A Bayesian inference-based feedback on car-following behavior
Improving the validity of traffic simulations with human driving input

R.B. Molenaar
A Bayesian inference-based feedback on car-following behavior
Improving the validity of traffic simulations with human driving input

MASTER OF SCIENCE THESIS

For the degree of Master of Science in Systems and Control at Delft University of Technology

R.B. Molenaar

February 25, 2014

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology
The work in this thesis was supported by the National Institute of Informatics, as part of their Global Lab project. Their cooperation is hereby gratefully acknowledged.
Traffic simulators are useful tools that can be used for entertainment, driver education, and research into human driving behavior and the effects of innovative traffic solutions. A traffic simulator provides a high degree of control, as every traffic situation can be recreated with exactly the same conditions for all test subjects. The downside of working with a simulator is the limited validity. When simulating a new situation, observing valid driving behavior from the test subjects requires valid simulated background traffic. In order to generate valid background traffic, realistic driving models are needed for that specific situation. This creates a vicious circle, as the realistic observations that are not observed yet are required to generate realistic driving models.

The purpose of this research is to generate realistic simulated traffic for new situations. Using the Bayesian inference method in an online feedback loop, observed driving behavior from test subjects is used to calibrate existing driving models. These models are fed back to the simulator, where they will be used to generate traffic with the information gathered from the subjects. The influence of implementing a Bayesian-inference based feedback loop on subjective validity of the simulated traffic is tested in various experiments.

The results show that a direct comparison, where the updated traffic and non-updated traffic are observed in the same session, results in an increase in subjective validity. An indirect comparison does not lead to an increase in subjective validity. The conclusion from the results is that, although the feedback loop leads to an increase in subjective validity, the effect is not strong enough to be noticeable when comparing traffic from two separate sessions.
# Table of Contents

## Acknowledgements

1 Introduction

1-1 Introduction ....................................................... 1
1-2 Motivation of research ........................................... 1
1-3 Goal ......................................................... 2
1-4 Current situation ............................................... 3
1-5 Research questions .............................................. 4
1-6 Further considerations ........................................... 5
1-7 Research approach ............................................... 5
1-7-1 Theory on longitudinal driving behavior ..................... 7
1-7-2 Methodology for calibrating longitudinal driving models ... 7
1-7-3 Experimental design for driving simulator ................. 7
1-7-4 Empirical findings obtained from experiments ............ 7

2 Car-following models

2-1 State definition .................................................. 10
2-2 Linear models ................................................... 11
  2-2-1 Linear General Motors model .......................... 11
  2-2-2 Helly model ............................................... 11
2-3 Nonlinear models ................................................. 12
  2-3-1 Gazis-Herman-Rothery model ......................... 12
  2-3-2 Intelligent Driver Model ............................... 12
  2-3-3 Action point model ................................... 14
  2-3-4 Collision avoidance model .......................... 14
  2-3-5 Latent class model ................................... 15
2-4 Model choice ..................................................... 17
2-5 Chapter summary ................................................ 17

Master of Science Thesis R.B. Molenaar
### Table of Contents

#### 3 Bayesian inference
- 3-1 Calculation of posterior distribution ........................................ 20
- 3-2 Cramér-Rao bound .............................................................. 21
  - 3-2-1 Constrained Cramér-Rao bound ...................................... 22
  - 3-2-2 Application of Cramér-Rao bound ................................... 22
- 3-3 Model comparison ............................................................... 23

#### 4 Research methodology
- 4-1 Overall methodology ......................................................... 25
- 4-2 Application of Bayesian method .......................................... 26
  - 4-2-1 Evidence for the GM model ........................................... 26
  - 4-2-2 Evidence for the IDM .................................................. 27
  - 4-2-3 Reaction time implementation ........................................ 28
  - 4-2-4 Adaptation of simulation environment ............................... 29
- 4-3 Testing performance of adaptation algorithm .......................... 29
  - 4-3-1 Experiments with synthetic data ..................................... 29
  - 4-3-2 Experiments with real data ........................................... 30
- 4-4 Testing subjective validity of adaptive bots ............................ 30
  - 4-4-1 Identifying a change in subjective validity ......................... 31
  - 4-4-2 Experiments with user interaction .................................. 31

#### 5 Experimental setup
- 5-1 Introduction ........................................................................ 35
- 5-2 Traffic simulator setup ......................................................... 35
- 5-3 Questionnaire design .......................................................... 36
- 5-4 Processing the questionnaires .............................................. 36
- 5-5 Assumptions ......................................................................... 37
  - 5-5-1 Observation models ...................................................... 37
  - 5-5-2 Observation interval ....................................................... 38
  - 5-5-3 Parameters and constraints ............................................. 38

#### 6 Result analysis and discussion
- 6-1 Research question 1 ............................................................ 42
- 6-2 Research question 2 ............................................................. 43
- 6-3 Research question 4 ............................................................. 45
- 6-4 Research question 5 ............................................................. 46
- 6-5 Parameter evolution ............................................................. 48
- 6-6 Discussion ............................................................................. 49
  - 6-6-1 Simulator ........................................................................ 49
  - 6-6-2 Participants ...................................................................... 51
  - 6-6-3 Experiment ....................................................................... 51
  - 6-6-4 Unexpected results .......................................................... 52
## Table of Contents

7 Conclusions ................................. 53
   7-1 Synthetic data .................................. 53
   7-2 Real data .................................. 54
   7-3 Interactive experiments ...................... 54

8 Further recommendations ...................... 55
   8-1 Length .................................. 55
   8-2 Adaptation .................................. 55
   8-3 Extending the algorithm ...................... 56
      8-3-1 Lane changing .......................... 56
      8-3-2 Extend model base ...................... 56
   8-4 Improve simulator ........................... 57
   8-5 High fidelity simulator ..................... 57

A Experiment remarks ........................ 59

B C# Implementation .......................... 63
   B-1 Main .................................. 63

C Session questionnaire ....................... 65
   C-1 Acceleration at intersection .............. 65
   C-2 Following distance ....................... 65
   C-3 Speed .................................. 66
   C-4 Reaction Time ............................ 66
   C-5 Open questions ........................... 67

D Final questionnaire .......................... 69

E Matlab .................................. 73
   E-1 Experiment 1 ............................... 73
   E-2 Experiment 2 ................................ 76

F Convex functions ............................ 79

   Bibliography ................................. 81

   Glossary .................................. 85
      List of Acronyms .......................... 85
      List of Symbols ........................... 85
List of Figures

1-1 Test track ................................................................. 3
1-2 Schematic representation of report outline .......................... 6

2-1 Car-following state definition ........................................ 10
2-2 Example of psycho-physical car-following behavior ............... 14
2-3 Latent class model structure .......................................... 15
2-4 Example of acceleration distribution per state ..................... 16

5-1 Example question showing the points assigned to each answer . 37

6-1 Speed pattern of lead car in synthetic experiment .................. 42
6-2 Plot showing the evolution of the parameters over time .......... 48
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Parameter estimates from different sources</td>
<td>13</td>
</tr>
<tr>
<td>4-1</td>
<td>The 3 different sessions performed in the color-based analysis of the experiment</td>
<td>31</td>
</tr>
<tr>
<td>6-1</td>
<td>Synthetic experiment results</td>
<td>43</td>
</tr>
<tr>
<td>6-2</td>
<td>Results of experiment 2</td>
<td>43</td>
</tr>
<tr>
<td>B-1</td>
<td>Table with input and output data types of main function</td>
<td>63</td>
</tr>
</tbody>
</table>
I would like to thank various people for their contributions to this research project; Prof.dr. Helmut Prendinger, for allowing me to be part of his lab for 6 months and for providing me with everything I needed to perform my experiments; Dr. Kugamoorthy and Mr. Madruga, for their endless support in the implementation and execution of the experiments.

Special thanks should be given to Prof. dr. Hans van Lint, my research project supervisor for his encouraging guidance and valuable support and to Dr. Mernout Burger, for his useful and constructive feedback on this project and for taking the time to finish this project with me.

Finally, I wish to thank my family and friends for their support, encouragement and feedback they have provided me with throughout my study.
Chapter 1

Introduction

1-1 Introduction

With an increasing amount of traffic occupying the road network, the need for innovative traffic solutions increases. A simulation can be used to evaluate the possible effects of new traffic solutions on driving behavior, without having to test them in a real life situation. This reduces the costs and safety risks of the evaluation and allows for full control of the situation, as very specific traffic conditions can be defined to simulate every possible situation. A traffic simulation comes with high degree of control, as the driving environment, specific driving scenarios and the driving behavior of the traffic in each situation can be adjusted. The situation can be kept identical for each test subject, which isolates the effects of the variables of an experiment.

The downside of using simulations is the limited validity. When discussing validity in this thesis, we distinguish two types of validity. Subjective validity describes how well the simulation experience matches the physical and visual experience of driving, while objective validity describes how well the behavioral response of the drivers participating in the simulation matches the response the drivers would have in a real-life situation.

The validity of the traffic simulation is very important. Whether the simulation is being used for educational purposes, such as driver training, or research in driving behavior, the validity of the simulation largely determines the relevance of the outcome. A strong correlation exists between the two types of validity. An increase in subjective validity is expected to lead to an increase in objective validity, since a driver is expected to behave more similar to 'the real thing' as the driving experience becomes more realistic.

1-2 Motivation of research

When driving simulators are used to assess how human drivers behave in new traffic situations, a subjectively realistic environment is essential to be able to observe valid driving
behavior. The simulated traffic that surrounds the user-controlled car forms an important role in modeling a realistic environment. The simulated traffic’s driving behavior is governed by microscopic driving models, which can be calibrated to fit different driving behaviors for different situations. For each situation, realistic simulated traffic can be obtained by setting the model parameters to match realistic human driving behavior in that situation. The problem is that realistic simulated background traffic is required in order to be able to observe a behaviorally valid response from the participants.

This problem is a vicious circle, as it is only possible to observe valid behavior if the background traffic is acting like the behavior that is being observed, which has not been modeled yet. To overcome this problem, a solution is proposed in this chapter. The simulated traffic is adapted to the observed behavior of the test subject during the experiment in (near)real-time as well as in between different runs. By continuously updating the driving behavior of the generated bots, using the observations gained from the subject, the subjective validity will increase (near) continuously. As a result of this, the objective validity will also increase.

1-3 Goal

The goal of this research is to develop a learning functionality for the traffic simulator currently used at the National Institute of Informatics, Tokyo (NII), to make the driving behavior of the simulated traffic more realistic. This is achieved by sampling and analyzing the driving behavior of a person who is participating as a driver in the traffic simulation. This information is used to perform modeling and parameter estimation, resulting in a model that describes human-like driving behavior. This model can then be used to adjust the driving behavior of the non-human drivers (bots) in the microscopic traffic simulator. A probability density function of each parameter and among the models is created to match the inter-driver heterogeneity of real traffic. Inter-driver heterogeneity is a concept referring to the differences in driving behavior between different people in a certain situation. The parameters for the non-human drivers will be selected from these probability density functions, based on their respective probability distributions.

Subjecting human drivers to realistic background traffic while participating in a simulator increases the behavioral validity of their driving. Assuming that the human behavior has not yet been measured or modeled for the specific simulation/situation, there is no model available to correctly define the bots such that they will exhibit the realistic driving behavior that has yet to be observed. This means that the participants can not be subjected to this realistic background traffic, which has a negative effect on the validity of their driving behavior as interactions with the surrounding traffic are a key factor in human driving behavior.

This problem can be solved by performing an online update of the simulated driving behavior, applying the identified models during the same simulation from which the observations are obtained. This will result in an incremental improvement in behavioral validity of the driving behavior observed from the participants.
The research is part of the Global Lab project at NII, in collaboration with Delft University of Technology (TU Delft). A traffic simulator is being developed at NII as part of the global lab project, which will be used to conduct the experiments. Before describing the changes to the simulator required for performing the experiments, it is important to know the current situation to which the changes will be made. The focus of this research lies in analyzing the longitudinal driving behavior of human drivers and using the information gathered from these drivers to adapt the models in the traffic simulator, adjusting the longitudinal driving behavior of the simulated cars, from now known as Artificial Intelligence (AI) cars, to that of the user driven cars. The description of the system will focus on this aspect, as this will be subject to change. The AI car traffic is modeled as a multi-agent system, where each car’s longitudinal driving behavior is generated using the Intelligent Driver Model (IDM).

In the traffic simulator at the NII, the parameters of the IDM are based on a calibration of a real world driving experiment, involving a platoon of cars following each other. The model parameters do not change with time, surroundings, or any other change in the simulation environment. Therefore, the model is not adapted to possible changes in the situation. The model is based on human driving behavior in a very specific situation, namely at the Shirosato Test Centre of the Japan Automobile Research Institute (JARI) [1]. The test site consists of a large oval on which a platoon of 15 cars were driven by 15 test subjects. Each individual’s driving behavior was used to calibrate an IDM model and these models were used to define the background traffic in the simulator. As the situation where the real world experiment was done, an open oval, does not really compare to the situation in the simulator, a city centre, the model is not suited to model human driving behavior in this simulator situation.

The simulated environment is an environment which allows users to drive among simulated traffic around a virtual 2.5 km loop through a city centre, which is modeled after the Yasukuni Dori, in Tokyo. The environment is constructed in such a way that different types of driving behavior will be observed in different parts of the loop. In the lower half of the loop, traffic will pass a series of signalized intersections. In this part of the loop, car-following behavior can be observed, as cars are moving in between the intersections in platoons. In the upper part of the loop, free driving behavior can be observed. This part of the loop has no intersections and will allow the drivers to drive freely, given that the traffic density isn’t too high. A schematic representation of the test track is shown in Figure 1-1.

Figure 1-1: Test track

14 Current situation

The research is part of the Global Lab project at NII, in collaboration with Delft University of Technology (TU Delft). A traffic simulator is being developed at NII as part of the global lab project, which will be used to conduct the experiments. Before describing the changes to the simulator required for performing the experiments, it is important to know the current situation to which the changes will be made. The focus of this research lies in analyzing the longitudinal driving behavior of human drivers and using the information gathered from these drivers to adapt the models in the traffic simulator, adjusting the longitudinal driving behavior of the simulated cars, from now known as Artificial Intelligence (AI) cars, to that of the user driven cars. The description of the system will focus on this aspect, as this will be subject to change. The AI car traffic is modeled as a multi-agent system, where each car’s longitudinal driving behavior is generated using the Intelligent Driver Model (IDM).

In the traffic simulator at the NII, the parameters of the IDM are based on a calibration of a real world driving experiment, involving a platoon of cars following each other. The model parameters do not change with time, surroundings, or any other change in the simulation environment. Therefore, the model is not adapted to possible changes in the situation. The model is based on human driving behavior in a very specific situation, namely at the Shirosato Test Centre of the Japan Automobile Research Institute (JARI) [1]. The test site consists of a large oval on which a platoon of 15 cars were driven by 15 test subjects. Each individual’s driving behavior was used to calibrate an IDM model and these models were used to define the background traffic in the simulator. As the situation where the real world experiment was done, an open oval, does not really compare to the situation in the simulator, a city centre, the model is not suited to model human driving behavior in this simulator situation.

The simulated environment is an environment which allows users to drive among simulated traffic around a virtual 2.5 km loop through a city centre, which is modeled after the Yasukuni Dori, in Tokyo. The environment is constructed in such a way that different types of driving behavior will be observed in different parts of the loop. In the lower half of the loop, traffic will pass a series of signalized intersections. In this part of the loop, car-following behavior can be observed, as cars are moving in between the intersections in platoons. In the upper part of the loop, free driving behavior can be observed. This part of the loop has no intersections and will allow the drivers to drive freely, given that the traffic density isn’t too high. A schematic representation of the test track is shown in Figure 1-1.
Introduction

1-5 Research questions

The main objective of the research presented in this thesis is to develop a method to increase the subjective validity of traffic simulators, by improving the longitudinal driving behavior of the simulated traffic with models based on the driving behavior that is displayed by participants in the simulator. The focus lies on the development and application of an algorithm and evaluating the effects of applying the method on the subjective validity experienced by participants of the simulation. The main research question that will be investigated in this thesis is:

*Can we improve the behavioral validity of driving simulator environments in an online Bayesian calibration-application loop that adapts the models of the background traffic to the behavior of test-subjects in the driving simulator based on limited real-time data?*

The research question is decomposed into several subquestions, in order to investigate each facet separately, before an answer can be given to the main research question. The subquestions are given below, including a description of how they fit in the research as a whole.

The first part of the research is to design an algorithm, using a Bayesian method proposed in [2], that is able to correctly identify a set of parameters and models, which are used to create a synthetic data set. The first sub question focuses on the suitability of the selected method, as presented in Chapter 3, to perform an identification on artificial data, which is generated using the selected models as described in Chapter 2.

**Research question 1:** *Is the Bayesian inference method able to correctly identify the set of models and parameters that are used to create a synthetically generated data set?*

Once it is established that the algorithm can perform as required with the synthetic data, the next step is to test the performance of the algorithm when supplied with real data, obtained from the traffic simulator in a previous experiment.

**Research question 2:** *Is the Bayesian inference method able to identify a model and parameter set that best fits the driving behavior that is observed from a human participant in the traffic simulator?*

Before applying a feedback loop in the simulator using the Bayesian method, another important issue is addressed. To find an improvement in subjective validity caused by the feedback loop, it is important to define a method to measure the relative improvement of subjective validity.

**Research question 3:** *How can an improvement in subjective validity be identified/proved?*

The second part of the research focuses on the effect of implementing the algorithm as a feedback loop in the simulator, where it will update the simulated traffic during the run of the simulation. Different combinations of updated and non-updated traffic are used, to investigate the effect of feeding back the observed driving behavior to the simulated traffic. In this part of the thesis, the following questions are investigated.

**Research question 4:** *Does updating the driving models of AI cars, by feeding back information gained from observing human driving behavior in the same simulation session, lead to such an improvement of the subjective validity of that driving behavior that it is noticeable in a direct comparison?*
Research question 5: Does updating the driving models of AI cars, by feeding back information gained from observing human driving behavior in the same simulation session, lead to such an improvement of the subjective validity of that driving behavior that it is noticeable when comparing over different sessions?

The information and conclusions obtained from investigating these sub questions will be used to answer the research question and gain insights into the influence of updating driving models on the perception of validity of the participants in traffic simulators. Finally, the impact of these results will be discussed and suggestions for future work are presented.

1-6 Further considerations

Besides answering the research questions, which form the backbone of the research, other factors are considered. These considerations are not the main focus of this research, but the choices and assumptions made do affect the outcome of the experiments. These considerations are shown below, together with a short explanation.

- *How can the Bayesian algorithm determine the proper application of a model?*

  All car-following models used in the experiments are parametrized models, whose output describes the acceleration of a car. The model relates the inputs the driver receives from the surroundings, such as position and velocity of its leader, to the acceleration of the car depending on the value of the parameters. The structure of the model and the physical meaning of the parameters limit the range of values that a certain parameter can have, in order for the model to operate as it is intended. Constraining the parameters to a certain range would solve this problem. The way to constrain the parameters is based on the identification algorithm.

- *How can an algorithm handle the possibility of limited observations?*

  Different ways of handling the possibility of limited observations come to mind:

  - Fix the undetermined parameters in the calibration algorithm, so that the available estimates do not deteriorate over time due to noise. The last available estimate of the parameter can then be used for all current models, until the data allows that parameter to be estimated accurately once more.
  
  - Design the experiment in such a way that all different driving behaviors will regularly, thus providing enough information for the calibration algorithm. The size of the data available to the calibration algorithm should be large enough to include the different types of driving behavior required to correctly calibrate the models.

1-7 Research approach

This report is organized according to the diagram presented in Figure 1-2, showing all the work that is performed in order to answer the research questions.

In the next section, the different chapters of the thesis will be described in more detail.
Figure 1-2: Schematic representation of report outline
1-7 Research approach

1-7-1 Theory on longitudinal driving behavior

One of the defining aspects of driving is the longitudinal driving behavior, as it has a high influence in traffic flows. A research that focuses on identifying car following behavior in human drivers has to start by understanding the car-following models themselves. In this chapter, the different models that will be used in the experiments are discussed.

1-7-2 Methodology for calibrating longitudinal driving models

Fitting observed driving behavior to car-following models is the main aspect of this research, finding a best matching model and parameter set to an observed data set. Applying the method to the traffic simulator brings some challenges with it. The maths behind the algorithm, the implementations of the models in the algorithm and a way to deal with the imposed challenges are all presented in this chapter.

1-7-3 Experimental design for driving simulator

An experiment is performed, in order to answer the research questions. The design of these experiments is vital to the reliability of the information and a good experiment design will help to gain more insights into the topics that are investigated. The methodology for the experiments, focused on answering the proposed research questions, is described in Chapter 4. The results of the experiments are recorded using questionnaires, where the participant’s experience is gathered using a series of open and multiple choice questions. The questionnaires are designed to relate the experiences of people to traffic engineering terms, without requiring knowledge about the subject of traffic models. The experimental setup, describing the assumptions made for these specific experiments, and the design of the questionnaires are discussed in Chapter 5.

1-7-4 Empirical findings obtained from experiments

In Chapter 6, the results of the experiments are presented. A detailed analysis of the results will show the influence of the implemented algorithm on the people’s driving experience, and the answers to the research questions are given, based on these result.
Chapter 2

Car-following models

At the core of a traffic simulation lie the traffic models describing the driving behavior in the simulation. The accuracy of this simulation is largely dependent on the traffic-flow model that is used to model the driving behavior of traffic. Different types of traffic models can be considered, each type has its own distinctive way of predicting traffic behavior based on a set of parameters.

A macroscopic traffic flow model describes the properties of traffic flow as a whole, comparable to a fluid dynamics model, by relating the different traffic flow characteristics to each other. Just as in a fluid stream, the properties of the traffic are expressed in mean speed, density and flow. Each car is considered a particle in the flow, where the cars share the same properties with the cars around them.

A microscopic traffic-flow model describes the driving behavior of each car individually. It usually consists of multiple sub models where each sub model describes a specific driving task, such as lane-changing and car-following. The lane-changing model represents the latitudinal driving behavior, describing under which circumstances vehicles change lanes. The car-following model describes the longitudinal driving behavior, a combination of free driving behavior, in which the vehicle accelerates to a desired free speed, and car-following behavior, where the vehicle’s acceleration is generally a function of it’s predecessor’s position and speed.

In this thesis the focus is on the models describing car-following behavior. Car-following behavior has a significant impact on traffic performance, and is therefore suited for this research. Different types of car-following models have been developed over the years, of which an overview is given in [3]. Car-following behavior can be observed from data by considering sets of cars driving in platoons, where the first car is the leader whose driving behavior is a measurable disturbance. The driving behavior of the followers can be defined by the various models presented in this chapter.
A standard notation will be used in the different models that will be discussed, in order to display their differences and similarities clearly. The standard notation, which will be used throughout this thesis, is shown in the equation below and matches the image in Figure 2-1:

\[
\begin{align*}
  x_i & \text{ is the position of vehicle } i, \text{ in [m]} \\
  v_i & \text{ is the speed of vehicle } i, \text{ in [m/s]} \\
  a_i & \text{ is the acceleration of vehicle } i, \text{ in [m/s}^2] \\
  \Delta x & \text{ is the relative position of vehicle } i \text{ with respect to vehicle } i-1 \\
  \Delta v & \text{ is the relative speed of vehicle } i \text{ with respect to vehicle } i-1 \\
  L_i & \text{ is the length of vehicle } i, \text{ in [m]} \\
  \tau_i & \text{ is the reaction time of driver } i, \text{ in [s]}
\end{align*}
\]

\[
(2-1)
\]

### 2.1 State definition

To analyze car-following behavior, platoons of \( n \) cars are observed (where \( n \geq 2 \)) such that the platoon contains at least 1 follower. To match the notation in the models, let \( x_i(t) \) and \( v_i(t) \) denote the position and speed of vehicle \( i \) at time \( t \), where the leader of the platoon is indicated by \( i = 1 \) and the last vehicle in the platoon by \( i = n \). The positions and velocities of the platoon are collected in the state of the platoon at time \( t \), defined by

\[
x = [x_1(t), v_1(t), \ldots, x_n(t), v_n(t)]'
\]

Platoons with more than 2 cars can be used for the analysis of multi-anticipative car-following behavior, in which the driving behavior of car \( i \) not only depends on car \( i-1 \), but also on the cars driving further upstream in the platoon.

As every driver has a different driving style, the parameters in each model need to be defined for every individual. Let \( \theta_i(t) \) denote the general parameter set used to describe driving behavior for driver \( i \) at time \( t \). As one would expect, human driving behavior is variable over time, as proved in [4]. This means that the parameters should be modeled as time-dependent values, as is done above.
2-2 Linear models

In this section, different types of linear car-following models are discussed. The models determine the acceleration of vehicle \( i \) at time \( t \) based on the information available at time \( t - \tau \). By delaying the available information by a certain time \( \tau \), the reaction time of a driver can be implemented in the model.

2-2-1 Linear General Motors model

The simplest linear car-following model is the linear General Motors (GM) model, as presented in [5]. It is based on the hypothesis that the acceleration is proportional to the relative speed, as shown in (2-2).

\[
a_i(t) = \alpha \Delta v_i^{(1)}(t - \tau) \quad \text{where} \quad \Delta v_i^{(1)} = v_{i-1} - v_i
\]

where the vector \( \theta_i = \{\alpha, \tau\} \) indicates the calibration parameters of the model. The assumption that the acceleration behavior is independent of the relative position is unrealistic [6]. Also, vehicles can be dragged along by vehicles that move at a higher speed, as the acceleration of vehicle \( i \) is solely defined by the difference in speed. A more general form of the GM model is given in Section 2-3-1, in which the influence of the relative distance and the speed of the driver itself is also included.

An extension to (2-2) is proposed by Bexelius in [7]. The influence of a second leader is included, leading to the ensuing multi-anticipative car-following model

\[
a_i(t) = \alpha_1 \Delta v_i^{(1)}(t - \tau) + \alpha_2 \Delta v_i^{(2)}(t - \tau) \quad \text{where} \quad \Delta v_i^{(2)} = v_{i-2} - v_i
\]

2-2-2 Helly model

The Helly model, proposed in [8], describes the acceleration of vehicle \( i \) as a function of the relative speed and distance to its leader. The proposed model introduces a factor \( D_i \) describing the desired following distance of driver \( i \) with an acceleration defined as

\[
a_i(t) = C_1 \Delta v(t - \tau) + C_2 \left[ \Delta x(t - \tau) - D_n(t) \right]
\]

\[
D_i(t) = \alpha + \beta v(t - \tau) + \gamma a_n(t - \tau)
\]

The calibration parameters are combined in \( \theta_i = \{C_1, C_2, \alpha, \beta, \gamma\} \). The implementation of the desired following distance factor proved a valuable improvement as the model performed well when fitted on observed data, reported in [8]. An extension on the Helly model was introduced in [4], modifying the model into a piecewise function, shown in (2-5). This extension introduced a deadzone, defined by \( c_1 \) and \( c_2 \), marking the unnoticeably small deviation of the distance headway \( S_1 \) for which a driver will adjust its acceleration. The acceleration towards the free speed \( v^0 \) with acceleration time \( \tau \) is also introduced, so that the model can also describe free-driving behavior.
\begin{equation}
    a_i(t) = \begin{cases}
        0 & c_1 < \Delta x_i(t) - S_1 < c_2 \\
        \min \left\{ a_1 \Delta v_i(t) + \beta_1 (\Delta x_i(t) - D_1), \frac{v_0 - v_i(t)}{\tau} \right\} & \text{elsewhere}
    \end{cases}
\end{equation}

\section{2-3 Nonlinear models}

The similarity between the different linear car-following models is large, as most models are derived from (2-2). The addition of nonlinearities in the models allows for the implementation of a more complex model. The increase in model complexity results in an increase in modeling power, but at the cost of a much more difficult identification, calibration and validation. The added model complexity also increases the chance of overfitting, which will lead to poor predictive performance of the model. The introduction of nonlinearities allows for a diversification of the models, as different models rely on different nonlinear terms to model the observed driving behavior.

\subsection{2-3-1 Gazis-Herman-Rothery model}

The Gazis-Herman-Rothery (GHR) model, as introduced in [9], was developed in the late fifties and early sixties. This model relates the acceleration of driver \(i\) to the relative velocity and relative distance with respect to its leader as well as its own velocity. The model is given by

\begin{equation}
    a_i(t) = cv_i^m(t) \frac{\Delta v(t - \tau)}{\Delta x_i(t - \tau)}
\end{equation}

where the calibration parameters are given by \(\theta_i = \{m_i, c_i, l_i\}\). The GHR model can be used as either a symmetric or an asymmetric model. The symmetric model describes both the acceleration and deceleration regime with the same \(\theta_i\), whereas the asymmetric model describes the different driving regimes using different parameter sets \(\theta_i^{\text{dec}}\) and \(\theta_i^{\text{acc}}\), where \(\theta_i^{\text{dec}}\) describes the deceleration parameter set and \(\theta_i^{\text{acc}}\) describes the acceleration parameter set.

A correct application of this model requires the parameters \(m\) and \(l\) to be calibrated for each particular network, one of the conclusions reported in [10]. A summary of different values found in the literature is presented in [3], where some sources made a distinction between different phases of driving behavior for parameter calibration. Some of these results are shown in Table 2-1, where 'dec' stands for deceleration and 'acc' stands for acceleration. It seems that either additional terms or a different model structure are required to achieve a more general description of straight line driving behavior.

\subsection{2-3-2 Intelligent Driver Model}

The Intelligent Driver Model (IDM) describes the acceleration of the vehicle \(i\) as a function of its velocity \(v_i\), the approach rate \(\Delta v_i\) and the gap \(s_i\) to the leading vehicle. The model is given by
Table 2-1: Parameter estimates from different sources

<table>
<thead>
<tr>
<th>Source</th>
<th>m</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chandler et al. (1958)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Herman and Potts (1959)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hoefs (1972) (dec no brk/dec brk/acc)</td>
<td>1.5/0.2/0.6</td>
<td>0.9/0.9/3.2</td>
</tr>
<tr>
<td>Treiterer and Myers (1974) (dec/acc)</td>
<td>0.7/0.2</td>
<td>2.5/1.6</td>
</tr>
<tr>
<td>Ozaki (1993) (dec/acc)</td>
<td>0.9/-0.2</td>
<td>1/0.2</td>
</tr>
</tbody>
</table>

\[
a_i(t) = \alpha_i \left[ 1 - \left( \frac{v_i(t)}{v_{0,i}} \right)^\delta - \left( \frac{s_i(t)^*}{s_{0,i}} \right)^2 \right] \tag{2-7}
\]

\[
s_i^* = s_0 + T v_i(t) + \frac{v_i(t)\Delta v_i(t)}{2\sqrt{\alpha i \beta}} \tag{2-8}
\]

where \( \theta_i = \{v_{0,i}, s_{0,i}, T_i, \alpha_i, \beta_i\} \) is the calibration parameter set and \( s_i^* \) is the desired gap for vehicle \( i \). All values of \( \theta_i \) are positive values, which means that calibrating the IDM will be a constrained optimization problem. The value of \( \delta \) can also be used for calibration, but is often fixed to \( \delta = 4 \) in the literature, for example in [11].

A closer examination of the IDM reveals that the acceleration function is a combination of two different acceleration behaviors. When the gap to the leader is large enough not to have an influence on the acceleration of vehicle \( i \), the equation reduces to

\[
a_{free} = \alpha \left[ 1 - \left( \frac{v_i}{v_{0,i}} \right)^\delta \right] \tag{2-9}
\]

which describes the free acceleration behavior of a car, with a decreasing magnitude until the desired driving speed \( v_0 \) is achieved. The second acceleration term describes the braking strategy, which will increasingly dominate the acceleration described by the IDM as the current gap \( s_i \) becomes smaller than the desired gap \( s_i^* \)

\[
a_{dec} = -\alpha \left( \frac{s_i^*}{s_{0,i}} \right)^2 \tag{2-10}
\]

Vehicles which are modeled using the IDM can not crash into their leader, hence the IDM is a collision-free model. This property of the IDM becomes clear when considering a situation in which a vehicle approaches a stopped or slow moving vehicle with a high relative speed \( \Delta v > 0 \). In such a case, the equilibrium part \( s_0 + T v \) of the desired minimum gap \( s^* \) can be neglected with respect to the non-equilibrium part. The term of the acceleration function governing the braking strategy, shown in (2-10) then becomes

\[
a_{dec} = \frac{(v_\Delta)^2}{4\beta s^2} \tag{2-11}
\]
which shows that $a_{\text{dec}} \to \infty$ as $s \to 0$, making a collision impossible. It is clear that the IDM shows unrealistic deceleration when the gap is much smaller than the desired gap. This shows that the IDM is a model that does not provide realistic driving behavior in near-collision situations.

When the IDM is applied in a discretized system collisions are possible, but they will only occur in realistic traffic situations if a very low sampling rate is used. In most simulations a sampling frequency of 10 Hz is maintained, which should be high enough to accurately approach the IDM’s continuous behavior. Even though it is unlikely, the possibility of a crash should not be overlooked when applying the IDM in a discretized system.

2-3-3 Action point model

The psycho-physical model, also known as the action-point model, describes car-following behavior on the basis of thresholds or action points. The idea is that drivers are unable to detect arbitrarily small changes in their stimuli, such as relative speed. Only when a certain threshold is reached, will a driver react by adjusting its own driving behavior, just as the addition of Tampére to the Helly model in (2-5). The regimes that are defined by these thresholds show in what situations the follower will react, which is shown in Figure 2-2, a figure adapted from [12]. An example of a psycho-physical car-following model is presented in the work of Wiedemann and Reiter[13].

2-3-4 Collision avoidance model

The collision avoidance model, as first introduced in [14], describes the relative distance between two vehicles rather than the acceleration of the following vehicle. The model specifies
a safe following distance, required for the following driver to avoid a collision with the vehicle ahead. The definition of the safe following distance is given by

$$\Delta x(t - \tau) = av_{n-1}^2(t - \tau) + \beta v_n^2(t) + \beta v(t) + b_0$$

where the calibration parameters are given in $$\theta = \{\alpha, \beta, \beta_l, b_0\}$$.

The Gipps model, presented in [15], is based on the CA model and is a widely used model for microscopic traffic simulations.

### 2-3-5 Latent class model

The latent class model that is reviewed here, as proposed by Koutsopoulos and Farah in [16], is a modeling framework that separates longitudinal driving behavior into different driving regimes, such as car-following, free-flowing, emergency braking, etc. The decision making of the driver in each regime is based on different stimuli from the surrounding environment. For instance in the car-following regime, the driver makes a decision to either accelerate (A), decelerate (D), or do nothing (DN) based on the stimuli coming from its leading vehicle(s).

The decisions are probabilistic (unlike the other models described in this thesis) matching the probabilistic driving behavior of regular traffic. The possible states of action in the free-flow regime are the same as in the car-following regime, but the stimulus can be different, such as the driver’s own desired speed [17]. The emergency regime can be assumed to consist of only one state, as deceleration is applied to avoid a collision. An illustration showing the model framework is seen in Figure 2-3, adapted from [16].

![Figure 2-3: Latent class model structure](image)

A driver can be in one of three discrete states: deceleration, do-nothing, or acceleration. This is unlike the other models described, which describe the acceleration behavior as a continuous function. The state of the driver is latent, it can not be observed. The only thing that is observed is the resulting action of the driver, expressed in the acceleration of the vehicle. It is possible for the driver to cause some acceleration/deceleration while being in the DN state. This is due to the difficulty the driver has keeping a constant speed, which requires
very accurate control of the accelerator, and the lack in ability of the driver to notice these small changes. An example of a probability distribution of each state is shown in Figure 2-4, indicating the probability of the value of acceleration corresponding to each state \( m \), where \( m = \{D, DN, A\} \).

![Figure 2-4: Example of acceleration distribution per state](image)

The state choice of driver \( i \) is modeled in the form of a discrete choice problem. The state the driver chooses is the state for which he maximizes his observed utility \( U \), where the utility of each state is dependent on the stimuli available to the driver. The utility of each state is defined as

\[
U_{DN}^i(t) = \epsilon_{DN}^i(t)
\]

\[
U_{A}^i(t) = V_{A}^i + \epsilon_{A}^i(t) = X_{A}^i \cdot \beta_{A} + \epsilon_{A}^i(t)
\]

\[
U_{D}^i(t) = V_{D}^i + \epsilon_{D}^i(t) = X_{D}^i \cdot \beta_{D} + \epsilon_{D}^i(t)
\]

where \( V_{i}^t(t) \) is the estimated utility at time \( t \) for driver \( i \) in each state \( (A, D, DN) \). \( \epsilon_i \) represents the random error term of driver \( i \) in choosing the correct state. \( \beta_i \) is a vector of parameters corresponding to the stimuli \( X_i^m(t) \). The choice of relevant explanatory variables depends on the state, as explained earlier. It is assumed that the error terms follow an independent and identically distributed (i.i.d) Gumbel distribution. The decision problem can be seen as a multinomial logit choice model, where the probability of each state for driver \( i \) is defined as

\[
P_{i}^{m} = \frac{e^{V_{i}^{m}}}{\sum_{m} e^{V_{i}^{m}}}
\]

Since the traffic situation is dynamic, drivers will continually re-evaluate their decisions, which is consistent with the notion of "partial short-term plan"[18], in which a driver executes a step,
and then re-evaluates the situation before deciding on the next step. The state of the driver
determines which acceleration model is used to describe the vehicle dynamics. The latent
class model allows for a very high degree of modeling freedom, as each state of each regime
can be modeled using a different acceleration model. The choice of acceleration model will
not be described, but it can be any acceleration function.

2-4 Model choice

In order to narrow down the amount of experiments needed for the research, two of the
presented models are selected to be used in the research. The first model is the IDM, described
in Section 2-3-2. This nonlinear model is already being used in the simulator and therefore
does not have to be implemented. By using this model, a comparison of updating the IDM
with the stock IDM is easy to perform. The main reason for choosing this model is that it is
already implemented in the simulator. The second model that will be used for the experiments
is the GM model, which is described in Section 2-2-1. In contrast to the IDM, which is a
nonlinear model with five model fitting parameters, the GM model is a linear model with just
one model fitting parameter. The implementation of the model in the software is relatively
simple as all information needed for the model is already available for the implemented IDM.
The main reason for choosing this model is the contrast with the IDM, which suits the
experiments well.

2-5 Chapter summary

In this chapter, different types of car-following models are discussed. The notation used in
the car-following models is introduced, and the state-definition of the car is explained as it
will be used during the experiment. An overview of some linear models and a few nonlinear
models is given, from which a choice is made for the experiments. Finally, the model selection
for the experiments is motivated.
Bayesian inference is a statistical method of updating prior probability estimates of a hypothesis using additional evidence. In the calibration of car-following models, prior information on the model parameters can be used when applying Bayesian inference as a calibration method. The prior information on the parameters can be utilized to rule out unrealistic estimation results, which may arise due to the lack of information in the data. For instance, a parameter represents desired free driving speed in a model, which is hard to identify using data that only contains observations of low speed congested driving behavior. The output of the Bayesian method represents a probability distribution of driving behavior, which can be used in a simulation to match the variation of driving behavior in a real-life situation.

A Bayesian approach to calibrating car-following models is proposed in [2]. The method makes use of Bayes’ rule to determine a probability distribution for the modeling parameters, based on a data set $D$ used for calibration. This probability distribution can be directly used to create a distribution in driving behavior in a simulation. Bayes’ rule can be used to express the posterior distribution of the parameters $\theta$, given the data $D$, or $\Pr(\theta \mid D)$ as

$$\Pr(\theta \mid D) = \frac{\Pr(D \mid \theta) \Pr(\theta)}{\Pr(D)}$$

(3-1)

where $\Pr(\theta)$ is the prior probability distribution of the parameter set $\theta$, in which prior information or knowledge on the possible values of the parameters can be contained, $\Pr(D \mid \theta)$ is the likelihood, indicating the compatibility of the data set $D$ with the hypothesis $\theta$, and $\Pr(D)$ is the normalization factor.

A Gaussian distribution is used to represent the prior probability density function of the parameters, as the resulting posterior distribution can then be represented analytically. This reduces the calculation time, as the result is not obtained using a numerical method which would require more computational effort. The prior probability can be constructed, using the known prior mean $\bar{\theta}$ and covariance matrix $\Sigma_{\text{prior}}$, as
\begin{align*}
\Pr(\theta) &= \exp \left( -\frac{1}{2}(\theta - \bar{\theta})^T \Sigma_{prior}^{-1}(\theta - \bar{\theta}) \right) / (2\pi)^{N/2}|\Sigma_{prior}|^{1/2} \tag{3-2} \\
\end{align*}

where \( N \) is equal to the number of parameters that are used in the model.

The likelihood is defined as

\begin{align*}
\Pr(D|\theta) &= \exp \left( -\frac{1}{2\sigma_l^2} \sum_{k=1}^{K} (v_p(k, \theta) - v_o(k))^2 \right) / (\sigma_l^2 2\pi)^{K/2} \tag{3-3} \\
\end{align*}

where \( v_p(k, \theta) \) is the predicted vehicle speed at time instance \( k \), based on a car-following model and the corresponding parameter set \( \theta \). \( K \) is the number of observations contained in the data set \( D \) and \( v_o(k) \) is the observed speed at time instant \( k \).

### 3-1 Calculation of posterior distribution

The expression for the probability of the posterior is obtained by substituting (3-2) and (3-3) into (3-1), resulting in

\begin{align*}
\Pr(\theta|D) &= \frac{1}{Z_p} \cdot \exp \left( -\frac{1}{2\sigma_l^2} \sum_{k=1}^{K} (v_p(k, \theta) - v_o(k))^2 \right) \cdot \exp \left( -\frac{1}{2} (\theta - \bar{\theta})^T \Sigma_{prior}^{-1}(\theta - \bar{\theta}) \right) \tag{3-4} \\
\end{align*}

where \( Z_p \) is a constant defined as

\begin{align*}
Z_p &= \sigma_l^K |\Sigma_{prior}|^{1/2}(2\pi)^{(K+N)/2} \cdot \Pr(D) \tag{3-5} \\
\end{align*}

The distribution defined in (3-4) can also be described using \( \theta^{MP} \) and \( \Sigma_{post} \), denoting the most probable parameter set and the corresponding covariance matrix respectively, knowing that the distribution is Gaussian shaped. We know this, because it is the result of a multiplication of two Gaussian distributions and therefore it is also a Gaussian distribution.

The most probable parameter set of the posterior distribution, \( \theta^{MP} \), can be found by maximizing the logarithm of (3-4)

\begin{align*}
\left( \theta^{MP} \right) &= \arg \max_{\theta} \ln \left( \Pr (\theta|D) \right) \tag{3-6} \\
&= \arg \max_{\theta} -E (\theta) \tag{3-7} \\
&= \arg \min_{\theta} E (\theta) \tag{3-8} \\
\end{align*}

where \( E(\theta) \) is defined as

R.B. Molenaar

Master of Science Thesis
$E(\theta) = K \ln(\sigma_l) + E_p(\theta) + E_l(\theta)$  \hfill (3-9)

in which the error in the likelihood, denoted as $E_l$, and the error in the parameter set, denoted as $E_p$, are defined as:

$$E_p(\theta) = \frac{1}{2} (\theta - \bar{\theta})^T \Sigma^{-1}_{prior} (\theta - \bar{\theta})$$  \hfill (3-10)

$$E_l(\theta) = \frac{1}{2\sigma_l^2} \sum_{k=1}^{K} (v_p(k, \theta) - v_o(k))^2$$  \hfill (3-11)

The terms $Z_p$ and $|\Sigma_{prior}|^{1/2}$ have not been included in (3-9) as they have no influence on the solution of (3-8). As $E(\theta)$ is a convex function for both models, proof of which is given in Appendix F, the minimization of (3-8) is found by solving the following condition:

$$\nabla_{\theta} E(\theta) = \Sigma_{prior}^{-1} (\theta - \bar{\theta}) + \nabla_{\theta} E_l(\theta) = 0$$  \hfill (3-12)

For describing the distribution of the posterior, the standard deviation of the likelihood function, denoted by $\sigma_l$, needs to be found. This is done by setting the derivative $\delta E / \delta \sigma_l$ to zero, defining $\sigma_l^2$ as

$$\sigma_l^2 = \frac{1}{K} \sum_{k=1}^{K} (v_p(k, \theta) - v_o(k))^2$$  \hfill (3-13)

The covariance matrix of the posterior distribution, $\Sigma_{post}$, is found using the Cramér-Rao bound, which will be described in the next section.

### 3-2 Cramér-Rao bound

The Cramér-Rao lower bound (CRLB) [19] is a theorem which defines the minimum estimation error, which defines a lower bound on the variance of estimators of deterministic parameters. The equation of the CRLB is given by

$$\Sigma(\theta) \geq \frac{1}{J_\theta}$$  \hfill (3-14)

where $J_\theta$ is the Fisher information defined by

$$J_\theta = -E \left[ \frac{\delta^2}{\delta \theta^2} \log f(\theta) \right]$$  \hfill (3-15)

The Fisher information shows the proportional rate of change in the likelihood function as the parameter set $\theta$ changes. This rate of change can be used as a measure for the variance of the parameter, as a rapid decrease in the likelihood function for a small change in a certain parameter indicates a high certainty (and therefore low variance) of that estimated parameter value.
One of the assumptions for the CRLB to hold is that the parameter space is an open subset of $\mathbb{R}^n$[20]. However, parameter constraints are required to guarantee the correct application of the car-following models. When parameter constraints are applied, forcing the parameters to lie within a closed subset of the original parameter space, the CRLB no longer applies. The application of the constraints changes the allowable parameter variations and therefore also the local structure of the log-likelihood function. A method is found in literature [21] to deal with the addition of constraints to the parameter space, a necessity for this application. This method is presented next.

### 3-2-1 Constrained Cramér-Rao bound

Let $\Theta_C \subset \mathbb{R}^n$ denote the constrained parameter space, in which the solution of (3-8), $\theta^{MP}$, is found. The constraint set $\Theta_C$ consists of regular points and non-regular points. The parameter set $\theta$ is considered to be a regular point of $\Theta_C$ if there are no active equality constraints in that point. Furthermore, the parameter set $\theta$ is considered to be a nonregular point if it is on one of the boundaries of $\Theta_C$ where at least one of the constraints is active.

Let $G_\theta$ denote the set of equality constraints active at the point $\theta$, where $G = [G^1, \cdots, G^k]^T$, where the vector functions are continuously differentiable. $\nabla G_\theta$ is the $k \times n$ gradient matrix of the function $G_\theta$, where $k$ is the number of linear equality constraints active at the point $\theta$.

The error covariance matrix $\Sigma$ of the constrained estimator satisfies the matrix inequality

$$\Sigma \succeq B_c$$

where

$$B_c = Q_\theta J_\theta^{-1}$$

in which $J_\theta$ is as defined in (3-15) and $Q_\theta$ is the $n \times n$ matrix given by

$$Q_\theta = I - J_\theta^{-1} [\nabla G_\theta]^T \left\{ [\nabla G_\theta] J_\theta^{-1} [\nabla G_\theta]^T \right\}^+ [\nabla G_\theta]$$

(3-18)

where $\{\ldots\}^+$ denotes the pseudo-inverse of the term in between brackets.

When the point $\theta$ is a regular point, i.e. no active equality constraints, (3-18) reduces to $Q_\theta = I$, since $\nabla G_\theta = 0$, showing that the covariance matrix of a constrained estimator (3-17) at a regular point is equal to the covariance matrix of an unconstrained estimator, as given in (3-14).

### 3-2-2 Application of Cramér-Rao bound

The covariance matrix of the posterior distribution can be calculated using the CRLB. The method used to calculate the lower bound for the constrained estimator returns the same results an unconstrained estimator in situations where no constraints are active. This means that the implementation of the method presented in Section 3-2-1 suffices to handle the distributions of both the regular and irregular points.

The Fisher information, (3-15), is obtained by calculating the Hessian of (3-9), giving
\[ J_\theta = \nabla^2_\theta E(\theta) = \Sigma_{\text{prior}}^{-1} + \nabla^2_\theta E_l(\theta) \] (3-19)

The derivation of \( \nabla^2_\theta E(\theta) \) is done for the GM and IDM in Section 4-2-1 and Section 4-2-2 respectively, as the derivation depends on the model equation.

The covariance matrix \( \Sigma_{\text{post}} \) corresponding to the obtained posterior distribution can then be calculated using the CRLB (3-14) as

\[ \Sigma_{\text{post}}(\theta^{MP}) = -(J_{\theta \theta}^{MP})^{-1} \] (3-20)

### 3-3 Model comparison

The Bayesian inference framework can also be used to compare different models, for which a method is presented in [22]. Let \( H_n \) denote a hypothesis \( n \), where each hypothesis represents a model \( n \) with the set of assumptions that are made for that model. A method for comparing two different hypotheses, \( H_1 \) and \( H_2 \) given the same dataset \( D \), is described next.

By making use of Bayes' theorem, the probability of model \( H_1 \) given the dataset \( D \), \( \Pr(H_1|D) \), can be related to the predictions about the data made by the model, \( \Pr(D|H_1) \), and the prior probability of theory \( H_1 \), \( \Pr(H_1) \), using

\[ \Pr(H_1|D) = \frac{\Pr(D|H_1) \Pr(H_1)}{\Pr(D)} \] (3-21)

Each model \( H_i \) has a vector of parameters \( \theta \). Each model is defined by two probability distributions. The first is the prior distribution \( \Pr(w|H_i) \) that shows what values the parameters are expected to have. The second is the conditional distributions \( \Pr(D|\theta,H_i) \) that represent the predictions the model \( H_i \) makes about the data \( D \), for each value of \( \theta \).

The posterior probability of the parameters \( \theta \) can be calculated using

\[ \Pr(\theta|D,H_i) = \frac{\Pr(D|\theta,H_i) \Pr(\theta|H_i)}{\Pr(D|H_i)} \] (3-22)

which can be read as

\[ \text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}} \] (3-23)

The maximum of the posterior shows the most probable values for the parameters, from now on denoted as \( \theta_{MP} \). Gradient-based methods can be used to find the maximum, but it can also be approximated by evaluating the Hessian at \( \theta_{MP} \), \( A = -\nabla^2 \log \Pr(\theta|D,H_i) \mid \theta_{MP} \), and taking the Taylor series expansion of the log posterior probability

\[ \Pr(\theta|D,H_i) \approx \Pr(\theta_{MP}|D,H_i) \exp(-\frac{1}{2} \Delta \theta^T A \Delta \theta) \] (3-24)

The probability density function of a multivariate Gaussian distribution can be written as \( \exp -\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) \) where \( c \) is a constant and \( \Sigma \) is the covariance. We can therefore conclude that the posterior distribution \( \Pr(\theta|D,H_i) \) can be locally approximated as a normal distribution with a covariance matrix \( A^{-1}[22] \).
The posterior probability of model $i$ is given by

$$
Pr(H_i|D) \propto Pr(D | H_i) Pr(H_i)
$$

where $Pr(D | H_i)$ is the evidence for theory $H_i$ and $Pr(H_i)$ is the subjective prior information, expressing how plausible the theories are before the acquired data is considered. By assigning equal priors $Pr(H_i)$, all theories $H_i$ are ranked by evaluating the evidence, giving

$$
Pr(H_i|D) \propto Pr(D | H_i)
$$

which appeared before as the normalizing constant in (3-1). The evidence $Pr(D | H_i)$ is a transportable quantity for comparing alternative models, independent of the model type or structure. Since $\int (Pr(w|D, H_i))dw = 1$, (3-1) can be rewritten as

$$
Pr(D | H_i) = \int Pr(D | \theta, H_i) Pr(\theta | H_i)d\theta
$$

Determining the value of (3-27) requires integration over the entire parameter space. Analytical computation of this value will require too much computational effort for the algorithm to run parallel to a traffic simulator, especially for models with larger parameter sets. The posterior distribution commonly has a strong peak around the most probable value for the parameters, $\theta_{MP}$. For single parameter models, the value of (3-27) can be approximated by multiplying the height of the peak of the approximated integrand $Pr(D|\theta, H_i) Pr(\theta | H_i)$ by the width of the peak $\sigma_{\theta|D}$, as done in [22]:

$$
\text{Evidence} \approx \text{Best fit likelihood} \times \text{Occam factor}
$$

When calibrating a model with multiple parameters, the Occam factor can be found from the determinant of the corresponding covariance matrix:

$$
Pr(D | H_i) \approx Pr(D | \theta_{MP}, H_i) \times Pr(\theta_{MP} | H_i) \det^{-1/2}(A(\theta_{MP})/2\pi)
$$

where $A = -\nabla^2 \log Pr(\theta|D, H_i) |_{\theta_{MP}}$, the same as in (3-24).

The model that predicts the data the best will have the highest best fit likelihood. The performance of each model is penalized by the Occam factor, which can have a value between 0 and 1, depending on the amount of parameters the model has. The Occam factor is named after Occam’s Razor [23] and is implemented to favor the model that uses the least amount of assumptions and entities to explain a certain phenomenon. A model with more parameters will have a lower Occam factor compared to a model with less parameters and a similar best fit likelihood, and therefore a lower evidence. The evidence for a model is therefore promoting a good fit as well as simplicity, thereby preventing over fitting.
In this chapter, the research strategy of this thesis is presented. Each research question is discussed, stating the hypothesis for that question and the methods that will be used to test this hypothesis. First, the overall methodology is discussed, in which the outline of the research is given. The application of the Bayesian method, described in Chapter 3, explains how the method is applied in the experiments, after which each research question is addressed.

4-1 Overall methodology

The algorithm that will be used to identify the observed driving behavior needs to be implemented into the existing software structure, so that it can run integrated with the traffic simulator. The implementation of the algorithm is tested before performing the experiments with human interaction, by providing the algorithm synthetic data as input, generated using known models and parameters. Noise is added to the synthetic data, in order to test the ability of the algorithm to deal with observation noise and process noise. The implementation is considered successful if the algorithm is able to find the same models and parameters, that were used to create the synthetic data set.

After the functionality of the algorithms is tested with the easiest form of identification, that of data which has been created with the same models that the algorithm tries to match the observed behavior with, the algorithm will be subjected to data taken from the simulator, where the ability of the algorithm to deal with actual measurements of human driving behavior will be analyzed. This step is significant compared to the first step, as this situation should prove that the models used for fitting the human driving behavior too, will actually be able to properly fit the observed data. When these tests prove successful, the algorithm should be able to identify the driving behavior of participants to the simulator, provided their driving behavior is varied enough to identify different types of driving behavior.

The next step is to find a way to deal with the lack of information on certain parameters because of the driving behavior displayed in the traffic simulator. The main idea is to make sure different types of driving behavior are observed during a simulation run, by adjusting
the environment and variation in traffic density in the simulator. In this specific case, as described in Section 1-4, the environment in the simulator already stimulates different types of driving behavior. On the long back straight, the main regime of driving is expected to be free driving. The other half of the loop, where the traffic is frequently interrupted by traffic lights, is expected to see more congested driving.

The test subjects participate in three driving sessions, where the results of the driving behavior identification is applied differently in each session. After every session the participants fill out a questionnaire, in which they grade the longitudinal driving behavior of the simulated traffic. The answers provided by the participants are used to evaluate the subjective validity of the observed driving behavior. From this information, the research questions can be answered.

4-2 Application of Bayesian method

In the previous sections the Bayes identification algorithm is described. This algorithm is used to fit the observed data to the different available models and compares the evidence for each model. The evidence of a car-following model, used in the Bayes algorithm, is directly derived from the equation that describes the car-following behavior. The models that are used in this research are the GM model and the IDM model, as introduced in Chapter 2. Both are stimulus-response models with a delay in the response, representing the reaction time of the driver. The evidence for each model is derived separately in Section 4-2-1 and Section 4-2-2.

4-2-1 Evidence for the GM model

We start by deriving the evidence for the GM model, of which a more detailed description and its properties is given in Section 2-2-1. The GM model describes the acceleration of a vehicle, delayed by a reaction time $\tau$, as a function of the relative speed to the vehicle leading it:

$$a(t,\theta) = \alpha \Delta v(t - \tau)$$  \hspace{1cm} (4-1)

where $a(t,\theta)$ is the acceleration at time $t$, dependent on the speed difference between the current car and its leader $\Delta v$ and a parameter set $\theta$. The parameter set used to describe the acceleration in the GM model is defined as

$$\theta = \{\alpha\}$$  \hspace{1cm} (4-2)

The prediction of the speed at time $t$ is calculated as

$$v_p(t, \theta) = v_o(t - \Delta t) + a(t - \Delta t, \theta)\Delta t$$  \hspace{1cm} (4-3)

Four terms are introduced to shorten notation: $v_p$, $v_q$, $v_r$ and $v_s$, which are defined as
\[ v_p = v_o(k - \Delta t) \]
\[ v_q = \Delta v_o(k - \Delta t - \tau) \]
\[ v_r = v_o(k - \tau) \]
\[ v_s = v_o(k) \]  

Using (4-3) and the introduced short notations (3-9), the GM model becomes

\[ E(\alpha) = K \ln (\sigma_l) + \frac{1}{2} (\alpha - \bar{\alpha}_{\text{prior}})^T \Sigma_{\text{prior}}^{-1} (\alpha - \bar{\alpha}_{\text{prior}}) + \frac{1}{2\sigma_l^2} \sum_{k=1}^{K} \left( v_p + \alpha \Delta t v_q - v_s \right)^2 \]  

The evidence for the GM model is analytically derived by computing the gradient, as defined in (3-12), resulting in

\[ \frac{\delta E(\alpha)}{\delta \alpha} = \frac{1}{\sigma_{\text{prior}}^2} (\alpha - \bar{\alpha}_{\text{prior}}) + \frac{\Delta t}{\sigma_l^2} \sum_{k=1}^{K} v_q \left( v_p + \alpha v_q \Delta t - v_s \right) \]  

where \( \bar{\alpha}_{\text{prior}} \) and \( \sigma_l^2 \) are the mean and variance of the prior parameter distribution. The Hessian of the evidence function of the GM model, introduced in (3-19), is given by

\[ \frac{\delta^2 E}{\delta \alpha^2} = \frac{1}{\sigma_{\text{prior}}^2} + \frac{\Delta t^2}{\sigma_l^2} \sum_{k=1}^{K} v_q^2 \]  

In order to calculate the evidence, the most probable parameter set \( \theta^{MP} \) is required, which is the solution of (3-12). The solution of (3-12) is found by numerical computation, by utilizing the Nelder-Mead algorithm [24]. The most probable parameter set is then used in (3-13) to calculate the value of \( \sigma_l^{MP} \).

With these results, (3-2), (3-3) and (3-19) can be calculated. All this information is then substituted into (3-29), which gives us the evidence of the GM model.

### 4-2-2 Evidence for the IDM

The evidence for the IDM, a model described in more detail in Section 2-3-2, is derived in a very similar way compared to the evidence derivation of the GM model in Section 4-2-1. The acceleration function of the IDM, with a reaction time \( \tau \), is described as

\[ a(t, \theta) = \alpha \left[ 1 - \left( \frac{v_r}{v_0} \right)^4 - \left( \frac{s^*(t - \tau)}{s(t - \Delta t - \tau)} \right)^2 \right] \]  

where

\[ s^*(t - \tau) = s_0 + Tv_r + \frac{v_r v_q}{2\sqrt{\alpha \beta}} \]
The prediction of the speed at time $t$ is then calculated as

$$v_p(t) = v_o(t - \Delta t) + \alpha \left[1 - \left(\frac{v_r}{v_0}\right)^4 - \left(\frac{s^*(t - \tau)}{s(t - \Delta t - \tau)}\right)^2\right] \Delta t$$

(4-10)

The error function of the IDM, defined in (3-9), that follows from these equations is

$$E(\theta) = K \ln (\sigma_l) + \frac{1}{2} \left(\theta - \tilde{\theta}_{\text{prior}}\right)^T \Sigma_{\text{prior}}^{-1} \left(\theta - \tilde{\theta}_{\text{prior}}\right) + \frac{1}{\sigma_v^2} \sum_{k=1}^{K} \left(v_p + \alpha \left[1 - \left(\frac{v_r}{v_0}\right)^4 - \left(\frac{s^*(t - \tau)}{s(t - \Delta t)}\right)^2\right] \Delta t - v_p\right)^2$$

(4-11)

The derivation of the Jacobian matrix, defined in (3-12), and Hessian matrix, as defined in (3-19) are performed in order to obtain the results needed for calculating the evidence, but the results are omitted here as the equations are too long to be displayed here. The results are available on request. After this, the same procedure is followed as with the GM model to calculate the evidence for the IDM.

### 4-2-3 Reaction time implementation

A necessary addition to the existing method is the implementation of a variable that represents the reaction time of drivers. For every driver, an average reaction time can be described, which models the time that elapses between receiving the stimulus and reacting on it. The simulator provides information, such as speed and position, every 0.1 seconds. The reaction time that can be found from the provided data can therefore only be determined to the nearest tenth of a second, and will be implemented as

$$a(t, \theta) = f(t - \tau, \theta)$$

(4-12)

where the acceleration at time $t$ will be determined by the acceleration function $f(t - \tau, \theta)$, using the data available to the driver at time $t - \tau$.

As the reaction time $\tau$ only appears as an index in the models, it is impossible to differentiate the model equations with respect to $\tau$. Unlike the other parameters, the reaction time can therefore not be used as a variable in the Bayesian inference method. Instead, the algorithm will run for each reaction time, computing the evidence each time. The implementation will lead to a variable sample length over the different reaction times, this is solved by performing the identification over a subset of the dataset defined as

$$D_{\text{sub}} = D(k)$$

(4-13)

where $k = \{n_\tau, n_\tau + 1, \ldots, (N - n_{\text{max}}) + n_\tau\}$ is the interval over which the algorithm will perform the identification, in which $N$ is the length of the original dataset $D$, $f$ is the frequency of the simulator, $n_\tau \cdot f$ is the current reaction time considered, and $n_{\text{max}} \cdot f$ is the maximum reaction time that the algorithm will run for.
The possible values for $\tau$ can be adjusted in the algorithm. In this case, a maximum of 2 seconds is chosen as the highest reaction time the algorithm will try. The interval is defined as $\tau = [0, 0.1, 0.2, \ldots, 2.0]$, so the algorithm will run 21 times for every model with each dataset it receives.

4-2-4 Adaptation of simulation environment

A set of models is defined, from which the simulated will randomly select a model to determine their longitudinal driving behavior. The parameters are drawn from the probability density function, which models the variety in driving behavior within the model. The initial definition of the models and parameter distributions that is used in the experiments is described in Section 5-5-3, but can be chosen differently to change the influence of the feedback. The adaptation of the simulation environment is achieved by updating this set of models. As described in Section 4-2, the Bayesian algorithm will calculate which model with parameter set and reaction time has the highest evidence for the provided dataset. This model, paired with the reaction time, will be added to the set of available models from which the simulated cars randomly draw their model with reaction time. The parameter distribution of the best fitting model will be updated with the posterior distributions found from the Bayesian algorithm. The simulated cars update their car-following model after every update to the model set, which means that the results of the identification are immediately affecting the simulator environment. The main code to implement the Bayesian algorithm with feedback loop in the simulator is presented in Appendix B.

4-3 Testing performance of adaptation algorithm

Before applying the implementation of the Bayesian algorithm in the simulator, the performance is tested. The algorithm is first tested with synthetic data, created with known models and parameter distributions, after which an experiment is done with real data obtained from the simulator.

4-3-1 Experiments with synthetic data

A data set is created using either of the two selected car-following models and a set of known parameters. This 'synthetic' data set can be used to evaluate the ability of the algorithm to identify the model and accompanying parameter set that best describes the longitudinal driving behavior data that is provided. This experiment is the first step in analyzing the performance of the algorithm, testing it with a controlled input. The use of synthetic data makes it possible to add very specific disturbances, allowing to test the influence of these disturbances on the algorithm, while keeping all other factors constant (ceteris paribus). The research question that is being investigated by performing these experiments is:

**Research question 1:** Is the Bayesian inference method able to correctly identify the set of models and parameters that are used to create a synthetically generated data set?

**Hypothesis 1:** Yes, it is possible to identify a set of models and parameters that are used to generate a data set.
As a very similar algorithm has already been applied to real-world traffic [2], it is expected that the method (implemented as described in Chapter 3) will be able to achieve correct predictions for generated data sets. The data set is created using the same models that are available to the algorithm. This experiment will prove that, if a driver’s behavior can be described using one of the available models with a certain parameter set, the algorithm will be able to correctly identify this. The conditions, under which this statement is correct, will be tested in this experiment. The effect of the data quality will be evaluated in this experiment, as different forms of noise can be added to the synthetic data.

4-3-2 Experiments with real data

In this experiment, the algorithm will be run using a data set obtained from a previous experiment with users, of which the results are presented in [25]. The ability of the algorithm to deal with actual measurements of human driving behavior, in a similar environment, will be analyzed. This is a significant step compared to the first experiment, as this experiment should prove that the algorithm will be able to properly fit the observed data to the longitudinal driving models that are available to it. When this experiment proves successful, the algorithm should be able to identify the driving behavior of the participants in the following experiments and provide this information to the simulator, where the generated AI traffic will be adjusted to the identified models and corresponding parameter sets.

Research question 2: Is the Bayesian inference method able to identify a model and parameter set that best fits the driving behavior that is observed from a human participant in the traffic simulator?

Hypothesis 2: Yes, it will be able to identify a model and matching parameter set from observed driving behavior.

As the implementation of the method has already proved to work with the data structure that the simulator provides, and that it works with real-world traffic [2], it is hypothesized that the proposed Bayesian method, implemented as described in Chapter 3, will be able to identify a model and matching parameter set based on a data set describing human car-following behavior. The accuracy of this prediction will be based on the length of the data set. If a shorter data set is used for learning, the implemented algorithm will have a new model set available sooner. However, if the data set is not long enough so that it contains all possible different driving regimes described by the model, a problem first addressed in Section 5-5-2, then it is possible that the model will be calibrated incorrectly.

4-4 Testing subjective validity of adaptive bots

In this section, the experiments with user-interaction are described. The focus of this thesis is to look at the effects of updating car-following behavior on the subjective validity of the simulated traffic. Before comparing the subjective validity of the updated and non-updated traffic, a method is defined to help identify the change in subjective validity and to maximize the information gained from the experiments.
4-4 Testing subjective validity of adaptive bots

<table>
<thead>
<tr>
<th>Session</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage updated cars</td>
<td>0%</td>
<td>50%</td>
<td>100%</td>
</tr>
<tr>
<td>Fraction green/orange</td>
<td>50/50</td>
<td>50/50</td>
<td>50/50</td>
</tr>
<tr>
<td>Expected difference answer</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 4-1: The 3 different sessions performed in the color-based analysis of the experiment.

4-4-1 Identifying a change in subjective validity

Research question 3: How can an improvement in subjective validity be identified/proved?
To identify a relative change in subjective validity, the updated and non-updated traffic need to be compared. Each participants will drive in three sessions after which their experiences will be recorded in a questionnaire, asking them about their opinion on the validity of the surrounding traffic. Three different session will be driven where, in each session, 50% of the cars are green colored and the other 50% are orange colored. Dividing the simulated traffic into two groups allows the application of different model sets to the different groups, which means two different models can be compared in a single simulator session.

Let the three different sessions be known as session A, B and C. The order of the three sessions is chosen randomly, to prevent learning effects to have a systematic influence on each of the sessions and to prevent other unknown influence of session order from showing up in the experiment results.

In session A, none of the cars will have their model parameters updated. The simulated cars will drive according to the IDM model that is currently implemented in the traffic simulator, as discussed in Section 1-4. In session B, one of the two groups of cars will have their model parameters updated, where the other cars’ longitudinal driving behavior will still be determined by the aforementioned IDM model. In session C, all of the cars will have their model parameters updated by the proposed algorithm. An overview of the different sessions is given in Table 4-1. The advantage of using such an experiment setup is that two different research questions can be answered.

4-4-2 Experiments with user interaction

The last two research questions will both be investigated by performing a single experiment with human participants, in which each participant will control a car in the simulated environment. The direct comparison is discussed first, after which the indirect comparison is covered. The effects are expected to be more noticeable in a direct comparison and the increase of subjective validity should therefore be easier to prove.

Research question 4: Does updating the driving models of AI cars, by feeding back information gained from observing human driving behavior in the same simulation session, lead to such an improvement of the subjective validity of that driving behavior that it is noticeable in a direct comparison?

Hypothesis 4: Yes, the subjective validity is expected to increase by feeding back the observed driving behavior.

The extra available information about the simulator situation (real driving behavior of humans in that specific situation) can be used to make the simulated traffic behave more like a
human would in such a situation. This increase in realism of the simulated traffic should have a positive effect on the behavioral validity of the driving behavior displayed by people participating in the simulation. Changes of human driving behavior over time can also be identified and matched in the simulated traffic, which should add to the subjective behavioral validity. The effect of updating the driving behavior of some of the AI cars is expected to be significant enough to notice, when comparing updated and non-updated AI cars side by side in the same traffic situation.

For this research question session B is considered, in which both the updated and non-updated AI cars can be observed side by side. By comparing the answers given by the participants, the difference in subjective validity between the updated AI cars and the non-updated AI cars can be found. It is expected that, if participants will detect any difference between the two cars, then it will show up first in this experiment. Having both driving styles next to each other allows for a direct comparison, making it easier to distinguish between the two of them. The same comparison between the green and orange cars will be made for session A and session C separately, in which the expected answer will be that there is no difference in subjective validity between the cars, as there is no actual difference in their driving behavior. The last row of Table 4-1 shows the hypotheses of Research Question 3 for each session.

To analyze whether the change in longitudinal driving behavior also leads to a noticeable difference across two separate sessions is investigated through the second research question coupled with this experiment:

**Research question 5:** Does updating the driving models of AI cars, by feeding back information gained from observing human driving behavior in the same simulation session, lead to such an improvement of the subjective validity of that driving behavior that it is noticeable when comparing over different sessions?

**Hypothesis 5:** Yes, considering the expected outcome of research question 4, and assuming the situations over which the comparison is made are identical.

The effect of updating the driving behavior of the AI cars with human driving behavior is expected to be significant enough to notice for participants, when comparing the driving behavior of updated AI cars in one session with non-updated AI cars in another session. Any change in the simulated environment, or any other factor that is of influence on the human driving behavior, should diminish the effect of the algorithm, as it is likely to overshadow the effect of the update algorithm.

To answer this research question, a comparison is made between session A and session C. The subjective validity of the longitudinal driving behavior without updates on the parameters and models, measured in session A, is compared with the subjective validity in session C, where the information gathered from the participants is used to update the distribution of models and corresponding parameter distributions. The hypothesis for this research question is that the users will be able to distinguish between them, as the car in front of them has a great influence on the participant’s driving behavior and a lot of interaction will take place. A change in model and/or parameter set, defining this driving behavior, is therefore expected to be distinguishable.

Research question 3 and 4 seem quite similar, but the difference is significant. If the difference in driving behavior can be appreciated when comparing over 2 different sessions, it will be possible to look at the difference in macroscopic traffic flow properties caused by adapting the
microscopic traffic models to the driving behavior observed from participants. If the difference is only noticeable in a direct comparison, the adapted and non-adapted cars will both be part of the same traffic situation. Both types of cars will influence the driving behavior of others, due to conflicts between the two types during the duration of the simulation.
5-1 Introduction

Experiments are performed as part of the research, in order to find answers to the research questions. The goal of performing the experiments is to see whether the designed algorithm functions as intended and to verify or falsify the hypotheses posed for each of the research questions. A good experiment design will aid in the performance of the experiment and in the usefulness of the information that is gathered from it. In this section, the experiment design is discussed. First, a short description of the current situation is given. The improvements that are achieved by implementing additions to the simulator will be analyzed by looking at the different results obtained from the current situation and the new situation with which the experiment is performed. The description of the current situation will be used as a reference in the evaluation of these experiment results.

5-2 Traffic simulator setup

The environment, in which the experiments with human interaction will be conducted, is described in this section. The system on which these experiments will be performed consists of a traffic simulator and a separate driving simulator for each person that is participating in the simulation. The traffic simulator simulates the artificial intelligence driven cars, AI cars, whose behavior is based on a set of models and parameters that is provided in the software, calculating their position and velocity every 100 ms. The traffic simulator receives the inputs from each of the driving simulators and calculates the next position and velocity of the corresponding user driven car in the simulation environment. All this information is passed to each of the driving simulators at every simulation step, giving the driving simulators the information they need to display the traffic surrounding the vehicle as well as the vehicle itself. An instance of the driving simulator is used for each participant, translating the information received from the traffic simulator to a visual image, displaying the current traffic situation. The input that the user provides to the driving simulator is fed back to the traffic simulator.
The hardware system consists of multiple PCs, communicating with each other through a local area network. The traffic simulator is running on one of the PCs, while the other PCs are running the driving simulators. The car in the simulator can be controlled using a keyboard or using a steering wheel and pedal setup that can be attached to the PC. A connection of driving simulators with the traffic simulator can also be achieved over wide area networks, allowing people in different geographical locations drive in the same environment simultaneously. The disadvantage of connecting through a wide area network is that a delay in communications is introduced, affecting the quality of the simulation experience for the participants. Therefore, all communication during the experiments is done through a local area network.

5-3 Questionnaire design

The purpose of the questionnaire is to evaluate the appreciation of the realism of car-following behavior from the experiences of people who are not familiar with the terms of traffic engineering. In this case, the goal is to evaluate how users rate the longitudinal driving behavior of the surrounding traffic. The questions in the questionnaire are designed with this in mind, asking the participants about their opinion of specific traffic situations that relate to longitudinal driving behavior.

Two different questionnaires are used in this research, the session questionnaire and the final questionnaire. The participants fill out the session questionnaire after each session, in which they will answer questions designed to analyse their subjective validity of the car-following behavior they observed around them. In the final questionnaire, the participants will be asked anonymous personal information, such as age, gender and driving experience. The final questionnaire also asks the participants to rate their experience with driving the car in the simulator and gives them an opportunity to state anything else they wish to add.

The longitudinal driving behavior is separated into four aspects, where each aspect is considered observable for a novice in the simulation. The four aspects are acceleration, following distance, speed and reaction time. The orange and green cars are both separately compared to real traffic in each of these four aspects, giving a total of eight questions. Each question in the session questionnaire can be answered by choosing one of 7 options, where the middle options indicates that the traffic in the simulator matches the real traffic.

5-4 Processing the questionnaires

A subjective validity score (SVS) is defined, in order to quantitatively compare the revealed preference of the observed car-following behavior. This score is a numerical representation of each observed driving behavior and will be used to compare the performance of different situations. The SVS is the sum of the scoring for each question, where the lowest score indicates the best performance. A more detailed description is given in the next paragraph.

The scoring given to each of these options is equal to the difference of the given answer and the value representing real life traffic. An example is shown in the sample question, below in (5-1), where the points are displayed in red. The points of the questions are added up for both types of traffic, representing the deviation of this type of traffic from real life traffic,
as graded by the participant. Statistical tests will be performed on these numbers, to see if participants valued either of the two types of traffic better than the other. The questionnaires are given in Appendix C.

<table>
<thead>
<tr>
<th>Grade according to the 1-7 scale below.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
</tr>
<tr>
<td>much slower</td>
</tr>
<tr>
<td>as fast as</td>
</tr>
<tr>
<td>much faster</td>
</tr>
</tbody>
</table>

Figure 5-1: Example question showing the points assigned to each answer

5-5 Assumptions

A few assumptions are made during the execution of these experiments, which are discussed in this section. The models, chosen to fit the observed driving behavior to, are discussed, as well as the choice of the length of the observation set. Finally, the assumptions and decisions made regarding the parameter sets are explained.

5-5-1 Observation models

IDM

The first model that was selected to be used in these experiments is the Intelligent Driver Model (IDM), which is described in detail in Section 2-3-2. The main reason for this selection is that the IDM is the current model that is being used to determine the car-following behavior of the AI cars. The non-linearity of the IDM makes the computation of the derivatives and Hessian more complex, but this is overcome by performing the symbolical computations with a computer. The first of two observation equations, used in these experiments, is

$$a(t, \theta) = \alpha \left[ 1 - \left( \frac{v_r}{v_0} \right)^4 - \left( \frac{s^* (t - \tau)}{s(t - \Delta t - \tau)} \right)^2 \right]$$ (5-1)

where

$$s^* (t - \tau) = s_0 + T v_r + \frac{v_r v_q}{2 \sqrt{\alpha \beta}}$$ (5-2)

GM

The second observation model chosen for this experiment is the General Motors (GM) model, of which a detailed introduction is given Section 2-2-1. This observation model is linear and chosen to contrast the first observation model, the IDM. The comparison of a simple model and a more complex model using Bayesian inference allows to evaluate the necessity of implementing a more complex model for fitting human driving behavior to. The acceleration equation of the GM model, and therefore the second observation equation, is given by

$$a(t, \theta) = \alpha \Delta v(t - \tau)$$ (5-3)
5-5-2 Observation interval

A time interval has to be chosen to define the length of each dataset that will be used as input for the algorithm. The length of the interval determines the properties of the feedback system. A shorter time interval means that the calculations can be performed in quicker succession and therefore lead to a quicker adaptation of the simulated traffic. However, the interval should be large enough to prevent the unobservability of some parameters. For instance, identifying the desired free driving speed parameter $v_0$ of the IDM is impossible from a data set that does not contain any observations of free driving. The length of the interval is chosen so that the participant will have been both in the area of the track with the traffic lights where accelerating and braking is stimulated, as well as in the upper part of the track where free driving can be observed. The length of the interval is set at 2000 data points, which is equal to a period of 200 seconds at a measurement interval of 0.1 seconds.

5-5-3 Parameters and constraints

The parameters determining the participants driving behavior are assumed to be subject to change over time, an assumption supported by recent publications in the field of microscopic traffic modelling. The evolution of the parameter set $\theta$ over time is modelled as a random walk, denoted as

$$\theta(t + \Delta t) = \theta(t) + \epsilon(t)$$ (5-4)

where $\epsilon(t)$ is a random noise signal defining the process noise, allowing the change of the parameters over time. As the parameter is modelled as a random walk, the probability distribution of the parameter is therefore given by a normal distribution.

The measurement noise in this experiment is considered negligible. As the experiment is performed with computers there is some measurement noise due to quantisation, but this will not have a significant influence in the presence of the process noise.

Initial parameter distributions

Prior distributions need to be assigned to the parameter sets of both models. The prior information on which these distributions will be based is taken from literature, as no prior experiments on parameter estimation has been performed on this traffic simulator. A choice has been made to use parameter distributions obtained from real-world experiments. The prior distribution of the GM model parameter set is based on the results published in [26]. The prior distribution is given by $N(\theta_{\text{prior}}, \sigma^2_{\text{prior}}) = N(0.3, 0.04)$.

The mean values of the prior distribution used for the parameter set of the IDM are based on the findings presented in [27] and adjusted to match the speed limit that is enforced in the simulation environment of 60km/h or 17m/s. The covariance matrix defined below is chosen largely based on intuition, as the results presented in literature did not contain useful
5-5 Assumptions

information. The prior distribution is defined as $N(\theta_{\text{prior}}, \Sigma_{\text{prior}})$ where

$$\theta_{\text{prior}}^T = (\alpha, \beta, T, s_0, v_0)^T = (1.4, 2.0, 1.5, 2.0, 17)^T$$  \hspace{1cm} (5-5)

$$\Sigma_{\text{prior}} = \begin{bmatrix} 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \end{bmatrix}$$  \hspace{1cm} (5-6)

The choices made in defining these initial parameter distributions are not all based on empirical evidence, as it was lacking in some aspects. Since no previous experiments regarding parameter fitting to car-following models have been conducted on the simulator that is used for these experiments, the distributions are based on experimental results of real-world studies. By choosing the observation interval sufficiently long, it is assumed that the system, consisting of the traffic simulator, the human driver and the feedback mechanism, is stable. The Bernstein-von Mises theorem[28] states that the posterior distributions for unknown quantities, the model parameters in this case, are effectively independent of the prior distribution once enough information about the unknown quantities is provided. The initial choice of the parameters will have a diminishing effect on the state of the system as time passes, which is required considering the unknown characteristics of human driving in this particular simulator.

Constraints

Even though participants are instructed to behave as they would in the real-world, some will still ignore these instructions and drive as though they are playing a video game. This requires constraints to be applied to the range of possible parameter values, in order to guarantee that the model will work as intended and to make sure that the system’s stability can guaranteed without supervision in case unrealistic driving behavior is observed. The constraints on both parameter sets are shown below

Constraints GM = $\begin{cases} 0.1 \leq \alpha \leq 1 \\ 0.8 \leq \alpha \leq 3 \\ 0.5 \leq \beta \leq 5 \end{cases}$

Constraints IDM = $\begin{cases} 0.3 \leq T \leq 5 \\ 0.5 \leq s_0 \leq 10 \\ 10 \leq v_0 \leq 25 \end{cases}$

The constraint for $\alpha$ is high, relative to the constraints applied in [29], as the simulated traffic seemed very sluggish at the suggested lower constraint proposed in this publication. After several trial runs, this minimum value for $\alpha$ was raised to 0.8. In contrast, the participants seemed to prefer a much shorter distance than the suggested minimum value of 4m. Therefore, also this constraint on the lower bound was altered.

Master of Science Thesis

R.B. Molenaar
Result analysis and discussion

In this chapter the results of the performed experiments, as described in Chapter 5, will be presented and discussed. First, the initial experiments that were performed with synthetic data are covered, in which the functionality of the algorithm is tested with data that is generated using known models. In this experiment all parameters used to create the data can be controlled, so that the performance of the algorithm can be investigated.

Second, the offline experiments will be presented, where the designed algorithm will be presented with a real data set from the simulator. The data set, containing the same data format that will be used in the final experiments, is taken from a previous research performed with the simulator and will be used to analyse the ability of the proposed algorithm to deal with a real data set.

Finally, the experiment that was conducted with human interaction will be presented. In this experiment, the driving behavior of the drivers is fed to the algorithm, the results of which will be used in the traffic simulator. Each of these experiments was designed to answer a research question, therefore they will be discussed according to each question they belong to.
6-1 Research question 1

Research question 1: *Is the Bayesian evidence method able to correctly identify the set of models and parameters that are used to create a synthetically generated data set?*

A script is written, using Matlab, to generate datasets with longitudinal driving behavior based on one of two car-following models. Each data set consists of data representing a pair of drivers, where one driver is following the other car.

The longitudinal driving behavior of the front driver is modelled using the speed patterns shown in Figure 6-1 and the following driver’s driving behavior is modeled with either the GM or IDM model. The scripts written to perform these experiments can be found in Appendix E.

The data is fed to the Bayesian algorithm, implemented in Matlab as described in Section 4-2, to see if it can correctly identify which one of the two models was used to generate the data. This is repeated several times, where the model and reaction time choice for the data generation is randomly chosen, in order to test the different possible scenarios.

The model and reaction time, that are used to generate the synthetic data, are randomly chosen. This data is provided to the identification algorithm, which will try to identify which model, parameters, and reaction time were used to generate the data. This test is repeated 10 times, to get different combinations of reaction times and models.

The results of the 10 experiments are shown below, in Table 6-1. The first column of the table shows which model, with reaction time $\tau$, is used to generate the data. The second column shows the output of the Bayesian algorithm implementation, with a model and reaction time $\tau$. The results show that the algorithm is capable of predicting which model is used to
generate the observed driving behavior, which was the purpose of performing the experiment. The addition of noise to the observed data leads to a deviation of the used reaction time $\tau$ in some cases, but these deviations are small.

### 6-2 Research question 2

Research question 2: *Is the Bayesian evidence method able to identify a model and parameter set that best fits the driving behavior that is observed from a human participant in the traffic simulator?*

In this experiment, the same algorithm as applied in Section 6-1 is used. The algorithm is provided with a dataset containing human driving behavior, observed in a previous experiment. The experiment was performed as part of a different research, but since the simulation environment is largely identical, the data is considered valid for the purpose of answering the research question. The Matlab implementation of the algorithm can be found in Appendix E.

The dataset consists of trajectory data of three users following each other. The driving behavior of driver2 and driver3 can be used to model their car-following behavior, while they follow driver1 and driver2 respectively.

The results of providing the same Bayesian algorithm as used in Section 6-1 are shown in the table below.

<table>
<thead>
<tr>
<th>Driver pair</th>
<th>Best fitting model (reaction time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>driver1-driver2</td>
<td>IDM (0.6)</td>
</tr>
<tr>
<td>driver2-driver3</td>
<td>GM (1.9)</td>
</tr>
</tbody>
</table>

*Table 6-2: Results of experiment 2*

The main result of this experiment is that the proposed algorithm is able to identify a car-following model from observed longitudinal driving behavior in the experimental environment. The results above show that the algorithm is not only able to fit a model, it fits the two models to two quite similar data sets.

### Table 6-1: Synthetic experiment results

<table>
<thead>
<tr>
<th>Session</th>
<th>Model used ($\tau$)</th>
<th>Identification results ($\tau$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GM(1.3)</td>
<td>GM(1.2)</td>
</tr>
<tr>
<td>2</td>
<td>IDM(1.2)</td>
<td>IDM(1.2)</td>
</tr>
<tr>
<td>3</td>
<td>GM(0.7)</td>
<td>GM(0.7)</td>
</tr>
<tr>
<td>4</td>
<td>IDM(1.6)</td>
<td>IDM(1.6)</td>
</tr>
<tr>
<td>5</td>
<td>GM(1.7)</td>
<td>GM(1.7)</td>
</tr>
<tr>
<td>6</td>
<td>IDM(1.0)</td>
<td>IDM(1.0)</td>
</tr>
<tr>
<td>7</td>
<td>IDM(1.3)</td>
<td>IDM(1.3)</td>
</tr>
<tr>
<td>8</td>
<td>GM(0.5)</td>
<td>GM(0.4)</td>
</tr>
<tr>
<td>9</td>
<td>GM(1.4)</td>
<td>GM(1.4)</td>
</tr>
<tr>
<td>10</td>
<td>GM(0.9)</td>
<td>GM(0.9)</td>
</tr>
</tbody>
</table>
This suggests that the algorithm is suited for the sort of data that is used in the following experiments and that both models are capable of achieving a better fit than the other model with the driving behavior that is being observed with the current simulation environment. Based on the results, it is assumed that neither of the models will be consequently outperformed by the other and therefore that neither model is redundant in the algorithm.

If one of the models would always have a higher evidence than the other model, it would mean that the identification algorithm output would always be that one model. The simulated traffic would be updated with the same model every time, meaning that there will be no model variation in the simulated traffic. The results from this experiment validate the model choice, as model heterogeneity in simulated traffic is a required feature in recreating a realistic traffic situation.
6-3 Research question 4

Research question 4: *Does updating the driving models of AI cars, by feeding back information gained from observing human driving behavior in the same simulation session, lead to such an improvement of the subjective validity of that driving behavior that it is noticeable in a direct comparison?*

For this research question, each session will be considered separately. For each of the three different sessions, the orange and the green traffic will be compared. In session A and C, both will be generated using similar distributions and parameters and in session B the orange traffic will be generated using the update algorithm specified and the green traffic will be generated using the non-updated parameters and models.

**RQ 4 T-test session A:**

Hypothesis A: *The subjective validity score (svs) of the orange (non-updated) cars does not differ significantly from the score given to the green (non-updated) cars in session A.*

\[ H_0 : \text{svs}_{\text{green}} = \text{svs}_{\text{orange}} \]  \hspace{1cm} (6-1)
\[ H_1 : \text{svs}_{\text{green}} \neq \text{svs}_{\text{orange}} \]  \hspace{1cm} (6-2)

A statistical test is performed to see if the zero-hypothesis can be refuted. The student t-test assumes the compared samples to have a normal distribution. The t-test is performed on a dataset of 25 values. A two-tailed paired t-test indicates that there is no significant difference between the subjective validity score of the orange cars (\( \mu = 0.788, \sigma = 0.551 \)) and the subjective validity score of the green cars (\( \mu = 0.962, \sigma = 0.513 \)) in session A, with \( t(25) = 2.060, p = 0.243 \).

The zero-hypothesis can not be refuted at a significance level of 0.05, because the value of \( p > 0.05 \). Therefore Hypothesis A is true, indicating that the participants see no difference in the subjective validity of the longitudinal driving behavior of the simulated traffic.

As the cars were generated using the same model and parameter distributions, this was expected. The fact that test drivers apparently are not able to distinguish between orange and green cars provides evidence for the validity of the chosen experimental setup.

**RQ 4 T-test session B:**

Hypothesis B: *The subjective validity score (svs) of the orange (updated) cars is significantly lower than the score given to the green (non-updated) cars in session B.*

\[ H_0 : \text{svs}_{\text{green}} = \text{svs}_{\text{orange}} \]  \hspace{1cm} (6-3)
\[ H_1 : \text{svs}_{\text{green}} > \text{svs}_{\text{orange}} \]  \hspace{1cm} (6-4)

A one-tailed paired t-test shows that the subjective validity score of the orange cars (\( \mu = 0.683, \sigma = 0.550 \)) is significantly lower than the score of the green cars (\( \mu = 0.971, \sigma = 0.589 \)) in session B, \( t(25) = 2.060, p = 0.0285 \).
As \( p < 0.05 \), the zero-hypothesis is refuted at a significance level of 0.05. The value of \( sv_{green} \) is significantly bigger than the value of \( sv_{orange} \), and therefore Hypothesis B is true. This test shows that the application of the update algorithm results in a significant increase in subjective validity when both types of traffic are observed side by side in the simulation.

**RQ 4 T-test session C:**

Hypothesis C: *The subjective validity score (svs) of the orange (updated) cars does not differ significantly from the score given to the green (updated) cars in session A.*

\[
H_0: sv_{green} = sv_{orange} \\
H_1: sv_{green} \neq sv_{orange}
\]  

A two-tailed paired t-test indicates that there is no significant difference between the subjective validity score of the orange cars (\( \mu = 0.856, \sigma = 0.562 \)) and the subjective validity score of the green cars (\( \mu = 1.067, \sigma = 0.546 \)) in session C, \( t(25) = 2.060, p = 0.137 \).

As \( p > 0.05 \), the zero-hypothesis can not be refuted at a significance level of 0.05. The value of \( sv_{green} \) is not significantly different from the value of \( sv_{orange} \). This means that Hypothesis C is true. The results here are as expected, just as in the test performed for session A, with both types of cars were modelled using the same model and parameter distributions.

**6-4 Research question 5**

Research question 5: *Does updating the driving models of AI cars, by feeding back information gained from observing human driving behavior in the same simulation session, lead to such an improvement of the subjective validity of that driving behavior that it is noticeable when comparing over different sessions?*

To answer this research question, a comparison is made between session A and session C. The subjective validity of the longitudinal driving behavior without updates on the parameters and models, measured in session A, is compared with the subjective validity in session C, where the information gathered from the participants is used to update the distribution of models and corresponding parameter distributions. The hypothesis for this research question is that the users will be able to distinguish between them, as the car in front of them has a great influence on the participant’s driving behavior and a lot of interaction will take place. A change in model and/or parameter set, defining this driving behavior, is therefore expected to be distinguishable.

**RQ 5 T-test**

*Hypothesis:* The subjective validity score (svs) of the orange and green cars from session C combined are significantly lower than the svs of the orange and green cars from session A combined.

\[
H_0: sv_{sessionA} = sv_{sessionC} \\
H_1: sv_{sessionA} > sv_{sessionC}
\]
A one-tailed paired t-test shows that the null-hypothesis $H_0$ cannot be rejected and means that there is significant difference between the combined svs of session A ($\mu = 1.750, \sigma = 0.768$) and the combined svs of session C ($\mu = 1.923, \sigma = 0.857$).

As $p > 0.05$, the zero-hypothesis can not be refuted at a significance level of 0.05. The participants are not able to distinguish in the subjective validity of the simulated traffic in session A and the simulated traffic in session C.

The results from this experiment contradict the hypothesis which states that, just as in Section 6-3, a significant increase is expected in the svs of the car-following behavior which is being updated with data gathered from human driving behavior. An examination of possible causes for this difference is discussed in Section 6-6-4, where various aspects of the experiment and their implications on the results are discussed.
6-5 Parameter evolution

A plot showing the evolution of the parameters over time is shown in Figure 6-2. This figure shows that the evolution of the parameters over time is reasonably similar for all three different sessions. A clear jump in parameters can be observed after the first adaptation. The values given at update 0 are the values that the simulated traffic starts with, in both the updated and non-updated situations. For the non-updated traffic, all cars in session A and half of the cars in session B, these values stay the same for the entire duration of the experiment. The shown values are the values obtained from the algorithm and represent the best fit for the observed driving behavior. The other half of the cars of session B and all the cars of session C are updated with the values obtained from the algorithm.

The graphs show the parameter set outputs from the algorithm, and is therefore an indication of the driving behavior of the driver that is being observed. The graphs show a similar development for each session. The identified set of parameters quickly deviates from the initial values in each of the sessions. After this jump the parameter evolution levels out for all sessions in a similar fashion. This indicates that the method quickly finds the parameters that match the driving behavior of the participants. The same results in session C suggest that the adaptation of the surrounding traffic, to the obtained car-following models, does not significantly change the way the participants drive in the simulator. The graph shows that the model parameters, and therefore their surroundings, changes with respect to the beginning situation. However, the results of the method after the first few runs does not really change. This means that the observed driving behavior stays similar, and that the drivings do not
respond significantly to a change in their environment.

No evidence of user adaptation is found from the obtained models and parameters. If (some) user adaptation would be happening, the parameters and models found in the beginning of the experiment, during which the traffic is similar to the simulated traffic in session A, would be different compared to the findings at the end of the experiment session, at which point the traffic is updated with the observations. If adaptation would occur, the findings of the beginning and end should be different. Since no significant difference is found, no conclusions can be drawn on the adaptation of the users to the change in the environment.

6-6 Discussion

During the experiments a few things stood out that were unexpected. Some of these things might have influenced the driving behavior of the participants and therefore the results of the experiments. Other observations merely serve as a suggestion for improvement when such an experiment is performed again. These observations and their effect on the experiment results are discussed in this section. An overview of user suggestions and complaints in shown in Appendix A To give a good overview of the different problems, they are grouped in three categories: simulator, participants and experiment. Each aspect is explained in more detail below and after each aspect, the effects on the observed results is discussed. Finally, a discussion of the unexpected results found in Section 6-4 is presented and the possible causes of these results are described.

6-6-1 Simulator

Mismatch

The traffic simulator models each lane as a line and each simulated car’s center of gravity follows this line. There is an xml file that defines the coordinates of these lines, which is separate from the visual representation of the lanes in the simulator environment. There is a small mismatch between the visual location of the lanes in the simulator and the coordinates defined in the xml file. As the simulated cars follow the lines defined in the xml file, the cars will leave their lane. This is hardly visible in when driving in a straight line, but it’s quite significant in the corners. This causes the simulated cars to leave their lanes in the corners, sometimes causing a collision with a participant.

Control

It is found from the experiments, that some of the participants have trouble driving the car the way they want. The steering is considered by some to be very sensitive, as they keep overshooting on the small steering correction they are trying to make. This leads to weaving behavior of the human drivers and some accidental lane-changes. This leads to a problem of the simulated cars getting stuck in the middle of the road, as described later in this section. The cars driven by the users are automatics, they are not required to make shift gears while driving. This lets the participants focus on controlling the car, but does have a downside.
The users have two pedals to control the speed of the car; a throttle and a brake. The car can be made to go in reverse by holding the brake, as the brake pedal can be used to impart a negative acceleration to the car. Since there is no delay once the car has come to a halt from a forward velocity, the participant will immediately start reversing if the brake pedals is pressed. On multiple occasions, the participants started reversing into the cars following them, as they tried to stop at a traffic light.

**Freezing cars**

The lane that the user car is currently in, is the lane whose line is closest to the user car’s center of gravity. If a user departs his lane and then returns to it, while following a simulated car, it is possible for the simulated car to observe the user’s car to be in front of him, while actually it is behind him. As the model ‘sees’ a car very close in front it will decelerate with maximum deceleration, as is described in the longitudinal driving models in 2. This happened on multiple occasions during the experiments, as the users sometimes struggle to maintain their lane. The only way to get the simulated car to start moving again, is for the user to change lane and continue driving. After a few seconds, the simulated car will move once more and participate in traffic as it did before.

**Lane changing**

The 'freezing' of simulated cars in the middle of the road, as described above, necessitates lane changing. By passing the car that is blocking your way, this situation can easily be overcome. The problem with passing the car is that a lane change is required. The driving simulator environment does not have side view mirrors, which limits the rear visibility. As it is hard to see if there is room for a lane change, chances of a collision when doing so are quite significant. Especially in the dense traffic situation that is used for the experiment. A simulated car blocking the lane often lead to a collision of the participant with other simulated traffic. As the simulator is not designed to model crashes in a realistic way, this often lead to unrealistic situations which caused distraction for the participant in the experiment.

**Effects**

The aforementioned problems with the simulator lead to distraction for the people participating in the experiment. For instance, an accident that happened in one of the sessions has a significant influence on the subjective validity. Or when a user needs to focus all their attention to accurately controlling the car, it does not allow them to experience the traffic around them. These factors cause quite a distraction to the participants and deter from the noticeable difference in longitudinal driving behavior of the AI cars, as these other, more drastic or sudden changes in their simulated environment, form a ‘noise’ that drowns out the effect of changing the car-following behavior of the cars around the participants. This reduces the noticeable effect of the update mechanism and therefore the results seen in the questionnaires. A recommendation for this problem is made in Section 8-4
6-6-2 Participants

Learning effect

Even though the participants get to practice before starting the three sessions, there is a total time limit of 1 hour for the entire experiment. This means that practice time is limited, and therefore some learning will take place during the experiment sessions. This means that the participants will be more used to the simulator at the end of the experiment than in the beginning and will therefore drive differently. The influence of the learning effect on the experiment results is minimized by doing the sessions in random order. A more detailed explanation of the learning effect is given in [30].

Amateur participants

The questions in the questionnaire are designed to relate quantities observable by novices to traffic theory quantities. The people participating in the experiment are not experts in traffic theory, and through the questionnaire they are able to provide meaningful feedback for this research. The users are observing the effects of the update algorithm in a direct comparison, as proved in Section 6-3. They are not able to notice an improvement when the comparison is made over different sessions. It is possible that, because of the lack of expert knowledge, the users are not able to make this comparison when a longer timespan is involved as the differences don’t stand out to them as much.

Effects

The effects of which subject participate in the experiment should not make any difference on the outcome, as the participants are supposed to be a representative group of normal traffic participants. The fact that they are not traffic experts may limit the amount of information that can be obtained from their experiences, which is partially solved by the questionnaire. The influence of the learning effect is minimized by randomizing the order of the sessions, but ideally there should be less of a learning curve. A recommendation for this is given in Section 8-1.

6-6-3 Experiment

Limited influence

It is very important to note that the subjective validity of the simulator is only in some way dependent on the longitudinal driving behavior of the surrounding traffic. It could be that the influence of the adapted models on the subjective validity is simply too small to make an impression that is noticeable when comparing the difference in traffic in two separate sessions, resulting in an unexpected conclusion in Section 6-4. Recommendations for solving this problem are given in Section 8-3-1 and Section 8-3-2, where some recommendations are aimed at trying to increase the influence of the adaptation while others focus on reducing the influence of other factors that can disturb the experiments.
Length of sessions

Due to the layout of the track, the minimum sample length is defined at 200 seconds. This ensures that the participants experience different driving regimes, so that the data provided to the method contains sufficient information. Since each participant’s time is limited to 1 hour, each session is distance limited to two laps, which takes about 10 minutes. This means that, with a group of 3 participants at the same time, a maximum of only 9 updates of the model occur within one experiment session. A recommendation for solving this is described in Section 8-1.

6-6-4 Unexpected results

The Bayesian algorithm is expected to increase the subjective validity of the simulated traffic. A comparison is made between updated traffic and non-updated traffic, where the subjective validity score of the updated traffic should be higher. Two different comparisons are made in this thesis. The result of Research question 4B, a direct comparison in the same session, shows a significant increase in subjective validity. The result of Research question 5, an indirect comparison made over different sessions, does not show a significant different in the svls values. The effects described above may have had an influence on the outcome of the research, as described below.

Traffic simulator

The problems with the traffic simulator caused irregularities between the different sessions, thus violating the ceteris paribus assumption. The distraction of the users, caused by these problems, reduces the noticeable effect of the update mechanism on the subjective validity of the simulated traffic. This could have caused the unexpected results of Research question 5, but also questions the result of Research question 4B, as the experiment assumptions were violated. A recommendation is made in Section 8-4 to solve these issues for later research.

Amateur participants

A different reason for the contradiction in results between RQ 4B and RQ 5 is the fact that the participants are not experts in traffic analysis. A novice might be able to spot the difference between a real Rembrand painting and a fake one, when compared side by side. It takes more knowledge of the subject to make this judgement when the paintings can only be observed apart, with a certain amount of time passing in between observations.

Limited influence

The limited influence of the updated car-following behavior on the subjective validity of simulated traffic could be, in combination with the amateur participants mentioned above, the reason why the results indicate no difference when compared over separate sessions. Increasing the influence, as suggested in Section 8-3.
Conclusions

In this chapter, the insights gained from the experiments and analysis of the results are discussed. A short summary of each part of the research is given, discussing the results and the consequences.

7-1 Synthetic data

The first experiment is performed on synthetic data. The main goal of this part of the research is to analyse the performance of the adapted method and its ability to deal with the data it will be provided with in later experiments.

The experiment conducted shows that the method works with the additions made to the algorithm, that is proposed in [2]. Additions are made to identify the reaction time of the driver whose driving is being observed. Additions are made to deal with parameter constraints, which are necessary to guarantee the correct functioning of the models in the simulator. The behavior of participants is an uncontrolled part of the experiment, any instability resulting from anomalous behavior should therefore be caught.

The results of the experiments show that the algorithm is capable of finding the model, parameter set and reaction time that are used to generate the data. The noise added to the generated data creates a deviation which, in some of the test cases, deviates the estimate from the actual reaction times.

The differences in fit between the different car-following models are big enough for the algorithm to correctly find the models that were used to create the data. This indicates that, when the driving observed in the simulator matches one of the models best, the algorithm will be able to find it.
7-2 Real data

The second experiment of the research is performed on offline data. The offline data is obtained from a different experiment that had been performed in the same simulator, and the data is a good representation of the data that will be available for the last two experiments.

The results obtained from this experiment show a significant step is taken towards implementing the algorithm in the simulator. By subjecting the method to real data, an indication of the performance in the interactive experiments is obtained. The data presented in the interactive experiments will be of the same form, although the driving behavior is expected to be different in some of the sessions.

The results show two different models are identified from the data of two drivers. This does not only prove that the algorithm can find a model and parameter set to match the observed driving behavior, but it also shows that both models have a place in the algorithm when applied to this simulator.

7-3 Interactive experiments

The last two experiments of the research look at the effect of interaction between the participants and the simulator. The first of the two experiments looks at a direct comparison of the two traffic types, side by side in the same simulator session. The second experiment focused on the indirect comparison of driving behavior, across different simulator sessions.

Research question 4

In this experiment, the three sessions will be looked at separately. The driving behavior of the two types of simulated traffic, distinguishable by their different colors, will be analysed in a direct comparison. The participants will be presented with a questionnaire after each session, asking them about their experience during the drive. They score each type of simulated traffic through answering a series of questions. The subjective validity score (svs) of both types will be looked at, in a series of three separate research questions.

Each zero-hypothesis states an equal svs for both types of traffic. In order to prove a significant difference in subjective validity, the zero-hypothesis has to be refuted. The results for the three sessions are as hypothesized. The implementation of the algorithm had the desired effect on the subjective validity of the simulated traffic.

Research question 5

The goal of this experiment is to see whether the updating algorithm applied to all cars in one session creates a significantly higher subjective validity for the participants than the session in which the cars are governed by the standard model. The comparison of session A and session C will give us these results. The zero-hypothesis is not refuted, i.e. there is not enough evidence to falsify the claim that the effect of updating the car-following behavior is noticeable over different sessions. Suggestions are made to improve the quality of the experiment environment, which is expected to lead to a different result.
While performing the experiments, a lot of things have been learned. In this chapter, recommendations are made for further research, based on the experiences of doing this research.

8-1 Length

The length of the total experiment for each participant was set at 1 hour. This was fixed to increase the amount of available test subjects, as most of them were PhD students and postdoc’s at the NII with a busy schedule. They were given compensation for their effort, which is another reason why the experiment time had to be limited. Each participant had to drive in three different sessions, answer a questionnaire about each session, and practice before the experiment started to get acquainted with the simulator setup. This resulted in a short session duration, at around 10 minutes, and a similar time to practice before the experiment.

The recommendation is to take more time with each participant. A 10 minute practice session meant that the participant were still adapting to the simulator during the experiment sessions, which should be prevented. By either taking more time or performing less sessions in one experiment, the user has more time to learn to drive in the simulator. This reduces the change of the user’s driving behavior during the sessions and increase the quality of the results.

A longer session time will allow the user more time to adapt to the changing traffic. By increasing the time of the updated sessions, session C in the experiment, it is expected that the parameters obtained from the identification will change over time, as the users adapt to the changing traffic around them.

8-2 Adaptation

In the analysis performed in Section 6-5, no evidence is found of adaptation from the users to the changing simulated traffic around them. But what would the a follow-up research
question be if it did? When such evidence is found in future research, the following research question is suggested:

*Do changes in the situation of the simulator lead to the same effects in traffic behavior as they do in real life, in an absolute or relative way?*

By finding an answer to this question, the possibility of using the simulator as a tool for advising on real-world traffic problems is researched. Start with a real-world situation in which measurements of microscopic driving behavior are available, before and after an alteration in the traffic situation. From these results it should be clear what the change in average driving behavior is induced by this change in the situation. Recreate the original situation in a simulator to recording people’s driving behavior, applying the same change to the situation, as was done in the real world, and record people’s driving behavior again. Then compare the change in driving behavior of the real-world situation and the change in driving behavior in the simulator situation. If the absolute changes or relative changes (relative since the simulator does not really match with the real-world) are similar, this indicates that the simulator can be used as a tool for analysis the effects of changes in traffic situations on the driving behavior of people.

### 8-3 Extending the algorithm

In the current implementation, the algorithm is quite limited in the amount of possible models. An extension of the model range of the algorithm, or the addition of the identification of lane-changing models, would increase the adaptive possibilities of the implementation. The increase in possibilities might allow the algorithm to increase the gain in subjective validity. Each extension is discussed in a bit more detail below.

#### 8-3-1 Lane changing

By adding the identification and feedback of lane changing models an additional degree of freedom is added to the method. Lane-changing is a major part of a car’s driving behavior and, due to possible interaction with the user driven cars, something that could increase the subjective validity of the simulated traffic significantly. A research question for this experiment could be:

*Will a change in lane changing behavior of simulated traffic increase the subjective validity of the traffic simulator?*

#### 8-3-2 Extend model base

The flexibility of the algorithm can be increased by adding more car-following models to the list of available models in the algorithm. This allows the algorithm with more options to fit the data to. A research question to go with this topic could be:

*Does adding more models to the algorithm lead to a more pronounced change in subjective validity?*
A more pronounced change in subjective validity might lead to a different conclusion for research question 4, given in Section 6-4, where a more significant change in driving behavior is required. Since the implementation of this method only includes two longitudinal driving models, a lot of the existing models are not included. It could be that a different choice in the models will change the outcome of this research, as the model that is not included might prove to be significantly better than the current models. This is excluded from the scope of this research, due to the limited time and limited access to the participants and experimental setup.

8-4 Improve simulator

As described in the discussion of the results in Section 6-6, there are a few things that could be improved on the simulator when it is used for experiments. There is a bug which causes the simulated cars to 'freeze' in the middle of the road when following users depart and re-enter the lane, that they are in, in a short time. Combine this with the mismatch between the xml-file and the visual world, through which the traffic simulator thinks the users are in a different lane than they actually are, and the result is that traffic freezing happens on a regular basis.

A different issue that should be addressed, especially when lane-changing is added to the algorithm, as suggested in Section 8-3-1, is the lack of side view mirrors. The users need the same information that they get when driving in a real car, to make the simulation experience as subjectively valid as possible. The addition of side view mirrors will provide the users with more information about the area behind and beside them, and allow them to participate in traffic more dynamically.

8-5 High fidelity simulator

The simulator used for this experiment is a low-fidelity simulator. The focus of the simulator lies on multi-user experiments, which has been utilized in the experiments performed in this research, but it uses a relatively simple setup. The participants sit behind a desk looking at a computer monitor, using a steering wheel and pedals controller setup to control the car. More complex simulators are available, that try to give the user the most realistic feeling possible. By recreating things such as acceleration of the participant, a car interior, and visuals all around the user, the feeling of the simulator should match the feeling of driving a car more closely. This should result in the participants driving more realistically in the simulator and therefore obtaining more realistic results. Since the costs of such a driving simulator are considerable, this was not an option for this research. It could lead to new insights, which is why it is suggested for further research.
Appendix A

Experiment remarks

In this appendix, the notes and unexpected events for each simulation session are summarised.

Session 1

Date: 16-07  Time: 12:30 - 13:30

- All drivers have difficulty controlling the car
- User 1: "Car sounds too high revving, sounds like you are going very fast."
- User 1 overcorrects the car, he has trouble to maintain his lane.
- Users drive backwards when trying to stop at traffic lights, as they are used to keep pressing the brake pedal.
- Non-overtaking orders slows the cars of the users down, as they get stuck behind AI cars.

Session 2

Date: 16-07  Time: 15:30 - 16:30

- User 3 crashed into user 1
- Users have difficulty maintaining their speed at their desired level and difficulty staying in their lane
- AI cars driving in lane 2 appear to be in lane 1 in the corners. When they remerge with lane 2 they scare users out of their lane.
- User 2 & 3 control the acceleration or the steering of the car, not both at the same time.
Session 3
Date: 16-07   Time: 17:00 - 18:00

- Green car crashed into User 3, due to a xml & sim mismatch (according to one of the simulation programmers)
- Overtaking is difficult, due to the lack of side view mirrors, but necessary, as AI cars sometimes freeze in the road and maintain that position until the user behind it overtakes them.
- Users sometimes do not see the traffic light, due to the user number balloon above the user car in front of them.
- Green car moved into the lane of User 3, pushing him out of the lane.
- Brake & reverse separation is advised, but not applied, due to lack of time.

Session 4
Date: 17-07   Time: 11:00 - 12:00

- Green car gets stuck in the middle of the road, User 3 had to change lane in order to overtake
- This happened twice more, once with a green AI car and User 1 and once with an orange IA car and User 3.
- User 3 has difficulty staying in a lane, he weaves slightly around the middle of the lane.
- User 2 and User 3 are half a lap behind User 1 at the end of the session, the result of the influence of the traffic lights on the session.

Session 5
Date: 17-07   Time: 12:30 - 13:30

- User 3 got motion sickness during the driving, aborted the first session about 80% in. Gave it another shot in the second session, but stopped as no improvement was made.
- User 3 did not notice orange cars during his driving session, as he was too distracted feeling sick.
- User 1 complains of non-realistic sound.
- User 3 did not fill out questionnaire 2, 3 and the big questionnaire.
- User 1 and 2 remark that braking is difficult without driving backwards.
Session 6

Date: 17-07  Time: 14:30 - 15:30

- Driving backwards as a result of continued braking causes a collision between two drivers.
- AI cars stuck in the lane cause confusion among User 2 and User 3, as they stay stationary behind the car.
- Fullscreen minimises on some occasions, without any clear indication what is causing it.
- Dive server crashed, for the second time today. Luckily not during a session.
- Car freezes in front of User 1, because of lane-change bug. Happens 3 times in total.
- Contact with an AI car freezes them, making it very annoying to pass.
- User 3 does not listen to instructions regarding maximum velocity.

Session 7

Date: 17-07  Time: 17:00 - 18:30

- Remarks were made that the session length is too short to form an opinion.
- User 3 has difficulty with the steering sensitivity.
- The distinction between genuine stopping behavior and freezing of cars is causing quite some confusion.
- Orange car rammed User 3 at intersection, as it tried to drive through him.
- Last sub session was restarted, as there were some connection issues.

Session 8

Date: 18-07  Time: 11:00 - 12:00

- User 2 has difficulty with the braking.
- User 2 got rear ended by an AI car.
- User 1 comments on the lack of side view mirrors.
Session 9

Date: 18-07  Time: 14:00 - 15:00

- User 1 gets rear ended by green AI car.
- User 2 gets rear ended by orange AI car.
- Session 3 = session 1 questionnaire and vice versa
- User 2 suggest a separation of the braking and reversing of the car.

Session 10

Date: 18-07  Time: 15:30 - 16:30

- User 2 hit User 1, knocking him to the other side of the road, twice
- users too occupied trying to control the car correctly to really notice their surroundings
- Orange car froze in the middle of the road.
- User 3 gets rear ended by green car.
Appendix B

C# Implementation

This function represents the backbone of the implementation, calling each function with the parameters done from this function. This is also the place where the values for $\theta_{prior}$ and $\sigma_{prior}$ are stored and updated. These values will be pushed towards the cars each time the identification cycle is run. The functions that are called from the main function are omitted from the appendix, as they are of little informative value. The codes are available on request.

B-1 Main

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Inputs</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>DenseVector:</td>
<td>$v_1$, $v_2$</td>
<td>-</td>
<td>$a$, $v_s$, $v_p$, $v_q$</td>
</tr>
<tr>
<td>Matrix</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>double:</td>
<td></td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Table B-1: Table with input and output data types of main function
public static void Main()
{
    DenseVector speedOne = new DenseVector(new[] {2.0, 3.1, 3.0, 2.9, 2.8});
    DenseVector speedTwo = new DenseVector(new[] {2.1, 3.2, 3.1, 3.0, 2.8});
    DenseVector ObservedSpeed = new DenseVector(new[] {2.0, 3.1, 3.0, 2.9, 2.8});
    //var speedOne = new DenseVector(new[] {2.0, 3.1, 3.0, 2.9, 2.8});
    DenseVector theta = new DenseVector(new[] {1.0, 2.0, 3.0, 4.0, 5.0});
    double SigP = 1;
    var invSigP2 = 1 / (Math.Pow(SigP, 2));
    double invSigL2 = 1;
    int dt = 1;
    double theta_MP = 3.0;
    double UpperSum =
    double LowerSum =
    double dvSum =
    DenseVector vs;
    DenseVector vp;
    DenseVector vq;
    var PTheta = new DenseVector(theta.Count);
    DenseMatrix VPred;
    DenseVector PD_Theta;
    double A;
    ReadData(theta, speedOne, speedTwo, dt, out UpperSum, out LowerSum,
    out dvSum, out vp, out vs, out vq, out VPred);
    FindThetaMp_GM(speedOne, speedTwo, vq, theta, theta_MP, dt, UpperSum,
    LowerSum, invSigP2, invSigL2, out theta_MP);  
    PTheta_GM(theta, theta_MP, SigP, 2, out PTheta);
    PD_Theta_GM(VPred, theta, invSigL2, speedOne, out PD_Theta);
    FindHessian(vq, invSigP2, dt, invSigL2, out A);
}
Appendix C

Session questionnaire

C-1 Acceleration at intersection

About the acceleration of green vehicles at intersections:
In your opinion, did the green vehicles around you accelerate (after slowing down/stopping at an intersection) slower, as fast, or faster than vehicles on a real road would?

Grade according to the 1-7 scale below.

\[
\begin{array}{cccccc}
\circ & \circ & \circ & \circ & \circ & \circ \\
\text{much} & \text{slower} & \text{as fast as} & \text{much} & \text{faster}
\end{array}
\]

About the acceleration of orange vehicles at intersections:
In your opinion, did the orange vehicles around you accelerate (after slowing down/stopping at an intersection) slower, as fast, or faster than vehicles on a real road would?

Grade according to the 1-7 scale below.

\[
\begin{array}{cccccc}
\circ & \circ & \circ & \circ & \circ & \circ \\
\text{much} & \text{slower} & \text{as fast as} & \text{much} & \text{faster}
\end{array}
\]

C-2 Following distance

About the acceleration of green vehicles at intersections:
In your opinion, did the green vehicles around you keep a smaller, as big as, or bigger gap to the vehicle in front of them than vehicles on a real road would?

Grade according to the 1-7 scale below.

About the acceleration of orange vehicles at intersections:
In your opinion, did the orange vehicles around you keep a smaller, as big as, or bigger gap
to the vehicle in front of them than vehicles on a real road would?

*Grade according to the 1-7 scale below.*

<table>
<thead>
<tr>
<th></th>
<th>Very Much Smaller</th>
<th>As Big As</th>
<th>Very Much Bigger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Much</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smaller</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C-3 Speed

*About the acceleration of green vehicles at intersections:*  
In your opinion, did the green vehicles around you drive slower, as fast as, or faster than vehicles in a comparable situation on a real road would?  
*Grade according to the 1-7 scale below.*

<table>
<thead>
<tr>
<th></th>
<th>Very Much Slower</th>
<th>As Fast As</th>
<th>Very Much Faster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Much</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slower</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*About the acceleration of orange vehicles at intersections:*  
In your opinion, did the green vehicles around you drive slower, as fast as, or faster than vehicles in a comparable situation on a real road would?  
*Grade according to the 1-7 scale below.*

<table>
<thead>
<tr>
<th></th>
<th>Very Much Slower</th>
<th>As Fast As</th>
<th>Very Much Faster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Much</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slower</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C-4 Reaction Time

*About the acceleration of green vehicles at intersections:*  
In your opinion, did the green vehicles around you react slower, as fast as, or faster to changes in the speed of the car in front of them than vehicles on a real road would?  
*Grade according to the 1-7 scale below.*

<table>
<thead>
<tr>
<th></th>
<th>Very Much Slower</th>
<th>As Fast As</th>
<th>Very Much Faster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Much</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slower</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*About the acceleration of orange vehicles at intersections:*  
In your opinion, did the green vehicles around you react slower, as fast as, or faster to changes ...
in the speed of the car in front of them than vehicles on a real road would? Grade according to the 1-7 scale below.

○ ○ ○ ○ ○ ○ ○
much slower as fast as much faster

C-5 Open questions

How did you experience the traffic situation in this session? Was it like the normal traffic situations that you might experience in the real world?
Appendix D

Final questionnaire

Part A: Participant information

Gender: male/female
Age: ___
Average car usage (hours driven per week): ___
Year of license issued: ___
Mileage driven last year (approximately): ___
Previous experience in a driving simulator: yes/no
Experience in playing video games: yes/no

Rate yourself as a driver

<table>
<thead>
<tr>
<th>What best describes you as a driver?</th>
<th>○ − − ○ − − ○ − − ○ − − ○ − − ○</th>
</tr>
</thead>
<tbody>
<tr>
<td>not observant</td>
<td>observant</td>
</tr>
<tr>
<td>What describes you best as a driver?</td>
<td>○ − − ○ − − ○ − − ○ − − ○ − − ○</td>
</tr>
<tr>
<td>slow</td>
<td>fast</td>
</tr>
<tr>
<td>What best describes you as a driver?</td>
<td>○ − − ○ − − ○ − − ○ − − ○ − − ○</td>
</tr>
<tr>
<td>not dangerous</td>
<td>dangerous</td>
</tr>
<tr>
<td>What best describes you as a driver?</td>
<td>○ − − ○ − − ○ − − ○ − − ○ − − ○</td>
</tr>
<tr>
<td>neglecting</td>
<td>attentive</td>
</tr>
<tr>
<td>What best describes you as a driver?</td>
<td>○ − − ○ − − ○ − − ○ − − ○ − − ○</td>
</tr>
<tr>
<td>nervous</td>
<td>calm</td>
</tr>
<tr>
<td>What best describes you as a driver?</td>
<td>○ − − ○ − − ○ − − ○ − − ○ − − ○</td>
</tr>
<tr>
<td>cautious</td>
<td>careless</td>
</tr>
</tbody>
</table>
Part B: Questions related to the driving experience in the simulation environment

Try to estimate how realistic, in your opinion, the below-mentioned aspects of your driving experience in the simulation felt.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>The feeling when steering</td>
<td></td>
</tr>
<tr>
<td>The feeling when decelerating</td>
<td></td>
</tr>
<tr>
<td>The feeling when accelerating</td>
<td></td>
</tr>
<tr>
<td>The feeling of the vehicle’s speed</td>
<td></td>
</tr>
<tr>
<td>The feeling of the vehicle’s lateral position</td>
<td></td>
</tr>
<tr>
<td>Your overall impression of the pictures/images</td>
<td></td>
</tr>
<tr>
<td>The feeling of depth in the pictures/images</td>
<td></td>
</tr>
<tr>
<td>The feeling of the size and the weight of the vehicle</td>
<td></td>
</tr>
<tr>
<td>Simulation experience in general</td>
<td></td>
</tr>
<tr>
<td>Were you sick during the time when you were driving in the simulator?</td>
<td></td>
</tr>
</tbody>
</table>

R.B. Molenaar  Master of Science Thesis
Part C: Questions related to the simulation environment

1. Did you experience any traffic situation that felt strange or unrealistic during your ride? In other words, did any other driver behave strangely or unrealistically at any time?
   
   (a) Yes (go to question 2a)
   (b) No (go to question 2b)

2. Follow up question on question 1:
   
   (a) If you answered yes to question 1. Try to describe the situation/situations
      
      • Describe the situation:
      • Do you think that this situation could appear in real life? Yes/No
   
   (b) If you answered no to question 1. In your opinion, could all situations that happened during you driving experience happen in real life?
      
      i. Yes
      ii. No

3. Which opinions do you have regarding other drivers’ behavior in connection with:
   
   • You changing lane?
   • Other drivers changing lane?
   • Other situations

Part D: General discussion

Please tell us your opinion about the experiment. Any comments you might have are welcome!
In this appendix, the MATLAB codes used for the first 2 experiments are shown.

**E-1 Experiment 1**

```matlab
clear all
clc

% Define bounds GM
lb_GM = [0.2];
ub_GM = [0.7];

% Define bounds IDM
lb_IDM = [0.3;0.5;0.3;0.5;10];
ub_IDM = [3;5;5;10;30];

% Define simulation parameters
dt = 0.1;
tend = 400;
sigPriorSq = 0.1;
thetaTrue_GM = 0.3;
thetaTrue_IDM = [0.75,1.7,1.6,2,20];
Sigma_GM = 0.01;
Sigma_IDM = diag([0.05,0.05,0.05,0.05,1]);

for jj = 1:10
    % Data generation
    Model = round(rand(1)); % 0 = IDM, 1 = GM
    real_lag = round(rand(1)*20)/10;
    if Model == 1
...
```matlab
[x_leader, x_follower] = test_script_2(Model, dt, tend, real_lag, thetaTrue_GM, 0, 0.5);
model_used = 'GM';
else
[x_leader, x_follower] = test_script_2(Model, dt, tend, real_lag, thetaTrue_IDM, 0, 0.5);
model_used = 'IDM';
end

%% Identification
% Run the Identification algorithm for all possible reaction times, from 0 to 2 seconds.
for ii = 1:20
    clear vs vp vr vq gap
    lag = ii/10 - 0.1;
    speed_leader = x_leader(2,:);
    speed = x_follower(2,:);
    gap_total = x_leader(1,:)-x_follower(1,:);
    lagsteps = round(lag/dt);
    for i = 1:length(speed)-1-lagsteps
        vs(i) = speed(i+1+lagsteps);
        vp(i) = speed(i+lagsteps);
        vr(i) = speed(i);
        vq(i) = speed_leader(i)-speed(i);
        gap(i) = gap_total(i);
    end
    N = length(vs);

    options = optimset('Display', 'off', 'Algorithm', 'active-set');

    % thetaMP_GM = fminsearch(@(a) Error_function_GM(a, thetaTrue_GM, Sigma_GM, vp, dt, vq, vs, 0)); % Unconstrained
    thetaMP_GM = fmincon(@(a) Error_function_GM(a, thetaTrue_GM, Sigma_GM, vp, dt, vq, vs), 0.3, [], [], [], [], lb_GM, ub_GM, options);
    % Constrained
    vpred_GM = vp+dt*thetaMP_GM*vq;
    summedError_GM = sum((vpred_GM-vs).^2);
    sigLSq_GM = (1/length(vs))*summedError_GM;
    pLogThetaMP_GM = (-0.5*(thetaMP_GM-thetaTrue_GM)'*(1/Sigma_GM)*(thetaMP_GM-thetaTrue_GM))-log(2*pi)*(N/2)-log(det(Sigma_GM))*0.5;
    pLogDThetaMP_GM = ((-1/(2*sigLSq_GM)))*sum((vp+dt*thetaMP_GM+vq-vs).^2)-log(sigLSq_GM*2*pi)*(N/2);
    p2LogDThetaMP_GM = ((-1/(2*sigLSq_GM)))*summedError_GM-log(sigLSq_GM*2*pi)*(N/2);
    A_GM = (1/Sigma_GM)*((1/sigPriorSq)+((dt)^2)/(sigLSq_GM))*sum(vq.^2);
    PDLog_GM = pLogDThetaMP_GM+pLogThetaMP_GM-0.5*log(det(A_GM/(2*pi)));```
\%thetaMP_IDM = fminsearch(@(a) Error_function_IDM(a,thetaTrue_IDM,dt,Sigma_IDM,sigPriorSq,vp,vq,vs,vr,gap),[0.75,1.7,1.6,2,35]);
thetaMP_IDM=fmincon(@(a) Error_function_IDM(a,thetaTrue_IDM,dt,Sigma_IDM,sigPriorSq,vp,vq,vs,vr,gap),[0.75,1.7,1.6,2,35],[],[],[],[],lb_IDM,ub_IDM,[],options);
vpred_IDM=vp+dt*(thetaMP_IDM(1)*(1-(vr./thetaMP_IDM(5)).^4-(thetaMP_IDM(4)+thetaMP_IDM(3).*vr-(vr.*vq)/(2*sqrt(thetaMP_IDM(1)*thetaMP_IDM(2))))./gap).^2);
summedError_IDM=sum((vpred_IDM-vs).^2);
sigLSq_IDM=(1/length(vs))*summedError_IDM;
pLogThetaMP_IDM=(-0.5*(thetaMP_IDM-thetaTrue_IDM)'*inv(Sigma_IDM)*(thetaMP_IDM-thetaTrue_IDM))'-log(2*pi)*(N/2)-log(det(Sigma_IDM))*0.5;
pLogDThetaMP_IDM=(-1/(2*sigLSq_IDM))*summedError_IDM-log(sigLSq_IDM*2*pi)*(N/2);
A_IDM=Hessian(thetaMP_IDM,sigLSq_IDM,vp,vq,vr,vs,vr,gap,dt);
A_constr=CramerRao(A_IDM,lb_IDM,ub_IDM,thetaMP_IDM);
PDLog_IDM=pLogDThetaMP_IDM+pLogThetaMP_IDM-0.5*log(abs(det(A_IDM/(2*pi))));
\% Save results for current RT in table
Resulttable(ii,:)=[PDLog_GM PDLog_IDM];
end
\% Find the (model/reaction time) combination with highest evidence
[ind_RT,ind_Model]=find(Resulttable==max(max(Resulttable)));
if ind_Model == 1
best_model = 'GM';
else
best_model = 'IDM';
end
display(['Model used for data generation is ' model_used ' with reaction time ' num2str(real_lag)]);
display(['Best model is ' best_model ' with reaction time ' num2str((ind_RT-1)/10)])
end
clear all
c
c%Load data
RQ_2_Step1

%constraints
lb_GM=[0.2];
ub_GM=[0.7];

lb_IDM=[0.3;0.5;0.3;0.5;10];
ub_IDM=[3;5;5;10;30];

leader = driver0;
follower = driver1;

for ii=1:20
    clear vs vp vr vq gap
    lag=ii/10−0.1;
dt=0.1;
tend=400;
sigPriorSq = 0.1;
thetaTrue_GM = 0.3;
thetaTrue_IDM=[0.75,1.7,1.6,2,20];
Sigma_GM = 0.01;
Sigma_IDM = diag([0.05,0.05,0.05,0.05,1]);

speed_leader = leader.speed;
speed = follower.speed;
PosZ).^2);
lagsteps=round(lag/dt);

for i = (20−lagsteps) : length(speed)−1−lagsteps;
    j=i−(19−lagsteps);
    vs(j)=speed(j+lagsteps);
    vp(j)=speed(j+lagsteps);
    vr(j)=speed(j);
    vq(j)=speed_leader(j)−speed(j);
    gap(j)=gap_total(j);
end
N=length(vs);

options = optimset('Display', 'off','Algorithm','active-set');

% Unconstrained
% thetaMP_GM = fminsearch(@(a) Error_function_GM(a,thetaTrue_GM, Sigma_GM,vp,dt,vq,vs),0);
% Constrained
thetaMP_GM=fmincon(@(a) Error_function_GM(a,thetaTrue_GM,Sigma_GM,vp
,dt,vq,vs,vr,gap),[0.75,1.7,1.6,2,35],[],[],[],lb_GM,ub_GM,[],options);
vpredd_GM = vp+dt*thetaMP_GM*vq;
summedError_GM = sum((vpredd_GM-vs).^2);
sigLSq_GM = (1/length(vs))*summedError_GM;
pLogThetaMP_GM =(-0.5*(thetaMP_GM-thetaTrue_GM)'*(1/Sigma_GM)*
(thetaMP_GM-thetaTrue_GM)-log(2*pi)*(N/2)-log(det(Sigma_GM))*0.5;
pLogDThetaMP_GM =((1/(2*sigLSq_GM))*sum((vp+dt*thetaMP_GM+vq-vs).^2))
-log(sigLSq_GM+2*pi)*(N/2);
p2LogDThetaMP_GM =((-1/(2*sigLSq_GM))*summedError_GM)-log(sigLSq_GM+2*pi)
*(N/2);
A_GM = (1/Sigma_GM)*((1/sigPriorSq)+(vq^2)/(sigLSq_GM))
PDLog_GM = pLogDThetaMP_GM+pLogThetaMP_GM-0.5*log(abs(det(A_GM)/(2*pi)))
A_collection{ii}=PDLog_GM;

% Unconstrained
thetaMP_IDM = fminsearch(@(a) Error_function_IDM(a,thetaTrue_IDM,dt,
Sigma_IDM,sigPriorSq,vp,vq,vs,vr,gap),[0.75,1.7,1.6,2,35]);
% Constrained
thetaMP_IDM=fmincon(@(a) Error_function_IDM(a,thetaTrue_IDM,dt,
Sigma_IDM,sigPriorSq,vp,vq,vs,vr,gap)
,[0.75,1.7,1.6,2,35],[],[],[],lb_IDM,ub_IDM,[],options);
vpredd_IDM = vp+dt*(thetaMP_IDM(1)*(1-(vr./thetaMP_IDM(5)))^4-
(thetaMP_IDM(4)+thetaMP_IDM(3).*vr-(vr.*vq)/(2*sqrt(thetaMP_IDM(1)*
thetaMP_IDM(2))))/gap).^2);
summedError_IDM = sum((vpredd_IDM-vs).^2);
sigLSq_IDM = (1/length(vq))*summedError_IDM;
pLogThetaMP_IDM =(-0.5*(thetaMP_IDM-thetaTrue_IDM)*inv(Sigma_IDM)*
(thetaMP_IDM-thetaTrue_IDM)'-log(2*pi)*(N/2)-log(det(Sigma_IDM))
+0.5;
pLogDThetaMP_IDM =((-1/(2*sigLSq_IDM))*summedError_IDM)-log(sigLSq_IDM
+2*pi)*(N/2);
A_IDM = Hessian(thetaMP_IDM,sigLSq_IDM,vp,vq,vs,vr,gap,dt);
A_constr=CramerRao(A_IDM,lb_IDM,ub_IDM,thetaMP_IDM);
PDLog_IDM = pLogDThetaMP_IDM+pLogThetaMP_IDM-0.5*log(abs(det(A_IDM/(2*
pi))));
A_collection{ii}=A_IDM;

if PDLog_GM>PDLog_IDM
    best_model = ['GM' int2str(ii)];
else
    best_model = ['IDM' int2str(ii)];
end
Resulttable(ii,:)=[PDLog_GM PDLog_IDM];
Resultlist{ii}=best_model;

% ConstraintReached = sum(lb_IDM==thetaMP_IDM);
if (ConstraintReached==1)
    display('Constraint reached')
    thetaMP_IDM
    A_IDM
end

[ind_RT, ind_Model] = find(Resulttable == max(max(Resulttable)));
if ind_Model == 1
    best_model = 'GM';
else
    best_model = 'IDM';
end

display(['Best model is ', best_model, ' with reaction time ', num2str((ind_RT - 1)/10)])
Convex functions

The Nelder-Mead algorithm is used to find the value of $\theta$ that gives the minimum value of the error function (3-9). The Nelder-Mead algorithm will only find the global minimum of convex functions, which is the case as is proved in this appendix.

A function $f$ is convex if the Hessian of the function is positive definite. The Hessian of the General Motors (GM) function, calculated in Section 4-2-1, is given by

$$\frac{\delta^2 E}{\delta \alpha^2} = \frac{1}{\sigma_{\text{prior}}^2} + \frac{\Delta t^2}{\sigma_l^2} \sum_{k=1}^{K} v_q^2$$  \hspace{1cm} (F-1)

For all real values of the variables, $\frac{\delta^2 E}{\delta \alpha^2} \geq 0$, as all terms in the equation are quadratic. As $\Delta t$ and $\sigma_l$ are non-zero, and $v_q$ can all be assumed to be non-zero for any reasonable dataset, we can conclude that $\frac{\delta^2 E}{\delta \alpha^2} > 0$. This proves that the Hessian of the GM function is positive definite and that the use of the Nelder-Mead algorithm is correct.

Similar proof can be derived for the IDM model, but this is excluded from the report as the calculations are too extensive.


Glossary

List of Acronyms

AI     Artificial Intelligence
JARI   Japan Automobile Research Institute
TU Delft Delft University of Technology
IDM    Intelligent Driver Model
GM     General Motors
GHR    Gazis-Herman-Rothery
CRLB   Cramér-Rao lower bound
NII    National Institute of Informatics, Tokyo
svs    subjective validity score

List of Symbols

$\theta^{MP}$ Most probable parameter set
$\theta_i$ Parameter set that defines car-following behavior of driver $i$
$\bar{\theta}$ Mean parameter values
$\Delta v$ Velocity difference between vehicle $i$ and vehicle $i-1$
$\Delta x$ Distance between vehicle $i$ and vehicle $i-1$
$Pr(\theta)$ Prior probability distribution of parameter set $\theta$
$Pr(\theta | D)$ Posterior probability distribution of parameter set $\theta$, with data set $D$
$Pr(D)$ Normalization factor
$Pr(D | \theta)$ Likelihood of data set $D$, given parameter set $\theta$

Master of Science Thesis

R.B. Molenaar
\( \Sigma \)  Covariance matrix
\( \tau_i \)  Reaction time of driver \( i \)
\( \Theta_C \)  Constrained parameter space
\( a_i \)  Acceleration of vehicle \( i \)
\( E \)  Estimator error
\( E_l \)  Estimator likelihood error
\( E_p \)  Estimator error in parameter set
\( J_\theta \)  Fisher information matrix
\( L_i \)  Length of vehicle \( i \)
\( v_i \)  Speed of vehicle \( i \)
\( v_o(k) \)  Observed speed at time \( k \)
\( v_p(k, \theta) \)  Predicted speed at time \( k \) with parameter set \( \theta \)
\( x_i \)  Position of vehicle \( i \)