MONASH University

On the simulation of arterial road networks

by

Samithree Rajapaksha

Supervisor: **A/Prof. Tim Garoni**

Associate Supervisor: **Dr. Joyce Zhang**

November 2020

A thesis submitted for the degree of Doctor of Philosophy
at Monash University
School of Mathematics

©Copyright

by

Samithree Rajapaksha

2020

Abstract

Traffic simulation models have become an increasingly important tool for investigating vehicular dynamics in arterial networks. In this thesis, we study three distinct, but related, topics in the field of traffic simulation. Specifically, we study calibration of traffic models via Approximate Bayesian Computation (ABC); the form of link travel time distributions, particularly in the presence of traffic lights; and the form of dwell time distributions for on-road public transport, in particular, for trams.

The ability of a simulation model to accurately predict real-world behaviour is significantly impacted by the values of the model parameters. We propose and study the use of Approximate Bayesian Computation (ABC) to calibrate traffic flow simulations. A key challenge in implementing ABC in practice is to identify summary statistics which are low-dimensional, to allow efficient computation, but also sufficiently informative, to allow accurate identification of appropriate parameters. We consider in some detail the problem of estimating the input and output rates of traffic on a single link, and show that the mean density at the upstream and downstream ends of the link provide accurate summary statistics. We also consider the estimation of turning probabilities in a grid network. Moreover, we show that ABC with regression adjustments provides a more accurate calibration method than simple rejection ABC.

Understanding the statistical distribution of link travel times is important for meso- and macro-scale modelling of vehicular dynamics in arterial road networks. Most studies related to travel time distributions have previously focussed on empirical estimations, and the results have tended to be somewhat inconclusive. To provide some physical intuition, we consider the travel time distribution of a simple stochastic transport model, the Nagel–Schreckenberg model (NaSch), paying particular attention to the special case of the Asymmetric Simple Exclusion Process (ASEP). We find that the simulated travel-time of the NaSch model is well approximated by a normal distribution as the link length becomes large, suggesting a central limit theorem holds. In the particular case of ASEP, we conjecture explicit forms for the mean and variance of this normal distribution, in terms of the input and output probabilities, and provide strong numerical evidence in their favour. Moreover, we also consider in detail the impact on the link travel time distribution of having traffic lights present at both the upstream and downstream intersections. In particular, we investigated the effects of varying signal parameters such as offset and splits. Our numerical results suggest that travel time distributions can be well described by a Gaussian mixture model, where the components in the mixture correspond roughly to the (random) number of cycles a vehicle must wait at the downstream intersection.

Accurately modelling the dwell time at bus or tram stops is an important challenge in simulations of on-road public transport. Via a combination of simulation and direct statistical survey, we study the distribution of the dwell time at selected tram stops in inner Melbourne. The empirical study strongly suggests that the dwell time at a given stop is well-described by a log normal distribution, with parameters that vary only weakly with congestion levels etc. Moreover, we also considered the question of possible correlations of the dwell times between consecutive trams and/or stops. We find, perhaps surprisingly, that the correlations in fact appear rather weak. As a practical consequence, we conclude that modelling each tram's dwell time, at each stop, as an independent log-normal random variable should provide an accurate method for generating random dwell times in simulation-based studies. We also provide estimates of the relevant log-normal parameters.

This thesis is dedicated to my loving parents and to my loving husband

Declaration

This thesis is an original work of my research and contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

.....
Signature

.....
Print Name

.....
Date

Acknowledgements

I am highly indebted to my supervisors, A/Prof. Tim Garoni and Dr. Joyce Zhang for their critique, advice, direction and inspiration which spurred me on from the beginning of this thesis preparation right to up till the end.

I wish to thank my milestone evaluation panel A/Prof. Jonathan Keith, A/Prof. Tianhai Tian, and Jeremy Burdan from VicRoads, for their constructive comments to improve my research findings. I also express my sincere thanks to all the other academic and administrative staff members of the School of Mathematics at Monash University, Melbourne, Australia. I deeply acknowledge Prof. Scott Sisson for guiding me to enhance my knowledge on Approximate Bayesian Computation. I'm grateful to Prof. Graham Currie for his guidance and supporting throughout my candidature.

I wish to express my profound gratitude to my industry sponsor VicRoads for their financial support, without which this study would not have been possible. I honour the School of Mathematics for providing me with the opportunity to study at this prestigious university. The financial assistance received from the School of Mathematics and The Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS) was undoubtedly helpful and paved the way for me to comfortably focus on my studies.

I am truly grateful to Senior Application Specialist, Simon Michnowicz at the Monash eResearch Centre, who supported me with queries related to supercomputing throughout my candidature. Also, I would like to acknowledge the National Computational Infrastructure, NCI, Monash University HPC/HTC Cluster, MonARCH and University of Melbourne HPC system, Spartan, made available for conducting the research reported in this thesis.

I would like to thank my peers Randika Ariyaratne, Helani Kottage, Somayeh Shiri, Sandun Goonaratne, Rachel Mence, Rejitha Nath, Puwasala Gamakumara, Udip Karunaratne, Chathu Hettiarachchi and Amila Goonaratne for supporting me throughout my candidature. I am truly grateful to my parents and to my sister for their unconditional love, enormous support and encouragement given to me while achieving the best things in my life. I am extremely grateful to my husband for his immense support and understanding throughout this candidature.

Contents

List of Figures **ix**

List of Tables **xvii**

List of Acronyms **xix**

1 Introduction **1**
1.1 Thesis outline 4

2 Traffic simulation models **5**
2.1 Overview 5
2.1.1 *Nagel Schreckenberg Model* (NaSch) 7
2.1.2 *Asymmetric Simple Exclusion Process* (ASEP) 8
2.1.3 *Cellular Automata Simulator for Arterial Roads* (CEASAR) 12

3 Calibration of traffic flow simulation models **14**
3.1 Overview 14
3.2 *Approximate Bayesian Computation* (ABC) for model calibration 17
3.2.1 ABC methods 18
3.3 Applying ABC for traffic model calibration 23
3.3.1 Informative summary statistics for ABC 24
3.3.2 Synthetic observed data for ABC 26
3.3.3 Parameter estimation using ABC 27
3.4 Discussion 45

4 Link travel time distributions **47**
4.1 Overview 47
4.1.1 Literature Review: link travel time distributions 48
4.2 Non-signalised links 51
4.2.1 Mean-field approximation 51
4.2.2 Simulations: ASEP 53
4.2.3 Simulations: NaSch 63
4.3 Signalised links 66
4.3.1 Models 66
4.3.2 System with a lower inflow: $\alpha = 0.1$ and $\beta = 0.9$ 72

4.3.3 System with a higher inflow: $\alpha = 0.5$ and $\beta = 0.9$ . . . 84
4.4 Discussion . . . 90

5 Simulating stopping behaviour of on-road transit vehicles 92
5.1 Overview . . . 92
5.2 Survey Data - Impact of Crowding on Streetcar Dwell Time [35] 95
5.3 Survey: The correlation structure of stopping behaviour of a tram 98
5.3.1 Survey method . . . 99
5.3.2 Tram bunching . . . 99
5.3.3 Marginal distributions of observables . . . 101
5.3.4 Correlation analysis . . . 107
5.4 Simulation . . . 113
5.4.1 Model . . . 113
5.4.2 System configuration . . . 114
5.5 Numerical results . . . 116
5.5.1 Correlation analysis . . . 119
5.6 Discussion . . . 126

6 Discussion 128

Appendix 130
Travel time distribution: signalised links . . . 130
Lower inflow: $\alpha = 0.1$ and $\beta = 0.9$ . . . 131
Higher inflow, $\alpha = 0.5$ . . . 135

Bibliography 139

List of Figures

2.1 Configuration of NaSch with $v_{max} = 3$. The number in the upper right corner is the speed of the vehicle. 7
2.2 Configuration of an ASEP with one dimensional lattice consists of L sites. Parameters α, β and p represent the traffic inflow probability, outflow probability and hopping probability respectively. 8
2.3 Exact phase diagram of Asymmetric Simple Exclusion Process (ASEP) with random sequential dynamics for $p = 1$. The density profile becomes flat on the dotted line $\alpha + \beta = 1$. 9
2.4 Exact phase diagram of Asymmetric Simple Exclusion Process (ASEP) with stochastic parallel dynamics for $p = 0.5$. The broken line is corresponds to $(1 - \alpha)(1 - \beta) = 1 - p$, where the density profile is flat [58]. 11

3.1 Joint posterior distribution of boundary parameters of an ASEP with random sequential updates when the observed parameters are in the low-density regime. Sub-figures (a), (c) and (e) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 100, 200 and 300 respectively. Sub-figures (b), (d) and (f) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 100, 200 and 300 respectively. The red dot represents the observed parameters, $\alpha = 0.2$, $\beta = 0.8$. 28
3.2 Joint posterior distribution of boundary parameters of an ASEP with random sequential updates when the observed parameters are in the high-density regime. Sub-figures (a), (c) and (e) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 100, 200 and 300 respectively. Sub-figures (b), (d) and (f) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 100, 200 and 300 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.2$. 29

3.3 Joint posterior distribution of boundary parameters of an ASEP with random sequential updates when the observed parameters are in the maximum-current regime. Sub-figures (a), (c) and (e) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 100, 200 and 300 respectively. Sub-figures (b), (d) and (f) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 100, 200 and 300 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.8$. 30

3.4 Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 1$ when the observed parameters are in the low-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.2$, $\beta = 0.8$. 32

3.5 Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 1$ when the observed parameters are in the high-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustment for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.2$. 33

3.6 Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 1$ when the observed parameters are in the maximum-current regime. Sub-figures (a) and (c) represents the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 1$, $\beta = 1$. 34

3.7 Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 0.5$ when the observed parameters are in the low-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.2$, $\beta = 0.8$. 35
3.8 Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 0.5$ when the observed parameters are in the high-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustment for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.2$. 36
3.9 Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 0.5$ when the observed parameters are in the maximum-current regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustment for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.7$, $\beta = 0.7$. 37
3.10 Joint posterior distribution of boundary parameters of NaSch when the observed parameters are in a low-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.2$, $\beta = 0.8$. 39
3.11 Joint posterior distribution of boundary parameters of NaSch when the observed parameters are in a maximum-current regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.4$, $\beta = 0.7$. 40

3.12 Joint posterior distribution of boundary parameters of NaSch when the observed parameters are in a high-density regime. Sub figures (a) and (c) represents the join posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.2$. 41
3.13 4×5 lattice network 42
3.14 Correlation between the average flow, and turning probabilities at Node 20 43
3.15 Correlation between the average speed, and turning probabilities at Node 20 43
3.16 Joint posterior distribution of east-bound right turning probability and left turning probability at the Node 20 44

4.1 Link Configuration: Parameters α, β and p_i represent the traffic inflow probability, outflow probability and cell dependent hopping probability respectively. 51
4.2 Link Configuration: Parameters α, β and q represent the traffic inflow probability, outflow probability and queue length respectively. 52
4.3 Travel time distribution of a vehicle in high-density regime ($\alpha = 0.8$ and $\beta = 0.2$) of an ASEP with deterministic updates. The blue curve represents the approximated travel time distribution. Empirical distributions are smoothed with non-parametric Gaussian kernel. 54
4.4 Travel time distribution of a vehicle in low-density regime ($\alpha = 0.1$ and $\beta = 0.6$) of an ASEP with deterministic updates. The blue and the golden curves represent the approximated travel time distribution based on the empirical queue length and queuing theory respectively. Empirical distributions are smoothed with non-parametric Gaussian kernel. 54
4.5 Mean comparison of approximated and empirical travel times of an ASEP for different density regions with varying link sizes. Error bars corresponding to one standard deviation are shown. Fitted line of NWLS estimates for empirical mean travel time is shown. 57
4.6 Variance comparison of approximated and empirical travel times of an ASEP for different density regions with varying link sizes. Error bars corresponding to one standard deviation are shown. Fitted line of NWLS estimates for empirical mean travel time is shown. 58
4.7 Travel time distribution of ASEP when the system is in the LD1: $\alpha = 0.1$ and $\beta = 0.3$. Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 60

4.8 Standardised travel time distribution of ASEP when the system is in the LD2 regime of an ASEP ($\alpha = 0.001$ and $\beta = 1$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 61
4.9 Standardised travel time distribution of ASEP when the system is in the LD2 regime ($\alpha = 0.1$ and $\beta = 0.6$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 61
4.10 Standardised travel time distribution of ASEP when the system is in the HD1 regime ($\alpha = 0.2$ and $\beta = 0.1$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 62
4.11 Standardised travel time distribution of an ASEP when the system is in the HD2 regime ($\alpha = 0.2$ and $\beta = 0.8$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 63
4.12 Travel time distribution of an ASEP when the system is in the MC regime ($\alpha = 0.4$ and $\beta = 0.7$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non parametric Gaussian kernel. 63
4.13 Travel time distribution of NaSch when the system is in very low-density region ($\alpha = 0.001$ and $\beta = 1$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 64
4.14 Travel time distribution of NaSch when the system is in low-density region ($\alpha = 0.1$ and $\beta = 0.6$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 64
4.15 Travel time distribution of NaSch when the system is in maximum-current region ($\alpha = 0.4$ and $\beta = 0.7$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 65
4.16 Travel time distribution of NaSch when the system is in high-density region ($\alpha = 0.8$ and $\beta = 0.2$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel. 65
4.17 Link Configuration: Parameters α and β represents the traffic inflow and outflow probabilities respectively. 66
4.18 Domain wall trajectories for $c = 90s$, $g_1 = g_2 = c/2$ and $d = 25s$. X-axis: position in the bulk link; Y-axis: time. Green and red vertical lines represent the traffic lights at upstream and downstream intersections. 69
4.19 Association between empirical queue length and the offset 73

4.20 Variability of estimated μ_i's in Gaussian mixture distribution with the offset. *App_*μ_1 and *App_*μ_2 represent the approximated mean travel time with free flow and with waiting for green times at the downstream intersection respectively. The inflow probability $\alpha = 0.1$ 74
4.21 Variability of estimated λ_i's in Gaussian mixture distribution with the offset. *App_*λ_1 and *App_*λ_2 represent the approximated mixing proportions of vehicles joins the green wave and the vehicles misses one cycle respectively. The inflow probability $\alpha = 0.1$ 75
4.22 variability of estimated σ_i's in Gaussian mixture distribution with the offset, The inflow probability $\alpha = 0.1$ 76
4.23 Travel time distribution: Case 45s-45s in low-density. Approximation represents the Gaussian mixture distribution by replacing μ_1, μ_2, λ_1 and λ_2 with the approximated values. Empirical distributions are smoothed with non-parametric Gaussian kernel. . . . 77
4.24 Travel time distribution: Case 45s-60s in low-density. Approximation represents the Gaussian mixture distribution with replacing μ_1, μ_2, λ_1 and λ_2 with the approximated values. Empirical distributions are smoothed with non-parametric Gaussian kernel. . . . 78
4.25 Histogram of travel time with low-density and offset 50s in case 45s-60s. Gaussian mixture distribution is shown by gold area and the non-parametric Gaussian kernel density estimation is shown by grey area. 79
4.26 Travel time distribution: Case 60s-45s in low-density. Approximation represents the Gaussian mixture distribution with replacing μ_1, μ_2, λ_1 and λ_2 with the approximated values. Empirical distributions are smoothed with non-parametric Gaussian kernel. . . . 81
4.27 Travel time distribution: Case 60s-60s in low-density. Approximation represents the Gaussian mixture distribution with replacing μ_1, μ_2, λ_1 and λ_2 with the approximated values. Empirical distributions are smoothed with non-parametric Gaussian kernel. . . . 82
4.28 Histogram of travel time in Case 45s-45s. Gaussian mixture distribution is shown in gold area. Non parametric Gaussian kernel density estimation is shown in grey area. 84
4.29 Histogram of travel time in Case 45s-60s. Gaussian mixture distribution is shown in gold area. Non-parametric Gaussian kernel density estimation is shown in grey area. 86
4.30 Histogram of travel time in Case 60s-45s. Gaussian mixture distribution is shown in gold area. Non-parametric Gaussian kernel density estimation is shown in grey area. 87
4.31 Histogram of travel time in Case 60s-60s. Gaussian mixture distribution is shown in gold area. Non-parametric Gaussian kernel density estimation is shown in grey area. 89

5.1 Melbourne tram network: routes and survey locations. Adapted from "Impact of Crowding on Streetcar Dwell Time" by Currie et al., 2013 [35] 95
5.2 Survey: Distribution of Passenger Flow Time at tram stops with a safety-zone. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel. 96
5.3 Survey: Distribution of Passenger Flow Time at tram stops with a curb-side. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel. 97
5.4 Survey: Distribution of Passenger Flow Time at tram stops with a Platform. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel. 97
5.5 Melbourne tram network: routes and survey locations 98
5.6 Survey: Occurrences of bunching at St Kilda Road. We consider any time-headway less than five seconds as bunched. 100
5.7 Survey: Distribution of Passenger Flow Time. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel shown in grey area. 101
5.8 Survey: Four Goodness-of-fit plots for gamma distribution and log normal distribution for time-headway at Leopold Street stop . . . 103
5.9 Survey: Four Goodness-of-fit plots for gamma distribution and log normal distribution for time-headway at Arthur Street stop 103
5.10 Survey: Distribution of time-headway. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel shown in grey area. 104
5.11 Survey: Travel time distribution of trams between stops. Gaussian Mixture Model curve and 95% confidence bands are included. Empirical distributions are smoothed with non-parametric Gaussian kernel shown in grey area. 106
5.12 Survey: Correlation of time-headway between consecutive stops. $HW_{t,\ stop\ name}$ represents the time-headway of tram t at a given stop. 107
5.13 Survey: Correlation of Passenger Flow Time between consecutive stops. $PFT_{t,\ stop\ name}$ represents the passenger flow time of tram t at a given stop 108
5.14 Survey: Correlation of time-headway between consecutive trams at each stop. $HW_{t,\ stop\ name}$ represents the time-headway of tram t at a given stop. 109
5.15 Survey: Correlation of passenger flow time between consecutive trams at each stop. $PFT_{t,\ stop\ name}$ represents the passenger flow time of tram t at a given stop. 111

5.16 Survey: Correlation between time-headway and passenger flow time of a tram at each stop. $HW_{t,\ stop\ name}$ and $PFT_{t,\ stop\ name}$ represents the time-headway and passenger flow time of tram t at a given stop. 112
5.17 Simplified flowchart of the simulation model. 113
5.18 Simulated route and stop locations on St Kilda Road 114
5.19 Comparison: Distribution of Passenger Flow Time. 95% confidence bands (yellow area) are included. Simulated and empirical distributions are smoothed with non-parametric Gaussian kernel which is shown in grey area. 117
5.20 Simulation: Distribution of time-headways. 95% confidence bands (yellow area) are included. Simulated and empirical distributions are smoothed with non-parametric Gaussian kernel which is shown in grey area. 118
5.21 Simulation: Correlation matrix of dwell times of trams between stops. Colour scheme representing the direction of the association is shown. 120
5.22 Simulation: Correlation matrix of time-headways of trams between stops. Colour scheme representing the direction of the association is shown. 120
5.23 Simulation: Correlation of dwell times between tram $t-1$ and t. Colour scale: dark blue to light blue represents the high positive correlation to slight positive correlation. 121
5.24 Simulation: Correlation of time-headways between tram $t-1$ and t. Colour scale: yellow to light blue represents the week negative correlation to week positive correlation. 123
5.25 Simulation: Correlation between dwell time and time-headway of a tram t at each stop. Colour scale: dark blue to light blue represents the high positive correlation to slight positive correlation. 124

1 Travel time distribution: Case 45s-45s and $\alpha = 0.1$. Fitted Gaussian mixture model is shown in gold area. 131
2 Travel time distribution: Case 45s-60s and $\alpha = 0.1$. Fitted Gaussian mixture model is shown in gold area. 132
3 Travel time distribution: Case 60s-45s and $\alpha = 0.1$. Fitted Gaussian mixture model is shown in gold area. 133
4 Travel time distribution: Case 60s-60s and $\alpha = 0.1$. Fitted Gaussian mixture model is shown in gold area. 134
5 Travel time distribution: Case 45s-45s and $\alpha = 0.5$ 135
6 Travel time distribution: Case 45s-60s and $\alpha = 0.5$ 136
7 Travel time distribution: Case 60s-45s and $\alpha = 0.5$ 137
8 Travel time distribution: Case 60s-60s and $\alpha = 0.5$ 138

List of Tables

2.1 Exact solutions of an ASEP with random sequential updates with $p = 1$ [43]. α and β are the inflow probability and outflow probability respectively. ρ_1, $\rho_{L/2}$, and ρ_L represents the average density at left-boundary, bulk and right-boundary respectively. 10

4.1 Non-linear WLS Regression Fit for empirical and approximated μ of travel time of an ASEP with varying link sizes. Formula: $\mu = b \times L^c$ 56
4.2 Non-linear WLS Regression Fit for empirical and approximated σ^2 of travel time of an ASEP with varying link sizes. Formula: $\sigma^2 = b \times L^c$ 60
4.3 Gaussian mixture distribution parameter estimation in low-density: case 45s-45s 77
4.4 Gaussian mixture distribution parameter estimation for travel time with low-density and offset 50s in case 45s-60s 79
4.5 Gaussian mixture distribution parameter estimation in low-density: case 45s-60s 80
4.6 Gaussian mixture distribution parameter estimation in low-density: case 60s-45s 81
4.7 Gaussian mixture distribution parameter estimation in low-density: Case 60s-60s 83
4.8 Gaussian mixture distribution parameter estimation in high-density: Case 45s-45s 85
4.9 Gaussian mixture distribution parameter estimation in high-density: Case 45s-60s 86
4.10 Gaussian mixture distribution parameter estimation in high-density: Case 60s-45s 88
4.11 Gaussian mixture distribution parameter estimation in high-density: Case 60s-60s 89

5.1 Survey: Log-normal distribution parameter estimation of passenger flow time at each stop type. 97
5.2 Survey: Log-normal distribution parameter estimation of passenger flow time at each surveyed stop. 102

5.3 Survey: Estimated parameters for time-headway distribution at surveyed routes in the Melbourne tram network. 105
5.4 Survey: Estimated Gaussian Mixture Model parameters for travel time distribution at surveyed routes in the Melbourne tram network 106
5.5 Survey: Correlation coefficient of time-headway of a tram between stops. 107
5.6 Survey: Correlation coefficient of passenger flow time of a tram between stops. 108
5.7 Survey: Correlation coefficient of time-headway between consecutive trams at each stop. 110
5.8 Survey: Correlation coefficient of passenger flow time between consecutive trams. 110
5.9 Survey: Correlation coefficient between time-headway and passenger flow time of a tram at each stop. 112
5.10 Multivariate Regression: Passenger Flow Time 115
5.11 Time dependent passenger arrival rates. Averaging over two surveyed stops (i.e. $r_t = (r_t^L + r_t^A)/2$) 115
5.12 Comparison: Log-normal distribution parameter estimation of passenger flow time at stop 05 and 06 118
5.13 Simulation: Gamma distribution parameter estimation of passenger flow time at stops 05 and 06 119
5.14 Simulation: Correlation between dwell time of consecutive trams. $Corr(DT_{(t-1,i)}, DT_{(t,i)})$, for $i = 01, 02, ..., 19$. $DT_{(t,i)}$ represents the dwell time of tram t at stop i. 122
5.15 Simulation: Correlation between time-headway of consecutive trams. $Corr(HW_{(t-1,i)}, HW_{(t,i)})$, for $i = 01, 02, ..., 19$. $HW_{(t,i)}$ represents the time-headway of tram t at stop i. 122
5.16 Simulation: Correlation between the dwell time and time-headway of a tram at each stop. $Corr(HW_{(t,i)}, DT_{(t,i)})$, for $i = 01, 02, ..., 19$. $DT_{(t,i)}$ represents the dwell time of tram t at stop i and $HW_{(t,i)}$ represents the time-headway of tram t at stop i. 125

List of Acronyms

ABC Approximate Bayesian Computation vii

ASEP Asymmetric Simple Exclusion Process vii

NaSch Nagel Schreckenberg Model vii

SPSA Simultaneous Perturbation Stochastic Approximation 17

GA Genetic Algorithm 17

MCMC Markov Chain Monte Carlo 18

MCMC-ABC Markov Chain Monte Carlo ABC 18

SMC Sequential Monte Carlo 18

SMC-ABC Sequential Monte Carlo ABC 46

RABC SMC-ABC replenishment 18

iid independent and identically distributed 16

CLT Central Limit Theorem 22

SCATS Sydney Coordinated Adaptive Traffic System 12

MTSS Main Traffic Signal System 45

SCOOT Split Cycle and Offset Optimization Technique 45

ACTS Adelaide Coordinated Traffic Signal System 45

AVL Automatic Vehicle Location 49

GPS Global Positioning System 48

CEASAR Cellular Automata Simulator for Arterial Roads vii

ITS Intelligent Transportation System 23

CA Cellular Automata 6

AVI Automatic vehicle identification 48

Introduction 1

Significant population growth poses serious challenges related to transportation in arterial networks across major cities worldwide. Simulation provides a relatively fast and cost-effective method to analyse and explore proposed modifications of transport systems, since it eliminates the costly constructions demanded by real time experimentation, and also minimises potential risks to travellers. However, the design and application of simulation models must be handled very carefully, in order to have confidence that the model outputs give reliable predictions. The objectives of this thesis are threefold: firstly, to identify a robust model calibration and validation procedure for traffic simulators; secondly, to accurately approximate the link travel time distribution, and assess the impact of traffic lights on it; thirdly, to simulate stopping behaviour of on-road public transport vehicles, incorporating dwell time dynamics at stops.

There exist a significant number of commercially available transport simulation models. These facilitate fundamental decision making on the layout of traffic infrastructure, such as evaluation of public transport priority schemes. Of the commercially available packages, CUBE, TransCAD and TRANSIM are commonly used for transportation planning, while other packages such as CORSIM, VISSIM, Paramics and DynaMIT are used for simulating transportation operations [1]. Traditional microscopic traffic simulation models are complex, and explore transit behaviours in great detail, while models such as cellular automata [2] reproduce the correct macroscopic behaviour based on a minimal description of microscopic interactions. These kinds of mesoscopic traffic simulation models are computationally efficient for large scale traffic flow simulations.

Traffic simulation models can simulate complex arterial networks, with different driving behaviours, and various traffic management strategies. Due to the complexity of arterial networks, these traffic simulators depend on a large number of free parameters. To have confidence in a model's predictions, such parameters need to be calibrated so that the simulated outputs reproduce relevant features of empirical data from actual transport networks. However, the simulated outputs may depend sensitively on the values chosen for the parameters. Moreover, for complex models, there are typically many choices of parameters which will reproduce the specified features of a given data set.

Although much effort has been put into the development and refinement of traffic simulation models, rather less attention has been given to developing statistically rigorous methodologies for their calibration and validation. It is generally considered the user's responsibility to identify and quantify the differences between observed and simulated measurements [3]. Many published traffic simulator validation studies are based on visual comparisons or comparisons of simple descriptive statistics [4, 5]. In visual validation, outputs of the simulated and the real systems are compared to determine whether or not they can be differentiated. The visualisation may be based on the animation modules available in most traffic simulation models or based on the plots of different outputs such as flows, speeds and queue lengths [6]. Popular statistical validation techniques among transport planners are Goodness of Fit measures, confidence intervals and statistical tests for mean comparisons [7, 8] such as two sample t–test, F-test, Mann-Whitney test, Kolmogrov-Smirnow and Wilcoxon Rank Sum tests.

Moreover, the traditional approach of calibrating parameters is to modify parameter values until the simulated output approximates the actual data. This can result in poor parameter choices and ignores the uncertainties remaining in the parameters [9]. Moreover, many standard statistical techniques are not easily applied due to non-stationarity [10]. Wanger et al claim that stochasticity is the reason many parameter estimation problems in traffic simulations fail [11]. Furthermore, the optimisation algorithms used in typical calibration processes are often computationally expensive.

Approximate Bayesian Computation (ABC) has gained much attention in recent years as an alternative to traditional parameter estimation techniques. ABC is a flexible and robust technique for model calibration. Instead of searching for "true" parameter values, ABC takes a Bayesian perspective, whereby one computes a (posterior) probability distribution on the space of possible parameter values, describing how likely the parameters are to reproduce outputs in agreement with empirical data. However, unlike classical Bayesian methods, ABC does not require explicit calculation of the likelihood function, and is therefore quite widely applicable, even to the calibration of complex

simulation models. As a proof of concept, in this thesis, we study in some detail the application of ABC methods to the calibration of some simple stochastic cellular automata traffic models. A key task in the application of ABC in practice is the identification of appropriate summary statistics. The study of simple cellular automata models, such as the asymmetric simple exclusion process (ASEP), for which many quantities are understood analytically, proves to be a very useful setting for considering appropriate summary statistics.

A fundamental question concerning traffic flow, with widespread practical applications in meso- and macro-scale traffic simulation, is to understand the statistical distribution of the travel time of individual vehicles on a single link. A wide variety of approaches have been taken to estimating link travel times, such as time series forecasting methods [12, 13], Kalman filtering [14, 15], Neural Networks [16–18] and empirical distribution estimations methods [19–22]. Most previous studies have tended to model travel time using uni-modal symmetric, continuous distributions, such as the normal distribution [23]. More recently, however, researchers have argued that travel time distributions are positively skewed [19–22], and possibly multi-modal. The existing literature on empirical studies therefore presents a variety of inconsistent conclusions, and tends not to provide clear physical intuition behind the proposed forms of travel time distributions. In order to develop physical intuition concerning the distribution of the link travel time, we again turn to simple cellular automata models, where explicit conjectures can be made, and compared with simulation. Based on a study of these simple models, we conclude that the travel time on links without traffic signals should generically exhibit Gaussian behaviour, at least when the size of the link is sufficiently large. In addition to being of practical interest, these conjectures concerning the cellular automata models are likely of theoretical interest to mathematical physicists in the field of statistical mechanics. Moreover, using an appropriate generalisation of these cellular automata to include traffic signals, we systematically study the effect of traffic signal parameters, such as offsets and splits, on the travel time distributions. We conclude that, in all cases, the travel time distribution is well-described by a Gaussian mixture model. In the case of low density, we also provide domain wall arguments which allow us to accurately predict, on simple physical grounds, the values of many of the parameter appearing in the mixture model.

Finally, we also study a very practical question related to the simulation of on-road public transport: what is the distribution of the amount of time a tram/bus *dwells* at a given stop? Many researchers have studied aspects of the dwell time of public transport vehicles, using various methods such as multivariate regression models [24–37], time series models [38] and decision trees [39]. The statistical distribution of the dwell time has also been studied, however the results are rather inconsistent, variously concluding the distribution

is best approximated by the: Pearson type 6 distribution [26]; Log-normal distribution [31, 40, 41]; and Wakeby distribution [41]. In an effort to resolve these discrepancies, we performed an empirical statistical survey of tram dwell times along two separate tram routes in inner Melbourne. Our data strongly suggest that the dwell time distribution of an individual tram is well-described by a log-normal distribution. To use a dwell time distribution in a simulation model however, it is *a priori* not sufficient merely to understand the marginal distribution of dwell time of an isolated tram, since the dwell times of consecutive trams could potentially be quite strongly dependent. We therefore also studied the correlations between consecutive trams and stops. Perhaps surprisingly, we find that the correlations are in fact typically quite weak. To a good approximation therefore, this justifies modelling dwell times as independent log-normal random variables in simulations. To our knowledge, the correlations in dwell times have not previously been studied. In order to better understand the physical mechanisms behind these observations, we also performed systematic simulations of a tram route in Melbourne, and studied both the marginal dwell time distributions and the dwell time correlations.

1.1 Thesis outline

In Chapter 2, we discuss existing traffic models, and give a detailed description of the models we employ throughout the thesis. In Chapter 3 we study the applicability of ABC in traffic simulation model calibration. In Chapter 4 we study the distribution of the link travel time, both in the absence and presence of traffic signals. Chapter 5 contains our study of the dwell time distribution, and related questions such as the headway distribution. Finally, Chapter 6 contains a discussion.

Traffic simulation models 2

2.1 Overview

Simulation is an essential tool employed by transport planers [42]. Traffic simulators can represent complex traffic networks, and mimic different driving and travel behaviours, such as the operation of different traffic management strategies.

Traffic simulation models can be classified into several categories based on different characteristics. We can categorise them into Microscopic, Macroscopic or Mesoscopic traffic simulators, depending on the level of detail each simulation model considers. Macroscopic simulation models view the traffic flow as a continuum, while Microscopic models focus on individual movements of vehicles and their interactions. Mesoscopic models have characteristics of both the macroscopic and microscopic behaviours. They deliberately ignore individual interactions by specifying the individual behaviour in terms of probability distributions [43]. Traditional microscopic traffic simulation models are complex and explore transit behaviour in great detail. Therefore, the usage of such simulation models for large networks is generally beyond reach. In contrast, Mesesocopic models are computationally efficient for large scale traffic simulations. These models can reproduce the correct macroscopic behaviour based on the minimal description of microscopic interactions.

Furthermore, traffic simulation models can be classified into two main

categories, whether the three main measurable: state; space; and time are discrete or continuous. It is possible to have any combination of discrete or continuous variables measured in the simulations, even though the discrete variables are often approximations [43]. Moreover, the dynamics of a traffic simulation can be either deterministic or stochastic.

These traffic models range from simple analytical models such as, Car-following [44–46], Gas-Kinetic [47, 48], Fluid dynamic [49–51], Coupled-map models [52, 53] and Cellular Automata [54–57] to detailed complex simulation software. There exist a significant number of commercially available transport simulation models. These facilitate fundamental decision making on the layout of traffic infrastructures, such as evaluation of public transport priority schemes. Of the commercially available packages, CUBE, TransCAD and TRANSIM are commonly used for transportation planning, while other packages such as CORSIM, VISSIM, Paramics and DynaMIT are used for simulating transportation operations [1].

Cellular Automata (CA) are models of physical systems in which both space and time are assumed to be discrete. They consist of a regular arrangement of cells. Each of the interacting units can have only a finite number of discrete states. A typical CA model has rule-based dynamics, and is well-suited for modelling traffic systems [43]. A fundamental stochastic CA model, capable of reproducing the main features of vehicular motion along a single lane, is the NaSch model, first introduced by Nagel and Schreckenberg in 1992 [54]. A special case of the NaSch model is the asymmetric simple exclusion process (ASEP), which is well-studied in both mathematical physics and probability theory.

In this chapter we recall the definition of the NaSch model which we use as our standard simulation model in this thesis. In particular, we also summarise relevant known analytical results concerning the ASEP. In addition, we also recall the definition of a network generalisation of NaSch, known as CEASAR, which provides a mesoscopic simulator for traffic flow on arterial roads.

2.1.1 NaSch

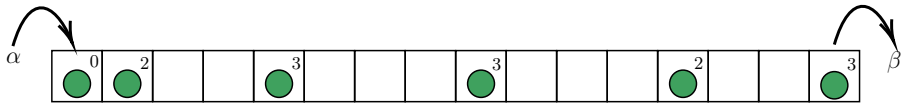


FIGURE 2.1: Configuration of NaSch with $v_{max} = 3$. The number in the upper right corner is the speed of the vehicle.

In NaSch model we consider a road segment divided into L cells. See Figure 2.1. A vehicle is inserted to the system with an inflow probability, $\alpha \in (0, 1)$ at the in-boundary (left-boundary) and removed from the out-boundary (right-boundary) with an outflow probability, $\beta \in (0, 1)$. A vehicle occupies one cell, takes $v_{max} + 1$ values of speed $v = 0, 1, 2, ..., v_{max}$ with a deceleration probability p. Each vehicle moves forward according to its velocity obeying NaSch dynamics (i.e. $l_n \rightarrow l_n + v_n$). Suppose l_n and v_n denote the position and the speed of the n^{th} vehicle, then $h_n = l_{n+1} - l_n - 1$ is the spatial headway of the n^{th} vehicle at time, t. At each time step $t \rightarrow t + 1$, the speed of the n^{th} vehicle is increased by one if it is not already in v_{max} (i.e. $v_n \rightarrow min(v_n + 1, v_{max})$), reduced to h_n if $v_n > h_n$ (i.e. $v_n \rightarrow min(v_n, d_n)$) or reduced by one, which happens randomly with probability p if $v_n > 0$ (i.e. $v_n \rightarrow max(v_n - 1, 0)$ with probability p) [54].

For $v_{max} = 1$, NaSch model has a special case called ASEP. ASEP can also be viewed as a road segment divided into L cells. A vehicle is inserted to the system with an inflow probability, $\alpha \in (0, 1)$ at the in-boundary (left-boundary) and removed from the out-boundary (right-boundary) with an outflow probability. $\beta \in (0, 1)$. A vehicle occupies one cell and moves forward a cell, with a probability of p, if the consecutive cell is unoccupied by a vehicle. See [58] for detailed explanation about ASEP.

2.1.2 ASEP

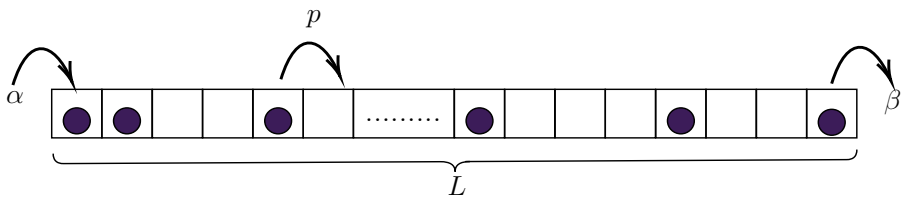


FIGURE 2.2: Configuration of an ASEP with one dimensional lattice consists of L sites. Parameters α, β and p represent the traffic inflow probability, outflow probability and hopping probability respectively.

ASEP is a stochastic traffic model defined on a one-dimensional lattice (see Figure 2.2). The lattice consists with L number of cells where, only one vehicle can occupy a cell. Vehicles can only hop into empty cells with a speed of $v_{max} = 1$. Vehicles are inserted from the left-boundary with probability $\alpha \in (0, 1)$ onto the first cell if the first cell is not occupied by a vehicle. Vehicles in the bulk move forward one cell with probability $p \in (0, 1)$ and exit from the right-boundary with probability $\beta \in (0, 1)$. The model is known for showing phase transitions among low-density, high-density and maximum-current phases with respect to the α, and β. When the vehicles are inserted with a lower probability $\alpha < 1/2$ and removed with a much higher probability $\beta > \alpha$, the system will result a low-density state with a fewer vehicles in the system. The vehicles will move forward with a free flow speed. Contrarily, when the vehicles are removed slowly $\beta < 1/2$ and inserted with a much higher probability $\alpha > \beta$, the system is in the high-density region with a higher number of vehicles in the system. The speed of the vehicles will be reduced. When the vehicles are inserted and removed sufficiently rapidly, $\alpha > 1/2$ and $\beta > 1/2$, the system is in the maximum-current region where the system has its maximum flow.

ASEP can have different update rules, random sequential, sublattice-parallel, ordered sequential and parallel [43]. In this study we only focused on ASEP with random sequential updates and parallel updates.

2.1.2.1 Random sequential updates

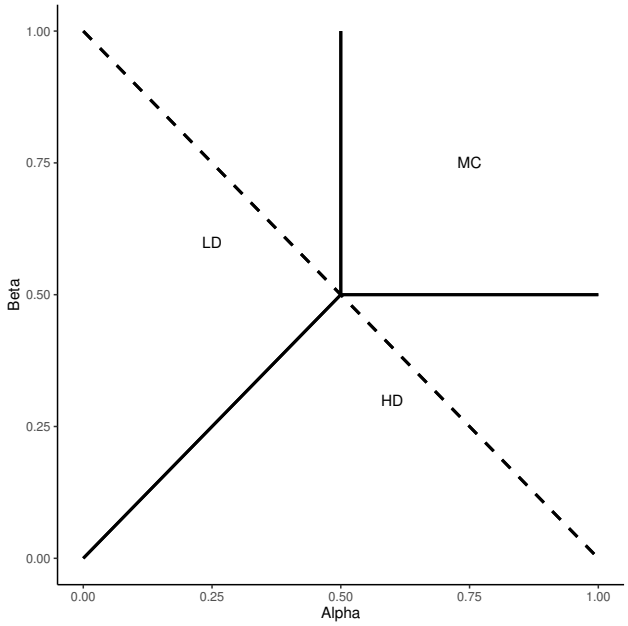


FIGURE 2.3: Exact phase diagram of Asymmetric Simple Exclusion Process (ASEP) with random sequential dynamics for $p = 1$. The density profile becomes flat on the dotted line $\alpha + \beta = 1$.

When the system has random sequential updates, in every iteration a vehicles will be picked randomly from a cell i, where $2 \leq i \leq L - 1$ and move one cell forward with probability p, if the $(i + 1)^{\text{th}}$ cell is not occupied. If $i = 1$ and the cell is occupied, then the vehicle will move to the cell $i = 2$. If $i = 1$ and the cell is not occupied, then a vehicle will be inserted to the first cell with probability α. If $i = L$ and the cell is occupied by a vehicle, then the vehicle will be removed from the system with probability β. $p < 1$ simply result in a re-scaling of time. Therefore, the phase diagram of the ASEP with random sequential updates does not significantly depend on p [59]. Therefore, throughout the thesis, when we use ASEP with random sequential updates, we consider $p = 1$. Figure 2.3 illustrates the exact phase diagram for $p = 1$.

Table 2.1: Exact solutions of an ASEP with random sequential updates with $p = 1$ [43]. α and β are the inflow probability and outflow probability respectively. ρ_1, $\rho_{L/2}$, and ρ_L represents the average density at left-boundary, bulk and right-boundary respectively.

Phase	**Condition**	ρ_1	$\rho_{L/2}$	ρ_L
Low-density	$\alpha < \beta, \quad \alpha < \frac{1}{2}$	α	α	$\frac{\alpha(1-\alpha)}{\beta}$
High-density	$\beta < \alpha, \quad \beta < \frac{1}{2}$	$1 - \frac{\beta(1-\beta)}{\alpha}$	$(1-\beta)$	$(1-\beta)$
Maximum-current	$\alpha > \frac{1}{2}, \quad \beta > \frac{1}{2}$	$1 - \frac{1}{4\alpha}$	$\frac{1}{2}$	$\frac{1}{4\beta}$

With random sequential updates, the system has a simple structure shows no correlation between cells. Hence the simple mean-field solutions are exact, and the Table 2.1 summarises the exact mean-field solutions for bulk density $\rho_{L/2}$ and boundary densities, ρ_1 and ρ_L of an ASEP with random sequential updates and $p = 1$ [43]. On the dotted line $\alpha + \beta = 1$, the density profile is flat. Therefore, the density profile fails to be injective in $[0, 1]^2$, if $\alpha + \beta = 1$ [60]. More information on the exact stationary state of an ASEP with sequential updates can be found in [43].

2.1.2.2 Parallel updates

When the system has fully parallel updates, the rules for hopping, insertion, and removal are applied simultaneously to all sites of the whole link. The parallel update usually produces the strongest correlations between the occupations of neighbouring cells.

Figure 2.4 illustrates the exact phase diagram for the ASEP with parallel dynamics for $p = 0.5$. Phase boundaries are at the critical probability, $\alpha_c, \beta_c = 1-\sqrt{1-p}$, where the transition is discontinuous in the density. The broken line given by, $(1-\alpha)(1-\beta) = 1-p$, the mean field solution is exact and the density profile is flat. For different combination of α and β ASEP has five different density regimes; LD1 ($\alpha < \beta < 1-\sqrt{1-p}$), LD2 ($\alpha < 1-\sqrt{1-p} < \beta$), HD1 ($\beta < \alpha < 1-\sqrt{1-p}$), HD2 ($\beta < 1-\sqrt{1-p} < \alpha$) and MC ($\alpha, \beta > 1-\sqrt{1-p}$). In LD1 regime, the density profile shows a purely exponential decay at the right boundary and in the LD2 regime, exponential decay is modulated by a power law. In HD1 regime, the density profile shows a purely exponential decay at the left boundary and in the HD2 regime, again the exponential decay is modulated by a power law. See [58, 61] for more details about an ASEP with parallel dynamics.

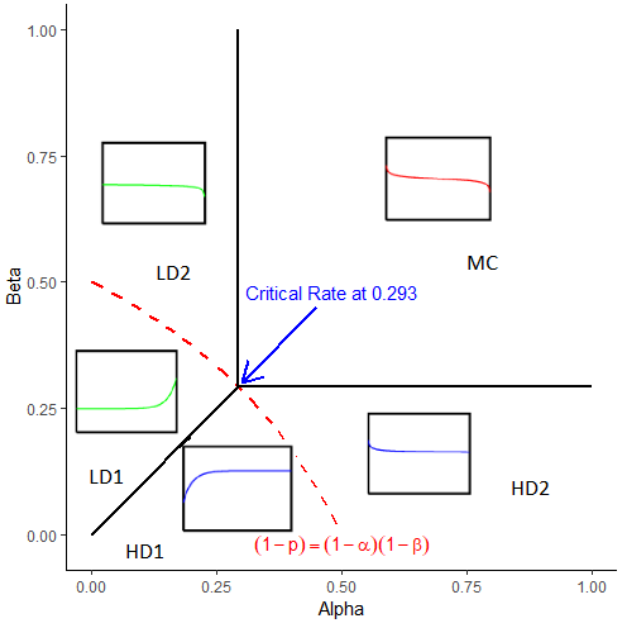


FIGURE 2.4: Exact phase diagram of Asymmetric Simple Exclusion Process (ASEP) with stochastic parallel dynamics for $p = 0.5$. The broken line is corresponds to $(1-\alpha)(1-\beta) = 1-p$, where the density profile is flat [58].

Exact asymptotic for the density profile of an ASEP with parallel updates and open boundary conditions can be found in [61]. The boundary densities were determined by the current J_L as Equation 4.24 in [61]:

$$\rho_1(\alpha,\beta,p) = 1 - \frac{1}{\alpha}J_L(\alpha,\beta,p)$$
$$\rho_L(\alpha,\beta,p) = \frac{1}{\beta}J_L(\alpha,\beta,p)$$

where,

$$J_L(\alpha,\beta,p) \underset{L\to\infty}{\sim} \begin{cases} \frac{1-\sqrt{1-p}}{2}, & \text{Maximum-current} \\ \frac{\alpha(p-\alpha)}{p-\alpha^2}, & \text{Low-density} \\ \frac{\beta(p-\beta)}{p-\beta^2}, & \text{High-density} \end{cases}$$

as defined in Equation 8.21, 8.23 and 8.24, [61]. Therefore, exact solutions for bulk density $\rho_{L/2}$ and boundary densities, ρ_1 and ρ_L of an ASEP with parallel updates can be derived as follows.

$$\rho_1(\alpha,\beta,p) \underset{L\to\infty}{\sim} \begin{cases} 1-\frac{1-\sqrt{1-p}}{2\alpha}, & \text{Maximum-current} \\ \frac{\alpha(1-\alpha)}{p-\alpha^2}, & \text{Low-density} \\ 1-\frac{\beta(p-\beta)}{\alpha(p-\beta^2)}, & \text{High-density} \end{cases}$$

$$\rho_L(\alpha,\beta,p) \underset{L\to\infty}{\sim} \begin{cases} \frac{1-\sqrt{1-p}}{2\beta}, & \text{Maximum-current} \\ \frac{\alpha(p-\alpha)}{\beta(p-\alpha^2)}, & \text{Low-density} \\ \frac{p-\beta}{p-\beta^2}, & \text{High-density} \end{cases}$$

$$\rho_{L/2}(\alpha,\beta,p) \underset{L\to\infty}{\sim} \begin{cases} \frac{1}{2}, & \text{Maximum-current} \\ \frac{\alpha(1-\alpha)}{p-\alpha^2}, & \text{Low-density} \\ \frac{p-\beta}{p-\beta^2}, & \text{High-density} \end{cases}$$

2.1.3 CEASAR

CEASAR was jointly developed by VicRoads, Monash University and the University of Melbourne. CEASAR is a mesoscopic network model which simulates individual movements of vehicles, while deliberately ignoring detailed geometry of roads and intersections. The model is agent-based, with individual vehicles moving according to behavioural rules which are influenced by the state of the surrounding traffic and randomness. CEASAR incorporates various traffic signal systems such as fixed cycle, self-organising signals, and realistic adaptive traffic signal including *Sydney Coordinated Adaptive Traffic System* (SCATS). Moreover, importantly CEASAR incorporates tram and bus priority processes, pedestrian signals and train level crossings.

CEASAR demands each vehicle makes a random decision about which link it wants to turn into at the approaching intersection with some turning probability. Such probabilities can be obtained from SCATS detector loop data or turning movement surveys if applied on a realistic network. Alternatively, CEASAR can be used to compute the fastest route according to real-time network traffic conditions, given the destination of a vehicle.

CEASAR has been used to model several practical scenarios, with the collaboration of VicRoads, and a number of theoretical studies. de Gier et al investigate how adaptive signal strategies can improve urban traffic flow using CEASAR [2]. Furthermore, Zhang et al studied the different types of adaptive traffic signal systems on Macroscopic Fundamental Diagrams (MFDs) of arterial road networks using CEASAR as the simulation tool [62]. CEASAR has also been used to identify the impact of disruptions on road networks, and the recovery process after the disruption is removed from the system [63]. A comprehensive description of the model can be found in [2, 62].

Calibration of traffic flow simulation models 3

3.1 Overview

Traffic simulation is an effective approach for modelling a collection of complex traffic scenarios relatively quickly and inexpensively. These simulation results can be utilised to identify the most promising scenarios, which can then be studied in greater detail with real-world trials. Due to the complexity of transport systems, refining a simulation model to accurately represent the real network is challenging. This refining process consists of three main phases: model verification, calibration and validation. The verification phase involves with the initial stage of investigation whether the conceptual framework of the model correctly determined the desired simulation model. The calibration phase involves with the estimation of input parameters. The empirical data can be used to refine the simulation model to adjust these input parameters, elaborating the behaviour of the real transport system. Finally, in the validation phase, simulation results are compared against the observed data to confirm whether the model accurately resembles the real transport system.

Traffic simulators require a large number of input parameters which should be defined explicitly. The classical approach to model calibration is to estimate the true parameters which explain the natural description of the transport system. Generally, the true parameters are estimated when the target discrepancy of performance measures are minimal. Standard measures of the performance used in the calibration of traffic models are the number of vehicles on a road

section [64–66], turning flows at intersections [7], link flows [64], maximum queue length [5, 64], speed [6, 67], travel time [64, 68] and origin-destination data [5]. However, with the recent development of computational methods established in Bayesian statistics: ABC, we do not have to rely on estimating the true parameters any more. Instead, we can approximate the likelihood of parameters and choose the parameters which are consistence with the real-world data. One can sensibly choose the correct parameters based on prior knowledge, not depending on any particular choice of parameters. ABC has identified as a flexible and comprehensible alternative to traditional model calibration methods which avoids the exact likelihood estimation by comparing summary statistics with the observed summary statistics.

The notion of efficient calibration techniques for complex traffic simulator models have been extensively studied over the past years and are focused on both analytical and simulation-based calibration techniques [66]. The common practice of calibration procedure is trial and error approaches due to the complexity associated with large scale traffic simulations and large amounts of unknown parameters. Generally, the process is to estimate the optimal values for sensitive parameters. Three main estimation techniques have been used for traffic simulator model calibration: least square estimation, maximum likelihood estimation and Bayesian methods [7]. Arguably the most extensively applied technique in calibration studies is the least-squares estimation technique.

Discrepancy measures like mean relative error [5, 6, 69], root mean square error [6, 64, 67, 70] and mean normalised error [64], correlation coefficient and Theil's inequality coefficient [6, 67] can be found in past researches related to calibrating traffic simulation models. Theil's inequality coefficient is one of the popular discrepancy measures, and it measures the relative error of the simulated output and is bounded $0 \leq U \leq 1$. $U = 0$ implies a perfect fit between the observed and simulated measurements and $U = 1$ implies the worst possible fit. The GEH statistic is another widely accepted experimental measure which accounts for the difference between simulated data and observed data [71]. Even though its mathematical structure is similar to a chi-squared test, GEH statistic is not a genuine statistical test. Instead, it is an empirical measurement that incorporates both relative and absolute deviations of the simulated value. Another model validation technique in parameter estimation is visual validation where the observed and simulated data were plotted to investigate any deviations from the real traffic network [4, 5].

Another method to evaluate the statistical significance of the simulated values is by using classical hypothesis testing for mean comparisons [65, 72]. Parametric tests such as two-sample t–test, F-test. Furthermore, non-parametric tests like Kolmogorov-Smirnov [65, 72] and Mann-Whitney Wilcoxon Rank

Sum test [65] were also used to test the goodness-of-fit of the simulated data with the real data. These two-sample tests assume the two sets of simulated and observed data are *independent and identically distributed* (iid). Therefore, these tests should be performed separately for each time-space measurement point, where the iid assumption holds, at least approximately. However, the measures of traffic networks depend strongly on time and are auto correlated. Rao et al. [72] have suggested a two-dimensional testing procedure involving a two-sample Kolmogorov–Smirnov (K-S) test and one-sample t-test for every observed and simulated measurement.

Rao and Owen [3] argued that a proper validation procedure should account for the auto correlated nature and the non-stationary property of simulated and empirical data accurately. However, due to an inability to find methods that can accurately represent the non-stationarity of a process, the authors claimed that a thorough consideration of non-stationary behaviour is not feasible. They suggested that employing a particular form of non-stationarity would yield reasonable estimates about the process. Therefore, by assuming the variance of simulated and empirical series are time-independent and are differenced stationary, they proposed an error analysis approach using univariate non-seasonal Autoregressive Integrated Moving Average (ARIMA) modelling, to validate traffic simulation models. They modelled the relative error component of the simulated and the empirical data using ARIMA process and checked the adequacy of the traffic simulator by considering the adequacy of the estimated model.

The maximum likelihood estimation methods which maximise the log-likelihood had also been applied in transport model calibration problems [42]. However, error terms in traffic data are not independent and are auto correlated. Therefore, the significant assumptions in maximum likelihood estimations are violated in traffic data. Hoogendoorn and Hoogendoorn suggested an approach to dealing with auto correlation; if the residual auto correlation coefficient of the maximum likelihood estimator significantly differs from zero, it is necessary to transform the model by eliminating the auto correlation [42]. Bayesian approaches, where estimating posterior probabilities of parameters based on the prior beliefs have been applied to calibrate traffic simulations such as CORSIM in a simple to moderate complex networks [9, 73].

Once the sensitivity parameters, reliable measures of performance and discrepancy measures have been identified, the traditional approach to calibration involves choosing an optimisation algorithm capable of solving the calibration problem. The majority of the simulated based optimisation algorithms are general-purpose algorithms which are not customised for a specific simulation scenario. These optimisation algorithms treat the simulator as a black box [10]. Widely accepted algorithms applied for traffic simulator calibrations

are, *Simultaneous Perturbation Stochastic Approximation* (SPSA)[66, 74, 75], *Genetic Algorithm* (GA) [65, 76, 77], Particle Filters and Kalman Filters [10].

3.2 ABC for model calibration

The problem of calibrating parameters for complex models has a long history, and there are a wide variety of tools available. As an alternative to classical parameter estimation techniques, ABC has gained attention in recent years as a flexible and comprehensible way of solving model calibration problems. The ABC technique initially became popular within population genetics, developed by Pritchard, Seielstad, Perez-Lezuan and Feldman in 1999 [78]. The ABC approach simulates data from a given model for different parameter values, and compares the output with the observed data. Parameter values which produce data closer to the observed data are retained to form an approximate posterior sample. In Bayesian inference, the posterior distribution contains the prior knowledge about the model parameters obtained by fitting the desired data generating process. More specifically, the prior knowledge about the model parameters, expressed through the prior distribution, are updated by empirical data through the model and estimate the posterior distribution which can be considered as the likelihood of the parameters given empirical data.

ABC were developed to estimate posterior distributions when likelihood functions are computationally intractable. ABC avoids exact likelihood estimation by comparing simulated data sets with the observed data. More specifically, a set of parameter points is first sampled from the prior distribution (i.e. based on the prior knowledge). Then, data is simulated from a data generating process for a given set of parameters. If the generated data is too different from the observed data, the sampled parameter values discarded. In precise terms, the parameters are accepted with a tolerance rather than demanding them to be unique. In order to compromise the curse of high dimensionality, ABC algorithms generally reduce the dimension using summary statistics. The parameters are accepted when simulated summaries are close to the observed summaries [79].

There are many forms of ABC algorithms with different level of complexity. The simplest among these is the ABC rejection sampling algorithm [78]. The ABC accept/reject algorithm was first introduced by Tavaré et al. [80] and extended by Pritchard et al. [81]. Beaumont et al. [82] introduced the first improvement of the ABC rejection sampling algorithm to correct the discrepancy between the simulated and the observed statistics by using local regression techniques.

Simple simulation-based rejection ABC algorithms are inefficient when the data and the summary statistic are of high dimension, which leads to a high rejection rate. In order to avoid the high rejection rate, *Markov Chain Monte Carlo ABC* (MCMC-ABC) has been introduced where more simulations are generated in the regions of high posterior probability [83]. Even though MCMC-ABC improves the acceptance rate significantly, it is still inefficient when the sampling and target distributions are mismatched. In order to alleviate the inefficiency of existing MCMC-ABC method, Sisson et al. introduced a novel approach, adapting the *Sequential Monte Carlo* (SMC) methods to the standard ABC algorithm [84]. Drovandi and Pettit extended the approach discussed in [84] to develop a robust algorithm, where the sequence of discrepancies is selected adaptively by the algorithm [85]. They named this method the *SMC-ABC replenishment* (RABC) algorithm.

Nott et al. [86] further improved the ABC algorithm by introducing a marginal adjustment method, which improves the estimate of the ABC joint posterior distribution obtained using regression-adjustment ABC method. They suggested that the proposed strategy allows applications with moderate to high dimensionality problems. More recently, Li et al. introduced a new approach, considering the marginal adjustment strategy proposed in [86] and the asymptotic normality of the Bayesian posterior. They estimate each bivariate distribution separately and then combine them to obtain the joint posterior distribution [87]. They argued that the new approach could be effectively implemented in high-dimension problems.

3.2.1 ABC methods

In Bayesian inference, we obtain the posterior distribution of input parameters declared through the prior distribution, according to the prior belief, are updated with the new evidence from observed data through the likelihood function of the model. Using Bayes' Theorem, the resulting posterior distribution, $p(\theta \mid y_{obs})$ can be defined by Equation 3.1 [88].

$$p(\theta \mid y_{obs}) = \frac{p(y_{obs} \mid \theta)p(\theta)}{p(y_{obs})} \approx p(y_{obs} \mid \theta)p(\theta) \tag{3.1}$$

where θ is the input parameter, y_{obs} is the observed data, $p(\theta)$ is the prior distribution for unknown parameter, $p(y_{obs} \mid \theta)$ is a likelihood function, $p(y_{obs})$ is the marginal likelihood. However, with the complex models, the closed-form of the $p(\theta \mid y_{obs})$ is not available. With the exact Bayesian inference, a common approach is to use simulations methods such as *Markov Chain Monte*

Carlo (MCMC) and SMC to calculate the necessary integrals [88]. However, if the exact likelihood function, $p(y_{obs} \mid \theta)$ is intractable, as in most of the traffic simulation models, implementation of standard Bayesian simulations are impossible. Among many likelihood-free Bayesian posterior approximation methods, ABC is identified as an effective and easily accessible method for large scale model calibration methods. The following sections describes the basic ABC methods that can be incorporate in traffic simulator model calibrations.

3.2.1.1 Standard rejection ABC

If we simulate from the joint prior distribution of parameters and data $p(\theta)p(y_{obs} \mid \theta)$, an exact match is possible between simulated data y_{sim} and observed data y_{obs}, when the simulated data is discrete. If the set of simulated parameter values θ^* gives the exact match of simulated and observed data ($y_{sim} = y_{obs}$), then θ^* is a sample from the posterior distribution.

$$p(\theta^* \mid y_{obs}) \propto p(y_{obs} \mid \theta^*)p(\theta^*) \tag{3.2}$$

However, in most applications, obtaining an exact match is impossible. Therefore, in standard rejection ABC the exact match is weakened to allow agreement to within a small distance $\epsilon > 0$ to the observed data. Hence the standard ABC rejection sampling algorithm is as follows:

(1) Sample (θ^i), i=1,2,...,n, from prior $p(\theta)$;

(2) Simulate psuedo-data y^i_{sim}, , i=1,2,...,n, from $p(y \mid \theta^i)$;

(3) Select (θ^i) such that: $d\{y_{obs}, y^i_{sim}\} \leq \epsilon$

Where, $d\{\cdot\}$ is a metric and the tolerance ϵ is arbitrarily small.

Therefore the approximate posterior is proportional to,

$$\int p(y^i_{sim} \mid \theta^i)p(\theta^i)I(d\{y_{obs} - y^i_{sim}\} \leq \epsilon)dy_{sim} \tag{3.3}$$

Where $I(\cdot)$ is the indicator function. Generally, the observed data y_{obs} is of high-dimension, and hence the rejection rate can be very high for small ϵ. Therefore, in order to improve the efficiency of the algorithm, y_{obs} can be replaced by some lower-dimensional summary statistic $S(y_{obs})$. If the summary

statistic is sufficient, then $p(\theta \mid y_{obs}) = p(\theta \mid S(y_{obs})$. However, for intractable likelihood functions, it is not typically possible to find low-dimensional sufficient statistics. Moreover, the choice of summary statistics is highly problem-specific and have to rely on low dimensional insufficient summaries [87]. There are several methods to select low-dimensional insufficient summaries as discussed in [89]. In this chapter we will discuss some summaries for traffic model calibration in Section 3.3. When we replace the observed data with some lower-dimensional summary statistic $S(y_{obs})$, the third step in the standard reject ABC algorithm becomes:

(3) Select (θ^i) such that: $d\{S(y_{obs}), S(y^i_{sim})\} \leq \epsilon$

Where, $S(\cdot)$ is a summary statistic $d\{\cdot\}$ is a metric and the tolerance ϵ is arbitrarily small.

The approximate posterior then becomes,

$$\int p(S(y^i_{sim}) \mid \theta^i) p(\theta^i) I(d\{S(y_{obs}), S(y^i_{sim})\} \leq \epsilon) dy_{sim} \tag{3.4}$$

Beaumont et al [82] improved the standard rejection ABC algorithm in two ways. For the first innovation, indicator is generalised to $K_h(d\{S(y_{obs}), S(y^i_{sim})\})$, where $K_\epsilon(\cdot)$ is a standard smoothing kernel. Then the approximate posterior distribution given summary statistics, $S(y_{obs})$ is:

$$p(\theta \mid S(y_{obs})) \propto \int p(S(y^i_{sim}) \mid \theta^i) p(\theta^i) K_h(d\{S(y_{obs}), S(y^i_{sim})\}) y_{sim} \tag{3.5}$$

For the second innovation, regression adjustments were used to weaken the effect of the discrepancy between $S(y_{sim})$ and $S(y_{obs})$ as summarised in the following section.

3.2.1.2 Rejection ABC with regression adjustments

For a parameter $\theta^i = (\theta^1, \ldots, \theta^m)^T$ of dimension m and the corresponding summary statistic $s = S(y_{sim}) = (s^1, \ldots, s^d)^T$ of dimension d, the weighted linear regression model is:

$$\theta^i = \alpha + \beta^T (S(y_{obs} - S(y^i_{sim}))) + \nu_i \tag{3.6}$$

Where ν_i's are the errors which are considered to be iid, β is a $d \times m$ matrix of regression coefficients and α is a $m \times 1$ vector of constants. Instead of fitting the model globally, a local linear regression model is fitted surrounding the $S(y_{obs})$. The posterior distribution can be constructed from the model 3.6 when $S(y^i_{sim}) = S(y_{obs})$. That means $(\alpha + \nu_1, \ldots, \alpha + \nu_n)$ is a sample from the posterior distribution $p(\theta^i \mid S(y_{obs}))$. The weighted least-squares estimates (α, β) minimises the model 3.6,

$$\sum_{i=1}^{n}(d\{\theta^i - (\alpha + \beta^T(S(y_{obs}) - S(y^i_{sim})))\})^2 K_h(d\{S(y_{obs}) - S(y^i_{sim})\}) \quad (3.7)$$

The linear regression adjusted vector denoting the resulting empirical residuals (ν_i), then the regression-adjusted vector becomes,

$$\theta^i_a = \theta^i - \widehat{\beta}^T(S(y_{obs} - S(y^i_{sim})) \approx \widehat{\alpha} + \nu_i \quad (3.8)$$

3.2.1.3 Asymptotic properties of ABC

The practical implication of standard ABC algorithms highly depends on the initial adjustments of the algorithm. It requires to correctly specify the metric $d\cdot$, summary statistics $S(\cdot)$ and the tolerance ϵ. Depending on the choices of these quantities, we can observe significant changes in the posterior approximation.

The comparison between simulated data and the empirical data is computationally feasible only when the dimension is reduced using a set of summary statistics. The dimension reduction by summary statistics, is crucial in ensuring better performances of the ABC algorithm. Ideally, the posterior approximation is better if when the summary statistics are sufficient. However, in practice for complex intractable likelihood functions, it is not achievable to obtain low-dimensional sufficient statistics. Moreover, the choice of summary statistics is highly problem-specific and have to rely on low dimensional insufficient summaries which capture enough information from the observed and simulated data. Several methods to select low-dimensional insufficient summaries are discussed in [89]. Furthermore, Fernhead suggested that choosing the number of summary statistics to be close to, or equal to, the number of parameters is required to achieve a good approximation of the posterior [90].

The choice of the metric also affects the performances of the ABC algorithm as it affects the spread as well as correlations of the accepted parameters. A scaling weights should be used in order to minimises the variability in the

distance calculation. When the Euclidean distance is used, the common practice is to standardised each summary statistic with a robust estimate of the standard deviation [91].

The specification of the tolerance also highly affects the approximation of the posterior and it depends on the considered ABC approach. The tolerance rate of a good ABC algorithm will increase as n increases. Furthermore, Scaling of the bandwidth, $\epsilon_n = O(1/\sqrt{n})$ is optimal when the regression adjustments not being used. However, the asymptotic results for ABC with regression adjustment are stronger. The ABC posterior and its mean can have the same asymptotic as the true ABC posterior and mean given the summaries, as n decreasing more slowly than $1/\sqrt{n}$. Fernhead argued that these strong results suggest that regression adjustment should be generally applied [90].

Frazier et al. studied the large sample properties of both posterior distributions and posterior means obtained from approximate Bayesian computation algorithms [92]. They characterise the rate of posterior concentration under mild regularity conditions on the underlying summary statistics. They argued that the limiting shape of the posterior distribution depends on both the rates which summary statistics converge to observed summary statistics and the tolerance rate used to accept parameter draws approaching zero. Tolerance declining to zero is not a rigid requirement for Bayesian consistency. However, the asymptotic normality of the resulting posterior does require the tolerance rate to decline to zero.

The most crucial finding of their study is that the existence of a binding function which holds identifiable conditions for summary statistics, to have the posterior concentration around the true parameter. If the summary statistics hold the law of large number and the existence of a binding function, $s(y_{sim})_{n,\theta} \to b(\theta)$, then the ABC posterior concentrates around the true parameter providing the bandwidth tends to zero, $\epsilon \to 0$ as $n \to \infty$.

They argued that the choice of the tolerance rate decides the desired limiting shape of the posterior. Assuming the summary statistics satisfy the *Central Limit Theorem* (CLT) at rate of $\frac{1}{\sqrt{n}}$,

- If $\sqrt{n}\epsilon_n \to 0$, then the ABC posterior will be approximately normal.
- If $\sqrt{n}\epsilon_n \to c > 0$, then it is hard to predict the limiting shape of the posterior.
- If $\sqrt{n}\epsilon_n \to \infty$, then the limiting shape will be determined by the ABC kernel.

The choice of kernel asymptotically has little impact as it only affects the shape of the ABC posterior [90]. For a detailed review on asymptotic and theoretical properties of ABC, see [90].

3.3 Applying ABC for traffic model calibration

With the recent development of *Intelligent Transportation System* (ITS) which provide more coordinated, and smarter transport network, we could get access to a massive amount of information related traffic volumes and densities of each link where a loop detector is installed. In reality, we could easily incorporate this information in traffic simulation model calibration processes. For traffic simulator model calibration using ABC, the only requirement is the availability of some informative summary statistics of the observed data from the desired network. Then the ABC algorithms will compare these observed summary statistics with the simulated summary statistics to approximate the likelihood of the model parameters.

The empirical data actually available to calibrate a traffic simulation model would typically take the form of a collection of distinct time series $(y_t^1)_{t=1}^T, ..., (y_t^n)_{t=1}^T$, where S_t^i is a vector of observed quantities at time t on day i.

The t index would typically iterate over a short time scale, e.g. seconds on minutes, and so y_t^i & y_{t+1}^i could potentially be strongly correlated. By contrast $(y_t^i)_{t=1}^T$ & $(y_t^j)_{t=1}^T$ can reasonably be assumed to be independent. Moreover, if each time series $(y_t^i)_{t=1}^T$ is observed during the same period of the day (e.g. morning peak hour) and on comparable day, e.g. typical weekday, then it is also reasonable to assure the $(y_t^i)_{t=1}^T$ are identically distributed. In this context, T corresponds to the duration of a fixed time window (e.g. 2 hours for morning peak hour), and we are not interested in T taking arbitrarily large values. By contrast, n is the number of days on which samples are taken and we would choose it to be as large as possible, subject practical constraints such as cost etc.

To model this empirical data collection process via simulation, one generates n independent runs of some stochastic simulation model, $(Z_t^1)_{t=1}^T, ..., (Z_t^n)_{t=1}^T$. The time series generated by the simulation model typically forms a Markov chain.

To study the feasibility of using ABC to calibrate a traffic simulation model, we therefore generate synthetic empirical data in the format described above. Specifically, the "observed" data we use in our implementations of ABC correspond to the realization $(y_t^1)_{t=1}^T, ..., (y_t^n)_{t=1}^T$ of n independent runs of a Markov chain with fixed parameters $\Theta = \theta_*$. E.g. in the ASEP simulations below $\Theta = (\alpha, \beta)$, where α is the input probability at the left boundary and β is the output probability at the right boundary.

In our implementations of ABC, to sample from the likelihood in step two we again generate data in the same format as above, using the same Markov chain simulator as used to generate the "observed" data, but with θ chosen randomly from a suitable prior distribution, rather than with θ fixed to its "true" value θ_*.

When actually implementing ABC, of course, we must summarise the data appropriately. In all cases below, our summary statistics are vectors of sample means, so that each coordinate of the summary is of the form

$$\hat{f}_n(Z) \coloneqq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T} f(Z_t^i) \tag{3.9}$$

where $Z \coloneqq \Big((Z_t^1)_{t=1}^T, ..., \Big((Z_t^n)_{t=1}^T\Big)$. From the independence of the set of time series $(Z_t^1)_{t=1}^T, ..., (Z_t^n)_{t=1}^T$, it follows that the law of large numbers and central limit theorem hold for $\hat{f}_n(Z)$ as $n \to \infty$, provided f has finite variance; this will be the case for all examples considered below. In particular, in the case that $(Z_t^i)_{t=1}^T$ is stationary, we have $\hat{f}_n(Z) \xrightarrow[n\to\infty]{} \mathbb{E}f(Z_1^i)$. If Z_t^i is an ergodic finite Markov chain, which is the case for all the examples we consider below, we can always ensure Z_t^i is as close to stationarity as desired by imposing a suitable burn in. We evaluate the feasibility of ABC in the traffic model calibration against two basic traffic models, ASEP and NaSch. Furthermore, we use ABC to estimate turning rates at a signalised intersection for 4×5 network. See Section 2.1.3.

3.3.1 Informative summary statistics for ABC

The macroscopic behaviour of the models we defined in Section 2.1.2 can be easily described by the different combinations of inflow probability α and outflow probability β when $p = 1$. That is, the densities at each cell ρ_i can be approximate by different values of α and β as summarised in Section 2.1.2.1

and 2.1.2.2. Additionally, we can easily reproduce the values of α and β based on the information related to cell densities ρ_i.

For ASEP with random sequential updates, it is known [43] that as $L \to \infty$:

$$\rho_1(\alpha,\beta) \sim \rho_u(\alpha,\beta) \coloneqq \begin{cases} \alpha, & \alpha \le 1/2, \alpha \le \beta \\ 1 - \dfrac{\beta(1-\beta)}{\alpha}, & \beta \le 1/2, \beta \le \alpha \\ 1 - \dfrac{1}{4\alpha} & \alpha, \beta \ge 1/2 \end{cases} \tag{3.10}$$

$$\rho_{L/2}(\alpha,\beta) \sim \rho_m(\alpha,\beta) \coloneqq \begin{cases} \alpha, & \alpha < 1/2, \alpha < \beta \\ 1-\beta, & \beta < 1/2, \beta < \alpha \\ \dfrac{1}{2} & \alpha, \beta \ge 1/2 \end{cases} \tag{3.11}$$

$$\rho_L(\alpha,\beta) \sim \rho_d(\alpha,\beta) \coloneqq \begin{cases} \dfrac{\alpha(1-\alpha)}{\beta}, & \alpha \le 1/2, \alpha \le \beta \\ 1-\beta, & \beta \le 1/2, \beta \le \alpha \\ \dfrac{1}{4\beta} & \alpha, \beta \ge 1/2 \end{cases} \tag{3.12}$$

Here, the subscripts “u “, “m“,“d“ refer to “upstream“, “mid-block“, and “downstream“. We note that $\rho_u(\alpha,\beta)$ and $\rho_d(\alpha,\beta)$ are continuous functions on $[0,1]^2$, whereas $\rho_m(\alpha,\beta)$ has a jump discontinuity along the (coexistence) line $\alpha = \beta$ with $0 \le \alpha,\ \beta < 1/2$.

Furthermore, we note that the map $S(\alpha,\beta) \coloneqq \Big(\rho_u(\alpha,\beta), \rho_d(\alpha,\beta)\Big)$ is injective on $[0,1]^2 \setminus \Big\{(\alpha,\alpha) \in [0,1]^2 : 0 \le \alpha < 1/2\Big\}$, but is two-to-one along the line $\alpha + \beta = 1$, where the density profile is known to be flat [43].

Therefore, for generic values of $(\alpha,\beta) \in [0,1]^2$, knowledge of $S(\alpha,\beta)$ suffices to recover (α,β). Moreover, if one can distinguish whether the system is in the high density or low density regime, then knowledge of $S(\alpha,\beta)$ uniquely identifies (α,β) for any $(\alpha,\beta) \in [0,1]^2$. This should be contrasted with the information available from knowledge of (ρ_u,ρ_m) or (ρ_m,ρ_d). For either of these choices, it is impossible to recover (α,β) for any point in the maximum current phase .

Simple asymptotic expressions (as $L \to \infty$) for the density profile are also known for ASEP with parallel updates, when $p = 1$:

$$\rho_1(\alpha,\beta) \sim \rho_u(\alpha,\beta) \coloneqq \begin{cases} \dfrac{\alpha-\beta+\alpha\beta}{\alpha(1+\beta)}, & \alpha \geq \beta \\ \dfrac{\alpha}{1+\alpha}, & \alpha \leq \beta \end{cases} \tag{3.13}$$

$$\rho_{L/2}(\alpha,\beta) \sim \rho_m(\alpha,\beta) \coloneqq \begin{cases} \dfrac{1}{1+\beta}, & \alpha > \beta \\ \dfrac{\alpha}{1+\alpha}, & \alpha < \beta \end{cases} \tag{3.14}$$

$$\rho_L(\alpha,\beta) \sim \rho_d(\alpha,\beta) \coloneqq \begin{cases} \dfrac{1}{1+\beta}, & \alpha \geq \beta \\ \dfrac{\alpha}{\beta(1+\alpha)}, & \alpha \leq \beta \end{cases} \tag{3.15}$$

Again note that $\rho_u(\alpha,\beta)$ and $\rho_d(\alpha,\beta)$ are continuous on $[0,1]^2$ whereas $\rho_m(\alpha,\beta)$ has discontinuity across the (coexistence) line $\alpha=\beta$. Note also that when $p=1$ for parallel updates, the maximum current phase is the single point $(\alpha,\beta)=(1,1)$.

Therefore, we can consider the average densities at the left-boundary ρ_1, at the bulk $\rho_{L/2}$ and at the right-boundary ρ_L are sufficient statistics for ASEP when $p=1$ irrespective to the update rule. These quantities are not explicitly defined for the models we described in Section 2.1.1, 2.1.3 and ASEP with parallel update and $p \neq 1$. However, we assume the average densities at the left-boundary ρ_1, at the bulk ρ_L and at the right-boundary ρ_L are sufficient statistics for ASEP with parallel update and $p \neq 1$ and NaSch as well. For the model we described in Section 2.1.3, we assume that the average speed V_l and average flow f_l at each link are informative statistics.

3.3.2 Synthetic observed data for ABC

In order to calibrate the models we described in Sections 2.1.2 and 2.1.1, we produced some synthetic observed data: $y_{obs} = (\rho_{1,obs},\ \rho_{L/2,obs}, \rho_{L,obs})$, fixing $\theta=(\alpha,\beta)$ in all three density regions: High-density, Low-density and Maximum-current. For the ASEP with random sequential updates and ASEP with parallel updates and $p=1$ we easily derived the synthetic observed data from the quantities defined in Section 2.1.2.1 and 2.1.2.2.

However, for the ASEP with parallel updates and $p \neq 1$ and NaSch, we simulate the synthetic observed data: $y_{obs} = (\mathbb{E}[\rho_1], \mathbb{E}[\rho_{L/2}], \mathbb{E}[\rho_L])$ from the model again fixing $\theta = (\alpha, \beta)$ in all three density regions: High-density, Low-density and Maximum-current. For calibrate 4×5 lattice network, we simulate synthetic observed data, $y_{obs} = (\mathbb{E}[V_l], \mathbb{E}[f_l])$, fixing $\theta = (TR_{20,r}, TR_{20,l})$. $TR_{20,r}$ and $TR_{20,l}$ are the right turning rate and the left turning rate at node 20. For the detailed description refer to the Section 3.3.3.5.

3.3.3 Parameter estimation using ABC

We calibrated all the models defined in Chapter 2, with the observed data mentioned in Section 3.3.2 utilising the Standard Rejection ABC algorithm with and without regression adjustments. The pseudo-data y_{sim} were simulated from each model described in Chapter 2. We approximate the joint posterior distribution of inflow probability α and outflow probability β in each model defined in Sections 2.1.2 and 2.1.1. Furthermore, We approximate the joint posterior distribution of the $TR_{20,r}$ and $TR_{20,l}$ at the signalised intersection (i.e. at the node 20) for 4×5 network using the model described in Section 2.1.3. We consider the informative summary statistics defined in Section 3.3.1. The systems always start with an empty state.

3.3.3.1 ASEP: Random sequential updates

Particularly, we first consider the most basic transportation model, ASEP with random sequential updates. We only consider the average density at left-boundary and right-boundary $S(y_{sim}) = (\mathbb{E}[\rho_1], \mathbb{E}[\rho_L])$ as summary statistics in order to keep the dimension equal to the number of parameters. We fixed the true parameter vector generating the synthetic observed data at all three density regimes: low-density $(\alpha_{obs}, \beta_{obs}) = (0.2, 0.8)$, high-density $(\alpha_{obs}, \beta_{obs}) = (0.8, 0.2)$ and maximum-current $(\alpha_{obs}, \beta_{obs}) = (0.8, 0.8)$. The synthetic observed data for the all three density regimes: Low-density regime $y_{obs} = (0.2, 0.2)$; High-density regime $y_{obs} = (0.8, 0.8)$; Maximum-current regime $y_{obs} = (0.69, 0.31)$ derived from Table 2.1.

We consider a one-dimensional lattice of length $L = 500$. We take $N = 10,000$ iid draws from uniform priors where $\alpha \sim U(0, 1)$ and $\beta \sim U(0, 1)$ representing all three density regimes. We simulate the system for fixed samples of size $T \in \{100, 200, 300\}$ and for 1×10^8 time steps ensuring that the stationarity was always reached. We choose the tolerance rates to be $1/T^{0.7}$.

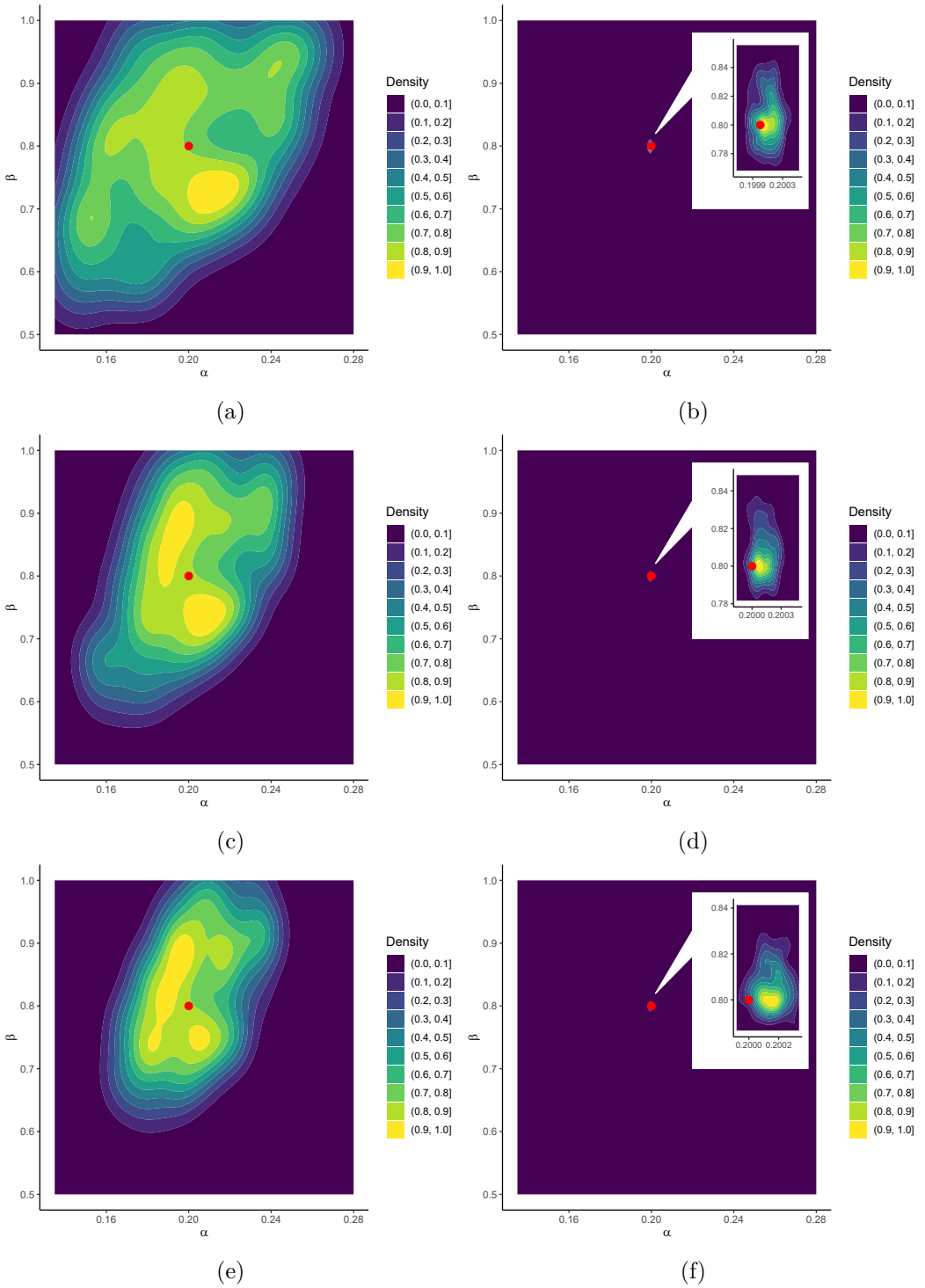


FIGURE 3.1: Joint posterior distribution of boundary parameters of an ASEP with random sequential updates when the observed parameters are in the low-density regime. Sub-figures (a), (c) and (e) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 100, 200 and 300 respectively. Sub-figures (b), (d) and (f) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 100, 200 and 300 respectively. The red dot represents the observed parameters, $\alpha = 0.2$, $\beta = 0.8$.

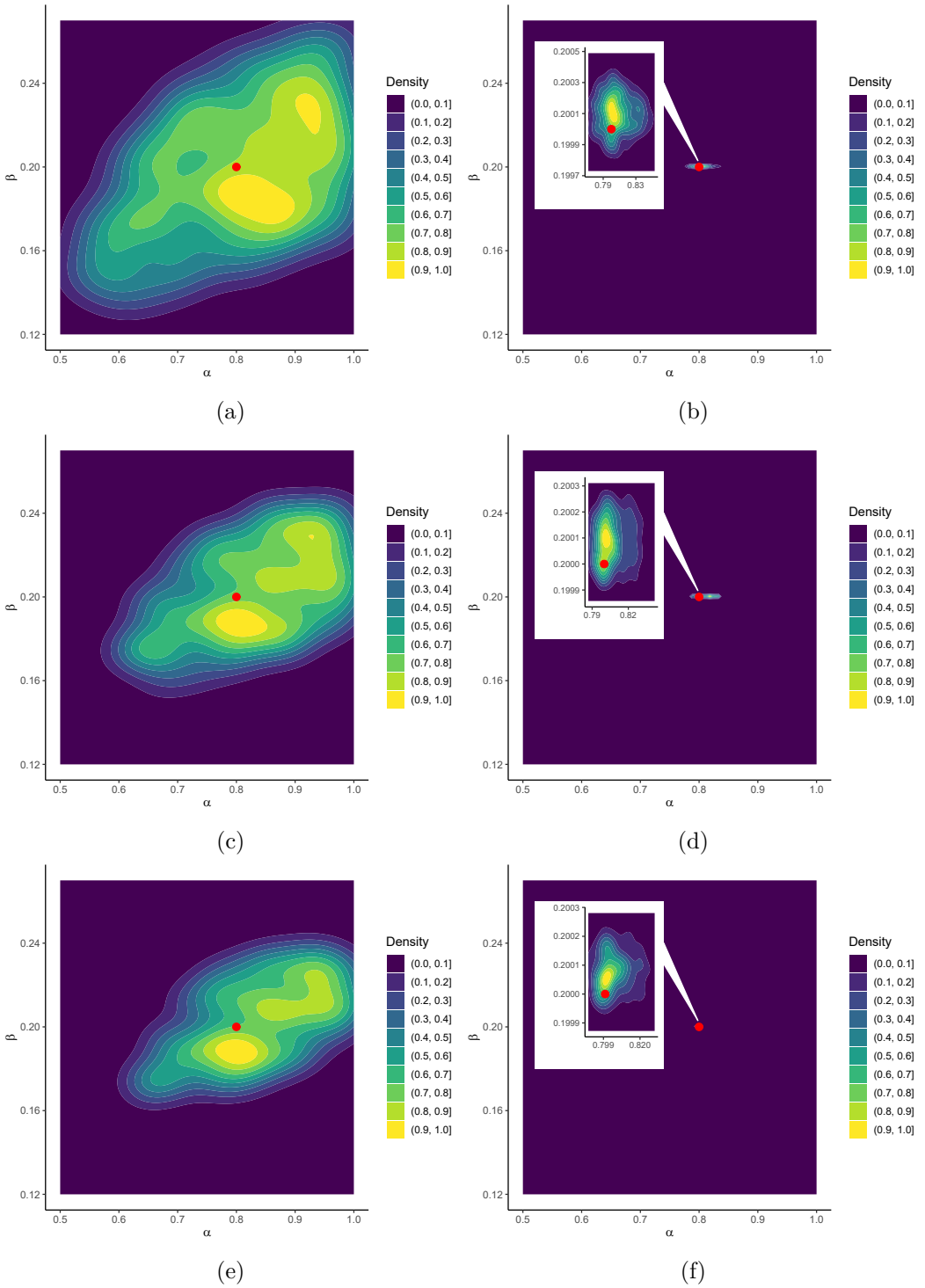


FIGURE 3.2: Joint posterior distribution of boundary parameters of an ASEP with random sequential updates when the observed parameters are in the high-density regime. Sub-figures (a), (c) and (e) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 100, 200 and 300 respectively. Sub-figures (b), (d) and (f) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 100, 200 and 300 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.2$.

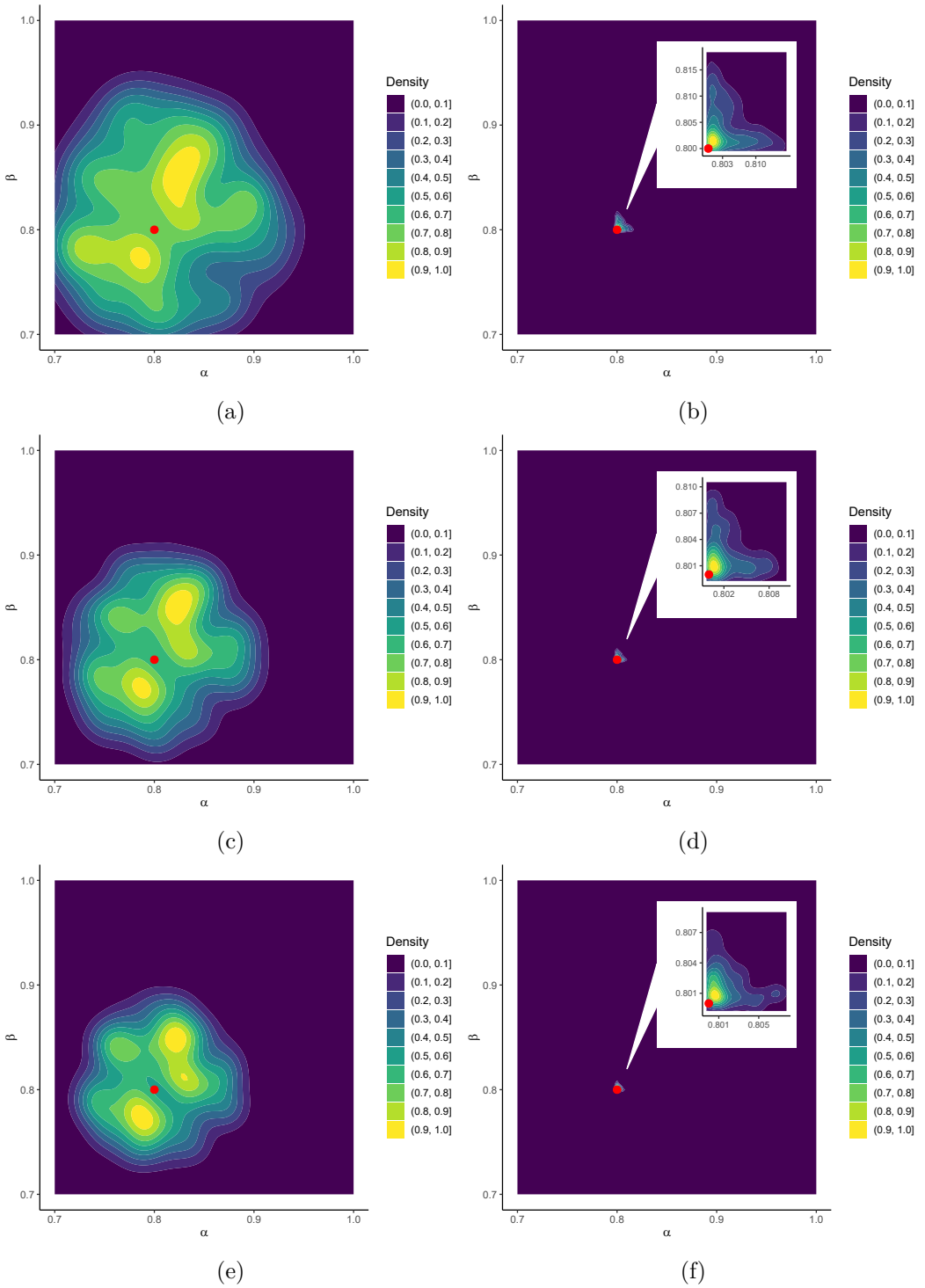


FIGURE 3.3: Joint posterior distribution of boundary parameters of an ASEP with random sequential updates when the observed parameters are in the maximum-current regime. Sub-figures (a), (c) and (e) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 100, 200 and 300 respectively. Sub-figures (b), (d) and (f) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 100, 200 and 300 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.8$.

Figures 3.1, 3.2 and 3.3 summarise the joint posterior distribution of boundary parameters of an ASEP with random sequential updates when the observed parameters are in the low-density regime, high-density regime and maximum-current regime for different sample sizes respectively. The concentration of the joint posterior distribution is increased with the sample size increases. Moreover, the regression adjustments incredibly increase the posterior mass around the true parameter vector. For larger sample sizes, the estimated joint posterior distribution with regression adjustments is identical to the observed parameters. Both the inflow probability α, and outflow probability β, can be estimated quite accurately with regression adjustments when the observed boundary parameters are in all three density regimes. However, when the true boundary parameters are in the high-density regime, the posterior distribution converges slowly.

3.3.3.2 ASEP: Parallel updates with $p = 1$

We consider the system described in Section 2.1.2 with a one-dimensional lattice of length $L = 200$ with parallel updates and $p = 1$. We only consider the average density at left-boundary and right-boundary $S(y_{sim}) = (\mathbb{E}[\rho_1], \mathbb{E}[\rho_L])$ as summary statistics in order to keep the dimension equal to the number of parameters. We fixed the true parameter vector generating the synthetic observed data at all three density regimes: low-density $(\alpha_{obs}, \beta_{obs}) = (0.2, 0.8)$, high-density $(\alpha_{obs}, \beta_{obs}) = (0.8, 0.2)$ and maximum-current $(\alpha_{obs}, \beta_{obs}) = (1, 1)$. The synthetic observed data for the all three density regimes: Low-density regime $y_{obs} = (0.167, 0.21)$; High-density regime $y_{obs} = (0.792, 0.833)$; Maximum-current regime $y_{obs} = (0.5, 0.5)$ derived from the quantities defined in Section 2.1.2.2.

We take $N = 1000$ iid draws from uniform priors where $\alpha \sim U(0, 1)$ and $\beta \sim U(0, 1)$ representing all three density regimes. We simulate the system for fixed samples of size $T \in \{30, 50\}$ and for 1×10^5 time steps ensuring that the stationarity was always reached. We choose the tolerance rates to be $1/T^{0.7}$.

Figures 3.4, 3.5 and 3.6 summarise the joint posterior distribution of α and β of the ASEP with parallel update and $p = 1$ when the observed parameters are in the low-density regime, high-density regime and maximum-current regime for different sample sizes respectively.

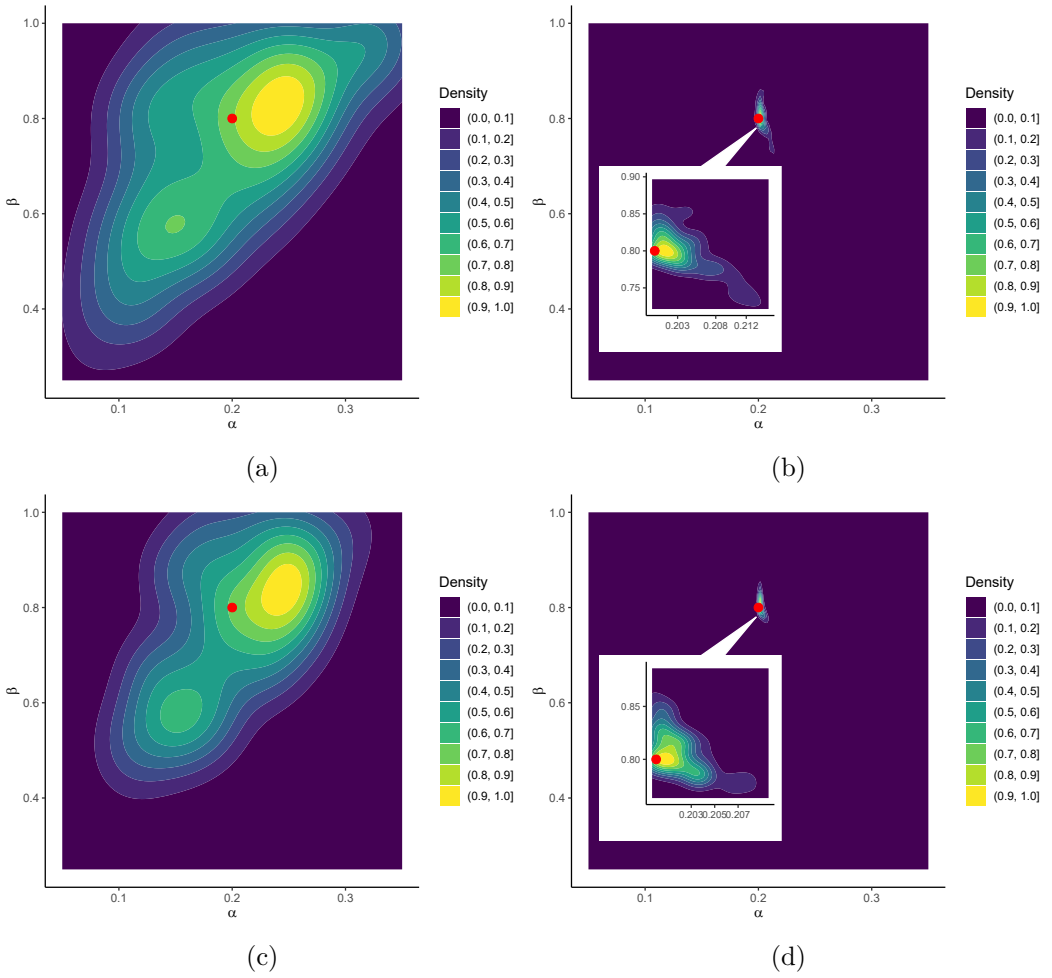


FIGURE 3.4: Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 1$ when the observed parameters are in the low-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.2$, $\beta = 0.8$.

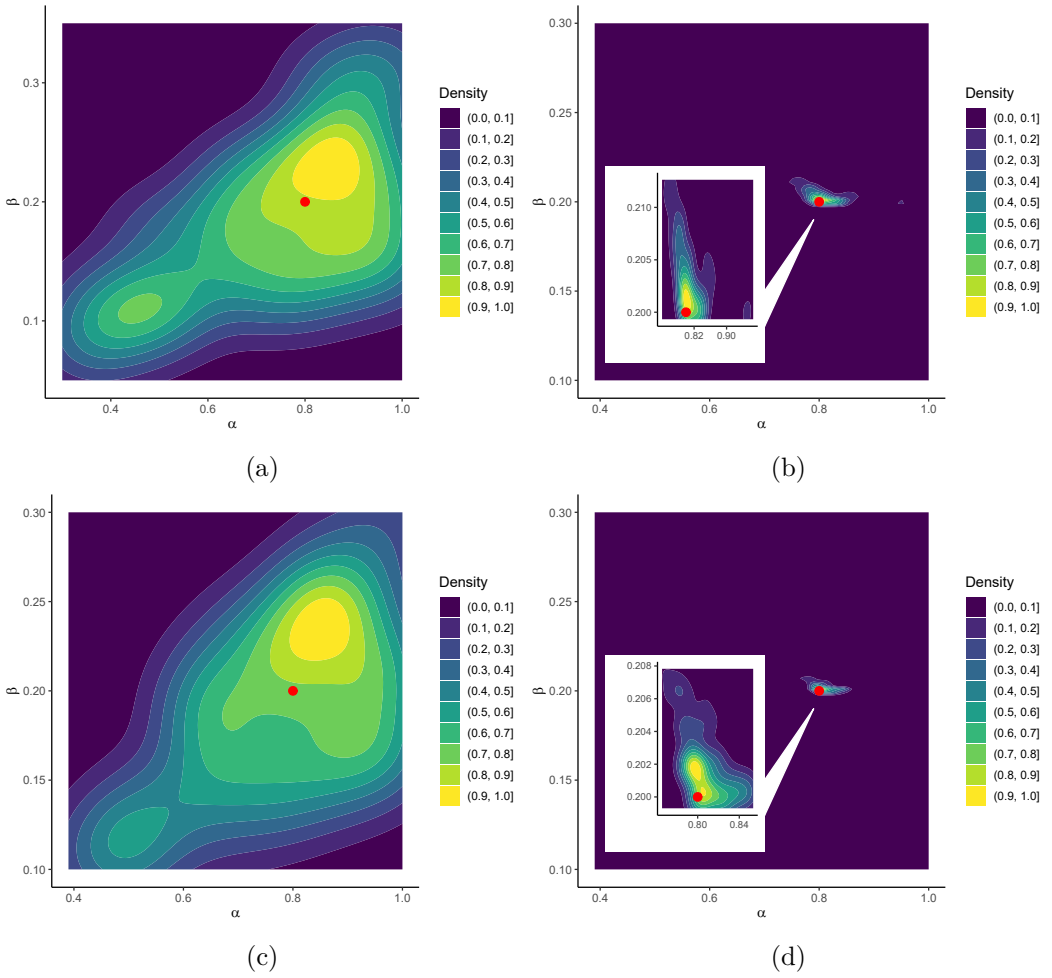


FIGURE 3.5: Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 1$ when the observed parameters are in the high-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustment for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.2$.

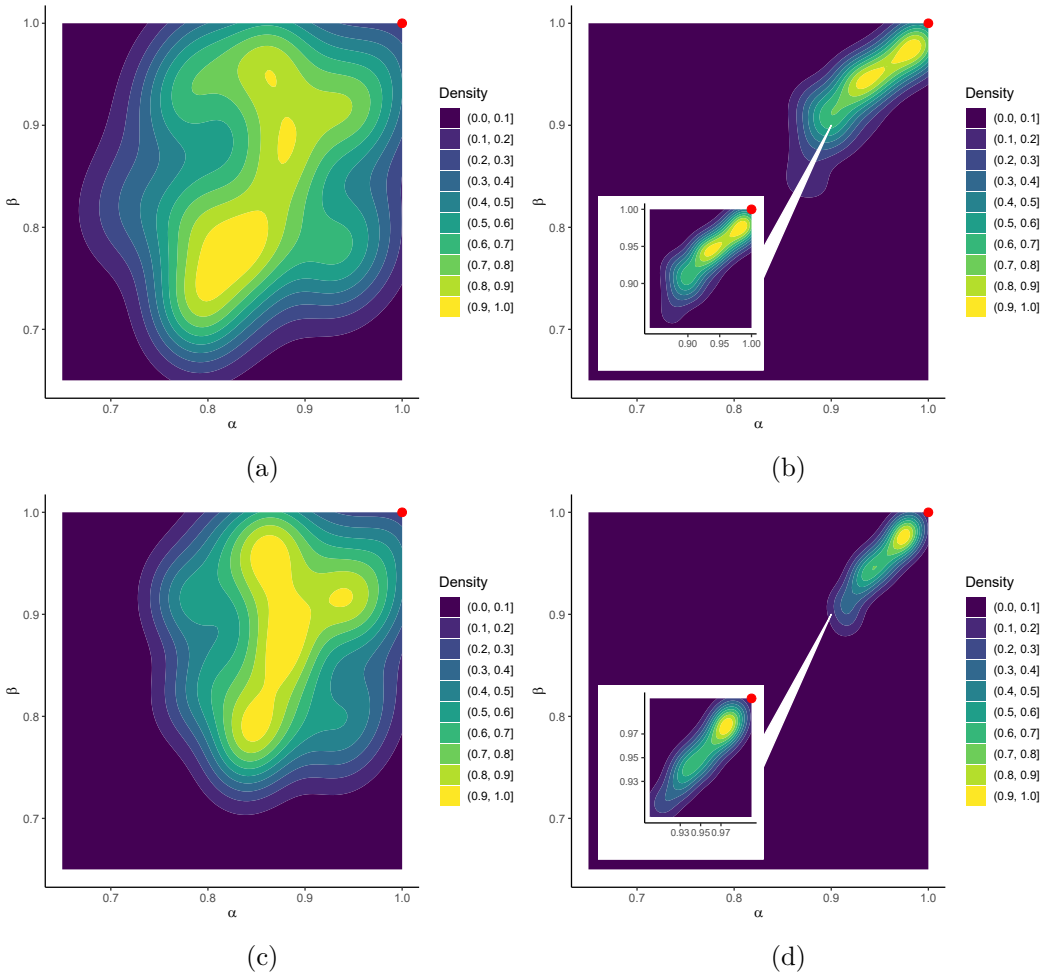


FIGURE 3.6: Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 1$ when the observed parameters are in the maximum-current regime. Sub-figures (a) and (c) represents the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 1$, $\beta = 1$.

When the observed boundary parameters are in the low-density regime, the estimated inflow probability, α is more accurate than the outflow probability, β. When the observed boundary parameters are in the high-density regime, the estimated outflow probability, β is more accurate than the inflow probability, α. This may be due to the boundary densities are more sensitive to the α when the system is in low-density regime and more sensitive to the β when the system is in the high-density regime. Compared to the low-density and high-density regime, when the observed boundary parameters are in the maximum-current regime of the ASEP with the parallel update and $p = 1$, the joint posterior distribution is significantly less accurate. However, the posterior mass around the true

parameter vector increases with the regression adjustments. Furthermore, the posterior concentration improves with the sample size increase. 3.6b and 3.6d.

3.3.3.3 ASEP: Parallel updates with $p = 0.5$

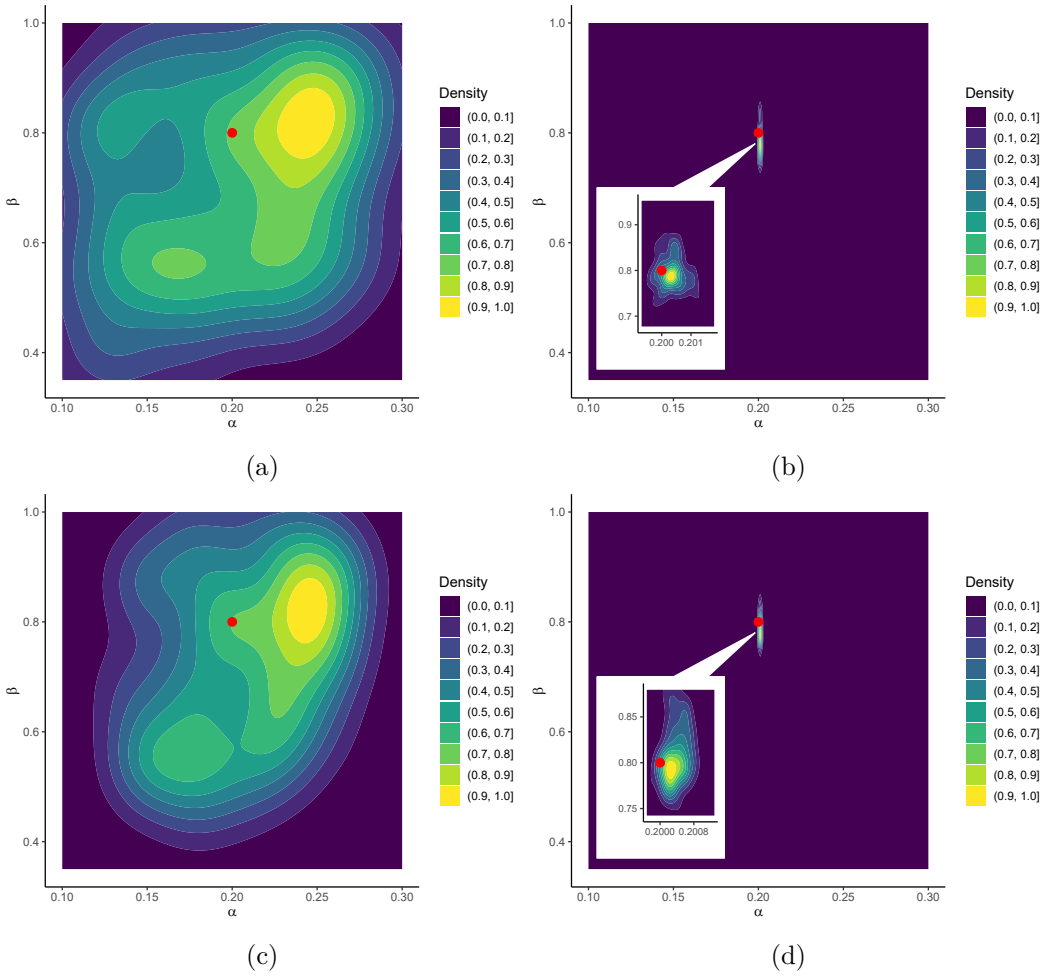


FIGURE 3.7: Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 0.5$ when the observed parameters are in the low-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.2$, $\beta = 0.8$.

We consider the system described in Section 2.1.2 with a one-dimensional lattice of length $L = 200$ with parallel updates and $p = 0.5$. We only consider the average density at left-boundary and right-boundary $S(y_{sim}) =$

$(\mathbb{E}[\rho_1], \mathbb{E}[\rho_L])$ as summary statistics in order to keep the dimension equal to the number of parameters. We fixed the true parameter vector generating the synthetic observed data at all three density regimes: low-density $(\alpha_{obs}, \beta_{obs}) = (0.2, 0.8)$, high-density $(\alpha_{obs}, \beta_{obs}) = (0.8, 0.2)$ and maximum-current $(\alpha_{obs}, \beta_{obs}) = (0.7, 0.7)$. The synthetic observed data for the all three density regimes: Low-density regime $y_{obs} = (0.348, 0.163)$; High-density regime $y_{obs} = (0.837, 0.652)$; Maximum-current regime $y_{obs} = (0.789, 0.211)$ simulated from the model described in Section 2.1.2 with parallel update and $p = 0.5$.

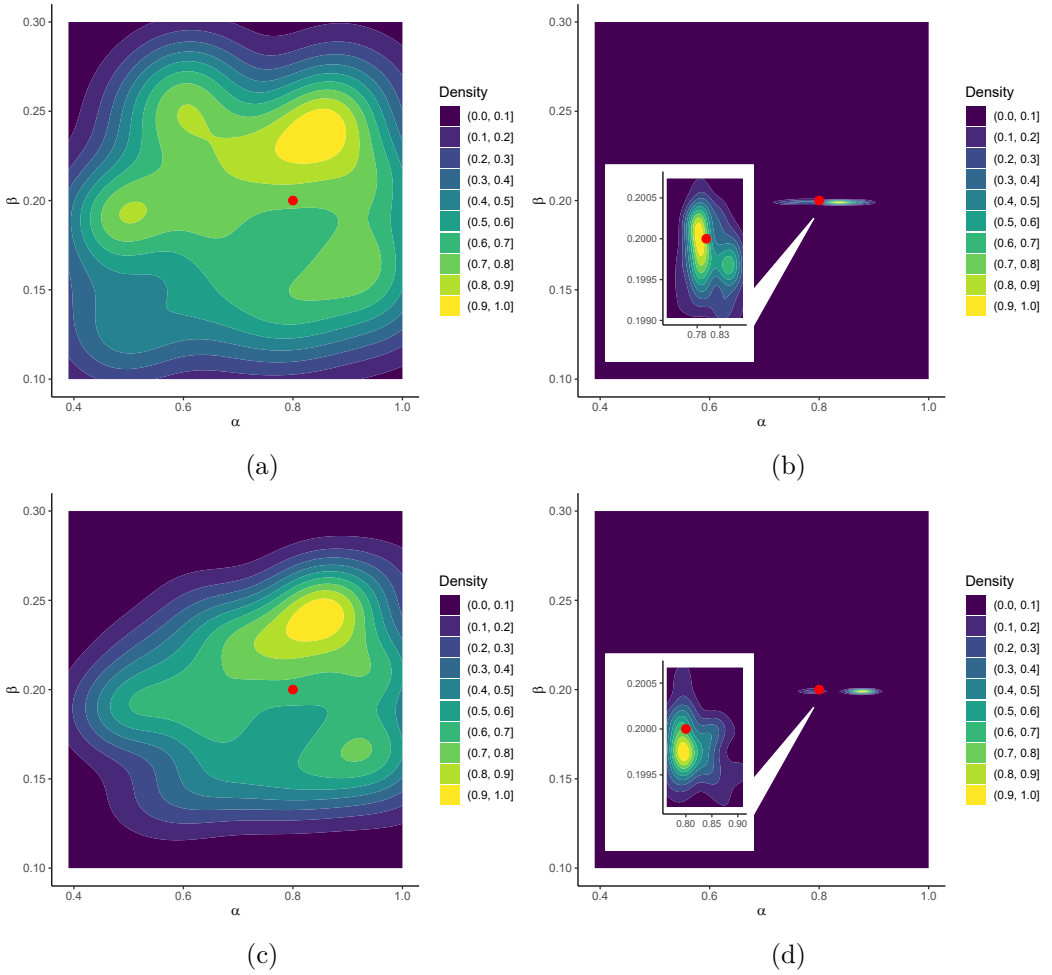


FIGURE 3.8: Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 0.5$ when the observed parameters are in the high-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustment for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.2$.

We take $N = 1000$ iid draws from uniform priors where $\alpha \sim U(0,1)$ and $\beta \sim U(0,1)$ representing all three density regimes. We simulate the system for fixed samples of size $T \in \{30, 50\}$ and for 1×10^5 time steps ensuring that the stationarity was always reached. We choose the tolerance rates to be $1/T^{0.7}$.

Figures 3.7, 3.8 and 3.9 summarise the joint posterior distribution of α and β of the ASEP with parallel updates and $p = 0.5$ when the observed parameters are in the low-density regime, high-density regime and maximum-current regime for different sample sizes respectively. Even for smaller sample sizes, estimated joint posterior distribution is much closer to the observed boundary parameters.

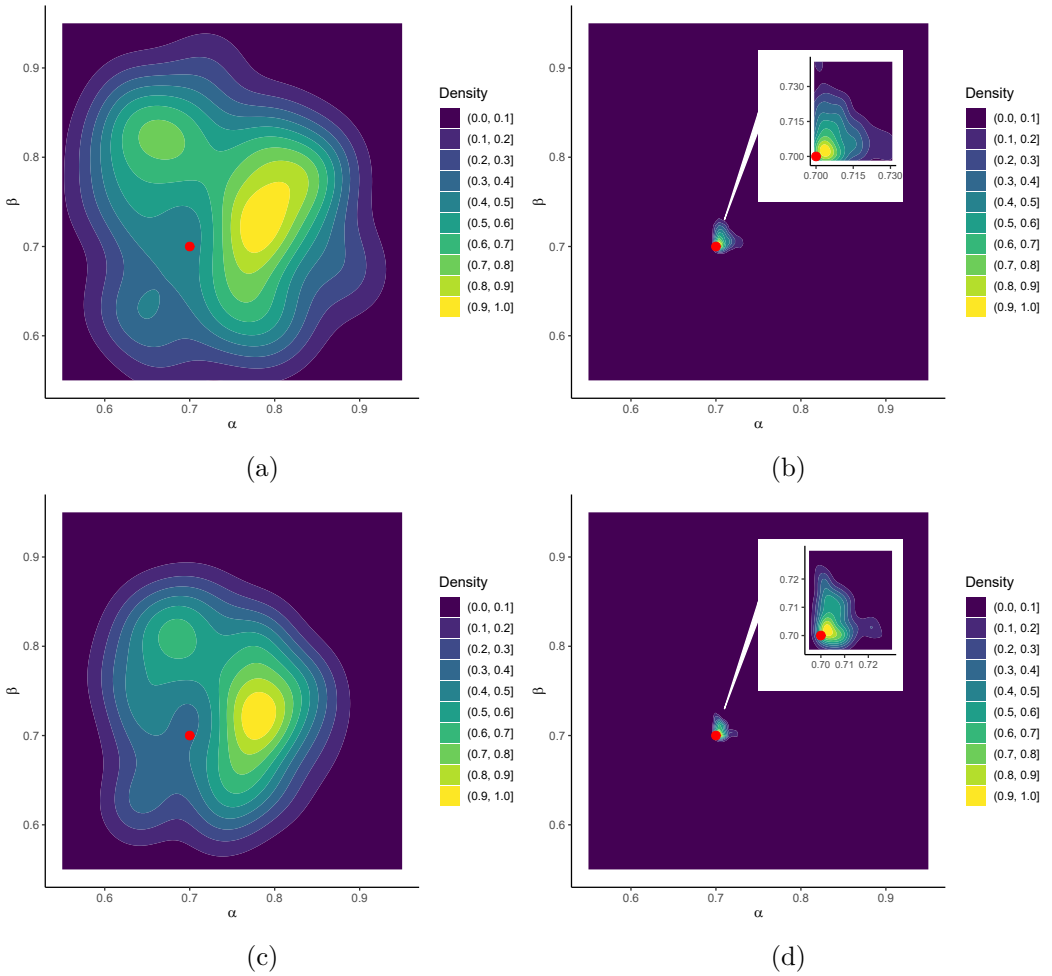


FIGURE 3.9: Joint posterior distribution of boundary parameters of an ASEP with parallel updates and $p = 0.5$ when the observed parameters are in the maximum-current regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustment for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.7$, $\beta = 0.7$.

Like in the ASEP with parallel updates and $p = 1$, the inflow probability α is accurately estimated in the low-density regime, and the outflow probability β is accurately estimated in the high-density regime. The maximum-current regime in the ASEP with $p = 1$ is just a single point where $\alpha = 1$ and $\beta = 1$. The system is close enough to the maximum-current regime if the inflow probability is large enough. Therefore, the density at the right-boundary is less sensitive to α. However, the maximum-current regime in the ASEP with $p < 1$ represent a range of α and β. Consequently, the density at the right-boundary is more sensitive to α. Correspondingly, the convergence rate of the posterior approximation given summary statistics is slower in the $p = 1$ case with respect to the $p < 1$ case. Therefore, compared to the maximum-current regime in the ASEP with parallel updates and $p = 1$, when the the observed boundary parameters are in the maximum-current regime of the ASEP with parallel updates and $p = 0.5$, the joint posterior distribution is accurately approximated.

The posterior mass around the true parameter vector is also increases with the regression adjustments. Furthermore, the posterior concentration is slightly improved with the sample size increase.

3.3.3.4 NaSch

Furthermore, we consider the system described in Section 2.1.1 with a one-dimensional lattice of length $L = 200$ cells. We consider the parameters, the maximum speed $v_{max} = 3$ and the random deceleration probability $d = 0.5$. When the observed parameters represent a low-density regime of NaSch model, we consider the average density at the middle and at the right-boundary $S(y_{sim}) = (\mathbb{E}[\rho_{L/2}], \mathbb{E}[\rho_L])$ as summary statistics. Otherwise, we consider the average density at left-boundary and right-boundary $S(y_{sim}) = (\mathbb{E}[\rho_1], \mathbb{E}[\rho_L])$ as summary statistics. The reason for not choosing ρ_1 as a summary statistic in low density is because when $v_{max} > 1$, vehicles may enter the system without occupying cell 1, making ρ_1 harder to estimate at low density. Two potential remedies exist for this: replace ρ_1 with an average of the occupation of the first v_{max} cells; or replace ρ_1 with the bulk density $\rho_{L/2}$. We choose the latter for simplicity.

We fixed the true parameter vector generating the synthetic observed data at three different density regimes: low-density regime$(\alpha_{obs}, \beta_{obs}) = (0.2, 0.8)$, maximum-current regime$(\alpha_{obs}, \beta_{obs}) = (0.4, 0.7)$ and high-density regime $(\alpha_{obs}, \beta_{obs}) = (0.8, 0.2)$. The synthetic observed data for these three density regimes: low-density regime $y_{obs} = (0.0813, 0.158)$; maximum-current regime $y_{obs} = (0.34, 0.31)$; high-density regime $y_{obs} = (0.886, 0.794)$ simulated from the model described

in Section 2.1.1 with $v_{max} = 3$ and $d = 0.5$. We take $N = 1000$ iid draws from uniform priors, where $\alpha \sim U(0, 1)$ and $\beta \sim U(0, 1)$. We simulate the system for fixed samples of size $T \in \{30, 50\}$ for 1×10^5 time steps ensuring that the stationarity was always reached. We choose the tolerance rates to be $1/T^{0.7}$.

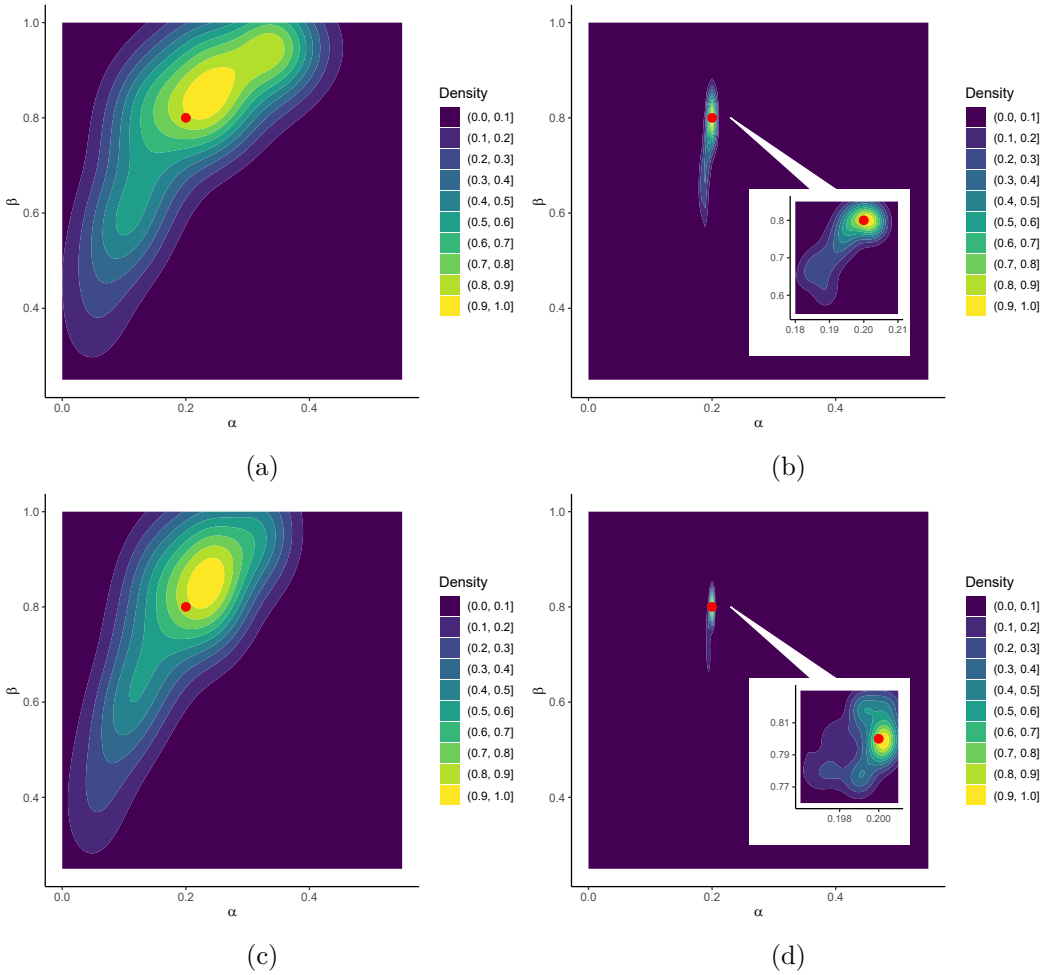


FIGURE 3.10: Joint posterior distribution of boundary parameters of NaSch when the observed parameters are in a low-density regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.2$, $\beta = 0.8$.

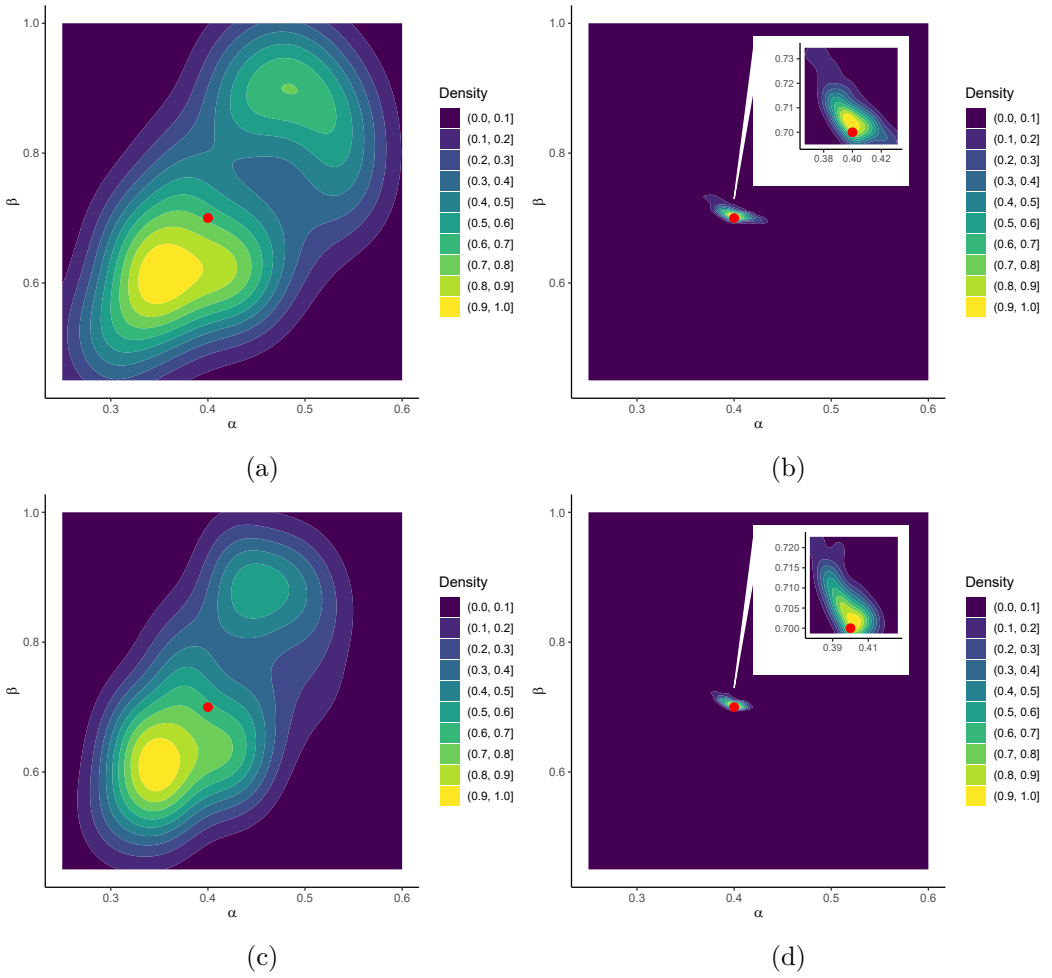


FIGURE 3.11: Joint posterior distribution of boundary parameters of NaSch when the observed parameters are in a maximum-current regime. Sub-figures (a) and (c) represent the joint posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub-figures (b) and (d) represent the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.4$, $\beta = 0.7$.

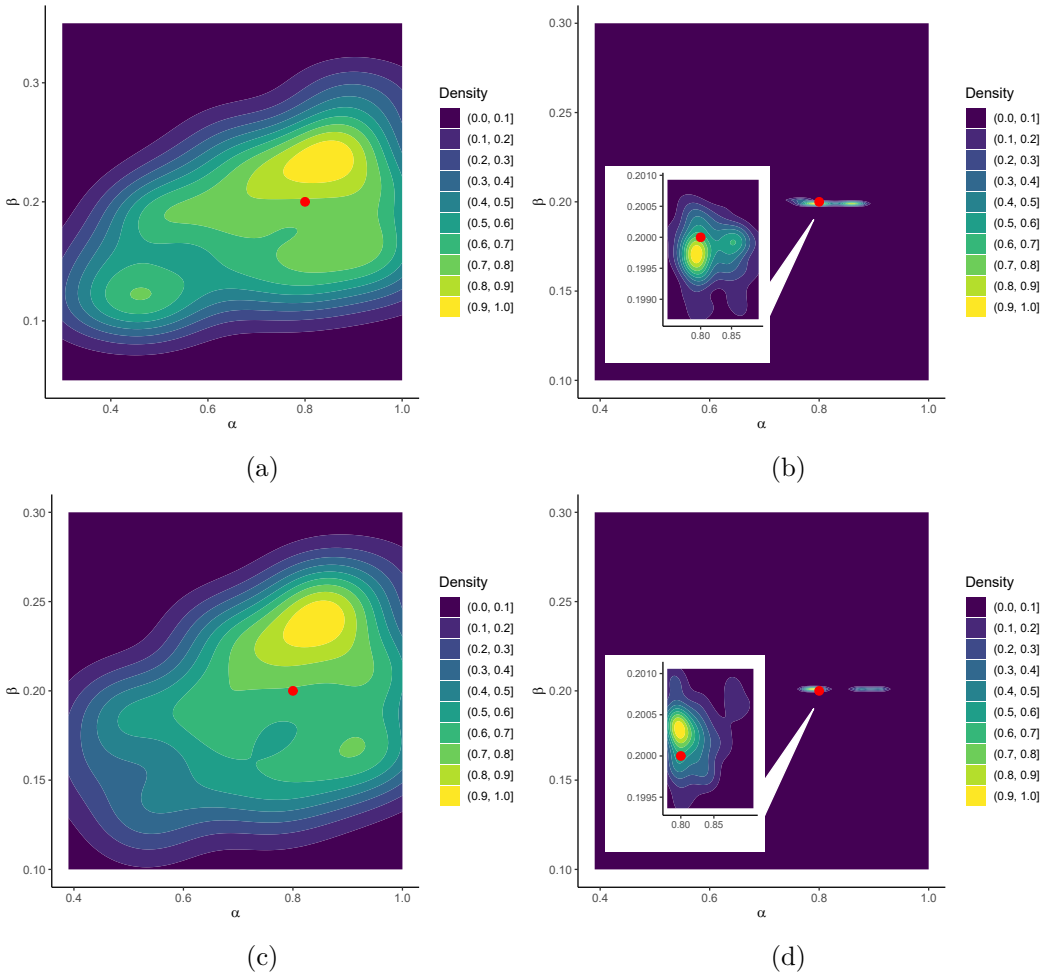


FIGURE 3.12: Joint posterior distribution of boundary parameters of NaSch when the observed parameters are in a high-density regime. Sub figures (a) and (c) represents the join posterior distribution of boundary parameters without regression adjustments for sample sizes 30 and 50 respectively. Sub figures (b) and (d) represents the join posterior distribution of boundary parameters with regression adjustments for sample sizes 30 and 50 respectively. The red dot represents the observed parameters, $\alpha = 0.8$, $\beta = 0.2$.

According to Figure 3.10, the ABC algorithm gives a more accurate estimate of the inflow probability α than the outflow probability β when the observed parameters are in a low-density regime. Regression adjustments greatly improves the accuracy of the posterior approximation. Figure 3.11 summarises the estimated joint posterior distribution of the inflow probability α and outflow probability β when the observed boundary parameters are in a maximum-current region. The basic ABC accept/reject algorithm with regression adjustment estimates both the boundary parameters accurately. When the observed boundary parameters are in a high-density regime, ABC algorithm estimates the outflow probability β more accurately than the inflow probability α as illustrated in Figure 3.12. Regression adjustments and the

larger samples improve the posterior concentration around the true parameter vector the same as before.

3.3.3.5 4×5 lattice network: turning rates

We consider a 4×5 network as illustrated in Figure 3.13 where We fixed all the turning movements at every intersection excluding node 20. We consider fixed cycle signals and independent, uniform boundary conditions. The true parameter vector generating the synthetic observed data is $(TR_{20,r,obs}, TR_{20,l,obs}) = (0.095, 0.142)$.

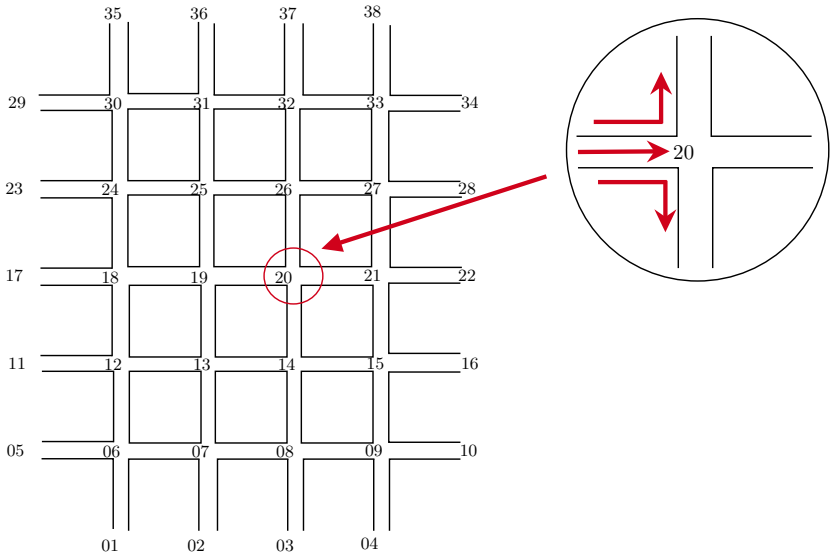
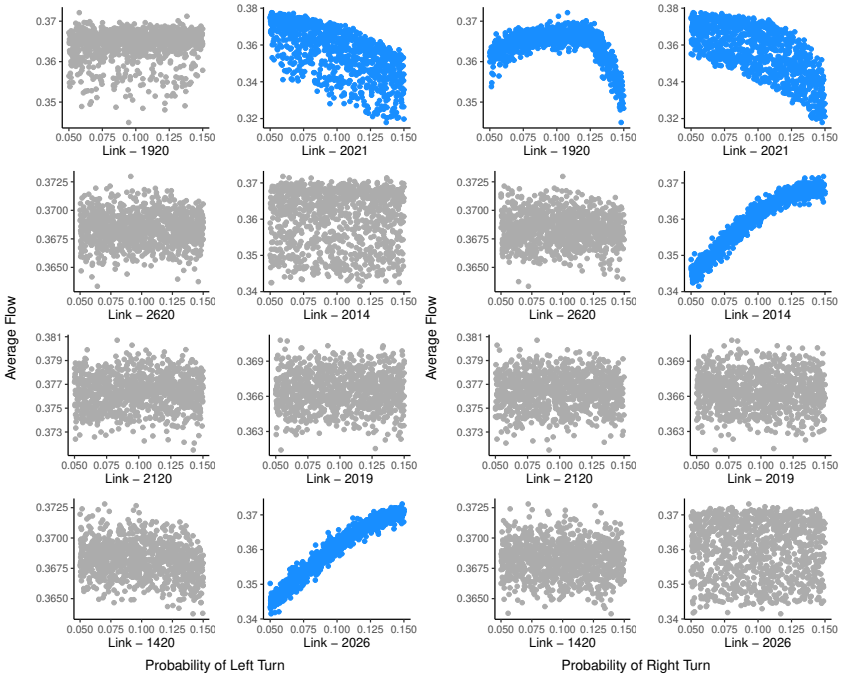


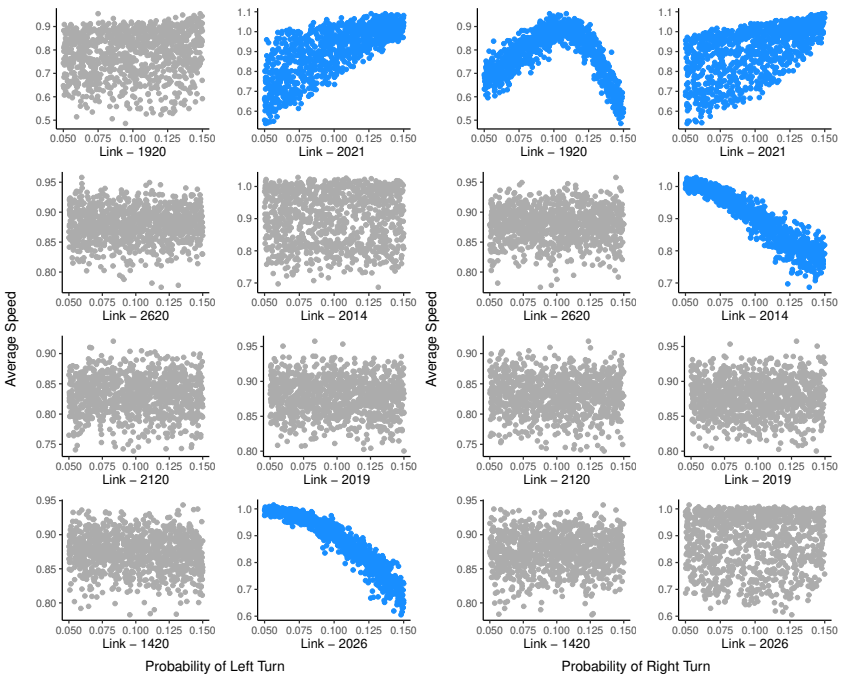
FIGURE 3.13: 4×5 lattice network

We simulated the network using the model defined in Section 2.1.3. We take $N = 10,000$ Monte Carlo draws from uniform priors where $TR_{20,r} \sim U(0.05, 0.15)$ and $TR_{20,l} \sim U(0.05, 0.15)$. We simulated the system for fixed sample size $T = 10$. We choose the tolerance rate to be $1/T^{0.7}$. We simulated the system for two hours ensuring that the stationarity was reached.



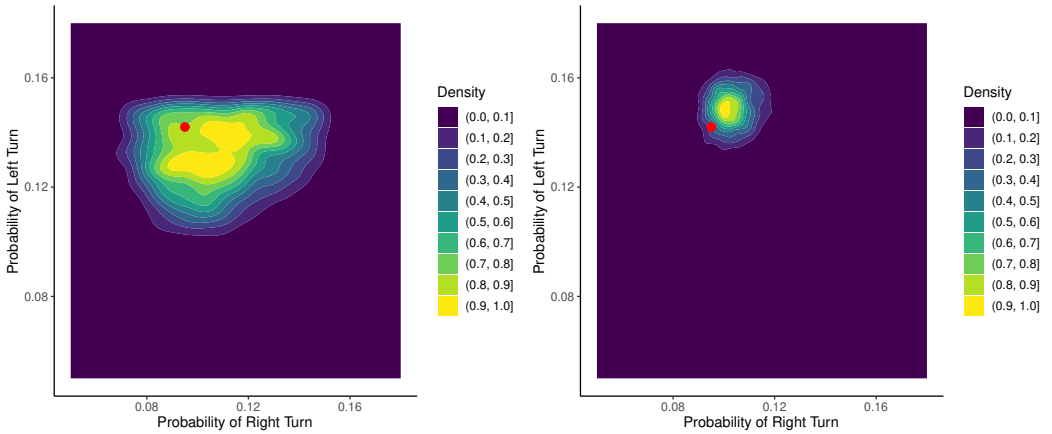
(a) Left turning probability (b) Right turning probability

FIGURE 3.14: Correlation between the average flow, and turning probabilities at Node 20



(a) Left turning probability (b) Right turning probability

FIGURE 3.15: Correlation between the average speed, and turning probabilities at Node 20



(a) Accept/reject ABC (b) With regression adjustments

FIGURE 3.16: Joint posterior distribution of east-bound right turning probability and left turning probability at the Node 20

We considered the average speed and average flow of neighbouring links as summary statistics, $S(y_{sim}) = (\mathbb{E}[V_l], \mathbb{E}[f_l])$.

Figures 3.14 and 3.15 summarise the relationship between summary statistics and turning probabilities. These plots provide a clear initiative to identify the most informative summary statistics for calibration. Traffic flow and speed of west-bound in-link, east north and west-bound out-links are informative to the turning probabilities. However, since the average flow on north-bound and south-bound out-links shows the strongest correlation with the turning rates, we chose the average flow on link 2026 and link 2014, $S(y_{sim}) = (\mathbb{E}[f_{2026}], \mathbb{E}[f_{2014}])$ as summary statistics.

Figure 3.16 summarises the estimated joint posterior distribution of the right turning rate and left turning rate. The regression adjustments approximate the joint posterior distribution of turning rates accurately even with a small sample size. To estimate the turning rates at a signalised intersection with ABC, traffic flow on left-bound and right-bound out links are much informative.

3.4 Discussion

There are many challenges inherited in validating and calibrating a large-scale traffic simulation model. One of the critical challenges in model calibration is constructing the network based on the geographical information and traffic signal information. Furthermore, we need to define traffic demands such as origin-destination matrix, turning rates at each intersection and frequency of pedestrian signals. Moreover, if the network consists of any public transport mode such as buses or trams, then we need to define the stopping behaviour of such transit vehicles as well. Likewise, there are many input parameters that we have to estimate in order to conduct an accurate simulation close to the actual network. The most challenging factor in the model calibration is, the high dimension of parameter vector of a typical traffic scenario. The existing classical calibration techniques fail to handle this curse of high-dimension.

ABC is a computational tool for parameter estimation in situations where the underlying likelihood functions are intractable which is applicable to many traffic simulation models. Instead of using likelihood function, the ABC methods quantify the similarities between observed data summaries and repeated simulations from the model. The only requirement to conduct a calibration utilising ABC is to have a data generating process and observed summary statistics to compare.

Nowadays, most of the busy traffic networks in the world are incorporating coordinated or adaptive traffic signal systems such as SCATS, *Main Traffic Signal System* (MTSS), *Split Cycle and Offset Optimization Technique* (SCOOT) and *Adelaide Coordinated Traffic Signal System* (ACTS). These types of traffic control systems can record traffic volumes for binned time intervals for each lane at each signalised intersection. Generally, there is a considerable amount of information regarding traffic volumes recorded in these types of systems which is useful in model calibration using ABC.

In this chapter, we studied the versatility of ABC in estimating the parameter space of numerical models simulating simple traffic flow updates. We consider ASEP with a few different update rules and NaSch in our study. We consider the basic ABC algorithm, accept/reject ABC with and without regression adjustments for the calibration. The results evident that the boundary parameters can be estimated accurately even with the most primitive ABC algorithm accept/reject with regression adjustments. Moreover, the summary statistics we consider, average density at the start and the end of the link that are easily measurable in practice with the loop detectors accurately calibrated the simple stochastic traffic models. Furthermore, with the average flow on

right-bound and left bound links, the turning rates also can be accurately approximated.

Complex traffic networks depend on high-dimension parameter space and the computational burden associated with calibration process are significant. However, ABC can be applied to more complex traffic simulations models, systematically, without relying on inefficient classical calibration techniques. More extensive networks can be calibrated as a combination of several ABC chunks which makes the process more efficient. Besides, the ABC approach estimates a distribution of parameters given observed summary statistics disregarding the classical perspective of having only one realisation for a particular traffic simulator calibration problem. ABC approximates the likelihood of the parameter where we can sensibly choose the correct parameters not depending on any particular choice of parameters.

Fortunately, there are recent works on much efficient ABC algorithms such as MCMC-ABC, *Sequential Monte Carlo ABC* (SMC-ABC), RABC and marginal adjusted ABC, which can be useful in calibrating much complex traffic simulation model. Incorporating some of these approaches into traffic simulator model calibration is a promising direction for future research in this area.

Link travel time distributions 4

4.1 Overview

One of the most important characteristics in modelling vehicular dynamics is link travel time distributions. It is evident that many researchers had different approaches to estimate link travel times such as time series forecasting methods [12, 13, 93, 94], Kalman filtering [14, 15], Neural Network [16–18, 95] non-parametric methods [96] and empirical distribution estimations methods [19–22]. Majority of these findings are subject to empirical data and cannot be generalised. A systematic framework which estimates link level travel time is essential for modelling vehicular dynamics in urban road networks. Researchers model travel time using a uni-modal symmetrical continuous distribution, such as normal distribution [23]. Many researchers argue that the travel time distributions are asymmetric and considerably positively skewed [19–22, 97] unlike normal distribution. In recent years researchers start to argue that the uni-modal distributions might not be sufficient to represent link travel time distributions [22, 98–105] as the travel time is highly sensitive to the traffic light at intersections as well.

4.1.1 Literature Review: link travel time distributions

Mazloumi et al. [106] showed that in narrower departure time windows, bus travel time distributions are best characterised by a normal distribution and for wider departure time windows, peak-hour travel time follows a normal distribution, while off-peak travel time follow a log-normal distribution using a *Global Positioning System* (GPS) data set for a bus route in Melbourne, Australia. Kieu et al. [20] analysed bus travel time data and recommended that the bus travel time on urban corridors can be well approximated by a log-normal distribution. Rahaman et al. [21] showed that the log-normal distribution performed significantly better up to about 7km length of links based on the bus travel time data collected in Calgary Transit System, Canada. Links longer than 7km, the normal distribution outperformed the log-normal distribution. The shape of the travel time distribution is significantly influenced by the congested level of the link. van Lint et al. [107] have identified four phases, free flow, congestion onset, congestion, and congestion dissolve with distinctively different shapes of travel time distributions.

However, it is evident that the travel time data consists of a reasonably longer upper tail which cannot be fully characterised by skewed distributions such as log-normal, Weibull or gamma distributions [22]. Susilawati et al. [22] argued that the travel time distributions can be characterised by very long upper tail and the Burr distribution can be considered as a leading candidate for capturing travel time variability. They studied two separate routes and estimated the travel time distributions for every link. Lengths of these links varied in between 135m-4008m. Their findings confirm that the travel time can be well approximated by two-parameter Burr distribution except for a few links in the inner city. However, Burr distribution is not a good candidate for travel time data as it is a power-law distribution which is first introduced to model income distributions. The distribution has a significantly longer upper tail which cannot be explicitly explained for travel time data.

Guo et al. [98] proposed a multi-state model utilising a Gaussian mixture distribution to model travel times on a link at San Antonio, Texas. They considered two states of travel times: trips having significantly shorter travel times and trips having significantly longer travel times. They argued that the longer travel times could be the delay caused by traffic incidents. Kazagli and Koutsopoulos [99] proposed a mixture of two log-normal distributions to represent travel time based on the *Automatic vehicle identification* (AVI) data collected from several corridors in central Stockholm, Sweden. They argued that one population represents the vehicles that travel through the network without any disturbances, and the other one represents the vehicles that are delayed for various reasons. Susilawati et al. [22] noticed some of the links they

studied have bi-modal travel time distributions and suspected the first cluster might represent the vehicles that travel with a free flow speed while the second cluster represents the vehicles with very long travel times. They argued that shorter links with queue accumulation at intersections due to traffic lights could be the reason for having two peaks in travel time. Ji et al. [101] also claimed that the link travel time is bi-modal with one mode corresponding to travels without delays and the other for travels with delays. Ma et al. [100] estimated the bus travel time distribution using *Automatic Vehicle Location* (AVL) data collected on two different bus routes over 6 months in Brisbane. They argued that the uni-modal distributions such as normal, log-normal, logistic, log-logistic and gamma models have relatively similar performance considering bus travel times on a route level. However, the link level bus travel time is multi-modal and can be well approximated by the use of mixture distributions such as Gaussian mixture models. They used up to three modes, which can be related to free flow, recurrent and non-recurrent traffic states, to describe the probability distributions of bus travel time. Chen et al [104] also fitted a mixture of Gaussian models to bus travel time distribution for different traffic states. Their results indicate that the travel time during peak hours can be well approximated by a mixture model with four components and the travel time during off-peak hours can be well approximated by two components. Chen, Yin and Sun [105] proposed a finite mixture of regression model with varying mixing proportions to estimate route travel time distributions based on AVI data. They incorporated the effect of signal timing in their study and concluded that the multi-state distributions provide a superior fit over the alternative uni-modal distributions. Qin and Yun [108] proposed a Bayesian approach based on particle filter framework to estimate link travel times based on Floating Car Data in Nanjing, China. They considered four different traffic states, non-stopped, stopped, stopped with delays and stopped twice or more, to exhibit travel times on a link experienced by vehicles.

Many researchers tend to incorporate link level travel time distributions to estimate the path level travel time distribution considering the correlation structure of link level travel time distributions [109–113]. Chen et al. [109, 110] introduced a copula-based method to model the dependent structure of travel times between links in a corridor. They first divided the desired arterial into two connected road segments; each consists of more than one link and applied a standard two-dimensional copula function selected from the family of bi-variate copulas to fit the travel time observations between two connected segments. They claimed that the link level travel time distributions are more complex and there exists a multi-modality phenomenon than the path level travel time distributions. They argued that the reason for this may be link travel times are more sensitive to intersections and to stop delays in shorter distances. However, they failed to explain the behaviour of the model parameters according to the observed travel times. Yu et al. [111]

also introduced a copula-based approach to estimate link-level and path-level travel time distributions by considering the phenomenon called Channelisation Section Spillover and the correlation of links. The model was assessed with travel time data on an arterial road in Hangzhou, China. They also failed to make any assumptions regarding input variables. Recently Qin et al. [112] proposed a pair copula construction approach to estimate travel time along a route using a more general and flexible structure of depicting dependence between consecutive link travel times. They introduced bi-variate copulas in each pair-copula to deal with the complex correlations between link travel times. However, they did not incorporate the effect on signal timing and offset setting at each intersection into the travel time estimation. Sen et al. [113] utilised the estimated link level travel time distributions to estimate the joint distribution of path travel time distribution. They considered the travel time volatility to quantify the dependent structure of the link travel times. They claimed that the Generalised hyperbolic distribution can be considered as the candidature distribution for the link level travel time based on the log likelihood values. None of these studies can be generalised as these results are subject to the route level characteristics and operating environments.

Recently, Wang et al. [114] analytically derived the travel time distribution through an intersection using the well-established LWR shock wave model. In this study they considered the stochasticity of free flow and assumed that the free flow pace follows a normal distribution. Consequently, they approximated the distribution of travel time to pass a signalised intersection from a certain location in the upstream link by incorporating the length of the road segment. They argued that the travel time distribution without delay, that is the vehicles not delayed at the intersection, should follow a normal distribution with some mean and standard deviation while the travel time distribution with delays also follows normal distribution with a slightly higher mean incorporating the delay imposed at the intersection and with the same standard deviation as travel times without delays. They only considered the signal time dynamics at the downstream junction rather than both upstream and downstream intersections. Hence, the results cannot be generalised to estimate the link travel time distribution explicitly. Luo et al. [102] proposed an adaptive clustering technique to estimate mean travel time in signalised road segments and they studied the multi-state properties of link travel times with the data collected from point-to-point detectors. However, they haven't stated any arguments regarding the underlying probabilistic distribution of the link travel time.

It is evident that these studies do not present consistent results. None of the results provide a clear initiative to the link travel time distributions. Furthermore, the link travel time distributions are highly sensitive to the traffic light dynamics at signalised intersections. Longer red time may cause vehicle

queue accumulation at the intersection and higher congestion on the upstream. Longer green time allows more vehicles to pass through the intersection and less congestion on the upstream. Moreover, the linking pattern of the traffic signals at consecutive intersections also has a significant impact on the delay imposed at the intersections. With the optimal offset many vehicles can join the green wave and hence shorter travel time can be expected.

In this chapter we propose a model to approximate link travel time[1] distribution for simple stochastic car following models based on Negative Binomial Distribution. We evaluate the proposed model against NaSch. In addition to that, we investigate the effect of traffic light dynamics at signalised intersections on link travel time distribution utilising NaSch model as well.

4.2 Non-signalised links

4.2.1 Mean-field approximation

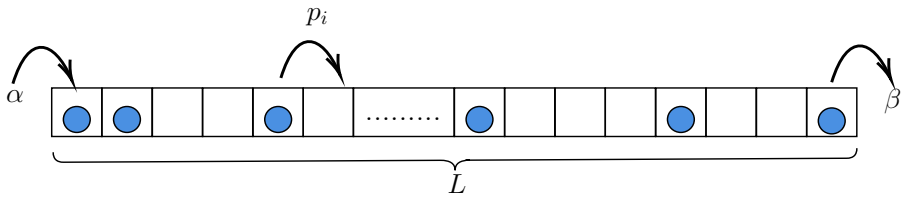


FIGURE 4.1: Link Configuration: Parameters α, β and p_i represent the traffic inflow probability, outflow probability and cell dependent hopping probability respectively.

To gain some physical intuition for the form of the travel time distribution, we begin by describing a simple *mean-field* type approximation for ASEP (with parallel updates), which explicitly neglects correlations between cells. The central, but simple, observation is that conditioned on cell $i+1$ being unoccupied, the number of steps that a vehicle will spend in cell i is a geometric random variable with success parameter p. Therefore, removing the conditioning, but neglecting correlations, we expect that to a good approximation the number of time steps a vehicle on cell i waits before moving to cell $i+1$ is geometric with (effective) parameter $p(1-\rho_{i+1})$, where ρ_{i+1} is the average density of cell $i+1$. Further approximating ρ_{i+1} by the bulk density ρ, it follows that, under

[1]The link travel time is the amount of time that elapses between when a vehicle first enters and link and when it leaves. In our discrete models, this is simply the number of iterations a vehicle spends on a link.

the above approximation, the travel time on a link of length L will be a sum of iid geometric random variables, each with parameter $p_{\text{eff}} := p(1-\rho)$. It is well-known that the sum of iid geometric random variables has a negative binomial distribution. Therefore, the above mean-field argument suggests that the travel time for ASEP should be well-approximated by random variable T_L which has negative binomial distribution with parameters (L, p_{eff}), i.e.

$$P(T_L = t) = \binom{t-1}{L-1}(1-p_{\text{eff}})^{t-L}p_{\text{eff}}^{L}, \qquad t = L, L+1, \ldots \tag{4.1}$$

In particular, it is elementary to show that T_L has mean L/p_{eff} and variance $L(1-p_{\text{eff}})/p_{\text{eff}}^2$. Moreover, since T_L is sum of iid random variables with finite variance, the central limit theorem implies that for large L, the distribution of T_L will be well approximated by a normal distribution, $N(L/p_{\text{eff}}, L(1-p_{\text{eff}})/p_{\text{eff}}^2)$. Thus, the above simple physical picture strongly suggests that the travel time on a single link of length L should have mean and variance proportional to L, and should be approximately normally distributed, at least when L is relatively large.

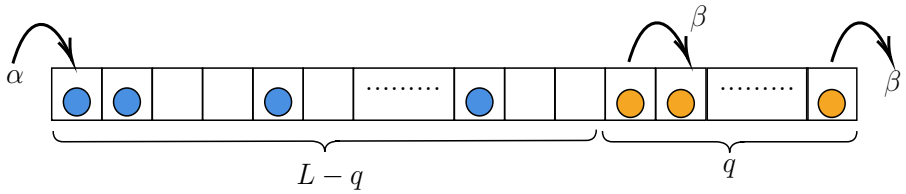


FIGURE 4.2: Link Configuration: Parameters α, β and q represent the traffic inflow probability, outflow probability and queue length respectively.

We now consider a refinement of the above picture for $p = 1$ at low-density. In this case, the system consists of a queued domain, of length Q, connected to the right boundary, and a free-flow domain, of length $L - Q$, connected to the left boundary; see Figure 4.2. Conditioned on a queue length of $Q = q$, the travel time of a vehicle through the free-flow domain is simply $L - q$. Since vehicles leave the queue with probability β, and $p = 1$, vehicles within the queue also move forward with probability β. Therefore, conditioned on $Q = q$, the number of time steps a vehicle remains in the system after it joins the queue has a negative binomial distribution with parameters (q, β). Moreover, the distribution of Q can be modelled by a single server $M/M/1$ queue in its steady state, conditioned on $Q \leq L$, where inter-arrival times and service times are exponentially distributed with rates ρ and $1 - \rho_L$ respectively [115], yielding

$$P(Q = q) = \frac{(1-\gamma)}{1-\gamma^{L+1}}\gamma^q, \qquad 0 \leq q \leq L, \tag{4.2}$$

where

$$\gamma := \frac{\rho_{L/2}}{(1-\rho_L)} = \frac{\alpha\beta}{\beta - \alpha + \alpha\beta}.$$

It follows that

$$P(T=t) = \frac{(1-\gamma)}{1-\gamma^{L+1}}(1-\beta)^{t-L}\sum_{q=0}^{L}\binom{t-L+q-1}{q-1}(\beta\gamma)^q \tag{4.3}$$

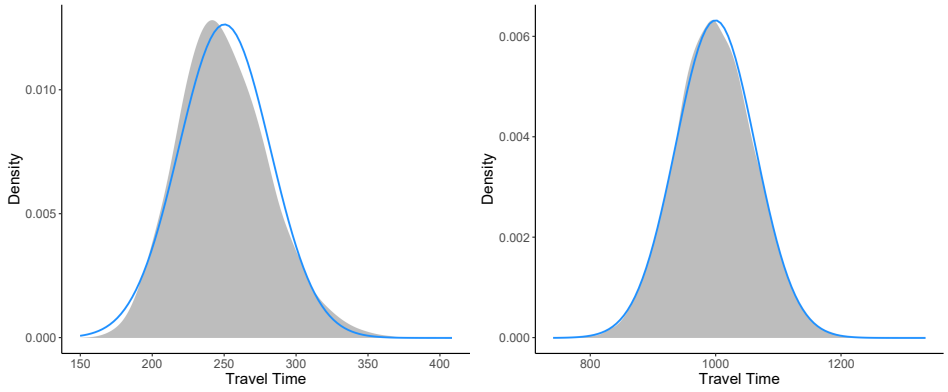
The Equation 4.1 with effective hopping probability is used to approximate the travel time of a vehicle throughout this chapter, unless the system is in the low-density regime with deterministic updates.

4.2.2 Simulations: ASEP

This section tests the accuracy of the approximations discussed in the previous section, by simulating ASEP with fully parallel dynamics, for both deterministic ($p = 1$) and stochastic case ($p < 1$) bulk updates. While some of these results, particularly those relating to $p = 1$, may not be of direct practical interest, they will likely be of theoretical interest within statistical mechanics. From that perspective, the approximations discussed in the previous section can be viewed as mathematical conjectures, and this section provides numerical evidence in their support. This is particularly the case for the generic Gaussian behaviour of the travel time, which suggests that a central limit theorem holds. In addition to being of theoretical interest, such a central limit theorem, if extended to more realistic models, would also be of practical interest.

4.2.2.1 Deterministic updates

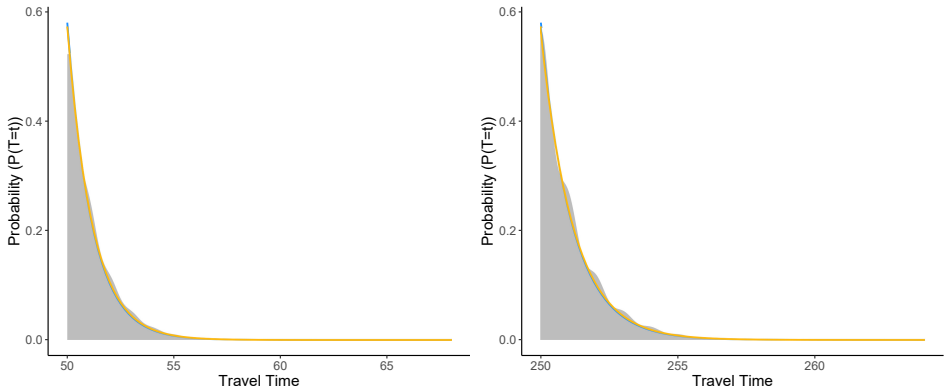
With deterministic updates ($p = 1$), two sub phases in low-density and high-density regimes, LD2 and HD2 vanish and the system has only three different phases: LD ($\alpha < \beta$), HD ($\alpha > \beta$) and MC($\alpha = \beta = 1$). We consider two different inflow probability and outflow probability combinations: $\alpha = 0.8$ and $\beta = 0.2$ corresponding to high-density, and $\alpha = 0.1$ and $\beta = 0.6$ corresponding to low-density. We simulated systems with various $L \leq 1000$ for approximately 5000 iterations, ensuring that stationarity was always reached. We recorded the number of time steps that a vehicle needs to get through the link.



(a) L=50 cells (b) L=200 cells

FIGURE 4.3: Travel time distribution of a vehicle in high-density regime ($\alpha = 0.8$ and $\beta = 0.2$) of an ASEP with deterministic updates. The blue curve represents the approximated travel time distribution. Empirical distributions are smoothed with non-parametric Gaussian kernel.

Figure 4.3 shows the simulated travel time distribution in high-density. Comparing $L = 50$ with $L = 200$, it is clear that the distribution is tending to a Gaussian, as argued in the previous section. Moreover, the simulated distribution is in excellent agreement with a Gaussian distribution with the parameters outlined in the previous section, using the analytical value of the bulk density given in Chapter 2. It would seem likely that this conjectured central limit theorem, with explicit forms for the shift and scale parameters used in the standardisation, could actually be proved rigorously in the case $p = 1$; this would be an interesting avenue for future work.



(a) L=50 cells (b) L=250 cells

FIGURE 4.4: Travel time distribution of a vehicle in low-density regime ($\alpha = 0.1$ and $\beta = 0.6$) of an ASEP with deterministic updates. The blue and the golden curves represent the approximated travel time distribution based on the empirical queue length and queuing theory respectively. Empirical distributions are smoothed with non-parametric Gaussian kernel.

Figure 4.4 shows the simulated travel time distribution in low-density. Also shown is the theoretical prediction (4.3); the agreement is clearly excellent.

4.2.2.2 Stochastic updates

With $p < 1$, the system can obtain all five regimes illustrated in Figure 2.4. As argued in Section 4.2.1, neglecting correlations, vehicles in the system with stochastic updates will move forward with an effective hopping probability $p_{\text{eff}} = p(1-\rho)$, and the travel time distribution should be approximately $N(L/p_{\text{eff}}, L(1-p_{\text{eff}})/p_{\text{eff}}^2)$ for large L. Moreover, the exact value of the bulk density ρ is known exactly [58], even when $p < 1$. We consider six different inflow and outflow probability combinations with $p = 0.5$ where the critical probability is $0.293(= 1 - \sqrt{1-p})$;

- $\alpha = 0.1$ and $\beta = 0.2$ in LD1 regime; $\alpha < \beta < 0.293$.
- $\alpha = 0.001$ and $\beta = 1$ in LD2 regime; $\alpha < 0.293 < \beta$.
- $\alpha = 0.1$ and $\beta = 0.6$ in LD2 regime; $\alpha < 0.293 < \beta$.
- $\alpha = 0.2$ and $\beta = 0.1$ in HD1 regime; $\beta < \alpha < 0.293$.
- $\alpha = 0.8$ and $\beta = 0.2$ in HD2 regime; $\beta < 0.293 < \alpha$
- $\alpha = 0.4$ and $\beta = 0.7$ in MC regime; $\alpha, \beta > 0.293$.

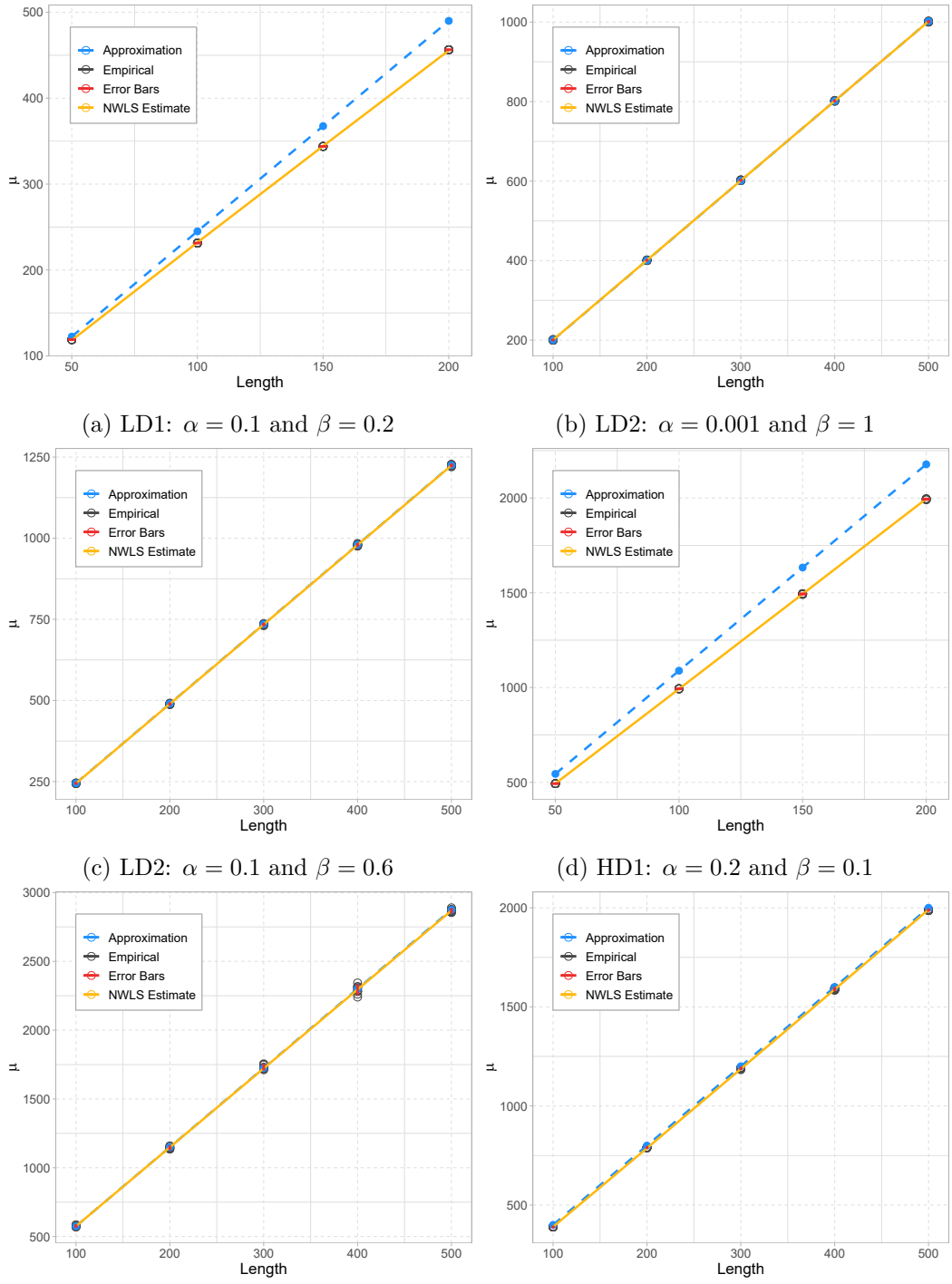
A very low-density regime with $\alpha = 0.001$ and $\beta = 1$ was considered to understand the travel time distribution of a vehicle when the interaction among vehicles is minimal. We simulated the system for all six cases for approximately 5000 time steps, ensuring that stationarity was always reached.

Both the empirical and approximated distributions are bell-shaped following similar distributions. Therefore, we can simply compare their estimated and approximated parameters μ and σ. Figures 4.5 and 4.6 summarise the approximated and empirical mean μ and variance σ^2 for the travel time distribution of an ASEP in all the five density regimes for varying sizes of links. To test the dependence of μ and σ with L, we fit non-linear weighted least square (NWLS) estimates for both approximated and observed means and variances incorporating the Levenberg-Marquardt algorithm [116]. See Tables 4.1 and 4.2.

Table 4.1: Non-linear WLS Regression Fit for empirical and approximated μ of travel time of an ASEP with varying link sizes. Formula: $\mu = b \times L^c$

		Empirical	Approximation
LD1: $\alpha = 0.1$, $\beta = 0.2$	b	$2.6284^{***}[\pm 0.014094]$	$2.450^{***}[\pm 9.126 \times 10^{-15}]$
	c	$0.9729^{***}[\pm 0.001066]$	$1^{***}[\pm 7.327 \times 10^{-16}]$
LD2: $\alpha = 0.001$, $\beta = 1$	b	$2.019^{***}[\pm 0.006358]$	$2.004^{***}[\pm 1.208 \times 10^{-14}]$
	c	$0.999^{***}[\pm 0.000534]$	$1^{***}[\pm 1.008 \times 10^{-}15]$
LD2: $\alpha = 0.1$, $\beta = 0.6$	b	$2.4349^{***}[\pm 0.009347]$	$2.450^{***}[\pm 4.09 \times 10^{-15}]$
	c	$1.0009^{***}[\pm 0.000668]$	$1^{***}[\pm 2.793 \times 10^{-16}]$
HD1: $\alpha = 0.2$, $\beta = 0.1$	b	$9.5708^{***}[\pm 0.016240]$	$10.89^{***}[\pm 3.521 \times 10^{-14}]$
	c	$1.00791^{***}[\pm 0.000368]$	$1^{***}[\pm 6.362 \times 10^{-16}]$
HD2: $\alpha = 0.8$, $\beta = 0.2$	b	$5.7804^{***}[\pm 0.056080]$	$5.75^{***}[\pm 9.905 \times 10^{-15}]$
	c	$0.9988^{***}[\pm 0.001663]$	$1^{***}[\pm 2.881 \times 10^{-16}]$
MC: $\alpha = 0.4$, $\beta = 0.7$	b	$3.68105^{***}[\pm 0.012360]$	$4^{***}[\pm 2 \times 10^{-16}]$
	c	$1.0126^{***}[\pm 0.000594]$	$1^{***}[\pm 2 \times 10^{-16}]$
– – –			
*** significance level is ≈ 0			

The mean and variance of the travel time are proportional to the system size L. The approximated means are almost identical to the empirical values in all density regions except for LD1 and HD1. Even, the fitted NWLS estimates for approximated and empirical means are mostly similar for all the cases except for LD1 and HD1 as in Table 4.1. The assumption of the flat density profile does not well capture the LD1 and HD1 case. Therefore, the local density profile has a slight increment at the right-boundary when the system is in LD1 regime, and a slight increment at the left-boundary when the system



(a) LD1: $\alpha = 0.1$ and $\beta = 0.2$

(b) LD2: $\alpha = 0.001$ and $\beta = 1$

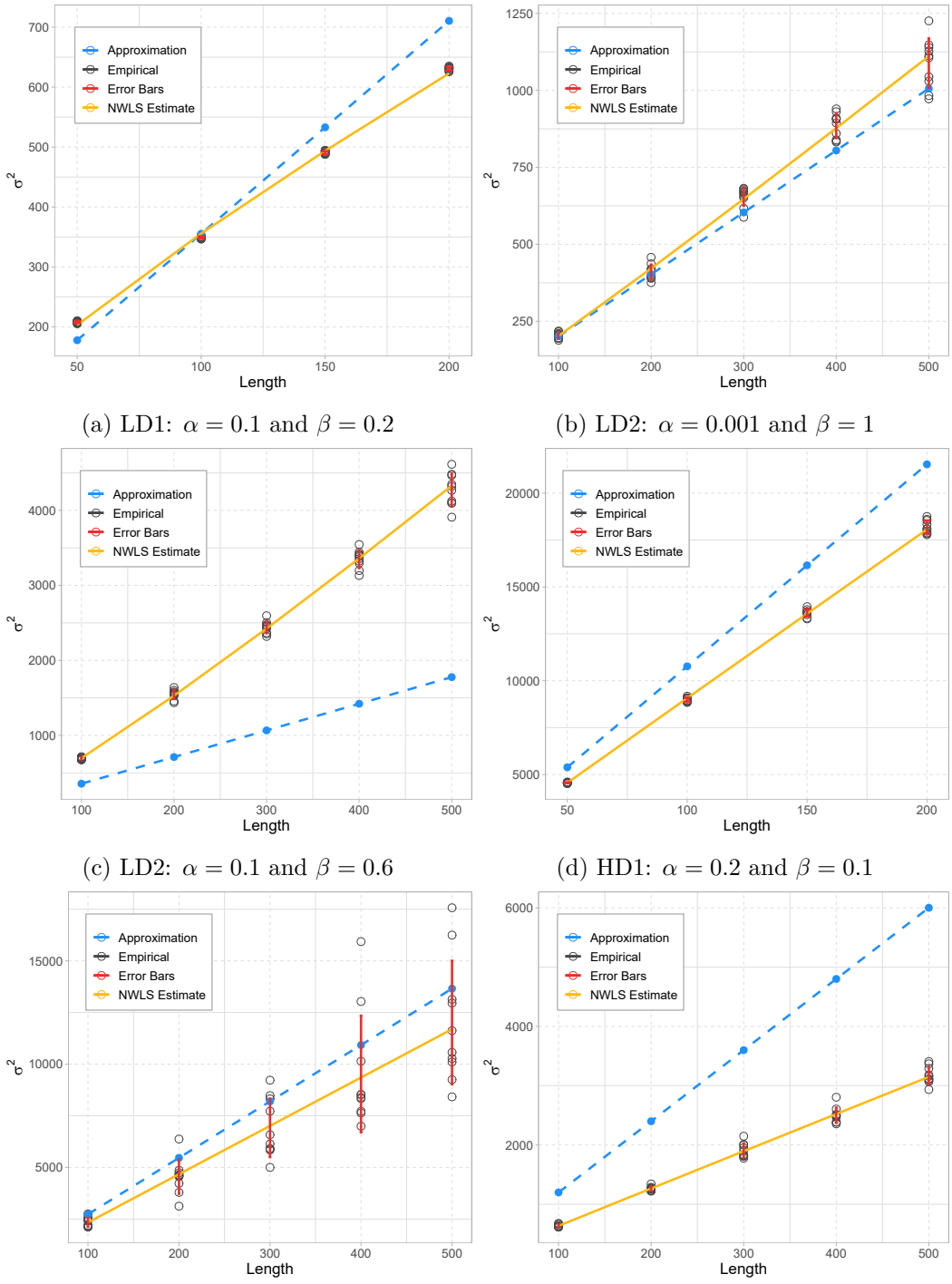
(c) LD2: $\alpha = 0.1$ and $\beta = 0.6$

(d) HD1: $\alpha = 0.2$ and $\beta = 0.1$

(e) HD2: $\alpha = 0.8$ and $\beta = 0.2$

(f) MC: $\alpha = 0.4$ and $\beta = 0.7$

FIGURE 4.5: Mean comparison of approximated and empirical travel times of an ASEP for different density regions with varying link sizes. Error bars corresponding to one standard deviation are shown. Fitted line of NWLS estimates for empirical mean travel time is shown.



(a) LD1: $\alpha = 0.1$ and $\beta = 0.2$

(b) LD2: $\alpha = 0.001$ and $\beta = 1$

(c) LD2: $\alpha = 0.1$ and $\beta = 0.6$

(d) HD1: $\alpha = 0.2$ and $\beta = 0.1$

(e) HD2: $\alpha = 0.8$ and $\beta = 0.2$

(f) MC: $\alpha = 0.4$ and $\beta = 0.7$

FIGURE 4.6: Variance comparison of approximated and empirical travel times of an ASEP for different density regions with varying link sizes. Error bars corresponding to one standard deviation are shown. Fitted line of NWLS estimates for empirical mean travel time is shown.

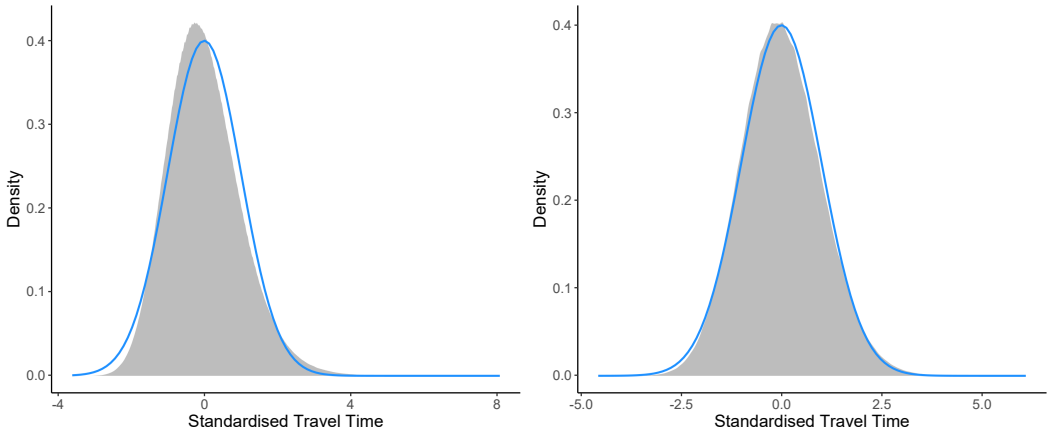
is in HD1 regime [58]. As a result of this queue accumulation, there exists a discrepancy between the empirical and theoretical means.

Mean travel time approximation is overestimated in HD1 regime since vehicles travel faster in the low-density region at the left-boundary than the assumed scenario. In LD1 regime, neglecting the queued region at the right-boundary, we expect the approximated means to be underestimated. However, in Figure 4.5a, approximated means are slightly overestimated. The reason for this might be the effective hopping probability is slightly lower than what we assume to be true. When we consider the fitted NWLS parameters in Table 4.1, the coefficient of L is slightly higher for empirical mean travel times than the approximated mean travel time while the exponent of L is slightly lower. The exponent of L should always be one. However, the NWLS fits for the exponent of L of the empirical travel time is slightly lower. Hence the coefficient of L gets slightly higher.

The approximated variances are mostly similar in the very low-density region in the LD2 regime (see Figure 4.6b). Having very low inflow probability and high outflow probability, the link has a uniform density profile. The interaction between vehicles is also minimal and the deceleration of vehicles is strictly due to the deceleration probability p. Therefore, in very low-density regime, vehicles have uniform hopping probabilities throughout. In the LD1 case, even though the approximated and empirical variances are much closer, the slope is different. NWLS estimates in Table 4.2 also reveal that the variances much differ, as the coefficient of L for approximated variances is lower than the empirical variances. The approximated and empirical variances are much closer in the HD1 and HD2 regime. However, the fitted NWLS parameters are much different.

In LD2 regime our model underestimated the variance while overestimated in the MC regime. See Figures 4.6c and 4.6f. However, the exponent terms of L in the NWLS estimate for approximated variances are slightly closer for all the density regimes except for LD1 (see Table 4.2).

We can expect a slightly skewed travel time distribution in the LD1 regime of the ASEP with stochastic updates due to the slight increase in the local density at the right boundary as shown in Figure 4.7a. However, as the size of the link increases, the increment in the density profile becomes negligible. Therefore, the travel time distribution of a vehicle in the LD1 regime of the ASEP with stochastic updates converges to a normal distribution as the link size increases. See Figure 4.7b.



(a) L=50 cells (b) L=250 cells

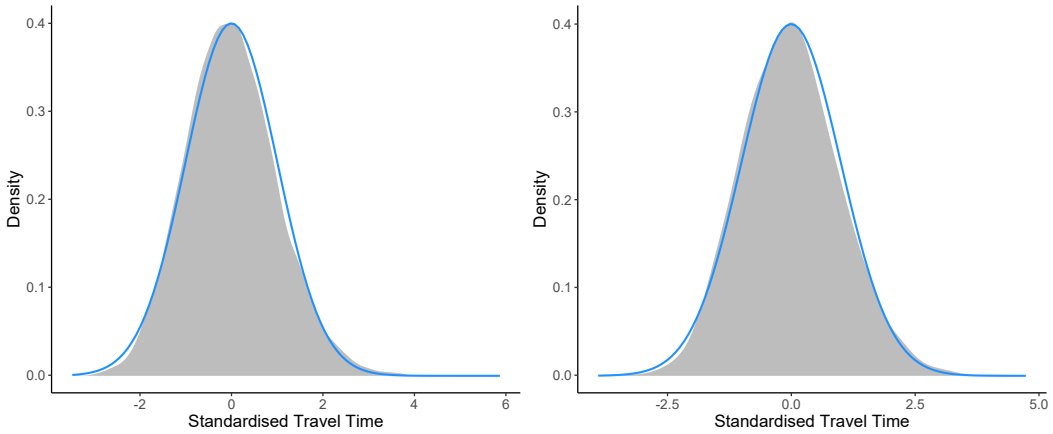
FIGURE 4.7: Travel time distribution of ASEP when the system is in the LD1: $\alpha = 0.1$ and $\beta = 0.3$. Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.

Table 4.2: Non-linear WLS Regression Fit for empirical and approximated σ^2 of travel time of an ASEP with varying link sizes. Formula: $\sigma^2 = b \times L^c$

		Empirical	Approximation
LD1: $\alpha = 0.1$, $\beta = 0.2$	b	$8.6041^{***}[\pm 0.243662]$	$3.552^{***}[\pm 0.00 \times 10^{-15}]$
	c	$0.8082^{***}[\pm 0.005806]$	$1^{***}[\pm 0.00 \times 10^{-16}]$
LD2: $\alpha = 0.001$, $\beta = 1$	b	$1.554^{***}[\pm 0.114800]$	$2.012^{***}[\pm 1.391 \times 10^{-14}]$
	c	$1.058^{***}[\pm 0.013330]$	$1^{***}[\pm 1.157 \times 10^{-15}]$
LD2: $\alpha = 0.1$, $\beta = 0.6$	b	$3.7204^{***}[\pm 0.148277]$	$3.552^{***}[\pm 3.989 \times 10^{-15}]$
	c	$1.1359^{***}[\pm 0.007475]$	$1^{***}[\pm 1.878 \times 10^{-16}]$
HD1: $\alpha = 0.2$, $\beta = 0.1$	b	$93.3491^{***}[\pm 1.359440]$	$107.7^{***}[\pm 4.036 \times 10^{-13}]$
	c	$0.9938^{***}[\pm 0.003420]$	$1^{***}[\pm 7.373 \times 10^{-16}]$
HD2: $\alpha = 0.8$, $\beta = 0.2$	b	$23.1890^{***}[\pm 4.530700]$	$27.31^{***}[\pm 7.951 \times 10^{-14}]$
	c	$1.0014^{***}[\pm 0.038300]$	$1^{***}[\pm 4.869 \times 10^{-16}]$
MC: $\alpha = 0.4$, $\beta = 0.7$	b	$6.5677^{***}[\pm 0.355548]$	$12^{***}[\pm 5.565 \times 10^{-14}]$
	c	$0.9932^{***}[\pm 0.009999]$	$1^{***}[\pm 7.756 \times 10^{-16}]$

– – –

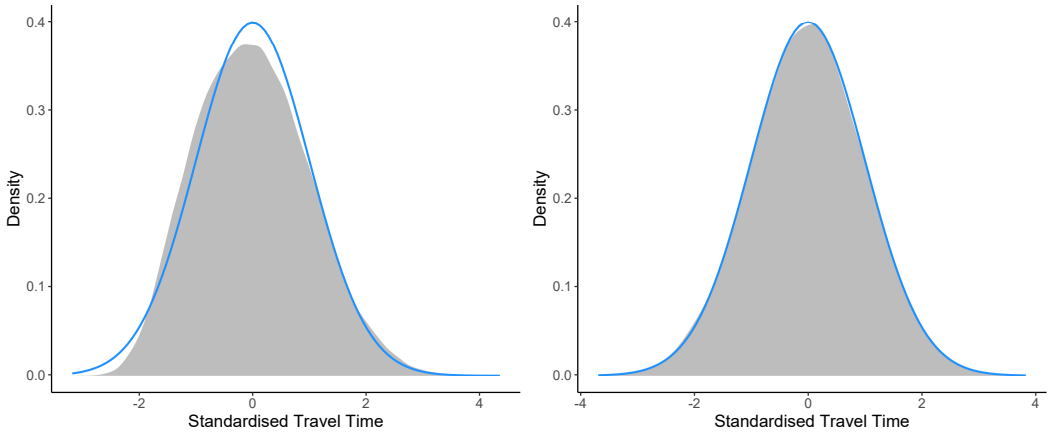
*** significance level is ≈ 0



(a) L=100 cells (b) L=500 cells

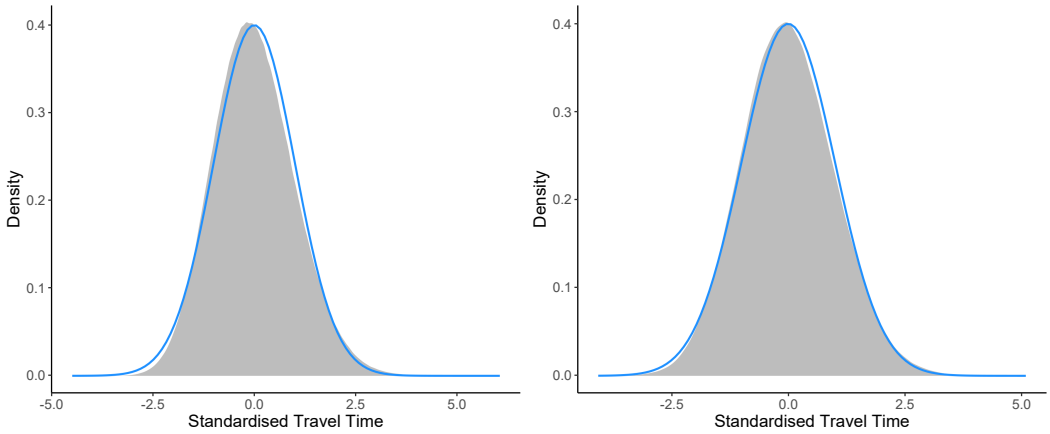
FIGURE 4.8: Standardised travel time distribution of ASEP when the system is in the LD2 regime of an ASEP ($\alpha = 0.001$ and $\beta = 1$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.

Figure 4.8b illustrates the standardised travel time distribution for an ASEP in the LD2 regime with a very low-density. The travel time distribution converges to a normal distribution irrespective of the link size since the local density profile is always flat. However, for smaller link sizes with a slightly higher density (i.e., $\alpha = 0.1$ and $\beta = 0.6$), the travel time distribution is slightly skewed, but converges to a normal distribution as the link size increases (see Figure 4.9).



(a) L=100 cells (b) L=500 cells

FIGURE 4.9: Standardised travel time distribution of ASEP when the system is in the LD2 regime ($\alpha = 0.1$ and $\beta = 0.6$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.

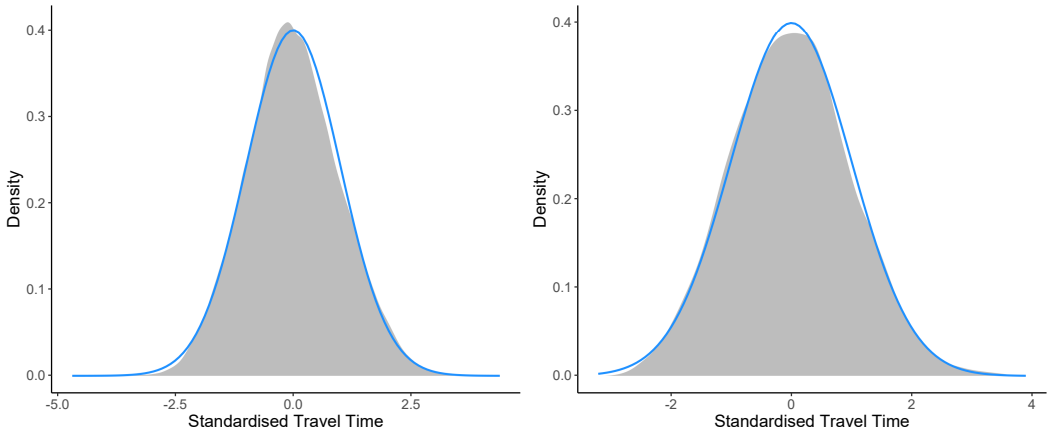


(a) L=50 cells (b) L=200 cells

FIGURE 4.10: Standardised travel time distribution of ASEP when the system is in the HD1 regime ($\alpha = 0.2$ and $\beta = 0.1$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.

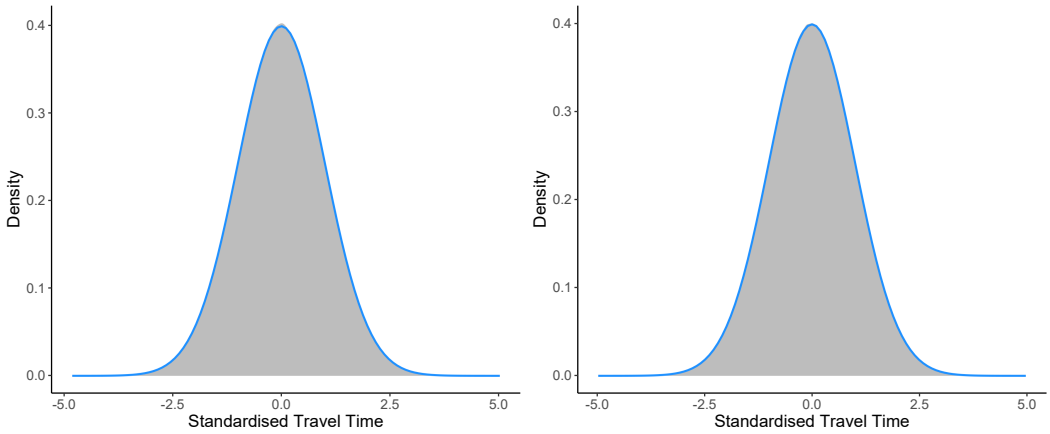
The local density profile of the ASEP with stochastic updates in the HD1 regime also has a slight increase at the left-boundary. For shorter links, the impact of heterogeneous density region significantly affects to the approximation of travel time distribution since we assume the density profile to be homogeneous. The empirical travel time distribution is slightly skewed as illustrated in Figure 4.10a. However, for longer links (i.e. L is large enough), the impact of the ignoring the heterogeneous region is insignificant and the travel time distribution converges to a normal distribution. See Figure 4.10b. Irrespective of the link size, in the HD2 regime where the heterogeneous density region is much smaller, the travel time distribution approximately follows a normal distribution as illustrated in Figure 4.11.

When the system is in the MC regime, the density profile has a horizontal integral shape where the density is slightly deviating from the bulk density at the boundaries. For longer links, the impacts of the two regions possibly cancel out due to the density in one side is higher and the other is lower, and the density profile is symmetrical. Therefore, as illustrated in Figure 4.12, the link travel time distribution converges to a normal distribution as the link size increases.



(a) L=100 cells (b) L=500 cells

FIGURE 4.11: Standardised travel time distribution of an ASEP when the system is in the HD2 regime ($\alpha = 0.2$ and $\beta = 0.8$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.



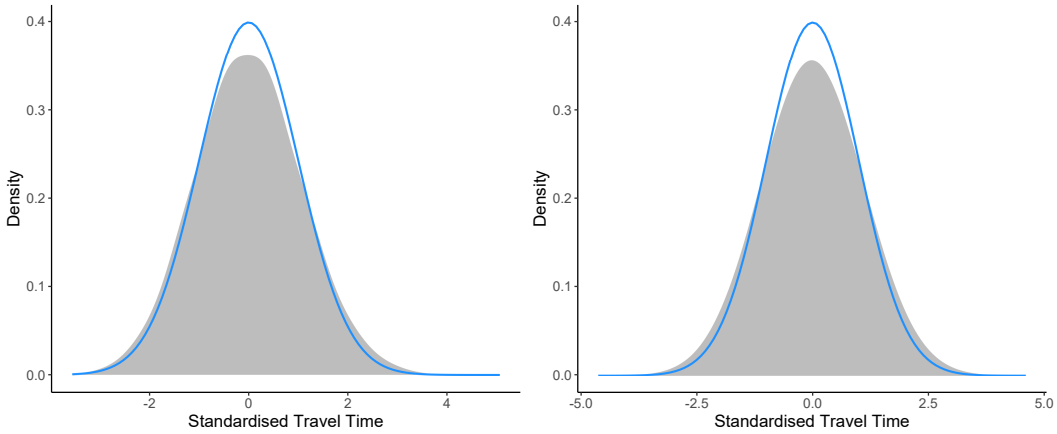
(a) L=100 cells (b) L=500 cells

FIGURE 4.12: Travel time distribution of an ASEP when the system is in the MC regime ($\alpha = 0.4$ and $\beta = 0.7$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non parametric Gaussian kernel.

4.2.3 Simulations: NaSch

We evaluate the travel time distribution of NaSch for two different density regions: LD ($\alpha < \beta$) and HD ($\beta < \alpha$). We simulate the NaSch model with $v_{max} = 3$ and $p = 0.5$ for different inflow and outflow probability combinations: $\alpha = 0.001$ and $\beta = 1$ representing a link with very low-density region, $\alpha = 0.1$ and $\beta = 0.6$ representing a low-density region, $\alpha = 0.4$ and $\beta = 0.7$

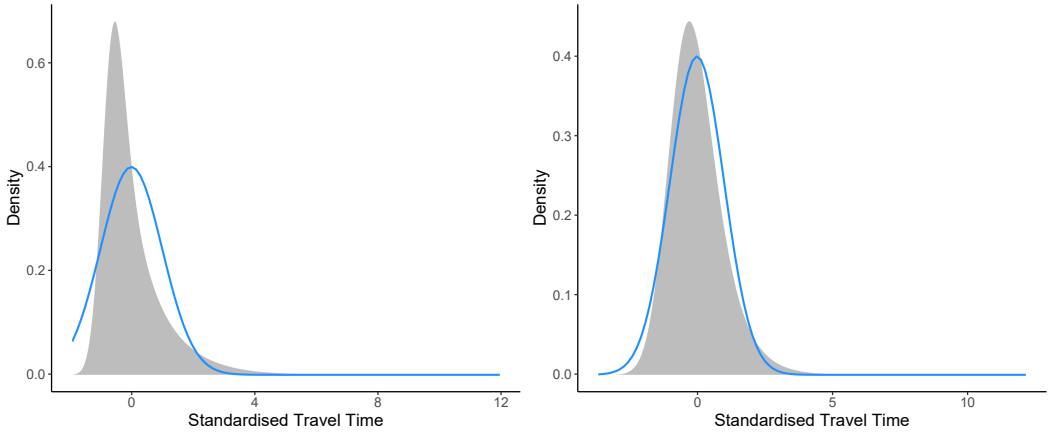
representing a maximum-current region and $\alpha = 0.8$ and $\beta = 0.2$ representing a high-density region.



(a) L=100 cells (b) L=1000 cells

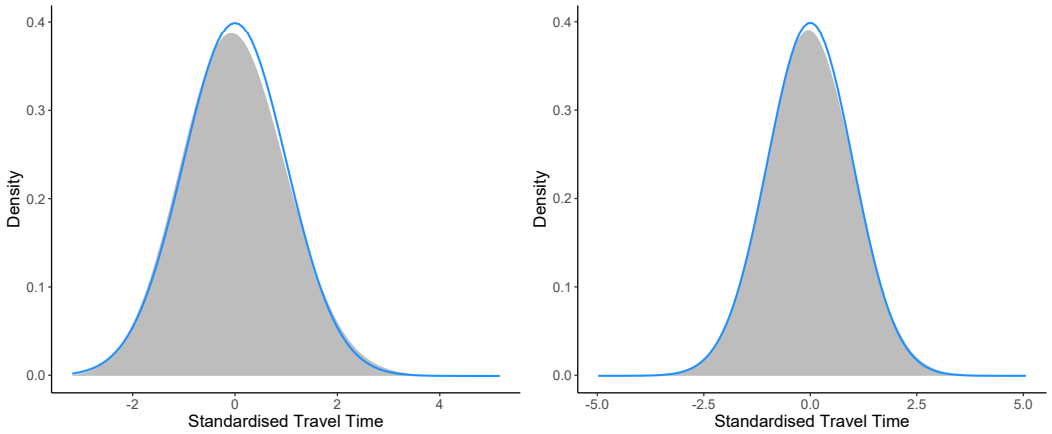
FIGURE 4.13: Travel time distribution of NaSch when the system is in very low-density region ($\alpha = 0.001$ and $\beta = 1$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.

We simulate all four systems for approximately 5000 seconds, ensuring that stationarity was always reached. We record the number of time steps that a vehicle remains in the system. Figures 4.13, 4.14, 4.15 and 4.16 illustrate the standardised travel time distributions for different density regions. When the links are congested, the travel time distribution converges to a normal distribution, irrespective of the link size.



(a) L=100 cells (b) L=1000 cells

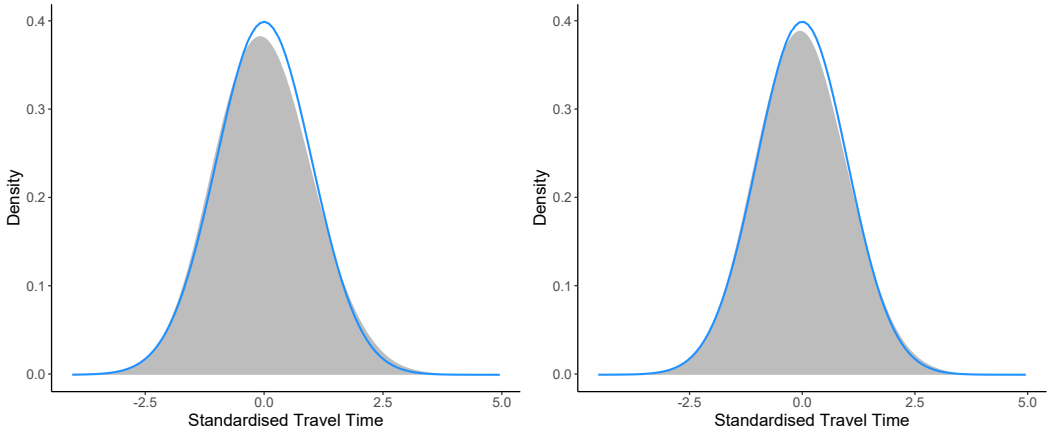
FIGURE 4.14: Travel time distribution of NaSch when the system is in low-density region ($\alpha = 0.1$ and $\beta = 0.6$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.



(a) L=100 cells (b) L=300 cells

FIGURE 4.15: Travel time distribution of NaSch when the system is in maximum-current region ($\alpha = 0.4$ and $\beta = 0.7$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.

When the link is not congested, with link sizes, the link travel time distribution is skewed. However, as the link size increases, the travel time distribution converges to a normal distribution as illustrated in Figure 4.14. In the very low-density regime where the interaction between vehicles is minimal, the travel time distribution is closer to a symmetric bell shaped distribution and converges to a normal distribution, as link sizes increase. See Figure 4.13. When the link is congested, the link travel time distribution converges to a normal distribution irrespective of the link size as represented in Figures 4.15 and 4.16.



(a) L=100 cells (b) L=300 cells

FIGURE 4.16: Travel time distribution of NaSch when the system is in high-density region ($\alpha = 0.8$ and $\beta = 0.2$). Blue curve represents the standard normal distribution. Standardised empirical distributions are smoothed with non-parametric Gaussian kernel.

4.3 Signalised links

As we discussed in Section 4.2, the travel time distribution on a link can be well approximated by a normal distribution when the impacts from the other factors such as signalised intersections, lane changing and road geometry are minimal. However, in practice, the effect of traffic lights on the link travel time cannot be neglected as it changes the travel time distribution dramatically. An extended period of time at a red light may cause vehicle queue accumulation at the intersection and heavy congestion on the upstream link, whereas prolonged green time allows more vehicles to pass through the intersection and less congestion on the upstream link. Moreover, the linking pattern of the traffic signals at consecutive intersections also has a significant impact on the delay imposed at the intersections. A less obvious observation is that poor offset choices lead to multi-modal travel time distributions.

As recent studies suggest, rather than having a uni-modal distribution to describe travel time, using different components to represent different traffic phases seems more reliable. In this section we present our findings on the distribution of link travel time with traffic lights at both upstream and downstream intersections. We consider two different congestion levels, low and high-density utilising the NaSch. In addition, we investigate the effect of varying offsets and green times at the intersections on the link travel time distribution.

4.3.1 Models

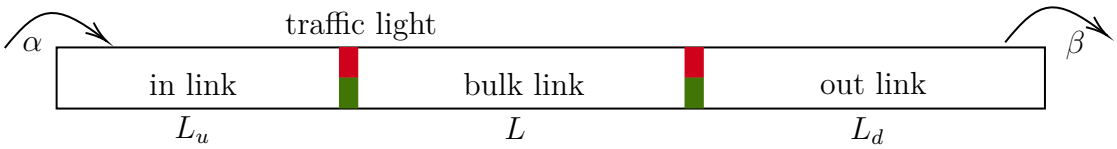


FIGURE 4.17: Link Configuration: Parameters α and β represents the traffic inflow and outflow probabilities respectively.

4.3.1.1 System configuration

We consider a link of total size $L_u + L + L_d$ cells. See Figure 4.17. Intersections with traffic lights are placed at locations L_u and $L_u + L$, allowing a bulk link

of length L in the middle of the system. A vehicle is inserted to the system with some inflow probability, $\alpha \in (0, 1)$ at the in-boundary (left-boundary) and removed from the out-boundary (right-boundary) with some outflow probability, $\beta \in (0, 1)$. A vehicle occupies one cell, take $v_{max} + 1$ values of speed $v = 0, 1, 2, ..., v_{max}$ with some deceleration probability p. Each vehicle is move forward according to its velocity obeying NaSch dynamics. Refer Section 2.1.1. We consider a fixed cycle length c at both intersections, green times g_1 and g_2, at upstream and downstream intersections, respectively, and offset d (the time delay of the start of the green light at the downstream intersection). We record the travel time of each vehicle passing through the bulk link. Furthermore, we record the queue length, that is the number of vehicles stopped during the red time at the downstream intersection, at the beginning of every green phase.

We set $L = 100$ cells, $L_u = L_d = 100$ cells. Aiming an average free flow speed of 50 km/h, the speed and the deceleration probability of the model were fixed at $v_{max} = 2$ cells/s and $p = 0.1$ respectively. We consider two different inflow probabilities, $\alpha = 0.1$ and 0.5. Outflow probability is $\beta = 0.9$ for both the cases. We consider a fix cycle length, $c = 90$s and four different green time combinations at upstream and downstream intersections, case $45s - 45s$ where $g_1 = 45s$ and $g_2 = 45s$, case $45s - 60s$ where $g_1 = 45s$ and $g_2 = 60s$, case $60s - 45s$ where $g_1 = 60s$ and $g_2 = 45s$ and case $60s - 60s$ where $g_1 = 60s$ and $g_2 = 60s$. We consider different offsets, where $d \in \{5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$ seconds.

4.3.1.2 Gaussian mixture distribution

A Gaussian mixture distribution is a multivariate distribution that consists of a weighted sum of Gaussian component densities [117]. Each component is defined by its mean μ and variance σ^2. The mixture is defined by a vector of mixing proportions. The probability density function of a Gaussian mixture model is given by

$$p(x \mid \lambda) = \sum_{i=1}^{M} \lambda_i g(x|\mu_i, \sigma_i), \ \sum_{i=1}^{M} \lambda_i = 1, \tag{4.4}$$

where x is a continuous data vector, $\lambda_i, i = 1, 2, ..., M$, are the mixture weights, and $g(x|\mu_i, \sigma_i), i = 1, 2, ..., M$ are the components of Gaussian densities.

The Gaussian mixture model (GMM) for modelling the travel time distribution assumes that vehicle travel times are from a Gaussian mixture distribution

with unknown parameters and uses the expectation-maximization (EM) algorithm to fit the mixture distribution. It was demonstrated in [100, 104] to well model travel time distributions in certain traffic conditions. However, none of the studies investigated the impacts of traffic signals including cycle length, phase splits and offset. In this section, we systematically study the parameters μ and λ in regards to cycle time c, splits (g_1, g_2) and offset d. In addition, we apply the domain wall model to approximate μ and λ as a function of the signal parameters. .

4.3.1.3 Domain wall model

The domain wall model is first introduced by [118] and has been successfully applied in modelling traffic flow behaviours in both stationary and transient states [63, 119] at a macroscopic level. A domain wall separates two regions of different densities. We can assume that the width of such a wall is small compared to the macroscopic system size, so that we can consider the position of the domain wall to be a single point. In the current study, we consider a deterministic variant of the domain wall model, instead of a stochastic version [63]. That means the movement of a domain wall is represented by a deterministic drift velocity instead of a (biased) random walk. Specifically, a wall $A|B$ separating domains A and B (A is upstream of B) with density[2] and flow[3] (ρ_A, J_A) and (ρ_B, J_B) respectively moves with a collective velocity of

$$v_{A|B} = \frac{J_A - J_B}{\rho_A - \rho_B}. \tag{4.5}$$

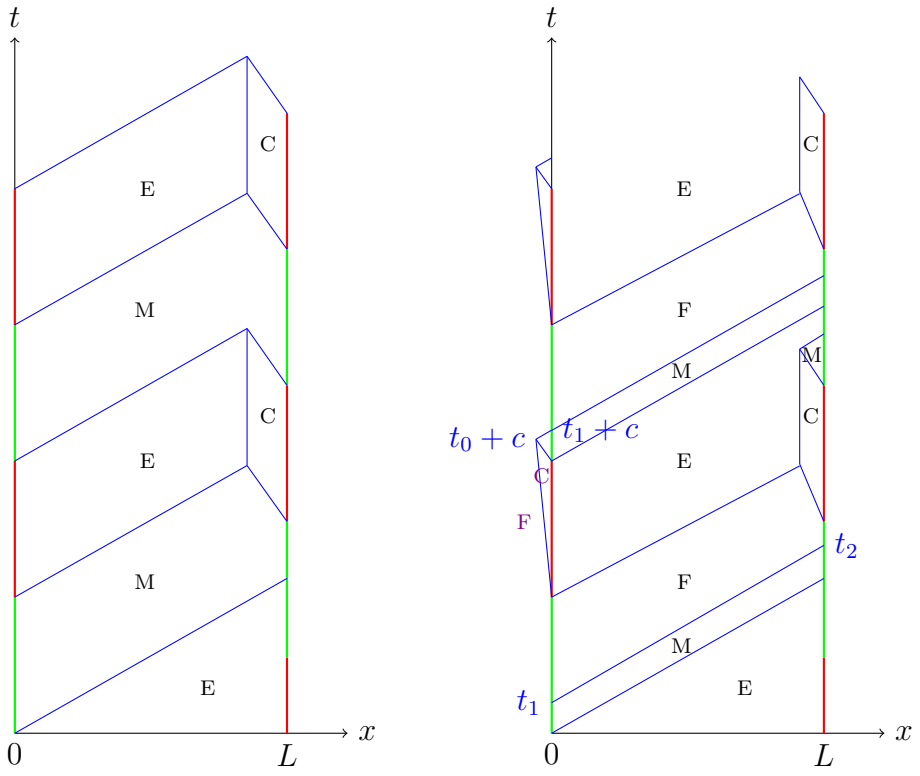
Note a negative $v_{A|B}$ implies that the wall moves upstream. A detailed description of the key properties of the domain wall model tailored for the traffic-light boundary conditions can be found in [119]. The remaining section explains how we approximate the GMM parameters using the deterministic domain wall model.

Now we describe the traffic flow during a couple of traffic cycles using the deterministic domain wall. We assume that the (bulk) link is initially in an empty regime, denoted by E with density $\rho_E = 0$ and flow $J_E = 0$, before the start of green phase at the upstream intersection. We will relax this assumption later. Once the green light starts at the upstream intersection, it creates a maximum-flow region, denoted by M characterised by maximum-flow J_c and critical density ρ_c, in the link due to the residual queue accumulated during the red time at the upstream intersection. Consequently, a domain wall $W_{M|E}$

[2]the number of vehicles occupying a unit length (e.g. meter) of roadway

[3]the number of vehicles passing a point per second

forms separating the regions M and E and drifts downstream. And eventually the wall will reach the downstream intersection, which takes $L/v_{M|E}$ seconds. Depending on the signal offset, the wall may reach the downstream intersection during the red or green phase. In Figure 4.18a, where the offset is 25s, and the green time and red time at both intersections are 45s, the wall $W_{M|E}$ arrives at the downstream intersection during the green phase. In the rest of the green time, some vehicles pass through the downstream intersection. Once the light turns red, it will create a congested region (i.e., a queue) denoted by C with flow $J_C = 0$ and density $\rho_C = 1$. A domain wall $W_{M|C}$ forms separating the regions M and C and moves upstream.



(a) One-phase argument

(b) Two-phase argument

FIGURE 4.18: Domain wall trajectories for $c = 90s$, $g_1 = g_2 = c/2$ and $d = 25s$. X-axis: position in the bulk link; Y-axis: time. Green and red vertical lines represent the traffic lights at upstream and downstream intersections.

In the meantime, when the upstream light turns red, an empty region forms in the bulk link due to the zero inflow. And the domain wall $W_{E|M}$, which separates the regions M and E, drifts towards the downstream intersection. The two walls, $W_{M|C}$ and $W_{M|E}$, eventually meet and merge together. As a result, the queue at the downstream intersection during the red time stops growing. When the downstream intersection starts a green light again, the

queue will start to resolve with the movement of a wall $W_{C|M}$. The start of green light at the upstream intersection will again create a maximum-flow region with a domain wall $W_{M|E}$. Because the green times are exactly the same, in the deterministic case, the domain walls, $W_{M|E}$ and $W_{C|E}$ will meet at the end of the queue, leading to a homogeneous traffic state M in the bulk link. This pattern will repeat in the following cycles. Now we can relax the assumption of having an initial empty state to having a residual queue at the downstream intersection which is sufficiently short so that the traverse pattern of the domain wall $W_{M|E}$ is not affected.

According to the domain wall dynamics discussed above, there are two possible batches of vehicles traversing the bulk link: without experiencing delay (i.e., through a *green wave*[4]); and delayed by red light once. The vehicles that travel in the first batch through the maximum-flow has an average travel time of

$$T_{\text{free}} = \frac{L}{v_{M|E}}, \tag{4.6}$$

where $v_{M|E}$ is the speed of of $W_{M|E}$, $v_{M|E} = \frac{J_c}{\rho_c}$. The delayed vehicles have a higher travel times due to the waiting time at the red light of the downstream intersection. The travel time of the vehicle waiting a whole red phase is $T_{\text{free}} + (c - g_2)$ seconds. The last vehicle passing by on the verge of the traffic light turning red has the minimum waiting time at the downstream intersection. Therefore, the minimum expected travel time of some vehicle missing a green phase is $\max(T_{\text{free}}, c + d - g_1)$ seconds. Hence, the average travel time of the vehicles in the second batch can be approximated by

$$T_{\text{delay}} = \frac{T_{\text{free}} + 2c + d - (g_1 + g_2)}{2}. \tag{4.7}$$

Furthermore, the proportion of the vehicles joining the green wave is given by

$$\tilde{\lambda}_1 = \frac{d + g_2 - T_{\text{free}}}{c}. \tag{4.8}$$

From the above statement, which is called *one-phase domain wall argument*, there are two batches of traffic corresponding to two components of the Gaussian mixture distribution. The undelayed traffic (or green wave traffic) with the average travel time $\tilde{\mu}_1 = T_{\text{free}}$ accounting $\tilde{\lambda}_1$; and the delayed traffic with $\tilde{\lambda}_2 = T_{\text{delay}}$ accounting $\tilde{\lambda}_2 = 1 - \tilde{\lambda}_1$.

We remark that the one-phase argument assumed that the residual queue at the the upstream intersection is sufficiently long to support a maximum-flow domain throughout the green phase. However, if the traffic demand

[4]a continuous traffic flow through upstream and downstream intersections without facing a red light

represented by the inflow rate α is low, and the accumulated traffic at the upstream intersection is not sufficient to supply maximum-flow traffic during the green time, then at some point, the M region will be replaced by a lower flow region, say F, which is determined by the value of α. This leads to a *two-phase domain wall argument*. Figure 4.18b shows the domain wall trajectories produced by the two-phase argument for the same system as Figure 4.18a.

Let ρ_0 and J_0, respectively, represent the density and flow of region F given the inflow rate α. Let q be the queue length immediately before the start of the green light at the upstream intersection. The queue grows as its upstream end propagates with wall $W_{F|C}$. The green light will create wall $W_{C|M}$, and the queue region disappears when the two walls meet. In the meantime, wall $W_{F|M}$ forms, specifically, t_0 seconds after the green light starts

$$t_0 = \frac{q}{v_{F|C} - v_{C|M}} \tag{4.9}$$

where $v_{C|M} = \frac{-J_c}{1-\rho_c}$ and $v_{F|C} = \frac{J_0}{\rho_0 - 1}$. The wall $W_{F|M}$ reaches the upstream intersection t_1 seconds after the start of the green light at the upstream intersection and propagates downstream with speed $v_{F|M}$. To calculate t_1, we know wall $W_{F|M}$ is $t_0 \cdot |v_{C|M}|$ upstream of the upstream intersection when it starts, and so it arrives at the upstream intersection at time

$$t_1 = t_0 + \frac{t_0 \cdot |v_{C|M}|}{v_{F|M}} \tag{4.10}$$

where $v_{F|M} = \frac{J_0 - J_c}{\rho_0 - \rho_c}$. Furthermore, it is expected to arrive at the downstream intersection at time

$$t_2 = t_1 + \frac{L}{v_{F|M}}. \tag{4.11}$$

Because the queue at the upstream intersection dissolves during the green phase and builds up during the red phase with a rate J_0, its length by the next green phase is $q = J_0(c - g_1)$. Depending on whether t_1 is before or after the end of the green light at the downstream intersection, we have two possible cases to discuss for the undelayed traffic.

1. If $t_1 \geq d + g_2$, then the amount of traffic that travels through the bulk link without delay is $J_c(d + g_2 - T_{\text{free}}) = \tilde{\lambda}_1 J_0 c$. Then, we have

$$\tilde{\lambda}_1 = \frac{J_c(d + g_2 - T_{\text{free}})}{J_0 c}. \tag{4.12}$$

2. Otherwise, the undelayed traffic consists of two parts: $J_c(t_2 - T_{\text{free}})$ (from the maximum-flow domain M) and $J_0(d + g_2 - T_{\text{free}} - (t_2 - T_{\text{free}}))$ (from the low-density domain F). Thus,

$$\tilde{\lambda}_1 = \frac{J_c(t_1 - T_{\text{free}}) + J_0(d + g_2 - t_2)}{J_0 c}. \tag{4.13}$$

Both domain wall arguments distinguish the traffic into two batches; undelayed and delayed. Therefore, according to the above statements, the travel time distribution can be approximated by the Gaussian mixture model with two components: the first one represents the undelayed vehicles, with mean $\tilde{\lambda}_1 = T_{\text{free}}$, deviation σ_1 and weight $\tilde{\lambda}_1$, while the second one represents the delayed vehicles with mean $\tilde{\mu}_2 = T_{\text{delay}}$, deviation σ_2 and weight $\tilde{\lambda}_2 = 1 - \tilde{\lambda}_1$. Because the domain wall model is deterministic, it is unable to estimate the deviations σ_1 and σ_2. Due to the same reason, it cannot approximate more than two components for the GMM.

4.3.2 System with a lower inflow: $\alpha = 0.1$ and $\beta = 0.9$

We simulated the system described in Section 4.3.1.1, considering a low inflow ($\alpha = 0.1$ and $\beta = 0.9$) for each distinct green time combination for different offsets (i.e., $d \in \{5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$). Each simulation consists of ten different independent simulations. We record the travel time of each vehicle and the queue length at downstream for 5000s, Figure 4.19, plots the average queue length at the start of green time for all the four green time combinations at the upstream and downstream intersections against the offset.

Having a longer red time creates a longer queue at the downstream intersection (Case 45s-45s and Case 60s-45s in Figure 4.19) and we can expect more delays in travel time. It is evident that a longer green time reduced the queue accumulation at downstream intersection allowing more vehicles to pass through. The proportion of vehicles pass through the downstream intersection without delaying at the red is maximum at the optimum offset. Correspondingly it is evident that having an offset closer to the optimum, helps to reduce the queue accumulation allowing more vehicles to join the green wave. Case 45s-45s and Case 60s-60s having the same green time at both the upstream and downstream intersections have similar queue accumulation after offset of 30s. In the case 45s-60s, having a longer green time at downstream intersection has shorter queues and has a range of optimal offsets from 40s to 50s. In other three cases, we have an optimal offset around 50s.

4.3.2.1 Fitting Gaussian mixture models

We fit a Gaussian mixture model to the travel time data for each case separately. λ_i, μ_i and σ_i represents the model parameters for each component $i = 1, 2, ...M$. Furthermore, we approximated μ_1, μ_2, λ_1 and λ_2 utilising the models we defined in Section 4.3.1.3. Here, the terms $App\_\mu_1, App\_\mu_2, App\_\lambda_1$ and App_λ_2 rep-

resents the approximated parameters for μ_1, μ_2, λ_1 and λ_2 respectively. When the travel time distribution consists of more than two peaks, we did not approximate the Gaussian mixture parameters. The model we proposed in Section 4.3.1.3 is reasonable only when we have two components.

In this section, all the figures related to travel time distribution shows two Gaussian mixtures curves, “Approximation" and “Gaussian Mixture". In the “Gaussian Mixture", all the parameters were left free while in the “Approximation", we fixed some of the parameters. We approximated, μ_1 by the Equation 4.6, μ_2 by the Equation 4.7, λ_1 by the Equation 4.13 and λ_2 by $1 - \lambda_1$.

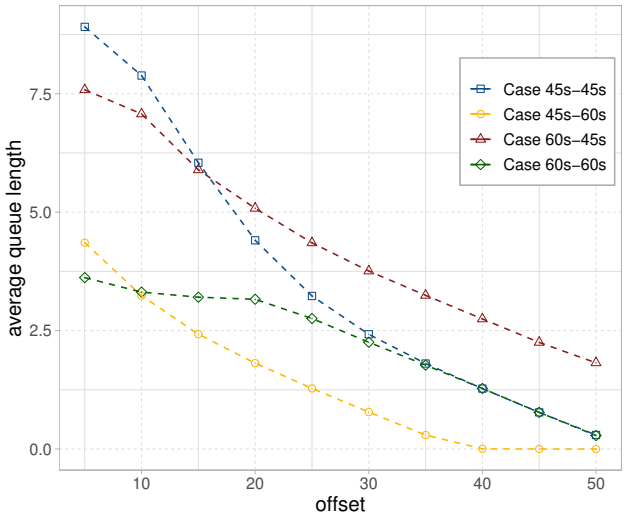
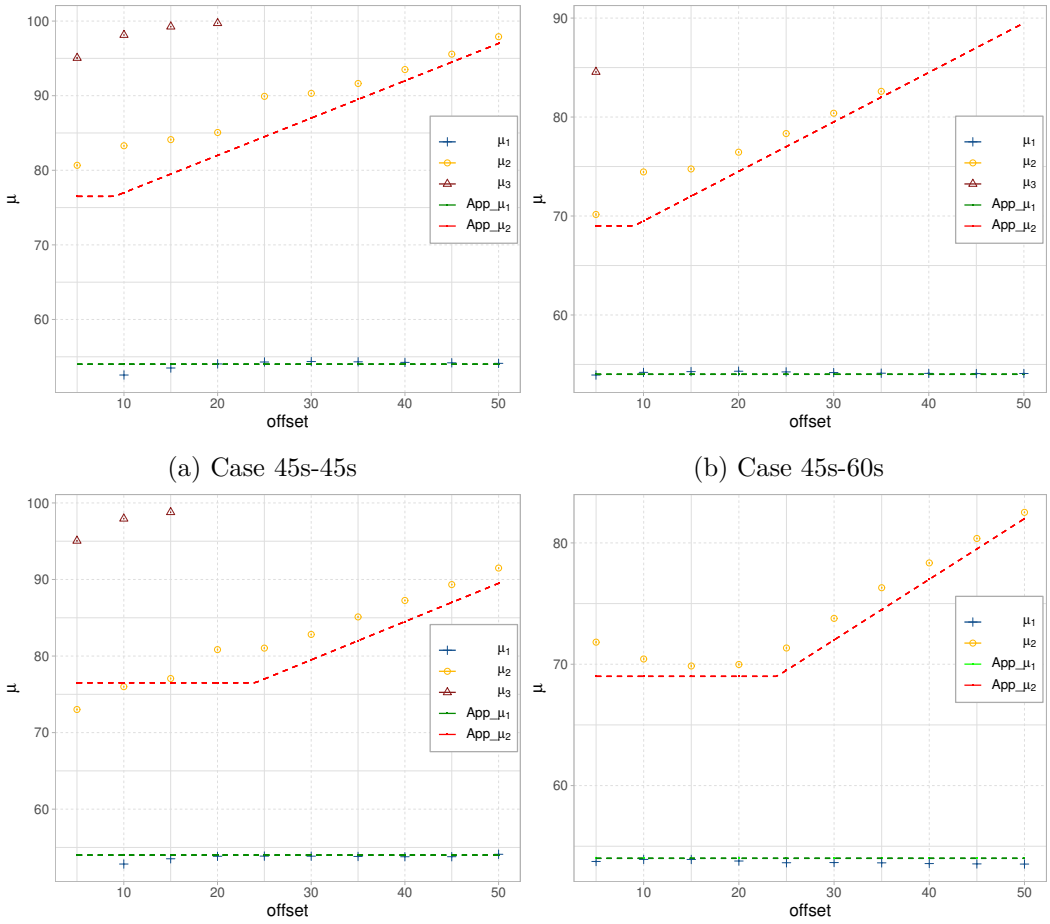


FIGURE 4.19: Association between empirical queue length and the offset

In Figure 4.20, we plot the means of each Gaussian mixture component, μ_i against the offset, d. When the system is less congested, vehicles will at most wait for one cycle to pass through the bulk link even with the worst offset ($d = 0$). The vehicles that join the green wave have an average travel time closer the travel time in free flow speed. In Case 60s-60s, having a longer green time at the downstream intersection, allows majority of the vehicles to pass the downstream intersection with less delays. Therefore, none of the vehicle waits for a full cycle at the downstream intersection. The mean travel time of the green wave is invariant with the offset while the proportion of vehicles join the green wave getting increases, until the offset closer to the optimal (See Figure4.21, the mixing proportion of Gaussian mixture, λ_i against the offset, d). Although, the mean travel time of the delayed vehicles getting increases with the offset while proportion getting decreases. At the optimal offset, all the vehicles join the green wave with having only a very small queue at the downstream intersection.



(a) Case 45s-45s

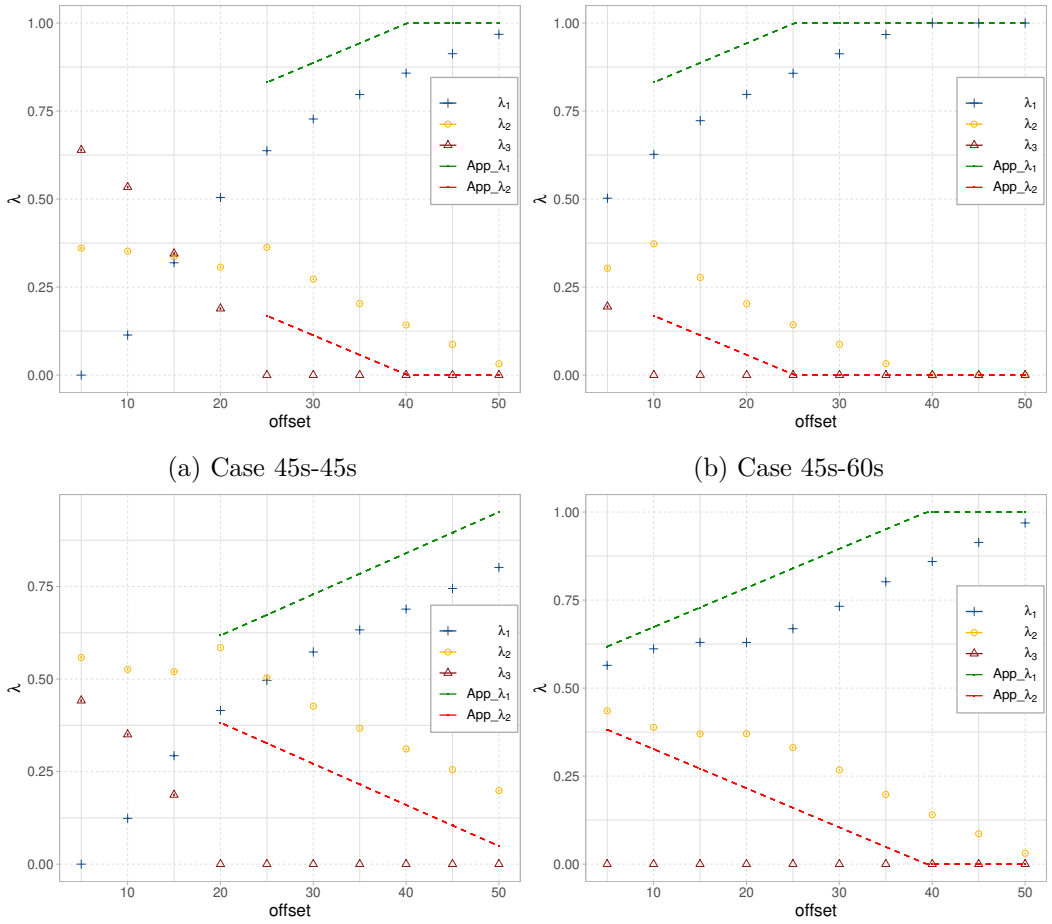
(b) Case 45s-60s

(c) Case 60s-45s

(d) Case 60s-60s

FIGURE 4.20: Variability of estimated μ_i's in Gaussian mixture distribution with the offset. App_μ_1 and App_μ_2 represent the approximated mean travel time with free flow and with waiting for green times at the downstream intersection respectively. The inflow probability $\alpha = 0.1$

The mean approximations, App_μ_1 and App_μ_2 (green and red dashed line in Figure 4.20) of the two clusters, in particular μ_1, are close to the estimated Gaussian mixture parameters, μ_1 and μ_2. The approximation of the mixing proportions, App_λ_1 and App_λ_2 are reasonably close to the estimated Gaussian mixture parameters, λ_1 and λ_2. λ_1 is over-estimated probably because we consider the deterministic domain wall model instead of the stochastic domain wall model.



(a) Case 45s-45s

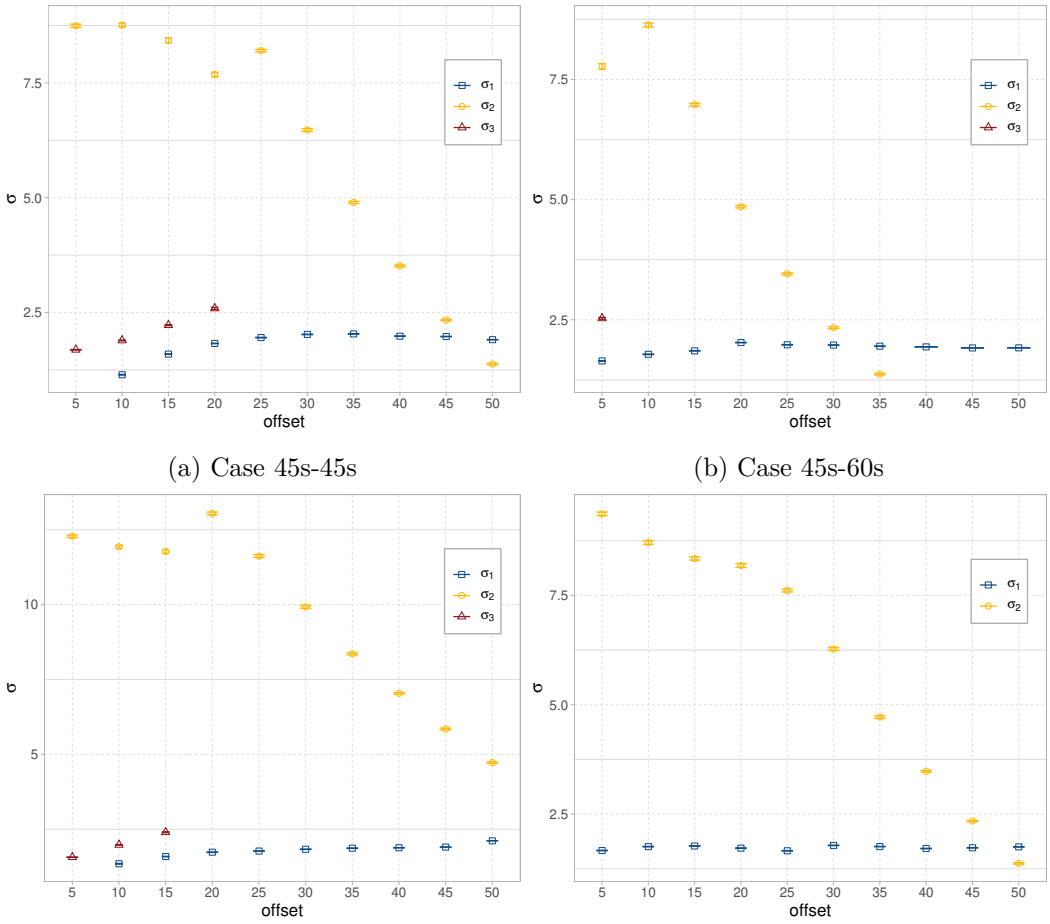
(b) Case 45s-60s

(c) Case 60s-45s

(d) Case 60s-60s

FIGURE 4.21: Variability of estimated λ_i's in Gaussian mixture distribution with the offset. App_λ_1 and App_λ_2 represent the approximated mixing proportions of vehicles joins the green wave and the vehicles misses one cycle respectively. The inflow probability $\alpha = 0.1$

In Figure 4.22, we plot the approximated standard deviations of the Gaussian mixture components against the offset. In all four cases we can see similar variances for the free-flow travel times. The standard deviation of the travel times for vehicles delayed for a whole cycle decreases with the offset.



(a) Case 45s-45s

(b) Case 45s-60s

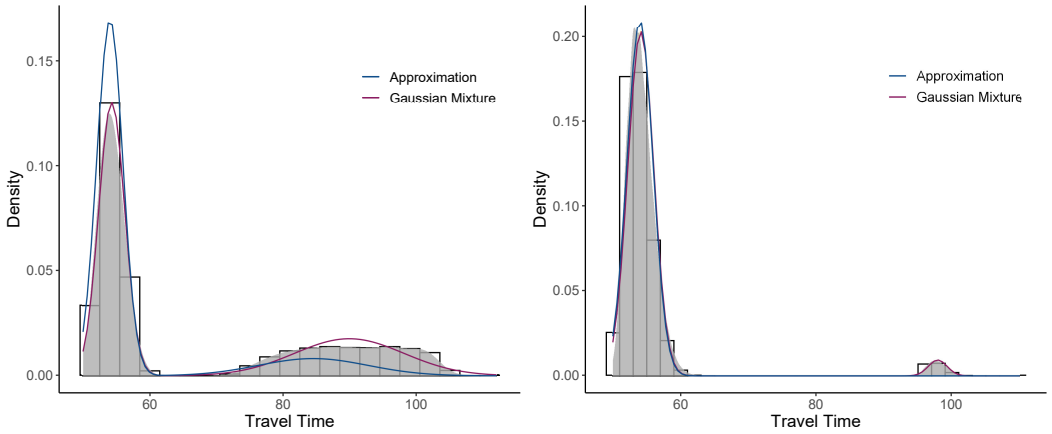
(c) Case 60s-45s

(d) Case 60s-60s

FIGURE 4.22: variability of estimated σ_i's in Gaussian mixture distribution with the offset, The inflow probability $\alpha = 0.1$

Case 45s-45s

Figure 4.23 illustrates the travel time distribution on the bulk link when the green time at both upstream and downstream intersections is 45s with offsets 25s and 50s respectively. The distribution is bi-modal, where the first cluster represent the vehicles travelling in the green wave while the second cluster represents the vehicles queued at the downstream intersection. When the offset getting closer to the optimum ($d = 50s$, best in $\{d\}$), all the vehicles could possibly pass the intersection without any delay and the second cluster will disappeared gradually. Table 4.3 summarises the fitted Gaussian mixture distribution parameters for the travel time for varying offsets in Case 45s-45s. It is evident that the with the offset getting closer to the optimal, the number of components in the Gaussian mixture decreases. In the first four offsets, ($d = 5s, 10s, 15s$ and $20s$), a significant number of vehicles have to wait a full



(a) offset=25s (b) offset=50s

FIGURE 4.23: Travel time distribution: Case 45s-45s in low-density. Approximation represents the Gaussian mixture distribution by replacing μ_1, μ_2, λ_1 and λ_2 with the approximated values. Empirical distributions are smoothed with non-parametric Gaussian kernel.

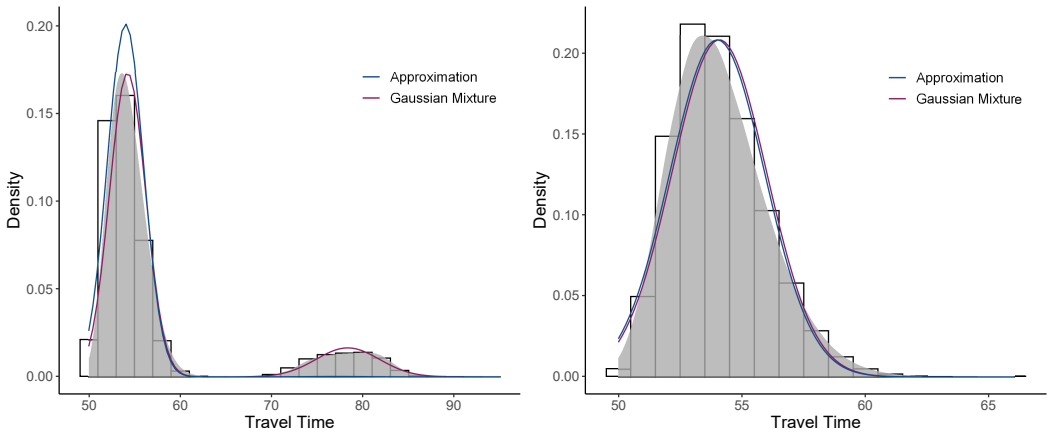
Table 4.3: Gaussian mixture distribution parameter estimation in low-density: case 45s-45s

offset	parameter	component 1	component 2	component 3	K-S test Sig.
5	λ	0	0.361 [±0.002]	0.639 [±0.002]	0.908414
	μ		80.667 [±0.064]	95.053 [±0.007]	
	σ		8.745 [±0.035]	1.689 [±0.005]	
10	λ	0.114 [±0.001]	0.352 [±0.002]	0.534 [±0.002]	0.960542
	μ	52.554 [±0.010]	83.285 [±0.066]	98.131 [±0.009]	
	σ	1.145 [±0.008]	8.761 [±0.043]	1.892 [±0.009]	
15	λ	0.319 [±0.001]	0.336 [±0.002]	0.345 [±0.002]	0.990336
	μ	53.495 [±0.010]	84.109 [±0.068]	99.251 [±0.013]	
	σ	1.594 [±0.007]	8.428 [±0.052]	2.225 [±0.012]	
20	λ	0.504 [±0.001]	0.307 [±0.002]	0.189 [±0.002]	0.947127
	μ	54.039 [±0.007]	85.065 [±0.068]	99.716 [±0.024]	
	σ	1.827 [±0.006]	7.682 [±0.049]	2.589 [±0.023]	
25	λ	0.637 [±0.002]	0.363 [±0.002]	0	0.290679
	μ	54.299 [±0.008]	89.911 [±0.047]		
	σ	1.954 [±0.006]	8.203 [±0.028]		
30	λ	0.727 [±0.002]	0.273 [±0.002]	0	0.463401
	μ	54.359 [±0.009]	90.311 [±0.038]		
	σ	2.024 [±0.006]	6.476 [±0.030]		
35	λ	0.797 [±0.001]	0.203 [±0.001]	0	0.106909
	μ	54.323 [±0.007]	91.636 [±0.034]		
	σ	2.035 [±0.006]	4.896 [±0.024]		
40	λ	0.858 [±0.001]	0.142 [±0.001]	0	0.102769
	μ	54.249 [±0.006]	93.504 [±0.028]		
	σ	1.986 [±0.005]	3.517 [±0.020]		
45	λ	0.913 [±0.001]	0.087 [±0.001]	0	0.069246
	μ	54.190 [±0.006]	95.568 [±0.024]		
	σ	1.977 [±0.005]	2.338 [±0.016]		
50	λ	0.968 [±0.001]	0.032 [±0.001]	0	0.086773
	μ	54.120 [±0.007]	97.903 [±0.027]		
	σ	1.909 [±0.004]	1.376 [±0.018]		

cycle for green time at the intersection, because the green time at the downstream intersection is shorter. Offset larger than 20s and closer to the optimal, the third component is getting slowly disappear as more vehicles join the green wave. With the best offset ($d = 50s$), most of the vehicles join the green wave and hence the second cluster of vehicles wait for green time at the downstream intersection completely disappeared. The non-parametric test, Kolmogorov–Smirnov test confirms that the fitted parameters of Gaussian mixture distribution for the travel time in Case 45s-45s are significant.

Case 45s-60s

The Figure 4.24 illustrates the travel time distribution of the link with longer green time at the downstream intersection (i.e. 60s) than the upstream intersection when the offset is 25s and 50s respectively. When the green and red time are similar, optimum offset is a single value. However, when they are different, optimum offset takes a range of values. Therefore, When the offset is equal to 40s or larger, all the vehicles join the green wave. However, a proportion of vehicles have a slightly longer travel time as they have to join the queue at the downstream junction. When the offsets are closer to the optimum, these two clusters are almost overlapped and not visible as in Figure 4.24. In the Case 45s-60s, the domain wall model over-estimated the $\lambda_1 = 1$. Hence λ_2 is under-estimated, which results in the disappearance of the second peak as in Figure 4.24.



(a) offset=25s (b) offset=50s

FIGURE 4.24: Travel time distribution: Case 45s-60s in low-density. Approximation represents the Gaussian mixture distribution with replacing μ_1, μ_2, λ_1 and λ_2 with the approximated values. Empirical distributions are smoothed with non-parametric Gaussian kernel.

However, if we considered the travel time in Case 45s-60s with offset 50s, actually consists with two clusters where the first cluster represents the vehicles

travel through with the free flow speed while the second cluster represents the vehicles delayed at the downstream intersection for green time, we can fit a Gaussian mixture model with two components as summarised in Table 4.4. The skewness in the distribution plot is due to these two distinct travel time clusters. Figure 4.25 represent the travel time distribution estimation with two components when the offset is 50s. Skewness in the travel time distribution can be well approximated by Gaussian mixture distribution with two components.

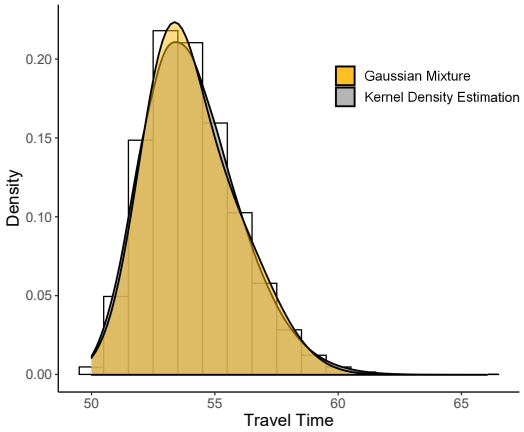


FIGURE 4.25: Histogram of travel time with low-density and offset 50s in case 45s-60s. Gaussian mixture distribution is shown by gold area and the non-parametric Gaussian kernel density estimation is shown by grey area.

Table 4.4: Gaussian mixture distribution parameter estimation for travel time with low-density and offset 50s in case 45s-60s

offset	parameter	component 1	component 2	K-S test Sig.
50	λ	0.541[±0.028]	0.459[±0.028]	0.9981
	μ	53.055[±0.027]	55.300[±0.112]	
	σ	1.260[±0.0174]	1.844[±0.037]	

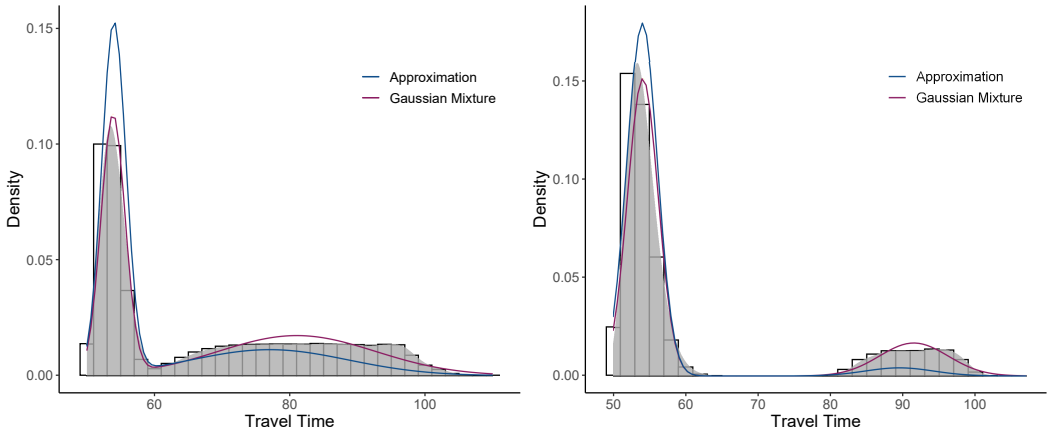
Table 4.5 summarises the fitted Gaussian mixture distribution parameters for the travel time for varying offset in case 45s-60s. The goodness-of-fit parameters also evident that the travel time can be well approximated by a Gaussian mixture distribution. Travel time distribution when the offset is 50s, Gaussian mixture with two components is better than having a one component. Goodness-of-fit test significance value is closer to 1 in Two component case. See Table 4.4.

Table 4.5: Gaussian mixture distribution parameter estimation in low-density: case 45s-60s

offset	parameter	component 1	component 2	component 3	K-S test Sig.
	λ	0.502 [±0.002]	0.303 [±0.002]	0.194 [±0.002]	
5	μ	53.933 [±0.008]	70.157 [±0.080]	84.555 [±0.025]	0.946484
	σ	1.644 [±0.006]	7.773 [±0.058]	2.530 [±0.019]	
	λ	0.627 [±0.002]	0.373 [±0.002]	0	
10	μ	54.203 [±0.007]	74.452 [±0.055]		0.663712
	σ	1.781 [±0.006]	8.635 [±0.040]		
	λ	0.723 [±0.001]	0.277 [±0.001]	0	
15	μ	54.277 [±0.007]	74.755 [±0.046]		0.259413
	σ	1.854 [±0.005]	6.975 [±0.027]		
	λ	0.797 [±0.001]	0.203 [±0.001]	0	
20	μ	54.318 [±0.008]	76.451 [±0.032]		0.236163
	σ	2.026 [±0.004]	4.853 [±0.025]		
	λ	0.857 [±0.001]	0.143 [±0.001]	0	
25	μ	54.247 [±0.008]	78.331 [±0.030]		0.138331
	σ	1.980 [±0.004]	3.458 [±0.020]		
	λ	0.913 [±0.001]	0.087 [±0.001]	0	
30	μ	54.181 [±0.007]	80.386 [±0.026]		0.154814
	σ	1.974 [±0.004]	2.335 [±0.019]		
	λ	0.967 [±0.001]	0.033 [±0.001]	0	
35	μ	54.116 [±0.007]	82.599 [±0.023]		0.106468
	σ	1.951 [±0.005]	1.366 [±0.017]		
	λ	1	0		
40	μ	54.097 [±0.006]			0.069893
	σ	1.936 [±0.004]			
	λ	1	0		
45	μ	54.080 [±0.006]			0.199747
	σ	1.914 [±0.004]			
	λ	1	0		
50	μ	54.086 [±0.006]			0.190069
	σ	1.916 [±0.004]			

Case 60s-45s

Figure 4.26 shows the travel time distribution for Case 60s-45s with offsets, 25s and 50s respectively. A longer green time at the upstream junction allows more vehicles into the bulk link. A shorter green time at downstream intersection lets fewer vehicles pass through. Hence, we can expect more queue accumulation at the downstream intersection. Due to the lower inflow probability, $\alpha = 0.1$, in this setting the longer green time at the upstream intersection did not drive the bulk link into congestion regime. The mass of the second cluster reduces as the offset increases similar to Case 45s-45s. In contrast, the mass of the second cluster is much larger than Case 45s-45s as shown in Figure 4.26, which is due to the increased congestion incurred by the longer green time at upstream intersection.



(a) offset=25s (b) offset=50s

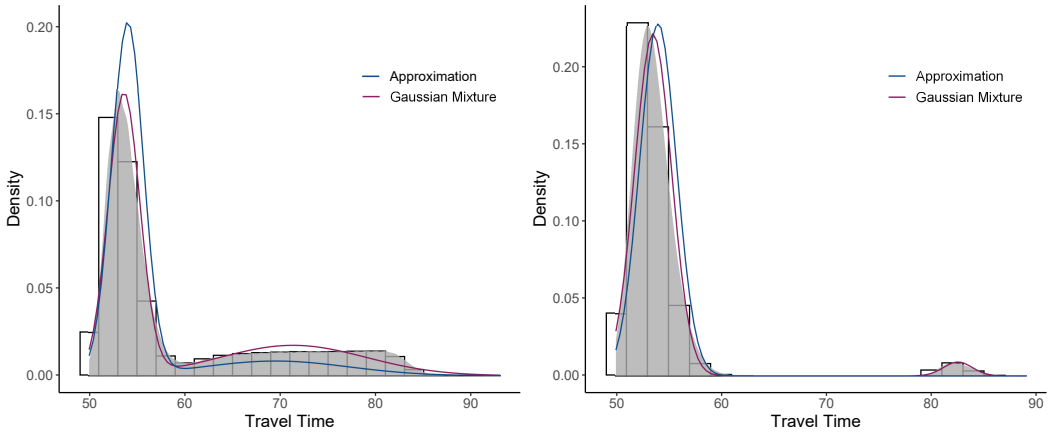
FIGURE 4.26: Travel time distribution: Case 60s-45s in low-density. Approximation represents the Gaussian mixture distribution with replacing μ_1, μ_2, λ_1 and λ_2 with the approximated values. Empirical distributions are smoothed with non-parametric Gaussian kernel.

Table 4.6: Gaussian mixture distribution parameter estimation in low-density: case 60s-45s

offset	parameter	component 1		component 2		component 3		K-S test Sig.
5	λ		0	0.558	[±0.002]	0.442	[±0.002]	0.318616
	μ			73.019	[±0.061]	95.050	[±0.008]	
	σ			12.280	[±0.042]	1.566	[±0.006]	
10	λ	0.124	[±0.001]	0.526	[±0.002]	0.350	[±0.002]	0.437613
	μ	52.833	[±0.015]	75.999	[±0.068]	97.939	[±0.013]	
	σ	1.347	[±0.012]	11.931	[±0.047]	1.978	[±0.011]	
15	λ	0.293	[±0.002]	0.520	[±0.002]	0.187	[±0.002]	0.602758
	μ	53.521	[±0.010]	77.061	[±0.068]	98.788	[±0.021]	
	σ	1.587	[±0.008]	11.772	[±0.053]	2.398	[±0.016]	
20	λ	0.415	[±0.002]	0.585	[±0.002]	0		0.245190
	μ	53.845	[±0.010]	80.841	[±0.064]			
	σ	1.733	[±0.007]	13.044	[±0.045]			
25	λ	0.496	[±0.002]	0.504	[±0.002]	0		0.495892
	μ	53.867	[±0.008]	81.038	[±0.055]			
	σ	1.771	[±0.007]	11.616	[±0.044]			
30	λ	0.573	[±0.002]	0.427	[±0.002]	0		0.651904
	μ	53.862	[±0.007]	82.825	[±0.055]			
	σ	1.830	[±0.006]	9.925	[±0.045]			
35	λ	0.633	[±0.001]	0.367	[±0.001]	0		0.304312
	μ	53.822	[±0.008]	85.110	[±0.050]			
	σ	1.867	[±0.005]	8.355	[±0.034]			
40	λ	0.689	[±0.002]	0.311	[±0.002]	0		0.166204
	μ	53.785	[±0.009]	87.261	[±0.040]			
	σ	1.881	[±0.005]	7.039	[±0.028]			
45	λ	0.745	[±0.001]	0.255	[±0.001]	0		0.100458
	μ	53.789	[±0.007]	89.336	[±0.037]			
	σ	1.904	[±0.005]	5.847	[±0.028]			
50	λ	0.801	[±0.001]	0.199	[±0.001]	0		0.218025
	μ	54.105	[±0.007]	91.495	[±0.038]			
	σ	2.114	[±0.005]	4.726	[±0.025]			

Table 4.6 summarises the fitted Gaussian mixture distribution parameters of the travel time for varying offset in Case 60s-45s. Due to the increased congestion in the link with the longer green time at upstream intersection, none of the vehicles join the green wave when the offset is 5s. With the offset increases, the proportion of vehicles joining the green wave increases. The proportion of vehicles waiting for the next green time at the downstream intersection decreases with the offset and disappear beyond the offset 20s.

Case 60s-60s



(a) offset=25s (b) offset=50s

FIGURE 4.27: Travel time distribution: Case 60s-60s in low-density. Approximation represents the Gaussian mixture distribution with replacing μ_1, μ_2, λ_1 and λ_2 with the approximated values. Empirical distributions are smoothed with non-parametric Gaussian kernel.

Figure 4.27 shows the travel time distribution of the link with equal but longer green times at both upstream and downstream intersection for offsets, 25s and 50s respectively. A longer green time at both upstream and downstream junction allows more vehicles to pass through the bulk link. Compared to the case 45s-45s, majority of the vehicles join the green wave. The second cluster of vehicles, the proportion of vehicles waiting for the green time at the downstream intersection get reduces.

Table 4.7 summarises the fitted Gaussian mixture distribution parameters of the travel time for varying offset in Case 60s-60s. Due to smaller α, the longer green time at the upstream intersection did not drive the bulk link into congestion regime. Furthermore, the longer green time at downstream intersection allows more vehicles to pass through. Therefore, unlike in the previous cases, none of the vehicle has to wait for a full cycle at the downstream intersection even with a smaller offset due to the longer green time. The

proportion of vehicles travelling with minor delays decreases with the offset increases and almost disappear at the offset 50s like Case 45s-45s.

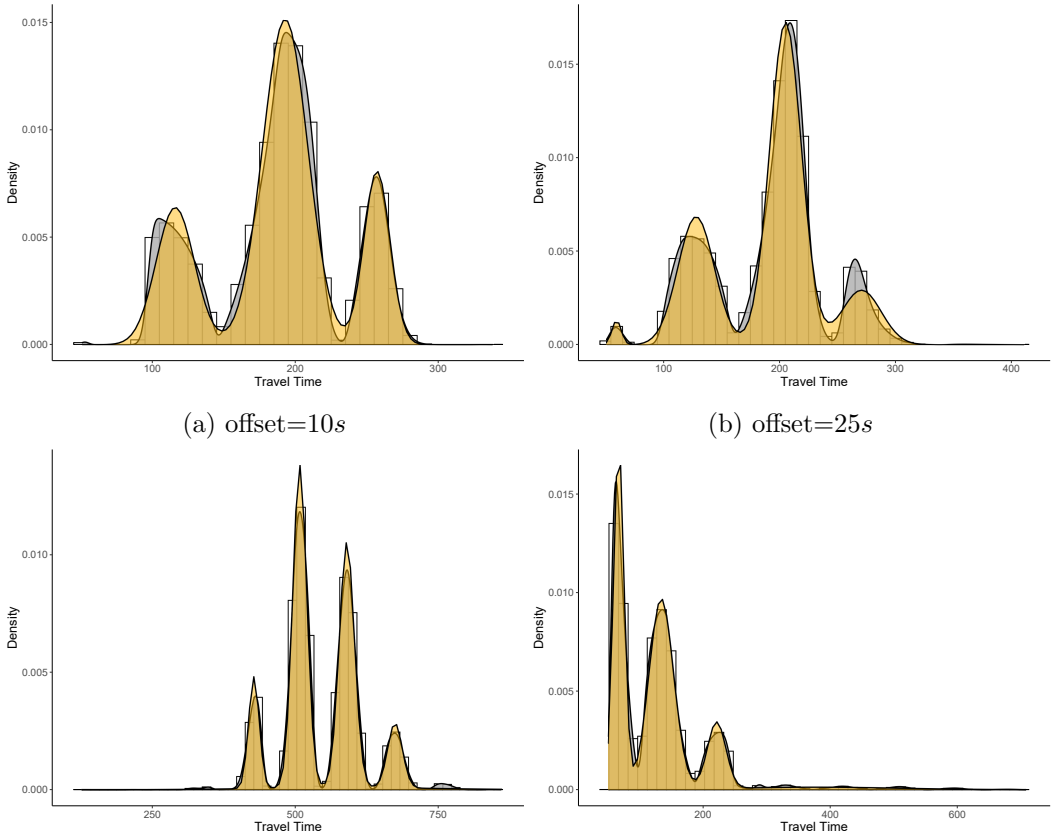
Table 4.7: Gaussian mixture distribution parameter estimation in low-density: Case 60s-60s

offset	parameter	component 1		component 2		component 3	K-S test Sig.
	λ	0.565	[±0.002]	0.435	[±0.002]	0	
5	μ	53.744	[±0.007]	71.820	[±0.058]		0.366059
	σ	1.667	[±0.007]	9.365	[±0.041]		
	λ	0.612	[±0.002]	0.388	[±0.002]	0	
10	μ	53.914	[±0.008]	70.433	[±0.063]		0.415263
	σ	1.755	[±0.006]	8.709	[±0.039]		
	λ	0.630	[±0.002]	0.370	[±0.002]	0	
15	μ	53.910	[±0.007]	69.857	[±0.061]		0.622529
	σ	1.770	[±0.006]	8.343	[±0.036]		
	λ	0.630	[±0.002]	0.370	[±0.002]	0	
20	μ	53.782	[±0.008]	69.979	[±0.052]		0.425102
	σ	1.721	[±0.006]	8.184	[±0.041]		
	λ	0.669	[±0.002]	0.331	[±0.002]	0	
25	μ	53.639	[±0.007]	71.338	[±0.047]		0.214528
	σ	1.660	[±0.005]	7.613	[±0.033]		
	λ	0.732	[±0.001]	0.268	[±0.001]	0	
30	μ	53.654	[±0.006]	73.780	[±0.041]		0.328552
	σ	1.784	[±0.005]	6.274	[±0.033]		
	λ	0.802	[±0.001]	0.198	[±0.001]	0	
35	μ	53.630	[±0.007]	76.305	[±0.031]		0.342055
	σ	1.757	[±0.004]	4.717	[±0.027]		
	λ	0.859	[±0.001]	0.141	[±0.001]	0	
40	μ	53.571	[±0.005]	78.351	[±0.029]		0.432673
	σ	1.712	[±0.005]	3.477	[±0.019]		
	λ	0.914	[±0.001]	0.086	[±0.001]	0	
45	μ	53.544	[±0.006]	80.366	[±0.026]		0.222793
	σ	1.728	[±0.005]	2.341	[±0.016]		
	λ	0.969	[±0.0005]	0.031	[±0.0005]	0	
50	μ	53.530	[±0.006]	82.519	[±0.023]		0.199886
	σ	1.749	[±0.004]	1.372	[±0.017]		

4.3.3 System with a higher inflow: $\alpha = 0.5$ and $\beta = 0.9$

We simulated the system described in Section 4.3.1.1, considering a high inflow probability ($\alpha = 0.5$ and $\beta = 0.9$) for each distinct green time combinations for different offsets (i.e., $d \in \{10, 20, 25, 30, 35, 40, 45, 50, 55\}$). Each case consists of ten different independent simulations. We record the travel time of each vehicle. Travel time distributions for all the four cases and different offsets are summarised in Appendix 6.

Case 45s-45s



(a) offset=10s (b) offset=25s

(c) offset=40s (d) offset=55s

FIGURE 4.28: Histogram of travel time in Case 45s-45s. Gaussian mixture distribution is shown in gold area. Non parametric Gaussian kernel density estimation is shown in grey area.

Figure 4.28 illustrates the estimated Gaussian mixture travel time distributions of the link for Case 45s-45s for selected offsets. Table 4.8 summarises the fitted parameters. Due to the large inflow probability, and both the upstream and downstream intersections having smaller green time, nearly all of the

vehicles delayed by at least one red light except for the optimal offset ($d = 55s$). When the offset is 55s, nearly 36.6% of the vehicles joined the green wave while there were a few vehicles in the green wave with other offset settings. Moreover, the off set of 55s has majority of the travel times are less than to 200s while in other offsets having peak travel time around 200s. We can consider the 55s is the optimal offset in this signal setting. With the offset 40s, nearly all the vehicles have a travel time roughly greater than to 350s. The bad offset settings make a great difference in the travel time as illustrated in Figures 4.28c and 4.28d. Furthermore, the travel time distributions can be well approximated by a Gaussian mixture distribution. The goodness-of-fit of the estimated parameters are also high as summarised in Table 4.8.

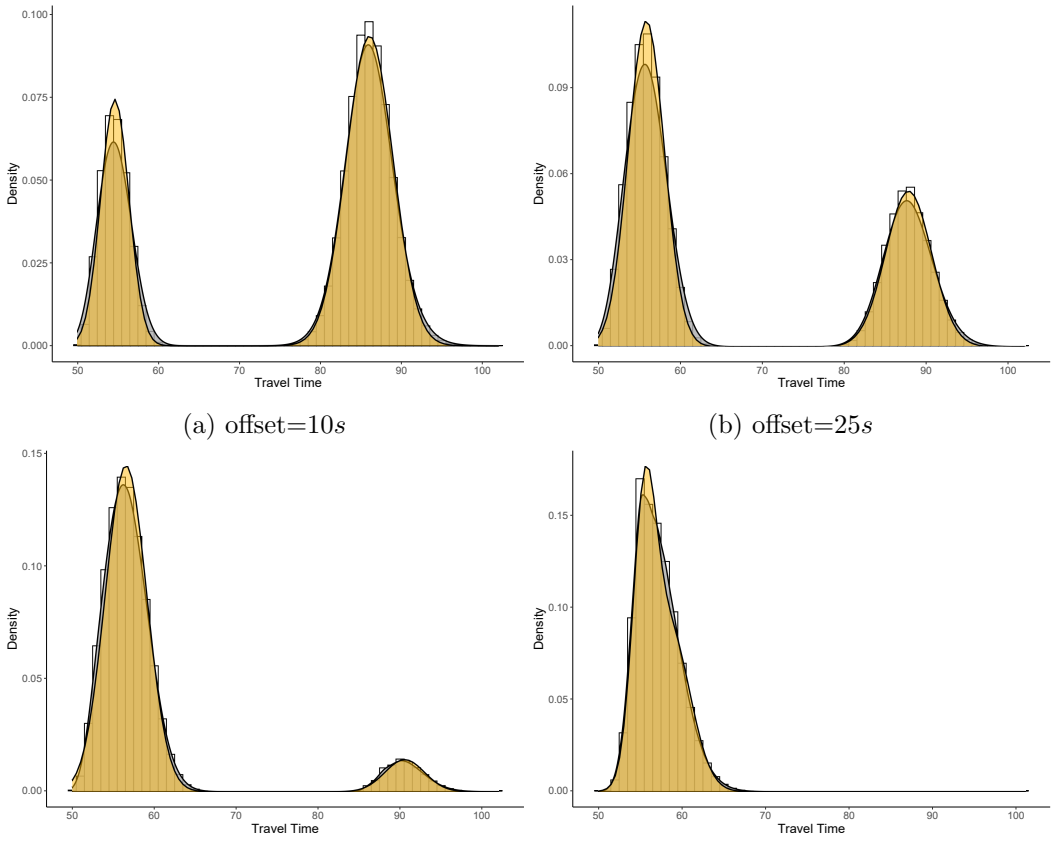
Table 4.8: Gaussian mixture distribution parameter estimation in high-density: Case 45s-45s

GMM-Component		Offset			
		10s	25s	40s	55s
Component 1	λ	0.2138[±0.00097]	0.0131[±0.000260]	0.1099[±0.00074]	0.3660[±0.00094]
	μ	116.06[±0.07401]	59.899[±0.048684]	427.65[±0.08602]	67.786[±0.03462]
	σ	13.364[±0.05286]	3.9991[±0.045533]	9.2357[±0.05843]	8.3835[±0.03089]
Component 2	λ	0.6018[±0.00138]	0.2686[±0.009374]	0.4216[±0.00162]	0.4701[±0.00109]
	μ	192.83[±0.03952]	128.27[±0.061989]	507.65[±0.03679]	134.55[±0.06873]
	σ	15.862[±0.02481]	15.646[±0.061204]	12.254[±0.03681]	19.411[±0.06385]
Component 3	λ	0.1843[±0.00089]	0.5936[±0.001153]	0.0879[±0.0006]	0.1223[±0.00045]
	μ	256.91[±0.04921]	205.21[±0.046909]	673.94[±0.11910]	221.29[±0.09853]
	σ	9.1119[±0.02099]	13.737[±0.034872]	12.750[±0.11852]	14.445[±0.04664]
Component 4	λ		0.1246[±0.000874]	0.3454[±0.00154]	0.0415[±0.00040]
	μ		270.46[±0.143832]	589.89[±0.04951]	377.86[±1.75560]
	σ		17.007[±0.062563]	13.204[±0.05491]	129.52[±1.12323]
Component 5	λ			0.0349[±0.00112]	
	μ			589.91[±2.30478]	
	σ			48.60[±2.19968]	
K-S test Sig.		0.79315	0.74486	0.78217	0.61536

Case 45s-60s

Figure 4.29 illustrates the travel time distribution of the link when the green time at the upstream intersection is shorter than the downstream intersection (i.e. Case 45s-60s). Even though the vehicle inflow probability is quite high, the shorter green time at the upstream intersection allows fewer vehicles into the bulk link. Moreover, having a longer green time at the downstream intersection allows more vehicles to pass through the intersection. Therefore, the bulk link is still in low-density regime. The travel time distribution of the bulk link is similar to the travel time distribution of the system when it is in the low-density with the same signal setting. See Figure 4.24. However, the proportion of vehicle join the green wave is relatively smaller when the inflow probability is high. The fitted Gaussian mixture distribution parameters for the selected offsets are summarised in Table 4.9. It is evident that the travel time distribution of the bulk link can be well approximated by a Gaussian

mixture in Case 45s-60s as well.



(a) offset=10s (b) offset=25s

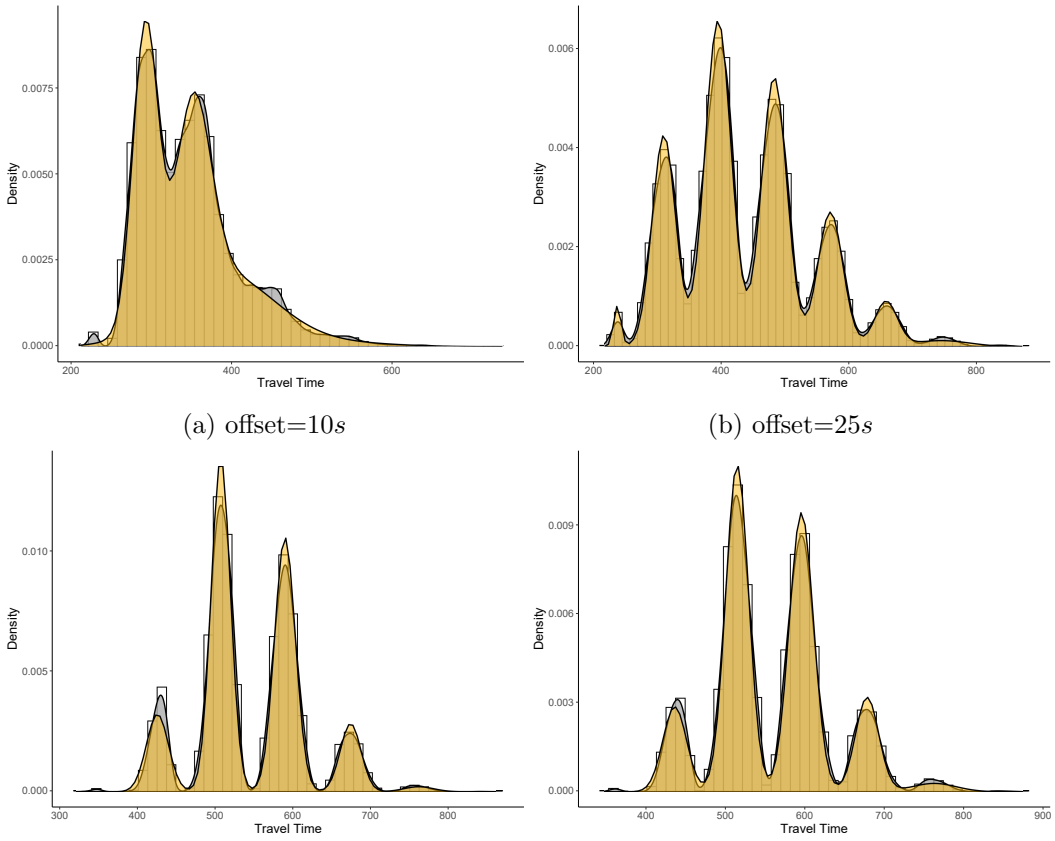
(c) offset=40s (d) offset=55s

FIGURE 4.29: Histogram of travel time in Case 45s-60s. Gaussian mixture distribution is shown in gold area. Non-parametric Gaussian kernel density estimation is shown in grey area.

Table 4.9: Gaussian mixture distribution parameter estimation in high-density: Case 45s-60s

GMM-Component		Offset			
		10s	25s	40s	55s
Component 1	λ	0.324[±0.00113]	0.622[±0.00109]	0.914[±0.00042]	0.445[±0.00789]
	μ	54.66[±0.00595]	55.87[±0.00473]	56.57[±0.00474]	55.48[±0.01056]
	σ	1.738[±0.00272]	2.191[±0.00363]	2.526[±0.00363]	1.364[±0.01062]
Component 2	λ	0.676[±0.00113]	0.377[±0.00109]	0.085[±0.00042]	0.555[±0.00789]
	μ	86.07[±0.01042]	87.78[±0.00679]	90.56[±0.01884]	58.44[±0.04014]
	σ	2.885[±0.00436]	2.790[±0.00471]	2.411[±0.01099]	2.503[±0.01271]
K-S test Sig.		0.77927	0.5031	0.57356	0.57126

Case 60s-45s



(a) offset=10s (b) offset=25s

(c) offset=40s (d) offset=55s

FIGURE 4.30: Histogram of travel time in Case 60s-45s. Gaussian mixture distribution is shown in gold area. Non-parametric Gaussian kernel density estimation is shown in grey area.

Figure 4.30 illustrates the travel time distribution of the link when the green time at the upstream intersection is longer than the downstream intersection (i.e. Case 60s-45s). Given that the vehicle inflow probability is relatively high, the upstream link is highly congested. Having a longer green time at the upstream intersection allows more vehicles into the bulk link, and causes more congestion. Nearly, all the vehicles have to wait for at least two cycles to pass through the downstream intersection. When the offset is 10s, majority of the vehicle pass the downstream intersection in less than 400s. However, for all the other offset settings it is relatively worse. The signal setting of the Case 60s-45s creates a complete gridlock[5] resulting longer travel time in the bulk link. It is quite evident that the distance between different peaks in the travel time distribution is closer to the cycle length. However, this phenomena is not

[5]Form of traffic congestion where continuous queues of vehicles block the entire link.

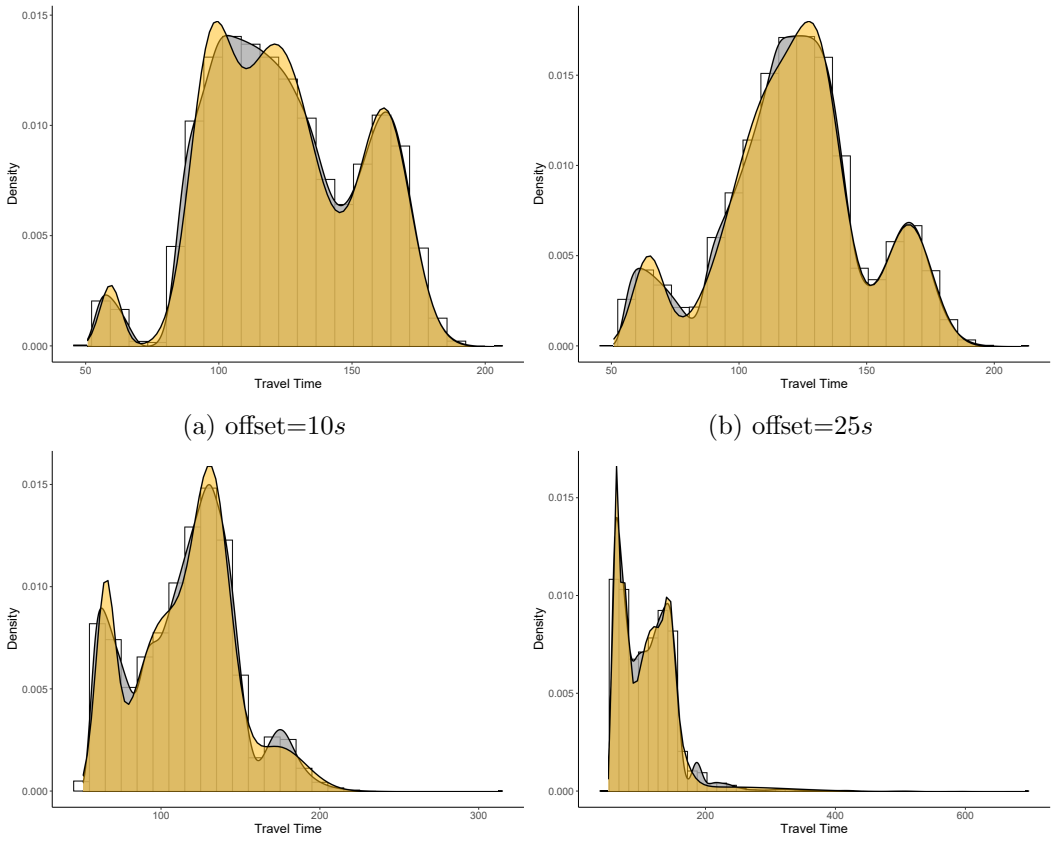
evident in the travel time distribution when the offset is closer to optimal (i.e. 10 s in Case 60s-45s), all these different peaks are trying to form one big peak mushing together.

Table 4.10: Gaussian mixture distribution parameter estimation in high-density: Case 60s-45s

GMM-Component		Offset			
		10s	25s	40s	55s
Component 1	λ	0.3437[±0.00212]	0.0121[±0.00015]	0.1144[±0.00089]	0.1069[±0.00085]
	μ	292.85[±0.10956]	238.52[±0.11617]	426.37[±0.07828]	436.89[±0.11831]
	σ	15.762[±0.09059]	5.9033[±0.07067]	14.186[±0.07961]	14.915[±0.07092]
Component 2	λ	0.3044[±0.00449]	0.1956[±0.00096]	0.4281[±0.00082]	0.3972[±0.00185]
	μ	352.56[±0.19546]	311.13[±0.09185]	507.64[±0.06001]	515.28[±0.04488]
	σ	31.668[±0.24811]	18.345[±0.08742]	12.356[±0.03233]	14.386[±0.05228]
Component 3	λ	0.3369[±0.00257]	0.3362[±0.00142]	0.3536[±0.00116]	0.3599[±0.00178]
	μ	390.09[±1.42728]	395.84[±0.10910]	589.91[±0.05700]	595.92[±0.06112]
	σ	64.469[±0.78102]	20.446[±0.10129]	13.367[±0.04947]	15.268[±0.05614]
Component 4	λ	0.0151[±0.00275]	0.2694[±0.00092]	0.0940[±0.00061]	0.1167[±0.00139]
	μ	544.61[±11.0953]	483.36[±0.11936]	674.11[±0.13924]	678.34[±0.14994]
	σ	64.335[±4.18825]	19.821[±0.06555]	13.266[±0.07297]	14.641[±0.13836]
Component 5	λ		0.1329[±0.00077]	0.0099[±0.00017]	0.0193[±0.00033]
	μ		571.11[±0.12348]	761.53[±0.42382]	762.59[±0.78090]
	σ		19.603[±0.09621]	19.074[±0.38398]	27.509[±0.46399]
Component 6	λ		0.0401[±0.00066]		
	μ		657.95[±0.42626]		
	σ		17.784[±0.19503]		
Component 7	λ		0.0137[±0.00022]		
	μ		744.08[±2.26895]		
	σ		44.588[±1.3278]		
K-S test Sig.		0.9999	0.98176	0.89573	0.85448

Case 60s-60s

Figure 4.31 illustrates the travel time distribution of the link when the green time at both the upstream intersection and the downstream intersection are 60s. The behaviour of the travel time distribution relatively closer to Case 45s-45s, since in both cases equal green time is implemented at the upstream and downstream intersection. Because of the longer green time compared to Case 45s-45s, the traffic flow through the bulk link increases. The means of the different components are also quite close. Hypothetically, If the red time reduces to zero, the travel time distribution will have only one peak as the non-signalised case. As we increase the green time at both intersections, all the peaks are mushing together form one peak. It is evident that the proportion of vehicles join the green wave is increases with the offset increases. We can consider the offset 55s is the optimum among the offsets we consider, even though there were a few vehicles having longer travel time.



(a) offset=10s (b) offset=25s

(c) offset=40s (d) offset=55s

FIGURE 4.31: Histogram of travel time in Case 60s-60s. Gaussian mixture distribution is shown in gold area. Non-parametric Gaussian kernel density estimation is shown in grey area.

Table 4.11: Gaussian mixture distribution parameter estimation in high-density: Case 60s-60s

GMM-Component		Offset			
		10s	25s	40s	55s
Component 1	λ	0.0279[±0.00041]	0.0748[±0.00061]	0.1446[±0.00162]	0.1238[±0.00261]
	μ	59.568[±0.02590]	64.547[±0.04417]	65.076[±0.04217]	61.874[±0.03333]
	σ	4.0001[±0.03131]	6.1146[±0.05032]	6.0760[±0.05088]	3.7291[±0.03304]
Component 2	λ	0.2089[±0.00479]	0.5753[±0.01781]	0.3542[±0.00715]	0.1758[±0.00333]
	μ	96.635[±0.00479]	113.89[±0.56916]	101.39[±0.29971]	73.749[±0.16045]
	σ	7.7466[±0.09608]	16.195[±0.23656]	17.378[±0.34337]	7.8479[±0.13489]
Component 3	λ	0.5127[±0.00642]	0.1979[±0.01711]	0.4044[±0.00591]	0.5603[±0.00129]
	μ	121.55[±0.16912]	132.06[±0.16114]	132.28[±0.10553]	120.81[±0.12282]
	σ	15.051[±0.19231]	8.5869[±0.25167]	11.547[±0.08543]	27.074[±0.06977]
Component 4	λ	0.2503[±0.00179]	0.1519[±0.00072]	0.0968[±0.00089]	0.0812[±0.00119]
	μ	162.37[±0.06806]	166.49[±0.06083]	173.56[±0.21203]	145.70[±0.07237]
	σ	9.5911[±0.04148]	9.0972[±0.03587]	17.737[±0.12359]	7.4196[±0.08993]
Component 5	λ				0.0589[±0.00068]
	μ				234.18[±1.18359]
	σ				99.342[±0.51207]
K-S test Sig.		0.99986	0.99999	0.999788	0.88772

When the system is highly congested, the behaviour of the travel time distribution is complex. Approximation of the Gaussian mixture model parameters analytically is relatively difficult. However, it can be concluded that a generally the travel time of a congested link can be well approximated by a Gaussian mixture as well. Among all cases, equal split in the high-density is the most complex one as it could lead to low-density, maximum-current and high-density on the bulk link.

4.4 Discussion

Travel time distributions have been widely used to characterise traffic flow dynamics. Many researchers argued that the travel time distributions are asymmetric and considerably positively skewed [19–22, 97] unlike normal distribution. Contrarily, some argued that the uni-modal distributions might not be sufficient to represent link travel time distributions [22, 98–105]. However, most of these studies are relied on empirical travel time data with limited sample sizes and do not present consistent results. Accurately modelling travel time distributions is advantageous in simulation-based studies greatly.

Moreover, the travel time on a link is highly sensitive to the signal settings at intersections like phase splits. The linking pattern of the traffic signals at consecutive intersections also has a significant impact on the travel time. Therefore, in this chapter, we studied the link travel time distributions and the impact of signalised intersection explicitly.

We proposed a model for the system consisting of a single-lane link without signals based on mean-field approximations to approximate link travel time distributions. We evaluated our models against the ASEP with both deterministic and stochastic updates. When the link is in the low-density regime, with deterministic updates, the system has two density regions and the travel time distribution is highly skewed. When the system is in the high-density regime with deterministic updates, the system has a single density regime with an effective hopping probability and observed that the approximated travel time distribution converges to a normal distribution.

For the ASEP with stochastic updates, the travel time distribution could be modelled by considering a single density region with an effective hopping probability. The distribution of travel time converges to a normal distribution irrespective of the link congestion. Our model approximates the centre of the normal distribution precisely. The travel time distribution on the NaSch model also suggests that a normal distribution can reasonably approximate

the link travel time. These approximations suggest that the link travel time distribution can be well approximated by a normal distribution when the effect of signalised intersection is absent.

When the signalised intersections are in effect, the general link travel time distribution is generally multi-modal. Only for a special case, when the link is not congested, and the linking pattern of the consecutive signalised intersections are optimal, the travel time distribution is uni-modal. We studied the travel time distribution for different signal settings and offsets for small and large inflow probabilities. We use deterministic domain wall to approximate μ_1, μ_2, λ_1 and λ_2. But the approximation is only for smaller inflow probabilities. Generally, the travel time distribution between signalised intersections can be well approximated by a Gaussian mixture. Furthermore, the skewed distributions we observed in practice, were also reproducible in the simulation model. The skewness can also be well approximated by Gaussian mixtures.

Our study on link travel time distributions adds debate on uni-modal and multi-modal travel time distributions. The results of the impact of the different offsets provide useful insight to the modelling travel time. We could incorporate the stochastic domain wall model to understand the proportion of vehicle joining the green wave more precisely which is a potential direction for future research.

Simulating stopping behaviour of on-road transit vehicles 5

5.1 Overview

To accurately simulate movements of on-road public transport vehicles, such as buses and trams, it is crucial to accurately model their stopping behaviour at stops. The stopping behaviour of mass transit modes is significantly affected by the *dwell time*: the time spent at each stop for the transfer of passengers, including the time necessary to open and close doors [37]. Empirical research [24, 25, 35, 37, 40] that has explored dwell time or passenger flow time [1] has found that the volume of passengers boarding and alighting has a major impact on the dwell time of a transit vehicle. However, there are many other factors, such as the location and the type of stop, the type of vehicle and on-board crowding, which significantly affect the dwell time. The service time per passenger on average extends between 1.5 to 6.0 seconds per alighting passenger and 1.5 to 8.0 seconds per boarding passenger [25].

Estimating the dwell time of transit vehicles has been extensively studied [24–37, 40, 41, 120]. Most studies related to dwell time have been limited to relating the mean dwell time to influential factors, by deploying multivariate regression models [24–37]. The results can be utilised to predict the time taken

[1] Time spent at each stop for the transfer of passengers

for passengers boarding and alighting when passenger demand data is available. However, these traditional data-driven regression approaches are unable to capture the interaction between transit vehicles and passengers precisely [121]. Moreover, regression approaches only capture average behaviour, and cannot predict anything about fluctuations. In order to understand fluctuations, which can have significant impacts in practice, one needs an understanding of the dwell time distribution.

Khoo [26] has estimated the empirical dwell time distributions for buses at different times of the day (peak hour, off peak hour), different platform crowding levels, and for different fare collection methods (cash/card system, conductor system). He argued that the dwell time of buses can be well approximated by Pearson type VI distribution, except at less crowded bus stops. Li et al. [40] have studied the dwell time distributions for front door alighting/boarding, rear door alighting/boarding and combined front and rear door alighting/boarding distributions separately. They argued that the empirical distributions in all cases can be well approximated by log-normal distributions. In their study on modelling dwell time for bus rapid transit stations, Li et al. [31] again argued that the dwell time at a stop follows log-normal distribution. Furthermore, they observed that the train dwell time distribution at short stop stations can also be approximated by a log-normal distribution [120]. Rashidi and Ranjitkar [41] have assessed three distribution functions for approximate bus dwell time: normal, log-normal and Wakeby distributions. They claimed that the Wakeby distribution approximates the dwell time better than the log-normal distribution, while the normal distribution is not a suitable approximation for dwell time.

The simulation package CORSIM models bus dwell times by relying on either average dwell times specified by users or embedded statistical distributions [122]. The VISSIM simulator has two methods for modelling dwell time at stops: dwell time distributions and advanced passenger models. Dwell time distributions are defined according to the normal distribution and empirical distributions [123]. In order to predict dwell time based on advanced passenger models, it is required to have access to passenger demand data. Therefore, it is important in practice to understand the dwell time distributions at each stop, to accurately simulate the stopping behaviour of on-road transit vehicles.

Another important aspect that we need to consider when modelling the stopping behaviour of a mass transit vehicle is the time-headway, which is the time interval between successive transit vehicles. Transport planners have devoted considerable effort to establish descriptive mathematical models of time-headway for different types of road networks such as suburban arterials [124], urban arterials [125] and highways [125–127]. Many theoretical distributions such as gamma, negative exponential, shifted exponential, Erlang, Johnson SB,

Johnson SU, Log-normal and Log-logistic have been fitted to time-headway data [124, 125]. However, to date most of the researchers have only looked at average behaviour of mass-transit time-headway [128–132] rather than studying the distributions. Lin and Ruan [133] proposed a time-headway regularity measure using the distribution of bus dispatching headway. They analysed AVL data from a selected route in Chicago and suggest that the time-headway can be well approximated by a Gamma distribution. Bellei and Gkoumas studied the time-headway distributions and service irregularity utilising a stochastic simulation model which represents a one-way transit line [129]. They argued that the time-headway distribution is likely to be uni-modal at the first stops in the route and bi-modal depicting the concentration of short and long alternated time-headway at the latter stops. Moreover, they studied the relationship between consecutive time-headway and argued that the underlying joint distribution is more likely a bi-variate uni-modal distribution at the first stops and a bi-variate multi-modal distribution at the latter stops. However, they did not focus on fitting theoretical distributions to their data.

A group of two or more closely spaced transit vehicles running along the same route (a phenomenon often called *bunching*) is highly likely in congested urban transit networks. If a transit vehicle is running late, its time-headway increases, causing more passengers to wait at downstream stops. Therefore, the vehicle will be further delayed due to the longer dwell times. Correspondingly, one expects, *a priori*, that the dwell time between consecutive vehicles are likely to be correlated. Moreover, the dwell time and time-headway of a vehicles are also likely to be correlated, as are the dwell times between consecutive stops. If the two stops are close, longer dwell time at the upstream stop might suggest a shorter dwell time at the next stop. Moreover, if there is similar passenger demand at consecutive stops, dwell time between stops tend to be positively correlated. Hence, we need to understand the nature of this correlation structure explicitly. To our knowledge, no previous studies of dwell time correlations have been undertaken.

Given the disagreement between previous studies of dwell time distributions [26, 31, 40, 41, 120], in this chapter we analysed the dwell time (or more precisely, the passenger flow time) distributions at selected tram stops in the inner network of Melbourne's CBD, using a data set gathered in [35]. Unfortunately, this data set did not allow us to study time-headway or correlations in the dwell time data. Therefore, we conducted a statistical survey at two distinct pairs of consecutive tram stops in Melbourne's CBD, located in two different tram routes, one of high frequency, the other of low frequency. Moreover, to gain intuition into the mechanisms underlying dwell time and time-headway distributions, we simulated a high frequency tram route utilising our empirical results, to understand the stopping behaviour of a tram more explicitly. Furthermore, we studied the correlation structure between trams

and between stops.

5.2 Survey Data - Impact of Crowding on Streetcar Dwell Time [35]

In 2013, Currie et al. conducted a survey to measure the passenger flow time of trams operated in Melbourne, in order to study the impact of crowding on tram dwell time [35]. The data set consist of 1,147 dwell time events from ten tram stops in the inner network (black dots in the Figure 5.1).

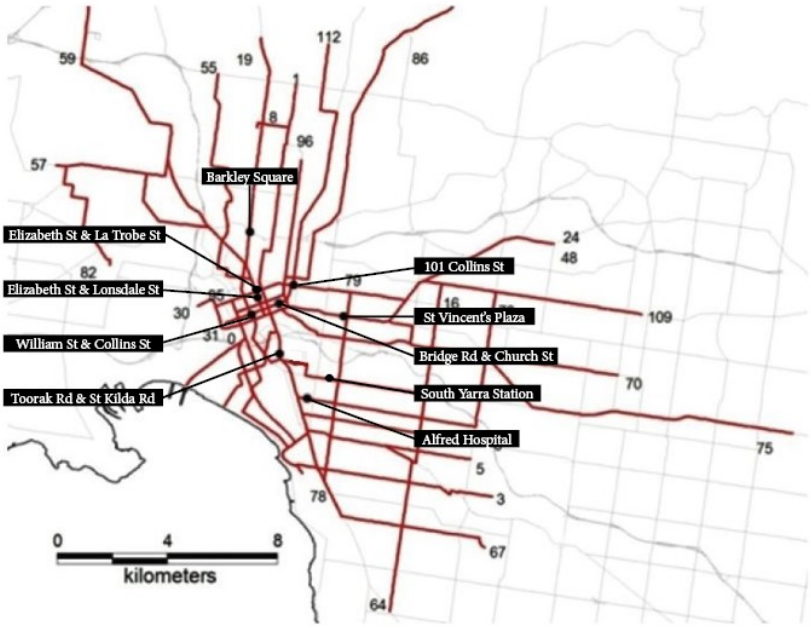


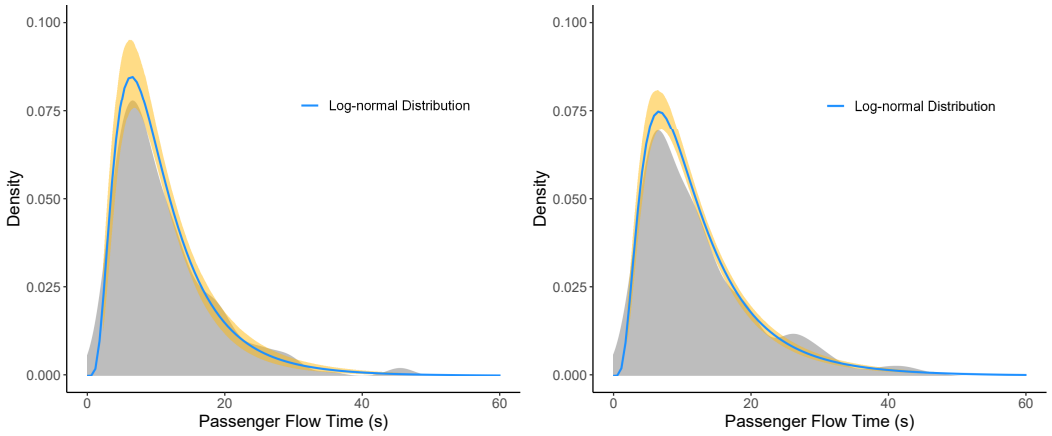
FIGURE 5.1: Melbourne tram network: routes and survey locations. Adapted from "Impact of Crowding on Streetcar Dwell Time" by Currie et al., 2013 [35]

Each stop was surveyed in the peak direction of travel for two hours in the morning (07:00-09:00) and three hours in the evening (15:00-18:00). For each passenger flow time measurement, corresponding boarding and alighting passenger demands were recorded. The authors explored the impact of both crowding and tram stop design on mean tram dwell time.

The data set is informative as it contains dwell time events from different tram stops, and it is capable of studying marginal dwell-time distributions.

The data set already exist, and we analysed it to understand the marginal passenger flow time distributions at selected tram stops in the inner network of Melbourne CBD. However, the data set is limited as it does not contain the events where the passenger demand is zero.

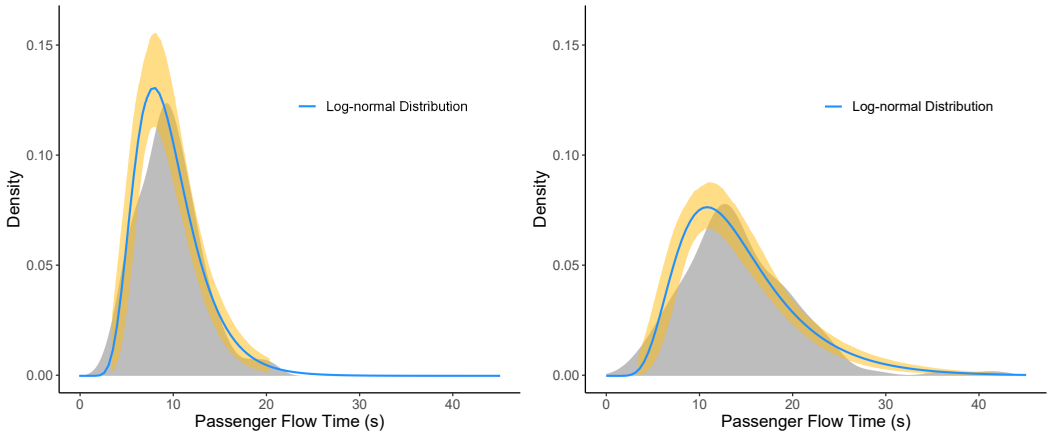
These tram stops include all three types of tram stop designs: curb-side, platform and safety-zone. Most of the tram stops in Melbourne's network are curb-side. Passengers have to wait on the curb and make their way through unguarded traffic to board trams which stop in the middle of the road. At safety zone stops, passengers are provided with a safety barrier to separate them from the traffic. Platform stops provide a waiting bay in the middle of the road, separating the passengers from general traffic.



(a) AM Peak (b) PM Peak

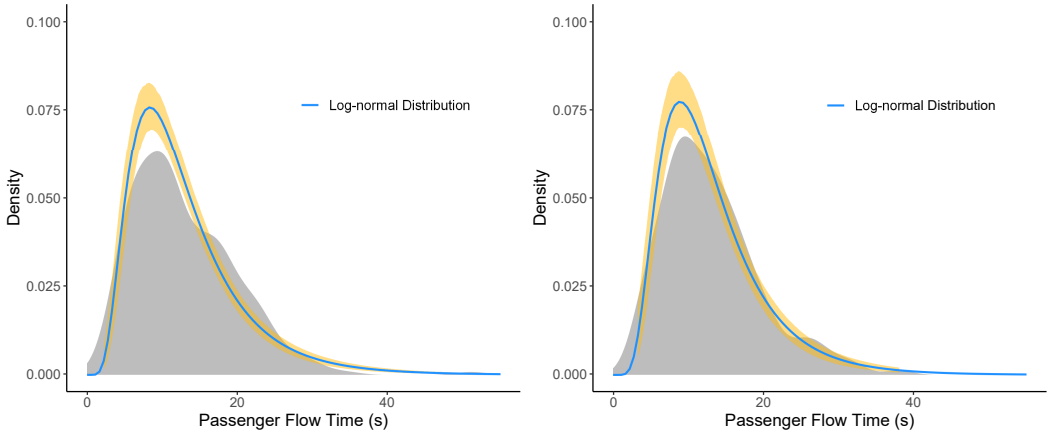
FIGURE 5.2: Survey: Distribution of Passenger Flow Time at tram stops with a safety-zone. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel.

Figures 5.2, 5.3 and 5.4 illustrate the passenger flow time distributions of each type of tram stop for both morning and afternoon peak hours along with the fitted log-normal distribution function and the 95% confidence band of the passenger flow time data. Table 5.1 summarises the estimated parameters and the corresponding goodness of fit test results based on the Kolmogorov–Smirnov test. All the surveyed stops are located in inner tram network. The passenger demand at these stops are fairly similar. The results show that the passenger flow time at each type of tram stop can be well approximated by a log-normal distribution with fairly universal parameters. Moreover, the universality of fitted parameters is quite remarkable: both μ and σ appear to depend only very weakly on the details of the stops.



(a) AM Peak (b) PM Peak

FIGURE 5.3: Survey: Distribution of Passenger Flow Time at tram stops with a curb-side. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel.



(a) AM Peak (b) PM Peak

FIGURE 5.4: Survey: Distribution of Passenger Flow Time at tram stops with a Platform. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel.

Table 5.1: Survey: Log-normal distribution parameter estimation of passenger flow time at each stop type.

		Log-normal parameter estimate		K-S test Sig.	Chi-square test Sig.
		$\log(\mu)$	$\log(\sigma)$		
Safety Zone	AM Peak	2.23[±0.0453]	0.611[±0.032]	0.8497	0.8448
	PM Peak	2.32[±0.034]	0.645[±0.024]	0.562	0.1632
Curb Side	AM Peak	2.2[±0.045]	0.36[±0.032]	0.6111	0.5006
	PM Peak	2.57[±0.044]	0.439[±0.031]	0.3467	0.1551
Platform	AM Peak	2.42[±0.036]	0.546[±0.026]	0.5014	0.0714
	PM Peak	2.45[±0.0369]	0.507[±0.027]	0.8558	0.2029

5.3 Survey: The correlation structure of stopping behaviour of a tram

The data set we analysed in Section 5.2 was not detailed enough to understand the correlation of passenger flow time between consecutive trams and consecutive stops. Moreover, the data set does not contain information about the time-headway of surveyed trams. Therefore, we conducted a survey on two consecutive tram stops, in two different tram routes in Melbourne CBD. We considered the Blyth Street stop and the Stewart Street stop on Sydney Road representing a low frequency tram route, and the Leopold Street stop and the Arthur Street stop on St Kilda Road representing a high frequency tram route as illustrated in Figure 5.5. Both the stops on St Kilda Road contain a safety-zone. At Sydney Road, both the surveyed stops are curb-side.

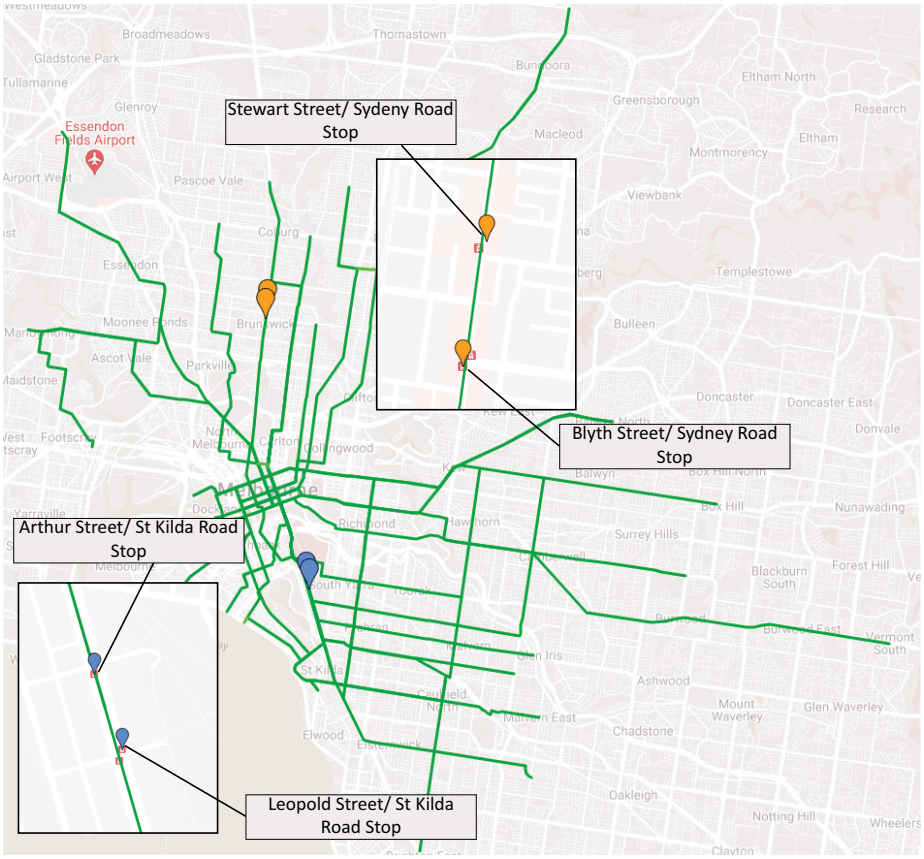


FIGURE 5.5: Melbourne tram network: routes and survey locations

5.3.1 Survey method

Both the routes were surveyed in the peak direction of travel, across five weekdays for 1.5 hours in the morning (08:00-09:30) at Sydney Road and in the evening (16:00-18:00) at St Kilda Road. Scheduled time-headway between trams was six minutes on Sydney Road. However, on St Kilda Road scheduled time-headway between trams was varied between one to five minutes as the surveyed stops captured multiple tram routes. For each passenger flow time observation, the following data was recorded:

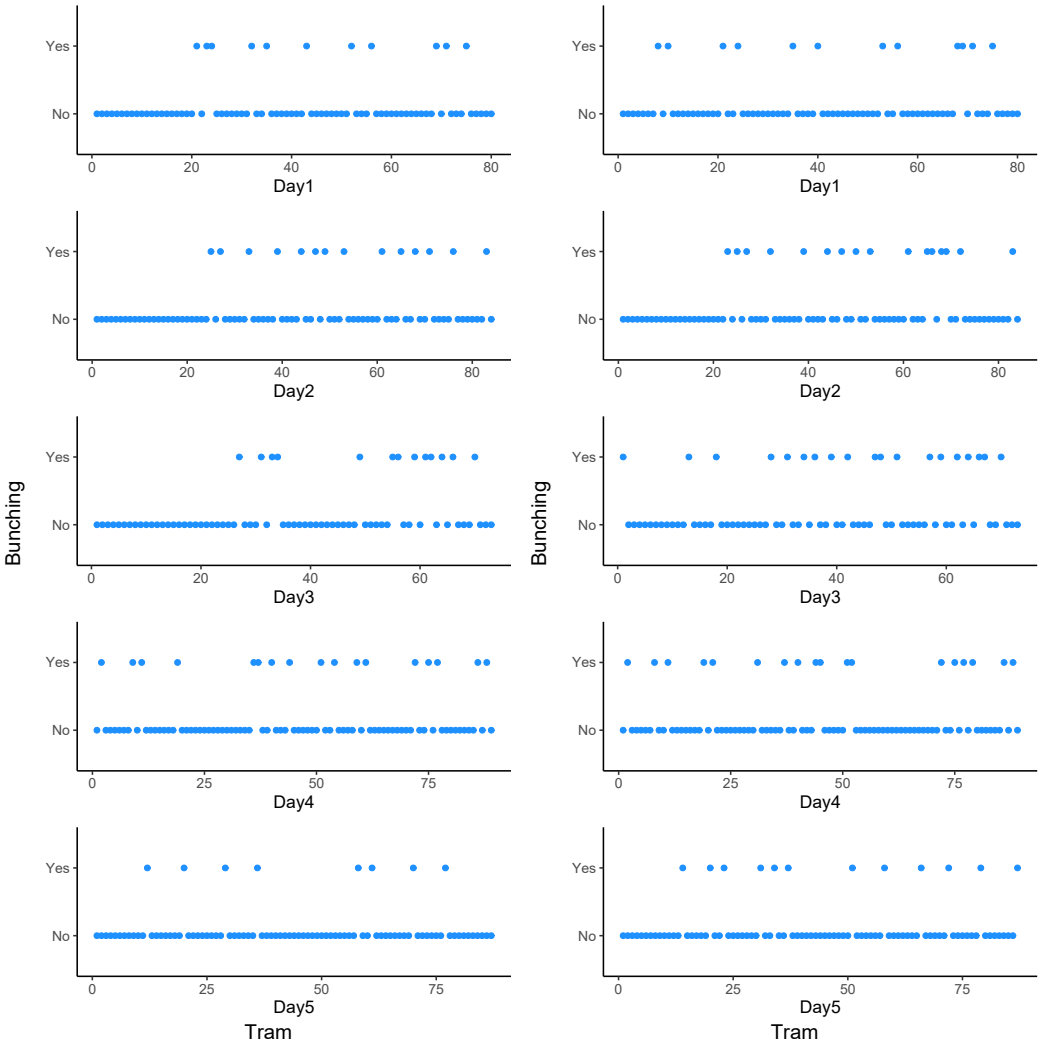
- Date
- Route No
- Tram No
- Arrival Time
- Passenger flow start time
- Passenger flow end time
- Departure time

5.3.2 Tram bunching

In St Kilda Road we experienced the tram bunching phenomena, where more than one tram arrived at the same time to serve passengers. We consider tram with a time-headway less than five seconds as a bunched tram. Figure 5.6 summarises how frequently tram bunching occurred during the St Kilda Road survey. However, we did not experience the bunching phenomena during the Sydney Road survey.

If a tram has to slow down for some reason, the time-headway between it and the preceding tram will be larger than the regular time-headway. More passengers will be waiting at the next stop, and tram will have a longer dwell time than usual. The time-headway between the preceding tram will be further increased. Comparably, the time-headway between the succeeding tram will be smaller than the expected time-headway. There will be fewer passengers at the stop and the succeeding tram will be faster than usual. Eventually, the two trams will meet along the route at some point. We noticed that most of the time when the tram bunching occurs, the succeeding tram serves for fewer passengers, and the tram is nearly empty.

When there are multiple routes operating on the same road segment as is in St Kilda Road, we could expect an increased number of trams barely following the scheduled time-headway. It is likely to occur bunching phenomena more frequently in St Kilda Road as illustrated in Figure 5.6. Therefore, we can expect a significant correlation between consecutive trams and consecutive stops in St kilda Road.



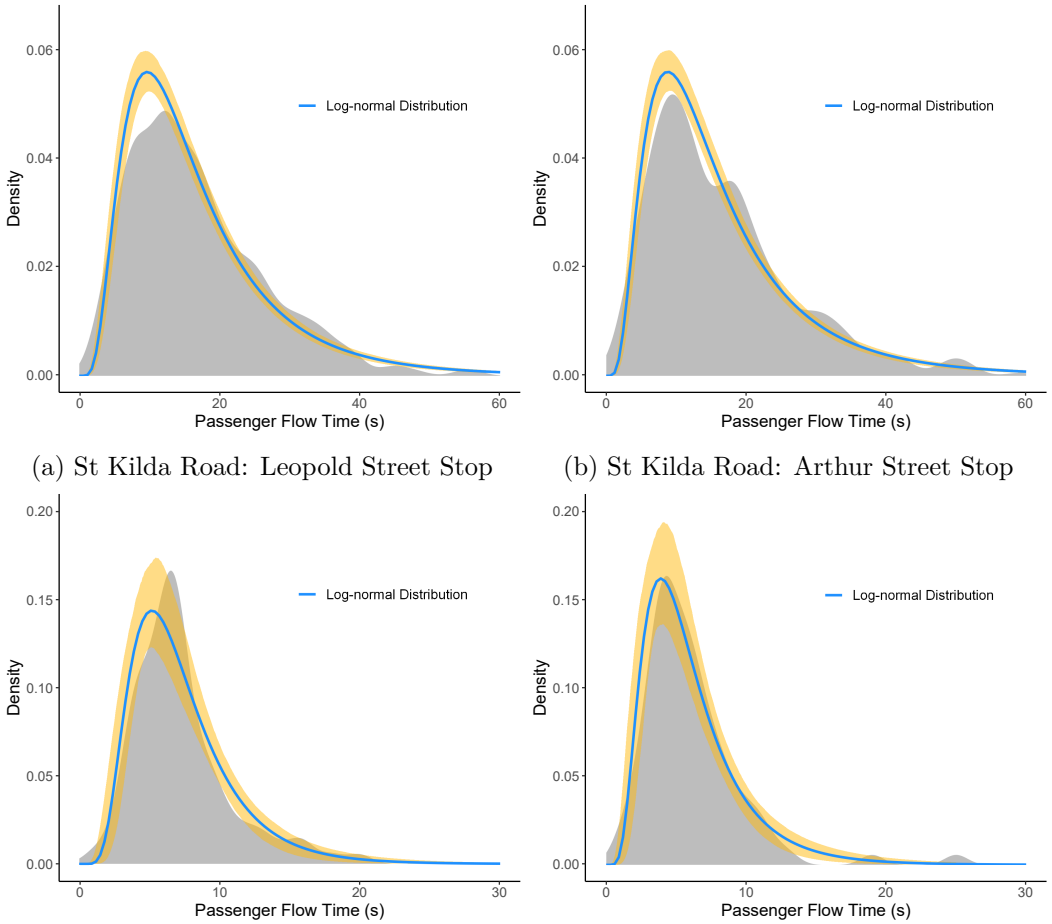
(a) Leopold street Stop

(b) Arthur street Stop

FIGURE 5.6: Survey: Occurrences of bunching at St Kilda Road. We consider any time-headway less than five seconds as bunched.

5.3.3 Marginal distributions of observables

5.3.3.1 Marginal distribution of passenger flow time



(a) St Kilda Road: Leopold Street Stop

(b) St Kilda Road: Arthur Street Stop

(c) Sydney Road: Stewart Street Stop

(d) Sydney Road: Blyth Street Stop

FIGURE 5.7: Survey: Distribution of Passenger Flow Time. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel shown in grey area.

We analyse the passenger flow time data we observed in order to compare approximated marginal distributions in Section 5.2 as a consistency test. Figure 5.7 illustrates the passenger flow time distributions at each surveyed stop. As observed in Section 5.2, passenger flow times at all four stops we surveyed are again well described by a log-normal distribution. Estimated σ's for passenger flow times at St Kilda Road stops are quite similar to the estimated σ's for passenger flow times at trams stops with a safety zone (see

Table 5.1). However, μ's are slightly higher as the St Kilda Road stops are extremely busy. Passenger flow times at Sydney Road stops show slightly lower μ's as the route is not as busy as the St Kilda Road route. Estimated σ's for passenger flow times at Sydney Road stops are again similar to the estimated σ's for passenger flow times at curb-side trams stops (see Table 5.1).

Table 5.2: Survey: Log-normal distribution parameter estimation of passenger flow time at each surveyed stop.

		Log-normal parameter estimate		K-S test Sig.	Chi-square test Sig.
		$\log(\mu)$	$\log(\sigma)$		
St Kilda Road	Leopold Street	2.645[±0.033]	0.613[±0.0234]	0.43	0.6147
	Arthur Street	2.604[±0.035]	0.656[±0.025]	0.3468	0.9465
Sydney Road	Stewart Street	1.874[±0.0514]	0.479[±0.0363]	0.09238	0.4543
	Blyth Street	1.657[±0.0594]	0.545[±0.042]	0.1119	0.3868

One of the main factors we noticed in Section 5.2 was estimated parameters, μ and σ of log-normal distribution for passenger flow time do not vary much with the type of the tram stop. However, when we compare the estimated parameters of passenger flow time distribution at St Kilda Road and Sydney Road, estimated σ's are quite similar, but the μ's are different. The difference in estimated μ's at Sydney Road is significant with respect to the St Kilda Road. This significant difference maybe because all the tram stops we consider in Section 5.2 are roughly as busier as St Kilda Road and are quite closer to each other. However, the stops we surveyed in Sydney Road are less busy and quite far from the rest of the surveyed stops. We can observe that the estimated μ for passenger flow time distribution does depend on the location and the passenger demand of the desired stop.

5.3.3.2 Marginal distribution of time-headway

We analyse the surveyed data in order to understand the time-headway distributions of trams on selected routes as well. The time-headway distribution of trams in St Kilda Road is positively skewed while in Sydney Road it is closer to symmetrical. Lin and Ruan suggested that the time-headway of selected routes in Chicago can be well approximated by a gamma distribution [133]. Therefore, we assess the goodness-of-fit of gamma distribution and log-normal distribution in describing the time-headway data collected in St Kilda Road stops.

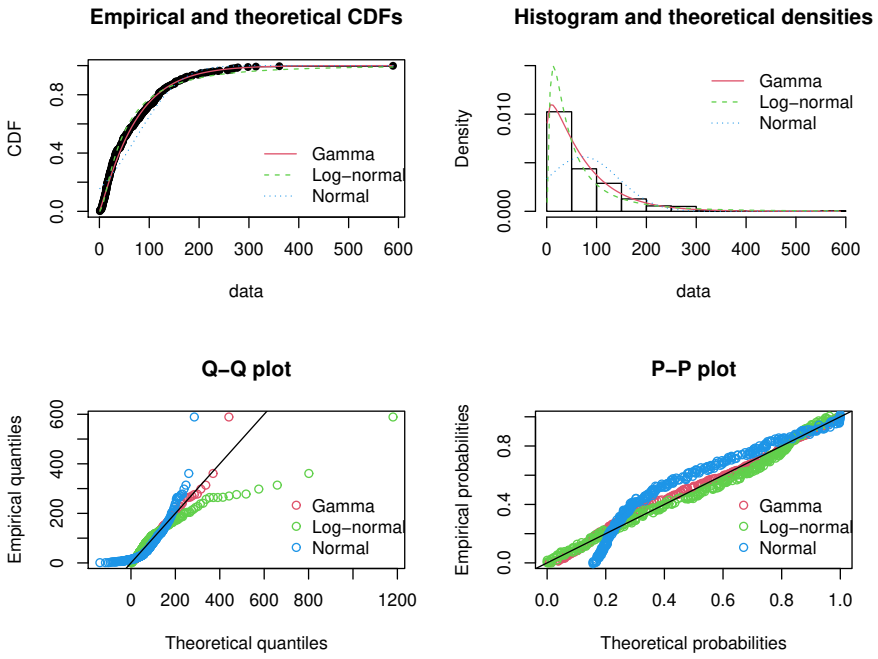


FIGURE 5.8: Survey: Four Goodness-of-fit plots for gamma distribution and log normal distribution for time-headway at Leopold Street stop

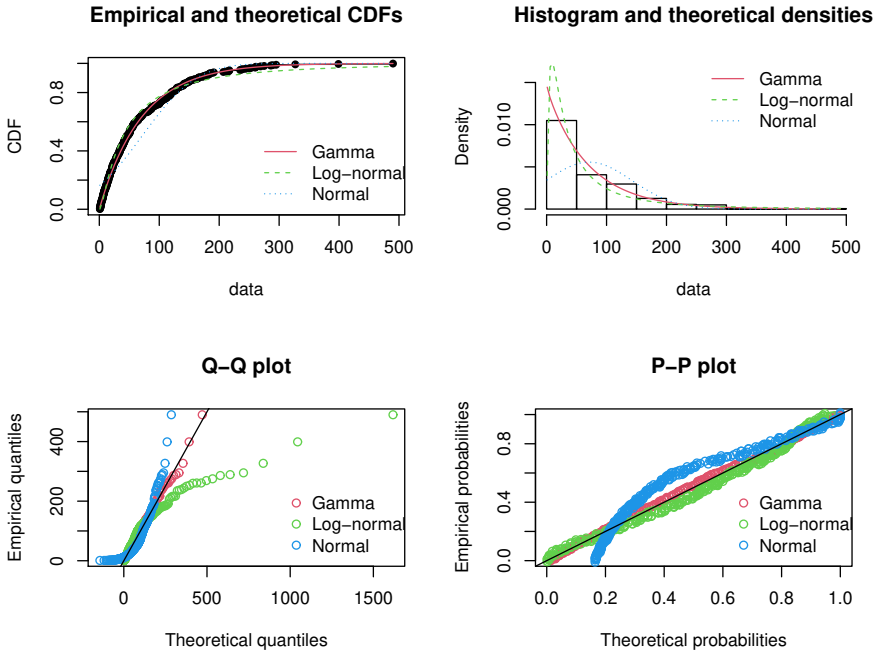
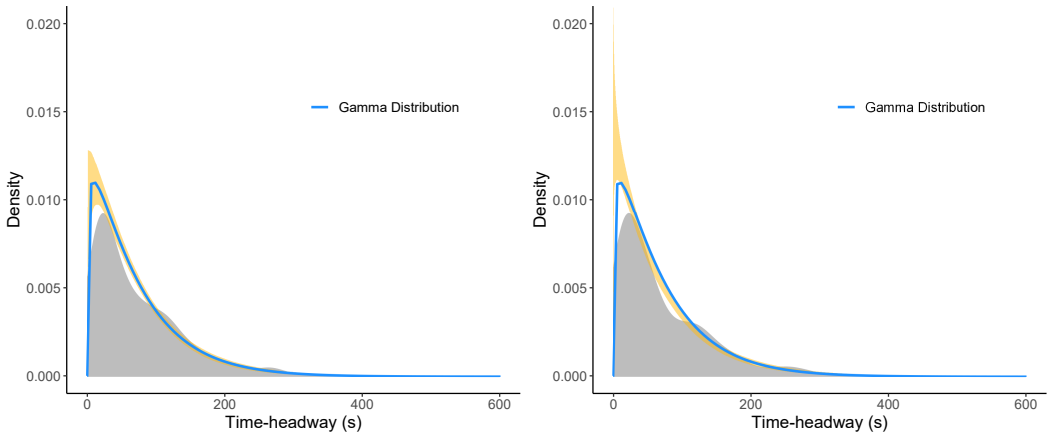


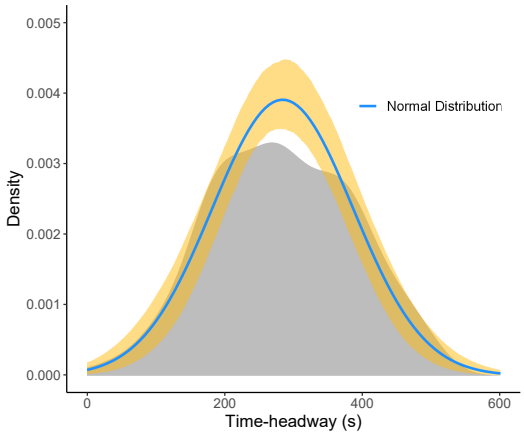
FIGURE 5.9: Survey: Four Goodness-of-fit plots for gamma distribution and log normal distribution for time-headway at Arthur Street stop

Figures 5.8 and 5.9 summarise the goodness-of-fit plots for gamma distribution and log-normal distribution for time-headway at Leopold Street stop and Arthur Street stop. The $Q-Q$ plot emphasises the lack-of-fit at the distribution tails while the $P-P$ plot emphasises the lack-of-fit at the centre of the distribution. In Figures 5.8 and 5.9, it is evident that the gamma distribution correctly describes both the centre and the right tail of the distribution. At St Kilda Road, time-headway between trams is highly skewed and well approximated by a gamma distribution as illustrated in Figures 5.10a and 5.10b. However, at Sydney Road, the time-headway distribution is more symmetric and appears to be described by a normal distribution as described in Figures 5.10c and 5.10d.

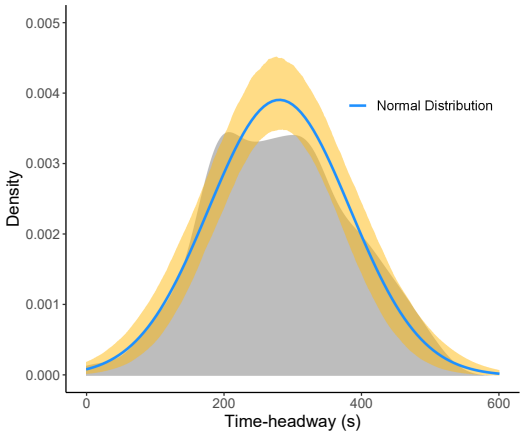


(a) St Kilda Road: Leopold Street Stop

(b) St Kilda Road: Arthur Street Stop



(c) Sydney Road: Stewart Street Stop



(d) Sydney Road: Blyth Street Stop

FIGURE 5.10: Survey: Distribution of time-headway. 95% confidence bands (yellow area) are included. Empirical distributions are smoothed with non-parametric Gaussian kernel shown in grey area.

At St Kilda Road trams are operated on dedicated lanes, so there is no impact from other traffic. In addition, St Kilda Road operates multiple tram

routes and has an irregular dispatching time-headway. When a route has an irregular dispatching time-headway, it is highly likely to occur tram bunching quite frequently (see Figure 5.6). Moreover, during the peak hour, St Kilda Road is highly crowded. The higher number of passengers will cause more delays at the stop and will increase the time-headway between preceding tram while decreasing the time-headway between the tram behind. In this way, we can expect a skewed time-headway distribution including a significant number of larger time-headways.

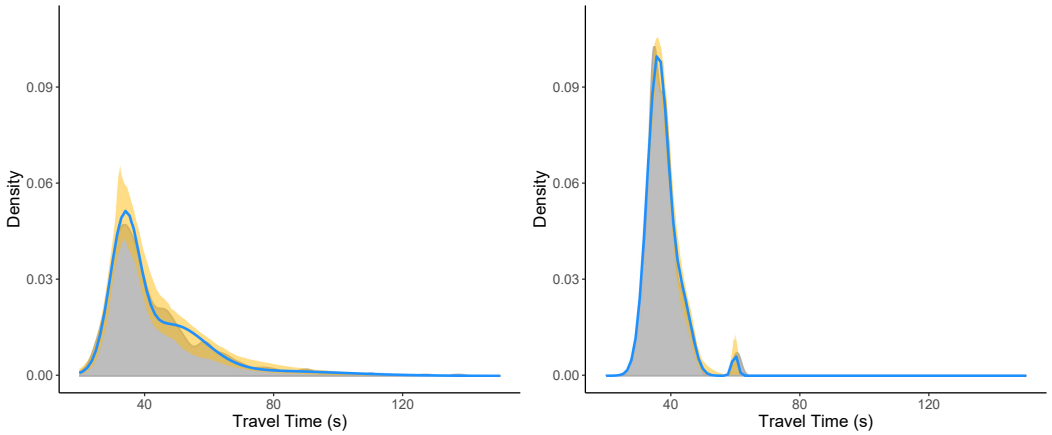
Even though the trams at Sydney Road operate under mixed traffic conditions, the observed time-headway distribution seems to be more symmetric than at St Kilda Road. In Sydney Road, a single tram route is operated with a regular time-headway of six minutes and are closely following the time table. Moreover, the passenger demand at Sydney Road is quite lower and trams are not likely to spend more time at the stops serving passengers. Therefore, the observed time-headway distributions at Sydney Road can be closely approximated by a normal distribution.

Table 5.3: Survey: Estimated parameters for time-headway distribution at surveyed routes in the Melbourne tram network.

St Kilda Road - Gamma Distribution				
	Shape	Rate	K-S test Sig.	Chi-square test Sig.
Leopold Street	1.1448[±0.07554]	0.01575[±0.00128]	0.1035	0.2129
Arthur Street	0.9859[±0.06392]	0.0139[±0.0012]	0.7324	0.3083
Sydney Road - Normal Distribution				
	Mean	SD	K-S test Sig.	Chi-square test Sig.
Stewart Street	283.83[±10.893]	102.18[±7.7023]	0.6202	0.5012
Blyth Street	280.5[±10.916]	102.39[±7.719]	0.6841	0.4928

5.3.3.3 Marginal distribution of link travel time

We analysed the observed travel time distributions at both St Kilda Road and Sydney Road. The empirical distribution at St Kilda Road is skewed with a long upper tail. The empirical distribution at Sydney Road is multimodal. This may be due to the traffic signals which operate between these consecutive stops. According to our findings in Chapter 3, we fitted the travel time distribution for trams between these two consecutive stops with Gaussian mixture models as illustrated in Figure 5.11.



(a) St Kilda Road (b) Sydney Road

FIGURE 5.11: Survey: Travel time distribution of trams between stops. Gaussian Mixture Model curve and 95% confidence bands are included. Empirical distributions are smoothed with non-parametric Gaussian kernel shown in grey area.

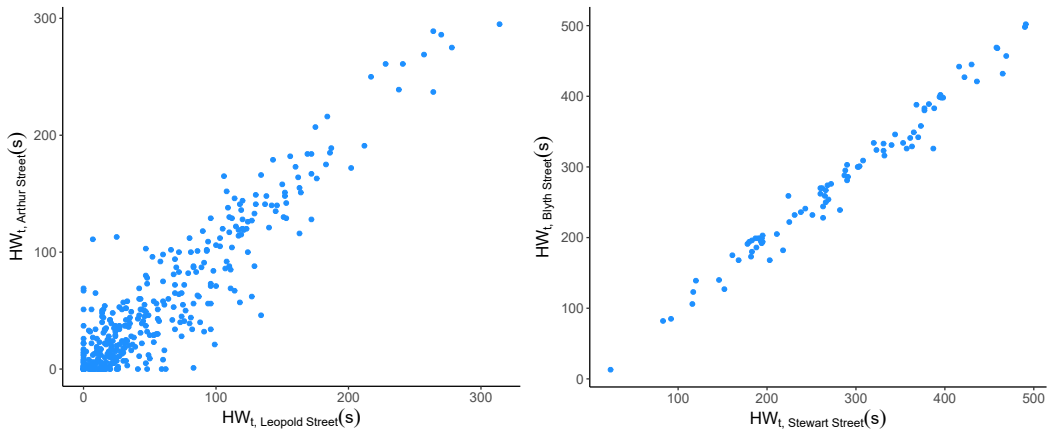
Table 5.4: Survey: Estimated Gaussian Mixture Model parameters for travel time distribution at surveyed routes in the Melbourne tram network

St Kilda Road					
	μ	σ	λ	K-S test Sig.	Chi-square test Sig.
Component 1	34.04[±0.47812]	4.48[±0.46566]	0.50[±0.062492]	0.604	0.8343
Component 2	48.35[±2.32012]	10.69[±1.43796]	0.41[±0.072338]		
Component 3	77.90[±10.15521]	23.66[±5.26319]	0.09[±0.039852]		
Sydney Road					
	μ	σ	λ	K-S test Sig.	Chi-square test Sig.
Component 1	36.03[±0.37610]	3.34[±0.20805]	0.84[±0.07993]	0.4919	0.5212
Component 2	43.88[±1.31467]	2.87[±0.50320]	0.15[±0.07997]		
Component 3	59.80[±0.27746]	0.88[±0.15309]	0.02[±0.00375]		

Table 5.4 summarises the estimated Gaussian Mixture Model parameters. The non-parametric test, Kolmogorov–Smirnov test and the Chi-squared test confirms that the fitted parameters of Gaussian Mixture Distribution for the travel times at both St kilda Road and Sydney Road are reasonably good fits. The majority of trams on Sydney Road could pass the intersection without any delay. However, on St kilda Road, half of the trams were delayed at the signalised intersection. In Sydney Road, estimated variances are very small compared to the estimated variances in St Kilda Road. When the route is less busy, the behaviour of trams are mostly similar. However, when the route becomes busier with a higher number of trams and higher passenger demand, travel behaviour will be changed drastically.

5.3.4 Correlation analysis

5.3.4.1 Correlation between two stops



(a) St Kilda Road (b) Sydney Road

FIGURE 5.12: Survey: Correlation of time-headway between consecutive stops. $HW_{t,\ stop\ name}$ represents the time-headway of tram t at a given stop.

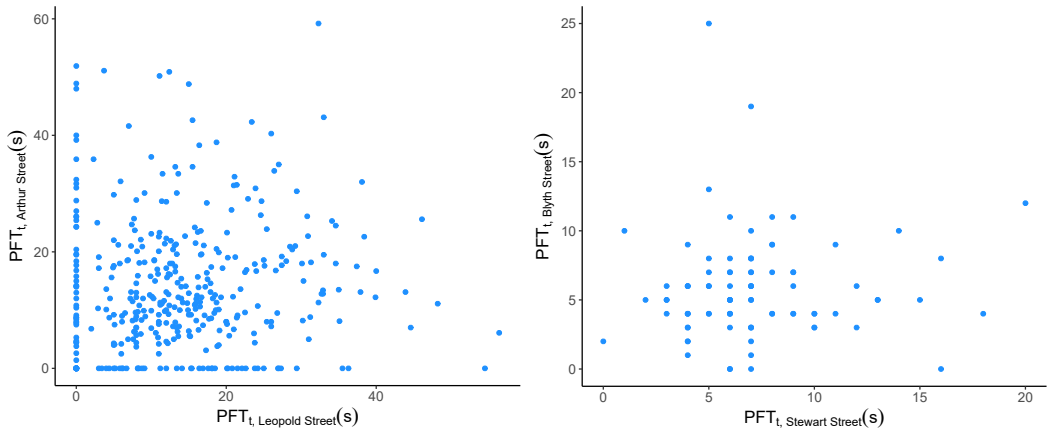
We denote $PFT_{t,\ stop\ name}$ as the passenger flow time of tram t at a given stop and $HW_{t,\ stop\ name}$ as the time-headway of tram t at a given stop. The Figure 5.12a illustrates the correlation of a tram's time-headway between Leopold Street stop $HW_{t,\ Leopold\ Street}$ and Arthur Street stop at St Kilda Road $HW_{t,\ Arthur\ Street}$. The Figure 5.12b illustrates the correlation of a tram's time-headway between Stewart Street stop $HW_{t,\ Stewart\ Street}$ and Blyth Street stop at Sydney Road $HW_{t,\ Blyth\ Street}$.

Time-headway between consecutive stops seems to be highly correlated (see Table 5.5). We can see the scheduled time table is maintained better on Sydney Road than on St Kilda Road. Sydney road is less busy and has a regular dispatching time-headways compared to St Kilda Road which operates multiple routes with irregular dispatching time-headways.

Table 5.5: Survey: Correlation coefficient of time-headway of a tram between stops.

	Correlation coefficient
St Kilda Road: $Corr(HW_{t,\ Leopold\ Street},\ HW_{t,\ Arthur\ Street})$	0.9237[±0.0196]
Sydney Road: $Corr(HW_{t,\ Stewart\ Street},\ HW_{t,\ Blyth\ Street})$	0.988[±0.0166]

The Figure 5.13a illustrates the correlation of a tram's passenger flow time between Leopold Street stop $PFT_{t,\ Leopold\ Street}$ and Arthur Street stop at St Kilda Road $PFT_{t,\ Arthur\ Street}$. The Figure 5.13b illustrates the correlation of a tram's passenger flow time between Stewart Street stop $PFT_{t,\ Stewart\ Street}$ and Blyth Street stop at Sydney Road $PFT_{t,\ Blyth\ Street}$.



(a) St Kilda Road (b) Sydney Road

FIGURE 5.13: Survey: Correlation of Passenger Flow Time between consecutive stops. $PFT_{t,\ stop\ name}$ represents the passenger flow time of tram t at a given stop

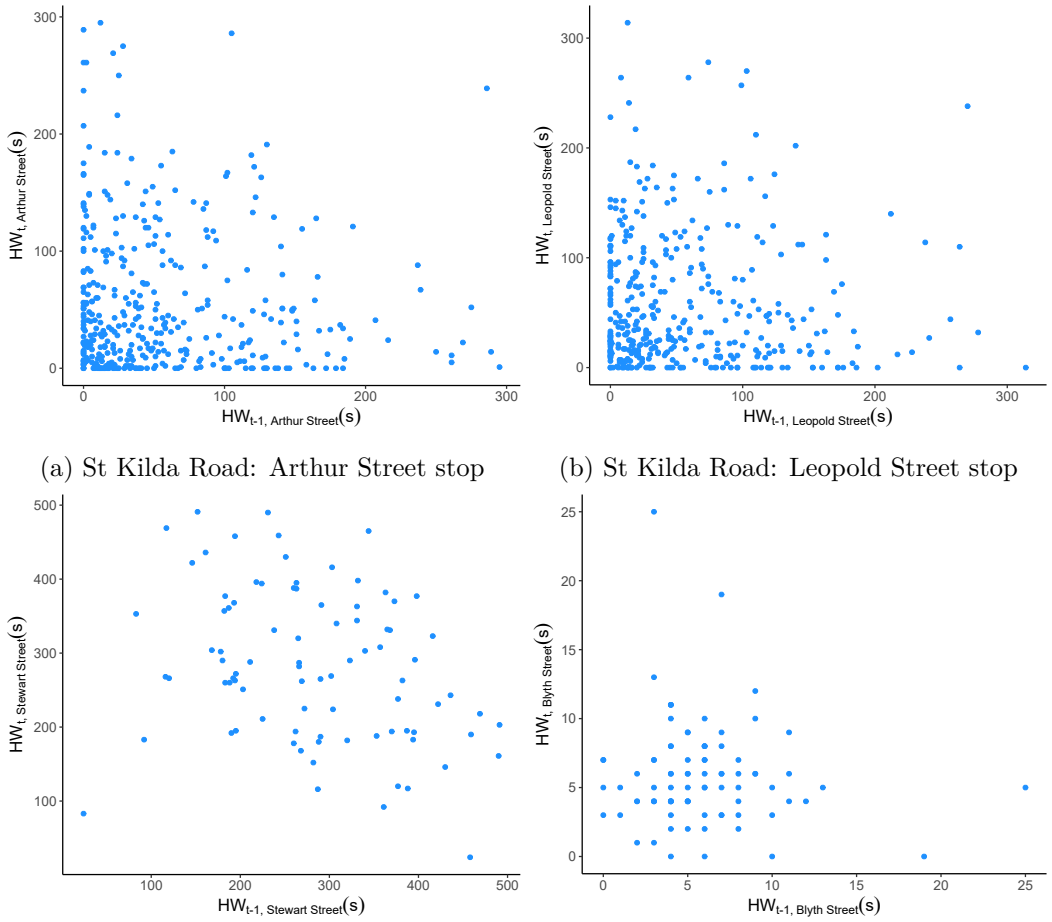
Table 5.6: Survey: Correlation coefficient of passenger flow time of a tram between stops.

	Correlation coefficient
St Kilda Road: $Corr(PFT_{t,\ Leopold\ Street},\ PFT_{t,\ Arthur\ Street})$	0.0594[±0.051]
Sydney Road: $Corr(PFT_{t,\ Stewart\ Street},\ PFT_{t,\ Blyth\ Street})$	0.0528[±0.108]

We can assume that the passenger demand at both surveyed stops in each route are quite similar as the stops are very closer to each other (Stewart Street stop and Blyth street stop are only 210m apart and Leopold Street stop and Arthur Street Stop are only 280m apart). Therefore, when a tram got delayed and closely followed the preceding tram, the first tram tends to serve more passengers at both the stops and have a higher passenger flow time while the second tram tends to serve fewer passengers and have a lower passenger flow time at both the stop. If $t-1^{th}$ tram and t^{th} tram are closely operating, then $PFT_{t-1,s_1} > PFT_{t,s_1}$ and $PFT_{t-1,s_2} > PFT_{t,s_2}$, where t represents the t^{th} tram and s_i represents the i^{th} stop in the route. Therefore, we could expect a positive correlation in passenger flow time of a tram between two consecutive stops.

However, the Figure 5.13 and the Table 5.6 illustrates that the passenger flow time of a tram between two consecutive stops does not show a significant association. At St Kilda Road we noticed that, when the tram becomes overcrowded as well as there's a tram closely behind, the tram operator chosen to bypass the following stop. This might be the reason we could not observe a significant spatial correlation in passenger flow time of a tram.

5.3.4.2 Correlation between two consecutive trams



(a) St Kilda Road: Arthur Street stop

(b) St Kilda Road: Leopold Street stop

(c) Sydney Road: Stewart Street stop

(d) Sydney Road: Blyth Street stop

FIGURE 5.14: Survey: Correlation of time-headway between consecutive trams at each stop. $HW_{t,\ stop\ name}$ represents the time-headway of tram t at a given stop.

If a tram got delayed and gets the time-headway bigger causing more passengers waiting at the stop, it will have a longer passenger flow time. Consequently, the tram closely following the delayed tram tend to have lower passenger

flow time due to the few number of passengers waiting at the stop. When we consider the nature of correlation between two consecutive trams, we can expect a negative correlation between passenger flow time of consecutive trams and time-headway of two consecutive trams. That is the correlation between PFT_{t-1,s_i} and PFT_{t,s_i} is negative. Moreover, the correlation between HW_{t-1,s_i} and HW_{t,s_i} is also negative.

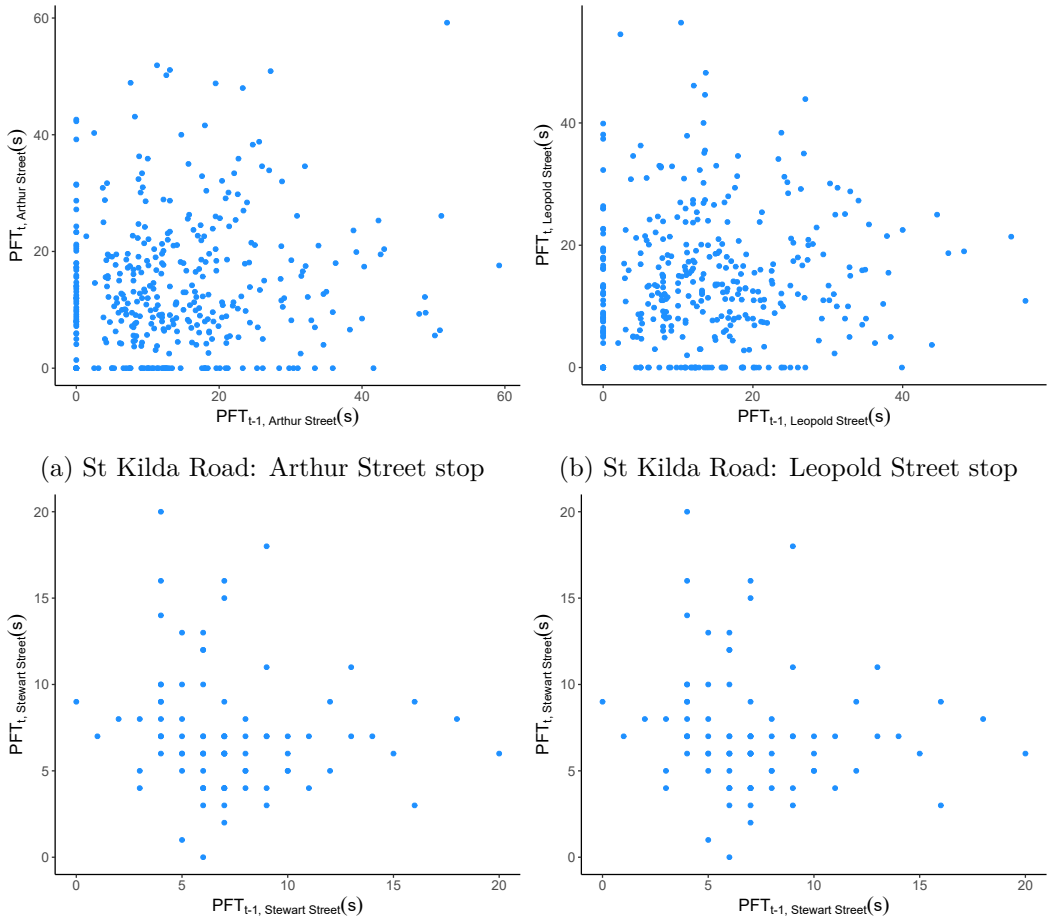
Table 5.7: Survey: Correlation coefficient of time-headway between consecutive trams at each stop.

		Correlation Coefficient
St Kilda Road	$Corr(HW_{t-1,\ Leopold\ Street},\ HW_{t,\ Leopold\ Street})$	$-0.0399[\pm 0.0511]$
	$Corr(HW_{t-1,\ Arthur\ Street},\ HW_{t,\ Arthur\ Street})$	$-0.0757[\pm 0.0509]$
Sydney Road	$Corr(HW_{t-1,\ Stewart\ Street},\ HW_{t,\ Stewart\ Street})$	$-0.308[\pm 0.1032]$
	$Corr(HW_{t-1,\ Blyth\ Street},\ HW_{t,\ Blyth\ Street})$	$-0.323[\pm 0.1027]$

The negative correlation is quite evident at Sydney Road as illustrated in Tables 5.7 and 5.8. See Figures 5.14 and 5.15. However, passenger flow time has a very week correlation. At St Kilda Road, it is evident that the correlation of time-headway between two consecutive trams is slightly negative as summarised in Table 5.7. However, there is no evidence of significant negative correlation of passenger flow time between consecutive trams at St Kilda Road. This might be due to the tram operators at St Kilda Road deliberately bypass some stops whenever there are no passengers to get off and when there is another tram closely behind.

Table 5.8: Survey: Correlation coefficient of passenger flow time between consecutive trams.

		Correlation Coefficient
St Kilda Road	$Corr(PFT_{t-1,\ Leopold\ Street},\ PFT_{t,\ Leopold\ Street})$	$0.0685[\pm 0.0509]$
	$Corr(PFT_{t-1,\ Arthur\ Street},\ PFT_{t,\ Arthur\ Street})$	$0.0825[\pm 0.0509]$
Sydney Road	$Corr(PFT_{t-1,\ Blyth\ Street},\ PFT_{t,\ Blyth\ Street})$	$-0.085[\pm 0.1081]$
	$Corr(PFT_{t-1,\ Stewart\ Street},\ PFT_{t,\ Stewart\ Street})$	$-0.128[\pm 0.1076]$



(a) St Kilda Road: Arthur Street stop

(b) St Kilda Road: Leopold Street stop

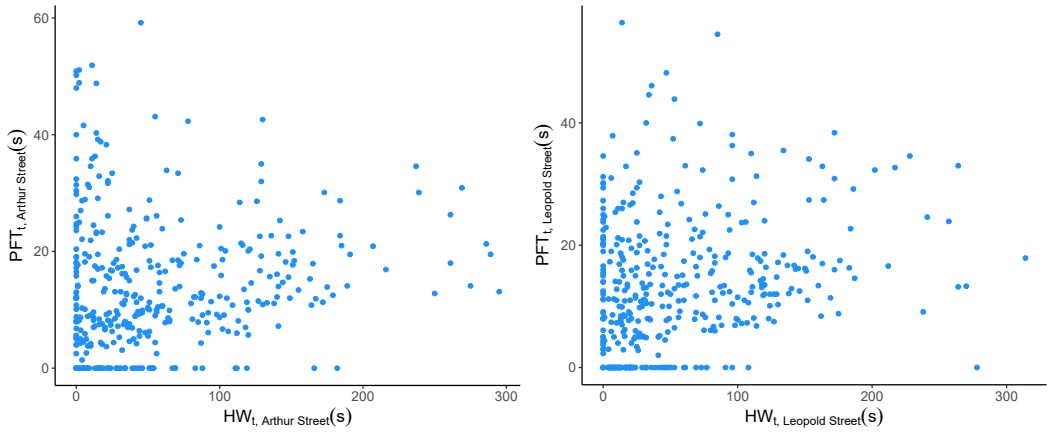
(c) Sydney Road: Stewart Street stop

(d) Sydney Road: Blyth Street stop

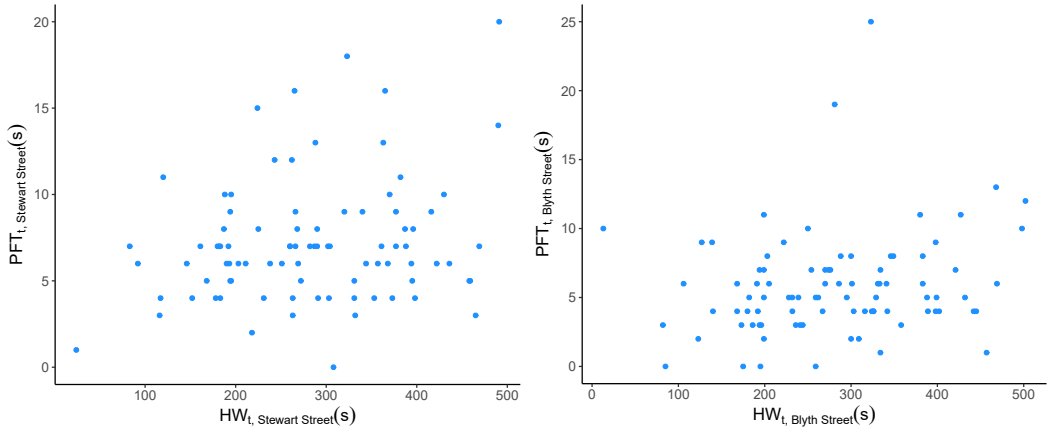
FIGURE 5.15: Survey: Correlation of passenger flow time between consecutive trams at each stop. $PFT_{t,\ stop\ name}$ represents the passenger flow time of tram t at a given stop.

5.3.4.3 Passenger flow time and time-headway correlation

If a tram has a shorter time-headway, there has been less time for passengers to arrive at the stop. Consequently, a tram tends to have a shorter passenger flow time when it's time-headway is shorter. Contrarily a tram tends to have a longer passenger flow time when it's time-headway is longer. Hence, we expect a positive correlation between time-headway and passenger flow time at each stop. As illustrated in Figure 5.16 and Table 5.9, the passenger flow time tends to have a week positive association with the time-headway at each of the four stops we surveyed.



(a) St Kilda Road: Arthur Street Stop (b) St Kilda Road: Leopold Street Stop



(c) Sydney Road: Stewart Street Stop (d) Sydney Road: Blyth Street Stop

FIGURE 5.16: Survey: Correlation between time-headway and passenger flow time of a tram at each stop. $HW_{t,\ stop\ name}$ and $PFT_{t,\ stop\ name}$ represents the time-headway and passenger flow time of tram t at a given stop.

Table 5.9: Survey: Correlation coefficient between time-headway and passenger flow time of a tram at each stop.

		Correlation Coefficient
St Kilda Road	$Corr(HW_{t,\ Leopold\ Street},\ PFT_{t,\ Leopold\ Street})$	0.2244[±0.0497]
	$Corr(HW_{t,\ Arthur\ Street},\ PFT_{t,\ Arthur\ Street})$	0.1072[±0.0507]
Sydney Road	$Corr(HW_{t,\ Blyth\ Street},\ PFT_{t,\ Blyth\ Street})$	0.2248[±0.1051]
	$Corr(HW_{t,\ Stewart\ Street},\ PFT_{t,\ Stewart\ Street})$	0.1959[±0.1057]

5.4 Simulation

We develop a simple stochastic simulation model describing a one-way tram line, which accounts dwell time at stops, time-headway between trams and passenger demand at each stop. We utilise this simulation model to understand the correlation structure and stopping behaviour of a tram operating in a high frequency, busy route with irregular dispatching time-headways. We utilise the information obtained from the survey at st Kilda Road, such as link travel time and passenger demand as inputs to the simulation model.

5.4.1 Model

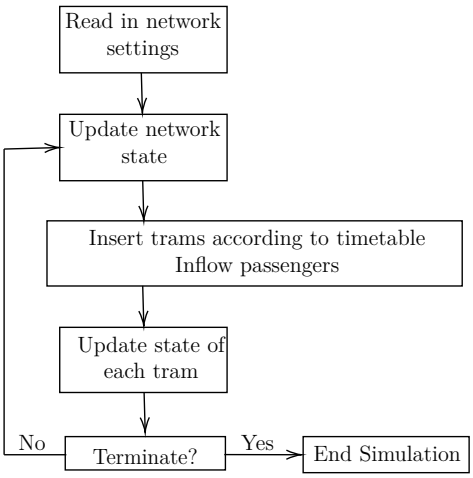


FIGURE 5.17: Simplified flowchart of the simulation model.

A stochastic model was constructed to study the stopping behaviour of trams on a dedicated lane. Figure 5.17 shows a flowchart of the proposed model. In the initialisation process, the model loads the input parameters that describe the route and the stop information such as the passenger demand. Each tram is dispatched from the depot according to a fixed time table. Passengers arrive randomly at each stop at a time dependent rate of r_t, where each passenger has a different destination. When a tram first arrives at a link l, its travel time on the link is randomly drawn from the empirical travel time distribution. At each stop O, trams dwell to serve passengers and the dwell time at each stop is determined by the number of boarding and alighting passengers at time t at stop O. The model records the time-headway and passenger flow time for each

tram at each stop. The simulation terminates when all the trams have been dispatched and have left the system.

5.4.2 System configuration

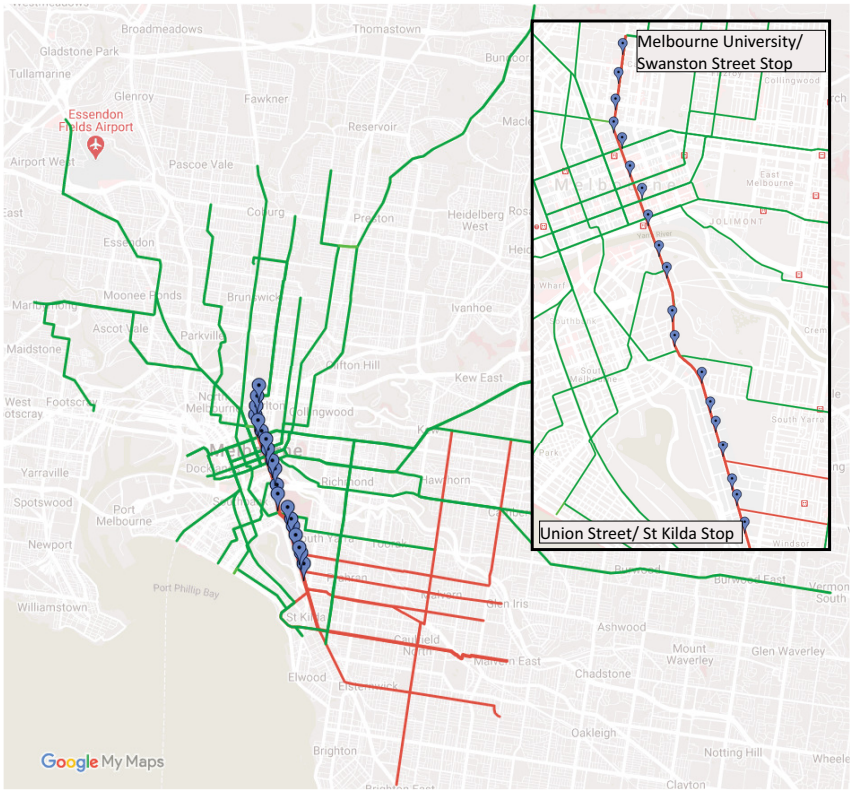


FIGURE 5.18: Simulated route and stop locations on St Kilda Road

We considered 96 trams, matching up with the scheduled timetable at Leopold street stop in St Kilda Road during the evening peak 4:00PM-6:00PM. 19 stops, numbered from 01 to 19, were considered where stop 01 represents the Union Street Stop, stop 12 represents the Flinders Street Station stop and stop 19 represents the Melbourne University/Swanston St Stop as described in Figure 5.18. We considered the stop 12 (Flinders Street stop) as the Major stop on the simulated route, where majority of the passengers alight. Stop 05 and 06 were considered as the two distinct stops we surveyed, Leopold Street stop and Arthur Street Stop. Random travel times between stops were sampled

from the empirical travel times between Leopold Street stop and Arthur Street stop assuming all the other links were functioning similar to the observed link. The maximum capacity of a tram was fixed at 300 passengers.

The passenger data is not measured in our survey. In principle, we could estimate passenger demand utilising the multivariate regression model from [24]. However, we did not have data for parameters c, d, e, f. Therefore, we performed a regression analysis for the passenger flow time depending only on the number of boarding and alighting passengers using the data from [24].

$$PFT = 5.98 + 0.69 \times Boarding + 0.55 \times Alighting \tag{5.1}$$

Table 5.10: Multivariate Regression: Passenger Flow Time

Multivariate Regression: Passenger Flow Time		
Multiple R-Squared	0.6623	
Adjusted R-Squared	0.6618	
Coefficients	Estimate	P-Value
Intercept	5.97717	≈ 0
Boarding	0.68692	≈ 0
Alighting	0.55240	≈ 0

During the survey, we noticed that the number of alighting passengers was comparatively low at Leopold Street stop and Arthur Street Stop on St Kilda Road. Therefore, we assumed that the passenger flow time in these two stops were determined largely by the number of boarding passengers. Assuming no alighting passengers at these two stops, we then approximated the passenger demand at each stop for fifteen-minute time windows using Equation (5.1). We estimated the passenger arrival rates for Leopold Street stop, r_t^L and Arthur Street stop, r_t^A in St Kilda Road separately assuming the passenger arrival rates are consistent within these time windows. We then estimated the time dependent passenger arrival rate, r_t, averaging over two surveyed stops (i.e. $r_t = (r_t^L + r_t^A)/2$) as summarised in Table 5.11.

Table 5.11: Time dependent passenger arrival rates. Averaging over two surveyed stops (i.e. $r_t = (r_t^L + r_t^A)/2$)

Time window	r_t pass/s	Time window	r_t pass/s
4:00-4:15PM	0.164	5:15-5:30PM	0.275
4:15-4:30PM	0.145	5:30-5:45PM	0.112
4:30-4:45PM	0.181	5:45-6:00PM	0.112
4:45-5:00PM	0.242	6:00-6:15PM	0.112
5:00-5:15PM	0.320	6:15-6:30PM	0.112

There were a no passenger flow time observations after 5:45PM as we did not conduct the survey after 5:45PM. Therefore, we assume the passenger arrival rate remains constant at the rate observed from 5:30PM-5:45PM until 6:30PM.

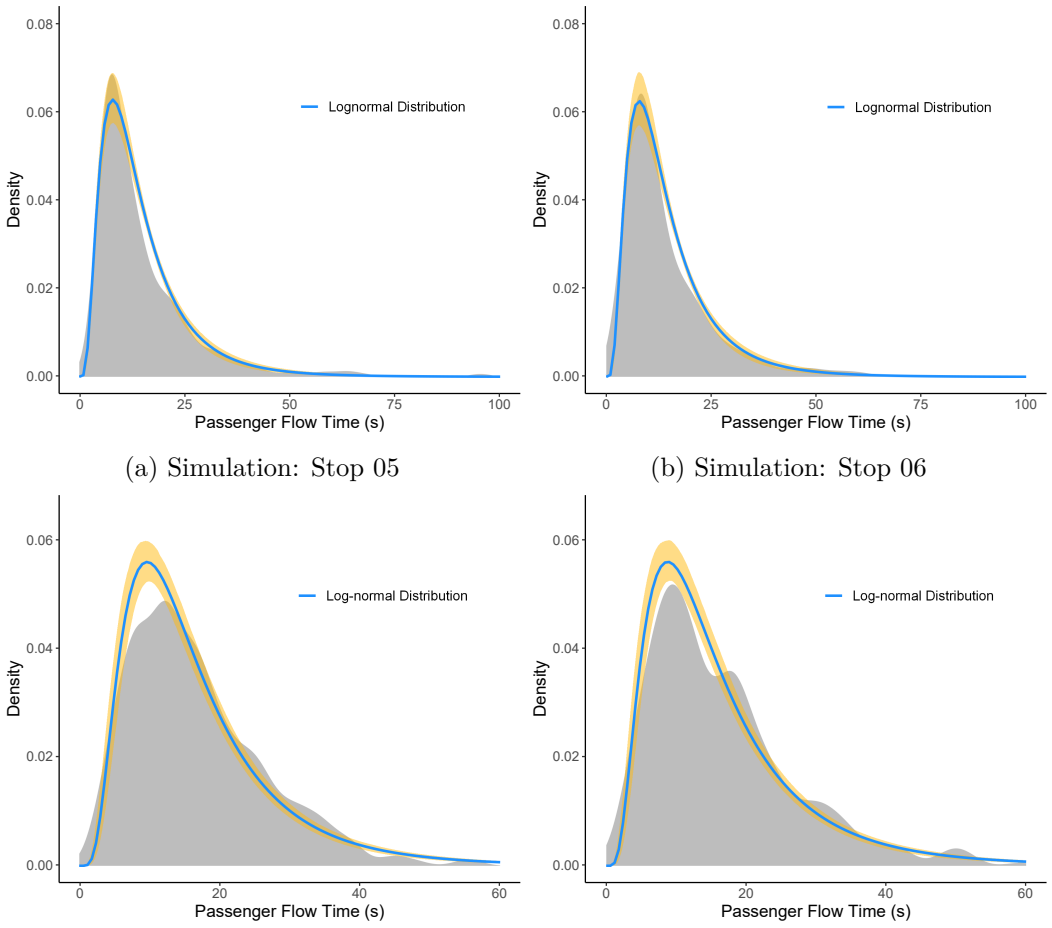
Union Street stop in St Kilda Road (Stop 01 in our simulation) is surrounded by many offices and therefore we consider it a busy stop with a higher demand. We assume a lower passenger demand at the High Street stop and Maubray Street stop (Stop 02 and 03 in our simulation) as they are located quite close together, dividing the passenger flow between them. Again, from Commercial Street stop to Arthur Street stop (Stop 04 to Stop 06 in our simulation), we considered high passenger demand as these stops are surrounded by many offices. All the stops after Toorak Road stop up to Flinders Street stop (Stop 12 in our simulation) were considered as less busy due to them being closer to the Flinders Street stop. Flinders Street stop is the main destinations of the passengers, as a majority of them get off at Flinders Street stop to catch trains during the evening peak time. Therefore, in the simulation, we consider stop 12 as the main stop for the majority of the passengers to get off. We considered a lower passenger demand after Flinders Street stop as these stops are less busy during the evening peak time. As a summary, we fixed the time-dependent passenger arrival rates, $0.5 \times r_t$, at stops 02, 03, 07-18 and, r_t, at stops 01, 04-06.

We determined the destinations for passengers from each stop according to a typical evening peak scenario. We considered most of the passengers who board before stop 12 have their destination set to stop 12: 60% of the passengers boarding before stop 12 alight at stop 12, and 40% of them alight uniformly at the other stops. Passengers who board after Stop 12 will alight uniformly at all the following stops. We simulated the system described in Section 5.4.1 for 2.5 hours in evening peak 4:00PM to 6:30PM for five days. We recorded the time-headway and dwell time for each tram at each stop.

5.5 Numerical results

We chose the parameters we described earlier in order to represents the tram route in St Kilda road from Union Street stop to Melbourne University/Swanston Street Stop. The passenger arrival rates we consider at stops 05 and 06 were mostly similar to the Leopold Street stop and Arthur Street stop respectively. Therefore, we compared the marginal distributions of passenger flow time and time-headway at both the simulated tram stops 05 and 06 with the two surveyed stops at St kilda Road: Leopold street stop and Arthur

street stop. According to the fitted regression model, in Equation5.1, time taken to open or close the tram doors is 2.99s. However, this value is slightly higher than the value estimated in [24]. Our model does not incorporate all the factors affect to the constant term in dwell time. Therefore, we derived passenger flow times from the dwell time data by considering the time taken to open or close the tram doors as 2.1s.



(a) Simulation: Stop 05

(b) Simulation: Stop 06

(c) Survey: Leopold Street stop

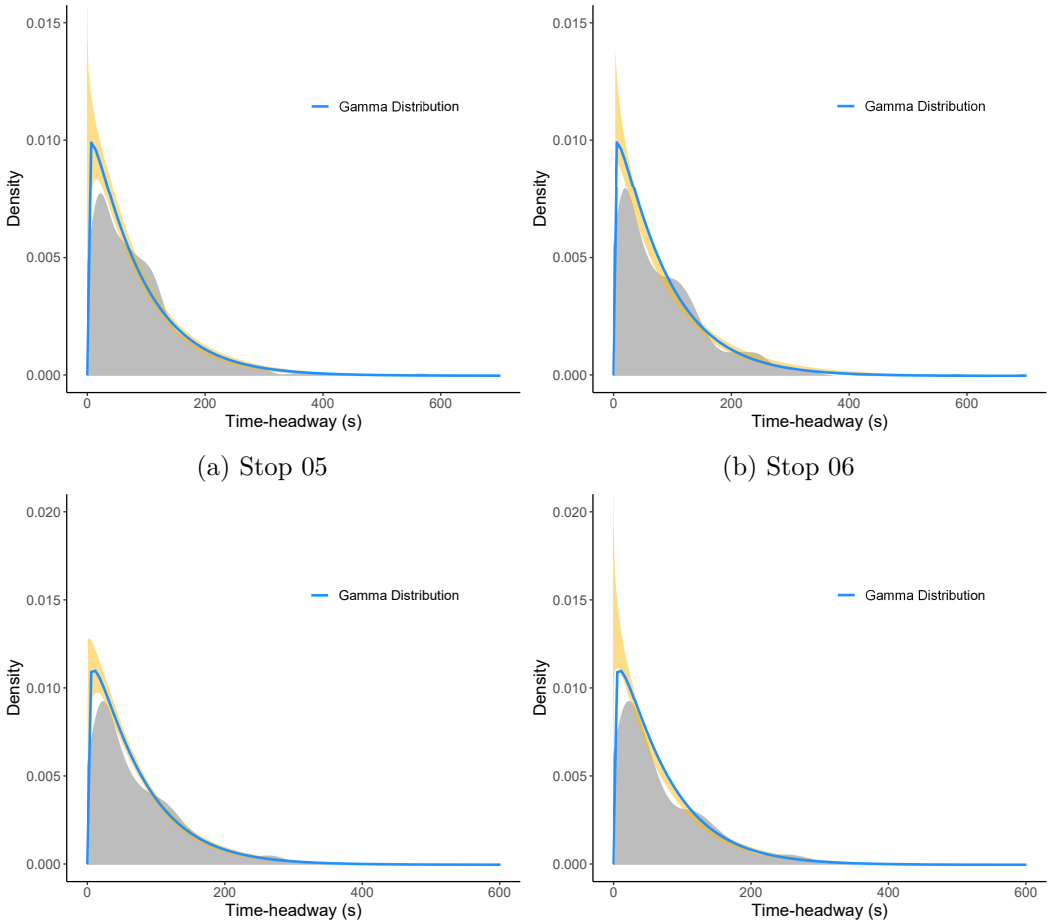
(d) Survey: Arthur Street stop

FIGURE 5.19: Comparison: Distribution of Passenger Flow Time. 95% confidence bands (yellow area) are included. Simulated and empirical distributions are smoothed with non-parametric Gaussian kernel which is shown in grey area.

We analysed the marginal distributions of passenger flow time at the simulated stops as summarised in Figure 5.19. We fitted log-normal distributions to the passenger flow times.

Table 5.12: Comparison: Log-normal distribution parameter estimation of passenger flow time at stop 05 and 06

Stop ID	$\log(\mu)$	$\log(\sigma)$	Chi-Square Sig.
Simulation: Stop 05	2.495[±0.0304]	0.649[±0.0215]	0.0005
Simulation: Stop 06	2.4911[±0.0312]	0.6572[±0.0220]	0.0005
Survey: Leopold Street	2.645[±0.033]	0.613[±0.0234]	0.6147
Survey: Arthur Street	2.604[±0.035]	0.656[±0.025]	0.9465



(a) Stop 05

(b) Stop 06

(c) Leopold Street

(d) Arthur Street

FIGURE 5.20: Simulation: Distribution of time-headways. 95% confidence bands (yellow area) are included. Simulated and empirical distributions are smoothed with non-parametric Gaussian kernel which is shown in grey area.

Log-normal distribution parameter estimations for stops 05 and 06 are summarised in Table 5.12. Even though the goodness-of-fit is poor, the estimated log-normal distribution parameters of the marginal passenger flow times distributions at stops 05 and 06 are mostly similar to the estimated log-normal distribution parameters of marginal distribution of passenger flow time

at the Leopold street stop and the Arthur street stop respectively. See Table 5.2. The marginal passenger flow time distributions of these two simulated stops are similar to what we observed in the survey at St Kilda Road. It appears that the dwell times we obtained from the simulations agrees with the empirical results.

Furthermore, we compared the marginal distributions of the simulated tram time-headways at stops 05 and 06 with the two surveyed stops at St Kilda Road. Figure 5.20 summarises the marginal time-headway distributions of trams at stops 05 and 06. We fitted gamma distribution to the time-headways at Stops 05 and 06. Estimated parameters for gamma distribution are summarised in Table 5.12. For both stops, the goodness-of-fit is high. The estimated parameters of the gamma distributions of marginal time-headway distribution at the simulated stops are also similar to the estimated gamma distribution parameters of marginal time-headway distributions at Leopold street stop and Arthur street stop. See Table 5.13.

Table 5.13: Simulation: Gamma distribution parameter estimation of passenger flow time at stops 05 and 06

Stop ID	Shape	Rate	Chi-Square Sig.
Stop 05	1.09008[±0.0841]	0.012868[±0.0012]	0.82059
Stop 06	0.957596[±0.07392]	0.010935[±0.0011]	0.384808
Stop 05	1.1448[±0.07554]	0.01575[±0.00128]	0.2129
Stop 06	0.9859[±0.06392]	0.0139[±0.0012]	0.3083

5.5.1 Correlation analysis

5.5.1.1 Correlation between two stops

Figure 5.21 summarises the correlation matrix of the dwell times of trams between stops along the route (i.e., $Corr(DT_{(t,i)}, DT_{(t,j)})$, where $DT_{(t,i)}$ is the dwell time of tram t at stop i. Dwell times at stops before stop 12 are positively correlated with the dwell time at the previous stop. We noticed the correlation decreases as the distance between stops increases. Dwell times at each pair of stops after stop 12 display a high positive correlation.

Figure 5.22 summarises the correlation matrix of the tram time-headways between stops along the route i.e., $Corr(HW_{(t,i)}, HW_{(t,j)})$, where $HW_{(t,i)}$ is the time-headway of tram t at stop i. The time-headways at each pair of stops display positive correlation and the strength of these correlations is

the strongest between consecutive stops. We observed that the correlation is monotonically decreases as the distance between stops increases.

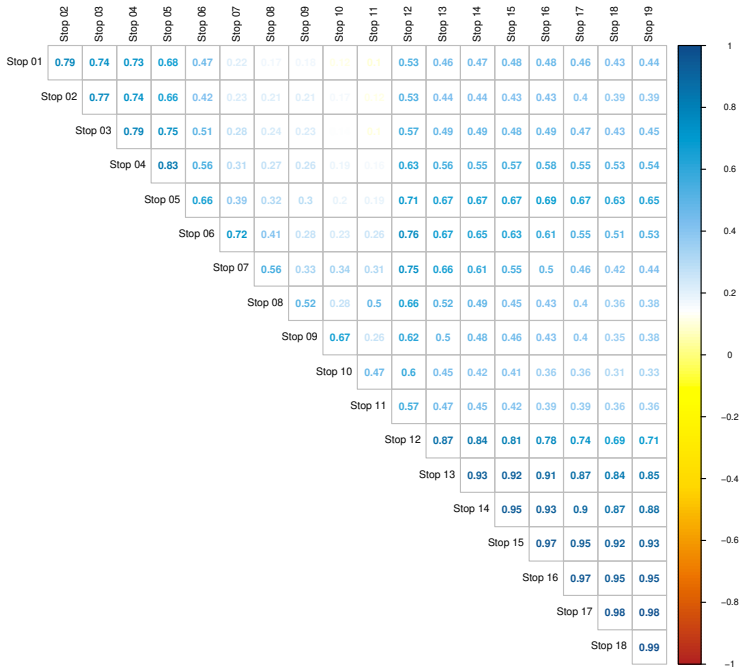


FIGURE 5.21: Simulation: Correlation matrix of dwell times of trams between stops. Colour scheme representing the direction of the association is shown.

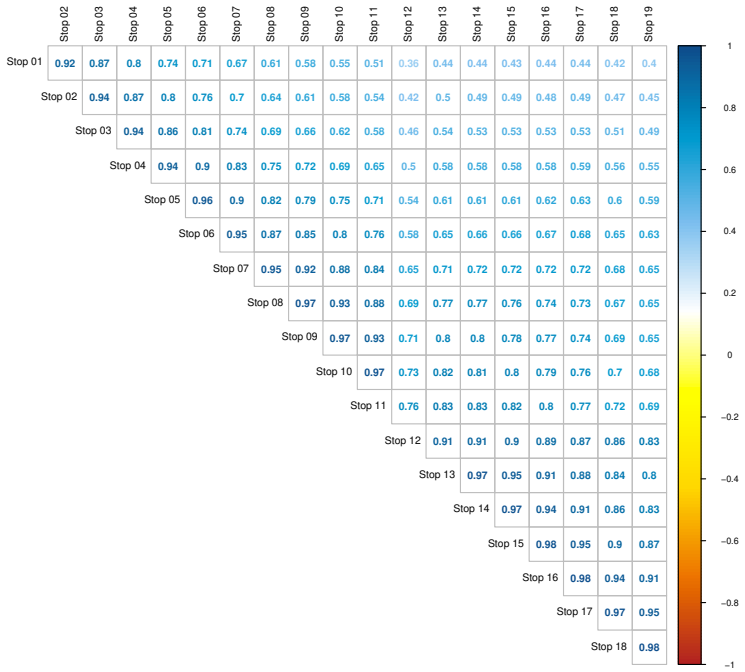


FIGURE 5.22: Simulation: Correlation matrix of time-headways of trams between stops. Colour scheme representing the direction of the association is shown.

5.5.1.2 Correlation between two consecutive trams

In this subsection we discussed the correlation of dwell times and time-headways between consecutive trams.

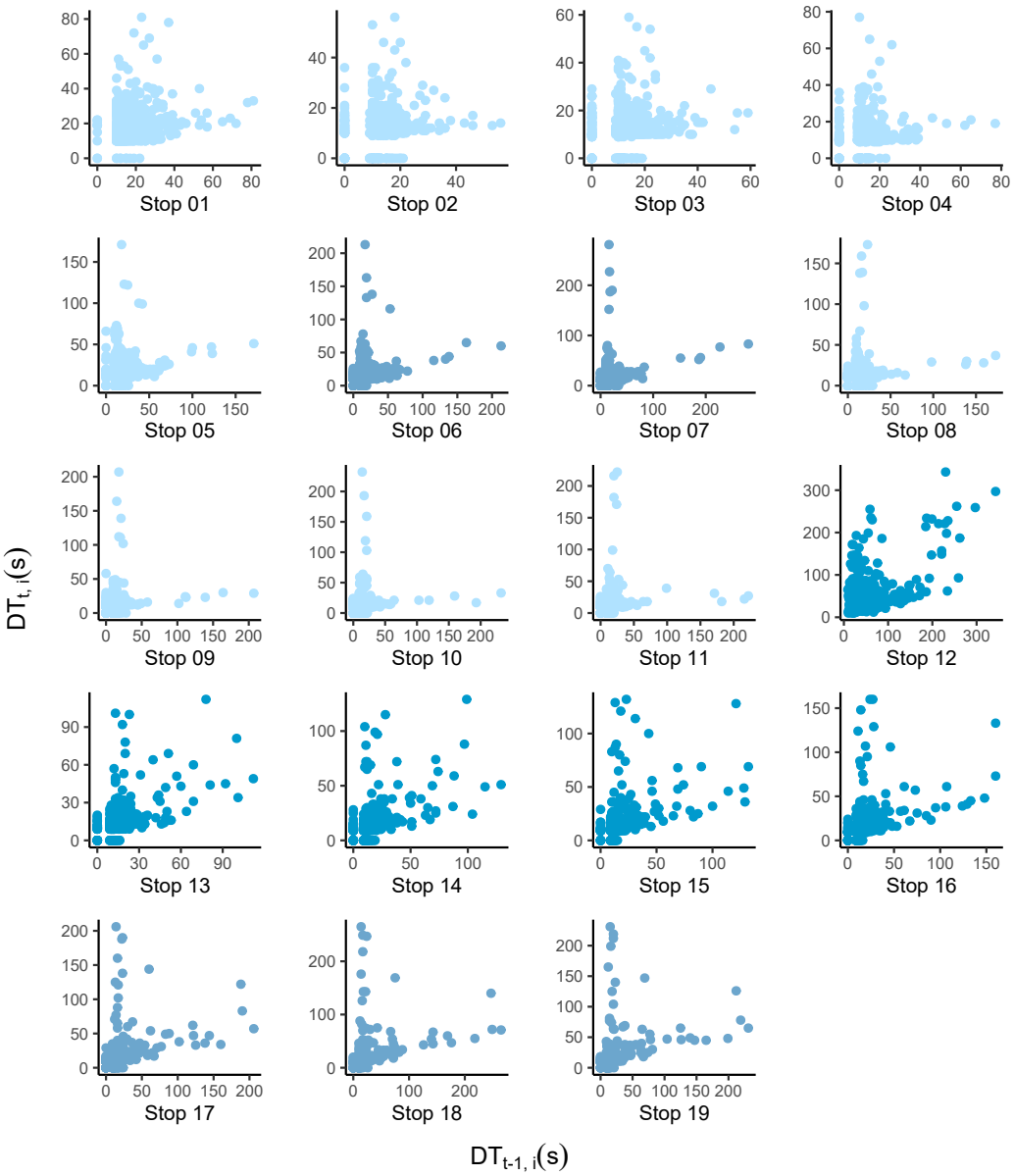


FIGURE 5.23: Simulation: Correlation of dwell times between tram $t-1$ and t. Colour scale: dark blue to light blue represents the high positive correlation to slight positive correlation.

Figure 5.23 summarises the correlation of dwell times between tram $t-1$ and t. Table 5.14 summarises the estimated correlation coefficients based on the Pearson's correlation coefficient. When the preceding tram has a longer

dwell time, the following one tends to have a shorter dwell time i.e. the dwell time between tram $t-1$ and tram t is negatively correlated.

Table 5.14: Simulation: Correlation between dwell time of consecutive trams. $Corr(DT_{(t-1,i)}, DT_{(t,i)})$, for $i = 01, 02, ..., 19$. $DT_{(t,i)}$ represents the dwell time of tram t at stop i.

	Correlation Coefficient	
$Corr(DT_{(t-1,01)}, DT_{(t,01)})$	0.175	[±0.04527]
$Corr(DT_{(t-1,02)}, DT_{(t,02)})$	0.151	[±0.04545]
$Corr(DT_{(t-1,03)}, DT_{(t,03)})$	0.164	[±0.04536]
$Corr(DT_{(t-1,04)}, DT_{(t,04)})$	0.112	[±0.04569]
$Corr(DT_{(t-1,05)}, DT_{(t,05)})$	0.194	[±0.04510]
$Corr(DT_{(t-1,06)}, DT_{(t,06)})$	0.231	[±0.04474]
$Corr(DT_{(t-1,07)}, DT_{(t,07)})$	0.225	[±0.04480]
$Corr(DT_{(t-1,08)}, DT_{(t,08)})$	0.152	[±0.04480]
$Corr(DT_{(t-1,09)}, DT_{(t,09)})$	0.133	[±0.04557]
$Corr(DT_{(t-1,10)}, DT_{(t,10)})$	0.123	[±0.04563]
$Corr(DT_{(t-1,11)}, DT_{(t,11)})$	0.134	[±0.04557]
$Corr(DT_{(t-1,12)}, DT_{(t,13)})$	0.576	[±0.03759]
$Corr(DT_{(t-1,13)}, DT_{(t,13)})$	0.521	[±0.03926]
$Corr(DT_{(t-1,14)}, DT_{(t,14)})$	0.497	[±0.03989]
$Corr(DT_{(t-1,15)}, DT_{(t,15)})$	0.456	[±0.04093]
$Corr(DT_{(t-1,16)}, DT_{(t,16)})$	0.417	[±0.04180]
$Corr(DT_{(t-1,17)}, DT_{(t,17)})$	0.358	[±0.04294]
$Corr(DT_{(t-1,18)}, DT_{(t,18)})$	0.327	[±0.04345]
$Corr(DT_{(t-1,19)}, DT_{(t,19)})$	0.344	[±0.04318]

Table 5.15: Simulation: Correlation between time-headway of consecutive trams. $Corr(HW_{(t-1,i)}, HW_{(t,i)})$, for $i = 01, 02, ..., 19$. $HW_{(t,i)}$ represents the time-headway of tram t at stop i.

	Correlation Coefficient	
$Corr(HW_{(t-1,01)}, HW_{(t,01)})$	-0.079	[±0.04609]
$Corr(HW_{(t-1,02)}, HW_{(t,02)})$	-0.109	[±0.04596]
$Corr(HW_{(t-1,03)}, HW_{(t,03)})$	-0.117	[±0.04591]
$Corr(HW_{(t-1,04)}, HW_{(t,04)})$	-0.114	[±0.04593]
$Corr(HW_{(t-1,05)}, HW_{(t,05)})$	-0.105	[±0.04597]
$Corr(HW_{(t-1,06)}, HW_{(t,06)})$	-0.077	[±0.04609]
$Corr(HW_{(t-1,07)}, HW_{(t,07)})$	0.023	[±0.04621]
$Corr(HW_{(t-1,08)}, HW_{(t,08)})$	0.134	[±0.04580]
$Corr(HW_{(t-1,09)}, HW_{(t,09)})$	0.154	[±0.04567]
$Corr(HW_{(t-1,10)}, HW_{(t,10)})$	0.182	[±0.04545]
$Corr(HW_{(t-1,11)}, HW_{(t,11)})$	0.203	[±0.04526]
$Corr(HW_{(t-1,12)}, HW_{(t,12)})$	0.137	[±0.04579]
$Corr(HW_{(t-1,13)}, HW_{(t,13)})$	0.208	[±0.04522]
$Corr(HW_{(t-1,14)}, HW_{(t,14)})$	0.170	[±0.04555]
$Corr(HW_{(t-1,15)}, HW_{(t,15)})$	0.119	[±0.04589]
$Corr(HW_{(t-1,16)}, HW_{(t,16)})$	0.069	[±0.04611]
$Corr(HW_{(t-1,17)}, HW_{(t,17)})$	0.041	[±0.04619]
$Corr(HW_{(t-1,18)}, HW_{(t,18)})$	0.013	[±0.04622]
$Corr(HW_{(t-1,19)}, HW_{(t,19)})$	0.006	[±0.04622]

When the route is busy with frequent trams and a large passenger demand, the negative correlation is weaker. During the st Kilda road survey we noticed that, when the trams are crowded and reached the capacity, there were many passengers failing to board. Therefore, the tram behind will have a similar number of passengers to board. Hence have a quite similar dwell time.

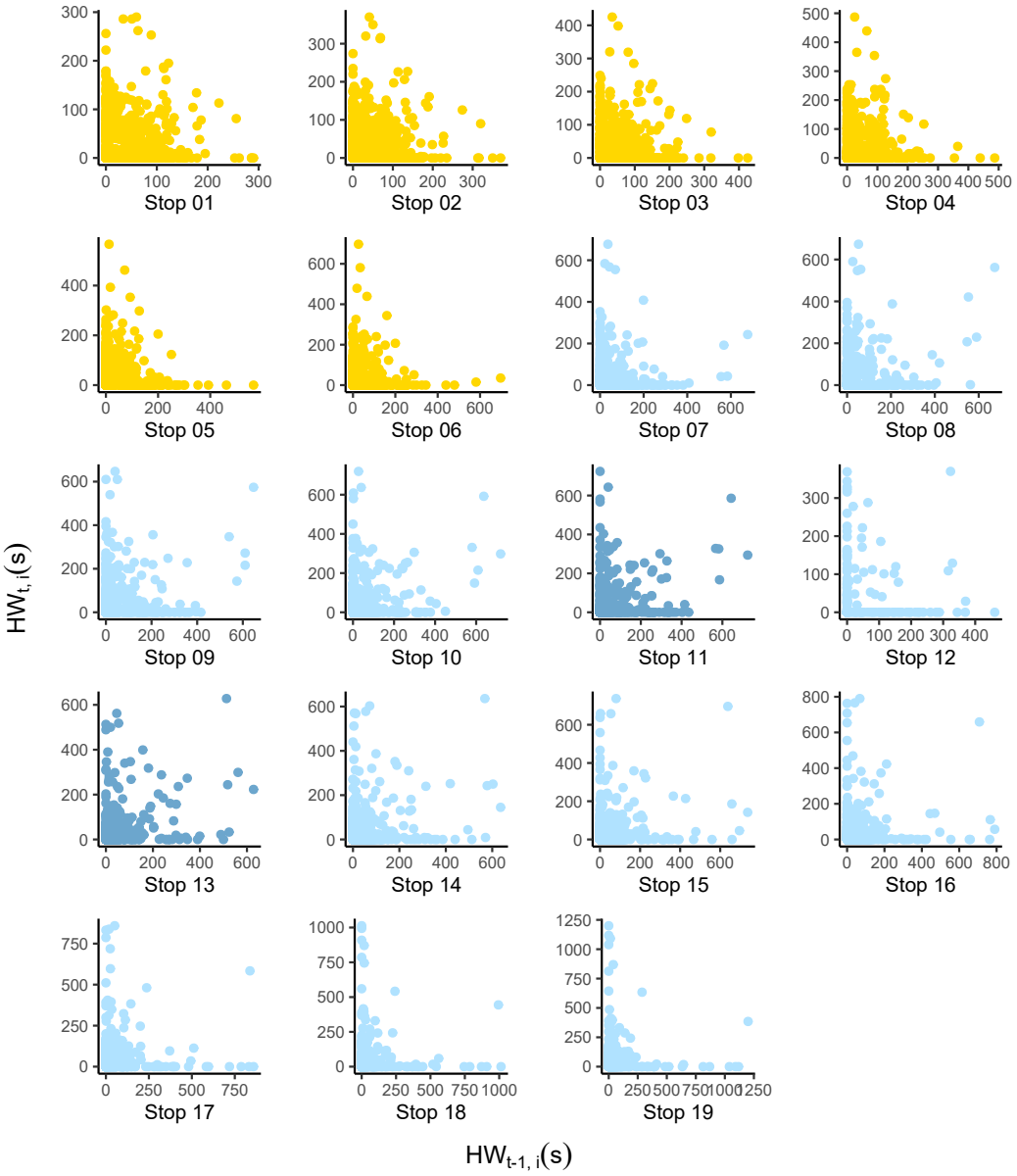


FIGURE 5.24: Simulation: Correlation of time-headways between tram $t-1$ and t. Colour scale: yellow to light blue represents the week negative correlation to week positive correlation.

Figure 5.24 summarises the correlation of time-headways between consecutive trams i.e. correlation of time-headways between tram $t-1$ and t. Table 5.15 summarises the estimated correlation coefficients based on the Pearson's correlation coefficient. As we expect, the negative correlation between time-headways of two consecutive trams is quite evident at the first six stops along the route. After stop 07, the time-headway of the preceding tram is weakly positively correlated to that of the succeeding tram.

5.5.1.3 Dwell time and time-headway correlation

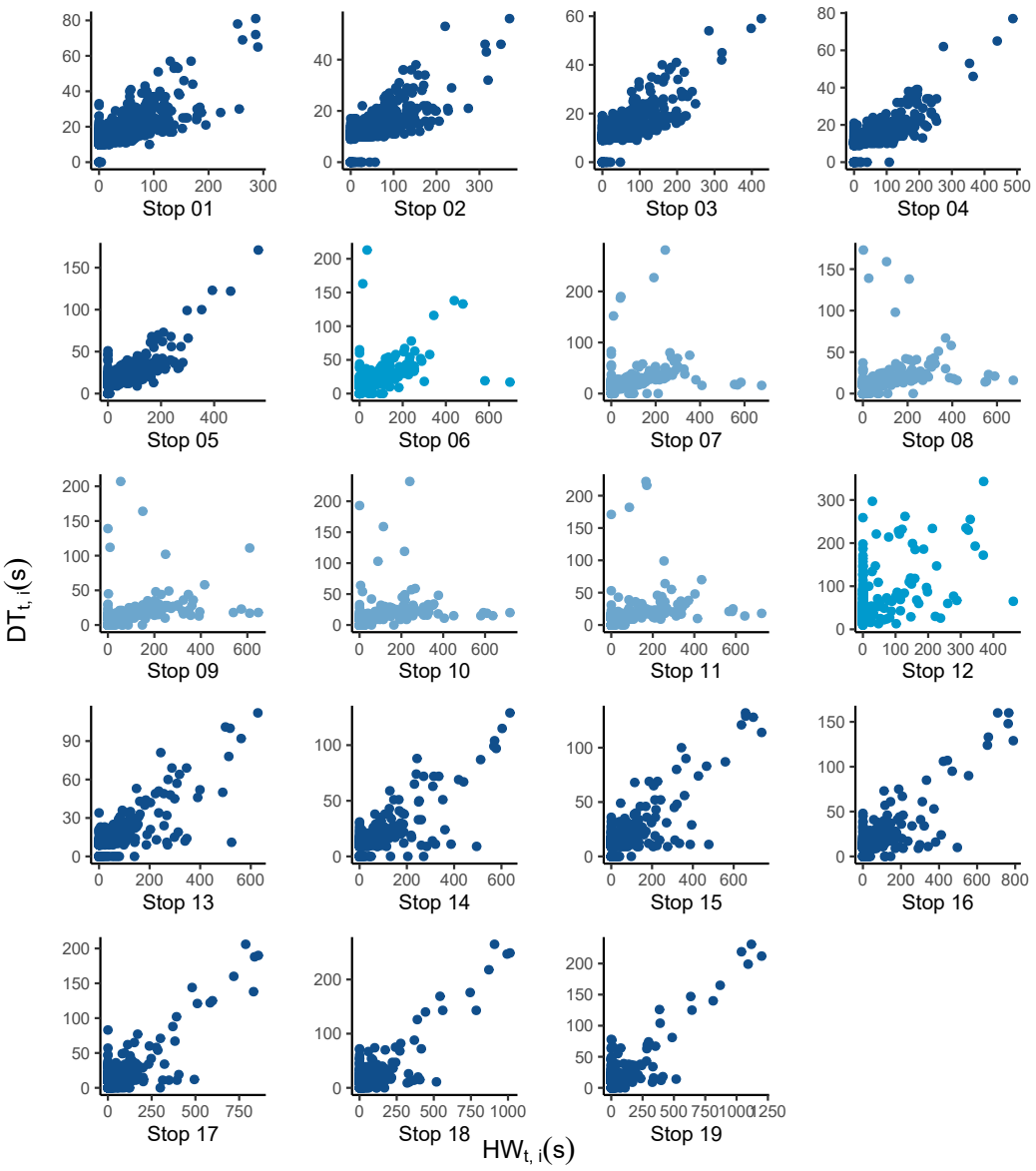


FIGURE 5.25: Simulation: Correlation between dwell time and time-headway of a tram t at each stop. Colour scale: dark blue to light blue represents the high positive correlation to slight positive correlation.

Table 5.16: Simulation: Correlation between the dwell time and time-headway of a tram at each stop. $Corr(HW_{(t,i)}, DT_{(t,i)})$, for $i = 01, 02, ..., 19$. $DT_{(t,i)}$ represents the dwell time of tram t at stop i and $HW_{(t,i)}$ represents the time-headway of tram t at stop i.

	Correlation Coefficient	
$Corr(HW_{(t,01)}, DT_{(t,01)})$	0.752	[±0.03029]
$Corr(HW_{(t,02)}, DT_{(t,02)})$	0.751	[±0.03035]
$Corr(HW_{(t,03)}, DT_{(t,03)})$	0.774	[±0.02913]
$Corr(HW_{(t,04)}, DT_{(t,04)})$	0.807	[±0.02716]
$Corr(HW_{(t,05)}, DT_{(t,05)})$	0.814	[±0.02670]
$Corr(HW_{(t,06)}, DT_{(t,06)})$	0.519	[±0.03931]
$Corr(HW_{(t,07)}, DT_{(t,07)})$	0.348	[±0.04311]
$Corr(HW_{(t,08)}, DT_{(t,08)})$	0.348	[±0.04311]
$Corr(HW_{(t,09)}, DT_{(t,09)})$	0.369	[±0.04272]
$Corr(HW_{(t,10)}, DT_{(t,10)})$	0.311	[±0.04370]
$Corr(HW_{(t,11)}, DT_{(t,11)})$	0.322	[±0.04353]
$Corr(HW_{(t,12)}, DT_{(t,12)})$	0.529	[±0.03903]
$Corr(HW_{(t,13)}, DT_{(t,13)})$	0.791	[±0.02815]
$Corr(HW_{(t,14)}, DT_{(t,14)})$	0.784	[±0.02856]
$Corr(HW_{(t,15)}, DT_{(t,15)})$	0.813	[±0.02678]
$Corr(HW_{(t,16)}, DT_{(t,16)})$	0.817	[±0.02649]
$Corr(HW_{(t,17)}, DT_{(t,17)})$	0.819	[±0.02637]
$Corr(HW_{(t,18)}, DT_{(t,18)})$	0.833	[±0.02542]
$Corr(HW_{(t,19)}, DT_{(t,19)})$	0.836	[±0.02525]

Figure 5.25 summarises the correlation between dwell time and time-headway of a tram for all the stops along the route i.e. $Corr(HW_{(t,i)}, DT_{(t,i)})$, $DT_{(t,i)}$ represents the dwell time of tram t at stop i and $HW_{(t,i)}$ represents the time-headway of tram t at stop i. Time-headway and dwell time are positively correlated at each stop. At all the stops up to 12, dwell time and time-headway are positively correlated, as a longer time-headway between trams results in more passengers waiting at the stop expecting a tram. At stop 12, the correlation between dwell time and time-headway is weaker. This is probably because Stop 12 only has a fewer passengers getting on while an exceptionally high number of passengers getting off. Therefore, the dwell time at stop 12 is largely due to the alighting passengers. The number of passengers getting off at the stop is not largely affected by the time-headway of a tram during the peak hour.

After stop 13, the strength of the correlation between dwell time and time-headway gradually decreases. This is because lower demand of passengers getting on and off were considered after stop 12. Most of the trams spend less dwell time at these stops due to the lower passenger demand. Most of the trams have shorter time-headways with a fewer passengers to serve. Therefore, the correlation between dwell time and time-headway becomes weaker except for a few numbers of cases with longer time-headways.

5.6 Discussion

To facilitate simulation-based studies, understanding of the stopping behaviour of mass transit vehicles is vital. The stopping behaviour is greatly affected by the time spent at each stop for serving passengers. The passenger demand at a particular stop might also be affected by the time-headway of the transit vehicle. Therefore, it is crucial to fully understand the stopping behaviour of mass transit vehicles by just looking at isolated transit vehicles and stops. Therefore, in this chapter, we study the correlation structure of the passenger flow time and time-headway between consecutive trams and consecutive stops.

We analysed a data set in [35] which contains passenger flow time events from ten different tram stops in the inner network of Melbourne CBD. These tram stops include all three types of tram stop designs: curb-side, platform and safety zone. The already existing data set is not detailed enough to understand the correlation structure between stops and between trams. Therefore, we surveyed two distinct pairs of consecutive tram stops in Melbourne's CBD, located on two different tram routes, one of high frequency and the other of low frequency. The results show that the marginal distribution of passenger flow time can be well approximated by a log-normal distribution with a reasonably universal σ. However, the μ appears to depend on the location of the stop. The marginal distribution of time-headway is symmetric when the route is less busy with a fewer number of trams. However, the marginal distribution is skewed and can be well approximated by gamma distribution when the route is busy with a high frequency of trams.

If a tram dwells for long at a stop, serving many passengers, then the next tram at that stop will have fewer passengers to pick up, implying a shorter dwell time. This would suggest a negative correlation between dwell time of consecutive trams at a given stop. Correspondingly, understanding of the marginal distribution of dwell time would be insufficient to simulate the movements of trams accurately. Therefore, understanding the conditional distribution of a tram's dwell time, conditional on the dwell time of the previous tram is also necessary.

The survey data we collected at two distinct pairs of consecutive tram stops and the data obtained from the simulation study were analysed to understand the correlation structure of passenger flow time/dwell time and time-headway between consecutive trams and between consecutive stops. The passenger flow time between consecutive stops appeared to be a weak positive correlation at both Sydney Road and St Kilda Road. In the simulation, the correlation of dwell time between stops is quite notable. The negative correlation of passenger

flow time between consecutive stops is quite evident at Sydney road while it is weaker at St Kilda Road. The negative correlation of dwell time between consecutive stops is also weaker in the simulation. The time-headway of a tram at two consecutive stops is positively correlated when the scheduled time table is maintained. The time-headway of two consecutive trams at any given stop is negatively correlated and is quite evident at Sydney road. However, the negative correlation is weaker at the St Kilda Road stops and at the simulated stops.

Even though we observed the correlation of dwell time between consecutive stops to be positive and between consecutive trams to be negative, the strength of these correlations is rather weak. The correlation of time-headway between two consecutive trams is also weak. Since the correlation structure appeared to be weaker than we expected empirically, modelling each tram's dwell time at each stop as an independent log-normal distribution should provide an accurate method for generating random dwell times at least in simulation-based studies similar to Melbourne CBD. Furthermore, the time-headway of a tram at any given stop can also be modelled using a gamma distribution when the route is busy with a high frequency of trams and with irregular dispatching of time-headways. We can model the time-headway of a tram at any given stop when the route is less busy with a fewer number of trams and with a regular dispatching time-headway using a normal distribution. This provides a clear initiative to practical implication of simulation studies on mass transit vehicles.

Future work includes studying the correlation structure of dwell time and time-headway between consecutive vehicles and consecutive stops along a route, which would provide more precise results. The simulation model can also be extended to model multiple routes rather than assuming a single route with irregular dispatching time-headways. Furthermore, the model can be extended to study the tram bunching phenomenon precisely.

Discussion 6

We conclude this thesis with a brief discussion of some theoretical and practical consequences of our work, and also outline some interesting possible avenues for future research.

Our study of ABC methods should be considered a proof of concept. By considering calibration of a simple stochastic model, ASEP, for which many exact results are known, we were able to identify appropriate summary statistics, and compare the effectiveness of simple rejection ABC vs regression adjusted ABC. This is only the tip of the ice-berg however. There are at least two directions in which this could be extended. Firstly, based on the success of the small examples studied here, it would be practically useful to apply regression-adjusted ABC to larger models. Secondly, there are many other variations and refinements of the ABC methodology, such as SMC-ABC, RABC, MCMC-ABC, and marginally adjusted ABC, which are potentially more computationally efficient than the regression-adjusted ABC implementation studied here. It would be of interest to again study such extensions in the setting of ASEP and NaSch.

Our study of travel times suggests extensions both on the very theoretical and very practical ends of the spectrum. On the theoretical end, we provide concrete conjectures for the form of a posited central limit theorem for the travel time of a parallel ASEP with open boundary conditions, with explicit expressions (in terms of the system size, and input and output probabilities) for the constants in the standardisation. Proving such results rigorously would be an interesting contribution to the field of non-equilibrium statistical mechanics. Moreover, a number of analogous variations of ASEP are also likely

to have central limit theorems for the travel time, e.g. random sequential instead of parallel updates, or periodic instead of open boundaries. Perhaps more ambitiously, it may be possible to prove the existence of a central limit theorem for the travel time of NaSch, albeit without explicit expressions for the standardisation constants.

On the practical end of the spectrum, our simulations of NaSch systems with traffic signals provides strong evidence that the travel time is generically given by a Gaussian mixture model. Moreover, in low density, our physical predictions for a number of the parameters characterising the mixture appear to agree well with the numerical results. However, it would be of great practical interest to perform a detailed empirical study of travel time in a real arterial road network. Such a study would shed further light on the observed dependencies of the travel time on the signal parameters, such as splits and offsets. It would also allow an empirical test of our physical predictions of the mixture parameters. Having a robust understanding of how to accurately model travel times in an arterial network using a Gaussian mixture model with physically interpretable parameters has the potential to be a very useful tool for traffic simulation.

Finally, perhaps the most practical outcome of the current work concerns the distribution of dwell times for trams. Based on both our empirical survey, and simulations, it seems justified to conclude that the dwell time of trams in a traffic simulator can safely be modelled as independent log-normal random variables at least in tram networks operated with priority processes similar to Melbourne CBD. The nature of the correlations between the dwell time of consecutive trams/stops has not previously been addressed. While the weakness of these correlations was perhaps somewhat unexpected, it is a very welcome observation from the perspective of accurately modelling dwell times in simulations. Moreover, the parameters in these log-normal distributions depend only weakly on the specifics of the network.

Appendix

Travel time distribution: signalised links

Lower inflow: $\alpha = 0.1$ and $\beta = 0.9$

Case 45s-45s

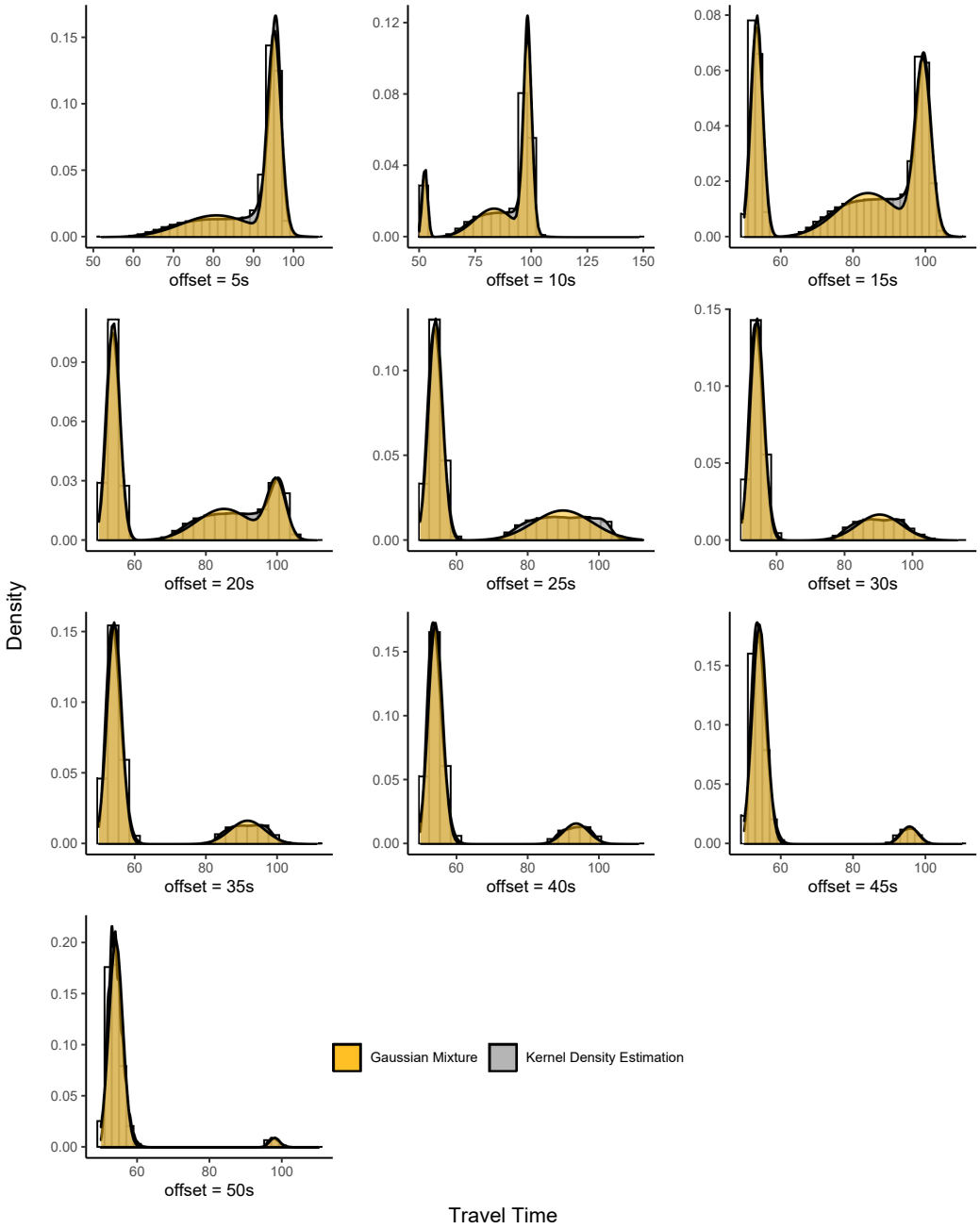


FIGURE 1: Travel time distribution: Case 45s-45s and $\alpha = 0.1$. Fitted Gaussian mixture model is shown in gold area.

Case 45s-60s

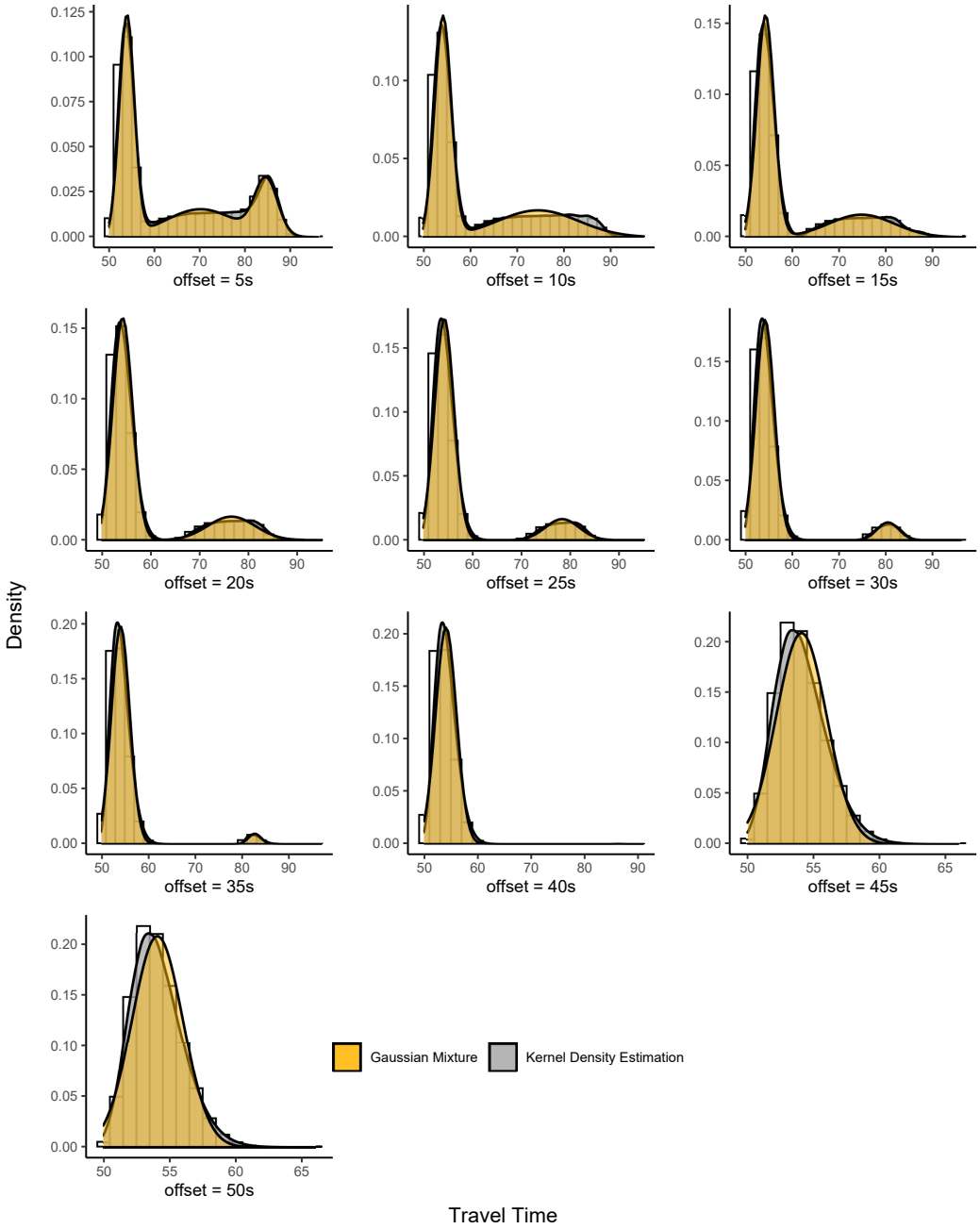


FIGURE 2: Travel time distribution: Case 45s-60s and $\alpha = 0.1$. Fitted Gaussian mixture model is shown in gold area.

Case 60s-45s

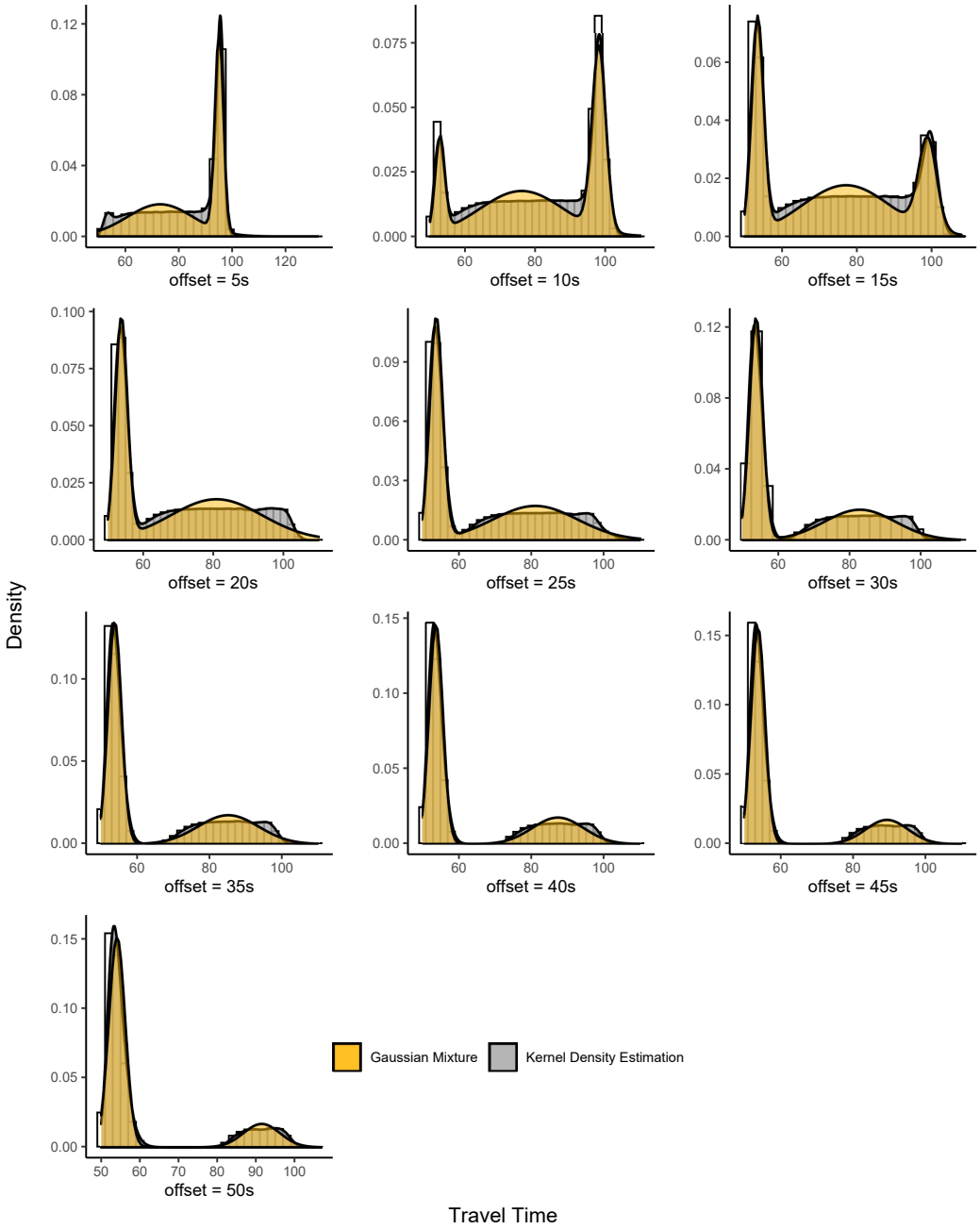


FIGURE 3: Travel time distribution: Case 60s-45s and $\alpha = 0.1$. Fitted Gaussian mixture model is shown in gold area.

Case 60s-60s

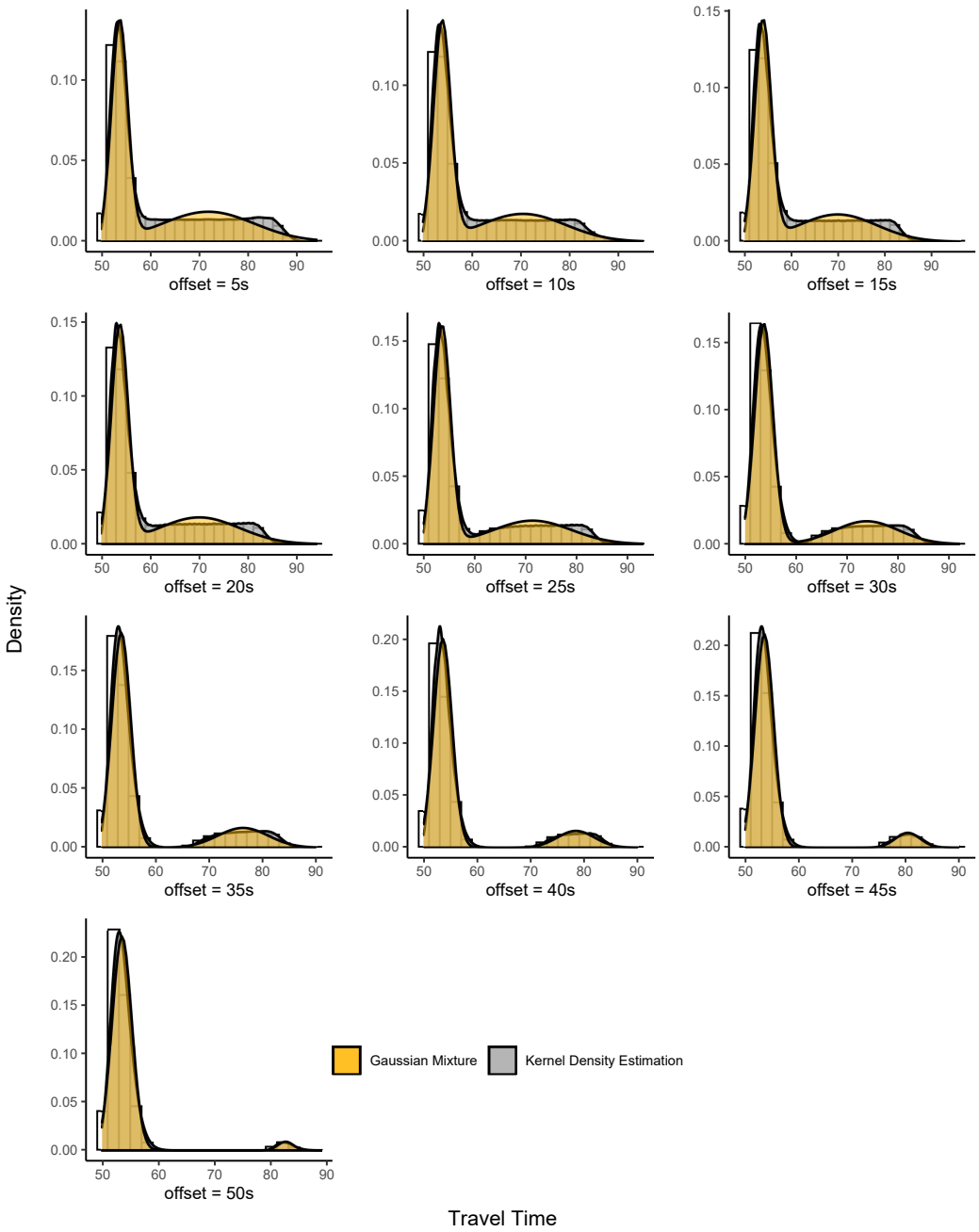


FIGURE 4: Travel time distribution: Case 60s-60s and $\alpha = 0.1$. Fitted Gaussian mixture model is shown in gold area.

Higher inflow, $\alpha = 0.5$

Case 45s-45s

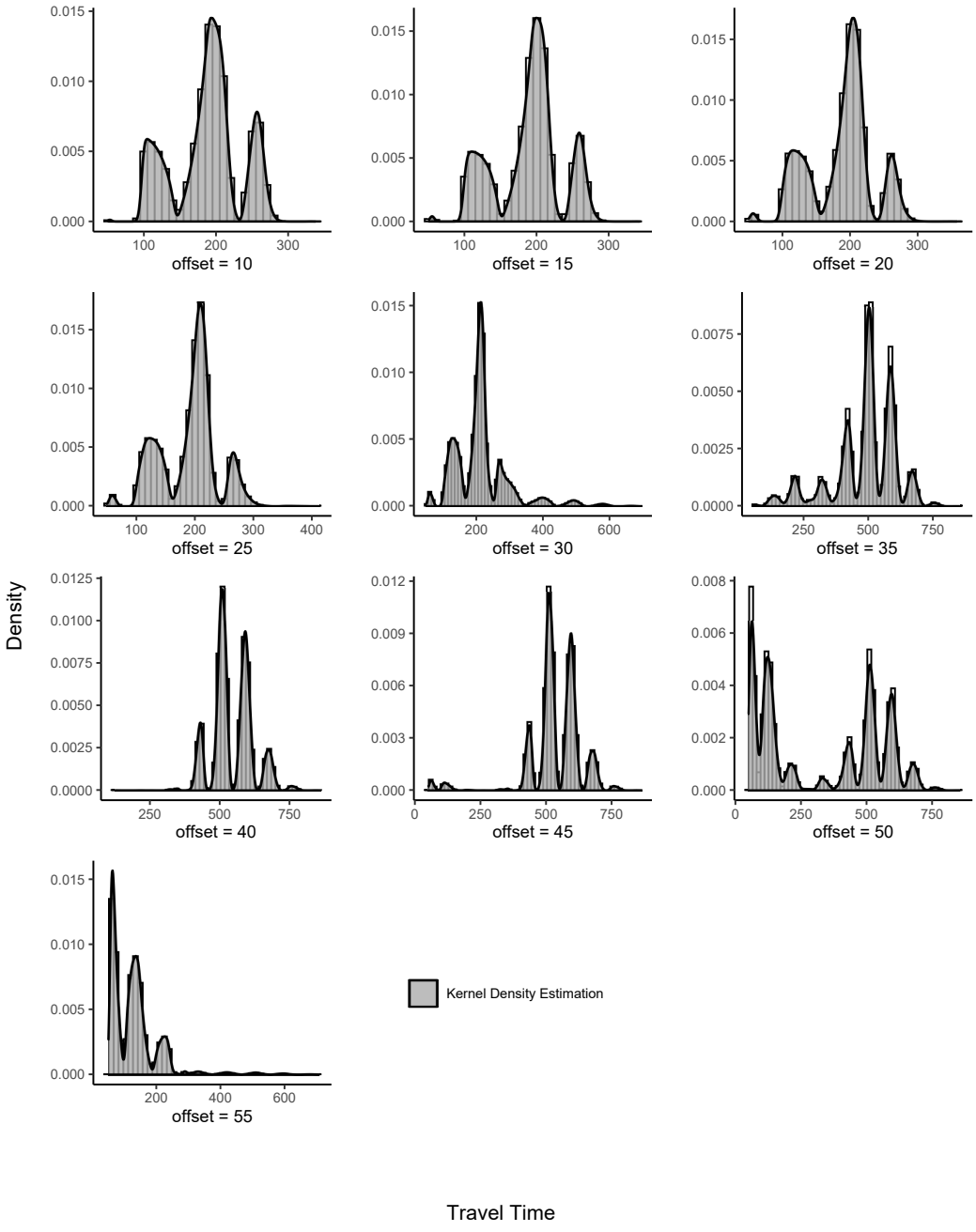


FIGURE 5: Travel time distribution: Case 45s-45s and $\alpha = 0.5$

Case 45s-60s

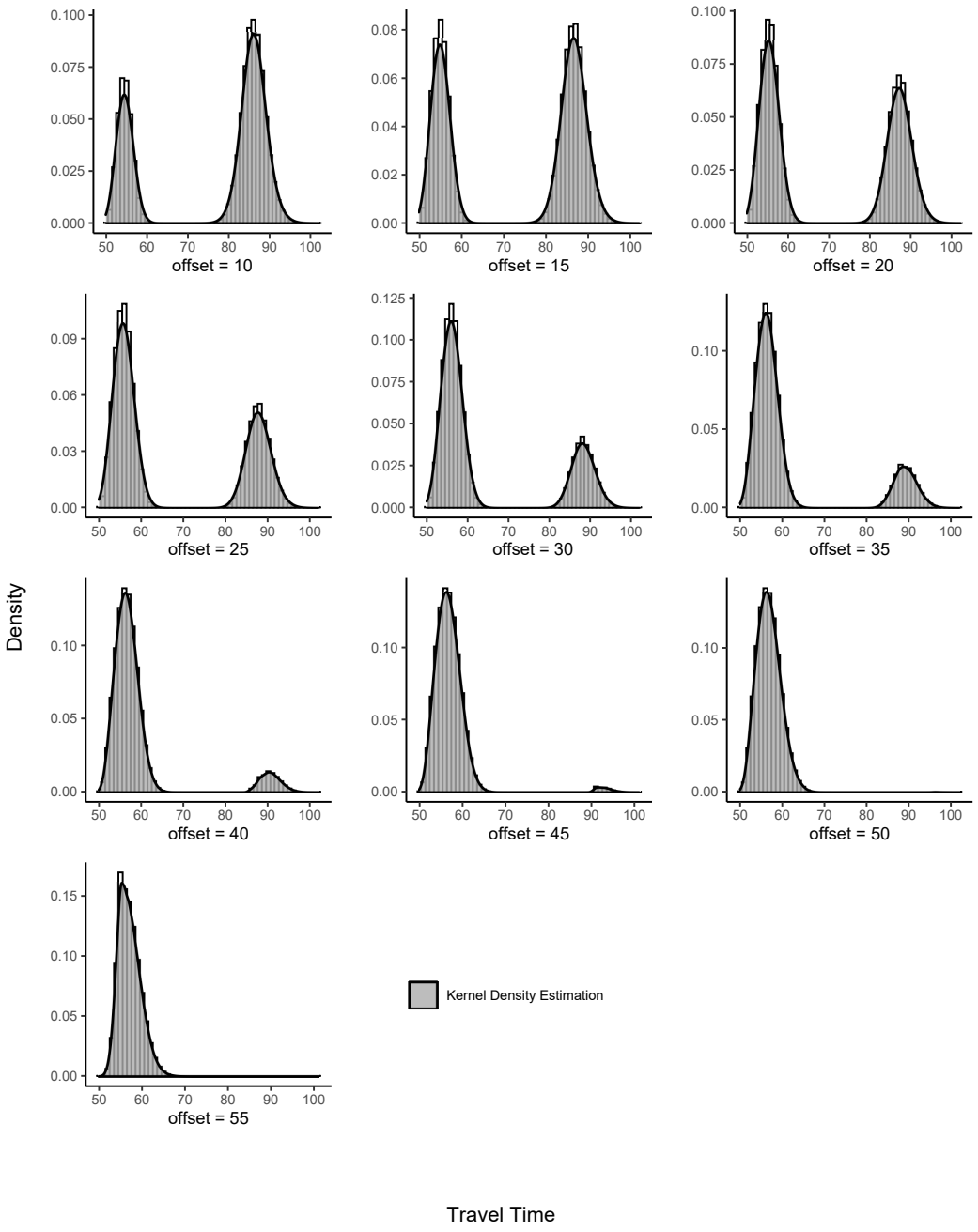


FIGURE 6: Travel time distribution: Case 45s-60s and $\alpha = 0.5$

Case 60s-45s

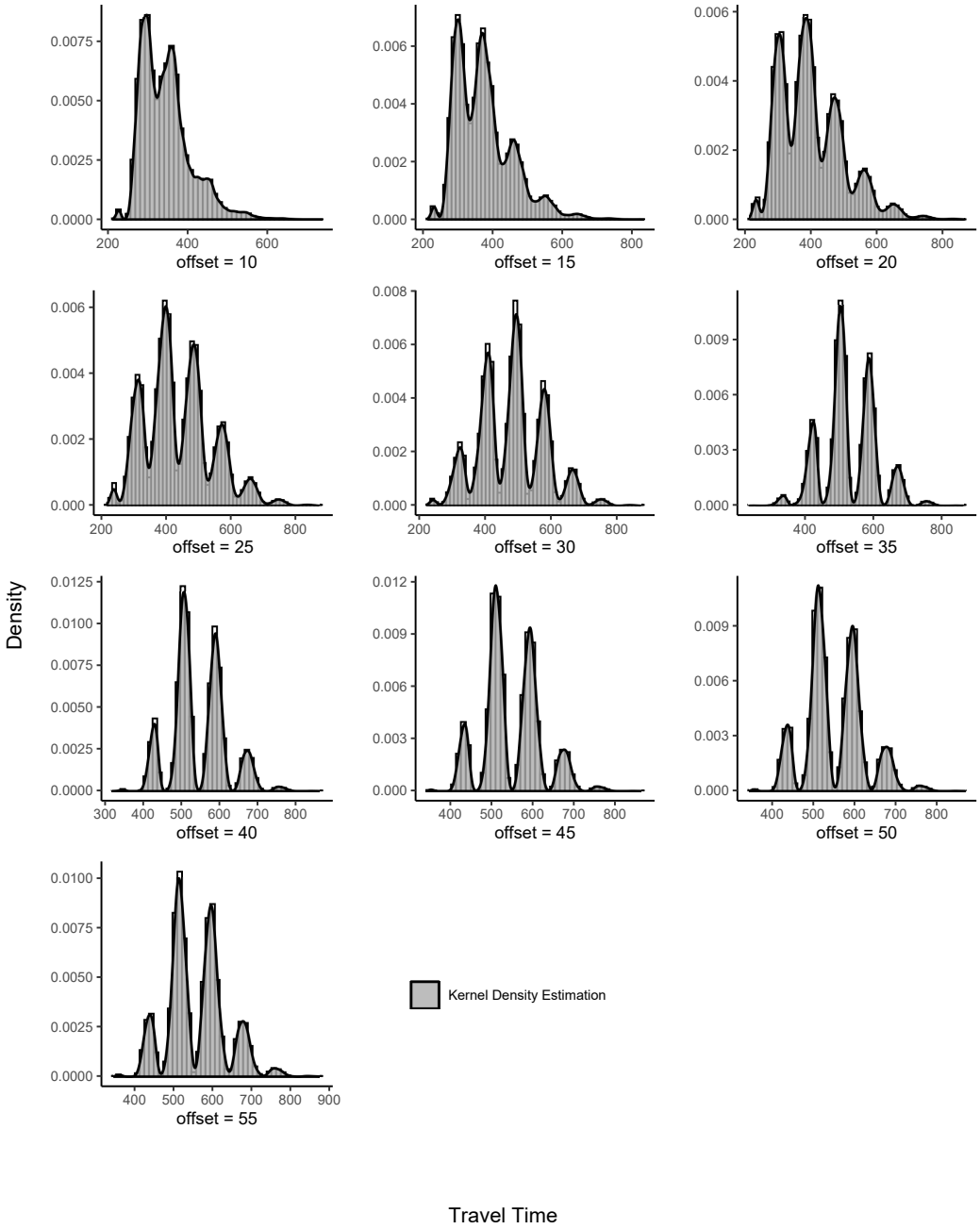


FIGURE 7: Travel time distribution: Case 60s-45s and $\alpha = 0.5$

Case 60s-60s

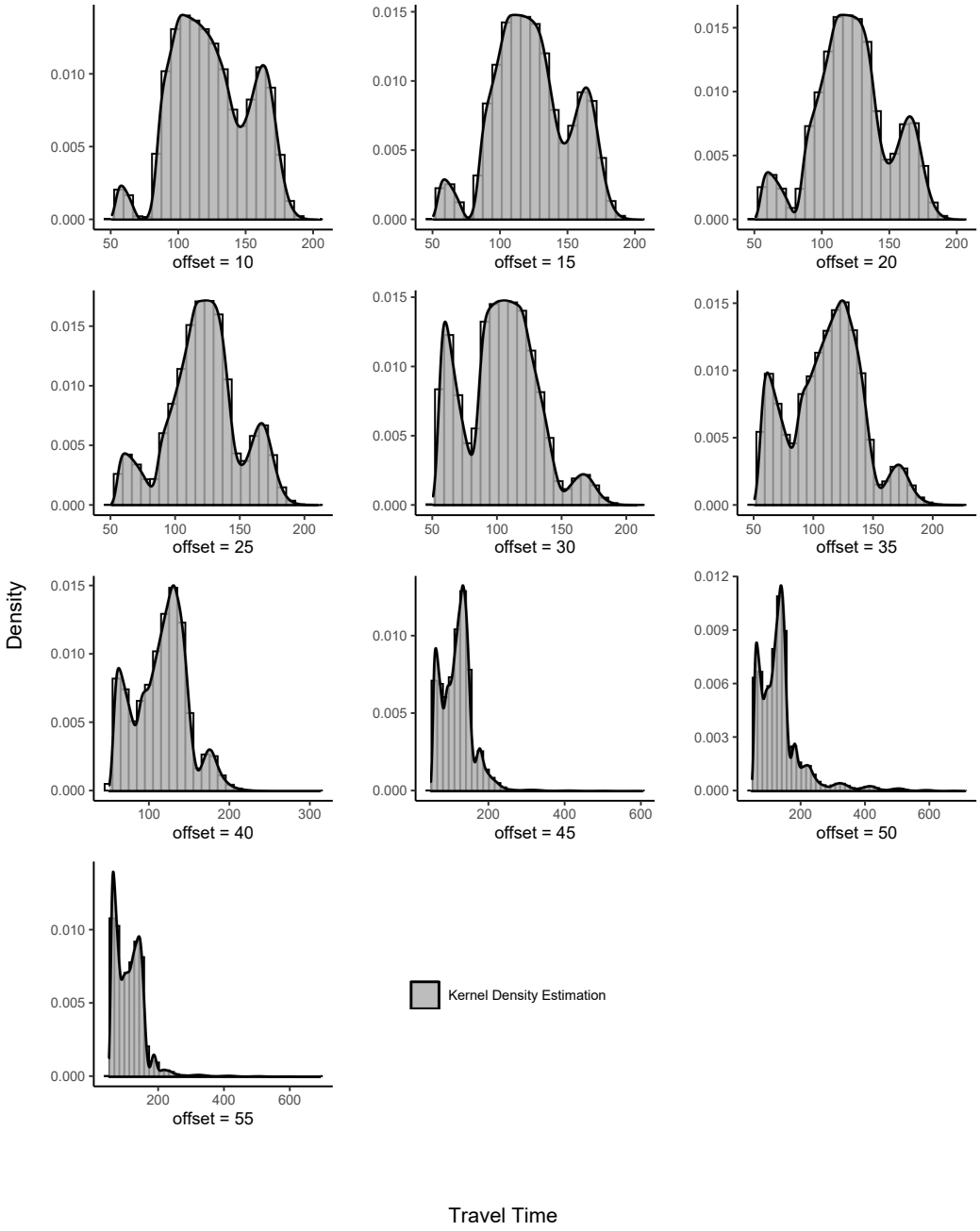


FIGURE 8: Travel time distribution: Case 60s-60s and $\alpha = 0.5$

Bibliography

[1] Michael D Fontaine et al. “Modeling and Simulation : Real - World Examples”. In: *Principles of Modeling and Simulation: A Multidisciplinary Approach*. Ed. by John A. Sokolowski and Catherine M. Banks. A JOHN WILEY & SONS, INC., 2009. Chap. 8.

[2] Jan De Gier, Timothy M Garoni, and Omar Rojas. “Traffic flow on realistic road networks with adaptive traffic lights”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2011.04 (2011), P04008.

[3] Lei Rao and Larry Owen. “Validation of High-Fidelity Traffic Simulation Models”. In: *Transportation Research Record* 1710.1 (2000), pp. 69–78.

[4] LR Rilett, KO Kim, and Bryan Raney. “Comparison of Low-Fidelity TRANSIMS and High-Fidelity CORSIM Highway Simulation Models with Intelligent Transportation System Data”. In: *Transportation Research Record: Journal of the Transportation Research Board* 1739.00 (2007).

[5] Baher Abdulhai, Himanshu Porwal, and Will Recker. “Short term freeway traffic flow prediction using genetically-optimized time-delay-based neural networks”. 1999.

[6] Tomer Toledo and Haris Koutsopoulos. “Statistical Validation of Traffic Simulation Models”. In: *Transportation Research Record* 1876.1 (2004), pp. 142–150.

[7] Winnie Daamen, Christine Buisson, and Serge P. Hoogendoorn, eds. *Traffic Simulation and Data Validation Methods and Applications*. 1st ed. Boca Raton: CRC Press, 2014.

[8] R. et al. JAYAKRISHNAN. "Calibration and path dynamics issues in microscopic simulation for advanced traffic management and information systems". In: *Transport Research Record* 1771.01 (2001), pp. 1–18.

[9] German Molina, M. J. Bayarri, and James O. Berger. "Statistical inverse analysis for a network microsimulator". In: *Technometrics* 47.4 (2005), pp. 388–398.

[10] Chao Zhang, Carolina Osorio, and Gunnar Flötteröd. "Efficient calibration techniques for large-scale traffic simulators". In: *Transportation Research Part B: Methodological* 97 (2017), pp. 214–239.

[11] Peter Wagner, Christine Buisson, and Ronald Nippold. "Challenges in Applying Calibration Methods to Stochastic Traffic Models". In: *Transportation Research Record: Journal of the Transportation Research Board* 2560 (2016), pp. 10–16.

[12] Xiayan Zhang and John A. Rice. "Short-term travel time prediction". In: *Transportation Research Part C: Emerging Technologies* 11.3 (2013), pp. 187–210.

[13] Chun-Hsin Wu, Jan-Ming Ho, and D. T. Lee. "Travel-time prediction with support vector regression". In: *IEEE Transactions on Intelligent Transportation Systems* 5.4 (2004), pp. 276–281.

[14] Lianyu Chu, Jun-Seok Oh, and Will Recker. "Adaptive Kalman filter based freeway travel time estimation". In: (Jan. 2005).

[15] F.W. Cathey and D.J. Dailey. "A prescription for transit arrival/departure prediction using automatic vehicle location data". In: *Transportation Research Part C: Emerging Technologies* 11.3 (2003), pp. 241–264.

[16] Guiyan Jiang and Ruoqi Zhang. “Travel time prediction for urban arterial road”. In: *Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems*. Vol. 2. 2003, 1459–1462 vol.2.

[17] C. D. Mark, A. W. Sadek, and D. Rizzo. “Predicting experienced travel time with neural networks: a PARAMICS simulation study”. In: *Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No.04TH8749)*. 2004, pp. 906–911.

[18] Dongjoo Park, R. Rilett Laurence, and Gunhee Han. “Spectral Basis Neural Networks for Real-Time Travel Time Forecasting”. In: *Journal of Transportation Engineering* 125.6 (1999), pp. 515–523.

[19] Y. Xue et al. “Empirical characteristics of transit travel time distribution for commuting routes”. In: *Transportation Research Board 90th Annual Meeting* (2011).

[20] Le-Minh Kieu, Ashish Bhaskar, and Edward Chung. “Public transport travel-time variability definitions and monitoring”. In: *Journal of Transportation Engineering* (2014).

[21] Md.Matiur Rahman, S.C. Wirasinghe, and Lina Kattan. “Analysis of bus travel time distributions for varying horizons and real-time applications”. In: *Transportation Research Part C: Emerging Technologies* (2018).

[22] Susilawati, Michael A. P. Taylor, and Sekhar V. C. Somenahalli. “Distributions of travel time variability on urban roads”. In: *Journal of Advanced Transportation* 47.8 (2013), pp. 720–736.

[23] M. A. P. TAYLOR. “Travel Time Variability—The Case of Two Public Modes”. In: *Transportation Science* 16.4 (1982), pp. 507–521.

[24] Graham Currie, Alexa Delbosc, and James Reynolds. “Modeling Dwell Time for Streetcars in Melbourne, Australia, and Toronto, Canada”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2275 (2012), pp. 22–29.

[25] Rajat Rajbhandari, Steven Chien, and Janice Daniel. “Estimation of Bus Dwell Times with Automatic Passenger Counter Information”. In: *Transportation Research Record* 1841.03 (2003), pp. 120–127.

[26] Hooi Ling Khoo. “Statistical Modeling of Bus Dwell Time at Stops”. In: *Journal of the Eastern Asia Society for Transportation Studies* 10 (2013), pp. 1489–1500.

[27] Stephen A Arhin and Errol C Noel. “Predicting Dwell Time by Bus Stop Type and Time of the Day”. In: *Journal of Civil & Environmental Engineering* 05.05 (2015).

[28] Martin Milkovits. “Modeling the Factors Affecting Bus Stop Dwell Time: Use of Automatic Passenger Counting, Automatic Fare Counting, and Automatic Vehicle Location Data”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2072.2072 (2008), pp. 125–130.

[29] Travis B Glick and Miguel A Figliozzi. “Measuring the Determinants of Bus Dwell Time: New Insights and Potential Biases”. In: July 2016 (2017), pp. 1–16.

[30] Grant Fletcher and Ahmed El-Geneidy. “Effects of Fare Payment Types and Crowding on Dwell Time”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2351.2351 (2013), pp. 124–132.

[31] Fazhi Li, Zhengyu Duan, and Dongyuan Yang. “Dwell time estimation models for bus rapid transit stations”. In: *Journal of Modern Transportation* 20.3 (2012), pp. 168–177.

[32] Emilio González, Manuel Romana, and Oscar Álvaro. “Bus Dwell-Time Model of Main Urban Route Stops”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2274.2274 (2012), pp. 126–134.

[33] Jianxia Xin and Shuyan Chen. “Bus Dwell Time Prediction Based on KNN”. In: *Procedia Engineering* 137 (2016), pp. 283–288.

[34] Cen Zhang and Jing Teng. “Bus Dwell Time Estimation and Prediction: A Study Case in Shanghai-China”. In: *Procedia - Social and Behavioral Sciences* 96.Cictp (2013), pp. 1329–1340.

[35] Graham Currie et al. “Impact of Crowding on Streetcar Dwell Time”. In: (2013).

[36] Graham Currie and James Reynolds. “Evaluating pay-on-entry versus proof-of-payment ticketing in light rail transit”. In: *Transportation Research Record* 2540.1 (2016), pp. 39–45.

[37] Zoi Christoforou, Ektoras Chandakas, and Ioannis Kaparias. “Investigating the Impact of Dwell Time on the Reliability of Urban Light Rail Operations”. In: *Urban Rail Transit* (2020), pp. 1–16.

[38] Soroush Rashidi and Prakash Ranjitkar. “Estimation of bu dwell time using univariate time series models”. In: *Journal of Advanced Transportation* 49 (2015), pp. 139–152.

[39] Soroush Rashidi, Prakash Ranjitkar, and Yuval Hadas. “Modeling Bus Dwell Time with Decision Tree-Based Methods”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2418.1 (2014), pp. 74–83.

[40] Min-Tang Li et al. “Simulation Model for Estimating Bus Dwell Time by Simultaneously Considering Numbers of Disembarking and Boarding Passengers”. In: *Transportation Research Record: Journal of the Transportation Research Board* 1971.1 (2006), pp. 59–65.

[41] Soroush Rashidi and Prakash Ranjitkar. “Estimation of bus dwell time using univariate time series models”. In: *Journal of the Eastern Asia Society for Transportation Studies* 10 (2013), pp. 1281–1291.

[42] Serge Hoogendoorn and Raymond Hoogendoorn. “Generic Calibration Framework for Joint Estimation of Car-Following Models by Using Microscopic Data”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2188 (2010), pp. 37–45.

[43] Andreas Schadschneider, Debashish Chowdhury, and Katsuhiro Nishinari. *STOCHASTIC TRANSPORT IN COMPLEX SYSTEMS FROM MOLECULES TO VEHICLES*. Elsevier, 2011.

[44] Denos C Gazis. “Mathematical theory of automobile traffic”. In: *Science* 157.3786 (1967), pp. 273–281.

[45] Peter G Gipps. “A behavioural car-following model for computer simulation”. In: *Transportation Research Part B: Methodological* 15.2 (1981), pp. 105–111.

[46] Shinya Kikuchi and Partha Chakroborty. “Car-following model based on fuzzy inference system”. In: *Transportation Research Record* (1992), pp. 82–82.

[47] Serge P Hoogendoorn and Piet HL Bovy. “Generic gas-kinetic traffic systems modeling with applications to vehicular traffic flow”. In: *Transportation Research Part B: Methodological* 35.4 (2001), pp. 317–336.

[48] Dirk Helbing and Martin Treiber. “Gas-kinetic-based traffic model explaining observed hysteretic phase transition”. In: *Physical Review Letters* 81.14 (1998), p. 3042.

[49] Dirk Helbing. “Improved fluid-dynamic model for vehicular traffic”. In: *Physical Review E* 51.4 (1995), p. 3164.

[50] Gabriella Bretti, Roberto Natalini, and Benedetto Piccoli. “A fluid-dynamic traffic model on road networks”. In: *Archives of Computational Methods in Engineering* 14.2 (2007), pp. 139–172.

[51] Serge P Hoogendoorn and Piet HL Bovy. “State-of-the-art of vehicular traffic flow modelling”. In: *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 215.4 (2001), pp. 283–303.

[52] Satoshi Yukawa and Macoto Kikuchi. “Coupled-map modeling of one-dimensional traffic flow”. In: *Journal of the Physical Society of Japan* 64.1 (1995), pp. 35–38.

[53] Kunihiko Kaneko. “Theory and applications of coupled map lattices”. In: *Nonlinear science: theory and applications* (1993).

[54] K. Nagel and M. Schreckenberg. “A CELLULAR AUTOMATON MODEL FOR FREEWAY TRAFFIC”. In: *JOURNAL DE PHYSIQUE I* 2 (12 1992), pp. 2221–2229.

[55] Nino Boccara. “Cellular Automata”. In: *Modeling Complex Systems* (2004), pp. 191–273.

[56] Bastien Chopard and Alexandre Masselot. “Cellular automata and lattice Boltzmann methods: a new approach to computational fluid dynamics and particle transport”. In: *Future Generation Computer Systems* 16.2-3 (1999), pp. 249–257.

[57] Boris S Kerner, Sergey L Klenov, and Dietrich E Wolf. “Cellular automata approach to three-phase traffic theory”. In: *Journal of Physics A: Mathematical and General* 35.47 (2002), p. 9971.

[58] Jan de Gier and Bernard Nienhuis. “Exact stationary state for an asymmetric exclusion process with fully parallel dynamics”. In: *Physical Review E* 59.5 (1999), p. 4899.

[59] Nikolaus Rajewsky et al. “The asymmetric exclusion process: Comparison of update procedures”. In: *Journal of statistical physics* 92.1-2 (1998), pp. 151–194.

[60] Bernard Derrida, Eytan Domany, and David Mukamel. “An exact solution of a one-dimensional asymmetric exclusion model with open boundaries”. In: *Journal of Statistical Physics* 69.3-4 (1992), pp. 667–687.

[61] Martin R Evans, Nikolaus Rajewsky, and Eugene R Speer. “Exact solution of a cellular automaton for traffic”. In: *Journal of statistical physics* 95.1-2 (1999), pp. 45–96.

[62] Lele Zhang, Timothy M. Garoni, and Jan de Gier. “A comparative study of Macroscopic Fundamental Diagrams of arterial road networks governed by adaptive traffic signal systems”. In: *Transportation Research Part B: Methodological* 49 (2013), pp. 1–23.

[63] Lele Zhang, Jan De Gier, and Timothy M. Garoni. “Traffic disruption and recovery in road networks”. In: *Physica A: Statistical Mechanics and its Applications* 401 (2014), pp. 82–102.

[64] Tomer Toledo et al. “Calibration and Validation of Microscopic Traffic Simulation Tools: Stockholm Case Study”. In: *Transportation Research Record: Journal of the Transportation Research Board* 1831.03 (2003), pp. 65–75.

[65] Seung-Jun Kim, Wonho Kim, and Larry R Rilett. “Calibration of microsimulation models using nonparametric statistical techniques”. In: *Transportation Research Record* 1935.1935 (2005), pp. 111–119.

[66] Ramachandran Balakrishna et al. “Calibration of Microscopic Traffic Simulation Models: Methods and Application”. In: *Transportation Research Record: Journal of the Transportation Research Board* 1999 (2007), pp. 198–207.

[67] Vincenzo Punzo and Fulvio Simonelli. “Analysis and comparison of microscopic traffic flow models with real traffic microscopic data”. In:

Transportation Research Record: Journal of the Transportation Research Board 1934.1 (2005), pp. 53–63.

[68] Byungkyu Park and J. Schneeberger. "Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System". In: *Transportation Research Record: Journal of the Transportation Research Board* 1856.03 (2003), pp. 185–192.

[69] Byungkyu Park and Hongtu Qi. "Development and Evaluation of a Procedure for the Calibration of Simulation Models". In: *Transportation Research Record: Journal of the Transportation Research Board* 1934 (2005), pp. 208–217.

[70] X Qin and H Mahmassani S. "Adaptive Calibration of Dynamic Speed-Density Relations for Online Network Traffic Estimation and Prediction Applications". In: 1876 (2004), p. 82–89.

[71] L. et.al CHU. "A Calibration Procedure for Microscopic Traffic Simulation". In: *Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems* 2 (2003), pp. 1574–1579.

[72] Lei Rao, L. Owen, and D. Goldsman. "Development and application of a validation framework for traffic simulation models". In: *1998 Winter Simulation Conference. Proceedings.* vol.2. Washington, DC, 1998, pp. 1079–1086.

[73] Alexis Boukouvalas et al. "Bayesian precalibration of a large stochastic microsimulation model". In: *IEEE Transactions on Intelligent Transportation Systems* 15.3 (2014), pp. 1337–1347.

[74] Jingtao Ma, Hu Dong, and H. Zhang. "Calibration of Microsimulation with Heuristic Optimization Methods". In: *Transportation Research Record: Journal of the Transportation Research Board* 1999 (2007), pp. 208–217.

[75] Jung-Beom Lee and Kaan Ozbay. “New Calibration Methodology for Microscopic Traffic Simulation Using Enhanced Simultaneous Perturbation Stochastic Approximation Approach”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2124 (2009), pp. 233–240.

[76] Tao Ma and Baher Abdulhai. “Genetic Algorithm-Based Optimization Approach and Generic Tool for Calibrating Traffic Microscopic Simulation Parameters”. In: *Transportation Research Record* 1800.1 (2002), pp. 6–15.

[77] Grant Schultz and L. Rilett. “Analysis of Distribution and Calibration of Car-Following Sensitivity Parameters in Microscopic Traffic Simulation Models”. In: *Transportation Research Record: Journal of the Transportation Research Board* 1876.1 (2004), pp. 41–51.

[78] Brandon M. Turner and Trisha Van Zandt. “A tutorial on approximate Bayesian computation”. In: *Journal of Mathematical Psychology* 56.2 (2012), pp. 69–85.

[79] Katalin Csilléry et al. “Approximate Bayesian Computation (ABC) in practice”. In: *Trends in Ecology and Evolution* 25.7 (2010), pp. 410–418.

[80] Simon Tavaré et al. “Inferring coalescence times from DNA sequence data”. In: *Genetics* 145.2 (1997), pp. 505–518.

[81] Jonathan K. Pritchard et al. “Population growth of human Y chromosomes: A study of y chromosome microsatellites”. In: 16.12 (1999), pp. 1791–1798.

[82] Mark A. Beaumont, Wenyang Zhang, and David J. Balding. “Approximate Bayesian computation in population genetics”. In: *Genetics* 162.4 (2002), pp. 2025–2035.

[83] Paul Marjoram et al. “Markov chain Monte Carlo without likelihoods.” In: *Proceedings of the National Academy of Sciences of the United States of America* 100.26 (2003), pp. 15324–8.

[84] S A Sisson, Y Fan, and Mark M Tanaka. “Correction for Sisson et al., Sequential Monte Carlo without likelihoods”. In: *Proceedings of the National Academy of Sciences* 106.39 (2009), pp. 16889–16889.

[85] C. C. Drovandi and A. N. Pettitt. “Estimation of Parameters for Macroparasite Population Evolution Using Approximate Bayesian Computation”. In: *Biometrics* 67.1 (2011), pp. 225–233.

[86] D. J. Nott et al. “Approximate Bayesian computation and Bayes linear analysis: Towards high-dimensional ABC”. In: *Journal of Computational and Graphical Statistics* 23.1 (2014), pp. 65–86.

[87] J. Li et al. “Extending approximate Bayesian computation methods to high dimensions via a Gaussian copula model”. In: *Computational Statistics and Data Analysis* 106 (2017), pp. 77–89.

[88] Scott A Sisson, Yanan Fan, and Mark Beaumont. *Handbook of approximate Bayesian computation*. CRC Press, 2018.

[89] Dennis Prangle. “Summary statistics in approximate Bayesian computation”. In: *arXiv preprint arXiv:1512.05633* (2015).

[90] Paul Fearnhead. “Asymptotics of ABC”. In: *Handbook of Approximate Bayesian Computation*. Chapman and Hall/CRC, 2018, pp. 269–288.

[91] Katalin Csilléry, Olivier François, and Michael GB Blum. “abc: an R package for approximate Bayesian computation (ABC)”. In: *Methods in ecology and evolution* 3.3 (2012), pp. 475–479.

[92] David T Frazier et al. “Asymptotic properties of approximate Bayesian computation”. In: *Biometrika* 105.3 (2018), pp. 593–607.

[93] B. Anil Kumar et al. “Bus travel time prediction: a log-normal autoregressive (AR) modelling approach”. In: *Transportmetrica. A, Transport science* 16.3 (2020), pp. 807–839.

[94] Chen Zhen and Fan Wei. “Data analytics approach for travel time reliability pattern analysis and prediction”. In: *Journal of Modern Transportation* 27.4 (2019), pp. 250–265.

[95] Kunpeng Zhang et al. “A novel generative adversarial network for estimation of trip travel time distribution with trajectory data”. In: *Transportation Research Part C: Emerging Technologies* 108 (2019), pp. 223–244.

[96] Mahmood Rahmani, Erik Jenelius, and Haris N. Koutsopoulos. “Non-parametric estimation of route travel time distributions from low-frequency floating car data”. In: *Transportation Research Part C: Emerging Technologies* 58 (2015), pp. 343–362.

[97] Muhammad Adnan et al. “Estimation of travel time distributions for urban roads using GPS trajectories of vehicles: a case of Athens, Greece”. In: *Personal and Ubiquitous Computing* (2020).

[98] Feng Guo, Hesham Rakha, and Sangjun Park. “Multistate Model for Travel Time Reliability”. In: *Transportation Research Record* 2188.1 (2010), pp. 46–54.

[99] Evanthia Kazagli and Haris N. Koutsopoulos. “Estimation of Arterial Travel Time from Automatic Number Plate Recognition Data”. In: *Transportation Research Record* 2391.1 (2013), pp. 22–31.

[100] Zhenliang Ma et al. “Modeling distributions of travel time variability for bus operations”. In: *Journal of Advanced Transportation* 50.1 (2016), pp. 6–24.

[101] Yuxiong Ji et al. “Estimation of Bimodal Urban Link Travel Time Distribution and Its Applications in Traffic Analysis”. In: *Mathematical Problems in Engineering* (2015).

[102] Xiaoqin Luo et al. “Grouped travel time estimation in signalized arterials using point-to-point detectors”. In: *Transportation Research Part B: Methodological* 130 (2019), pp. 130–151.

[103] Jabari Saif Eddin, Prassas Elena, and Xu Zeng. “Applying Finite Mixture Models to New York City Travel Times”. In: *Journal of transportation engineering* 146.5 (2020), pp. 2473–2893.

[104] Shukai Chen and Daniel(Jian) Sun. “A multistate-based travel time schedule model for fixed transit route”. In: *Transportation Letters* 11.1 (2019), pp. 33–42.

[105] Peng Chen, Yin Kai, and Jian Sun. “Application of Finite Mixture of Regression Model with Varying Mixing Probabilities to Estimation of Urban Arterial Travel Times”. In: *Transportation Research Record* 2442.1 (2014), pp. 96–105.

[106] Ehsan Mazloumi, Graham Currie, and Geoffrey Rose. “Using GPS Data to Gain Insight into Public Transport Travel Time Variability”. In: *Journal of Transportation Engineering-asce - J TRANSP ENG-ASCE* 136 (Jan. 2009).

[107] J. W. C. van Lint and H. J. van Zuylen. “Monitoring and Predicting Freeway Travel Time Reliability: Using Width and Skew of Day-to-Day Travel Time Distribution”. In: *Transportation Research Record* 1917 (1 2005), pp. 54–62.

[108] Wenwen Qin and Meiping Yun. “Estimation of Urban Link Travel Time Distribution Using Markov Chains and Bayesian Approaches”. In: *Journal of Advanced Transportation* 2018 (2018).

[109] Min Chen et al. “A copula-based approach for estimating the travel time reliability of urban arterial”. In: *Transportation Research Part C: Emerging Technologies* 82 (2017), pp. 1–23.

[110] Peng Chen et al. “Modeling arterial travel time distribution by accounting for link correlations: a copula-based approach”. In: *Journal of Intelligent Transportation Systems* 23.1 (2019), pp. 28–40.

[111] Y. Yu et al. “Copula-Based Travel Time Distribution Estimation Considering Channelization Section Spillover”. In: *IEEE Access* 8 (2020), pp. 32850–32861.

[112] Wenwen Qin, Xiaofeng Ji, and Feiwen Liang. “Estimation of urban arterial travel time distribution considering link correlations”. In: *Transportmetrica A: Transport Science* 16.3 (2020), pp. 1429–1458.

[113] Sen Luan et al. “Modeling travel time volatility using copula-based Monte Carlo simulation method for probabilistic traffic prediction”. In: *Transportmetrica A: Transport Science* 0.0 (2019), pp. 1–24.

[114] Shuling Wang, Wei Huang, and Hong K. Lo. “Traffic parameters estimation for signalized intersections based on combined shock wave analysis and Bayesian Network”. In: *Transportation Research Part C: Emerging Technologies* 104 (2019), pp. 22–37.

[115] John F Shortle et al. *Fundamentals of queueing theory*. Vol. 399. John Wiley & Sons, 2018.

[116] Douglas M Bates and Donald G Watts. *Nonlinear regression analysis and its applications*. Vol. 2. Wiley New York, 1988.

[117] Douglas A Reynolds. “Gaussian Mixture Models.” In: *Encyclopedia of biometrics* 741 (2009).

[118] Anatoly B Kolomeisky et al. “Phase diagram of one-dimensional driven lattice gases with open boundaries”. In: *Journal of Physics A: Mathematical and General* 31.33 (1998), p. 6911.

[119] Lele Zhang et al. “Behaviour of traffic on a link with traffic light boundaries”. In: *Physica A: Statistical Mechanics and its Applications* 503 (2018), pp. 116–138.

[120] D. Li et al. “Train Dwell Time Distributions at Short Stop Stations”. In: 2014.

[121] Qiang Meng and Xiaobo Qu. “Bus dwell time estimation at bus bays: A probabilistic approach”. In: *Transportation Research Part C: Emerging Technologies* 36 (2013), pp. 61–71.

[122] *TRAFED User ’ s Guide*. English. Version Version 6.0. Federal Highway Administration. 2006. 35 pp.

[123] *VISSIM User Manual*. Version Version 4.10. PTV Vissim. 2005. 1 p.

[124] Jinhwan Jang et al. “Modeling of time headway distribution on sub-urban arterial: Case study from South Korea”. In: *Procedia-social and behavioral sciences* 16 (2011), pp. 240–247.

[125] Ali S Al-Ghamdi. “Analysis of time headways on urban roads: case study from Riyadh”. In: *Journal of Transportation Engineering* 127.4 (2001), pp. 289–294.

[126] Serge P Hoogendoorn and Hein Botma. “Modeling and estimation of headway distributions”. In: *Transportation research record* 1591.1 (1997), pp. 14–22.

[127] Serge P Hoogendoorn and Piet HL Bovy. “New estimation technique for vehicle-type-specific headway distributions”. In: *Transportation Research Record* 1646.1 (1998), pp. 18–28.

[128] Scott A Hill. “Numerical analysis of a time-headway bus route model”. In: *Physica A: Statistical Mechanics and its applications* 328.1-2 (2003), pp. 261–273.

[129] Giuseppe Bellei and Konstantinos Gkoumas. “Transit vehicles’ headway distribution and service irregularity”. In: *Public transport* 2.4 (2010), pp. 269–289.

[130] Gaohua Guo et al. “Headway-based evaluation of bus service reliability”. In: *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE. 2011, pp. 1864–1868.

[131] Man Zhang et al. “Tailored Wakeby-type distribution for random bus headway adherence ratio”. In: *Transportation Research Part C: Emerging Technologies* 86 (2018), pp. 220–244.

[132] Young-Ji Byon et al. “Bunching and headway adherence approach to public transport with GPS”. In: *International Journal of Civil Engineering* 16.6 (2018), pp. 647–658.

[133] M Ruan et al. “Probability-based bus headway regularity measure”. In: *IET intelligent transport systems* 3.4 (2009), pp. 400–408.