REALISTIC OFF-RAMP COUPLING CONDITIONS FOR MACROSCOPIC HIGHWAY NETWORK MODELS

A Dissertation
Submitted to
the Temple University Graduate Board

in Partial Fulfillment
of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY

by
Najmeh Salehi
August, 2020

Examinining Committee Members:

Benjamin Seibold, Advisory Committee Chair, Mathematics
Daniel B. Szyld, Mathematics
Isaac Klapper, Mathematics
David Futer, Mathematics (Graduate Chair)
Paola Goatin, External Member, INRIA Sophia Antipolis - Méditerranée
ABSTRACT

REALISTIC OFF-RAMP COUPLING CONDITIONS FOR
MACROSCOPIC HIGHWAY NETWORK MODELS

Najmeh Salehi
DOCTOR OF PHILOSOPHY
Temple University, August, 2020
Dr. Benjamin Seibold, Chair

Transportation systems are among the critical infrastructures in every society. In order to design robust and reliable transportation networks, one needs to have a solid understanding of the behavior of the traffic flow in these networks. Many studies have been conducted to describe, control and predict the traffic flow on the networks. However, there are still some shortcomings in the existing literature that need to be addressed. For example, there are currently no satisfactory macroscopic coupling models for off-ramps on the highways. Specifically, existing models have fundamental modeling shortcomings, and model-fitting and validation of coupling models with traffic data have received little attention. To this end, this thesis will address some of the existing gaps in the literature of macroscopic traffic flow modeling by developing new coupling conditions for off-ramps on highways. This dissertation contributes to the existing literature in the following aspects: modeling, analysis, and validation with data. From a modeling point of view, there are two sets of coupling conditions in the literature for the off-ramp: FIFO (First In, First Out) and non-FIFO. Under the classical FIFO coupling conditions, a clogged off-ramp yields zero flux through the junction. Clearly, on multi-lane highways this is unrealistic, as a queue forming from the off-ramp will generally be restricted to the right-most lane, and vehicles that do not wish to exit can pass the queue. Moreover, the issue with the non-FIFO coupling conditions is that they lead
to spurious re-routing of vehicles. To remedy these issues, we develop a new coupling model by using a vertical queue at the junction. The vertical queue keeps track of the excess vehicles of a certain type (exiting vs. non-exiting) that may join the congested traffic by more than the other vehicle type does. From the analysis point of view, the introduction of the vertical queue as well as the requirement of the model to preserve the split ratios, lead to some differences from the existing models in the literature that renders proving the well-posedness of the model a non-trivial task. In this dissertation, we undertake this task and establish the well-posedness of the model. Specifically, we show that there exists a unique solution that is continuously dependent on the initial data. Finally, we use the data generated from a microsimulator to validate our model and compare it with the existing models. Specifically, we establish micro-simulation representations of the off-ramp scenarios, and describe how to systematically extract macro quantities from the results of the microsimulator. Then, we compare the results of the macroscopic models with the macro quantities extracted from the microsimulator.
ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Dr. Benjamin Seibold. He has always been very patient and enthusiastic and I have found his advice and support invaluable. I feel very fortunate and privileged for being a student of such a great mentor and mathematician. Without his guidance, this research project would not have been possible.

I would also like to thank my thesis committee members, Dr. Daniel B. Szyld, Dr. Isaac Klapper, Dr. David Futer and Dr. Paola Goatin. I would also like to thank all the Temple math faculty. My knowledge of mathematics has grown so much in the past years and I owe that to the professors of the many classes I have taken as a graduate student at Temple. In particular, I would like to thank Dr. Yury Grabovsky, Dr. Wei-Shih Yang, and Dr. Cristian E. Gutierrez for all the knowledge I learned from them.

I want to thank my wonderful and supportive family, and all my friends.
To My Family.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>DEDICATION</td>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>xii</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>OFF-RAMP COUPLING CONDITIONS</td>
<td>6</td>
</tr>
<tr>
<td>2.1</td>
<td>Modeling and Computational Foundations</td>
<td>8</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Riemann Problem and Cell Transmission Model</td>
<td>9</td>
</tr>
<tr>
<td>2.1.2</td>
<td>Generalized Riemann Problem</td>
<td>11</td>
</tr>
<tr>
<td>2.1.3</td>
<td>Split Ratio</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>Existing Models</td>
<td>13</td>
</tr>
<tr>
<td>2.2.1</td>
<td>FIFO Model</td>
<td>13</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Non-FIFO Model</td>
<td>14</td>
</tr>
<tr>
<td>2.3</td>
<td>New FIFOQ Model</td>
<td>16</td>
</tr>
<tr>
<td>2.4</td>
<td>Cell Transmission Discretization</td>
<td>20</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Approximation Along Edges</td>
<td>20</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Treatment of Queues</td>
<td>21</td>
</tr>
<tr>
<td>2.5</td>
<td>Model Comparison and Numerical Examples</td>
<td>22</td>
</tr>
<tr>
<td>2.6</td>
<td>Model Extensions</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>ANALYSIS OF THE FIFOQ MODEL</td>
<td>34</td>
</tr>
<tr>
<td>3.1</td>
<td>Preliminaries and Riemann Solver</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Important Characteristics of the FIFOQ Model</td>
<td>40</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Self-Consistency of the Model</td>
<td>41</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Flux Maximization</td>
<td>44</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1.1 Greenshields fundamental diagram and its corresponding velocity diagram. ........................................... 3
1.2 Triangular fundamental diagram and its corresponding velocity diagram. ............................................... 3

2.1 A road network consisting of three roads and a single junction. The picture illustrates a diverge junction where one outgoing road, $I_3$, is an off-ramp. $\alpha_2$ and $\alpha_3$ are the split ratios (see Section 2.1.3) ............................................................... 6
2.2 Demand and supply functions on incoming and outgoing roads, respectively. .......................................... 10
2.3 Interpretation/derivation of two models with queues. .......................................................... 17
2.4 The quantity $\mu$ is removed from the flux into the node, therefore at most one queue is active at each instant in time. ............ 18
2.5 Model comparison: Time evolution of the solution of the FIFO model. .................................................. 25
2.6 Model comparison: Time evolution of the solution of the non-FIFO model. .......................................... 26
2.7 Model comparison: Time evolution of the solution of the FIFOQ model. ............................................ 27
2.8 More complex example: Time evolution of the FIFOQ solution. ......................................................... 30
2.9 Two examples of how road sharing ratios are determined based on road geometry. .............................. 31
2.10 As queue 3 increases, road sharing ratio $c_2$ decreases. ........................................................... 33

3.1 A positive speed rarefaction wave hits a positive speed Big shock on the outgoing road, to generate a negative speed Big shock. 50
3.2 A piecewise constant approximation of the initial data. ........................................................ 55
3.3 On the incoming road, interaction of a small rarefaction wave or a small shock with a Big shock with negative speed can produce a positive speed Big shock. ........................................ 64
3.4 On the outgoing roads, interaction of a small rarefaction wave or a small shock with a Big shock with positive speed can produce a negative speed Big shock. ........................................ 64
3.5 Construction of "generalized tangent vectors" on the roads. .. 79

4.1 An example of how to correct queue length calculations. ... 97
4.2 A ring road with 10 edges. ........................................ 97
4.3 Fundamental diagram for a one-lane road network. ............ 98
4.4 Speed-density relationship diagram for a one-lane road network. 99
4.5 Fundamental diagram for a three-lane road network. ........... 101
4.6 Speed-density relationship diagram for a three-lane road network. 101
4.7 Fundamental diagram for a four-lane road network. ............ 102
4.8 Speed-density relationship diagram for a four-lane road network. 102
4.9 A network with an exit-only off-ramp. ............................ 103
4.10 Exit-only off-ramp network results for $t = 10s$. ............... 105
4.11 Exit-only off-ramp network results for $t = 15s$. ............... 106
4.12 Exit-only off-ramp network results for $t = 25s$. ............... 106
4.13 Exit-only off-ramp network results for $t = 80s$. ............... 107
4.14 Exit-only off-ramp network results for $t = 100s$. .......... 107
4.15 Exit-only off-ramp network results for $t = 110s$. .......... 108
4.16 Exit-only off-ramp network results for $t = 250s$. .......... 108
4.17 Exit-only off-ramp network results for $t = 260s$. .......... 109
4.18 Exit-only off-ramp network results for $t = 290s$. .......... 109
4.19 Exit-only off-ramp network results for $t = 300s$. .......... 110
4.20 Total type 2 vehicle tallies for the exit-only off-ramp scenario. 111
4.21 Total type 3 vehicle tallies for the exit-only off-ramp scenario. 112
4.22 A network with a non-exit-only off-ramp. ........................ 113
4.23 Non-exit-only off-ramp network results for $t = 10s$. .......... 114
4.24 Non-exit-only off-ramp network results for $t = 15s$. .......... 115
4.25 Non-exit-only off-ramp network results for $t = 25s$. .......... 115
4.26 Non-exit-only off-ramp network results for $t = 80s$. .......... 116
4.27 Non-exit-only off-ramp network results for $t = 100s$. ....... 116
4.28 Non-exit-only off-ramp network results for $t = 110s$. ....... 117
4.29 Non-exit-only off-ramp network results for $t = 250s$. ....... 117
4.30 Non-exit-only off-ramp network results for $t = 260s$. ....... 118
4.31 Non-exit-only off-ramp network results for $t = 290s$. ....... 118
4.32 Non-exit-only off-ramp network results for $t = 300s$. ....... 119
4.33 Total type 2 vehicle tallies for the non-exit-only off-ramp scenario. 120
4.34 Total type 3 vehicle tallies for the non-exit-only off-ramp scenario. 121
LIST OF TABLES

2.1  Total vehicle flow after $t = 25$ min.  . . . . . . . . . . . . . . . . . 23
CHAPTER 1

INTRODUCTION

Transportation systems are among the critical infrastructures in every society. Designing robust and reliable transportation networks requires a solid understanding of the behavior of the traffic flow in these networks. Mathematical modeling of traffic flow on road network via partial differential equations (PDE) has been studied since 1950s, when Lighthill and Whitham [59] developed a model to describe the traffic flow based on methods from fluid dynamics. Since then, many models have been proposed to describe the vehicular traffic flow. These models can be broadly categorized into three classes: microscopic models, macroscopic models, and mesoscopic models. The microscopic models are the most detailed models. They simulate single vehicle units based on car-following and lane-changing theories. Therefore, the dynamic variables of the models represent microscopic properties like the position and velocity of single vehicles. In contrast to the microscopic models, macroscopic models are less concerned about the behavior of individual cars and focus more on the aggregate behavior of the vehicles. Mesoscopic models fill the gap between microscopic and macroscopic models. These models combine the properties of both microscopic and macroscopic simulation models. They describe the traffic entities at a high level of detail, however the behavior and interaction of the vehicles are modeled at a lower level of detail.

Each of these traffic flow modeling approaches has its own advantages and
disadvantages that make it more beneficial in some situations and less so in other situations. In this dissertation, we focus on the macroscopic way of modeling the traffic flow. For a comprehensive review and comparison of various modeling approaches to describe the vehicular traffic flow dynamics, see [42].

Macroscopic models describe the evolution of traffic on roads using the macroscopic variables such as density and average velocity. The macroscopic model that contains a single continuity equation which is based on the conservation of the vehicles is called the “first order” model. The Lighthill-Whitham-Richards (LWR) model is a first order macroscopic model and represents the traffic behavior through the continuity equation:

$$\rho_t + f_x = 0,$$

where $\rho(x, t)$ is the vehicle density in location $x$ and time $t$, and $f(\rho, v)$ is the flux function given by $\rho v$. The average velocity $v(x, t)$ is uniquely given by the density. In fact, velocity is a decreasing function of the density, $v = v(\rho)$ and hence, $f = f(\rho) = \rho v(\rho)$.

The empirical relationship between flux and density can be graphically shown in the fundamental diagram (FD). Fundamental diagram presents the equilibrium relationship between the flow rate $f$ and the density $\rho$. This relationship has been first introduced by Greenshields in [32]. Figures 1.1 and 1.2 show two customary fundamental diagrams.
When modeling vehicular traffic flow on highway networks via macroscopic models, suitable coupling conditions at the network intersections are crucial. The coupling conditions describe the flow balance of vehicles between incoming and outgoing roads. Many studies have been conducted to describe, control and predict the traffic flow on networks. However, there are still some shortcomings in the existing literature that need to be addressed. For example, there are currently no satisfactory macroscopic coupling models for off-ramps on the highways. In this study we consider a network of an off-ramp junc-
tion, namely, one junction with one incoming and two outgoing roads such that one of the outgoing roads is an off-ramp. The evolution of traffic flow on each network edge is usually described in a lane-aggregated fashion using a single-class LWR model. At off-ramps, split ratios (i.e., what percentage of traffic exits the highway) are prescribed that can be drawn from historic data. In Chapter 2, we first present the existing models and highlight their shortcomings. Specifically, we consider a 1-in-2-out junction with a clogged off-ramp. In this situation, we show that the classical FIFO (First In, First Out) coupling conditions yield unrealistic results, in that a clogged off-ramp yields zero flux through the junction. As a remedy, non-FIFO conditions have been proposed in the literature [56]. However, we show that those lead to spurious re-routing of vehicles. We then introduce new coupling conditions to address the shortcomings of the existing models by using a vertical queue at the junction. We then discuss how the new model, called FIFO with queue (FIFOQ), can be used in conjunction with a cell transmission method (CTM). CTM is a numerical method used to predict the macroscopic traffic behavior on roads by evaluating the flow and density at finite number of intermediate points at different time steps [16]. The three models are then systematically compared in a representative example describing a first forming and then clearing off-ramp queue. The example illustrates how the new model remedies the weaknesses of both existing models.

The introduction of vertical queues in this model as well as the requirement of the model to preserve constant split ratios lead to some key differences with the previous models in the literature. After developing the model, a natural concern is its well-posedness. In Chapter 3, we mathematically analyze these differences. Specifically, we prove the well-posedness of the model. In this regard, existence of the solution is proved by construction using the wave-front tracking algorithm [4, 41]. For the proof of continuous dependence of the solution to the initial data and its uniqueness, we use the technique of generalized tangent vectors [3, 5].

In Chapter 4, we investigate the performance of our proposed model with
data. Specifically, we use artificially generated data from a microsimulator called SUMO (Simulation of Urban MObility) to validate our models. SUMO is a free, open, microscopic and space-continuous road traffic simulation suite designed to handle large road networks. We use the data generated by SUMO as a proxy for real traffic data. In this chapter, we first describe the fundamentals of microscopic models, car-following models and lane-changing models. Then we explain the specific car-following and lane-changing model used in this chapter and their parameters. We then establish micro-simulation representations of the off-ramp scenarios. We also discuss how to systematically extract macro quantities, including vehicle density and queue length, from the results of the microsimulator. Finally, we compare the results of the macroscopic models with the macro quantities extracted from the microsimulator.
CHAPTER 2

OFF-RAMP COUPLING CONDITIONS

One way to model vehicular traffic flow on a highway network is to describe intersections as nodes (vertices) of a graph, and the roads between them as edges of the graph. The flow on an edge in a given direction is then described via a lane-aggregated Lighthill-Whitham-Richards (LWR) model [60, 71], and suitable coupling conditions are formulated on each node that describe the flow balance of vehicles between incoming and outgoing edges. The classical FIFO coupling conditions were introduced by Daganzo [17]. Mathematical proofs of well-posedness of the resulting hyperbolic conservation law network flow were provided in [40, 36, 11, 25] and references therein.

Figure 2.1: A road network consisting of three roads and a single junction. The picture illustrates a diverge junction where one outgoing road, $I_3$, is an off-ramp. $\alpha_2$ and $\alpha_3$ are the split ratios (see Section 2.1.3)

A key modeling shortcoming of FIFO conditions is that, as shown in section
2.2.1, a clogged off-ramp will result in zero flow through the node. Clearly, on multi-lane highways this is unrealistic, as a queue forming from the off-ramp, if it is not too long, will be restricted to the right-most lane, and vehicles that do not wish to exit can pass the queue (on the left side of the queue). To remedy this shortcoming, non-FIFO coupling conditions were proposed [56]. These allow for a nonzero flow past a clogged off-ramp, however, at the expense of violating route-choices of drivers: As shown in section 2.2.2, some vehicles that in reality form a queue waiting to exit will instead be accounted for as flow continuing along the highway. Other more recent off-ramp models possess similar re-routing effects [57].

What is needed are coupling conditions that remedy the shortcoming of FIFO, but without producing spurious re-routing of vehicles. One possible way to achieve this is by formulating multi-commodity models, such as by Daganzo [18] or Bressan and Nguyen [2]. These models replace LWR by a system of conservation laws, i.e., they explicitly track different types of vehicles. This simplifies some challenges in formulating diverge models, e.g., no split ratio must be provided. The practical challenge is that, to a traffic operator or real-time simulation, it is not known in advance which vehicles intend to exit the highway and which do not. Instead, typically split ratios tend to be known from historic data. We therefore consider here the situation in which a single-class lane-aggregated LWR model is to be used on the edges, as for instance done in the Mobile Millennium project [63].

This study presents how this goal can be achieved by the introduction of a vertical queue that keeps track of the excess vehicles of a certain type (exiting vs. non-exiting) that may join congested traffic by more than the other vehicle type does. While vertical queues have been used in macroscopic traffic models for the “storage” of vehicles that wish to enter the network [21, 2], their use for the purpose of balancing splitting flows is novel. A vertical queue could be present on either of the two out-roads (the highway or the ramp), however, it is always designed to be minimal, i.e., two queues cannot be active simultaneously.
This chapter is organized as follows. First, the mathematical definitions and notations are introduced, and a discussion of split ratios vs. turn ratios is provided. Then, the existing coupling models, FIFO and non-FIFO, are presented, and their modeling shortcomings discussed. After that, the new model is presented, including how it can be implemented in a cell-transmission model (CTM) framework. The three models are then systematically compared in a representative example describing a first forming and then clearing off-ramp queue. The example illustrates how the new model remedies the weaknesses of both existing models. The chapter closes with further discussion and extension of the new model.

2.1 Modeling and Computational Foundations

We consider models for road networks that are represented by a directed graph, whose edges represent the roads (all lanes going in one direction aggregated), and nodes (vertices) represent the intersections, also called junctions. An edge \( i \) of the network is an interval \( I_i = [a_i, b_i] \). This study specifically focuses on 1-in-2-out nodes modeling highway off-ramps. Therefore, the discussion is restricted to one node with three edges: one in-road \( (I_1) \) and two out-roads \( (I_2 \text{ and } I_3) \).

On each edge \( i \), the evolution of the traffic density, \( \rho_i(x,t) \), is described by the LWR model

\[
\frac{\partial}{\partial t} \rho_i + \frac{\partial}{\partial x} f(\rho_i) = 0, \quad \text{where } (x,t) \in I_i \times \mathbb{R}^+. \tag{2.1}
\]

The flux function \( f = f(\rho) = \rho v(\rho) \) encodes the fundamental diagram (FD) of traffic flow, where \( v = v(\rho) \) is the bulk velocity of traffic. In this dissertation, we use the Greenshields flux

\[
f(\rho) = v^{\max} \rho \left( 1 - \frac{\rho}{\rho^{\max}} \right),
\]

corresponding to an affine linear density–velocity relationship, with \( v^{\max} \) the speed limit and \( \rho^{\max} \) the jamming density. However, it is important to stress
that the model, its CTM implementation, and its analysis, apply to any concave down flux function, including triangular FDs. The critical density at which the flow is maximized is denoted with $\sigma$.

### 2.1.1 Riemann Problem and Cell Transmission Model

The key building block for finite volume discretizations of the LWR model (2.1) is the Riemann problem (RP), which is a Cauchy problem with initial data

$$\rho(x, 0) = \rho_0(x) = \begin{cases} 
\rho_L & x < 0 , \\
\rho_R & x \geq 0 . 
\end{cases}$$

By standard theory of hyperbolic conservation laws [58], the RP (with concave flux) has the following unique entropy solution:

- If $\rho_L < \rho_R$, then $f'(\rho_L) > f'(\rho_R)$, and then the solution
  $$\rho(x, t) = \begin{cases} 
\rho_L & x < st , \\
\rho_R & x \geq st . 
\end{cases}$$
  consists of a shock, i.e., a traveling discontinuity in which vehicles brake (e.g., the upstream end of a traffic jam). The shock speed $s = \frac{f'(\rho_L) - f'(\rho_R)}{\rho_R - \rho_L}$ is given by the secant slope in the FD. This value is derived from the Rankin-Hugoniot conditions $s(\rho_R - \rho_L) = f(\rho_R) - f(\rho_L)$, that describe the relationship between the states on both sides of a shock wave [70].

- If $\rho_L \geq \rho_R$, then $f'(\rho_L) \leq f'(\rho_R)$, and the solution
  $$\rho(x, t) = \begin{cases} 
\rho_L & x < f'(\rho_L)t , \\
(f')^{-1}\left(\frac{x}{t}\right) & f'(\rho_L)t \leq x < f'(\rho_R)t , \\
\rho_R & x \geq f'(\rho_R)t , 
\end{cases}$$
  is a rarefaction wave, in which vehicles gradually accelerate.
The RP is the key building block of the Godunov method [31], which divides each edge into finite volume cells, and updates the average density in each cell by the numerical fluxes across cell boundaries. Those fluxes are the RP solutions, evaluated at the cell interface. The cell transmission model (CTM) [16] is equivalent to the Godunov method, applied to the LWR model.

An important conceptual interpretation of the Godunov fluxes is in terms of supply and demand functions [37]. Given a concave down flux function \( f(\rho) \) with critical density \( \sigma \), the demand and supply functions

\[
\gamma_d(\rho) = \begin{cases} 
0 & 0 \leq \rho < \sigma \\
0 & \sigma < \rho \leq 1 
\end{cases}, \quad \text{and} \quad \gamma_s(\rho) = \begin{cases} 
0 & \rho < \sigma \\
\sigma & \sigma < \rho < 1 
\end{cases},
\]

are the non-decreasing and non-increasing components of the flux function, respectively.

\[ (2.2) \]

The flux of vehicles through an interface between two cells is then the maximum possible value that does not exceed the demand (on road capacity) imposed by the upstream cell (L), and the supply (of road capacity) that the downstream cell (R) provides:

\[
F(\rho_L, \rho_R) = \min (\gamma_d(\rho_L), \gamma_s(\rho_R)).
\]

\[ (2.3) \]
For the implementation of a Godunov scheme, respectively CTM, the interface flux (2.3) is all that is needed. However, the full solution of the RP can also be constructed, as follows. On each cell (upstream and downstream), there are two possibilities: if the flux (2.3) matches the flux \( f(\rho) \) in that cell, then the constant state remains as it is; otherwise, a new state \( \hat{\rho} \) emerges at the cell interface that reproduces the interface flux, i.e., \( f(\hat{\rho}) = F \). Of the two solutions that this equation generally possesses, the one is chosen that results in a wave that travels away from the cell interface, that is: the congested state \( \hat{\rho} > \sigma \) on the upstream cell; and the free-flow state \( \hat{\rho} < \sigma \) on the downstream cell.

### 2.1.2 Generalized Riemann Problem

In a Godunov/CTM discretization of a road network, a network node can be treated in the same fashion. A generalized Riemann Problem (GRP) is given by a constant density state near the junction and on each edge. For a general node, the demands of all in-roads and the supplies of all out-roads are computed, and by a route choice matrix that determines how the in-fluxes wish to distribute into the out-fluxes, the resulting vehicle flows are constructed so that they never exceed their respective supply/demand values [25].

In this study, we focus on the 1-in-2-out case. The GRP considers a constant density \( \rho_1 \) on the in-road \( I_1 \) (highway), a constant density \( \rho_2 \) on the out-road \( I_2 \) (highway), and a constant density \( \rho_3 \) on the out-road \( I_3 \) (off-ramp). A node coupling model (or coupling condition) \( \Phi \) is then a mapping from those three densities to three new fluxes \( \Phi(\rho_1, \rho_2, \rho_3) = (\Gamma_1, \Gamma_2, \Gamma_3) \), where the flux on the in-road matches the sum of the two out-road fluxes, \( \Gamma_1 = \Gamma_2 + \Gamma_3 \).

Note that by the same construction as in the simple RP, one also obtains three new states \( (\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3) \) emanating at the node’s position on each edge, but as we are interested in Godunov/CTM discretizations, the construction of the fluxes \( \Gamma_i \) suffices.
2.1.3 Split Ratio

In a 1-in-2-out node, drivers make route choice decisions. Therefore, realistic coupling models must require an additional piece of information. This is commonly assumed to be a “split ratio” that prescribes what ratios of the incoming flow proceeds onto which of the two out-roads. For a general node with multiple in- and out-roads, a split ratio matrix is needed [25]. In the 1-in-2-out case, the split ratios are given by two numbers $\alpha_2$ and $\alpha_3$, with $\alpha_2 + \alpha_3 = 1$, corresponding to the two out-roads $I_2$ and $I_3$, respectively.

While mathematically, the notion of a split ratio (matrix) is easy to accept, its practical/modeling rationale in fact calls for a careful discussion. If both out-roads are in free-flow, then what truly determines how many vehicles exit at a given time is the composition of the incoming traffic flow into “type 2” (intending to continue on the highway) and “type 3” (intending to exit) vehicles. Unfortunately, this “type ratio” is generally not known. (The type may be known for a few vehicles; and future V2X connectivity \(^1\) may substantially increase that knowledge; but for now, the type ratio is not known.) Therefore, a historic “exit ratio” is used as a proxy for the unknown “type ratio”. Assuming that traffic behaves relatively similar from week to week, and assuming that exit ratios evolve slowly (relative to the flow dynamics) in time, historic data on how many vehicles have exited at a certain time of day can be used to define the (quasi-constant-in-time) split ratio $\alpha_2 + \alpha_3 = 1$.

The problem with this approach, i.e. using the historic exit ratio as a proxy for the unknown type ratio, is that the type ratio and the exit ratio are not necessarily the same. As detailed below, they are identical if both out-roads are in free-flow, or if the split happens in a way that passing of other vehicles is impossible. However, at a highway off-ramp, neither of these assumptions needs to be satisfied. As an extreme example, consider an off-ramp that is completely clogged ($\rho_3 = \rho_3^{\max}$) due to an incident. Hence, no vehicle flow

\(^1\)Vehicle-to-everything (V2X) communication is the passing of information from a vehicle to any entity that may affect the vehicle, and vice versa.
occurs on the ramp, and type 3 vehicles will start to queue up on the highway. However, a multi-lane highway generally allows type 2 vehicles to pass this queue (to some extent, cf. the empirical study in [64]). Consequently, the type ratio of vehicles that are upstream of the ramp will change and gradually shift towards more and more type 3 vehicles. However, this is not due to an actual change in upstream traffic, but rather due to the clogged (downstream) off-ramp.

In the following, we present two classical coupling models, that both fail to capture this situation correctly (for different reasons). We then present a new model that remedies the problems.

2.2 Existing Models

As above, we consider a GRP at a 1-in-2-out node, i.e., states $\rho_1$, $\rho_2$, and $\rho_3$ are given. Using equations (2.2), we obtain the in-road demand $\gamma_1^d$, and the out-road supplies $\gamma_2^s$ and $\gamma_3^s$. Then, using the split ratio $\alpha_2 + \alpha_3 = 1$, the “partial demands” [43] are given as

$$\gamma_{21}^d = \alpha_2 \gamma_1^d \quad \text{and} \quad \gamma_{31}^d = \alpha_3 \gamma_1^d.$$ 

Below we always assume that traffic actually “splits”, i.e., $\alpha_2 > 0$ and $\alpha_3 > 0$.

2.2.1 FIFO Model

The FIFO coupling model [17] is based on the assumption that the actual exit ratio equals the split ratio under all circumstances. Hence, the new fluxes satisfy

$$\Gamma_2 = \alpha_2 \Gamma_1 \quad \text{and} \quad \Gamma_3 = \alpha_3 \Gamma_1.$$ 

As described above, the GRP requires that the new fluxes do not exceed the respective supplies/demands on the edges, i.e., $0 \leq \Gamma_1 \leq \gamma_1^d$, $0 \leq \Gamma_2 \leq \gamma_2^s$, and $0 \leq \Gamma_3 \leq \gamma_3^s$. Using (2.4), the latter two conditions can be re-written as
$0 \leq \Gamma_1 \leq \frac{1}{\alpha_2} \gamma_2^s$ and $0 \leq \Gamma_1 \leq \frac{1}{\alpha_3} \gamma_3^s$. Maximizing the flux through the node under those constraints determines the fluxes as

$$\Gamma_1 = \min \left( \frac{\gamma_1^d}{\alpha_1}, \frac{\gamma_2^s}{\alpha_2}, \frac{\gamma_3^s}{\alpha_3} \right), \quad \Gamma_2 = \alpha_2 \Gamma_1, \quad \Gamma_3 = \alpha_3 \Gamma_1.$$ 

The FIFO model ensures that the resulting fluxes are always distributed according to the prescribed split ratio. Therefore, in the case of a clogged off-ramp $I_3$ (but free-flow $I_2$), FIFO would result in zero flow through the node, $\Gamma_1 = \Gamma_2 = \Gamma_3 = 0$. Clearly, this is unrealistic on multi-lane highways that allow vehicles to by-pass queues (to some extent), and vehicles waiting to pass through the node do not necessarily hold up all traffic. In other words, highways are clearly not FIFO.

### 2.2.2 Non-FIFO Model

Using partial demands, the out-fluxes in FIFO can equivalently be written as

$$\Gamma_j = \min \left( \frac{\gamma_j^d}{\alpha_j \gamma_2^s}, \frac{\alpha_j \gamma_3^s}{\alpha_3} \right), \quad j = 2, 3.$$ 

The idea of the non-FIFO model, proposed in [56], is to associate the supply constraints of each out-road only with the flux on that respective road, leading to the model

$$\Gamma_2 = \min(\gamma_{21}^d, \gamma_2^s), \quad \Gamma_3 = \min(\gamma_{31}^d, \gamma_3^s), \quad \Gamma_1 = \Gamma_2 + \Gamma_3.$$ 

A physical interpretation of this model is that, before reaching the node, drivers are already presorted according to the respective out-road that they plan to take. Then, respective portions of the road width are allocated to the drivers according to the split ratio. Thus, each type of driver can pursue their destination without impediment from the other type.

By construction, the non-FIFO model does not incur the blockage problem that FIFO incurs. In the case of a clogged off-ramp (i.e., $\gamma_3^s = 0$), one has $\Gamma_3 = 0$ (as it has to be), but $\Gamma_2 > 0$ in general. Type 2 vehicles can pass
the off-ramp queue, and the flux $\Gamma_2$ is determined solely by the demand and supply of the highway segments.

Unfortunately, the non-FIFO model suffers from a different modeling problem. It assumes that the split ratios are given and independent of the actual density states and fluxes. However, because the out-fluxes are generally not distributed according to the split ratio (i.e., $\frac{\Gamma_2}{\Gamma_3} \neq \frac{\alpha_2}{\alpha_3}$), vehicles of one type will actually become more prevalent upstream of the node than vehicles of the other type. The non-FIFO model ignores this fact: once the clogged off-ramp becomes free-flow again, vehicles of type 3 will start flowing again; however, the fact that a queue of type 3 vehicles should be present is ignored. This results in a spurious re-routing of vehicles: vehicles that should have taken the off-ramp will be allocated to out-flow on the highway instead.

To recap, the non-FIFO model remedies the unrealistic blockage imposed by the FIFO model. However, it does so at the expense of unphysical re-routing effects. It should be stressed that under certain circumstances, drivers in reality may in fact change their route choices based on the actual traffic state; however, this is not always possible, and a model should not produce re-routing as an unwanted side effect. To describe intentional route changing, models that allow the split ratios to depend on the traffic state have been proposed [35]. However, these models do not remove the fundamental flaws of FIFO and non-FIFO. A methodology that would resolve the re-routing problem is to explicitly track and evolve the split ratios (or more accurately: the type ratios) as they move along the incoming edge, as proposed for instance in [18, 2]. In addition to being computationally substantially more demanding, such explicit multi-commodity models are held back by the aforementioned fundamental challenge that the route choices of vehicles are generally not accessible in advance.

The question is therefore, what can be done to remedy the problems of both models (FIFO and non-FIFO) within the framework of single-class, lane-aggregated, macroscopic models? The new model should still be based on historic split ratios; it should allow for passing of queued vehicles; it should not lose any vehicles (that are waiting in queues); and it should account for
the type of queued vehicles. As the formation of a queue, caused by reduced
supply of one out-road, introduces non-local-in-time effects (vehicles queued
up initially may still wait later in time), imbalances in the vehicle type com-
position among queued vehicles must, in some form, be tracked. Next, we
construct a model that does so, with the minimal amount of additional infor-
mation stored, namely the excess of vehicles of a certain type, relative to the
actual split ratio.

2.3 New FIFOQ Model

We now derive a new model that remedies both shortcomings of FIFO and
non-FIFO, called FIFOQ (“FIFO with Queue”). As argued above, the proper
evolution of waiting and passing vehicles requires the model to be augmented
by some additional variable that accounts for the composition of backed-up
vehicle types. We choose to introduce a local vertical queue as that additional
variable. Because vertical queues improperly capture the non-local impact of
true vehicle jamming, we derive a model that minimizes the impact of the
vertical queue, so that its sole purpose is to account for excess vehicles of a
certain type. It should be stressed that the use of vertical queues in traffic
models is not novel at all. For instance, vertical queues are a common means
to implement vehicles that enter the network [21, 2, 8, 7, 6]. In contrast, the
usage of queues for the purpose of properly tracking vehicle types upstream of
an off-ramp is a novel concept.

The fundamental dynamics of a vertical queue are that its rate of change
equals inflow into its upstream end, \( \Gamma_{in} \), minus outflow out of its downstream
front, \( \Gamma_{out} \), i.e., \( \dot{m} = \Gamma_{in} - \Gamma_{out} \), where \( m \) denotes the number of vehicles stored
in the queue [78].

To derive the new model, we start with a preliminary setup: a non-FIFO
model with queues. Recall that the non-FIFO model can be interpreted as
traffic being presorted and vehicle types given a share of the in-road width
according to the split ratio. We therefore think of separating the in-road into
two parts and consider two independent 1-in-1-out nodes with queues $m_2$ and $m_3$, see Figure 2.3(a). These queues could now be used, for example, to prevent backward going shocks on the in-road.

However, this is not an adequate model because it does not guarantee a preservation of the split ratio. Moreover, two queues may form simultaneously.

![Figure 2.3: Interpretation/derivation of two models with queues.](image)

(a) A non-FIFO model with queues.  
(b) New model: 1-in-2-out FIFO node with queue (“FIFOQ”).

We therefore extend the model by adding additional components. As shown in Figure 2.3(b), a 1-in-2-out FIFO node serves to sort the vehicles into two groups based on their destinations. Next downstream, a free-flow section follows, which consists of two independent pipes with flux functions $f_{1i} = v_{1i}^\text{max} \rho(1 - \frac{\rho}{\rho_{1i}^\text{max}})$ for $i = 2, 3$, where $c_2$ and $c_3$ are the road sharing ratios of different vehicle types. In many situations, one highway lane will be associated with the off-ramp, while the remaining lanes are associated with the flow past the off-ramp traffic, and these geometric considerations could be incorporated into the model. Here, we use a simplifying assumption, namely that the road is divided precisely according to the split ratio, i.e., $c_2 = \alpha_2$ and $c_3 = \alpha_3$. In Section 2.6, we will discuss the case in which the road sharing ratios are different from the split ratios, namely $c_j \neq \alpha_j$, for $j \in \{2, 3\}$. The reason why we can assume this two-pipe region to be in free-flow is that the two queues, $m_2$ and $m_3$, can be used to absorb any congestion that emanates from the two out-roads. Of course, these individual components are for model derivation purposes only; in the end, the whole model is “collapsed to zero length” to yield a single coupling condition.
Thanks to the queues, this model allows for traffic to pass through the junction while still respecting the split ratio, even when one out-road is clogged. However, the model can further be improved. In the case when both out-roads are clogged (or provide sufficiently low supply), this current model would develop two queues, one in each pipe. In turn, it would never create a congested state on the in-road. That is unsatisfactory, as the impact that a real traffic jam would have on the road conditions further upstream would not be seen. We therefore modify the model to allow for backward propagating shocks on the in-road, as long as that congested state is composed according to the split ratio. We therefore introduce a quantity $\mu$ that is removed from the flux into the node, so that at most one queue is active at any instant in time, see figure 2.4. That one active queue then tracks the excess of vehicles of a certain type that are more prevalent near the off-ramp than the split-ratio would dictate. Let $\bar{\Gamma}_1, \bar{\Gamma}_{12}, \bar{\Gamma}_{13}, \bar{m}_2,$ and $\bar{m}_3,$ denote the fluxes and queue values before the removal of $\mu$. We assume that the two-pipe region is in free-flow initially. Thus, the supplies of the pipes, $\gamma_{12}$ and $\gamma_{13}^s$, are always maximal.

Because we here assume that traffic divides into two pipes whose width ratio equals the split ratio, we simply have $\bar{\Gamma}_1 = \gamma_1^d$ (if the pipe widths were different, $\bar{\Gamma}_1 < \gamma_1^d$ could occur). Moreover, $\bar{\Gamma}_{12} = \alpha_2\bar{\Gamma}_1 = \alpha_2\gamma_1^d$ and $\bar{\Gamma}_{13} = \alpha_3\bar{\Gamma}_1 = \alpha_3\gamma_1^d$.

Collapsing the pipes to have length zero implies that the flux out of them
equals the flux into them. Therefore, the fluxes into the 1-in-1-out nodes (or queues) are $\hat{\Gamma}_{12}$ and $\hat{\Gamma}_{13}$, respectively. If there are no queues, the out-road fluxes are

$$\Gamma_2 = \min (\alpha_2 \gamma^d_1, \gamma^s_2) \quad \text{and} \quad \Gamma_3 = \min (\alpha_3 \gamma^d_1, \gamma^s_3).$$

(2.5)

In turn, if a queue is active, i.e., $\hat{m}_i > 0$, then the corresponding out-flux is maximal, i.e., $\Gamma_i = \gamma^s_i$. The evolution of the queues is described by the differences in the fluxes: $\dot{\hat{m}}_2 = \hat{\Gamma}_{12} - \Gamma_2 = \alpha_2 \hat{\Gamma}_1 - \Gamma_2$ and $\dot{\hat{m}}_3 = \hat{\Gamma}_{13} - \Gamma_3 = \alpha_3 \hat{\Gamma}_1 - \Gamma_3$.

As by the prior discussion, we now determine $\mu$ such that at most one queue is active at any time. We do this because, when both out-roads are in the congestion regime, having two queues at the junction will never lead to a congested $I_1$, which is unrealistic. Now, if one queue is active, for example $\hat{m}_2 > 0$, and the other queue starts filling, i.e., $\hat{m}_3 > 0$, then $\mu = \hat{m}_2 / \alpha_3$. With $\Gamma_1 = \gamma^d_1 - \mu$, we obtain for the in-flux

$$\Gamma_1 = \min \left( \gamma^d_1, \frac{\gamma^s_2}{\alpha_3} \right).$$

Conversely, if no queue is active, i.e., $\hat{m}_2 = \hat{m}_3 = 0$ initially and both queues start stacking up the vehicles, we have

$$\mu = \min \left( \frac{\dot{\hat{m}}_2}{\alpha_2}, \frac{\dot{\hat{m}}_3}{\alpha_3} \right).$$

(2.6)

By combining equations (2.8) and (2.10), we have

$$\mu = \min \left( \gamma^d_1 - \min \left( \gamma^d_1, \frac{\gamma^s_2}{\alpha_2} \right), \gamma^d_1 - \min \left( \gamma^d_1, \frac{\gamma^s_3}{\alpha_3} \right) \right)$$

$$= \gamma^d_1 - \min \left( \gamma^d_1, \max \left( \frac{\gamma^s_2}{\alpha_2}, \frac{\gamma^s_3}{\alpha_3} \right) \right),$$

and with $\Gamma_1 = \gamma^d_1 - \mu$, we obtain for the in-flux

$$\Gamma_1 = \min \left( \gamma^d_1, \max \left( \frac{\gamma^s_2}{\alpha_2}, \frac{\gamma^s_3}{\alpha_3} \right) \right).$$

Therefore the complete model reads as:
If \( m_2 = m_3 = 0 \):

\[
\begin{align*}
\Gamma_1 &= \min \left( \gamma_d^1, \max \left( \frac{\gamma_s^2}{\alpha_2}, \frac{\gamma_s^3}{\alpha_3} \right) \right), \\
\Gamma_2 &= \min \left( \alpha_2 \gamma_1^d, \gamma_2^s \right), \\
\Gamma_3 &= \min \left( \alpha_3 \gamma_1^d, \gamma_3^s \right), \\
\Gamma_1 &= \min \left( \gamma_d^1, \frac{\gamma_s^2}{\alpha_2} \right), \\
\Gamma_2 &= \gamma_2^s, \\
\Gamma_3 &= \min \left( \alpha_3 \gamma_1^d, \gamma_3^s \right), \\
\Gamma_1 &= \min \left( \gamma_d^1, \frac{\gamma_s^2}{\alpha_2} \right), \\
\Gamma_2 &= \min \left( \alpha_2 \gamma_1^d, \gamma_2^s \right), \\
\Gamma_3 &= \gamma_3^s,
\end{align*}
\]

If \( m_2 > 0 \):

\[
\begin{align*}
\Gamma_2 &= \gamma_2^s, \\
\Gamma_3 &= \min \left( \alpha_3 \gamma_1^d, \gamma_3^s \right), \\
\Gamma_2 &= \min \left( \alpha_2 \gamma_1^d, \gamma_2^s \right), \\
\Gamma_3 &= \gamma_3^s,
\end{align*}
\]

If \( m_3 > 0 \):

\[
\begin{align*}
\Gamma_2 &= \gamma_2^s, \\
\Gamma_3 &= \min \left( \alpha_3 \gamma_1^d, \gamma_3^s \right), \\
\Gamma_2 &= \min \left( \alpha_2 \gamma_1^d, \gamma_2^s \right), \\
\Gamma_3 &= \gamma_3^s,
\end{align*}
\]

\[
\begin{align*}
\dot{m}_2 &= \alpha_2 \Gamma_1 - \Gamma_2, \\
\dot{m}_3 &= \alpha_3 \Gamma_1 - \Gamma_3.
\end{align*}
\]

By construction, this model never generates more than one queue to be active (given that at most one queue is active initially): if \( m_3 > 0 \), then the model dynamics automatically imply that \( \dot{m}_2 = 0 \), and vice-versa. Therefore, as soon as one queue becomes active, the other one remains constant at zero. Moreover, in the case of a singular split ratio, such as \( \alpha_3 = 0 \), we let \( \mu = \frac{\dot{m}_2}{\alpha_2} \), in which case the model reduces to the standard FIFO model for a 1-in-1-out node.

### 2.4 Cell Transmission Discretization

We now describe how the new model is discretized into a CTM, by suitably augmenting the Godunov scheme [31] with a treatment of the queue evolution.

#### 2.4.1 Approximation Along Edges

Let space and time be discretized via a regular grid, where

(i) \( \Delta x \) is the cell size;
(ii) $\Delta t$ is the time step, adhering to the necessary Courant–Friedrichs–Lewy (CFL) stability condition \cite{13}, $\Delta t \leq \frac{\Delta x}{\max |f(\rho)|}$, where the maximum is taken over all flux functions in the network; and

(iii) $(x_j, t^n) = (j \Delta x, n \Delta t)$ are the space-time grid points, and $\rho^n_j$ denotes the (average) density on cell $j$ at time $t_n$.

Using $\lambda = \frac{\Delta t}{\Delta x}$, the Godunov scheme along an edge reads as

$$\rho^{n+1}_j = \rho^n_j - \lambda(F(\rho^{n+1}_j, \rho^n_j) - F(\rho^n_{j-1}, \rho^n_j)),$$  \hspace{1cm} (2.7)

where, for internal cell boundaries, the numerical flux

$$F(\rho_L, \rho_R) = \begin{cases} 
\min_{\rho_L \leq \rho \leq \rho_R} f(\rho) & \text{if } \rho_L \leq \rho_R, \\
\max_{\rho_{\text{fr}} \leq \rho \leq \rho_R} f(\rho) & \text{if } \rho_R \leq \rho_L,
\end{cases}$$

equals the standard CTM flux. At terminal cells of an edge, i.e., adjacent to a network node, the in-flux or out-flux in equation (2.7) are replaced by the respective flux $\Gamma_i$ defined by the coupling model. Moreover, at boundaries that are not connected to a node, ghost cells are used (cf. \cite{11}).

### 2.4.2 Treatment of Queues

The proper time-stepping of the model with queues must take into account that a queue may deplete during a time step, see \cite{21}. For simplicity of notation, we describe the situation of a 1-in-1-out node with a vertical queue $m$, i.e., $\dot{m} = \Gamma_1 - \Gamma_2$. At each time, we must determine the new length of the queue. If the queue is filling, the increment is simply added to $m^n$. However, if the queue is emptying, we must calculate the time of queue depletion, $\bar{t} = t^n + \frac{m^n}{\Gamma_2 - \Gamma_1}$, and compare it with the time $t^{n+1} = t^n + \Delta t$. We have the following cases:

- If $\Gamma_1^n \geq \Gamma_2^n$, then: $m^{n+1} = m^n + \Delta t(\Gamma_1^n - \Gamma_2^n)$.
- If $\Gamma_1^n < \Gamma_2^n$, then: $m^{n+1} = \begin{cases} 
m^n + \Delta t(\Gamma_1^n - \Gamma_2^n) & \text{if } \Delta t \leq \bar{t} - t^n, \\
0 & \text{if } \Delta t > \bar{t} - t^n.
\end{cases}$
We now show the important property that if a queue empties in a given time step, it remains empty until the end of the step. If $\dot{m} < 0$, the GRP at the junction has a switching point when $m = 0$. Thus, we consider not only the initial states $\rho_1, \rho_2$ and the final states $\tilde{\rho}_1, \tilde{\rho}_2$, but also intermediate states $\bar{\rho}_1, \bar{\rho}_2$, i.e. the states at $\bar{t}$. While the queue is still emptying, we have $\Gamma_1 = f(\tilde{\rho}_1)$ and $\Gamma_2 = \gamma_2^s(\rho_2)$. Then, at time $\bar{t}$, when $m = 0$, we consider the GRP with initial states $\bar{\rho}_1$ and $\bar{\rho}_2$. Since by the definition of the demand function we have: $f(\bar{\rho}_1) \leq \gamma_1^d(\bar{\rho}_1)$, and because the queue was depleting, we will have $f(\bar{\rho}_1) < \gamma_2^s(\bar{\rho}_2)$. Therefore, we have that $f(\bar{\rho}_1) \leq \min(\gamma_1^d(\bar{\rho}_1), \gamma_2^s(\bar{\rho}_2))$. Thus, the queue does not begin to fill again.

We also need to modify the Godunov scheme in the case of an emptying queue. We divide the time step into two sub-intervals, $(t^n, \bar{t})$ and $(\bar{t}, t^{n+1})$, where $\Delta t_a = \bar{t} - t^n$ and $\Delta t_b = t^{n+1} - \bar{t}$. Then, we solve two different RPs [21]. For $\Delta t_a$, we solve the classical Godunov scheme, and for $\Delta t_b$, we solve another RP with fluxes as given by the case when $m = 0$. Thus, the total fluxes over the full time step add up to

$$\Gamma_1^{n+1} = \frac{\Delta t_a}{\Delta t} f_1(\rho_1) + \frac{\Delta t_b}{\Delta t} \min(\gamma_1^d(\bar{\rho}_1), \gamma_2^s(\bar{\rho}_2)),$$

$$\Gamma_2^{n+1} = \frac{\Delta t_a}{\Delta t} \gamma_2^s(\rho_2) + \frac{\Delta t_b}{\Delta t} \min(\gamma_1^d(\bar{\rho}_1), \gamma_2^s(\bar{\rho}_2)).$$

2.5 Model Comparison and Numerical Examples

We construct two specific examples that highlight the key differences of the three models: FIFO, non-FIFO, and FIFO with queue (FIFOQ). In this example, the off-ramp is set to be completely clogged at the initial time ($t = 0$), and after some fixed time it is (artificially) set to free-flow. Then, at the final time, the total tally of vehicles exiting and passing the off-ramp is taken.

We use the Greenshields flux function with a critical density of 40 veh/km/lane, and a maximum velocity of 100 km/h, which yields a capacity of 2000 veh/h/lane.
<table>
<thead>
<tr>
<th>Model</th>
<th>Total in-flux</th>
<th>Total out-flux on $I_2$</th>
<th>Total flux on $I_3$</th>
<th>Ratio of out-fluxes</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIFO</td>
<td>133 veh</td>
<td>111 veh</td>
<td>22 veh</td>
<td>5:1</td>
</tr>
<tr>
<td>non-FIFO</td>
<td>196 veh</td>
<td>174 veh</td>
<td>22 veh</td>
<td>7.81:1</td>
</tr>
<tr>
<td>FIFOQ</td>
<td>200 veh</td>
<td>167 veh</td>
<td>33 veh</td>
<td>5:1</td>
</tr>
</tbody>
</table>

Table 2.1: Total vehicle flow after $t = 25$ min.

For the three edges, we use

\[
\rho_1^{\text{max}} = \rho_2^{\text{max}} = 320 \text{ veh/km} \quad \text{and} \quad \rho_3^{\text{max}} = 80 \text{ veh/km},
\]

\[
v_1^{\text{max}} = v_2^{\text{max}} = v_3^{\text{max}} = 100 \text{ km/h},
\]

representing a 4-lane highway and a single-lane off-ramp (for simplicity with the same speed limit as the highway). Moreover, a split ratio of $\alpha_2 = \frac{5}{6}$ and $\alpha_3 = \frac{1}{6}$ is assumed, and the initial densities are $\rho_1 = 0.4\rho_1^{\text{max}}$, $\rho_2 = 0$, and $\rho_3 = \rho_3^{\text{max}}$, representing a free-flow highway with a completely clogged off-ramp. This situation is let to evolve until $t = 9$ min. Then, we perform an idealized maneuver (for the sake of simplicity) in which we set $\rho_3 = 0$, i.e., we instantaneously remove all vehicles from the off-ramp. We then let this new situation evolve until $t = 25$ min. At that final time, we tally the total (time-integrated) fluxes $I_2$ and $I_3$ onto the highway and the off-ramp, respectively.

Table 2.1 shows the total vehicle fluxes, obtained by the different models. At the end of this experiment, the queue that forms in the FIFOQ model has emptied. One can clearly see the key shortcomings of the existing models. The FIFO model respects split ratios, but it predicts zero total flux up until $t = 9$ min, resulting in severely reduced highway flux. In turn, the non-FIFO model produces reasonable highway flows, but it fails to respect the split ratio: due to the negligence of the off-ramp queue that arises in reality, it falsely re-allocates vehicles that intended to exit, into highway flow.

The new FIFOQ model remedies both of those shortcomings. Unlike FIFO, it produces a nonzero highway flux before $t = 9$ min. However, unlike non-FIFO, it tracks the accumulation of type 3 vehicles via the growing vertical
queue $m_3$. Then, after $t = 9$ min, this queue depletes, releasing those queued type 3 vehicles onto the off-ramp. Therefore, the total off-ramp flux produced by the FIFOQ model is noticeably larger than with the other two models; which is more realistic, because it remedies the spurious re-routing of the non-FIFO model.
(a) \( t = 1.5 \) min

(b) \( t = 6 \) min

(c) \( t = 9.5 \) min

(d) \( t = 15 \) min

(e) \( t = 25 \) min

Figure 2.5: Model comparison: Time evolution of the solution of the FIFO model.
Figure 2.6: Model comparison: Time evolution of the solution of the non-FIFO model.
Figure 2.7: Model comparison: Time evolution of the solution of the FIFOQ model.
Figures 2.5, 2.6, and 2.7 show the time evolution (at five representative snapshots in time) of the solutions produced by the three models, computed by a highly resolved Godunov/CTM discretization (with proper queue treatment). For the FIFOQ model, Figure 2.7 shows the magnitude of the queue $m_3$ in the lower queue box. The blue arrows visualize the flows into and out of the respective queues. For consistency, empty queue boxes are also shown for the models without queues.

For both classical models, Figures 2.5(a) and 2.5(b), as well as Figures 2.6(a) and 2.6(b), show the initial backwards propagating shock that arises due to the clogged off-ramp. Clearly, with FIFO, the level of congestion on the in-road is much larger (in fact, completely jammed) than with non-FIFO. For the FIFOQ model (Figures 2.7(a) and 2.7(b)) the accumulation of type 3 vehicles is instead tracked in the queue.

Then, for all models, Figures 2.5(c), 2.6(c), and 2.7(c) show the system state right after the clearing of the off-ramp (the idealized maneuver). For both FIFO and non-FIFO, the flow turns maximal, resulting in a rarefaction fan on the in-road. In turn, in the FIFOQ model the queue $m_3$ has started to decrease.

Finally, for both FIFO and non-FIFO, Figures 2.5(d) and 2.5(e), as well as Figures 2.6(d) and 2.6(e), show the gradual approach of the system towards a uniform state on the in-road. In contrast, for FIFOQ the in-road has remained in free-flow the whole time. Figures 2.7(d) and 2.7(e) show the shrinking and eventual depletion of the queue. Right after $t = 25$ min, the outflow states will change to what the other two models would also yield as $t \to \infty$. Note that, in Figure 2.7, because queue 3 absorbs all congestion from $I_3$ and since $I_2$ is in the free-flow regime, we do not see any congestion on $I_1$.

In order to demonstrate the capabilities of the FIFOQ model, we study it in a more complicated second example. We use the same road and split parameters as the previous example, but modify the initial and boundary data as follows: $\rho_1 = 0.3 \rho_1^{\text{max}}$, $\rho_2 = 0.3 \rho_2^{\text{max}}$, and $\rho_3 = 0.88 \rho_3^{\text{max}}$, representing a free-flow highway with a congested but not clogged off-ramp. Boundary conditions
of the same values are prescribed on the in-road and off-ramp. Moreover, the highway queue is initialized with \( m_2 = 17 \text{ veh} \). This situation is let to evolve until \( t = 9 \text{ min} \). Then, the outflow conditions at the ramp are changed to maximum capacity \((\rho_3 = \sigma_3)\), and the solution is tracked until \( t = 25 \text{ min} \). Figure 2.8 shows time evolution of the solution. Figure 2.8(a) shows how \( m_2 \) is decreasing and a shock travels backwards on the in-road. In this example, the queue \( m_2 \) depletes, and queue \( m_3 \) starts forming. Figure 2.8(b) shows the situation soon after that instance. Figure 2.8(c) shows the state of the system after the outflow conditions of the ramp have been changed. A rarefaction arises on the ramp, and once the density at the beginning of the ramp has decreased sufficiently, the queue \( m_3 \) shrinks again. Figures 2.8(d) and 2.8(e) show the shrinking and eventual depletion of the off-ramp queue.

### 2.6 Model Extensions

The proposed FIFOQ model [73] assumes that the road sharing ratios, \( c_j \), for \( j \in \{2, 3\} \), are equal to the split ratios \( \alpha_j \), for \( j \in \{2, 3\} \). However, this is not an accurate representation of reality. Road sharing ratios are quantities that are determined by the geometry of the road networks and can be affected by the congestion on the roads and queued up vehicles. Therefore, to improve the model, we need to extend the FIFOQ model by relaxing the assumption that the road sharing ratios are equal to the split ratios. We first assume that road sharing ratios are constant and different from the split ratios (see Figure 2.9). The derivation of the extended model for the cases where the split ratios are different from the road sharing ratios is similar to the derivation of the FIFOQ model. However, for sake of completeness, we repeat it here.

Let \( \tilde{\Gamma}_1, \tilde{\Gamma}_{12}, \tilde{\Gamma}_{13}, \tilde{m}_2, \) and \( \tilde{m}_3 \), denote the fluxes and queue values before the removal of \( \mu \). We assume that the two-pipe region (see Figure 2.3, part b) is in free-flow initially. Thus, the supplies of the pipes, \( \gamma_{12}^s \) and \( \gamma_{13}^s \), are always maximal.
Figure 2.8: More complex example: Time evolution of the FIFOQ solution.
(a) Cars that want to leave the highway, will use the right most lane: $c_2 + c_3 > 1$ 
(b) Exit Only Junction: $c_1 + c_2 = 1$.

Figure 2.9: Two examples of how road sharing ratios are determined based on road geometry.

We have $\tilde{\Gamma}_1 = \min(\gamma_1^d, \min(\frac{c_3}{\alpha_3}, \frac{c_2}{\alpha_2}), f_{1\text{max}})$, where $f_{1\text{max}} = \frac{v_{\text{max}} \rho_{\text{max}}}{4}$. Note that if we assume that the road is divided precisely according to the split ratio, i.e., $c_2 = \alpha_2$ and $c_3 = \alpha_3$, our model will reduce to the model proposed in [73]. Moreover, $\tilde{\Gamma}_{12} = \alpha_2 \tilde{\Gamma}_1 = \alpha_2 \min(\gamma_1^d, \min(\frac{c_3}{\alpha_3}, \frac{c_2}{\alpha_2}), f_{1\text{max}})$ and $\tilde{\Gamma}_{13} = \alpha_3 \tilde{\Gamma}_1 = \alpha_3 \min(\gamma_1^d, \min(\frac{c_3}{\alpha_3}, \frac{c_2}{\alpha_2}), f_{1\text{max}})$.

Collapsing the pipes to have zero length implies that the flux out of them equals the flux into them. Therefore, the fluxes into the 1-in-1-out nodes (or queues) are $\tilde{\Gamma}_{12}$ and $\tilde{\Gamma}_{13}$, respectively. If there are no queues, the out-road fluxes are:

$$\Gamma_2 = \min \left( \alpha_2 \min(\gamma_1^d, \min(\frac{c_3}{\alpha_3}, \frac{c_2}{\alpha_2}), f_{1\text{max}}), \gamma_2^s \right), \quad (2.8)$$

$$\Gamma_3 = \min \left( \alpha_3 \min(\gamma_1^d, \min(\frac{c_3}{\alpha_3}, \frac{c_2}{\alpha_2}), f_{1\text{max}}), \gamma_3^s \right). \quad (2.9)$$

In turn, if a queue is active, i.e., $\tilde{m}_i > 0$, then the corresponding out-flux is maximal, i.e., $\Gamma_i = \gamma_i^s$. The evolution of the queues is described by the differences in the fluxes: $\dot{\tilde{m}}_2 = \tilde{\Gamma}_{12} - \Gamma_2 = \alpha_2 \tilde{\Gamma}_1 - \Gamma_2$ and $\dot{\tilde{m}}_3 = \tilde{\Gamma}_{13} - \Gamma_3 = \alpha_3 \tilde{\Gamma}_1 - \Gamma_3$.

As by the prior discussion, we now determine $\mu$ such that at most one queue is active at any time. If one queue is active, for example $\tilde{m}_2 > 0$, and the other queue is filling, i.e., $\tilde{m}_3 > 0$, then $\mu = \frac{\hat{m}_3}{\alpha_3}$.
Now, with \( \Gamma_1 = \min(\gamma_1^d, \min(\frac{c_3}{\alpha_3}, \frac{c_2}{\alpha_2}) f_{1\text{max}}) - \mu \), we obtain for the in-flux
\[
\Gamma_1 = \min \left( \min(\gamma_1^d, \min(\frac{c_3}{\alpha_3}, \frac{c_2}{\alpha_2}) f_{1\text{max}}), \frac{\gamma_3^s}{\alpha_3} \right).
\]

Conversely, if no queue is active, i.e., \( \tilde{\dot{m}}_2 = \tilde{\dot{m}}_3 = 0 \), we have
\[
\mu = \min \left( \frac{\tilde{\dot{m}}_2}{\alpha_2}, \frac{\tilde{\dot{m}}_3}{\alpha_3} \right). \tag{2.10}
\]

We define \( s^* = \min(\frac{c_2}{\alpha_2}, \frac{c_3}{\alpha_3}) \) and \( f^* = \min(\gamma_1^d, s^* f_{1\text{max}}) \). Now by combining equations (2.8) and (2.10), we have
\[
\mu = \min \left( f^* - \min \left( f^*, \frac{\gamma_2^s}{\alpha_2}, \frac{\gamma_3^s}{\alpha_3} \right), \right.
\]
\[
\left. f^* - \min \left( f^*, \max \left( \frac{\gamma_2^s}{\alpha_2}, \frac{\gamma_3^s}{\alpha_3} \right) \right) \right).
\]

and with \( \Gamma_1 = f^* - \mu \), we obtain for the in-flux
\[
\Gamma_1 = \min \left( f^*, \max \left( \frac{\gamma_2^s}{\alpha_2}, \frac{\gamma_3^s}{\alpha_3} \right) \right).
\]

Therefore, the extended FIFOQ model when road sharing ratios are constant different from split ratios, i.e. \( c_j \neq \alpha_j \) for \( j \in \{2, 3\} \), will be:

- If \( m_2 = m_3 = 0 \),
  \[
  \begin{align*}
  \Gamma_1 & = \min \left( f^*, \frac{\gamma_2^s}{\alpha_2}, \frac{\gamma_3^s}{\alpha_3} \right), \\
  \Gamma_2 & = \min \left( \alpha_2 f^*, \gamma_2^s \right), \\
  \Gamma_3 & = \min \left( \alpha_3 f^*, \gamma_3^s \right), \\
  \Gamma_1 & = \min \left( f^*, \frac{\gamma_3^s}{\alpha_3} \right), \\
  \Gamma_2 & = \min \left( \alpha_2 f^*, \gamma_2^s \right), \\
  \Gamma_3 & = \gamma_3^s,
  \end{align*}
  \]

- If \( m_2 > 0 \),
  \[
  \begin{align*}
  \Gamma_1 & = \gamma_2^s, \\
  \Gamma_2 & = \min \left( \alpha_2 f^*, \gamma_2^s \right), \\
  \Gamma_3 & = \gamma_3^s,
  \end{align*}
  \]

- If \( m_3 > 0 \),
  \[
  \begin{align*}
  \Gamma_1 & = \gamma_3^s, \\
  \Gamma_2 & = \gamma_2^s, \\
  \Gamma_3 & = \min \left( \alpha_2 f^*, \gamma_2^s \right),
  \end{align*}
  \]

\( \tilde{\dot{m}}_2 = \alpha_2 \Gamma_1 - \Gamma_2 \),

\( \tilde{\dot{m}}_3 = \alpha_3 \Gamma_1 - \Gamma_3 \).
Again, by construction, this model never generates more than one queue to be active. Note that if we assume that $c_2 = \alpha_2$ and $c_3 = \alpha_3$, then $s^* = 1$ and $f^* = \gamma_1^d$ and our model reduces to the model proposed earlier [73].

Although in this case, the new assumption places more limitations on the fluxes, the extended model has the same properties as the FIFOQ model. Specifically, one can show that for the extended version of the model, the junction will never experience an infinite cascade of the events; it will maximize the fluxes through the junction; and after the complete depletion of a queue, it will not immediately fill up again. See chapter 3.2 for more details.

As the next step of the extension of the model, we consider the case when the road sharing ratios are decreasing functions of the queue lengths, see Figure 2.10. In this figure, when the queue in the rightmost lane becomes sufficiently long, drivers who wish to exit but are not in the rightmost lane will be caught by surprise and will need to slow down to move into the right lane, thus they will block the cars behind them.

![Figure 2.10: As queue 3 increases, road sharing ratio $c_2$ decreases.](image)

When the length of the queue $j$, for $j \in \{2, 3\}$ increases and vehicles start occupying not only the designated lane but also the adjacent lanes, the road sharing ratio for the vehicles of type $k$ for $k \in \{2, 3\} \setminus \{j\}$ decreases consequently.

Note that, in the derivation of this extended version, the quantity $s^*$ is equal to $\min(\frac{c_2(m_3)}{\alpha_2}, \frac{c_3(m_2)}{\alpha_3})$. Thus, $f^*$ will also be affected and it will introduce more restriction on the fluxes. The rest of the procedure is similar to the above discussion.
CHAPTER 3

ANALYSIS OF THE FIFOQ MODEL

In this chapter, we present a mathematical analysis of the FIFOQ model proposed in the previous chapter. The introduction of the vertical queue as well as the fact that the model preserves the constant split ratios, lead to some key differences with the existing models in the literature. In this regard, finding a weak entropy solution to the Cauchy problem for the junction is a non-trivial task. Finding a weak entropy solution has been studied for simple networks with at most two incoming and outgoing roads as well as complex networks [12, 19, 28, 29]. There are also some studies that consider buffers in their models [26, 37]. However, these models do not consider the split ratios. Our approach is to construct the solution using the wave-front tracking algorithm [4, 41, 15]. To this end, we first estimate the initial data on the roads with a piecewise constant approximation that satisfies some specific properties. Then we consider the Riemann problem at the junction which is a Cauchy problem with constant initial data on each road. Then, a unique solution is determined using the new coupling conditions in conjunction with the Riemann solver at the junction. The Riemann solver maps the initial data to the solution. Proving the existence of wave-front tracking solution involves estimating the number of waves, the number of interactions and the total variation of the
solution. However, due to the existence of the queue and dependence of the model to the constant split ratios, these estimates are not straightforward for our model. Therefore, similar to other studies [26, 29, 27, 20], we rely on the total variation of fluxes. In order to obtain the needed bounds on the total variation which are necessary for proving the existence of the solution, we investigate some properties of the Riemann solver. Then, we consider all the possible scenarios on the road and prove that the total variation of the fluxes is uniformly bounded. Then, we use these results to prove existence of the wave-front tracking solutions and obtain the entropy admissible solution to (3.3). For the continuous dependence of the solution on the initial data as well as the uniqueness of the solution, the concept of generalized tangent vector is used [3, 5].

3.1 Preliminaries and Riemann Solver

For the network of one incoming and two outgoing roads, the vehicle density \( \rho_i = \rho_i(x, t) \) on road \( I_i \) for \( i = 1, 2, 3 \), is evolved via the following conservation law:

\[
\partial_t \rho_i(x, t) + \partial_x f_i(\rho_i(x, t)) = 0 \quad \text{for } i \in \{1, 2, 3\}, \tag{3.1}
\]

where \( \rho_i \in [0, \rho_{i}^{\text{max}}] \). One can scale the equations such that: \( \rho_{1}^{\text{max}} = 1 \). Moreover, the flux functions have the following properties:

\[
f_i''(\rho) < 0 \quad \text{and} \quad f_i(0) = f_i(\rho_i^{\text{max}}) = 0, \quad \text{for } i \in \{1, 2, 3\}. \tag{3.2}
\]

In this study, we visualize all examples using the Greenshields flux function, namely \( f_i(\rho) = v_i^{\max} \rho \left( 1 - \frac{\rho}{\rho_i^{\text{max}}} \right) \), which attains its maximum at \( \sigma_i \), for \( i \in \{1, 2, 3\} \). However, the analysis holds true for any flux functions that satisfies the criteria in (3.2).

Now, consider a junction \( J \), located at \( x = 0 \), with one incoming road \( I_1 \) (where \( x \in (\infty, 0) \)) and two outgoing roads \( I_2 \) and \( I_3 \) (where \( x \in [0, \infty) \)), which is augmented with vertical queues \( m_j \) for \( j \in \{2, 3\} \). We have the
following generalized Cauchy problem \(^1\) that describes the evolution of the densities on the roads and the dynamic of the possible queue:

\[
\begin{align*}
\partial_t \rho_i(x, t) + \partial_x f_i(\rho_i(x, t)) &= 0, \\
\dot{m}_j(t) &= \alpha_j f_1(\rho_1(0^-, t)) - f_j(\rho_j(0^+, t)),
\end{align*}
\]

with the initial conditions:

\[
\begin{align*}
\rho_i(x, 0) &= \rho_{i,0}(x), & i &= 1, 2, 3, \\
m_j(0) &= m_{j,0}, & j &= 2, 3.
\end{align*}
\]

(3.3) (3.4)

Where \(\rho_{i,0} \in [0, \rho_{i,\text{max}}]\) and \(m_{2,0} \cdot m_{3,0} = 0\). In order to find a unique solution for the above problem, we need appropriate coupling conditions. Accordingly, we need to know the value of the densities near the junction:

\[
\rho_1(t, 0^-), \rho_2(t, 0^+) \text{ and } \rho_3(t, 0^+), \text{ for } t \geq 0.
\]

These coupling conditions depend on the split ratios, \(\alpha_j\), for \(j = 2, 3\).

Moreover, we define the Riemann solver \(\mathcal{RS}_m\)

\[
\mathcal{RS}_m : \prod_{i=1}^3 [0, \rho_{i,\text{max}}] \to \prod_{i=1}^3 [0, \rho_{i,\text{max}}],
\]

at the junction, such that:

\[
\mathcal{RS}_m(\rho_{1,0}, \rho_{2,0}, \rho_{3,0}) = (\rho_1, \rho_2, \rho_3).
\]

The Riemann Solver is a map that associates to each Riemann data \((\rho_{1,0}, \rho_{2,0}, \rho_{3,0})\) at the junction \(J\), a vector \((\rho_1, \rho_2, \rho_3)\) that satisfies some properties that are elaborated in the following.

We investigate the FIFOQ model described in section 3.1, where the road sharing ratios are equal to the split ratios, namely \(c_j = \alpha_j\) for \(j \in \{2, 3\}\). For this model, we showed that, depending on the availability of a queue at the junction, the model consists of three different modes. As discussed in

\(^1\)We call this problem “Cauchy problem” because it has no boundary conditions and “generalized”, because it has interface conditions at \(x = 0\) and ODEs for the queues.
section 2.3, mechanism of the model is such that there is at most one active queue at any instant in time. On the incoming road $I_1$, the flux is obtained by computing the minimum of the available demand $\gamma^d_1$ on the incoming road and the effective supplies, $\frac{\gamma^s_j}{\alpha_j}$ for $j \in \{2, 3\}$, on the outgoing roads. On the outgoing road $I_j$ for $j \in \{2, 3\}$, the flux is obtained by calculating the minimum of corresponding partial demand and available supply. Consequently, when there exists an active queue $j$, for $j \in \{2, 3\}$, the flux is equal to the available supply on the road. Therefore, the FIFOQ model reads as:

Mode 1: If $m_2 = m_3 = 0$:

$$
\begin{align*}
\Gamma_1 &= \min \left( \gamma^d_1, \max \left( \frac{\gamma^s_2}{\alpha_2}, \frac{\gamma^s_3}{\alpha_3} \right) \right), \\
\Gamma_2 &= \min \left( \alpha_2 \gamma^d_1, \gamma^s_2 \right), \\
\Gamma_3 &= \min \left( \alpha_3 \gamma^d_1, \gamma^s_3 \right).
\end{align*}
$$

(3.6)

Mode 2: If $m_2 > 0$:

$$
\begin{align*}
\Gamma_1 &= \min \left( \gamma^d_1, \frac{\gamma^s_2}{\alpha_2} \right), \\
\Gamma_2 &= \gamma^s_2, \\
\Gamma_3 &= \min \left( \alpha_3 \gamma^d_1, \gamma^s_3 \right).
\end{align*}
$$

(3.7)

Mode 3: If $m_3 > 0$:

$$
\begin{align*}
\Gamma_1 &= \min \left( \gamma^d_1, \frac{\gamma^s_3}{\alpha_3} \right), \\
\Gamma_2 &= \min \left( \alpha_2 \gamma^d_1, \gamma^s_2 \right), \\
\Gamma_3 &= \gamma^s_3.
\end{align*}
$$

(3.8)

Starting with the initial condition

$$(\rho_{1,0}, \rho_{2,0}, \rho_{3,0}, m_{2,0}, m_{3,0}),$$

where $\rho_{1,0} = \rho_1(0, 0^-)$ and $\rho_{j,0} = \rho_j(0, 0^+)$ for $j = 2, 3$, the flux on road $I_i$ near the junction is equal to $f_i(\rho_{i,0})$, for $i = 1, 2, 3$. As the junction updates the fluxes on the roads for $t = 0^+$, the solution will be:

$$\mathcal{RS}_m(\rho_{1,0}, \rho_{2,0}, \rho_{3,0}) = (\rho_1, \rho_2, \rho_3),$$

(3.9)

where the solution $\rho_i$ of $f_i(\rho_i) = \Gamma_i$, for $i \in \{1, 2, 3\}$ should be chosen such that:
• on the incoming road $I_1$, the solution to the Riemann problem \(^2\) is a wave $(\rho_{1,0}, \rho_1)$ with a negative speed.

• on the outgoing road, $I_j$, $j \in \{2, 3\}$, the solution to the Riemann problem is a wave $(\rho_j, \rho_{j,0})$ with a positive speed.

**Lemma 1.** The Riemann solver $\mathcal{RS}_m$, satisfies the consistency condition:

$$\mathcal{RS}_m(\mathcal{RS}_m(\rho_1, \rho_2, \rho_3)) = \mathcal{RS}_m(\rho_1, \rho_2, \rho_3),$$

where $\rho_i \geq 0$ for $i \in \{1, 2, 3\}$.

**Proof.** Consider the initial condition $(\rho_{1,0}, \rho_{2,0}, \rho_{3,0}) \in \prod_{i=1}^3 [0, \rho_{i,\text{max}}]$ on the roads. Given

$$\mathcal{RS}_m(\rho_{1,0}, \rho_{2,0}, \rho_{3,0}) = (\rho_1, \rho_2, \rho_3),$$

$$\mathcal{RS}_m(\rho_1, \rho_2, \rho_3) = (\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3),$$

we want to show that:

$$(\rho_1, \rho_2, \rho_3) = (\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3). \quad (3.10)$$

First, consider the case in which $\rho_1 \in [0, \sigma_1]$. This implies that the initial condition on road $I_1$ was in the free-flow regime. Having $\rho_1$ in the free-flow regime implies $\gamma_1^d(\rho_{1,0}) \leq \max\left(\frac{\gamma_2^s(\rho_{2,0})}{\alpha_2}, \frac{\gamma_3^s(\rho_{3,0})}{\alpha_3}\right)$, and two sub-cases will arise. Namely, the initial demand on $I_1$, $\gamma_1^d(\rho_{1,0})$, is either less than both of the two effective supplies on the outgoing roads, or it is greater than one of the effective supplies. For the first sub-case, we have

$$\gamma_1^d(\rho_{1,0}) \leq \min\left(\frac{\gamma_2^s(\rho_{2,0})}{\alpha_2}, \frac{\gamma_3^s(\rho_{3,0})}{\alpha_3}\right),$$

which implies that $\rho_2$ and $\rho_3$ are in the free-flow regime. Therefore, because the supply function is non-increasing, we have $\gamma_j^s(\rho_{j,0}) \leq \gamma_j^s(\rho_j)$ for $j \in \{2, 3\}$.

Moreover, because the wave on $I_1$ leaves the junction with a negative speed and

---

\(^2\)The Riemann problem is defined as extending $I_1$ to $(-\infty, +\infty)$ with states $\rho_{1,0}$ for $x < 0$ and $\rho_1$ for $x > 0$. See Chapter 2 for more details.
the demand function is a non-decreasing function, we have \( \gamma_1^d(\rho_{1,0}) \geq \gamma_1^d(\rho_1) \).
Based on these inequalities, and using equation (3.6), we have \( \hat{\rho}_1 \in [0, \sigma_1] \) and is equal to \( \rho_1 \) and for the densities on the outgoing roads, we have \( \hat{\rho}_j = \rho_j \) for \( j \in \{2, 3\} \), and statement (3.10) is satisfied.

For the second sub-case, without loss of generality, assume that:

\[
\frac{\gamma_2^s(\rho_{2,0})}{\alpha_2} \leq \frac{\gamma_1^d(\rho_{1,0})}{\alpha_1} \leq \frac{\gamma_3^s(\rho_{3,0})}{\alpha_3}.
\]

This additional assumption implies that \( \rho_1 \) and \( \rho_3 \) are in the free-flow regime and \( \rho_2 \) is in the congestion regime (queue 2 is active and absorbs the excess vehicles). Moreover, similar to the discussions for the first case, we have \( \gamma_1^d(\rho_{1,0}) = \gamma_1^d(\rho_1), \ \gamma_2^s(\rho_2) \leq \gamma_2^s(\rho_{2,0}) \) and \( \gamma_3^s(\rho_3) \geq \gamma_3^s(\rho_{3,0}) \). Therefore, the Riemann solver maps \((\rho_1, \rho_2, \rho_3)\) to \((\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3)\), where \( \hat{\rho}_i = \rho_i \), for \( i = 1, 2, 3 \) and satisfies (3.10).

Now, consider the case when \( \rho_1 \) is in the congestion regime. \( \rho_1 \in [\sigma_1, \rho_1^{\text{max}}] \) implies that:

\[
\gamma_1^d(\rho_{1,0}) \geq \max \left( \frac{\gamma_2^s(\rho_{2,0})}{\alpha_2}, \frac{\gamma_3^s(\rho_{3,0})}{\alpha_3} \right),
\]

therefore, \( \rho_2 \) and \( \rho_3 \) are in the congestion regime. Note that, for supplies on the outgoing roads we have:

\[
\gamma_2^s(\rho_j,0) = f_j(\rho_j) = \Gamma_j = \gamma_2^s(\rho_{j,0}), \quad \text{for } j = 2, 3.
\]

Similarly, for this data, the demand on the incoming road is greater than or equal to the effective supplies on the out-roads. Therefore, the Riemann solver maps \((\rho_1, \rho_2, \rho_3)\) to itself and this finishes the proof.

The FIFOQ model is designed such that, at any instant in time, at most one queue is active. As mentioned in the introduction of the quantity \( \mu \), in the case when both out-roads are clogged or provide sufficiently low supply, the FIFOQ model allows for backward propagating shocks on the in-road, and the congested state is composed according to the split ratios. For instance, if both queues are initially empty and they both have the tendency to fill up,
the one with lower effective supply is chosen to stack up vehicles. In this case, \( \mu \) is calculated according to:

\[
\mu = \gamma_1^d - \min \left( \gamma_1^d, \max \left( \frac{\gamma_2^s}{\alpha_2}, \frac{\gamma_3^s}{\alpha_3} \right) \right),
\]

and when there exists an active queue \( j \), for \( j \in \{2, 3\} \), the quantity \( \mu \) will be:

\[
\mu = \frac{\alpha_k \gamma_1^d}{\alpha_k} - \min \left( \alpha_k \gamma_1^d, \gamma_k^s \right) = \gamma_1^d - \min \left( \alpha_k \gamma_1^d, \gamma_k^s \right),
\]

where \( k \in \{2, 3\} \setminus \{j\} \).

Whether or not \( \mu = 0 \) plays an important role in finding the solution. For example, in the presence of an active queue \( j \) for \( j \in \{2, 3\} \), when \( \mu = 0 \), states on roads \( I_i \) for \( i \in \{1, 2, 3\} \setminus \{j\} \), are in the free-flow regime. On the other hand, if \( \mu \neq 0 \), then the whole junction is in the congestion regime. Moreover, if \( \mu \neq 0 \), when an active queue depletes, the other queue will start filling up. However, if during the depletion of the active queue, \( \mu = 0 \), then the whole junction will be in the free-flow regime after the complete depletion of the queue.

### 3.2 Important Characteristics of the FIFOQ Model

In this subsection we highlight some of the main characteristics of the FIFOQ model that are caused by the introduction of the vertical queue at the junction. We first show that the model captures all of the events and smoothly transitions between different modes. After that, we show that the model with the Riemann solver introduced in the previous section will maximize the fluxes through the junction. Finally, we show that starting with a constant initial data, the model will never produce infinite number of waves on the roads in a finite time interval. The latter case is extended for any type of initial data and discussed in section 3.3.
3.2.1 Self-Consistency of the Model

Because the FIFOQ model, unlike many other coupling models in the literature, involves a distinction of cases, one needs to verify that the definition of the cases does not lead to the intrinsic inconsistencies. To this end, we prove the self-consistency of the model. By self-consistency of the model, we mean that the model can organically transition between these modes. In other words, the model captures the sequence of mode changes without missing any events.

Let us explain the mode changing through some examples. For the first example, consider the case in which the model initializes with mode 1. Now, it is possible that as the junction updates the fluxes, queue \( j \) for \( j \in \{2, 3\} \) starts filling up. Thus the model should transition to mode \( j \). Or consider the case where the model is in mode 1 and a wave on a road hits the junction and causes the activation of queue \( j \) for \( j \in \{2, 3\} \). Again, this leads to a transition to mode \( j \). An interesting fact about the model is that, in both cases, after substituting the values of demand and supplies in the equation of mode 1, it is organically propelled to mode \( j \). In other words, in the case when both queues were idle and one queue starts to fill up, the equations for mode 1 become equivalent to the equations for mode \( j \).

As another example, consider the case when queue \( j \) is active and the system is in mode \( j \), for \( j \in \{2, 3\} \). Now after the complete depletion of the queue \( j \), as described below in Lemma 2, different scenarios can arise. Therefore, the model should change to a new mode. In fact, as the length of queue \( j \) hits zero, the model will choose the next mode that, according to the values of demand and supplies, does not lead to a contradiction. The next Lemma highlights the admissible and inadmissible mode transitions of the model for this case.

**Lemma 2.** After the complete depletion of an active queue \( j \) with \( j \in \{2, 3\} \), it will not immediately fill up again. In other words, either queue \( k \in \{2, 3\} \setminus \{j\} \), starts filling up, or both queues will remain empty.
\textbf{Proof.} Without loss of generality, assume that queue 3 is active and it is depleting. Let us assume at time $t = \bar{t}$, queue 3 hits zero. We use superscript $-$ and $+$, to refer to the times in intervals $[\bar{t} - \epsilon, \bar{t})$ and $(\bar{t}, \bar{t} + \epsilon]$, for a small $\epsilon > 0$, respectively. We have $\dot{m}_3^- < 0$, therefore, the model is in mode 3, and we have:

$$
\begin{align*}
\Gamma_1^- &= \min \left( \gamma_1^d(\rho_1^-), \frac{\gamma_2^s(\rho_2^-)}{\alpha_2} \right), \\
\Gamma_2^- &= \min \left( \alpha_2 \gamma_1^d(\rho_1^-), \gamma_2^s(\rho_2^-) \right), \\
\Gamma_3^- &= \gamma_3^s(\rho_3^-).
\end{align*}
\tag{3.11}
$$

After the complete depletion of queue 3, choosing mode 3 again leads to a contradiction. This is because, after substituting the values for supplies and demand into the equations of mode 3, we have:

$$
\dot{m}_3^+ = \alpha_3 \Gamma_1^+ - \Gamma_3^+ < 0,
\tag{3.12}
$$

which is in contradiction with the fact that $m_3^+ = 0$ and it cannot deplete, because it is already empty. Therefore, it should switch either to mode 1 or 2.

For the first case (switching to mode 1), we have:

$$
\begin{align*}
\Gamma_1^+ &= \min \left( \gamma_1^d(\rho_1^-), \max \left( \frac{\gamma_2^s(\rho_2^-)}{\alpha_2}, \frac{\gamma_3^s(\rho_3^-)}{\alpha_3} \right) \right), \\
\Gamma_2^+ &= \min \left( \alpha_2 \gamma_1^d(\rho_1^-), \gamma_2^s(\rho_2^-) \right), \\
\Gamma_3^+ &= \min \left( \alpha_3 \gamma_1^d(\rho_1^-), \gamma_3^s(\rho_3^-) \right).
\end{align*}
\tag{3.13}
$$

First, consider the situation that:

$$
\frac{\gamma_2^s(\rho_2^-)}{\alpha_2} > \frac{\gamma_3^s(\rho_3^-)}{\alpha_3},
\tag{3.14}
$$

therefore, we have: $\Gamma_1^+ = \Gamma_1^-$ and $\Gamma_2^+ = \Gamma_2^-$. We show, by contradiction, that the quantity $\mu$, before the complete depletion of the queue, is equal to zero. Assume that $\mu \neq 0$. We have: $\Gamma_1^- = \frac{\gamma_2^s(\rho_2^-)}{\alpha_2}$. Therefore, dynamic of queue 3, before the complete depletion, is:

$$
\dot{m}_3^- = \alpha_3 \frac{\gamma_3^s(\rho_2^-)}{\alpha_2} - \gamma_3^s(\rho_3^-) < 0,
$$
which is in contradiction with equation (3.14). Therefore, before the complete depletion of the queue, quantity $\mu$ is zero. Thus, there is no congestion on $I_1$ due to the vehicles who want to enter the outgoing road $I_2$. Moreover, we have: $\gamma_1^d(\rho_1^-) < \frac{\gamma_2^s(\rho_2^-)}{\alpha_2}$. Since $\dot{m}_3^- < 0$, therefore $\alpha_3\gamma_1^d(\rho_1^-) < \gamma_3^s(\rho_3^-)$. And for the fluxes on the roads we have:

$$\Gamma_1^+ = \gamma_1^d(\rho_1^-), \quad \Gamma_2^+ = \alpha_2\gamma_1^d(\rho_1^-), \quad \text{and} \quad \Gamma_3^+ = \alpha_3\gamma_1^d(\rho_1^-),$$

and, because the fluxes on the outgoing roads are equal to their corresponding partial demands, no queue forms and the whole junction is in the free-flow regime. Therefore, mode 3, under the condition of statement in (3.14), will change to mode 1.

Now, consider the situation that:

$$\gamma_2^s(\rho_2^-) < \frac{\gamma_3^s(\rho_3^-)}{\alpha_3} \quad \text{(3.15)}$$

therefore fluxes will be:

$$\Gamma_1^+ = \min \left( \gamma_1^d(\rho_1^-), \frac{\gamma_3^s(\rho_3^-)}{\alpha_3} \right),$$

$$\Gamma_2^+ = \Gamma_2^-,$$

$$\Gamma_3^+ = \min \left( \alpha_3\gamma_1^d(\rho_1^-), \gamma_3^s(\rho_3^-) \right) = \alpha_3\Gamma_1^+ \quad \text{(3.16)}$$

Now, consider the situation that $\Gamma_1^+ = \gamma_1^d(\rho_1^-)$. We distinguish the situations in which $\mu = 0$ and $\mu \neq 0$.

If, before the complete depletion of the queue, $\mu = 0$, namely, $\gamma_1^d(\rho_1^-) < \frac{\gamma_2^s(\rho_2^-)}{\alpha_2}$, then by assumption (3.15), $\Gamma_j^+ = \alpha_j\gamma_j^d(\rho_j^-)$ for $j \in \{2, 3\}$. This implies that, $m_2^+ = m_3^+ = 0$ and the whole junction is in the free-flow regime. In other words, mode 3 changes to mode 1. However, if before the complete depletion of the queue, $\mu \neq 0$, then $\Gamma_2^+ = \Gamma_2^- = \gamma_2^s(\rho_2^-)$. So we have:

$$\dot{m}_2^+ = \alpha_2\Gamma_1^+ - \Gamma_2^+ = \alpha_2\gamma_1^d(\rho_1^-) - \gamma_2^s(\rho_2^-) > 0.$$  

Therefore, after the complete depletion of queue 3, queue 2 starts forming. In this case, the equations for mode 1 are equivalent to the equations for mode 2. Thus, the mode changes from 3 to 2.
On the other hand, if \( \gamma_d^1(\rho_1) > \frac{\gamma_s^3(\rho_3)}{\alpha_3} \), then \( \Gamma_1^+ = \frac{\gamma_s^3(\rho_3)}{\alpha_3} \), while \( \Gamma_2^+ = \Gamma_2^- \). Under this assumption, \( \gamma_d^1(\rho_1) \) must be greater than \( \frac{\gamma_s^3(\rho_3)}{\alpha_2} \), otherwise, we get to a contradiction with equation (3.15). Therefore, \( \Gamma_2^+ = \gamma_s^3(\rho_2^-) \) and \( \dot{m}_2^+ > 0 \). Thus, after the complete depletion of queue 3, queue 2 starts stacking up. Even though the mode changes from 3 to 1, the equations for mode 1 are equivalent to the equations for mode 2. Thus, mode 3 changes to mode 2.

Regarding switching immediately to mode 2, conditions that do not lead to any contradictions are those studied in the cases where mode 1 organically changes to mode 2. In fact, queue 2 starts forming, if quantity \( \mu \neq 0 \) before the complete depletion of queue 3. Thus, queue 2 has the tendency to form, but due to the existence of queue 3, it waits until queue 3 depletes completely. Fluxes, before the complete depletion of the queue, are:

\[
\Gamma_1^- = \frac{\gamma_s^3(\rho_3^-)}{\alpha_2}, \quad \Gamma_2^- = \gamma_s^3(\rho_2^-), \quad \text{and} \quad \Gamma_3^- = \gamma_s^3(\rho_3^-),
\]

and the fluxes, after the complete depletion of the queue, are the same as equations in (3.16). Therefore, given the current supplies and demand, substituting these values into the mode equations (3.6) to (3.8), if two mode equations are satisfied, the mode that indicates an active queue is chosen. For example, if both mode equations for modes 1 and 2 are satisfied, we choose mode 2. In fact, in this situation, based on the values for demand and supplies on the roads, equations for mode 1 are equivalent to equations for mode 2. This finishes the proof.

\[\square\]

**Remark 1.** For the cases where mode switches from 3 to 2, since \( \frac{\gamma_s^3(\rho_3^-)}{\alpha_2} < \min \left( \gamma_1^d(\rho_1^-), \frac{\gamma_s^3(\rho_3)}{\alpha_3} \right) \), after the complete depletion of queue 3, the flux on incoming road \( I_1 \) increases. A similar argument holds true for switching from mode 2 to mode 3.

### 3.2.2 Flux Maximization

In this section, we show that the FIFOQ model with the Riemann solver introduced in previous section, maximizes the flux through the junction. We
first show that, when there is no active queue at the junction, the FIFOQ model reduces to the FIFO model. We have the following lemma:

**Lemma 3.** When there is no active queue, the FIFOQ model reduces to the FIFO model.

**Proof.** Because there is no active queue, every vehicle that enters the junction can freely get through the junction to get into their desired road. Therefore, we have: \( \alpha_2 \Gamma_1 = \Gamma_2 \) and \( \alpha_3 \Gamma_1 = \Gamma_3 \). Moreover, the demand on the incoming road, is less than or equal to the effective supplies, \( \frac{\gamma_j^d}{\alpha_j} \) for \( j = 2, 3 \), on the outgoing roads, and we have:

\[
\Gamma_1 = \gamma_1^d = \min(\gamma_1^d, \max(\frac{\gamma_2^d}{\alpha_2}, \frac{\gamma_3^d}{\alpha_3})) = \min(\gamma_1^d, \frac{\gamma_2^d}{\alpha_2}, \frac{\gamma_3^d}{\alpha_3}),
\]

\[
\Gamma_j = \alpha_j \gamma_1^d = \min(\alpha_j \gamma_1^d, \gamma_j^d), \quad j = 2, 3,
\]

(3.17)

Which is equivalent to the FIFO model.

Using lemma 3 we are ready to prove the following lemma.

**Lemma 4.** The model provides a solution that maximizes the fluxes through the junction.

**Proof.** To prove this lemma, we consider two cases: when there is an active queue, and when there is no active queue. For the case when there is an active queue, without loss of generality, let us assume that \( m_3 > 0 \) and \( m_2 = 0 \). The evolution of the queue 3 will give us \( \dot{m}_3 = \alpha_3 \Gamma_1 - \Gamma_3 \). According to the conservation of vehicles, we have: \( \Gamma_1 = \Gamma_{13} + \Gamma_2 \), where \( \Gamma_{13} \) is the flux into the active queue 3 and we have \( \Gamma_{13} = \alpha_3 \Gamma_1 \). Thus, for the flux on \( I_2 \) we have: \( \Gamma_2 = \alpha_2 \Gamma_1 \). Moreover, the flux on each road must be less than or equal to its demand or supply.

Therefore, in order to maximize the fluxes through the junction, we need to solve the following linear maximization problem:

Maximize \( \Gamma_1 \)

subject to \( 0 \leq \Gamma_1 \leq \gamma_1^d, \)

\( \Gamma_j = \alpha_j \gamma_1^d \), \( j = 2, 3, \quad j = 2, 3 \),
which is coupled with the following independent linear maximization problem for road $I_3$:

\[
\text{Maximize } \Gamma_3 \\
\text{subject to } 0 \leq \Gamma_3 \leq \gamma_s^3.
\]

The solution to these optimization problems is $\Gamma_1 = \min(\gamma^d_1, \frac{\gamma_s^2}{\alpha_2})$ and $\Gamma_3 = \gamma_s^3$ and so $\Gamma_2 = \alpha_2 \Gamma_1 = \min(\alpha_2 \gamma_1^d, \gamma_2^s)$. Which is exactly the statement that was provided by the FIFOQ model.

For the case when there is no active queue, according to lemma 3, the FIFOQ model reduces to the FIFO model, which is known to maximize the fluxes through the junction [25]. This completes the proof.

\[\Box\]

### 3.2.3 Immunization from Infinite Cascade of Events

Most of the existing models in the literature have self-similarity solutions. This means that, for these models, the solution can be written as $\rho(x,t) = R(\xi)$. In the FIFOQ model, based on the rate of changes in the queue length, we can analyze if the model has a self-similarity solution or not. Specifically, if $\dot{m}_2 = \dot{m}_3 = 0$, then the model has a self-similarity solution which includes the queues. If $m_j \geq 0$ and $\dot{m}_j > 0$ for $j \in \{2, 3\}$, then we have self-similarity solution on the edges. However, for queue $j$, we have: $m_j(t) = m_j + \dot{m}_j t$, which increases by time. Finally, if $m_j > 0$ and $\dot{m}_j < 0$, then there is no self-similarity solution. This is true because the queue $m_j$ will hit zero in finite time, which will trigger new states. The aim here is to show that for this case, waves that arise and interact with the junction will never lead to an infinite cascade of events. By infinite cascade of events we mean the potentially plausible situation where an initially finite number of waves generate a series of infinite number of waves on the roads in a finite time interval. We begin our study by considering a simple case in which the initial data is constant. In other words, we consider the generalized Riemann problem (GRP), where the densities on the roads are constant values. We need the following definitions:
Definition 1. As introduced in [29, 22], density $\rho_1$ ($\rho_j$ for $j \in \{2, 3\}$) on incoming road $I_1$ (outgoing road $I_j$ for $j \in \{2, 3\}$), is called a good datum, if $\rho_1 \in [\sigma_1, \rho_{1\max}]$ ($\rho_j \in [0, \sigma_j]$), and a bad datum, otherwise.

Definition 2. A discontinuity $(\rho_l, \rho_r)$ on road $I_i$ for $i \in \{1, 2, 3\}$, is called a Big shock if, $\rho_l \in [0, \sigma_i)$ and $\rho_r \in [\sigma_i, \rho_{i\max}]$. In other words, a Big shock connects a state in the free-flow regime to a state in the congestion regime.

Due to the assumption of GRP, for time $t > 0$, the waves present on the roads are either generated from the junction or produced through the interaction of two waves which had come out of the junction. In order to see whether the infinite cascade of events can happen or not, we need to identify the conditions under which, after interaction of two waves on the roads, the new generated wave moves toward the junction. Therefore, we investigate all possible interactions of the admissible waves that are generated by the junction, on the incoming and outgoing roads. To this end, we treat rarefaction and shock waves as explained in the wave-front tracking algorithm. Namely, if the Riemann problem admits a rarefaction wave, we split it to produce a sequence of jumps of size $\frac{1}{\nu}$, where $\nu \in \mathbb{N}$ is a fixed number. Each of the rarefaction jumps generated in this method moves with a speed that is determined by its corresponding connected states using the Rankin-Hugoniot conditions explained in Section 2.1.1. Solutions regarding the shock waves are exact and no modifications occur. In this way we can construct an approximate solution $\rho_\nu(t, x)$ of the exact solution $\rho(t, x)$ until time $t_1$, when at least two waves interact with each other. Now at the time of interaction, we can again follow the above procedure and construct the solution until the next interaction, for which, we repeat this whole process again and so on. To prove the existence of the solution, we need to prove that this procedure can be done in any time $t$ and $\rho_\nu$ has finite number of discontinuities and satisfies the following statements:

$$\|\rho_\nu\|_{L^\infty} \leq \|\rho\|_{L^\infty},$$
$$\|\rho_\nu - \rho\|_{L^\infty} \leq \frac{1}{\nu}.$$  

(3.18)
In this subsection, due to the simplicity of the initial data, without loss of generality, one can approximate the rarefaction wave with a single jump. This rough approximation does not contradict any of the results of the following lemmas and it gives the reader the convenience of working with the interacting waves while preserving the actual merit of the exact solution. Through the proofs of lemmas, the reader will see that the results still hold true if the value of $\nu$ increases.

Now considering all of the above discussions, we have the following Lemma:

**Lemma 5.** If a new wave that is generated due to the interaction of waves, which originally had come out of the junction, moves towards the junction, then it is a Big shock.

**Proof.** For the incoming road $I_1$, only waves with negative speed can be generated by the junction. There are three possible types of waves, namely, Big shock, small shock and small rarefaction. Both small shock and small rarefaction should take their values from the congestion regime. Now, among all possible interactions of these three types of waves on $I_1$, the only way to produce a wave that moves toward the junction is through the interaction of a small rarefaction with a Big shock. In this interaction, the new wave, which is a Big shock, may move toward the junction. Note that, the bad data of the parent Big shock is preserved due to this interaction and it is passed to the new Big shock.

Similarly, on the outgoing roads $I_2$ and $I_3$, three types of waves can arise from the junction: Big shock, small shock and small rarefaction. Note that, the small shock and the small rarefaction take values in the free-flow regime. Now, considering all of the possible interactions of these three type of waves on the outgoing roads, the only situation that can produce a wave with a negative speed is when a small rarefaction hits a Big shock. Thus the newly generated wave, which is a Big shock, may move toward the junction while it preserves the bad datum of its Big shock parent. $\square$

Based on the proof of the previous lemma, small rarefaction is an essential
part of generating Big shocks that have the potential to move toward the
collision. In the following lemma, the circumstances under which the collision
can produce small rarefaction waves on the roads are illustrated.

**Lemma 6.** Given the GRP assumptions, there will never be a wave on the
outgoing roads that moves toward the collision.

**Proof.** On the outgoing roads, all waves that are coming out of the collision,
leave the collision with a positive speed. Thus, based on the results in the
proof of Lemma 5, the only way to generate a wave with a negative speed
on the outgoing roads, is through the interaction of a rarefaction wave and a
Big shock. We prove that this scenario does not happen. According to the
assumption of GRP, the only mechanism that may lead to new states on the
roads after \( t = 0^+ \), is the complete depletion of a queue. Therefore, without
loss of generality, we assume that \( m_{3,0} > 0 \) and queue 3 depletes. Now, as seen
in Figure 3.1, to generate a Big shock on \( I_2 \), we need to have: \( \rho_{2,0} \in [\sigma_2, 1] \)
and \( \rho_2 \), the state on \( I_2 \) after the first update of the fluxes from collision, to
be in the free-flow regime, namely \( \rho_2 \in [0, \tau(\rho_{2,0})] \), where \( \tau(\rho_{2,0}) \neq \rho_{2,0} \), is the
density level that shares the same flux with \( \rho_{2,0} \). Now, since queue 3 is active
and \( I_2 \) provides enough supply for its corresponding demand, the model is in
mode 3 and \( I_1 \) is in the free-flow regime. Therefore, we have the following
equations:

\[
\text{for } t = 0^+, \quad \Gamma_1 = \min \left( \gamma_1^d(\rho_{1,0}), \frac{\gamma_2^s(\rho_{2,0})}{\alpha_2} \right) = \gamma_1^d(\rho_{1,0}) , \\
\Gamma_2 = \min \left( \alpha_2 \gamma_1^d(\rho_{1,0}), \gamma_2^s(\rho_{2,0}) \right) = \alpha_2 \gamma_1^d(\rho_{1,0}) , \\
\Gamma_3 = \gamma_3^s(\rho_{3,0}).
\]

Thus, a Big shock is generated on \( I_2 \) which connects a free-flow state to the
congested state \( \rho_{2,0} \).
Figure 3.1: A positive speed rarefaction wave hits a positive speed Big shock on the outgoing road, to generate a negative speed Big shock.

Note that, after the junction updates the fluxes on the roads, the demand on incoming road $I_1$ does not change. This is true since if $\rho_{1,0}$ is in the free-flow regime, then $\gamma_1^d(\rho_{1,0}) = f_1(\rho_{1,0})$. Moreover, because the flux does not change on $I_1$, $\Gamma_1 = f_1(\rho_{1,0})$ implies that $\rho_1 = \rho_{1,0}$ and $\gamma_1^d(\rho_1) = \gamma_1^d(\rho_{1,0})$. On the other hand, if $\rho_{1,0}$ is in the congestion regime, $\Gamma_1 = \gamma_1^d(\rho_1) = f_1(\sigma_1)$ implies that $\rho_1 = \sigma_1$ and $\gamma_1^d(\rho_1) = \gamma_1^d(\sigma_1) = \gamma_1^d(\rho_{1,0})$. Using similar arguments, we can show that the supply on $I_3$ has not changed. Now, $\dot{m}_3 = \alpha_3 \gamma_1^d(\rho_{1,0}) - \gamma_3^s(\rho_{3,0}) < 0$ implies that $\gamma_1^d(\rho_1) < \gamma_3^s(\rho_{3,0})$. Therefore, the partial demand on $I_1$ for road $I_3$ is less than its corresponding supply.

Now, assume that at $t = \bar{t}$, queue 3 hits zero, and thus new states will be introduced on the roads after the complete depletion of the queue 3. We have:

$$\Gamma_1 = \min \left( \gamma_1^d(\rho_1), \max \left( \frac{\gamma_2^s(\rho_2)}{\alpha_2}, \frac{\gamma_3^s(\rho_3)}{\alpha_3} \right) \right) = \gamma_1^d(\rho_{1,0}),$$

$$\Gamma_2 = \min \left( \alpha_2 \gamma_1^d(\rho_1), \gamma_2^s(\rho_2) \right),$$

$$\Gamma_3 = \min \left( \alpha_3 \gamma_1^d(\rho_1), \gamma_3^s(\rho_3) \right) = \alpha_3 \gamma_1^d(\rho_1).$$

Therefore, we have:

$$\gamma_1^d(\rho_1) = \gamma_1^d(\rho_{1,0}) < \frac{\gamma_2^s(\rho_{2,0})}{\alpha_2} = \frac{f_2(\rho_{2,0})}{\alpha_2} < \frac{f_2(\sigma_2)}{\alpha_2} = \frac{\gamma_2^s(\rho_2)}{\alpha_2},$$

which implies $\Gamma_2 = \alpha_2 \gamma_1^d(\rho_1)$. However, in order to produce a new state on $I_2$, $\Gamma_2$ should be equal to $\gamma_2^s(\rho_2)$. But, $\Gamma_2 = \alpha_2 \gamma_1^d(\rho_1) = \alpha_2 \gamma_1^d(\rho_{1,0}) = \Gamma_2$, and state on road $I_2$ will not change. Thus, it is not possible for the Big shock to be followed by a small rarefaction wave.
Lemma 7. The GRP assumption implies that a rarefaction wave that follows a Big shock can only appear on the incoming road $I_1$.

Proof. Given the GRP assumption, after the junction updates the fluxes on roads at $t = 0$, a rarefaction wave on the incoming road $I_1$ can only be generated when there is an increase in the flux on the road due to the complete depletion of a queue. So, $I_1$ experiences a small rarefaction only if the queue length hits zero. This is the situation described in Remark 1, where mode transitions from 3 to 2, or from 2 to 3 (i.e. when immediately after one queue empties, the other queue starts filling up).

We are now ready to prove the following theorem:

Theorem 8. Beginning with GRP, the junction will never suffer from an infinite cascade of events.

Proof. Based on the above discussions, infinite cascade of events cannot happen on the outgoing roads. Therefore, we need to check if finite events on $I_1$ can cause infinitely many events in a finite time interval. We investigate the situation in which a small rarefaction hits a Big shock. Note that, a small rarefaction can only be generated as a result of mode transition from mode 3 to 2 or from 2 to 3. Without loss of generality, we consider the transition from mode 3 to 2. Assume that the queue $m_{3,0}$ is non-zero and $\rho_{0,1}$, the state on the incoming road $I_1$, is bad datum namely, $\rho_{1,0} \in [0, \sigma_1]$. There are two possible cases: $\mu \neq 0$ and $\mu = 0$. For the first case, queue 3 depletes only if:

$$\frac{\gamma^s_2}{\alpha_2}(\rho_{2,0}) < \min\left(\frac{\gamma^d_1(\rho_{1,0})}{\alpha_1}, \frac{\gamma^s_3(\rho_{3,0})}{\alpha_3}\right).$$

(3.19)

Note that, the supply on roads $I_3$ and $I_2$ does not change and we have:

$$\frac{\gamma^s_j}{\alpha_j}(\rho_{j,0}) = \frac{\gamma^s_j}{\alpha_j}(\rho_j), \text{ where } \rho_j \text{ is the state on road } I_j, j = 2, 3, \text{ for } t = 0^+.$$
Therefore, when queue 3 becomes empty (at $t = \bar{t}$), the mode changes from 3 to 2. For the flux on $I_1$, we have:

$$\bar{\Gamma}_1 = \min\{\gamma_1^d(\rho_1), \frac{\gamma_3^s}{\alpha_3}(\rho_{3,0})\},$$

therefore a new rarefaction appears on $I_1$. Now, there are two possible cases. Either $\bar{\Gamma}_1 = \gamma_1^d(\rho_1)$ or, $\bar{\Gamma}_1 = \frac{\gamma_3^s}{\alpha_3}(\rho_{3,0})$. In the first case, $\bar{\rho}_1 = \sigma_1$ and interaction of the rarefaction and Big shock will produce a new positive speed Big shock. After the interaction of this new Big shock with the junction (at $t = \bar{\bar{t}}$), the state on $I_1$ will remain in the free-flow regime. Moreover, the state on $I_3$ will be in the free-flow regime and a new Big shock will be generated on $I_3$ that leaves the junction. The state on $I_2$ does not change and queue 2 continues absorbing the excess vehicles, however the rate of stacking up has decreased.

This is true because:

for $t \in (\bar{\bar{t}}, \bar{\bar{t}})$,

$$\bar{\bar{m}}_2 = \alpha_2 \bar{\Gamma}_1 - \bar{\Gamma}_2 = \alpha_2 \gamma_1^d(\rho_1) - \gamma_2^s(\rho_{2,0})$$

$$> \alpha_2 \gamma_1^d(\rho_{1,0}) - \gamma_2^s(\rho_{2,0}) = \alpha_2 \bar{\bar{\rho}}_1 - \bar{\bar{\bar{\rho}}}_2 = \bar{\bar{m}}_2 \quad \text{for } t > \bar{\bar{t}},$$

and based on (3.19), $\bar{\bar{m}}_2$ is positive.

Now, for $\bar{\Gamma}_1 = \frac{\gamma_3^s}{\alpha_3}(\rho_{3,0})$, two sub-cases will arise. First, if $f_1(\rho_{1,0}) \geq \frac{\gamma_3^s}{\alpha_3}(\rho_{3,0})$, then the interaction of the rarefaction and the Big shock on $I_1$ will not produce a Big shock with positive speed. Therefore, no waves on $I_1$ can hit the junction. Second, if $f_1(\rho_{1,0}) < \frac{\gamma_3^s}{\alpha_3}(\rho_{3,0})$, then the new Big shock, generated due to the interaction of the small rarefaction and the previous Big shock, can move toward the junction. When this Big shock hits the junction, even though the rate of stacking up the vehicles decreases, it still remains positive. This is true because $f_1(\rho_{1,0}) < \frac{\gamma_3^s}{\alpha_3}(\rho_{3,0})$, we have:

$$\bar{m}_2 = \alpha_2 \bar{\Gamma}_1 - \bar{\Gamma}_2 = \alpha_2 f_1(\rho_{1,0}) - \gamma_2^s(\rho_{2,0})$$

$$< \alpha_2 \frac{\gamma_3^s}{\alpha_3}(\rho_{3,0}) - \gamma_2^s(\rho_{2,0}) = \alpha_2 \bar{\bar{\bar{\rho}}}_1 - \bar{\bar{\bar{\rho}}}_2 = \bar{\bar{m}}_2,$$

and from (3.19), we know that $\bar{\bar{m}}_2 > 0$.

For the case with $\mu = 0$, both $I_1$ and $I_3$ are in the free-flow regime. Because there is no other factor that affects the behavior of the queue, the situation
does not change after this point and the junction will remain in mode 2. This completes the proof.

**Corollary 9.** Beginning with the GRP assumption, the number of depletion and formation of queues, is finite.

### 3.3 Well-posedness

This section is concerned with the existence of solutions to the generalized Cauchy problem (3.3). Given an initial condition \( \rho_{i,0} \) for \( i = 1, 2, 3 \), with sufficiently small total variation, an admissible weak solution is defined as follows:

**Definition 3.** A solution \((\rho_1, \rho_2, \rho_3, m_2, m_3)\), where \( \rho_i \in C([0, \infty); L^1_{\text{loc}}(I_i)) \) for \( i = 1, 2, 3 \), and \( m_j \in W^{1,\infty} \) for \( j = 2, 3 \), is a weak solution of (3.3) if:

- \( \rho_i(t, x) \) is an entropy admissible solution to (3.1) if it satisfies the Kruzkov entropy admissibility condition:
  \[
  \int_0^T \int_{I_i} (|\rho_i - k_i| \phi_t + \text{sgn}(\rho_i - k_i)(f_i(\rho_i) - f_i(k_i))\phi_x) \, dt \, dx \geq 0,
  \]
  for every \( k_i \in [0, \rho_i^{\text{max}}] \) and every smooth, non-negative test function \( \phi : [0, T] \times I_i \to \mathbb{R} \) with compact support on \([0, T] \times I_i \setminus \{0\}\).

- for a.e. \( t > 0 \):
  \[
  \dot{m}_j = \alpha_j f_1(\rho_1(t, 0^-)) - f_j(\rho_j(t, 0^+)), \quad j = 2, 3,
  \]
  where \( \rho_i \)'s are bounded variation entropy admissible solutions.

The key result to this section is the following theorem.

**Theorem 10.** Assume that there exists a constant \( C \) such that, for initial conditions \( \rho_{i,0} \in L^1 \), \( i = 1, 2, 3 \), we have:

\[
\text{Total Variation } \{\rho_{i,0}\} \leq C, \quad (3.20)
\]

then the Cauchy problem (3.3) and (3.4), for all \( T > 0 \), admits a weak solution \((\rho_1, \rho_2, \rho_3, m_2, m_3)\) such that:
• \( \rho_i \) is a weak entropy solution of \( \partial_t \rho_i + \partial_x f(\rho_i) = 0 \), where \( \rho_i(t, x) \in [0, T] \times I_i, \quad i = 1, 2, 3, \)

• For a.e. \( t \in [0, T] \), \( \dot{m}_j(t) = \alpha_j f_1(\rho_1(x, 0^-)) - f_j(\rho_j(x, 0^+)), \quad j = 2, 3, \)

• \( \rho_i(0, x) = \rho_{i,0}(x), \quad i = 1, 2, 3, \)

• \( m_j(0) = m_{j,0}, \quad \text{for} \ j = 2, 3, \)

• \( \mathcal{RS}_m(\rho_1(t, 0^-), \rho_2(t, 0^+), \rho_3(t, 0^+)) = (\rho_1(t, 0^-), \rho_2(t, 0^+), \rho_3(t, 0^+)), \quad \text{for a.e.} \ t \in [0, T]. \)

The remainder of this section is devoted to prove Theorem 10 and construct a solution to the generalized Cauchy problem. We use the wave-front tracking method to construct approximate solutions. Note that, to provide an admissible solution to the Cauchy problem there are other methods such as the generalized Lax formula [2]. However, for our model, the wave-front tracking algorithm is more appropriate.

3.3.1 Wave-front Tracking Approximations

In this subsection we describe a procedure to construct piecewise constant approximations via wave-front tracking algorithm [15, 41]. As described in [4], an \( \epsilon \)-approximation wave-front tracking solution of the system of conservation laws (3.1) is a piecewise constant function \( \rho = \rho(x, t) \), such that its jumps are located along finitely many straight lines in the \((t, x)\)-plane and approximately satisfy the Rankin-Hugoniot conditions. See [25] and [27] for general theory and applications on networks.
To construct a wave-front tracking solution, for each road $I_i$ with $i \in \{1, 2, 3\}$, define a sequence of piecewise constant functions $\rho_{(i,0),\nu}$, see Figure 3.2, with a uniformly bounded total variation (condition (3.20)) such that:

$$\lim_{\nu \to \infty} \rho_{(i,0),\nu} = \rho_{i,0}$$

in $L^1_{loc}(I_i; [0, \rho_{i,max}])$.

For each $\nu \in \mathbb{N} \setminus \{0\}$, $\rho_{(i,0),\nu}$ has a finite number of discontinuities. The algorithm starts by setting $t = 0$ and solving the Riemann problem at junction $J$ and updating the fluxes on the roads. Then, every rarefaction wave is approximated with a finite number of jumps, also called rarefaction shocks, of size at most $\frac{1}{\nu}$, that move with a speed near the characteristic speed of the connected states. Furthermore, if $\sigma_i$ for $i = 1, 2, 3$, is in the range of rarefaction shock, then its speed is zero. We repeat this procedure for each event, namely, every time there is an interaction between waves, or between waves and the junction, as well as when an active queue becomes empty.

Here, we provide the definition of the $\epsilon$-approximate wave-front tracking solution as in [26].

**Definition 4.** For small $\epsilon > 0$, $\rho_\epsilon = (\rho_{1,\epsilon}, \rho_{2,\epsilon}, \rho_{3,\epsilon})$ and $m_{j,\epsilon}$ for $j \in \{2, 3\}$, are considered as an $\epsilon$-approximate wave-front tracking solution to (3.3), provided that:

1. For every $i = 1, 2, 3$, the time-map $\rho_{i,\epsilon}: [0, \infty) \to L^1_{loc}(I_i; [0, \rho_{i,max}])$ is a continuous map.
2. \( m_{j,\epsilon}(0) = m_{j,0} \) for \( j \in \{2,3\} \) and \( m_{j,\epsilon} \) is a time-map such that \( m_{j,\epsilon} : [0,\infty) \to W^{1,\infty}([0,\infty)) \).

3. As a function of two variables, \( \rho_{i,\epsilon} = \rho_{i,\epsilon}(t, x) \) for \( i = 1, 2, 3 \), is piecewise constant, with finitely many discontinuities occurring along the straight lines in the \((t, x)\) – plane. Furthermore, jumps of \( \rho_{i,\epsilon}(t, x) \) are of two types: shocks and rarefaction waves, which are: \( \mathcal{J}_i(t) = \mathcal{S}_i(t) \cup \mathcal{R}_i(t) \).

4. Along each shock, \( x(t) = x_{i,\beta}(t), \beta \in \mathcal{S}_i(t) \), for \( i = 1, 2, 3 \), we have: 
\[
\rho_{i,\epsilon}(t, x_{i,\beta}(t)--) < \rho_{i,\epsilon}(t, x_{i,\beta}(t)+).
\]
Moreover,
\[
\left| \dot{x}_{i,\beta}(t) - \frac{f(\rho_{i,\epsilon}(t, x_{i,\beta}(t)-)) - f(\rho_{i,\epsilon}(t, x_{i,\beta}(t)+))}{\rho_{i,\epsilon}(t, x_{i,\beta}(t)-) - \rho_{i,\epsilon}(t, x_{i,\beta}(t)+)} \right| \leq \epsilon.
\]

5. Along each rarefaction front \( x(t) = x_{i,\beta}(t), \beta \in \mathcal{R}_i(t) \), for \( i = 1, 2, 3 \), we have: 
\[
\rho_{i,\epsilon}(t, x_{i,\beta}(t)+) < \rho_{i,\epsilon}(t, x_{i,\beta}(t)-) < \rho_{i,\epsilon}(t, x_{i,\beta}(t)+) + \epsilon.
\]
Moreover,
\[
\dot{x}_{i,\beta}(t) \in [f'(\rho_{i,\epsilon}(t, x_{i,\beta}(t)-)), f'(\rho_{i,\epsilon}(t, x_{i,\beta}(t)+))].
\]

6. For \( i = 1, 2, 3 \), \( \|\rho_{i,\epsilon}(0, \cdot) - \rho_{0,\epsilon}(\cdot)\|_{L^1(I_i)} < \epsilon \).

7. For a.e. \( t > 0 \), \( \mathcal{RS}_m(\rho_{1,\epsilon}(t, 0^-), \rho_{2,\epsilon}(t, 0^+), \rho_{3,\epsilon}(t, 0^+)) = (\rho_{1,\epsilon}(t, 0^-), \rho_{2,\epsilon}(t, 0^+), \rho_{3,\epsilon}(t, 0^+)) \).

8. For a.e. \( t > 0 \) and \( j \in \{2,3\} \), \( \dot{m}_{j,\epsilon} = \alpha_j f_1(\rho_{1,\epsilon}(t, 0^-)) - f_j(\rho_{j,\epsilon}(t, 0^+)) \).

### 3.3.2 Bound on the Total Variation of the Fluxes

The total variation of the fluxes \( f_i \) on road \( I_i \) at time \( t \) is defined as:
\[
TV_{f_i}(t) = \sup \left\{ \sum_{\ell=1}^{N} |f_i(\rho(x_\ell + \epsilon, t)) - f_i(\rho(x_\ell - \epsilon, t))| \right\}, \text{ for small } \epsilon > 0,
\]
where \( N \geq 1 \), and the points \( x_\ell \) are the discontinuities of \( f_i \). Based on this definition, total variation is the summation of the magnitudes of the jumps in each of the fluxes on the roads.
**Theorem 11.** For a junction of one incoming road and two outgoing roads, the total variation of fluxes for any time $\bar{t} \geq 0$ satisfies the following inequality:

$$TV_f(\bar{t}^+) \leq TV_f(\bar{t}^-) + Rf(\sigma) , \quad (3.21)$$

where $f(\sigma) = \max\{f_i(\sigma_i) | i = 1, 2, 3\}$, quantity $R = \max\{3, \frac{\alpha_2}{\alpha_3}, \frac{\alpha_3}{\alpha_2}\}$.

In order to prove this theorem, we need some lemmas.

**Lemma 12.** When a wave $(\rho_{i,\ell}, \rho_{i,r})$ on a road $I_i$ for $i = 1, 2, 3$, hits the junction at time $\bar{t}$, the total variation of fluxes satisfies the following inequality

$$TV_f(\bar{t}^+) \leq TV_f(\bar{t}^-) + R[\Gamma_i] , \quad (3.22)$$

where $[\Gamma_i]$ is the magnitude of the jump of the fluxes on the road related to the wave.

**Proof.** Depending on the availability of a queue, one can categorize the arising situations into two general cases: when there exists an active queue, and when both queues are empty. **First**, consider the case when there exists an active queue. The queue can either deplete or stack up. In order to categorize the possible situations, we define a sterile case as a scenario in which a wave that hits the junction only affects the rate of the change of the queue on its corresponding road and does not produce any waves on the other roads. We call all of the other cases as non-sterile cases. In the following, we investigate the non-sterile and sterile subcases that arise. Without loss of generality, assume that queue 3 is active i.e., $m_3 \neq 0$.

We first show that, if a wave, $(\rho_{1,\ell}, \rho_{1,r})$, interacts with the junction from incoming road $I_1$, it will not increase the total variation of the fluxes. After the interaction of a Big shock or small shock on $I_1$ with junction, we have:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) = |\Gamma_2^+ - \Gamma_2^-| - |\Gamma_1^+ - \Gamma_1^-| = (\alpha_2 - 1)[\Gamma_1] = -\alpha_3[\Gamma_1] < R[\Gamma_1]. \quad (3.23)$$

$^3$To simplify the notations, in order to represent the magnitude of the jumps instead of $[|\Gamma_i|]$, we use $[\Gamma_i]$. 

This is true because, based on the FIFOQ model, due to the existence of the queue before and after the interaction of the wave (i.e. system is in mode 3), we have $\Gamma_2^\pm = \alpha_2 \Gamma_1^\pm$. Moreover, after the interaction of a shock wave, no new wave will come out of the junction on $I_1$. Thus, only the state on the outgoing road $I_2$ will be affected and the difference of the total variation of the fluxes follows equation (3.23) which is in agreement with (3.22).

Now, when a rarefaction wave on $I_1$ hits the junction, although a new wave will be generated on the incoming road, the total variation of fluxes will not increase. Note that, for this case we must have: $\gamma_2^s(\rho_2^-) > \alpha_2 \gamma_1^d(\rho_1^-)$. Now, according to the FIFOQ model, after the interaction of the wave $(\rho_{1,\ell}, \rho_1, \ell)$ with the junction at $t = \bar{t}$, we have:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) = |\Gamma_1^+ - f_1(\rho_{1,\ell})| + |\Gamma_2^+ - \Gamma_2^-| - |f_1(\rho_{1,\ell}) - \Gamma_1^-| = |\gamma_1^d(\rho_{1,\ell}) - \frac{\gamma_2^s(\rho_2^-)}{\alpha_2}| + |\gamma_2^s(\rho_2^-) - \alpha_2 \gamma_1^d(\rho_1^-)| - |\gamma_1^d(\rho_{1,\ell}) - \gamma_1^d(\rho_1^-)|$$

$$= -\alpha_3(\gamma_2^s(\rho_2^-) - \gamma_1^d(\rho_1^-)) < 0 ,$$

that satisfies (3.22).

Now, assume that a wave on $I_2$ hits the junction. Note that, waves that move toward the junction on $I_2$, belong to the class of events where either quantity $\mu \neq 0$ or it becomes non-zero after the interaction of the wave with the junction. Again, since the active queue corresponds to $I_3$, according to the model, flux on $I_1$ is proportional to the flux on $I_2$ and thus we have:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) = |\Gamma_1^+ - \Gamma_1^-| - |\Gamma_2^+ - \Gamma_2^-| = (\frac{1}{\alpha_2} - 1)|\Gamma_2^+ - \Gamma_2^-| = (\frac{\alpha_3}{\alpha_2})[\Gamma_2] < R[\Gamma_2] .$$

Although, after the interaction of a wave on $I_2$ with junction, the total variation of the fluxes increases, it still remains bounded and the statement (3.22) holds true.

Now, consider the case when a wave on $I_3$ hits the junction. Due to the existence of the queue, the state on $I_3$ is in the congestion regime. Therefore,
the waves are of two types, small shock and small rarefaction. Investigating all possible scenarios reveals that they are all sterile cases. In other words, after the interaction of the wave on $I_3$ with the junction, only queue 3 is affected. So, we have:

$$TV_f(\tilde{t}^+) - TV_f(\tilde{t}^-) = -[\Gamma_3^+ - \Gamma_3^-] = -[\Gamma_3] < R[\Gamma_3] ,$$

therefore, statement (3.22) holds.

Now that we considered all of the possible cases with an existing active queue, we investigate the second case, where no active queue is available when a wave hits the junction. In this case, after the interaction of the wave with the junction, states on the roads may change and one queue may start filling up. We show that, after interaction of a wave with the junction, the difference of the total variation of the fluxes before and after the interaction, is bounded and we have:

$$TV_f(\tilde{t}^+) - TV_f(\tilde{t}^-) \leq 0 . \quad (3.24)$$

To show this, we split the scenarios into two cases: (i) when, after the interaction of a wave on $I_i$ with junction, no new wave is generated on the road $I_i$, and (ii) when a new wave is produced on $I_i$. In the case (i), after the interaction of the wave, at least for one of the outgoing roads, we have: $\Gamma_j^\pm = \alpha_j \Gamma_1^\pm$. Note that, before the interaction, this statement is true for both outgoing roads (since the system is in mode 1). If, after the interaction of the wave on incoming road $I_1$ with the junction, no queue starts filling up, then $\Gamma_j^\pm = \alpha_j \Gamma_1^\pm$ for $j = 2, 3$, and we have: $TV_f(\tilde{t}^+) - TV_f(\tilde{t}^-) = 0$. On the other hand, if after the interaction of the wave on $I_1$ with the junction, without loss of generality, queue 3 starts filling up, then, according to the model, we have:

$$TV_f(\tilde{t}^+) - TV_f(\tilde{t}^-) = [\Gamma_2] + [\Gamma_3] - [\Gamma_1]$$

$$= (\alpha_2 - 1)[\Gamma_1] + [\Gamma_3] = -\alpha_3[\Gamma_1] + [\Gamma_3] = -\dot{m}_3^+ \leq 0 ,$$

which satisfies statement (3.24).
Now, if a wave on an outgoing road $I_j$ for $j \in \{2, 3\}$, hits the junction and causes its corresponding queue to stack up vehicles, then:

$$TV_j(\bar{t}^+) - TV_j(\bar{t}^-) = -[\Gamma_j] < 0 \quad \text{for } j \in \{2, 3\}.$$  

We now consider case (ii). We claim that this case can only happen on the incoming road $I_1$, through the interaction of a rarefaction wave with the junction. This is true because, in order to satisfy the conditions related to this case (namely having no active queue) on the outgoing roads $I_j$ for $j \in \{2, 3\}$, they should be in the free-flow regime. Therefore, waves that can hit the junction can only be Big shocks. However, interaction of a Big shock on $I_j$ with the junction, activates queue $j$ and does not introduce a new wave on $I_j$ for $j \in \{2, 3\}$. We studied these cases in the above discussions.

Now, assume that a rarefaction wave $(\rho_{1,\ell}, \rho_{1,r})$ on $I_1$, moves toward the junction. In order to generate a new wave on $I_1$ after the interaction of $(\rho_{1,\ell}, \rho_{1,r})$ with the junction we need to have:

$$\gamma_d^d(\rho_{1,r}) < \frac{\gamma_s^2(\rho_2^-)}{\alpha_2} < \frac{\gamma_s^3(\rho_3^-)}{\alpha_3} < \gamma_d^d(\rho_{1,\ell}) . \quad (3.25)$$

This implies that, after the interaction of this wave with the junction, a new wave is generated on $I_1$ as well as on the outgoing roads, quantity $\mu$ becomes non-zero and queue 2 starts filling up. Therefore, after substituting the values based on the FIFOQ model and simplifications, the difference between the total variation of the fluxes satisfies the following inequality:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) = |\Gamma_1^+ - f_1(\rho_{1,r})| + |\Gamma_2| + |\Gamma_3| - |\Gamma_1^- - f_1(\rho_{1,r})|$$

$$= -\alpha_2 \left( \frac{\gamma_s^3(\rho_3^-)}{\alpha_3} - \frac{\gamma_s^2(\rho_2^-)}{\alpha_2} \right) < 0 .$$

This completes the proof.

\[\square\]

**Corollary 13.** For $\bar{t} > 0$, the time when a wave on road $I_i$ for $i \in \{1, 2, 3\}$, hits the junction, the following statement holds for the total variation of the fluxes:

$$TV_f(\bar{t}^+) \leq C_i TV_f(\bar{t}^-). \quad (3.26)$$
Proof. Based on the previous lemma, we know that:

\[
TV_f(\bar{t}^+) \leq TV_f(\bar{t}^-) + R[\Gamma_i] \\
= TV_f(\bar{t}^-) - R[\Gamma_i] + (R + 1)R[\Gamma_i] \\
\leq \max\{R + 1, 1\}(RTV_f(\bar{t}^-) - R[\Gamma_i]) + \max\{R + 1, 1\}R[\Gamma_i] \\
= R(R + 1)TV_f(\bar{t}^-),
\]

let \( C_j = R(R + 1) \), then (3.26) follows.

Hence, for the events related to the interaction of the waves with the junction, the total variation of fluxes before the interaction is proportional to the total variation of the fluxes after the interaction and the proportionality coefficient depends on split ratios and number of roads. Now, we are ready to prove Theorem 11.

Proof of Theorem 11. The total variation of fluxes at time \( \bar{t} \), can change only if any of the following events happens:

(i) \( \bar{t} = 0 \), initial updates happen

(ii) two waves hit each other on a road at time \( \bar{t} \)

(iii) a wave on a road hits the junction at time \( \bar{t} \)

(iv) the length of an active queue hits zero at time \( \bar{t} \)

In the first case, after \( \bar{t} = 0 \), as the junction updates the fluxes on the roads, regardless of the type of activity of the queue, at most three waves can leave the junction. Since the jump of the fluxes cannot exceed \( f(\sigma) \), equation (3.21) is true at \( \bar{t} = 0 \). Note that, in the case when we start with the GRP, \( TV_f(\bar{t}^-) = 0 \). In the second case, when two waves hit each other on the road, the total variation of the fluxes decreases [25] and so the statement holds true. For the third case, based on Lemma 12, the statement (3.21) is established.

And finally, for the fourth case, without loss of generality assume that queue 3 was active and hits zero. After the complete depletion of the queue,
based on Lemma 2, either no queue forms or the other queue start stacking up. For both cases, we show that the difference of the total variation of the fluxes before and after the complete depletion of a queue is bounded, and we have:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) \leq 2f(\sigma).$$

(3.27)

Regarding the events after the complete depletion of an active queue, we first consider the case where no queue forms. This means $m_2^+ = m_3^+ = 0$. Moreover, quantity $\mu = 0$ before the depletion of the queue, therefore after the complete depletion, only state on $I_3$ will change. Thus, one Big shock will be generated on $I_3$ and we have:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) = |\Gamma_3^+ - \Gamma_3^-| = |\alpha_3\gamma_3^+(\rho_3^-) - \gamma_3^-(\rho_3^-)| = -\dot{m}_3^- \leq f(\sigma),$$

that satisfies (3.27).

For the second case, assume that as $m_3$ hits zero, queue 2 starts forming. Now, based on the relationship between the supplies and demand in the FIFOQ model, either (i) we only have a rarefaction wave on $I_1$ or, (ii) we also have a shock wave on $I_3$. To be more precise, if after the complete depletion of the queue, the quantity $\mu$ was not zero, then waves will appear on both roads $I_1$ and $I_3$, otherwise, only $I_1$ experiences a new wave. For case (i) we have:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) = |\Gamma_1^+ - \Gamma_1^-| = |\gamma_1^-(\rho_1) - \gamma_1^+(\rho_2^-)| = \alpha_3\dot{m}_3^- = 1\alpha_2\dot{m}_2^- \leq f(\sigma),$$

and for case (ii), after substitution of the supplies and demand values and simplifications, we have:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) = |\Gamma_1^+ - \Gamma_1^-| + |\Gamma_3^+ - \Gamma_3^-| = \dot{m}_2^- - \dot{m}_3^- \leq 2f(\sigma),$$

This completes the proof.

Remark 2. The total variation of the fluxes for the events when queue $j \in \{2, 3\}$ hits zero, satisfies the following inequality:

$$TV_f(\bar{t}^+) - TV_f(\bar{t}^-) \leq \dot{m}_k^+ - \dot{m}_j^-,$$

(3.28)

where $k \in \{2, 3\} \setminus \{j\}$. Note that, for the case in which, after the complete depletion of a queue, no queue forms, we have $\dot{m}_k^+ = 0$. 
3.3.2.1 Finiteness of Queue Related Events

As discussed earlier in the proof of Theorem 11, the cause for the increase in the total variation of the fluxes on the roads, is not limited to the interaction of the waves with increasing fluxes with the junction. In fact, complete depletion of an active queue can also lead to an increase in the total variation of the fluxes. More precisely, if immediately after the complete depletion of a queue, the other queue starts filling up, a wave with increasing flux will be generated on the incoming road. In addition, as a queue $j$ for $j \in \{2, 3\}$ empties, if the other queue does not activate, although the flux on outgoing road $j$ decreases, the total variation of the fluxes increases.

In this subsection, we prove that the queue-related events are finite, namely the formation and depletion of the queues can happen finitely many times. To this end, we investigate the mechanisms that lead to the depletion of an active queue. Moreover, for any sequence of piecewise constant $\rho_{0, \nu}$ with finitely many discontinuities that satisfies the following conditions:

\[
\text{Total Variation } \{\rho_{0, \nu}\} \leq \text{Total Variation } \{\rho_0\} \leq C, \\
||\rho_{0, \nu} - \rho_0||_{L^1} < \frac{1}{\nu}, \quad \text{for every } \nu \in \mathbb{N},
\]

(3.29)

where $\rho_0 \in L^1$ satisfies property (3.20), we prove that an infinite cascade of queue related events will never happen. We need the following two lemmas to characterize the conditions under which a queue $j$ with $\dot{m}_j > 0$ starts depleting.

**Lemma 14.** Every Big shock that moves toward the junction, either itself or its parent, comes from the piecewise constant initial data $\rho_{0, \nu}$, or is produced due to the interaction of a Big shock leaving the junction with a small shock or small rarefaction.

**Proof.** Based on the definition of Big shock, in order to generate a Big shock on the roads via interaction of the waves, one of the interacting (parent) waves must be a Big shock. On the incoming road $I_1$, interaction of a small rarefaction or a small shock respectively from right and left with a Big shock leaving
the junction, may produce a Big shock with a positive speed. Note that, rarefaction waves on the incoming road originated from the junction, are either due to a mode change from $j$ to $k$, where $j \in \{2, 3\}$ and $k \in \{2, 3\} \setminus \{j\}$, or generated due to the interaction of a small rarefaction wave on $I_j$ for $j \in \{2, 3\}$, with the junction.

Figure 3.3: On the incoming road, interaction of a small rarefaction wave or a small shock with a Big shock with negative speed can produce a positive speed Big shock.

On the outgoing roads, the interaction of a small rarefaction or a small shock respectively from left and right with a Big shock with positive speed, may produce a Big shock that moves toward the junction. The rarefaction in this case is coming out of the junction and is produced due to the interaction of a rarefaction wave with positive speed on $I_1$ with the junction.

Figure 3.4: On the outgoing roads, interaction of a small rarefaction wave or a small shock with a Big shock with positive speed can produce a negative speed Big shock.
Remark 3. Small rarefaction waves moving toward the junction, correspond to $\rho_{0,\nu}$. Namely, either they are exactly available as discontinuities in $\rho_{0,\nu}$, or they are produced via interaction of themselves.

Lemma 15. Small shocks on the roads moving toward the junction, can only be generated through interaction with each other or with small rarefaction waves.

Proof. Interaction of two small shocks produces a new small shock. On the incoming road $I_1$, if a positive speed small rarefaction could catch up to a positive speed small shock moving in front of it, then a small shock will be generated. Moreover, if a positive speed small shock interacts with a positive speed small rarefaction, then after the interaction, a small shock will be produced. Similarly for the outgoing roads, if a small shock and a small rarefaction, both moving toward the junction, could interact on the outgoing road, then a new small shock will be generated.

By using the results from the previous lemmas, we characterize the circumstances under which an active queue depletes in the following lemma. In addition to the cases where we start with an active depleting queue, we study the mechanisms that can cause an active absorbing queue deplete. Moreover, we will see under what conditions, the other queue (which was idle) starts forming.

Lemma 16. Consider the situation where queue $j$ is active and $\dot{m}_j > 0$ for $j \in \{2, 3\}$. The scenarios that lead to a mode change from $j$ to $k$ with $k \in \{2, 3\} \setminus \{j\}$, include a wave corresponding to the initial piecewise constant data $\rho_{0,\nu}$.

Proof. Without loss of generality, assume that queue 3 is active and stacking up. Therefore, we have: $\dot{m}_3 = \alpha_3 \Gamma_1 - \Gamma_3 > 0$. In order to change the sign of $\dot{m}_3$ to negative, we should either decrease the magnitude of $\Gamma_1$, or increase the value of $\Gamma_3$. We investigate each of these scenarios by considering two possible conditions, namely, when $\mu \neq 0$ and when $\mu = 0$.
For the case with $\mu \neq 0$, a Big shock on $I_1$, small shock on $I_2$ and small rarefaction on $I_3$ are the candidates that after the interaction with the junction, can lead to a change in the dynamic of the queue. In the following, we consider each of these interactions.

For the case in which a Big shock on $I_1$ interacts with the junction, after this interaction, the states on $I_1$ and $I_2$, are in the free-flow regime. Moreover, a Big shock on $I_2$ comes out of the junction. If $I_1$ remains in the free-flow regime until the complete depletion of the queue, the whole junction will be in the free-flow regime after the complete depletion of the queue and thus no (new) queue will form. To avoid this situation, we need to create conditions under which the quantity $\mu$ becomes non-zero, namely, the outgoing road $I_2$ becomes congested before the complete depletion of queue 3. Now, since $I_2$ is in the free-flow regime, the only way to create congestion on this road is through the interaction of a Big shock on $I_2$ with the junction. Therefore, a new Big shock with negative speed on $I_2$ that hits the junction, should be generated. To this end, either the Big shock on $I_1$ should be followed by a small rarefaction where both move toward the junction or, according to Lemma 14, the Big shock on $I_2$ that had left the junction should hit a small shock and produce a new Big shock that moves toward the junction. Now, if the junction is kept in the congestion regime, after the depletion of queue 3, queue 2 will form. Note that, both small rarefaction and small shock correspond to the waves from the initial data. In other words, they are either exactly the discontinuities existing in the initial data or the result of interactions of these discontinuities with each other.

Now, we consider the case in which a small shock on $I_2$ interacts with the junction. Because we have an absorbing queue 3, a small shock on $I_2$ that hits the junction can change the dynamic of the active queue via decreasing the flux on the incoming road. To be precise, when queue 3 is active and $\mu \neq 0$, the flux on the incoming road is equal to the effective supply on the outgoing road $I_2$. Thus, as the supply decreases due to a small shock that hits the junction, the flux on $I_1$ will decrease and the behavior of the queue may
change. As the junction maintains its current congestion regime through the
complete depletion of queue 3, queue 2 starts filling up. Note that, based on
Lemma 15, any small shock that interacts with the junction corresponds to
the waves from the initial data. Namely, they are either exactly available as
discontinuities in the initial data, or created due to the interactions of these
discontinuities with each other.

Next, we consider the case in which a small rarefaction on $I_3$ interacts
with the junction. After this interaction, $\Gamma_3$ increases and the behavior of the
queue may change. Therefore, after the complete depletion of queue 3, while
the junction is held in the congestion regime, queue 2 starts filling up. Note
that, according to Remark 3, any small rarefaction that interacts with the
junction corresponds to the waves from the initial data.

Now, we consider the cases with $\mu = 0$. In this situation, a small shock
on $I_1$, a Big shock on $I_2$ and a small rarefaction on $I_3$ are the candidates to
interact with the junction to enforce queue 3 toward depletion.

Consider the case when a small shock on $I_1$ interacts with the junction.
After this interaction, a new small shock will appear on $I_2$ and due to the
decrease in the flux on $I_1$, queue 3 may start depleting. In order to have an
immediate activation of queue 2 after depletion of queue 3, quantity $\mu$ should
become non-zero. To this end, similar to the discussion above for road $I_1$,
before complete depletion of queue 3, either a small rarefaction on $I_1$ or a
Big shock on $I_2$, should hit the junction. Now, if the junction stays in the
congestion regime, after the complete depletion of queue 3, queue 2 forms.
Note that the small shock that initiates this entire scenario, corresponds to
the waves from the initial data.

Next, we consider the case with a Big shock on $I_2$ interacting with the
junction. This interaction can decrease the value of the flux on $I_1$ and therefore
cause queue 3 to deplete. After the interaction of the Big shock on $I_2$ with
the junction, quantity $\mu$ becomes non-zero and if the junction remains in the
congestion regime, queue 2 will form after the depletion of queue 3. Note that,
a Big shock on $I_2$ is either from the initial data or generated according to
Lemma 14, via interaction of a Big shock with a small rarefaction or a small shock, both of which correspond to the waves from the initial data.

Next, we consider the case with a small rarefaction on $I_3$ interacting with the junction. In this case, $\Gamma_3$ increases and therefore the dynamic of the queue may change. In order to have a mode change from 3 to 2, conditions similar to what we discussed earlier for the case with a small shock on $I_1$, are needed. Note that the small rarefaction that initiates this entire scenario, corresponds to the set of waves from the initial data. This completes the proof.

Lemma 16 emphasizes that the behavior of the queues will change only if there are some specific discontinuities in the initial data. Thus, for each of these desired scenarios, we need a specific piece of information from the initial data. Otherwise, the scenario fails to happen. Now, using the above discussions, we have the following lemma:

**Lemma 17.** Beginning from an initial piecewise constant data $\rho_{0,\nu}$ that satisfies the condition (3.29), the number of queue formation and depletion events is finite.

*Proof.* Let us define $N_q(t)$ as the number of formation and depletion events of the queues until time $t$. To have $\lim_{t \to T} N_q(t) = \infty$, where $T$ can be infinity, we need a sequence of infinite number of depletion and formation of the queues. We show that such sequence of events does not exist.

If for $t = 0^+$, there exists an active queue absorbing the congestion and, in spite of the interactions of the waves with the junction, continues stacking up the vehicles, then $N_q(t) = 1$ for $t \geq 0$.

If there is an active queue, one possible situation that increase $N_q$, is a mode change from $j$ to $k$, where $j \in \{2, 3\}$ and $k \in \{2, 3\} \setminus \{j\}$. Otherwise, the system either goes to mode 1, or stays in mode $j \in \{2, 3\}$. In both cases, $N_q$ does not increase. Based on Lemma 16, all possible cases that can lead to the depletion of an active queue and formation of the other queue rely on the waves from the initial data. Therefore, the number of queue events...
generated from an immediate mode change from $j$ to $k$, where $j \in \{2, 3\}$ and $k \in \{2, 3\} \setminus \{j\}$, is finite. Note that, in case of a mode change from $j$ to $k$, where $j \in \{2, 3\}$ and $k \in \{2, 3\} \setminus \{j\}$, $N_q(t)$ increases by 2 units.

Moreover, if the initial active queue $j \in \{2, 3\}$ empties and both queues stay idle after the complete depletion of queue $j$, then $N_q$ increases by 1 unit.

Now in order to create conditions that lead to the activation of one of the queues and increase the value of $N_q$, we need scenarios similar to the ones in case $\mu = 0$, that are discussed in the proof of Lemma 16. Therefore, all possible cases that can lead to the formation of a queue rely on the waves from the initial data. Thus, in this category, the number of queue events is finite. This completes the proof.

**Corollary 18.** Lemma 17 implies that, there exists a time $T^* > 0$, such that $N_q(T^*)$ is constant and:

\[ N_q(t) = N_q(T^*) , \quad \text{for } t \geq T^* . \]

This is true because for the time after $T^*$, either one queue stacks up forever, or both queues remain idle and no queue forms again and so $N_q(t)$, for $t > T^*$, remains constant.

### 3.3.3 Existence of a Wave-front Tracking Approximation

In order to prove the existence of a wave-front tracking solution, we should prove that for each instant in time, the number of discontinuities of a piecewise constant approximate solution produced via the wave-front tracking algorithm, is finite. In this subsection we prove that, as time passes, the number of discontinuities of the wave-front tracking approximation will never go to infinity.

**Lemma 19.** For $t \geq 0$ and $\nu \in \mathbb{N} \setminus \{0\}$, $\rho_{\nu}$ which is produced based on the wave-front tracking algorithm, provides an $\frac{1}{\nu}$-approximate solution to (3.3).
Proof. For $t \geq 0$, let us define quantity $N_\nu(t)$ to be the number of discontinuities of the approximation solution $\rho_\nu(t,.)$ on all roads. We need to show that, for each instant in time, we have: $N_\nu(t) < \infty$. Now beginning with a finite number of discontinuities related to $\rho_{0,\nu}$, we show that $N_\nu(t)$ will never blow up.

When two waves interact on the roads, the number of discontinuities decreases. In fact, discontinuities on the roads can increase only if new waves come out of the junction. New waves can be generated on the roads from the junction, only if waves hit the junction or queues deplete. As discussed in the previous subsections, the number of interactions of the waves with the junction is bounded by $N(0)$ which is the number of discontinuities of $\rho_{0,\nu}$, and number of depletion of the queues is also bounded by $N_q(T^*)$. Therefore, $N_\nu$ is bounded and we have:

$$N_\nu(t) \leq R(N(0) + N_q(T^*)) .$$

3.3.4 Uniform Bound on the Total Variation of Fluxes

Regarding the above discussions, we can conclude that in the case when we start with an initial piecewise constant data with finite number of discontinuities, the total variation of fluxes is uniformly bounded and we have the following theorem.

**Theorem 20.** Beginning from an initial data that satisfies the condition (3.29), we have:

$$TV_f(t) \leq \kappa TV_f(0^-) \quad \text{for} \ t \geq 0 ,$$

(3.30)

where $\kappa$ is a constant value that depends on the initial data and quantity $C_J$.

**Proof.** Total variation of the fluxes can increase due to the interaction of the waves with the junction and depletion of the active queues. Let us define
$N_e(t)$ as the number of events that can increase the total variation of the fluxes through time $t$.

For the non-sterile cases, when a wave hits the junction, it produces new waves on the roads. In the proofs of Theorem 11 and Lemma 12, the situations where after the interaction of a wave with the junction, the total variation of the fluxes increases, are identified. Moreover, in Lemma 14 and 15, as well as Remark 3, we show that all the waves initiating these situations correspond to the initial data namely, if we did not have some specific discontinuities in the initial data, those scenarios could never happen. Let $N(0)$ to be the number of the waves existing in the initial data. Since the initial data consists of finite number of discontinuities and therefore finite number of waves, i.e. $N(0) < \infty$, in the worst case scenario, the number of the events that increase the total variation of the fluxes, is finite. Using the result from Corollary 18, we have:

$$N_e(t) \leq N(0) + N_q(t) \quad \text{for } t \geq 0 .$$

Now by Corollary 13 and Remark 2, we have:

$$TV_f(t^+) \leq C_J TV_f(t^-) + TV_f(t^-) + [\dot{m}]
= (C_J + 1)TV_f(t^-) + [\dot{m}]
\leq (C_J + 1)^{N(0)} TV_f(0^-) + N_q(t)[\dot{m}^*]
\leq (C_J + 1)^{N(0)} TV_f(0^-) + N_q(T^*)[\dot{m}^*]
= \kappa_1 TV_f(0^-) + \kappa_2 ,$$

where $[\dot{m}] = \dot{m}_k^+ - \dot{m}_j^-$ for $j \in \{2, 3\}$ and $k \in \{2, 3\} \setminus \{j\}$, and $[\dot{m}^*]$ is the maximum possible value of the upper bound in equation (3.28). Moreover, $\kappa_1 = (C_J + 1)^{N(0)}$ and $\kappa_2 = N_q(T^*)[\dot{m}^*]$ are values do not depend on time. Note that, one can choose a constant quantity $\kappa$ such that $\kappa_1 TV_f(0^-) + \kappa_2 = \kappa TV_f(0^-)$. In this case $\kappa$ depends on $\kappa_i$ for $i = 1, 2$ and constant value of $TV_f(0^-)$. Now, using this new constant, one can see the above inequality statement as:

$$TV_f(t^+) \leq \kappa TV_f(0^-) .$$
Now, inspired by the result in Theorem 20 and based on the results from Remark 2 and Corollary 13, we are ready to discuss the boundedness of the total variation of the fluxes for the cases where the number of discontinuities goes to infinity. Namely, where $\nu$ approaches infinity.

To this end, we first start with an example where there exist an active absorbing queue which despite all interactions of the waves with the junction, continues stacking up the vehicles. In Theorem 21 we show that, regardless of the type of the initial data, when we have a permanent active absorbing queue, the total variation of fluxes remains uniformly bounded.

**Theorem 21.** Without loss of generality, assume that for all $t > 0$, queue 3 is active and stacking up the vehicles, i.e. $m_3 > 0$ and $\dot{m}_3 > 0$. Then, the total variation of the fluxes is uniformly bounded.

**Proof.** $m_3(t) > 0$ and $\dot{m}_3(t) > 0$ lead to the introduction of some new constraints on the states of the roads. Specifically, since

$$m_3 = \alpha_3 \Gamma_1 - \Gamma_3 = \alpha_3 \min \left( \frac{\gamma_1^d(\rho_1)}{\alpha_2}, \frac{\gamma_2^s(\rho_2)}{\alpha_3} \right) - \gamma_3^s(\rho_3) > 0,$$

and no interaction of the waves with the junction can change the dynamic of the queue 3, all waves that hit the junction should satisfy additional constraints. Therefore, there exists $\rho_i^* \in (0, \rho_{\text{imax}})$ for $i \in \{1, 2, 3\}$ such that $\rho_1 \in (\rho_1^*, \rho_{\text{imax}})$, $\rho_2 \in [0, \rho_2^*)$ and $\rho_3^* = \sigma_3 \leq \rho_3$. As discussed in the proof of Theorem 11, if a wave interacts with the junction at time $t$, we have:

$$TV_f(\bar{t}^+) \leq TV_f(\bar{t}^-) + R \times [\Gamma_j],$$

where $[\Gamma_j]$ is the magnitude of the jump of the fluxes corresponding to the interacting wave, and $R = \max\{3, \frac{\alpha_2}{\alpha_3}, \frac{\alpha_3}{\alpha_2}\}$. Moreover, Lemma 12 indicates that the waves on $I_1$ or $I_3$ that interact with the junction do not increase the total variation of the fluxes. Therefore, the only waves that can increase the total variation of the fluxes are the waves that come towards the junction from $I_2$. Now, due to the permanent activity of queue 3, all of the waves that come from $I_3$ belong to the class of sterile cases that only affect the rate of stacking
up of the vehicles. Therefore, without loss of generality, we can assume that state on $I_3$ is fixed, so $\rho_3 = \rho_{3,0}$ and, $\gamma_{3}^s(\rho_3) = \gamma_{3}^s(\rho_{3,0})$.

As mentioned above, the only waves that can increase the total variation of fluxes are the ones that hit the junction from $I_2$. Thus, statement (3.31) reduces to:

$$TV_f(t^+) - TV_f(t^-) = \frac{\alpha_3}{\alpha_2} [\Gamma_2],$$

(3.32)

where $t = \tilde{t}$ is the time of the interaction and $[\Gamma_2] = |f_2(\rho_{2,r}) - f_2(\rho_{2,l})|$ is the magnitude of the jump of the fluxes corresponding to the wave $(\rho_{2,l}, \rho_{2,r})$ that hits the junction.

Beginning with a wave-front tracking approximation $\rho_{\nu,0}$ of $\rho_0 \in L^1$ that satisfies the properties mentioned in (3.29) and based on the result of the Lemma 19 that the number of discontinuities of the approximation solution remains finite, we consider the worst case scenario in which all of the waves and discontinuities belong to $I_2$ and move towards the junction. Now, for all waves (including Big shock, small shock and rarefaction wave) that move toward the junction and for all $t > 0$ we have:

$$TV_f(t^+) = TV_f(t^-) + \frac{\alpha_3}{\alpha_2} [\Gamma_2],$$

(3.33)

where $[\Gamma_2]$ is the jump of the fluxes corresponding to the wave that hits the junction at time $t$. Therefore, we have:

$$TV_f(t^+) = TV_f(t^-) + \frac{\alpha_3}{\alpha_2} [\Gamma_2]$$

$$\leq TV_f(t^-) + \frac{\alpha_3}{\alpha_2} \max\{[\Gamma_2^*]; \hat{t} \in [0, t]\}$$

(3.34)

$$\leq TV_f(0^+) + \frac{N(0)}{\alpha_2} \frac{\alpha_3}{\alpha_2} [\Gamma_2^*],$$

where $N(0)$ is the number of discontinuities at $t = 0$, and $[\Gamma_2^*]$ is the maximum of the jump of the fluxes corresponding to the waves that hit the junction until time $t$. Note that, $N(0)$ and $[\Gamma_2^*]$ are of order $O(\nu)$ and $O(1)$, respectively. Therefore, $N(0) \times [\Gamma_2^*]$ is of order $O(1)$. Thus, as $\nu$ approaches infinity, $\frac{N(0)}{\alpha_2} \frac{\alpha_3}{\alpha_2} [\Gamma_2]$ approaches to a constant value $C^*$, that only depends on the con-
stant split ratios and we have:

$$TV_f(t) \leq TV_f(0) + C^\ast,$$  \hspace{1cm} (3.35)

therefore, the total variation of the fluxes is uniformly bounded. \hfill \square

In the earlier theorems and lemmas, we showed that the total variation of the fluxes for each single event is bounded. We also discussed that having finite number of discontinuities in the initial data guarantees a uniform bound on the total variation of the fluxes. In Lemma 21 we proved the boundedness of the total variation of fluxes for the case in which there exists a permanent absorbing queue and number of discontinuities approaches infinity. However, scenarios on the roads are not limited to these cases. There are cases where an increase in $\nu$ causes an increase in the number of depletion and formation of the queues. Theorem 22 deals with these situations and proves the uniform boundedness of the total variation of the fluxes.

**Theorem 22.** The total variation of the flux function $f$, that satisfies the criteria (3.2), is uniformly bounded.

**Proof.** Inspired by the proofs of the earlier theorems and lemmas, we are now ready to prove that, regardless of the increase in the number of discontinuities on the roads, interaction of the waves with the junction and formation and depletion of the queues, the total variation of the fluxes remains uniformly bounded.

In the proof of Theorem 20, we showed that the number of events that increase the total variation of the fluxes is bounded by an order of $O(\nu)$. Moreover, in Lemma 19, we showed that the number of discontinuities will not blow up and has an upper bound of order of $O(\nu)$. And finally, results in Remark 2 and Corollary 13, imply that the total variation of the fluxes corresponding to the queue related events and the interaction of the waves with the junction, has an upper bound of order $O(\frac{1}{\nu})$. Therefore, using inequalities
(3.22) and (3.28), we have:

\[
TV_f(t^+) \leq TV_f(t^-) + R \times [\Gamma^*_j] + R \times [\dot{m}]
\]

\[
\leq TV_f(t^-) + R \times ( [\Gamma^*] + [\dot{m}^*] )
\]

\[
\leq TV_f(0^+) + R \times \underbrace{N_e(t)}_{O(1)} \times ( [\Gamma^*] + [\dot{m}^*] ) .
\]

(3.36)

where \([\Gamma^*]\) is the maximum jump of the fluxes related to the waves that hit the junction until time \(t\), and \([\dot{m}^*]\) is the maximum possible value of the upper bound in equation (3.28). Now, as \(\nu\) approaches infinity, \(N_e(t) \times ( [\Gamma^*] + [\dot{m}^*] )\) becomes \(C^*\), a constant value of \(O(1)\). Therefore, the total variation of the fluxes is uniformly bounded and we have:

\[
TV_f(t^+) \leq TV_f(0^+) + C^* .
\]

(3.37)

\[\square\]

### 3.3.5 Existence of the Solution to the Cauchy Problem

Here we give a proof for Theorem 10.

**Proof of Theorem 10.** Consider an \(\frac{1}{\nu}\)-approximate wave-front tracking solution \((\rho_\nu, m_\nu)\) in the sense of definition 4 (where \(\epsilon = \frac{1}{\nu}\)), to problem (3.3). Based on Theorem 20, the total variation of the fluxes, is uniformly bounded and for a constant \(C^*\), we have:

\[
TV_f(t^+) \leq TV_f(0^+) + C^* .
\]

(3.38)

Similar to the discussions in [29] (Theorem 8) and [26] (Theorem 1), the condition in equation (3.38) is enough to find \(\rho = (\rho_1, \rho_2, \rho_3)\), such that \(\rho_i\) on road \(I_i\), for \(i = 1, 2, 3\), is an entropy admissible solution to problem (3.3). For \(t \in [0, T]\), \(m_\nu\) is uniformly bounded. Thus, according to the Arzelà-Ascoli theorem, there exists a subsequence \(m_{\nu_k}\), such that: \(m_{\nu_k} \to m\). Now, since \(m_{\nu_k}\) is a relatively compact subset of \(L^1([0, T^*])\), by the Eberlein-Šmulian theorem, it is relatively weakly sequentially compact. Namely, there exists a subsequence
of $m_{νk}$ and some $m'$ in $L^1([0,T^*])$ such that $m'_{νk}$ converges to $m'$ weakly. Thus, $m' = m$ in weak sense.

\textbf{Corollary 23.} For $t > T^*$, either one queue, with a constant rate of stacking up, continues absorbing the congestion or, queues stay idle. Thus, for $t > T^*$, if there exists an active queue, the value of $m$ is a linear function of time, namely, $m(t) = At + B$, where $A = \dot{m}(T^*)$ and $B = m(T^*) - T^*\dot{m}(T^*)$. And when both queues are idle, it remains zero, i.e., $m(t > T^*) = m(T^*) = 0$.

\subsection{Continuous Dependence on the Initial Data and Uniqueness}

In this subsection, we prove the continuous dependence of the solution to the initial data by using the generalized tangent vector technique [4, 3]. The generalized tangent vector technique is a well-known method to investigate the continuous dependence in the traffic modeling and the LWR-based dynamic network loading problems [25, 26, 29, 34, 33]. Specifically, this method has been used by Garavello et al. [26] to show the Lipschitz continuous dependence to the initial data for a road network model augmented with a single buffer at the junction. Here, after a brief introduction, we adapt the concept of generalized tangent vector to accommodate the FIFOQ model. In developing such a method, we need to have in mind that there are two main differences between the FIFOQ model and the models presented in [26, 34]. The first difference is that the FIFOQ model preserves the split ratios. The other difference is that, it is possible in the FIFOQ model that when a queue hits zero, the other queue starts forming. In the following, we describe how these differences are addressed.

In general, continuous dependence is not a guaranteed property. In fact, there are situations where the Lipschitz continuous dependence does not hold. In [25, 27, 12], the authors present a counterexample to the Lipschitz continuous dependence in case of a specific Riemann solver, for a junction of two incoming and two outgoing roads. A traditional way to obtain a Lipschitz
semigroup of solutions is to construct suitable approximate solutions, while controlling their distance variations over time \[69\]. Note that, as shown in \[14\] and \[53\], for a scalar conservation law, the entropic solutions constitute a contractive semigroup in $L^1$. In fact, one can directly compare two solutions $\rho$ and $\rho'$ by proving that the $L^1$ distance of them namely,

$$
||\rho(t,.) - \rho'(t,.)||_{L^1},
$$

(3.39)

is non-increasing. The aim is to show that this property holds for our model.

Consider a curve $\gamma : [0,1] \to L^1(\mathbb{R})$, that takes its values in the set of piecewise constant functions with finite number of discontinuities, where these discontinuities are located at $x_1(\theta) < x_2(\theta) < ... < x_N(\theta)$. Then, $\gamma$ admits as tangent vector $(\nu, \xi)(\theta) \in L^1(\mathbb{R}) \times \mathbb{R}^N$, where:

$$
\nu(\theta,x) = \lim_{h \to 0} \frac{\gamma(\theta+h,x) - \gamma(\theta,x)}{h}, \quad \text{for a.e. } x,
$$

$$
\xi_i(\theta) = \lim_{h \to 0} \frac{x_i(\theta+h) - x_i(\theta)}{h}, \quad i = 1,...,N.
$$

(3.40)

In fact, we have:

$$
\dot{\gamma}(\theta) = (\nu, \xi)(\theta).
$$

(3.41)

Then, the norm of the tangent vector $(\nu, \xi)(\theta)$ is defined by:

$$
||(\nu, \xi)(\theta)|| = ||\nu(\theta)||_{L^1} + \sum_{i=1}^{N} |\xi_i(\theta)||\gamma(\theta,x_i^+) - \gamma(\theta,x_i^-)|,
$$

(3.42)

which measures any infinitesimal $L^1$ displacement of the curve $\gamma$.

Now for piecewise functions $\rho$ and $\rho'$ in $L^1$, we can define the distance between them by:

$$
d(\rho, \rho') = \inf_{\Omega(\rho,\rho')} \int_0^1 ||\dot{\gamma}(\theta)|| d\theta,
$$

(3.43)

where $\Omega(\rho,\rho')$ is the set of all curves $\gamma$ that admit piecewise tangent vectors and $\gamma(0) = \rho$ and $\gamma(1) = \rho'$. Because the distance $d$ can easily coincide with $L^1$ distance, (3.39) is bounded by a $L^1$ length of the curve $\gamma(t) : \theta \to \rho^\theta(t,.)$.

Thus, if one proves that the $L^1$ length of the curve $\gamma$ for each time $t$ is bounded
by the $L^1$ length of the curve that joins $\rho(0)$ to $\rho'(0)$, the boundedness of $L^1$ distance of $\rho(t)$ to $\rho'(t)$ follows.

To elaborate this, let us study some properties and a lemma. First, we can show that the distance $d$ in (3.43), can be obtained by only using the curves that have tangent vectors with a zero $L^1$ component.

**Lemma 24.** For $\rho$ and $\rho' \in L^1$, let $\tilde{\Omega}(\rho, \rho')$ be the set of all curves $\gamma : [0, 1] \rightarrow L^1(\mathbb{R})$, such that $\gamma(0) = \rho$ and $\gamma(1) = \rho'$, and admit piecewise smooth tangent vector $(\nu, \xi)$, where $\nu \equiv 0$. Then we have:

$$\inf_{\tilde{\Omega}(\rho, \rho')} \int_0^1 ||\dot{\gamma}(\theta)||d\theta = \inf_{\Omega(\rho, \rho')} \int_0^1 ||\dot{\gamma}(\theta)||d\theta = d(\rho, \rho') = ||\rho - \rho'||_{L^1} . \quad (3.44)$$

**Proof.** Proof is provided in [27], chapter 4. \qed

In Lemma 24, because the norm of $\dot{\gamma}$ spans exactly the area contained between the graphs of $\rho$ and $\rho'$, we have $\int_0^1 ||\dot{\gamma}(\theta)||d\theta = ||\rho - \rho'||_{L^1} .

The aim here is to prove that the tangent vector is bounded. We start with two initial data $\rho(0)$ and $\rho'(0)$ and concentrate on the wave-front tracking $\rho_\nu(t)$ and $\rho'_\nu(t)$. Now, for every $\gamma_0 \in \Omega(\rho_0, \rho'_0)$ with tangent vector $(\nu, \xi)_0(\theta)$, we want to show that:

$$||(\nu, \xi)_t(\theta)|| \leq ||(\nu, \xi)_0(\theta)|| , \quad (3.45)$$

where $(\nu, \xi)_t(\theta)$ is the corresponding tangent vector of $\gamma_t$ which is the time evolution of $\gamma_0$. Considering $\Omega_t$, the set of all of the evolution curves of $\gamma_0$, we have:

$$d(\rho_\nu(t), \rho'_\nu(t)) = \inf_{\Omega(\rho_\nu(t), \rho'_\nu(t))} \int_0^1 ||\dot{\gamma}(\theta)||d\theta \leq \inf_{\Omega_t} \int_0^1 ||\dot{\gamma}(\theta)||d\theta = \inf_{\Omega_t} \int_0^1 ||(\nu, \xi)_t(\theta)||d\theta \leq \inf_{\Omega(\rho_0(0), \rho'_0(0))} \int_0^1 ||(\nu, \xi)_t(\theta)||d\theta = d(\rho_\nu(0), \rho'_\nu(0)) . \quad (3.46)$$

Now, since distance $d$ coincides with the $L^1$ metric, by establishing (3.45), we can prove the Lipschitz dependence of the solution to the initial data.
3.3.6.1 Bound on the Tangent Vectors

Earlier we discussed that the tangent vector $\xi$ for a given piecewise constant function $f(x) : [a, b] \rightarrow R$, where $a$ and $b$ can be possibly $-\infty$ and $\infty$, respectively, is defined as shifts of discontinuities of $f(.)$. Let us assume $\{x_i\}_{i=1}^N$, where $a = x_0 < x_1 < x_2 < \cdots < x_N < x_{N+1} = b$, are the positions of these discontinuities and $\{f_i\}_{i=1}^N$ represent the value of $f$ on each interval $(x_{i-1}, x_i), i = 1, \ldots, N$.

![Figure 3.5: Construction of “generalized tangent vectors” on the roads.](image)

Here, according to the result of Lemma 24, we can ignore the vertical perturbation and only consider the horizontal shifts, namely $\xi_i$. Moreover, we extend the concept of tangent vector such that the tangent vector $\xi$ represents the shift in the discontinuities of $\rho(x)$ as well as the shift in the number of vehicles stacked up in the queue. Therefore, $\xi = (\xi_1, \xi_2, \xi_3, \xi_{m_2}, \xi_{m_3})$. Although $\xi_i, i = 1, 2, 3$ and $\xi_{m_j}, j = 2, 3$, do not have the same unit, to unify the quantity corresponding to the shift, we use the same symbol. The norm of the tangent vector $\xi$ is then defined as follows:

$$\|\xi\| = \sum_i |\xi_i| |\rho_i| + \sum_j |\xi_{mj}| \quad \text{for } i \in \{1, 2, 3\}, \text{ and } j \in \{2, 3\}.$$ \hspace{1cm} (3.47)

The norm of the tangent vector for each road component is equal to the magnitude of the shift $\xi_i$ multiplied by the size of the jump of the density. This implies that the unit of the norm of the tangent vector for the road component represents the number of vehicles on the road that are increased/decreased due
to the applied shifts. Thus, in terms of norms, all components of the tangent vector represent the same quantity, namely the number of vehicles. Hence, they have the same unit.

Now, consider a junction $J$ with one incoming and two outgoing roads. Assume that the interacting wave $(\rho_{i,l}, \rho_{i,r})$ on road $I_i$ for $i \in \{1, 2, 3\}$, hits the junction and generates new waves $(\rho_{j,l}^n, \rho_{j,r}^n)$ on some roads $I_j$ for $j \in \{1, 2, 3\}$. If we apply a shift of $\xi_i$ to the wave $(\rho_{i,l}, \rho_{i,r})$, it will cause a shift to the wave $(\rho_{j,l}^n, \rho_{j,r}^n)$, for $j \in \{1, 2, 3\}$. In addition, if there exists an active queue, the shift will also be translated to a shift in the number of the vehicles stacked up at the queue. A shift may introduce a time-lag or a time-lead to the new events (i.e. generating the next waves as well as queue formation and depletion events). In the following lemma, we show that the amount of this time-lag or time-lead is preserved. This lemma also characterizes the relationship between the initial shift of the interacting wave, and the shifts produced as a result of the initial shifts. We use $\tau_i$ to represent the time-lag/lead on roads $I_i$, for $i \in \{1, 2, 3\}$. Similarly, $\tau_{mj}$ represents the time-lag/lead for events on queue $j$, for $j \in \{2, 3\}$.

**Lemma 25.** The time-lag or time-lead is preserved, in other words:

$$\tau_1 = \tau_2 = \tau_3 = \tau_{m2} = \tau_{m3},$$

where in the above statement, for $i = 1, 2, 3$ and $j = 2, 3$, we have:

$$\tau_i = \frac{\xi_i \left( \rho_{i,l}^+ - \rho_{i,r}^- \right)}{\Gamma_1^+ - \Gamma_i^-} \quad \text{and} \quad \tau_{mj} = \frac{\xi_{mj}}{|m_j|}.$$

**Remark 4.** Although in statement (3.48), we expressed $\tau$s for all roads and both queues, we meant if they are applicable. As an example, consider a sterile-case where a shifted wave on the outgoing road $I_3$ hits the junction and causes a shift to its corresponding queue. For this scenario, because $\tau_1$, $\tau_2$ and $\tau_{m2}$ are not available, the statement (3.48) will reduce to:

$$\tau_3 = \tau_{m3}.$$
We follow this convention for all scenarios and only include the $\tau$s that are actually available.

Remark 5. Events that can lead to a change in the tangent vector are: waves interacting with each other, waves hitting the junction, and queue depletion events. We look at these events “locally”, namely for a short period of time. For instance, if a wave hits the junction and causes the active queue to deplete and after the complete depletion, new waves leave the junction, we split this scenario into two separate local events: “wave hits the junction” and “waves are coming out of the junction due to the complete depletion of the queue”.

Lemma 26. By substituting the values of $\tau_i$ and $\tau_{mj}$ in (3.48), we have the following result:

$$\left| \xi_j (\rho_j^+ - \rho_j^-) \right| = q_{ij} \left| \xi_i (\rho_i^+ - \rho_i^-) \right|, \quad \text{where } q_{ij} = \left| \frac{\Gamma_j^+ - \Gamma_j^-}{\Gamma_i^+ - \Gamma_i^-} \right|,$$  \hfill (3.49)

$$\left| \xi_{mj} \right| = q_{ij} \left| \xi_i (\rho_i^+ - \rho_i^-) \right|, \quad \text{where } q_{ij} = \left| \frac{\eta_{mj}}{\Gamma_i^+ - \Gamma_i^-} \right|. \hfill (3.50)$$

The statement in (3.49) is related to the case when a shifted wave hits the junction and introduces new shifts to the waves that are generated from the junction, while statement (3.50) corresponds to the effects of a shifted wave on the available active queue at the junction.

Due to the infinitesimal magnitude of the shift, the order of the events will not change and the sequential occurrence of them will be preserved. For instance, after the interaction of a shifted wave, in terms of type, speed and state, the newly generated waves will have the same qualities as the case with no shift. Note that, the relationships characterized in the previous lemma are local. In other words, they consider the events that happen in a short time interval after the interaction of the wave with the junction.

In order to show the boundedness of the tangent vectors, one needs to prove that the multiplication factors $q_{ij}$, that appear in Lemma 26, remain uniformly bounded, regardless of the number of the interactions of the waves. To this end, we have the following theorem:
Theorem 27. The multiplication factors, \( q_{ij} \), and consequently the norm of the tangent vectors introduced at (3.47), are uniformly bounded.

Before proving this theorem, we need the following lemma for the case of two waves interacting with each other on the road:

Lemma 28. Consider two waves on the road with speeds \( s_1 \) and \( s_2 \) that hit each other and generate a new wave with speed of \( s_3 \). If the first and second wave are shifted by \( \xi_1 \) and \( \xi_2 \), respectively, the shift corresponding to the new wave, \( \xi_3 \), has the following property:

\[
\xi_3 = \lambda \xi_1 + (1 - \lambda) \xi_2 ,
\]  
\( (3.51) \)

where \( \lambda = \frac{s_3 - s_2}{s_1 - s_2} \in [0, 1] \), and:

\[
\xi_3[\rho_i] = \xi_1[\rho_1] + \xi_2[\rho_2] ,
\]  
\( (3.52) \)

where \([\rho_i]\) for \( i = 1, 2, 3 \), represents the change in the densities corresponding to the \( i \)th wave.

Proof. The proof is provided in [25].

Remark 6. Using Lemma 28 and applying the triangle inequality, one can show that, the interaction of two waves on the road decreases the norm of the tangent vector.

We are now ready to prove Theorem 27.

Proof of Theorem 27. Events that can affect the tangent vector have been introduced in Remark 5. In fact, due to the existence of the queue at the junction, in addition to the wave interactions, depletion of the queues will also influence the tangent vector. Studying the events that happen at the junction along with all their consequences can be very complicated, that’s why, as mentioned in Remark 5, we split the events and consider them locally. Therefore, we treat the scenarios that happen in the road network system as a set of these local events.
By investigating all possible local events and using the results of Theorem 11 and Lemma 26, we have the following worst-case scenario matrix. In this matrix, each of the elements represent an upper bound for the corresponding multiplication factor \( q_{ij} \). In fact, quantity \( Q_{ij} \) is either equal or greater than the multiplication factor \( q_{ij} \). The inequality cases happen when after depletion of queue \( j \), congestion on \( I_j \) for \( j \in \{2, 3\} \) is resolved (the quantity \( \mu = 0 \)) and queue \( k \), with \( k \in \{2, 3\} \setminus \{j\} \), starts forming.

The worst-case scenario matrix is as follows:

\[
\begin{bmatrix}
1 & 2 & 3 & m_2 & m_3 \\
1 & \alpha_2 & \alpha_3 & \alpha_2 & \alpha_3 \\
2 & \frac{1}{\alpha_2} & 1 & \frac{\alpha_3}{\alpha_2} & 1 & \frac{\alpha_3}{\alpha_2} \\
3 & \frac{1}{\alpha_3} & \frac{\alpha_2}{\alpha_3} & 1 & \frac{\alpha_2}{\alpha_3} & 1 \\
m_2 & \frac{1}{\alpha_2} & 1 & \frac{\alpha_3}{\alpha_2} & 1 & \frac{\alpha_3}{\alpha_2} \\
m_3 & \frac{1}{\alpha_3} & \frac{\alpha_2}{\alpha_3} & 1 & \frac{\alpha_2}{\alpha_3} & 1
\end{bmatrix}
\]

(3.53)

Given constant \( \alpha_2 \) and \( \alpha_3 \), the elements of matrix \([Q_{ij}]\) are known and bounded and there exists a constant \( K \) such that \(|Q_{ij}| \leq K\), for \( i, j \in \{1, 2, 3, m_2, m_3\} \). Moreover, for each time \( t \) we have:

\[
||((\nu, \xi)_t)|| \leq K||((\nu, \xi)_{t-})||.
\]

(3.54)

Moreover, elements of the matrix \([Q_{ij}]\) have the following property:

\[
Q_{ij} \cdot Q_{jk} = Q_{ik} \quad \forall \ i, j, k = 1, 2, 3,
\]

(3.55)

Thus along with the result in Remark 6, one can prove that, no matter how many interactions occur on the roads and at the junction, the norm of the tangent vectors are always bounded. As studied in detail in the previous section, the mechanism that can change the direction of the wave that leaves the junction is through its interaction with the waves on the roads. However, this interaction decreases the norm of the tangent vector. Now, if the new generated wave moves toward the junction, it will then follows (3.55). Thus, the norm of the tangent vector is uniformly bounded.
CHAPTER 4

MODEL VALIDATION WITH DATA

In this chapter, we investigate the performance of our proposed model with data. We use artificially generated data from a microsimulator called SUMO to validate our model. SUMO is a free, open, microscopic and space-continuous road traffic simulation suite designed to handle large road networks. We use the data generated by SUMO as a proxy for real traffic. For all of the experiments in this chapter, we use the car-following model developed by Krauss [52].

We first describe the fundamentals of microscopic models, car-following models and lane-changing models. Then we explain the specific car-following and lane-changing models used in this chapter and the parameters of these models. Finally, we conduct validation and comparison experiments for our macroscopic models.

4.1 Microscopic Models

The main idea behind microscopic models of traffic flow is to describe the movements of each vehicle based on the positions and velocities of its neighboring vehicles. In general, two dynamical processes of car-following and lane-changing need to be described. Because the car-following models tend
to be more sophisticated than the lane-changing models [52], we first describe the car-following models on a single lane. Afterwards, we will describe the lane-changing models. We use $x_i$ and $v_i$ to denote the positions and velocities of vehicles on a single lane road. We assume that the index $i$ increases as we move downstream.

To develop a car-following model, we can assume that a change of velocity is performed only under the condition that the current velocity is different from some desired velocity, $V_{\text{des}}$. The desired velocity $V_{\text{des}}$ is determined based on factors such as safety and legal considerations. The simplest model of how a driver tries to proceed toward the desired velocity is as follows:

$$\frac{dv_i(t)}{dt} = \frac{v_{\text{des}} - v_i(t)}{\tau}. \quad (4.1)$$

This equation is interpreted as a stimulus–response model and $\frac{1}{\tau}$ is referred to as the sensitivity. Almost all of the car-following models can be described using the simple model in equation (4.1) through an appropriate choice of $v_{\text{des}}$ and $\tau$.

### 4.1.1 Car-Following Models

The underlying assumption of most of the car-following models is that the movement of a vehicle $i$ is completely described by the movement of the vehicle it is following, i.e. vehicle $i+1$. Because the steady state velocity of all vehicles need to be equal to each other, the desired velocity of vehicle $i$ can be assumed to be equal to the current velocity of its preceding vehicle. Thus:

$$\frac{dv_i(t)}{dt} = \frac{v_{i+1}(t) - v_i(t)}{\tau}. \quad (4.2)$$

In this study, we use the car-following model proposed by Krauss [52]. In this model, two types of vehicle motions are considered. The first type is the free motion of the vehicle, which is limited by some maximum velocity $v_{\text{max}}$:

$$v \leq v_{\text{max}}.$$
In this setting, the maximum velocity can be interpreted as the desired velocity of the driver.

The second type of vehicle motion is due to the interaction with other vehicles. The main objective in this type of motion is to avoid collision with other vehicles. Therefore, this model assumes that the driver always chooses a velocity that is limited by a safety velocity $v_{\text{safe}}$:

$$v \leq v_{\text{safe}}.$$ 

The model also, reasonably, assumes that the acceleration of the vehicle is bounded:

$$-b \leq \frac{dv}{dt} \leq a .$$

For some positive $a$ and $b$. The model can be formulated in discrete time steps as follows:

$$v(t + \Delta t) \leq \min(v_{\text{max}}, v(t) + a\Delta t, v_{\text{safe}}),$$

where $v_{\text{safe}}$ is computed such that $v_{\text{safe}} > v(t) - b\Delta t$. To characterize the safety velocity $v_{\text{safe}}$, we consider a pair of vehicles: a leader in position $x_l$ and velocity of $v_l$, and a follower in position $x_f$ and velocity of $v_f$. Given the length of vehicles $l$, the gap between the two cars can be computed as:

$$g = x_l - x_f - l .$$

In order to avoid collisions, the gap $g$ needs to be non-negative. Given some non-negative desired gap $g_{\text{des}}$, the gap $g$ always remains non-negative if it is greater than $g_{\text{des}}$ once and satisfies the dynamical inequality:

$$\frac{dg}{dt} \geq \frac{g_{\text{des}} - g}{\tau_{\text{des}}} .$$ (4.3)

To verify that collision cannot occur under this condition, note that if $g = 0$, then the time derivative of $g$ is always non-negative if $g_{\text{des}}$ is non-negative.

The condition in equation (4.3) can be derived more intuitively. Consider a pair of vehicles: a leader with velocity of $v_l$, and a follower with velocity of $v_f$ and a gap $g$ from the leader. The velocity of the follower is considered safe.
if, under any conditions, the driver can bring the vehicle to a complete stop to avoid colliding with the leader. In other words, given the driver reaction time of $\tau$, and the breaking distance of $d(v)$, we should have:

$$d(v_f) + v_f\tau \leq d(v_l) + g . \quad (4.4)$$

To avoid inverting the function $d(v)$, Taylor series expansion of $d(v)$ around the point $\bar{v} = \frac{v_f + v_l}{2}$ is used. Therefore, we have:

$$d'(\bar{v})v_f + v_f\tau \leq d'(\bar{v})v_l + g . \quad (4.5)$$

Considering a car decelerating from velocity $v$ to 0 with some rate $\dot{v} = -b(v)$, we have:

$$d'(v) = -\frac{d}{dv} \int_{v}^{0} \frac{v'}{b(v')}dv' = \frac{v - v_l}{b(v)} . \quad (4.6)$$

Therefore, we can write the safety condition as:

$$v_l - v_f \geq \frac{v_l\tau - g}{b(v) + \tau} . \quad (4.7)$$

These conditions lead to the following set of updating equations that completely characterize the car-following model:

$$v_{\text{safe}}(t) = v_l(t) + \frac{g(t) - g_{\text{des}}(t)}{\tau_b + \tau} , \quad (4.8)$$

$$v_{\text{des}}(t) = \min(v_{\text{max}}, v(t) + a\Delta t, v_{\text{safe}}(t)) , \quad (4.9)$$

$$v(t + \Delta t) = \max(0, v_{\text{des}}(t) - \eta) , \quad (4.10)$$

$$x(t + \Delta t) = x(t) + v\Delta t . \quad (4.11)$$

In these equations, the desired gap is chosen as $g_{\text{des}}(t) = \tau v_l(t)$, and $\tau$ is the reaction time of the driver. The time scale $\tau_b$ is set to be $\tau_b = \frac{\bar{v}}{h}$. The random perturbation $\eta$ represents driver imperfection.
4.1.2 Lane-Changing Models

A complete description of a microsimulator needs to include both a car-following model as well as a lane-changing model. Lane-changing is a fundamental driving behavior that occurs when a driver moves from one lane to a different lane, with the objective of improving driving conditions or leaving a road. Because lane-changing involves both longitudinal and lateral movements, it has major impacts on both traffic safety [77, 61, 65] as well as traffic flow patterns [50, 62, 10, 54].

In general, the analytical lane-changing models can be categorized into two classes. The first class of lane-changing models are the rule-based models. These models aim at designing a set of rules to describe the lane-changing behaviour of the drivers. Gipps (1986) [30] introduced one of the classical lane-changing models with a two-level framework that includes a series of rules. However, his proposed model was limited and did not accommodate many important issues such as driver heterogeneity and drivers’ dynamic behavior. To this end, many researchers developed new lane-switching models to address these issues. For example Hidas (2002,2005) [38, 39] developed lane-changing models by categorizing the lane changes into three classes of free, cooperative, and forced lane changes. Kesting et al. (2007) [51] proposed a model called Minimizing Overall Braking Induced by Lane change (MOBIL), that integrates the advantages and disadvantages of potential lane-changing with into an acceleration function in the context of a car-following model. Laval and Leclercq (2008) [55] develop a model that uses lane-specific macroscopic quantities inside a microscopic lane-changing model to describe various traffic situations.

The other class of lane-changing models are the discrete choice-based models. These models try to develop driver utility functions to describe their lane-changing behavior. One of the first models in this area was developed by Ahmed et al. (1996) [1]. They model heterogeneous drivers by considering utility functions for potential lane-changing activities. Following this paper, many
researchers proposed similar lane-changing models that accommodate various factors related to the vehicle itself or the surrounding vehicles. For example, Toledo et al. (2003) [76] developed a comprehensive utility function to consider both mandatory and discretionary lane-changes. Their results show the importance of considering the trade-off between mandatory and discretionary lane-changes. Sun and Elefteriadou (2011,2012) [75, 74] studied the effect of various driver characteristics (such as aggressiveness, gender, and alertness) on their lane-changing behavior. They extracted the most important factors to develop utility functions to describe lane-changing behavior of the drivers.

In this study, we use the default lane-changing model of SUMO, developed by Jakob Erdmann [23]. This model is a rule-based model and has two main objectives. First, in each simulation step, it determines the lane-changing decision of vehicles based on the vehicle route as well as traffic conditions. The model also determines the velocity changes for the vehicle itself and the obstructing vehicles to support the execution of the desired lane-change.

A vehicle under consideration (called the ego vehicle) can only change lanes if there is enough space on the target lane, and it is not too close to its leading and following vehicles on the target lane. A vehicle is said to have a blocking leader (follower) if it is too close to the leading (following) vehicle on the target lane.

Based on the motivation behind the lane-change, this model considers four different types of lane-changes:

- Strategic lane-change
- Cooperative lane-change
- Tactical lane-change
- Regulatory lane-change

In each simulation step, for every vehicle, the following substeps are undertaken:
1. The preferred successor lanes (called bestLanes) are computed.

2. Safe velocities are computed under the assumption that the vehicle remains on the current lane, and considering lane-changing related speed requests in the previous step.

3. The lane-changing request is computed.

4. Either undertake the lane-changing maneuver, or compute the speed requests for the next simulation step.

In the following, the four type of lane-changing motivations are discussed in the order of their priority beginning with the most important.

A **strategic lane-change** is undertaken when a vehicle needs to change lanes to maintain its route. This type of lane-change happens if the vehicle’s current lane is not connected to the next edge on the vehicle’s route. In this situation, the vehicle is said to be on a dead lane. For every lane on the current edge, SUMO keeps track of the sequence of lanes that can be followed without lane-changing until a dead-end is reached or a maximum distance is achieved. These lanes are labelled as “bestLanes”. For every lane on the current edge, SUMO also computes the occupation, which is the traffic density along the bestLanes, as well as the bestLaneOffset, which is the index difference from the strategically advisable lanes. In each step, SUMO also determines if a strategic lane-changing maneuver is urgent. This happens when a vehicle approaches a dead-end lane and it is done by checking if the following inequality is satisfied:

\[ d - o < \text{LookAheadSpeed} \times |\text{BestLaneOffset}| \times f. \]  \hspace{1cm} (4.12)

In this equation, \( d \) is the distance to the end of the dead-end lane, \( o \) is the occupation, and \( \text{LookAheadSpeed} \) is the presumed speed while approaching the end of the dead-end lane. Moreover, \( f \) is a coefficient based on the time needed to execute a lane-changing maneuver. It is equal to 10 for a lane-change
to the left, and 20 for a lane-change to the right. If a lane-change is deemed to be urgent it will be executed with the highest priority.

If a desired strategic lane-change is not possible due to a blocking leader or follower, speed adjustment requests are placed to make the lane change possible in the following steps. To determine the speed adjustments, based on the speeds of vehicles involved, gaps and remaining time to reach the end of a dead lane, a hierarchy of cases are considered. For each case, speed adjustment requests are made accordingly to facilitate the lane change in the following steps. These speed adjustments can be for the ego vehicle itself or its surrounding vehicles.

The speed adjustments computed here are integrated with the maximum safe speed computations for the car-following model. In other words, the effective safe speed is set to be equal to the minimum of the safe speed computed by the car-following model and all of the speed adjustment requests.

**Cooperative lane-changing** refers to the maneuvers that are executed only for the purpose of facilitating another vehicle’s lane-changing toward its lane. In this lane-changing model, the vehicles are informed if they are blocking another vehicle from their intended lane-change. In this case, if there are no strategic motivations against it, the ego vehicle may change lanes to create space for the other vehicle. If the ego vehicle cannot successfully execute a cooperative lane change, it adjusts its own speed to increase the probability of success in the next simulation steps. However, it does not request other vehicles to adjust their speed.

**Tactical lane changing** happens when a vehicle tries to avoid following a slow leader. This type of lane-change involves a trade-off between the speed gain from the lane-change and the amount of effort it takes to execute the lane-changing maneuver. To achieve this, for each vehicle, the model considers a signed variable named speedGainProbability. The sign of this variable indicates the direction of the desired lane change (-1 for right, 1 for left), and its magnitude represents the amount of expected benefit from the tactical lane change. If the magnitude of speedGainProbability exceeds a threshold, then
the tactical lane change is attempted. In case of a successful tactical lane change, speedGainProbability is reset to zero. In each simulation step, for each lane changing direction \(d\), the potential speed gain \(g\) is computed. If this value is positive, speedGainProbability is incremented by \(d \times g\). If \(g\) is not positive and speedGainProbability has the same sign as \(d\), then it is halved.

**Regulatory lane-changing** is a type of lane-changing maneuver that is executed due to the obligation of the drivers to clear overtaking lanes if they are not using them to overtake other vehicles. In jurisdictions with right-handed driving, the left lane(s) are designated as overtaking lanes. In the lane-changing model considered in this study, for each vehicle, a variable called keepRightProbability is maintained to model regulatory lane-changing behavior. This variable is decremented over time and, if it exceeds a lower threshold, a regulator lane-change maneuver is attempted.

The four lane-changing motivations described above are combined in the following hierarchical decision making framework:

1. Urgent strategic lane change in direction \(d\) is needed: change (strategic)
2. A lane change in direction \(d\) would create an urgent condition: stay (strategic)
3. The ego vehicle is blocking another vehicle with urgent strategic lane change: change (cooperative)
4. speedGainProbability above threshold and its sign matches \(d\): change (tactical)
5. keepRightProbability above threshold and \(d = -1\): change (regulatory)
6. non-urgent strategic change in direction \(d\) needed: change (strategic)

In this framework, the first statement that applies to the vehicle is requested. In each simulation step, each vehicles considers changing lanes to the right first and if no change to the right is executed, then a lane change to the left is considered.
4.2 Parameter Selection

For the car-following model, we consider a car length of 5 meters with a gap of 2 meters. Maximum velocity is considered to be $27.75 \frac{m}{s}$. Maximum and minimum acceleration limits are considered to be $20 \frac{m}{s^2}$ and $-20 \frac{m}{s^2}$, respectively. Regarding the lane-changing model, we disabled the regulatory lane-changing feature of the model. This is done because in our macroscopic models, we do not assume any regulations regarding the overtaking lanes.

4.3 Extraction of Macroscopic Quantities from Microsimulation Results

In this section, we discuss the extraction of macroscopic quantities, vehicle density and queue length, from the microsimulator results. The extraction of vehicle density is done via a kernel density estimation approach with a customized treatment of the boundary points to alleviate the underestimation of vehicle densities due to lack of data in these areas. The extraction of the queue is done using a rule-based approach that mainly uses the speed of each vehicle as well as its distance from the leading vehicle to determine if it belongs to the queue. In the following subsections, we explain each of these approaches in more detail. Note that, when using the kernel density estimation method to estimate the vehicle densities on $I_1$, we do not include the queued vehicles in our calculations.

4.3.1 Kernel Density Estimation

To be able to use the results of the microsimulator to validate and compare the macroscopic models, we need to convert the data from the microsimulator into a form that is compatible with the results of these models. Given trajectory data $x_i(t)$ for each vehicle $i$ from the microsimulator, we denote the position of vehicles at time $t$ with \{${x_1(t), x_2(t), \ldots, x_N(t)}$\}. The aim is to
estimate a smooth density function $\rho(x, t)$ based on these trajectories. In this section, we assume an unbounded road and use a non-parametric statistical method called kernel density estimation (KDE) to estimate a density function from the trajectory data. In the next section, we will discuss how to modify this method and apply it to the more realistic case of bounded road segments.

The KDE method, also known as the Parzan-Rosenblatt window method [72, 68], treats the data points as distributions and uses the summation of these individual distributions as the estimate. Given a set of points $x_1, x_2, \ldots, x_N$ sampled from an unknown density function $\rho(x)$, KDE method computes the comb function

$$c(x) = \sum_{j=1}^{N} \delta(x - x_j) ,$$

using this function, the estimate of the density function on the whole road is given as:

$$\hat{\rho}(x) = \int_{\mathbb{R}} K(x - y)c(y)dy = \sum_{j=1}^{N} K(x - x_j) , \quad (4.13)$$

where $K$ is a smoothing kernel function. In this study, we use a Gaussian kernel function

$$K(x) = \frac{1}{\sqrt{2\pi h}} e^{-\frac{x^2}{2h^2}}, \quad x \in \mathbb{R} , \quad (4.14)$$

In this equation, the bandwidth parameter $h$ is a smoothing parameter. This parameter needs to be chosen such that a trade-off is reached between accuracy and smoothness. In other words, the bandwidth parameter needs to be small enough so that the estimator captures the patterns in the data, and large enough to give a smooth estimator [67, 66, 9, 45]. In this study we choose $h$ the smallest value so that equidistant vehicles generate an almost constant $\hat{\rho}$ for $\rho > \frac{1}{5}\rho^{\text{max}}$. For the microsimulator data in this study, this method leads to $h = 34.3$ meters. For more details on how to select an appropriate value for the bandwidth parameter $h$, see [9, 45].
4.3.1.1 Boundary Correction

The estimate in equation (4.13) works well if there are sufficient number of data points on both sides of point $x$. Therefore, due to lack of data points beyond the boundaries of the road segments, equation (4.13) underestimates $\hat{\rho}(x)$ for the points near the boundaries.

There are many studies that attempt to address the boundary effect [44, 46, 47, 48]. These studies can be classified into two general approaches. The first approach is to convert the region into an unbounded region using a transformation. The other approach is to use reflection to generate ghost data points. In this study, we use the reflection method described in [49] and [24]. We assume, without loss of generality, that the data points are sorted in an increasing order, i.e. $x_1 \leq x_2 \leq \cdots \leq x_N$. We first describe this method for the left boundary. We use $d$ to denote the average distance of the interior points near $x_1$. Then we add ghost points $x^*_1, x^*_2, \ldots, x^*_k$ to the left of $x_1$ using reflection at $a^* = x_1 - \frac{d}{2}$. In other words,

$$x^*_j = 2a^* - x_j.$$  

Similarly, ghost points are added to the right of the right boundary. Then, we use our estimate in equation (4.13) using this augmented data set to get a corrected estimation of the density function.

Note that, on our network with one incoming road $I_1$ and two outgoing roads, $I_2$ and $I_3$, the boundary correction method described in this section is only applied to the left boundary of $I_1$ and right boundaries of $I_2$ and $I_3$. The reason for this is that, for example, for the right boundary of $I_1$, we can use the data from the out-road $I_2$. Similarly for the left boundaries of $I_2$ and $I_3$, the data from $I_1$ can be used.

4.3.2 Extraction of the Queue Length

To determine whether or not a vehicle belongs to the queue, we use a rule-based approach that mainly uses its speed and distance from the leading
vehicle. Specifically, if both of the following conditions are satisfied, then the vehicle, regardless of the lane it occupies, is considered to be in the queue:

- The speed of the vehicle is less than a threshold speed.
- The distance of the vehicle from its leader is smaller than a threshold distance.

We also assume that once a vehicle enters the queue, it is considered to be in the queue until it enters the off-ramp. Based on our speed and distance analysis experiments, the appropriate threshold values for speed and distance are chosen as $5\text{m/s}$ and $15\text{m}$.

In the FIFOQ model, the vertical queue is only for modeling purpose and, regardless of its length, does not have any physical presence on the network. Therefore, the queue does not occupy any portion of $I_1$. However, this is not in agreement with the microsimulator, in which the queue does occupy a portion of $I_1$ as it grows. In order to compensate for this discrepancy, when extracting the queue length from the microsimulator results, we subtract the number of vehicles that would have been present in the location of the queue, if there was no queue on $I_1$. To estimate this quantity, we use the average vehicle density on the lanes that are adjacent to the rightmost lane (the lane that is affected by the queue) on $I_1$. Figure 4.1 shows an example of how to calculate this quantity. In this figure, the six red cars are the ones that satisfy the threshold conditions and therefore considered to be in the queue. Therefore, the queue length before correction is 6 veh. In order to estimate the amount that needs to be subtracted from the queue length, we calculate the average (per lane) number of vehicles on the lanes that are adjacent to the rightmost lane. To achieve this, we consider the eight blue cars in the region between the two vertical dashed lines. For this region, the average number of vehicles per lane is $\frac{8}{3}$ veh. Therefore, the corrected queue length is $6 - \frac{8}{3}$ veh.
4.4 Experiments for the Off-ramp Model

In this section, we design and conduct experiments to investigate the performance of our proposed FIFOQ model. We compare the performance of the FIFOQ model both with the results of SUMO simulation as well as the existing macroscopic models, i.e. FIFO and non-FIFO models.

4.4.1 Derivation of the Fundamental Diagram

To obtain the fundamental diagram, we consider a ring road with 10 edges. This network is shown in Figure 4.2. In this network, each edge is 1000 meters long with a speed limit of \(27.75 \frac{m}{s}\) meters.
To obtain a fundamental diagram, we observe the car density, velocity and flux values for different number of cars introduced to the system. Specifically, we enter a given number of cars to the system. We then let the system run for 2000 seconds and, for each edge, report the car density, velocity and flux values. We repeat this process with different number of cars to obtain the complete fundamental diagram. Specifically, we start by 0 cars in the system and iteratively increase the number of cars by 5. Figures 4.3 and 4.4 show the obtained results for a ring road consisting of one-lane roads. As observed from Figure 4.3 the obtained fundamental diagram is of triangular shape.

Figure 4.4 shows the relationship between speed and density. As seen in this figure, for densities less than or equal to the critical density, the speed remains fixed and equal to the maximum speed. For densities beyond the critical density, speed starts decreasing until it becomes zero when density is around $0.14\frac{veh}{m}$.

Figure 4.3: Fundamental diagram for a one-lane road network.
Next, we use the linear regression method to fit a model to the data. Based on Figure 4.3, we assume a triangular shape for the relationship between density and flux:

\[
f(\rho) = \begin{cases} 
  a\rho, & \text{for } \rho \leq \sigma, \\
  b - c\rho, & \text{for } \rho > \sigma.
\end{cases}
\]

In this formula, \( \sigma = \frac{b}{a+c} \) is the critical density that leads to the maximum flux value. To find the parameters \(a, b\) and \(c\), we employ a least square fitting approach using the data obtained from the SUMO micro-simulator. We denote this triangular FD with \( f_{a,b,c}(\rho) \), and given the microsimulator data \((\rho_{\text{data}}, f_{\text{data}}) = \{(\rho_j, f_j), j = 1, 2, \ldots, n\}\), we first estimate \(\sigma\) with \(\rho_j\) with \(j = \arg\max_j f_j\). In other words, \(\sigma\) is chosen to be equal to \(\rho_j\), where \(j\) is the index that leads to maximum value of \(f_j\). To estimate the remaining parameters, we solve the following optimization problem:

\[
\min_{a,b,c} \sum_{j=1}^{n} |f_{a,b,c}(\rho_j) - f_j|^2, 
\]

Subject to the constraint \( \frac{b}{a+c} = \sigma \). For a given value of \(\sigma\), we can show that this optimization problem, after substituting \(b = \sigma(a + c)\), is a convex
optimization problem and can be solved using the first order conditions to obtain $a$ and $c$. Then, an estimate of $b$ can be found by using the equation $b = \sigma(a + c)$. For the network of one-lane edges, we obtain $\sigma = 0.029 \frac{veh}{m}$, $a = 27.749 \frac{m}{s}$, $b = 1.005 \frac{veh}{s}$, and $c = 6.992 \frac{m}{s}$. These results are in line with the assumption about maximum speed $v^{max} = 27.75 \frac{m}{s}$. This is because in the free-flow region of FD, all of the vehicles can achieve the maximum speed and we have $f(\rho) = v^{max} \rho \approx 27.749 \rho$. Moreover, the density level which leads to zero flux, i.e. the jam traffic density, is given by

$$\rho^{max} = \frac{\text{number of lanes}}{\text{car length + safety gap}} = \frac{\text{number of lanes}}{7m}. \quad (4.16)$$

This is in line with the result of SUMO for the one-lane road, because $\rho^{max} = \frac{1}{7} \approx \frac{b}{c}$.

Figures 4.5 and 4.6 show the results for a network consisting of three-lane roads. Following the same approach as the network with one-lane roads, we fit a triangular FD model to this data using simple linear regression models. For this network, we obtain $\sigma = 0.087 \frac{veh}{m}$, $a = 27.750 \frac{m}{s}$, $b = 3.016 \frac{veh}{s}$, and $c = 7.001 \frac{m}{s}$. These results are in line with the assumption about maximum speed $v^{max} = 27.75 \frac{m}{s}$. This is because in the free-flow region of FD, all of the vehicles can achieve the maximum speed and we have $f(\rho) = v^{max} \rho = 27.750 \rho$. Moreover, the jam traffic density, is also in line with our existing assumptions as we have $\rho^{max} = \frac{3}{7} \approx \frac{b}{c}$. 
Figures 4.5 and 4.6 show the results for a network consisting of three-lane roads. Following the same approach as the networks with one-lane and three-lane roads, we fit a triangular FD model to this data using simple linear regression models. For this network, we obtain $\sigma = 0.1157 \frac{veh}{m}$, $a = 27.750 \frac{m}{s}$, $b = 4.018 \frac{veh}{s}$, and $c = 6.987 \frac{m}{s}$. These results are in line with the assumption about maximum speed $v_{max} = 27.75 \frac{m}{s}$. This is because in the free-flow region of FD, all of the vehicles can achieve the maximum speed and we have $f(\rho) =$
\( v_{\text{max}} \rho = 27.750 \rho \). Moreover, the jam traffic density, is also in line with our existing assumptions as we have \( \rho^{\text{max}} = \frac{4}{T} \approx \frac{b}{c} \).

Figure 4.7: Fundamental diagram for a four-lane road network.

Figure 4.8: Speed-density relationship diagram for a four-lane road network.
4.4.2 A Clogged Off-ramp Scenario for a Network with an Exit-only Off-ramp

We consider a network with one incoming road $I_1$, an out-road $I_2$, and an off-ramp $I_3$. We assume that the incoming road has four lanes, while the out-road and off-ramp have three and one lanes, respectively. Moreover, $I_1$ is 1500 meters, $I_2$ and $I_3$ are both 500 meters. Figure 4.9 shows a screen shot of part of this network in the SUMO environment. For this network we have the road sharing ratios as $c_2 = \frac{3}{4}$, and $c_3 = \frac{1}{4}$. We also assume that the split ratios are equal to the road sharing ratios, i.e. $\alpha_2 = \frac{3}{4}$, and $\alpha_3 = \frac{1}{4}$.

We consider a scenario in which the off-ramp is initially completely clogged and gets unclogged after 20 seconds. Initial densities of $I_1$ and $I_2$ are assumed to be 0.08 veh/m and 0.055 veh/m, respectively. New vehicles are generated at the left boundary of $I_1$ with a rate of 2.22 veh/s. This will create and maintain a car density of $0.08 = \frac{2.22}{27.75} \text{ veh/m}$ on road $I_1$. Note that, for the free-flow regime, we have $f(\rho) = v_{\text{max}} \rho$. Figures 4.10 to 4.19 show the vehicle positions in the microsimulator as well as macroscopic quantities of FIFO, non-FIFO, and FIFOQ models for various points in time. In these figures, the top left and top right plots show the vehicle positions on $I_1$ and $I_2$, respectively. The plot on the second row shows the vehicle positions on $I_3$. The left and right plots on
the third row show the computed vehicle densities on \( I_1 \) and \( I_2 \), respectively. The bottom left plot shows the queue sizes, \( m_3 \), for SUMO and FIFOQ models. Note that, FIFO and non-FIFO models do not consider queues. Finally, the bottom right plot shows the vehicle density on \( I_3 \).

Figures 4.10 and 4.11 show the results at times \( t = 10s \) and \( t = 15s \). As seen in these figures, in the beginning of the simulation, the off-ramp is completely clogged and vehicles are entering the system through \( I_1 \) and moving towards the junction. The densities obtained by FIFOQ and non-FIFO are identical and capture the densities on the roads reasonably well. However, the FIFO model shows shock waves on \( I_1 \) and \( I_2 \) moving away from the junction. Moreover, because the junction is completely clogged, the FIFO model predicts zero flux through the junction. In addition, because the off-ramp is completely clogged, some of the vehicles start forming a queue at the junction. The length of this queue is captured reasonably well by the FIFOQ model.

Figure 4.12 shows the results at time \( t = 25s \), i.e. 5 seconds after the execution of the manoeuvre to unclog the off-ramp. As a result of the manoeuvre, a rarefaction wave on \( I_3 \) moves towards the junction. This wave is captured by all of the macroscopic models.

Figure 4.13 shows the results at time \( t = 80s \). According to this figure, at around \( t = 80s \), the rarefaction wave that was generated by the the execution of the manoeuvre to unclog the off-ramp, gets close to the junction and is about to hit the junction. After the rarefaction wave hits the junction, it becomes unclogged. Therefore, the FIFO model predicts nonzero flux through the junction. Figure 4.14, which shows the state of the system at time \( t = 100s \), shows this effect. As seen in this figure, the FIFO model shows a rarefaction wave on \( I_1 \) moving away from the junction.

Figure 4.15 shows results at time \( t = 110s \). According to this figure, the non-FIFO model shows a shock wave on \( I_3 \) moving away from the junction. The other two models do not show this effect at this time. As we will see later, for the FIFOQ and FIFO models, this effect happens at a later time.

Figure 4.16 shows the results at time \( t = 250s \). Based on this figure, the
FIFO model shows that the rarefaction wave catches up to the shock wave on $I_1$. And we can see that at time $t = 260s$, as shown in Figure 4.17, a shock wave is generated that moves toward the junction. At time $t = 290s$, as shown in Figure 4.18 this wave is about to hit the junction. Figure 4.19 shows the results after the wave hits the junction. As seen in this figure, at this time, both FIFOQ and FIFO models show a wave similar to the one non-FIFO model showed in Figure 4.15. After this point, all three macroscopic models show the same density estimates and match the results from SUMO reasonably well.

Figure 4.10: Exit-only off-ramp network results for $t = 10s$. 
Figure 4.11: Exit-only off-ramp network results for $t = 15s$.

Figure 4.12: Exit-only off-ramp network results for $t = 25s$. 
Figure 4.13: Exit-only off-ramp network results for $t = 80s$.

Figure 4.14: Exit-only off-ramp network results for $t = 100s$. 
Figure 4.15: Exit-only off-ramp network results for $t = 110s$.

Figure 4.16: Exit-only off-ramp network results for $t = 250s$. 
Figure 4.17: Exit-only off-ramp network results for $t = 260$ s.

Figure 4.18: Exit-only off-ramp network results for $t = 290$ s.
Next, we compare the models in term of the number of vehicles that move through the junction over time. Figure 4.20 shows the cumulative number of type 2 vehicles (i.e. vehicles intending to use $I_2$) that have passed through the junction over time. According to this figure, FIFOQ and non-FIFO models closely match the results of SUMO. However, the FIFO model shows different numbers for times up to around 90 seconds. This is the time interval during which the offramp is completely clogged and the FIFO model prescribes zero flux through the junction. Therefore, the FIFO model cannot accurately capture the number of type 2 vehicles passing through the junction.

Figure 4.19: Exit-only off-ramp network results for $t = 300s$. 
Figure 4.20: Total type 2 vehicle tallies for the exit-only off-ramp scenario.

Figure 4.21 shows the cumulative number of type 3 vehicles (i.e. vehicles intending to use $I_3$) that have passed through the junction over time. According to this figure, FIFOQ and FIFO models closely match the results of SUMO. However, the non-FIFO model shows different numbers for times beyond around 110 seconds. Therefore, the non-FIFO model cannot accurately capture the number of type 2 vehicles passing through the junction.
Based on Figures 4.20 and 4.21, the FIFOQ model is the only model that accurately captures the total number of vehicles passing through the junction for the type 2 and type 3 vehicles. This clearly shows the failure of the traditional models and the success of the FIFOQ model.

4.4.3 A Clogged Off-ramp Scenario for a Network with a Non-exit-only Off-ramp

In this section, a network with a non-exit-only Off-ramp is considered. Figure 4.22 shows the screen shot of this network in SUMO environment. As seen in this figure, roads $I_1$ and $I_2$ both have four lanes, and $I_3$ has one lane. The rightmost lane of $I_1$ can be used by both type 2 and type 3 vehicles. Therefore, the road-sharing ratios are $c_2 = 1$, and $c_3 = \frac{4}{5}$. We also assume that the split ratios are $\alpha_2 = \frac{3}{4}$, and $\alpha_3 = \frac{1}{4}$. Note that, in this scenario, because the split ratios are not equal to the road sharing ratios, we use the extended version of the FIFOQ model described in section 2.6.
Figure 4.22: A network with a non-exit-only off-ramp.

For this network, we run the the same simulation scenario used in the network with an exit-only off-ramp, as described in section 4.4.2. and Figures 4.23 to 4.32 show the obtained results for this network. The results for the network with a non-exit-only off-ramp are mostly similar to the network with an exit-only off-ramp. Figures 4.23 and 4.24 show the results at times $t = 10$ and $t = 15s$. Based on these figures, in the beginning of the simulation, the off-ramp is completely clogged and the densities obtained by FIFOQ and non-FIFO are identical and close to the densities from the SUMO simulation. However, the FIFO model shows shock waves on $I_1$ and $I_2$ moving away from the junction. Moreover, because the junction is completely clogged, the FIFO model shows zero flux through the junction. In addition, because the off-ramp is completely clogged, some of the vehicles start forming a queue at the junction. The length of this queue is captured reasonably well by the FIFOQ model.

Figure 4.25 shows that, after the execution of the manoeuvre, a rarefaction wave on $I_3$ moves towards the junction, which is captured by all of the macroscopic models.

Figure 4.26 shows the results at time $t = 80s$. As seen in this figure, at around $t = 80s$, the rarefaction wave that was generated by the the execution of the manoeuvre to unclog the off-ramp, gets close to the junction and is about to hit the junction. After the rarefaction wave hits the junction, it becomes unclogged. Therefore, the FIFO model predicts nonzero flux through
the junction. Figure 4.27, which shows the state of the system at time $t = 100s$, shows this effect. As seen in this figure, the FIFO model shows a rarefaction wave on $I_1$ moving away from the junction.

Figure 4.28 shows results at time $t = 110s$. According to this figure, the non-FIFO model shows a shock wave on $I_3$ moving away from the junction. The other two models do not show this effect at this time. As we will see later, for the FIFOQ and FIFO models, this effect happens at a later time.

Figure 4.29 shows the results at time $t = 250s$. Based on this figure, the FIFO model shows that the rarefaction wave catches up to the shock wave on $I_1$. And we can see that at time $t = 260s$, as shown in Figure 4.30, a shock wave is generated that moves toward the junction. At time $t = 290s$, as shown in Figure 4.31 this wave is about to hit the junction. Figure 4.32 shows the results after the wave hits the junction. As seen in this figure, at this time, both FIFOQ and FIFO models show a wave similar to the one non-FIFO model showed in Figure 4.28. After this point, all three macroscopic models show the same density estimates and match the results from SUMO reasonably well.

![Figure 4.23: Non-exit-only off-ramp network results for $t = 10s$.](image)
Figure 4.24: Non-exit-only off-ramp network results for $t = 15s$.

Figure 4.25: Non-exit-only off-ramp network results for $t = 25s$. 
Figure 4.26: Non-exit-only off-ramp network results for $t = 80s$.

Figure 4.27: Non-exit-only off-ramp network results for $t = 100s$. 
Figure 4.28: Non-exit-only off-ramp network results for $t = 110\,\text{s}$.

Figure 4.29: Non-exit-only off-ramp network results for $t = 250\,\text{s}$.
Figure 4.30: Non-exit-only off-ramp network results for $t = 260s$.

Figure 4.31: Non-exit-only off-ramp network results for $t = 290s$. 
Next, we investigate the number of vehicles that move through the junction over time. Figure 4.33 shows the cumulative number of type 2 vehicles that have passed through the junction over time. According to this figure, FIFOQ and non-FIFO models closely match the results of SUMO. However, the FIFO model shows different numbers for times up to around 90 seconds. This is the time interval during which the off-ramp is completely clogged and the FIFO model prescribes zero flux through the junction. Therefore, the FIFO model cannot accurately capture the number of type 2 vehicles passing through the junction.
Figure 4.33: Total type 2 vehicle tallies for the non-exit-only off-ramp scenario.

Figure 4.34 shows the cumulative number of type 3 vehicles that have passed through the junction over time. According to this figure, FIFOQ and FIFO models closely match the results of SUMO. However, the non-FIFO model shows different numbers for times beyond around 110 seconds. Therefore, the non-FIFO model cannot accurately capture the number of type 2 vehicles passing through the junction.
Based on Figures 4.33 and 4.34, the FIFOQ model is the only model that accurately captures the total number of vehicles passing through the junction for the type 2 and type 3 vehicles. This clearly shows the failure of the traditional models and the strength of the FIFOQ model.
CHAPTER 5

CONCLUSIONS

In this dissertation, we develop new coupling conditions for off-ramps on highways to address some of the existing modeling issues in this area. Specifically, under the classical FIFO coupling conditions, a clogged off-ramp leads to the unrealistic result of having zero flux through the junction. Moreover, the issue with the non-FIFO coupling conditions is that they lead to spurious re-routing of vehicles. To address these issues, the new FIFOQ model uses a vertical queue at the junction to keep track of the excess vehicles of a certain type (exiting vs. non-exiting) that may join the congested traffic by more than the other vehicle type does. Our results show that the FIFOQ model resolves both of the aforementioned shortcomings of FIFO and non-FIFO models.

From a mathematical analysis point of view, our results establish the well-posedness of the new FIFOQ model. In other words, we show that there exists a unique solution that is continuously dependent on the initial data. We use the wave-front tracking algorithm to establish the existence of the solution. For the proof of continuous dependence of the solution to the initial data and its uniqueness, we use the technique of generalized tangent vector.

Finally, we investigate the performance of the FIFOQ model with data. Specifically, we establish micro-simulation representations of the off-ramp scenarios, and describe how to systematically extract macroscopic quantities (vehicle density and queue length) from the results of the microsimulator. Then,
we compare the results of the macroscopic models with the macro quantities extracted from the microsimulator. The results of these experiments highlight the shortcomings of the existing models as well as the ability of the new FIFOQ model to reflect the conditions on the roads more realistically.

Our results show that the new FIFOQ model is in fact mathematically well-posed, and produces the structurally desirable properties that the existing models lack. We therefore would like to advocate that network simulations based on the LWR model consider implementing the new FIFOQ model, based on the CTM discretizations presented in section 2.4.

This dissertation addresses some of the gaps existing in the literature of macroscopic traffic flow modeling of off-ramps on highways. However, there are still some limitations that need to be addressed in the future. One key issue that needs to be studied is that, in reality, the road sharing ratios depend on the queue length. In this dissertation, we have addressed this issue from the modeling point of view. Specifically, the extended FIFOQ model assumes that the road sharing ratios are dependent on the queue length. While this extended model turns out to work well in simulation, future studies should investigate it both from an analytical and data validation perspectives. From the mathematical analysis point of view, the well-posedness proof of the FIFOQ model assumes that the road sharing ratios are constant and equal to the split ratios. Proving the well-posedness of the extended FIFOQ models, in which the road sharing ratios are not necessarily equal to the split ratios, needs to be explored in the future. From the data validation point of view, in this dissertation we have used data from a microsimulator to validate the FIFOQ model under the assumption that the road sharing ratios are constant but not necessarily equal to the split ratios. However, data validation of the extended FIFOQ model, with road sharing ratios that are dependent on the queue length, has not been explored. Therefore, the use of data to find suitable choices for the road sharing ratios as functions of the queue length as well as validation of the extended FIFOQ model with non-constant road sharing ratios are suggested as possible avenues for future research efforts in this
area. The future work in this area should also include a validation with real roadway data. The data required for this task would need to include vehicle flows with accurate spatio-temporal resolutions that exceed what is currently publicly available. However, future data sets can have this needed resolution.
REFERENCES


