

# Applicability and Application of Microscopic Traffic Simulations

**Florian Knorr**

Dissertation

zur Erlangung des akademischen Grades  
Doktor der Naturwissenschaften

vorgelegt der

Fakultät für Physik  
Universität Duisburg-Essen

im

Februar 2013

Universität Duisburg-Essen



# Applicability and Application of Microscopic Traffic Simulations

UNIVERSITÄT  
DUISBURG  
ESSEN

*Open-Minded*

Florian Knorr  
Fakultät für Physik  
Universität Duisburg-Essen

Dissertation  
zur Erlangung des akademischen Grades  
*Doktor der Naturwissenschaften*

Februar 2013

Gutachter 1: Prof. Dr. Michael Schreckenberg

Gutachter 2: Prof. Dr. Andreas Schadschneider

Tag der Disputation: 12. Juli 2013



MEINEN ELTERN

*It is good to have an end to journey towards; but it is the journey that matters in the end.*

Ursula K. Le Guin

## Acknowledgments

First and foremost, I would like to thank my academic advisor, Prof. Dr. Michael Schreckenberg, not only for working with me during the past three years but also for his open door policy: the door to his office stood almost always open, so that one could easily stop by and discuss a problem with him. I am also grateful to Prof. Dr. Schreckenberg for giving me a high degree of freedom in my research but for providing the necessary guidance at the same time. In particular, the results of chapter 6 would not have been obtained without his continuous encouragement.

For fruitful discussions and patient explanations I thank Prof. Dr. Boris Kerner (Daimler AG), Dr. Arne Kesting (formerly: Dresden University of Technology, now: TomTom), Prof. Dr. Andreas Schadschneider (University of Cologne), and Dr. Martin Treiber (Dresden University of Technology). Their technical expertise and helpfulness allowed me to gain a deeper insight in various topics of the physical aspects of traffic.

I also wish to thank all former and current members of the Chair of Physics of Transport and Traffic, who have been very supportive, for making the past years so enjoyable. I thank Lars Habel for preparing the empirical data in an easily accessible format and for providing some of his scripts to access these data. For careful proofreading I thank Gerhard Hermanns and Peter Hemmerle.

I am also grateful to the state of North Rhine-Westphalia and the European Union's Fund for Regional Development, which sponsored my research within the ZIEL 2 program "automotive.NRW".

Lastly, I would like to thank my family for all their love and encouragement.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Traffic – why physicists care . . . . .	1
1.2	Outline of the thesis . . . . .	3
<b>2</b>	<b>Traffic Flow Fundamentals</b>	<b>5</b>
2.1	Traffic Flow Observables . . . . .	5
2.2	The Fundamental Diagram . . . . .	6
2.3	Traffic Breakdown . . . . .	9
2.4	Phase Transitions and Traffic Phases . . . . .	11
2.5	Three-Phase Traffic Theory . . . . .	12
2.5.1	Three-Phase Traffic Theory and Earlier Modeling Approaches . . . . .	13
<b>3</b>	<b>Modeling Approaches and Breakdown Reproducibility</b>	<b>15</b>
3.1	Microscopic Traffic Models . . . . .	17
3.1.1	Intelligent Driver Model . . . . .	18
3.1.2	Nagel-Schreckenberg Model . . . . .	19
3.1.3	Comfortable Driving Model . . . . .	20
3.2	Empirical Data and Model Testing . . . . .	22
3.3	Simulation Setup . . . . .	24
3.3.1	Modeling Open Boundaries and Ramps . . . . .	24
3.3.2	Calibration . . . . .	26
3.4	Results . . . . .	27
3.4.1	Reproducibility of Upstream Boundary Conditions . . . . .	30
3.4.2	Propagation Velocity of Jammed Traffic . . . . .	31
3.4.3	Influence of Truck Length and Downstream Boundary Condition . . . . .	32
3.5	Discussion . . . . .	33

---

<b>4</b>	<b>Traffic Phases and Phase Transitions in the CDM</b>	<b>37</b>
4.1	Earlier Results . . . . .	38
4.2	Simulation Setup – Open Boundaries . . . . .	39
4.3	Spatiotemporal Dynamics . . . . .	40
4.4	Classification of Traffic Phases . . . . .	42
4.4.1	Analyzing Single Vehicle Data . . . . .	46
4.4.2	Simulated versus Empirical Data . . . . .	49
4.5	Deceleration Behavior of the CDM . . . . .	52
4.6	Discussion . . . . .	54
<b>5</b>	<b>Optimizing Traffic Flow by Vehicular Communication</b>	<b>57</b>
5.1	Radio Wave Propagation Fundamentals . . . . .	58
5.1.1	Propagation Models . . . . .	58
5.1.2	Noise, Interference, and Signal Reception . . . . .	61
5.2	Simulating Inter-Vehicle Communication . . . . .	62
5.3	Evaluating Traffic Information and Improving Traffic Flow . . . . .	63
5.3.1	Beacons for Congestion Warning . . . . .	64
5.3.2	Vehicle Motion and Adapted Driving Behavior . . . . .	66
5.3.3	Simulation . . . . .	67
5.4	Conclusion . . . . .	76
<b>6</b>	<b>Knowledge in VANETs: a Random Walk Approach</b>	<b>79</b>
6.1	Model Description . . . . .	80
6.2	Determining the Fraction of Known Particles . . . . .	81
6.2.1	Analytic Solution of the Symmetric Case . . . . .	81
6.2.2	Analytic Solution of the Asymmetric Case . . . . .	84
6.2.3	The Asymmetric Case and the Density of Local Maxima . . . . .	85
6.3	Simulations for Vehicular Networks . . . . .	85
6.4	Discussion . . . . .	86
<b>7</b>	<b>Conclusions and Outlook</b>	<b>89</b>
<b>A</b>	<b>Hypergeometric Series</b>	<b>91</b>
<b>B</b>	<b>The FOTO-method</b>	<b>95</b>
<b>C</b>	<b>CDM with Extensions for VANETs</b>	<b>99</b>
	<b>References</b>	<b>103</b>
	<b>Anlagen</b>	<b>112</b>
	Abstract . . . . .	112
	Zusammenfassung . . . . .	113
	Erklärung . . . . .	114
	Lebenslauf . . . . .	115

## 1.1 Traffic – why physicists care

Certainly, traffic is an important subject that affects us all: when we commute to and from work, or when we are stuck in a traffic jam, we are part of traffic. But what makes traffic an interesting research topic, and why is it physicists who show a particular interest in traffic and related systems?

At first glance, it may not be obvious why physicists show so much interest in vehicular traffic. Motion, acceleration, and deceleration of an object such as a vehicle are topics, which are well understood since Isaac Newton. So, why still care about traffic? Actually, things get far more complicated when we consider several vehicles and the interaction between them.

This interaction, which forces some vehicles to brake and allows others to accelerate, leads to collective effects. Finally, when we speak of “particles” instead of vehicles, traffic becomes a dynamical system of interacting particles. This formulation better illustrates the relation between traffic and other physical systems; molecules in a liquid or electrons in a solid are just other examples of interacting many-particle systems.

These analogies attracted the interest of physicists, who studied traffic with the same methods as other physical systems. Together with Robert Herman, who is today better known for his theoretical prediction of the cosmic microwave background radiation [1], later Nobel laureate Ilya Prigogine wrote in 1971:

Traffic flow can be measured, in principle, in the same way as can be measured any flow of fluid; also, the relevant parameters that determine the value of the flow (such as density of traffic, or number of lanes on a highway) can be identified and varied. [115, p. xv]

Of course, one should not overstrain these analogies. There are some distinct differences between vehicles on a road and molecules in a liquid. Vehicles are driven by humans, who all have different attitudes and exhibit different driving behaviors (e.g., some drivers may deliberately ignore a speed limit whereas others are overcautious). Molecules, on the other hand, always obey the laws of physics!

Yet the variety of driving characteristics does not change the qualitative features of vehicular traffic ([56, ch. 20] and [62, 120]). Despite this inter-driver variability, there are, after all, some goals which are common to all drivers and which dominate their collective motion: drivers try to avoid collisions, and they want to reach their destination in a reasonable time. These goals may serve as basic principles of any traffic flow theory.

All of the above observations justify the study of traffic with the tools of statistical mechanics. Indeed, treating traffic as a physical system has allowed physicists to gain a deeper insight into the various aspects of (vehicular) traffic. It is, for instance, a common observation that interactions between vehicles become stronger and more frequent with increasing traffic density. Drivers then have to accelerate and to brake permanently to keep a desired distance to the vehicle ahead. Finally, when traffic becomes too dense, subsequent braking events may lead to a breakdown of traffic flow: congestion occurs!

Generally, when a slight variation of the system's parameters (e.g., density) abruptly changes the system's behavior—resulting in a very characteristic change of one (or more) observable(s)—physicists usually speak of a phase transition. In fact, there is evidence that such a breakdown of traffic flow is a first-order phase transition [87, 113]. There is also evidence of self-organized criticality [6, 101] and the formation of shock waves [21, 92, 121] to name but a few examples illustrating the variety of physical phenomena that traffic exhibits.

During the past decades, traffic research has produced a multitude of mathematical and computational models to study traffic flow on several levels of detail. The quality of these models has been steadily improved, not least because of the increased availability of empirical and experimental data (e.g., provided by stationary detectors, aerial photographs, or recorded vehicle trajectories). We are now able to understand (at least some of) the principles and mechanisms responsible for a traffic breakdown and the resulting spatiotemporal traffic dynamics. This knowledge also enables us to control, to improve, and to simulate traffic flow. As we will see, computers have become a very valuable tool of traffic research. (For one of the earliest applications of computers to the study of traffic flow, see [30, 31].)

It is no surprise that computer simulations, which have become an integral part of modern research, also play a central role in the study of traffic flow. As traffic phenomena are the result of hundreds or even thousands of interacting vehicles, real-world experiments are often intractable for practical or financial constraints. Hence, the impact of new technologies (e.g., such as on-ramp metering) can more easily be studied with the help of traffic simulations. This, however, requires the underlying traffic model to correctly reproduce the empirical features of traffic and to allow for a fine grained analysis.

The intention of this thesis is to investigate both aspects. We want to investigate the ability of selected traffic models to reproduce the dynamics of real traffic flow. Afterwards, we want to use the results to study the influence of vehicle-to-vehicle communication on traffic flow, in particular its potential to reduce congestion.

## 1.2 Outline of the thesis

This thesis starts with a review of the most important concepts of traffic flow in **chapter 2**. We describe how empirical data are measured via loop detectors and how the fundamental observables—flow  $J$ , vehicle density  $\rho$ , and velocity  $v$ —are related. We also explain the term “traffic breakdown” and show that it can be interpreted as a boundary-induced phase transition, known from other non-equilibrium systems. This, finally, leads us to Kerner’s three-phase traffic theory, of which the basic ideas we will briefly explain. Notwithstanding the obvious analogies between traffic and other physical systems, we have to note that the terms “phase” and “phase transition” are used in a less strict manner than in the physical literature.

In **chapter 3** we investigate how well three selected microscopic traffic models can reproduce such a breakdown (i.e., the transition from free flow to congested traffic). To allow not only for a qualitative but also for a quantitative analysis, our investigation is based on empirical traffic data. To study the breakdown’s spatiotemporal dynamics, we trace its propagation over several detector cross-sections a few kilometers apart. Besides the overall agreement with the empirical data, we also study some modeling related aspects (e.g., the length of vehicles or the setup of the downstream boundary) and their influence on the results.

From the three models, we select the *comfortable driving model* (CDM) for a more detailed analysis in **chapter 4**. There, we study the spatiotemporal patterns resulting from different inflow and outflow probabilities of an open road segment. Based on time series of local measurements, the local traffic states are assigned to different traffic phases according to Kerner’s three-phase traffic theory. For this classification we use the rule-based FOTO-method, which was developed by Kerner et al. [69]. Our analysis shows that the model is indeed able to reproduce three qualitatively different traffic phases: free flow (F), synchronized traffic (S), and wide moving jams (J). In addition, we investigate the likelihood of transitions between these three traffic phases. We show that a transition from free flow to a wide moving jam often involves a sequence of two transitions; first from free to synchronized flow and then from synchronized flow to a wide moving jam. This, again, is in agreement with empirical results which indicate that the so called F→S transition (from free flow to synchronized traffic) is much more likely than a direct F→J transition. For an even deeper analysis we also study single vehicle data, where we will find evidence that the states classified as “synchronized” in the CDM are not entirely jam-free. The chapter concludes with a statistical analysis of the CDM’s deceleration behavior.

In **chapter 5** we address a more practical problem. We present a strategy to reduce traffic congestion with the help of vehicle-to-vehicle communication. Periodically emitted status messages are used to analyze traffic flow and to warn other drivers of a possible traffic breakdown. Drivers who receive such a warning are told to keep a larger gap to their predecessor. By doing so, they are less likely the source of perturbations, which can cause a traffic breakdown. We analyze the proposed strategy via computer simulations and investigate which fraction of communicating vehicles is necessary for

a beneficial influence on traffic flow to become observable. We show that penetration rates of 10 % and less can have a significant influence on traffic flow and travel times. In addition to applying a realistic mobility model (CDM), we further increase the degree of realism by the use of empirical traffic data from loop detectors on a German Autobahn and a realistic radio wave propagation model (Nakagami- $m$ ).

Motivated by heterogeneous traffic flow with communicating and non-communicating vehicles, we address a combinatorial problem in **chapter 6**. If communicating vehicles can detect the vehicle ahead (and behind) by front (and rear) sensors, the information gathered by communicating vehicles may considerably exceed the information about non-communicating vehicles. This additional information may be of use for traffic surveillance or traffic state estimation. Therefore, we ask how many non-communicating vehicles are on average detected by communicating ones. To answer this question, we map the problem to a discrete random walk in one dimension with a fixed number of steps. This mapping allows us to determine the detected vehicles for both open and periodic boundary conditions as well as for the case where communicating vehicles detect both or only one neighboring vehicle(s). In the random walk picture, where the two vehicle types (communicating and non-communicating) stand for steps in opposite directions, non-communicating vehicles are detected whenever the resulting path has a corner. For open boundary conditions, it turns out that a simple mean field approximation yields the exact result if communicating vehicles can detect only the vehicle ahead.

## Traffic Flow Fundamentals

Gone are the days when Prigogine and Herman complained of “the absence of adequate experimental data” [115, p. 91]. A large amount of experimental and empirical data has been collected since then (e.g., [3, 35, 54, 56, 82, 131, 139]). From the analysis of these data many characteristic features of traffic could be derived and theories be built upon. In this chapter we will give an overview of the most important findings.

### 2.1 Traffic Flow Observables

Currently, stationary loop detectors are still the most common source of traffic data. In their simplest form, loop detectors count the number of passing vehicles, measure their velocity, and determine the fraction of time that the detector is occupied by a vehicle. These data are usually aggregated over intervals of duration  $T$ , which we will indicate by using the notation  $\langle \cdot \rangle_T$ . (Later, we will forgo this explicit distinction for the sake of better readability when it is clear that aggregated data are considered.) Two fundamental traffic observables, traffic flow and velocity, are directly accessible from these local measurements, and the third, traffic density, can be estimated from the previous two.

- The **traffic flow**  $J$  is given by the number of vehicles  $N$  that passed the detector’s cross-section in the interval  $[t, t + T)$ :

$$\langle J(t) \rangle_T = \frac{N}{T}. \quad (2.1)$$

Although  $T$  may be significantly smaller, the flow is typically measured in terms of vehicles per hour (vehicles/h).

- Similarly, the average **velocity** or the time-mean speed  $v(t)$  of the passing vehicles is given by the arithmetic mean

$$\langle v(t) \rangle_T = \frac{1}{N} \sum_{i=1}^N v_i, \quad (2.2)$$

where  $v_i$  denotes the velocity of the  $i$ th vehicle that passed the detector in the interval  $[t, t + T)$ .

- The **traffic density**  $\rho$ , the third fundamental observable, must be measured along a road segment and cannot be directly obtained from the local measurements of a stationary detector. However, it can be estimated either from the previous two observables or from the occupancy  $O_D$ , which denotes the fraction of time that the detector is occupied by passing vehicles. Both methods are based on the hydrodynamic relation

$$J = \rho v. \quad (2.3)$$

The first estimate of  $\rho$  requires solving equation (2.3) for  $\rho$  and then substituting  $J$  and  $v$  for  $\langle J(t) \rangle_T$  and  $\langle v(t) \rangle_T$ , respectively:

$$\langle \rho(t) \rangle_T = \frac{\langle J(t) \rangle_T}{\langle v(t) \rangle_T}. \quad (2.4)$$

As the average velocity  $\langle v(t) \rangle_T$  tends to overestimate the actual velocity on the road [73], the above equation underestimates the density's actual value.

Another method of determining the density uses the **occupancy**:

$$O_D(t) = \frac{1}{T} \sum_{i=1}^N \tau_i. \quad (2.5)$$

Here,  $\tau_i$  denotes the time the  $i$ th vehicle occupies the detector. Assuming uniform vehicle lengths and the validity of equation (2.3), one can show the proportionality of density and occupancy [38]:

$$\langle \rho(t) \rangle_T = O_D(t) \rho_{\max}. \quad (2.6)$$

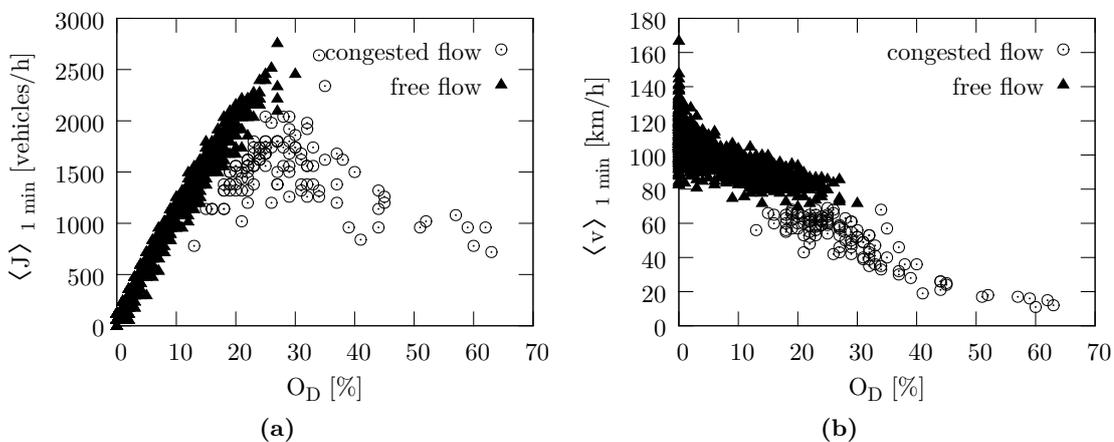
The constant of proportionality  $\rho_{\max}$  denotes the maximum traffic density with a value of between 115 vehicles/km/lane and 200 vehicles/km/lane [41, 97].

As detectors aggregate these data simultaneously, it is possible to study the relation between the three observables  $\rho$ ,  $J$ , and  $v$ .

## 2.2 The Fundamental Diagram

The interdependence between traffic flow and vehicle density is best described by a diagram. The resulting *fundamental diagram*<sup>1</sup> is of great importance in the analysis of traffic flow. Due to the hydrodynamic relation (2.3), any graphical relation between two of the three traffic observables is an equivalent representation of the fundamental diagram. Figure 2.1 shows an empirical fundamental diagram as obtained on a German Autobahn.

<sup>1</sup>Note that some authors, including Kerner [56], use the term “fundamental diagram” to indicate a unique, i.e., functional, relation between density and flow. In the following, we will not distinguish between functional and non-functional relations, and we will use the terms “fundamental diagram” and “flow–density relation” synonymously (e.g., see [123, p. 228]).

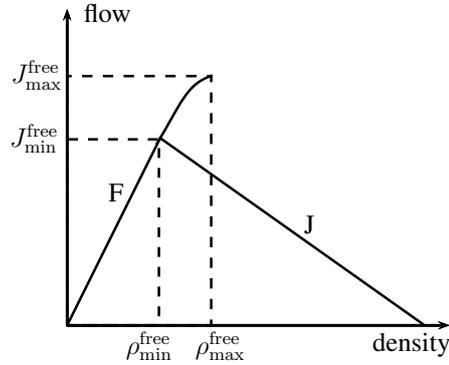


**Figure 2.1** Two equivalent representations of an empirical fundamental diagram as recorded on the left lane by detectors on Autobahn A44: (a) shows the relation between occupancy (i.e., density) and flow, whereas (b) relates the average velocity to the occupancy. Data points with an average velocity of 70 km/h and above are associated with free flow. At lower average velocities, the corresponding measurement is attributed to congested traffic. (The concrete choice of the velocity threshold is to some extent arbitrary but should be between 60 km/h and 80 km/h [15, 69, 138].)

The depicted fundamental diagram exhibits some characteristic features [123]: (i) The maximum flow of approximately 2500 vehicles/h/lane is reached at an occupancy of  $O_D \approx 20\%$ . (ii) The flow in congested traffic usually does not exceed 1800 vehicles/h/lane. (iii) Data points attributed to free flow in figure 2.1a seem to lie on a straight line, whereas congested data points are scattered widely in a two-dimensional region—especially when the value of  $O_D$  lies between 20% and 30%.

Based on the observed distribution of data points, various idealized shapes of the fundamental diagram have been proposed in the past [39, 123]. A prominent example of a theoretical model for the fundamental diagram is shaped like an inverted  $\lambda$  [86], as depicted in figure 2.2.

For low and high densities, its shape is easy to understand. At low densities,  $\rho < \rho_{\min}^{\text{free}}$ , all vehicles can travel at their desired or maximum speed, and additional vehicles linearly increase the total flow. From the fundamental relation (2.3) follows that the slope of the line F marks the vehicles' maximum velocity. At high densities,  $\rho > \rho_{\max}^{\text{free}}$ , on the other hand, vehicles hinder each other and force others to slow down. In this density regime, the flow is a monotonic decreasing function of traffic density. The slope of the congested line J in figure 2.2 corresponds to the typical propagation velocity of a jam's downstream front. The intersection point of both curves (F and J) allows for a very descriptive interpretation, too; it indicates the transition point from jammed to free traffic where vehicles leave the jam. Interestingly, in this outflow region of a jam, the flow rate is considerably below its maximum value  $J_{\max}^{\text{free}}$ . According to empirical

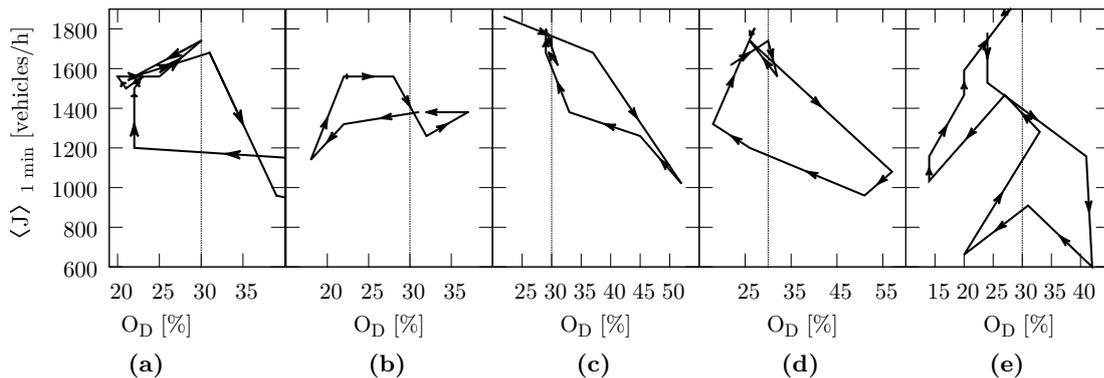


**Figure 2.2** Schematic sketch of a fundamental diagram showing a metastable region between  $\rho_{\min}^{\text{free}}$  and  $\rho_{\max}^{\text{free}}$ . At lower densities, free flow predominates (line F), whereas at higher densities congestion occurs (line J).

results [67], the ratio  $J_{\max}^{\text{free}}/J_{\min}^{\text{free}}$  is approximately 1.5.

At intermediate densities,  $\rho_{\min}^{\text{free}} < \rho < \rho_{\max}^{\text{free}}$ , there exists a metastable range. In this metastable regime, traffic flow may choose either of the two branches (congested or free flow). Thereby, flow can take more than one value at a given density<sup>1</sup>. A possible explanation for this behavior is that in the range  $\rho_{\min}^{\text{free}} < \rho < \rho_{\max}^{\text{free}}$  the drivers' desire to travel as fast as possible and their tendency to hinder each other compete.

The choice of which branch to take, however, does not only depend on the values of flow and density but also on the system's history. Figure 2.3 shows this hysteresis behavior. In all cases, the maximum flow is observed at an occupancy below 30%.

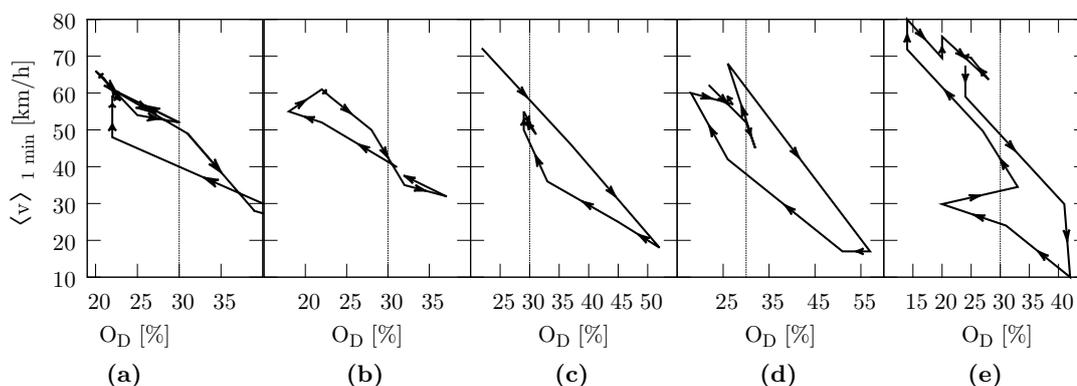


**Figure 2.3** Time-traced flow–occupancy data from two different days (November 2 and 4, 2010) collected on the German Autobahn A44. (Note the different ranges of the abscissae. For the sake of clarity, the position  $O_D = 30\%$  is indicated by a thin, dotted line.)

<sup>1</sup>According to the three-phase traffic theory (see section 2.5), there even exists a two-dimensional region of steady states in the flow–density plane.

The flow can spontaneously switch from free (upper branch) to congested (lower branch) traffic in this density range for no obvious reason, which means that the flow drops considerably. Since this decrease in the flow rate corresponds to an effective reduction of the road's capacity, one often calls this phenomenon *capacity drop*. Before high flows can be restored, the occupancy has to fall below 30%, where the flow rate takes relatively low values at first. Hence, maximum flow can only be reached from the free flow branch and not from the congested one.

What happens when traffic switches from free flow to a congested state? According to equation (2.3), a decrease of the flow with a simultaneous increase of the occupancy must be associated with a drastic drop of the velocity. Figure 2.4, which shows the evolution of the average velocities  $\langle v \rangle_{1 \text{ min}}$  for the time-traced flows of figure 2.3, confirms this conclusion.



**Figure 2.4** Time-traced velocities corresponding to the flows of figure 2.3. (Note the different ranges of the abscissae. Again, the position  $O_D = 30\%$  is indicated by a thin, dotted line.)

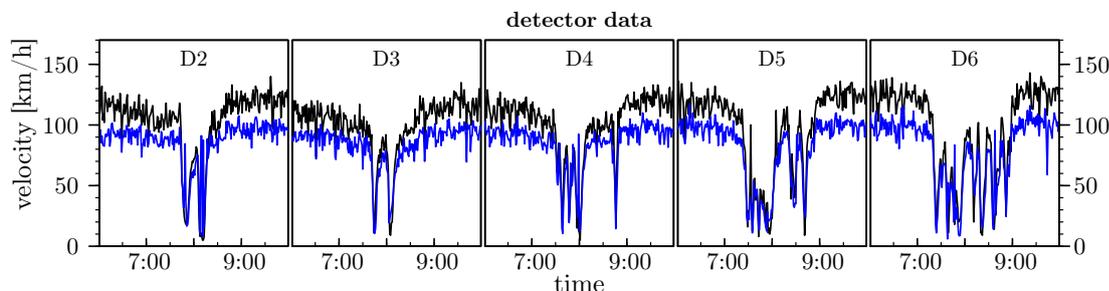
As one can see, the average velocity may drop for more than 50 km/h (see figure 2.4d) from one minute to the next. Due to the abrupt change of flow and velocity, this transition is often called a “traffic breakdown”. From a driver’s perspective this breakdown simply means that he or she gets stuck in congested traffic. The average flow in these cases still measures 1000 vehicles/h/lane. Consequently, dozens or even hundreds of vehicles must be involved. A traffic breakdown, therefore, is not only a collective phenomenon of great practical importance, but it is also interesting from a physical point of view.

The following section will give a brief overview of empirical and theoretical results with respect to traffic breakdown. Its reproducibility with microscopic traffic models will be the subject of the next chapter.

## 2.3 Traffic Breakdown

We have seen that a traffic breakdown describes a sharp drop of the average velocity on a stretch of road (for another example, see figure 2.5). The obvious question of what

causes a breakdown was unanswered for a long time and has been intensively studied.



**Figure 2.5** Time series of several subsequent detector cross-sections showing a breakdown. The black curve shows the velocity time series for the left lane, and the blue curve shows the corresponding time series for the right lane. The time series are ordered by the location of the corresponding detector from upstream to downstream (left to right). (For the detectors’ positions, see figure 3.3.)

The study of single vehicle trajectories reconstructed from aerial photographs [144] suggested that a breakdown occurs spontaneously (i.e., without an accident or construction site). In addition to that, the trajectories showed that the breakdown creates a traffic jam which propagates—just like a wave—against the driving direction at a constant velocity.

The Nagel-Schreckenberg model [102], one of the earliest and probably most simplistic microscopic traffic models that could reproduce these empirical findings, incorporated a stochastic component to reflect the drivers’ incapability to maintain a constant velocity. In [102], it is concluded that, after a critical vehicle density is reached, small perturbations (e.g., a single driver hitting the brakes too hard) suffice to cause a traffic breakdown, forcing all vehicles to slow down.

Empirical evidence for this assumption was given by an intriguingly simple experiment [105, 132]: drivers were asked to maintain a constant velocity and constant space gap while driving on a circular road. After several minutes, the formation of jammed vehicle clusters traveling against the driving direction could be observed—exactly as predicted by traffic models incorporating a probabilistic component such as the Nagel-Schreckenberg model.

Although this experiment proves that traffic jams can, in principle, emerge anywhere on an uncongested road, in reality, traffic breakdowns usually occur in the vicinity of bottlenecks (e.g., ramps, lane closures, or uphill gradients). Consequently, breakdown phenomena at bottlenecks have been intensively studied (e.g., see [22, 51, 57, 68, 125]). Even though the interpretation of the data has sometimes led to partly contradicting conclusions ([68] versus [22] versus [57] and [51, 57] versus [125]), one can identify three ingredients, as Helbing et al. [42] call it, which are necessary for a breakdown to occur:

1. a high traffic volume,
2. a bottleneck (i.e., a spatial inhomogeneity), and

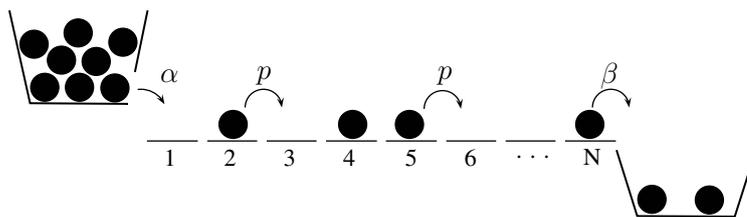
3. a temporary perturbation (e.g., a lane change or natural fluctuations in driving behavior, as discussed earlier).

When these requirements coincide, a breakdown is likely<sup>1</sup> to occur, and freely flowing traffic switches to a congested state with very different characteristics. Such behavior resembles phase transitions in physical systems and raises the question of whether a traffic breakdown represents a phase transition in traffic flow.

## 2.4 Phase Transitions and Traffic Phases

Using the terminology of phase transitions with respect to traffic flow is problematic because traffic definitely is a system far from equilibrium (e.g., non-vanishing net currents). Therefore, the well-known concepts of equilibrium statistical physics cannot be applied.<sup>2</sup>

Yet there are one-dimensional driven particle systems, which show transitions that are similar to a traffic breakdown. The prime example of such systems is the *totally asymmetric simple exclusion process* (TASEP) as sketched in figure 2.6.



**Figure 2.6** Visualization of the totally asymmetric exclusion process (TASEP) with open boundaries, which is a good example of a non-equilibrium (transport) process. Particles enter the system from the left boundary with probability  $\alpha$ . In the bulk, particles move with probability  $p$  to the right if the next site is empty. Particles in the right-most site (labeled  $N$ ) leave the system with probability  $\beta$ . (The variables  $\alpha$ ,  $\beta$ , and  $p$  being probabilities implies that the system evolves in discrete time. In continuous time, the variables denote transition rates rather than probabilities.)

The system consists of  $N$  sites that can be occupied by one particle at most. Particles enter the system from the left boundary with probability  $\alpha$ , they move to an unoccupied neighboring site to their right with probability  $p$ , and they leave the system via the right boundary with probability  $\beta$ . The bulk dynamics of the TASEP and of related systems is governed by the probabilities with which particles enter or leave the system [23, 37, 85, 114], and the resulting phase diagram reveals distinct phases: a low density phase (LD), a high density phase (HD), and a maximum current phase (MC). These phases are separated by first- (LD $\leftrightarrow$ MC and HD $\leftrightarrow$ MC) and second-order (LD $\leftrightarrow$ HD) phase

<sup>1</sup>The breakdown is a probabilistic phenomenon whose probability increases with the traffic volume [15, 54, 111, 112].

<sup>2</sup>For a short introduction to non-equilibrium physics, see the article by Mallick [95].

transitions, respectively. Therefore, the observed phase transitions are called boundary-induced [87]. Depending on the inflow and outflow probabilities of the system, a local perturbation may move either along the flow of particles or in the opposite direction. Compared with vehicular traffic [2], the latter case may be interpreted as a traffic jam propagating upstream. Similarly, the shock, which marks a discontinuity in the density profile, can be seen as the jam's upstream front. In fact, some of the models that are used to describe traffic flow are very similar to the TASEP and related particle hopping models. Hence, we can indeed identify a traffic breakdown as a boundary-induced phase transition [113].

At this point, we have to note that the term “traffic phase” or just “phase” is used in a less strict manner in the traffic-related literature. There, a (traffic) phase denotes a spatiotemporal traffic pattern with very characteristic features [56, Appendix A]. In this sense, one has to understand the term “phase” in the expression “three-phase traffic theory” which denotes the currently most popular and probably best-established traffic theory.

## 2.5 Three-Phase Traffic Theory

The meticulous analysis of empirical traffic data by Kerner and colleagues has led to the development of the three-phase theory of traffic (exhaustively presented in the books by Kerner [54, 56]). Kerner found that the obvious classification of traffic in free flow and congested flow is too coarse-grained because his studies showed a rich variety of spatiotemporal patterns of congested traffic. According to the three-phase traffic theory, congested traffic subdivides into two phases: “synchronized traffic” and “wide moving jams”. Therefore, this theory is based on and culminates in the identification of distinct spatiotemporal patterns of congested traffic with very characteristic properties. The following points summarize the most important properties of the three traffic phases [54, 56]:

**Free flow (F)** is the prevalent traffic phase at low vehicle densities. When the vehicle density is low, vehicle interactions are scarce and weak. Therefore, vehicles can travel at their desired speed without hindering each other. This results in the high average velocities of free flow.

As long as the vehicle interactions are weak, traffic flow increases linearly with traffic density. At higher densities and with stronger interactions, vehicles have to adapt their velocity, which leads to the curvature at the upper end of the line F in figure 2.2.

**Wide moving jams (J)** represent regions of high traffic density and (very) low vehicle speeds. These regions are clearly delimited in space from any neighboring traffic phase by two sharp fronts. At these fronts, density and vehicle speeds change considerably.

The term “moving” means that the jammed region propagates upstream on the road through other traffic phases and even through bottlenecks. While propagating, the

downstream front moves at a constant velocity of approximately  $-15$  km/h [120, 139]. (Slightly higher speeds, ranging from  $-17$  km/h to  $-24$  km/h, were observed by Koshi et al. [86] and Iwasaki [45].)

**Synchronized flow (S)**, from which wide moving jams emerge spontaneously [53], are characterized by intermediate vehicle speeds. Compared with free flow, average velocities are lower in synchronized flow, and yet the flow rate may be nearly as high as in free flow [52, 60].

Unlike wide moving jams, synchronized flow cannot propagate through bottlenecks. Often, it is even fixed at one [52].

This traffic phase is described as “synchronized” because the average velocities on different lanes of a multi-lane road tend to approach each other, i.e., to synchronize.

The analysis of empirical observations has led to several conclusions in the framework of the three-phase traffic theory, which we want to summarize here and to which we will refer in chapter 4. The fundamental hypothesis of the three-phase traffic theory states that the “steady states of the synchronized phase cover a two-dimensional region in the flow–density plane” [54, p. 46]. This could explain the wide scattering of congested states, which can be seen in figure 2.1a. From this follows, however, that a functional relationship between vehicle flow and vehicle density such as equation (2.3) does not exist.

With respect to the transitions between traffic phases, the three-phase traffic theory states that wide moving jams do not emerge from free flow [52]. Wide moving jams spontaneously form from synchronized flow, which is called an  $S \rightarrow J^1$  transition. Similarly, synchronized flow emerges spontaneously from free flow ( $F \rightarrow S$  transition) as a result of a local disturbance. Therefore, a traffic breakdown is always associated with an  $F \rightarrow S$  transition.

### 2.5.1 Three-Phase Traffic Theory and Earlier Modeling Approaches

Furthermore, the development of the three-phase traffic theory followed from Kerner’s assessment that earlier traffic models cannot explain nor adequately describe the breakdown phenomenon and several related spatiotemporal features of traffic flow [54, ch. 10]. He claims that the correct description of these phenomena was a unique feature of his theory, and he speaks of a large number of earlier models as models within the *fundamental diagram approach*. The term *fundamental diagram approach* is to make clear that the steady states of models within this approach lie on a curve in the flow–density plane (or as in figure 2.2: on two curves), as opposed to the two-dimensional plane of steady states postulated by the three-phase traffic theory.

The main criticism Kerner raises against the fundamental diagram approach is that the fundamental diagram combines temporally (and spatially) averaged data [54, ch. 10]. Hence, it does not contain information about the spatial and temporal evolution of

---

<sup>1</sup>Note that both traffic flow and the wide moving jam phase are abbreviated by the letter ‘J’. The flow is set in italics, whereas the traffic phase is set roman.

traffic flow. Moreover, according to Kerner, models within this approach foresee only one desired space gap for vehicles following another one driving with a fixed velocity. Therefore, Kerner concludes that models based on a fundamental diagram cannot correctly reproduce important features of traffic flow, in particular, synchronized flow.

In the following, we will use the terms “model with a fundamental diagram” and “model within the fundamental diagram approach” synonymously to stress that such a model is not supposed to reproduce synchronized flow.

Yet it should be mentioned that studies by Schönhof and Helbing [124, 125] and by Treiber et al. [142] showed that traffic models with a fundamental diagram can also produce a rich variety of spatiotemporal traffic patterns, of which they attributed some to the synchronized flow phase. In addition, such models also offer an explanation for the wide scattering of congested traffic states in the empirical fundamental diagram (cf. figure 2.1): Treiber et al. [143] suggested that drivers choose their time headway (i.e., the temporal headway to the leading car) depending on the current traffic state. This variance-driven time gap produces widely scattered data points in the flow–density plane, too, as they could show for several models within the fundamental diagram approach.

## Modeling Approaches and Breakdown Reproducibility

During the past 50 years, numerous models to describe vehicular traffic have been proposed (e.g., see the review articles [20, 41, 100, 103] and the recent book by Schadschneider et al. [123]). Based on the approach to the complex many-body system traffic, these models can be further classified as macroscopic or microscopic. Macroscopic models treat traffic flow by analogy to the flow of a compressible fluid and, hence, are restricted to study the collective dynamics instead of the individual vehicles' motion. Microscopic models, on the other hand, explicitly model vehicle-vehicle interactions and keep track of every single vehicle. Depending on the area of application, either approach may be favored—possible criteria for the evaluation are the desired level of detail, the computational tractability, or the ability to reproduce particular empirical features of traffic flow.

The pursuit of finding better and better models has led to a multitude of macroscopic and microscopic traffic models. In the field of microscopic models alone, one currently counts more than one hundred such models [16], and their number is still increasing. These models, however, may be subdivided again. Traffic cellular automata [94], in which both space and time are discrete variables, represent a prominent subclass of microscopic traffic models. Together with their rule-based dynamics, they are closely related to the particle hopping models known from non-equilibrium physics (see section 2.4). In contrast, car-following models are continuous in space and time. In this subclass, a vehicle's motion is governed by a differential equation—usually involving the position, the velocity, and the acceleration of the preceding vehicle (for a comprehensive overview of the various modeling approaches, see [12, 123]).

The question of what characterizes a “good” model is not clear though as several levels of detail can be considered: inter-vehicle dynamics (how a vehicle interacts with its immediate predecessor), traffic dynamics (whether the model exhibits the known traffic patterns), or traffic statistics (whether the model is able to reproduce empirical lane

---

This chapter is largely based on and taken from the author's article [T77] (see page 114).

usage and headway distributions). This distinction is necessary as a model performing well in one of these fields is not guaranteed to behave equally well in the others.

Despite the large number of models, there are relatively few studies which actually compare these models with empirical traffic data. Especially in the field of traffic dynamics, which will be the focus of this chapter, there is a lack of comparative studies.

The only comparisons between empirically observed and simulated traffic dynamics that we are aware of were carried out by Treiber et al. [139], Popkov et al. [113], and Kerner et al. [59]. The first two articles, however, focused on single-lane dynamics and did not provide quantitative results. Yet lane changes have an important influence on traffic dynamics, as Kerner and Klenov [61] found by analyzing vehicle trajectories: lane changes between neighboring lanes can be responsible for the emergence (and dissolution) of congested traffic states. The article by Kerner et al. [59] offers a very detailed discussion of traffic dynamics and a qualitative comparison of empirical data with two models based on Kerner's three-phase traffic theory. The authors, however, do not provide a quantitative analysis nor do they discuss the influence of the various model parameters on their results.

It is a very interesting and important question whether traffic models are able to predict the transition from one traffic phase to another. Investigating this question is challenging as such a (phase) transition occurs spontaneously and is not restricted to a certain location on the road. Moreover, it is not clear how to assess the performance of a model compared with empirical traffic data.

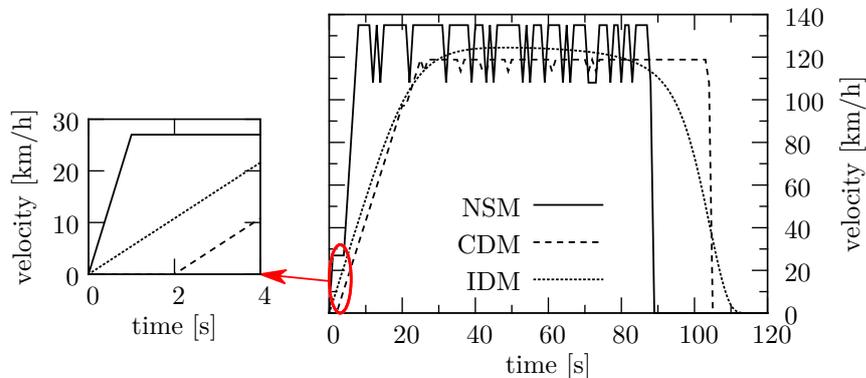
Our analysis of traffic dynamics and the comparison with simulation results is based on detector data. As detector data represent locally aggregated information, they allow no explicit statement on the spatial extent and propagation of traffic states. For this reason, the data of a single detector do not suffice to study the spatiotemporal dynamics. Analyzing the time series of a sequence of neighboring detectors removes this restriction. Therefore, we have chosen a highway section with a sufficient number of detectors for our study. To examine the spatiotemporal traffic dynamics, we have selected two microscopic traffic models that gave good results in previous studies [16, 81]. As a reference, we have also included the Nagel-Schreckenberg model (NSM) [102], which is a rather simplistic traffic cellular automaton (CA) [94].

In this chapter, we will discuss the ability of the three selected traffic models to reproduce the spatiotemporal dynamics of traffic flow. To assess the quality of the results, we will present methods that allow both a qualitative and a quantitative analysis. In this context, we will also discuss some modeling aspects (e.g., boundary setup and vehicle length) and their influence on the observed traffic pattern. For this purpose, we modeled a section of a German Autobahn and used empirical detector data that show a traffic breakdown in the morning peak hour. As we explicitly wanted to consider multi-lane dynamics in heterogeneous traffic, we chose models with asymmetric lane changing rules to mimic the lane changing behavior that is found not only on a German Autobahn but also in most other European countries.

### 3.1 Microscopic Traffic Models

We start with a review of the three microscopic models that we selected for the comparison: the *intelligent driver model* (IDM) [139, 141], the well-known Nagel-Schreckenberg model (NSM) [102], and the *comfortable driving model* (CDM) [79, 83], which is also known as *brake-light model* (BLM).

Although all three models are microscopic ones, they pursue different modeling approaches. The single vehicle dynamics already reveals some of the differences between them. For each model, figure 3.1 shows the velocity profile of a vehicle starting at rest and approaching a parked vehicle 3 km ahead.

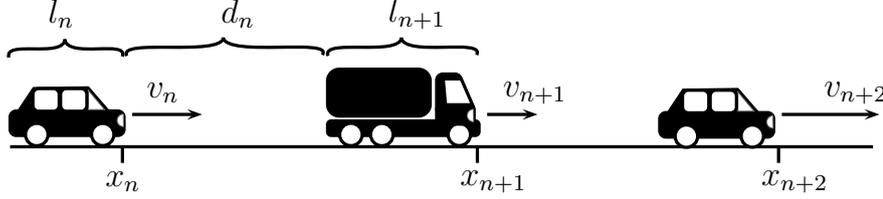


**Figure 3.1** Velocity profiles for a vehicle starting from rest with a standing vehicle 3 km ahead for the NSM, CDM, and IDM. The enlarged figure detail on the left shows the influence of the probabilistic acceleration behavior of the CDM and the NSM. As the vehicles accelerate only with a given probability, it takes 2 s until the vehicle in the CDM starts moving. The same effect causes the slight pause in the NSM’s acceleration process after 1 s and the velocity fluctuations after the vehicle has reached its maximum velocity.

The temporal discretization of the cellular automata models manifests itself in discontinuities of the corresponding velocity profile. The delayed acceleration in the CDM is due to a so-called slow-to-start rule. According to this rule, a vehicle accelerates only with a relatively low probability, when starting from rest. When approaching the parked vehicle, the IDM’s vehicle initiates a smooth braking process with realistic deceleration values. In both CA models, the vehicle comes to rest within one or two time steps (corresponding to 1 s or 2 s) after driving at maximum speed—thus, resulting in unrealistically high deceleration rates. It has to be noted, however, that the resulting high deceleration rates can be observed in all models on roads with multiple lanes and on-ramps. For the CDM, we will investigate the frequency of such unrealistic braking maneuvers more closely in section 4.5, where we will find that they are rather rare.

For the subsequent review of the models, we will refer to the notation of figure 3.2, where vehicles are labeled in ascending order from upstream to downstream. The length

of a vehicle, its current position, and its current velocity are denoted by  $l$ ,  $x$ , and  $v$ , respectively.



**Figure 3.2** Graphical presentation of the variables that are common to all traffic models considered here: the vehicle length  $l$ , its velocity  $v$ , and its position  $x$ . The so-called distance (headway)  $d_n$  is an abbreviation for the expression  $d_n = x_{n+1} - l_{n+1} - x_n$ .

### 3.1.1 Intelligent Driver Model

In contrast to the NSM and CDM, the IDM is a car-following model in continuous space and time<sup>1</sup>. As explained earlier, the motion of vehicles in a car-following model is governed by a differential equation. In the case of the IDM, the rate of change of the  $n$ th vehicle's velocity depends on its current velocity  $v_n$  compared with the vehicle's desired or maximum velocity  $v_n^{\max}$  and on its distance to the preceding vehicle  $d_n$  compared with its desired distance  $d_n^*$  [139]:

$$\frac{dv_n}{dt} = a \left( 1 - \left( \frac{v_n}{v_n^{\max}} \right)^\delta - \left( \frac{d_n^*(v_n, \Delta v_n)}{d_n} \right)^2 \right). \quad (3.1)$$

The desired distance  $d_n^*$  is calculated from the current velocity, the velocity difference to the preceding vehicle  $\Delta v_n = v_{n+1} - v_n$ , and the desired acceleration and deceleration rates,  $a$  and  $b$ , respectively [141, ch. 11]:

$$d_n^*(v_n, \Delta v_n) = d_0 + \max \left( 0, v_n T - \frac{v_n \Delta v_n}{2\sqrt{ab}} \right). \quad (3.2)$$

As one can see,  $d_n^*$  is bounded from below by a minimum headway  $d_0$ . Together with  $d_0$ , the safety time gap  $T$  determines the distance between vehicles in homogeneous traffic flow. The acceleration exponent  $\delta$  in equation (3.1) controls the acceleration rate of a vehicle approaching its maximum velocity  $v_n^{\max}$ . (The smaller it is, the sooner the vehicle chooses an acceleration rate below the maximum value of  $a$ .)

Altogether, equation (3.1) leads to a very smooth driving behavior with almost always realistic values of the deceleration. Typical values for the parameters' values are given in table 3.1.

A general model for lane changes in the IDM, called MOBIL (minimizing overall braking induced by lane change), is presented in [71, 140]. This model offers rules for both symmetric and asymmetric lane changes.

<sup>1</sup>Although the model is continuous in time, the numerical evaluation of the underlying differential equation requires a discretization of time.

**Table 3.1** Typical values of the IDM's parameters taken from [71]. The maximum velocities were rounded to integer values though. The last row also contains a value for the temporal discretization that we used for the numerical evaluation of equation 3.1. (The numerical solution is an approximation to the exact solution, which requires  $\Delta t \rightarrow 0$ .) In contrast to CA models, the value of  $\Delta t$  can be adapted without, at the same time, changing the values of other parameters (e.g., velocity).

parameter		value	unit
$d_0$	minimum distance	2.0	m
$a$	acceleration	1.5	m/s <sup>2</sup>
$b$	desired deceleration	2.0	m/s <sup>2</sup>
$T$	safety time gap	1.2	s
$v_{\text{car}}^{\text{max}}$	max. velocity car	$34.0 \pm 20\%$	m/s
$v_{\text{truck}}^{\text{max}}$	max. velocity truck	$23.0 \pm 20\%$	m/s
$l_{\text{car}}$	car length	4.0	m
$l_{\text{truck}}$	truck length	12.0	m
$\delta$	acceleration exponent	4	
$\Delta t$	temporal discretization	0.25	s

### 3.1.2 Nagel-Schreckenberg Model

As in all cellular automaton models, space and time in the NSM are discrete. One can think of the spatial discretization as space being divided in cells of length  $\Delta x$ , where each cell can either be occupied by a vehicle or empty. Consequently, each vehicle's position is a multiple of the spatial discretization  $\Delta x$ . Vehicle motion in the NSM is the result of the simultaneous application of a set of rules to each vehicle. Each application corresponds to an advance in time of  $\Delta t$ . The NSM's set of rules comprises three steps [102]:

1. **Speed adaptation:** A vehicle  $n$  adapts its speed. This means that it brakes if a collision with the leading car is imminent (i.e., if  $v_n(t) > d_n(t)$ ), and it accelerates as long as it has not reached its maximum speed  $v_n^{\text{max}}$  (i.e., if  $v_n^{\text{max}} > v_n(t)$ ), otherwise:

$$v_n(t+1) \leftarrow \min(v_n(t) + 1, v_n^{\text{max}}, d_n(t)). \quad (3.3)$$

2. **Dawdling:** The introduction of a probabilistic component  $0 < p < 1$  produces random speed fluctuations. This step is to reflect the drivers' inability to maintain a constant velocity (see section 2.3 and [105, 132]):

$$v_n(t+1) \leftarrow \max(v_n(t+1) - 1, 0) \quad \text{with probability } p. \quad (3.4)$$

3. **Vehicle motion:** Finally, the vehicle moves with the velocity that was determined in the previous steps:

$$x_n(t+1) \leftarrow x_n(t) + v_n(t+1). \quad (3.5)$$

Table 3.2 summarizes typical values of the parameters used by the NSM. A comparison with table 3.1 shows that vehicles in the NSM are considerably longer than in the IDM. The reason is that the NSM's values already include the minimum headway that a vehicle takes in dense traffic. Hence, the term vehicle length is somewhat misleading as it rather denotes the minimum space required by a vehicle, and a proper comparison with the IDM should take into account the IDM's minimum distance  $d_0$  as well.

**Table 3.2** Typical values of the NSM's parameters taken from [102, 122]. We did, however, double the length of slow vehicles to better mimic the physical properties of trucks. Moreover, the vehicles' length denotes the minimum space required by a vehicle. Therefore, it is larger than the distance between a vehicle's front and rear end.

parameter		value	unit
$p$	dawdling probability	0.4	
$\Delta x$	spatial discretization	7.5	m
$\Delta t$	temporal discretization	1.0	s
$v_{\text{car}}^{\text{max}}$	max. velocity car	37.5	m/s
$v_{\text{truck}}^{\text{max}}$	max. velocity truck	22.5	m/s
$l_{\text{car}}$	car length	7.5	m
$l_{\text{truck}}$	truck length	15.0	m

In the following, we will use the asymmetric lane changing rules for the NSM that are described in [80, 122]. (Several other asymmetric rule sets have been proposed. For an overview, see the review article [20].)

### 3.1.3 Comfortable Driving Model

The CDM, originally called *brake-light model*, is an advancement of the Nagel-Schreckenberg cellular automaton with extensions for anticipatory driving. Thereby, the CDM enables a vehicle to react more carefully to the preceding one. (In the NSM any preceding vehicle is ignored, unless a collision is imminent.) The model includes anticipatory effects by considering the status of the preceding vehicle's brake light  $b_{n+1}$ , which may be on (i.e.,  $b_{n+1}(t) = 1$ ) or off (i.e.,  $b_{n+1}(t) = 0$ ), by anticipating its velocity  $v_{\text{anti}}(t) = \min(v_{n+1}(t), d_{n+1}(t))$ , and by calculating the effective distance  $d_n^{\text{eff}}(t)$  as

$$d_n^{\text{eff}}(t) = d_n(t) + \max(v_{\text{anti}}(t) - d_{\text{safe}}, 0) \quad (3.6)$$

where the parameter  $d_{\text{safe}}$  governs the effectiveness of the anticipation.

Similar to the NSM, vehicle motion results from the simultaneous application of several rules that are explained in the following:

1. **Acceleration:** In the first step, a vehicle tries to accelerate to its maximum velocity. To avoid unnecessary acceleration, it checks the status of its own and the

preceding vehicle's brake light and compares its time headway  $t_h(t) = d_n(t)/v_n(t)$  to a velocity-dependent interaction horizon  $t_s(t) = \min(v_n(t), h)$ .

$$v_n(t+1) \leftarrow \begin{cases} \min(v_n^{\max}, v_n(t) + 1), & \text{if } b_n(t) = b_{n+1}(t) = 0 \text{ or } t_h(t) \geq t_s(t), \\ v_n(t), & \text{otherwise.} \end{cases} \quad (3.7)$$

2. **Braking:** Here, the vehicle checks whether it actually has to brake and updates the status of its brake light. The function  $\Theta(\cdot)$  denotes the Heaviside step function.

$$v_n(t+1) \leftarrow \min(d_n^{\text{eff}}(t), v_n(t+1)) \quad (3.8)$$

$$b_n(t+1) \leftarrow 1 - \Theta(v_n(t+1) - v_n(t)) \quad (3.9)$$

3. **Determination of randomization parameter<sup>1</sup>  $p$ :**

$$p \leftarrow \begin{cases} p_b, & \text{if } b_{n+1}(t) = 1 \text{ and } t_h(t) < t_s(t), \\ p_0, & \text{if } v_n = 0, \\ p_d, & \text{otherwise.} \end{cases} \quad (3.10)$$

4. **Dawdling:** In the CDM, this step influences both the vehicle's velocity and the state of its brake light. Let  $\xi$  be a (pseudo-)random number, uniformly distributed in  $[0, 1]$ :

$$v_n(t+1) \leftarrow \begin{cases} \max(v_n(t+1) - 1, 0), & \text{if } \xi < p, \\ v_n(t+1), & \text{otherwise.} \end{cases} \quad (3.11)$$

$$b_n(t+1) \leftarrow \begin{cases} 1, & \text{if } \xi < p \text{ and } p = p_b, \\ b_n(t+1), & \text{otherwise.} \end{cases} \quad (3.12)$$

5. **Vehicle motion:**

$$x_n(t+1) \leftarrow x_n(t) + v_n(t+1) \quad (3.13)$$

From the definition of the anticipated velocity  $v_{\text{anti}}(t)$  follows that in the CDM a vehicle's motion does not only depend on the velocity of and the distance to the leading vehicle but also on the distance of the latter to its own predecessor. This allows vehicles to accept time headways of below 1 s when driving at high velocities (also see section 4.1).

Typical values for the parameters described above are given in table 3.3.

The CDM's model description [79, 83] also provides a set of asymmetric lane changing rules.

In an earlier study [81], which analyzed the distribution of time headways in various traffic cellular automata, the CDM showed a better agreement with empirical data than simpler models such as the NSM. A short review and possible criticism of this study are given in section 4.1.

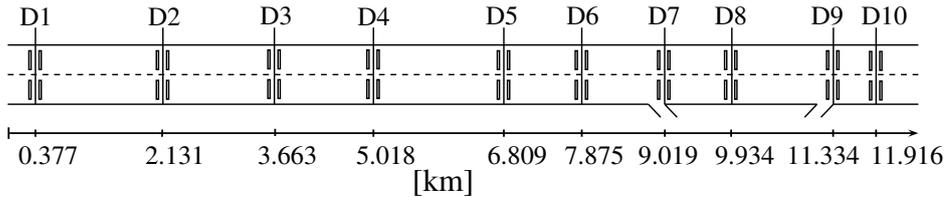
<sup>1</sup>For the meaning of the three dawdling probabilities  $p_0$ ,  $p_d$ , and  $p_b$ , see the explanation to equation (5.6) in section 5.3.2.

**Table 3.3** Typical values of the CDM’s parameters taken from [83]. Similar to the NSM, we doubled the length of slow vehicles. We also increased the maximum velocity of both cars and trucks by 3 m/s to obtain more realistic values with respect to a German Autobahn, but the absolute difference in maximum velocities was preserved.

parameter		value	unit
$d_{\text{safe}}$	effectiveness of anticipation	7	
$h$	brake lights’ range of influence	6	
$p_b$	dawdling probability (over-deceleration)	0.94	
$p_0$	dawdling probability (slow-to-start)	0.5	
$p_d$	dawdling probability (standard)	0.1	
$\Delta x$	spatial discretization	1.5	m
$\Delta t$	temporal discretization	1.0	s
$v_{\text{car}}^{\text{max}}$	max. velocity car	33.0	m/s
$v_{\text{truck}}^{\text{max}}$	max. velocity truck	25.5	m/s
$l_{\text{car}}$	car length	7.5	m
$l_{\text{truck}}$	truck length	15.0	m

## 3.2 Empirical Data and Model Testing

To investigate the spatiotemporal traffic dynamics of the selected models, we used data from ten detector cross-sections on the German Autobahn A44 between the cities of Unna and Werl. A schematic sketch of the 13 km-long, two-lane highway segment is depicted in figure 3.3.



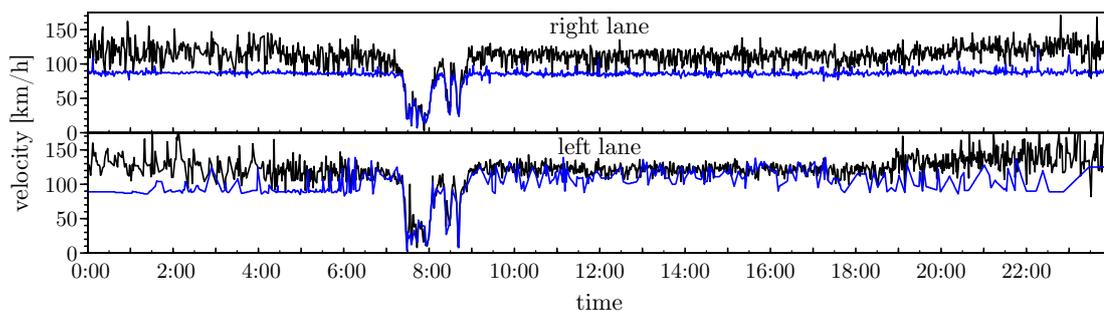
**Figure 3.3** Schematic sketch of the highway section considered. The detectors are labeled as D1, . . . , D10.

This section is well suited for our analysis as it contains a large number of detectors and a single off-ramp and single on-ramp at its downstream end, which serve as bottlenecks. Both ramps can potentially trigger a traffic breakdown. The upstream cross-sections allow measurements of the traffic patterns generated at the bottlenecks without perturbations by additional bottlenecks.

The detectors on this section distinguish two vehicle classes, namely “trucks” and “cars”. Yet there is no strict rule that this classification is based on [18]. For each vehicle class, the detectors measure the number and the average velocity of all vehicles that

pass the corresponding cross-section per 1 min-interval.

For a realistic modeling of open boundary conditions, which includes the inflow and outflow via the two ramps, we used the detector data of November 4, 2010. These data show a spontaneous traffic breakdown in the morning peak hour traffic. The time series of detector D5, as depicted in figure 3.4, is exemplary for all detectors upstream the on-ramp: after the breakdown occurred at approximately 7:15 a.m., upstream vehicles approaching the on-ramp had to slow down causing a congested traffic pattern that travels upstream. The inflow from the downstream boundary was approximately 36 000 vehicles during the entire day with additional 7000 vehicles joining the road via the on-off-ramp system. Vehicles leaving the system as required by the boundary conditions were selected on the basis of their vehicle type according to the empirical data. Similarly, the measured data determined the type of entering vehicles. Vehicles entering the road via the on-ramp increase the probability of a breakdown as they are likely to provoke perturbations that are the cause of a breakdown.



**Figure 3.4** The velocity time series of detector D5 separated by lane and vehicle type. The time series of cars is depicted with a solid black line. The corresponding time series of trucks is depicted with a solid blue line. These time series are exemplary for all detectors upstream of the bottleneck. The breakdown occurs at approximately 7:15 a.m., it affects both lanes, and the subsequent congested traffic pattern lasts approximately 90 min. Microscopic traffic models should be able to reproduce this breakdown and the subsequent traffic pattern. (Missing data points, usually resulting from the lack of trucks in dilute traffic in the left lane, were approximated by a linear interpolation for better readability.)

By aggregating the same information as the real detectors, the simulations yield an equivalent set of data. To compare the resulting data, we propose two very intuitive approaches. Let  $V_{e/s} = \{v_{e/s}(t_1), v_{e/s}(t_2), \dots, v_{e/s}(t_n)\}$  denote the empirical (subscript e) and the simulated (subscript s) time series of a detector. By regarding the two sequences of length  $n$  as  $n$ -dimensional vectors, Brockfeld et al. [17] suggested using the 1-norm  $L_1$

$$L_1 = \sum_{i=1}^n |v_e(t_i) - v_s(t_i)| \quad (3.14)$$

as a direct error measure. For the similarity analysis of two time series, however, the

shape of the series is, in general, more important than their absolute values. Otherwise, a poorly chosen upper boundary for vehicle speeds (see tables 3.1, 3.2, and 3.2) would result in a large error margin even if the traffic dynamics (i.e., the sequence and duration of traffic phases) is correctly reproduced. Hence, the different values of the maximum velocities in free flow and random fluctuations should be filtered, so that they do not enter the comparison. Therefore, and according to [33], we suggest a normalization of the series before calculating the 1-norm of equation (3.14). The normalization of a time series  $X = \{x_1, x_2, \dots, x_n\}$  with mean  $\mu(X)$  and standard deviation  $\sigma(X)$  is obtained by the affine transformation

$$x_i \rightarrow \frac{x_i - \mu(X)}{\sigma(X)}. \quad (3.15)$$

One should mention that the 1-norm is just one way of determining the distance between two time series. In general, any symmetric and positive-definite metric for which the triangle inequality holds could be applied [33].

Anyway, the just described method compares the absolute values of two time series and, therefore, requires a proper normalization of the time series' values. If one only wants to know whether the two time series show a similar temporal evolution, a normalization is not necessary. A classical approach to analyze time series is the decomposition of the series into three components [93]: a trend component, a seasonal component, and a noise component. Thus, if both series are similar, this means that the trend and seasonal components of the two series are similar. Consequently, to assess the similarity between the empirical time series and the one obtained from our simulations, we can study the time series' residuals  $R(t_i)$ :

$$R(t_i) = v_e(t_i) - v_s(t_i). \quad (3.16)$$

The residuals denote the difference between the observed and the simulated value at each time step. If the model's time series reproduces the empirical data, the calculation of the residuals cancels the trend and seasonal components and leaves only the noise term. Under the assumption of white noise, the residuals are expected to behave randomly and to be uncorrelated. Similarly, it is also possible to determine the correlation between the two series. In this case, the degree to which the simulated time series reproduces the empirical data can be determined by the cross-correlation of the two series.

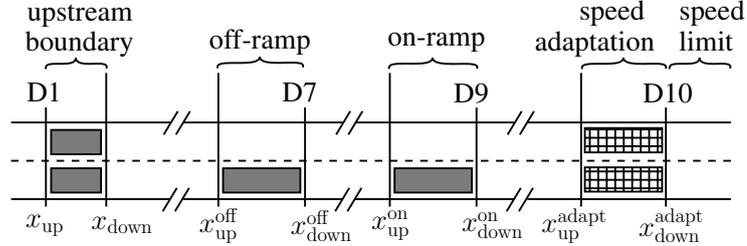
### 3.3 Simulation Setup

#### 3.3.1 Modeling Open Boundaries and Ramps

The modeling of boundaries and ramps still offers several degrees of freedom and can significantly influence the resulting traffic patterns. Apart from the already mentioned boundary-induced phase transitions (section 2.4), the boundary setup can affect and limit the spectrum of observable traffic states [8].

Therefore, and for better reproducibility of our results, we want to give a concise description of the boundary and on-ramp setup. To insert new vehicles, we define an

entrance section of length  $l = x_{\text{down}} - x_{\text{up}}$ , where  $x_{\text{down}}$  ( $x_{\text{up}}$ ) stands for the downstream (upstream) end of the entrance section (see figure 3.5).



**Figure 3.5** Graphical representation of the boundary setup. Gray rectangles mark the segments where vehicles are inserted or removed. The checked rectangles mark the regions where vehicles decelerate (if necessary) so that they pass detector D10 with the velocity required by the speed limit. (For the sake of completeness, we give the following numerical values:  $x_{\text{up}} = 0.377$  km,  $x_{\text{down}} = 0.4895$  km,  $x_{\text{up}}^{\text{off}} = 8.793$  km,  $x_{\text{down}}^{\text{off}} = 9.018$  km,  $x_{\text{up}}^{\text{on}} = 11.108$  km,  $x_{\text{down}}^{\text{on}} = 11.334$  km,  $x_{\text{up}}^{\text{adapt}} = 11.466$  km, and  $x_{\text{down}}^{\text{adapt}} = 11.916$  km.)

A vehicle entering the system is placed in the largest gap of the entrance section if the insertion is safe. The insertion is safe if neither the inserted vehicle's deceleration nor the following vehicle's deceleration falls below a threshold  $b_{\text{th}} = -1 \text{ m/s}^2$ . (In the IDM, one has to check the other lane as well due to the so-called longitudinal-transverse coupling [71].) After the insertion, the new vehicle adapts its velocity to the one which is obtained by averaging the leading and the following vehicle's current velocity. This strategy can be applied both to the on-ramp and to the upstream boundary.

Similarly, vehicles leaving the road via the off-ramp are selected from an exit section. The insert and exit sections representing the on-ramp and the off-ramp are restricted to the rightmost lane, whereas the upstream boundary's entrance section spans both lanes.

To mimic the traffic state at the downstream boundary, we decided to apply a dynamic speed limit from the position of detector D10 to the end of the road. The speed limit equals the maximum velocity that detector D10 measured during the previous aggregation interval. (As the CA models require an integer value for the speed limit, we converted the value by rounding it up.) To allow for a smooth deceleration with respect to the speed limit, we defined a deceleration section upstream of detector D10. Vehicles within the deceleration section gradually reduce their velocity such that they pass detector D10 with the desired velocity.

Vehicles are inserted according to the empirical data of detector D1. Consequently, the open boundary's entrance section starts at the position of detector D1. For the length of the entrance section we used a value of 112.5 m, which is half the value we used for the on-ramp's (off-ramp's) entrance (exit) section. As there are no detectors directly measuring the vehicles entering or leaving the road via the ramps, we determine these values by comparing the values of detectors D7 and D9 with each other. Detector D7 is situated immediately downstream of the off-ramp and detector D9 immediately

upstream of the on-ramp. Thus, if detector D9 measured more vehicles than detector D7, the surplus is inserted in the on-ramp section. Otherwise, the additional vehicles are removed from the exit section. In principle, it would also be possible to treat the D7- and D9-detector data separately. In this approach, however, we might remove a vehicle from the exit section and insert a new one in the entrance section at the same time. As each insertion represents a perturbation of traffic flow, which we want to minimize, this approach is less favorable. Another difficulty is that an insertion or deletion may fail (e.g., if the safety constraints prohibit an insertion or if there is no vehicle in the exit section). In such a case, the insertion or deletion of a vehicle is retried in subsequent time steps until it succeeds so that the total number of vehicles is conserved.

Moreover, trucks were not allowed to overtake in our simulation for two reasons. First, with asymmetric lane changing rules slower vehicles are required to stay in the right lane by legislation. Second, the examined road section is equipped with dynamic traffic signs which may impose an overtaking ban for trucks if traffic demand is high. Consequently, trucks were inserted only in the right lane at the upstream boundary. This is supported by the fact that, in the empirical data, the share of trucks in the left lane, averaged over all detectors, was below 3.6%. (The highest share of trucks was observed at detector D6 with 7.4%.) Figure 3.4 supports this finding: for long periods, the trucks' time series has a piecewise linear shape. The fact that the applied linear interpolation is so clearly visible indicates a low number of trucks.

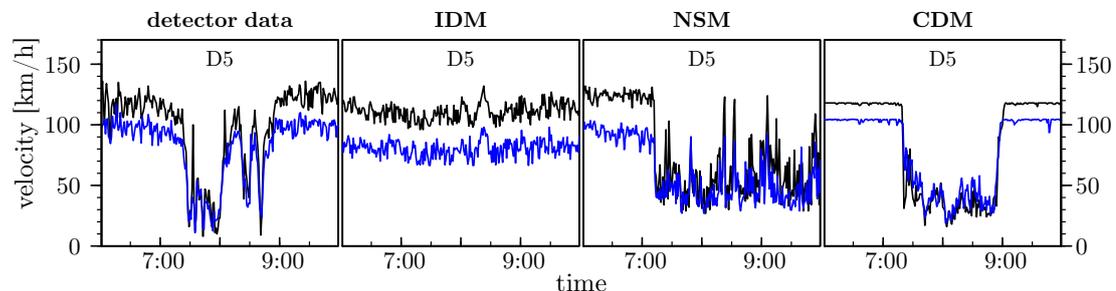
### 3.3.2 Calibration

In contrast to previous publications [16, 17, 117], we tried to avoid an automated model calibration, as such a calibration does not only question the reliability of the obtained values [16] but can also produce parameter sets whose the degree of realism can sometimes be doubted (see, e.g., the optimized parameter sets in [117]).

Due to various factors (e.g., weather, road characteristics, or different drivers), one cannot forgo a calibration completely though. For the necessity of a calibration see figure 3.6, which shows the actual time series at detector D7 and the velocity time series of the uncalibrated models at the same position. As one can see, the NSM overestimates the duration of the congested traffic following the breakdown, whereas the IDM does not show a breakdown or congested traffic at all.

If the properties of the vehicles (i.e., length and maximum velocity) are to remain unchanged, the calibration is obvious in the case of the NSM because the model has only one parameter left to calibrate (see table 3.2). This parameter controls the frequency of random fluctuations in traffic flow. Hence, we slightly reduced its value from 0.4 to 0.32.

Despite its larger number of parameters, a calibration of the IDM was straightforward as well: we increased a driver's desired time headway from 1.2 s to 1.5 s. Thereby, we reduced the road's effective capacity. (This change is still in good agreement with empirical data [70].) In this case, however, we found that many drivers were no longer able to enter the main road via the on-ramp due to the previously defined safety criteria. Therefore, we changed the threshold of the acceptable deceleration of following vehicles to  $b_{\text{th}} = -50 \text{ m/s}^2$ . (Higher values of  $b_{\text{th}}$  did not yield satisfying results.)



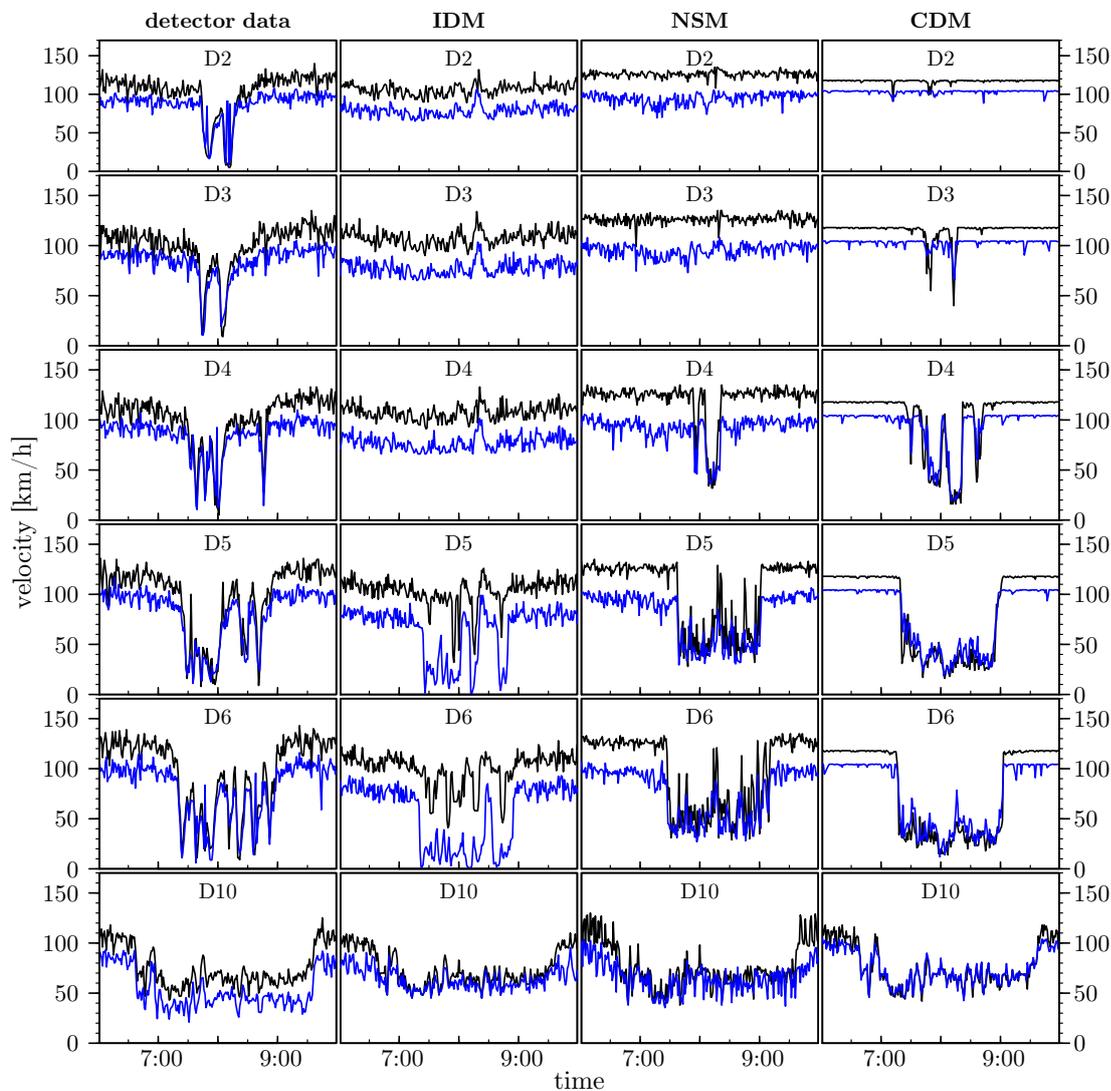
**Figure 3.6** The empirical time series (left) and the time series produced by the three models with the original sets of parameters. The time series of the left lane is colored black and the corresponding time series of the right lane is colored blue.

### 3.4 Results

A first evaluation of the models' behavior is best carried out by a graphical inspection of the resulting time series. Apart from the empirical time series, figure 3.7 shows the corresponding time series of several detector cross-sections for all models during the morning peak hour. The empirical time series of detectors upstream of the bottleneck (i.e., detectors D2 – D7) show a breakdown at 7:15 a.m., and the resulting congested traffic pattern lasts until 9:00 a.m. During this period, the average velocity on the faster left lane abruptly drops below the free flow velocity on the slower right lane, and the velocity synchronizes across lanes. The temporal extent of congested traffic as well as the severity of the breakdown decrease with growing distance to the bottleneck (D6 → D2).

All models' time series show a very good agreement with the empirical data at the downstream boundary (detector D10). This is, of course, a consequence of the dynamic speed limit which is imposed on the downstream boundary (see section 3.3.1). Upstream of the bottleneck region, however, deviations between the models become obvious. The IDM's time series does not show a synchronization between the two lanes. Although one can observe a breakdown in either lane, traveling in the left lane is always faster than in the right lane. This is a result of the very aggressive attitude ( $b_{\text{th}} = -50 \text{ m/s}^2$ ) with which drivers from the on-ramp enter the right lane of the main road. Greater values of  $b_{\text{th}}$ , however, led to the suppression of the breakdown. In contrast, both cellular automaton models reproduce the lane-synchronization following the traffic breakdown even for  $b_{\text{th}} = 0 \text{ m/s}^2$ . This follows from the models' weaker safety constraints, which also result in the very unrealistic deceleration behavior observed in figure 3.1.

Apart from the lack of synchronization across lanes in the IDM, another fundamental difference is the origin of velocity fluctuations in free traffic flow: the CDM and NSM incorporate random fluctuations by a stochastic component, whereas the fluctuations in the IDM result from the heterogeneity of the traffic (i.e., from the uniformly distributed maximum velocities of vehicles; see table 3.1). For the first two models the strength of velocity fluctuations is governed by their spatial discretization, which explains why the

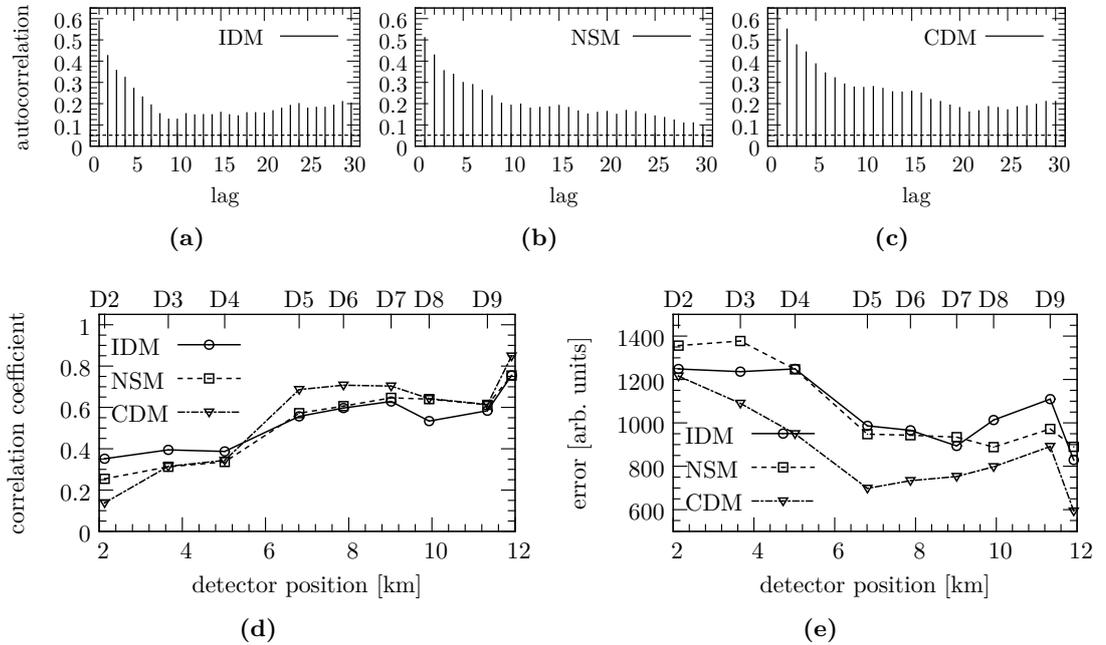


**Figure 3.7** A comparison between empirical time series (first column) and the corresponding ones obtained with the IDM, NSM, and CDM (second to fourth column), respectively. From top to bottom, the rows show the time series for detectors D2, D3, D4, D5, D6, and D10. The average velocities for the left (right) lane are plotted with a black (blue) line. Please note that, for the sake of clarity, we limited the depicted time to the interval from 6 a.m. to 10 a.m. Our analysis, however, is based on the time series of the entire day.

fluctuations in the CDM's time series are less pronounced than for the NSM.

As one can see from figure 3.7, the duration of congestion (i.e., low average velocities) decreases in the upstream direction in both the empirical data and the data generated by the traffic models. All models underestimate the breakdown's spatial extent though. At detector D2, free flow is restored in all models, whereas the empirical data still show the influence of the breakdown.

Similar to the graphical inspection, the numerical analysis does not reveal qualitative differences in the models' velocity time series. Figures 3.8a–c show the residuals' autocorrelation for each model at detector D5, which is representative for the other detectors upstream of the on-ramp.



**Figure 3.8** Autocorrelation of the time series' residuals at detector D5 for (a) the IDM, (b) the NSM, and (c) the CDM. Assuming white noise, the 95%-confidence interval is plotted as a dashed horizontal line. The cross-correlation between the empirical time series and the models' time series is shown in (d). The 1-norm according to equation (3.14) is given in (e).

Ideally, the residuals are uncorrelated, yielding an autocorrelation coefficient close to zero. All models show a positive correlation ( $> 0.5$ ) for the residuals, which decreases with increasing time-lag. Hence, the assumption of white noise for the residuals has to be rejected, and a perfect reproduction of the empirical data could not be observed.

Figure 3.8d shows the correlation between the models' time series and the empirical data for each detector location. Due to the imposed downstream boundary conditions, the correlation coefficient's maximum value ( $> 0.7$ ) is obtained at detector D10 for all models. For all models the correlation coefficient remains above 0.5 up to detector D5

indicating a satisfying reproduction of the empirical data. The decrease in correlation with increasing distance to the ramps reflects the fact that no model could successfully reproduce the spatial extent of the breakdown, as already discussed.

Probably the most intuitive comparison is obtained by the calculation of the 1-norm after transforming the time series according to equation (3.15). Remember that our primary quality criterion is the reproduction of the breakdown and the subsequent congested traffic pattern. Hence, we expect the 1-norm to indicate to what extent the corresponding model achieves this goal. Figure 3.8e shows the 1-norm calculated for each detector and averaged over both lanes. At detector D2 (i.e., at kilometer 2.131), the error for all models is about 1300 (measured in arbitrary units) as all models fail to reproduce the breakdown at this position. Again, due to the boundary conditions, the best accordance with the empirical measurements is observed at detector D10. The qualitative behavior between the detectors D1 and D10 is similar for all detectors. Even the absolute values of the NSM and the IDM are nearly identical; only the CDM's values are considerably below the values of the previous models and indicate a better performance of this model.

### 3.4.1 Reproducibility of Upstream Boundary Conditions

So far, we have seen that all investigated models are, in principle, able to reproduce the spatiotemporal traffic dynamics. As the simulations used real detector data, it is worth discussing to what extent the empirical boundary conditions could actually be reproduced. In section 3.3.1 we described how vehicles were inserted into the road segment, and we stated that a failed insertion had to be repeated in subsequent time steps until it succeeded. Of course, one would expect an insertion to always be successful as the empirical data prove that real traffic can satisfy the observed demand. In the simulations, we saw that insertions repeatedly failed, especially during the peak hour period: table 3.4 gives the maximum number of cars waiting for insertion or removal at the on-ramp or the off-ramp, respectively.

**Table 3.4** Overview of the length of the on- and off-ramp queues (i.e., the number of vehicles waiting for insertion or removal).

model	max. queue length [veh] at		max. queue duration [min] at	
	on-ramp	off-ramp	on-ramp	off-ramp
IDM	139	41	82	22
NSM	63	42	14	9
CDM	55	42	10	41

Moreover, we have recorded the maximum time for either ramp during which the queue of vehicles waiting for insertion or removal was above zero. For instance, the maximum number of vehicles waiting for insertion in the NSM was 63 and the time span of a non-empty queue was 14 min. The results of table 3.4 confirm our previous

statement that the stricter safety conditions of the IDM make vehicle insertion more difficult. The maximum queue length for the IDM was twice as large as those for the two cellular automaton models. For all models, the queue length at the on-ramp reached values considerably above zero. Whether a vehicle insertion fails depends to a large extent on each model's safety conditions. Furthermore, the queue length's growth reflects, at least partially, the fact that the models do not perfectly mimic human driving behavior. In real traffic, for example, it has been found that the time headway of vehicles traveling at 110 km/h can even drop below 0.5 s [3]. Such values are impossible for all models due to their safety constraints. (In real traffic, such headways pose an actual risk, as drivers cannot react to actions of the preceding vehicle in a timely manner.)

### 3.4.2 Propagation Velocity of Jammed Traffic

Another characteristic property of traffic flow is the propagation velocity of traffic jams. A jam is a moving traffic pattern that travels against the direction of traffic flow. In chapter 2 we have already mentioned that a traffic jam is delimited by two sharp fronts between which a high vehicle density and low average velocities predominate. There is also a physical interpretation [2, 37, 113] for the velocities that both fronts propagate with: the propagation velocity of the upstream front corresponds to the *shock velocity*. The shock (i.e., the upstream front) marks the transition region, where vehicles approaching the jam have to slow down and where the vehicle density increases abruptly. The velocity that the jam (i.e., its center of mass) propagates with is given by the velocity of the downstream front. For traffic jams, which travel upstream, the sign of this *collective velocity* has to be negative. This can be seen from figure 3.7, where the traffic breakdown occurs the later, the further upstream the corresponding detector is located.

To measure both the jam propagation (i.e., collective) velocity and the shock velocity, we have to find the times when these fronts pass a detector. As the upstream and downstream fronts of a traffic jam are associated with sharp discontinuous changes of the average velocity, we define the time when the shock reaches the detector as the moment when the average velocity first drops below a given threshold  $v_{\text{th}}$ . Similarly, when the average velocity first exceeds  $v_{\text{th}}$  again, we assume that the downstream jam front has passed the detector. The numerical values of the propagation velocity of both the shock and the traffic jam can be obtained from the quotient of the distance between subsequent detectors and the time that passed between the corresponding events being registered at either detector. In accordance with Rehborn et al. [120], we have set  $v_{\text{th}} = 30$  km/h. As the temporal resolution of 1 min is relatively coarse, we averaged over the values of  $v_{\text{th}}$  from 27 km/h to 33 km/h and obtained the results of table 3.5.

Although the propagation velocities vary strongly between models and between successive detectors, all models correctly predict a negative propagation velocity (i.e., the fronts move upstream against the direction of traffic flow). Due to the relatively small inter-detector distances and due to an approximate error of 30 s in the temporal detection of jam fronts, the error margins are quite large though. In principle, the error can be decreased by considering two detectors far away from each other (e.g.,

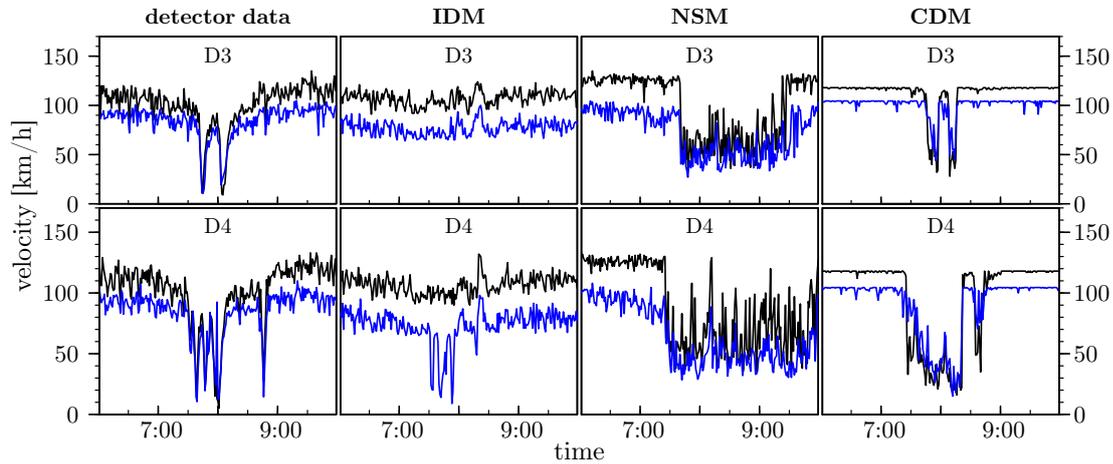
**Table 3.5** Averaged breakdown propagation velocities. The velocities are given in units of km/h.

detectors	shock velocity for				jam propagation velocity for			
	real data	NSM	IDM	CDM	real data	NSM	IDM	CDM
D7→D6	-23	-7	-5	-2	-15	-17	-24	-9
D6→D5	-12	-32	-14	-7	-11	-28	-3	-7
D5→D4	-11	-3	NA	-3	-13	-3	NA	-3

D7→D2): for the empirical data set we found the propagation velocity to be  $-13$  km/h for D7→D2 and D6→D2, which is in good agreement with the value of  $(-15 \pm 2)$  km/h found by Rehborn and colleagues [120]. (A similar analysis of the simulation data was not possible as the jammed traffic pattern did not even reach detector D3.)

### 3.4.3 Influence of Truck Length and Downstream Boundary Condition

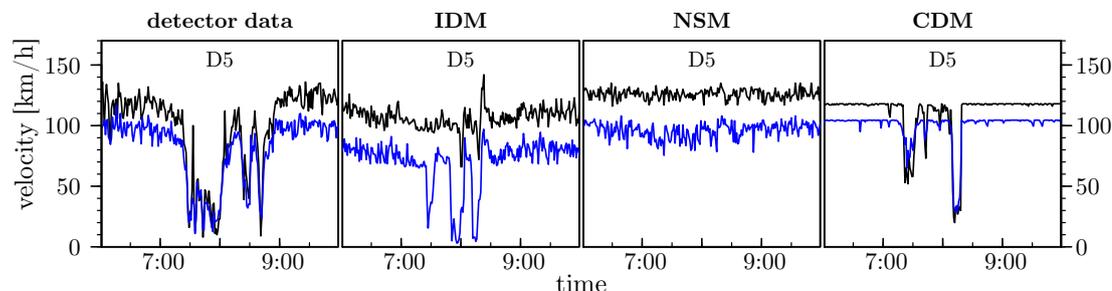
Finally, we demonstrate the influence of vehicle length and the modeling of the downstream boundary on our results. The vehicle length directly influences the spatial extent of the breakdown. Figure 3.9 shows the velocity time series of simulation runs where we increased the length of trucks by 7.5 m for the NSM, by 1 m for the IDM, and by 1.5 m for the CDM.

**Figure 3.9** After increasing the length of trucks, the breakdown can be observed at detectors further upstream (cf. figure 3.7). (The time series of the left lane is colored black, and the corresponding time series of the right lane is colored blue.)

A comparison with figure 3.7 proves that this slight change of the vehicle properties results in a better agreement with the empirical data. For the NSM not only the

spatial extent of the breakdown but also the duration of the subsequent congested traffic increases. Due to the NSM's spatial discretization, the vehicle length can only be altered by multiples of 7.5 m, which corresponds to a massive change in the road's maximum capacity.

Compared to the spatial extent of the breakdown, its origin is a more fundamental modeling aspect. On the downstream boundary we had imposed a variable speed limit to mimic the velocities observed in real traffic. Hence, it is worth investigating whether the breakdown (and the subsequent traffic pattern) resulted from the imposed speed limit or whether it resulted from the vehicle interactions in the on-ramp region. Therefore, we repeated the simulations without the speed limit and, thereby, guaranteed free flow at the downstream boundary. The resulting exemplary time series of detector D5 is depicted in figure 3.10.



**Figure 3.10** The time series of detector D5. When we impose free flow at the downstream boundary, the intensity of the breakdown and the duration of congestion decreases in all models. For the NSM the breakdown even completely disappears. (The time series of the right (left) lane is colored blue (black).)

The NSM's time series no longer exhibits a breakdown, whereas the IDM's and CDM's time series still do. Consequently, the source of the breakdown in the NSM was our choice of the downstream boundary's setup, whereas the breakdowns in the CDM and IDM are caused by the on-ramp boundary. However, for the latter two models the shape of the time series changed, too: the duration of congestion decreased. Therefore, even if it does not select the traffic state (i.e., congested or freely flowing), the downstream boundary does determine its duration.

### 3.5 Discussion

Our analysis shows that all investigated models can reproduce the breakdown and the associated spatiotemporal traffic dynamics satisfyingly—at least after minor calibrations. Although the three models are microscopic ones, this result is quite surprising as there are still substantial differences between the models (e.g., continuous versus discrete in space, realistic versus unlimited braking capacities, or deterministic versus stochastic approach). In particular, the NSM did not fall back behind the other two, more recent models, despite its very simplistic and coarse-grained approach.

Similarly, all models showed the same shortcomings. Compared with the empirical data, they underestimated the spatial extent of the breakdown. In real data, the breakdown could be detected more than 8 km upstream of the on-ramp, which none of the models could reproduce. We saw, however, that the spatial extent of the breakdown can be effectively adjusted by the lengths of the different vehicle types.

Adjusting the length of vehicles, whose number is given by the boundary conditions, has a direct impact on the vehicle density though. Thereby, it also directly influences the probability with which vehicles can enter the road from the on-ramp. In this context, we could observe qualitative differences between the models: in peak-hour traffic (i.e., a high vehicle density), vehicles could more easily enter the main road via the on-ramp in the NSM and the CDM than in the IDM. The better reproduction of the empirical inflow rates is a direct consequence of the NSM's and CDM's unlimited braking capacity. The IDM shows a realistic deceleration behavior but makes higher demands on a safe entry. Therefore, it is more difficult for vehicles to enter in high density phases (e.g., peak hour traffic). Hence, the length of vehicles cannot be chosen freely to adjust the spatial extent of the congested traffic pattern.

In contrast to the CDM and IDM, we found that the breakdown's occurrence in the NSM did strongly depend on the modeling of the downstream boundary. Without applying a speed limit to the downstream boundary, we could no longer observe a breakdown in the NSM. From a physical point of view, this is a major difference between the models. The consequences for practical applications, however, are less severe: for large scale traffic simulations (e.g., [36, 136]), the downstream boundary conditions are given by the traffic state of the neighboring road segment and cannot be chosen arbitrarily. Therefore, if the primary concern is the reproduction of the current traffic state, its physical origin may be of secondary importance.

As a whole, we rate the findings of this chapter as very important because we did not only show that the traffic models can *in principle* show a traffic breakdown, but we also showed that the breakdown occurs at the same inflow and outflow rates at which the breakdown was observed in *real traffic*.

In addition, the methods presented to model the open-boundary highway segment with ramps are formulated in such a general manner that they can be used with or easily transferred to a large number of microscopic traffic models.

As all models gave good results compared with the empirical data, it is not easy to make a definite recommendation for one of them. The highest simulation speeds (measured by the time needed to run the simulation) were observed with the NSM. Aspects related to secure driving (e.g., to study accident-avoidance systems), which require a realistic acceleration and deceleration behavior, are best answered with the IDM. If the focus is rather on a realistic description of the spatiotemporal traffic dynamics, then the CDM is a good alternative. It is a fast, easily extendable, and still realistic model.

The CDM showed the least error according to the 1-norm (figure 3.8e) at all detectors and the strongest correlation with the empirical data from detectors D6 to D10. Besides that, the model also performed very well in reproducing the empirical

boundary conditions (table 3.4) and showed a synchronization across lanes, unlike the IDM. It is important to stress that these results were obtained without calibrating the model's parameters. By working with the unchanged parameter set, most of the CDM's original results [79, 83] remain unaffected.

Therefore, we decided to further investigate and use the CDM in the following chapters. By this choice, we expect scientifically sound statements on how alternations in driving behavior influence traffic flow.



## Traffic Phases and Phase Transitions in the CDM

In the previous chapter 3, we concluded that all investigated models can reproduce the traffic breakdown satisfyingly. Yet all three models were developed more than a decade ago, and they are based on the *fundamental diagram approach*. This means that the steady states of these models lie on one-dimensional curves in the flow–density plane.

In section 2.5.1 we have explained that such a relation between flow rate and vehicle density contradicts Kerner’s fundamental hypothesis of the three-phase traffic theory. According to Kerner, there exists a two-dimensional region of steady states in the flow–density plane: the “synchronized flow”-phase S.

Therefore, we will provide a more detailed analysis of the CDM’s dynamics in this chapter. In particular, we will focus on the spatiotemporal traffic patterns resulting from different inflow and outflow probabilities of an open road segment. Based on time series of local measurements, the local traffic states are assigned to the different traffic phases of Kerner’s three-phase traffic theory. For the CDM or any other microscopic traffic model, this analysis is particularly interesting because different scales have to be considered. On the one hand, microscopic traffic models are based on the assumption that a vehicle’s motion is governed by the next (or the next two) vehicles ahead. Hence, from a single vehicle’s perspective it makes no difference in these models whether there are only two or two thousand vehicles on the road. In contrast to this microscopic modeling approach, many empirical studies of traffic focus on macroscopic phenomena. Emerging behavior in traffic (e.g., traffic breakdown, jam formation) requires the coordinated motion of hundreds of vehicles (see the recent review article [12], which also discusses current challenges on traffic modeling). It is far from self-evident that microscopic models provide good results with respect to macroscopic traffic phenomena. In this context our investigation and classification of traffic patterns has to be understood. We thoroughly investigate the macroscopic traffic patterns resulting from the CDM’s microscopic dynamics, with special attention being paid to the identification of synchronized traffic.

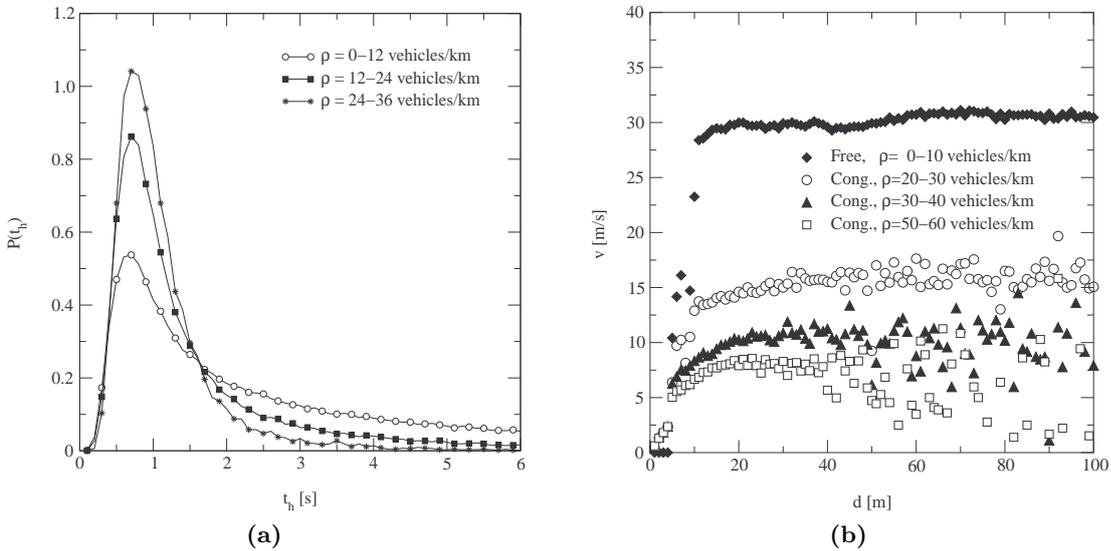
---

This chapter is largely based on and taken from the author’s article [T78] (see page 114).

## 4.1 Earlier Results

First, we will review and discuss some earlier results of the CDM. In 2004, Knospe et al. [81] proposed an empirical test for traffic cellular automata models. The test was based on empirical single vehicle data [82, 108, 134] collected by loop detectors. To assess the quality of the investigated models, they studied how well these models could reproduce the empirical data.

The selected data comprised the distribution of time headways and the relation between spatial headway and velocity, which are depicted in figure 4.1. The time headway  $t_h$  denotes the time that elapsed after a vehicle passed the detector's cross-section until the next vehicle reached the detector.



**Figure 4.1** Examples of the data that the test of Knospe et al. was based on. The empirically observed time headway distribution is given in (a). The so-called optimal velocity function, shown in (b), relates the distance  $d$  between successive vehicles and their current velocity  $v$ . (Both figures were taken from Knospe et al. [81].)

Knospe et al. compared these data with data from traffic simulations on a one-lane road, where data was gathered in the same way as local detectors do. After a careful calibration, which resulted in most of the values of table 3.3, they found a very good agreement between the empirical data and the one obtained from simulations with the CDM. An important finding of their test was that the CDM, in contrast to the other investigated models, could reproduce time headways below 1 s, as observed in real traffic flow (see figure 4.1a).

Despite these important results, we agree with Kerner [54, p. 204] that this test does not suffice for a full assessment of the model's quality. As the comparison focused on the statistical distribution of traffic variables (independent from the detector's position), its

informative value regarding the spatiotemporal evolution of traffic variables is limited. In particular, a classification into traffic phases and the study of spatiotemporal traffic patterns require taking into account the (temporal) relation of traffic variables.

Knospe et al. were aware of this problem and studied the time series of flow  $J$  and density  $\rho$ . To identify synchronized traffic, they suggested to use the cross-correlation  $cc(J, \rho)$ :

$$cc(J, \rho) = \frac{\langle \rho(t)J(t + \tau) \rangle - \langle \rho(t) \rangle \langle J(t + \tau) \rangle}{\sqrt{\text{Var}(\rho(t))} \times \sqrt{\text{Var}(J(t + \tau))}}, \quad (4.1)$$

where  $\text{Var}(\cdot)$  denotes the variance (e.g.,  $\text{Var}(\rho(t)) = \langle \rho(t)^2 \rangle - \langle \rho(t) \rangle^2$ ). Since the synchronized phase covers a two-dimensional region of steady states, they concluded that the cross-correlation of flow and density should vanish in synchronized traffic (i.e.,  $cc(J, \rho) \approx 0$ ). This method allowed Knospe et al. to identify phases of synchronized traffic in the CDM.

In this context, we would like to point out again that working with the traffic density requires special care. In section 2.1 we have already explained that a direct measurement of density is not possible by local measurements. Moreover, estimating the density via the detector's occupancy is erroneous when the average velocity is low due to the measurement process by induction loops. (In jammed traffic, when a vehicle occupies the detector during the entire probe interval, the detector is likely to assume an occupancy of  $O_D = 0$ .) This may explain the criticism expressed by Kerner et al. [65], who claim that states of low velocities at low densities, which were found by Knospe et al., follow from erroneous density estimations involving a large systematic error.

Therefore, we chose another method to study the spatiotemporal dynamics of the CDM. Before we explain this method, we will present the general setup of our simulations.

## 4.2 Simulation Setup – Open Boundaries

As we use a similar setup to Barlovic et al. [8], we give only a brief summary of the simulation setup. (Minor modifications were necessary because each vehicle occupies  $l_{\text{car}} > 1$  sites.) We consider a one-lane road of  $N$  sites (or cells), on which vehicles move from left to right. The left boundary or entrance section consists of the leftmost  $v_{\text{car}}^{\text{max}} + l_{\text{car}} + 1$  sites, where  $v_{\text{car}}^{\text{max}}$  denotes the vehicles' maximum velocity. Vehicles, which enter the road from the left boundary with probability  $\alpha$ , are inserted with velocity  $v_{\text{car}}^{\text{max}}$  at position  $x_{\text{insert}} = \min(v_{\text{car}}^{\text{max}} + l_{\text{car}} + 1, x_{\text{last}} - v_{\text{car}}^{\text{max}})$ , where  $x_{\text{last}}$  is the rear position of the vehicle closest to the entrance section. Then, all vehicles, including the newly inserted one, move according to the CDM's rules of motion (section 3.1.3). If a vehicle is not able to leave the entrance section, it is removed afterwards. This insertion strategy allows for higher inflow rates at the left boundary compared with the obvious insertion strategy, which places a vehicle in the leftmost site with probability  $\alpha$  if this site is empty.

The right boundary or exit section is modeled as follows: before other vehicles move forward, the rightmost site is occupied with probability  $\beta$ , and it is cleared after the

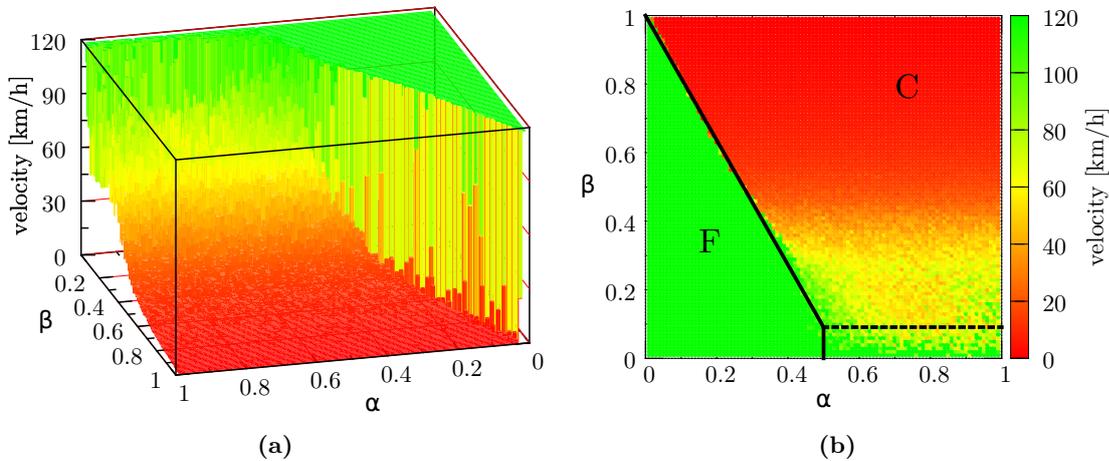
vehicles have moved. Moreover, a vehicle is removed from the road if it will reach the rightmost site or beyond during the next update step by maintaining its current velocity.

One can interpret the probability  $\alpha$  as the inflow of vehicles to the road segment, and the probability  $\beta$  determines the strength of local perturbations caused at the downstream boundary. These perturbations may, for instance, result from vehicles that enter the road via an on-ramp and occur randomly in front of vehicles on the main road. Consequently, high values of  $\beta$  result in a low exit probability of the system.

The following results were obtained on a road consisting of  $N = 5001$  sites ( $\approx 7.5$  km). The first  $2 \times 10^4$  from a total of  $2.5 \times 10^4$  time steps were discarded in each simulation run to exclude transient behavior. The vehicles' properties were taken from table 3.3, and we restricted the simulations to use one vehicle type only, namely "cars".

### 4.3 Spatiotemporal Dynamics

We have analyzed the vehicle dynamics for all combinations of entrance and exit boundary conditions resulting from a step size of 0.01 (i.e.,  $\alpha, \beta \in [0.01, 0.02, \dots, 0.99]$ ). The physical phase diagram resulting from these measurements is depicted in figure 4.2.

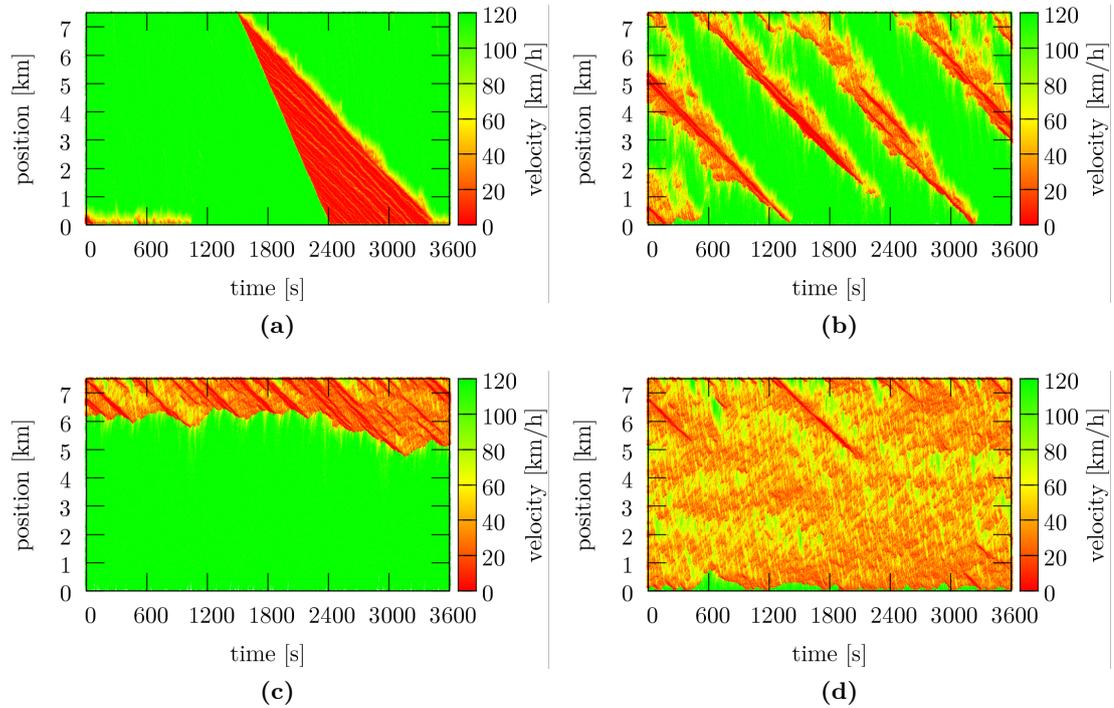


**Figure 4.2** (a) The 3D representation of the phase diagram illustrates very well that the transition from free flow to congested traffic is discontinuous. For better readability, we have added straight lines to the equivalent 2D representation in (b), which represents a stylized phase diagram. (The units of the z-axis were obtained by using the conversion factors of table 3.3 according to which a site measures 1.5 m and an update step corresponds to 1 s.)

It shows the average velocity of vehicles in the middle segment of the road, which excludes the first and the last  $N/3$  sites of the road (cf. Namazi et al. [106]). As one can see, especially from figure 4.2a, free flow (F) and congested traffic (C) are separated by a

very sharp, discontinuous transition for  $\beta \gtrsim 0.1$ . In figure 4.2b, this transition is marked by a solid, diagonal line between the regions F and C. By analogy with investigations of the Nagel-Schreckenberg model [113], we have set apart the rectangular area in the bottom right corner of figure 4.2b. Although this area is not free of congestion, the highest flow rates can be observed there. In free flow (F) all vehicles move with their maximum velocity. For a more detailed study of congested states and the corresponding traffic phases, it is, however, necessary to investigate the vehicles' microscopic and spatiotemporal dynamics.

To illustrate this, figure 4.3 shows four spatiotemporal plots of the entire road during a one-hour interval (i.e., 3600 consecutive update steps) for four distinct combinations of inflow and outflow probabilities. All configurations were taken from the region C in figure 4.2b.



**Figure 4.3** Examples of different spatiotemporal patterns resulting from different inflow ( $\alpha$ ) and outflow ( $\beta$ ) probabilities: (a)  $\alpha = 0.86, \beta = 0.09$ , (b)  $\alpha = 0.42, \beta = 0.32$ , (c)  $\alpha = 0.3, \beta = 0.47$ , and (d)  $\alpha = 0.38, \beta = 0.41$ . (The entrance section is located at position  $x = 0$  and the exit section at  $x = 7.5$  km.)

The pattern of figure 4.3a is taken from the rectangular area in figure 4.2b. In this area high inflow rates coincide with a low probability of the exit section being blocked. The figure shows two phenomena which are characteristic for this combination of boundary conditions: (i) Due to the high inflow rates, random velocity fluctuations are likely to occur close to the left (upstream) boundary. These fluctuations cause congested

traffic propagating backwards and, thereby, reduce the effective inflow probability. Consequently, the flow of the remaining vehicles corresponds to the outflow of congested traffic, where no other jams occur (see figure 4.3a during the first 1000 time steps). (ii) Jams, which often, but not necessarily, occur at the right boundary (at  $t \approx 1500$  s in figure 4.3a), have a sharp upstream front. This is again a consequence of the high inflow rates, because any local perturbation immediately affects the following vehicle and propagates upstream.

Figure 4.3b depicts several waves of congested traffic traveling upstream at a nearly constant velocity. These stop-and-go waves are known even from the most simplistic traffic cellular automata (e.g., the Nagel-Schreckenberg model [102] or the VDR model [7, 8]).

A localized congested pattern is presented in figure 4.3c. Relatively low exit probabilities (i.e., high values of  $\beta$ ) constantly provoke jams at the right boundary. The inflow probability, however, does not suffice to supply enough vehicles for the jams to propagate to the left boundary. Hence, the jams get saturated and end close to the exit section.

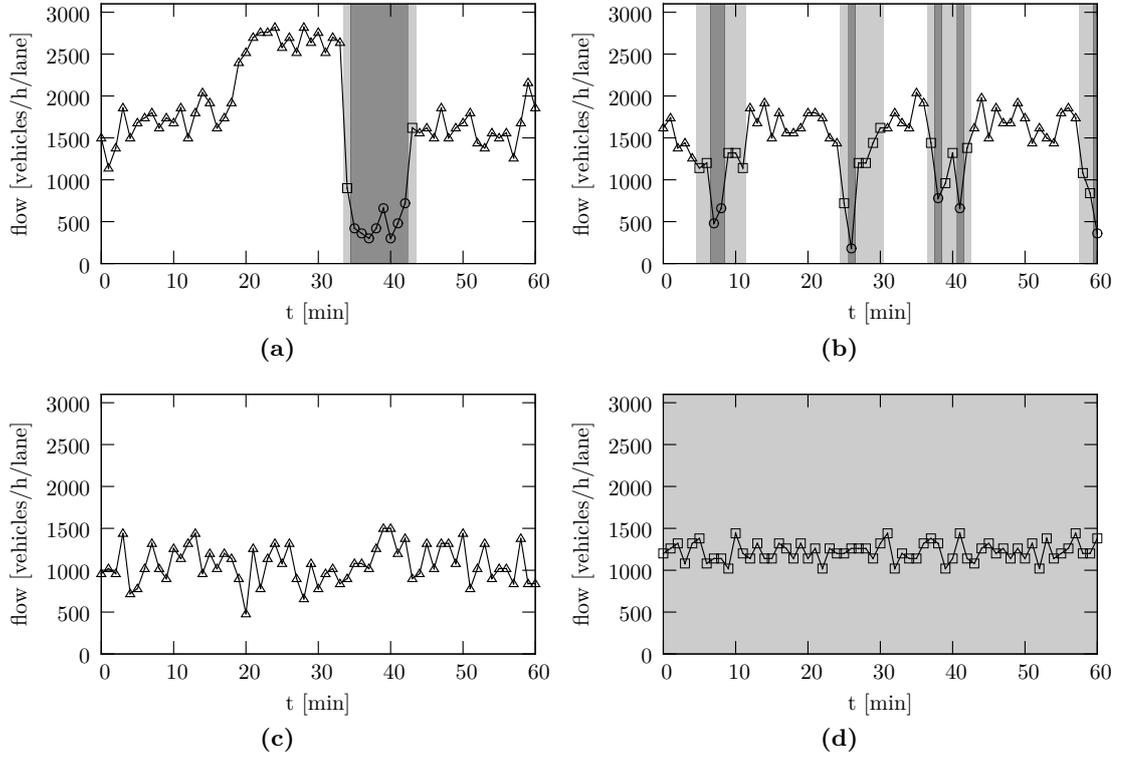
Quite interesting is the pattern of figure 4.3d. Here, nearly the entire road is covered by a state of intermediate velocities (30 km/h to 70 km/h) and relatively high flows ranging from 1020 vehicles/h to 1440 vehicles/h (see figure 4.4d).

Such a state has not been observed for simpler models [8] and has led to different interpretations [48, 66] for the CDM and a subsequent model [47] (see section 4.6). The question arises: To which traffic phase(s) should one assign this spatiotemporal pattern? And more general: How can such an assignment be done for any of the above traffic patterns?

## 4.4 Classification of Traffic Phases

Kerner et al. [58, 69] have presented a method called “FOTO” (Forecasting of Traffic Objects) that can be used to identify traffic states. The method uses data aggregated by a local detector (i.e., velocity and traffic flow). Based on a set of rules, it decides whether the local traffic state is *free flow* (F), *synchronized traffic* (S), or *wide moving jam* (J). The underlying set of rules can be summarized as follows: (i) if the average velocity is high, free flow predominates, (ii) if both the average velocity and the flow rate are low, a wide moving jam passes the detector, and (iii) if at medium velocities the flow is still high, then the corresponding traffic phase is “synchronized flow”. It is important to note that the classification of traffic states results from the simultaneous analysis of both the flow rate and the average velocity because the analysis of only one variable usually does not suffice to identify a traffic phase. (In addition to this rough description, we review the FOTO-method’s principles of operation in appendix B.)

In combination with a method called “ASDA” (Automatische Stau-Dynamik Analyse; Automatic Tracking of Moving Jams), it is even possible to track the propagation of traffic phases detected by FOTO [56]. An advantage of both FOTO and ASDA is that they „perform without any validation of model parameters in different environmental



**Figure 4.4** The time series of the flow rate from local measurements for the patterns depicted in figure 4.3. Data was collected by a detector positioned in the middle of the road at kilometer 3.75. Figures (a)–(d) correspond to the spatiotemporal plots of figure 4.3.

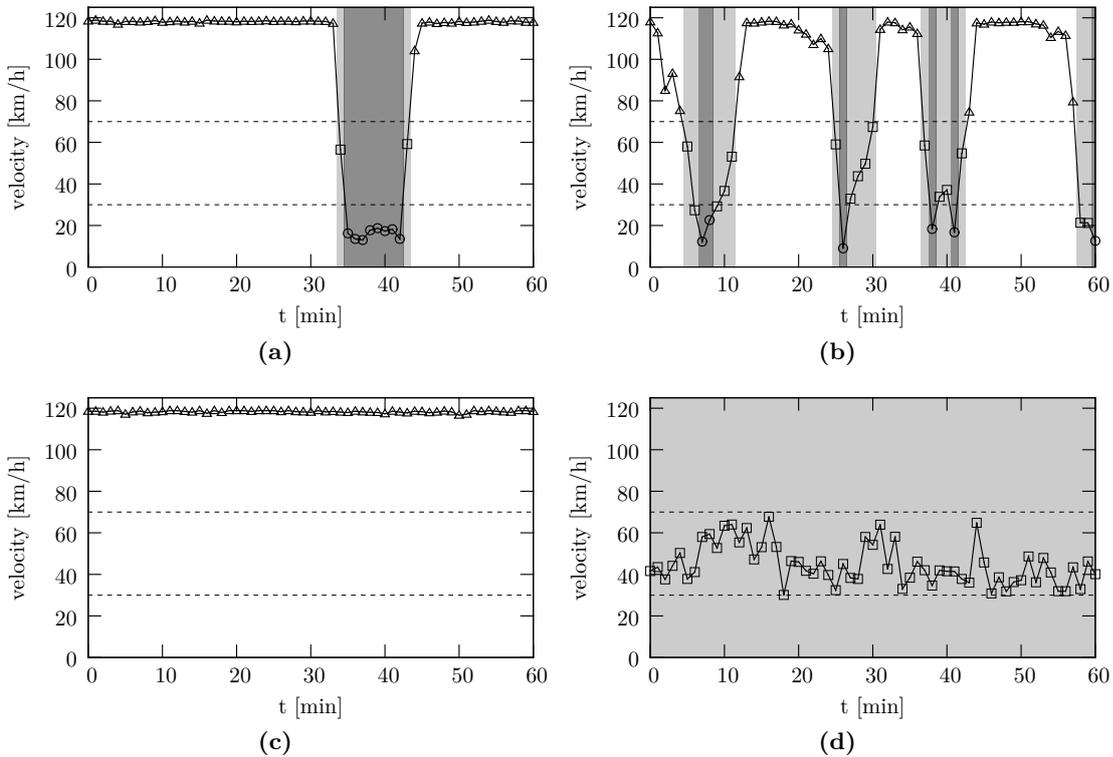
States of free flow are depicted with a white background color and triangles ( $\triangle$ ) as data points. Synchronized traffic is shown with a light gray background color and rectangular symbols ( $\square$ ). A dark gray background color and circular symbols ( $\circ$ ) indicate wide moving jams. In section 4.4 we explain how this classification was done.

From figures (a) and (b) it becomes evident that there is no unique flow rate above which the traffic flow breaks down in the CDM. In (a) the flow rate reaches values considerably above 2500 vehicles/h/lane before congestion sets in, whereas it barely exceeds 2000 vehicles/h/lane in (b) before a breakdown occurs. Also note that figures (c) and (d) are assigned to different traffic phases, even though the flow rates are at the same level.

(The corresponding velocity time series is given in figure 4.5.)

and traffic conditions” [69]. Hence, we could apply the FOTO-method to our simulation results without modifications.

For this purpose, we have positioned a detector in the middle of the road (at site 2500), which gathers the same data as its real-world equivalent (i.e., flow and velocity aggregated over 60 subsequent time steps). This allows a more detailed look at the vehicle dynamics. The velocity time series and the resulting assignment to a traffic phase for the patterns of figure 4.3 are given in figure 4.5.



**Figure 4.5** The velocity time series from local measurements for the patterns depicted in figure 4.3. Data was collected by a detector positioned in the middle of the road at kilometer 3.75.

Based on the local measurements of traffic flow and vehicle velocity, a classification of the current traffic state was performed according to the FOTO-method. States of free flow are depicted with a white background color and triangles ( $\triangle$ ) as data points. Synchronized traffic is shown with a light gray background color and rectangular symbols ( $\square$ ). A dark gray background color and circular symbols ( $\circ$ ) indicate wide moving jams. For better readability, dashed horizontal lines indicate the values of 30 km/h and 70 km/h, respectively. (Note that these lines are not associated with the classification of traffic states.)

As expected, the drastic drop of the average velocity in figure 4.3a manifests as a

wide moving jam, which lasts for eight minutes. Note that the detector data do not show an abrupt transition from free flow to a wide moving jam ( $F \not\rightarrow J$ ). First, we can observe a transition from free flow to a synchronized phase ( $F \rightarrow S$ ) before a wide moving jam is detected ( $S \rightarrow J$ ). Similarly, the recovery to free flow is achieved by a sequence of two transitions ( $J \rightarrow S$  and  $S \rightarrow F$ ).

The wave-like structures of figure 4.3b can easily be identified in figure 4.5b as well. All these waves show at least one aggregation interval that is identified as a jam. Again, we find the following sequence of transitions:  $F \rightarrow S \rightarrow J \rightarrow S \rightarrow F$ .

As the congested states of figure 4.3c are located close to the exit boundary, the detector in the middle of the road measures only free flow (see figure 4.5c).

More interesting is the time series of figure 4.5d, which belongs to the spatiotemporal pattern of figure 4.3d. We have already seen (figures 4.4c and 4.4d) that the flow rates measured at the detector are approximately the same for the patterns of figures 4.3c and 4.3d. The measurements of figure 4.5d are assigned to the synchronized phase S, whereas the measurements of figure 4.3c belong to free flow. The difference between the two measurements becomes evident in the velocity time series. In figure 4.5d, all detected velocities are between 30 km/h and 70 km/h, and thus they are distinctly lower than the average velocities of figure 4.5c. The maximum change during subsequent measurements is slightly below 29 km/h (from  $t = 43$  min to  $t = 44$  min). The absolute value of these changes in the average velocity are relatively high compared with real traffic flow. Kerner [52, p. 6], for example, reports fluctuations in the range of  $\pm 10\%$  at average speeds of 65 km/h. Yet all our measurements are assigned to synchronized traffic (S) by the FOTO-method. This observation is in agreement with Kerner [55], who attributes a self-sustaining character to synchronized traffic. The traffic patterns at the detector's position cover nearly the entire road (except for few jams close to the exit region), as a comparison with figure 4.3d shows. Consequently, we expect the time series and, thus, the classification of traffic phases of figure 4.5d to be representative for the entire road—independent of the detector's position.

To get a more quantitative result on the likelihood of the transitions between traffic phases, we have analyzed the time series of all simulations that are not labeled “F” in figure 4.2b. The exclusion of the free flow states (F) has two reasons: (i) In the free flow regime, where vehicles can move without hindrance, we do not expect any transitions to occur. (ii) In very dilute traffic (i.e., a small value of  $\alpha$ ), the application of FOTO is likely to produce erroneous results due to very low vehicle flows. When very few vehicles enter the road, a detector will detect no vehicle most of the time. However, when it does detect a vehicle, there is a high probability that no vehicle was detected during the previous sampling interval. This in turn, leads to a measurement of high velocity (the single vehicle travels at maximum velocity) following a measurement with very low velocity (no vehicle), which will be interpreted as a  $J \rightarrow F$  transition by the FOTO-method (see appendix B).

In addition to the detector in the middle of the road, which we have used in the analysis of figure 4.5, we have added another detector close to the right boundary at position  $x = 7.2$  km, where more congestion is expected due to its proximity to the exit

section. We analyzed the time series of both detectors and classified the observed traffic states. The classification was performed both with the standard set of rules of FOTO, which we have used up to now and which comprises four distinct rules as described in appendix B, and with an extended set of rules comprising 13 different rules [69], which offers a better distinction between the states J and S.

Table 4.1 shows the resulting probabilities of observing a given transition. As we have restricted our analysis to the congested regime of figure 4.2b, transitions from or to state J make up at least 70 % of all transitions at either detector location.

**Table 4.1** Probabilities for transitions from one traffic phase to another. We used both the standard and the extended rule set of FOTO. (As a consequence of rounding, the probabilities do not necessarily add up to 100 %.)

transition	probability [%] for the detector at			
	$x = 3.75$ km		$x = 7.2$ km	
	standard	extended	standard	extended
J $\rightarrow$ F	0.2	0.2	0.0	0.3
J $\rightarrow$ S	44.1	45.3	35.4	37.6
S $\rightarrow$ F	6.3	5.2	14.8	12.5
S $\rightarrow$ J	42.8	44.0	34.9	36.8
F $\rightarrow$ S	5.0	3.7	14.4	11.5
F $\rightarrow$ J	1.6	1.7	0.4	1.3

We found that transitions from a jammed state to free flow (J  $\rightarrow$  F) are very unlikely ( $\leq 0.3$  %). Similarly, F  $\rightarrow$  J transitions occur with a probability of less than 2 %. More importantly, it has to be noted that transitions from free flow to synchronized traffic (F  $\rightarrow$  S) are more than two to three times more likely than F  $\rightarrow$  J transitions. At the detector at position  $x = 7.2$  km, they are even more than eight times more likely—independent of the rule set used.

This is again in good agreement with empirical data. According to Kerner, spontaneous F  $\rightarrow$  J transitions cannot be observed in real traffic, but wide moving jams J always emerge from synchronized flow [54, 56, 61].

Concerning the different sets of rules, we can say that they led to slightly different quantitative results, but they did not change the qualitative character of our results.

#### 4.4.1 Analyzing Single Vehicle Data

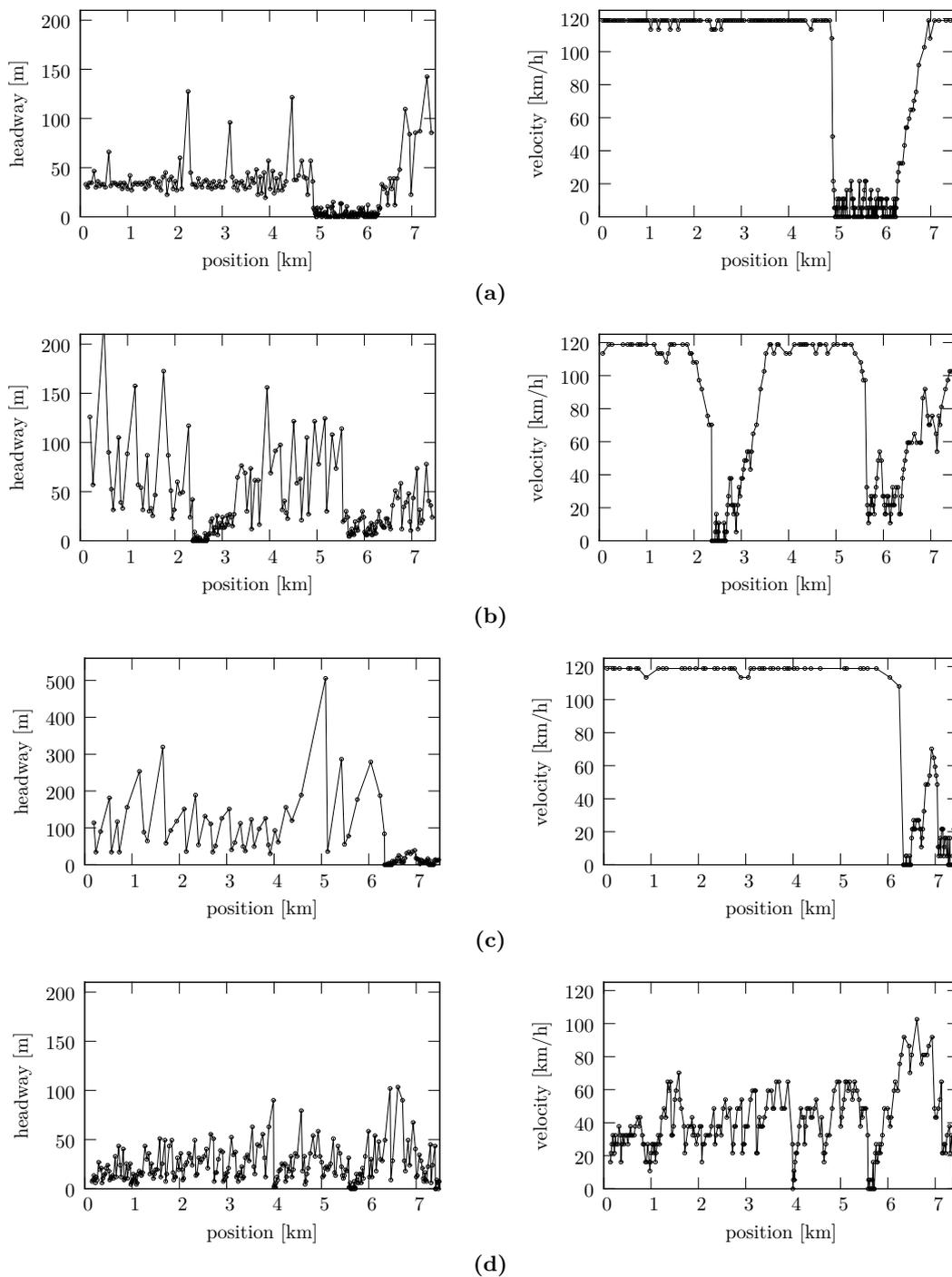
The findings of the previous section might lead to the conclusion that the CDM can, indeed, reproduce all the traffic phases proposed by Kerner. This would be a clear contradiction to the three-phase traffic theory, as the model under consideration violates the fundamental hypothesis of this theory. Therefore, a closer look on the observed traffic

patterns, whose analysis suggested the existence of three distinct phases, is necessary. Even though we used Kerner’s own FOTO-method, some of the objections raised against the approach by Knospe et al. (see section 4.1) might apply to our approach as well. The usage of data aggregated over intervals of 1 min masks to some extent the inter-vehicle dynamics. Therefore, we felt it necessary to inspect the single vehicle data, too.

Figure 4.6 shows a snapshot of the vehicles’ spatial headways and velocities at a fixed time  $t$  for the traffic patterns of figure 4.3. As in figures 4.3a and 4.5a, the massive jam surrounded by free flow is clearly visible in figure 4.6a. Within the jam, both the vehicles’ headways (i.e., the bumper-to-bumper distance) and velocities are equal or close to zero. A similar observation can be made in figure 4.6b, where two jam waves can be identified. Although, in the second jam wave, at kilometer 6, the velocity does not drop completely to zero. The dilute traffic of figure 4.6c is characterized by large headways (note the different scaling) and high velocities, except for the downstream boundary, where the effect of the induced perturbations becomes visible as congestion.

Figure 4.6d requires a very careful interpretation. The snapshot shows the headways and velocities of vehicles in a traffic pattern that FOTO identified as synchronized flow (figure 4.5d). We clearly see one stopped vehicle at kilometer 4 and a few (11) stopped vehicles at kilometer 5.7. The question of whether it is still justified to speak of the pattern of figure 4.6d as synchronized traffic cannot be answered definitively. According to the macroscopic definition of the synchronized phase, which says that any state of congested traffic that is not a wide moving jam is synchronized flow [54, p. 21], we might call the observed traffic pattern “synchronized” with confidence. This implies that we do not consider 11 stopped vehicles as a wide moving jam. On the other hand, Kerner defines synchronized traffic as traffic flow “with no significant stoppage” [54, p. 23]. Here, the question is what constitutes a significant stoppage. Possible criteria are the number of stopped vehicles, the spatial extent of the stopped vehicles, or the duration of the stoppage. As already mentioned, the number of stopped vehicles was 11, which seems to be a relatively low value compared with the number of vehicles in a wide moving jam. (In figure 4.6a we count 68 stopped vehicles, and another 34 vehicles move at 1 site per time step ( $\cong 5.4$  km/h) at time  $t = 30$  min.) Moreover, the distance between the first and last stopped vehicle in figure 4.6d was 132 m. It has to be mentioned that between these two vehicles there were also five moving ones. Similar to the total number of stopped vehicles, we consider the spatial extent as small. Therefore, we also determined the time the 11 vehicles had to wait until they could move again. The average waiting time was 14 s, and the longest waiting time was 25 s. As this is probably less time than one spends waiting at a red traffic light, the author judges this waiting time as rather small, too.

Studying the vehicles’ time headways allows a more objective analysis: Kerner et al. [65] reported that one finds regions of interrupted flow within wide moving jams. These flow interruptions are characterized by maximum time headways of  $t_{h,\max} \geq 20$  s between two vehicles. So, for each stopped vehicle, we recorded the time headways that a local detector at the corresponding vehicle’s position would have measured. The average time headway of the next ten following vehicles was 4 s. Yet, once, a time headway of



**Figure 4.6** Snapshots of the vehicles' spatial headways (left) and their current velocities (right) at a fixed time  $t$ . The figures (a)–(d) correspond to the patterns depicted in figures 4.3 and 4.5. The snapshot was taken at  $t = 30$  min. Note that figure (c) uses a different scale for the headway axis.

33 s was observed. All other time headways were below 8 s. Consequently, based on Kerner's microscopic criterion  $t_{h,\max} \geq 20$  s, we would have to reject the classification of the FOTO-method for some time intervals. Moreover, we would have to conclude that the pattern of figure 4.3d is not entirely jam-free, even if no wide moving jam could be identified via FOTO.

At this point, it has to be mentioned that the three-phase traffic theory also treats so-called narrow moving jams [54, p. 259], which, in contrast to wide moving jams, consist merely of an upstream and downstream front. These narrow moving jams can either grow into wide moving jams or disappear completely, and they are associated with the synchronized flow phase. Hence, if the aforementioned sequence of stopped vehicles represents a narrow moving jam, which its length of 132 m suggests, then the FOTO-method's classification as synchronized flow is still correct.

Considering all the previous points, we feel that a final answer on how to interpret the traffic pattern of figure 4.3d is not possible. On the one hand, one might argue that the single large time headway was a statistical fluctuation rather than a proof of the existence of a wide moving jam. On the other hand, one might object that only single vehicle data are a reliable source for the analysis of traffic patterns because detector data provide incomplete information due to the aggregation process. Yet the sequence of stopped vehicles covers less than 2% of the considered road segment. Therefore, the identification of the observed stoppage as a wide moving jam appears not mandatory, and its identification as a narrow moving jam seems justified as well.

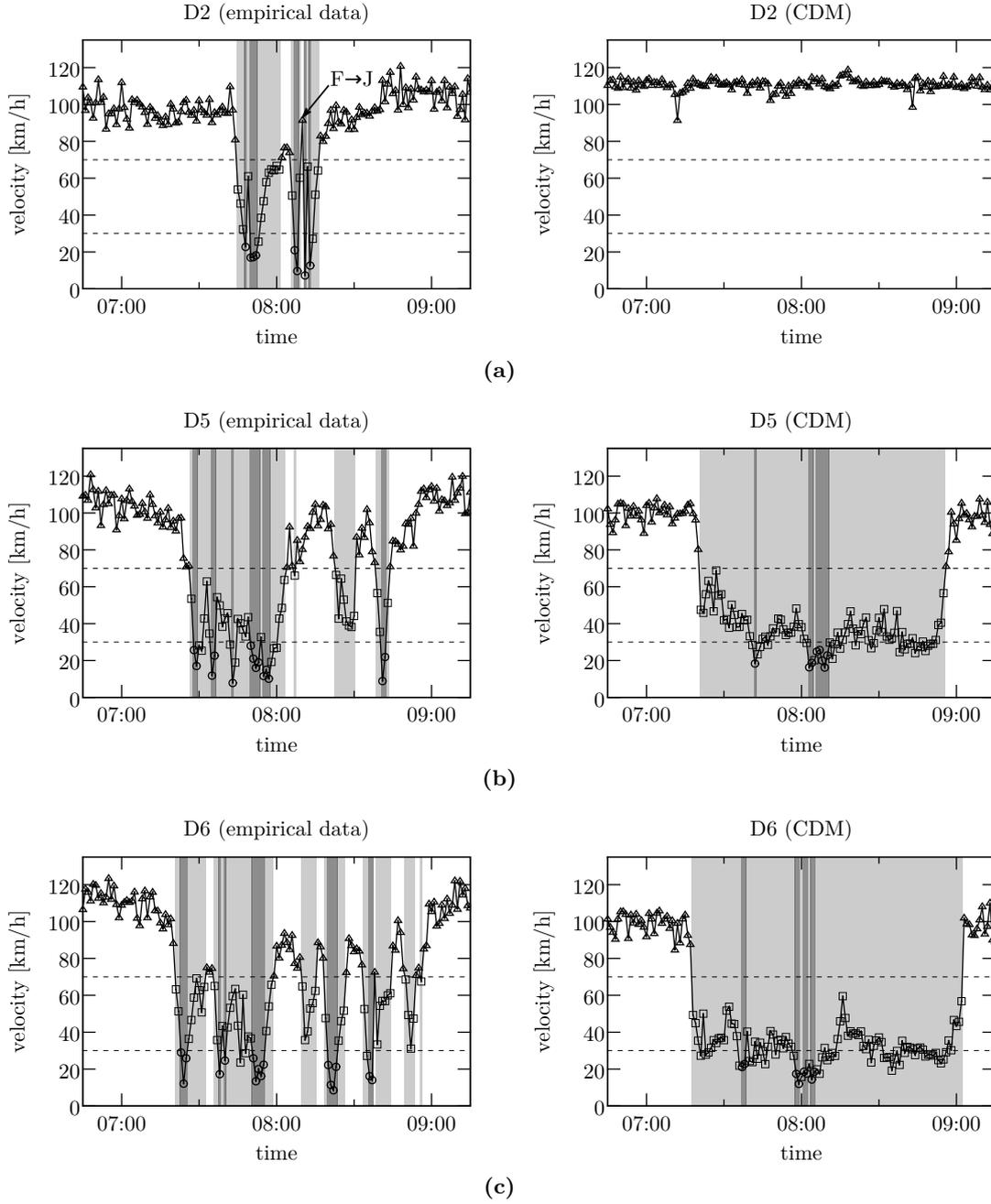
#### 4.4.2 Simulated versus Empirical Data

Keeping this uncertainty in mind, we now use the FOTO-method to classify the observed and the simulated traffic data of the previous chapter, in particular the time series of figure 3.7. In contrast to the simplistic setup of section 4.2, the real road segment is a two-lane road. To apply the FOTO-method to multi-lane roads, averaging the velocity and the flow rate over all lanes is common practice [56, 69]. So, we averaged the vehicle flow across lanes, and we calculated the average velocity by a weighted average of the velocity of trucks and vehicles on both lanes.

The resulting classification of traffic states is given in figures 4.7a, 4.7b, and 4.7c for the detector cross-sections D2, D5, and D6, respectively (see figure 3.3). The plots on the left show the averaged empirical data, and the plots on the right show the detectors' time series from simulations with the CDM.

The comparison between real and simulated data confirms what figure 3.7 already suggested: the CDM overestimates the temporal extent of congested traffic during the morning peak hour. At detectors D5 and D6, an uninterrupted sequence of congested traffic could be found between 7:20 a.m. and 8:55 a.m. for the CDM. The real time series at these locations show free flow during a 10 min-interval at approximately 8:10 a.m.

Table 4.2, which lists the number of intervals assigned to the traffic phase J, F, and S, confirms this observation as well, but allows for a quantitative characterization. On the one hand, the simulation results overestimate the occurrence of congested traffic, but they underestimate the number of intervals with wide moving jams (J).



**Figure 4.7** A comparison of the traffic state classification according to the FOTO-method for both the empirical (left) and the simulated (right) time series.

States of free flow are depicted with a white background color and triangles ( $\Delta$ ) as data points. Synchronized traffic is shown with a light gray background and rectangular symbols ( $\square$ ). A dark gray background color and circular symbols ( $\circ$ ) indicate wide moving jams.

**Table 4.2** Overview of the traffic phases found in the time from 6:00 a.m. to 10:00 a.m. at the detectors downstream of the off-ramp.

detector	empirical data			simulated data		
	number of 1 min-intervals classified as					
	J	S	F	J	S	F
D6	16	44	177	8	97	136
D5	14	37	190	8	87	146
D4	5	27	209	6	32	203
D3	9	13	219	0	7	234
D2	8	19	214	0	0	241

Similarly, the insufficient spatial extent of the congested traffic in our simulations becomes obvious for the detectors D3 and D2, for which congested traffic states were found only during seven and zero 1 min-intervals, respectively. These intervals were classified as synchronized traffic (S) without exception. The real time series, however, do not only show congestion during 21 (D3) and 27 (D2) measurement intervals, but they also show phases of wide moving jams.

The comparison between the empirical and simulated time series of figure 4.7 as well as with the time series of figure 4.5 is revealing for another reason: in figures 4.5a and 4.5b, the transition  $F \rightarrow S$  and the subsequent  $S \rightarrow J$  transition were only one or two minutes apart. Hence, one might have assumed that the intermediate step S is a mere artifact of the aggregation process. In figures 4.7b and 4.7c, however, we see that the state classified as S by the FOTO-method lasts for more than 10 min in the CDM before an  $S \rightarrow J$  transition follows. This proves that the CDM is also able to exhibit long-lasting intervals of synchronized flow that precede a wide moving jam. On the other hand, we observe such short-lasting S-phases in the time series of the real detectors in figures 4.7b and 4.7c. Altogether, we can conclude that the CDM is also able to reproduce long-lasting intervals of synchronized flow in addition to the short-lasting ones of figures 4.5a and 4.5b, which are both consistent with empirical data.

Finally, it should be noted that we found  $F \rightarrow J$  transitions in the empirical data (see the arrow-marked  $F \rightarrow J$  transition at 8:11 a.m. in figure 4.7a). To ensure that this effect did not result from the chosen averaging process, we also applied the FOTO-method to the raw data of each lane. The analysis of the raw data (not depicted) revealed  $F \rightarrow J$  transitions as well. This observation, which contradicts Kerner's three-phase theory, is rather an evidence of an imperfection of the FOTO-method than of the theory itself. Remember that the traffic phases are spatiotemporal phenomena, whereas the FOTO-method relies on local information only. The benefit of the FOTO-method is that it provides several criteria on which the analysis of traffic states can be based on. This is important to note because many previous studies appear to be based on a subjective evaluation of traffic patterns as we will discuss in section 4.6.

## 4.5 Deceleration Behavior of the CDM

All previous comparisons showed a good agreement between the CDM and empirical data. An obvious criticism of the CDM, which we have already mentioned in section 3.1, is the (potentially) unrealistic deceleration behavior.<sup>1</sup> The CDM’s maximum braking capacity is bounded only by the vehicles’ maximum velocity  $v^{\max}$  and the temporal discretization  $\Delta t$ . For our choice of parameters (table 3.3) this leads to a maximum deceleration of  $33 \text{ m/s}^2$ . Such values are not only unrealistic, their occurrence may also be a sign of a lack of realism in the description of vehicular dynamics. Braking maneuvers requiring the maximum deceleration correspond to the situation where a vehicle travels at maximum velocity and has to come to rest within one time step when approaching a slow vehicle ahead. In reality, we expect that the vehicle is aware of a slow predecessor early enough so that it initiates a smooth braking process.

To answer the question of how often unrealistic braking maneuvers occur in the CDM, we measured their frequency in different simulation setups. In all cases, we studied a circular road (i.e., periodic boundary conditions) of a total length of 7.5 km (5001 cells) with either one or two lanes. For both setups, we analyzed the vehicles dynamics for various vehicle densities  $\rho$ . Moreover, we performed the simulations with a share of trucks of either 0 % or 10 %.

Each simulation comprised  $11 \times 10^6$  time steps from which the first  $1 \times 10^6$  were discarded. After each time step, we determined the acceleration or deceleration of each vehicle.

The resulting statistics is depicted in figure 4.8. The plots show the relative frequency of acceleration and deceleration events with a given acceleration and deceleration rate, respectively. The results from simulations where all vehicles were fast ones (i.e., 0 % trucks) are given on the left (figures 4.8a and 4.8c). Analogously, the plots on the right (figures 4.8b and 4.8d) give the corresponding distributions where 10 % of all vehicles were trucks. The upper two plots show the results for one-lane roads, whereas the lower plots give the results for two-lane roads.

In all cases (i.e., one-lane or two-lane road, 0 % or 10 % share of trucks, various densities), the distributions have a peak in the range from  $-1.5 \text{ m/s}^2$  to  $1.5 \text{ m/s}^2$ —values that are in good agreement with our expectations and with reality.

To estimate in how many cases the deceleration must be called “unrealistic”, we assumed an upper bound for a vehicle’s braking capacity of  $9 \text{ m/s}^2$  [110]. Vehicles did not exceed this value in 99.5 % of all cases in all simulation setups. Even if we limit the maximum braking capacity at  $3 \text{ m/s}^2$ , deceleration rates stay below this threshold in more than 96.5 % of all cases.

Consequently, unrealistic braking events are rather scarce, although they do occur. Events requiring an emergency braking with maximum deceleration happen with a probability of less than  $10^{-7}$ , if they occur at all.

A closer look reveals that the highest probability for unrealistic braking maneuvers is

---

<sup>1</sup>It has to be said that the unrealistic deceleration behavior is no inherent feature of traffic cellular automata in general, as some more recent models show [91, 127].



at intermediate traffic densities of 20 vehicles/km/lane and 26 vehicles/km/lane. In this density regime, vehicle interactions are frequent and strong. In denser traffic, vehicles mutually hinder each other such that velocities and, therefore, deceleration rates are low. In dilute traffic, on the other hand, vehicles hardly interact but travel with maximum velocity. Consequently, they rarely need to brake.

In heterogeneous traffic with two types of vehicles, however, even in dilute traffic vehicles approaching a slower one have to brake more frequently and stronger, which can be seen on the right of figure 4.8. This effect is particularly pronounced on the one-lane road where there is no possibility to overtake trucks.

In homogeneous traffic (i.e., 0% trucks), there was not even a single unrealistic braking maneuver observed in dilute traffic (i.e.,  $\rho = 6$  vehicles/km/lane and 13 vehicles/km/lane) on the single-lane road. On the two-lane road, the maximum deceleration observed was  $13 \text{ m/s}^2$ . Hence, although the CDM shows unrealistic values for a vehicle's deceleration, these events make up only a small fraction of all vehicle updates. The benefit of this inadequacy is that vehicle motion in the CDM is accident-free.

## 4.6 Discussion

In this chapter we thoroughly investigated the dynamics of the CDM. We started our investigation with a very simple simulation setup, which represented a one-lane road with open boundaries. By varying the inflow and outflow probabilities into and out of a simplified road segment, we were able to obtain a phase diagram of the CDM (see figure 4.2). The sharp line separating free and congested flow in the phase diagram may be evidence of a first-order phase transition.

Beyond the distinction between congested and free traffic flow, we aimed for a more detailed analysis of the congested phase. In particular, we wanted to study congested states with regard to Kerner's three-phase traffic theory. According to this theory, the CDM is not expected to reproduce synchronized traffic [66]. We found, however, traffic patterns (figure 4.3d) that showed several characteristics of synchronized traffic. For the further analysis of these traffic patterns, we had to investigate the spatiotemporal dynamics on the road (see figure 4.3). Besides the traffic phases free flow (F) and wide moving jams (J), we also found states with medium average velocities and relatively high flow rates which can be interpreted as a first sign of synchronized flow (S).

Notwithstanding these findings, it is still not clear whether the CDM reproduces the synchronized phase in the sense of Kerner's three-phase traffic theory. All we can say is that we found three clearly distinguishable states or "phases".

With respect to the difficulty of characterizing traffic states, let us refer to two articles by Kerner et al. [66] and Jiang and Wu [48]. In the article by Kerner et al. [66], in which the CDM was investigated, the spatiotemporal pattern that we found in figure 4.3d was classified as an "oscillating moving jam". It appears that the same pattern was classified as synchronized flow by Jiang and Wu [48, Figs. 3+4], who analyzed an extension of the CDM [47]. (Their extended model incorporates some of the findings of

the three-phase traffic theory.) The fact that the seemingly identical traffic pattern led to different interpretations by different authors illustrates the difficulty of classifying traffic states.

Therefore, we decided to apply a rule-based method to classify traffic phases. The FOTO-method, which we used for this purpose, certainly has some limitations in assessing the spatiotemporal dynamics as it is based on local measurements, but it provides “hard” and objective criteria. The application of such a rule-based technique, thereby, facilitates the comparability of our results with both empirical data and other models.

The classification (figure 4.5) shows that the CDM is able to reproduce at least some aspects of Kerner’s three-phase traffic theory correctly. In particular, the pattern of figure 4.5d was indeed identified as synchronized flow by the FOTO-method.

Yet some qualitative differences remain as the single vehicle data of figure 4.6 prove. In contrast to models based on three-phase traffic theory, the fluctuations of velocity (figures 4.5d and 4.6d) are considerably higher in the CDM (e.g., cf. [66, Fig. 9]). This is a consequence of the different modeling approach. In early traffic cellular automata (e.g., [8, 102]), a vehicle always accelerates when it is safe to do so. In the CDM this effect is partially compensated by an anticipatory component. In models based on the three-phase traffic theory (e.g., the models presented in [66]), on the other hand, a vehicle may accept any gap if its velocity is within a given range that is determined by the models’ rules of motion.

For more quantitative statements on unrealistic braking maneuvers, we also analyzed the rates at which vehicles decelerate in section 4.5. As the CDM avoids collisions only by allowing for a practically unlimited braking capacity, we studied how often braking maneuvers occur that require unrealistically high deceleration rates. We found that the probability for such maneuvers is low ( $< 0.5\%$ ) but not negligible. If the inter-vehicle dynamics, where realistic deceleration rates are desirable, play only a minor role, we consider the CDM as a model that satisfyingly reproduces the traffic dynamics on various scales.

However, let us emphasize that we fully acknowledge fundamental conceptual differences between models based on the three-phase traffic theory and models within the fundamental diagram approach. From a more practical point of view, these differences apparently do not result in fundamental differences of the traffic patterns observed. Therefore, the application of the CDM for traffic simulations (e.g., the simulation of large scale traffic networks [136]) is definitively justified.



## Optimizing Traffic Flow by Vehicular Communication

The previous chapters presented vehicular traffic as an interesting research topic and discussed the ability of computer models to reproduce real traffic dynamics. We saw that, from a physical point of view, a traffic breakdown corresponds to a phase transition from free flow to congested traffic.

There is, however, another and probably more common perception of congestion. Traffic congestion is an annoyance; it costs time, fuel, and, thus, money. In 2010, for instance, the average US auto commuter spent an additional 34 hours in his car and wasted 14 gallons of fuel simply as a result of congestion [126]. A comparison with 1982 (14 hours, 6 gallons) shows that the growing traffic volume poses an ever increasing problem to our transportation systems.

Simply extending the road network is not suited to fight congestion due to spatial, financial, and environmental constraints. Recent progress in the area of information and communication technology, however, promises to make today's transportation systems not only more efficient, but also safer, more reliable, and more convenient. All these modern approaches to improve today's transport networks are subsumed under the general term Intelligent Transportation Systems (ITSs).

One especially promising field in this context is vehicular communication. By vehicular or vehicle-to-vehicle communication we mean information exchange between neighboring vehicles. For example, a vehicle initiating an emergency braking could emit a warning message so that following vehicles would be able to timely adapt their velocity.

The required technology for this purpose is very common. Similar to wireless computer networks (WLAN), vehicles equipped with a communication device can establish a wireless communication network to send and to receive messages. Such Vehicular Ad Hoc Networks (VANETs) are considered a central part of ITSs. In principle, VANETs could enable all parties involved in traffic (e.g., vehicles, traffic lights, or road

---

This chapter is largely based on and taken from the author's articles [T74, T76] (see page 114).

side units) to exchange information and to coordinate their behavior. As no underlying infrastructure is required and as message exchange is carried out with low latency times, we also consider VANETs as an excellent tool to reduce congestion in the context of ITSs.

In this chapter, we will address the question of how vehicular communication can be applied to make transportation more efficient. Before we discuss the application of vehicle-to-vehicle communication to traffic flow optimization, let us review the physical foundations of vehicular (i.e., wireless) communication.

## 5.1 Radio Wave Propagation Fundamentals

This review is limited to the essential aspects of radio wave propagation theory that are necessary to understand the settings used in our simulations and to judge their applicability to vehicular networks.

The characteristics and properties of a radio channel limit both the transmission range and the amount of data which can be transmitted per time interval. In real-world scenarios, radio waves are subject to reflection, scattering, and diffraction. As any object in the vicinity of a receiving or transmitting node (i.e., a vehicle) can possibly act as a source of either of the aforementioned effects, a realistic treatment of radio propagation would have to take into account the nodes' environment. Considering the high mobility of nodes in the case of vehicular networks with relative velocities up to 400 km/h, this approach is nearly intractable. Therefore, various mathematical models have been proposed to mimic the properties of radio wave propagation [119, 128].

### 5.1.1 Propagation Models

In general, propagation or path loss models describe a functional dependence between the strength  $P_t$  of a transmitted signal and the received signal power  $P_r$  at a distance  $d$  of the sender:

$$P_r(P_t, d). \quad (5.1)$$

This functional dependence, called path loss, is the most important magnitude in the analysis and modeling of wireless radio channels.

#### 5.1.1.1 Free Space Propagation Model

The free space or Friis propagation model assumes a clear line-of-sight between sender and receiver. Moreover, it ignores any multipath effects such as diffraction, scattering, or reflection. In this case  $P_r(P_t, d)$  is given by [29, 119]:

$$P_r(P_t, d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2} \quad (5.2)$$

for isotropically transmitting and receiving antennae.  $G_t$  and  $G_r$  denote dimensionless gain factors of the transmitting and receiving antenna. The signal's wavelength is given by  $\lambda$  [m].

Due to the simplifications made for the derivation of equation (5.2), the free space model can only serve for estimating an upper bound of signal strength at a given distance in most real-world scenarios.

### 5.1.1.2 Two Ray Ground Reflection Model

A more realistic propagation model takes into account a ground-reflected propagation path in addition to the line-of-sight path of the Friis model. This results in a faster decay of the received power and leads to the formula of the two ray ground reflection model [119]:

$$P_r(P_t, d) = \frac{P_t G_t G_r}{d^4} h_t^2 h_r^2, \quad (d \gg \sqrt{h_t h_r}). \quad (5.3)$$

The variables  $h_t$  and  $h_r$  denote the heights of the transmitting and receiving antenna, respectively. Here, the signal decays with the inverse fourth power of the distance and is independent of the wavelength.

At short distances, the contribution of the ground-reflected path should be neglected, and the free space model should be used [26, 119]. A crossover distance  $d_c$

$$d_c = \frac{4\pi}{\lambda} h_t h_r \quad (5.4)$$

guarantees the continuity of  $P_r(P_t, d)$ .

### 5.1.1.3 Nakagami- $m$ Model

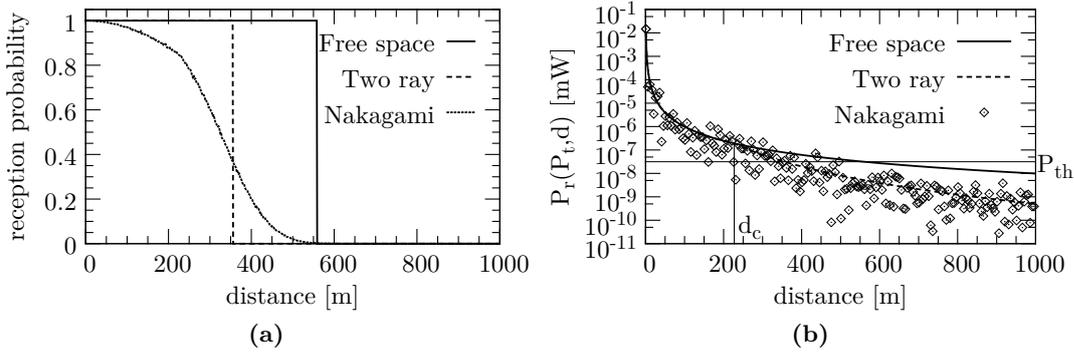
Both the Friis and the two ray ground reflection model are deterministic models. This leads to an unrealistically sharp cutoff in reception range (see figure 5.1a). Yet multipath propagation causes the received signal to fluctuate rapidly at short time and length scales. As a consequence of these fluctuations, the reception can no longer be guaranteed at short distances, but the reception also becomes possible over distances that are larger than in the deterministic case (see figure 5.1b).

Probabilistic propagation models can reproduce this behavior by employing a probability distribution. From the distribution's density function one obtains the reception probability. A probabilistic propagation model, which has proven to give a realistic description of signal propagation, is the Nakagami- $m$  model [104]. Its probability density function (pdf)  $p(x)$  is given by:

$$p(x) = \frac{2m^m x^{2m-1}}{\Gamma(m)\Omega^m} e^{-mx^2/\Omega}, \quad x, \Omega \geq 0, m \geq \frac{1}{2} \quad (5.5)$$

where  $\Gamma(\cdot)$  denotes the Gamma function and  $m$  is a shape factor which can be used to model the distribution's width. The distribution's second moment  $\langle x^2 \rangle$  is given by  $\Omega$ . For  $m = 1$ , the pdf of the Nakagami model reduces to a Rayleigh distribution. The Nakagami model can therefore be understood as a generalization of the Rayleigh distribution.

When the Nakagami- $m$  model is applied to calculate signal powers, we can identify  $\langle x^2 \rangle$  as the expected power at distance  $d$ . The expected power in turn can be estimated via equations (5.2) or (5.3).



**Figure 5.1** A comparison of the Friis, free space, and Nakagami path loss models: (a) The deterministic nature of the first two models becomes apparent by the sharp change of reception probability at a fixed distance. For the Nakagami model the average reception probability shows a continuous transition from one to zero with increasing distance. (b) The stochasticity of the Nakagami model, however, causes large fluctuations of the received signal strength at all distances.

#### 5.1.1.4 Propagation Models for Highway Scenarios

Due to its importance for the analysis of the radio channel and of communication protocols, propagation models have become an intensively studied topic (see, e.g., [14, 32, 49, 129]). There are, however, only few empirical measurements evaluating the path loss characteristics in real-world VANET scenarios. We will summarize some of the findings with respect to highway traffic, which is the subject of our own research.

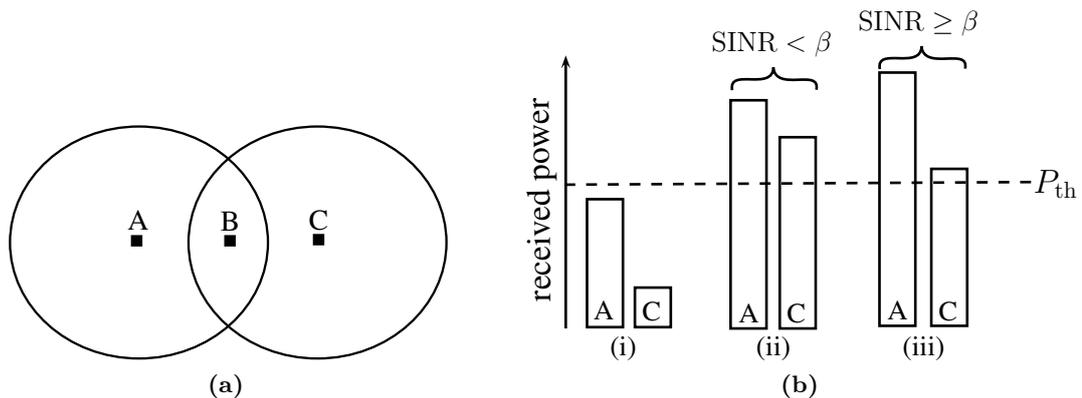
Kunisch and Pamp [88] performed measurements of the car-to-car radio channel in rural, highway, and urban environments. As a result of varying traffic densities, their measurements include reception characteristics for both line-of-sight and non-line-of-sight communication. The authors report that path loss can be described by an adjusted two ray model. Measurements by Kåredal et al. [49] confirm the applicability of the two ray propagation model for rural environments but suggest using a power law to model path loss for highway scenarios. It should be noted, however, that on the highway where Kåredal et al. performed measurements ground reflection may have been obstructed by a wall separating the different travel directions.

Yin et al. [146] analyzed empirical data collected on freeways around Detroit, USA. They report that the Nakagami model is well suited to fit the observed path loss if different shape factors are used for small and large distances, respectively. At low distances (100 m and less), the shape factor should be chosen from the interval [1.0, 1.8] and from [0.7, 1.0] for larger distances. Similarly, Taliwal et al. [133] found a good

agreement between the observed channel characteristics for highway environments and the Nakagami distribution. According to their measurements the shape factor lies between 0.5 and 1 and the expected power follows the two ray ground reflection model with a crossover distance of about  $d_c = 160$  m.

### 5.1.2 Noise, Interference, and Signal Reception

Signal strength is not the only factor that governs the connectivity in wireless networks. Due to the lack of a central authority in wireless networks, coordinated communication between senders and receivers cannot be guaranteed. Therefore, several senders may simultaneously try communicating with the same receiver. An example of such a scenario, known as *hidden node problem*, is illustrated in figure 5.2a. The communication range of the two senders  $A$  and  $C$  is shown as circles with the corresponding node in its center. Both nodes can communicate with node  $B$  which is within their communication range, but they are not aware of each other. So, each of them might start sending messages assuming  $B$  is ready to receive.



**Figure 5.2** An illustration of the hidden node problem. In the configuration depicted in (a) neither sender  $A$  nor  $C$  are able to determine whether  $B$  is already receiving. In the case of simultaneous transmission, the ability of  $B$  to receive either of the signals depends both on the absolute signal power, which has to exceed a threshold value  $P_{th}$ , and on the ratio of signal powers. In figure (b) there is no reception in cases (i) and (ii). Signal  $A$  can only be received if its signal strength divided by the sum of all remaining signal powers (in the depicted case: only  $C$ ) is larger than  $\beta$ .

Reception is likely to fail in this situation, unless one signal is much stronger than the other. A receiver is able to capture the strongest of several interfering signals if it is clearly distinguishable from the others. For this to happen the signal to interference plus noise ratio (SINR) must be above some threshold  $\beta$ . The SINR measures the strength of a given signal relative to all other interfering signals including thermal noise ( $\propto k_B T$ ) [34].

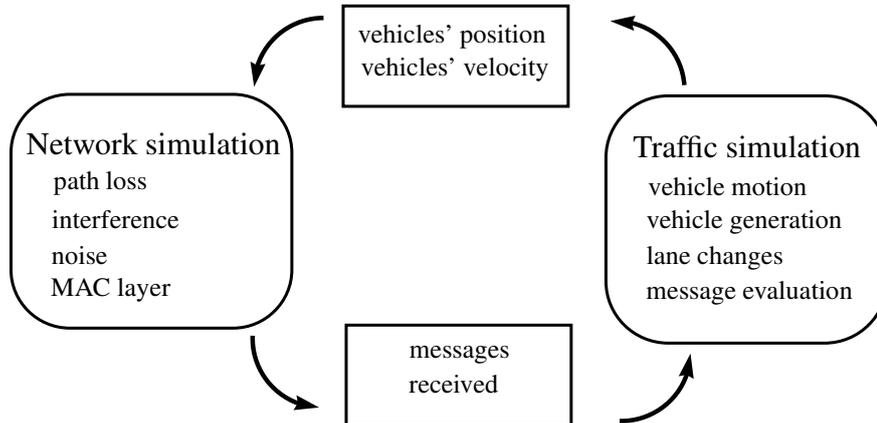
Figure 5.2b summarizes the conditions for successful reception, as described in the last two sections: both signal strength as well as the SINR have to be above threshold values for successful communication.

## 5.2 Simulating Inter-Vehicle Communication

For both practical and economic reasons much of the work in the field of VANET-research is done via computer simulations as a large number of equipped vehicles would be necessary in real-world experiments. For the sake of completeness, we give a brief overview of current challenges and approaches regarding the simulation of VANETs.

Many current VANET simulations rely on static vehicle traces generated by a traffic simulator, which they pass to a network simulator (e.g., ns2 [26] or JiST/SWANS [9–11]). The shortcoming of this approach is that vehicles cannot change their driving behavior or route choice in response to the received messages. Taking into account the interaction between vehicle motion and message exchange requires the combination of both a realistic traffic flow model and a communication model.

The network simulations in the rest of this thesis used the JiST/SWANS platform [9–11, 72] which combines a discrete event simulation engine with a *scalable wireless ad hoc network simulator* (SWANS). Extensions by Ibrahim and Weigle [44] enable the network simulator to interact with the underlying mobility model. The resulting bidirectional coupling is illustrated in figure 5.3.



**Figure 5.3** For vehicles to react to the received messages and for the correct prediction of the signal reception statistics, a coupling between the network and traffic simulations is necessary. The microscopic traffic simulation calculates the position and velocity of each vehicle on the road. This information is required by the network simulation to calculate which vehicles successfully receive a sent message. Such a received message, in turn, may influence the vehicles’ motion.

### 5.3 Evaluating Traffic Information and Improving Traffic Flow

Our approach to improve traffic flow via vehicular communication is based on two points: (i) vehicles autonomously estimate the current traffic situation and broadcast this information. (ii) based on this estimate and if required by the traffic state, vehicles adapt their driving behavior. We admit that the first point, VANET-based traffic state estimation, has already been investigated earlier. Due to the nature of VANETs, where the topology is rapidly changing, exchanging messages with a constant transmission power and at a fixed transmission interval is likely to show an increase in package collisions at large vehicle densities. Therefore, it was suggested to adapt either the transmission range [4] or the transmission interval [130] depending on the local vehicle density. To estimate the local density, Artimy [4] proposed analyzing the vehicle's mobility pattern. Similarly, Sommer et al. [130] suggested analyzing the number of package collisions and the number of neighboring vehicles for this purpose. They showed that these adaptive schemes allow for a more efficient and reliable use of the network resources.

The focus of the previous studies, however, was on the detection of congestion or of large vehicle densities. Applications aiming at easing congestion, however, require more information than knowledge about the local vehicle density. Several studies (e.g., [89, 107, 145]) suggested using VANETs to identify congested roads and to offer alternative, less congested routes. In addition to the risk of causing congestion on the alternative routes, these strategies become effective only after a traffic breakdown has already occurred.

More innovative approaches search to influence the driving behavior that causes the breakdown. Kerner et al. [64] suggested that the exchanged messages should comprise the length of a recommended space gap that vehicles are to maintain. Vehicles that adapt their space gap according to the recommendation are less likely to provoke a traffic breakdown. Similarly, Fekete et al. [27] proposed that vehicles adapt their own velocity to the average velocity of neighboring vehicles, which they determine via inter-vehicle communication, thus reducing inhomogeneities in traffic flow. The previous two works studied the influence of their strategies for relatively high penetration rates of communicating vehicles of 100 % and 60 %, respectively. Such rates will be reached only long after a successful market introduction of VANET devices [96]. Moreover, it is not clear when, if at all, vehicles are to alter driving behavior.

Therefore, we developed a VANET-based strategy to reduce traffic jams, which explicitly accounts for human reaction time and becomes effective at low penetration rates. Our approach aims to identify "critical" road segments and to prevent a traffic jam before it actually occurs. In this context, we will call a road segment critical when a breakdown is likely to occur on it. Therefore, a purely density-based analysis does not suffice. For, as we have already seen (see figure 2.2), the traffic state is not definitely determined by the traffic density. With traffic breakdowns being caused by misbehavior of human drivers in dense traffic, we suggest using periodically emitted

short messages, so-called *beacons*, to analyze the current traffic state. On the basis of this analysis, communicating vehicles may change their driving behavior in dense traffic. After changing their driving behavior, these vehicles inform following ones about the discovery of a critical road segment, and they are less likely the trigger of a traffic breakdown. This makes our method different from the aforementioned, previous approaches that guide vehicles to less congested routes to escape congestion. In addition to that, our simulations do not only use very realistic radio propagation and mobility models but also empirical traffic data from a German Autobahn.

### 5.3.1 Beacons for Congestion Warning

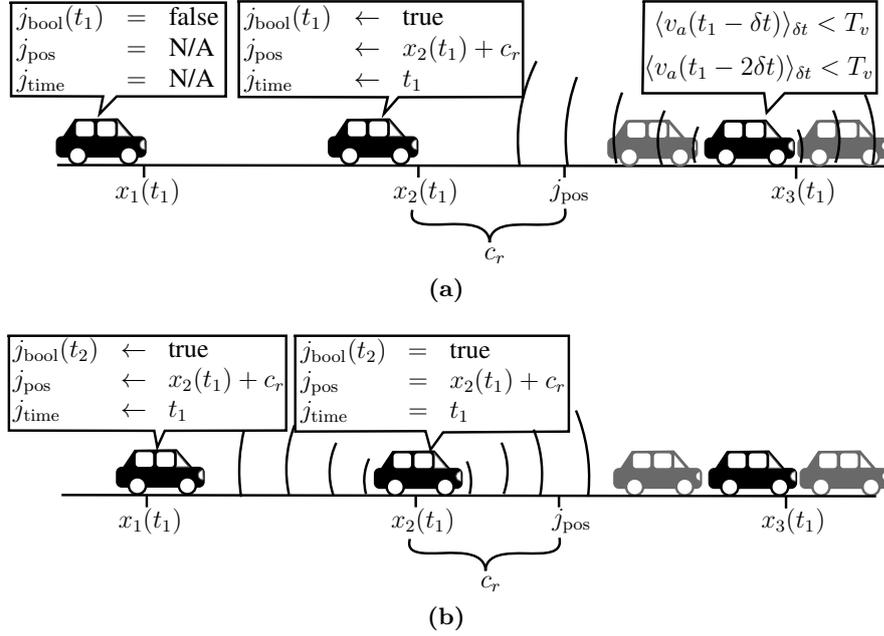
Our approach to use vehicle-to-vehicle communication to ease congestion is solely based on beacon messages. Beacon messages are periodically broadcast status messages containing a vehicle's position, its velocity and acceleration, a unique vehicle identifier, and a time stamp.<sup>1</sup> This information suffices to estimate the local state of traffic [4, 130]. In our approach we suggest expanding a beacon's content by two additional variables: a position  $j_{\text{pos}}$  and a time stamp  $j_{\text{time}}$  marking a critical road segment. By employing the data that are already sent in regular beacons and merely extending it with two additional data fields ( $j_{\text{pos}}$  and  $j_{\text{time}}$ ), we do not create overhead due to additional packets in the network. The size of a beacon usually does not exceed 100 bytes; however, to pay attention to expected future demands (e.g., security) we fix the beacon size to 500 bytes in our following considerations to demonstrate that our approach is able to work with those beacon sizes.

From the beacons a vehicle receives during an interval  $[t, t + \delta t)$ , it may calculate the average velocity  $\langle v_a(t) \rangle_{\delta t}$  of all transmitting vehicles ahead. When the average velocity drops below a given threshold  $T_v$  for two successive intervals at time  $t$  ( $\langle v_a(t - \delta t) \rangle_{\delta t} < T_v$  and  $\langle v_a(t - 2\delta t) \rangle_{\delta t} < T_v$ ), it marks the segment as critical by setting  $j_{\text{time}} \leftarrow t$  and  $j_{\text{pos}} \leftarrow x(t) + c_r$ , where  $x(t)$  denotes the vehicle's current position at time  $t$  and  $c_r$  an average communication range. Vehicles receiving a notification about such a critical traffic condition append the information to their own beacons (see figure 5.4a).

Following vehicles use this information to decide whether to change their driving behavior. For this decision, a vehicle has to judge the relevance of the information about a critical road segment. Such information is considered as relevant if the vehicle is close to the said segment, if it is approaching it, and if the information is sufficiently up-to-date (see figure 5.4b). To quantify the proximity to the road segment and the timeliness of the information, we introduce temporal and spatial thresholds  $T_s$  and  $T_t$ , respectively. A Boolean variable  $j_{\text{bool}}$  indicates whether the vehicle has to change its driving behavior. Thereby, the aforementioned conditions can be written as:

$$\begin{aligned} j_{\text{bool}}(t) &\leftarrow \text{false} \\ \text{if } (t - j_{\text{time}} < T_t \text{ and } 0 < j_{\text{pos}} - x(t) < T_s): \\ j_{\text{bool}}(t) &\leftarrow \text{true} \end{aligned}$$

<sup>1</sup>Periodic beacons that we use for message exchange are essential for most safety applications [109] and have been extensively studied. As the technical aspects and challenges are not in the focus of this section, we refer to the article by Mittag et al. [99] and the references therein for further information.



**Figure 5.4** How the proposed strategy works: when a vehicle (at  $x_2(t_1)$ ) detects that the average velocity of downstream vehicles  $\langle v_a(t_1 - \delta t) \rangle_{\delta t}$  is smaller than the threshold  $T_v$  and if also  $\langle v_a(t_1 - 2\delta t) \rangle_{\delta t} < T_v$ , it seeks to change its driving behavior ( $j_{\text{bool}}(t_1) \leftarrow \text{true}$ ). It also appends this information to its own beacons. Thereupon, a following vehicle [formerly at  $x_1(t_1)$ ] that receives a beacon from the vehicle formerly at  $x_2(t_1)$  changes its driving behavior, provided that this information is still considered as up-to-date (i.e., not older than  $T_t$ ). (a) Critical road condition is detected at time  $t_1$ . (b) Upstream vehicles are warned at time  $t_2$  ( $t_2 > t_1$ ) provided that the warning is still up-to-date ( $t_2 < T_t + t_1$ ).

Consequently,  $T_t$  and  $T_s$  define a temporal and spatial horizon of relevance for traffic information. This helps to reduce the load on the radio channel; the congestion warning is of value only for a limited number of vehicles in a large network and, thus, need not be broadcast outside this area. For simulations we set  $T_t = 30$  s and  $T_s = 3$  km. As will be discussed later, we found our results to be relatively insensitive to the concrete choice of  $T_s$  and  $T_t$ .

The threshold below which a traffic condition is classified as critical is set to  $T_v = 81$  km/h. This value is 10.8 km/h (3 m/s) below the maximum velocity of trucks in our simulation (see table 3.3) and indicates that vehicle interactions and perturbations are strong enough to make even slow vehicles (i.e., trucks) slow down. Empirical observations showing that maximum traffic flow is reached at approximately 80 km/h support this choice of parameter (see, e.g., [137]). Finally, we set  $\delta t = 1$  s and  $c_r = 150$  m. The variable  $c_r$  estimates the distance to the critical road segment. Compared with

using the average position of transmitting downstream vehicles, choosing a constant value for  $c_r$  proved as particularly advantageous at low penetration rates when only a single communicating vehicle is sensed downstream.

### 5.3.2 Vehicle Motion and Adapted Driving Behavior

For the simulation of vehicle motion, we used the CDM again. Here, we describe the adaptations that we made for a change of driving behavior in case of approaching a critical road segment. A detailed review of the model with the proposed modifications can be found in appendix C.

To motivate our modifications, let us review the meaning of the CDM's probabilistic variable  $p$  of equation (3.10) that causes minor fluctuations in a vehicle's velocity, reflecting the drivers' incapability to maintain a constant velocity. The original CDM distinguishes three cases when determining  $p$  for a vehicle labeled  $n$ :

$$p \leftarrow \begin{cases} p_b, & \text{if } b_{n+1}(t) = 1 \text{ and } t_h(t) < t_s(t), \\ p_0, & \text{if } v_n(t) = 0, \\ p_d, & \text{otherwise.} \end{cases} \quad (5.6)$$

In the first case ( $p = p_b$ ), the vehicle  $n$ , following the preceding car  $n + 1$  with time headway  $t_h(t)$ , reacts to the predecessor's activated brake light ( $b_{n+1}(t) = 1$ ) if it is within an interaction horizon  $t_s(t)$  ( $t_h(t) < t_s(t)$ ). This case was introduced to appropriately mimic the drivers' tendency to overreact in dense traffic. This phenomenon, sometimes called "over-deceleration effect" [54, p. 260], results from a driver's finite response time to the preceding car's brake lights. If the time headway is too small (i.e.,  $t_h(t) < t_s(t)$ ) the following driver performs an unnecessarily strong braking maneuver to avoid collision. The second case ( $p = p_0$ ), known as slow-to-start rule, effects a reduced acceleration rate for vehicles starting from rest. The last case models the random fluctuations observed even in free traffic flow. From these explanations, we can already conclude the relation  $p_d < p_0 < p_b$ .

In the modified model, we require vehicles which changed their driving behavior (i.e.,  $j_{\text{bool}}(t) \leftarrow \text{true}$ ), to keep a larger gap to the preceding vehicle. Such an increased gap is created simply by not accelerating until the gap is large enough. An additional gap, whose length we denote by  $g_n(t)$ , is considered as sufficiently large if it is larger than the distance traveled during the driver's reaction time  $t_{\text{reaction}}$  of 1 s ( $t_{\text{reaction}} = 1 \text{ s}$ )<sup>1</sup>, i.e., if  $g_n(t) > v_n(t) \times t_{\text{reaction}}$ . It is intuitive that a larger gap decreases the probability for the occurrence of an over-deceleration effect. (For a detailed discussion of the impact of driving behavior on traffic flow, we refer to a recent article by Kerner [50].)

Consequently, we modified the calculation of the randomization parameter  $p$  by introducing a fourth parameter  $p_c$  ( $p_0 < p_c < p_b$ ) to model the situation described

<sup>1</sup>The reaction time does not only differ from person to person but also from situation to situation for a single person [43]. Therefore, our choice for the driver reaction time must be understood as an average value. This value is in good agreement with empirical data that report the median of the driver reaction time is below 1 second [97, p. 42]. We admit, however, that any other choice would not have been possible in the CDM due to the model's temporal discretization.

above:

$$p \leftarrow \begin{cases} p_b, & \text{if } b_{n+1}(t) = 1 \text{ and } t_h(t) < t_s(t) \\ & \text{and } g_n(t+1) \leq v_n(t) \times t_{\text{reaction}}, \\ p_c, & \text{if } b_{n+1}(t) = 1 \text{ and } t_h(t) < t_s(t) \\ & \text{and } g_n(t+1) > v_n(t) \times t_{\text{reaction}}, \\ p_0, & \text{if } v_n(t) = 0 \\ p_d, & \text{otherwise.} \end{cases} \quad (5.7)$$

For the simulations, we adopted the values for  $p_b$ ,  $p_0$ , and  $p_d$  from section 3.1.3 and set  $p_c = 0.8 \times p_b$ .

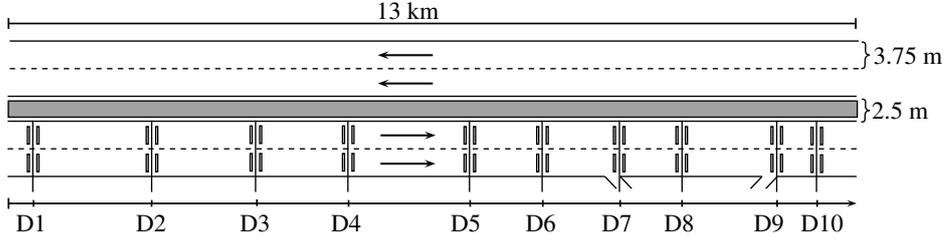
### 5.3.3 Simulation

We used computer simulations to assess the impact of the proposed strategy, and we made a big effort to provide a high degree of realism. Therefore, our simulation combined a realistic traffic model (the CDM, see chapters 3 and 4), empirical traffic data, and a realistic path loss model (the Nakagami- $m$  model, see section 5.1.1.4). (At this point, it should be mentioned that some very recent studies [14, 98] promise an even better path loss modeling by explicitly taking into account vehicles as obstacles for signal propagation.)

#### 5.3.3.1 Simulation Setup

As scenario we used the same 13 km-long two-lane highway segment as in section 3.2. In contrast to section 3.2, however, we also simulated the traffic flow of the opposite driving direction as well. This is necessary to put a realistic load on the radio channel because vehicles traveling in the opposite direction use the same channel to communicate. For simplicity, we used identical vehicle inflow rates for both driving directions, but we omitted the on-ramp and the off-ramp for the newly added, opposite driving direction. Therefore, flow in each direction is approximately 36 000 vehicles per day with additional 7000 vehicles joining the road via the on-off-ramp system in one of the two directions. Vehicles leaving the system, as required by the boundary conditions, were selected randomly without preference for (non-)communicating vehicles. Therefore, the off-ramp has no influence on the proposed strategy, but it is necessary to partially compensate for vehicles entering via the on-ramp. Similarly, entering vehicles are communicating according to the penetration rate of communicating vehicles. The lane width was set to 3.75 m with additional 2.5 m between opposite driving directions. The geometry of the entire road segment is given in figure 5.5.

Every radio equipped vehicle shares an identical set of physical parameters that are summarized in table 5.1. All values are in agreement with the draft standard [24] by the ASTM. As data rate, we chose the minimum of the mandatory rates 3, 6, and 12 Mbits/s. For modeling radio propagation and signal fading we used the Nakagami- $m$  model, which we chose on grounds of the good agreement with real-world measurements



**Figure 5.5** Schematic sketch of the highway segment used for simulations. The locations of loop detectors are labeled as D1, . . . , D10.

described in section 5.1.1.4. For communication distances of 80 m and above, we set  $m = 0.75$ . For smaller distances, where a clear line-of-sight is likely to exist, we set  $m = 1.5$ .

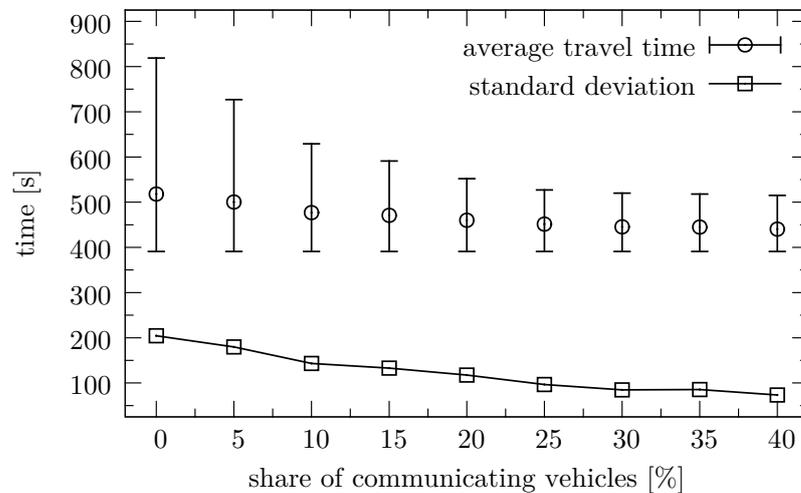
The medium access control (MAC) layer was implemented according to the ASTM-standard [24], too. Status messages have a constant size of 500 byte and are broadcast with a frequency of 4 Hz. Note that both message size and frequency fulfill the communication requirements for safety applications presented by the US Department of Transportation [109].

**Table 5.1** Properties of radio devices and radio channel.

<b>PHY</b>	
fading	Nakagami- $m$
interference	cumulative noise & capture effect
transmission power	17 dBm ( $\approx 50.12$ mW)
reception threshold	-81 dBm ( $\approx 7.94 \times 10^{-9}$ mW)
sensitivity threshold	-91 dBm ( $\approx 7.94 \times 10^{-10}$ mW)
noise	-99 dBm ( $\approx 1.26 \times 10^{-10}$ mW)
antenna gain	0 dBm (1 mW)
SINR	10
data rate	3 Mbits/s
<b>MAC</b>	
beacon size	500 B
beacon interval	250 ms
jitter	4 %

### 5.3.3.2 Results

To assess the impact of vehicle communication on traffic flow, we analyzed the travel times for the considered highway segment and the given boundary conditions. Travel time is probably the most intuitive quality measure in this context. For commuters travel time is even the most important determinant of the highway quality of service [40] as they consider travel time as “lost”. The microscopic simulation recorded the time it took each vehicle to pass the entire stretch of highway considered. Figure 5.6 shows the average travel times for different penetration rates of communicating vehicles. The results were averaged over at least five independent simulation runs per penetration rate. The average travel time drops from 518 s without vehicle-to-vehicle communication to 440 s when 40 % of all vehicles are equipped with communication devices. This corresponds to a travel time reduction of more than 15 %. Error bars mark the positions of the 0.1- and 0.9-quantiles, respectively, which means 80 % of all vehicles were able to cross the highway segment within the error-bar-indicated interval. The lower bound is mainly delimited by the minimal travel time following from the vehicles’ maximum velocity (see table 3.3). The upper bound, on the other hand, reflects the traffic dynamics and the existence of breakdowns. Here, the benefits of vehicle-to-vehicle communication become more obvious; with one fourth of all vehicles being able to communicate, the 0.9-quantile drops for more than 35 % from 819 s to 527 s.

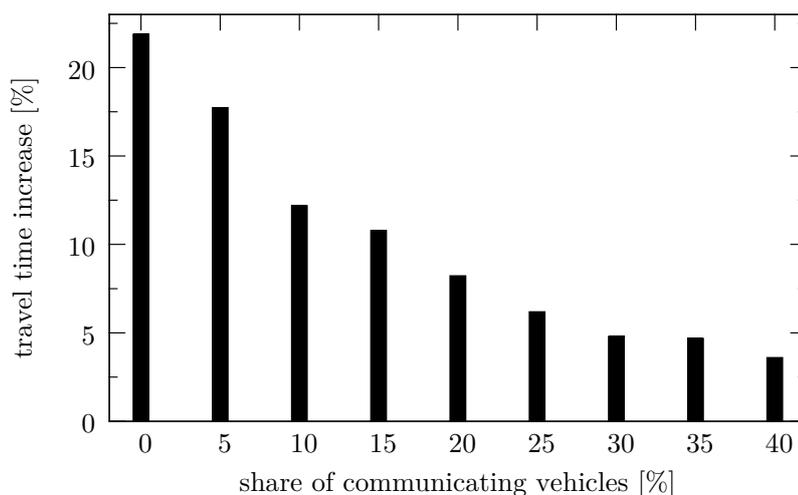


**Figure 5.6** Average travel times for different penetration rates. Error bars indicate the 0.1- and 0.9-quantiles. The lower curve shows the corresponding standard deviation.

In addition, the lower curve in figure 5.6 shows the standard deviation of observed travel times. The standard deviation can be understood as a measure of travel time reliability [40]. It shows a similar dependence on the penetration rate; with one in four vehicles being able to communicate, the standard deviation’s value measures only 96 s compared with more than 200 s when communication is turned off.

To determine the average increase of time delay in each scenario (i.e., for each

penetration rate), we calculated an optimal travel time. To do so, we started a vehicle at each second of the simulated day on an empty road and recorded the corresponding travel time. Taking into account the share of trucks and trucks' increased travel time due to their reduced maximum velocity (see table 3.3), one obtains an optimal average travel time of 425 s. This time serves as a reference value in figure 5.7, which shows how much the scenarios deviate from the ideal condition with no interactions.

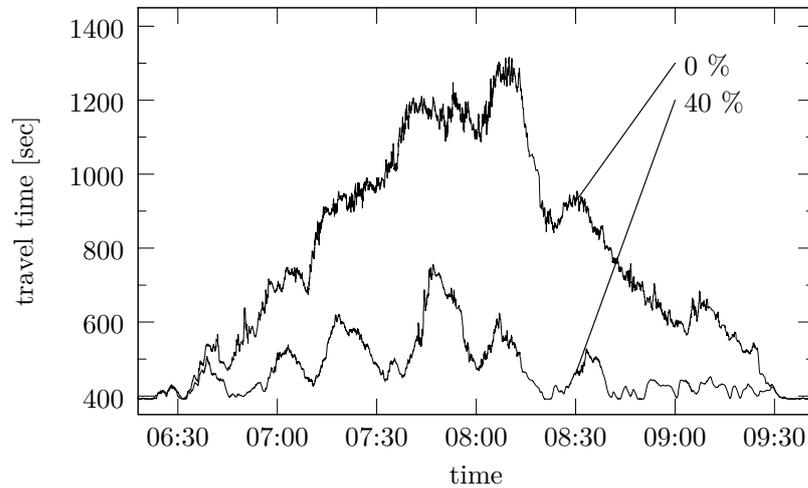


**Figure 5.7** Average travel times compared with an ideal travel time without vehicle interactions. This time is not expected to be reached because, even in dilute traffic, interactions between vehicles in the merging region of the on-ramp casually cause braking maneuvers. Nevertheless, for a penetration rate of 40 % the increase in average travel time is only 3.6 % compared with the ideal travel time.

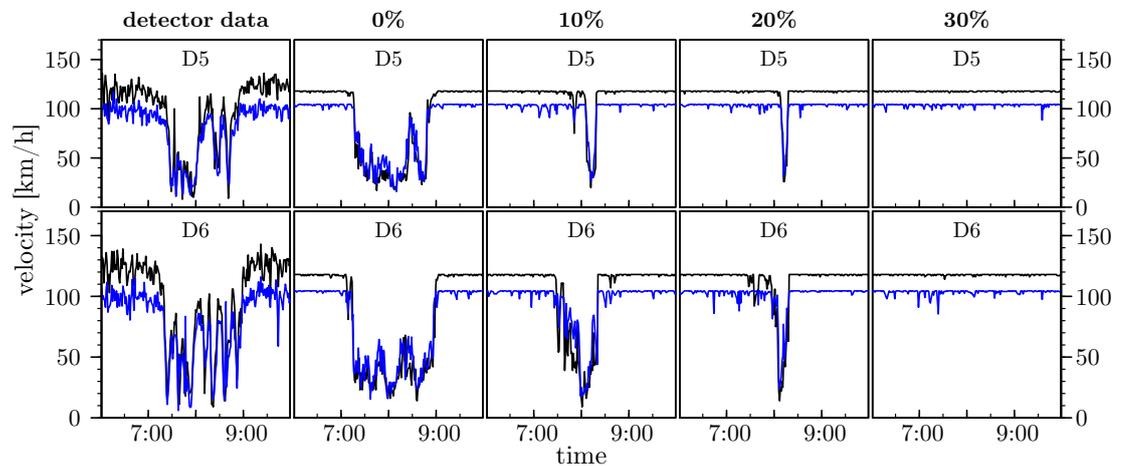
The temporal traffic dynamics is illustrated in figure 5.8 where travel times of single vehicles are plotted against the hour of day. The application of vehicle communication significantly reduces the duration and the severity of the congestion observed in the time from 6:30 a.m. to 9:30 a.m. This result is confirmed by the analysis of the stationary detector data shown in figure 5.9.

An even better visualization can be obtained from the space–time–velocity plots of figure 5.10. The figure illustrates very well the decrease of both the spatial and the temporal extent of congestion for an increasing share of communicating vehicles. Upstream of the on-ramp, i.e., for positions  $x > 11$  km, the average velocities show homogeneous vertical stripes for some time intervals. This follows from the speed limit that was applied (see section 3.3.1) to the end of the road segment and that causes all vehicles to move at the same speed. Consequently, all plots of figure 5.10 show a similar spatiotemporal pattern for  $x > 11$  km.

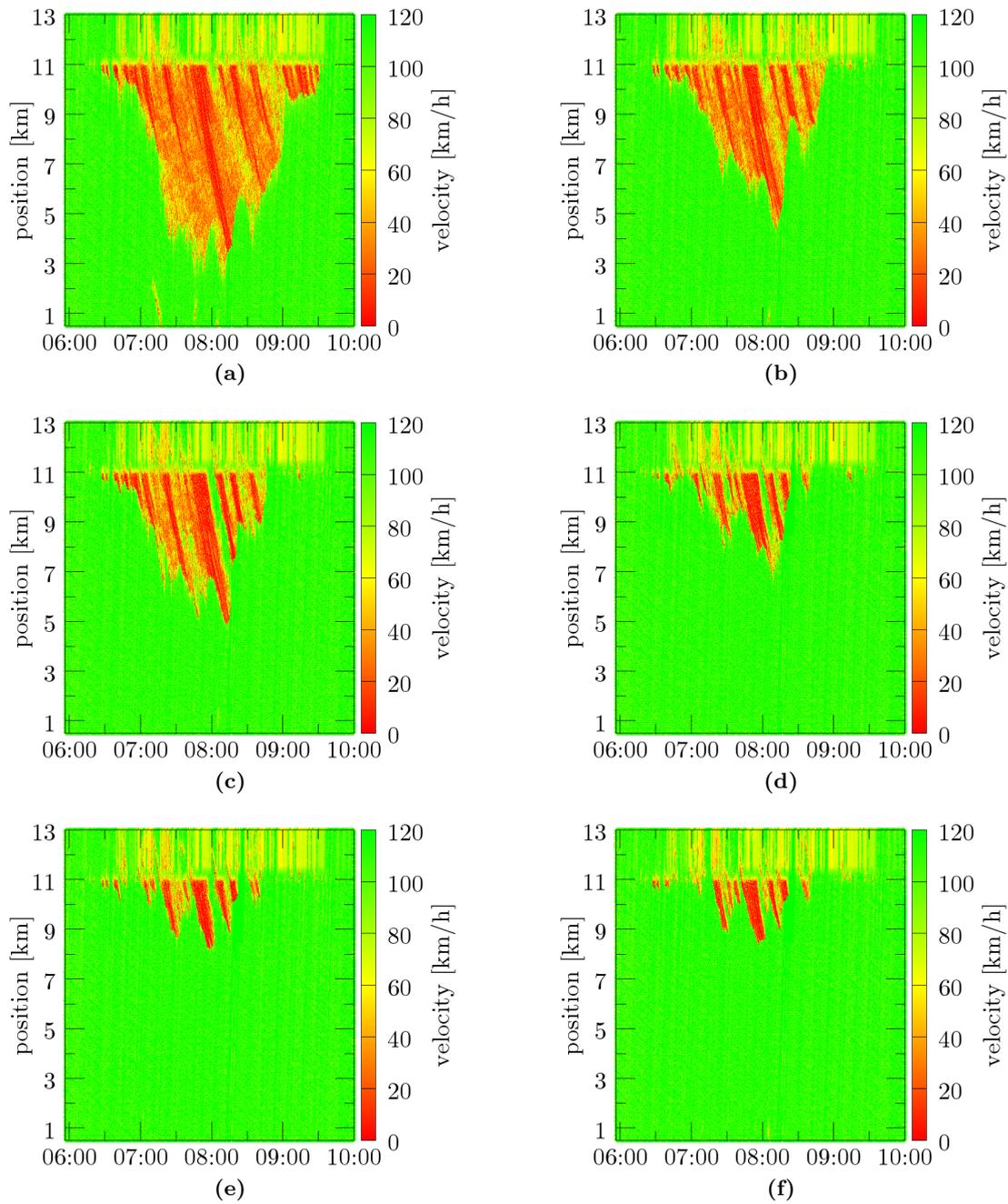
The speed limit might also explain why the congested traffic does not entirely vanish at high rates of communicating vehicles. Therefore, we repeated our simulations for an idealized highway scenario. This idealized scenario studied a two-lane highway



**Figure 5.8** The evolution of single vehicle travel times during the morning peak hour. For better readability, the plot shows only data from vehicles of type “car”, and a moving average over ten subsequent data points is applied.



**Figure 5.9** Exemplary time series of the two detector locations D5 (top) and D6 (bottom) at various rates of communicating vehicles. The left column shows the empirical time series. The next columns (from left to right) show the simulated time series for 0%, 10%, 20%, and 30% communicating vehicles. As can easily be seen, the extent of the breakdown decreases with an increasing rate of communicating vehicles.



**Figure 5.10** The spatiotemporal evolution of the average velocity for a rate of communicating vehicles of (a) 0%, (b) 10%, (c) 20%, (d) 30%, (e) 35%, and (f) 40%.

segment of length 18 km (12 000 cells) with a 300 m-long on-ramp located 1.5 km from the downstream boundary. The on-ramp inflow was kept constant at 450 vehicles/h during the entire simulation time of 6.5 h. To simulate peak hour traffic, the demand of the main road was varied as follows (see figure 5.11a): the inflow of the initially empty main road was 1000 vehicles/h/lane for the first 30 min. Next, the inflow was linearly increased to 1400 vehicles/h/lane during a 2 h-interval. For the following 3 h, the inflow rate was linearly decreased to 1000 vehicles/h/lane where it was kept constant until the simulation ended. The share of trucks was kept constant at 10 % in our idealized peak hour scenario. Thus, a total of 15 000 vehicles entered the main road from the upstream boundary and an additional 2925 vehicles from the on-ramp. The maximum value of 1400 vehicles/h/lane could be reached in all simulations presented below (i.e., jams did not propagate to the upstream boundary).

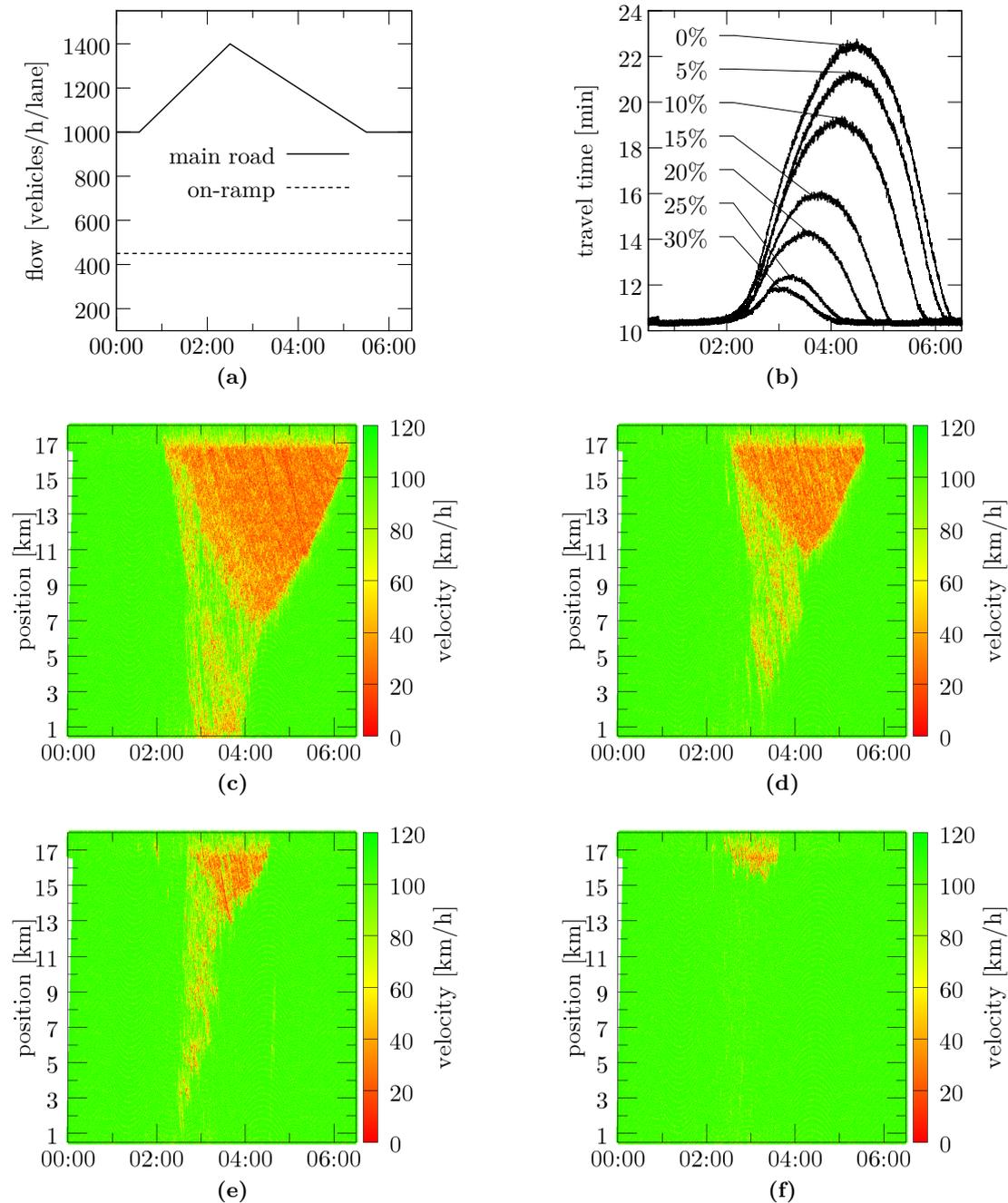
Figure 5.11 shows the spatiotemporal dynamics of all vehicles moving on the road with the on-ramp.

One can identify the onset of congestion in the on-ramp area after approximately 2.5 h (i.e., after the main road inflow reached its maximum). As a result of the decreasing flow after 2.5 h, free flow was restored after maximally 6 h in all cases. Without vehicle-to-vehicle communication (figure 5.11c) the impact of the on-ramp induced traffic breakdown could be observed even 10 km upstream of the on-ramp. With an increasing percentage of equipped vehicles (figures 5.11d, 5.11e, and 5.11f), both the spatial and temporal extent of regions with reduced velocity decreased significantly. This means that both the number of vehicles affected by the on-ramp induced breakdown and the degree of affection decreased as well. With 30 % of all vehicles being equipped, the peak hour induced perturbations nearly vanished and became tightly localized around the on-ramp. Due to the merging process and the high vehicle flow, we do not expect perturbations to disappear completely. To quantify the positive impact of IVC on traffic flow, we recorded the travel time of each vehicle (see figure 5.11b). If no vehicle is able to communicate, the travel time more than doubled for some vehicles from approximately 10 min in dilute traffic to more than 22 min. With penetration rates of 15 % and above, it took no vehicle longer than 16 min to pass the same distance. With free flow at the downstream boundary, the results suggest that rates of communicating vehicles of 30 % and above practically restore free flow on the entire road except for a narrow region around the on-ramp.

Thus, all results suggest that the proposed strategy is indeed able to reduce congestion and vehicular communication is an appropriate mean to optimize traffic flow.

### 5.3.3.3 Dependence of the Results on the Choice of Parameters

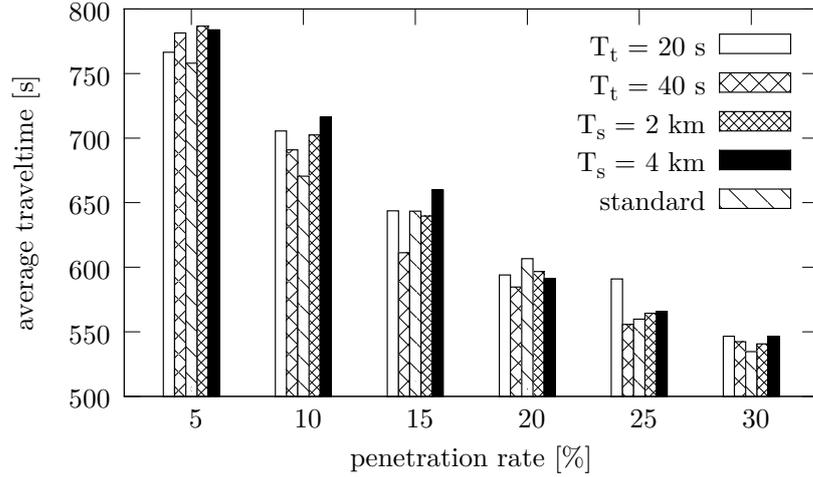
Another important aspect is the sensitivity of our results and, consequently, of the proposed strategy on the choice of newly introduced parameters. In section 5.3.1, the variables  $T_s$  and  $T_t$  were introduced to define an area of relevance for traffic information. Only vehicles within this spatiotemporal area change their driving behavior in response to received messages. To check the sensitivity of our results on the choice of  $T_s$  and  $T_t$ , we modified the original values of  $T_s = 3$  km and  $T_t = 30$  s and analyzed travel



**Figure 5.11** Figure (a) shows the temporal evolution of the flow in our idealized peak hour scenario. The resulting evolution of the single-vehicle travel times is given in figure (b). (The maximum velocity was 108 km/h for cars and 81 km/h for trucks.  $T_v$  was set to 65 km/h.)

Examples of the spatiotemporal vehicle dynamics for (c) 0%, (d) 10%, (e) 20%, and (f) 30% of communicating vehicles.

times in the interval from 6:00 a.m. to 9:00 a.m., in which most fluctuations occur (see figures 5.8 and 5.9, left column). As shown in figure 5.12, the qualitative behavior of travel times is conserved when changing  $T_s$  and  $T_t$  for  $\pm 1/3$  of their original value.

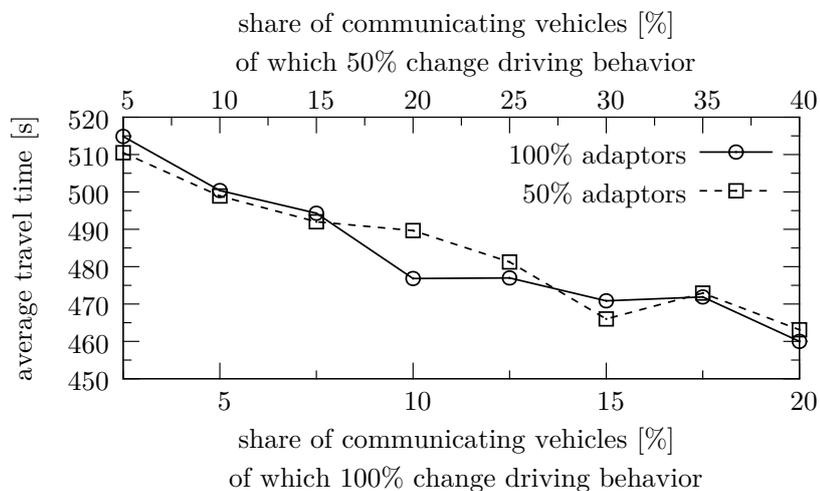


**Figure 5.12** Average travel times in the time from 6:00 a.m. to 9:00 a.m. for different values of  $T_s$  and  $T_t$ . The average travel time obtained with the original values of  $T_s$  and  $T_t$  serves as reference and is labeled “standard”.

Lee and Kim [90] made a similar observation when they studied the dissolution of traffic jams after introducing additional local interactions between vehicles. For the local interactions that they introduced, each vehicle evaluated the velocity of a single downstream vehicle at distance  $d$ . They found their results to be independent of the concrete choice of  $d$  over a broad range of values.

Although we made no suggestion how the proposed strategy can be put into practice, we wanted to assess its success if some vehicles or drivers do not (or cannot) follow the recommended change of driving behavior. For this purpose, we repeated the simulations and made half of the communicating vehicles ignore any recommendation on the driving behavior. However, these vehicles did correctly broadcast messages, as described earlier. We compared the average travel times from these simulations to the previous simulation runs. In figure 5.13, we plotted the average travel times obtained for different shares of communicating vehicles. At each data point, we also show the average travel time that we obtained when the share of communicating vehicles was only half as large but when all of these vehicles were able to adapt their driving behavior.

Here, we found a very good agreement between the values for a given penetration rate (e.g., 40%) and the results obtained for half this penetration rate (e.g., 20%) when all communicating vehicles follow the recommendation to change the driving behavior. Even the largest deviation found at a penetration rate of 20% (corresponding to 10% with all vehicles changing their driving behavior) is below half of a standard deviation (see figure 5.6). Hence, the success of our strategy depends crucially on the drivers’



**Figure 5.13** A comparison of the average travel times for various rates of communicating vehicles where we distinguished two cases: (i) all communicating vehicles that should change their driving behavior actually do change their driving behavior, and (ii) only half of the vehicles that should do so adapt their driving behavior by setting  $j_{\text{bool}} \leftarrow \text{true}$ .

willingness (or capability) to follow the recommendation, whereas the benefit from better connectivity was found to be negligible.

## 5.4 Conclusion

In this chapter, we have presented a method to reduce congestion and improve traffic flow via vehicle-to-vehicle communication. Solely based on periodic beacon messages and using only velocity and position as a source of traffic state estimation, the proposed method makes minimal requirements regarding the technical implementation.

The impact of the proposed strategy was evaluated by simulations employing a bidirectionally coupled simulator. The traffic simulator used empirical loop detector data, which show a breakdown of traffic flow during the morning peak hour and which the simulator is able to reproduce.

In contrast to several previous studies where route choice was altered in response to an already existing jam, our approach becomes effective before a jam occurs. For this purpose, our approach requires a traffic state analysis which does not only take into account vehicle density; by evaluating vehicles' position and velocity, vehicles are able to decide when and where to change their driving behavior. The success of this approach can be seen from figure 5.9; with a sufficient number of vehicles applying our strategy, the original traffic breakdown, seen in both the empirical and the simulated data, can no longer be observed at penetration rates of 20% at the detectors' positions in figure 5.9.

Simulations showed that even low penetration rates suffice to considerably improve traffic flow. With only one in ten vehicles being able to communicate, the increase in travel time sinks from 22 % to 12 % compared with the case without communication. Only slight improvements were found for penetration rates of 30 % and above.

All of the aforementioned points suggest VANETs are an adequate means to increase traffic efficiency and to improve traffic state estimation.



## Knowledge in VANETs: a Random Walk Approach

A combinatorial problem of heterogeneous traffic flow with communicating and non-communicating vehicles was the starting point for the analysis presented in this chapter. As shown in the previous chapter, information exchange between vehicles is expected to enhance traffic safety and traffic stability notably [19, 54, 63, 64, 90]. After market introduction, however, only a small fraction of vehicles will be equipped with the necessary communication devices. On the other hand, local sensors (e.g., radar or lidar) enable vehicles to gather information about the preceding and succeeding vehicle. Consequently, a communicating vehicle may not only broadcast information about itself but also information about the vehicle driving behind or ahead. Hence, the number of vehicles of which the position is known may be significantly higher than the actual number of communicating vehicles alone. This additional information may be used to improve traffic surveillance or traffic safety.

Not only for traffic related applications it is an interesting question how many vehicles are on average known (i.e., either actively communicating or being detected by a neighboring communicating one) for a given share of communicating vehicles on the road.

To answer this question, we first generalize the problem by translating the sequence of vehicles into a one-dimensional particle chain with two types of particles, to which we will either refer as “active” or “passive”. The traffic flow example in mind also explains our choice to speak of active and passive particles. When studying exclusion processes [23], one would rather speak of occupied and empty sites and in the context of zero range processes (e.g., [25]) of sites and particles, respectively.

For random configurations of active and passive particles we determine the average number of passive particles neighboring an active one by mapping the system to a one-dimensional (1D) discrete random walk. This mapping even allows for an interpretation which is closer to physics. If one thinks of the different particle types as changes in the height profile of a 1D surface by  $\pm 1$ , then the fraction of passive particles next to an active one corresponds to the surface’s extremal-point density (see figure 6.1b). In

---

This chapter is largely based on and taken from the author’s article [T75] (see page 114).

general, such extremal-point densities allow to study the dynamics of non-equilibrium surface fluctuations and have applications to an even broader range of research. For a detailed discussion of the dynamics of rough surfaces, the density of local extrema, and analytic solutions, we refer to the very comprehensive article [135] by Toroczkai et al.

For our example of heterogeneous traffic, we will examine the case where communicating vehicles are equipped with front and rear sensors or with front sensors only. In the picture of active and passive particles, this corresponds to the situation where a passive particle can be detected by both its left and its right active neighbor or to the situation where only an active particle to its left can detect it. It turns out that in the latter case a mean field approximation reproduces the exact result. Before we proof this, let us formulate the problem in a more formal way.

## 6.1 Model Description

To analyze the problem presented in the introduction, we start with a one-dimensional lattice with  $N$  sites. Each lattice site contains either an active or a passive particle (represented by  $\bullet$  or  $\circ$ ). Let the number of active (passive) particles be  $A$  ( $P$ ) and  $N = P + A$ . Then it follows directly that the probability of a randomly selected site to contain an active (passive) particle is  $p = A/N$  ( $1 - p = P/N$ ).

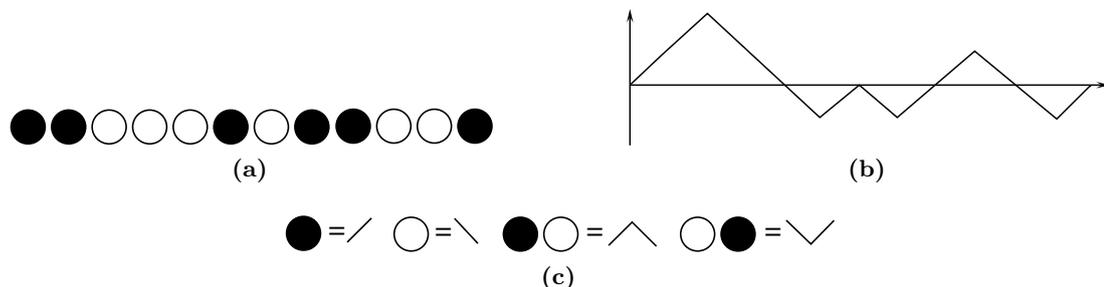
Active particles are always assumed to be known. For a passive particle to be known it has to neighbor an active one. Here we distinguish two cases:

1. **Symmetric case:** A passive particle is called known if at least one of its two neighbors is active.
2. **Asymmetric case:** A passive particle is called known only if its left neighbor is active.

The knowledge about the system (i.e., the particle chain) obviously depends on the average number of known passive particles in a given configuration. Figure 6.1a depicts a sample configuration with  $N/2 = A = P = 6$ . In this configuration all but one (three) particle(s) are known in the symmetric (asymmetric) case, independent of the boundary condition.

In general, the number of known particles does not only depend on the amount of active particles but also on their distribution on the lattice. For instance, to have full knowledge of the system (i.e., for all particles to be known) the minimum share of active particles is  $p = 1/3$  in the symmetric case ( $\circ \bullet \circ \bullet \circ$ ) and  $p = 1/2$  in the asymmetric case ( $\bullet \circ \bullet \circ$ ).

At first sight, the usage of only two particle types, representing communicating and non-communicating vehicles, appears as a rough approximation of real traffic flow as we ignore the inter-vehicle distance. We admit that the introduction of additional empty lattice sites would better reflect the actual situation on a real road or in traffic cellular automata (e.g., [102]), as considered in section 6.3. But the approximation is justified since the detection range of vehicle sensors of about 200 m is larger than the typical



**Figure 6.1** The presented model consists of a chain with two types of particles: active ones and passive ones. (a) Active (passive) particles are represented by  $\bullet$  ( $\circ$ ). The number of known particles comprises all active particles and all passive particles next to an active one. In the sample configuration of (a) there are 11 particles out of 12 known in the symmetric case and 9 out of 12 in the asymmetric case. (b) The model can also be viewed as a 1D random walk where the different types of particles correspond to steps in opposite directions. In this picture passive particles become known (i.e., they are detected) whenever the direction of the walk is reversed. Hence, the fraction of known passive particles is also related to the density of extremal points of the corresponding random walk. Figure (c) illustrates the conversion between the two representations.

space gap between two vehicles [81, 86]. Besides, the case where the inter-vehicle spacing exceeds the sensors' range is of little interest because in this case the contribution of vehicular communication to traffic safety and efficiency will be marginal at best.

## 6.2 Determining the Fraction of Known Particles

In this section we determine the fraction of known particles  $k_{a,s}^{o,p}$  as a function of the total particle number  $N$  and the number of active particles  $A$ . The upper index in  $k_{a,s}^{o,p}$  specifies whether we consider open (o) or periodic (p) boundaries. The lower index denotes the symmetric (s) or asymmetric (a) case.

As a passive particle becomes known when it is next to an active one, we have to count all  $\bullet\circ$  (and  $\circ\bullet$ ) sequences among all  $\binom{N}{A}$  possible configurations in the asymmetric (symmetric) case. We will start to derive an analytic solution for the symmetric case for periodic and open boundary conditions. The extension to the asymmetric case is straight forward from there. For the latter case it actually turns out that a simple mean field approximation yields the exact result. The mean field approximation can be represented by a single-valued function  $k_{a,s}(p)$  with  $p = A/N$ .

### 6.2.1 Analytic Solution of the Symmetric Case

It is helpful to interpret the occupied lattice as a one-dimensional random walk where each active particle corresponds to a step to the left and each passive particle to a step

to the right or vice versa. Figure 6.1b illustrates this analogy. In this picture passive particles are discovered whenever the direction of the walk is reversed (i.e., when a corner occurs). Care must be taken, however, when a passive particle neighbors two active particles ( $\bullet \circ \bullet$ ). In this case only one of the two corresponding corners stands for a newly discovered particle.

Hence, the function  $k_s^o(N, A)$  can be calculated by determining the number of paths with  $R$  corners and weighing it by a factor proportional to  $R$ . Thereby, our model is similar to a problem by Feynman and Hibbs [28]: to calculate the kernel of a relativistic particle moving in  $1 + 1$  dimension, Feynman suggested summing over all possible paths with  $A$  ( $P$ ) steps to the left (right) and weighing each path with a factor depending on the number of corners.

To count the configurations which have exactly  $R$  corners, we have to distinguish four cases:

1. The first particle is active and the last one is passive ( $\bullet \cdots \circ$ ).
2. The first particle is passive and the last one is active ( $\circ \cdots \bullet$ ).
3. Both the first particle and last particle are active ( $\bullet \cdots \bullet$ ).
4. Neither the first particle nor the last particle are active ( $\circ \cdots \circ$ ).

In analogy to Jacobson and Schulman, who presented a solution [46] to Feynman's problem, we introduce the function  $\Phi_{xy}(R)$ . The function returns the number of configurations/paths with  $R$  corners where the first particle is of type  $x$  and the last one is of type  $y$ . For  $x = \bullet$  and  $y = \circ$  one obtains [46]:

$$\Phi_{\bullet\circ}(R) = \binom{A-1}{\frac{1}{2}(R-1)} \binom{P-1}{\frac{1}{2}(R-1)} \quad (6.1)$$

with odd  $R$  and  $0 \leq R-1 \leq 2 \min(A-1, P-1)$ . Due to the problem's symmetry it immediately follows  $\Phi_{\bullet\circ}(R) = \Phi_{\circ\bullet}(R)$ .

For  $\Phi_{\bullet\bullet}(R)$  we obtain:

$$\Phi_{\bullet\bullet}(R) = \binom{P-1}{\frac{1}{2}R-1} \binom{A-1}{\frac{1}{2}R} \quad (6.2)$$

for even  $R$  and  $2 \leq R \leq \min(2A-2, 2P)$  from which follows

$$\Phi_{\circ\circ}(R) = \binom{A-1}{\frac{1}{2}R-1} \binom{P-1}{\frac{1}{2}R} \quad (6.3)$$

with  $R$  even and  $2 \leq R \leq \min(2P-2, 2A)$ .

To count passive particles in a  $\bullet \circ \bullet$  configuration only once, one of the corresponding corners must be skipped. With periodic boundaries the probability  $p_{\bullet\bullet\bullet}(N, A)$  for such

a configuration to occur on a lattice with  $N$  sites and  $A$  active particles is (see, e.g., [116])

$$p_{\bullet\circ\bullet}(N, A) = \frac{A-1}{N-2}. \quad (6.4)$$

For convenience, we introduce

$$n = N \binom{N}{A}.$$

Then it directly follows that the average fraction of known particles  $k_s^o$  as a function of  $N$  and  $A$  is:

$$k_s^o(N, A) = \frac{1}{n} \sum_R \sum_{x,y \in \{\circ, \bullet\}} \Phi_{xy}(R) \times \left( R + A - \frac{R}{2} p_{\bullet\circ\bullet}(N+1, A) \right) \quad (6.5)$$

where we assumed the  $\Phi_{xy}(R)$  to return zero if  $R$  is not in the set of valid values as specified in equations (6.1)–(6.3). The  $1/n$ -term normalizes the probability such that  $k(N, N) = 1$ . The  $+1$ -term in the first argument of  $p_{\bullet\circ\bullet}(\cdot, \cdot)$  is a consequence of the open boundary conditions. As with open boundaries the leftmost and rightmost particles have only one neighboring site, the probability for a  $\bullet\circ\bullet$  sequence to occur is identical to the one on a periodic lattice with  $N+1$  sites where the additional site is occupied by a passive particle. By using the normalization  $\sum_R \sum_{x,y \in \{\bullet, \circ\}} \Phi_{xy}(R) = \binom{N}{A}$ , we can rewrite equation (6.5) as

$$k_s^o(N, A) = \frac{A}{N} + \left( \frac{1}{n} - \frac{1}{2n} \frac{A-1}{N-1} \right) \sum_R \sum_{x,y \in \{\circ, \bullet\}} R \times \Phi_{xy}(R). \quad (6.6)$$

The sums in equation (6.6) can be converted to hypergeometric series (see appendix A) and turn out to be particular cases of Gauss's hypergeometric theorem. Thereby, the previous formula can be reduced to a simple rational function depending only on the total number of particles and the number of active particles:

$$k_s^o(N, A) = \frac{A(A^2 + A - 3NA + 3N^2 - 2N)}{N^2(N-1)}. \quad (6.7)$$

The extension to periodic boundaries is straight forward. When closing the chain to form a ring, one additional corner originates for all  $\Phi_{\bullet\circ}(R)$  and  $\Phi_{\circ\bullet}(R)$  configurations. In this case one has to replace  $R \rightarrow R+1$  for the weighing factor in round brackets in equation (6.5) and  $\bullet\circ\bullet$  sequences need no special treatment which leads to

$$k_s^p(N, A) = \frac{A(A^2 + 3A - 3AN + 3N^2 + 2 - 6N)}{N(N-1)(N-2)}. \quad (6.8)$$

Finally, let us compare the error of a simple mean field approximation with the exact result. One obtains the mean field approximation by noting that the only configuration in which a passive particle remains unknown is when both its neighbors are passive

particles as well. The probability for this configuration is  $(1 - p)^3$ . In all other cases, i.e.,  $(1 - (1 - p)^3)$ , a passive particle is known, which yields

$$k_s(p) = k_s(A/N) = p^3 + 3p(1 - p) \quad (6.9)$$

The larger is the system (i.e., the number of lattice sites  $N$ ), the better is the agreement between equation (6.7)/(6.8) and (6.9). For, in the thermodynamic limit  $N \rightarrow \infty$  and  $p = \text{const.}$  (i.e.,  $N \approx (N - 1)$  and  $A^i/N^j \rightarrow 0$  for  $0 < i < j$  and  $A < N$ ), the mean field approximation converges to  $k(N, A)$  as  $k(N, A) = k(p) + O(N^{-1})$ .

### 6.2.2 Analytic Solution of the Asymmetric Case

In the asymmetric case a passive particles becomes known only if its left neighbor is active. (We restrict the discussion to this case, although the results are also valid if a passive particle requires an active particle to its right in order to be known.)

With a similar reasoning as in the previous section, one can derive the fraction of known particles as a function of the lattice size  $N$  and the number of active particles  $A$ . In general, every second corner in the corresponding random walk stands for a known passive particle. Care must be taken for the  $\Phi_{xy}(R)$  configurations with  $x \neq y$ : these configurations have an odd number of corners. Moreover, there is at least one known passive particle for each  $\Phi_{\bullet\circ}(R)$  configuration which is not guaranteed for  $\Phi_{\circ\bullet}(R)$  configurations. This leads to

$$k_a^o(N, A) = \frac{1}{n} \sum_R \left[ (\Phi_{\bullet\bullet}(R) + \Phi_{\circ\circ}(R)) \left( \frac{R}{2} + A \right) + \Phi_{\bullet\circ}(R) \left( \frac{R+1}{2} + A \right) + \Phi_{\circ\bullet}(R) \left( \frac{R-1}{2} + A \right) \right]. \quad (6.10)$$

Evaluating the sums gives

$$k_a^o(N, A) = k_a^o(p) = p(2 - p) \quad (6.11)$$

for open boundaries. This is exactly the same result that a mean field approximation yields, as one can easily verify: in the asymmetric case, the only configuration in which a passive particle remains unknown is the one in which the particle to its left is passive ( $\circ\circ$ ). The probability for this sequence is  $(1 - p)^2$  and, consequently,

$$k_s(p) = 1 - (1 - p)^2 = p(2 - p). \quad (6.12)$$

A similar calculation for periodic boundary conditions yields:

$$k_a^p(N, A) = \frac{A(2N - A - 1)}{N(N - 1)}. \quad (6.13)$$

### 6.2.3 The Asymmetric Case and the Density of Local Maxima

We would like to discuss briefly the relation between the fraction of known passive particles and the density of local maxima in the random walker's path, which we have already mentioned in the introductory section. For the microscopic characterization of a rough surface, its extremal-point density may be of similar importance as is the width of the surface for its macroscopic characterization [135].

Comparing figures 6.1a and 6.1b shows: each  $\bullet\circ$  configuration in figure 6.1a results in a local maximum in figure 6.1b. Hence, we obtain the average density of local maxima  $\rho_{\max}^o$  by setting  $A = 0$  on the right-hand side of equation (6.10):

$$\rho_{\max}^o(N, A) = \frac{A(N - A)}{N^2} = p(1 - p). \quad (6.14)$$

As we average over all configurations, the density of local maxima and minima is identical ( $\rho_{\max}^o = \rho_{\min}^o$ ). An analogous calculation for periodic boundary conditions yields

$$\rho_{\max}^p(N, A) = \rho_{\min}^p(N, A) = \frac{A(N - A)}{N(N - 1)}. \quad (6.15)$$

(Note that the interpretation of periodic boundary conditions for the random walker's path is somewhat difficult as the two ends of the 1D surface have different heights if  $A \neq P$ .)

## 6.3 Simulations for Vehicular Networks

Finally, we validated our theoretical predictions with simulations of a vehicular communication network on a circular one-lane road of radius 1.5 km. Each communicating vehicle was assumed to send periodic status messages with a frequency of 4 Hz. (The full set of parameters was taken from table 5.1.) These messages comprised the vehicle's position and velocity as well as the position and the velocity of the preceding (and following) vehicle(s) in the asymmetric (symmetric) case. For realistic communication modeling we used the probabilistic Nakagami- $m$  model (section 5.1.1.3). Vehicle motion was simulated using the CDM again. The road was initialized with densities ranging from 5% up to 65%. Higher densities were omitted because at a density of 65% the average bumper-to-bumper distance already is below 4.1 m and a large traffic jam spans the entire road. Thus, from a practical point of view it is not necessary to know the position and velocity of each vehicle to estimate the traffic dynamics at such high vehicle densities. For our analysis we initialized the road homogeneously (i.e., vehicles were positioned at the same distance from each other) and recorded the communication statistics of a 60s interval after a relaxation time of 240s. We averaged over at least five independent runs for each density.

To assess the validity of the theoretical calculations for real-world applications, we compared the simulations' results to the corresponding mean field approximation. Here we distinguished between global and local knowledge. Global knowledge is the

aggregated information from all communicating vehicles at a given time. A central node or server to which all vehicles send the available information might possess such global knowledge. Analogously, we refer to the knowledge of a single vehicle as local knowledge.

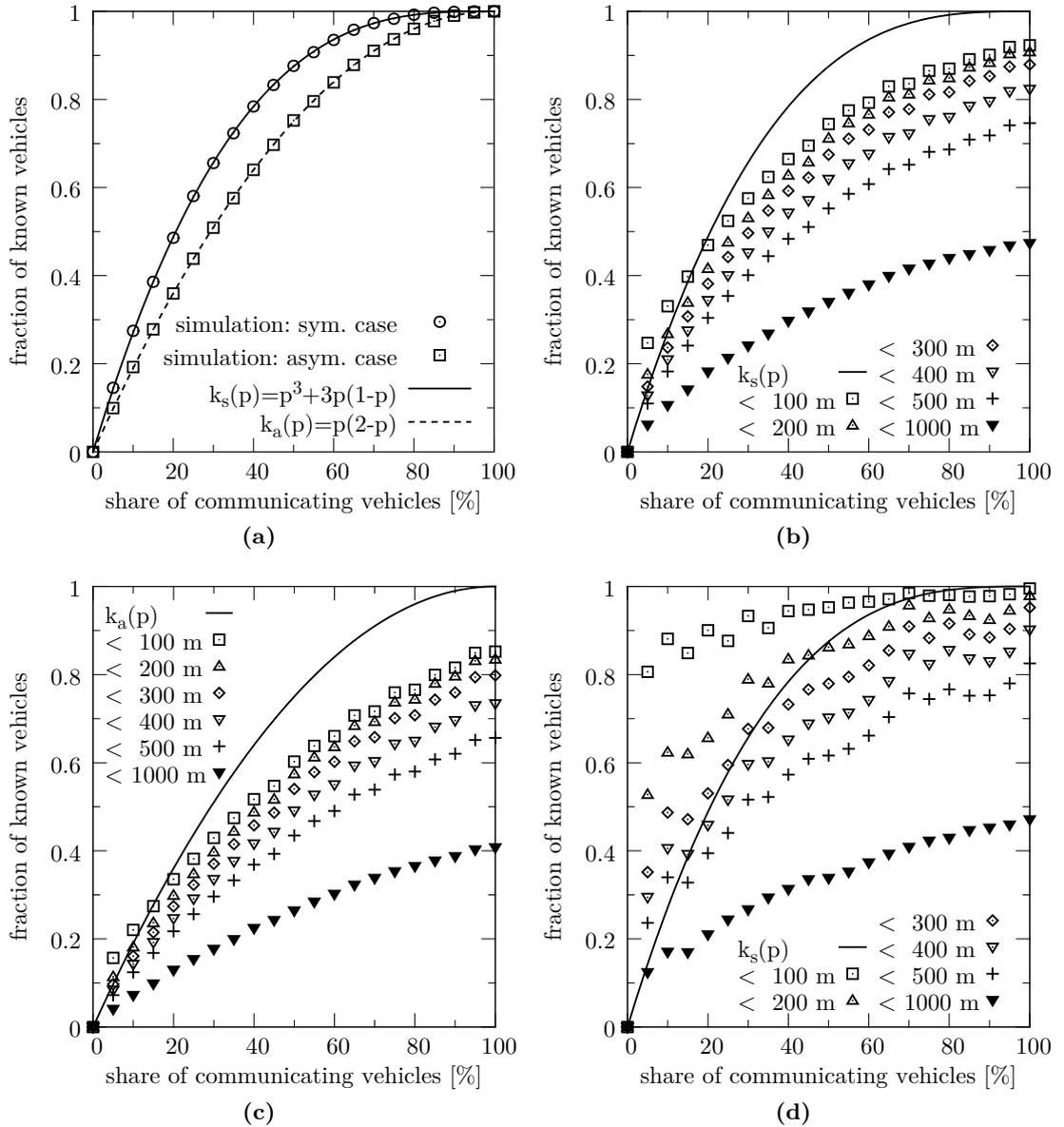
In figure 6.2a the percentage of known vehicles is shown for the entire system (i.e., the global knowledge). The functional relations (6.9) and (6.11) serve as references. The good agreement is not surprising as we assumed each vehicle is able to successfully transmit its knowledge to a central node.

A realistic treatment of communication between vehicles, however, is likely to deteriorate these results for two reasons. The communication range is limited, and, in the case of simultaneous transmissions, closer senders are preferred due to a stronger signal strength at the receiver. Hence, the fraction of known particles will decrease the larger is the distance between two vehicles. Figures 6.2b and 6.2c show a comparison between the average knowledge of single vehicles and the predictions by equations (6.9) and (6.11). We divided neighboring vehicles in different categories containing all vehicles closer than 100 m, 200 m, . . . , 500 m, and 1000 m, respectively. For each range we determined the average knowledge a communicating vehicle had in the symmetric (figure 6.2b) as well as in the asymmetric case (figure 6.2c). Even if all vehicles were communicating, the fraction of known vehicles stayed below 93 %, which results from the limited communication range and package collisions.

For low rates of communicating vehicles, the number of known vehicles was above the analytic prediction though. To understand this behavior, see figure 6.2d where we averaged only over a fixed vehicle density of  $\rho = 9\%$  (symmetric case). At low densities, there are only few vehicles within the neighborhood of communicating vehicles. From these vehicles, however, we assumed that the vehicle's sensors always detect the immediate predecessor and successor. Consequently, at low densities, where these two vehicles make up a large part of the neighboring vehicles, the fraction of known vehicles is large as well. This relatively large fraction at low vehicle densities increases the overall fraction of known vehicles when averaging over the entire density range, as seen in figure 6.2b.

## 6.4 Discussion

In this chapter we developed a model of a one-dimensional system with two types of particles, namely active and passive particles. By mapping the model to a one-dimensional random walk, we could derive formulas to determine the fraction of known particles within the system depending on the system size  $N$  and the number of active particles  $A$ . As expected, the functions are strictly increasing for  $A \rightarrow N$ . For low fractions of active particles (i.e.,  $p = A/N \ll 1$ ) the dominating term in equations (6.7) and (6.8) is  $3A/N$  for the symmetric case. This, in turn, means that each active particles discovers approximately two passive particles. For  $p = 0.1$  and  $N \rightarrow \infty$  the number of discovered passive particles per active particle is slightly above 1.7 as can be obtained from equation (6.9). Similarly, with an asymmetry in particle discovery the



**Figure 6.2** (a) Under the assumption that all communicating vehicles (with an overall share of  $p$ ) successfully transmit their knowledge to a central instance, this aggregated information is very well described by the functions  $k_a(p)$  (dashed line) and  $k_s(p)$  (solid line) derived in sections 6.2.1 and 6.2.2. The simulation results were averaged over various vehicle densities in the range  $5\% \leq \rho \leq 65\%$ . The average fraction of known vehicles from a single vehicle's perspective is given both for (b) the symmetric and (c) the asymmetric case. At small densities, the average knowledge may be above the theoretical prediction. This follows from the fact that there are only few vehicles in the neighborhood of a communicating one and that the immediate leader and follower are always assumed to be known. This can be seen from figure (d) which shows the fraction of known vehicles for a fixed density of  $\rho = 9\%$ .

term  $2A/N$  contributes most to the sums in equations (6.11) and (6.13) for  $A \ll N$ . In the latter case we could also show that a simple mean field approximation is exact for open boundaries. In all cases mean field approximations are in good agreement with the exact result for large  $N$  ( $N > 100$ ) as the error decreases with  $N^{-1}$ . Furthermore, in systems with open boundaries, the knowledge is lower than in systems with periodic boundaries as the active particles occupying the boundaries have only one neighbor to discover instead of two.

The comparison between the analytic results and our simulations of a vehicular communication network, which motivated our analysis, showed that realistic communication modeling decreases the fraction of known vehicles compared with the theoretical prediction. The analytic solution still can serve as an upper bound in most situations.

## Conclusions and Outlook

Vehicular traffic is not only an important example of transport processes. It is essential for modern, dynamic national economies, but it also causes severe problems (e.g., air pollution, congestion, accidents).

A first step to ease or to solve these problems is to understand traffic. Mathematical and computer models may contribute to a deeper understanding of traffic. For this purpose, it is, however, necessary that such models reproduce the empirical findings satisfyingly.

In this spirit, we studied the spatiotemporal dynamics of three distinct microscopic traffic models with a clear focus on the comfortable driving model (CDM). In chapter 4, we analyzed the spatiotemporal dynamics of the CDM with respect to the observed traffic phases and the transitions between these phases. We felt such an analysis was necessary as many newly presented traffic models claim to reproduce the synchronized phase of Kerner's three-phase traffic theory simply by looking at space–time–velocity plots (e.g., [48, 127]). Hence, we used the rule-based FOTO-method, which still has some limitations in assessing the spatiotemporal dynamics as it is based on local measurements. Yet it provides hard and objective criteria for this purpose. By applying the FOTO-method, we found that the obtained results are in both quantitative and qualitative agreement with the three-phase traffic theory. Notwithstanding this agreement, we do not say that the CDM reproduces the synchronized phase, but we say that we found three clearly distinguishable states or “phases”.

The comparison with empirical data showed that all investigated models could reproduce a traffic breakdown observed in real-world measurements. Taking into account the large number of vehicles involved, this is a very fascinating result! Despite their short-ranged interaction, which is limited to one or two vehicles ahead, these models succeed in reproducing the macroscopic features of traffic flow involving a much larger number of vehicles.

Yet we have to ask critically whether we do not oversimplify things. All models that we investigated in chapter 3 and most current microscopic models treat vehicles as elementary building blocks of the complex system “traffic”. But the building block

“vehicle” can be further subdivided into the vehicle with its well-known mechanical restrictions (e.g., braking capacity, limited acceleration rate) and the vehicle’s driver who may show different attitudes (e.g., timid, cautious, aggressive, or selfish) [13]. Although some of these aspect have already been incorporated in traffic models (e.g., optimistic versus pessimistic driving behavior in [91] or polite versus impolite, when changing lanes [71]), these attitudes are not really a characteristic of the individual driver but are assigned in response to external factors. Does this level of detail suffice?

The good agreement found in chapters 3 and 4 might suggest to answer this question with “yes” as the single-driver effects seem to average out. Indeed, with respect to the modeling of traffic flow, this attitude is justifiable. Studies comparing the car-following behavior of only a few vehicles with simulated data, however, found that, even after calibrating the microscopic traffic models, it is not possible to suppress clear deviations between the empirical and simulated data due to differences in the drivers’ behavior [17, 70, 118]. With respect to these results, we want to refer to a very recent article presenting a traffic model that treats “driver-vehicles subsystems as a living, hence complex, system” [13]. For a more general overview of the challenges in the modeling of traffic flow and crowd movements, where the “human factor” plays an even more important role, see the recent review article [12].

Notwithstanding the limitations just mentioned, the microscopic modeling approach is very well suited to describe traffic flow. Probably its biggest advantage is the fine-grained analysis it offers: starting from a single-vehicle basis, it allows to study which interactions on the single-vehicle level lead to macroscopic effects. Moreover, it enables us to modify the motion of selected vehicles as desired. We did this, for example, in chapter 5, where we studied the influence of vehicular communication. For communicating vehicles, a subset of all vehicles, we slightly modified the rules of motion and measured the resulting impact on traffic flow. We saw that few vehicles ( $\approx 10\%$ ) behaving differently can improve traffic flow a lot.

The heterogeneity of traffic flow was also subject of the previous chapter 6. On a rather abstract level we saw similarities between a sequence of vehicles, a 1D surface, and a 1D random walk. The obtained formulas for knowledge in heterogeneous traffic were derived under some simplified assumptions (e.g., unlimited sensor range). Yet we could show that the fraction of known vehicles is considerably higher than the fraction of sensor-equipped vehicles in heterogeneous traffic.

## Hypergeometric Series

In chapter 6 we have encountered sums of the form

$$\sum_{k=0}^n k \binom{n}{k} \binom{s}{t+k} \quad \text{with } n, s, t \in \mathbb{N} \quad (\text{A.1})$$

to determine the fraction of known particles (see, e.g., equation (6.6)). We will show how such expressions, which involve binomial coefficients and factorials, can be simplified by identifying them as hypergeometric series.

In general, a power series

$$\sum_{k \geq 0} c_k x^k \quad (\text{A.2})$$

is called hypergeometric if  $c_0 = 1$  and if the ratio of successive coefficients is a rational function of  $k$ , i.e.,

$$\frac{c_{k+1}}{c_k} = \frac{P(k)}{Q(k)} \quad (\text{A.3})$$

with polynomials  $P$  and  $Q$  in  $k$ . (The requirement  $c_0 = 1$  is not very rigid: provided that  $c_0 \neq 0$ , we can always achieve  $c_0 = 1$  by factoring out the leading term, as we will show.)

When working with hypergeometric series, it is often more convenient to convert the series into hypergeometric notation  ${}_pF_q$  [84]:

$${}_pF_q \left[ \begin{matrix} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{matrix} \middle| x \right] := \sum_{k=0}^{\infty} c_k x^k = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdot (a_2)_k \cdots (a_p)_k}{(b_1)_k \cdot (b_2)_k \cdots (b_q)_k \cdot k!} x^k \quad (\text{A.4})$$

where we introduced the Pochhammer symbol

$$(a)_k := a(a+1)(a+2) \cdots (a+k-1). \quad (\text{A.5})$$

By comparing the ratio of two successive terms of  ${}_pF_q$ , we find

$$\frac{c_{k+1}x^{k+1}}{c_kx^k} = \frac{(k+a_1)(k+a_2)\cdots(k+a_p)}{(k+b_1)(k+b_2)\cdots(k+b_q)(k+1)}x. \quad (\text{A.6})$$

From this representation, it directly follows how to convert a hypergeometric series to hypergeometric notation: after calculating  $c_{k+1}/c_k$  and factorizing the numerator and the denominator in terms of  $k$ , we can immediately read off the *upper parameters*  $a_1, a_2, \dots, a_p$  as well as the *lower parameters*  $b_1, b_2, \dots, b_q$ .

Before evaluating equation (A.1), let us first consider the simpler sum

$$\sum_{k=0}^n \binom{n}{k} \binom{s}{t+k} = \binom{s}{t} \sum_{k=0}^n \binom{n}{k} \binom{s}{t+k} / \binom{s}{t} \quad \text{with } n, s, t \in \mathbb{N}. \quad (\text{A.7})$$

In the last step we have factored out the term  $\binom{s}{t}$  so that the sum is a valid hypergeometric series with  $c_0 = 1$  and  $x = 1$ . Yet the factorization does not change ratio of successive coefficients. Calculating  $c_{k+1}/c_k$  yields

$$\frac{c_{k+1}}{c_k} = \frac{\binom{n}{k+1} \binom{s}{t+k+1}}{\binom{n}{k} \binom{s}{t+k}} = \frac{(k-n)(k-s+t)}{(k+1)(k+t+1)}. \quad (\text{A.8})$$

A comparison with equation (A.6) immediately leads to

$$\sum_{k=0}^n \binom{n}{k} \binom{s}{t+k} = \binom{s}{t} \times {}_2F_1 \left[ \begin{matrix} -n, t-s \\ t+1 \end{matrix} \middle| 1 \right]. \quad (\text{A.9})$$

Let us now consider equation (A.1). Obviously, due to  $c_0 = 0$ , it is not a valid hypergeometric series, and the quotient  $c_{k+1}/c_k$  is not defined for  $k = 0$ . We can, however, apply the same steps as before if we shift the summation index by one and consider

$$\sum_{k=0}^{n-1} (k+1) \binom{n}{k+1} \binom{s}{t+k+1} = n \binom{s}{t+1} \sum_{k=0}^{n-1} \frac{k+1}{n} \binom{n}{k+1} \binom{s}{t+k+1} / \binom{s}{t+1}.$$

Again, we calculate the ratio  $c_{k+1}/c_k$  for the previous sum, which gives

$$\frac{c_{k+1}}{c_k} = \frac{(k+2) \binom{n}{k+2} \binom{s}{t+k+2}}{(k+1) \binom{n}{k+1} \binom{s}{t+k+1}} = \frac{(k-n+1)(k-s+t+1)}{(k+1)(k+2+t)}. \quad (\text{A.10})$$

Finally, we obtain

$$\sum_{k=0}^n k \binom{n}{k} \binom{s}{t+k} = n \binom{s}{t+1} \times {}_2F_1 \left[ \begin{matrix} -n+1, t-s+1 \\ t+2 \end{matrix} \middle| 1 \right]. \quad (\text{A.11})$$

Using the identities (A.9) and (A.11), we are now able to simplify both series by applying

Gauss's (hypergeometric) theorem, which we give here without proof [5, pp. 2–3]:

$${}_2F_1 \left[ \begin{matrix} a, b \\ c \end{matrix} \middle| 1 \right] = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \quad \text{if } \operatorname{Re}(c-a-b) > 0. \quad (\text{A.12})$$

Hence, equation (A.1) reduces to a simple rational function of  $s$ ,  $t$ , and  $n$ :

$$\sum_{k=0}^n k \binom{n}{k} \binom{s}{t+k} = n \frac{(n+s-1)!}{(s-t-1)!(n+t)!}, \quad (\text{A.13})$$

where we used that for a positive integer  $m \in \mathbb{N}$  the Gamma function  $\Gamma(m+1)$  reduces to a simple factorial (i.e.,  $\Gamma(m+1) = m!$ ).

This knowledge can now be applied to evaluate the sums of chapter 6. Equation (6.6), for example, contains the sum

$$\sum_R \sum_{x,y \in \{\circ, \bullet\}} R \times \Phi_{xy}(R). \quad (\text{A.14})$$

We now take  $x = y = \circ$ , and we remind that

$$\Phi_{\circ\circ}(R) = \binom{A-1}{\frac{1}{2}R-1} \binom{P-1}{\frac{1}{2}R} \quad (\text{A.15})$$

with  $R$  even and  $2 \leq R \leq \min(2P-2, 2A)$ , i.e.,  $1 \leq R/2 \leq \min(P-1, A)$ . Consequently, for  $x = y = \circ$ , equation (A.14) reduces to

$$\sum_{k=0}^{\min(P-1, A)} 2k \binom{A-1}{k-1} \binom{P-1}{k} \quad (\text{A.16})$$

which can be rewritten with equation (A.13) as

$$\sum_{k=1}^{\min(P-1, A)} 2k \binom{A-1}{k-1} \binom{P-1}{k} = (P-1) \frac{(P+A-3)!}{(A-1)!(P-2)!} \quad (\text{A.17})$$

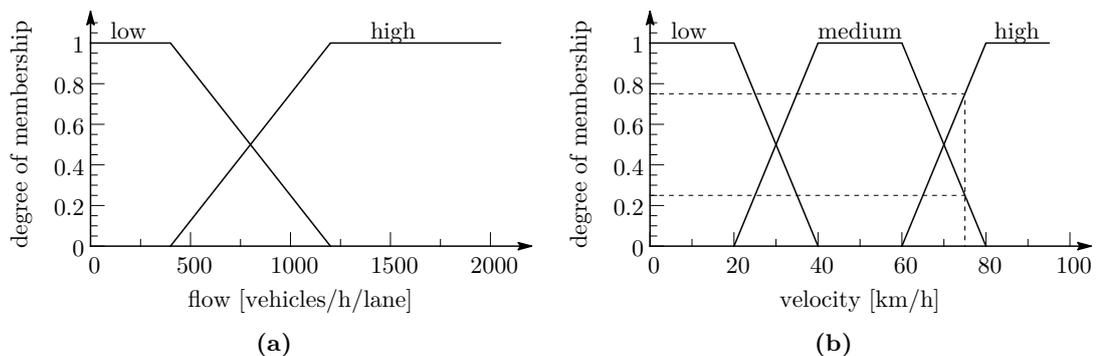
in either case (i.e.,  $P-1 \leq A$  and  $P-1 > A$ ). With a similar reasoning the remaining terms of equation (A.14) can be evaluated, altogether leading to:

$$\sum_R \sum_{x,y \in \{\circ, \bullet\}} R \times \Phi_{xy}(R) = \frac{2(P+A-1)!}{(P-1)!(A-1)!}. \quad (\text{A.18})$$



## The FOTO-method

For the sake of completeness, we review the FOTO-method, which was developed by Kerner and colleagues [58, 69] and which we used in chapter 4. The method allows a classification of locally measured traffic data according to the three-phase traffic theory. The classification is based on a set of rules. It uses the aggregated data provided by a local detector (i.e., velocity and flow) and decides to which traffic state the measured combination of observables most likely belongs. As the traffic phases S and J both denote phases of congested traffic, the distinction is not always obvious. Therefore, the set of rules employs a fuzzification of the input parameters as shown in figure B.1.



**Figure B.1** Illustration of the fuzzification process. For each value of (a) flow and (b) velocity one can read off the associated degree of membership to the classes “low”, “high”, and, in the case of the velocity, “medium”. The dashed lines in (b) illustrate that one value (75 km/h) can be member of more than one class.

The fuzzification process transforms the measured value into fuzzy values which denote the degree of membership to the given classes. Hence, one value may belong to more than one class. This is illustrated in figure B.1b, where fuzzifying the velocity of 75 km/h shows that this velocity is both a “medium” and a “high” velocity. Yet the degree

of membership to the class “high” is larger (0.75) than to the class “medium” (0.25).

Based on this fuzzification of flow and velocity, the classification of traffic states via the FOTO-method works as follows [56, 69]:

- F1** If the measured average velocity is classified as “high”, then the associated traffic phase is free flow (F)—independent of the flow rate’s value.
- F2** If the measured average velocity is a member of the class “medium”, the associated traffic phase corresponds to synchronized traffic (S).
- F3** If the measured average velocity is “low” but the flow is “high”, the associated traffic phase is synchronized traffic (S).
- F4** If both the measured average velocity and the flow are “low”, the associated traffic phase should be classified as a wide moving jam (J).

Due to non-exclusive memberships, a pair of flow and average velocity may match more than one of the above criteria. In this case, one has to chose the traffic state with the highest degree of membership of both velocity and flow. To illustrate this classification, table B.1 gives several examples of how the FOTO-method works.

**Table B.1** Examples of how the FOTO-method processes the input variables velocity and flow. For both observables the degree of membership to the classes “low”, “high”, and, for the velocity, also to the class “medium” is determined. Then, the agreement with the criteria F1–F4 is checked. The largest value for F1–F4 determines the choice of the traffic phase (gray-shaded cells). To distinguish the criteria F2 and F3, the table contains two separate columns labeled  $S_2$  and  $S_3$ , which refer to F2 and F3, respectively.

$\langle v \rangle$	$\langle J \rangle$	velocity			flow		traffic phase			
		low	medium	high	low	high	J	$S_2$	$S_3$	F
80	1260	0	0	1	0	1	0	0	0	1
71	1290	0	0.45	0.55	0	1	0	0.45	0	0.55
27	900	0.65	0.35	0	0.375	0.625	0.375	0.35	0.625	0
66	1230	0	0.7	0.3	0	1	0	0.7	0	0.3
43	1050	0	1	0	0.1875	0.8125	0	1	0	0
13	540	1	0	0	0.825	0.175	0.825	0	0.175	0
25	630	0.75	0.25	0	0.7125	0.2875	0.7125	0.25	0.2875	0

The agreement with criteria F3 and F4, which requires evaluating flow and velocity simultaneously, is checked in a two-step process. First, for both the average velocity and the average flow the corresponding degrees of membership are determined. Then, the minimum degree of membership of either of the two variables determines the extent to which the criterion is met. This can, for example, be seen from the last row of table B.1: the criterion F4 requires both velocity and flow to be “low”. The degree of membership

to the class “low” is 0.75 for the velocity but only 0.7125 for the flow. Consequently, the degree of matching the criterion F4 is  $\min(0.75, 0.7125) = 0.7125$ . Finally, the traffic phase for which the best matching was observed is chosen.

Kerner et al. [69] also presented an extended version of the FOTO-method comprising 13 different rules instead of the four F1–F4. As the extended rule set provides for a better discrimination between synchronized flow and wide moving jams only in very rare situations and as the mechanism is not different to the one presented here, we refer to the corresponding sections in [69] or [56].





## CDM with Extensions for VANETs

Although the CDM has already been presented in section 3.1.3, we provide here a full review of the CDM that includes the extensions for traffic flow optimization via vehicular communication (see sections 5.3.1 and 5.3.2). According to these extensions, communicating vehicles are to keep a larger gap when they are notified about a critical road segment (i.e., if  $j_{\text{bool}} = 1 = \text{true}$ ). The additional gap is denoted as  $g_n(t)$ . By setting  $g_n(t) = 0$  below, the original rules of motion of the CDM are restored.

Remember that the CDM is a traffic cellular automaton where space is discretized in units of 1.5 m and time in intervals of 1 s. The position and the velocity of a vehicle labeled as  $n$  are given by  $x_n(t)$  and  $v_n(t)$ , respectively (see figure 3.2). Vehicles are labeled downstream in ascending order such that the vehicle in front of  $n$  is labeled  $n + 1$ . For convenience, the additional variables  $d_n(t)$  and  $t_h(t) = d_n(t)/v_n(t)$  are introduced to denote the spatial and temporal headway between the vehicles  $n$  and  $n + 1$ . The model includes anticipatory effects by considering the status of the preceding vehicle's brake light  $b_{n+1}(t)$ , by anticipating its velocity  $v_{\text{anti}}(t) = \min(v_{n+1}(t), d_{n+1}(t))$ , and by calculating an effective distance  $d_n^{\text{eff}}(t)$  as

$$d_n^{\text{eff}}(t) = d_n(t) + \max(v_{\text{anti}}(t) - d_{\text{safe}}, 0)$$

where  $d_{\text{safe}}$  governs the effectiveness of the anticipation (usually  $d_{\text{safe}} = 7$ ).

The rules of motion in our modified model read as follows (the temporal discretization of 1 s allows to treat all magnitudes as dimensionless because conversions do not change a variable's numerical value):

### 1. Acceleration:

$$v_n(t+1) \leftarrow \begin{cases} \min(v_n^{\text{max}}, v_n(t) + 1), & \text{if } b_n(t) = b_{n+1}(t) = 0 \text{ or } t_h(t) \geq t_s(t), \\ v_n(t), & \text{otherwise.} \end{cases} \quad (\text{C.1})$$

## 2. Adaptation of $g_n$ :

$$v_n(t+1) \leftarrow \begin{cases} v_n(t), & \text{if } v_n(t+1) > (d_n^{\text{eff}}(t) - g_n(t)) \\ & \text{and } g_n(t) > 0 \text{ and } v_n(t+1) > v_n(t), \\ v_n(t+1), & \text{otherwise.} \end{cases} \quad (\text{C.2})$$

$$g_n(t+1) \leftarrow \begin{cases} \max(d_n^{\text{eff}}(t) - v_n(t+1), 0), & \text{if } v_n(t+1) > (d_n^{\text{eff}}(t) - g_n(t)) \\ & \text{and } g_n(t) > 0, \\ g_n(t) & \text{otherwise.} \end{cases} \quad (\text{C.3})$$

## 3. Braking:

$$v_n(t+1) \leftarrow \min(d_n^{\text{eff}}(t) - g_n(t+1), v_n(t+1)) \quad (\text{C.4})$$

$$b_n(t+1) \leftarrow 1 - \Theta(v_n(t+1) - v_n(t)) \quad (\text{C.5})$$

## 4. Determination of randomization parameter $p$ :

$$p \leftarrow \begin{cases} p_b, & \text{if } b_{n+1}(t) = 1 \text{ and } t_h(t) < t_s(t) \text{ and } g_n(t+1) \leq v_n(t)t_{\text{reaction}}, \\ p_c, & \text{if } b_{n+1}(t) = 1 \text{ and } t_h(t) < t_s(t) \text{ and } g_n(t+1) > v_n(t)t_{\text{reaction}}, \\ p_0, & \text{if } v_n(t) = 0, \\ p_d, & \text{otherwise.} \end{cases} \quad (\text{C.6})$$

## 5. Dawdling:

$$v_n(t+1) \leftarrow \begin{cases} \max(v_n(t+1) - 1, 0), & \text{if } \xi < p, \\ v_n(t+1), & \text{otherwise.} \end{cases} \quad (\text{C.7})$$

$$b_n(t+1) \leftarrow \begin{cases} 1, & \text{if } \xi < p \text{ and } p = p_b, \\ b_n(t+1), & \text{otherwise.} \end{cases} \quad (\text{C.8})$$

## 6. Reaction to warning message:

$$g_n(t+1) \leftarrow \begin{cases} \min(7l_{\text{car}}, d_n(t) - v_n^{\text{max}}), & \text{if } j_{\text{bool},n}(t) = 1 \text{ and } d_n(t) > v_n^{\text{max}}, \\ g_n(t+1), & \text{if } j_{\text{bool},n}(t) = 1 \text{ and } d_n(t) \leq v_n^{\text{max}}, \\ 0, & \text{if } j_{\text{bool},n}(t) = 0. \end{cases} \quad (\text{C.9})$$

## 7. Vehicle motion:

$$x_n(t+1) \leftarrow x_n(t) + v_n(t+1) \quad (\text{C.10})$$

In the first step, a vehicle tries to accelerate to its maximum velocity. To avoid unnecessary acceleration, it checks the status of its own and the preceding vehicle's

---

brake light and compares its time headway  $t_h(t)$  with a velocity-dependent interaction horizon  $t_s(t) = \min(v_n(t), h)$ .

In the next step, the vehicle's velocity and the additional gap  $g_n(t)$  are adapted to avoid braking. If maintaining the current value of  $g_n(t)$  requires braking (first case in equation (C.2)), any acceleration of the previous step is undone. Then, the new value of  $g_n$  is calculated so that the vehicle does not have to brake or, if this is not possible, so that the deceleration is minimized.

Afterwards, the vehicle checks the safety constraints and brakes, if necessary. Subsequently, it updates the status of its brake light.

Then, the determination of the randomization parameter  $p$  follows (see section 5.3.2). The function  $\Theta(\cdot)$  denotes the Heaviside step function. If a (pseudo-)random number  $\xi$ , uniformly generated in  $[0, 1]$ , is smaller than the randomization parameter  $p$ , the velocity is reduced by one unit. The last two steps, where the new value of  $g_n$  is calculated (step 6) and the vehicle moves (step 7), are interchangeable. The length of  $g_n$  is limited to a multiple of a passenger car's length  $l_{\text{car}}$  to avoid unrealistically large gaps in dense traffic.



## References

1. Alpher, R. A. and R. Herman. “[Evolution of the Universe](#)”. *Nature* 162 (1948): 774–775 (see p. 1).
2. Antal, T. and G. M. Schütz. “[Asymmetric exclusion process with next-nearest-neighbor interaction: Some comments on traffic flow and a nonequilibrium reentrance transition](#)”. *Phys. Rev. E* 62 (1 2000): 83–93 (see pp. 12, 31).
3. Appert-Rolland, C. “[Experimental study of short-range interactions in vehicular traffic](#)”. *Phys. Rev. E* 80 (3 2009): 036102 (see pp. 5, 31).
4. Artimy, M. “[Local Density Estimation and Dynamic Transmission-Range Assignment in Vehicular Ad Hoc Networks](#)”. *IEEE Trans. Intell. Transp. Syst.* 8, no. 3 (2007): 400–412 (see pp. 63, 64).
5. Bailey, W. N. “Generalized Hypergeometric Series”. Cambridge tracts in mathematics and mathematical physics. New York: Stechert-Hafner service agency, 1964 (see p. 93).
6. Bak, P. “How nature works: the science of self-organized criticality”. New York: Copernicus, 1996 (see p. 2).
7. Barlovic, R., L. Santen, A. Schadschneider, and M. Schreckenberg. “[Metastable states in cellular automata for traffic flow](#)”. *Eur. Phys. J. B* 5, no. 3 (1998): 793–800 (see p. 42).
8. Barlovic, R., T. Huisinga, A. Schadschneider, and M. Schreckenberg. “[Open boundaries in a cellular automaton model for traffic flow with metastable states](#)”. *Phys. Rev. E* 66 (4 2002): 046113 (see pp. 24, 39, 42, 55).
9. Barr, R. “An efficient, unifying approach to simulation using virtual machines”. PhD thesis. 2004 (see p. 62).
10. Barr, R., Z. J. Haas, and R. van Renesse. “[JiST: An efficient approach to simulation using virtual machines](#)”. *Softw. Pract. Exper.* 35, no. 6 (2005): 539–576 (see p. 62).

11. Barr, R., Z. J. Haas, and R. van Renesse. “Scalable Wireless Ad hoc Network Simulation”. In: *Handbook on theoretical and algorithmic aspects of sensor, ad hoc wireless, and peer-to-peer networks*. Ed. by Wu, J. CRC Press, 2005. Chap. 19, 297–311 (see p. 62).
12. Bellomo, N. and C. Dogbe. “On the Modeling of Traffic and Crowds: A Survey of Models, Speculations, and Perspectives”. *SIAM Rev.* 53, no. 3 (2011): 409–463 (see pp. 15, 37, 90).
13. Belloquid, A., E. De Angelis, and L. Fermo. “Towards the modeling of vehicular traffic as a complex system: a kinetic theory approach”. *Math. Models Methods Appl. Sci.* 22 (2012): 1140003 (see p. 90).
14. Boban, M., T. T. V. Vinhoza, J. Barros, M. Ferreira, and O. K. Tonguz. “Impact of Vehicles as Obstacles in Vehicular Ad Hoc Networks”. *IEEE J. Sel. Areas Commun.* 29, no. 1 (2011): 15–28 (see pp. 60, 67).
15. Brilon, W., J. Geistefeld, and M. Regler. “Reliability of freeway traffic flow: a stochastic concept of capacity”. In: *Proc. 16th Int. Symp. Transp. & Traffic Theory*. Maryland, 2005 (see pp. 7, 11).
16. Brockfeld, E., R. Kühne, A. Skabardonis, and P. Wagner. “Toward Benchmarking of Microscopic Traffic Flow Models”. *Transp. Res. Rec.* 1852, no. 1 (2003): 124–129 (see pp. 15, 16, 26).
17. Brockfeld, E. and P. Wagner. “Calibration and Validation of Microscopic Traffic Flow Models”. In: *Traffic and Granular Flow '03*. Ed. by Hoogendoorn, S. P., S. Luding, P. H. L. Bovy, M. Schreckenberg, and D. E. Wolf. Berlin: Springer, 2005, 67–72 (see pp. 23, 26, 90).
18. Bundesministerium für Verkehr, Bau und Stadtentwicklung. “Technische Lieferbedingungen für Streckenstationen: TLS 2002”. Bremerhaven: Wirtschaftsverl. NW, Verlag für Neue Wiss., 2002 (see p. 22).
19. Car 2 car communication consortium. 2012. URL: <http://www.car-to-car.org/> (see p. 79).
20. Chowdhury, D., L. Santen, and A. Schadschneider. “Statistical physics of vehicular traffic and some related systems”. *Phys. Rep.* 329, no. 4–6 (2000): 199–329 (see pp. 15, 20).
21. Coifman, B. *Time Space Diagrams For Thirteen Shock Waves*. Tech. rep. Inst. Transp. Studies, UC Berkeley, 1997 (see p. 2).
22. Daganzo, C. F., M. J. Cassidy, and R. L. Bertini. “Possible explanations of phase transitions in highway traffic”. *Transp. Res. A* 33, no. 5 (1999): 365–379 (see p. 10).
23. Derrida, B. and M. R. Evans. “The asymmetric exclusion model: exact results through a matrix approach”. In: *Nonequilibrium Statistical Mechanics in One Dimension*. Ed. by Privman, V. Cambridge: Cambridge University Press, 1997. Chap. 14, 277–304 (see pp. 11, 79).
24. E 2213-03. *Standard Specification for Telecommunications and Information Exchange Between Roadside and Vehicle Systems — 5 GHz Band Dedicated Short Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. 2003 (see pp. 67, 68).

25. Evans, M. R. “Phase transitions in one-dimensional nonequilibrium systems”. *Braz. J. Phys.* 30, no. 1 (2000): 42–57 (see p. 79).
26. Fall, K. and K. Varadhan. *The ns manual*. 2010. URL: <http://www.isi.edu/nsnam/ns/> (see pp. 59, 62).
27. Fekete, S. P., B. Hendriks, C. Tessars, A. Wegener, H. Hellbrück, S. Fischer, and S. Ebers. “Methods for Improving the Flow of Traffic”. In: *Organic Computing – A Paradigm Shift for Complex Systems*. Ed. by Müller-Schloer, C., H. Schmeck, and T. Ungerer. Vol. 1. Basel: Springer, 2011, 447–460 (see p. 63).
28. Feynman, R. P. and A. R. Hibbs. “Quantum Mechanics and Path Integrals”. New York: McGraw-Hill, 1965, 34–36 (see p. 82).
29. Friis, H. T. “A Note on a Simple Transmission Formula”. *Proc. IRE* 34, no. 5 (18, 1946): 254–256 (see p. 58).
30. Gazis, D. C. “The Origins of Traffic Theory”. *Oper. Res.* 50, no. 1 (2002): 69–77 (see p. 2).
31. Gazis, D. C. and R. S. Foote. “Surveillance and Control of Tunnel Traffic by an On-line Digital Computer”. *Transport. Sci.* 5, no. 3 (1969): 255–275 (see p. 2).
32. Giordano, E., R. Frank, G. Pau, and M. Gerla. “CORNER: a step towards realistic simulations for VANET”. In: *Proc. 7th ACM int. workshop on Vehicular InterNETworking. VANET '10*. New York: ACM, 2010: 41–50 (see p. 60).
33. Goldin, D. and P. Kanellakis. “On similarity queries for time-series data: Constraint specification and implementation”. In: *Principles and Practice of Constraint Programming — CP '95*. Ed. by Montanari, U. and F. Rossi. Vol. 976. Lecture Notes in Computer Science. Springer, 1995, 137–153 (see p. 24).
34. Gupta, P. and P. R. Kumar. “The capacity of wireless networks”. *IEEE Trans. Inf. Theory* 46, no. 2 (2000): 388–404 (see p. 61).
35. Gurusinghe, G. S., T. Nakatsuji, Y. Azuta, P. Ranjitkar, and Y. Tanaboriboon. “Multiple Car-Following Data with Real-Time Kinematic Global Positioning System”. *Transp. Res. Rec.* 1802, no. 1 (2002): 166–180 (see p. 5).
36. Hafstein, S., R. Chrobok, A. Pottmeier, M. Schreckenberger, and F. Mazur. “A High-Resolution Cellular Automata Traffic Simulation Model with Application in a Freeway Traffic Information System”. *Comput.-aided civ. infrastruct. eng.* 19, no. 5 (2004): 338–350 (see p. 34).
37. Hager, J. S., J. Krug, V. Popkov, and G. M. Schütz. “Minimal current phase and universal boundary layers in driven diffusive systems”. *Phys. Rev. E* 63 (5 2001): 056110 (see pp. 11, 31).
38. Hall, F. L. “Traffic stream characteristics”. In: *Traffic flow theory: A state of the art report – revised monograph on traffic flow theory*. Ed. by Gartner, N. H., C. Messer, and A. K. Rathi. 1997. Chap. 2 (see p. 6).
39. Hall, F. L., B. L. Allen, and M. A. Gunter. “Empirical analysis of freeway flow-density relationships”. *Transp. Res. A* 20, no. 3 (1986): 197–210 (see p. 7).
40. Hall, F., S. Wakefield, and A. Al-Kaisy. “Freeway Quality of Service – What Really Matters to Drivers and Passengers?” *Transp. Res. Rec.* 1776 (2001): 17–23 (see p. 69).

41. Helbing, D. “Traffic and related self-driven many-particle systems”. *Rev. Mod. Phys.* 73, no. 4 (2001): 1067–1141 (see pp. 6, 15).
42. Helbing, D., M. Treiber, A. Kesting, and M. Schönhof. “Theoretical vs. empirical classification and prediction of congested traffic states”. *Eur. Phys. J. B* 69, no. 4 (2009): 583–598 (see p. 10).
43. Hugemann, W. “Driver Reaction Times in Road Traffic”. In: *11. EVU-Jahrestagung*. Portoroz, 2002 (see p. 66).
44. Ibrahim, K. and M. C. Weigle. “ASH: Application-aware SWANS with Highway mobility (poster)”. In: *Proc. IEEE INFOCOM Workshop on mob. netw. for veh. environ.* 2008 (see p. 62).
45. Iwasaki, M. “Empirical analysis of congested traffic flow characteristics and free speed affected by geometric factors on an intercity expressway”. *Transp. Res. Rec.* no. 1320 (1991): 242–250 (see p. 13).
46. Jacobson, T. and L. S. Schulman. “Quantum stochasticity: the passage from a relativistic to a non-relativistic path integral”. *J. Phys. A* 17 (1984): 375–383 (see p. 82).
47. Jiang, R. and Q.-S. Wu. “Cellular automata models for synchronized traffic flow”. *J. Phys. A* 36, no. 2 (2003): 381 (see pp. 42, 54).
48. Jiang, R. and Q.-S. Wu. “Spatial-temporal patterns at an isolated on-ramp in a new cellular automata model based on three-phase traffic theory”. *J. Phys. A* 37, no. 34 (2004): 8197 (see pp. 42, 54, 89).
49. Kåredal, J., N. Czink, A. Paier, F. Tufvesson, and A. Molisch. “Path loss modeling for vehicle-to-vehicle communications”. *IEEE Trans. Veh. Technol.* 60, no. 1 (2011): 323–328 (see p. 60).
50. Kerner, B. S. “Complexity of spatiotemporal traffic phenomena in flow of identical drivers: Explanation based on fundamental hypothesis of three-phase theory”. *Phys. Rev. E* 85 (3 2012): 036110 (see p. 66).
51. Kerner, B. S. “Empirical Features of Congested Patterns at Highway Bottlenecks”. *Transp. Res. Rec.* 1802 (2002): 145–154 (see p. 10).
52. Kerner, B. S. “Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks”. *Phys. Rev. E* 65 (4 2002): 046138 (see pp. 13, 45).
53. Kerner, B. S. “Experimental Features of Self-Organization in Traffic Flow”. *Phys. Rev. Lett.* 81, no. 17 (1998): 3797–3800 (see p. 13).
54. Kerner, B. S. “Introduction to Modern Traffic Flow Theory and Control: The Long Road to Three-Phase Traffic Theory”. Berlin: Springer, 2009 (see pp. 5, 11–13, 38, 46, 47, 49, 66, 79).
55. Kerner, B. S. “Modern approaches to basic traffic modeling: Three-Phase Traffic Theory”. *Transp. Res. Circular E-C* 149 (2011): 22–42 (see p. 45).
56. Kerner, B. S. “The Physics of Traffic”. Berlin: Springer, 2004 (see pp. 2, 5, 6, 12, 42, 46, 49, 96, 97).
57. Kerner, B. S. “Theory of Breakdown Phenomenon at Highway Bottlenecks”. *Transp. Res. Rec.* 1710 (2002): 136–144 (see p. 10).

58. Kerner, B. S., M. Aleksic, and U. Denneker. “Verfahren und Vorrichtung zur Verkehrszustandsüberwachung”. Patent DE19944077C2. Feb. 7, 2002 (see pp. 42, 95).
59. Kerner, B. S., L. Klenov Sergey, and A. Hiller. “Empirical test of a microscopic three-phase traffic theory”. *Nonlinear Dyn.* 49, no. 4 (2007): 525–553 (see p. 16).
60. Kerner, B. S. and S. L. Klenov. “Microscopic theory of spatial-temporal congested traffic patterns at highway bottlenecks”. *Phys. Rev. E* 68 (3 2003): 036130 (see p. 13).
61. Kerner, B. S. and S. L. Klenov. “Phase transitions in traffic flow on multilane roads”. *Phys. Rev. E* 80 (5 2009): 056101 (see pp. 16, 46).
62. Kerner, B. S. and S. L. Klenov. “Spatial-temporal patterns in heterogeneous traffic flow with a variety of driver behavioural characteristics and vehicle parameters”. *J. Phys. A* 37, no. 37 (2004): 8753 (see p. 2).
63. Kerner, B. S., S. L. Klenov, and A. Brakemeier. “Enhancing traffic efficiency and safety through the use of wireless vehicle communication”. *Traf. Eng. & Cont.* 51, no. 6 (2010): 217–222 (see p. 79).
64. Kerner, B. S., S. L. Klenov, and A. Brakemeier. “Testbed for wireless vehicle communication: a simulation approach based on three-phase traffic theory”. In: *IEEE Intell. Veh. Symp.* 2008: 180–185 (see pp. 63, 79).
65. Kerner, B. S., S. L. Klenov, A. Hiller, and H. Rehborn. “Microscopic features of moving traffic jams”. *Phys. Rev. E* 73 (4 2006): 046107 (see pp. 39, 47).
66. Kerner, B. S., S. L. Klenov, and D. E. Wolf. “Cellular automata approach to three-phase traffic theory”. *J. Phys. A* 35, no. 47 (2002): 9971 (see pp. 42, 54, 55).
67. Kerner, B. S. and H. Rehborn. “Experimental features and characteristics of traffic jams”. *Phys. Rev. E* 53 (2 1996): R1297–R1300 (see p. 8).
68. Kerner, B. S. and H. Rehborn. “Experimental Properties of Phase Transitions in Traffic Flow”. *Phys. Rev. Lett.* 79 (20 1997): 4030–4033 (see p. 10).
69. Kerner, B. S., H. Rehborn, M. Aleksic, and A. Haug. “Recognition and tracking of spatial-temporal congested traffic patterns on freeways”. *Transp. Res. C* 12, no. 5 (2004): 369–400 (see pp. 3, 7, 42, 44, 46, 49, 95–97, 112, 113).
70. Kesting, A. and M. Treiber. “Calibrating Car-Following Models by Using Trajectory Data: Methodological Study”. *Transp. Res. Rec.* 2088 (2008): 148–156 (see pp. 26, 90).
71. Kesting, A., M. Treiber, and D. Helbing. “General Lane-Changing Model MOBIL for Car-Following Models”. *Transp. Res. Rec.* 1999 (2007): 86–94 (see pp. 18, 19, 25, 90).
72. Kliot, G. *Technion Extensions of the JiST/SWANS Simulator*. <http://www.cs.technion.ac.il/~gabik/Jist-Swans/>. Sept. 2010 (see p. 62).
73. Knoop, V., S. P. Hoogendoorn, and S. P. van Zuylen. “Empirical Differences Between Time Mean Speed and Space Mean Speed”. In: [TGF2007]: 351–356 (see p. 6).

79. Knospe, W., L. Santen, A. Schadschneider, and M. Schreckenberg. “[A realistic two-lane traffic model for highway traffic](#)”. *J. Phys. A* 35, no. 15 (2002): 3369–3388 (see pp. 17, 21, 35).
80. Knospe, W., L. Santen, A. Schadschneider, and M. Schreckenberg. “[Disorder effects in cellular automata for two-lane traffic](#)”. *Physica A* 265 (1999): 614–633 (see p. 20).
81. Knospe, W., L. Santen, A. Schadschneider, and M. Schreckenberg. “[Empirical test for cellular automaton models of traffic flow](#)”. *Phys. Rev. E* 70 (1 2004): 016115 (see pp. 16, 21, 38, 81).
82. Knospe, W., L. Santen, A. Schadschneider, and M. Schreckenberg. “[Single-vehicle data of highway traffic: Microscopic description of traffic phases](#)”. *Phys. Rev. E* 65 (5 2002): 056133 (see pp. 5, 38).
83. Knospe, W., L. Santen, A. Schadschneider, and M. Schreckenberg. “[Towards a realistic microscopic description of highway traffic](#)”. *J. Phys. A* 33, no. 48 (2000): L477 (see pp. 17, 21, 22, 35).
84. Koepf, W. “Hypergeometric summation: an algorithmic approach to summation and special function identities”. Advanced lectures in mathematics. Wiesbaden: Vieweg, 1998 (see p. 91).
85. Kolomeisky, A. B., G. M. Schütz, E. B. Kolomeisky, and J. P. Straley. “[Phase diagram of one-dimensional driven lattice gases with open boundaries](#)”. *J. Phys. A* 31, no. 33 (1998): 6911 (see p. 11).
86. Koshi, M., M. Iwasaki, and I. Ohkura. “Some findings and an overview on vehicular flow characteristics”. In: *Proc. 8th Int. Symp. Transp. & Traffic Theory*. Ed. by Hurdle, V. F., E. Hauer, and G. Stewart. Toronto: Toronto University Press, 1981: 403–426 (see pp. 7, 13, 81).
87. Krug, J. “[Boundary-induced phase transitions in driven diffusive systems](#)”. *Phys. Rev. Lett.* 67, no. 14 (1991): 1882–1885 (see pp. 2, 12).
88. Kunisch, J. and J. Pamp. “[Wideband Car-to-Car Radio Channel Measurements and Model at 5.9 GHz](#)”. In: *IEEE 68th Conf. Veh. Technol.* 2008: 1–5 (see p. 60).
89. Lakas, A. and M. Chaqfeh. “[A novel method for reducing road traffic congestion using vehicular communication](#)”. In: *IWCMC*. 2010: 16–20 (see p. 63).
90. Lee, H. K. and B. J. Kim. “[Dissolution of traffic jam via additional local interactions](#)”. *Physica A* 390, no. 23–24 (2011): 4555–4561 (see pp. 75, 79).
91. Lee, H. K., R. Barlovic, M. Schreckenberg, and D. Kim. “[Mechanical Restriction versus Human Overreaction Triggering Congested Traffic States](#)”. *Phys. Rev. Lett.* 92 (23 2004): 238702 (see pp. 52, 90).
92. Lighthill, M. J. and G. B. Whitham. “[On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads](#)”. *Proc. Royal Soc. A* 229 (1955): 317–345 (see p. 2).
93. Madsen, H. “Time Series Analysis”. Boca Raton: Chapman and Hall/CRC, 2007. Chap. 3 (see p. 24).
94. Maerivoet, S. and B. De Moor. “[Cellular automata models of road traffic](#)”. *Phys. Rep.* 419, no. 1 (2005): 1–64 (see pp. 15, 16).

95. Mallick, K. “Some recent developments in non-equilibrium statistical physics”. *Pramana* 73, no. 3 (2009): 417–451 (see p. 11).
96. Matheus, K., R. Morich, I. Paulus, C. Menig, A. Lübke, B. Rech, and W. Specks. “Car-to-Car Communication – Market Introduction and Success Factors”. In: *5th Eur. Congr. and Exhib. on Intell. Transp. Syst. and Serv.* 2005 (see p. 63).
97. May, A. D. “Traffic flow fundamentals”. Englewood Cliffs: Prentice-Hall, 1990 (see pp. 6, 66).
98. Meireles, R., M. Boban, P. Steenkiste, O. Tonguz, and J. Barros. “Experimental Study on the Impact of Vehicular Obstructions in VANETs”. In: *2nd IEEE Veh. Networking Conf. (VNC 2010)*. Jersey City, NJ: IEEE, 2010: 338–345 (see p. 67).
99. Mittag, J., F. Thomas, J. Härri, and H. Hartenstein. “A comparison of single- and multi-hop beaconing in VANETs”. In: *Proc. 6th ACM VANET’09*. New York: ACM, 2009: 69–78 (see p. 64).
100. Nagatani, T. “The physics of traffic jams”. *Rep. Prog. Phys.* 65, no. 9 (2002): 1331–1386 (see p. 15).
101. Nagel, K. and M. Paczuski. “Emergent traffic jams”. *Phys. Rev. E* 51 (4 1995): 2909–2918 (see p. 2).
102. Nagel, K. and M. Schreckenberg. “A cellular automaton model for freeway traffic”. *J. Phys. I* 2, no. 12 (1992): 2221–2229 (see pp. 10, 16, 17, 19, 20, 42, 55, 80).
103. Nagel, K., P. Wagner, and R. Woesler. “Still Flowing: Approaches to Traffic Flow and Traffic Jam Modeling”. *Op. Res.* 51, no. 5 (2003): 681–710 (see p. 15).
104. Nakagami, M. “The m-distribution - A General Formula of Intensity Distribution of Rapid Fading”. In: *Statistical Methods in Radio Wave Propagation*. Ed. by Hoffman, W. G. Oxford: Pergamon, 1960, 3–36 (see p. 59).
105. Nakayama, A., M. Fukui, M. Kikuchi, K. Hasebe, K. Nishinari, Y. Sugiyama, S.-i. Tadaki, and S. Yukawa. “Metastability in the formation of an experimental traffic jam”. *New J. Phys.* 11, no. 8 (2009): 083025 (see pp. 10, 19).
106. Namazi, A., N. Eissfeldt, P. Wagner, and A. Schadschneider. “Boundary-induced phase transitions in a space-continuous traffic model with non-unique flow-density relation”. *Eur. Phys. J. B* 30, no. 4 (2002): 559–570 (see p. 40).
107. Narzt, W., U. Wilflingseder, G. Pomberger, D. Kolb, and H. Hortner. “Self-organising congestion evasion strategies using ant-based pheromones”. *Intell. Transp. Syst., IET* 4, no. 1 (2010): 93–102 (see p. 63).
108. Neubert, L., L. Santen, A. Schadschneider, and M. Schreckenberg. “Single-vehicle data of highway traffic: A statistical analysis”. *Phys. Rev. E* 60 (6 1999): 6480–6490 (see p. 38).
109. NHTSA. *Vehicle Safety Communications Project - Final Report*. Tech. rep. DOT HS 810 591. Washington, DC: USDOT, 2006 (see pp. 64, 68).
110. Nicklin, R. C. “Kinematics of tailgating”. *Phys. Teach.* 35 (1997): 78–79 (see p. 52).
111. Ögüt, K. S. and J. H. Banks. “Stability of Freeway Bottleneck Flow Phenomena”. *Transp. Res. Rec.* 1934 (2005): 108–115 (see p. 11).

112. Persaud, B., S. Yagar, and R. Brownlee. “Exploration of the Breakdown Phenomenon in Freeway Traffic”. *Transp. Res. Rec.* 1634 (1998): 64–69 (see p. 11).
113. Popkov, V., L. Santen, A. Schadschneider, and G. M. Schütz. “Empirical evidence for a boundary-induced nonequilibrium phase transition”. *J. Phys. A* 34, no. 6 (2001): L45 (see pp. 2, 12, 16, 31, 41).
114. Popkov, V. and G. M. Schütz. “Steady-state selection in driven diffusive systems with open boundaries”. *Europhys. Lett.* 48, no. 3 (1999): 257 (see p. 11).
115. Prigogine, I. and R. Herman. “Kinetic Theory of Vehicular Traffic”. New York: American Elsevier, 1971 (see pp. 1, 5).
116. Pulkkinen, O. and J. Merikoski. “Cluster size distributions in particle systems with asymmetric dynamics”. *Phys. Rev. E* 64 (5 2001): 056114 (see p. 83).
117. Punzo, V. and F. Simonelli. “Analysis and Comparison of Microscopic Traffic Flow Models with Real Traffic Microscopic Data”. *Transp. Res. Rec.* 1934, no. 1 (2005): 53–63 (see p. 26).
118. Ranjitkar, P., T. Nakatsuji, and M. Asano. “Performance Evaluation of Microscopic Traffic Flow Models with Test Track Data”. *Transp. Res. Rec.* 1876 (2004): 90–100 (see p. 90).
119. Rappaport, T. S. “Wireless Communications: Principles and Practice”. 2nd ed. Upper Saddle River: Prentice Hall, 2002 (see pp. 58, 59).
120. Rehborn, H., S. L. Klenov, and J. Palmer. “Common traffic congestion features studied in USA, UK, and Germany based on Kerner’s three-phase traffic theory”. In: *IEEE Intell. Veh. Symp.* 2011: 19–24 (see pp. 2, 13, 31, 32).
121. Richards, P. I. “Shock Waves on the Highway”. *Oper. Res.* 4, no. 1 (1956): 42–51 (see p. 2).
122. Rickert, M., K. Nagel, M. Schreckenberg, and A. Latour. “Two lane traffic simulations using cellular automata”. *Physica A* 231, no. 4 (1996): 534–550 (see p. 20).
123. Schadschneider, A., D. Chowdhury, and K. Nishinari. “Stochastic Transport in Complex Systems: From Molecules to Vehicles”. Oxford: Elsevier Science, 2010 (see pp. 6, 7, 15).
124. Schönhof, M. and D. Helbing. “Criticism of three-phase traffic theory”. *Transp. Res. B* 43, no. 7 (2009): 784–797 (see p. 14).
125. Schönhof, M. and D. Helbing. “Empirical Features of Congested Traffic States and Their Implications for Traffic Modeling”. *Transport. Sci.* 41, no. 2 (2007): 135–166 (see pp. 10, 14).
126. Schrank, D., T. Lomax, and B. Eisele. “Urban mobility report 2011”. Texas Transportation Institute, The Texas A&M University System, 2011 (see p. 57).
127. Shang, H.-Y. and Y. Peng. “A new three-step cellular automaton model considering a realistic driving decision”. *J. Stat. Mech.* 2012, no. 10 (2012): P10001 (see pp. 52, 89).
128. Singh, P. K. and K. Lego. “Comparative Study of Radio Propagation and Mobility Models in Vehicular Adhoc Network”. *Int. J. Comp. Appl.* 16, no. 8 (2011): 37–42 (see p. 58).

129. Sommer, C. and F. Dressler. “Using the Right Two-Ray Model? A Measurement based Evaluation of PHY Models in VANETs”. In: *17th ACM Int. Conf. MobiCom*. Las Vegas: ACM, 2011 (see p. 60).
130. Sommer, C., O. K. Tonguz, and F. Dressler. “Adaptive Beaconing for Delay-Sensitive and Congestion-Aware Traffic Information Systems”. In: *2nd IEEE VNC*. IEEE, 2010: 1–8 (see pp. 63, 64).
131. Sparmann, U. “Spurwechselforgänge auf zweispurigen BAB-Richtungsfahrbahnen”. Forschung Straßenbau und Straßenverkehrstechnik. Bonn: Bundesministerium für Verkehr, 1978 (see p. 5).
132. Sugiyama, Y., M. Fukui, M. Kikuchi, K. Hasebe, A. Nakayama, K. Nishinari, S. Tadaki, and S. Yukawa. “Traffic jams without bottlenecks – experimental evidence for the physical mechanism of the formation of a jam”. *New J. Phys.* 10, no. 3 (2008): 033001 (see pp. 10, 19).
133. Taliwal, V., D. Jiang, H. Mangold, C. Chen, and R. Sengupta. “Empirical determination of channel characteristics for DSRC vehicle-to-vehicle communication”. In: *Proc. 1st ACM int. workshop on Vehicular ad hoc networks*. VANET '04. New York: ACM, 2004: 88 (see p. 60).
134. Tilch, B. and D. Helbing. “Evaluation of single vehicle data in dependence of the vehicle-type, lane, and site”. In: *Traffic and Granular Flow '99*. Ed. by Helbing, D., H. J. Herrmann, M. Schreckenberg, and D. E. Wolf. Berlin: Springer, 2000: 333–338 (see p. 38).
135. Toroczkai, Z., G. Korniss, S. Das Sarma, and R. K. P. Zia. “Extremal-point densities of interface fluctuations”. *Phys. Rev. E* 62 (1 2000): 276–294 (see pp. 80, 85).
136. *Traffic information system “autobahn.NRW”*. 2011. URL: <http://www.autobahn.nrw.de/> (see pp. 34, 55).
137. Transp. Res. Board. “Highway capacity manual”. Nat. Res. Council (U.S.), 1994. Chap. 3 (see p. 65).
138. Treiber, M. and D. Helbing. “An adaptive smoothing method for traffic state identification from incomplete information”. In: *Interface and transport dynamics: Computational Modelling*. Berlin: Springer, 2003, 343–360 (see p. 7).
139. Treiber, M., A. Hennecke, and D. Helbing. “Congested traffic states in empirical observations and microscopic simulations”. *Phys. Rev. E* 62 (2 2000): 1805–1824 (see pp. 5, 13, 16–18).
140. Treiber, M. and A. Kesting. “Modeling Lane-Changing Decisions with MOBIL”. In: [TGF2007]: 211–221 (see p. 18).
141. Treiber, M. and A. Kesting. “Verkehrsdynamik und -simulation”. Berlin: Springer, 2010 (see pp. 17, 18).
142. Treiber, M., A. Kesting, and D. Helbing. “Three-phase traffic theory and two-phase models with a fundamental diagram in the light of empirical stylized facts”. *Transp. Res. B* 44, no. 8–9 (2010): 983–1000 (see p. 14).
143. Treiber, M., A. Kesting, and D. Helbing. “Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps”. *Phys. Rev. E* 74 (1 2006): 016123 (see p. 14).

- 
144. Treiterer, J. *Investigation of traffic dynamics by aerial photogrammetry techniques*. Final Report. Ohio State University, 1975 (see p. 10).
  145. Wedel, J. W., B. Schünemann, and I. Radusch. “[V2X-Based Traffic Congestion Recognition and Avoidance](#)”. In: *Proc. 10th Int. Symp. on Pervasive Syst., Algorithms, and Networks*. ISPAN '09. Washington, DC: IEEE Comput. Soc., 2009: 637–641 (see p. 63).
  146. Yin, J., G. Holland, T. ElBatt, F. Bai, and H. Krishnan. “[DSRC Channel Fading Analysis from Empirical Measurement](#)”. In: *1st Int. Conf. on Communications and Networking in China*. *ChinaCom*. 2006: 1–5 (see p. 60).

## Abstract

Traffic flow is a very prominent example of a driven non-equilibrium system, which shows a very complex spatiotemporal dynamics. A characteristic phenomenon of traffic dynamics is the spontaneous and abrupt drop of the average velocity on a stretch of road leading to congestion.

To assess the quality of three selected microscopic traffic models (the Nagel-Schreckberg model (NSM), the intelligent driver model (IDM), and the comfortable driving model (CDM)), we study their ability to reproduce such a traffic breakdown, whose spatiotemporal dynamics we investigate as well. Our analysis is based on empirical traffic data from stationary loop detectors showing a spontaneous breakdown on a German Autobahn. We then present several methods to assess the results and to compare the models with each other. In addition, we will also discuss some important modeling aspects and their impact on the resulting spatiotemporal pattern.

For the CDM, which gave good results in this assessment, we analyze the spatiotemporal patterns resulting from different inflow and outflow rates with open boundary conditions. Based on time series of local measurements, the local traffic states are assigned to different traffic phases according to Kerner's three-phase traffic theory. For this classification we use the rule-based FOTO-method, which was developed by Kerner et al. [69]. Our analysis shows that the model is indeed able to reproduce three qualitatively different traffic phases: free flow, synchronized traffic, and wide moving jams. This is surprising because traffic models with a fundamental diagram, such as the CDM, are not expected to reproduce the synchronized phase.

By virtue of this overall good agreement with empirical findings, we chose the CDM to investigate via computer simulations how traffic congestion can be reduced with the help of vehicle-to-vehicle communication. As the reasons for a traffic breakdown are perturbations caused by human drivers in dense traffic, we propose using periodically emitted status messages to analyze traffic flow and to warn other drivers of a possible traffic breakdown. Drivers who receive such a warning are told to keep a larger gap to their predecessor. By doing so, they are less likely the source of perturbations, which can cause a traffic breakdown. We show that penetration rates of 10% and less can have significant influence on traffic flow and travel times.

Finally, we address a rather practical problem of heterogeneous traffic consisting of communicating and non-communicating vehicles. If communicating vehicles can detect the vehicle ahead (and behind) by front (and rear) sensors, we give exact solutions for the average number of detected vehicles.

## Zusammenfassung

Die Dynamik des Verkehrsflusses zeigt unterschiedliche, komplexe räumliche und zeitliche Muster. Ein bekanntes Phänomen ist beispielsweise der abrupte und plötzliche Abfall der Durchschnittsgeschwindigkeit auf einem Streckenabschnitt, der zu Stauungen führt.

Zur Beurteilung der Qualität dreier ausgewählter Verkehrsmodelle (dem Nagel-Schreckberg Modell (NSM), dem Intelligent Driver Model (IDM) und dem Comfortable Driving Model (CDM)) wird deren Fähigkeit einen solchen Zusammenbruch des Verkehrsflusses zu reproduzieren untersucht. Für eine möglichst realistische Untersuchung werden dazu reale Zählschleifendaten herangezogen, die eben einen solchen Zusammenbruch zeigen. Es werden mehrere Methoden vorgestellt, mit denen die Übereinstimmung der empirischen mit den Simulationsdaten beurteilt werden kann. Zusätzlich wird der Einfluss einiger Modellierungsaspekte auf die Ergebnisse diskutiert.

Für das CDM, welches eine gute Übereinstimmung mit den Echtdaten zeigte, werden in einem zweiten Schritt die raum-zeitlichen Verkehrsmuster untersucht, die sich aus unterschiedlichen Ein- und Ausflussraten bei offenen Rändern ergeben. Auf Grundlage der Zeitreihen lokaler Messungen werden dann die Verkehrszustände den Verkehrsphasen der Kernerschen Drei-Phasen-Verkehrstheorie zugeordnet. Für diese Zuordnung wird die regelbasierte FOTO-Methode verwendet [69]. Auch diese Analyse zeigt, dass das CDM alle drei unterschiedlichen Verkehrsphasen (Freifluss, synchronisierter Verkehr und Stau) reproduzieren kann. Diese Beobachtung ist überraschend, da diese Eigenschaft von Verkehrsmodellen mit einem Fundamentaldiagramm wie dem CDM nicht erwartet wird.

Aufgrund dieser insgesamt guten Übereinstimmung mit empirischen Daten wurde das CDM auch für simulative Untersuchungen von Strategien zur Verkehrsflussoptimierung mittels Fahrzeug-Fahrzeug Kommunikation verwendet. Da der Zusammenbruch des Verkehrsflusses aus von Fahrern verursachten Störungen resultiert, wird vorgeschlagen die Verkehrslage durch regelmäßig gesendete Kurznachrichten zu analysieren und Fahrer vor einem drohenden Zusammenbruch des Verkehrs zu warnen. Fahrer, die eine solche Warnung erhalten, halten daraufhin einen größeren Abstand zum Vordermann ein und verursachen dadurch mit geringerer Wahrscheinlichkeit Fluktuationen, die den Verkehr zusammenbrechen lassen. Es zeigt sich, dass schon Durchdringungsraten von 10 % kommunizierender Fahrzeuge einen deutlich messbaren Einfluss auf den gesamten Verkehrsfluss haben.

Schließlich betrachten wir heterogenen Verkehr, bestehend aus kommunizierenden und nicht-kommunizierenden Fahrzeugen, genauer. Unter der Annahme, dass kommunizierende Fahrzeuge ihren Vorder- und Hintermann mit Hilfe von Sensoren orten können, fragen wir, wie viele nicht-kommunizierende Fahrzeuge so im Durchschnitt geortet werden.

## Erklärung

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbständig, ohne fremde Hilfe und ohne Benutzung anderer als den angegebenen Quellen angefertigt zu haben. Alle aus fremden Werken direkt oder indirekt übernommenen Gedanken sind als solche gekennzeichnet.

Die vorliegende Dissertation wurde in keinem anderen Promotionsverfahren eingereicht.

Mit dieser Arbeit strebe ich die Erlangung des akademischen Grades “Doktor der Naturwissenschaften” (Dr. rer. nat.) an.

Duisburg, den 13. Februar 2013

---

Florian Knorr

## Teilpublikationen

- T74. Knorr, F., D. Baselt, M. Schreckenberg und M. Mauve. „[Reducing Traffic Jams via VANETs](#)“. *IEEE Trans. Veh. Technol.* 61, no. 8 (2012): 3490–3498 (siehe S. 57).
- T75. Knorr, F. und M. Schreckenberg. „[Counting the corners of a random walk and its application to traffic flow](#)“. *J. Phys. A* 45, no. 31 (2012): 315001 (siehe S. 79).
- T76. Knorr, F. und M. Schreckenberg. „[Influence of inter-vehicle communication on peak hour traffic flow](#)“. *Physica A* 391, no. 6 (2012): 2225–2231 (siehe S. 57).
- T77. Knorr, F. und M. Schreckenberg. „[On the reproducibility of spatiotemporal traffic dynamics with microscopic traffic models](#)“. *J. Stat. Mech.* 2012, no. 10 (2012): P10018 (siehe S. 15).
- T78. Knorr, F. und M. Schreckenberg. „[The comfortable driving model revisited: traffic phases and phase transitions](#)“. *J. Stat. Mech.* 2013, no. 7 (2013): P07002 (siehe S. 37).

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.