



Stochastic Systems

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

A Diffusion-Based Analysis of a Multiclass Road Traffic Network

Michel Mandjes, Jaap Storm

To cite this article:

Michel Mandjes, Jaap Storm (2021) A Diffusion-Based Analysis of a Multiclass Road Traffic Network. *Stochastic Systems* 11(1):60–81. <https://doi.org/10.1287/stsy.2019.0065>

This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “*Stochastic Systems*. Copyright © 2021 The Author(s). <https://doi.org/10.1287/stsy.2019.0065>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Copyright © 2021, The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

A Diffusion-Based Analysis of a Multiclass Road Traffic Network

Michel Mandjes,^{a,b,c} Jaap Storm^{d,e}

^aKorteweg-de Vries Institute for Mathematics, University of Amsterdam, 1098 XH Amsterdam, Netherlands; ^bEURANDOM, Eindhoven University of Technology, 5600 MB Eindhoven, Netherlands; ^cFaculty of Economics and Business, Amsterdam Business School, University of Amsterdam, 1001 NL Amsterdam, Netherlands; ^dDepartment of Mathematics, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, Netherlands; ^eDepartment of Mathematics and Computer Science, Technische Universiteit Eindhoven, 5600 MB, Eindhoven, Netherlands
Contact: m.r.h.mandjes@uva.nl (MM); p.j.storm@tue.nl,  <https://orcid.org/0000-0003-3165-4528> (JS)

Received: December 19, 2019

Revised: April 24, 2020

Accepted: July 7, 2020

Published Online in Articles in Advance:
February 17, 2021

<https://doi.org/10.1287/stsy.2019.0065>

Copyright: © 2021 The Author(s)

Abstract. This paper studies a stochastic model that describes the evolution of vehicle densities in a road network. It is consistent with the class of (deterministic) kinematic wave models, which describe traffic flows based on conservation laws that incorporate the macroscopic fundamental diagram (a functional relationship between vehicle density and flow). Our setup is capable of handling multiple types of vehicle densities, with general macroscopic fundamental diagrams, on a network with arbitrary topology. Interpreting our system as a spatial population process, we derive, under natural scaling, fluid, and diffusion limits. More specifically, the vehicle density process can be approximated with a suitable Gaussian process, which yield accurate normal approximations to the joint (in the spatial and temporal sense) vehicle density process. The corresponding means and variances can be computed efficiently. Along the same lines, we develop an approximation to the vehicles' travel time distribution between any given origin and destination pair. Finally, we present a series of numerical experiments that demonstrate the accuracy of the approximations and illustrate the usefulness of the results.



Open Access Statement: This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as "Stochastic Systems. Copyright © 2021 The Author(s). <https://doi.org/10.1287/stsy.2019.0065>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>."

Funding: This research is partly funded by the Netherlands Organization for Scientific Research (NWO) Gravitation project Networks [Grant 024.002.003].

Keywords: conservation laws • diffusion approximation • functional central limit theorem • functional law of large numbers • fundamental diagram • road traffic networks • traffic flow theory • travel time

1. Introduction

A substantial body of literature focuses on describing and predicting the dynamics of vehicles on road traffic networks. A broad range of traffic flow models has been developed, each of them focusing on specific aspects. Their ultimate goal lies in the development of mechanisms that effectively control streams of vehicles.

The majority of the existing traffic flow models is of a deterministic nature. These range from macroscopic models (typically in the form of partial differential equations that describe vehicles as continuous flows that obey physical laws) to microscopic models (incorporating the driving behavior pertaining to individual vehicles). However, as was pointed out in Qu et al. (2017), besides physical laws, traffic flows are also strongly affected by various microscopic variables, such as the different perceptions, moods, responses, and driving habits of individual car drivers. This realization has led to the consensus (Qu et al. 2017, section 1) that such microscopic variables should be modeled as random variables. Thus, to accurately describe streams of vehicles in a road traffic network, *stochastic* traffic flow models are needed. In addition, as argued in great detail in Jabari and Liu (2012), probabilistic traffic flow models are particularly useful from the viewpoint of traffic simulation, estimation, and online control.

In relation to the choice of a suitable probabilistic model, as was pointed out in Jabari and Liu (2012), a frequently used type of stochastic traffic flow models are macroscopic deterministic models with added noise. These, however, result in an inconsistency with the corresponding deterministic models in the case of nonlinear dynamics and possibly lead to the undesirable feature of negative sample paths. Truncation of the noise, to prevent negative sample paths from happening, would only exacerbate the previously mentioned inconsistency. Turning to microscopic stochastic traffic flow models allows for explicit modeling of stochastic behavior. Examples are cellular automata models (Maerivoet and De Moor 2005), with a seminal example being the celebrated Nagel-Schreckenberg model, and car following models (van Wageningen-Kessels et al. 2015, section 3). However, because of the complexity of the underlying dynamics, for such models,

the computation of the vehicle density distribution at a given time is typically intractable for instances of a realistic size.

An approach frequently relied on in macroscopic models, based on conservation laws, makes use of the so-called *fundamental diagram*, which describes the functional relation between the vehicle density and velocity. The origin of such *kinematic wave models*, which succeed in replicating empirical traffic phenomena, lies in the seminal papers by Lighthill and Whitham (1955) and Richards (1956). In the kinematic wave framework, the flow of traffic mass is modeled by a conservation law, with the additional feature that the velocity of mass at a position in space is a function of the density of this mass. Obviously, to facilitate accurate performance predictions, having a handle on the precise shape of the fundamental diagram is crucial. In particular, researchers try to reproduce physical phenomena like the *scatter* observed in empirical studies, the *drop in capacity* at the onset of congestion, *hysteresis* while accelerating and decelerating, and the effect of multiple types of vehicles (Chanut and Buisson 2003, van Wageningen-Kessels et al. 2015).

Although traditionally the focus was on single-class models (Drake et al. 1967, Smulders 1990, Daganzo 1994), in which all vehicles are essentially exchangeable, in practice, multiple classes should be distinguished. This explains why more recently deterministic multiclass models (Wong and Wong 2002, Chanut and Buisson 2003, Logghe and Immers 2008) have been developed, which capture the heterogeneity of vehicles and drivers. However, these models come with new complications, in that they do not automatically satisfy the *anisotropic* (informally meaning that information travels slower than the fastest class of vehicles) and even *hyperbolic* (informally meaning that any single vehicle has only local influence on the system dynamics) properties of conservation laws. As discussed in Daganzo (1995b) and van Wageningen-Kessels et al. (2015), realistic traffic flow models should obey these properties.

The main conclusion of the previous information is that there is a clear demand for stochastic multiclass traffic flow models that are flexible enough to cover various generic physical phenomena. The discrete-space stochastic model proposed in Jabari and Liu (2012, 2013), is referred to by Qu et al. (2017) as one of the few models that provides an explicit expression for the vehicle density distribution. In addition, it is consistent with kinematic wave models, and it aligns with the previously mentioned physical effects that are inherent to the fundamental diagram. Although this model has these attractive properties, several significant improvements can be made. In the first place, the focus is exclusively on the class of fundamental diagrams that involve just single-type vehicle densities, not covering the practically relevant multitype case. In the second place, the model considers isolated road segments rather than more general networks. The third issue relates to the underlying analysis that is based on a diffusion approach, leading to a Gaussian approximation of vehicle densities. Although the claimed diffusion, under the proposed scaling, is correct, there is a need for addressing the formal details of the mathematical underpinnings. Finally, in Jabari and Liu (2012, 2013), the focus is fully on describing the distribution of the vehicle densities, whereas no attention is paid to computing the distribution of travel times between given origin-destination pairs.

Our work concerns a rigorous analysis of a generalized version of the cellular stochastic traffic flow model that was introduced and analyzed in Jabari and Liu (2012, 2013), remedying the issues identified in the previous paragraph. More specifically, the contributions of our paper are the following:

- Unlike in the single-class model of Jabari and Liu (2012, 2013), we focus on a setup capable of handling a *multiclass* macroscopic fundamental diagram. Although in our presentation we focus on a road segment consisting of multiple cells, we point out how the analysis naturally extends to networks of arbitrary size and structure.
- Moreover, we embed the model in a rigorous mathematical framework that allows us to appeal to the theory of *spatial population processes* Kurtz (1981). We show how this well-developed machinery facilitates the formal establishment, under a natural scaling, of *fluid* and *diffusion limits* that are consistent with kinematic wave traffic flow theory.
- The fluid limit provides accurate *approximations* for the means of the vehicle densities, whereas the diffusion limit can be used to approximate the corresponding correlations (both in the spatial and temporal sense). Informally, these fluid and diffusion approximations can be seen as a law of large numbers and a central limit theorem, respectively, at the sample-path level. Our results allow for generic macroscopic fundamental diagrams (fulfilling a mild regularity assumption) and provide explicit expressions for means and (co)variances that can be numerically evaluated in an efficient manner.
- Whereas Jabari and Liu (2012, 2013) focus on vehicle densities only, we point out how diffusion results can be used to produce accurate approximations of the distribution of the *travel time* experienced by vehicles moving through the network.

• The accuracy of our approximation procedure is demonstrated through a series of *numerical experiments*. In addition, we show that our methodology is able to reproduce known traffic phenomena such as forward propagation, backward moving traffic jams, and shockwave formation, with a per-phenomenon consistent estimate for the variances and covariances of traffic densities.

The results obtained in this paper can be used to numerically evaluate the impact of various traffic control mechanisms. For instance, as imposing a speed limit effectively means an adaptation of the fundamental diagram, we can assess the efficacy of such a measure. Importantly, the framework used is highly flexible, in that the fundamental diagram needs to fulfill only mild regularity assumptions. This flexibility can be exploited when studying the impact of technological developments on traffic dynamics. For instance, one can evaluate scenarios in which a certain fraction of the cars is self-driving; an obvious prerequisite is the availability of (a proxy of) the corresponding fundamental diagram.

This paper is organized as follows. To shed light on the properties stochastic traffic flow models should fulfill, we first provide in Section 2 a brief account of kinematic wave models. We pay attention to this class of models because, although being essentially standard for many readers with a background in transportation research, researchers with a background in operations research and applied probability may be less familiar with it. Building on kinematic wave models, we define in Section 3 our multiclass counterpart of the discrete-space model introduced in Jabari and Liu (2012, 2013). Importantly, the setup proposed is highly flexible and thus covers a wide variety of macroscopic fundamental diagrams proposed in the literature. Then, in Section 4, we define our scaling and prove the fluid and diffusion limit results for the vehicle densities, yielding a Gaussian approximation for the corresponding distribution. Then it is shown in Section 5 how related diffusion results lead to an approximation for the travel time distribution. Numerical experiments are provided in Section 6, showing the method's accuracy and its capability to reproduce well-known phenomena. Section 7 presents an account of extensions, generalizations, and ramifications. Concluding remarks are presented in Section 8.

2. Theory of Kinematic Wave Models

The stochastic traffic flow model, which will be the main object of study in this paper, will be defined in Section 3. There we define our model in such a way that the mean dynamics obey a conservation law, which is the fundamental modeling principle in every kinematic wave traffic flow model. To provide the reader with the necessary background, we review in this section the basics of kinematic wave models and corresponding conservation laws.

2.1. Kinematic Wave Models and Conservation Laws

Kinematic wave models are a class of macroscopic traffic flow models that consider a road segment without any intermediate sources or sinks as a continuous space and study the propagation of vehicle density (defined as the number of vehicles per unit length of road) over the road segment. The absence of sources and sinks implies that traffic density should be preserved over the segment, which suggests the use of conservation laws to describe the evolution of traffic density over the segment as *fluid*.

In our framework, we consider m different types of vehicles. Let, for $j \in \{1, \dots, m\}$, the type- j vehicle density be given by $\rho_j(x, t)$ and the type- j velocity by $v_j(x, t)$, where t denotes time and x the position on the segment. Conservation of mass (or density) in an arbitrary part of the road segment $[x_1, x_2]$, in a time interval $[t_1, t_2]$, is expressed as

$$\int_{x_1}^{x_2} \rho_j(x, t_2) dx - \int_{x_1}^{x_2} \rho_j(x, t_1) dx = \int_{t_1}^{t_2} \rho_j(x_1, t) v_j(x_1, t) dt - \int_{t_1}^{t_2} \rho_j(x_2, t) v_j(x_2, t) dt, \quad (1)$$

with $j \in \{1, \dots, m\}$ (LeVeque 1992, section 2.1). Here it is implicit that traffic moves from x_1 to x_2 .

Now, in traffic flow theory, there is the fundamental and empirically backed assumption that the velocity $v_j(x, t)$ is a function of *all* the type densities $\rho_k(x, t)$, where $k \in \{1, \dots, m\}$. In the literature, this functional relationship is referred to as the (*macroscopic*) *fundamental diagram* (MFD), usually expressed as

$$q_j(\rho(x, t)) = \rho_j(x, t) v_j(\rho(x, t)),$$

where q_j is the flow of type- j density, and ρ is the length- m vector of traffic densities. Combining (1) with the MFD allows the derivation (LeVeque 1992, section 2) of the following system of conservation laws:

$$\frac{\partial \rho_j(x, t)}{\partial t} + \frac{\partial q_j(\rho)}{\partial \rho} \frac{\partial \rho(x, t)}{\partial x} = 0, \quad \rho_j(x, 0) = \rho_{j,0}(x), \quad (2)$$

where $\partial q_j(\rho)/\partial \rho$ is to be interpreted as the gradient of q_j and $\rho_{j,0}(x)$ is some initial vehicle density (as a function of the position x).

2.2. The Macroscopic Fundamental Diagram

As mentioned in the introduction, the fundamental diagram describes flow as a function of the vehicle density. The generally accepted view is that flow is an (approximately) concave function of the overall density, which is zero on the boundaries of some compact domain and positive on its interior. The reasoning behind this is simple: when traffic density is low, vehicles are not obstructed and can thus drive at maximum speed. As the density increases, vehicles will be obstructed to a greater extent because of congestion, thus lowering their velocity. Ultimately, if the traffic density becomes too high, then vehicles do not have enough space in front of them, resulting in zero velocity and flow.

The first fundamental diagram, for single-class traffic, can be found in a pioneering study by Greenshields (1934). Since then, researchers have tried to find functional forms that reproduce various traffic phenomena as well as possible (Maerivoet and De Moor 2005, section 5). For single-class traffic, three notable MFDs can be found in Daganzo (1994), Drake et al. (1967), and Smulders (1990). In recent years, to capture effects caused by mixed vehicle traffic, multiclass MFDs have been proposed (Wong and Wong 2002, Benzoni-Gavage and Colombo 2003, Chanut and Buisson 2003, Zhang et al. 2006, Ngoduy and Liu 2007, Logghe and Immers 2008).

2.3. Solutions for Conservation Laws, Relation with Stochastic Traffic Flow Model

The functional forms for ρ and q that are used in practice are typically not differentiable, so that the problem in (2) has no strong solutions. One therefore resorts to weak solutions, satisfying additional conditions (e.g., Rankine-Hugoniot and entropy conditions) to ensure uniqueness and correct physical behavior. For a mathematical treatment of conservation laws see, for instance, Evans (2010, section 3.4 and chapter 11). Whereas it is often possible to show existence of weak solutions of (2), they rarely have explicit analytical expressions. Therefore, for practical applications, one usually relies on numerical methods. In traffic flow theory, the most widely used method is the Godunov method (LeVeque 1992, section 13), which has many attractive features, assuming that the Courant-Friedrichs-Lewy (CFL) condition (LeVeque 1992, equations (10.55) and (13.11)) is satisfied (Chanut and Buisson 2003, paragraph above equation (18)).

The Godunov method exactly solves (2) with an approximate, piece-wise constant, initial condition. The x axis, which is the road segment in our case, is discretized into d cells (which can be done according to a variable mesh), and the initial density is taken constant in each cell, by integrating the (m -dimensional) initial condition $\rho(x, 0)$ between the boundaries of each cell. This defines a set of $d - 1$ Riemann problems, the solution of which describes the flow between each of the cell boundaries exactly. As such, one can describe the exact evolution of the system with the equation for $i \in \{2, \dots, d\}$,

$$\rho(x_i, t + \Delta t) = \rho(x_i, t) + \frac{\Delta t}{\Delta x_i} (q_{i-1} - q_i), \quad (3)$$

where $\rho(x_i, t)$ is now an m -dimensional vector of densities at time t in the i th cell, Δt and Δx_i are, respectively, the lengths of the discrete time steps and the mesh width of cell i , and q_i is the (m -dimensional) vector of flows on the boundary from cell i to cell $i + 1$, given by the solution of the corresponding Riemann problem. To be able to use the Godunov scheme, the challenge is to give a complete description of discrete flux-function q_i . In general, there is no explicit solution for the associated Riemann problem, but there are conditions under which a solution is guaranteed for small time steps (Evans 2010, theorem 4 in section 11).

We finish this section by mentioning that an approximate solution of (2), using the Godunov method, is also known as the cell transmission model (CTM). This concept, from Daganzo (1994), relates to a discrete-time, discrete-space traffic flow model. One can show that (2) can be recovered from the CTM by letting the cell lengths go to 0 (Jabari and Liu 2012). In our paper, we take an arbitrary (possibly multiclass) MFD as our starting point. We use the associated discrete flux function to define a stochastic traffic flow model, such that the mean dynamics are given by the CTM. In Section 4, we approximate the variance of this process through a diffusion limit, which allows for studying the distributions of vehicle densities in various types of dynamics, as will be illustrated in the numerical experiments of Section 6.

3. Model Definition

In this section, we introduce the stochastic traffic flow model as a strong Markov process, for which we subsequently define the corresponding state space and the infinitesimal generator. To keep the exposition as

clear as possible, in this paper we primarily concentrate on a single road segment. However, the model, the underlying mathematical theory, and scaling results we present for this road segment setup carry over to a general network setting, as will be pointed out in Section 7 (along with various other extensions). We conclude the section by giving two relevant examples of kinematic wave models that we extend into stochastic traffic flow models and that will be used later in the paper when numerically illustrating our results.

3.1. State Space

Consider a road segment without any intermediate sources or sinks, on which m different types of vehicles are defined. We divide the road in d cells, where ℓ_i is the length of cell i , for $i \in \{1, \dots, d\}$. We denote by $X_{ij}(t)$ the number of type- j vehicles, for $j \in \{1, \dots, m\}$, that are in cell i at time $t \in [0, \infty)$. In the context of our road-traffic model, it is evident that the number of type- j vehicles in cell i has an upper bound, which we denote X_{ij}^{jam} . As a consequence, $X_{ij}(t) \in \{0, 1, \dots, X_{ij}^{\text{jam}}\}$. In the sequel, we denote $X(t)$ for the md -dimensional random vector with entries $X_{ij}(t)$.

Equivalently to $X_{ij}(t)$, we can consider the type- j vehicle density in cell i at time t , which we denote $\rho_{ij}(t)$. It is defined as

$$\rho_{ij}(t) := \frac{X_{ij}(t)}{\ell_i}, \quad (4)$$

taking values in $\{0, 1/\ell_i, 2/\ell_i, \dots, X_{ij}^{\text{jam}}/\ell_i\}$. In the literature, $\rho_{ij}^{\text{jam}} := X_{ij}^{\text{jam}}/\ell_i$ is referred to as the *jamming density* (of type- j vehicles in cell i , that is).

3.2. Transition Rates

To describe the flows of type- j vehicles between the cells, we assume the existence of a discrete flux function, obtained from a general (possibly multiclass) MFD by solving the Riemann problem. We proceed by formally introducing such a discrete flux function, describing the flow from cell i to cell $i+1$ for each of the types j . Denote q_{ij}^{max} for the largest possible flow of type- j vehicles between the neighboring cells i and $i+1$. The assumption below says that the flow from cell i to cell $i+1$ is a function of the densities in both cells, that is, of $(\rho_{ik})_k$ and $(\rho_{i+1,k})_k$, where the first and last cell have to be handled slightly differently. Moreover, the function is dependent on i , that is, on the position of a cell, to model cell-dependent MFDs. This is particularly useful if one wishes to incorporate heterogeneity between the cells because of, for example, differences in road geometry and maximum velocity.

Assumption 1 (Macroscopic Fundamental Diagram). *For $i \in \{1, \dots, d-1\}$, we assume that there is a function*

$$\tilde{q}_i : \prod_{j=1}^m \left([0, \rho_{ij}^{\text{jam}}] \times [0, \rho_{i+1,j}^{\text{jam}}] \right) \rightarrow \prod_{j=1}^m [0, q_{ij}^{\text{max}}], \quad (5)$$

which is Lipschitz-continuous with Lipschitz constant $K_i \in (0, \infty)$. We also assume that there is a function

$$\tilde{q}_0 : \prod_{j=1}^m \left([0, \infty) \times [0, \rho_{1j}^{\text{jam}}] \right) \rightarrow \prod_{j=1}^m [0, q_{0j}^{\text{max}}], \quad (6)$$

which is Lipschitz-continuous with Lipschitz constant $K_0 \in (0, \infty)$; likewise, we assume that there is a function

$$\tilde{q}_d : \prod_{j=1}^m \left([0, \rho_{dj}^{\text{jam}}] \times [0, \infty) \right) \rightarrow \prod_{j=1}^m [0, q_{dj}^{\text{max}}], \quad (7)$$

which is Lipschitz-continuous with Lipschitz constant $K_d \in (0, \infty)$. It is, in addition, assumed that the j th component of $\tilde{q}_i$ is 0 whenever, for $j \in \{1, \dots, m\}$, the j th argument is 0; all components of q_i are 0 whenever, for some $j \in \{1, \dots, m\}$, the $(m+j)$ th argument is $\rho_{i+1,j}^{\text{jam}}$.

The Lipschitz property is natural to assume, because by considering the density as a continuum, it entails that the flow should be a continuous function of the density and that the change in flow should be bounded when adding or removing a vehicle from a cell. The last part of the assumption ensures that if there are no cars in the origin cell, then there can be no flow between the two cells involved; likewise, if the destination cell is full.

Using Assumption 1, we proceed by defining the infinitesimal generator that makes $\{X(t), t \geq 0\}$, or equivalently $\{\rho(t), t \geq 0\}$, a Markov process. Consider two neighboring cells i and $i+1$, and let the current

density be ρ . For $i \in \{1, \dots, d-1\}$ and $j \in \{1, \dots, m\}$, we let the time until a car of type j passes the boundary between cells i and $i+1$ be an exponentially distributed random variable with rate

$$q_{i,j}(\rho) := \tilde{q}_{i,j}((\rho_{ik})_k, (\rho_{i+1,k})_k).$$

To model arrivals at cell 1, we introduce independent Poisson processes with rate λ_j , for $j \in \{1, \dots, m\}$. We let the time until a type- j vehicle enters at cell 1, for $j \in \{1, \dots, m\}$, be an exponentially distributed random variable with rate

$$q_{0,j}(\rho) := \min \left\{ \lambda_j, \sup_{x \in (\mathbb{R}^+)^m} \tilde{q}_{0,j}(x, (\rho_{1k})_k) \right\}.$$

The interpretation of this formula is that cars of type j constantly arrive at rate $\lambda_j > 0$ but that the flow is bounded because of the vehicles already present in cell 1. Similarly, we let the time until a type- j vehicle leaves cell d , for $j \in \{1, \dots, m\}$, be an exponentially distributed random variable with rate

$$q_{d,j}(\rho) := \min \left\{ \nu_j, \sup_{x \in (\mathbb{R}^+)^m} \tilde{q}_{d,j}((\rho_{dk})_k, x) \right\},$$

where $\nu_j > 0$. The interpretation is that vehicles of type j depart at a rate that depends on the densities in cell d but that is bounded by some maximal outflow rate ν_j . We assume that, given ρ , all the exponentially distributed times, as defined previously, are independent.

From the previous statements, it follows that the evolution of the number of type- j vehicles in cell i , with $i \in \{1, \dots, d\}$ and $j \in \{1, \dots, m\}$ is given, for a given initial state $X(0)$, by

$$X_{ij}(t) = X_{ij}(0) + Y_{i-1,j} \left(\int_0^t q_{i-1,j}(\rho(s)) ds \right) - Y_{i,j} \left(\int_0^t q_{i,j}(\rho(s)) ds \right), \quad (8)$$

where $Y_{i,j}(t)$ are independent unit-rate Poisson processes for $i \in \{0, \dots, d\}$ and $j \in \{1, \dots, m\}$. To see this, the Poisson processes $Y_{i,j}$ are counting how many cars of type j pass between the boundary between cells i and $i+1$. To measure the impact on the density, we scale these processes so that they have jumps of size $1/\ell_i$. By using (4), we have the analogous process

$$\rho_{ij}(t) = \rho_{ij}(0) + \frac{1}{\ell_i} Y_{i-1,j} \left(\int_0^t q_{i-1,j}(\rho(s)) ds \right) - \frac{1}{\ell_i} Y_{i,j} \left(\int_0^t q_{i,j}(\rho(s)) ds \right). \quad (9)$$

We have thus constructed the dm -dimensional Markov process $\{\rho(t), t \geq 0\}$.

3.3. Example MFDs

To illustrate the flexibility of Assumption $\sim$ Assumption: MFD, we now give two concrete examples of MFDs and their associated discrete flux-functions $\tilde{q}$. In Section 6, we will use these examples.

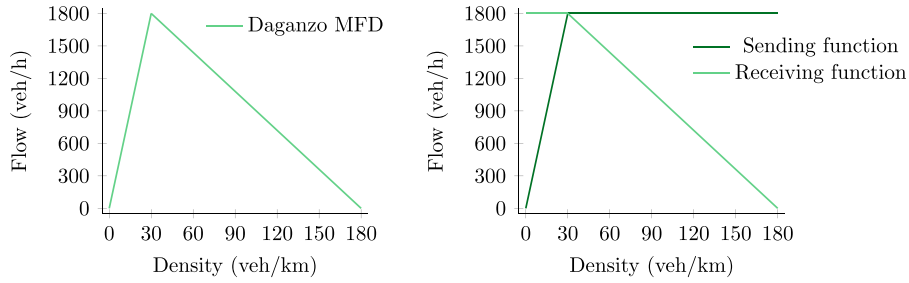
Example 1 (Daganzo MFD). Our first example is a discrete flux function for a single-class CTM that was proposed in Daganzo (1994). The flux function describes the flow between two neighboring cells i and $i+1$. Because traffic flows in one direction, cell i will be sending traffic and cell $i+1$ will be receiving traffic. We define the *sending* function $S(\cdot)$ and the *receiving* function $R(\cdot)$ by

$$S(\rho) := v^f \rho \wedge q^{\max}, \quad R(\rho) := w(\rho^{\text{jam}} - \rho) \wedge q^{\max}.$$

Here v^f is the velocity at which vehicles drive when they do not experience congestion (called the *free-flow* regime), and w is the velocity at which traffic jams move upstream. These specific functions are based on the assumption that vehicles drive with their maximum velocity v^f when vehicle density is low, whereas the flow decreases linearly when congestion occurs. The flow between the cells is now given by

$$\tilde{q}^D(\rho_i, \rho_{i+1}) = \min\{S(\rho_i), R(\rho_{i+1})\},$$

where ρ_i is the density of cell i . In Figure 1, we have plotted an example with $v^f = 60$ km/h, $w = 12$ km/h, $q^{\max} = 1800$ veh/h, and $\rho^{\text{jam}} = 180$ veh/km, with the MFD on the left and the sending and receiving functions that characterize the discrete flux function on the right.

Figure 1. Daganzo MFD (Left) and Corresponding Sending and Receiving Function, Determining the Associated Discrete Flux Function (Right)

It is not hard to check that $\tilde{q}^D$ satisfies Assumption 1. Because both $S(\cdot)$ and $R(\cdot)$ are Lipschitz, the minimum of $S(\cdot)$ and $R(\cdot)$ is Lipschitz as well. Moreover, when $\rho_i = 0$ or $\rho_{i+1} = \rho^{\text{jam}}$, then $\tilde{q}^D$ equals zero.

The paradigm of a sending and receiving function can be generalized to obtain other single-class discrete flux functions. More precisely, the functions $S(\cdot)$ and $R(\cdot)$ can be every nonnegative Lipschitz continuous function with compact support. Moreover, it can be shown that every single-class MFD has a discrete flux function, obtained by solving the Riemann problem associated to the Godunov scheme, which can be written as the minimum of a sending and receiving function.

Example 2 (Chanut and Buisson MFD). This MFD was proposed in Chanut and Buisson (2003) for a multilane, two-class traffic flow model and is a generalization of Smulders' fundamental diagram (Smulders 1990). In the model, vehicle classes are differentiated by velocities and lengths, and therefore, they influence the overall density and flow differently. We first present the MFD for the kinematic wave model and present the discrete flux function afterward.

For simplicity we refer to the first class of vehicles as *cars* and to the second class of vehicles as *trucks*. In the model, the jamming density is given by

$$\rho^{\text{jam}}(\rho_1, \rho_2) := \frac{N}{\rho_1 L_1 + \rho_2 L_2} (\rho_1 + \rho_2),$$

where ρ_1, ρ_2 are the car and truck densities, L_1 and L_2 are the respective vehicle lengths, and N is the number of lanes. The critical density, which distinguishes the free-flow regime and the congestion regime, and which is parameterized by $\beta \in [0.2, 0.5]$, is given by

$$\rho^c(\rho_1, \rho_2) := \beta \rho^{\text{jam}}(\rho_1, \rho_2).$$

In the free-flow regime, defined as $\rho_1 + \rho_2 \leq \rho^c(\rho_1, \rho_2)$, the vehicle velocity starts at the maximum velocity v_i^f when $\rho_1 + \rho_2 = 0$ and decreases linearly in $\rho := \rho_1 + e\rho_2$, with $e := L_2/L_1$, until the critical velocity v^c is reached (at the critical density ρ^c). In congestion, both vehicle classes have the same velocity, and it is assumed that the total flow decreases linearly in ρ . The quantity ρ can be interpreted as the density in terms of passenger-car equivalents (PCEs); one truck is equivalent to e cars. In the congested regime, the flow of vehicles, measured in PCEs, can thus be expressed as

$$q_{\text{pce}}(\rho_1, \rho_2) := C \frac{\rho^{\text{jam}}(\rho_1, \rho_2) - (\rho_1 + \rho_2)}{\rho^{\text{jam}}(\rho_1, \rho_2) - \rho^c(\rho_1, \rho_2)},$$

where $C := v^c \rho^c(\rho_1, 0)$. If $\rho_1 + \rho_2 \leq \rho^c(\rho_1, \rho_2)$ (in free-flow, that is), then the flow of both vehicle classes is now given by

$$\begin{aligned} q_1 &= \rho_1 v_1, & v_1 &= v_1^f - \left(v_1^f - v^c \right) \frac{\rho_1 + \rho_2}{\rho^c(\rho_1, \rho_2)}; \\ q_2 &= \rho_2 v_2, & v_2 &= v_2^f - \left(v_2^f - v^c \right) \frac{\rho_1 + \rho_2}{\rho^c(\rho_1, \rho_2)}, \end{aligned}$$

whereas if $\rho_1 + \rho_2 > \rho^c(\rho_1, \rho_2)$ (in congestion, that is), then

$$q_1 = \rho_1 v, \quad q_2 = \rho_2 v, \quad v = \frac{q_{\text{pce}}(\rho_1, \rho_2)}{\rho_1 + e\rho_2}.$$

To use this MFD in a discrete model, one has to determine the associated discrete flux function by solving the Riemann problem for every possible combination of two neighboring cell densities. This discrete flux function is given in Chanut and Buisson (2003) and can be written in terms of the sent density and the received density, as in Example 1, except for the case where the upstream cell is in the free-flow regime, the downstream cell is in the congested regime, and the shock discontinuity propagates at a negative speed.

We conclude this example by summarizing the discrete flux function. Let Δ and Ω denote the sent and received density, respectively. If a cell is in the free-flow regime, then they are given by

$$\Delta = \Delta_1 + e\Delta_2, \quad \Omega = C, \quad \Delta_i = \rho_i v_i, \quad i = 1, 2,$$

whereas if a cell is in the congested regime, then

$$\Delta = \Delta_1 + e\Delta_2 C, \quad \Omega = (\rho_1 + e\rho_2)v, \quad \Delta_i = \frac{\rho_i}{\rho_1 + e\rho_2} C, \quad i = 1, 2.$$

Now, if for two neighboring cells, the upstream cell is in free flow and the downstream cell is in congestion, then one has to consider the speed s of the shock discontinuity, which is given by

$$s = \frac{(q_1^u + eq_2^u) - (q_1^d + eq_2^d)}{(\rho_1^u + e\rho_2^u) - (\rho_1^d + e\rho_2^d)},$$

where the upper indexes u and d denote whether the quantities q_i and ρ_i belong to the upstream or downstream cell. If $s < 0$, then the sending-receiving concept does not apply, as the flow between the two cells is given by an intermediate state of densities, according to the solution of the Riemann problem; then we have

$$q_1 = v \frac{q_1^u - s\rho_1^u}{v - s}; \quad q_2 = v \frac{q_2^u - s\rho_2^u}{v - s},$$

where v is the velocity of vehicles in the downstream cell.

In any other configuration of two neighboring states, for example, if either the upstream cell and downstream cell are both in the free-flow regime or if $s > 0$, then the flow is given by the following sending-receiving scheme: if $\Delta^u \leq \Omega^d$, then we have

$$q_1 = \Delta_1^u; \quad q_2 = \Delta_2^u,$$

whereas if $\Delta^u > \Omega^d$, then

$$q_1 = \frac{\rho_1^u}{\rho_1^u + e\rho_2^u} \Omega^d; \quad q_2 = \frac{\rho_2^u}{\rho_1^u + e\rho_2^u} \Omega^d.$$

4. Scaling Limits

Now that we have defined our model, we proceed by analyzing it. As exact analysis is infeasible, we do so under a certain scaling. Concretely, we show that after an appropriate centering and normalization, the md -dimensional process $\rho(\cdot)$ converges to a Gaussian process as the scaling parameter n goes to ∞ . In this way, we can find an approximation for the distribution of $(\rho(t_1), \dots, \rho(t_k))$ at times $0 \leq t_1 < \dots < t_k < \infty$ (corresponding to a mdk -dimensional normally distributed random vector). The scaling we consider concerns both the lengths of the cells and time, that is, we work with $\ell_i \mapsto n\ell_i$ for $i \in \{1, \dots, d\}$, and $t \mapsto nt$. We thus obtain a sequence of processes $\{\rho^n(t)\}_n$.

A first main result of this section is that $\{\rho^n(\cdot)\}_n$ converges almost surely, as $n \rightarrow \infty$, to a limiting process $\bar{\rho}(\cdot)$. This convergence is usually referred to as a *first order approximation*, or a *fluid limit*, and can be seen as a law of large numbers at the path level. This fluid limit is consistent with the CTM associated to the function we choose in Assumption 1, as its dynamics coincide with (3). As a second main result of this section, we prove that the sequence of processes $\hat{\rho}^n(\cdot) := \sqrt{n}(\rho^n(\cdot) - \bar{\rho}(\cdot))$ converges in distribution, as $n \rightarrow \infty$, to a Gaussian process, which is a *second-order approximation* or a *diffusion limit*. This result is essentially a central limit theorem at the path level.

The scaling that we impose can intuitively be understood as follows. For our results, we want to exploit principles underlying the classical law of large numbers and central limit theorem, analogously to when one considers a large sample of (approximately independent and identically distributed) random variables. Hence, we need a large number of vehicles to be present in each of the cells. To achieve this, we scale the length of the cells by a factor of n , where we let n grow large. However, the discrete flux functions defined in (5)–(7) have a compact codomain. This means that, to keep the flow across cells invariant under the scaling of the cell lengths, we need to scale the number of vehicles that jump per unit time, or equivalently, scale time. Put differently, without scaling these fluxes, we have the undesired effect that the processes

$$\frac{1}{n\ell_i} Y_{ij} \left(\int_0^t q_{ij}(\rho(s)) ds \right)$$

converge almost surely to the zero process as $n \rightarrow \infty$. We observe that, to get a nondegenerate limiting process, leaving the flux over finite time intervals invariant, we should scale time by a factor n .

In this section, we first state and prove the fluid limit, and then move to the diffusion limit. In the sequel we denote by $\rho^n(t)$ the process $\rho(nt)$, where the lengths of cells are $n\ell_i$, for $i \in \{1, \dots, d\}$. As a closing comment, all results in this section can be translated into results for $X(\cdot)$, because of (4).

4.1. Fluid limit

To prove both the first- and second-order approximations for $\rho(t)$ (and hence for $X(t)$), we use two theorems from (Kurtz 1981, chapter 8). These two theorems are written in terms of population processes, but directly translate to the setting we consider. To illustrate this, consider the scaled processes $\rho^n(t)$. By (9), we have

$$\begin{aligned} \rho_{ij}^n(t) &= \rho_{ij}^n(0) + \frac{1}{n\ell_i} Y_{i-1,j} \left(\int_0^{nt} q_{i-1,j} \left(\frac{\rho(s)}{n} \right) ds \right) - \frac{1}{n\ell_i} Y_{ij} \left(\int_0^{nt} q_{ij} \left(\frac{\rho(s)}{n} \right) ds \right) \\ &= \rho_{ij}^n(0) + \frac{1}{n\ell_i} Y_{i-1,j} \left(n \int_0^t q_{i-1,j}(\rho^n(s)) ds \right) - \frac{1}{n\ell_i} Y_{ij} \left(n \int_0^t q_{ij}(\rho^n(s)) ds \right), \end{aligned}$$

where the second equality is because of change variables (with s being replaced by ns), and the second line agrees with Kurtz (1981, equation (8.6)). Hence, we only need to verify the assumptions of Kurtz (1981, theorem 8.1) to establish the fluid limit. We state the theorem and verify these assumptions in the proof. Here, u.o.c. means uniformly on compacts.

Theorem 1 (Fluid Limit). *Under Assumption 1 and $\lim_{n \rightarrow \infty} \rho^n(0) = \bar{\rho}(0)$ for some $\bar{\rho}(0)$, we have*

$$\rho^n(\cdot) \xrightarrow{\text{a.s.}} \bar{\rho}(\cdot), \quad \text{u.o.c.,}$$

where, for $j \in \{1, \dots, m\}$ and $i \in \{1, \dots, d\}$, the process $\bar{\rho}(\cdot)$ satisfies the integral equation

$$\bar{\rho}_{ij}(t) = \bar{\rho}_{ij}(0) + \frac{1}{\ell_i} \int_0^t q_{i-1,j}(\bar{\rho}(s)) ds - \frac{1}{\ell_i} \int_0^t q_{ij}(\bar{\rho}(s)) ds. \quad (10)$$

As said before, we want to appeal to Kurtz (1981, theorem 8.1). Therefore, we verify the assumptions of this theorem. Because, by Assumption 1, the functions $\tilde{q}_i$ are Lipschitz on a compact domain, $\sup_x q_{ij}(x)$ exists, for every $i \in \{0, \dots, d\}$ and $j \in \{1, \dots, m\}$. Hence, by the fact that there is a finite number of possible transitions, the first assumption (Kurtz 1981, equation (8.10)) directly follows:

$$\sum_{i=0}^d \sum_{j=1}^m \sup_{x \in \mathbb{R}_+^{dm}} q_{ij}(x) \leq dm \max_{\substack{i \in \{0, \dots, d\} \\ j \in \{1, \dots, m\}}} \sup_{x \in \mathbb{R}_+^{dm}} q_{ij}(x) < \infty.$$

The other assumption, that is, Kurtz (1981, equation (8.11)), is written in our setting as

$$\left\| \sum_{i=1}^d \sum_{j=1}^m \frac{1}{\ell_i} [(q_{i-1,j}(\rho_1) - q_{ij}(\rho_1)) - (q_{i-1,j}(\rho_2) - q_{ij}(\rho_2))] e_{ij} \right\| \leq K \|\rho_1 - \rho_2\|,$$

for all ρ_1, ρ_2 in the state space of $\rho(\cdot)$. This property is an immediate consequence of the Lipschitz assumption that we imposed on $\tilde{q}_i$.

Remark 1. We explain how the integral Equation (10) is consistent with the Godunov-based approximate solution of (2). Consider the time interval $[0, t]$ and take a partition π , such that $0 = t_0 < t_1 < \dots < t_k = t$, and so that the mesh $|\pi| := \sup_{l \in \{1, \dots, k\}} |t_l - t_{l-1}| \leq \Delta t$, where Δt follows from the so-called CFL condition (compare with LeVeque 1992, equation (13.11)) that makes the Godunov scheme stable. Using (3) one can write

$$\begin{aligned} \rho_{ij}^G(t) &= \rho_{ij}^G(t_k - 1) + \frac{t_k - t_{k-1}}{\ell_i} [q_{i-1,j}(\rho(t_{k-1})) - q_{i,j}(\rho(t_{k-1}))] \\ &= \rho_{ij}^G(0) + \frac{1}{\ell_i} \sum_{l=1}^k (t_l - t_{l-1}) [q_{i-1,j}(\rho(t_{l-1})) - q_{i,j}(\rho(t_{l-1}))], \end{aligned}$$

where $\rho_{ij}^G(t)$ is now the type- j density in cell i , given by the solution of the Godunov method, as emphasized by the superscript G . Now, we interpret $\rho_{ij}^G(t)$ as a continuous function by linearly interpolating between the values at the t_l , $l = 1, \dots, k$. Letting $k \rightarrow \infty$ such that $\|\pi\| \rightarrow 0$, or instead letting $\Delta t \rightarrow 0$, we have that the previous expression for $\rho_{ij}(t)$ converges to the expression for $\bar{\rho}_{ij}(t)$ in (10), uniformly on compact time intervals.

To conclude, the results of this section show that the fluid limit is consistent with kinematic wave models. In Section 6.1, we show that $\bar{\rho}(\cdot)$ is consistent with the mean dynamics of our model. The added value of our stochastic model, relative to existing deterministic models, lies in the fact that we also have a handle on the likelihood of fluctuations around the fluid limit, using the diffusion limit that we establish in Section 4.2.

4.2. Gaussian Diffusion Limit

We now turn to the diffusion limit by considering the scaled and centered process $\hat{\rho}^n$, defined as

$$\hat{\rho}^n(t) := \sqrt{n}(\rho^n(t) - \rho(t)).$$

When establishing the diffusion limit, it is practical to work with a vector representation of the processes involved, for which we introduce the notation

$$F(\rho(t)) := \sum_{i=1}^d \frac{1}{\ell_i} \sum_{j=1}^m (q_{i-1,j}(\rho(t)) - q_{i,j}(\rho(t))) e_{ij},$$

where $e_{ij} \in \mathbb{Z}^{dm}$ is the vector that has a one on the (i, j) th coordinate and zeros everywhere else. This immediately gives us the following vector representation for $\bar{\rho}(\cdot)$:

$$\bar{\rho}(t) = \bar{\rho}(0) + \int_0^t F(\bar{\rho}(s)) ds. \quad (11)$$

To formulate a weak limit for $\hat{\rho}^n(t)$, we want to use Kurtz (1981, theorem 8.2). This theorem, however, requires the existence of $\partial F(x)$, the matrix of partial derivatives of F , as a continuous and bounded operator, for which the Lipschitz assumption that we imposed on $\tilde{q}_i$ is not enough. In the proof of Kurtz (1981, theorem 8.2), the existence of ∂F is used for two statements. The first of these statements is that the matrix-valued ordinary differential equation

$$\frac{\partial}{\partial t} \Phi(t, s) = \partial F(\rho(t)) \Phi(t, s), \quad \Phi(s, s) = I, \quad (12)$$

has a unique solution. The second statement is that the mapping $P: D_{\mathbb{R}^d}[0, \infty) \rightarrow D_{\mathbb{R}^d}[0, \infty)$, given by

$$P\xi(t) := \xi(t) + \int_0^t \Phi(t, s) \partial F(\bar{X}(s)) \xi(s) ds,$$

is continuous in the Skorohod topology. With the use of standard arguments from differential equations, the result from Kurtz (1981, theorem 8.2) remains true under the weaker condition that ∂F exists in the weak sense and is bounded, which follows from the Lipschitz property that we imposed on $\tilde{q}_i$. For completeness, we summarize these arguments.

Because the $\tilde{q}_i$ are Lipschitz, $F(\cdot)$ is Lipschitz as well, so that ∂F exists in the weak sense, and its component functions are in L^∞ (Evans 2010, theorem 5.8.4 and its remark). Obviously, every weak derivative of $F(\cdot)$ equals the *strong* derivative of $F(\cdot)$ at every point where $F(\cdot)$ is differentiable. Moreover, as $F(\cdot)$ is Lipschitz, the integral Equation (11) has a unique strong solution, so that we can consider $t \mapsto \partial F(\bar{\rho}(t))$ as a (deterministic) function of t .

Combining both observations, we can (and will) from now on let ∂F be the matrix that is given by the partial derivatives of F , whenever they exist; whenever the corresponding partial derivative does not exist, we use $K = \max_{k \in \{1, \dots, (d+1)m\}} \kappa_k$ instead, with κ_k denoting the Lipschitz constant corresponding to the k th coordinate of $F(\cdot)$. That is, ∂F is the matrix of the partial derivatives of F in the weak sense, and we emphasize that (12) should be interpreted accordingly from now on. We then have that the operator $t \mapsto \partial F(\bar{\rho}(t))$ is a deterministic, bounded, almost everywhere continuous function. Therefore, we can apply Carathéodory's theorem (Hale 1969, theorem 5.1) to find that there exists a solution to (12), which is unique because of Hale (1969, theorem 5.3). Furthermore, as ∂F is bounded, the continuity of the function P still holds. (It is noted that simply assuming the condition of Kurtz (1981, theorem 8.2) will not lead to a satisfactory solution, because various frequently used MFDs from the literature do not obey this property.)

We are now ready to state our version of Kurtz (1981, theorem 8.2), the validity of which follows from the previous considerations. We let $Z^n(\cdot) \Rightarrow Z(\cdot)$ denote weak convergence of processes $Z^n(\cdot)$ to a process $Z(\cdot)$, in the space of càdlàg functions with the Skorohod topology.

Theorem 2 (Diffusion Approximation). *Under Assumption 1 and $\lim_{n \rightarrow \infty} \sqrt{n} |\rho^n(0) - \rho_0| = 0$ for some ρ_0 , we have that $\hat{\rho}^n(\cdot) \Rightarrow \hat{\rho}(\cdot)$, where $\hat{\rho}_{ij}(t)$, $i \in \{1, \dots, d\}$, $j \in \{1, \dots, m\}$, obeys the stochastic integral equation*

$$\begin{aligned} \hat{\rho}_{ij}(t) = & \frac{1}{\ell_i} \int_0^t \nabla q_{i-1,j}(\bar{\rho}(s)) \cdot \hat{\rho}(s) ds - \frac{1}{\ell_i} \int_0^t \nabla q_{i,j}(\bar{\rho}(s)) \cdot \hat{\rho}(s) ds \\ & \frac{1}{\ell_i} \int_0^t \sqrt{q_{i-1,j}(\bar{\rho}(s))} dB_{i-1,j}(s) - \frac{1}{\ell_i} \int_0^t \sqrt{q_{i,j}(\bar{\rho}(s))} dB_{i,j}(s), \end{aligned} \quad (13)$$

where $\nabla q_{i-1,j}$ and $\nabla q_{i,j}$ are gradients that are to be interpreted in the weak sense, corresponding with our redefinition of ∂F , where $\mathbf{x} \cdot \mathbf{y}$ denotes the Euclidean inner product for vectors $\mathbf{x}$ and $\mathbf{y}$, and where the $B_{i,j}(\cdot)$ are independent standard Brownian motions.

The limiting process $\hat{\rho}(\cdot)$ thus satisfies the linear stochastic integral equation (13). It is known that this class of stochastic integral equations allows an explicit solution, which is a Gaussian process with a known mean and covariance (Karatzas and Shreve 2012, section 5.6). To state these results, we first introduce some notation.

Let $Q(\rho(t))$ be a vector of length $(d+1)m$ with entries $q_{i-1,j}$, $i \in \{1, \dots, d\} + 1$, and $j \in \{1, \dots, m\}$, ordered lexicographically, that is, $Q_{(i-1)m+j} = q_{i-1,j}$. Now, let H be the $dm \times (d+1)m$ matrix with $H_{kl} := \mathbb{1}_{\{k=l\}} - \mathbb{1}_{\{k+m=l\}}$ for $k \in \{1, \dots, dm\}$ and $l \in \{1, \dots, (d+1)m\}$. In addition, L is a dm -dimensional diagonal matrix, with the k th diagonal element being $1/\ell_i$ if $\lceil k/m \rceil = i$, for $k \in \{1, \dots, dm\}$ and $i \in \{1, \dots, d\}$. It is readily checked that with this new notation we have $F(\rho(t)) = LH Q(\rho(t))$. The process $\hat{\rho}(\cdot)$ thus satisfies the (vector-valued) stochastic differential equation

$$d\hat{\rho}(t) = \partial F(\bar{\rho}(t)) \hat{\rho}(t) dt + LH \Sigma(\bar{\rho}(t)) dB(t),$$

where $B(t)$ is a length $(d+1)m$ vector of independent standard Brownian motions and where $\Sigma(\bar{\rho}(t))$ is the $(d+1)m \times (d+1)m$ diagonal matrix with the square roots of $Q(\bar{\rho}(t))$ on the diagonal. The corresponding mean vector and covariance matrix are defined as

$$M(t) := \mathbb{E}[\hat{\rho}(t)], \Gamma(s, t) := \text{cov}(\hat{\rho}(s), \hat{\rho}(t)) = \mathbb{E}[(\hat{\rho}(s) - M(s))(\hat{\rho}(t) - M(t))^T].$$

In addition, $V(t) := \text{Var}[\hat{\rho}(t)] = \Gamma(t, t)$. As in Karatzas and Shreve (2012, section 5.6, problems 6.1, 6.2), with $\bar{\Phi}(s) := \Phi(s, 0)$, these allow explicit expressions:

$$\begin{aligned} M(t) &= \bar{\Phi}(t) \left[M(0) + \int_0^t \bar{\Phi}^{-1}(s) ds \right], \\ \Gamma(t, s) &= \bar{\Phi}(s) \left[V(0) + \int_0^{t \wedge s} \bar{\Phi}^{-1}(u) LH \Sigma(\bar{\rho}(u)) (\bar{\Phi}^{-1}(u) LH \Sigma(\bar{\rho}(u)))^T du \right] \bar{\Phi}^T(t). \end{aligned} \quad (14)$$

Moreover, $M(t)$ and $V(t)$ solve the linear (matrix) differential equations

$$\begin{aligned} \frac{dM(t)}{dt} &= \partial F(\bar{\rho}(t)) M(t), \\ \frac{dV(t)}{dt} &= \partial F(\bar{\rho}(t)) V(t) + V(t) (\partial F(\bar{\rho}(t)))^T + LH \Sigma(\bar{\rho}(t)) (LH \Sigma(\bar{\rho}(t)))^T. \end{aligned} \quad (15)$$

At the beginning of this section, we expressed that our objective was to find an approximation for the distribution of $(\rho(t_1), \dots, \rho(t_k))$ for $0 \leq t_1 < t_2 < \dots < t_k < \infty$. With the fluid limit $\bar{\rho}(\cdot)$ and diffusion limit $\hat{\rho}(\cdot)$, we have succeeded in doing so. In Section 6, we present numerical examples that assess the accuracy of the resulting approximation.

Remark 2. It is known that the accuracy of the diffusion approximation degrades in the vicinity of nonsmooth points of the transition rate functions. The reason is that in such a setting the natural fluctuations of the process are such that it resides a nonnegligible fraction on time on both sides of the nonsmooth point, such that the local dynamics on both sides matter. For more background on this phenomenon, we refer to Mandelbaum et al. (1998) and Massey and Pender (2013).

5. Travel Times

$$\{T_{i,i+k,j}(t) > x\} = \{Y_{i+k,j}(t+x) < Y_{i-1,j}(t)\}, \quad x > 0. \quad (16)$$

By this identity, a sequence of probabilities $\mathbb{P}(T_{i,i+k,j}(t) > x_n)$, for $x_n > 0$ and $n \in \{1, \dots, N\}$, can be derived from the joint distribution of the random vectors $Y(t), Y(t+x_1), \dots, Y(t+x_N)$. In this section, we approximate the joint distribution of these random vectors with a Gaussian distribution, using fluid and diffusion limits for $Y(\cdot)$, similar to the ones we derived in Section 4, under the same scaling. We thus find an approximation of the distribution of $T_{i,i+k,j}(t)$, for $i \in \{1, \dots, d\}$, $k \in \{0, \dots, d-i+1\}$, and $j \in \{1, \dots, m\}$.

At first sight, it may look restrictive to require that type- j vehicles do not overtake each other within a cell. Noticing, however, that individual vehicles within a class are not systematically faster or slower than one another, it is anticipated that the approximation of the travel time distribution is reasonably accurate. Moreover, when overtaking behavior within a class of vehicles occurs often, our setup allows the class to be split into two classes, say, a fast and a slow one.

Initially, one would think that the limit results for $Y(\cdot)$ directly follow from the ones we derived for $\rho(\cdot)$ in Section 4; with the vector Y ordered lexicographically, we have

$$\rho(t) = LX(0) + LH Y(t), \quad (17)$$

in accordance with (9). The results for $\rho(\cdot)$, however, cannot be translated into results for $Y(\cdot)$, because of the easily verified fact the matrix H is singular. A simple illustration of this is that when $Y(t)$ satisfies (17), then so does $Y(t) + c$ for every $c \in \mathbb{R}$. As a consequence, we have to follow a different approach.

Before we present the fluid and diffusion limits for $Y(\cdot)$, we make two final observations. First, the diffusion approximation can be established along the lines of the proof of Kurtz (1981, theorem 8.2), but we choose to give a different, more concrete proof, which illustrates how the redefinition of the operator ∂F , and the corresponding weak solutions naturally follow from taking the limit. Second, the point processes $Y_{i-1,j}(\cdot)$ have intensity $q_{i-1,j}(\rho(t))$ at time t , which by (17) is also a function of $Y(t)$. Moreover, as L and H have bounded norm, this function is clearly still Lipschitz in $Y(t)$. To simplify the notation, from now on we simply say that each $Y_{i-1,j}(t)$ has intensity given by $h(Y(t))$ at time t , for a Lipschitz continuous function $h(\cdot)$. When we apply the scaling that we introduced in Section 4, we denote the intensity by $h^n(\cdot)$ to emphasize that the argument is reduced by a factor n .

We now present the first result, which is the counterpart of Theorem 1.

Proposition 1. Consider the sequence of scaled processes $Y^n(\cdot)$, where $Y^n(t) := \frac{1}{n}Y(nt)$. Under Assumption 1 and if $X^n(0) \rightarrow \bar{X}(0)$, we have, as $n \rightarrow \infty$, that

$$Y^n(\cdot) \xrightarrow{\text{a.s.}} \bar{Y}(\cdot), \quad \text{u.o.c.}$$

where $\bar{Y}(\cdot)$ satisfies

$$\bar{Y}(t) = \int_0^t h(\bar{Y}(s)) \, ds.$$

To prove the proposition, we can replicate the proof of Kurtz (1981, theorem 8.1). The conditions are met, as $h(\cdot)$ is Lipschitz in $Y(\cdot)$, and $h(\cdot)$ has compact support and is therefore uniformly bounded.

We immediately present the diffusion limit for $Y(\cdot)$, which is the counterpart of Theorem 2. Here ∂h is the weak derivative of h , defined similarly as ∂F in Section 4. Also, $\bar{\Sigma}(\bar{Y}(t))$ is the $(d+1)m \times (d+1)m$ diagonal matrix with the square roots of $h(\bar{Y}(t))$ on the diagonal.

Theorem 3. Consider the sequence of centered and scaled processes $\{\hat{Y}^n(\cdot)\}_n$, where $\hat{Y}^n(t) := \sqrt{n}(Y^n(t) - \bar{Y}(t))$. Under Assumption 1 and if $\sqrt{n}|X^n(0) - \bar{X}(0)| \rightarrow 0$, we have, as $n \rightarrow \infty$, that

$$\hat{Y}^n(\cdot) \Rightarrow \hat{Y}(\cdot),$$

where $\hat{Y}(\cdot)$ satisfies

$$d\hat{Y}(t) = \partial h(\bar{Y}(t))\hat{Y}(t)dt + \bar{\Sigma}(\bar{Y}(t))dB(t), \quad \hat{Y}(0) = 0,$$

with $B(\cdot)$ a $(d+1)m$ -dimensional vector of independent standard Brownian motions.

We begin by rewriting $\hat{Y}^n(\cdot)$ by adding and subtracting the compensator of $Y^n(\cdot)$,

$$\begin{aligned} \hat{Y}^n(t) &= n^{1/2} \left(Y^n(t) - \frac{1}{n} \int_0^{nt} h^n(Y(s))ds + \frac{1}{n} \int_0^{nt} h^n(Y(s))ds - \int_0^t h(\bar{Y}(s))ds \right) \\ &= n^{-1/2} M^n(t) + \int_0^t \Psi^n(s) \hat{Y}^n(s) ds, \end{aligned}$$

where we define the sequences of martingales $\{M^n(\cdot)\}_n$ and processes $\{\Psi^n(\cdot)\}_n$ by

$$M^n(t) := Y(nt) - \int_0^{nt} h^n(Y(s))ds, \quad \Psi_{kl}^n(t) := \frac{h_k(Y^n(s)) - h_k(\bar{Y}(s))}{Y_l^n(s) - \bar{Y}_l(s)},$$

where $k, l \in \{1, \dots, (d+1)m\}$. In differential notation, we thus have

$$d\hat{Y}^n(t) = \Psi^n(t)\hat{Y}^n(t)dt + n^{-1/2}dM^n(t),$$

from which we derive, with $\zeta^n(t) := \int_0^t \Psi^n(s)ds$ and $\hat{Z}^n(t) := \exp(-\zeta^n(t))\hat{Y}^n(t)$,

$$d\hat{Z}^n(t) = n^{-1/2} \exp(-\zeta^n(t))dM^n(t). \quad (18)$$

Our derivation consists of three steps.

- In the first step, we wish to show that $\exp(-\zeta^n(\cdot))$ converges almost surely, u.o.c. to $\int_0^\cdot \partial h(\bar{Y}(s))ds$. By Proposition 1, $Y^n(t) \xrightarrow{\text{a.s.}} \bar{Y}(t)$ u.o.c. As the function $h(\cdot)$ is almost everywhere differentiable, we can rewrite the integral in the definition of $\zeta^n(\cdot)$ up to the null set where the derivative does not exist. Combining these arguments, and using that ∂h is uniformly bounded, an application of the dominated convergence yields, for $t \geq 0$,

$$\zeta^n(t) \xrightarrow{\text{a.s.}} \zeta(t) := \int_0^t \partial h(\bar{Y}(s))ds. \quad (19)$$

By continuous mapping and Jacod and Shiryaev (2003, theorem VI.2.15) (bearing in mind that $x \mapsto -\exp(-x)$ is increasing and continuous), we have

$$\sup_{s \leq t} |\exp(-\zeta^n(t)) - \exp(-\zeta(t))| \xrightarrow{\text{a.s.}} 0.$$

- In the second step we show that the martingales $n^{-1/2}M^n(\cdot)$ converge weakly to a $(d+1)m$ -dimensional scaled Brownian motion. We want to apply the martingale functional central limit theorem (MFCLT) (Jacod and Shiryaev 2003, theorem VIII.3.22). To this end, we verify the corresponding conditions. For every n , $M^n(\cdot)$ is a compensated time-inhomogeneous Poisson process with uniformly bounded rate, implying that $n^{-1/2}M^n(\cdot)$ is locally square integrable, which is the first condition. Moreover, it is immediate that the jumps $n^{-1/2}(M^n(t) - \lim_{s \uparrow t} M^n(s)) \leq n^{-1/2} \rightarrow 0$, which by Jacod and Shiryaev (2003, equation (VIII.3.5)) is sufficient for the second condition.

Finally we need to verify that the normalized predictable quadratic covariation matrix, that is, $\langle n^{-1/2}M^n \rangle_t$, converges in probability to a deterministic limit. We may assume, without loss of generality, that $\langle M_k^n, M_l^n \rangle_t = 0$, because we can construct the processes $Y(\cdot)$ as independent unit-rate Poisson processes on a product space and

apply an absolutely continuous change of measure so that $Y(\cdot)$ has the required intensity (cf. Brémaud 1981, section VI.2). By construction and an application of Girsanov's theorem, we have $\langle M_k^n, M_l^n \rangle_t = 0$ almost surely. To conclude our argument, we have

$$\begin{aligned}\langle n^{-1/2}M^n \rangle_t &= n^{-1}\langle M^n \rangle_t = \text{diag}\left\{n^{-1} \int_0^{nt} h^n(Y(s))ds\right\} \\ &= \text{diag}\left\{\int_0^t h(Y^n(s))ds\right\} \xrightarrow{\text{a.s.}} \text{diag}\left\{\int_0^t h(\tilde{Y}(s))ds\right\},\end{aligned}$$

where the convergence is a consequence of Proposition 1. By the MFCLT, it now follows that $n^{-1/2}M^n(\cdot)$ converges weakly to a scaled Brownian motion $\tilde{B}(\cdot)$ with $\langle \tilde{B} \rangle_t = \text{diag}\{\int_0^t h(\tilde{Y}(s))ds\}$.

• For the third and final step, we combine previous results to conclude weak convergence of the stochastic integrals in Equation (18). We want to use Jacod and Shiryaev (2003, theorem VI.6.22) for which we need the so-called P-UT property, which follows from our second step and Jacod and Shiryaev (2003, proposition 6.13). We use Slutsky's lemma to obtain joint weak convergence of $\exp(-\zeta^n(\cdot))$ and $n^{-1/2}M^n(\cdot)$ and conclude that $Z^n(\cdot) = \int_0^\cdot n^{-1/2} \exp(-\zeta^n(u)) dM^n(u)$ converges weakly to $\int_0^\cdot \exp(-\zeta(u)) d\tilde{B}(u)$. Finally, by applying Slutsky again in combination with continuous mapping, we obtain the weak convergence of $\hat{Y}^n(\cdot)$ to a process $\hat{Y}(\cdot)$ defined through

$$d\hat{Y}(t) = \partial h(\tilde{Y}(t))\hat{Y}(t)dt + d\tilde{B}(t), \quad \hat{Y}(0) = 0. \quad (20)$$

The stated result follows directly from (20), using the definition of $\tilde{B}(\cdot)$.

We can now approximate the travel time distribution by combining (16) with Proposition 1 and Theorem 3. In the next section, this procedure will be illustrated through a series of examples.

6. Numerical Examples

Now that we have established fluid and diffusion limits for both $\rho(\cdot)$ and $Y(\cdot)$, we illustrate their usefulness by a series of numerical experiments. These reproduce traffic phenomena like *forward propagation*, *backward moving jams*, and *shockwave formation*. We start the section, however, by assessing the validity and accuracy of the approximation.

6.1. Generating Diffusion-Based Approximations

In our limit theorems, we scaled both the length of the cells and time by a factor n . However, this parameter n is artificial, in that in a real-world situation, such a scaling parameter, obviously does not have a physical meaning. In this section, we point out how our limiting results (for n large) can be converted into approximations (in which n should not appear). We then focus on accuracy of the resulting approximation by comparing it with simulation-based estimates.

6.1.1. Approximation; Role of Scaling Parameter n . We proceed by providing an approximation for the distribution of $\rho(\cdot)$, based on the diffusion limit that we established in Section 4. By an explicit calculation, we then show that this approximation is independent of the scaling parameter n , as it should.

In Theorems 1 and 2, we state limiting results for the processes $\rho^n(nt)$ and $\hat{\rho}^n(t) := \sqrt{n}(\rho^n(nt) - \bar{\rho}(t))$. These show that

$$\rho^n(n \cdot) \xrightarrow{\text{a.s.}} \bar{\rho}(\cdot) \quad \text{and} \quad \hat{\rho}^n(\cdot) \Rightarrow \hat{\rho}(\cdot).$$

Rewriting gives, for n large, an expression for $\rho^n(\cdot)$ in terms of a fluid limit $\bar{\rho}(\cdot)$ and a (zero-mean) Gaussian perturbation around it:

$$\rho^n(nt) \approx \bar{\rho}(t) + \frac{1}{\sqrt{n}}\hat{\rho}(t).$$

To obtain the prelimit process $\rho(\cdot)$ back on the left-hand side, we reverse the scaling, by slowing time down by a factor n and dividing cell lengths by ℓ_i/n . This gives the distributional approximation

$$\rho(t) \approx \bar{\rho}(t/n) + \frac{1}{\sqrt{n}}\hat{\rho}(t/n); \quad (21)$$

here, $\rho(\cdot)$ corresponds to cells of length ℓ_i , whereas the processes on the right-hand side correspond to cells of length ℓ_i/n . Now we can show that the right-hand side of (21) actually does not depend on n . For $i \in \{1, \dots, d\}$ and $j \in \{1, \dots, m\}$, we have

$$\begin{aligned}\bar{\rho}_{ij}(t/n) &= \bar{\rho}_{ij}(0) + \frac{n}{\ell_i} \int_0^{t/n} q_{i-1,j}(\bar{\rho}(s)) ds - \frac{n}{\ell_i} \int_0^{t/n} q_{i,j}(\bar{\rho}(s)) ds \\ &= \bar{\rho}_{ij}(0) + \frac{n}{\ell_i} \frac{1}{n} \int_0^t q_{i-1,j}\left(\bar{\rho}\left(\frac{s}{n}\right)\right) ds - \frac{n}{\ell_i} \frac{1}{n} \int_0^t q_{i,j}\left(\bar{\rho}\left(\frac{s}{n}\right)\right) ds.\end{aligned}$$

Now, setting $\bar{\rho}^\circ(t) := \bar{\rho}(\frac{t}{n})$, we get

$$\bar{\rho}^\circ(t) = \bar{\rho}^\circ(0) + \frac{1}{\ell_i} \int_0^t q_{i-1,j}(\bar{\rho}^\circ(s)) ds - \frac{1}{\ell_i} \int_0^t q_{i,j}(\bar{\rho}^\circ(s)) ds.$$

For $\hat{\rho}(\cdot)$, the calculation is similar. For $i \in \{1, \dots, d\}$ and $j \in \{1, \dots, m\}$,

$$\begin{aligned}\frac{1}{\sqrt{n}} \hat{\rho}_{ij}(t/n) &= \frac{1}{\sqrt{n}} \frac{n}{\ell_i} \int_0^{t/n} \nabla q_{i-1,j}(\bar{\rho}(s)) \cdot \hat{\rho}(s) ds - \frac{1}{\sqrt{n}} \frac{n}{\ell_i} \int_0^{t/n} \nabla q_{i,j}(\bar{\rho}(s)) \cdot \hat{\rho}(s) ds \\ &\quad + \frac{1}{\sqrt{n}} \frac{n}{\ell_i} \int_0^{t/n} \sqrt{q_{i-1,j}(\bar{\rho}(s))} dB_{i-1,j}(s) - \frac{1}{\sqrt{n}} \frac{n}{\ell_i} \int_0^{t/n} \sqrt{q_{i,j}(\bar{\rho}(s))} dB_{i,j}(s) \\ &= \frac{n}{\ell_i} \frac{1}{n} \int_0^t \nabla q_{i-1,j}\left(\bar{\rho}\left(\frac{s}{n}\right)\right) \cdot \frac{1}{\sqrt{n}} \hat{\rho}\left(\frac{s}{n}\right) ds - \frac{n}{\ell_i} \frac{1}{n} \int_0^t \nabla q_{i,j}\left(\bar{\rho}\left(\frac{s}{n}\right)\right) \cdot \frac{1}{\sqrt{n}} \hat{\rho}\left(\frac{s}{n}\right) ds \\ &\quad + \frac{1}{\sqrt{n}} \frac{n}{\ell_i} \frac{1}{\sqrt{n}} \int_0^t \sqrt{q_{i-1,j}\left(\bar{\rho}\left(\frac{s}{n}\right)\right)} dB_{i-1,j}(s) - \frac{1}{\sqrt{n}} \frac{n}{\ell_i} \frac{1}{\sqrt{n}} \int_0^t \sqrt{q_{i,j}\left(\bar{\rho}\left(\frac{s}{n}\right)\right)} dB_{i,j}(s),\end{aligned}$$

where we have used the scale-invariance of Brownian motion, that is, $B(s/n) \stackrel{d}{=} B(s)/\sqrt{n}$. By setting $\hat{\rho}^\circ(t) := \hat{\rho}(t/n)/\sqrt{n}$, we get

$$\begin{aligned}\hat{\rho}^\circ(t) &= \frac{1}{\ell_i} \int_0^t \nabla q_{i-1,j}(\bar{\rho}^\circ(s)) \cdot \hat{\rho}^\circ(s) ds - \frac{1}{\ell_i} \int_0^t \nabla q_{i,j}(\bar{\rho}^\circ(s)) \cdot \hat{\rho}^\circ(s) ds \\ &\quad + \frac{1}{\ell_i} \int_0^t \sqrt{q_{i-1,j}(\bar{\rho}^\circ(s))} dB_{i-1,j}(s) - \frac{1}{\ell_i} \int_0^t \sqrt{q_{i,j}(\bar{\rho}^\circ(s))} dB_{i,j}(s).\end{aligned}$$

In conclusion, we end up with the distributional approximation

$$\rho(\cdot) \approx \bar{\rho}^\circ(\cdot) + \hat{\rho}^\circ(\cdot), \quad (22)$$

which is independent of n . More specifically, we have that $\rho(\cdot)$ is approximately a Gaussian process, where

$$\mathbb{E}[\rho(t)] \approx \bar{\rho}^\circ(t), \text{Var}[\rho(t)] \approx \text{Var}[\hat{\rho}^\circ(t)], \quad (23)$$

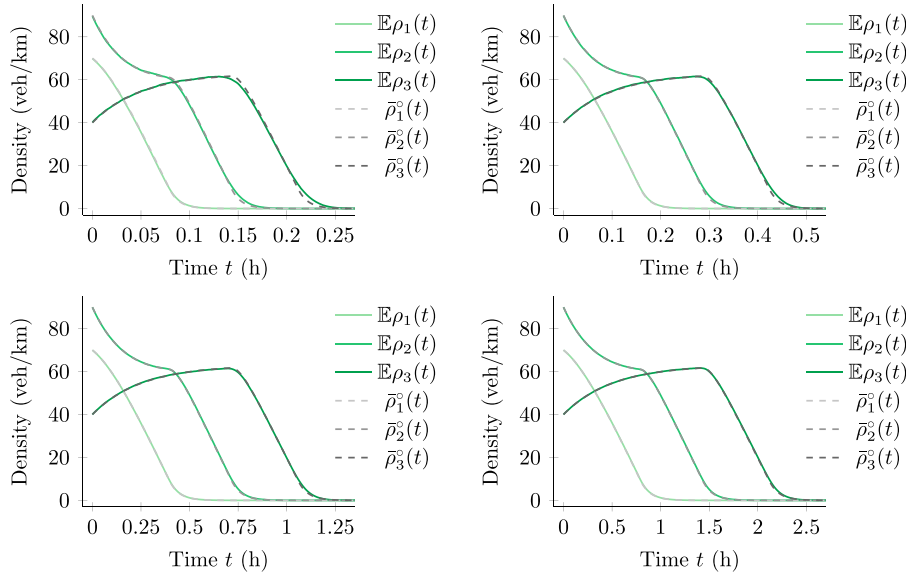
with $\bar{\rho}^\circ(\cdot)$ being deterministic and $\hat{\rho}^\circ(\cdot)$ a zero-mean Gaussian process.

6.1.2. Validity and Accuracy of the Approximation. We proceed by validating the approximation (22) by simulation. We use the Daganzo MFD (cf. Example 1 in Section 3.3) with parameters $v^f = 100$ km/h, $w = 20$ km/h, $q^{\max} = 1,800$ veh/h, $\rho^{\text{jam}} = 105$ veh/km, $\lambda = 0$ veh/h, and $\nu = 900$ veh/h. We consider a segment with $d = 3$ cells, each of them having length $\ell_i = \ell$, where $\ell \in \{1, 2, 5, 10\}$ km, to illustrate that the accuracy of the approximation improves when ℓ grows.

For the simulation, we consider a scenario that is initialized with $\rho(0) = (70, 90, 40)^\top$ and runs on the time-interval $[0, 1,000]$ in seconds. We estimate the mean and standard deviation of $\rho(k\Delta t)$, for $\Delta t = \ell$ and $k = 0, 1, \dots, 1,000$, using 1,000 simulated samples. We compare this with the approximated mean and standard deviation. We rely on (23), where we use (15) to evaluate the variance.

The comparison of the means is given in Figure 2, whereas the comparison for the standard deviations can be found in Figure 3. Both figures show that the difference between the curves becomes smaller when ℓ increases.

Figure 2. Comparison of $\rho_i^\circ(t)$ and $\mathbb{E}\rho_i(t)$



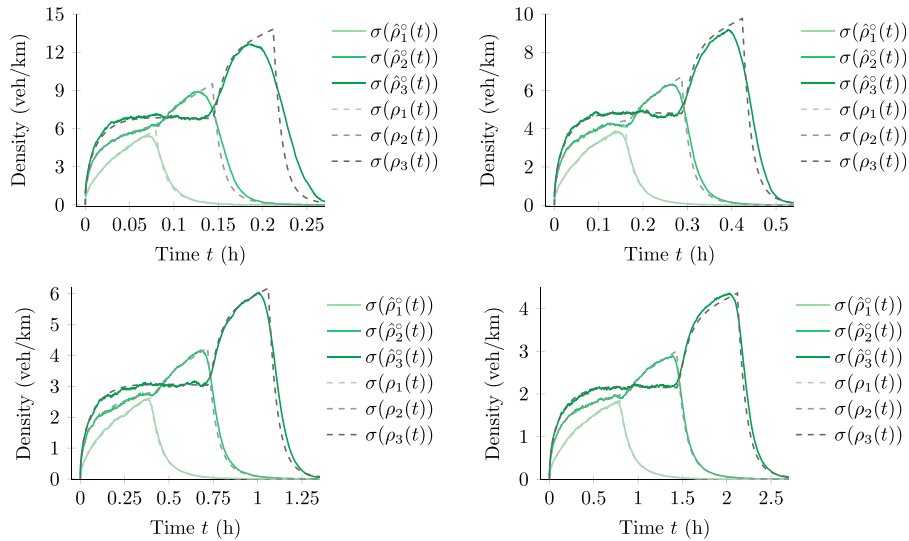
Note. The latter was estimated using simulation, for $i = 1, 2, 3$, and $\ell = 1$ (top left), $\ell = 2$ (top right), $\ell = 5$ (bottom left), and $\ell = 10$ (bottom right).

6.2. Numerical Experiments

In the remainder of this section, we illustrate the use of the limit theorems through a series of numerical experiments. Traffic researchers have tried to replicate various phenomena in traffic and did so quite well with deterministic kinematic wave models, thus recovering the fluid limit. In this section, we illustrate the usefulness of our results, in that, in addition to replicating these phenomena with the fluid limit, we are also capable of providing an approximation of the associated vehicle density variance and covariance (where we have covariances in both the spatial and temporal sense). In addition, we demonstrate our procedure to approximate travel time distributions.

For our experiments, we take the multiclass MFD from Chanut and Buisson (2003) (corresponding with Example 2 in Section 3.3), with parameters that are used in the numerical experiments of the original paper, with a kilometer taken as unit length. We have $v_1^f = 108$ km/h, $v_2^f = 79.2$ km/h, $v^c = 61.2$ km/h, $L_1 = 0.0065$ km, $L_2 = 0.0165$ km, $N = 3$, and $\beta = 0.25$. We refer to class 1 vehicles as *cars* and to the class 2 vehicles as *trucks*.

Figure 3. Comparison of $\sigma(\hat{\rho}_i^\circ(t)) := \sqrt{\text{Var}[\hat{\rho}_i^\circ(t)]}$ and $\sigma(\rho_i(t)) := \sqrt{\text{Var}[\rho_i(t)]}$



Note. The latter was estimated using simulation, for $i = 1, 2, 3$ and $\ell = 1$ (top left), $\ell = 2$ (top right), $\ell = 5$ (bottom left), and $\ell = 10$ (bottom right).

To obtain the Gaussian approximation of the mean and variance, we jointly solve the differential form of (10) and (15) using a numerical solver. To obtain the matrix ∂Q , one takes, for each regular point of Q , the derivative of the discrete flux functions. In every nonregular point, one assigns the maximum velocity of a vehicle to the components that are not differentiable, so that ∂Q is bounded by a Lipschitz constant. Additionally, to obtain the travel time distribution, we use the analog of (14) for $\hat{Y}$, so that we can evaluate the tail probabilities in (16) based on the joint normal approximation of $Y_{i+k,j}(t+x)$ and $Y_{i-1,j}(t)$.

Example 3 (Forward Propagation). In the first experiment, we use the results of Theorem 1 and Theorem 2 to study forward propagation. In the experiment, the first five cells initially have a density of 60 veh/km of which 20% are trucks, and we let the system evolve over a time interval of 2,000 seconds, with $\Delta t = 2$ seconds. The experiment is taken from Chanut and Buisson (2003, figure 3), where $d = 100$, but with $\ell_i = 600$ m for each i . Under these parameters, the Godunov scheme is stable (cf. LeVeque 1992, section 10.6 and equation (13.11)).

Figure 4 shows the results of our experiment, with from left to right, the total vehicle densities, the car densities, and the truck densities, where the dotted lines show the corresponding 95% confidence intervals. The trucks propagate slower on average than the cars, and they dissipate slower from the initial density; they are more clumped.

Example 4 (Travel Times for the Setting of Example 3). To illustrate our approximation of the travel time distribution, we consider the time that it takes a vehicle in cell 10 at time $t = 200$ seconds, to reach cell 50. To be precise, we evaluate (16) with Y replaced by $\hat{Y}$, $t = 200$ seconds, $i = 10$, $k = 39$, and $x \in \{0, 2, 4, \dots, 2,000\}$, in the setting of the previous forward-propagation experiment.

First, in Figure 5, we plotted the mean evolution of $\hat{Y}$ with its associated 95% confidence interval, at different points in time. The vertical gray lines are associated with the position of cell 10 and cell 50, whereas the horizontal gray lines correspond with the peak of the car and truck density in cell 10 at time $t = 200$ s. From panel (b), we see that at time $t = 200$ seconds, $\hat{Y}_{10,1} \approx 93$, and $\hat{Y}_{10,2} \approx 17$. From panels (c)–(f), we observe that the cars arrive at cell 50 (roughly) between $t = 1,050$ seconds and $t = 1,150$ seconds and the trucks (roughly) between $t = 1,300$ seconds and $t = 1,400$ seconds.

In Figure 6, we plotted both the approximated distribution $\hat{F}_{T_j}(\cdot)$ and density $\hat{f}_{T_j}(\cdot)$ of the travel time of type- j vehicles ($j = 1, 2$), as derived from $\hat{Y}$ in Figure 5. We remark that one can empirically show that the travel time distribution for cars is close to normal. However, regarding the travel time distribution for trucks, a QQ-plot reveals a distribution with significantly lighter tails. In general, from (16), we observe that the asymptotic normality of the counting processes $Y_{i,j}(\cdot)$ does not imply asymptotic normality of the travel times.

Example 5 (Backward Moving Jam). This experiment illustrates our results in the case of a backward moving traffic jam. This phenomenon in traffic is seen when there is a large density on a given road segment, with the front dissipating and the back absorbing arriving cars. The jam thus seems to be moving backward on the road; detailed background on moving jams can be found in, for example, Kerner (2004).

In our numerical experiments, we set $d = 20$, with $\ell_i = 1$ km for every i , with initial density of 300 veh/km in cells 8 up to (and including) 12, and an initial density of 88 veh/km in the other cells. Moreover, we set $\lambda_1 = 4,800$ veh/h, $\lambda_2 = 960$ veh/h, and $\nu = 1.2$, and we simulate for 1,500 seconds with $\Delta t = 1$ second.

Figure 7 shows the solution of the differential equations for $\bar{\rho}^\circ(\cdot)$, summed over the vehicle types, with the error bars indicating the standard errors corresponding to each cell; the i th panel corresponds with the

Figure 4. Solution Differential Equations for $\rho(t)$ and $\text{Var}[\hat{\rho}(t)]$ for Sum of Densities (Left), Car Densities (Center), and Truck Densities (Right): Example 3—Forward Propagation

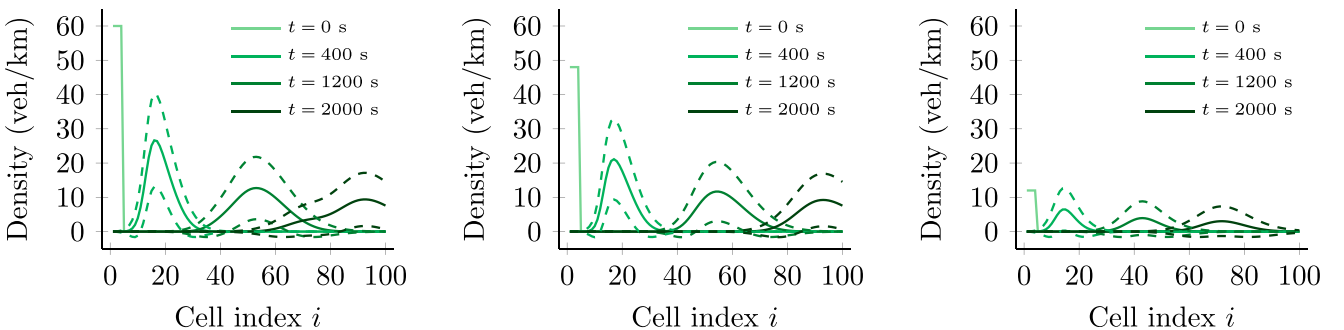
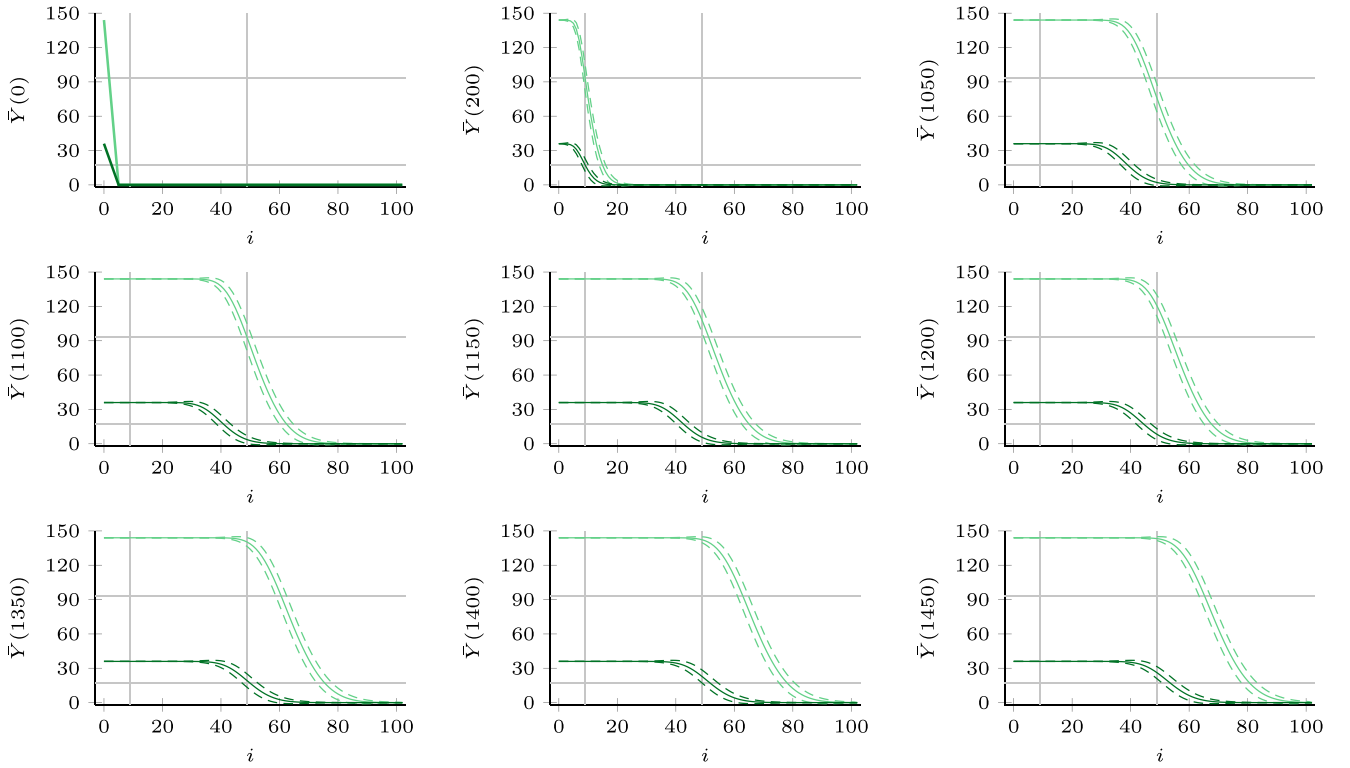


Figure 5. Evolution of $\bar{Y}$



Note. Dashed line, 95% confidence interval, given by $\text{Var}[\hat{Y}(t)]$; light green line, $\bar{Y}_{i,1}(t)$; dark green line, $\bar{Y}_{i,2}(t)$ for $t \in \{0, 200, 1,050, 1,100, 1,150, 1,200, 1,350, 1,400, 1,450\}$.

situation after $300(i-1)$ seconds, for $i \in \{1, \dots, 6\}$. We conclude from the pictures that the approximation succeeds in replicating the phenomenon of the jam moving backward.

It is also observed that the vehicle densities in the cells in front of the traffic jam have a particularly large variability. This is because vehicles approach the jam relatively fast, resulting in a high in-flow rate for the cells in front of the jam. This behavior is consistent with a large scatter in the MFD at the onset of congestion. The cells behind the jam (with index between 12 and 20, that is) also have increased in-flow (and later out-flow), because of the dissipation of the traffic jam. As a consequence, the variability (in terms of the standard error) of the vehicle density increases, until the flow stabilizes, from which point on the standard error goes down again.

Example 6 (Shocks). The final experiment illustrates how our results replicate shock waves, an important topic in the theory of conservation laws (LeVeque 1992, section 7). In this case, a fast small density reaches a larger and slower downstream density. In traffic flow theory, such shocks have been studied since Richards (1956).

In our experiment, we take $d = 40$ with $\ell_i = 0.6$ km for each i . We simulate for 500 seconds with $\Delta t = 1$ second. Initially, cells 20 up to 24 have a density of 350 veh/km, and cells 5 up to 9 have a density of

Figure 6. Approximated Cumulative Distribution Function (Where $\hat{F}_j(x) := \mathbb{P}(T_{10,49,j}(200) < x)$) and Corresponding Density ($\hat{f}_j(x)$) of Travel Times for Types $j = 1, 2$, from Cell 10 to Cell 50, for Cars Leaving Cell 10 at Time 200 Seconds

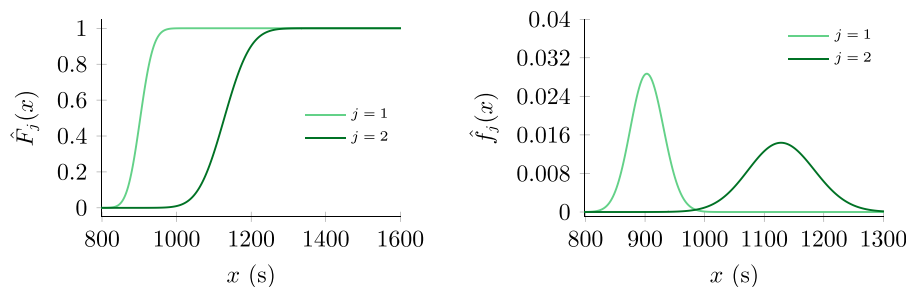
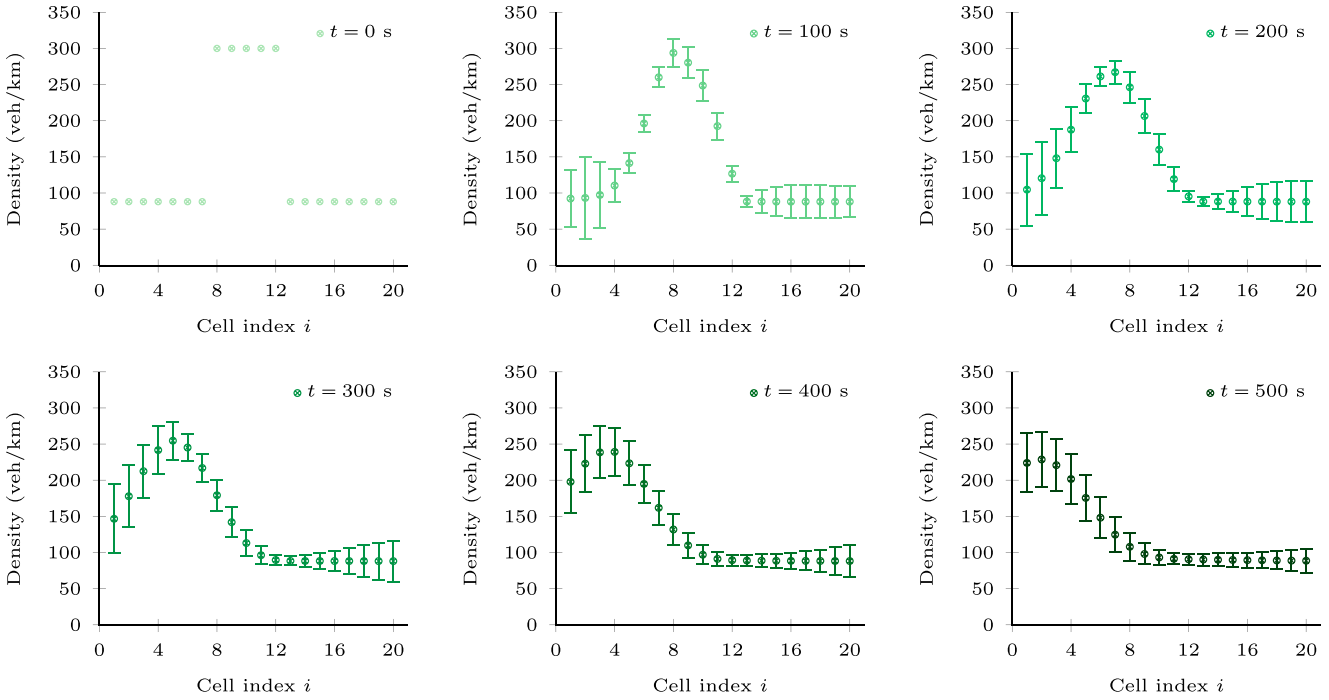
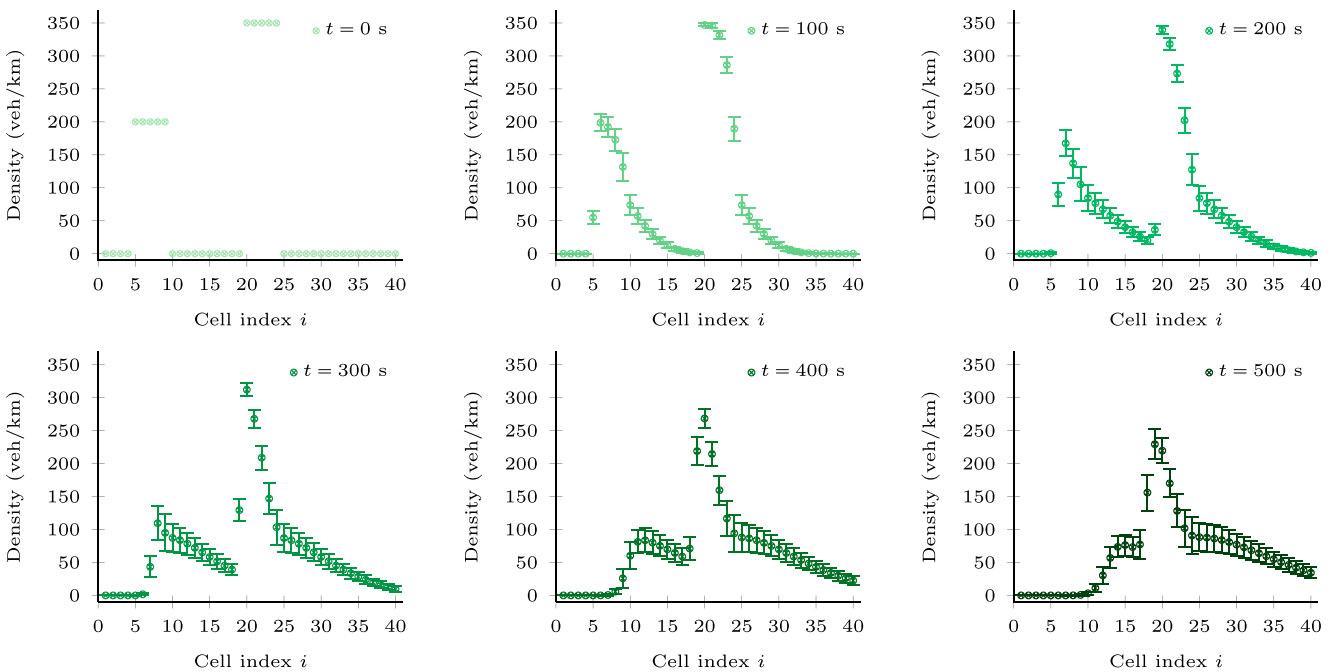


Figure 7. Solution Differential Equations for $\rho(t)$ and $\text{Var}[\hat{\rho}(t)]$: Example 5—Backward Moving Jam

200 veh/km, with every other cell being empty initially. Figure 8 depicts the evolution of the densities. The small density catches up with the large density, and one can see how the densities merge at the back of the jamming density, whereas the front of the jamming density dissipates.

Remark 3. Importantly, it is now straightforward to apply our results in a control context. For example, suppose that one wants to prevent congestion in cell i . By lowering the maximum speed in cells $1, \dots, i-1$, we alter the shape of the MFD in these cells. With our methods, we can assess to what extent this decreases the vehicle density in cell i . Moreover, the corresponding variance allows us to determine the maximum velocity to be imposed in cells

Figure 8. Solution Differential Equations for $\rho(t)$ and $\text{Var}[\hat{\rho}(t)]$: Example 6—Shocks

$1, \dots, i-1$ to make sure that the vehicle density of cell i remains below some predefined acceptable level with probability of at most $\alpha \in (0, 1)$.

7. Extensions, Generalizations, and Ramifications

In this section, we discuss a number of possible and relevant extensions, generalizations, and ramifications. To make the exposition as clear as possible, to this point, we worked with the basic model of a single road segment and a simple form for the infinitesimal generator of the underlying Markov chain. As we will make clear, however, a considerably broader class of systems can be dealt with analogously.

7.1. Networks

In the setup presented, we restricted ourselves to a single segment consisting of cells, each of them endowed with its own discrete flux function. This flux function, say from cell i to cell $i+1$, was assumed to depend on the vehicle densities in the sending cell i , as well as the destination cell $i+1$. On inspecting the framework in Kurtz (1981, chapter 8), however, one sees that it is allowed to consider flux functions that depend on the vehicle densities in *all* cells in the segment. This observation makes it straightforward to generalize our framework to road segments, each consisting of a number of cells, that are embedded in a network (with nodes that connect the first and last cells of the segments). By extending Assumption 1, so that each $\tilde{q}_i$ is a function of the densities in an arbitrary set of cells, the fluid limit and diffusion limit results carry over, as well as the results on the travel times.

As a special case of the network structure discussed previously, on-ramps and off-ramps (to and from our road segment) can be covered. Considering a segment consisting of d cells, if a ramp has to be placed adjacent to cell i , then this can be done by adding an extra cell, outside the segment, next to cell i . Depending on the direction of flow in the new cell, traffic will either merge from cell $i-1$ and the new cell into cell i , or it will diverge from cell $i-1$ into cell i and the new cell, thus encoding whether it concerns an on-ramp or an off-ramp.

Of course, it is not immediately clear that, for a general network and an arbitrary MFD, there exists a unique solution of the corresponding Riemann problem (now on a network), that in addition satisfies the regularity properties of Assumption 1. In Daganzo (1995a, section 2.3), it is explained how an arbitrary network can be captured by a network that only connects two upstream cells to one downstream cell, or vice versa; this reduction property simplifies the problem of finding a discrete flux function considerably. In particular, it features traffic merging from two cells into one or traffic diverging from one into two cells (see Daganzo 1995a for a schematic representation), making it particularly suitable for modelling on-ramps and off-ramps on a segment. In the remainder of Daganzo (1995a), the discrete flux function for the Daganzo MFD (i.e., our first example in Section 3.3) is found. Moreover, under this simplification of the network structure, one can solve the single-class case under fairly general conditions on the MFD (cf. Garavello and Piccoli 2006, chapters 4 and 5). For the multiclass case, to the best of the authors' knowledge, no general results exist, but there are solutions for specific MFDs (see Ngoduy 2010 and references therein). Likewise, for other proposed multiclass MFDs, which are often functions with strong regularity properties, one would envisage that solutions to the network Riemann problem can be found.

The generality of our model and the flexibility of the framework of Kurtz (1981, chapter 8) make it possible to include behavior of vehicles that is specific to merging and diverging situations. More specifically, because discrete flux functions in our model can be cell dependent, we can model merging and diverging behavior of vehicles (or even more complicated flow structures), using special merge and diverge flux functions. As long as these functions satisfy the regularity conditions in Assumption 1, the results from Sections 4 and 5 go through.

Finally, we can also explicitly model the routing of the vehicles through the network. In traffic flow theory, there are two major ways of routing vehicles on networks. The first is to route deterministically. This can be done by working with origin-destination pairs, where every vehicle has a fixed route from the origin to the destination (Garavello and Piccoli 2006, chapter 7). In our setup, we can incorporate this mechanism by treating vehicles using a specific route as a type, with suitable boundary conditions (describing the arrivals at the origin and departures at the destination). The second approach is to route probabilistically (Daganzo 1995a, Garavello and Piccoli 2006, section 5); think of the turning rates of vehicles at junctions. In the case where a number of flows merges into one flow at a junction and when the system is saturated at the point of merging, a probability distribution is used to describe the contribution of the merged flows to the total flow

out of the junction. This probability distribution controls the rate at which vehicles from each of the merging cells yield to vehicles from the other cells.

7.2. Generalized Transition Rates

In practice, the model's stochastic dynamics need not be homogeneous in time. One would, for instance, wish to accommodate a daily pattern with rush hours and more quiet periods. This can be realized by making the arrival rate (to cell 1) and departure rates (from cell d) time dependent. To preserve the results in Sections 4 and 5, we can scale the rates of $Y(\cdot)$ by a factor n (instead of scaling time), besides the scaling of the cell lengths; in the setup we studied in the previous sections, in which the Poisson processes do not depend on time, scaling time and scaling the rates of $Y(\cdot)$ are equivalent. To preserve the time-dependent arrival pattern, however, we let the scaled process counting the number of type- j arrivals to cell 1 be given by

$$\frac{1}{n\ell_i} Y_{0,j} \left(n \int_0^t \lambda_j(s) \wedge \sup_{x \in \mathbb{R}_+^k} \tilde{q}_{0,j}(x, (\rho_{1k}^n(s))_k) ds \right), \quad j \in \{1, \dots, m\},$$

where $\lambda_j(s)$ is the intensity of the type- j Poisson arrival process at time $s \geq 0$. Clearly, $\lambda_j(s)$ does not depend on $\rho^n(t)$, so that the minimum of these functions is still Lipschitz in $\rho^n(t)$ by Assumption 1.

Additionally, we can let the discrete flux function from Assumption 1 be stochastic itself, as was suggested in Qu et al. (2017). As long as the fluid limit of the associated Poisson processes is deterministic and continuous, the results from Sections 4 and 5 remain valid, albeit with a different limit. One typically gets that in the fluid limit all quantities are replaced by their time-average counterpart, whereas in the diffusion limit one obtains a larger variance, as a consequence of the uncertainty that has been added to the model. One possible way to obtain a stochastic MFD would be to periodically resample the MFD's parameters. Another option is to let the model's transition rates depend on exogenous influences (for instance the weather). This could be achieved by introducing an external (e.g., Markovian) background process, where the state of this process determines the current value of the transition rates. We refer to, for example, Jansen et al. (2019) for a derivation of fluid and diffusion limits in a related type of networks (viz. a network of Markov-modulated infinite-server queues). In other contexts, a similar methodology has been used, in, for example, Huang et al. (2016) and Spreij and Storm (2019).

Finally, we can relax of the assumption of the cell transition times being exponentially distributed to being phase type distributed. This is a useful extension, as any nonnegative distribution can be approximated arbitrarily closely by an appropriately chosen phase type distribution. It means that the time until a vehicle jumps to the next cell consists of multiple phases, where the time spent in each of the phases should be exponentially distributed. The transition rates should depend on the number of vehicles in each phase, both in the sending cell and the receiving cell, aggregated over the phases (to reflect the per-cell vehicle density). Importantly, such an extension still fits the framework developed in Kurtz (1981), so that the results from Sections 4 and 5 carry over.

8. Concluding Remarks

In this paper, we developed a Markovian model for vehicle densities in a multiple-type road traffic network, with dynamics that are consistent with state-of-the-art traffic-flow models. Because this model does not allow any explicit analysis, we resorted to an asymptotic framework. More concretely, we established a fluid and diffusion limit: scaling the lengths of the cells and time by n and appropriately centering and normalizing the vehicle densities, the resulting process converges, as $n \rightarrow \infty$, to a Gaussian process. This diffusion limit can be used to produce an approximation for the vehicle density distribution in time. Along the same lines, an approximation for the travel time distribution was developed. In a set of numerical experiments, we concluded that the resulting approximations are highly accurate; in addition, we showed that our model is capable of reproducing various known traffic phenomena.

Acknowledgments

The authors thank Dieter Fiems (University of Ghent) for helpful suggestions regarding the travel time analysis and Sandjai Bhulai (Vrije Universiteit Amsterdam) for providing useful feedback.

References

- Benzoni-Gavage S, Colombo R (2003) An n -populations model for traffic flow. *Eur. J. Appl. Math.* 14(5):587–612.
- Brémaud P (1981) *Point Processes and Queues*. Springer Series in Statistics (Springer, New York).

- Chanut S, Buisson C (2003) Macroscopic model and its numerical solution for two-flow mixed traffic with different speeds and lengths. *Transportation Res. Record* 1852(1):209–219.
- Daganzo C (1994) The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Res. Part B: Methodological* 28(4):269–287.
- Daganzo C (1995a) The cell transmission model, part II: Network traffic. *Transportation Res. Part B: Methodological* 29(2):79–93.
- Daganzo C (1995b) Requiem for second-order fluid approximations of traffic flow. *Transportation Res. Part B: Methodological* 29(4):277–286.
- Drake J, Schofer J, May A (1967) A statistical analysis of speed-density hypotheses. *Highway Res. Record* 154:53–87.
- Evans L (2010) *Partial Differential Equations* (American Mathematical Society, Providence, RI).
- Garavello M, Piccoli B (2006) *Traffic Flow on Networks*, vol. 1 (American Institute of Mathematical Sciences on Applied Mathematics, Springfield, MO).
- Greenshields B (1934) The photographic method of studying traffic behavior. *Proc. 13th Annu. Meeting of the Highway Research Board*, 382–399.
- Hale J (1969) *Ordinary Differential Equations*. Pure and Applied Mathematics: A Series of Texts and Monographs (Wiley-Interscience, New York).
- Huang G, Jansen H, Mandjes M, Spreij P, De Turck K (2016) Markov-modulated Ornstein-Uhlenbeck processes. *Adv. Appl. Probabilities* 48(1):235–254.
- Jabari S, Liu H (2012) A stochastic model of traffic flow: Theoretical foundations. *Transportation Res. Part B: Methodological* 46(1):156–174.
- Jabari S, Liu H (2013) A stochastic model of traffic flow: Gaussian approximation and estimation. *Transportation Res. Part B: Methodological* 47(1):15–41.
- Jacod J, Shiryaev A (2003) *Limit Theorems for Stochastic Processes*, vol. 288 (Springer Science & Business Media, New York).
- Jansen H, Mandjes M, de Turck K, Wittevrongel S (2019) Diffusion limits for networks of Markov-modulated infinite-server queues. *Performance Evaluation* 135:102039.
- Karatzas I, Shreve S (2012) *Brownian Motion and Stochastic Calculus*, vol. 113 (Springer Science & Business Media, New York).
- Kerner B (2004) *The Physics of Traffic: Empirical Freeway Pattern Features, Engineering Applications, and Theory* (Springer, New York).
- Kurtz T (1981) *Approximation of Population Processes*, vol. 36 (SIAM, Philadelphia).
- LeVeque R (1992) *Numerical Methods for Conservation Laws*, vol. 132 (Springer, New York).
- Lighthill M, Whitham G (1955) On kinematic waves. I: Flood movement in long rivers. II: A theory of traffic flow on long crowded roads. *Proc. Royal Soc.* 229A(1178):281–345.
- Logghe S, Immers L (2008) Multi-class kinematic wave theory of traffic flow. *Transportation Res. Part B: Methodological* 42:523–541.
- Maerivoet S, De Moor B (2005) Traffic flow theory. Preprint, submitted July 15, 2005, <https://arxiv.org/abs/physics/0507126>.
- Mandelbaum A, Massey W, Reiman M (1998) Strong approximations for Markovian service networks. *Queueing Systems* 30(1-2):149–201.
- Massey W, Pender J (2013) Gaussian skewness approximation for dynamic rate multi-server queues with abandonment. *Queueing Systems* 75(2-4):243–277.
- Ngoduy D (2010) Multiclass first-order modelling of traffic networks using discontinuous flow-density relationships. *Transportmetrica* 6(2):121–141.
- Ngoduy D, Liu R (2007) Multiclass first-order simulation model to explain non-linear traffic phenomena. *Physica A: Statistical Mechanics and Its Applications* 385:667–682.
- Qu X, Zhang J, Wang S (2017) On the stochastic fundamental diagram for freeway traffic: Model development, analytical properties, validation, and extensive applications. *Transportation Res. Part B: Methodological* 104:256–271.
- Richards P (1956) Shock waves on the highway. *Oper. Res.* 4(1):42–51.
- Smulders S (1990) Control of freeway traffic flow by variable speed signs. *Transportation Res. Part B: Methodological* 24(2):111–132.
- Spreij P, Storm PJ (2019) Diffusion limits for a Markov modulated binomial counting process. *Probability Engrg. Inform. Sci.* 34(2):235–257.
- van Wageningen-Kessels F, Van Lint H, Vuik K, Hoogendoorn S (2015) Genealogy of traffic flow models. *EURO J. Transportation Logistics* 4(4):445–473.
- Wong G, Wong S (2002) A multi-class traffic flow model—An extension of lwr model with heterogeneous drivers. *Transportation Res. Part A Policy Practice* 36(9):827–841.
- Zhang P, Liu RX, Wong S, Dai SQ (2006) Hyperbolicity and kinematic waves of a class of multi-population partial differential equations. *Eur. J. Appl. Math.* 17(2):171–200.