to analyze city networks. Tiaprasert et al. (2017) developed an extended multiclass cell transmission model with better modeling of first-in-first-out and overtaking behaviors. Zhu and Ukkusuri (2018) and Chen et al. (2020) used multiclass cell transmission models to study how variable speed limit control for AVs would affect surrounding legacy vehicle traffic.

The multiclass kinematic wave theory has previously been studied outside of the context of AVs (Logghe and Immers, 2008; Jin, 2012; van Wageningen-Kessels et al., 2014). The multiple classes could be motivated by heterogeneity in physical vehicle characteristics (Van Lint et al., 2008; Liu et al., 2015) or driver behavior (Van Aerde and Rakha, 1995). van Wageningen-Kessels et al. (2009, 2010) used a Lagrangian transformation to solve the problem. To the best of our knowledge, a LTM for efficiently solving other multiclass kinematic wave theory problems has yet to be proposed. Although the motivation and context for this paper is focused on mixed legacy and automated vehicle traffic, the ideas may be relevant to the multiclass kinematic wave theory in other contexts.

# 3 Multiclass kinematic wave theory

Consider a length of road from $x = 0$ to $x = L$. Let $N^m(t, x)$ be the cumulative count of vehicles of class $m$, i.e the number of vehicles that have passed point $x$ at or before time $t$. The total cumulative counts are given by $N(t, x) = \sum\limits_{m \in \mathcal{M}} N^m(t, x)$ where $\mathcal{M}$ is the set of all vehicle classes. Let $q^m(t, x) = \frac{\partial N^m(t,x)}{\partial t}$ and $k^m(t, x) = -\frac{\partial N^m(t,x)}{\partial x}$ be the flow and density of class $m$, respectively. For completeness we define the total flow and density as $q(t, x) = \sum\limits_{m \in \mathcal{M}} q^m(t, x)$ and $k(t, x) = \sum\limits_{m \in \mathcal{M}} k^m(t, x)$, respectively. Assume that the cumulative counts for each class satisfy a conservation law, i.e.

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

which results in a conservation law for the total cumulative counts also. Let

$$\mathbf{q}(t, x) = \mathbf{f}(\mathbf{k}(t, x)) \tag{2}$$

be the vector-valued flow-density relationship which specifies the flow of each class as a function of the class-proportions of density $\frac{k^m(t,x)}{k}$ and total density. For the results in Section 3, $\mathbf{f}(\mathbf{k})$ may be a general flow-density relationship with any concave shape. Assume that $\mathbf{f}(\mathbf{k})$ is not separable by class, so that the partial differential equations cannot be solved separately. For instance, flow may be 0 when the total density $\sum_{m \in \mathcal{M}} k^m(t,x)$ is equal to some maximum density value, meaning the flow of one class depends on the density of other classes. Assume first-in-first-out (FIFO) behavior within the link.

The problem is to find $N^m(t,x)$ for any $(t,x)$ point given initial and boundary conditions. Initial condition $\mathbf{c}(0,x)$ specifies the initial cumulative counts at time $t=0$. Upstream boundary conditions $\mathbf{c}(t,0)$ and downstream boundary conditions $\mathbf{c}(t,L)$ specify restrictions on entering and exiting flow, respectively. These conditions are generally relevant. LTM (Yperman et al., 2005) uses only upstream and downstream boundary conditions, and the generalized Lax-Hopf formula (Claudel and Bayen, 2010a,b) is constructed for similar initial and boundary conditions, albeit for a single class. For formality, we observe that $\mathbf{c(t,x)}$ can be defined everywhere, except with $\mathbf{c}(t,x) = \boldsymbol{\infty}$ where $(t,x)$ is not an initial point ($t=0$), upstream point ($x=0$), or downstream point ($x=L$). The class-specific values are necessary to propagate class proportions according to the specified flow-density relationship $\mathbf{f}(\mathbf{k})$. For instance, Levin and Boyles (2016) proposed a flow-density relationship for mixed human and automated-vehicle traffic flow that varies with the proportion of automated vehicles.

Claudel and Bayen (2010a) observed that the single-class kinematic wave theory can be written as a Hamilton-Jacobi partial differential equation:

$$\frac{\partial N(t,x)}{\partial t} - \psi\left(-\frac{\partial N(t,x)}{\partial x}\right) = 0 \tag{3}$$

which leads to an exact formula for solving $N(t,x)$ (for a single class):

$$N(t,x) = \inf_{u,T \geq 0} \{c(t-T, x+Tu) + T\varphi^*(u)\} \tag{4}$$

where $c(t,x)$ defines the conditions imposed on the solution, and $\varphi^*(u)$ is the Legendre-Fenchel transform of the flow-density relationship. For a triangular flow-density relationship, this formula simplifies to Newell (1993)'s method (Jin, 2015). Unfortunately, neither the Hamilton-Jacobi theory nor the method of characteristics are designed to propagate class proportions in time and space, in addition to flows and densities. Hence, Levin and Boyles (2016) proposed a cell transmission model to find an approximate solution. However, due to numerical errors in such approximations, an exact solution is preferable.

## 3.1 Class proportion regions

Assume that the order of vehicles does not change within $x \in [0,L]$, i.e. each vehicle can be labeled with a cumulative count number that remains constant within $x \in [0,L]$. This is equivalent to assuming first-in-first-out behavior. In reality, this assumption is broken by lane-changing. However, within the kinematic wave theory where vehicle speed is a function of density, this assumption is reasonable. This assumption ensures that class proportions move with vehicles. Shockwaves occur in the standard (single class) kinematic wave theory as the separation of regions of different density. Some shockwaves move with vehicles, and others move backwards in space, meaning that vehicles can cross a shockwave boundary. In the multiclass kinematic wave theory, shockwaves separate regions of different total density as well as regions of different class proportions.

Since class proportions move with vehicles, the class proportions can be described with respect to vehicle labels. Consider $I$ different regions of class proportions. Region $i$ is defined as the vehicles between vehicle $b_{i-1}$ and $b_i$ which are the boundary vehicles. For instance, a region might be defined as the vehicles that enter the upstream boundary between a specified time, i.e. $b_{i-1} = N(t_{i-1}, 0)$ and $b_i = N(t_i, 0)$ where $t_{i-1} < t_i$. Within a region, vehicle class proportions are constant. However, note that class proportions are not all-or-nothing within a region. It is fully possible to have $k^1(t,x) > 0$ and $k^2(t,x) > 0$ with a corresponding flow-density relationship $\mathbf{f}(\mathbf{k})$. The key defining property is that the proportion $\frac{k^m(t,x)}{\sum\limits_{m'} k^{m'}(t,x)}$ is constant throughout each region. For example, this could be used to model a region of density with 50% legacy vehicles and 50% AVs, uniformly distributed throughout the region.

**Proposition 1.** *When traffic flow obeys first-in-first-out behavior, class proportion regions move with vehicles.*

*Proof.* Suppose that at $x = 0$, the class proportions between $b_1$ and $b_2$ are $\frac{N^m(t_2,0)-N^m(t_1,0)}{N(t_2,0)-N(t_1,0)}$ where $N(t_1, 0) = b_1$ and $N(t_2, 0) = b_2$. At any points $(t'_1, x'_1)$ and $(t'_2, x'_2)$ such that $x'_1 \leq L$ and $x'_2 \leq L$ with $N(t'_1, x'_1) = b_1$ and $N(t'_2, x'_2) = b_2$,

$$\frac{N^m(t'_2, x'_2) - N^m(t'_1, x'_1)}{N(t'_2, x'_2) - N(t'_1, x'_1)} = \frac{N^m(t_2, 0) - N^m(t_1, 0)}{N(t_2, 0) - N(t_1, 0)} \tag{5}$$

due to FIFO. □

Each region has an unique flow-density relationship $q(t, x) = f_i(k(t, x))$ that is specific to the class proportions of region $i$. When class proportions are constant within regions, the class-dependent flow-density relationship $\mathbf{f}(\mathbf{k})$ can be simplified to a scalar function $q(t, x) = f_i(k(t, x))$ that describes the flow within that region. Similarly, the initial and boundary conditions can be written as a scalar $c(t, x)$, with the $b_i$ labels indicating the boundary vehicles between different class proportion regions. Figure 1 shows space-time trajectories of vehicles of two classes as they enter a queue. The speed of the shockwave separating the queue from the uncongested region behind it varies based on the flow-density relationship of the vehicle class. However, the class regions move with the traffic. Consequently, the multiclass cumulative counts $N^m(t, x)$ can be simplified to tracking a single cumulative count $N(t, x)$ while also tracking the size and location of class proportion regions as function of time.

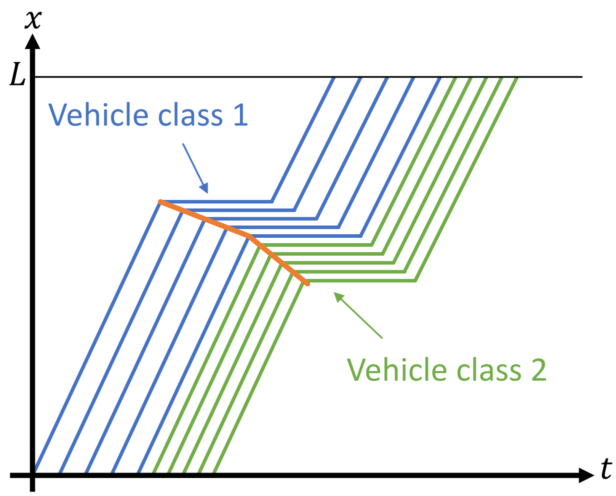



Figure 1: Illustration of space-time diagram for vehicles of two different classes, with different flow-density relationships.

### 3.2 Space- and time-varying flow-density relationship

The standard kinematic wave theory is typically used with a space-varying flow-density relationship (varying across links due to capacity and other factors). A space- and time-varying flow-density relationship is admissible also. The flow-density relationship $f_i$ is included via the Legendre-Fenchel transform

$$\varphi_i^*(u_i) = \sup_{k \in \mathrm{Dom}(f_i)} \{ku_i + f_i(k)\} \tag{6}$$

corresponding to region $i$ (Claudel and Bayen, 2010a). Notice that the boundaries of these space- and time-varying regions depend on vehicle trajectories. This is not limiting because the trajectory of the lower bound of region $i$ (here lower bound refers to the first vehicle of region $i$ which has the lowest cumulative count label among all vehicles in $i$) depends on the movements of vehicles ahead of it, i.e. vehicles in region $i-1$. Hence the location of the boundary between regions $i-1$ and $i$ can be known before computing the cumulative counts within region $i$. Figure 2 illustrates how the class regions might affect the search along characteristics. Passing the boundary between regions 1 and 2, a different congested wave speed is used for the characteristic speed. This relationship is formalized in Proposition 2.

**Proposition 2.** *Let $c(t,x)$ describe the initial or boundary conditions. The Lax-Hopf formula for a flow-density relationship that varies for $I$ different space-time regions is*

$$N(t,x) = \inf_{u_i \in [-v_i, w_i], T_i \in \mathbb{R}_+} \left\{ c\left(t - \sum_{i=1}^{I} T_i, x + \sum_{i=1}^{I} T_i u_i\right) + \sum_{i=1}^{I} T_i \varphi_i^*(u_i) \right\} \tag{7}$$

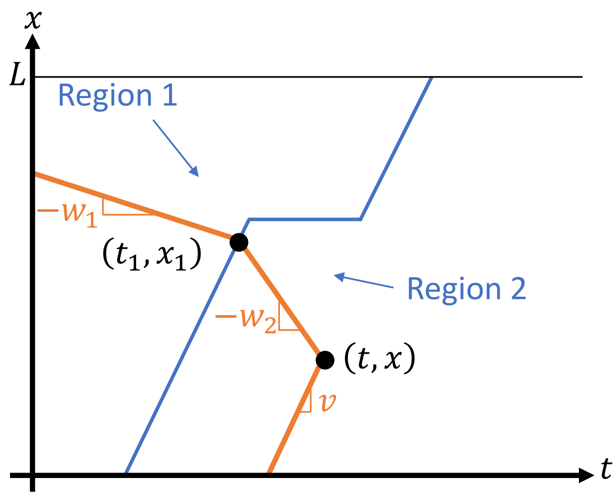



Figure 2: Illustration of searching along characteristics to evaluate $N(t, x)$.

*such that the point* $\left(t - \sum_{j=I}^{i+1} T_j, x + \sum_{j=I}^{i+1} T_j u_j\right)$ *corresponds to the trajectory of the boundary vehicle* $b_i$*.*

*Proof.* By induction on $I$. **Basis:** By Theorem 3.1 of Claudel and Bayen (2010a), the Lax-Hopf formula for a single region is

$$N(t,x) = \inf_{u_1 \in [-v_1, w_1], T_1 \in \mathbb{R}_+} \{c(t - T_1, x + T_1 u_1) + T_1 \varphi_1^*(u_1)\} \tag{8}$$

$$= \inf_{u_1 \in [-v_1, w_1], T_1 \in \mathbb{R}_+} \left\{ c\left(t - \sum_{i=1}^{1} T_i, x + \sum_{i=1}^{1} T_i u_i\right) + \sum_{i=1}^{1} T_i \varphi_i^*(u_i) \right\} \tag{9}$$

**Inductive step:** Given the space-time trajectory of vehicle $b_I$, the Lax-Hopf formula for region $I+1$ is

$$N(t,x) = \inf_{u_{I+1} \in [-v_{I+1}, w_{I+1}], T_{I+1} \in \mathbb{R}_+} \{c(t - T_{I+1}, x + T_{I+1} u_{I+1}) + T_{I+1} \varphi_{I+1}^*(u_{I+1})\} \tag{10}$$

$$= \inf_{u_{I+1} \in [-v_{I+1}, w_{I+1}], T_{I+1} \in \mathbb{R}_+} \{b_I + T_{I+1} \varphi_{I+1}^*(u_{I+1})\} \tag{11}$$

because the lower boundary of region $I+1$ has a cumulative count number of $b_I$. The space-time trajectory of $b_I$ is determined by the Lax-Hopf formula for $I$ regions, i.e. equation (7), which then yields

$$N(t,x) = \inf_{u_i \in [-v_i, w_i], T_i \in \mathbb{R}_+} \left\{ c\left(t' - \sum_{i=1}^{I} T_i, x' + \sum_{i=1}^{I} T_i u_i\right) + \sum_{i=1}^{I} T_i \varphi_i^*(u_i) + T_{I+1} \varphi_{I+1}^*(u_{I+1}) \right\} \tag{12}$$

with $(t', x')$ on the space-time trajectory of $b_I$. Equation (12) can be simplified to

$$N(t,x) = \inf_{u_i \in [-v_i, w_i], T_i \in \mathbb{R}_+} \left\{ c\left( t - \sum_{i=1}^{I+1} T_i, x + \sum_{i=1}^{I+1} T_i u_i \right) + \sum_{i=1}^{I+1} T_i \varphi_i^*(u_i) \right\} \tag{13}$$

for $I+1$ regions. $\square$

Notice that Proposition 2 does not depend on the shape of the flow-density relationship (other than requiring concavity). As discussed in Claudel and Bayen (2010a), Jensen's inequality simplifies the search over any curve connecting $(t, x)$ to a boundary point $(t', x')$ such that $c(t', x')$ is defined.

Given an arbitrary $(t, x)$ point, the Lax-Hopf formula (7) can detect when $(t, x)$ lies outside of region $i$. Proposition 3 shows that if $(t, x)$ is outside of region $i$, then calculating $N(t, x)$ through equation (7) using the first $i$ regions will return a value greater than $b_i$, indicating that $(t, x)$ lies beyond region $i$. It may not return the correct value of $N(t, x)$, but the relationship $N(t, x) > b_i$ is sufficient to determine that regions $i+1$, $i+2$, $i+3$, ... need to be included to obtain the correct $N(t, x)$ value.

**Proposition 3.** *Suppose that $(t, x)$ lies outside of region $i$, i.e. $N(t, x) > b_i$. Then the Lax-Hopf formula (7) using regions $1 \ldots i$ will return a value $N'(t, x) > b_i$.*

*Proof.* Since vehicle $b_i$ is within region $i$, the trajectory of $b_i$ can be determined by the Lax-Hopf formula (7) using the first $i$ regions. Therefore, if $N(t, x) > b_i$, the Lax-Hopf formula (7) must return a (possibly incorrect) value $N'(t, x) > b_i$ since the vehicle label at $(t, x)$ is greater than $b_i$. $\square$

# 4 Multiclass Newell's method

For the remainder of this paper, like Newell (1993) we assume that the flow-density relationship $f_i(k)$ for each region $i$ has the triangular shape

$$f_i(k) = \min\left\{vk, -w_i(k - K)\right\} \tag{14}$$

where $K$ is the jam density, $w_i$ is the congested wave speed, and $v$ is the free flow speed. Sections 4–4.2 and 5 apply for any triangular flow-density relationship. Only the numerical results in Sections 4.3 and 6 use a specific shape of the flow-density relationship from Levin and Boyles (2016).

To further simplify the problem, we make assumptions based on the expected nature of our intended context, i.e. a mixture of human and automated vehicles. Since $K$ depends on the physical length of vehicles, we assume that $K$ is independent of the class proportions. We also assume that $v$ is independent of the class proportions because it is based on the speed limit for the road. A single free flow speed value of $v$ simplifies the search among forward-moving characteristics, but the method could easily be generalized to a region-specific free flow speed. However, $w_i$ is expected to vary based on class proportions. Levin and Boyles (2016) predicted that higher $w_i$ values were likely for automated vehicles, resulting in correspondingly higher capacities. Let $k_i^c$ be the critical density of region $i$ corresponding to capacity $Q_i = vk_i^c$. The critical densities and capacities also depend on the class proportions because they vary with the congested wave speed. The Legendre-Fenchel transform of $f_i(k)$ is

$$\varphi_i^*(u) = \sup_{k \in \mathrm{Dom}(f_i)} \{ku + f_i(k)\} = vk_i^c + uk_i^c \tag{15}$$

defined for wave speeds $u \in [-v, w_i]$. As observed by Newell (1993) and Himpe et al. (2016), since $\varphi_i^*(u)$ is linear, the infimum in equation (7) occurs at a characteristic wave speed of $u = -v$ or $u = w_i$.

## 4.1 Two-region multiclass Newell's method

To develop intuition about the general solution method, we first present the multiclass Newell's method for two regions. The upper bound of region 1 is $b_1$, so region 1 is active for vehicles $[-\infty, b_1]$, and region 2 is active for vehicles $(b_1, \infty]$. As illustrated in Figure 2 and proven in Proposition 2, the search along backwards-moving characteristics may involve two different congested wave speeds $w_1$ and $w_2$ corresponding to regions 1 and 2 of the flow-density relationship. Since the free flow speed is uniformly $v$, forward-moving characteristics always move at speed $v$, which avoids the need to adjust the speed of the forward-moving characteristics when passing region boundaries for this problem.

The objective is to find $N(t, x)$ for an arbitrary point $(t, x) \in \mathbb{R}_+ \times [0, L]$. With two class-proportion regions, $(t, x)$ could fall into either of the two regions. It is not known a priori which region $(t, x)$ belongs to, as that depends on the trajectory of the boundary vehicle $b_1$. By Proposition 3, the first step is to calculate $N(t, x)$ using only region 1:

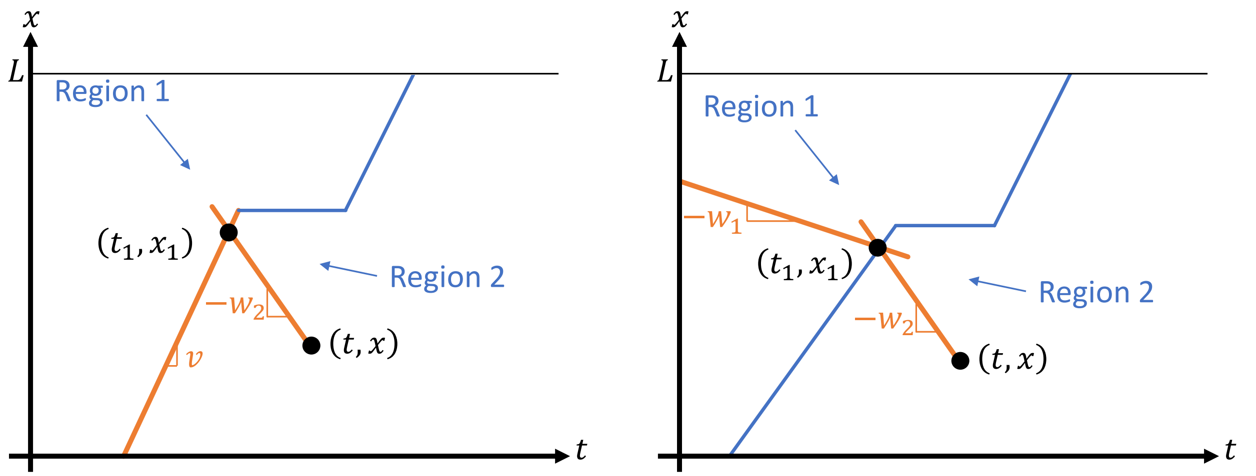
$$N(t, x) = \inf_{u_1 \in [-v, w_1], T_1 \in \mathbb{R}_+} \{c(t - T_1, x + T_1 u_1) + T_1 \varphi_1^*(u_1)\} \tag{16}$$

This can be evaluated using Newell's method, tracing the forward- and backward-moving characteristics to the known condition $c(t, x)$. If the returned value of $N(t, x)$ satisfies $N(t, x) \leq b_1$, then $(t, x)$ is in region 1, and the returned $N(t, x)$ value is correct. Otherwise, $N(t, x)$ must be calculated using region 2.

Because the triangular flow-density relationship is used, there are two possible characteristic speeds to check: $v$ and $w_i$ (Newell, 1993). Figure 2 illustrates that the problem is to find where the trajectory of vehicle $b_1$ intersects the line of slope $-w_2$ at point $(t_1, x_1)$ satisfying

$$x_1 - x = -w_2(t_1 - t) \tag{17}$$

The point $(t, x)$ is where we want to find the cumulative count $N(t, x)$. $(t_1, x_1)$ is the point on the trajectory of the boundary vehicle $b_1$ intersecting a line of slope $-w_2$ from $(t, x)$. However, the trajectory of $b_1$ may not be known exactly. That trajectory is determined as the set of $(t, x)$ points such that $N(t, x) = b_1$ — which can be evaluated using the Lax-Hopf formula (7). Figure 3 illustrates this search for $(t_1, x_1)$ along the two characteristic lines of




(a) Uncongested characteristic

(b) Congested characteristic

Figure 3: Illustration of searching along characteristics to evaluate $(t_1, x_1)$.

slope $v$ and $-w_1$ that determine $N(t_1, x_1)$. These result in two systems of equations that can be evaluated separately.

#### 4.1.1 Uncongested characteristic for $(t_1, x_1)$

The problem is to find $(t_1, x_1)$ such that $N(t_1, x_1) = b_1$, $(t_1, x_1)$ lies on the characteristic of slope $-w_2$ passing through $(t, x)$, and that $(t_1, x_1)$ lies on a characteristic of slope $v$ such that $c(t_0, x_0) = b_1$ and

$$x_0 - x_1 = v(t_0 - t_1) \tag{18}$$

Equations (17) and (18) together have variables $(t_0, x_0)$ and $(t_1, x_1)$. However, the point $(t_0, x_0)$ can be simplified by splitting it into two cases: either $(t_0, x_0)$ is part of an initial condition with $t_0 = 0$, or $(t_0, x_0)$ is part of an upstream condition with $x_0 = 0$. Since $N(t, x)$ is constant along a characteristic of slope $v$ (Newell, 1993), in either case the remaining variable of $x_0$ or $t_0$ can be determined by evaluating where $c(t_0, x_0) = b_1$. That results in two equations (17) and (18) with two variables $(t_1, x_1)$, which can be solved to obtain solution $(t_1^{\mathrm{u}}, x_1^{\mathrm{u}})$ where the superscript u is used to denote the solution from the uncongested characteristic.

#### 4.1.2 Congested characteristic for $(t_1, x_1)$

When a second congested characteristic of slope $-w_1$ determines $(t_1, x_1)$, the second equation to determine point $(t_0, x_0)$ on the initial or boundary condition becomes

$$x_0 - x_1 = -w_1(t_0 - t_1) \tag{19}$$

Since $N(t, x)$ increases at the rate of $K\Delta x$ along backwards-moving characteristics (Newell, 1993), $N(t_1, x_1) - N(t_0, x_0) = K(x_0 - x_1)$. The point $(t_1, x_1)$ satisfies $N(t_1, x_1) = b_1$, but since $x_1$ is a variable, the point $(t_0, x_0)$ cannot be determined by $b_1$ alone. Instead, a third equation can be written:

$$c(t_0, x_0) + K(x_0 - x_1) = b_1 \tag{20}$$

The system of equations (17), (19), and (20) has 3 variables — $(t_1, x_1)$ and $(t_0, x_0)$ with $t_0 = 0$ (initial condition) or $x_0 = L$ (downstream boundary condition). Solving this system yields $(t_1^{\mathrm{c}}, x_1^{\mathrm{c}})$ where the superscript c is used to denote the solution from the congested characteristic. If multiple solutions exist, the correct solution minimizes $b_1 + K(x' - x_1)$ due to the infimum in the Lax-Hopf formula (7).

### 4.1.3 Combining $(t_1^{\mathrm{u}}, x_1^{\mathrm{u}})$ and $(t_1^{\mathrm{c}}, x_1^{\mathrm{c}})$

Due to the infimum, the correct solution is the characteristic with the lowest $N(t, x)$ value. Since $b_1$ is a constant vehicle label, we have two possible values for the point where the trajectory of $b_1$ intersects the line with slope $-w_2$ through $(t', x')$: $(t_1^{\mathrm{u}}, x_1^{\mathrm{u}})$ and $(t_1^{\mathrm{c}}, x_1^{\mathrm{c}})$. Since $\frac{\partial N}{\partial t} \geq 0$ and $\frac{\partial N}{\partial x} \leq 0$, and the congested characteristic through $(t', x')$ has a negative slope of $-w_2$, $N(t, x)$ is nondecreasing along that characteristic. Therefore, the correct value of $(t_1, x_1)$ for the trajectory of vehicle 1 is $t_1 = \max\{t_1^{\mathrm{u}}, t_1^{\mathrm{c}}\}$ since that corresponds to a later arrival time of vehicle $b_1$, or equivalently, a lower value of $N(t, x)$ along the characteristic of slope $-w_2$.

There are two particular corner cases associated with the value $(t_1, x_1)$. First, observe that $x_1 > x$ (and $t_1 < t$). Otherwise, point $(t, x)$ is in region 1, and that case was handled already. Second, it is possible that $x_1 > L$, i.e. the intersection of the backwards-moving characteristic of slope $-w_2$ with the trajectory of vehicle $b_1$ is outside of the range $x \in [0, L]$. In that case, $c(t, x)$ is the limiting condition, and the standard Lax-Hopf formula for region 2 can be used:

$$N(t, x) = \inf_{u_2 \in [-v, w_2], T_2 \in \mathbb{R}_+} \{c(t - T_2, x + T_2 u_2) + T_2 \varphi_2^*(u_2)\} \tag{21}$$

Once the correct value of $(t_1, x_1)$ is determined, with $x < x_1 \leq L$, then the maximum value of $N(t, x)$ based on the congested characteristic is $b_1 + K(x - x_1)$ (Newell, 1993) since $N(t_1, x_1) = b_1$ and the cumulative count increases by $K\Delta x$ along a characteristic moving at speed $-w_2$ in region 2. The second bound on $N(t, x)$ is based on the forward moving characteristic: finding point $(t_0^{\mathrm{u}}, x_0^{\mathrm{u}})$ such that $c(t_0^{\mathrm{u}}, x_0^{\mathrm{u}})$ is defined (or less than $\infty$) and $(t_0^{\mathrm{u}}, x_0^{\mathrm{u}})$ is on a line of slope $v$ passing through $(t, x)$, i.e. $x_0^{\mathrm{u}} - x = v(t_0^{\mathrm{u}} - t)$. Since the cumulative count does not change along this characteristic, $N(t, x)$ can be calculated as

$$N(t, x) = \min\{N(t_1, x_1) + K(x - x_1), N(t_0^{\mathrm{u}}, x_0^{\mathrm{u}})\} \tag{22}$$

## 4.2 Multi-region multiclass Newell's method

The problem is to find points $(t_i, x_i)$ such that $N(t_i, x_i) = b_i$. As suggested in Section 4.1, these points form a system of linear equations as they are connected via characteristic lines of slope $-w_i$. With multiple regions,the calculation of the backwards-moving characteristic is more complicated because there are multiple characteristics. The objective is to find a point $(t_i, x_i)$ on the characteristic of slope $-w_{i+1}$ through point $(t_{i+1}, x_{i+1})$ such that $N(t_i, x_i) = b_i$. We first discuss how to handle a downstream boundary condition (defined for $x = L$), then discuss the changes needed for an initial condition (defined for $t = 0$) in Section 4.2.3.

### 4.2.1 Downstream boundary condition

For a downstream boundary condition, the objective is to find $T_i \geq 0$ for all $i \in [1, I]$ such that

$$t_0 = t - \sum_{i=1}^{I} T_i \tag{23}$$

and

$$x_0 = x + \sum_{i=1}^{I} T_i w_i = L \tag{24}$$

where $x_0 = L$ because backwards moving waves intersect the downstream boundary of $L$. The intermediate points where region $i$ ends are labeled $(t_i, x_i)$ and can be found by

$$t_i = t - \sum_{j=i+1}^{I} T_j \tag{25}$$

and

$$x_i = x + \sum_{j=i+1}^{I} T_j w_j \tag{26}$$

Proposition 4 formally proves that defining $(t_i, x_i)$ through $T_i$ as in equations (25) and (26) achieves a point $(t_i, x_i)$ on the line of slope $-w_{i+1}$ through point $(t_{i+1}, x_{i+1})$.

**Proposition 4.** *For $i \in [2, I-1]$, point $(t_{i-1}, x_{i-1})$ is on the line with slope $-w_i$ through point $(t_i, x_i)$. Point $(t, x)$ is on the line with slope $-w_I$ through point $(t_{I-1}, x_{I-1})$.*

*Proof.* Point $(t_{I-1}, x_{I-1})$ is defined by equations (25) and (26) as

$$t_{I-1} = t - \sum_{i=I}^{I} T_i = t - T_I \tag{27}$$

and

$$x_{I-1} = x + \sum_{i=I}^{I} T_i w_i = x + T_I w_I \tag{28}$$

Therefore $t - t_{I-1} = T_I$ and $x - x_{I-1} = -T_I w_I$, so $(t_{I-1}, x_{I-1})$ is on a line of slope $-w_I$ through $(t, x)$. Similarly, as defined by equations (25) and (26),

$$t_i - t_{i-1} = t - \sum_{j=i+1}^{I} T_j - t + \sum_{j=i}^{I} T_j = T_i \tag{29}$$

and

$$x_i - x_{i-1} = x + \sum_{j=i+1}^{I} T_j w_j - x - \sum_{j=i}^{I} T_j w_j = -T_i w_i \tag{30}$$

so $(t_{i-1}, x_{i-1})$ is on a line of slope $-w_i$ through $(t_i, x_i)$. □

Newell (1993) showed that along characteristics moving at speed $-w_i$, the cumulative count increases at the rate of $K\Delta x$. The objective is for intermediate points $(t_i, x_i)$ to satisfy

$$N(t_i, x_i) = b_i \tag{31}$$

To achieve this, we impose the constraint

$$c(t_0, x_0) + K(x_0 - x_i) \geq b_i \tag{32}$$

The $\geq$ term in constraint (32) comes from the fact that $N(t_i, x_i)$ is the minimum of uncongested (upstream) conditions and congested (downstream) conditions. However if $c(t_0, x_0) + K(x_0 - x_i) < b_i$ then $N(t_i, x_i) < b_i$ also, which violates condition (31). Notice that if $c(t_0, x_0) \geq b_i$, then constraint (32) admits $x_i = 0$ (meaning $\sum_{j=1}^{i} T_j = 0$). The intuitive interpretation is that if $c(t_0, x_0) \geq b_i$, then regions $j = 1 \ldots i$ are not part of the congested characteristic as the region boundary vehicle $b_i$ passed $x = L$ before the congested characteristic reaches $x = L$. To achieve a linear program, we assume that $c(t_0, L)$ is linear, i.e. a constant flow constraint, such as a capacity. A piecewise linear definition of $c(t_0, L)$ can be handled by treating each piece separately via the inf-morphism property.

When $(t_i, x_i)$ is determined by upstream (uncongested) conditions, constraint (32) may not hold with equality. Since $N(t, x)$ is constant along uncongested characteristics of speed $v$ (Newell, 1993), we can find the smallest value $\hat{t}_i$ such that $c\left(\hat{t}_i, 0\right) = b_i$, then trace a line of slope $v$ to $(t_i, x_i)$:

$$N(t_i, x_i) \geq c\left(\hat{t}_i, 0\right) = b_i \tag{33}$$

with

$$\hat{t}_i = t_i - \frac{x_i}{v} \tag{34}$$

As discussed in Section 4.1.3, the correct value of $t_i$ is the larger one. Combining equations (33) and (34) yields

$$t_i \geq \hat{t}_i + \frac{x_i}{v} \tag{35}$$

meaning that vehicle $b_i$ can intersect the line of slope $w_{i+1}$ at time $t_i$ or greater (if congested conditions determine $b_i$'s trajectory). If $b_i$ does not cross the line $x = 0$ during $t \in [0, \infty)$, there should be a maximum value $\hat{x}_i \in [0, L]$ such that $c\left(0, \hat{x}_i\right) = b_i$. (If $\hat{x}_i < 0$, then $\hat{t}_i > 0$. If $\hat{x}_i > L$, then region $i$ is never active.) These two possibilities are illustrated in Figure 4. If we have an $\hat{x}_i \in [0, L]$ and not a $\hat{t}_i \in [0, \infty)$, then equation (33) can be rewritten as

$$N(t_i, x_i) \geq c\left(0, \hat{x}_i\right) = b_i \tag{36}$$

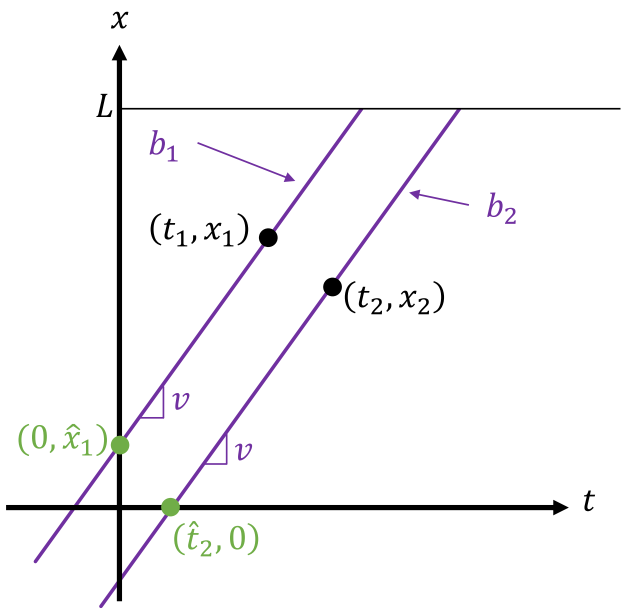



Figure 4: Illustration of boundary vehicles crossing the initial or upstream conditions.

Equation (36) is equivalent to

$$t_i \geq \frac{x_i - \hat{x}_i}{v} \tag{37}$$

because $vt_i$ is the minimum time for $b_i$ to traverse the distance $x_i - \hat{x}_i$ at free flow speed $v$.

Since $N(t_i, x_i)$ is defined by the minimum value of $N$ from tracing the forward and backward characteristics, the objective is to find the minimum value of $t_i$ subject to constraints (32) and (35) or (37) (whichever is appropriate). This leads to the following linear program:

$$\min \quad \sum_{i=1}^{I} t_i \tag{38a}$$

$$\text{s.t.} \quad t_i = t - \sum_{j=i+1}^{I} T_j \qquad \forall i \in [0, I] \tag{38b}$$

$$x_i = x + \sum_{j=i+1}^{I} T_j w_j \qquad \forall i \in [0, I] \tag{38c}$$

$$x_0 = L \tag{38d}$$

$$c(t_0, x_0) + K(x_0 - x_i) \geq b_i \qquad \forall i \in [1, I-1] \tag{38e}$$

$$\begin{cases} t_i \geq \hat{t}_i + \frac{x_i}{v} & \text{if } \hat{t}_i \in [0, \infty) \\ t_i \geq \frac{x_i - \hat{x}_i}{v} & \text{if } \hat{x}_i \in [0, L] \end{cases} \qquad \forall i \in [1, I-1] \tag{38f}$$

$$T_i \geq 0 \qquad \forall i \in [1, I] \tag{38g}$$

Constraint (38f) is noteworthy as it encapsulates two constraints. As discussed in the derivation of equations (35) and (37), either a $\hat{t}_i \in [0, \infty)$ or a $\hat{x}_i \in [0, L]$ defines the point at which vehicle $b_i$ passes the initial or upstream boundaries. Whichever is appropriate should be used.

Once intermediate points $(t_i, x_i)$ are known, then $N(t, x)$ can be determined as the minimum of several constraints. First, $N(t, x)$ is upper-bounded by the forward-moving characteristic of speed $v$ passing through point $(t, x)$.

$$N(t, x) \leq c\left(t - \frac{x}{v}, 0\right) \tag{39}$$

The calculation of the backwards moving characteristic is more complicated because $(t_i, x_i)$ has two possible values: $N(t_i, x_i) = b_i$, or $x_i = L$ with $N(t_i, x_i) > b_i$. In the latter case, $T_i = 0$ (region $i$ is not active). Their effect on $N(t, x)$ can be summarized as follows. If $x_{I-1} < L$, then

$$N(t, x) \leq b_{I-1} + K(x_{I-1} - x) \tag{40}$$

because along characteristics moving at speed $-w_{I-1}$ in region $I-1$, the cumulative count increases at the rate of $K\Delta x$ (Newell, 1993). Regardless of which regions are active,

$$N(t, x) \leq c(t_0, x_0) + K\left((x_{I-1} - x_0) + \sum_{i=0}^{I} K(x_i - x_{i+1})\right) = c(t_0, x_0) + K(x_0 - x) \tag{41}$$

Unfortunately, if $I$ regions are considered but $N(t, x) < b_I$, then constraint (38e) will cause linear program (38) to be infeasible. One possible algorithm for avoiding the infeasibility is to solve $J^\star = \max\{J \in [1, I] : \text{linear program (38) is feasible}\}$. This requires solving a series of linear programs for each point $(t, x)$. Instead, we suggest an alternative by allowing constraints (38e) and (38f) to be exceeded, but penalizing the excess. Rewrite constraint (38e) as

$$c(t_0, x_0) + K(x_0 - x_i) - b_i \geq -\lambda_i \tag{42}$$

where $\lambda_i \geq 0$ is the amount by which constraint (38e) is exceeded. Similarly, rewrite constraints (38f) as

$$\begin{cases} t_i - \hat{t}_i - \frac{x_i}{v} \geq -\lambda_i & \text{if } \hat{t}_i \in [0, \infty) \\ t_i - \frac{x_i - \hat{x}_i}{v} \geq -\lambda_i & \text{if } \hat{x}_i \in [0, L] \end{cases} \tag{43}$$

Then objective (38a) is modified to

$$\min \sum_{i=1}^{I} t_i + \beta \sum_{i=1}^{I} \lambda_i \tag{44}$$

If $\beta$ is sufficiently large, then $\lambda_i = 0$ is optimal if linear program (38) is feasible. If not, then the revised linear program will be feasible but the optimal solution will have $\lambda_i > 0$, which will indicate that $(t_i, x_i)$ was not found because $(t, x)$ occurs before region $i$. Linear program (38) can thus be written as

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{I} t_i + \beta \sum_{i=1}^{I} \lambda_i && && \text{(45a)}\\
\text{s.t.} \quad & t_i = t - \sum_{j=i+1}^{I} T_j && \forall i \in [0, I] && \text{(45b)}\\
& x_i = x + \sum_{j=i+1}^{I} T_j w_j && \forall i \in [0, I] && \text{(45c)}\\
& x_0 = L && && \text{(45d)}\\
& c(t_0, x_0) + K(x_0 - x_i) - b_i \geq -\lambda_i && \forall i \in [1, I-1] && \text{(45e)}\\
& \begin{cases} t_i - \hat{t}_i - \frac{x_i}{v} \geq -\lambda_i & \text{if } \hat{t}_i \in [0, \infty) \\ t_i - \frac{x_i - \hat{x}_i}{v} \geq -\lambda_i & \text{if } \hat{x}_i \in [0, L] \end{cases} && \forall i \in [1, I-1] && \text{(45f)}\\
& T_i \geq 0 && \forall i \in [1, I] && \text{(45g)}\\
& \lambda_i \geq 0 && \forall i \in [1, I] && \text{(45h)}
\end{aligned}
$$

The number of variables in linear program (45) scales linearly with the number of class proportion regions considered. After solving linear program (45), $N(t, x)$ is the maximum value satisfying several constraints, including (39) and (41). Constraint (40) is modified as follows: find the largest active region $\tilde{I}$ with $\lambda_{\tilde{I}} = 0$, i.e.

$$
\tilde{I} = \max \{ i \in [1, I] : \lambda_i = 0 \} \tag{46}
$$

If $\lambda_i > 0$, then constraints (38e) or (38f) do not hold, meaning that region $i$ occurs after point $(t, x)$. Then if $\tilde{I} > 0$,

$$
N(t, x) \leq b_{\tilde{I}-1} + K \left( x_{\tilde{I}-1} - x \right) \tag{47}
$$

#### 4.2.2 Piecewise-linear downstream boundary conditions

Linear program (45) can be modified for a piecewise-linear downstream boundary condition by solving with each piece separately with the additional constraint

$$
t^{\flat} \leq t_0 \leq t^{\sharp} \tag{48}
$$

where $t^{\flat}$ and $t^{\sharp}$ are the lower and upper bounds of $t$ for which the condition applies, respectively. Then, using the inf-morphism property (Claudel and Bayen, 2010a), the correct solution is the one that obtains the lower $N(t, x)$ value. However, two additional issues arise due to the addition of constraint (48). First, the value of $(t_i, x_i)$ in the optimal solution may not have either (45e) or (45f) as an active constraint, i.e. $N(t_i, x_i) > b_i$, to achieve feasibility with constraint (48). This issue is illustrated in Figure 5(a). If $w_2 > w_1$, the correct value

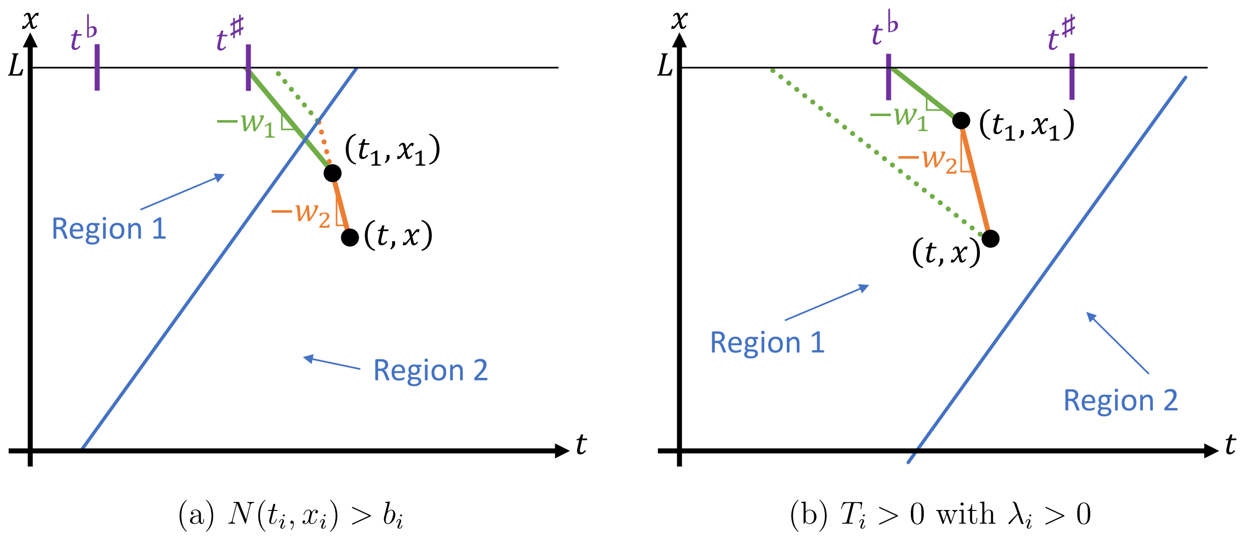



Figure 5: Illustration of issues caused by the addition of constraint (48). Dotted lines indicate the optimal characteristics without constraint $t_0 \in \left[t^\flat, t^\sharp\right]$.

of $(t_1, x_1)$ may be too high for a line of slope $-w_1$ to reach the boundaries of the constraint $t_0 \le t^\sharp$. In this case, constraint (40) should only be applied when constraints (45e) or (45f) are active. Second, due to the shallowness of a slope $w_i$, it is possible to have $T_{i+1} > 0$ while $\lambda_{i+1} > 0$, which means that $N(t_i, x_i) < b_i$. Figure 5(b) illustrates this as $N(t_1, x_1) < b_1$ yet because $w_1 < w_2$, $T_2 > 0$ is needed to achieve $t_0 \ge t^\flat$. In this case, $(t_0, x_0)$ should not be achievable with constraints (38e) and (38f) in place. This results in a incorrect (too small) value of $c(t_0, x_0)$ being applied to constraint (41). In this case, constraint (41) on $N(t, x)$ should not be used.

### 4.2.3 Initial condition

Linear program (45) can be used with a slight modification to handle an initial condition (defined for $t = 0$). Replace constraint (45d) with the following:

$$t_0 = 0 \tag{49}$$

The remainder of the constraints are sufficiently general to hold for both a downstream and initial condition.

## 4.3 Numerical demonstration of multiclass Newell's method

We demonstrate the multiclass calculation of $N(t, x)$ by considering a 2-class scenario of legacy and automated vehicles. Using the triangular flow-density relationship derived by Levin and Boyles (2016), the congested wave speed for region $i$ with an automated vehicle proportion of $p_i^{\mathrm{AV}}$ is

$$w_i = \frac{\ell}{p_i^{\mathrm{AV}} \tau^{\mathrm{AV}} + \left(1 - p_i^{\mathrm{AV}}\right) \tau^{\mathrm{HV}}} \tag{50a}$$

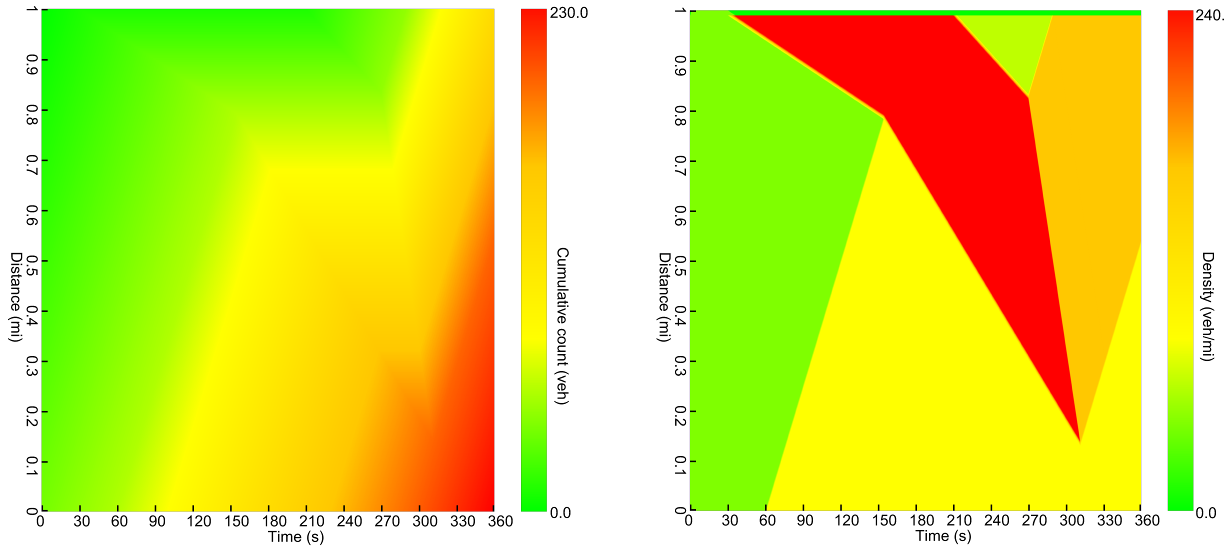
with capacity of

$$Q = \frac{v}{v\left(p_i^{\mathrm{AV}}\tau^{\mathrm{AV}} + \left(1i - p_i^{\mathrm{AV}}\right)\tau^{\mathrm{HV}}\right) + \ell} \tag{50b}$$

where $\tau^{\mathrm{HV}}$ and $\tau^{\mathrm{AV}}$ are the reaction times of legacy and automated vehicles, respectively, and $\ell = \frac{1}{K}$ is the vehicle length. For this demonstration, we used specific values of $\tau^{\mathrm{HV}} =$ 1.5s, $\tau^{\mathrm{AV}} =$ 0.25s, $v =$ 30mi/hr, and $K =$ 240veh/mi. We consider a 1-mile length of road with a red traffic signal (0 flow) from $t = 30$s to $t = 210$s. We defined two regions: region 1, for vehicles $[0, 50]$, with 0% automated vehicles, and region 2, for vehicles $(50, \infty)$, with 100% automated vehicles. Therefore, capacities for regions 1 and 2 are 1800vph and 4800vph, respectively. The initial condition $c(0, x)$ was set at a density of 40veh/mi, with $c(0, x) = 40 - kx$. Upstream entering flow was 1200veh/hr for $t \in [0, 60)$ then 2400veh/hr for $t \in [60, \infty)$. These parameters are not intended to suggest representative values for automated vehicle flow. Rather, the purpose is to provide a demonstration of the multiclass Newell's method for solving $N(t, x)$.

The solution method described in Section 4.2 was implemented in Java, using IBM's CPLEX 12.6 to solve linear programs. At a time interval of 1s and a spatial interval of 0.01mi, computing the cumulative count map for $t \in [0, 360]$s and $x \in [0, 1]$mi required 41.48s on a desktop computer with an Intel Core i5–8600K at 3.60GHz with 16GB of memory. Computing the density map required 77.34s. Density was approximated by $k(t, x) \approx \frac{N(t,x) - N(t,x+\epsilon)}{\epsilon}$.

The cumulative counts are shown in Figure 6(a) and the densities are shown in Figure 6(b). There are several interesting patterns of note. First, from $t = 60$s to $t = 360$s, entering density is 80veh/mi, which exceeds the critical density of region 1 (60veh/mi) but is well within the uncongested regime of region 2. Consequently, the shockwave separating the queue from the entering flow changes speed at $(155\text{s}, 0.79\text{mi})$. Downstream, and after the traffic signal turns green at $t = 210$s, the queue initially dissipates at a rate of 1800veh/hr, with a density of 60veh/mi, until the 50th vehicle has passed. Thereafter, the queue dissipates at the rate of 4800veh/hr, or a density of 160veh/mi, due to the higher capacity of region 2.




(a) $N(t, x)$ (b) $k(t, x)$

Figure 6: Illustration of computed cumulative count map and density map from multiclass traffic flow scenario

# 5 Multiclass link transmission model

We now apply the multiclass Newell's method into a multiclass version of Yperman et al. (2005)'s link transmission model for dynamic network loading. This paper focuses on developing a link transmission model for a mixture of human and AV flows with a changing flow-density relationship. Dynamic traffic assignment for a single-class of vehicles is well-established in the literature, and the methods in this paper can be used with existing node models (Tampère et al., 2011) and dynamic traffic assignment algorithms (Levin et al., 2015). For more information on dynamic network loading, we refer the reader to Yperman (2007) which describes the single-class link transmission model for link flow modeling as well as its connection to node models. In Section 6, we will demonstrate a dynamic traffic assignment using the multiclass link transmission model on a city network, combined with standard node models for traffic signals, merges, and diverges. As with other link models, the multiclass link transmission model must calculate sending and receiving flows, which are then used by the node models to calculate entering and exiting flow and the upstream and downstream ends of each link. The link transmission model stores the upstream and downstream boundary conditions up to time $t$. Cumulative count values within the link are not stored.

## 5.1 Class proportion regions

Consider discretized time with time step $\Delta t$. The class proportion regions are defined separately for each link. The class proportion region boundaries could be chosen in several ways. To work with both continuous and discrete flow, we define the region boundaries to coincide with time steps. Flow entering between time $t$ and $t + \Delta t$ is defined as one class proportion region, with boundaries $N(t, 0)$ and $N(t + \Delta t, 0)$. We assume that $N^m(t, 0)$ and $N^m(t + \Delta t, 0)$ are known, then we can calculate the average entering flow between $[t, t + \Delta t)$ as

$$q^m(t, 0) \approx \frac{N^m(t + \Delta t, 0) - N^m(t, 0)}{\Delta t} \tag{51}$$

We then create a new class proportion region applying to the vehicles entering between $t$ and $t + \Delta t$ with proportions of $\frac{q^m(t,0)}{\sum\limits_{m' \in \mathcal{M}} q^{m'}(t,0)}$. This implicitly assumes that class proportions are uniformly mixed between $t$ and $t + \Delta t$. If desired, a higher resolution of class proportion regions could be implemented.

However, when working with continuous flows, class proportions at the upstream end of each link between time $t$ and $t + \Delta t$ are assumed to become uniformly distributed. This is not a problem at centroids, but violates FIFO at the upstream ends of internal links. Although there are fewer mixing points in this LTM than the multiclass CTM (Levin and Boyles, 2016), mixing still occurs. This is actually a general problem with dynamic network loading of multi-commodity flow (Carey et al., 2014). For other papers, the commodities are typically paths or destinations, and vehicle class becomes an additional identifier in this paper. Bar-Gera and Carey (2017) studied methods of ensuring FIFO holds in continuous dynamic network loading models. Although addressing this issue is outside the scope of this paper, it is worth studying in future work.

## 5.2 Sending flow

The sending flow is the maximum flow that could exit in one time step, and the receiving flow is the maximum flow that could enter in one time step. The sending flow $S(t)$ can be calculated by

$$S(t) = N(t + \Delta t, L) - N(t, L) \tag{52}$$

where $N(t + \Delta t, L)$ is the maximum value found based on constraints within the link. (The actual flow may be less than the sending flow depending on downstream constraints from the node model.) Although $N(t, L)$ is stored, $N(t + \Delta t, L)$ must be calculated. Since $L$ is the downstream boundary, the only characteristic that can reach a boundary condition is the uncongested characteristic, which always travels at speed $v$ regardless of the class proportions. Therefore,

$$N(t + \Delta t, L) = N\left(t + \Delta t - \frac{L}{v}, 0\right) \tag{53}$$

Due to equation (53), only $\frac{L}{v}$ time steps of the upstream cumulative count need to be stored.

Sending flow is also constrained by capacity, which is not a constant value because it depends on the class proportions. Given class proportion regions $0 \ldots I$, the first class proportion region of consideration, $i^\flat$, is the first region active on the time interval $[t, t + \Delta t)$:

$$i^\flat = \arg\max_{i \in [0, I]} \{b_{i-1} \leq N(t, L)\} \tag{54}$$

Then the capacity constraint $\tilde{Q}$ can be constructed using the following algorithm. Add up $Q_{i^\flat}$ of vehicles until either the class region boundary $b_{i^\flat}$ or the time step $\Delta t$ is reached. If the class region boundary is reached, then continue with the next region $i^\flat + 1$ until 1 time step has elapsed. This calculation is defined by Algorithm 1. $\tilde{Q}$ and $\tilde{\tau}$ are state accumulation variables. $\tilde{b}$ is the region lower bound, which could be $b_{i-1}$ or $N(t, L)$. For each class region $i$, that region can apply for up to $\tau = \frac{b_i - \tilde{b}}{Q_i}$ time (after which the class region transitions to $i+1$). If the remaining time $\Delta t - \tilde{\tau} > \tau$, then add $Q_i \tau$ to the capacity variable $\tilde{Q}$, update $\tilde{b}$ and $\tilde{\tau}$, and continue with region $i + 1$. Otherwise, add $Q_i(\Delta t - \tilde{\tau})$ capacity and terminate. The sending flow calculation can then be summarized by

$$S(t) = \min \left\{ \tilde{Q}, N\left(t + \Delta t - \frac{L}{v}, 0\right) - N(t, L) \right\} \tag{55}$$

**Algorithm 1** Calculating capacity for sending flow

$\tilde{Q} \leftarrow 0$
$\tilde{b} \leftarrow N(t, L)$
$\tilde{\tau} \leftarrow 0$
**for** $i = i^\flat$ to $I$ **do**
  $\tau = \frac{b_i - \tilde{b}}{Q_i}$
  **if** $\tilde{\tau} + \tau \geq \Delta t$ **then**
    $\tilde{Q} \leftarrow \tilde{Q} + Q_i \left(\Delta t - \tilde{\tau}\right)$
    **return** $\tilde{Q}$
  **else**
    $\tilde{Q} \leftarrow \tilde{Q} + Q_i \tau$
    $\tilde{\tau} \leftarrow \tilde{\tau} + \tau$
    $\tilde{b} \leftarrow b_i$
  **end if**
**end for**

## 5.3 Receiving flow

Like the sending flow, the receiving flow $R(t)$ is calculated as

$$R(t) = N(t + \Delta t, 0) - N(t, 0) \tag{56}$$

$N(t + \Delta t, 0)$ requires tracing a congested characteristic, which has variable speed. The method for tracing this characteristic is defined in Section 4.2. However, there are some useful implementation results to be discussed here. First, we prove in Proposition 5 that only $\frac{L}{w^\flat}$ time steps of downstream cumulative counts need to be stored, which determines the memory requirements. Second, observe that if a class region $i$ ends with boundary vehicle $b_i$ such that $N\left(t + \Delta t - \frac{L}{w^\flat}, L\right) \geq b_i$, then class region $i$ is no longer relevant to calculating $N(t + \Delta t, 0)$, and can be discarded from memory. The capacity constraint for the receiving flow can be estimated based on the class proportions of upstream link sending flows.

**Proposition 5.** *Downstream cumulative counts must be stored for the previous $\frac{L}{w^\flat} - \Delta t$ time, where $w^\flat$ is the minimum congested wave speed.*

*Proof.* To calculate $N(t+\Delta t, 0)$, the multiclass Newell's method follows a congested characteristic of speeds $w_i$ to a time $t + \Delta t - \sum_{i=1}^{I} T_i$ such that $\sum_{i=1}^{I} T_i w_i = L$. Since $w_i \geq w^\flat$, $\sum_{i=1}^{I} T_i$ can be upper-bounded by $\sum_{i=1}^{I} T_i \leq \frac{L}{w^\flat}$ therefore

$$t + \Delta t - \sum_{i=1}^{I} T_i \geq t + \Delta t - \frac{L}{w^\flat} \tag{57}$$

which requires looking back at most $\frac{L}{w^\flat} - \Delta t$ in time. □

## 5.4 Iterative algorithm for receiving flow

The repeated use of solvers to calculate $N(t+\Delta t, 0)$ can limit the computational performance of large dynamic network loading models. In this section we present a polynomial time algorithm based on iteratively solving a system of linear equations. The multiclass link transmission model benefits from certain simplifications that are not true in the general multiclass Newell's method. First, since only the sending and receiving flows are calculated, $N(t+\Delta t, x)$ is calculated for $x = L$ or $x = 0$. When $x = L$, the solution is simply looking backwards along an uncongested wave of speed $v$. When $x = 0$, the solution involves looking backwards along a congested wave with varying speed. However, due to the construction of the region boundaries, $N(t+\Delta t)$ is in region $I+1$ where region $I$ has boundary vehicle $b_I = N(t, 0)$. The second simplification is therefore that the class boundary region is known with certainty when calculating $N(t+\Delta t, 0)$. This suggests that $N(t+\Delta t, 0)$ can be calculated by iteratively solving a system of linear equations. The iteration occurs because the intermediate points $(t_i, x_i)$ satisfying $N(t_i, x_i) = b_i$ could be determined either by congested or uncongested conditions. The first iteration can assume that all $(t_i, x_i)$ points are determined by congested conditions, then revise them to use uncongested conditions if they result in the contradiction that $N(t_i, x_i) = b_i$ yet $N(t_i, x_i) \leq c\left(t_i - \frac{x_i}{v}, 0\right)$.

We assume here that $I \geq 3$. If $I < 3$, then the method of Section 4.1 may be used instead. For the first iteration, assume that congested conditions determine all points $(t_i, x_i)$. Therefore, $N(t+\Delta t, 0)$ is determined by following a congested characteristic back to the downstream boundary condition $c(t, L)$. This results in the following system of equations:

$$c(t_0, L) + K(L - x_1) = b_1 \tag{58a}$$

$$K(x_i - x_{i+1}) = b_{i+1} - b_i \qquad \forall i \in [1, I-1] \tag{58b}$$

$$-w_1(t_1 - t_0) = x_1 - L \tag{58c}$$

$$-w_{i+1}(t_{i+1} - t_i) = x_{i+1} - x_i \qquad \forall i \in [1, I-1] \tag{58d}$$

$$-w_{I+1}(t + \Delta t - t_I) = 0 - x_I \tag{58e}$$

The simplified solution involves removing equations (58b), (58c), and (58d) to obtain a system of equations in only three variables: $t_0$, $t_I$, and $x_I$. These equations can be removed because the difference in cumulative count between boundary vehicles $b_{i+1}$ and $b_i$ is constant, so $K(x_i - x_{i+1})$ is also constant. Since $w_{i+1}$ is given, the difference $t_{i+1} - t_i$ can be determined as

$$t_{i+1} - t_i = \frac{x_{i+1} - x_i}{-w_{i+1}} = \frac{b_{i+1} - b_i}{K w_{i+1}} \tag{59}$$

Equations (58a) and (58b) can be combined into

$$c(t_0, L) + K(L - x_I) = b_I \tag{60}$$

Equations (58b) and (59) can be used to find $t_I - t_0$ as

$$t_I - t_0 = \frac{b_1 - c(t_0, L)}{K w_1} + \sum_{i=2}^{I} \frac{b_i - b_{i-1}}{K w_i} \tag{61}$$

When $c(t_0, L)$ is linear (as it is for LTM), equations (58e), (60), and (61) form a system of three equations with variables $t_0$, $t_I$, and $x_I$. Assume that $c(t_0, L) = c_0 + q_0 t_0$. Solving this system results in the solution

$$t_0 = \frac{\frac{b_I - c_0 - KL + K w_{I+1}(t + \Delta t)}{K w_{I+1}} + \frac{c_0 - b_1}{K w_1} - \sum_{i=2}^{I} \frac{b_i - b_{i-1}}{K w_i}}{\frac{q_0}{K w_{I+1}} + 1 - \frac{q_0}{K w_1}} \tag{62}$$

with $t_I$ and $x_I$ found through equations (61) and (60), respectively. Using equations (58c) and (59), all intermediate points $(t_i, x_i)$ can be found.

The prior solution to $(t_i, x_i)$ points assumed that all points were determined by congested conditions. It is easy to check whether any given $(t_i, x_i)$ point should be determined by uncongested conditions instead: $c\left(t_i - \frac{x_i}{v}, 0\right) < b_i$. Let $\hat{i}$ indicate the furthest vehicle position determined by uncongested conditions, i.e.

$$\hat{i} = \max\left\{ i \in [1, I] : c\left(t_i - \frac{x_i}{v}, 0\right) < b_i \right\} \tag{63}$$

Then the true value of $(t_{\hat{i}}, x_{\hat{i}})$ is on a line of slope $v$ from a point $t'$ such that $c(t', 0) = b_{\hat{i}}$. Given $c(t, 0)$, that $t'$ can be uniquely determined and can be hereafter treated as a constant. This works because the cumulative count does not change along a uncongested characteristic and is therefore independent of the length of that characteristic. We now have a revised system of linear equations:

$$v(t_{\hat{i}} - t') = x_{\hat{i}} - 0 \tag{64a}$$

$$0 - x_I = -w_{I+1}(t + \Delta t - t_I) \tag{64b}$$

$$K(x_i - x_{i+1}) = b_{i+1} - b_i \qquad \forall i \in \left[\hat{i}, I - 1\right] \tag{64c}$$

$$x_{i+1} - x_i = -w_{i+1}(t_{i+1} - t_i) \qquad \forall i \in \left[\hat{i}, I - 1\right] \tag{64d}$$

As before, we want to simplify equations (64) to have only variables $t_{\hat{i}}$, $x_{\hat{i}}$, $t_I$, and $x_I$. Equations (64c) can be simplified to

$$K\left(x_{\hat{i}} - x_I\right) = b_I - b_{\hat{i}} \tag{65}$$

Using equation (59), equation (64d) can be simplified to

$$t_I - t_{\hat{i}} = \sum_{i=\hat{i}+1}^{I} \frac{b_i - b_{i-1}}{K w_i} \tag{66}$$

Equations (64a), (64b), (65), and (66) now form a system of four equations with four variables. The solution is

$$t_{\hat{i}} = \frac{\frac{b_I - b_{\hat{i}}}{K} + vt' + w_{I+1}(t+\Delta t) - w_{I+1} \sum\limits_{i=\hat{i}+1}^{I} \frac{b_i - b_{i-1}}{K w_i}}{v + w_{I+1}} \tag{67}$$

with $t_I$, $x_I$, and $x_{\hat{i}}$ found by equations (66), (64b), and (65), respectively. Intermediate $(t_i, x_i)$ points can be calculated by equations (64c) and (64d). As before, if some of these intermediate points should be instead determined by uncongested conditions, calculate a new $\hat{i}$ using equation (63) and solve equations (64) again. The algorithm to calculate $N(t+\Delta t, 0)$ using iterative systems of linear equations is summarized in Algorithm 2.

**Algorithm 2** Iterative algorithm for calculating the receiving flow for the multiclass LTM

1: Set $\hat{i} \leftarrow 1$
2: **if** $I + 1 - \hat{i} \leq 1$ **then**
3:     Solve $N(t, 0)$ using Newell (1993)'s method (1 region)
4: **else if** $I + 1 - \hat{i} \leq 2$ **then**
5:     Solve $N(t, 0)$ using 2-region multiclass Newell's method (Section 4.1)
6: **end if**
7: **if** $\hat{i} = 1$ **then**
8:     Solve equations (58e), (60), and (61) for $N(t, 0)$
9: **else**
10:     Solve equations (64a), (64b), (65), and (66) for $N(t, 0)$
11: **end if**
12: **if** $\left\{i \in [1, I] : c\left(t_i - \frac{x_i}{v}, 0\right) < b_i\right\} = \varnothing$ **then**
13:     **return** $b_I + K x_I$
14: **else**
15:     $\hat{i} \leftarrow \max\left\{i \in [1, I] : c\left(t_i - \frac{x_i}{v}, 0\right) < b_i\right\}$
16:     **go to** line 2
17: **end if**

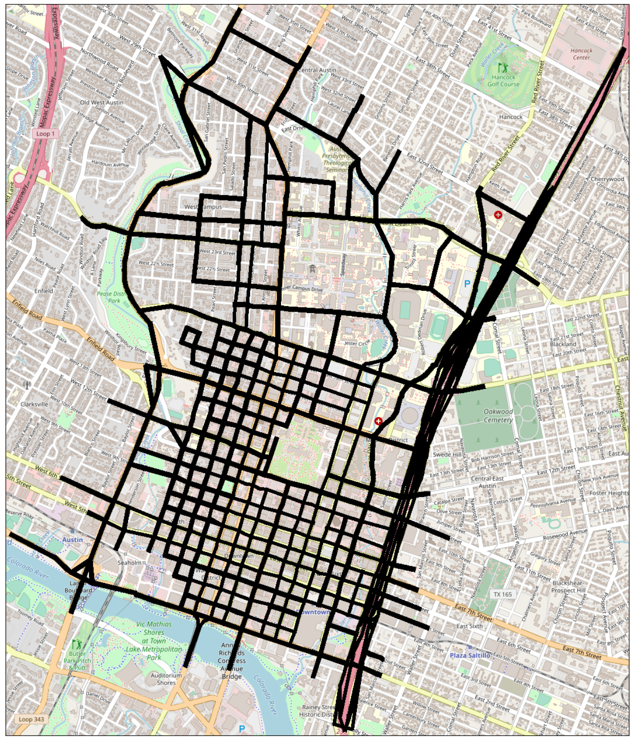


Figure 7: Downtown Austin city network

# 6 Numerical results from dynamic traffic assignment

We implemented the multiclass LTM in a dynamic traffic assignment model that discretizes flow and tracks the location, path, and driver type (legacy or automated) of individual vehicles with a time step of $\Delta t = 15$s. By discretizing vehicles, we can reduce the FIFO issues discussed in Section 5.1. The sending and receiving flows calculated by the multiclass LTM are used to determine the propagation of vehicles along links and through nodes. This same software package (albeit with a different link flow model) has been used for numerical results in previous studies (e.g. Levin and Boyles, 2016; Levin et al., 2017). We conducted a demonstration of dynamic traffic assignment using the multiclass LTM on the downtown Austin city network, shown in Figure 7. This network was constructed by the Network Modeling Center at The University of Texas at Austin for project work for the City of Austin, and was calibrated to match observations from 2010. The network includes 171 zones, 546 nodes, and 1247 links, and has a demand of around 62,836 vehicles over 2 hours.

## 6.1 Convergence properties of dynamic traffic assignment

We solved dynamic traffic assignment using the method of successive averages. Normal convergence of dynamic traffic assignment was observed with the multiclass LTM in use. For instance, Figure 8 shows the gap with respect to iteration for a scenario with 50% AVs.

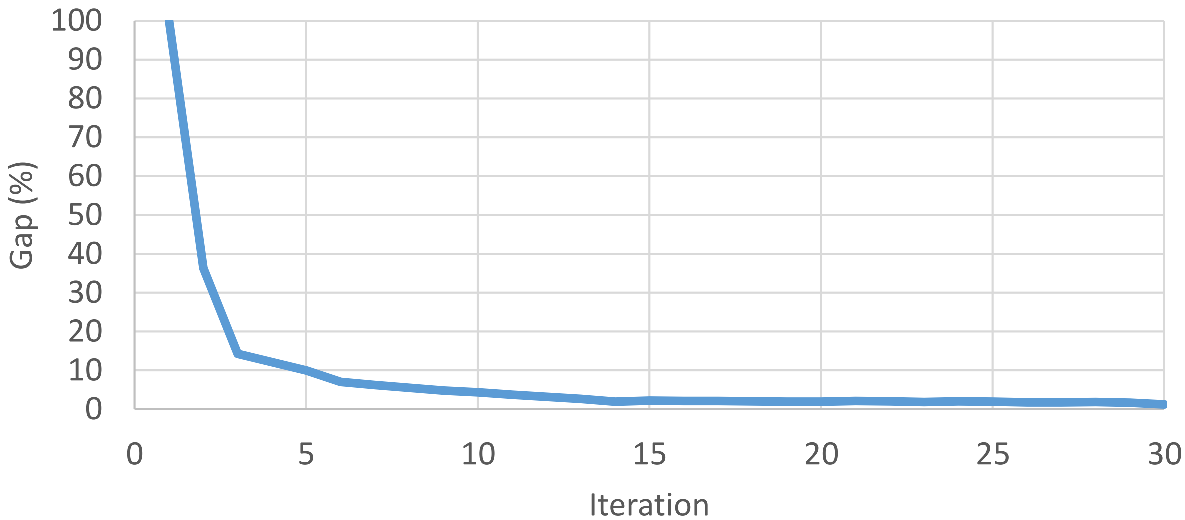



Figure 8: Convergence of dynamic traffic assignment using the multiclass LTM

This gap is defined as

$$\text{gap} = \frac{\text{TSTT} - \text{SPTT}}{\text{TSTT}} \tag{68}$$

where TSTT is the total system travel time and SPTT is the shortest path travel time. A gap of under 2% was achieved after 30 iterations, which is typical for this network and for dynamic traffic assignment (Levin et al., 2015). We note that due to the discontinuities in travel times caused by LTM, the existence and uniqueness properties of dynamic user equilibrium cannot be analytically established. The average computation time per iteration was 19.2s on a desktop with an Intel Core i5-9400 processor clocked at 2.90 GHz with 24.0GB of memory. Most of the additional computational effort of Algorithm 2 compared with LTM occurs when the number of class proportion regions is large. Efforts to reduce the number of regions whenever possible are important for achieving a lower computation time. For example, one new region is usually created for each link per time step. However, if the entering flow is zero, then a new region does not have to be created for that link. If the entering flow for a link has a AV proportion identical to the last region, then a new region also does not need to be created there.

## 6.2 Effects of AV market penetration on traffic congestion

Next, we evaluated the sensitivity of network travel times with respect to AV market penetration. Link capacities and congested wave speeds were scaled according to equation (50). To retain as much of the original network parameters as possible, link capacities were not set equal to the output of equation (50b). Instead, we calculated $Q$ for $p_i^{\text{AV}}$ and with 0 AV market penetration, then scaled the calibrated link capacities proportionally. Formally, if $\tilde{Q}_a$ is the link capacity of link $a$, and $Q\left(p_i^{\mathrm{AV}}\right)$ is the capacity calculated by equation (50b), then the new link capacity for region $i$ is chosen as $\tilde{Q}_a \frac{Q\left(p_i^{\mathrm{AV}}\right)}{Q(0)}$. Congested wave speeds were then determined by the triangular shape of the flow-density relationship, with a fixed jam density. We used reaction times of $\tau^{\mathrm{HV}} = 1\mathrm{s}$ and $\tau^{\mathrm{AV}} = 0.5\mathrm{s}$. For each AV market penetration, 10 Monte Carlo simulations were performed to reduce the effects of stochasticity in the demand. Because flow is discretized, each vehicle is randomly determined to be either a legacy or automated vehicle with probability given by the overall AV market penetration.

The results presented here assume that intersection capacity increases with the AV market penetration as specified by the flow-density relationship (50). It is possible that intersection capacity may be higher for AVs due to signal-free intersection controls (Dresner and Stone, 2004), but such controls were not included in these results. It is also possible that AVs might instead behave more cautiously at intersections, resulting in lower intersection capacities than predicted here. These results are also highly dependent on the shape of the flow-density relationship for AVs, and it is possible that AVs might reduce capacity instead (Shladover et al., 2012; Calvert et al., 2017; James et al., 2019).

Figure 9 shows the average travel time as a function of the city-wide AV market penetration. Error bars show the standard deviation of travel times from the Monte Carlo simulations. Since legacy and automated vehicles share all links, the average travel times do not vary based on the driver type. Interestingly, the average travel times remained relatively constant between 10% and 50% AVs. We note that Braess (1968) and Daganzo (1998) demonstrated that increases in capacity could result in higher travel times due to selfish route choice, which was actually demonstrated for cooperative adaptive cruise control by Melson et al. (2018). The lack of improvement could also be due to intersection capacity limitations and queue spillback. Although in these results, the intersection capacity increases with the time- and space-dependent AV market penetration, queue spillback could reduce the flow improvements achieved by capacity increases at intersections. At 60% and higher AV market penetrations, travel times decreased consistently.

Of course, the effects that AVs have on congestion depend both on the network and the legacy and automated vehicle driving parameters. The results in Figure 9 are compared with other results in the literature. For freeways, Shi and Prevedouros (2016) and Yu et al. (2019) observed a similar pattern as shown in Figure 9: speeds increased only slightly at low market penetrations but had an increasing impact at higher market penetrations. However, they predicted more continuous effects than shown in Figure 9. Intersection controls could be causing the difference. Ghiasi et al. (2017) predicted even larger increases in capacity at low market penetrations, but without network results. Patel et al. (2016) found that congestion decreased steadily with increasing AV proportions even at low market penetrations, but Levin and Boyles (2016) observed that traffic signals were a major bottleneck for the downtown Austin network.

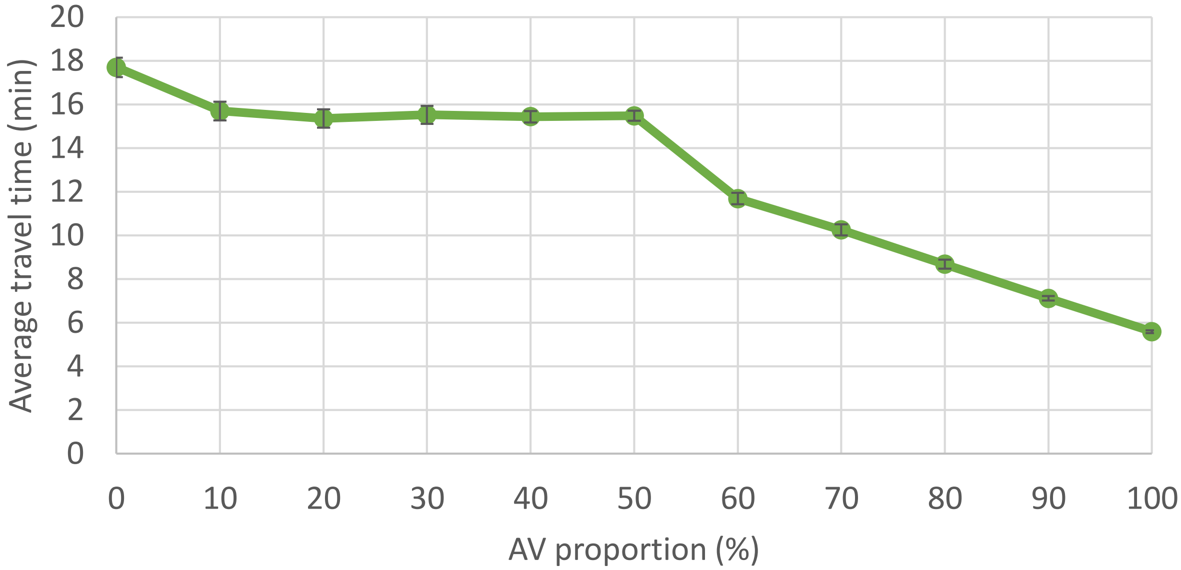



Figure 9: Effects of AV market penetration on travel times

# 7 Conclusions

This paper extended Newell (1993)'s method for a multiclass kinematic wave theory. Like the original, the proposed method finds exact solutions under triangular flow-density relationships. We started by exploring the multiclass kinematic wave theory, then recognizing that when vehicles obey first-in-first-out behavior, that the class-specific cumulative counts $N^m(t, x)$ can be adequately described by a single aggregate cumulative count $N(t, x)$ and class proportion regions that move with the traffic flow. We wrote a corresponding multiclass Lax-Hopf formula (Claudel and Bayen, 2010a). Restricting ourselves to triangular flow-density relationships with varying capacities and congested wave speeds, we developed a linear program to implement a multiclass Newell (1993)'s method to find exact solutions to the multiclass kinematic wave theory.

With the ultimate goal of a multiclass link transmission model, we developed a more efficient iterative algorithm for calculating receiving flows. The multiclass link transmission model was demonstrated on the downtown Austin city network. Dynamic traffic assignment was observed to converge normally, and computation times on this city network were easily manageable. Numerical results show a non-linear decrease in overall congestion due to greater AV market penetration.

There are many future opportunities for this research. The multiclass LTM could be used for more accurate city-scale predictions of traffic flow and congestion with varying proportions of automated vehicles to prepare for their future use on public roads. The presented solution method is fairly complex, and it may be possible to use the problem structure to achieve a more efficient solution method. The flow-density relationship for autonomous vehicles may not follow a triangular shape, so extending the results in Sections 4 and 5 to more general flow-density relationships would be valuable. It may be possible to use Lagrangian coordinates and/or variational theory (Daganzo, 2006; Leclercq et al., 2007; Laval and Leclercq, 2013) to develop a more efficient solution method. The predictions of this model should also be verified against microsimulation. Multiclass flow that does not strictly adhere to the first-in-first-out property on links is more challenging to model, and could be the subject of future work. Although the standard LTM (Yperman et al., 2005) has been used in system optimal dynamic traffic assignment (Levin, 2017; Chakraborty et al., 2018), the complexity of the multiclass Newell's method makes formulating the system optimal dynamic traffic assignment problem challenging for multiclass flow. Finally, other multiclass kinematic wave theory problems might be able to adapt the concepts and methods presented here.

# Acknowledgements

The authors gratefully acknowledge the support of the National Science Foundation, Award No. 1935514.

# References

H. Bar-Gera and M. Carey. Representation requirements for perfect first-in-first-out verification in continuous flow dynamic models. *Transportation Research Part B: Methodological*, 100:284–301, 2017.

M. Beckmann, C. McGuire, and C. B. Winsten. Studies in the economics of transportation. Technical report, 1956.

P.-D. Braess. Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, 12(1): 258–268, 1968.

S. Calvert, W. Schakel, and J. Van Lint. Will automated vehicles negatively impact traffic flow? *Journal of Advanced Transportation*, 2017, 2017.

M. Carey, H. Bar-Gera, D. Watling, and C. Balijepalli. Implementing first-in–first-out in the cell transmission model for networks. *Transportation Research Part B: Methodological*, 65: 105–118, 2014.

S. Chakraborty, D. Rey, E. Moylan, and S. T. Waller. Link transmission model-based linear programming formulation for network design. *Transportation Research Record*, 2672(48): 139–147, 2018.

D. Chen, S. Ahn, M. Chitturi, and D. A. Noyce. Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles. *Transportation research part B: methodological*, 100:196–221, 2017.

R. Chen, T. Zhang, and M. W. Levin. Effects of variable speed limit on energy consumption with autonomous vehicles on urban roads using modified cell-transmission model. *Journal of Transportation Engineering, Part A: Systems*, 146(7):04020049, 2020.

Z. Chen, F. He, L. Zhang, and Y. Yin. Optimal deployment of autonomous vehicle lanes with endogenous market penetration. *Transportation Research Part C: Emerging Technologies*, 72:143–156, 2016.

Y.-C. Chiu, J. Bottom, M. Mahut, A. Paz, R. Balakrishna, T. Waller, and J. Hicks. Dynamic traffic assignment: A primer. *Transportation Research E-Circular*, (E-C153), 2011.

C. G. Claudel and A. M. Bayen. Lax–Hopf based incorporation of internal boundary conditions into hamilton–jacobi equation. part i: Theory. *IEEE Transactions on Automatic Control*, 55(5):1142–1157, 2010a.

C. G. Claudel and A. M. Bayen. Lax–Hopf based incorporation of internal boundary conditions into hamilton-jacobi equation. part ii: Computational methods. *IEEE Transactions on Automatic Control*, 55(5):1158–1174, 2010b.

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

C. F. Daganzo. The cell transmission model, part ii: network traffic. *Transportation Research Part B: Methodological*, 29(2):79–93, 1995.

C. F. Daganzo. Queue spillovers in transportation networks with a route choice. *Transportation Science*, 32(1):3–11, 1998.

C. F. Daganzo. On the variational theory of traffic flow: well-posedness, duality and applications. *Networks and Heterogeneous Media*, 1(4):601–619, 2006.

L. Davis. Effect of adaptive cruise control systems on mixed traffic flow near an on-ramp. *Physica A: Statistical Mechanics and its Applications*, 379(1):274–290, 2007.

K. Dresner and P. Stone. Multiagent traffic management: A reservation-based intersection control mechanism. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 530–537. IEEE Computer Society, 2004.

A. Ghiasi, O. Hussain, Z. S. Qian, and X. Li. A mixed traffic capacity analysis and lane management model for connected automated vehicles: A markov chain method. *Transportation Research Part B: Methodological*, 106:266–292, 2017.

S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.

S. Gong and L. Du. Cooperative platoon control for a mixed traffic flow including human drive vehicles and connected and autonomous vehicles. *Transportation research part B: methodological*, 116:25–61, 2018.

W. Himpe, R. Corthout, and M. C. Tampère. An efficient iterative link transmission model. *Transportation Research Part B: Methodological*, 92:170–190, 2016.

R. M. James, C. Melson, J. Hu, and J. Bared. Characterizing the impact of production adaptive cruise control on traffic flow: an investigation. *Transportmetrica B: Transport Dynamics*, 7(1):992–1012, 2019.

W.-L. Jin. A kinematic wave theory of multi-commodity network traffic flow. *Transportation Research Part B: Methodological*, 46(8):1000–1022, 2012.

W.-L. Jin. Continuous formulations and analytical properties of the link transmission model. *Transportation Research Part B: Methodological*, 74:88–103, 2015.

A. Kesting and M. Treiber. How reaction time, update time, and adaptation time influence the stability of traffic flow. *Computer-Aided Civil and Infrastructure Engineering*, 23(2): 125–137, 2008.

A. Kesting, M. Treiber, M. Schönhof, and D. Helbing. Extending adaptive cruise control to adaptive driving strategies. *Transportation Research Record*, 2000(1):16–24, 2007.

J. A. Laval and L. Leclercq. The hamilton–jacobi partial differential equation and the three representations of traffic flow. *Transportation Research Part B: Methodological*, 52:17–30, 2013.

L. Leclercq, J. A. Laval, and E. Chevallier. The lagrangian coordinates and what it means for first order traffic flow models. In *Transportation and Traffic Theory 2007. Papers Selected for Presentation at ISTTT17Engineering and Physical Sciences Research Council (Great Britain) Rees Jeffreys Road FundTransport Research FoundationTMS ConsultancyOve Arup and Partners, Hong KongTransportation Planning (International) PTV AG*, 2007.

M. W. Levin. Congestion-aware system optimal route choice for shared autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, 82:229–247, 2017.

M. W. Levin and S. D. Boyles. Effects of autonomous vehicle ownership on trip, mode, and route choice. *Transportation Research Record: Journal of the Transportation Research Board*, (2493):29–38, 2015.

M. W. Levin and S. D. Boyles. A multiclass cell transmission model for shared human and autonomous vehicle roads. *Transportation Research Part C: Emerging Technologies*, 62: 103–116, 2016.

M. W. Levin, M. Pool, T. Owens, N. R. Juri, and S. T. Waller. Improving the convergence of simulation-based dynamic traffic assignment methodologies. *Networks and Spatial Economics*, 15(3):655–676, 2015.

M. W. Levin, H. Fritz, and S. D. Boyles. On optimizing reservation-based intersection controls. *IEEE Transactions on Intelligent Transportation Systems*, 18(3):505–515, 2017.

M. J. Lighthill and G. B. Whitham. On kinematic waves. ii. a theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 229, pages 317–345. The Royal Society, 1955.

H. Liu, J. Wang, K. Wijayaratna, V. V. Dixit, and S. T. Waller. Integrating the bus vehicle class into the cell transmission model. *IEEE Transactions on Intelligent Transportation Systems*, 16(5):2620–2630, 2015.

H. Liu, X. Kan, S. E. Shladover, X.-Y. Lu, and R. E. Ferlis. Impact of cooperative adaptive cruise control on multilane freeway merge capacity. *Journal of Intelligent Transportation Systems*, 22(3):263–275, 2018.

Z. Liu and Z. Song. Strategic planning of dedicated autonomous vehicle lanes and autonomous vehicle/toll lanes in transportation networks. *Transportation Research Part C: Emerging Technologies*, 106:381–403, 2019.

S. Logghe and L. H. Immers. Multi-class kinematic wave theory of traffic flow. *Transportation Research Part B: Methodological*, 42(6):523–541, 2008.

M. Makridis, K. Mattas, and B. Ciuffo. Response time and time headway of an adaptive cruise control. an empirical characterization and potential impacts on road capacity. *IEEE Transactions on Intelligent Transportation Systems*, 21(4):1677–1686, 2019.

C. L. Melson, M. W. Levin, B. E. Hammit, and S. D. Boyles. Dynamic traffic assignment of cooperative adaptive cruise control. *Transportation Research Part C: Emerging Technologies*, 90:114–133, 2018.

V. Milanés, S. E. Shladover, J. Spring, C. Nowakowski, H. Kawazoe, and M. Nakamura. Cooperative adaptive cruise control in real traffic situations. *IEEE Transactions on intelligent transportation systems*, 15(1):296–305, 2013.

G. F. Newell. A simplified theory of kinematic waves in highway traffic, part i: General theory. *Transportation Research Part B: Methodological*, 27(4):281–287, 1993.

C. Nowakowski, J. O'Connell, S. E. Shladover, and D. Cody. Cooperative adaptive cruise control: Driver acceptance of following gap settings less than one second. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 54, pages 2033–2037. SAGE Publications Sage CA: Los Angeles, CA, 2010.

R. Patel, M. W. Levin, and S. D. Boyles. Effects of autonomous vehicle behavior on arterial and freeway networks. *Transportation Research Record: Journal of the Transportation Research Board*, (2561), 2016.

J. Ploeg, B. T. Scheepers, E. Van Nunen, N. Van de Wouw, and H. Nijmeijer. Design and experimental evaluation of cooperative adaptive cruise control. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 260–265. IEEE, 2011.

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

W. J. Schakel, B. Van Arem, and B. D. Netten. Effects of cooperative adaptive cruise control on traffic flow stability. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 759–764. IEEE, 2010.

L. Shi and P. Prevedouros. Autonomous and connected cars: Hcm estimates for freeways with various market penetration rates. *Transportation Research Procedia*, 15:389–402, 2016.

S. E. Shladover, D. Su, and X.-Y. Lu. Impacts of cooperative adaptive cruise control on freeway traffic flow. *Transportation Research Record*, 2324(1):63–70, 2012.

S. E. Shladover, C. Nowakowski, X.-Y. Lu, and R. Ferlis. Cooperative adaptive cruise control: Definitions and operating concepts. *Transportation Research Record*, 2489(1): 145–152, 2015.

A. Talebpour and H. S. Mahmassani. Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C: Emerging Technologies*, 71:143–163, 2016.

C. M. Tampère, R. Corthout, D. Cattrysse, and L. H. Immers. A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B: Methodological*, 45(1):289–309, 2011.

K. Tiaprasert, Y. Zhang, C. Aswakul, J. Jiao, and X. Ye. Closed-form multiclass cell transmission model enhanced with overtaking, lane-changing, and first-in first-out properties. *Transportation Research Part C: Emerging Technologies*, 85:86–110, 2017.

M. Van Aerde and H. Rakha. Multivariate calibration of single regime speed-flow-density relationships. In *Proceedings of the 6th 1995 Vehicle Navigation and Information Systems Conference*, pages 334–341, 1995.

J. Van Lint, S. P. Hoogendoorn, and M. Schreuder. Fastlane: New multiclass first-order traffic flow model. *Transportation Research Record*, 2088(1):177–187, 2008.

F. van Wageningen-Kessels, J. Van Lint, S. Hoogendoorn, and C. Vuik. Multiple user classes in the kinematic wave model in lagrangian coordinates. In *Proc. Traffic Granular Flow*, pages 1–7, 2009.

F. van Wageningen-Kessels, H. Van Lint, S. P. Hoogendoorn, and K. Vuik. Lagrangian formulation of multiclass kinematic wave model. *Transportation research record*, 2188(1): 29–36, 2010.

F. van Wageningen-Kessels, H. Van Lint, S. P. Hoogendoorn, and K. Vuik. New generic multiclass kinematic wave traffic flow model: Model development and analysis of its properties. *Transportation Research Record*, 2422(1):50–60, 2014.

Z. Vander Laan and K. F. Sadabadi. Operational performance of a congested corridor with lanes dedicated to autonomous vehicle traffic. *International Journal of Transportation Science and Technology*, 6(1):42–52, 2017.

J. Vander Werf, S. E. Shladover, M. A. Miller, and N. Kourjanskaia. Effects of adaptive cruise control systems on highway traffic flow capacity. *Transportation Research Record*, 1800(1):78–84, 2002.

J. Wang, S. Peeta, and X. He. Multiclass traffic assignment model for mixed traffic flow of human-driven vehicles and connected and autonomous vehicles. *Transportation Research Part B: Methodological*, 126:139–168, 2019.

L. Ye and T. Yamamoto. Impact of dedicated lanes for connected and autonomous vehicle on traffic flow throughput. *Physica A: Statistical Mechanics and its Applications*, 512: 588–597, 2018.

L. Ye, T. Yamamoto, and T. Morikawa. Heterogeneous traffic flow dynamics under various penetration rates of connected and autonomous vehicle. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 555–559. IEEE, 2018.

I. Yperman. *The link transmission model for dynamic network loading*. PhD thesis, KU Leuven, 2007.

I. Yperman, S. Logghe, and B. Immers. The link transmission model: An efficient implementation of the kinematic wave theory in traffic networks. In *Proceedings of the 10th EWGT Meeting, Poznan, Poland*, 2005.

H. Yu, S. Tak, M. Park, and H. Yeo. Impact of autonomous-vehicle-only lanes in mixed traffic conditions. *Transportation research record*, 2673(9):430–439, 2019.

F. Zhu and S. V. Ukkusuri. Modeling the proactive driving behavior of connected vehicles: A cell-based simulation approach. *Computer-Aided Civil and Infrastructure Engineering*, 33(4):262–281, 2018.

W.-X. Zhu and H. Zhang. Analysis of mixed traffic flow with human-driving and autonomous cars based on car-following model. *Physica A: Statistical Mechanics and its Applications*, 496:274–285, 2018.

A. K. Ziliaskopoulos and S. T. Waller. An internet-based geographic information system that integrates data, models and users for transportation applications. *Transportation Research Part C: Emerging Technologies*, 8(1):427–444, 2000.