University of Amsterdam

Masters Thesis

An Agent-Based Gap-Filling Model for the Simulation of Heterogeneous Traffic Flow in Developing Countries

Author:
Simon Alger

Supervisor:
Prof. Michael H Lees

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in Computational Science

in the

Section Computational Science
Informatics Institute

August 2015

Declaration of Authorship

I, Simon Alger, declare that this thesis, entitled 'An Agent-Based Gap-Filling Model for the Simulation of Heterogeneous Traffic Flow in Developing Countries' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at the University of Amsterdam.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date: 1 August 2015

"What I cannot create, I do not understand."

Richard P. Feynman

UNIVERSITY OF AMSTERDAM

Abstract

Faculty of Science
Informatics Institute

Master of Science in Computational Science

An Agent-Based Gap-Filling Model for the Simulation of Heterogeneous Traffic Flow in Developing Countries

by Simon Alger

The majority of the world's population lives in developing countries, where traffic often consists of many different types of vehicles and is characterised by a disordered driving style. By contrast, traffic flow in most of the developed world consists of a small number of vehicle types and an orderly, rule-based driving style prevails. Relatively little research has been done to model developing world traffic. Most of the limited work on the subject has adapted lane-based models used for traffic in the developed world, and ignored the question of why lane discipline has become dominant in developing world traffic. The aim of this thesis is therefore to investigate the relationship between vehicle heterogeneity, lane discipline and traffic efficiency by means of a simulation study.

Due to the lack of suitable existing models, an agent-based model for disordered heterogeneous traffic flow is proposed. Based on the principle of gap-filling found in crowd dynamics models, vehicle agents move along the roadway by maximising speed in their chosen direction while simultaneously avoiding collisions. The model is shown to be capable of reproducing the characteristic phenomena of developing world traffic, such as unstructured lane changes and accumulation of smaller vehicles at the front of queues.

By modifying the model to handle lane-based flow, it was found that lane discipline neither increases nor decreases traffic efficiency for homogeneous streams, and significantly decreases traffic efficiency in highly heterogeneous streams. These results support the expectation that disordered driving may have emerged in developing world traffic because of the higher efficiency it allows when the traffic stream consists of a heterogeneous mix of vehicles of different sizes and speeds. Having demonstrated the high fidelity of the gap-filling microscopic model, future work could focus on integrating it into a city-wide simulation which can be used to choose between potential infrastructure designs, estimate level of service, and determine the effectiveness of traffic controls.

Acknowledgements

I would like to thank the following people for their assistance with this project.

My supervisor, Michael Lees, without whose thoughtful advice, constructive criticism and enthusiasm this project would not exist in its current form.

My mother, Janine Alger, for her encouragement, patience and unwavering support.

Finally, I would like to thank the Erasmus Mundus *ema2sa* programme for providing me with the financial support to pursue my Masters degree.

Contents

List of Figures

List of Tables

List of Algorithms

Abbreviations

ANVS	**A**verage **N**ormalised **V**ehicle **S**peed
ARS	**A**uto-**r**ick**s**haw
LCV	**L**ight **C**ommercial **V**ehicle
MTW	**M**otorised **T**wo-**W**heeler
PHMM	**P**arallel **H**ybrid **M**esoscopic **M**odel
TTC	**T**ime **T**o **C**ollision
VVP	**V**elocity **V**ector **P**oint

Chapter 1

Introduction

The rise of widespread motorised road transport has totally altered traditional ways of life in nearly every country over the course of the twentieth century. It has affected the way members of society go about their daily lives, especially the means by which they access employment, goods and services. It has resulted in consequences unforeseen in both their nature and magnitude, such as the proliferation of suburbia due to fast and affordable commuting, substantial rises in the rate of accidental death, and harm to the environment through pollution and construction of infrastructure. The economic benefits are clear: by substantially reducing the cost and time of transporting goods and people, businesses now have access to a larger number of employees, consumers enjoy a wider array of merchants, and workers are more affordably able to reach employers.

Motorised vehicle transport, however, has not been evenly distributed among the nations of the world. Industrialised countries have much higher rates of vehicle ownership and the accompanying economic and social opportunities than developing nations. The composition and nature of traffic in these two groups of countries have evolved to be markedly different due to their varying economic, social, environmental and cultural milieux. While developed nations typically feature a highly homogeneous mix of vehicles which move in a disciplined fashion according to well-defined rules, traffic in many parts of the developing world consists of a highly heterogeneous vehicle mix, with chaotic traffic flow in which drivers exercise relatively little discipline. The reasons behind why these two distinct modes of traffic exist is largely unconsidered in the literature. Addressing this deficiency is the goal of the present work.

Although a great deal of research has focussed on understanding and modelling traffic flow in the context of the developed world, there exists a relative dearth of research on traffic in the developing world. That research which does attempt to model conditions prevalent in developing nations has been undertaken by modification of models

developed for the industrialised context, leaving no models suitable for studying the relationship between composition of modalities and driving discipline. This work makes two contributions to the limited literature on developing world traffic: to present and evaluate a new model for disordered traffic flow, and to use this model to gain insight into the two disparate modes into which global traffic has evolved.

* * *

As can be seen in Figure 1.1, there is a strong positive correlation between per capita household consumption and car ownership. High rates of car use are thus largely the preserve of rich, developed countries. As has been mentioned, the prevailing distribution of vehicle types found on a region's roads will be influenced by various factors. In many urban parts of Europe, where roads were built to be narrow before the advent of motorised transport, cars are generally smaller and comprise a lower fraction of the traffic mix. By contrast, in many parts of the USA and Australia where the majority of infrastructure development has occurred since automobiles became widespread, roads are wider. A consequence of this is that car use is higher and the cars used are generally larger: sport utility vehicles (SUVs) and pick-up trucks are substantially more common.

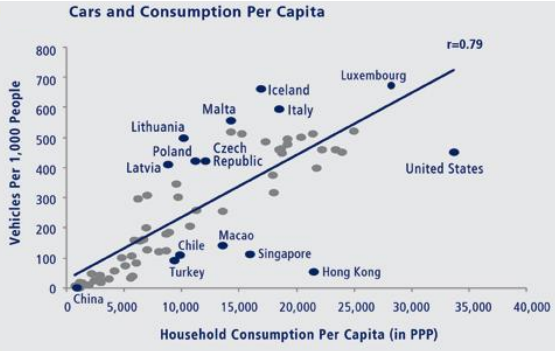

FIGURE 1.1: Per capita car ownership and household consumption. Source: World Bank

Urban planning also affects, and is affected by, the mix of vehicles in use in a particular region. In wealthy but size-constrained islands such as Singapore, Hong Kong and Macao, car use is far less than predicted by their rates of household consumption. Due to the high human density and reliable public transport systems in these regions, people living there have not required cars to the same extent as elsewhere. Conversely, high rates of car ownership in countries such as the US have led to the rise of suburbs and the related phenomenon of urban sprawl. Geographical factors also play an important role: due to the flat topography of many parts of western and northern Europe, such as Belgium, the Netherlands, Denmark and other parts of Scandinavia, bicycle use is very common. In the Netherlands, for example, 25% of all trips are made by bicycle.

Perhaps due in part to their lower per capita wealth, many developing nations such as India, Pakistan, China, Indonesia and Malaysia do not have rates of car use as high as in richer countries. Nevertheless, citizens of these developing countries meet their needs for transportation through a variety of cheaper modalities, both motorised and non-motorised. Particularly common among these modalities is two- and three-wheeled vehicles. Studies have shown a 65% prevalence of such vehicles to be typical for urban roads in India [1, 2]. These can be motorised, as in the case of motorcycles, scooters, mopeds and auto-rickshaws, or non-motorised, in the case of bicycles and cycle-rickshaws.

Just as with the composition of the vehicle mix, several factors may play a role in the different driving behaviours of various regions. Despite the varying cultures, norms and social attitudes in nations which have relatively homogeneous car-dominated traffic, the driving style in these regions tends to be uniformly disciplined and rule-based. By contrast, nations as diverse as India, Vietnam and Indonesia, which have a more heterogeneous vehicle mix dominated by smaller two-wheeled vehicles, tend to have traffic characterised by lack of lane discipline. While it is beyond the scope of this work to account for the historical, political, cultural and social factors that may lead regions to develop disparate patterns of driving behaviour, it is striking that there exist almost no regions with a highly heterogeneous vehicle mix and a high degree of lane discipline, nor regions that have a homogeneous car-dominated vehicle mix with low lane discipline.

It seems natural to suppose, then, that there may be a causal relationship between a high degree of vehicle heterogeneity and the emergence of lane disorder. Perhaps it is the case that when there is a range of vehicles of varying sizes on the road, it is advantageous to the individual driver to disregard lane discipline. Perhaps this is only true for certain classes of vehicles, such as those which are smaller and more manoeuvrable. Finally, perhaps there is a beneficial effect on the overall traffic network when vehicles in a heterogeneous mix decide to forgo lane discipline. The present work is motivated by these questions and seeks to use modelling and simulation to elucidate a causal relationship between vehicle heterogeneity and lane discipline.

In comparison to the body of work on traffic modelling in general, the literature on developing world traffic is limited, and has only received academic attention relatively recently. The first attempts to model such traffic consisted of adapting models and paradigms from conventional traffic flow theory by using traditional traffic flow models, but with a heterogeneous mix of vehicles. However, such approaches ignored the qualitatively different driving behaviours found in developing world traffic, where lane discipline is largely absent and smaller vehicles use gaps between other vehicles to advance along the roadway. Later models partially resolved this deficiency by allowing

vehicles to occupy any transverse position along the roadway, but retained the highly limiting assumption from conventional traffic theory that vehicles have a single leader. Empirical analysis of developing world traffic has shown this not to be the case.

The most recent model of developing world traffic was published in 2013. Its central assumption is that disordered heterogeneous traffic flow functions like fluid flowing through a porous medium. Smaller vehicles move through *pores* – spaces defined by the boundaries of other vehicles – to advance along the roadway. This model is able to replicate some functional phenomena observed in developing world traffic. However, it employs a layer of mathematical abstraction that is difficult to specify in practice and, being a macroscopic model, lacks the resolution required to fully interrogate individual vehicle behaviour. Due to the aggregate terms in which the model is expressed, it is not possible to directly translate its macroscopic formulation into a microscopic model that would allow ready assessment of vehicle trajectories.

Owing to the absence of suitable microscopic models for disordered heterogeneous traffic flow, it proved necessary for the purpose of this project to first formulate a new model for developing world traffic, in order to analyse the relationship between vehicle heterogeneity and lane discipline. The goals of this research project are therefore two-fold: 1) to formulate and evaluate a model for developing world traffic that accurately reproduces its characteristic lack of lane discipline, and 2) to use this model to investigate how traffic efficiency is affected by varying degrees of lane discipline and vehicle heterogeneity.

The model that has been developed is agent-based and draws on principles that have found successful application in the modelling of crowd dynamics. The basis of the model is that individual vehicle-agents will attempt to navigate through a stream of traffic by a process of *gap-filling*, achieved by moving in the direction that is as close as possible to a pre-defined preferred direction, while avoiding imminent collisions with other vehicles. In this way, the model avoids the shortcomings of previous microscopic models which assume that vehicles have a single leader upon which they base movement considerations, such as overtaking. The microscopic and agent-based nature of the model permits detailed analysis of individual vehicle trajectories and speeds, which is not possible in the macroscopic porous flow model.

The proposed model was tested by means of repeated simulation of varying traffic scenarios, and was found to replicate functional phenomena of developing world traffic accurately. The model was modified to accommodate the presence of lane discipline, and using this modification the effect of lane discipline on traffic efficiency was evaluated for a straight, urban, one-way link. It was found that for a homogeneous stream of cars, there was no significant difference in traffic efficiency between vehicles obeying lane discipline and those which did not. However, for a highly heterogeneous mix of vehicles,

as is typically found on developing world roads, it was found that lane discipline has a significant negative effect on traffic efficiency, both for the individual driver and the traffic stream as a whole. These results indicate that the emergence of lane disorder on roads in the developing world need not be reflective of differing cultural norms, but may have resulted from the higher efficiency of such a regime when the vehicle mix is highly heterogeneous.

These results, however, should not be regarded as definitive. The proposed gap-filling model, while able to accurately reproduce *functional* phenomena of developing world traffic, could not be calibrated due to a dearth of suitable data. Additionally, the model could be improved to take into account the expected speed variations of other vehicles, to more closely mimic the trajectory planning process of actual drivers. Finally, the only traffic geometry considered was that of a one-way link; a different relationship between vehicle heterogeneity, lane discipline and traffic efficiency may prevail for alternative configurations, such as intersections. Nevertheless, this work represents a novel approach to modelling of developing world traffic by introducing a chaotic, crowd-inspired dynamic model, and using this model to understand the fundamental relationships of traffic conditions that are a daily reality for many millions of people.

This report is structured as follows. Chapter 2 gives an overview of literature relevant to traffic modelling in general, the characteristics of developing world traffic, and a detailed account of current heterogeneous traffic flow models. Chapter 3 proceeds to give a full specification of the new gap-filling model, which is subsequently evaluated in Chapter 4 by means of functional validation and sensitivity analysis. Chapter 5 addresses the research question and discerns the effect of lane discipline and vehicle heterogeneity on traffic efficiency. Chapter 6 explores possibilities for future work, and discusses how the model might be adapted and integrated into other modelling frameworks. Finally, Chapter 7 summarises and presents the conclusions of this research project.

An electronic supplement to this thesis, consisting of a short video demonstration of simulation output, is available at `http://gapfi.s3-website-us-east-1.amazonaws.com`. This supplement is permanently archived on the Internet Archive and available at `http://web.archive.org/web/gapfi.s3-website-us-east-1.amazonaws.com/`.

Chapter 2

Review of Relevant Literature

This chapter provides a treatment of existing literature in the field of traffic modelling, with a focus on that which is relevant to heterogeneous traffic in particular. The chapter is divided into three sections. The first section gives an overview of traffic models in general. The next section provides a description of the characteristics of traffic in developing countries, describing its unique features and contrasting it with traffic in developed countries. The final section presents and examines in detail various models of heterogeneous traffic.

The goal of this chapter is to acquaint the reader with traffic modelling methodologies that are relevant to the problem of modelling traffic in developing countries, and to set the stage for the new model that will subsequently be presented.

2.1 Overview of traffic modelling and simulation

Traffic modelling is a highly active field of research due to the ubiquity of road transport in the everyday life of many millions of people, and the inherent difficulty of modelling a complex system consisting of thousands of locally-interacting agents. The scholarly literature on the subject is both diverse and voluminous, reflecting not only the large number of competing models used to describe the fundamental processes underlying traffic flow, but also the wide range of traffic scenarios that need to be modelled. What follows in this section is necessarily a highly condensed overview of this expansive literature.

One of the earliest attempts to model the flow of traffic was the fundamental diagram (FD), proposed by Bruce Greenshields in 1935 at an annual meeting of the American Highway Research Board [3]. The FD is a curve relating the average speed of vehicles on a stretch of road to the density of those vehicles, the latter being defined as vehicles per

unit length. The first FD presented by Greenshields posited an inverse linear relationship between speed and density, as shown in Figure 2.1 [3]. This is based on the simple intuition that, as the density of vehicles on a stretch of road increases, the speed of those vehicles decreases. This in turn follows from the assumption that each vehicle will attempt to maintain a relatively constant following time between itself and the vehicle immediately ahead, meaning that the following distance (which is inversely linearly correlated to density) will increase linearly with vehicle speed.

From the speed-density relationship it is possible to derive similar curves for speed-versus-flow, and density-versus-flow, where flow is the temporal rate at which vehicles pass through a cross-section of the roadway and thus has units of vehicles per unit time. The governing relationship between flow, speed and density in this simplified regime follows easily: flow is the product of speed and density. This is known as the *fundamental relation* of traffic flow [4]. The simplicity of the FD, together with its acceptable level of accuracy in certain scenarios, has led to its enduring use in traffic and infrastructure management to this day, allowing transportation planners to ensure that road networks are able to accommodate the traffic demand required of them.

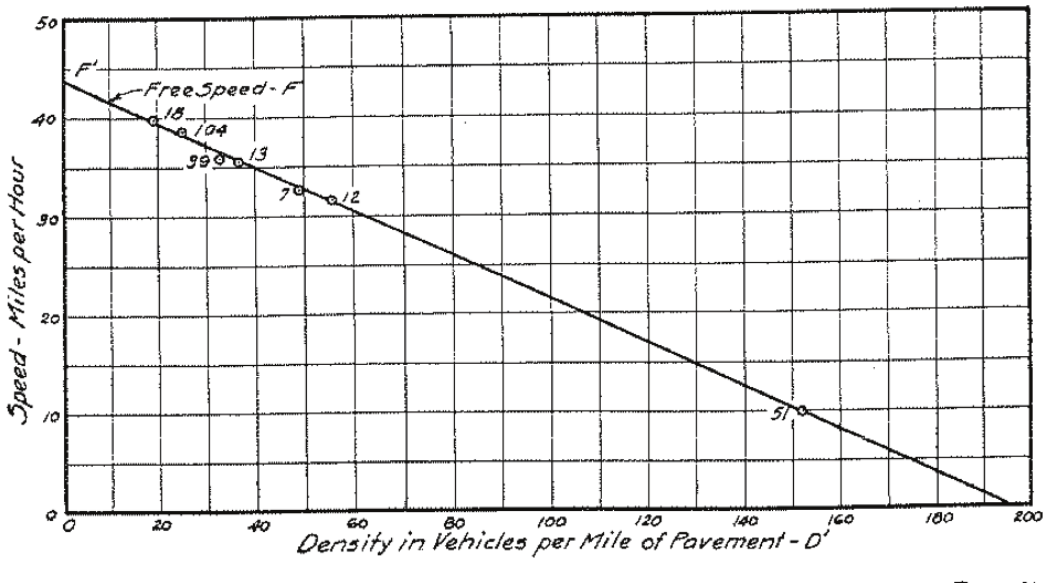

FIGURE 2.1: One of the first fundamental diagrams, proposed by Greenshields in 1935, predicts an inverse linear relationship between traffic speed and density. [3]

In the decades since the FD was first proposed there has been much research on the precise form of the curves, with various shapes being proposed. Greenshields' original linear speed-density relation gives rise to a parabolic density-flow curve but other forms that have been proposed include, inter alia, bell-shaped [5], parabolic-linear [6] and bi-linear [7]. The emphasis on the precise shape of the flow-density curve for various scenarios owes to its importance to transportation planners: flow is the rate at which vehicles pass through a stretch of road and therefore relates to traffic demand, while

the density of vehicles on a stretch can be controlled by the geometric parameters of the road (in particular, road width). Fundamental diagrams therefore have been, and continue to be, highly useful tools in the design of road networks that meet demand requirements while providing acceptable levels of service to road users.

Despite the simplicity and usefulness of the FD, plots of observed speed-flow and flow-density values generally show varying degrees of scatter around the predicted values. Since the FD fails to take into account vehicle dynamics, there are a number of specific phenomena that occur which the simple FD does not adequately predict. The precise nature of these phenomena is not within the scope of this chapter but they can include capacity drop following congestion and hysteresis [8]. Additionally, FDs are not applicable to various alternative traffic scenarios where vehicle dynamics become important, such as intersections, traffic circles, traffic lights, and special-use lanes, among others. Due to these limitations, a great number of traffic models have been proposed in the decades following first publication of the fundamental diagram.

Dynamic traffic models are generally classified according to their scale, with the most common designations in literature being micro-, meso- and macroscopic. To illustrate the relationships between these models and the chronology of their development, a genealogy is shown in Figure 2.2. While not every model depicted in the genealogy will be discussed, the schematic conveys both the breadth and depth of research into traffic models. Each of the three broad categories of models will now be discussed in turn.

2.1.1 Microscopic traffic models

Microscopic traffic models describe traffic flow at the scale of the individual vehicle. They describe the individual behaviour of each vehicle, as well as the nature of the interactions between vehicles. With reference to Figure 2.2, existing microscopic models are generally either *car-following* (CF) or *cellular automata* (CA) models. Strictly speaking, car-following models describe only the *interactions* between vehicles and make use of a separate model to describe free flow behaviour. However, in practice, these models encompass the complete range of vehicle behaviours. As can be seen from the genealogy, most research into microscopic models has been in the context of CF behaviour; indeed, CF models have been called "most important representatives of microscopic traffic flow models" [4, p. 165].

The basis of CF models is to relate the distance between a vehicle's speed to the distance between itself and the vehicle immediately ahead. In particular, each vehicle ensures that the distance to the vehicle immediately in front of it ("gap distance") is never less than a critical value ("safety distance"), which varies in proportion with the vehicle's

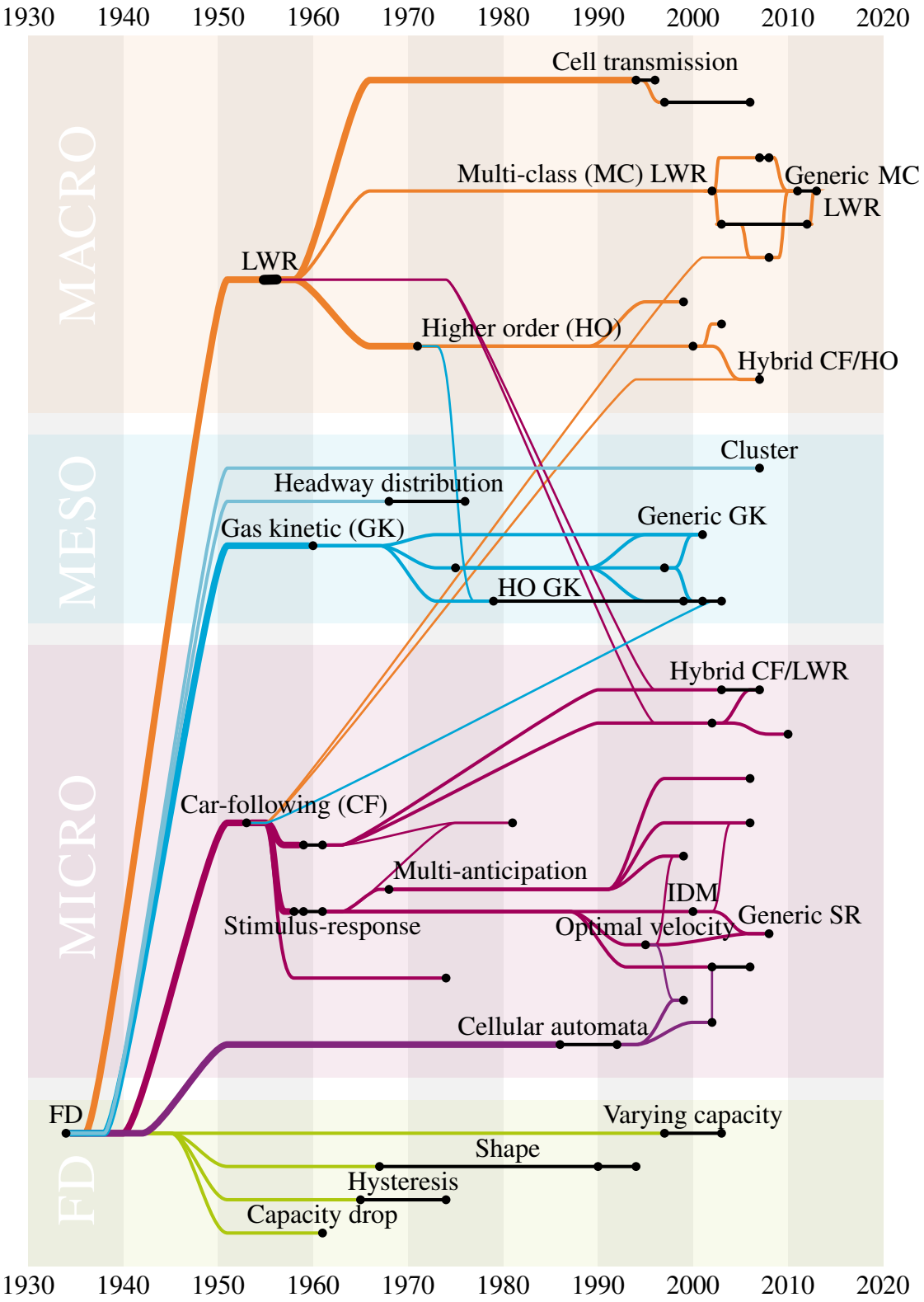

FIGURE 2.2: Genealogy of traffic flow models. Models are indicated by dots, coloured lines indicate descent and black lines between dots indicate that the same (or a very similar) model has been published multiple times. [8]

speed. The relationship between safety distance and vehicle speed is in turn determined by a "safety time", which is generally fixed. The precise way in which the speed of each vehicle varies with the gap distance accounts for the diversity of car-following models.

In the earliest CF models, vehicle speed is instantaneously adjusted in accordance with gap distance. This is clearly physically unrealistic, as it would give rise to infinite acceleration. Various higher order models correct this deficiency by including acceleration. An important example of such a model is the Optimal Velocity Model (OVM). This model posits that for every value of gap distance there is an optimal velocity for the following vehicle, which in turn is dependent on a number of parameters. Two examples of optimal velocity (v_{opt}) functions are shown in Figure 2.3. The instantaneous rate at which actual vehicle velocity approaches optimal velocity is governed by an adaptation time τ. The rate of change of vehicle speed $\dot{v}$ (i.e. acceleration) is updated according to equation 2.1 [4, p. 177], where s is gap distance. As can be seen, OVMs, and indeed CF models in general, introduce a number of parameters (most of which have been omitted in this brief description) which are not directly measurable and thus present barriers to direct model calibration.

$$\dot{v} = \frac{v_{opt}(s) - v}{\tau} \tag{2.1}$$

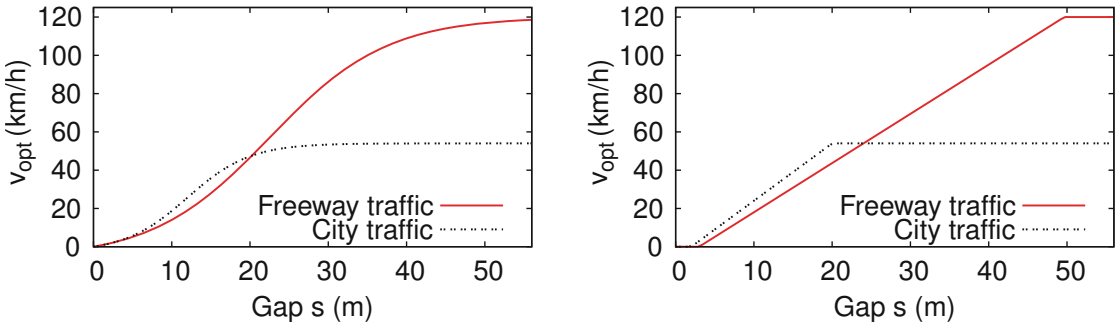

FIGURE 2.3: Plots of hyperbolic (*left*) and max-min (*right*) optimal velocity functions (v_{opt}) for OVM models. [4]

While falling outside the scope of this review, there are other subclasses of CF models, such as those known as the General Motors CF models, which define vehicle acceleration in terms of the speed difference between successive vehicles and gap size according to a power law relation, while not taking into account a vehicle's own speed.

The other major class of microscopic models makes use of cellular automata to capture traffic dynamics in a more general fashion. These models work by discretising space, time and therefore related state descriptors such as speed and flow. The roadway is discretised into a number of spatial cells which are each assigned an integer value, representative of the state of the roadway at that point. In the simplest CA models two values are used:

a value of 1 could mean a cell is occupied by a vehicle, and a value of 0 that it is empty [4][1]. In more advanced models, such as that of Nagel-Schreckenberg, this integer value represents the speed of a vehicle occupying that space [9]. Time is advanced in steps, with the diversity of CA models lying in the precise method of update. The simplest CA model, called "Rule 184", uses a 1-dimensional lattice of values and deterministically alters the value of a cell based on its own value and that of its two adjacent cells according to a rule table [4]. The Nagel-Schreckenberg model employs stochasticity to extend Rule 184 by adding a "dawdling" probability to capture drivers' propensity to accelerate less than is possible when gaining speed, or brake more than necessary when coming to a stop. Indeed it is the case with many CA models that they require a degree of noise or stochasticity to give stable and non-trivial results [4].

Despite their relative simplicity, CA models have been able to replicate various observed traffic phenomena, such as the spontaneous formation of traffic jams in the Nagel-Schreckenberg model, shown in Figure 2.4. Their simplicity makes them easy to implement and fast to compute, although the latter advantage diminishes in importance as computing power grows cheaper and more abundant. The discretised grid-like nature of CA traffic models makes them ineffective for modelling individual vehicle dynamics with high fidelity, nor are CA models particularly robust with respect to initial conditions or model parameters [4]. For these reasons and others, car-following models have emerged as the dominant form of microscopic traffic model.

The primary strength of microscopic models in general lie in their ability to model how individual vehicles affect flow of traffic. This leads to their being well-suited to applications where heterogeneity of vehicles is important, individual driving style and human reaction time is of interest, or where analysis of the specific interactions between vehicles is desired.

2.1.2 Macroscopic traffic models

Macroscopic traffic models attempt to model traffic as a continuum flow and, for this reason, are often known as *continuum* or *hydrodynamic models*. They describe the relationships between the dynamical variables flow, speed and density which are locally aggregated but vary with time and position in the stream, thus allowing the traffic flow to be viewed as a field. This property makes macroscopic models well-suited to studying phenomena that occur in fluid flow, such as the propagation of waves, or the development of congested areas (bottlenecks) [4].

[1] In the case of the Nagel-Schreckenberg model a cell value of zero indicates a stationary vehicle and an empty cell represents an unoccupied stretch of road space. Thus it is not strictly necessary that the cells in a CA contain only integers; any computable entity will suffice. Integers and and a non-numeric value to indicate unoccupied space are the most common arrangement.

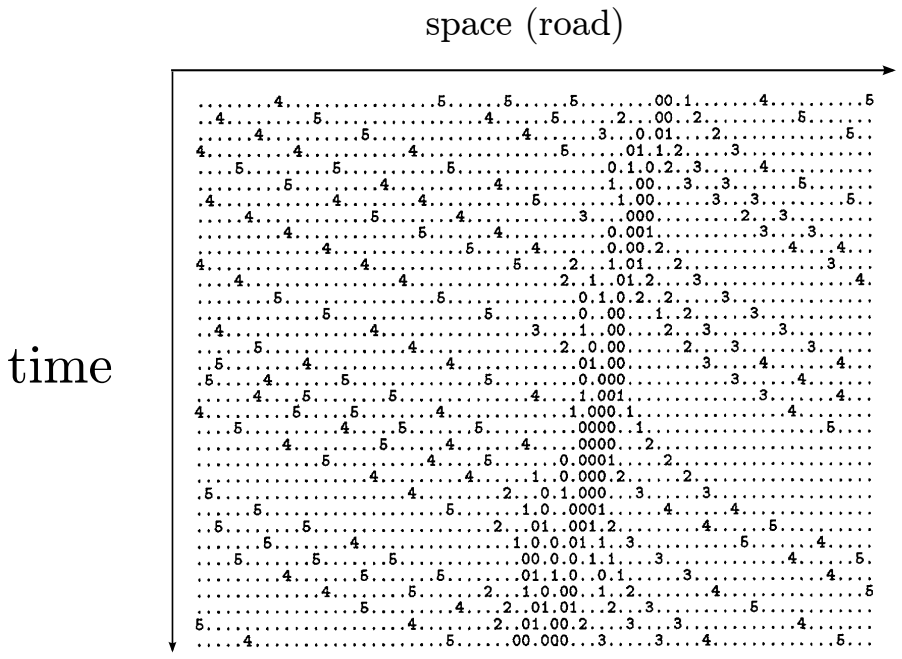

FIGURE 2.4: Temporal evolution of the 1-dimensional Nagel-Schreckenberg cellular automata model showing the formation and back-propagation of a traffic jam. Integers represent vehicle velocities and periods indicate unoccupied road space. [9]

Much like physical continuum models, macroscopic traffic models require a continuity equation to describe the temporal evolution of the system. Specifically, vehicle density is modelled as a function of flow differential (inflow and outflow of the control stretch), with the macroscopic vehicle speed calculated from the fundamental relation (also known as the hydrodynamic relation in the context of fluid flow). The specific form of the continuity equation will depend on the precise geometry of the road network being modelled. The continuity equation together with the fundamental relation can be closed by specifying flow or local speed and subsequently solved to provide the field functions for speed, density and flow as a function of time and position on the road space. A continuity equation and the fundamental relation are components of all macroscopic models; the diversity of macroscopic models lies in the precise nature of the density and flow equations [8].

A popular and powerful macroscopic model is the Lighthill-Whitham-Richards (LWR) model, proposed independently in the mid-1950s by Lighthill [10] and Richards [11]. Whereas conventional continuity formulations use differential equations for density as a function of differences in flow, the LWR model describes flow as a static function of density. Such a flow-density relation is a fundamental diagram, as described earlier. There are many specific varieties of LWR models that differ primarily in the nature of

the flow-density fundamental diagram[2]. Specifically, the flow-density or speed-density relation describes the *equilibrium* flow or speed as a function of vehicle density, with an assumption of the basic LWR model being that equilibrium flow or speed is instantaneously achieved. Thus under this assumption the flow at a point in space and time is given by equation 2.2 where Q is flow and ρ is density.

$$Q(x,t) = Q_e(\rho(x,t)) \tag{2.2}$$

An equivalent functional relationship holds for equilibrium speed. Combining the flow-density relationship with the continuity equation and the assumption that equilibrium flow is instantaneously achieved, the LWR model can be described mathematically by the single differential equation 2.3 [4].

$$\frac{\partial \rho}{\partial t} + \frac{\mathrm{d}Q_e(\rho)}{\mathrm{d}\rho}\frac{\partial \rho}{\partial x} = 0 \tag{2.3}$$

The assumption of instantaneous adjustment to equilibrium vehicle speed implies infinite vehicle acceleration, a severe shortcoming of the basic LWR model. Various higher-order models have been proposed to address this deficiency, and are currently considered the macroscopic "models of choice" [4, p. 127] for most purposes as they are relatively simple and can be rapidly computed [8]. These models introduce an acceleration equation to control the approach of vehicle velocity to equilibrium velocity, similar to the case of higher-order car-following models. With reference to the genealogy of Figure 2.2, some of these higher-order models have recently been combined with aspects of car-following models to produce hybrid macro-micro models [12]. However, because these hybrid models produce solutions for only the macroscopic dynamical variables (and thus lack the specificity of typical microscopic models) they are generally classified as macroscopic.

Another important variety of LWR-based models which have been developed in recent years is the multi-class (MC) LWR models. Like traditional LWR models they use fundamental diagrams to relate flow and density. However, they admit different classes of vehicles, each with its own (potentially different) fundamental diagram. The addition of separate classes of vehicles attempts to account for the heterogeneity of traffic composition which is poorly captured by traditional macroscopic models due to the inherent assumption of local homogeneity in any continuum [8]. Many MC LWR models also make use of an "effective density" in place of the standard density variable to account for the differential effect of some vehicles on traffic flow. For example, a small number of

[2]Note that, via the fundamental relation, any unique flow-density diagram corresponds to unique speed-flow and speed-density curves. The choice of whether to explicitly specify one diagram over its two equivalent curves is usually determined by the model assumptions and parameters.

buses can lead to congestion whereas a larger number of passenger cars or motorcycles would be required for the same level of congestion. This weighting of vehicles by their contribution to traffic flow is generally achieved by use of Passenger Car Units (PCUs) which will be discussed in detail in section 2.3.

In general, macroscopic models are favoured by transportation planners for their simplicity and ease of implementation, particularly since planners are largely interested in the macroscopic quantities of density, flow and speed. However, macroscopic models are ill-suited to applications where the effects of interest are difficult to capture macroscopically, such as interactions between vehicles, vehicle heterogeneity, lane changes, and so on.

2.1.3 Mesoscopic traffic models

Mesoscopic traffic models combine aspects of microscopic and macroscopic models. While the distinction is somewhat arbitrary and subjective, many mesoscopic models are viewed as either primarily macroscopic with microscopic components, or vice versa. For example, in a mesoscopic model that is primarily microscopic, the parameters of the model (such as vehicle speed and acceleration profiles) may depend on macroscopic quantities such as density. These are known as *local-field models*, due to their assumption of the existence of field-like properties, at least on a local scale. Conversely, a mesoscopic model that is fundamentally macroscopic in nature may use microscopic rate equations to determine the rate of vehicle flow into and out of the continuum. These are known as *master equations* [4].

In addition to these two classes of mesoscopic models, there are models which are neither primarily microscopic nor macroscopic. *Gas-kinetic traffic models* are the oldest mesoscopic models, which borrow principles from gas kinetics to model the flow of vehicles in a traffic stream. Rather than modelling each vehicle individually, probability distributions of density and speed are used, together with a notion of idealised "collisions", to produce expected speed and density values. Gas-kinetic models were introduced in the early 1960s, with revivals in the mid-1970s and mid-1990s. In recent years attention in the area of mesoscopic models has shifted to *parallel-hybrid models*.

Parallel-hybrid models are a type of multi-scale model that couple microscopic models and macroscopic models simultaneously. Microscopic models are used to for selected critical sections of a traffic network (such as intersections, traffic lights, on-/off-ramps), with the rest of the network modelled macroscopically. In this way it is possible to gain the high fidelity and specificity of microscopic models with the speed and large

scale of macroscopic models. This is achieved by two complementary processes known as *aggregation* and *disaggregation*, illustrated schematically in Figure 2.5 [4].

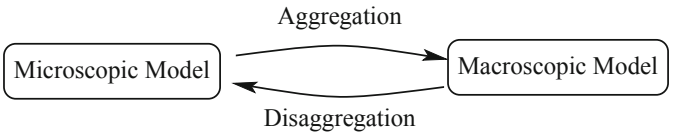

FIGURE 2.5: In mesoscopic models, the complementary processes of aggregation and disaggregation connect microscopic components to macroscopic components. [4]

The dynamical variables of density, speed and flow in macroscopic models are based on locally-aggregated information. Aggregation, then, is the process of calculating these macroscopic quantities by local aggregation of microscopic information. It is important that the regions over which this aggregation is done are of an appropriate scale both spatially and temporally. These regions should be *large* relative to the microscopic scale, so that sufficiently many vehicles are included, yet *small* relative to the spatial and temporal scale of the macroscopic events of interest, such as traffic jams or changes in traffic flow. Aggregation can be achieved mathematically by means of a weighted average using *kernel-based moving averaging*, where the weight of a vehicle's microscopic properties in the calculation of average macroscopic quantities at a point is determined by its spatial and temporal distance from the macroscopic point of interest. In practice many simulation tools make use of *virtual detectors* which are designed to replicate the output of physical traffic measurement tools. The extraction of macroscopic quantities from such data has been well-studied over the preceding decades [4].

The process of disaggregation establishes microscopic quantities from macroscopic variables. This is a more complicated procedure since macroscopic quantities inherently contain less information than those on the microscopic scale. Nevertheless, by allowing certain assumptions, this can be done. For example, setting the speeds of individual vehicles from a macroscopic speed variable can be achieved by assuming a particular distribution. A similar process can be followed for spacing vehicles longitudinally based on macroscopic density by assuming, for example, a negative exponential distribution for the distance between vehicles (as is commonly done). Nevertheless, such assumptions have been described as "ad-hoc" and "generally ... [not] well justified" [4, p. 69].

2.2 Characteristics of traffic in developing countries

Road traffic in developing regions of the world, particularly in the Indian subcontinent and south-east Asia, differs substantially from that found in industrialised countries in Europe and North America [1, 13–17]. Three notable characteristics of such traffic are

(1) greater heterogeneity of vehicle types, (2) reduced or absent lane discipline, and (3) a combination of motorised and non-motorised vehicles. In addition, in most urban areas vehicular traffic flow is also affected by a high degree of pedestrian traffic, lack of discipline at intersections, vehicles parking in the streets, overflow of nearby commercial activities onto the road space and, in some cases, narrow streets [13].

Motorised and non-motorised vehicles share the same roadway without segregation or sorting. Motorised vehicles include buses, trucks, light commercial vehicles, auto-rickshaws (three-wheeled vehicles), and various motorised two-wheelers such as motor-cycles, scooters and mopeds. Non-motorised vehicles include cycle-rickshaws, bicycles and animal-drawn carts [13]. It is clear that these vehicles differ substantially in their size, free speed, acceleration, manoeuvrability and driver behaviour. It has long been the case that many of the non-motorised vehicle modalities were viewed as undesirable by authorities and needed to be slowly discouraged over time, this goal being present in Calcutta's metropolitan transportation plan as recently as 1986 [13]. Since then authorities have adjusted their stance, and now seek to integrate motorised and non-motorised modalities, perhaps due to pressure to move towards environmentally and economically sustainable development [13].

Aside from the heterogeneity and unique types of vehicles, the other major difference between the traffic in developing and in industrialised countries is the former's lack of lane discipline [1, 13, 14]. Lane markings are often not present on the road surface and are frequently disregarded when they are present. Vehicle dynamics are characterised by a high degree of lateral movement. Vehicles do not move in single file and each vehicle has not one but possibly several leaders [13]. Speed variance between nearby vehicles can be high due both to their heterogeneity and also the significant amount of overtaking. On undivided bidirectional roads, vehicles regularly cross the centreline to overtake. Vehicle horns are typically used in short bursts when necessary to alert other drivers to the vehicle's presence [18].

Lane discipline as exhibited on industrialised roads encompasses a range of behaviours, most notably never overtaking a vehicle in the same lane and minimising time spent travelling over a lane marking. In this sense overtaking can be viewed as a 'procedure' that has a clear beginning (prior to initiating lateral movement) and clear end (cessation of lateral movement after having moved into a new lane). Overtaking in many developing countries, however, typically does not occur as a well-defined manoeuvre, but rather through a process of continuous weaving and lateral movement that has been described as 'seepage' [19]. Under this regime, vehicles that wish to advance will attempt to fill any gaps in the traffic stream without regard to lane boundaries. This occurs both when the vehicles being passed are moving or stationary.

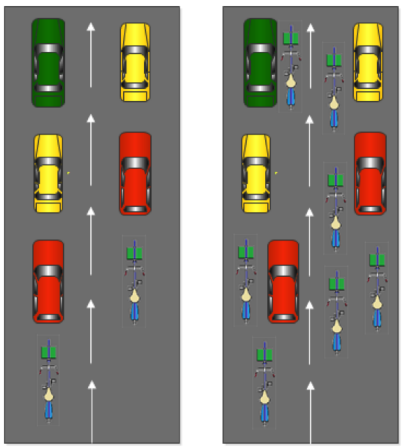

FIGURE 2.6: In overtaking manoeuvres as found in traffic in industrialised nations, vehicles never occupy the same longitudinal position as another vehicle in the same lane, as shown in figure (a). In traffic in developing nations as shown in figure (b), overtaking occurs by seepage of vehicles through any available lateral gaps. [19]

When referring to seepage of motorcycles in particular, the phenomenon is not unique to developing countries but is also found in industrialised nations where it is referred to variously as *lane sharing* or *lane splitting*. Most research into this practice has focussed on its safety, with mixed conclusions. Some studies have found it beneficial to the safety of motorcyclists because it reduces the chance of rear end collisions between motorcycles and larger vehicles [20], whereas others have found it can increase the likelihood of accidents [21, 22]. Nevertheless, little attention has been paid to the origins of widespread seepage in developing countries nor whether it is advantageous to the individual driver or to the traffic network at large [19].

Perhaps the single traffic scenario where the combined effect of vehicle heterogeneity and lack of lane discipline is most apparent is in the formation of queues [14], such as at stop signs or traffic lights. Smaller vehicles exploit their size and manoeuvrability to move ahead of larger vehicles and thus occupy positions nearer the front of the queue. This can be viewed as a type of seepage where the vehicles being passed are stationary. The process has been referred to as *creeping* as well as *percolation* [14], by analogy with fluid filtering through a porous medium.

To illustrate some of the unique behaviours just described, Figure 2.7 shows a collection of images depicting typical traffic scenarios on roads with disordered heterogeneous traffic. Images (a) and (b) show gridlock congestion at unsignalised intersections. Image (c) shows the end product of queue percolation, where smaller vehicles have exploited lateral gaps between larger vehicles to arrive at the front of the queue. Image (d) shows a highly congested road with several different vehicle types, illustrating that the percolation effect is not unconditional: it requires a path of sufficiently large gaps in order

for smaller vehicles to find their way to the front of the queue. Image (e) depicts a roundabout with a mix of different motorised and non-motorised vehicles, in addition to pedestrians weaving their way between the vehicles. Image (f) shows an undivided bidirectional road and illustrates the 'friction' introduced by having undisciplined streams of traffic flow in opposite directions without any segregation. Together with image (d), it again illustrates the gap-filling nature of this type of traffic: there are no gaps left in the stream that could be filled by a vehicle moving into them. Image (g) shows an example of significant encroachment of pedestrians and commercial activity onto the roadway. Finally, image (h) shows a bus crossing the centreline on an undivided bidirectional road in order to overtake a truck, thereby forcing vehicles travelling in the opposite direction to move into the shoulder of the road.

2.3 Heterogeneous traffic models

Despite the large amount of research that has been performed to produce the traffic models described in section 2.1, and the fact that disordered heterogeneous traffic is a daily reality for many millions of people around the world, relatively little research activity has focused on the modelling of such traffic. This section will review the existing literature in the relatively nascent field of heterogeneous traffic modelling.

On a point of nomenclature, the heterogeneity referred to in 'heterogeneous traffic' is typically that of the vehicle composition, indicating a range of vehicle modalities, as opposed to the relatively homogeneous mix found on roads in industrialised nations. Furthermore, some works [1, 2, 23] on modelling of traffic in developing countries use 'heterogeneous traffic' as short-hand to refer to the gamut of characteristics of such traffic, despite vehicle heterogeneity only forming a single component thereof. Other works make this distinction clearer by using the term 'disordered heterogeneous traffic' to refer to the combination of absent lane discipline and a heterogeneous vehicle mix [14]; the somewhat vaguer 'mixed traffic' has also been used for this designation [13]. In this report the term 'disordered heterogeneous traffic' as just described will be used most frequently due to its clarity. However, it should be emphasised that developing world traffic is characterised by more than just the combination of absent lane discipline ('disorder') and a heterogeneous mix of vehicles.

The earliest attempts at modelling lane-based heterogeneous flow took place in the context of classical traffic flow theory, where the heterogeneous vehicle mix was homogenised by the use of Passenger Car Units (PCUs), also known as Passenger Car Equivalents (PCEs), the latter term being more common in the Indian traffic literature. PCUs are essentially weighting coefficients that specify the relative effect of a vehicle class on some

FIGURE 2.7: Various road scenarios illustrating traffic conditions in developing countries. See the text for a full description. Assembled from various online video and photo sharing sources.

macroscopic traffic quantity, such as average headway or (more commonly) average vehicle speed and traffic flow. Passenger cars form the index of any system of PCU values and therefore, by definition, always have a PCU of 1. For example, a simple PCU derivation based on equivalence of flow for n different vehicle classes might have the form of equation 2.4 [14].

$$q_{total} = \sum_{i=1}^{n} PCU_i q_i + E \tag{2.4}$$

Using empirical data for flow (q), regression analysis may then be used on equation 2.4 to determine the coefficients PCU_i that minimise the error term E.

There exists a plethora of different techniques for calculating PCU values, which has led to a patchwork of different PCU sets for almost every different traffic condition that one may seek to model. Models based on PCU are generally much more accurate when used for lane-based traffic, due to the significant interaction effects and non-linearities introduced by lane disorder. PCUs first found widespread use for modelling heterogeneous traffic on Western roads and became a staple tool for transportation planners. In recent years as research activity on disordered heterogeneous traffic has increased, numerous attempts have been made to adapt PCUs for use in this context [24, 25]. Such endeavours have led to significantly differing PCU values that are found to vary with most parameters of the models in which they are used. For this reason they are generally avoided when accurate simulation of disordered heterogeneous traffic is needed [13, 14]. Nevertheless, the use of PCUs does present to transportation planners a simple and readily applicable tool for estimation of roadway capacities and they therefore still find widespread use in both developed and developing world contexts [24].

Research on lane-based heterogeneous flow has generally resulted in mesoscopic and macroscopic models. In 1999 a multi-lane multi-class phase-space density (MLMC PSD) model was proposed for modelling multi-lane heterogeneous flow [26]. The model characterises interactions between a fast-moving vehicle approaching a slow-moving vehicle, while using a lane change probability parameter to resolve the faster vehicle's hindrance. Similar to other gas-kinetic models, MLMC PSD results in a probabilistic phase-space density function which gives the density distribution $\rho_u^j(x, v, v^o, t)$ for each vehicle class u, in lane j, at longitudinal position x, with velocity v, with optimal velocity v^o, at time t. Most macroscopic models for lane-based heterogeneous flow have been multi-class extensions of the basic LWR model as discussed in section 2.1.2.

There have been several works addressing the microscopic simulation of disordered heterogeneous traffic flow, specifically that found in developing countries. Almost all of

these models employ car-following logic [1, 16, 17, 27–30], although a cellular automata approach has been attempted [31]. The majority of these car-following models have been implemented using standard traffic software designed for lane-based flow, with lane disorder approximated by an increased lane-changing probability. Those models that were implemented without the aid of traditional traffic simulation software adopt similar but varying approaches to modelling lane disorder. All of these car-following models assume that each vehicle has a single leader on which its overtaking considerations are based, which has been shown empirically to be untrue in the case of developing world traffic [13, 14]. They also cannot capture seepage due to their use of overtaking manoeuvres which, when invoked, move vehicles from their current lane to the target lane in an ordered fashion, with no dwelling on lane markers nor exploitation of lateral gaps between vehicles. Because vehicles in real disordered heterogeneous flow undergo a large amount of lateral movement, "it is inappropriate to use lane-based vehicle interaction models" [13, p. 2].

The first work specifically addressing disordered heterogeneous traffic without the use of car-following models was published in 2011 [14]. It proposes a macroscopic model which likens the traffic stream to a porous medium, where each vehicle moves through a series of pores defined by the gaps between other vehicles, as shown in Figure 2.8.

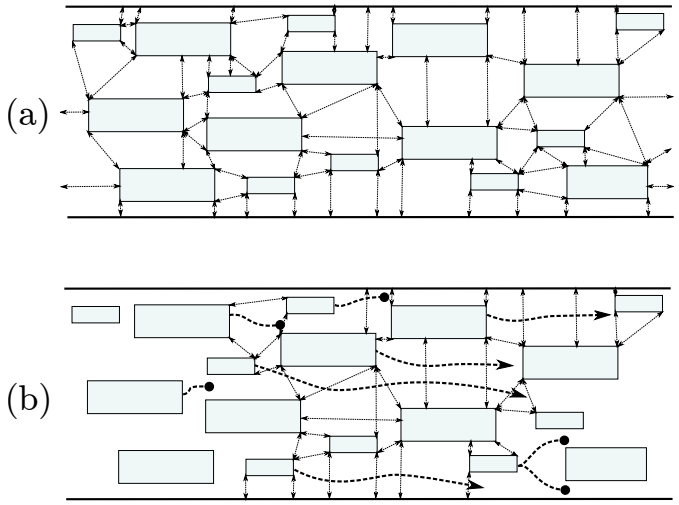

FIGURE 2.8: A traffic system can be regarded as a porous medium, with (a) a network of pores defined by the gaps between vehicles, where (b) vehicles advance through these pores. [14]

The model makes use of a statistical pore size distribution to quantify the size of pores, while each vehicle class has a defined minimum pore size through which it is able to move. Kinematic behaviour is governed by speed-density relationships for each vehicle class, with the assumption that density is additive across different vehicle classes. The size-exclusion action of the pores, together with the speed-density relations for each vehicle class, give rise to a set of simultaneous partial differential equations which can be

approximately solved by a finite difference scheme to give a modified speed-density relation for the entire traffic stream. Having solved for the modified speed-density relation, analysis of macroscopic variables such as speed, density and flow can proceed as with conventional macroscopic traffic models. It has been shown that the model is able to reproduce the functional phenomenon of percolation in queues in a two-vehicle system (cars and scooters).

The porous flow model as proposed, however, assumes that the pore size distribution is constant with respect to time. While this may be approximately true in the case of traffic streams that have reached equilibrium and are uninterrupted, in practice the pore size distribution would be constantly changing with respect to time. The authors note that specifying and empirically verifying a mechanism for temporal evolution of the pore size distribution would be "notoriously difficult" [14, p. 17]. In addition to this concern, the porous flow model suffers from the lack of specificity that is characteristic of any macroscopic model.

Subsequent work published in 2014 calibrated the above porous flow model using videographic analysis of urban Indonesian traffic [32]. In May 2015 a simplified mesoscopic model of queue percolation was proposed [19], using sorting of vehicles' expected travel times over a stretch of road (based on maximum vehicle speed) and the vehicles' PCU values to determine each vehicle's position in the queue. The model does not account for dynamic interactions between vehicles. To date there have been no published microscopic models of disordered heterogeneous traffic flow which do not make the assumption of car-following behaviour.

Chapter 3

Description of Agent-based Gap-filling Model

This chapter provides a comprehensive description of an agent-based gap-filling microscopic traffic model for the simulation of disordered heterogeneous traffic flow. The chapter is divided into three sections. The first section (3.1) provides an introduction to and motivation for the development of the model. The next section (3.2) covers mathematical and algorithmic details, thereby providing a full specification of the dynamical model. Finally, section 3.3 discusses the initial and boundary conditions that have been used in the present implementation.

3.1 Introduction

From the literature on developing-world traffic reviewed in the previous chapter, it is clear that a limiting assumption of existing microscopic models of disordered heterogeneous traffic flow is that of vehicles obeying car-following dynamics. Empirical analysis of such traffic has concluded that vehicles have not one but multiple leaders, and therefore do not base their movement considerations on a single vehicle but rather on the presence of sufficient available space in which to move. Recent models have taken a view that likens such traffic flow to movement of fluid through a porous medium. While the assumptions of the porous flow analogy are likely closer to reality than the assumption of car-following behaviour, the porous flow model and its concomitant need for a pore size distribution introduces an abstraction that is inherently difficult to quantify and, being a macroscopic property, provides little insight into the behaviour of individual vehicles in a traffic stream.

Inspired by the porous flow approach, and borrowing techniques from the modelling of crowd dynamics where individual agents show little to no coordination, the model proposed here likens disordered heterogeneous traffic flow to a gap-filling process. Vehicle agents act selfishly and with little coordination to advance along the road space by exploiting available gaps between other vehicles.

Modelling vehicles as decision-making agents permits a high level of resolution that is not possible in the macroscopic porous flow model. Properties such as individual vehicle speeds, spacing and trajectories can readily be assessed in the agent-based framework. By adjustment of agents' preferred direction of motion, it is possible to alter the model to handle arbitrarily complex traffic scenarios using the same basic dynamic model and agent decision-making logic. This will be demonstrated for the case of traffic lights and conversion to lane-based traffic. While the simulation environment can theoretically accommodate bi-directional roads, intersections, roundabouts, highway on-/off-ramps and so forth without significant change to the underlying model, throughout this paper the road environment is assumed to be that of a mid-block urban section, with traffic flowing in one direction.

Note that, since each vehicle is modelled as an agent, the text that follows frequently refers to the decision-making processes of *vehicles* as if they were themselves *agents* and, in the context of the model, uses these terms interchangeably. Furthermore, unless stated otherwise, it is assumed that vehicles are travelling directly to the right (in the direction of the positive X-axis) in the subsequent derivations, algorithms and diagrams.

3.2 Description of Model

Traditional microscopic car-following models make use of three disparate dynamic sub-models to specify overall vehicle behaviour [4]:

1. an *acceleration model* for longitudinal movement,
2. a *lane-changing model* for transverse movement, and
3. a *decision model* for other discrete-choice scenarios, such as using a priority lane.

The present model unites the first two of these dynamic sub-models under a common principle and, in its first iteration, does not consider discrete-choice scenarios. The model regards each vehicle on the road as a decision-making agent. Each agent continuously decides where on the road space it moves based on various imposed mechanical and physical constraints. In the temporally-discretised model proposed here, this decision

making takes place at every time step. Each agent is programmed with a preferred direction of motion, which can in principle be either constant or dynamic. At each time step each agent evaluates the space of all possible velocities[1] that may be followed for that time step based on various constraints, primarily maximum vehicle acceleration/deceleration and turning rate. The minimum time to collision of each possible velocity is calculated, and that velocity which avoids collision within a threshold period of time *and* advances the agent furthest in its preferred direction of travel is selected. In this way, agents move in their preferred direction at the maximum possible speed, while avoiding collisions with other vehicles. This process is the genesis of the gap-filling behaviour.

The rationale behind using time to collision (TTC) has its roots in both crowd modelling and conventional traffic flow theory. From the perspective of crowd modelling, a rule-based motion planning system for simulating (high-density) human crowds has been proposed [33], which assumes that individual agents navigate through a crowd in a preferred direction (such as toward an exit) by planning their movements so as to avoid collisions with other agents within a threshold value of time. This has proven to be an effective strategy for agents to move in a collision-minimal environment, even when two streams of agents are moving in opposing directions across one another. Similarly, in conventional traffic flow theory, a vehicle following a leader vehicle is assumed to adjust its following distance so as to maintain a fixed 'following time', leading to larger vehicle spacing at higher speeds. This assumption is manifest in the negative slope of the speed-density fundamental diagram, as well as in the "two-second rule" promoted by road authorities to encourage safe driving.

It should be noted that the collision avoidance model does not guarantee a collision-free environment. Collisions can still occur due to unexpected actions of other vehicles. For example, if one vehicle suddenly pulls in front of another, where the rear vehicle does not have sufficient time to brake or swerve, a collision can occur. Another example would be a leader vehicle dramatically reducing its speed in a short period of time, resulting in the follower vehicle not being able to brake sufficiently to avoid a collision.

Having discerned the commonalities between the dynamics of chaotic crowds and disordered traffic flow, the principal contribution made by the present work is to adapt the rule-based motion-planning system for crowds into the domain of traffic flow, which features vehicles with vastly different geometric and kinematic behaviour than individuals in a crowd. To this end, a detailed kinematic analysis has been undertaken in order to ensure that vehicle driving mechanics are sufficiently realistic. This fidelity becomes particularly important when vehicle densities are high, such as during queue

[1]The terms *velocity* and *speed* are referred to henceforth in their proper physical sense, as vector and scalar quantities, respectively.

formation or overtaking/seepage, since a hallmark of disordered heterogeneous traffic is the exploitation of geometric gaps in the roadway.

The general structure of the temporal update step of the model is given in algorithm 1. The details of TTC calculation have been omitted (but are given in full in section 3.2.3), and the velocity selection step has been simplified for the sake of clarity (the full velocity selection algorithm is given in section 3.2.2).

Algorithm 1: Temporal update

Data: $S^t = \{v_1^t, ..., v_n^t\}$
Result: $S^{t+1} = \{v_1^{t+1}, ..., v_n^{t+1}\}$
begin
  for $v_i^t \in S^t$ **do**
    $t_{crit} \longleftarrow$ vehicle's time to collision threshold
    $P \longleftarrow$ points in vehicle's discretised velocity space
    $P' \longleftarrow$ points in P sorted in descending order of optimality
    for $p \in P'$ **do**
      calculate $TTC(p)$ `// minimum time to collision of velocity` p
      if $TTC(p) \geq t_{crit}$ **then**
        $p^* \longleftarrow p$ `// chosen velocity`
        break
    if $TTC(p) < t_{crit} \; \forall \; p \in P'$ **then**
      $p^* \longleftarrow \arg\max_p TTC(p)$
    set position and orientation of v_i^{t+1} based on p^* and v_i^t

The next three sections give full mathematical and algorithmic details on the specification of the model. Section 3.2.1 discusses vehicle kinematics and presents the derivation for (1) a vehicle's velocity space at each time step, and (2) equations of motion for updating the vehicle's orientation and position based on the chosen velocity. Section 3.2.2 covers the decision-making algorithm for the selection of the optimal velocity from the velocity space. Finally, the details of the time-to-collision calculation are elucidated in section 3.2.3. These three sections form the core of the model. A fourth section (3.2.4) explains the modification that allows switching between disordered and lane-based flow.

3.2.1 Kinematics

This section will consider vehicle kinematics, as well as mechanical and ergonomic constraints, to derive the vector space of possible velocities and equations of motion for a vehicle moving along an arbitrary path.

Consider the trajectory of a vehicle as it moves across a length of road. The path traced out by any point affixed to the vehicle (for instance, its centroid) is a parametric plane curve, with the x- and y-components thereof a function of time. The second temporal derivative of this parametric curve is the vehicle's acceleration vector as a function of time. Since acceleration can never be mathematically undefined at any time[2], this plane curve must be at least twice differentiable. As such, at each point on the curve it is possible to specify an osculating circle: a plane circle that touches the curve at the point in question and whose normal line is perpendicular to the tangent of the curve at the same point [34]. In this way an arbitrary trajectory can be decomposed into infinitesimal circular arcs, each with a radius specifying the 'sharpness' of the curve at that point. Lower radii indicate sharper turns, higher radii wider turns, with a radius approaching infinity corresponding to straight-line motion.

The shape of a vehicle's trajectory at any point, and thus the radius and location of the osculating circle, is determined by the relative orientations of the front and rear wheels, which are in turn controlled by the driver using either the steering wheel or handlebars, depending on the type of vehicle. For two- or three-wheeled vehicles, the centre of the osculating circle is determined by the intersection of lines extended from the axle of the front wheel and rear wheel(s). Vehicles with more than three wheels make use of Ackermann steering geometry, as shown in Figure 3.1, to ensure that the lines parallel to the axle of each wheel all intersect at a single point. If this were not the case, wheel slipping would occur.

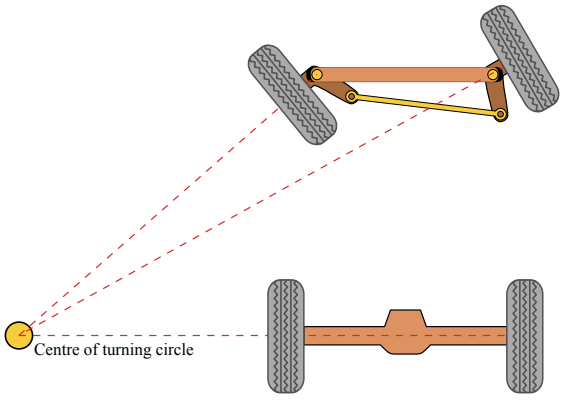

FIGURE 3.1: Ackermann steering geometry allows vehicles with three or more independent-axle wheels to turn without wheel slippage, by ensuring that the wheel axle centrelines meet at a common point. [35]

Consider the vehicle represented in Figure 3.2, which is undergoing circular motion over a small period of time Δt. The positions and orientations of the vehicle are shown at time t and $t + \Delta t$. The motion takes place within the global XY-coordinate system.

[2]By Newton's second law of motion, a body's acceleration is proportional to the resultant external force applied. While this resultant force can be zero, it can never be undefined.

Suppose another coordinate system is affixed to the vehicle, with its x-axis aligned with the vehicle's central longitudinal axis, and its origin, **c**, located along the rear axle of the vehicle. This coordinate system will translate and rotate as the vehicle moves along its circular path. Label this coordinate system xy at time t and $x'y'$ at time $t + \Delta t$, as shown in the figure.

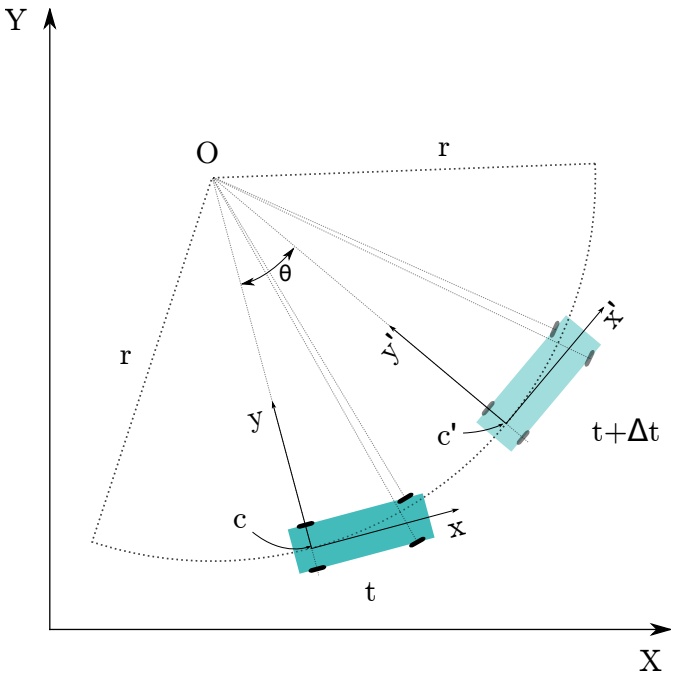

FIGURE 3.2: A vehicle moving along a circular arc.

The model assumes that at each time step, a driver controls the value of two quantities related to the vehicle's motion over the following time step:

- the vehicle's scalar speed, v, and
- the radius of the vehicle's turn, r.

These two values are determined directly by the vehicle's controls: speed is governed by use of the accelerator pedal or throttle grip, and radius of turn by the position of the steering wheel or handlebars. These values are also both subject to constraints:

- the vehicle's scalar speed, v, is controlled by the maximum acceleration and deceleration of the vehicle. The speed chosen for time step t, v_t, therefore must satisfy $v_t \in [v_{t-1} + a_{min}(v_{t-1})\Delta t, v_{t-1} + a_{max}(v_{t-1})\Delta t]$ where the values $a_{min}(v), a_{max}(v)$ are respectively the minimum and maximum accelerations possible for the vehicle when travelling at speed v.
- the radius of the vehicle's turn, r, is governed at low speeds by the minimum turning radius of the vehicle, which corresponds to the steering wheel or handlebars

being turned as far in one direction as possible. At higher speeds, however, inertial forces on the occupants are the limiting factor.

In order to generate the velocity space of the vehicle at time t so that a suitable velocity can subsequently be selected from this space, the speed and radius of turn chosen by the driver need to be expressed as a velocity in xy-coordinates (the $x'y'$-coordinate system cannot be used since its location and orientation is dependent on the velocity chosen). To do this, point $\mathbf{c'}$ needs to be expressed in xy-coordinates.

Denote the velocity chosen at time t as $\mathbf{v_t}$, corresponding to a speed v_t (equal to the vector norm of $\mathbf{v_t}$), with scalar components in the x- and y-directions denoted v_t^x and v_t^y, respectively. We seek to define the space of all velocities $\mathbf{v_t}$ which are feasible based on acceleration/deceleration and turning rate constraints.

The *distance* moved by the origin of the vehicle's travelling coordinate system from time t to $t + \Delta t$ is equal to the arc length s along the circular path from $\mathbf{c}$ to $\mathbf{c'}$. The *displacement* of the origin is equal to the chord length of the circular segment from c to c'. The length of an arc of radius r subtended by an angle θ is $s = r\theta$. This is equal to the distance moved by the vehicle over this time step, which is $v_t \Delta t$. Mathematically,

$$\begin{aligned} s &= r\theta \\ s &= v_t \Delta t \\ \therefore \theta &= \frac{v_t \Delta t}{r}. \end{aligned}$$

In a circle of radius r, the length d of a chord subtended by an angle of θ is $d = r\sqrt{2 - 2\cos\theta}$. However, angle θ has been shown to be proportional to the size of the time step Δt, which is assumed to be small, so a small-angle approximation can be made. The small-angle approximation for $cos\theta$ is $1 - \frac{\theta}{2}$, so we have

$$\begin{aligned} d &= r\sqrt{2 - 2\cos\theta} \\ &= r\sqrt{2 - 2\left(1 - \frac{\theta}{2}\right)} \\ &= r\theta \\ &= s. \end{aligned}$$

Therefore, under the small-angle approximation, $d = s$, i.e. chord length and arc length are equal. This implies that the distance from $\mathbf{c}$ to $\mathbf{c'}$ is equal to $v_t \Delta t$. Using the cosine rule it is possible to show that the sine of the angle between a line from $\mathbf{c}$ to

$\mathbf{c}'$ relative to the x-axis is $\frac{v_t \Delta t}{2r}$. By similar triangles it follows that the y-component of the velocity vector, v_t^y is $\frac{(v_t)^2 \Delta t}{2r}$, with the x-component $v_t^x = \sqrt{(v_t)^2 - (v_t^y)^2}$ from Pythagoras' theorem. The vehicle's chosen speed v_t and radius of turn r can then be expressed as a velocity vector in xy-coordinates by means of equations 3.1.

$$\begin{aligned} v_t^y &= \frac{(v_t)^2 \Delta t}{2r} \\ v_t^x &= \sqrt{(v_t)^2 - (v_t^y)^2} \end{aligned} \tag{3.1}$$

Having expressed v_t and r in terms of the components of a velocity vector $\mathbf{v_t}$ in xy-coordinates, it is now possible to translate the constraints on v_t and r to constraints on $\mathbf{v_t}$. As mentioned before, constraints on the speed v_t are imposed by the acceleration and deceleration of the vehicle, and constraints on radius of turn r are due to minimum turning radius at low speed and inertial forces at high speed.

For a body undergoing circular motion at velocity v_t and radius r, inertial forces give rise to a radial acceleration of magnitude $a_r = \frac{(v_t)^2}{r}$. A 2015 study on passenger cars on highways in China found that the threshold value of comfortable radial acceleration is $a_{r,max} = 1.8 m/s^2$ [36]. It is reasonable to assume that this value will be similar for occupants in other types of vehicle, as it is primarily related to human comfort and not the stability of the vehicle[3]. A vehicle's minimum turning radius is a function of its geometry and wheel base, and therefore differs by vehicle type. Table 3.1 shows the minimum turning radii for the six vehicle classes used in this simulation study.

TABLE 3.1: Minimum turning radius by vehicle class

Vehicle type	**Minimum turning radius r_{min} (m)**	**Source**
Bus	12.4	American Association of State Highway and Transportation (AASHTO)
Truck	11.6	AASHTO
Light commercial vehicle	7.79	Unavailable; interpolated from truck and passenger car on basis of vehicle length
Passenger car	6.4	AASHTO
Auto-rickshaw	2.88	Manufacturer (Lovson)
Motorised two-wheeler	1.56	Manufacturer (Bajaj)

[3]Most vehicles can in fact withstand radial accelerations several times higher than $1.8 m/s^2$ before either wheel slipping or toppling occurs.

The mechanical and physical constraints on v_t and r are given by the set identities 3.2.

$$\begin{aligned} v_t &\in [v_{t-1} + a_{min}(v_{t-1})\Delta t,\ v_{t-1} + a_{max}(v_{t-1})\Delta t] \\ r &\in \left[\max(r_{min}, \frac{(v_t)^2}{a_{r,max}}), \infty\right) \end{aligned} \tag{3.2}$$

Combining the second constraint of identities 3.2 with the transformation equations 3.1, Figure 3.3 shows the space of all velocity vectors that satisfy turning radius requirements across a range of speeds. The tangential and radial velocities referred to in the axis labels are exactly equivalent to v_t^x and v_t^y of the xy-coordinate system, respectively, but have been renamed to make their meaning more apparent[4]. Note that at very low speeds, the radial velocity is also low, since slow-moving vehicles cannot turn rapidly. As speed increases so does radial velocity, with a clear transition around $12km/h$. This indicates a change from the regime where turning radius is limited by the vehicle's minimum turning radius, to the regime where inertial forces reach the comfort threshold and become the limiting factor for the vehicle's turning radius. This implies that if one were to accelerate a passenger car from rest, with the steering wheel turned to its maximum extent in either direction, the inertial forces would become 'uncomfortable' as the car's speed reaches $12km/h$, and drivers will instinctively avoid making turns sharper than this.

Figure 3.3 also gives an example of a passenger car travelling at a speed of $40km/h$, whose velocity vector space (when also taking acceleration/deceleration constraints into account) is shaded in red.

Once a velocity has been selected from the feasible region (as is explained in the next section), the vehicle's position and orientation needs to be updated for the following time step. To do this, the chosen velocity needs to be transformed from xy-coordinates to global XY-coordinates by means of rotation. Equations 3.3 show the update rules for translation and rotation, respectively. The rotational update rule is obtained from combining the rotation of the vehicle over the time Δt, $\theta = \frac{v_t \Delta t}{r}$, with $v_t^y = \frac{(v_t)^2 \Delta t}{2r}$, which was given earlier.

$$\begin{aligned} \mathbf{c}_{t+1} &= \mathbf{c}_t + \mathbf{v_t}\Delta t \\ \theta_{t+1} &= \theta_t + \frac{2v_t^y}{v_t} \end{aligned} \tag{3.3}$$

[4]While it is true that a vehicle can have only radial *acceleration* and never radial *velocity* (without wheel slippage), this is the case only when the coordinate system rotates and translates with the vehicle. Since we are attempting to express the velocity of the vehicle over a finite time step in the stationary xy-coordinate system, *apparent* radial velocity is possible, with the magnitude of this velocity proportional to the size of the time step.

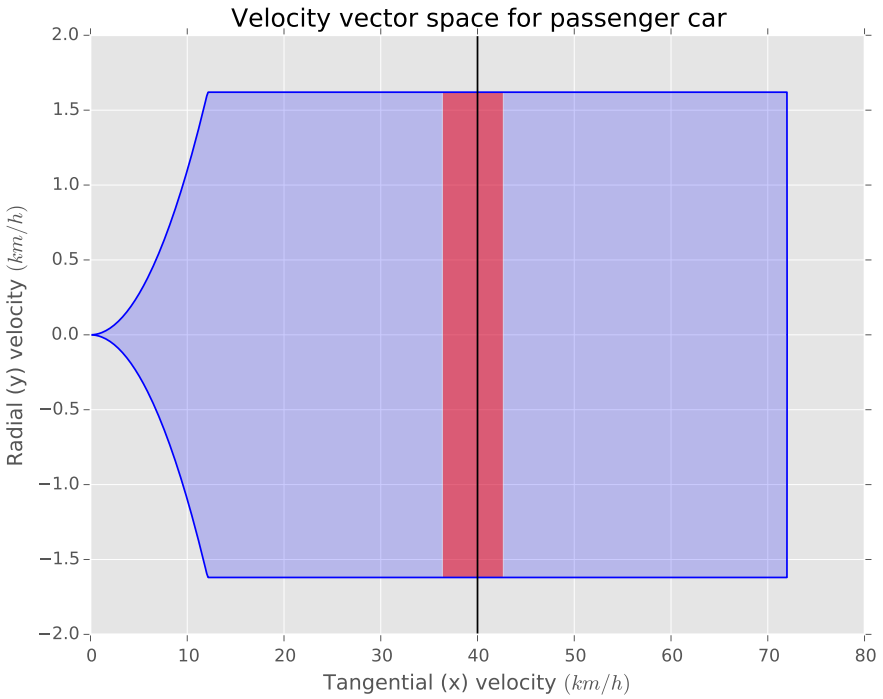

FIGURE 3.3: Velocity vector space of a passenger car for a time step $\Delta t = 0.5$ seconds. The blue shaded region indicates the set of all velocity points that satisfy turning radius constraints at each speed. For a hypothetical passenger car travelling at $40 km/h$ (vertical black line), the red shaded region indicates the set of all velocity vector points that additionally meet acceleration/deceleration constraints at this speed.

3.2.2 Velocity vector selection

The selection of the velocity vector from the feasible velocity space, such as the red region of Figure 3.3, is the crux of the microscopic model. It can be summarised heuristically as "choose the velocity that avoids collisions within some threshold period of time, and that has the largest component in the preferred direction of movement".

This process consists of four steps:

- Discretise the velocity vector space into a number of velocity vector points (VVPs).
- For each VVP, calculate the minimum TTC with other vehicles or obstacles.
- Eliminate all VVPs where TTC is less than threshold t_{crit}.
- From the remaining VVPs, choose that which is considered optimal. This usually means the VVP which has the largest component in the preferred direction of movement.

In practice, calculating the minimum TTC of a particular VVP is a relatively expensive operation as it requires iterating over all other vehicles and performing several pairwise

geometrical operations in order to predict the time to collision. These are described in section 3.2.3. Nevertheless, since the optimality of a VVP does not depend on the precise value of its TTC, but only requires that its TTC exceeds the threshold t_{crit}, it is possible to significantly speed up VVP selection by first sorting VVPs in descending order of optimality, then iterating over the sorted list until a VVP has a TTC above the threshold value. This operation is described in pseudocode in Algorithm 2. The binary nature of the TTC constraint provides a significant speed advantage over other possible algorithms which require knowing the precise value of TTC. However, there are also potential drawbacks of the present approach, particularly that TTC may be used as a proxy for the size of a gap. Under some circumstances, larger gaps may be more desirable than smaller gaps.

The calculation of the time to collision of a particular velocity vector point can be significantly optimised by the use of two parameters which control shortcut processing, and are described more fully in section 3.2.3. The 'minimum threshold' halts collision checking when a TTC below a minimum threshold is found. For example, if the minimum threshold is set to 4 seconds and a collision is found to occur within 3.5 seconds, then the TTC algorithm will halt immediately and not check for any more collisions, even though these may occur sooner than 3.5 seconds. Conversely, the 'maximum threshold' instructs the algorithm not to perform detailed collision checking if a cheap preliminary check indicates that the TTC must be above this threshold value.

Algorithm 2: Velocity vector point (VVP) selection

Data: V = continuous velocity vector space
Result: v^* = optimal velocity vector in V
begin
  $v_{\text{list}} \longleftarrow$ sample velocity points from V
  $v_{\text{list}}^{\text{sorted}} \longleftarrow v_{\text{list}}$ sorted descending by optimality
  $t_{crit} \longleftarrow$ vehicle's time to collision threshold in seconds
  for $v_i \in v_{list}^{sorted}$ **do**
    $TTC_i \longleftarrow$ TTC of v_i using a minimum and maximum threshold of t_{crit}
    if $TTC_i \geq t_{crit}$ **then**
      return $v^* = v_i$
  if $TTC_i < t_{crit} \; \forall \; v_i \; \in v_{list}^{sorted}$ **then**
    for $v_i \in v_{list}^{sorted}$ **do**
      $TTC_i \longleftarrow$ TTC of v_i using a min threshold of 0 and max threshold of t_{crit}
    return $v^* = v_i : \arg\max_i TTC_i$

The method for selecting the optimal VVP is based on the notion that an agent should select the velocity which has the largest component in its preferred direction of motion. Since the preferred direction of motion is specified in global XY-coordinates, VVPs need to be expressed in these coordinates, and not in the original xy-coordinates in which

they are generated. This is done by rotation of the xy-coordinate system so that it aligns with the XY-coordinates. If multiple VVPs have the same largest component in the preferred direction, that which would result in the vehicle's new orientation being closest to straight is selected, by means of the rotational update rule given in equation 3.3.

Care needs to be taken when discretising the velocity space in order to avoid problematic vehicle behaviour. With reference to the velocity space of a hypothetical passenger car travelling at $40km/h$ (shown as the red-shaded region in Figure 3.3), this vector space appears to be rectangular in shape. However, considering that its left- and right-hand edges are generated by the equation $v_t^x = \sqrt{(v_t)^2 - (v_t^y)^2}$, it is clear these edges are in fact circular arcs, but merely appear to be straight due to the vastly different scales of v_t^x and v_t^y. If the velocity space is discretised using a rectangular grid in xy-coordinates, the slight circularity of these edges can be lost. The importance of this circularity is illustrated by means of an example, shown in Figure 3.4. The velocity spaces, transformed into global XY-coordinates, are shown for a vehicle in three different orientations. The first, figure (a), shows a vehicle which is aligned with the global coordinate axes. In this case, selection of an optimal VVP v^* proceeds as described in the previous paragraph, with the agent selecting a velocity that results in maximum speed while keeping straight. Next, suppose that an obstruction (such as the presence of another vehicle) caused the agent to momentarily rotate a small amount, shown in figure (b). In this case, the agent selects the VVP with the largest component in the preferred direction of motion. However, the VVP selected would result in the vehicle following a velocity corresponding to maximum speed, but also maximum *rate of turn*, causing it to lurch to the right, giving rise to the situation of figure (c). By the same logic, the agent selects a VVP corresponding to a maximum left turn, returning it to the scenario of figure (b). In this way, a rotational oscillation is created that does not cease in the absence of external interference (such as another obstruction being presented).

It is therefore critical to ensure that the velocity space is discretised so that the circularity of the left- and right-hand edges is preserved. This can be achieved by uniformly sampling v_t values from its range (given in equation 3.2), and then sampling v_t^x, v_t^y values based on the constraints of r. This procedure is shown in algorithm 3.

The use of a velocity selection algorithm that selects the velocity point with the largest component in the preferred direction of motion, while avoiding imminent collisions, obviates the need for separate specification of acceleration and lane-changing sub-models. Under this regime, vehicles will accelerate at the maximum rate permitted at their current speed, as specified by the acceleration function. In order to provide an upper bound on vehicle speed, the specification of an acceleration function which decreases to zero at

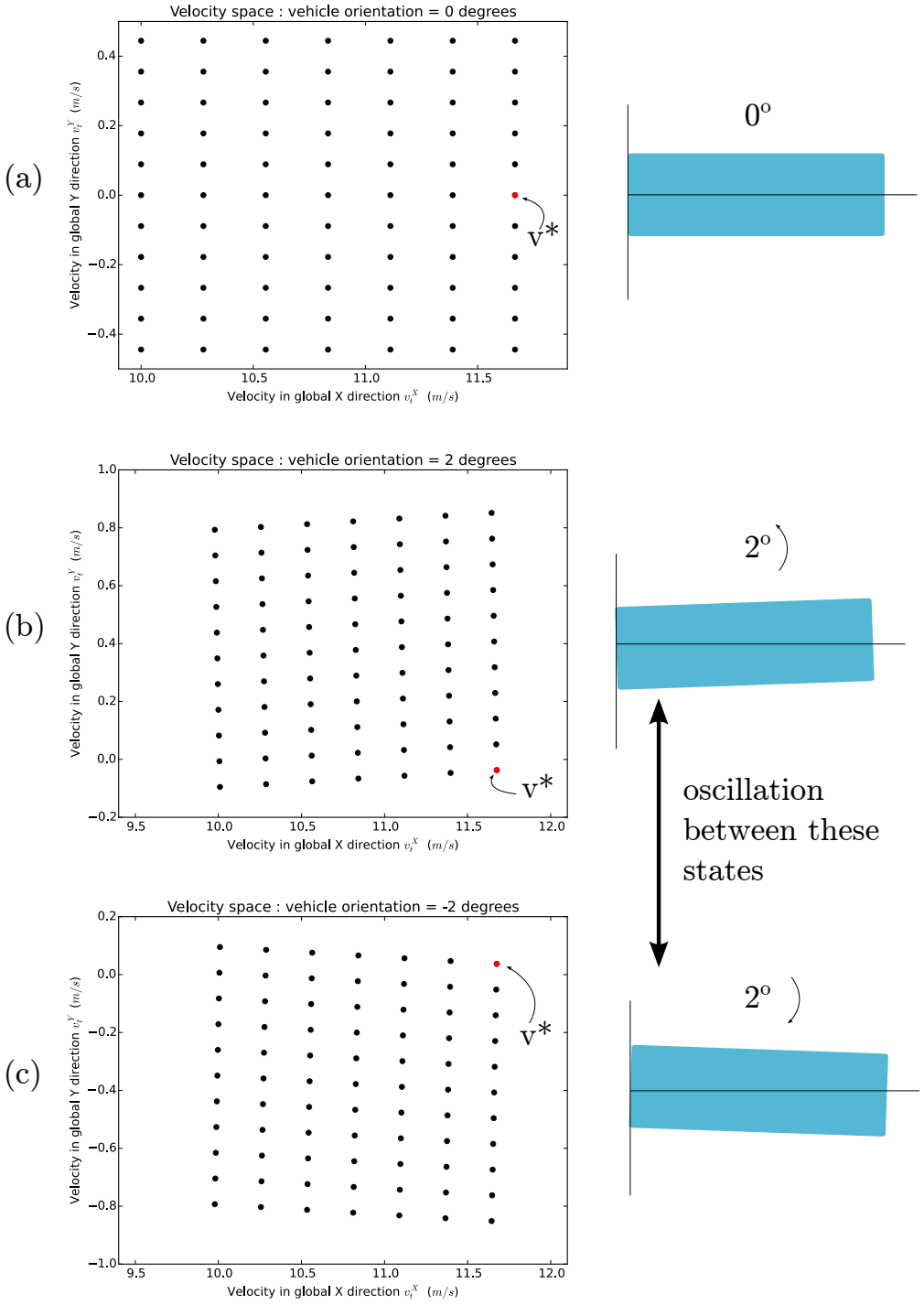

FIGURE 3.4: Mechanism behind rotational oscillation of a vehicle which may occur due to an improperly discretised velocity space.

Algorithm 3: Discretisation of velocity space

Data: V = continuous velocity vector space; (j, k) = size of grid
Result: v_{list} = list of points $v \in V$
begin
  $v_{\text{list}} \longleftarrow$ empty list
  for $v_t = j$ *equally-spaced points in* $[v_{t-1} + a_{min}(v_{t-1})\Delta t,\ v_{t-1} + a_{max}(v_{t-1})\Delta t]$ **do**
    $r = \max(r_{min}, \frac{(v_t)^2}{a_{r,max}})$
    for $v_t^y = k$ *equally-spaced points in* $\left[-\frac{(v_t)^2\Delta t}{2r}, \frac{(v_t)^2\Delta t}{2r}\right]$ **do**
      $v_t^x = \sqrt{(v_t)^2 - (v_t^y)^2}$
      append (v_t^x, v_t^y) to v_{list}
  return v_{list}

some value of speed is necessary. Further details on the acceleration functions used in the accompanying simulation study are given in section 4.1.1. As just mentioned, separate specification of a lane-changing model is also not necessary. Agents will independently move around other agents in order to avoid collisions and, due to the self-righting nature of the velocity selection algorithm, then return to an orientation aligned with their preferred direction.

3.2.3 Calculation of time to collision

The accurate calculation of time to collision between pairs of vehicle agents forms an important part of the model. The present model differs from traditional crowd simulation models because of the agents' rectangular shape (as opposed to circular) and from traditional microscopic traffic models because these rectangles can be rotated with respect to the road. Both of these factors make TTC computation geometrically more complex and therefore computationally more expensive. Because of the exploitation of small gaps in traffic and the high degree of interaction between vehicles exhibited in developing world traffic, the need for accurate kinematic and geometric modelling of vehicles is considerably higher than when modelling homogeneous lane-based traffic.

While computing the time to collision for a pair of moving circles is a simple and computationally cheap vector calculation, the geometric fidelity of approximating a rectangle with a circle is low. For the vehicles used in the accompanying simulation study, the maximum ratio of length to width is 4.12, in the case of a bus. Using a circumscribed circle approximation, the width of the vehicle would be exaggerated by a factor of 4.2 and its area by a factor of 3.4. Even for the vehicle with the lowest aspect ratio, its width would be exaggerated by a factor of 2.1 if using circumscribed circles.

Computing time to collision for a pair of moving, arbitrarily-rotated rectangles is a computationally expensive operation due to the discontinuities introduced by the perpendicular edges. The algorithm chosen for computing TTC for rectangles is modular and independent of the model *per se*. The method chosen in the present implementation is to compute the minimum time until any one of the corners of either of the rectangles intersects an edge of the other rectangle. If the rectangles are already overlapping, TTC is returned as zero.

As mentioned, calculating TTC for rectangles is computationally expensive and the precision of such a check is unnecessary for agents that are separated by a large distance and thus will not collide in a time horizon similar to the threshold TTC t_{crit}. For this reason, the model makes use of a hybrid collision checking algorithm. A preliminary check computes the TTC by approximating the rectangles as circumscribed circles, which will always provide an underestimate of the time-to-collision. If this value lies above a specified threshold t_{upper}, the detailed rectangle-based TTC is not computed, thus avoiding unnecessary computation. A minimum threshold, t_{lower}, is also used to halt TTC computation early if a computed TTC value is less than t_{lower}. Thus, the returned TTC value is only necessarily precise when it lies in the range $[t_{lower}, t_{upper}]$. For the purposes of the present model, VVPs are only assessed on whether their TTCs lie above the threshold t_{crit}, so these thresholds are set to be $t_{lower} = t_{upper} = t_{crit}$.

One characteristic of disordered heterogeneous flow that is observed in practice, and relates to seepage, is that vehicles moving transversely across the roadway do not consider the movements of vehicles that are behind them. Thus, a vehicle may move directly in front of another vehicle, with the expectation that the rear vehicle will adjust its speed and possibly direction in order to avoid a collision. To capture this behaviour in the model, each agent, when checking the TTC of a particular velocity against the other vehicles on the road, considers only those vehicles whose centroid lies in front of its own centroid. The benefit of this approach is that it allows for automatic recovery from collisions. If two vehicles collide, only the rear of these two vehicles detects the collision, and this vehicle will stop moving. Since the forward vehicle does not detect a collision, it will continue moving away from the rear vehicle, thus resolving the collision.

The TTC computation algorithm is given in algorithm 4.

The present model does not account for acceleration of other agents. When calculating the TTC of a pair of agents, it is assumed that the velocities of all agents stays constant. This, of course, is not the case in reality. Without the *expectation* that other vehicles will change their speed, a follower vehicle would accelerate behind a leader vehicle until the headway between the two is nearly zero. It is only if the follower considers that the leader vehicle might slow down that it will maintain a non-zero following distance. In

Algorithm 4: Computation of time to collision for a single agent

Data: v = velocity vector point (VVP); w_{list} = list of other agents and obstacles
t_{lower} = lower TTC threshold; t_{upper} = upper TTC threshold;
Result: TTC_{min} = minimum TTC
begin
  $TTC_{min} \longleftarrow \infty$
  for $w_i \in w_{list}$ **do**
    if *location of* w_i *is ahead of current agent in preferred direction* **then**
      $TTC^{circ} \longleftarrow$ TTC using circumscribed circles
      if $TTC^{circ} \geq t_{upper}$ **then**
        $TTC_{min} = \min(TTC_{min},\ TTC^{circ})$
        skip checking w_i, proceed to w_{i+1}
      $TTC^{rect} \longleftarrow$ TTC using rectangles
      $TTC_{min} = \min(TTC_{min},\ TTC^{rect})$
      if $TTC_{min} < t_{lower}$ **then**
        break out of loop
  return TTC_{min}

practice, vehicles adjust their following distance based on their knowledge of the possible decelerations of other vehicles, so as to keep their time-to-collision above a comfortable threshold. Implementing acceleration in the model would require significantly more complex geometrical operations for calculating TTC than for a constant velocity model. For this reason, the constant velocity assumption was made.

In order to avoid the problem of near-zero following distance without introducing the complexity of an acceleration-based TTC computation, an approximation was used. When an agent computes the time-to-collision with another agent, it considers two separate velocity vectors for the other agent: the agent's true velocity, $\mathbf{v_b}$, and a velocity of reduced magnitude but in the same direction, $\alpha\,\mathbf{v_b}$, where $0 \leq \alpha \leq 1$. The constant α can be considered the degree to which the follower agent expects the leader agent to maintain its current velocity. The use of the constant α introduces the problem of selecting a value for it. If it is too high, then vehicles will presumably collide more frequently; if it is too low, vehicles will move too slowly and space themselves further apart. Details of how sensitivity analysis was used to determine a suitable value of α are given in section 4.3. Based on that analysis, a value of $\alpha = 0.5$ was used for all simulations.

There is also the need to specify a threshold TTC value t_{crit} for each vehicle. Based on the worst-case scenario of an agent approaching a stationary object (or another vehicle stopped in a queue), the agent should be able to reduce its speed in time so as to avoid collisions. From basic equations of motion, if a body is travelling with a speed v_t and

accelerates at a rate of a_{min} (which is negative), the time taken to become stationary is given by equation 3.4.

$$t_{crit} = \frac{v_t}{|a_{min}|} \tag{3.4}$$

It is therefore possible to dynamically update the vehicle's TTC threshold based on the vehicle's speed and its acceleration due to maximum braking, a_{min}.. This approach is adopted in the model as it has the dual advantage that the vehicle will always be able to avoid collision with a stationary object, and that the need for calibration of t_{crit} is obviated.

When one vehicle passes another side-by-side, there needs to be a gap between the vehicles to avoid contact damage. Videographic analysis of vehicle behaviour has revealed an apparent preference for the size of this transverse gap which increases approximately linearly with speed, and has a minimum and maximum threshold, dependent on vehicle type [2]. A similar preference for maintaining a gap in the longitudinal direction is observed, even when vehicles are stationary. This gap is purely due to driver preference and is not related to maintaining a safe following distance, as it is present even when vehicles do not move (such as in a queue). Therefore when TTC computation is performed, a smal lateral and longitudinal clearance is added to the dimensions of the vehicle. Note that this clearance is only added for the agent performing the velocity selection: the clearance requirements of other agents need to be enforced by those agents themselves.

As a consequence of different vehicle types having varying preferences for lateral clearance, a modelling anomaly can arise when a vehicle with low clearance preference (e.g. a motorcycle) passes near to a vehicle with a higher clearance preference (e.g. a bus). If the lateral distance between the two vehicles is larger than the clearance preference of the first vehicle but less than that of the second vehicle, the latter will sense that it has collided and immediately halt. It is not completely clear from visual inspection of real-world disordered heterogeneous flow precisely how vehicles resolve this problem in practice. In reality it is likely that clearance requirements are somewhat elastic, and exist in a range dependent on the prevailing traffic situation. For the sake of simplicity, the model assumes the transverse clearance requirements to be the same for all vehicle classes and at all speeds.

3.2.4 Modification for lane-based traffic

The gap-filling model that has been presented in this chapter is designed to replicate the behaviours of a heterogeneous mix of vehicles which are not practising lane discipline.

However, lane discipline is not a single process or activity and instead encompasses a range of activities from respecting lane boundaries to practising lane courtesy (such as moving to a 'slow' lane when not overtaking). However, there are two basic elements of lane discipline which are notably absent in disordered heterogeneous traffic. These are

- never occupying the same longitudinal position as another vehicle in the same lane, and
- minimising time spent on lane boundary markers.

In order to accommodate lane discipline within the context of the gap-filling model, each of these two elements of lane discipline are addressed separately. With respect to never having two vehicles side-by-side in the same lane, a modification is made to the TTC prediction algorithm. When computing the time to collision between an agent and another agent in front of the first, the width of the second agent is artificially adjusted so that it occupies the full width of the lane, as shown in Figure 3.5. This still allows conventional overtaking to occur, as demonstrated in image (b), but prevents seepage into the gaps between vehicles.

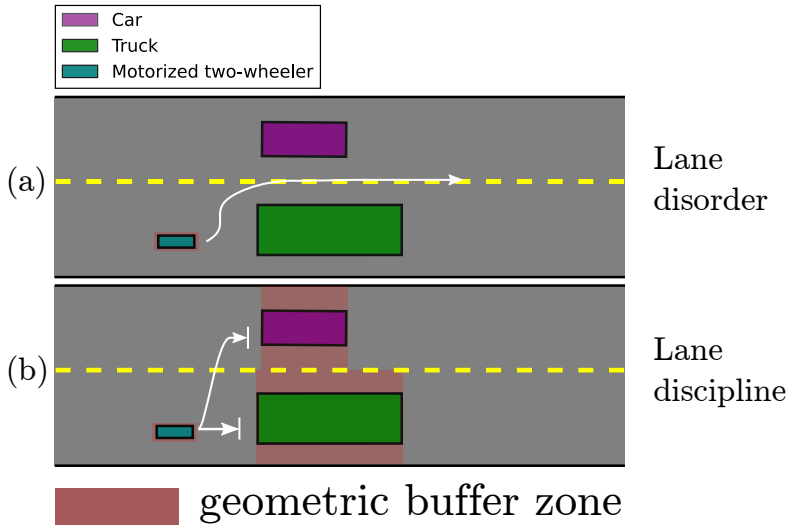

FIGURE 3.5: The use of artificial geometric buffer zones around agents allows lane-based behaviour to be incorporated into the gap-filling model by eliminating lateral gaps between vehicles.

In order to implement the second element of lane discipline identified above, the velocity selection algorithm of each agent is modified to select velocities that move the agent towards the centre of its current lane. So as not to conflict with agents' attempts to change lanes, this centre-seeking behaviour is only invoked when the agent's forward motion is not hindered by any other agents.

3.3 Initial and Boundary Conditions

Having given the dynamic details of the gap-filling model in section 3.2, this section will review the initial and boundary conditions that have been used together with an implementation of the model to simulate disordered heterogeneous traffic flow.

For initial placement of vehicles onto a simulated roadway, it is possible to either use a predefined arrangement by specifying the type and position of each vehicle, or to use a 'warm-up' stretch of road. The latter approach has been adopted for the simulations that were performed for this report. This warm-up stretch is located before the stretch of road under study. Under this arrangement, the centroid of each of the n agents in the simulation are initialised to XY-coordinates given by the equations

$$X_i = \begin{cases} \sum_{k=1}^{i-1}(X_k - \frac{l_{i-1}}{2} - \frac{l_i}{2} - b_X) & : 2 \leq i \leq n \\ -\frac{l_1}{2} & : i = 1 \end{cases} \tag{3.5}$$

$$Y_i = \mathrm{u}(0 + \frac{w_i}{2} + b_Y, \text{road width} - \frac{w_i}{2} - b_Y) \tag{3.6}$$

where l_i, w_i are the agent's length and width (in the direction of the road's length and width, respectively), u(a, b) is a uniform random number in the range $[a, b]$ and b_X, b_Y are defined buffers in the X and Y-directions, respectively.

In the present proof-of-concept implementation of the model for the purpose of studying its applicability, the boundary conditions for the left- and right-hand sides of the stretch of road under study have been made periodic. That is, vehicles exiting the right-hand side of the road stretch re-enter on the left-hand side. The advantage of this approach is that it permits a smaller stretch of road to be used, which makes for easier analysis and reduced computational burden due to the lower numbers of agents being simulated at any one time. This is appropriate for an initial proof-of-concept implementation. However, if the goal were to simulate a highly specific scenario, such as vehicle arrival times being distributed according to a particular probability distribution, then periodic boundary conditions would not be appropriate.

In addition to the free-flow scenario described in the previous paragraph, queueing behaviour forms an important part of disordered heterogeneous traffic flow. The present implementation of the gap-filling model makes use of a *variable-period stop*: a fictitious traffic control element which acts as a stop sign until all vehicles have come to a halt, then temporarily disappears until all vehicles have passed, at which point it reappears as a stop. The time between subsequent reappearances is not necessarily constant as

is the case for a simple traffic light, but rather dependent on the time taken for queue formation and settling to occur, hence the *variable-period* descriptor. The advantage of using this control element is that it readily allows studying of queue behaviour when used in conjunction with periodic road boundaries, because the need to use a longer road with multiple stops is obviated, thus reducing computational burden because of the lower number of agents. Its variable-period characteristic means that temporal evolution of various queue properties can be studied, for example whether a group of vehicles reaches a steady-state ordering after having undergone queue formation multiple times.

The stopping requirement of the variable-period stop is implemented in the model by decreasing the free speed of vehicles as they approach the stop. This is done so as not to exceed the maximum permissible deceleration of each vehicle. Therefore, when calculating the range of speeds that a vehicle may follow at each time step, the speed is capped at the maximum speed that will allow deceleration to zero velocity by the time the vehicle reaches the stop. From equations of motion, if a vehicle is a distance s from a stop and has a minimum acceleration of a_{min} (which is negative), then its speed is capped at

$$v_{t,max} = \sqrt{-2a_{min}s}.$$

Chapter 4

Model Evaluation

Having presented in the previous chapter a model for the simulation of disordered heterogeneous traffic flow, this chapter provides an evaluation of the model. This will be presented in three parts. Section 4.1 validates the model by comparing aspects of its behaviour to functional phenomena observed in real disordered heterogeneous traffic, followed by section 4.2, which evaluates the model's behaviour when adapted to simulate lane-based traffic. Finally, section 4.3 presents a brief sensitivity analysis for selecting the model's collision avoidance parameter.

4.1 Model Validation

Traditionally, traffic models have been validated by their ability to approximate empirically-observed macroscopic data, or their ability to reproduce functional phenomena observed in actual traffic flow. Owing to the heterogeneity of vehicle mix and disordered driving behaviour typical of developing world traffic, adaptation of conventional traffic monitoring techniques, such as loop detectors, to the developing world context has proved problematic. Because of this, there have been recent advances in using videographic recording together with sophisticated frame analysis software to gain accurate macroscopic data on developing world traffic flow [37–39]. However, the field of automated videographic analysis of disordered heterogeneous traffic is too nascent for extensive structured data on such traffic flow to have been published. Owing to this dearth of structured data, the approach adopted in this paper is that of functional validation: to show that the present model is able to reproduce the functional phenomena of percolation and seepage, as well as to verify its behaviour through the use of conventional speed-density fundamental diagrams.

Model validation proceeds with an overview of the test data and test configurations in sections 4.1.1 and 4.1.2, followed by a brief visual inspection of the model output in section 4.1.3. Detailed functional analyses of queue percolation, seepage dynamics and speed-density relations follows in sections 4.1.4, 4.1.5 and 4.1.6, respectively.

4.1.1 Test Data

The guiding principle behind the development of the model is that local interactions between vehicles are responsible for the emergent traffic phenomena of seepage and percolation found in disordered heterogeneous traffic. However, the local nature of these interactions, and their strong dependence on the precise geometry of the gaps between vehicles, entails a relatively high degree of geometric and kinematic fidelity in order to ensure that these dynamic interactions are reflective of actual disordered heterogeneous traffic flow. For this reason, accurate vehicle geometries, free speeds, acceleration characteristics and, where applicable, traffic compositions have been obtained from literature and used for all the simulations that follow.

The present simulation study has been limited to six classes of vehicle, all of them motorised: bus, truck, light commercial vehicle (LCV), passenger car, auto-rickshaw (ARS), and motorised two-wheeler (MTW). This simplification was made on pragmatic grounds and does not affect the validity of the gap-filling model: the model applies to any object whose turning and acceleration characteristics can be specified.

Table 4.1 shows the six classes of vehicles just mentioned, along with their physical dimensions and free speeds [1]. The *free speed* of a vehicle is the speed to which it will accelerate and preferentially remain in the absence of any need to reduce its speed. In the present model it is assumed that a vehicle never exceeds its free speed, although in reality this may temporarily occur, for example, when overtaking another vehicle. The exact free speed of each vehicle in a class is drawn from a truncated normal distribution with parameters as specified in the table.

TABLE 4.1: Vehicle geometry and free speed characteristics

Vehicle class	Length (m)	Width (m)	Free speed (km/h)			
			Mean	Std.dev	Min	Max
Bus	10.3	2.5	53.01	7.2	38.61	67.41
Truck	7.2	2.5	51.50	6.6	38.30	64.70
LCV	5.0	1.9	50.30	7.7	34.90	65.70
Car	4.2	1.7	58.90	14.3	30.30	87.50
ARS	2.6	1.4	44.90	7.7	29.50	60.30
MTW	1.8	0.6	45.05	12.4	20.25	69.85

The maximum acceleration and deceleration values of the six vehicle classes are given in Table 4.2 [1, 40]. Because of various physical and mechanical factors such as engine output, gearbox design and air resistance, maximum acceleration is dependent on vehicle speed. Maximum deceleration is also dependent on a number of factors, most notably road surface conditions. The present model assumes a speed-invariant maximum deceleration.

The free speed of each vehicle is implemented by means of the acceleration function. At speeds below an individual vehicle's free speed, its maximum acceleration is as specified in Table 4.2. At speeds equal to or above its free speed, the vehicle's maximum acceleration is zero. Additionally, the vehicle's minimum speed is also set to zero so that vehicles cannot reverse on the road.

TABLE 4.2: Maximum vehicle acceleration and deceleration

Vehicle class	Acceleration (m/s^2)			Maximum deceleration (m/s^2)
	0 - 20 km/h	20 - 40 km/h	40+ km/h	
Bus	0.89	0.45	0.33	0.88
Truck	0.79	0.45	0.33	0.88
LCV	0.82	0.60	0.35	1.71
Car	1.50	1.30	1.00	1.71
ARS	1.01	0.58	0.34	1.16
MTW	1.35	1.03	0.37	1.59

Based on the dimensions of the vehicles classes, it is possible to establish a size ordering. For the dimensions given above, this ordering happens to be invariant whether any one of vehicle length, width, or area is used. The ordering in descending size is: bus, truck, LCV, car, ARS and MTW. A convenient grouping of these vehicle classes, based on size and dynamic characteristics, that will be used subsequently in this report is buses and trucks as *large* vehicles, LCV and cars as *medium* vehicles, and ARS and MTW as *small* vehicles.

The relative prevalence of each vehicle class that is typical of an urban developing world environment (in this case, India) is given in Table 4.3 [1]. These frequency values are used to compose representative samples of urban developing world traffic.

Unless otherwise specified, all simulations have been performed with a timestep $\Delta t = \frac{1}{4}$ seconds.

TABLE 4.3: Typical vehicle mix composition

Vehicle class	Fraction of vehicles (%)
Bus	5.2
Truck	2.6
LCV	3.1
Car	26.2
ARS	10.5
MTW	52.4

4.1.2 Test Configurations

The test configurations used in this simulation study are shown in Figure 4.1. Periodic boundaries are used for all the simulations in this report. Vehicles are initialised to left of the first periodic boundary (the left-hand side of the road surface as shown in the figure). In the freeflow scenario, vehicles simply continue travelling along the roadway. reappearing on the left-hand boundary after crossing the right-hand boundary. The other scenario employs a variable-period stop which was described on page 41. Frame (b.1) shows vehicles having approached a variable-period stop, which is coloured red. Once all vehicles are stationary to within a certain tolerance, the stop turns off and vehicles begin to move: frame (b.2). Once all vehicles have passed the position of the stop, it reappears: frame (b.3). Vehicles then queue at the stop as before: frame (b.4).

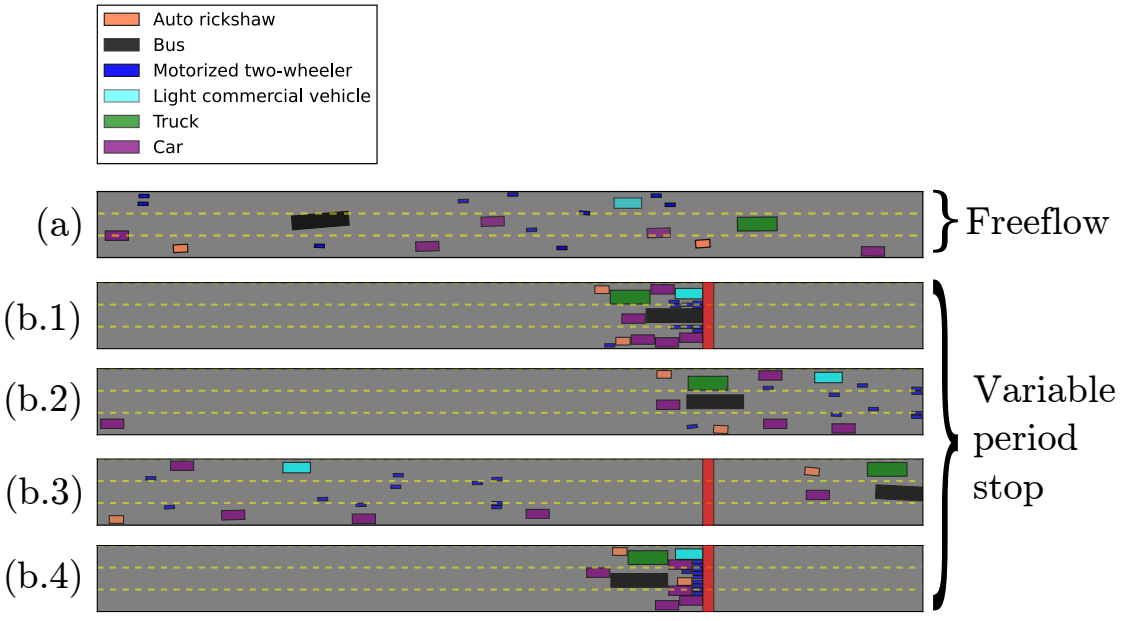

FIGURE 4.1: The two test configurations used in this simulation study: freeflow and variable-period stop. See the text for a full explanation.

In addition to the two test configurations just described, two different starting conditions

are typically used in this simulation study: one where vehicles begin in a largest-to-smallest ordering, the other where vehicles begin in a smallest-to-largest order (using the packing method described in section 3.3). These two starting scenarios are viewed as opposing ends of a spectrum of all possible starting positions. In the former scenario, smaller vehicles would need to weave in between larger vehicles in order to get to the front whereas, in the second scenario, smaller vehicles are already in front of larger vehicles and therefore do not need to 'percolate' in order to reach the front of a queue. Note that, in their starting positions, no two vehicles occupy the same longitudinal position on the roadway (with a buffer space between the front and rear bumpers of adjacent vehicles). Initial tranverse positions on the roadway are chosen so that each vehicle begins in the middle of one of the available lanes (this initial condition is enforced whether vehicles follow lane discipline or not).

4.1.3 Visual Inspection

Perhaps the best way to analyse the functional validity of the model is by direct inspection of the visualisation output. To this end the reader is referred to the video demonstration contained in the electronic supplement to this thesis, the details of which are given on page 5.

Figure 4.2 shows a number of frames from the visualisation of a single simulation run. This simulation was performed using a variable-period stop (VPS) and largest-to-smallest starting positions. Note the presence of larger vehicles at the front of the traffic stream in frame (1). The car (coloured purple) at the very front of the stream has reached that position by accelerating past the bus, which started ahead of the car. Frame (2) shows the first vehicles having reached the VPS and decelerated to become stationary. Frame (3) shows more vehicles approaching the stop and, where possible, percolating to the front. Frame (4) shows the percolation more fully developed, with most MTWs having moved to the front of the queue and halted, with the exception of the single MTW that was lagging behind the stream in frame (3), due to it having been randomly assigned a lower free speed. Frame (5) shows the final position of vehicles at the stop, with the lagging MTW having moved to near the front of the queue. Finally, Frame (6) shows the position of the vehicles at the end of the third stop. Note how smaller vehicles have now nearly completely percolated to the front of the queue, and how the queue is in fact longer after the third stop than after the first. This will be discussed later.

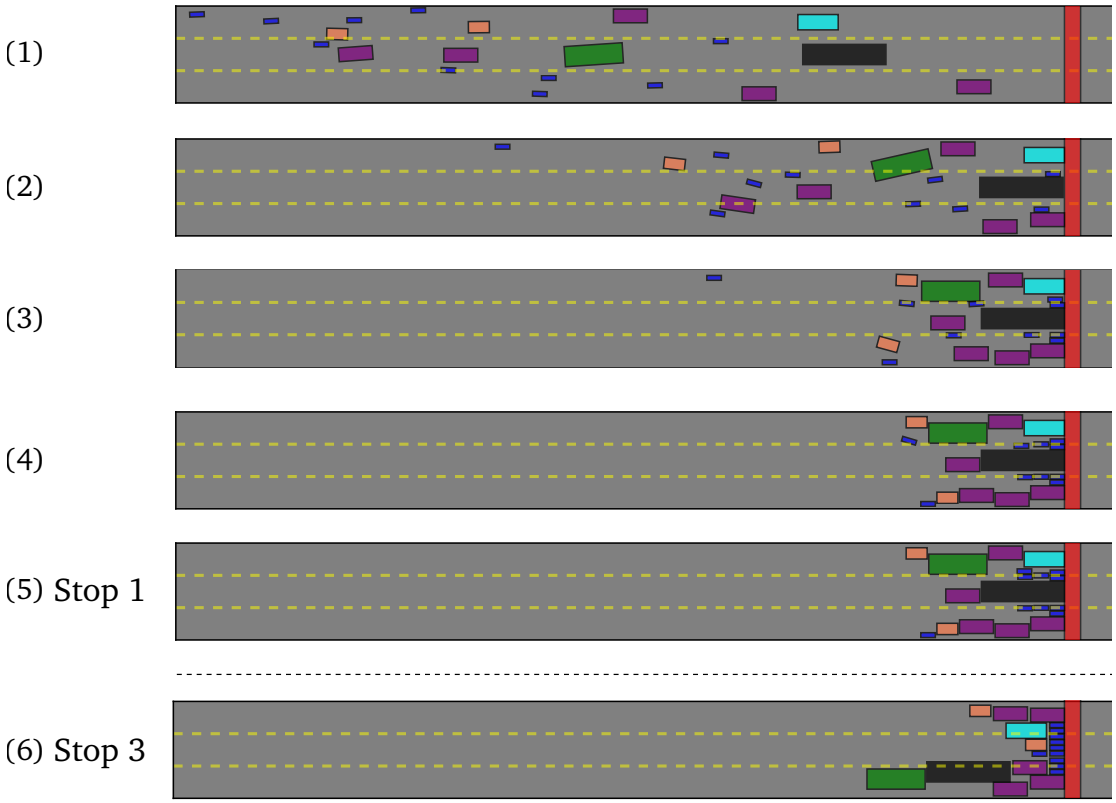

FIGURE 4.2: Frames from a single simulation run showing the formation of queues, in which smaller vehicles percolate to the front.

4.1.4 Queue Percolation

While the previous section provided visual insight into the dynamic behaviour of vehicles in the model, it represents the output of only a single simulation. This section, and the two that follow, will present a statistical analysis of the aggregate output of many simulations in order to ascertain the model's general behaviour. One of the goals of the model is to reproduce the functional phenomena of queue percolation and seepage. In this section we consider the former.

Percolation of smaller vehicles to the front of queues is inherently a dynamic phenomenon: it is the *movement* of smaller vehicles through the gaps between larger vehicles. Percolation, therefore, cannot be assessed by a single snapshot of vehicle positions once they have reached a stop. For example, if larger vehicles were much faster than smaller vehicles they would generally reach the stops before smaller vehicles, and percolation, while possibly still occurring, would not be evident from the vehicles' final stationary positions. It is therefore necessary to consider the change in vehicles' positions from the time before any vehicles reached the stop, until the time when all vehicles have become stationary.

Figures 4.3 and 4.4 show the mean positional rank of each vehicle class at the start of each simulation, and at the first four stops, for starting scenarios of largest-to-smallest (Figure 4.3) and smallest-to-largest (Figure 4.4). Positional rank is measured by the relative location of the front bumper of each vehicle, scanning from right to left (i.e. positional rank decreases in the vehicles' preferred direction of motion)[1]. Each figure represents the average of 50 simulation runs. For the sake of clarity, the thickness of each line is correlated with the size of the vehicle class it represents. Each simulation consists of 20 vehicles distributed approximately according to the frequencies given in Table 4.3, on a road of width 150m. Thus, vehicle ranks can range from 1, for the vehicle at the front, to 20, for the rearmost vehicle, with ties (such as at stops) being resolved in the usual fashion.

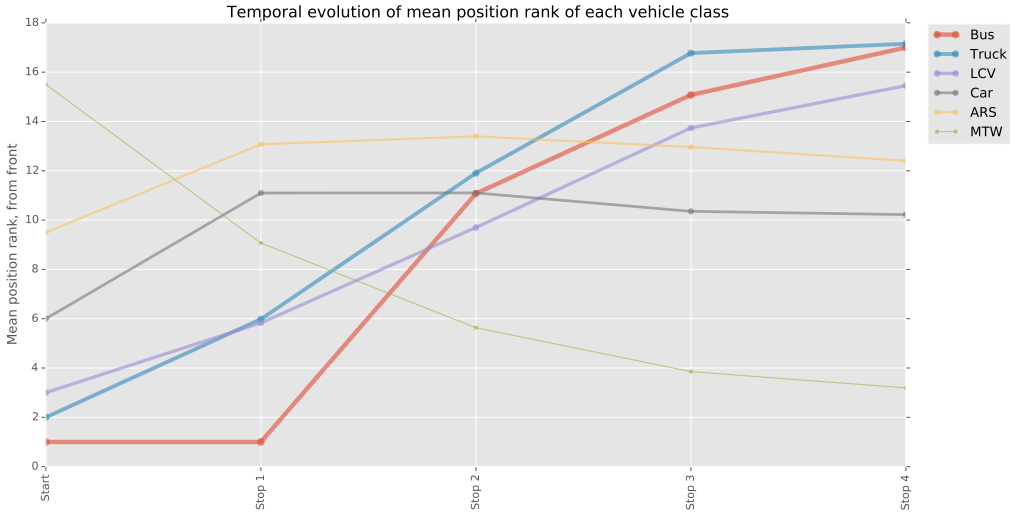

FIGURE 4.3: Evolution of mean positional rank of each vehicle class at each stop from simulation start to the fourth stop, with *largest-to-smallest* inital ordering of vehicles. See the text for the simulation parameters.

With reference to the first figure, the largest-to-smallest ordering of the vehicles means that larger vehicles have a lower rank than smaller vehicles at the start of the simulation. If percolation of smaller vehicles were indeed occurring, we would expect the positional rank of smaller vehicles to decrease as they move to the front, at the expense of larger vehicles which are displaced to more rearward positions. Figure 4.3 shows this to indeed be the case. By the first stop it is already evident that, on average, MTWs are improving their average rank, and by the second stop they are on average the vehicle class with the lowest mean rank. By the fourth stop vehicles are on average ordered in near-perfect inverse correlation to their size. The two clear exceptions are the consistently higher rank of trucks relative to buses, and that of ARS to cars, despite the former

[1]Here the magnitude of a vehicle's rank is referred to in its strictly numerical sense. Thus, a higher rank refers to a vehicle being further back in the traffic stream.

vehicles' smaller dimensions in comparison to the latter, in each case. This behaviour can be attributed to the former vehicles' lower accelerations (refer to Table 4.2) relative to the latter. Because of this, trucks and ARS are passed by the faster-accelerating buses and cars, respectively, and thus find themselves at more rearward positions in the queue. It is therefore the case that a vehicle's propensity to reach the front of queues depends on both its *size* (smaller vehicle results in more gaps that can be filled), and its *manoeuvrability* (which allows it to pass other vehicles).

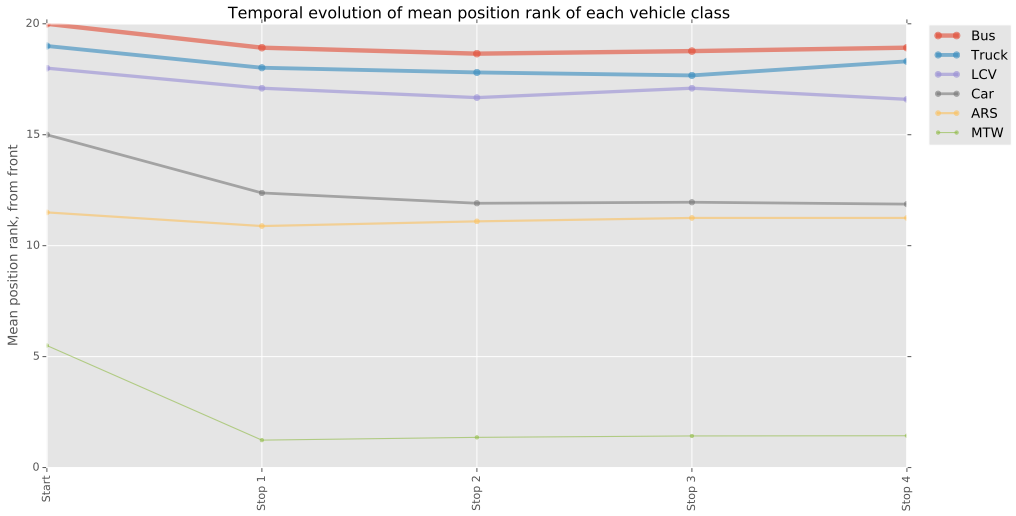

FIGURE 4.4: Evolution of mean positional rank of each vehicle class at each stop from simulation start to the fourth stop, with *smallest-to-largest* inital ordering of vehicles. See the text for the simulation parameters.

Having shown that a largest-to-smallest starting scenario results in percolation of smaller vehicles, we now consider a smallest-to-largest starting scenario to show that it is indeed smaller vehicles' size and manoeuvrability that allows them to advance to the front of queues, and to exclude the possibility that there is some other mechanism that results in vehicles which start at the back initially reaching the front of queues. Figure 4.4 shows the evolution of the mean positional rank of each vehicle class for the smallest-to-largest starting scenario. As can be seen, there is no change in the mean relative orderings of the vehicle classes from the initial starting configuration to each subsequent stop. The initial drop in each of the six curves from 'Start' to 'Stop 1' is due to vehicles which were initially one behind the other catching up alongside other vehicles and occupying the same longitudinal position on the road at the stop. For example, frame (6) of Figure 4.2 shows a number of MTWs at the front of the queue which all share the same rank of 1.

Considered together, Figures 4.3 and 4.4 establish the existence of an equilibrium ordering of queued vehicles in disordered heterogeneous traffic flow. This ordering consists of generally smaller and more manoeuvrable vehicles at the front of queues and larger, less

manoeuvrable vehicles at the back. The first figure shows that if this ordering is not the case, the traffic stream will approach this ordering in only a few stops (by the process of repeated percolation). The second figure shows that when this ordering is attained, it perpetuates itself. This behaviour matches precisely the queueing behaviour seen in real disordered heterogeneous traffic and the model has therefore been validated in this respect.

We now examine an aspect of interest to transportation planners and consider whether the formation of this equilibrium ordering improves queueing efficiency in some way. It might be expected that as queues move from a non-equilibrium ordering and 'settle' into an equilibrium ordering over the course of multiple stops, that queue packing becomes more efficient as vehicles repeatedly fill more of the remaining unused road space. Two metrics that can be used to describe the packing efficiency of a queue are the queue's length and its areal density. For a given sum of vehicle areas a_v, queue areal density ρ is $\rho = \frac{a_v}{l_q \times w_r}$ where l_q and w_r are queue length and road width, respectively. Areal density and queue length are thus inversely proportional for a given road width and set of vehicles, so we consider queue length as it is a more intuitive metric. In this case, a more efficiently-packed queue would correspond to lower queue length.

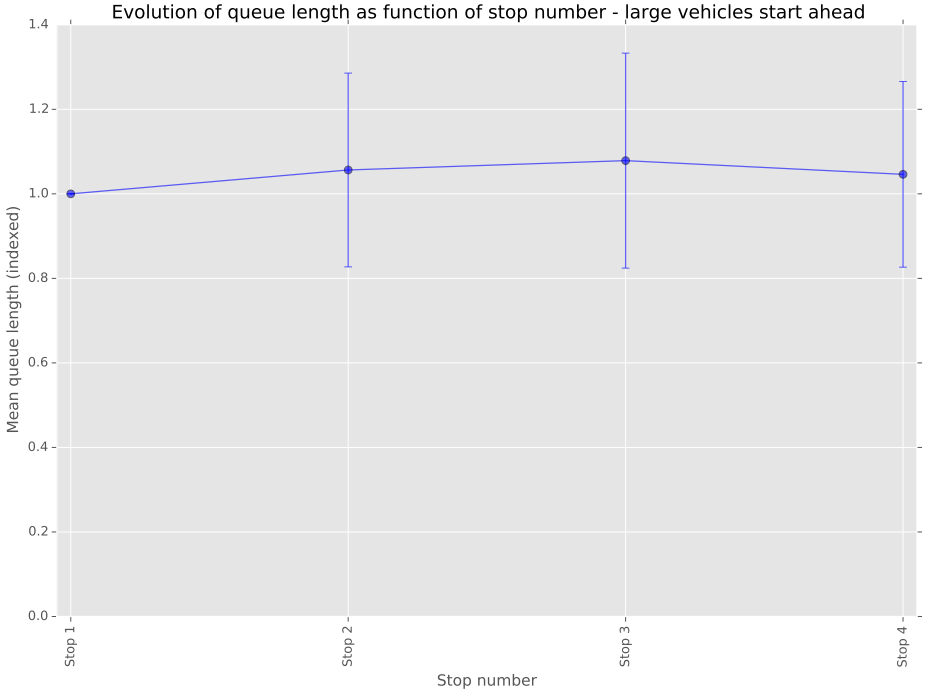

FIGURE 4.5: Evolution of queue length in largest-to-smallest starting scenario. Values are indexed by the length of the first queue in each simulation, with error bars showing standard deviation between results for multiple simulation runs.

Figure 4.5 shows the average length of queue at each of the first four stops for a largest-to-smallest starting scenario, and the same simulation parameters that are described on

page 49 and were used for Figures 4.3 and 4.4. The queue length for each simulation is indexed to that of the first queue, and the figure thus shows how the queue length changes with each stop (note that there can be no queue length prior to the first stop as vehicles have not yet queued). From the figure, queue length appears to, on average, increase after the first stop and taper off to a fixed ratio of the length of the first stop. This is opposite to the expectation that queue length would decrease as vehicles sort into their equilibrium ordering.

The reason for this increase in queue size is a combination of the relationship between vehicle size, acceleration and free speed. If a group of smaller vehicles approaches a stop that is already populated by mostly larger vehicles, the length of the resulting queue would be less than the queue resulting from a scenario where a group of larger vehicles approaches a stop which is populated by smaller vehicles. This is because smaller vehicles will fill the gaps between larger vehicles and thus improve their position in the queue, something that larger vehicles cannot do. An analogous physical example is pouring pebbles into a jar, followed by sand. The sand fills the gaps between the pebbles and the resulting height of sand and pebbles is less than if the sand were poured first, followed by the pebbles.

Once smaller vehicles have moved through gaps towards the front of a queue that was populated by larger vehicles, two scenarios can occur: if the smaller vehicles are able to reach the next stop before the larger vehicles, the resulting queue will be longer than in the scenario where the larger vehicles reach the next stop before the smaller vehicles. For the particular vehicle accelerations, free speeds and the distance between subsequent stops used in these simulations (150m), the former is the case. The distance between stops is short enough that smaller vehicles with their higher accelerations and greater manoeuvrability (although lower free speeds) are able to reach the next stop before larger vehicles. To verify this, Figure 4.6 shows the indexed queue lengths for a smallest-to-largest starting scenario, with the same result.

4.1.5 Seepage Dynamics

Besides queue percolation, the other functional phenomenon that is typical of disordered heterogeneous flow is seepage, where vehicles do not pass other vehicles by moving into a separate lane, but rather pass by continuously weaving between vehicles. The reproduction of seepage in the model is best appreciated by visualising simulation output. For this, the reader is again referred to the electronic supplement.

This section will provide a statistical analysis of seepage in the model. Since this is an understudied phenomenon, there exist no standard metrics for quantification of the

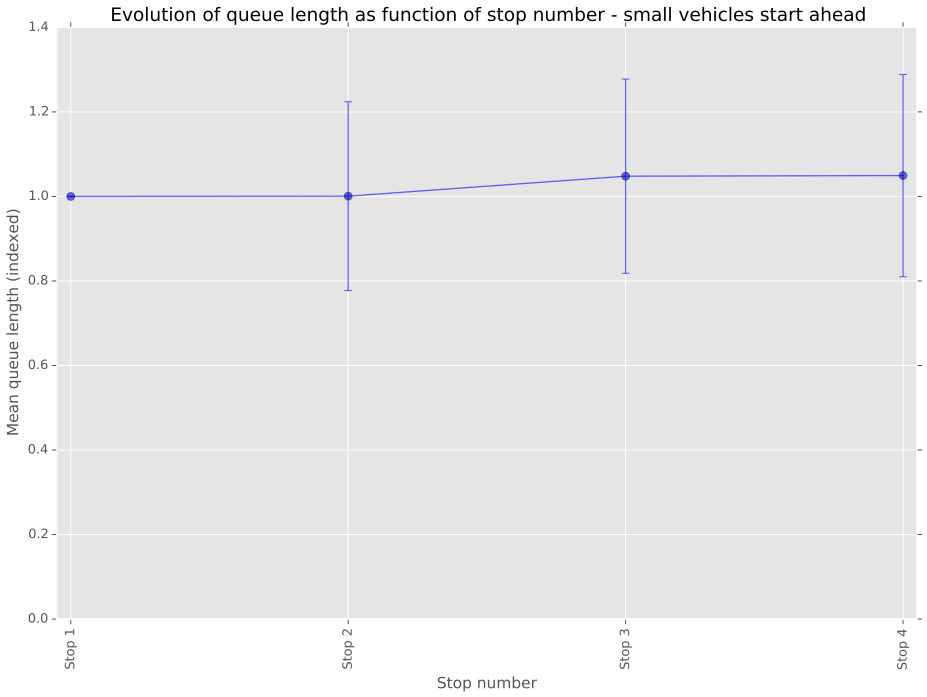

FIGURE 4.6: Evolution of queue length in smallest-to-largest starting scenario. Values are indexed by the length of the first queue in each simulation, with error bars showing standard deviation between results for multiple simulation runs.

degree of seepage that occurs in a traffic stream. We proceed by quantifying seepage in terms of *weaving* – that is, the distance a vehicle travels transverse to the road axis, per unit of distance travelled in the longitudinal direction. This yields a dimensionless, positive number with higher values indicating a greater degree of weaving. The presence of seepage in the model will be quantified by converting this ratio to an equivalent number of lane changes per unit distance, and comparing this to values for conventional developed world traffic, obtained from literature.

Figure 4.7 shows the degree of weaving for each vehicle class in both starting scenarios. As can be seen, the degree of weaving performed by vehicles in the gap-filling model is several times higher than that which is typical of urban developed-world traffic. Even though the first 15% (measured by time) of data is discarded for each simulation, there is a clear difference in the amount of weaving performed by each vehicle class depending on its starting position. In general, for each vehicle class there appears to be more weaving when that vehicle class starts nearer the back than compared to when it starts nearer the front. Regardless of starting scenario there appears to be a general trend for smaller vehicles to weave more than larger vehicles. While MTWs and ARS show lower *mean* levels of weaving than larger vehicles, the *variance* in these values is higher. This is because these smaller vehicles will generally be near the front of the traffic stream as was shown earlier, and therefore need to do less passing. The mean measurement of

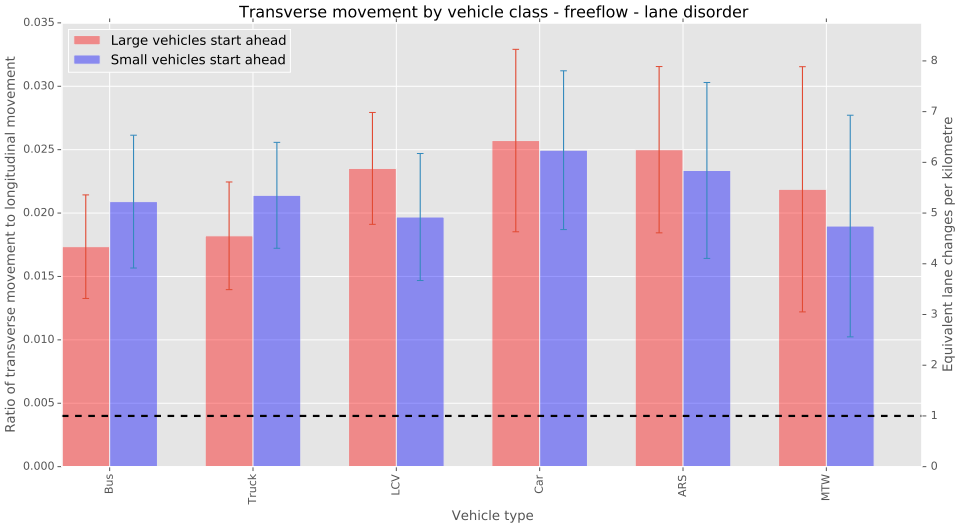

FIGURE 4.7: The degree of weaving (ratio of transverse to longitudinal movement) for each vehicle class in freeflow simulations of 20 agents distributed according to Table 4.3, on a 3-lane road of width 12m and length 150m with periodic boundaries. Each value is the average of 50 simulations, with error bars indicating the standard deviation between all vehicles aggregated within each class, aggregated from separate simulation runs. To eliminate transient effects from starting conditions, the first 15% of each simulation is discarded. The right-hand vertical axis is a rescaling of the primary axis in units of lane changes per kilometre, based on a lane width of 4m. The horizontal broken line indicates an approximate typical number of lane changes per kilometre for urban developed-world traffic [41].

their weaving is therefore reduced because it includes those vehicles which have no other vehicles to pass.

Finally, we can also examine the degree of weaving that occurs in between stops in a multiple stop scenario. Figure 4.8 shows, for each vehicle class, the average[2] degree of weaving performed by each vehicle in each interval from the beginning of the simulation to the time all vehicles have reached the fourth stop. Both starting scenarios are shown.

Examining plot (a), it is apparent that smaller vehicles which start at the back need to do more weaving in order to reach the front. This amount of weaving decreases as they move to the front of the stream. Conversely, larger vehicles which are initially in front do very little weaving, whereas as they fall to the back of the queue they need to weave more in order to avoid collisions with vehicles in front of them. This confirms one of the conclusions drawn from Figure 4.7, that vehicles nearer the rear of the stream tend to weave more than those nearer the front. Plot (b) shows the case when smaller vehicles start in front. In this case MTWs initially perform a small degree of weaving as they pass one another in order to fill the available lateral space on the road, after which their

[2]Each line represents the temporal average of the degree of weaving in the given interval for each vehicle, averaged across all vehicles in the class. These lines are then averaged across multiple simulation runs to yield the lines shown in the figure.

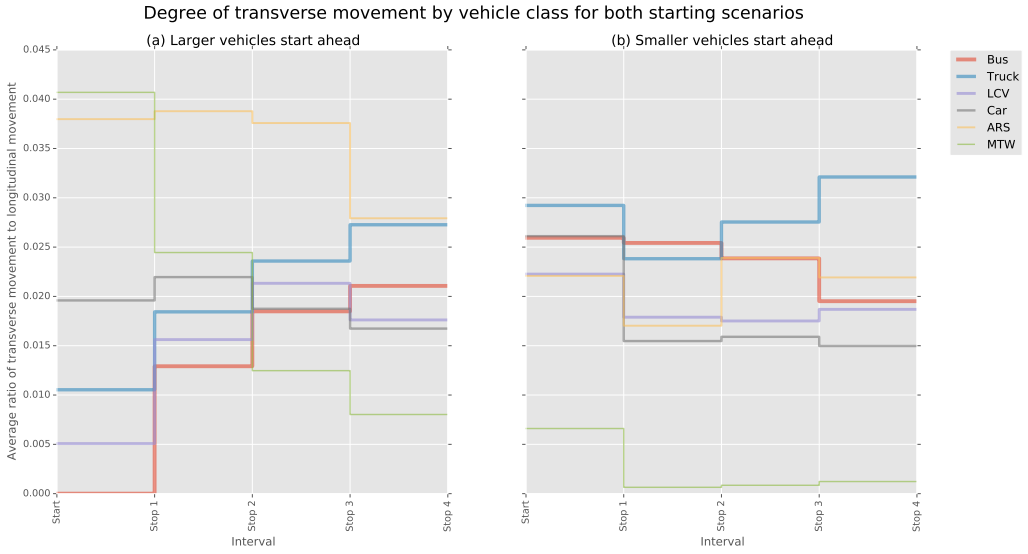

FIGURE 4.8: Degree of transverse movement by vehicle class in variable-period stop scenario, for both starting configurations. Data is obtained from the same simulations as Figures 4.3 and 4.4.

average degree of weaving is close to zero. Conversely, larger vehicles behind the MTWs are forced to do more weaving in order to avoid collisions with the smaller vehicles in front of them.

It can therefore be concluded that the model is valid with respect to its reproduction of seepage dynamics, as vehicles undergo significantly more lateral movement than is typical of conventional developed world traffic, and the degree of lateral movement is dependent on the need to pass other vehicles, whether to avoid collisions or to fill gaps in queues.

4.1.6 Speed-density Relations

The final facet of model validation is to verify the functional form of the fundamental diagrams it produces. Due to the dearth of accurate combined macroscopic and microscopic data for actual developing world traffic, previous works proposing new models for such traffic have sought to reproduce the functional form of speed-density, density-flow and speed-flow relations [1, 14]. Exact calibration of models to these fundamental diagrams has proven problematic, due to the lack of data containing both microscopic parameters (free speeds, accelerations, measures of weaving propensity, traffic composition, and so on) and macroscopic data (speed, density and flow values). These data have typically been collection in isolation. Recent advances in software-based videographic traffic analysis appear promising for producing such combined datasets.

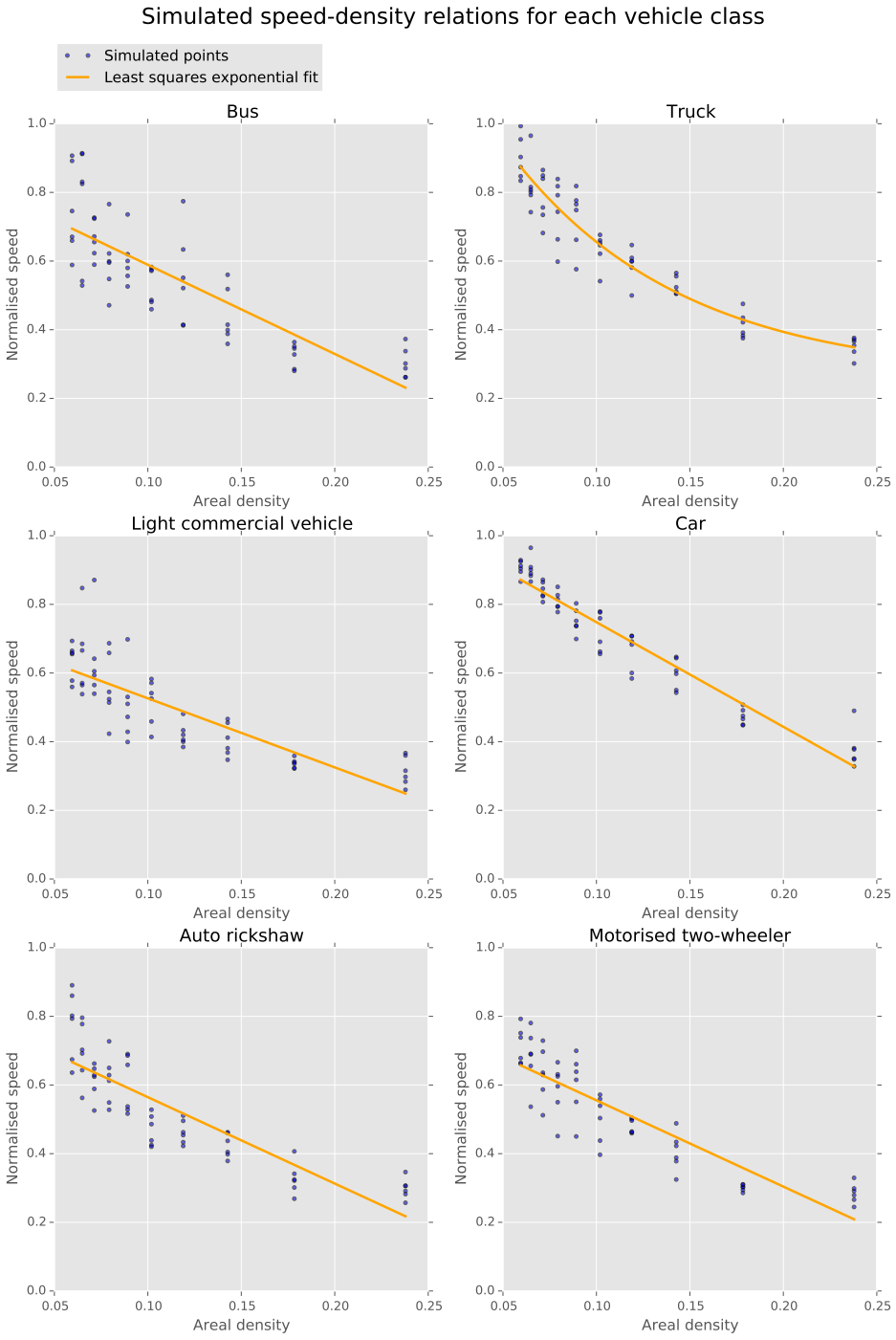

FIGURE 4.9: Simulated speed-density relations by vehicle class for a freeflow simulation of 20 vehicles with the same parameters as Figure 4.7, with a largest-to-smallest starting configuration. Varying areal vehicle densities were implemented by varying the road width. The first 15% of each simulation is discarded to reduce transient starting effects. Each point represents the output of a single simulation run.

In this section only speed-density relations will be shown as the other two fundamental diagrams can be derived from the fundamental speed-density-flow relation. Figure 4.9 shows simulated speed-density values for each vehicle class in a traffic stream of 20 vehicles with frequencies approximately equal to that of Table 4.3. To remove the inherent variance in the speeds of each vehicle (as vehicle speeds are drawn from a statistical distribution), each vehicle's speed is normalised according to its free speed. Thus a normalised speed of 1 indicates a vehicle travelling at full speed without hindrance. An exponential trendline of the form $y = ae^{bx} + c$ is added to each plot as it is able to fit straight-line data as well as the curvature of the data in some diagrams. Areal density is used in place of linear density, as is recommended in the literature for heterogeneous traffic flow [2, 14]. Density values are varied by altering the road width while keeping the number of vehicles constant: the effect of this will be discussed shortly. Road width was varied from a maximum of 12 metres to 2 metres, the minimum width that would permit the widest vehicle to pass.

As can be seen, for all vehicle classes speed decreases with increasing density. The functional form of the fundamental diagram is therefore reproduced by the model. Reversing the starting configuration does not affect the fundamental diagram, as shown in Figure 4.10 for a smallest-to-largest starting order.

Achieving varying vehicle densities by varying the road width introduces the problem of reducing the space available for overtaking. This is likely responsible for the flattening of the speed-density curves at higher densities: when road width is sufficiently small that vehicles in a certain class cannot overtake, decreasing road width any further will not effect the speed of that vehicle class. To investigate this effect, fundamental diagrams for a homogeneous mix of cars were produced. Figure 4.11 shows the speed-density relation for a constant number of cars when road width is varied (from 10m to 2m), whereas Figure 4.12 shows the same when road width is held constant (at 8m) and the number of cars is varied. By comparing the two figures it is clear that in the former case of variable road width there is a flattening out of speed values at higher densities, as compared to the constant road width case. Furthermore, the vertical broken line on Figure 4.12 shows the density at which the number of cars equals the maximum that could fit side-by-side along the roadway. It is apparent that at densities below this value, cars are able to travel at values very near to their free speeds, whereas at densities above this value a reduction in speed occurs due to reduced overtaking ability. Nevertheless, no matter which method for varying densities is used to produce these fundamental diagrams, it has been shown that the model is capable of reproducing the functional form of these well-established analytical tools.

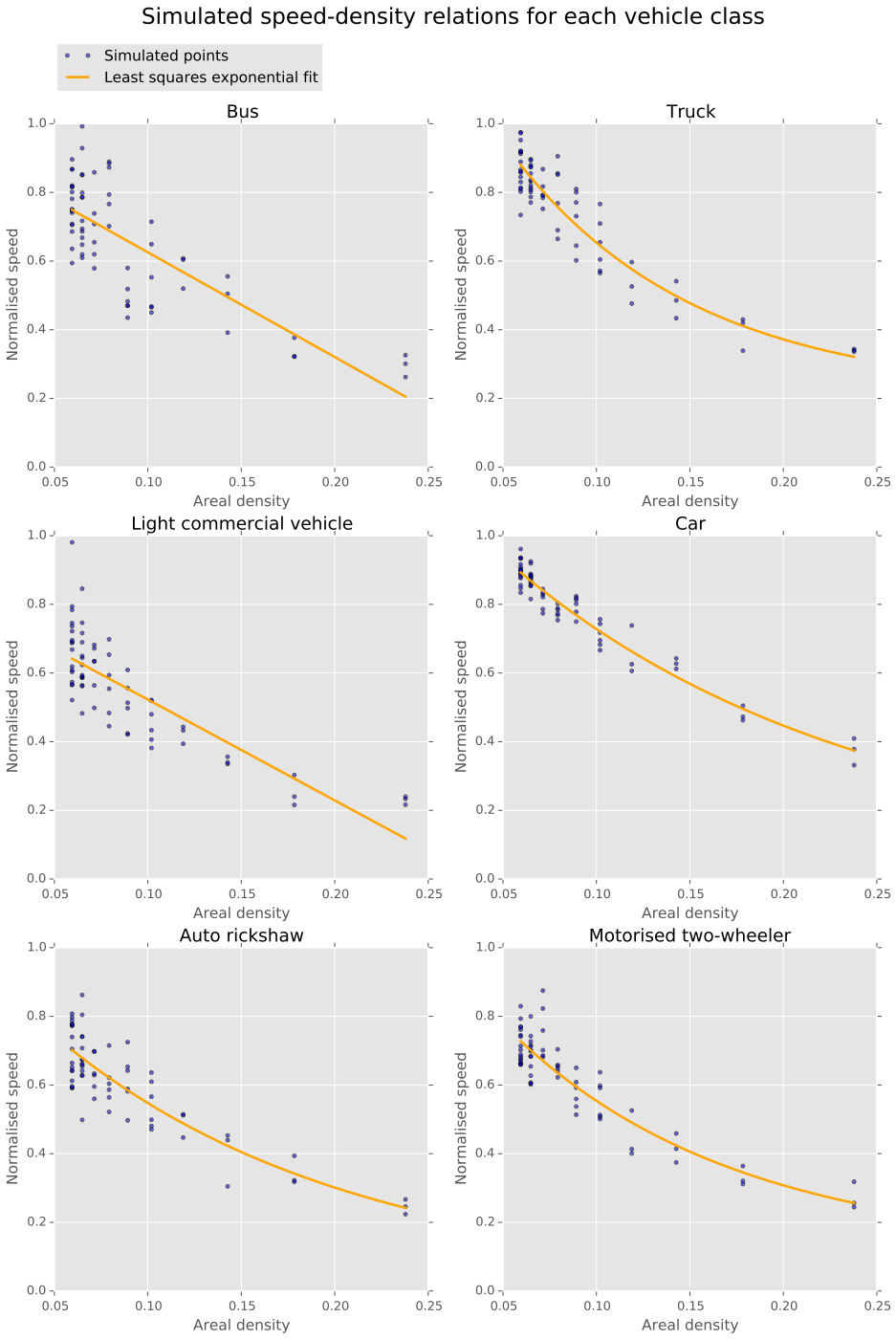

FIGURE 4.10: Simulated speed-density relations by vehicle class for a freeflow simulation of 20 vehicles with the same parameters as Figure 4.7, with a smallest-to-largest starting configuration. Varying areal vehicle densities were implemented by varying the road width. The first 15% of each simulation is discarded to reduce transient starting effects. Each point represents the output of a single simulation run.

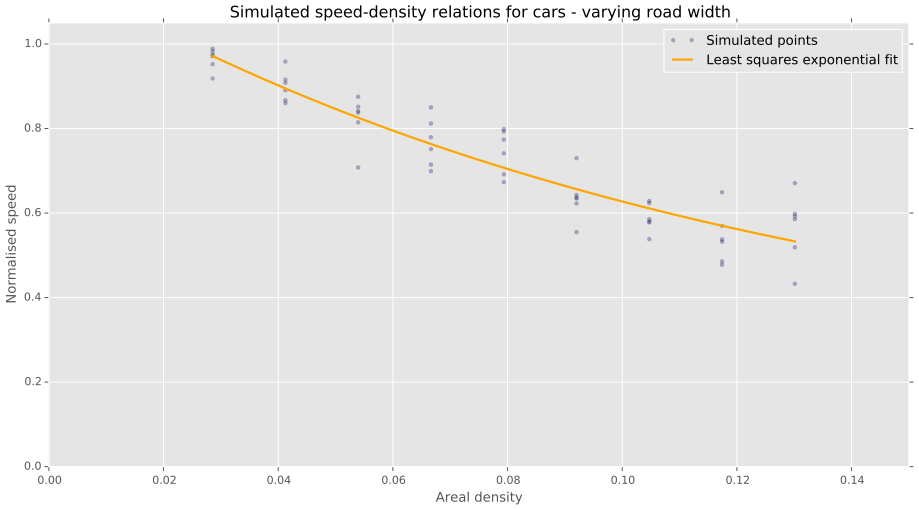

FIGURE 4.11: Speed-density relation of a homogeneous car mix where density is varied by varying road width.

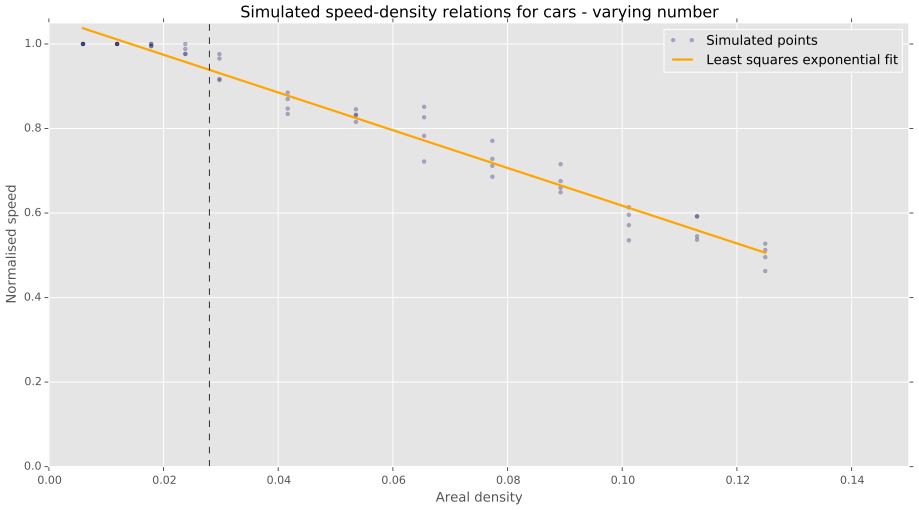

FIGURE 4.12: Speed-density relation of a homogeneous car mix where density is varied by varying the number of cars. The broken vertical line indicates the density at which the number of cars equals the maximum that can fit side-by-side on the road.

4.2 Lane Discipline

While the present model is designed for simulating disordered heterogeneous traffic, a minor modification described in section 3.2.4 allows it to incorporate the most pertinent aspects of lane discipline, for the purpose of comparison with lane disorder. This section will evaluate the characteristics of the simulation when all vehicles follow lane discipline.

Under the lane discipline assumptions, two vehicles in the same lane cannot have overlapping longitudinal coordinates, i.e. lane sharing/splitting is not permitted. This means

that, during queue formation, smaller vehicles cannot move through the gaps between larger vehicles if it would entail lane splitting. Figure 4.13 shows the evolution of a queue simulation where all vehicles follow lane discipline, beginning with a largest-to-smallest starting configuration. As can be seen, vehicles stack one behind the other, rather than in the more packed arrangement of Figure 4.2. Because vehicles following lane discipline are not able to stop alongside each other, it might be expected that lane discipline results in longer queues.

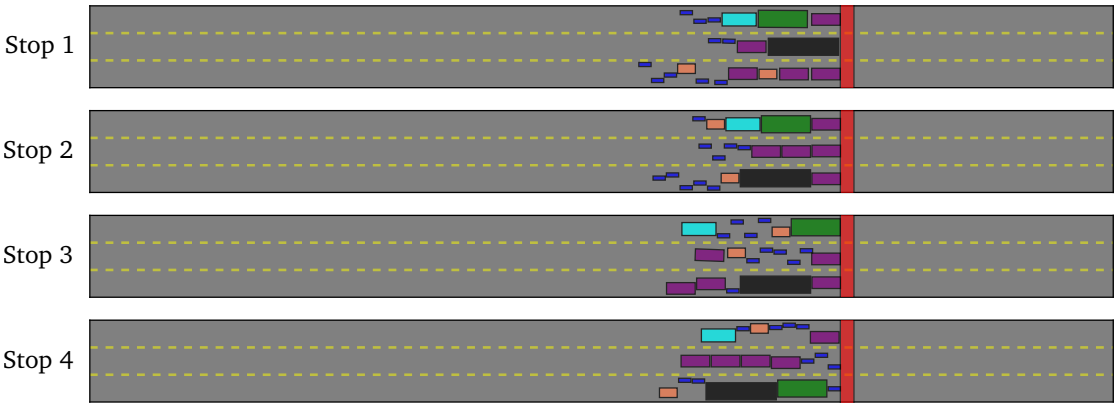

FIGURE 4.13: An example of queue formation when all vehicles follow lane discipline, starting from a largest-to-smallest configuration.

Figure 4.14 shows the queue length for the same mix of 20 vehicles (with type frequencies approximating Table 4.3), as depicted in Figure 4.13, for the cases where (i) no vehicles follow lane discipline and (ii) all vehicles follow lane discipline. As can be seen, lane-disciplined queues are significantly longer than queues with lane disorder, as would be expected. The magnitude of this difference would increase as the number of available lanes decreases. It is also apparent that the average length of lane-disciplined queues increases somewhat with the number of stops. This is believed to be a magnified version of the increase that was previously noted for lane-disordered queues: as larger vehicles get displaced to the back of queues, their greater size increases the queue length as there are no smaller vehicles to fill the gaps around the larger vehicles. The characteristic of lane discipline increasing queue length, due to vehicles queueing one behind the other, magnifies the effect.

We now seek to quantify the degree of lateral movement that occurs in the model when vehicles follow lane discipline. Figure 4.15 shows the degree of transverse movement by vehicle class. As can be seen by comparison with Figure 4.7, there is significantly more transverse movement under the lane discipline mode than when lane disorder prevails. When vehicles are not forced to stay in discrete, defined lanes they are free to adjust their lateral positions only so much as is required to avoid collisions. However when lane discipline is followed, a vehicle wishing to pass another vehicle has to move a full lane width in the transverse direction (3.5-4m). While vehicles under this regime are

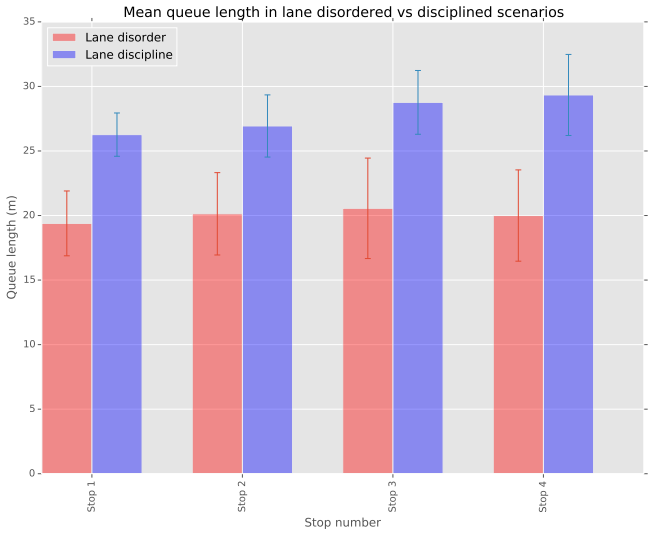

FIGURE 4.14: Queue length in lane disciplined and lane disordered settings. The largest-to-smallest starting condition was used for all simulations. A road width of 12m with 3 lanes and road length of 150m with periodic boundaries was used, as shown in Figure 4.13. Each value is the average of 30 simulation runs with error bars indicating the standard deviation between simulations.

technically obeying the stated assumptions of lane discipline, this frequent lane-switching behaviour is not typical of real lane-disciplined traffic streams where vehicles tend to exhibit a certain degree of 'lane inertia'. Future adaptations of the present gap-filling model for lane discipline could include a 'distress' factor that causes vehicles not to switch lanes unless their velocity is sufficiently reduced by the presence of a forward vehicle.

4.3 Sensitivity Analysis

The only free parameter contained in the model is α, which is used in the collision avoidance algorithm. Its function is fully explained in section 3.2.3. It expresses the degree of expectation that an agent has that other agents will maintain their current speeds. A value of 1 indicates that an agent assumes all other agents will maintain their current speeds, and it will therefore adopt following distances consistent with this expectation. A value of 0 corresponds to an agent assuming that all other agents may immediately become stationary, and would thus presumably lead to fewer collisions yet higher following distances.

While the term 'following distance' is useful in an intuitive sense, it lacks precise meaning in the context of non-lane-based traffic. This analysis will therefore focus on the sensitivity of the collision avoidance algorithm to the parameter α.

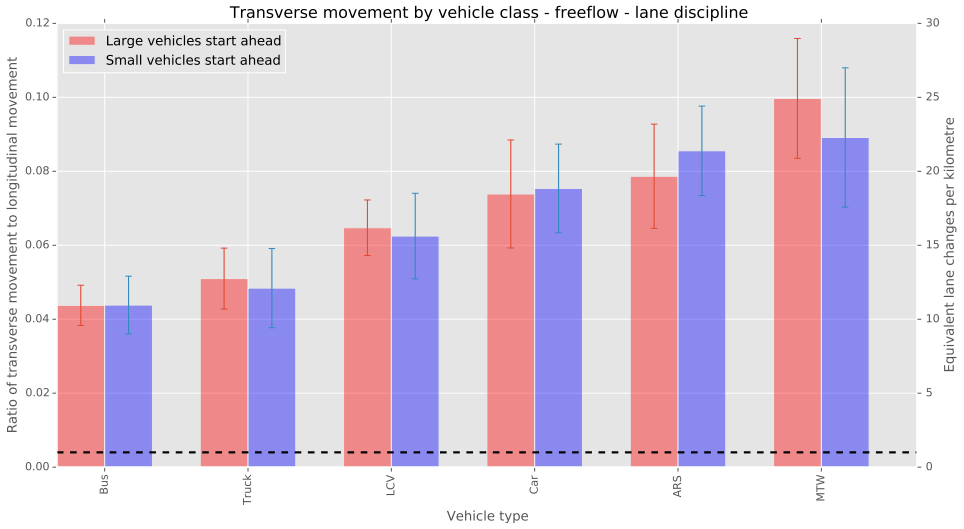

FIGURE 4.15: The degree of weaving (ratio of transverse to longitudinal movement) for each vehicle class in freeflow simulations of 20 agents distributed according to Table 4.3, on a 3-lane road of width 12m and length 150m with periodic boundaries, when vehicles are following lane discipline. Each value is the average of 50 simulations, with error bars indicating the standard deviation between all vehicles aggregated within each class, aggregated from separate simulation runs. To eliminate transient effects from starting conditions, the first 15% of each simulation is discarded. The right-hand vertical axis is a rescaling of the primary axis in units of lane changes per kilometre, based on a lane width of 4m. The horizontal broken line indicates an approximate typical number of lane changes per kilometre for urban developed-world traffic [41].

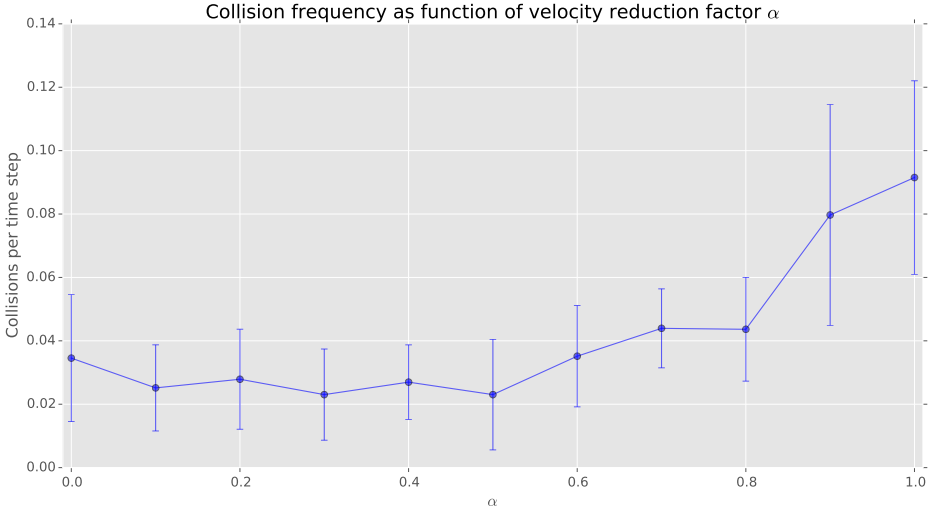

FIGURE 4.16: The effect of the velocity reduction factor α on frequency of collisions. The simulation setup was a 100m long road stretch with periodic boundaries, 7m wide, containing 10 cars following lane disorder. Each data point is the average of 10 simulations and error bars show standard deviation between simulation runs.

Figure 4.16 shows the frequency at which collisions occur as α is varied over its range of possible values. As can be seen and was mentioned earlier in Chapter 3, the collision avoidance algorithm does not guarantee a collision-free environment. Collisions can occur in the present model for two reasons: if a vehicle pulls in front of another and 'cuts it off', or if a vehicle suddenly reduces it speed. The former cause is largely unavoidable, while the second cause can be remedied by appropriate selection of α, or by use of a TTC model that includes acceleration.

It is desirable to choose a value of α that is as large as possible (so as to reduce following distances and lead to more efficient road usage) while simultaneously keeping collision frequency low. Based on this requirement, and with reference to Figure 4.16, a value of $\alpha = 0.5$ was used for all simulations described in this report.

Chapter 5

Relationship Between Lane Discipline and Traffic Efficiency

The motivation behind this project is to discern the effects of lane discipline on traffic flow in the developing world context. Due to a lack of suitable microscopic models that take into account the disorder in such traffic flow, it proved necessary to first devise a new dynamic model for accurately simulating these interactions. This model has been presented in Chapter 3 and validated and evaluated in Chapter 4. This chapter will now proceed to use this model to examine the interplay between lane discipline and traffic efficiency.

For someone accustomed to the strict discipline of roads in developed countries, it is perhaps natural to contemplate, upon observing traffic in the developing world, whether the efficiency of such traffic may be improved by the enforcement of developed world traffic rules, particularly rules relating to lane discipline. After all, it is clear that the coordination introduced by traffic lights is responsible for avoiding the gridlock situations shown in images (a) and (b) of Figure 2.7 on page 19, as evidenced by the traffic congestion that occurs when traffic lights occasionally fail on Western roads. Similarly, obedience to lane discipline introduces an analogous, implicit form of coordination: the knowledge that a vehicle can continue moving in its lane and consider only the actions of the vehicle directly in front of it, without having to consider possible collisions with encroaching vehicles. Indeed, total failure to adhere to lane discipline on Western roads is frequently described as reckless driving, considered a serious offence, and is usually severely punished. It is therefore natural to explore whether the enforcement of lane discipline on roads in the developing world may afford similar benefits. The experiments and analysis presented in this chapter constitute a first attempt at answering this question.

There exist many different measures of traffic efficiency, the suitability of which for a given purpose depends on the goals of the analysis. For the present investigation, the metric of average normalised vehicle speed (ANVS) will be used. As defined earlier in this report, it is calculated by dividing a vehicle's speed by its free speed, and averaging temporally and also possibly across a group of vehicles. This metric has the advantage that it is simultaneously able to capture elements of the state of traffic flow from a macroscopic perspective, and from an individual driver's perspective. From the macroscopic view, an ANVS of less than unity indicates some degree of congestion in the network, with the amount of congestion increasing as ANVS approaches zero. From each vehicle's perspective, ANVS not only describes the level of service a vehicle receives from the road network, but also motivates its actions in pursuit of higher ANVS. The importance of this last point is paramount in environments where new enforcement of lane discipline might prove problematic (for example, for cultural or historical reasons), because vehicles as agents will selfishly choose actions that maximise their individual utility. This fact of selfish maximisation of ANVS is evident even on roads in the developed world when, in heavy bumper-to-bumper traffic conditions, some vehicles frequently switch between lanes in an attempt to increase ANVS, despite this being to the detriment of the overall speed of the stream because of the friction involved in switching lanes.

The terms *AVNS* and *speed* will be used interchangeably throughout this chapter.

Before answering the question of whether disordered heterogeneous traffic could be made more efficient by adhering to lane discipline, we will consider the role of lane discipline in homogeneous traffic flow. Consider a perfectly homogeneous mix of vehicles. In the following experiments, a vehicle mix consisting solely of passenger cars travelling in the freeflow environment will be used. Now suppose that this stream of cars is completely obeying lane discipline. In this sense, a lane disciplined homogeneous mix of cars can be thought of as an idealised representation of developed world traffic. The question is now posed, "would it be advantageous to some of these cars to forgo lane discipline and instead selfishly exploit gaps in the traffic stream?". The answer is not immediately clear *a priori*. On the one hand, it might be the case that by ignoring lane discipline these cars are able to move through gaps that would have otherwise been unavailable to them, thus increasing their ANVS. However, it might actually be the case that the extra friction introduced by moving haphazardly in between other cars actually ends up slowing these vehicles down to a greater extent than they would have been if they had obeyed lane discipline. Finally, one can ask whether the situation would be different if it is a minority or majority of cars that decide to forgo lane discipline, and whether this would depend on the density of cars on the road. Simulation can be used to answer these questions.

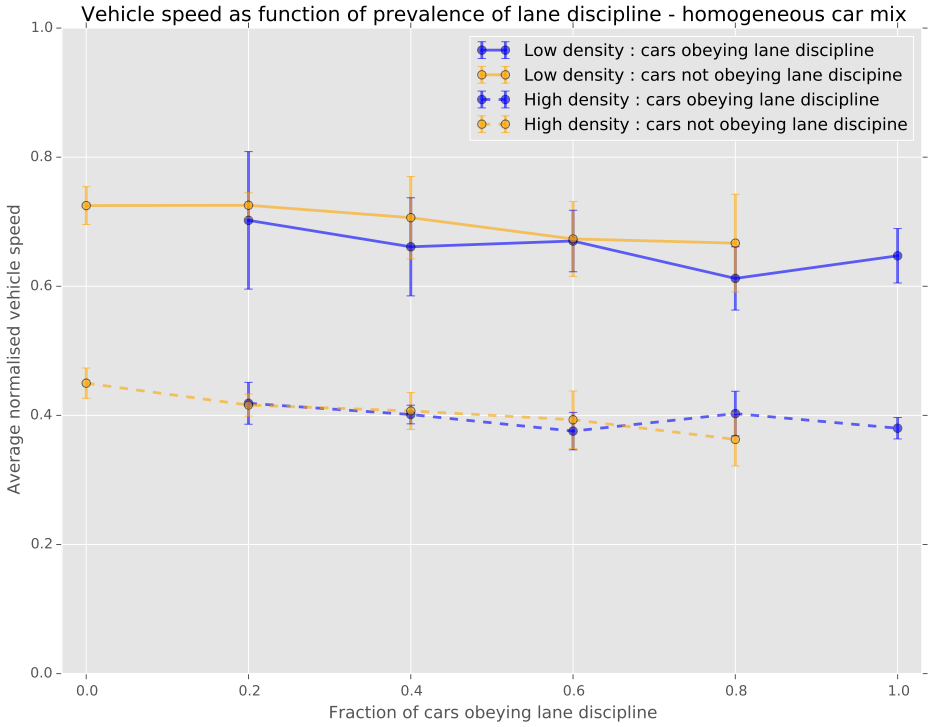

FIGURE 5.1: Normalised vehicle speed as function of prevalence of lane discipline for a homogeneous car mix. Solid lines correspond to a low density of cars, broken lines to a high density. Values are averaged over 30 simulation runs with error bars showing standard deviation between runs.

Figure 5.1 visualises car speeds in the above scenarios. Two broad cases are considered: one where the density of cars is low (solid lines), and one where the car density is high (broken lines). The abscissa shows varying levels of lane discipline. For example, a value of 0.2 indicates that 20% of cars are obeying lane discipline while the other 80% are not. The blue lines in each case indicate the ANVS of the vehicles which are obeying lane discipline, and the orange lines that of the vehicles not obeying lane discipline. It is apparent that at high car densities, ANVS is considerably lower than at low levels of density, as might be expected. The important result contained in this figure is that, irrespective of car density, there does not appear to be any significant difference in speed between those vehicles obeying lane discipline and those which are not. This is the case whether it is a minority or majority of cars that forgo lane discipline. Therefore, it can be concluded that for a homogeneous stream of cars, there does not appear to be any significant benefit in speed by forgoing lane discipline, nor any significant benefit by retaining it.

We now turn our attention to a heterogeneous mix of vehicles. Consider a stream of all six vehicle types used in this study, occurring in frequencies approximately equal to those contained in Table 4.3. If one were to consider six different levels of lane discipline

(as was just done for cars) for each of the six vehicle types, the problem would quickly grow to intractable dimensionality. To avoid this, the six vehicle types are grouped into three broader groups on the basis of size, and five scenarios constructed based on whether or not these groups obey lane discipline. These scenarios are specified in Table 5.1.

TABLE 5.1: Vehicle groups following lane discipline, by experimental scenario

	Experimental scenario				
Vehicle Group	1	2	3	4	5
Large (Bus, Truck)	False	True	False	False	True
Medium (LCV, Car)	False	False	True	False	True
Small (ARS, MTW)	False	False	False	True	True

The five scenarios can thus be summarised as: no vehicles obeying lane discipline, some vehicle groups obeying lane discipline, or all vehicles obeying lane discipline. Figure 5.2 shows the ANVS for each vehicle group in each of these scenarios. All blue bars correspond to scenario 1, orange bars correspond to scenarios 2, 3 and 4, and green bars correspond to scenario 5. Dark shaded bars correspond to the speed of the vehicle group specified on the horizontal axis, and light shaded bars correspond to the speed of the two vehicle groups other than the one specified on the horizontal axis. The plot therefore shows, for each vehicle group, in each of the scenarios, the average speed of that group and the average speed of the other groups.

The first observation that can be made based on the two bars on the left of the plot is that, for all vehicles taken together, speed decreases significantly when lane discipline is introduced. This is contrary to the finding for the homogeneous stream of cars where lane discipline was found to have no overall effect on vehicle speed. Examining the bars for each of the vehicle groups in turn reveals that all vehicle groups are slowed down by the introduction of lane discipline, whether is done just for the group in question or for all vehicles.

It is clear that large and medium vehicles are slower than the small vehicles, whether or not they are following lane discipline, which is to be expected. What might not be expected, however, is that introducing lane discipline into the small vehicle group has a significant negative effect on the speeds of all other vehicles. Compare the large difference between the light-shaded blue and light-shaded orange bars for the small group, to the much smaller differences between the light-shaded blue and light-shaded orange bars in each of the large and medium groups. Thus enforcing lane discipline for small vehicles

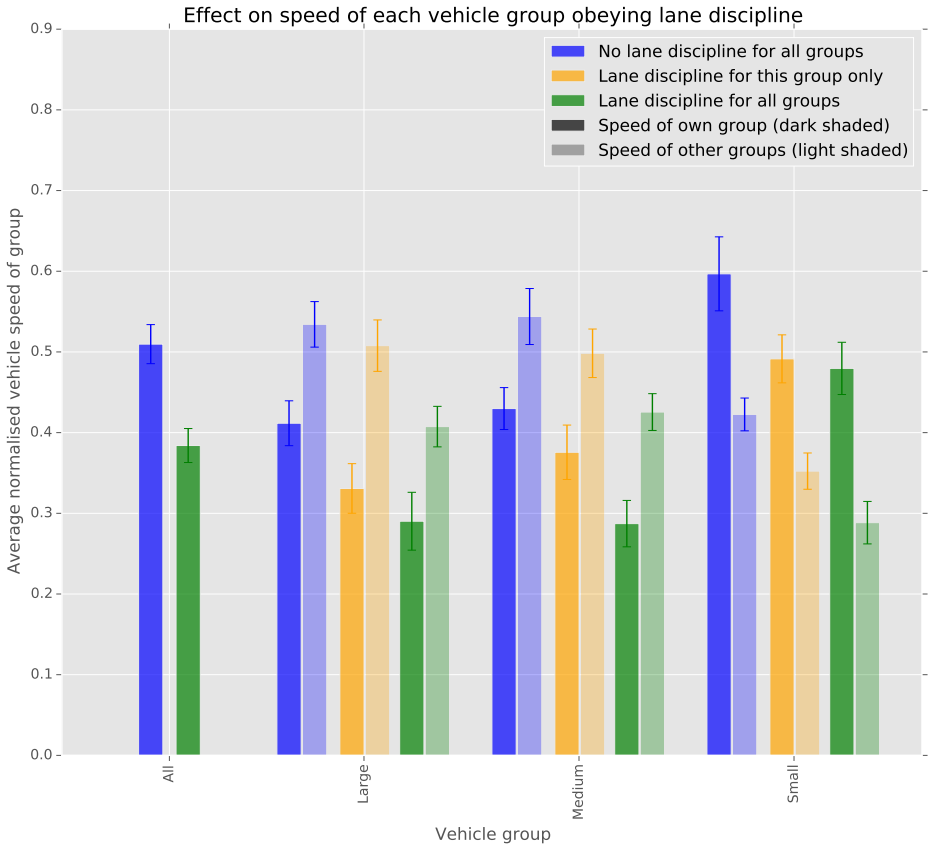

FIGURE 5.2: Effect on AVNS for each vehicle group in a highly heterogeneous stream obeying lane discipline. Simulations contained 20 vehicles distributed according to Table 4.3, with a minimum of 2 vehicles per class. All results are the average of 20 simulations, with error bars indicating standard deviation between multiple simulations.

ends up slowing down all other vehicles, even when those vehicles are allowed to move disorderedly. This is likely due to two factors: (i) the high prevalence of small vehicles on the road (approx. 60%), and (ii) that smaller vehicles do not significantly impede the flow of larger vehicles if the smaller vehicles are allowed to weave in between the larger vehicles.

It is therefore clear for a highly heterogeneous vehicle mix that, whether considering all vehicle groups together or each group in isolation, lane discipline results in uniformly lower speeds for all vehicles. If we regard the vehicles represented in Figure 5.2 as selfish agents seeking to maximise ANVS, they would all choose to forgo lane discipline. They need not, in fact, be selfish: the average speed of all *other* vehicles decreases when a vehicle group chooses to obey lane discipline. It is thus individually *and* collectively advantageous for vehicles in a highly heterogeneous stream not to obey lane discipline.

It should be noted that these results apply only to a single link where all vehicles are travelling in the same direction. The applicability of these findings is limited because

other traffic environments, such as intersections, have not been considered. Nevertheless, taking this result together with the finding for homogeneous streams, it appears that there are two distinct modes of traffic operation: for homogeneous streams, the presence of lane discipline; for heterogeneous streams, the absence of lane discipline. The presence or absence of lane discipline in each of these cases would appear to have arisen, in part, out of vehicles seeking to maximise their speed. In homogeneous streams where vehicles are all the same size and fill most of the lane, there is no benefit to disregarding lane discipline, thus vehicles choose to obey lane discipline for (presumably) other reasons. Conversely, in highly heterogeneous streams, there is a significant speed benefit for those vehicles which disregard lane discipline, hence why disordered flow may have emerged as the dominant traffic mode under such conditions.

To the extent that typical developed world traffic is heterogeneous through the presence of cars, motorcycles and scooters, the benefits of relaxing some of the requirements of lane discipline have already been recognised by authorities. As was mentioned in Chapter 2, the practice of lane sharing with motorcycles is becoming increasingly popular, with many local and national governments now legalising the practice. While the literature is unclear on the overall safety of this practice, it is certainly the case that lane sharing *per se* increases traffic efficiency.

Chapter 6

Possibilities for Future Work

This chapter will briefly discuss a few aspects of the present work that could form the basis for future research in this area. These are considered in three separate sections: improvements to the basic model (6.1), incorporating the present microscopic model into a mesoscopic framework (6.2), and adapting the model to simulate alternative traffic scenarios (6.3).

6.1 Improvement of basic model

The gap-filling model of disordered heterogeneous traffic that has been proposed in this report could be improved in a number of ways.

The improvement likely to yield the biggest change would be adapting the model to include acceleration. This would mean that, when calculating the time-to-collision with other vehicles, possible variations of the speed of these vehicles are taken into account. As has been mentioned in Chapter 3, it is precisely this expectation of variation in the speed of other vehicles that leads to non-zero following distances. If a vehicle were to assume that a second vehicle which is ahead of the first would maintain its current velocity indefinitely, the first vehicle would accelerate until it is immediately behind the second vehicle, leaving no space in between. By including acceleration in the model, the first vehicle would be able to account for possible deceleration of the second vehicle, which would allow it to adjust its following distance as necessary to ensure time-to-collision is kept above the threshold value. The knowledge of the acceleration/deceleration properties of other vehicles would in practice be gained by a driver's own experience, but in the model it might be assumed that these are known precisely. Additional parameters such as driver 'cautiousness' could be included to specify the degree of acceleration/deceleration a vehicle expects of other vehicles.

The present model incorporates automatic adjustment of an agent's threshold time-to-collision, t_{crit}, based on the agent's own speed. Other (additional) adjustment mechanisms could be considered in future work. For example, the use of anisotropic TTC thresholds may hold value. The motivation for such a scheme is that when, in order to pass another agent, a follower agent adopts a velocity that is not aligned with its preferred direction of motion, the follower agent adopts this velocity only temporarily, and therefore should not require that the TTC of this velocity be as high as for velocities which are more fully in the preferred direction of travel. Therefore, each potential velocity direction in the agent's coordinate system could have a different TTC threshold associated with it, with the highest values being for velocities in the direction of motion, and considerably lower values for velocities further away from this preferred direction.

To simulate lane discipline in the gap-filling model, geometric buffer zones around vehicles were used to prevent vehicles occupying overlapping longitudinal positions when in the same lane (i.e. lane splitting). While this proved to be effective, it was found during evaluation of the model that this led to levels of transverse movement significantly higher than for the lane disordered version, and higher than is typical of lane disciplined traffic. This is due to the dynamic model still being essentially gap-filling, and therefore when an agent is hindered by a forward vehicle, it will attempt to fill an adjacent gap in order to increase its speed. This gap-filling tendency could be dampened in the lane disciplined scenario by introducing a 'distress' factor that controls how likely an agent is to pursue a lateral velocity, based on how much its speed is hindered. For example, it might be specified that obstructions need to reduce a vehicle's speed to 75% of its free speed before the vehicle will turn out of its lane in order to move around the obstruction.

The current model adopts a binary approach to TTC values and the TTC threshold: velocities with TTC values below the threshold are not considered at all, unless there are *no* velocities above the threshold in which case that with the highest TTC value is chosen. Subsequent work on the model might investigate the use of velocity selection algorithms that use the continuous value of TTC, instead of just considering whether TTC is above the threshold, to select an appropriate velocity. In this way, TTC could function as a kind of 'sonar' to help the agent identify larger gaps in the traffic stream which might allow the agent to maximise its speed over the long term.

Finally, future work on the model could investigate how to accommodate transverse vehicle clearance requirements that change with speed and with the type of vehicle. During testing it was found that using heterogeneous clearances led to irregular behaviour. For example, if a vehicle with low clearance requirement approached a vehicle with a high clearance requirement, the latter vehicle may regard a collision as having occurred and immediately halt, despite this not being the case. One possible resolution that would

allow dynamic transverse clearances to be implemented would be to have these clearances be elastic, albeit with a preferred clearance size. If an agent discovers that there is another agent within its clearance limits, it can gradually attempt to move away so that the preferred minimum clearance is attained.

6.2 Integration into parallel-hybrid mesoscopic models

An important potential use for the present gap-filling microscopic model is to combine it with parallel-hybrid mesoscopic models (PHMMs). As was mentioned in Chapter 2, PHMMs couple microscopic and macroscopic models in order to simulate large scale traffic networks, while still maintaining the fidelity and specificity afforded by microscopic models.

A widely-used example of such a mesoscopic framework is MATSim. This software allows agent-based simulation of large-scale traffic networks, such as entire cities consisting of millions of agents [42]. While it does not simulate direct physical vehicle interactions as a microscopic model would, it makes use of a highly modular approach which allows for the addition of modules which perform detailed handling of agent interactions, such as calculating travel times for links, determining vehicle behaviour at controlled and uncontrolled intersections, among others. For example, the default handling in MATSim of vehicles travelling along a single link follows first-in-first-out (FIFO) logic, whereby the first vehicle to enter a link is the first vehicle to exit. While this may not be a poor assumption when a homogeneous mix of vehicles with similar free speeds is the case, it becomes severely limiting when a heterogeneous mix consisting of vehicles of different sizes, free speeds and manoeuvrability is being modelled.

Recent work has recognised this shortcoming and used a simple model based on vehicle free speeds, together with PCU factors based on vehicle areas, to model the variation in vehicle travel times along a single link [43]. A high-fidelity microscopic model such as that developed in this project would greatly improve the accuracy of this computation, due to it accounting for the complex interactions between vehicles in disordered heterogeneous flow, which cannot be fully captured by macroscopic methods such as the use of PCU factors. Network configurations and demand data have already been compiled and implemented in MATSim for the Indian cities of Bangalore and Patna, making this a ripe avenue for future research.

6.3 Microsimulation of alternative traffic scenarios

While the simulations performed for this project focussed solely on the simplest case of a one-way link, the model can be readily applied to alternative traffic configurations, such as bi-directional flow and intersections (both controlled and uncontrolled). Simulations of crowd models using similar collision avoidance principles to the present gap-filling model have shown that agents are able to pass through streams of other agents travelling in opposing directions. This is illustrated in Figure 6.1.

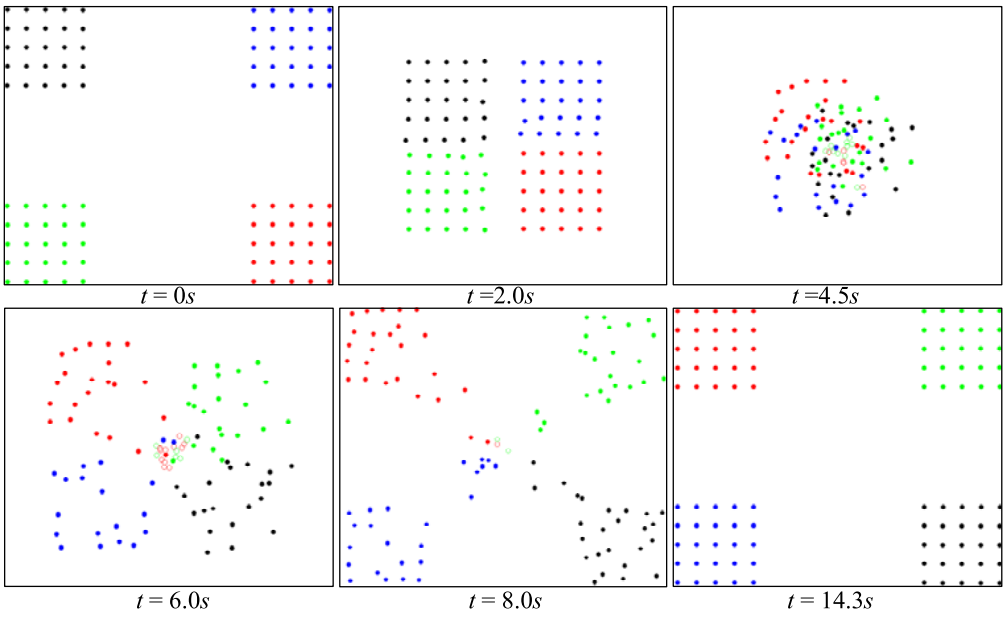

FIGURE 6.1: Simulation of a crowd model showing agents crossing disordered streams of other agents by using a collision avoidance approach. [33]

Such a scenario could easily be simulated using the present model by have multiple groups of vehicle agents that differ only in their initial location and preferred direction of motion. An approach like this could be used to model bi-directional traffic, uncontrolled intersections and traffic circles, such as those shown in frames (a), (b), (e) and (f) of Figure 2.7 on page 19.

Chapter 7

Conclusion

This report has presented a novel agent-based gap-filling model for the simulation of heterogeneous traffic flow as found in many developing countries. This model has been validated with respect to the functional phenomena unique to developing world traffic, and subsequently been used to determine a causal relationship between vehicle heterogeneity, lane discipline and traffic efficiency.

While there has been substantial research over the past 70 years on modelling and simulation of car-based traffic that is typical of developed countries, there has only recently been research activity in the field of modelling developing world traffic. Traffic in developing countries such as India, Pakistan, Vietnam, Thailand, Malaysia, China and Indonesia is comprised of a large range of vehicles, both motorised and non-motorised. This traffic is also characterised by a high degree of disorder, where vehicles do not adhere to the principles of lane discipline. Most attempts to model such traffic have focussed on adapting models developed for homogeneous lane-based traffic in Western countries, and changing either the composition of the vehicle mix or numerical parameters. These attempts have proven of limited value due to their not accounting for the qualitatively different driving modes typical of developing world traffic. The most recent published model for such traffic likens the traffic stream to a porous medium, where smaller vehicles move through pores defined by the boundaries of larger vehicles. While this model employs more realistic assumptions of vehicle movement, its macroscopic scale and abstract parameterisation precludes analysis of individual vehicle trajectories and speeds.

The gap-filling model presented in this work is based on the principle of collision avoidance found in dynamic crowd models. By deriving accurate kinematic relations for each class of vehicle based on physical and mechanical properties, as well as a physics engine tailored to the approximately rectangular geometry of most vehicles, the model employs

an agent-based framework consisting of vehicles which use a rule-based velocity selection process to navigate through a stream of traffic at maximum speed while seeking to avoid collisions. Due to a lack of suitable data on the macroscopic and microscopic properties of developing world traffic, it was not possible to calibrate or validate the model quantitatively. Instead, the model was validated against functional aspects of developing world traffic, and found to reproduce these phenomena accurately.

By employing a modification, the model is able to accommodate lane-based traffic. This modification was made to permit investigation of the effect that the enforcement of lane discipline might have on developing world traffic. Simulations under this scenario found that vehicles were obeying the requirements of lane discipline, although their degree of movement between lanes was considerably higher than for typical Western lane-based traffic. Nevertheless, it need not be the case that if developing world drivers were forced to obey lane discipline their lane-changing rates would reduce to that of typical developed world traffic.

The effects of lane discipline and lane disorder were then determined on, separately, homogeneous and heterogeneous traffic streams travelling in an urban, one-way link. By use of simulation it was found that, for a homogeneous stream of cars, there was no significant difference in traffic efficiency (measured by each vehicle's speed as a fraction of its free speed), regardless of whether the density of vehicles was low or high. For a heterogeneous mix of vehicles in proportions typically found on roads in the developing world, it was found that imposition of lane discipline had a significant *negative* effect on traffic efficiency, both for the individual driver and the traffic stream as a whole. It can therefore be concluded that the practice of lane disorder in developing world traffic is accompanied by significant increases in traffic efficiency compared to the lane-disciplined scenario, and this may be responsible for the dominance of lane disorder in environments with highly heterogeneous traffic streams.

The present work could be extended in a number of ways. The underlying model might be improved by taking vehicle accelerations into account during the collision avoidance process. By altering the preferred direction of motion for some vehicles, alternative traffic scenarios, such as intersections, could be simulated. Since the aforementioned relationship between vehicle heterogeneity, lane discipline and traffic efficiency was only investigated on a single one-way link, these modifications could be used to investigate whether the relationship holds for intersections. Finally, the microscopic model could be integrated into a parallel-hybrid mesoscopic model to accurately simulate traffic flow on a larger, city-wide scale.

This work makes the following contributions to the limited literature on disordered heterogeneous traffic modelling. By introduction of a microscopic gap-filling model

instead of a modified conventional traffic model as has been the focus of previous works, a new modelling paradigm for disordered heterogeneous traffic has been established. This gap-filling model should be considered the first iteration of a class of models which view the apparently chaotic interactions of disordered heterogeneous traffic flow as the emergent behaviour resulting from individual agents following simple rules in a complex system. Finally, this work also poses the question of why lane disorder has emerged in traffic in the developing world and answers this by means of simulation.

It is hoped that the reader is left with an appreciation of the intricacies of modelling disordered heterogeneous traffic flow, and how new modelling tools are necessary to understand this form of traffic that is a daily reality for many millions of people.

Bibliography

[1] V Thamizh Arasan and Reebu Zachariah Koshy. Methodology for modeling highly heterogeneous traffic flow. *Journal of Transportation Engineering*, 131(7):544–551, 2005.

[2] V Thamizh Arasan and G Dhivya. Measuring heterogeneous traffic density. In *proceedings of the International conference on Sustainable Urban Transport and Environment, World Academy of Science, Engineering and Technology, Bangkok*, pages 342–346, 2008.

[3] BD Greenshields, Ws Channing, Hh Miller, et al. A study of traffic capacity. In *Highway research board proceedings*, volume 1935. National Research Council (USA), Highway Research Board, 1935.

[4] Martin Treiber and Arne Kesting. Traffic flow dynamics. *Traffic Flow Dynamics: Data, Models and Simulation, Springer-Verlag Berlin Heidelberg*, 2013.

[5] Joseph S Drake, Joseph L Schofer, and Adolf D May Jr. A statistical analysis of speed-density hypotheses. in vehicular traffic science. *Highway Research Record*, (154), 1967.

[6] Stef Smulders. Control of freeway traffic flow by variable speed signs. *Transportation Research Part B: Methodological*, 24(2):111–132, 1990.

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

[8] Van Wageningen-Kessels, Serge P Hoogendoorn, V Vuik, JWC Van Lint, et al. Traffic flow modeling: a genealogy. 2014.

[9] Kai Nagel and Michael Schreckenberg. A cellular automaton model for freeway traffic. *Journal de physique I*, 2(12):2221–2229, 1992.

[10] Michael J Lighthill and Gerald Beresford Whitham. On kinematic waves. ii. a theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society*

of London A: Mathematical, Physical and Engineering Sciences, volume 229, pages 317–345. The Royal Society, 1955.

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

[12] Emmanuel Bourrel and Vincent Henn. Mixing micro and macro representations of traffic flow: a first theoretical step. In *Proceedings of the 9th meeting of the Euro Working Group on Transportation*, pages 610–616, 2002.

[13] Sarosh Khan and Pawan Maini. Modeling heterogeneous traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, (1678):234–241, 1999.

[14] Rahul Nair, Hani S Mahmassani, and Elise Miller-Hooks. A porous flow approach to modeling heterogeneous traffic in disordered systems. *Transportation Research Part B: Methodological*, 45(9):1331–1345, 2011.

[15] GEETAM Tiwari, JOSEPH Fazio, and S Pavitravas. Passenger car units for heterogeneous traffic using a modified density method.

[16] K Venkatesan, A Gowri, and R Sivanandan. Development of microscopic simulation model for heterogeneous traffic using object oriented approach. *Transportmetrica*, 4(3):227–247, 2008.

[17] Tom V Mathew and Padmakumar Radhakrishnan. Calibration of microsimulation models for nonlane-based heterogeneous traffic at signalized intersections. *Journal of Urban Planning and Development*, 136(1):59–66, 2010.

[18] Shreerup Goswami, Subrata Kumar Nayak, Akula Chandra Pradhan, and Surjendu Kumar Dey. A study on traffic noise of two campuses of university, balasore, india. 2011.

[19] Amit Agarwal and Gregor Lämmel. Modeling seepage behavior of smaller vehicles in mixed traffic conditions using an agent based simulation. *Unpublished*, 2015. URL https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2015/15-09/2015-05-12_seepage_submitted_ctrg.pdf.

[20] Hugh H Hurt, JV Ouellet, and DR Thom. *Motorcycle Accident Cause Factors and Identification of Countermeasures: Appendix*, volume 2. The Administration, 1981.

[21] Myra Sperley and Amanda Joy Pietz. Motorcycle lane-sharing: Literature review. Technical report, 2010.

[22] David D Clarke, Pat Ward, Craig Bartle, and Wendy Truman. In-depth study of motorcycle accidents. Technical report, Road safety research report, 2004.

[23] V Thamizh Arasan and P Vedagiri. Estimation of saturation flow of heterogeneous traffic using computer simulation. *Indian Institute of Technology, Chennai*, 2006.

[24] V Thamizh Arasan and Shriniwas S Arkatkar. Derivation of capacity standards for intercity roads carrying heterogeneous traffic using computer simulation. *Procedia-Social and Behavioral Sciences*, 16:218–229, 2011.

[25] Muhammad Adnan. Passenger car equivalent factors in heterogenous traffic environment-are we using the right numbers? *Procedia engineering*, 77:106–113, 2014.

[26] Serge Hoogendoorn and Piet Bovy. Gas-kinetic model for multilane heterogeneous traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, (1678):150–159, 1999.

[27] KVR Ravishankar and Tom V Mathew. Vehicle-type dependent car-following model for heterogeneous traffic conditions. *Journal of transportation engineering*, 137(11): 775–781, 2011.

[28] Mahendrakumar Metkari, Anuj Budhkar, and Akhilesh Kumar Maurya. Development of simulation model for heterogeneous traffic with no lane discipline. *Procedia-Social and Behavioral Sciences*, 104:360–369, 2013.

[29] Asaithambi Gowri, Kanagaraj Venkatesan, and Ramaswamy Sivanandan. Object-oriented methodology for intersection simulation model under heterogeneous traffic conditions. *Advances in Engineering Software*, 40(10):1000–1010, 2009.

[30] Manraj Singh Bains, Balaji Ponnu, and Shriniwas S Arkatkar. Modeling of traffic flow on indian expressways using simulation technique. *Procedia-Social and Behavioral Sciences*, 43:475–493, 2012.

[31] Ch Mallikarjuna and K Ramachandra Rao. Cellular automata model for heterogeneous traffic. *Journal of Advanced Transportation*, 43(3):321–345, 2009.

[32] Lasmini Ambarwati, Adam J Pel, Robert Verhaeghe, and Bart van Arem. Empirical analysis of heterogeneous traffic flow and calibration of porous flow model. *Transportation research part C: emerging technologies*, 48:418–436, 2014.

[33] Muzhou Xiong, Michael Lees, Wentong Cai, Suiping Zhou, and Malcolm Yoke Hean Low. Analysis of an efficient rule-based motion planning system for simulating human crowds. *The visual computer*, 26(5):367–383, 2010.

[34] Eric W. Weisstein. Osculating circle. URL http://mathworld.wolfram.com/OsculatingCircle.html. Visited on 2015/01/12.

[35] Andy Dingley. Ackermann steering geometry. URL https://commons.wikimedia.org/wiki/File:Ackermann_turning.svg. Visited 2015/01/12.

[36] Jin Xu, Kui Yang, YiMing Shao, and GongYuan Lu. An experimental study on lateral acceleration of cars in different environments in sichuan, southwest china. *Discrete Dynamics in Nature and Society*, 2015, 2015.

[37] Rijurekha Sen, Andrew Cross, Aditya Vashistha, Venkata N Padmanabhan, Edward Cutrell, and William Thies. Accurate speed and density measurement for road traffic in india. In *Proceedings of the 3rd ACM Symposium on Computing for Development*, page 14. ACM, 2013.

[38] R Asha Anand, Lelitha Vanajakshi, and Shankar C Subramanian. Traffic density estimation under heterogeneous traffic conditions using data fusion. In *Intelligent Vehicles Symposium (IV), 2011 IEEE*, pages 31–36. IEEE, 2011.

[39] L Ambarwati, AJ Pel, RJ Verhaeghe, and B Van Arem. Empirical analysis of heterogeneous traffic flow. In *Proceedings of the Eastern Asia Society for Transportation Studies (EASTS), Vol. 9, Taipei, Taiwan, September 9-12, 2013*, 2013.

[40] Akhilesh Kumar Maurya and Prashant Shridhar Bokare. Study of deceleration behaviour of different vehicle types. *International Journal for Traffic and Transport Engineering*, 2(3):253–270, 2012.

[41] VL Knoop, SP Hoogendoorn, Y Shiomi, and C Buisson. Quantifying the number of lane changes in traffic: An empirical analysis (poster). In *91st Annual Meeting Transportation Research Board, Washington, USA, 22-26 January 2012*, 2012.

[42] M Balmer, K Meister, M Rieser, K Nagel, Kay W Axhausen, Kay W Axhausen, and Kay W Axhausen. *Agent-based simulation of travel demand: Structure and computational performance of MATSim-T*. ETH, Eidgenössische Technische Hochschule Zürich, IVT Institut für Verkehrsplanung und Transportsysteme, 2008.

[43] Amit Agarwal. Agent based simulation of the travel demand for patna city, india. 2012.