University of California
Santa Barbara

Analyzing and Influencing Traffic Networks with Mixed Autonomy

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Electrical and Computer Engineering

by

Daniel Arthur Lazar

Committee in charge:

Professor Ramtin Pedarsani, Chair
Professor Konstadinos Goulias
Professor Joao Hespanha
Professor Jason Marden

June 2021

The Dissertation of Daniel Arthur Lazar is approved.

Professor Konstadinos Goulias

Professor Joao Hespanha

Professor Jason Marden

Professor Ramtin Pedarsani, Committee Chair

June 2021

Analyzing and Influencing Traffic Networks with Mixed Autonomy

by

Daniel Arthur Lazar

To Grace

Acknowledgements

I'd like to first thank my advisor, Ramtin. I have learned so much during these past five years, in a large part due to your mentorship. You have taught me how to search for the interesting core of a general problem, how to mathematically formulate my thoughts, and how to communicate them. You have also advocated for me in so many ways, including connecting me to internships and other opportunities, and I thank you for this.

I'd like to thank the people who worked, often behind the scenes, to make all this possible, including Val and the CCDC and ECE staff, and the people who support our work by maintaining the facilities in which we work.

I'm so grateful to my academic mentors and collaborators. Dorsa in many ways has co-advised me, and I will forever be grateful to her for impressing upon me her way of thinking about engineering problems, as well as spending so much time training me in giving effective presentations. Sam was an invaluable mentor in my early stages in graduate school, and help guide my written communication in papers. I would also like to thank my committee members, Professors Goulias, Hespanha, and Marden.

Erdem is the best collaborator a person can ask for. He is extremely sharp, attentive, and has an inverted sleep schedule from mine so that don't have to worry about editing a manuscript at the same time. Even beyond our great technical conversations, he's been a great travel partner. I'll never forget the time we almost got flattened by that huge truck in Tulum, or pretending we could tell the difference between cheeses to appease the owner at the cheese factory in Toledo.

I would be remiss if I left out my other collaborators and the many people who I had great technical discussions with. This includes, but is not limited to, Mark, Woody, Mohit, Kabir, Nick, and Philip.

My internships were very important in developing other important technical and professional skills outside of the ones emphasized in grad school. I'm grateful to Shahab, Gabe, and Charlie at Apple, and Hongyuan, Kshitiz, Jack, and Ashkan at Facebook. Thank you to Jerry for initially encouraging me to go to grad school.

I made so many valuable friendship in graduate school, and I can't believe that we met only five years ago. Sharad, Simone, Henrique, Amanda, Raphael, Rachel, Ahmed, Mohammed, Timmie, Murat, Liz, Alex, Emmanuel, Melanie, Matina, Angela, Evan, Kevin, Sana, Guosong, and so many others – you were the ones who got me through this. Thank you also to my dear friends Trevor, Elana, Aron, and Frank, for providing much needed support during these years.

When I started graduate school, I thought I was putting my life on hold. I didn't realize the many opportunities for growth that existed that existed outside of academics – some fun, and some painful. I ended up learning as much outside of my studies than I did in my studies, and I want to thank everyone who took part in that growth. I want to specifically thank my friends and fellow fighters in the union, my bandmates, my political collaborators, and those working to make our department a better place to work.

I'd like to thank my mother, who taught me an understanding of people, and who has offered so much support over the years. I'd like to thank my father, who instilled in me a love for engineering and scientific pursuits. I hope to follow his example of being uncompromising when it comes to doing what is right. Thank you David, for leading by example in thinking about the big questions, though we often come to different conclusions. Amanda, being able to talk to you about anything has been so important to me.

Finally, I'd like to thank Grace for always being by my side, even from across the state. I couldn't have done any of this without you. I love you, and here's to many more.

Curriculum Vitæ

Daniel Arthur Lazar

Education

2021	Ph.D. in Electrical and Computer Engineering (Expected), University of California, Santa Barbara.
2018	M.S. in Electrical and Computer Engineering, University of California, Santa Barbara.
2014	B.S. in Electrical Engineering, Washington University in Saint Louis.

Journal Publications

1. **D. A. Lazar***, Erdem Bıyık*, Dorsa Sadigh, and Ramtin Pedarsani, "Learning How to Dynamically Route Autonomous Vehicles", to appear, *Transportation Research Part C*.
2. Erdem Bıyık*, **D. A. Lazar***, Ramtin Pedarsani, and Dorsa Sadigh, "Incentivizing Efficient Equilibria in Traffic Networks", to appear, *IEEE Transactions on Control of Network Systems*.
3. **D. A. Lazar**, and Ramtin Pedarsani, "Optimal Tolling for Multitype Mixed Autonomous Traffic Networks", in *IEEE Control and Systems Letters*, 2020.
4. **D. A. Lazar**, S. Coogan, R. Pedarsani, "Routing for Traffic Networks with Mixed Autonomy," in *IEEE Transactions on Automatic Control*, 2020.

Proceedings of Refereed Conferences

1. **D. A. Lazar**, and Ramtin Pedarsani, "Anonymous Tolling for Traffic Networks with Mixed Autonomy", submitted, *IEEE Conference on Decision and Control*
2. W. Z. Wang*, M. Beliaev*, E. Bıyık*, **D. A. Lazar**, R. Pedarsani, and D. Sadigh, "Emergent Prosociality in Multi-Agent Games Through Gifting", accepted, *International Joint Conferences on Artificial Intelligence*, 2021.
3. M. Beliaev, E. Bıyık, **D. A. Lazar**, W. Z. Wang, D. Sadigh, and R. Pedarsani, "Incentivizing Routing Choices for Safe and Efficient Transportation in the Face of the COVID-19 Pandemic", *ACM/IEEE International Conference on Cyber-Physical Systems*, 2021.
4. **D. A. Lazar**, and R. Pedarsani, "Optimal Tolling for Multitype Mixed Autonomous Traffic Networks", *American Control Conference*, 2021.
5. **D. A. Lazar**, S. Coogan, and R. Pedarsani, "Optimal tolling for heterogeneous traffic networks with mixed autonomy", *IEEE Conference on Decision and Control*, 2019.

6. E. Bıyık, **D. A. Lazar**, D. Sadigh, and R. Pedarsani, "The green choice: Learning and influencing human decisions on shared roads", *IEEE Conference on Decision and Control*, 2019.
7. **D. A. Lazar**, K. Chandrasekher, R. Pedarsani, and D. Sadigh, "Maximizing road capacity using cars that influence people", *IEEE Conference on Decision and Control*, 2018.
8. E. Bıyık*, **D. Lazar**, R. Pedarsani, and D. Sadigh, "Altruistic autonomy: Beating congestion on shared roads", *Workshop on the Algorithmic Foundations of Robotics*, 2018.
9. **D. A. Lazar**, S. Coogan, and R. Pedarsani, "The price of anarchy for transportation networks with mixed autonomy", *American Control Conference*, 2018.
10. **D. A. Lazar**, S. Coogan, and R. Pedarsani, "Capacity modeling and routing for traffic networks with mixed autonomy", *IEEE Conference on Decision and Control*, 2017.

Abstract

Analyzing and Influencing Traffic Networks with Mixed Autonomy

by

Daniel Arthur Lazar

As publicly available cars gain semi-autonomous capabilities and companies promise to deliver fully autonomous vehicles in the near future, it is important to understand the effects that these vehicles will have on traffic networks. If all vehicles on the road are autonomous and are properly coordinated, road capacity and intersection throughput can be greatly increased. However, in the medium-term future roads will exist in a state of *mixed autonomy*, meaning they will be shared between human-driven and autonomous vehicles.

The properties of this regime are not yet well-established. Preliminary studies show that even if autonomous vehicles can increase road capacity, converting human-driven vehicles to autonomous vehicles can *increase* the overall delay experienced by travelers. This phenomenon stems from the nature of this *sociotechnical* system, containing humans who are making selfish routing choices rather than choices that are socially optimal. Accordingly, we must understand the nature of these choices in order to design efficient mixed autonomous traffic networks.

This thesis theoretically characterizes areas of inefficiency and opportunity in mixed autonomy and offers control algorithms with theoretical guarantees or benchmarks. Specifically, it studies problems including 1) formulating a framework for understanding and bounding the inefficiency due to selfish routing, 2) establishing toll structures which can decrease or eliminate this inefficiency, including tolls which differentiate between the different types of vehicles to varying degrees, 3) actively learning human preferences for

time/money tradeoffs to determine optimal pricing for an autonomous ride-hailing service in the presence of human drivers who minimize their travel time, 4) decongesting roads by using Reinforcement Learning to route autonomous vehicles in a way which leverages the reaction of the surrounding uncontrolled human-driven vehicles, and 5) designing a controller for the low-level actions of autonomous vehicles to form platoons in the presence of vehicles driven by humans. In all these settings, we provide theoretical guarantees or benchmarks so that a system designer can intelligently choose whether to and how to implement each suggested scheme. In this fashion, we can ensure that autonomous vehicles will not worsen, and will rather improve, the performance of transportation networks in the years to come.

Contents

List of Figures

Chapter 1

Introduction

The addition of autonomous vehicles to public road networks opens up a number of challenges and opportunities. For the forseeable future, traffic networks will exist in a state of *mixed autonomy*, meaning they will be shared by human-driven and autonomous vehicles. In this regime, properly designed and controlled autonomous vehicles can increase road capacity, stabilize traffic flow, and route in a way to reduce congestion on roads. The interactions between autonomous vehicles and human-driven vehicles are not always easy to predict, and they take place on a number of scales, ranging from vehicle-level interactions to routing-level interactions. If autonomous vehicles are not carefully designed with these interactions in mind, instead of decreasing congestion, their presence may instead end up *worsening* it.

In mixed autonomy, each interaction between autonomous and human-driven vehicles yields an opportunity for using autonomous vehicles to positively *influence* human drivers in a way toward improving a societal goal, such as reducing traffic congestion. We can then consider possible interventions such as designing autonomous vehicles to intelligently form platoons to increase road capacity, directly controlling the routing of autonomous vehicles to influence the routing of human drivers, or using tolling to influ-

ence the routing of both vehicle types. Each of these interventions incur a cost, be it in altering infrastructure, coordination, or even unfair economic effects. It is therefore important to fully understand the potential gain from each proposed intervention so that it can be weighed against the costs of implementing that intervention.

In this work we study problems in both high-level (routing) and low-level (vehicle control) interactions. In these, for specific possible interventions, we seek to do the following: 1) model the interaction between human-driven and autonomous vehicles, 2) propose an algorithm for guiding this interaction to improve a societal goal, and 3) provide theoretical guarantees about the range of possible outcomes of this intervention, or provide theoretical baselines for understanding the empirical performance of the intervention. By understanding these key elements, a system designer can make informed decisions on how to guide these interactions in mixed autonomy.

1.1 Contributions

The main contributions of this thesis are as follows.

- Chapter 2: We propose a model for road capacity in mixed autonomy which encapsulates vehicle platooning, when autonomous vehicles travel closely to one another. In our model, the capacity of a road is a function of the fraction of vehicles on a road that are autonomous. Using this model, we formulate a *congestion game* formulation of mixed autonomous vehicle flow, where the latency on each road is a function of the flow of each vehicle type on that road. This formulation can be used to study optimal routing for traffic networks with mixed autonomy, as well as *Wardrop Equilibria*, in which the users of the network are modeled as selfish decision makers who choose the shortest route available to them.

- In this same chapter we study the *Price of Anarchy* (PoA) for traffic networks with mixed autonomy. Using the total delay experienced by all users as the *social cost*, which measures the quality of a routing, the Price of Anarchy is an upper bound, over a class of congestion games, of the ratio of the social cost in the worst-case Wardrop Equilibrium to that of the routing which minimizes the social cost. Classic results bound the Price of Anarchy for congestion games with a single vehicle type, and in very limited circumstances, multiple vehicle types. We show that in general, the PoA in mixed autonomy can be *unbounded*. Moreover, we provide novel PoA bounds in mixed autonomy which depend on the asymmetry in how the vehicle types contribute to road congestion. We also prove a lower bound on the Price of Stability in this setting.

- Chapter 3: Having established the potential downsides of allowing uncontrolled selfish routing, we study using tolling to improving the social cost of equilibrium routing. We show that inefficiency due to selfish routing can be eliminated using differentiated tolls, in which the different vehicle types pay different tolls. This solution may be costly due to the logistical, privacy, and fairness challenges of levying separate tolls to the different vehicle types. Because of this, we study *anonymous tolls*, in which all vehicle types pay the same toll. We show that anonymous tolls have fundamentally limited usefulness in this setting; nonetheless we provide anonymous tolls which improve worst-case guarantees on the Price of Anarchy. To avoid this fundamental limitation, we provide a tolling scheme which is *almost* anonymous, and has only an infinitesimal distinction between the tolls for the different vehicle classes, and we show that this tolling scheme completely eliminates inefficiency due to selfish routing in our considered setting. We also provide a lower bound on the Price of Anarchy in this setting with variable marginal cost tolls

applied.

- Chapter 4: We consider the setting of a network shared between users of an autonomous ride-hailing service and human drivers in which a social planner controls the prices offered to users of the ride-hailing service. We model how autonomous service users choose between routes with different prices and latencies, and we develop an algorithm to learn the preferences of the users. With these learned preferences, we formulate a method for choosing prices for the ride hailing service to maximize a social objective. To understand the efficacy of this scheme, we develop theoretical benchmarks for the range of possible outcomes in this setting, and, using these benchmarks, show the efficacy of our algorithm.

- Chapter 5: Though the above studies focus on controlling network *equilibria*, it is not straightforward to control a network to move from one equilibrium to another. We consider a setting in which a social planner has direct control over the routing of the autonomous vehicles, and could potentially use this control to achieve efficient equilibria and mitigate congestion due to network shocks such as traffic collisions. To investigate this, we study a dynamic routing game, in which the routes of the arriving flow of autonomous vehicles can be controlled by a central planner, and in response the human drivers dynamically choose their routes selfishly. Again we develop theoretical benchmarks to understand the maximum benefit which could be accrued if all routing was perfectly controlled. We use Deep Reinforcement Learning to find a policy which routes autonomous vehicles such that it guides human reaction to route in a way which decreases congestion. We show the efficacy of these learned policies, and show large congestion in the absence of them.

- Chapter 6: Finally, we investigate how to use micro-level interactions to positively affect global behavior. Specifically, we analyze the optimal lane assignment and

ordering of vehicles in the mixed autonomy setting. Using this, we design the low-level controller of autonomous vehicles to reorder themselves in the presence of vehicles driven by humans in a manner to maximize the capacity of the road. We showcase the efficacy of our algorithm in simulation, and compare its results to the theoretically optimal vehicle configuration.

1.2 Related Works

In this section we summarize key prior work in the realm of mixed autonomy, and for works which we extend to the realm of mixed autonomy. The relevant chapters of this thesis contain more detailed surveys of the literature. The review below highlights some key relevant works, but is not a comprehensive literature review of these large topics.

Many studies have established the deleterious effects of road congestion, including the negative impact of ride-hailing services [3, 4]. There are a number of methods which have been proposed to minimize or eliminate congestion, including ramp metering, road tolling, variable speed limits, and synchronized intersection cycles [5–9]. Other works consider selectively revealing network state information to a fraction of users of a traffic network in order to strategically influence their routing choices [10, 11].

The framework of *congestion games* has been developed in order to understand properties of route choice in traffic networks, or other resources susceptible to congestion, such as computer networks [12]. This framework allows the study of both the socially optimal routing as well as the equilibria that emerge when the users of the network are selfish and wish to minimize their personal cost [13–17]. This framework has been used to characterize the Price of Anarchy (PoA) for a class of games, bounding the maximum gap between the social cost of an equilibrium routing and the socially optimal routing [18–21]. A framework for calculating the Price of Anarchy was establish for a

subclass of games which include multiple vehicle types, each influencing and experiencing congestion differently [22]. However, the assumptions of this framework fail to capture the case of networks with mixed autonomy.

One important ramification of this quantity is determining how useful it would be to implement some control or influence over user routing; one example of such a scheme is *tolling*. If the PoA is large, a system designer would wish to control routing somehow; if the PoA is bounded to a small value, the suboptimal state reached by the selfish equilibrium would not be much worse than it would be if it were optimally controlled. In our work we find the PoA to be large in the case of mixed autonomy; we therefore examine control schemes such as tolling. Tolling schemes which completely eliminate inefficiency due to selfish routing have been established as far back as the 1950s, with many later developments to handle some cases of networks with multiple vehicle types, users with different values of time vs money, or users with unknown parameters [23–26].

Another method for influencing routing is when a system designer controls the routing of some fraction of users, and the remaining users route selfishly in response. This is known as a Stackelberg Game, and it has been studied in many settings including congestion games and routing games using more complicated congestion models [27–29].

This dissertation deals largely with the effects of autonomous vehicles on traffic networks. The study of controlling platoons (closely space groups) of vehicles dates back to the 1970s, in the study of how to space vehicles in an automated guideway transit systems [30]. As autonomous technology advanced, multitudes of works studied how to control platoons of passenger vehicles. Some relevant works study the effects of autonomous technologies on throughput and mobility on urban roads [31, 32]. Other works characterize how the presence of autonomous vehicles changes traffic dynamics [33], extending well known models for vehicle flow of a single type [34]

Many relevant works study the effects and possibilities of integrating vehicles with

autonomous technology in among with vehicles that are driven manually by humans. Some works investigate how autonomous vehicles can be used to stabilize traffic flow [35,36]. Other works seek to control platoons to improve traffic flow [37] or even decongest highways [38] or localize traffic disturbances [39].

Despite many works showing the potential benefits of autonomy, other works show that the presence of vehicles with autonomous technology can *worsen* congestion. For example, the existing autonomous technology installed in vehicles on the market can amplify traffic disturbances [40]. Moreover, even if autonomous vehicles improve road usage efficiency, converting some human-driven vehicles to more efficient autonomous vehicles can actually *worsen* congestion [41]. This motivates the study of how much worse congestion can be in this mixed-autonomy setting, and how to mitigate this congestion.

Chapter 2

Mixed Autonomous Traffic Networks: Modeling and Analysis

In this chapter[1] we model traffic networks with mixed autonomy and use this model to analyze the inefficiency due to selfish routing in the mixed autonomy setting. Specifically, we introduce two capacity models for roads with mixed autonomy and use these models to construct a family of latency functions for the roads, which relate the flow of vehicles on the roads to the delay experienced by drivers on these roads. We show that in the mixed autonomy setting, the inefficiency due to selfish routing can be much greater than in the case of a single vehicle type. We therefore provide bounds on this inefficiency, and analyze the tightness of these bounds.

[1]This chapter is in part adapted, with permission, from the works "Capacity modeling and routing for traffic networks with mixed autonomy", previously published by IEEE Conference on Decision and Control, © 2017 IEEE, "The price of anarchy for transportation networks with mixed autonomy", previously published by American Control Conference, © 2018 IEEE, and "Routing for traffic networks with mixed autonomy" previously published by IEEE Transactions on Automatic Control, © 2020 IEEE, and is joint work with Sam Coogan and Ramtin Pedarsani [42–44].

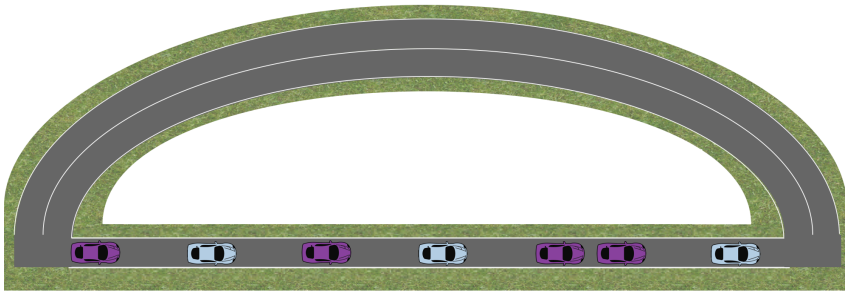
2.1 Introduction

In recent years, automobiles are increasingly equipped with autonomous and semi-autonomous technology, which has potential to dramatically decrease traffic congestion [45]. Specifically, autonomous technologies enable *platooning*, in which these vehicles automatically maintain short headways between them via adaptive cruise control (ACC) or cooperative adaptive cruise control (CACC). ACC uses sensing such as radar or LIDAR to maintain a specific distance to the preceeding vehicle with faster-than-human reaction time, and CACC augments this with vehicle-to-vehicle communications.

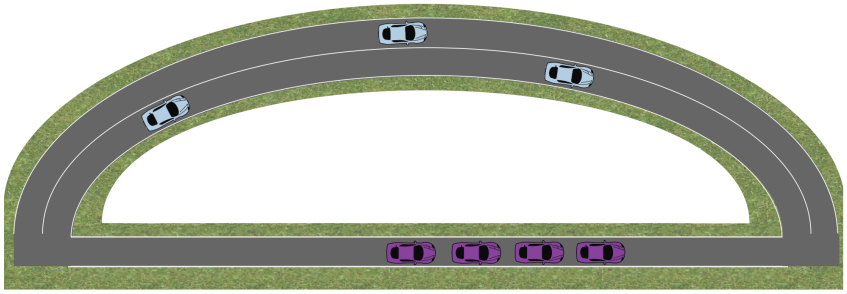
When all vehicles are autonomous, the use of platooning has the potential to increase network capacity as much as three-fold [31] by enabling synchronous acceleration at green lights [46]. However, the presence of human-driven vehicles – leading to *mixed autonomy* – makes much of these benefits unclear.

Moreover, even in the absence of autonomous capabilities, it is well known that if drivers route selfishly and minimize their individual traffic delays, this does not in general minimize *overall* traffic delay. Understanding the extent of this phenomenon can help city planners – if selfish routing does not adversely affect travel delay too much, then it may not be necessary to try to control vehicle flow using schemes such as tolling. Alternatively, if selfishness can lead to much worse road delay, then a city planner may wish to try to control human routing decisions. See Figure 6.1 for an example of selfish routing and optimal routing in mixed autonomy.

The ratio between traffic delay under worst-case selfish routing and optimal routing is called the Price of Anarchy (PoA) and is well understood for networks with only human-driven vehicles [19–21, 47, 48]. Many such works also bound the bicriteria, which quantifies, for any given volume of vehicle flow demand, how much additional flow can be routed optimally to result in the same overall latency as the original volume of traffic



(a) When vehicles route selfishly, vehicles pack onto a congested road.



(b) In optimal routing, only autonomous vehicles are sent onto the road most amenable to platooning.

Figure 2.1: A social planner can decrease overall travel times by make routing decisions that utilize autonomous vehicles' ability to platoon, and choosing different routes for human-driven vehicles (blue) and autonomous vehicles (purple).

routed selfishly. Other studies have bounded the PoA with multiple modes of transportation [22, 49]. However, these prior works require assumptions that do not capture vehicle flow on roads shared between human-driven and autonomous vehicles, leaving open the question of the Price of Anarchy in mixed autonomy. In fact, *we show that these previous results do not hold, and the PoA for roads with mixed autonomy can in general be unbounded!*

Motivated by this observation, in we provide novel bounds on the PoA and bicriteria that depend on the degree to which platooning affects road delay and on the degree of the polynomial describing road delay. To do so, we use two models that describe road capacity as a function of the fraction of vehicles on the road that are autonomous; each model corresponds to a different assumption regarding the technology that enables

platooning. We use these capacity models with a known polynomial road delay function, and, for this class of latency functions, we bound the Price of Anarchy and bicriteria. We develop two mechanisms for bounding the PoA, which yield bounds that are tighter depending on platoon spacing and polynomial degree. In our development we provide the main elements of our proofs and defer proofs of the lemmas to the appendix.

In our formulation, the benefit due to the presence of autonomous vehicles is limited to platoon formation, and the probability that each vehicle is autonomous is independent of the surrounding vehicles. While we acknowledge that autonomous vehicles yield other benefits such as smoothing traffic shockwaves, we consider platooning because it is a mature technology that is commercially available. Further, if autonomous vehicles actively rearrange themselves to form platoons, the resulting capacity falls between the two capacity models presented here, as investigated in Chapter 6.

Motivating Example. To show that the PoA bounds previously developed for roads with only one type of vehicle (*i.e.* no autonomous vehicles) do not hold, we present an example of a road network with unbounded PoA (Fig. 2.2). Consider a network of two parallel roads, with road latency functions $c_1(z_1^1, z_1^2) = 1$ and $c_2(z_2^1, z_2^2) = \zeta z_2^1$. On each road i the latency is a function of the human-driven flow (z_i^1) and the autonomous flow (z_i^2) on that road. Suppose we have $\frac{1}{\zeta}$ units of human-driven vehicle flow and 1 unit of autonomous traffic demand to cross from node s to node s', with $\zeta \geq 1$.

Optimal routing puts all human-driven cars on the top road and all autonomous cars on the bottom road; when vehicles route selfishly they all end up on the bottom road. This yields a price of anarchy of $\zeta + 1$. The bicriteria is also $\zeta + 1$, as $\zeta + 1$ times as much traffic, optimally routed, yields the same total cost as the original amount of traffic at Wardrop Equilibrium.[2] This examples leads us already to our first proposition, which

[2]Though in this case autonomous vehicles do not affect road delay, other examples in Section 2.3 also

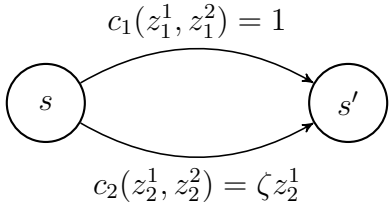


Figure 2.2: Road network with price of anarchy and bicriteria that grow unboundedly with ζ when considering $1/\zeta$ units of human-driven flow and 1 unit of autonomous flow demand, with $\zeta \geq 1$. Function arguments z_i^1 and z_i^2 respectively denote human-driven and autonomous vehicle flow on a road.

lays the foundations for the contributions of this chapter:

Proposition 2.1. *The Price of Anarchy and bicriteria are in general unbounded in mixed autonomy.*

Motivated by this proposition, we develop the notion of the *degree of asymmetry* of a road and use this, in conjunction with the degree of the polynomial cost function, to parameterize the bound on the Price of Anarchy. We also lower bound a more optimistic quantity of inefficiency due to selfish routing, the Price of Stability, in this setting. To summarize, we

1. show that previous PoA results do not hold for mixed autonomy,
2. develop a realistic class of polynomial cost functions for traffic of mixed autonomy,
3. develop two mechanisms for bounding the PoA using this cost function with both capacity models,
4. bound the Price of Anarchy and bicriteria, analyze the bound tightness, and
5. provide a lower bound for the Price of Stability in mixed autonomy.

yield an unbounded PoA with both vehicle types affecting road delay.

2.1.1 Related Works

Congestion Games and Wardrop Equilibria. Our work is related to the optimal traffic assignment problem, *e.g.* [13], which studies how to optimally route vehicles on a network when the cost (*i.e.* delay) on a road link is a function of the flow of vehicles that travel on that link. We are concerned specifically with the relationship between optimal traffic assignment and Wardrop Equilibria, which occur when drivers choose their paths selfishly. For a survey on literature on Wardrop Equilibria, see [50]; [51] describes other notions of equilibria. Classic works on Wardrop Equilibria and the associated tools for analyzing them include [16, 17, 52].

In an important development, Smith [16] establishes the widely used *Variational Inequality* and uses it to describe flows at Wardrop Equilibrium, in which all users sharing an origin and destination use paths of equal cost and no unused path has a smaller cost. For any feasible flow z and equilibrium flow $\hat{z}$, the Variational Inequality dictates that

$$\langle c(\hat{z}), \hat{z} - z \rangle \leq 0 \ , \tag{2.1}$$

where z is a vector describing vehicle flow on each road, $c(z)$ maps a vector of flows to a vector of the delay on each road, and $\langle \cdot, \cdot \rangle$ denotes the inner product of two terms. Note that in the absence of an assumption about the *monotonicity* of c (see the following section for a definition), the Variational Inequality is a necessary but not sufficient condition for equilibria [51]. The Variational Inequality is fundamental for establishing our PoA bound.

Multiclass Traffic. Some previous works consider traffic assignment and Wardrop Equilibria with multiclass traffic, meaning traffic with multiple vehicle types and transportation modes that affect and experience road latency differently (*e.g.* [14, 15, 53]).

Florian [15] demonstrates how to calculate equilibria for a multimodal system involving personal automobiles and public transportation. They use a relaxation that assumes public transportation will take the path that would be shortest in the absence of cars. In the case of mixed autonomy, this is not a fair assumption.

Dafermos [53] assumes the Jacobian of the cost function is symmetric and positive definite. Similarly, Hearn *et. al.* [14] deal with a monotone cost function, *i.e.* satisfying the property

$$\langle c(z) - c(z'), z - z' \rangle \geq 0 \tag{2.2}$$

for flow vectors z and z'.

However, traffic networks with mixed autonomy are in general nonmonotone. To see this, consider a network of two roads with costs $c_1(z_1^1, z_1^2) = 3z_1^1 + z_1^2 + t_1$ and $c_2(z_2^1, z_2^2) = 3z_2^1 + 2z_2^2 + t_2$, where t_1 and t_2 are constants denoting the free-flow latency on roads 1 and 2. This corresponds to a road in which autonomous vehicles can platoon closely and another road on which they cannot platoon as closely. The Jacobian of the cost function is as follows:

$$\begin{bmatrix} 3 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 3 & 2 \end{bmatrix},$$

which is not symmetric, and the vector $z = \begin{bmatrix} -1 & 2 & 0 & 0 \end{bmatrix}^T$ demonstrates that it is also not positive semidefinite. Monotonicity is closely related to the positive (semi-) definiteness of the Jacobian of the cost function. To show that the monotonicity condition is violated as well, consider that there are 2 units of human-driven flow demand and 3 units autonomous flow demand. With one routing in which all human-driven flow is on

the first road and all autonomous flow is on the second and another routing with these reversed, we find that the monotonicity condition is violated.

Similarly, Faroukhi *et. al.* [54] prove that in heterogeneous routing games with cost functions that are continuously differentiable, nonnegative for feasible flows, and non-decreasing in each of their arguments, then at least one equilibrium is guaranteed to exist. These mild conditions are satisfied in our setting. For heterogeneous games with two types, they further prove a necessary and sufficient condition for a potential function (and therefore unique equilibrium) to exist. However, the condition required can be considered a relaxation of the condition that the Jacobian of the cost function be symmetric. While broader than strict symmetry, this condition is still not satisfied in mixed autonomy. Notably, they describe tolls that, when applied, yield a cost function that satisfies this condition.

As described above, these previous works in multiclass traffic require restrictive assumptions and therefore do not apply to the case of mixed autonomy. In fact, in the case of mixed autonomy, the routing game is not formally a proper congestion game, as it cannot be described with a potential function. Nonetheless, in this paper we adapt tools developed for such games to derive results for mixed autonomous traffic.

Price of Anarchy. There is an abundance of research into the Price of Anarchy in nonatomic congestion games, originating in [55] and codified in [19–21, 47, 48]. In [21], the authors develop a general tool for analyzing Price of Anarchy in nonatomic congestion games. Though their development is specific to monotone cost functions, in this paper we broaden it to cost functions that are not necessarily monotone. Also relatedly, we find that in the case of no asymmetry, our Price of Anarchy bound for polynomial cost functions simplifies to the classic bound in [20, 47].

The previously mentioned works consider primarily single-type traffic. Perakis [22]

considers PoA in multiclass traffic using nonseparable, asymmetric, nonlinear cost functions with inelastic demand. However, they restrict their analysis to the case that the Jacobian matrix of the cost function is positive semidefinite. Similarly, Chau and Sim [49] consider the PoA for multiclass traffic with *elastic* demand with symmetric cost functions and positive semidefinite Jacobian of the cost function. As demonstrated earlier, these assumptions are violated in the case of mixed autonomy.

The study of multitype congestion games is related to the field of *weighted congestion games*, for which tools are available to bound the Price of Anarchy [56]. In this setting, each flow type is assigned a weight, and the total flow on an edge is the weighted sum of the flow types on that edge. However, the weight for each type is uniform across the network. Our setting is more general; a single flow (*i.e.* vehicle) type may have congestion effects that differ based on the edge (*i.e.* road) it travels upon. In fact, our bound on PoA on networks constructed from roads with the first capacity model can be viewed as an extension of the PoA for weighted congestion games to games in which the weights for the players are link-dependent.

A relevant work studies the Price of Anarchy for a broader class of delay functions than the ones we study [57]. They provide Price of Anarchy bounds which are polynomial in the number of source-destination pairs for each vehicle type; our bounds do not depend on this parameter and are therefore arbitrarily tighter in our case of interest. In the affine case their bound simplifies to 2 times the degree of asymmetry, which we define below; our bound is 4/3 times the degree of asymmetry.

Autonomy. In one line of research, autonomous vehicles are controlled to *locally* improve traffic by smoothing out stop-and-go shockwaves in congested traffic [30, 35, 36, 58–63], optimally sending platooned vehicles through highway bottlenecks [64, 65], and simultaneously accelerating platooned vehicles at signalized intersections [31, 46]. Other papers

investigate fuel savings attained using autonomous vehicles [66–69] or jointly controlling vehicles on a highway to localize and eliminate traffic disturbances [39]. Some works consider optimally routing and rebalancing a fleet of autonomous vehicles [70], though these generally consider a simpler model for road latency, in which capacitated roads have constant latency for flows below their capacity, and all roads are considered to be in this regime.

Some works have related models for road capacity and throughput under mixed autonomy, in particular [32, 46]. [71] provides a capacity model which assumes that all autonomous vehicles are platooned in periodic platoons, each with the same number of vehicles and the same number of human-driven vehicles between platoons. In contrast, we consider two capacity models: one in which autonomous vehicles can maintain a short headway behind any vehicle they follow, and one in which autonomous vehicles are placed randomly as the result of a Bernoulli process and only platoon opportunistically. The capacity model in [32] shares key features with ours and was published contemporaneously to our model [42]. Another work shows that autonomy can *increase* the total delay experienced by users [72].

2.2 Modeling

Consider a network described by a graph with n edges, where edges correspond to roads. We consider nonatomic flow, meaning each driver controls an infinitesimally small fraction of vehicle flow. Let m denote the number of vehicle types, where the latency on each edge is a function of the flow of each vehicle type on that edge. We generally use i to index roads and j to index vehicle types. For integer x, let $[x]$ denote $\{1, 2, \ldots, x\}$. Let $\langle v, w \rangle$ denote the inner product between the vectors v and w.

Consider inelastic flow demand, with demand described by $\{s_r, s'_r, \beta_r, j_r\}_{r \in [R]}$, where

for each *commodity* r in the set of commodities $[R]$, s_r is that commodity's source node, s'_r is the destination node, β_r is the flow demand, and j_r is the vehicle type of the commodity. Thus, each commodity describes the flow demand of a specific vehicle type between a specific source and destination node.

We use z_i^j to describe the flow of vehicle of type i on road j. Then, for road i, z_i is a column vector describing the flow of all vehicle types on road i:

$$z_i := \begin{bmatrix} z_i^1 & z_i^2 & \dots & z_i^m \end{bmatrix}^T .$$

Let z describe the flow over the entire network:

$$z = \begin{bmatrix} z_1^T & z_2^T & \dots & z_n^T \end{bmatrix}^T .$$

We can alternately describe the flow through a graph in terms of the flow on each route. Let $\mathcal{P}_r$ denote the set of simple paths available to commodity r. Let $f_p^j \geq 0$ denote the flow of vehicle type j on path p. Then a routing f is feasible if $\sum_{p \in \mathcal{P}_r} f_p^{j_r} = \beta_r$ for commodities $r \in [R]$. Since there is a one-to-one correspondence between path flows and edge flows, a characterization of feasible routings can be applied to edge routings z as well. Thus, a feasible routing z refers to a routing which satisfies the commodity flow demands and graph constraints. We use $\mathcal{Z}$ to denote the set of feasible routings. We use $\mathcal{I}_p$ to denote the set of edges in path p.

We refer to the flow vector z as a *routing* or a *strategy*. We use $\mathcal{Z} \subseteq \mathbb{R}^{mn}_{\geq 0}$ to denote the set of feasible routings, meaning routings that route all flow demand from their origin nodes to their destination nodes while respecting conservation of flow in the network.

Let $c_i^j : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$ denote the delay experienced on road i by vehicle type j, which is a function of the flow of all vehicle types on that road. We assume that all vehicle

types experience the same delay on a road; we can then write a vector-valued function which outputs the delay experienced by each vehicle type on road i, $c_i : \mathbb{R}^m_{\geq 0} \to \mathbb{R}^m_{\geq 0}$, as

$$c_i(z_i) = \mathbf{1}_m c_i^1(z_i) \ ,$$

where $\mathbf{1}_m$ denotes the m-dimensional column vector of ones. We then define the delay function for the network, $c : \mathbb{R}^{mn}_{\geq 0} \to \mathbb{R}^{mn}_{\geq 0}$ as

$$c(z) := \begin{bmatrix} c_1(z_1)^T & c_2(z_2)^T & \dots & c_n(z_n)^T \end{bmatrix} \ .$$

The social cost, which is the *aggregate delay* experienced by all users of the network, is then

$$C(z) := \langle c(z), z \rangle \ .$$

A social planner then wishes to find the *socially optimal routing*, which is the feasible routing that minimizes the social cost, and therefore solves the following optimization:

$$\min_{z \in \mathcal{Z}} C(z) \ .$$

We are concerned with the social cost of user equilibria in this setting. This means routings for which no user type would wish to change their routing, where users experience the cost of a route as a sum of the latency of each road in their path. Specifically, we consider Wardrop Equilibrium, which can be specified in terms of path routing.

Definition 2.1. *A flow is at* Wardrop Equilibrium *if no user can decrease their cost by switching routes. Mathematically, a flow f is at Wardrop Equilibrium if*

$$\forall r \in [R],\ \forall p, p' \in \mathcal{P}_r,\ f_p^j > 0 \implies \sum_{i \in \mathcal{I}_p} c_i^j(z_i) \leq \sum_{i \in \mathcal{I}_{p'}} c_i^j(z_i) \ .$$

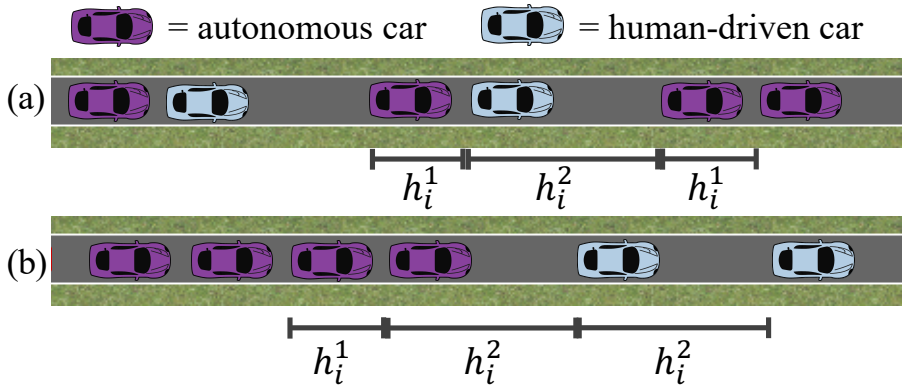


Figure 2.3: Capacity models 1 and 2. In capacity model 1 (top), autonomous cars can platoon behind any vehicle, and therefore take up length h_i^1 when traveling at the free-flow velocity. In capacity model 2 (bottom), autonomous vehicles can only platoon behind other autonomous vehicles; in that case they take up length h_i^1, but if following a human-driven vehicle, they take up length h_i^2. Human-driven vehicles always take up length h_i^2.

In words, for each commodity, if a path has positive flow, then no other path available to that commodity can have lower cost, where the cost is relative to the user type of the commodity. As mentioned above, in the setting without tolling, we consider the cost experienced by all users on a road to be equal. In the following sections we develop models for road capacity and road delay in order to construct the cost functions.

2.2.1 Capacity Models

We model the capacity of a road under two assumptions: 1) the spacing between two cars depends only on the vehicle type of the following vehicle and 2) the spacing between two vehicle types depends on the vehicle types of both vehicles. The first assumptions corresponds to a setting in which vehicles with autonomous technology do not platoon cooperatively, and instead treat all other vehicles interchangeably, and the second corresponds to the setting in which autonomous vehicles actively platoon coordinate acceleration and deceleration. We first develop the first capacity model.

Let d_i denote road length times the road's nominal velocity and let h_i^j denote the

nominal space taken up by vehicle type j, which is the sum of the length of the vehicle and the headway that vehicle maintains at free-flow velocity. We define the capacity of a road as the number of vehicles that can travel on a road at the road's nominal velocity. The capacity will be a function of the fraction of the total vehicles of each vehicle type and is calculated by taking the product of the length of the road and the free-flow velocity and dividing it by the average space taken up by a car on the road. We formalize this in the following proposition.

Proposition 2.2. *Assume that vehicles maintain a nominal spacing independent of the vehicle type they follow, and let h_i^j denote the nominal space occupied by vehicle of type j on road i. Then the capacity is*

$$\theta_i(z_i) = \frac{d_i \sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} h_i^j z_i^j} \ . \tag{2.3}$$

We justify this proposition by noting that under this assumption, the average space maintained by a vehicle, as the number of vehicles grows large, is a weighted combination of the spacing of each vehicle type, with weights determined by the fraction of total vehicles that are of that vehicle type. Note that this capacity model does not depend on the ordering of the vehicles.

We next consider the setting in which the spacing between a pair of vehicles depend on the types of both the vehicles. For this model we consider two vehicle types[3], the first one autonomous, which maintains a nominal spacing of h_i^1 when traveling behind other autonomous vehicles, and a nominal spacing of h_i^2 when traveling behind human-driven vehicles. We consider the second vehicle type to be human-driven vehicles, which maintain a nominal spacing of h_i^2 when traveling behind any vehicle type. We formalize

[3]Since the second capacity model only incorporates two vehicle types, any networks which include some roads with capacity model 1 and some with capacity model 2 will have two vehicle types. Networks which only include roads with capacity model 1 can have more than two vehicle types.

the capacity in the following proposition.

Proposition 2.3. *Consider a setting of two vehicle types on road i, where type 1 maintains a nominal spacing of h_i^1 when traveling behind other vehicles of type 1, and in all other cases, vehicles maintain a nominal spacing of h_i^2. Assume that vehicles are placed on a road as the result of a Bernoulli process. Then the capacity is*

$$\theta_i(z_i) = \frac{d_i}{(\frac{z_i^1}{z_i^1+z_i^2})^2 h_i^1 + (1 - (\frac{z_i^1}{z_i^1+z_i^2})^2) h_i^2} \ . \tag{2.4}$$

Proof. For this capacity model, we assume the vehicles are placed as the result of a Bernoulli process with parameter $\frac{z_i^1}{z_i^1+z_i^2}$. Consider M vehicles, each with length L, with s_q denoting the headway of vehicle q. Note that the front vehicle will have $s_q = 0$. The expected total space taken up is, due to linearity of expectation,

$$\begin{aligned}\mathbb{E}[\sum_{m=1}^{M} L + s_q] &= ML + \sum_{q=1}^{M-1} \mathbb{E}[s_q] \\ &= (M-1)((\frac{z_i^1}{z_i^1+z_i^2})^2 h_i^1 + (1 - (\frac{z_i^1}{z_i^1+z_i^2})^2) h_i^2) + L \ .\end{aligned}$$

Then, as the number of vehicles grows, the average space occupied by a vehicle approaches $(\frac{z_i^1}{z_i^1+z_i^2})^2 h_i^1 + (1 - (\frac{z_i^1}{z_i^1+z_i^2})^2) h_i^2$, yielding the above expression. □

Figure 2.3 provides an illustration of the technology assumptions. To make the meaning of nominal vehicle spacing more concrete, we offer one way of calculating spacing: let L denote vehicle length, $\tau_{h,i}$ and $\tau_{a,i}$ denote the reaction speed of human-driven and autonomous vehicles, respectively. Let v_i be the nominal speed on road i, which is likely the road's speed limit. Then we consider $h_i^1 = L + v_i \tau_{a,i}$ and $h_i^2 = L + v_i \tau_{h,i}$.[4]

[4]Our theoretical results hold even in the case that $h_i^1 \geq h_i^2$ on some roads and $h_{i'}^1 < h_{i'}^2$ on others.

2.2.2 Delay Model

We now propose a model, similar to Bureau of Public Roads (BPR) model [15,73,74], for the road delay incurred by mixed traffic resulting from the capacity models derived above.

Assumptions 2.1. *In the remainder of this chapter, we assume the following relationship between the flow of vehicles on a road and the delay on the road:*

$$c_i(z_i) = t_i(1 + \rho_i\Big(\frac{\sum_{j\in[m]} z_i^j}{\theta_i(z_i)}\Big)^{\sigma_i}) \ . \tag{2.5}$$

Here, t_i denotes the free-flow delay on road i, and ρ_i and σ_i are model parameters. Typical values for ρ_i and σ_i are 0.15 and 4, respectively [74]. However, our solution methodology is valid for any parameters such that $t_i \geq 0$, $\rho_i \geq 0$ and $\sigma_i \geq 1$.

Remark 2.1. *This model of delay function assumes that the density of vehicles on a road remains low enough that the vehicle flow does not enter the congested regime, in which delay* increases *as flow* decreases*. In the absence of this assumption the Price of Anarchy is trivially infinite.*

Remark 2.2. *This choice of cost functions implies that road delay is separable, meaning that the vehicles on one road do not affect those on another. We bound the Price of Anarchy and bicriteria for some nonseparable affine cost functions in another work [43].*

The class of cost functions we consider are not monotone, meaning they do not necessarily satisfy (2.2), but are elementwise monotone, defined below:

Definition 2.2. *A cost function* $c : R^{mn}_{\geq 0} \to R^{mn}_{\geq 0}$ *is* elementwise monotone *if it is non-decreasing in each of its arguments,* i.e. $\frac{\mathrm{d}c_i^j(z_i)}{\mathrm{d}z_{i'}^{j'}} \geq 0 \ \forall i, i' \in [n] \wedge \forall j, j' \in [m]$.

2.3 Price of Anarchy Analysis

In this section we bound the PoA and bicriteria of traffic networks with mixed autonomy. As established in the introduction of this chapter, the PoA is in general unbounded in traffic networks with mixed autonomy. However, we can establish a bound for the PoA by parameterizing it as described below.

Definition 2.3. *The* degree of asymmetry *on a road is the maximum ratio of road space utilized by a car of one type to a car of another type on the same road, while traveling at nominal velocity. The* maximum degree of asymmetry*, k, is the maximum of the above quantity over all roads in the network. Formally,*

$$k := \max_{i\in[n],j,j'\in[m]} h_i^j / h_i^{j'} \ . \tag{2.6}$$

Note that this quantity is always greater than or equal to 1. Also, we do not assume that one vehicle type affects delay more than another type on all roads. For example, autonomous vehicles may require shorter headways than human-driven vehicles on highways but longer headways on neighborhood roads to maintain safety for pedestrians.

Definition 2.4. *The* maximum polynomial degree*, denoted σ, for a road network with cost functions in the form* (2.5) *is the maximum degree of a polynomial denoting the cost on all roads in the network: $\sigma = \max_{i\in[n]} \sigma_i$.*

We use $\mathcal{C}_{k,\sigma}$ to denote the class of cost functions of the form (2.5), with maximum degree of asymmetry k and maximum polynomial degree σ, with cost functions using θ_i from either capacity mode 1 in (2.3) or capacity model 2 in (2.4). Let

$$\xi(\sigma) := \sigma(\sigma+1)^{-\frac{\sigma+1}{\sigma}} \ . \tag{2.7}$$

Note that for $\sigma \geq 1$, $\xi(\sigma) < 1$. With this, we present our first bound.

Theorem 2.1. *Consider a class of nonatomic congestion games with cost functions drawn from $\mathcal{C}_{k,\sigma}$. Let $\hat{z}$ be an equilibrium and z^* be a social optimum for this game. Then,*

$$C(\hat{z}) \leq \frac{k^\sigma}{1-\xi(\sigma)} C(z^*) \ .$$

Proof. Given any road cost function c (and social cost C) and equilibrium $\hat{z}$, we define an aggregate cost function ℓ (and social cost L) parameterized by the original equilibrium flow $\hat{z}$ and the cost function c. We also define a corresponding equilibrium flow $\hat{f}$, also parameterized by $\hat{z}$. This allows us to combine flow of all vehicle types on a road into one flow type in the aggregate function so we can bound the Price of Anarchy for the aggregate cost function. This means finding the relationship between the social cost at the equilibrium routing $\hat{f}$ and a socially optimal routing f^*, both relative to the aggregate cost function ℓ. We then find the relationship between the optimal routing for the aggregate cost function to that of the original cost function. Formally, the steps of the proof are:

$$\begin{aligned} C(\hat{z}) &= L(\hat{f}(\hat{z}); \hat{z}, c) && (2.8)\\ &\leq \frac{1}{1-\xi(\sigma)} L(f^*; \hat{z}, c) && (2.9)\\ &\leq \frac{1}{1-\xi(\sigma)} k^\sigma C(z^*) \ . && (2.10) \end{aligned}$$

We begin by introducing the tool with which we bound the PoA in (2.9). We then define ℓ and $\hat{f}$ such that (2.8) holds and show that $\hat{f}$ is an equilibrium for ℓ. We discuss the structure of the tool used to bound the PoA and provide an intuitive explanation of how the chosen structure of ℓ leads to a tighter PoA bound than an alternative choice. We

then provide lemmas corresponding to inequality (2.9), which bounds the PoA of this new cost function, and (2.10), which relates the social cost of optimal routing under ℓ to that of the original cost function, c. We defer proofs of the lemmas to the appendix.

We first introduce a general tool that we use for our results by extending the framework established by Correa *et. al.* [21], which relies on the Variational Inequality to bound the PoA. We use the following parameters:

$$\begin{aligned} \beta(c, z') &:= \max_{z \in \mathbb{R}^{mn}_{\geq 0}} \frac{\langle c(z') - c(z), z \rangle}{\langle c(z'), z' \rangle} \ , \\ \beta(\mathcal{C}) &:= \sup_{c \in \mathcal{C}, z' \in \mathcal{Z}} \beta(c, z') \ , \end{aligned} \tag{2.11}$$

where 0/0=0 by definition, and $\mathcal{C}$ is the class of network cost functions being considered. Then,

Lemma 2.1. *Let $\hat{z}$ be an equilibrium of a nonatomic congestion game with cost functions drawn from a class $\mathcal{C}$ of elementwise monotone cost functions.*

(a) If z^* is a social optimum for this game and $\beta(\mathcal{C}) < 1$, then

$$C(\hat{z}) \leq (1 - \beta(\mathcal{C}))^{-1} C(z^*) \ .$$

(b) If q^* is a social optimum for the same game with $1 + \beta(\mathcal{C})$ times as much flow demand of each type, then

$$C(\hat{z}) \leq C(q^*) \ .$$

The lemma and proof are nearly identical to that of Correa *et. al.* [21], extended to encompass nonmonotone, yet elementwise monotone, cost functions.

We now explain our choice of ℓ and $\hat{f}$ that yields (2.8) then provide an intuitive

explanation for this choice. Recall that we define

$$\hat{z} = \begin{bmatrix} \hat{z}_1^1 & \hat{z}_1^2 & \dots & \hat{z}_1^m & \hat{z}_2^1 & \hat{z}_2^2 & \dots & \hat{z}_n^m \end{bmatrix}^T .$$

We first define a flow aggregator function which combines the flow in $\hat{z}$ of each type on a road. We define $\hat{f}_i : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$, where

$$\hat{f}_i(\hat{z}_i) = \sum_{j \in [m]} \hat{z}_i^j \, . \tag{2.12}$$

We define a new cost function ℓ that is a mapping from flow vector (with one flow for each road) to road latencies, *i.e.* $\ell : \mathbb{R}^n_{\geq 0} \to \mathbb{R}^n_{\geq 0}$. We define ℓ so that it has the same road costs with flow $\hat{f}$ as c does with flow $\hat{z}$. Note, however, that c is a mapping from flows, with m flow types per road, to road latencies, again with each road represented m times ($c : \mathbb{R}^{mn}_{\geq 0} \to \mathbb{R}^{mn}_{\geq 0}$). However, ℓ represents each road once.

To formally define ℓ, we first define a latency function parameterized by another latency function and a flow. First define an intermediary function $v_i^{j'} : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that

$$v_i^{j'}(\hat{z}_i) = \begin{cases} 0 & j = 0 \\ \sum_{j \in [j']} \hat{z}_i^j & j' \in [m-1] \\ \infty & j = m \end{cases} \tag{2.13}$$

We formally define ℓ, which depends on the equilibrium flow being considered, $\hat{z}$. This cost function is defined for both capacity models. If road i has the first capacity

type, then

$$\ell_i(f_i;\hat{z}_i,c_i) := \begin{cases} t_i(1+\rho_i\big(\sum_{j\in[j'-1]} h_i^j \hat{z}_i^j + h_i^{j'}(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)\big)^{\sigma_i}) & \text{for } v_i^{j'-1} \le f_i \le v_i^{j'} \end{cases} . \tag{2.14}$$

The second capacity model only incorporates two vehicle types. For roads using capacity model 2, if $h_i^1 \le h_i^2$, ℓ_i is defined as follows.

$$\ell_i(f_i;\hat{z}_i,c_i) := \begin{cases} t_i(1+\rho_i(\frac{h_i^2 f_i}{d_i})^{\sigma_i}) & f_i \le \hat{z}_i^2 \\ t_i(1+\rho_i(\frac{h_i^2(f_i)^2 - (h_i^2-h_i^1)(f_i-\hat{z}_i^1)^2}{d_i f_i})^{\sigma_i}) & f_i > \hat{z}_i^2 \end{cases} . \tag{2.15}$$

If $h_i^1 > h_i^2$,

$$\ell_i(f_i;\hat{z}_i,c_i) := \begin{cases} t_i(1+\rho_i(\frac{h_i^1 f_i}{d_i})^{\sigma_i}) & f_i \le \hat{z}_i^1 \\ t_i(1+\rho_i(\frac{h_i^1(f_i)^2 - (h_i^1-h_i^2)(f_i-\hat{z}_i^2)^2}{d_i f_i})^{\sigma_i}) & f_i > \hat{z}_i^1 \end{cases} . \tag{2.16}$$

We defined the social cost for a network with cost function ℓ:

$$L(f;\hat{z},c) := \sum_{i\in[n]} f_i \ell_i(f_i;\hat{z}_i,c_i) . \tag{2.17}$$

Finally, we use $\mathcal{L}_\sigma$ to denote the set of functions L parameterized by cost function c with maximum polynomial degree σ.

With all this established, we relate equilibria for networks with cost function c and networks with cost function ℓ in the following proposition.

Proposition 2.4. *For a network with latency function c and Wardrop Equilibrium $\hat{z}$, and with $\hat{f}$, $v_i^{j'}$, ℓ, and L defined as in* (2.12), (2.13), (2.14), (2.15), (2.17),

1. $L(\hat{f}(\hat{z});\hat{z},c) = C(\hat{z})$, *and*
2. *if $\hat{z}$ is a Wardrop Equilibrium for c, then $\hat{f}$ is a Wardrop Equilibrium for ℓ.*

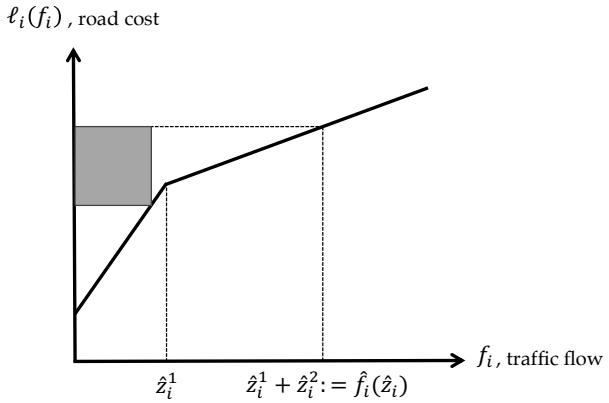


Figure 2.4: Illustration of the geometric interpretation of the parameter $\beta(\mathcal{L})$ where $\mathcal{L}$ represents the class of aggregate cost functions. Parameter $\beta(\mathcal{L})$ is an upper bound on the ratio between the size of the shaded rectangle and the dashed rectangle. This is an upper bound over all choices of $\ell_i \in \mathcal{L}$ and $\hat{z}_i$ with nonnegative elements.

In our proof, when bounding $\beta(\mathcal{L}_\sigma)$, we can bound (2.11) by only considering a single road i. Key in this is that without loss of generality, on this road, we assign vehicle type indices such that ℓ_i adds the vehicle types in decreasing order of how much they congest. This means the we assign vehicle type indices such that $h_i^1 \geq h_i^2 \geq \ldots \geq h_i^m$, so the most congesting vehicle type is added first, then the second-most congesting type, and so on.

To provide some intuition as to why we add the "costly" vehicle type first, consider the affine case with the first capacity model with two vehicle types. Correa *et. al.* give a geometric interpretation of the parameter $\beta(\mathcal{C})$ when cost are separable, meaning road latency only depends on one element of the flow vector. They show that for any cost function drawn from $\mathcal{C}$, $\beta(\mathcal{C})$ provides an upper bound on the ratio of the area of a rectangle above the cost function curve to the area of a rectangle enclosing it, where the enclosing rectangle has one corner at the origin. See Figure 2.4 for an illustration.

This interpretation provides the intuition that the more convex a function can be, the greater $\beta(\mathcal{C})$ can grow. Thus, to make our bound as tight as possible in our case, when considering $\beta(\mathcal{L})$, we add the costly vehicle type first. In the affine case with the first capacity model, this makes the class of cost functions concave. Then the element of this

class that maximizes the size of the interior rectangle relative to the exterior rectangle minimizes the concavity of the function by setting $\hat{z}_i^1 = 0$ or $\hat{z}_i^2 = 0$. Thus, the PoA bound does not depend on the degree of asymmetry. Though this exact interpretation does not apply for $\sigma > 1$ or for the second capacity model, the intuition is nonetheless useful.

With this intuition, we now present inequalities (2.9) and (2.10) as a lemmas, which we prove in the appendix.

Lemma 2.2. *Consider a nonatomic congestion game with road cost functions of the form* (2.14) *or* (2.15)*, parameterized by c with maximum polynomial degree σ. Then, for socially optimal routing f^* serving the same flow demand as $\hat{f}(\hat{z})$,*

$$L(\hat{f}(\hat{z}); \hat{z}, c) \leq \frac{1}{1 - \xi(\sigma)} L(f^*; \hat{z}, c)$$

where $\xi(\sigma) = \sigma(\sigma+1)^{-\frac{\sigma+1}{\sigma}}$.

Lemma 2.3. *Let c be a cost function from the class $\mathcal{C}_{k,\sigma}$. Consider L as defined in* (2.12)*,* (2.13)*,* (2.14)*,* (2.15)*,* (2.17)*, and let f^* be a socially optimal routing for L serving the same flow demand as $\hat{z}(\hat{f})$. Let z^* be a socially optimal routing for C, serving the same flow demand as $\hat{z}$. Then,*

$$L(f^*; \hat{z}, c) \leq k^\sigma C(z^*) \, .$$

We prove Lemma 2.2 by bounding $\beta(\mathcal{L})$ for the class of aggregate cost functions and applying Lemma 2.1. We analyze the structures of c and ℓ to prove Lemma 2.3. The proofs for these lemmas are provided in the appendix to this chapter. With these lemmas, the theorem is proved.

□

Remark 2.3. *In this proof outline above, L is parameterized by a specific equilibrium of c, denoted $\hat{z}$. Though c may have multiple equilibria, the proof construction applies for any equilibrium of c. Thus, for any $c \in \mathcal{C}_{k,\sigma}$, for each equilibrium of this cost function, this methodology can be applied, thereby bounding the cost of any equilibrium.*

Note that for $k = 1$ (*i.e.* no asymmetry), the PoA bound simplifies to those in [20,47]. If the cost functions are affine and there is no asymmetry, this reduces to the classic $\frac{4}{3}$ bound [19]. We characterize the tightness of this bound in the following corollary:

Corollary 2.1. *Given a maximum polynomial degree σ, the PoA bound is order-optimal with respect to the maximum degree of asymmetry k.*

We provide an example proving the corollary later in this section. When considering road networks with low asymmetry, we can establish another bound.

Theorem 2.2. *Consider a class of nonatomic congestion games with cost functions drawn from $\mathcal{C}_{k,\sigma}$. Let $\hat{z}$ be an equilibrium and z^* a social optimum for this game. If $k\xi(\sigma) < 1$, then*

$$C(\hat{z}) \leq \frac{1}{1 - k\xi(\sigma)} C(z^*) \, .$$

Proof. To prove this theorem, instead of going through an aggregate cost function we directly find $\beta(\mathcal{C})$ for our class of cost functions and apply Lemma 2.1. We do this in two two lemmas: we first find a relationship between the parameter $\beta(c, v)$ and the road capacity model $\theta_i(x_i, y_i)$, then we bound the resulting expression.

Lemma 2.4. *For cost functions of the form* (2.5)*, the parameter $\beta(\mathcal{C})$ is bounded by*

$$\beta(\mathcal{C}) \leq \max_{i \in [n], \hat{z}_i, z_i \in \mathbb{R}^m_{\geq 0}} \frac{\sum_{j \in [m]} z_i^j}{\sum_{j \in [m]} \hat{z}_i^j} \Big(1 - \big(\frac{\theta_i(q_i)(\sum_{j \in [m]} z_i^j)}{\theta_i(z_i)(\sum_{j \in [m]} \hat{z}_i^j)} \big)^{\sigma} \Big) \, .$$

Lemma 2.5. *For capacities of the forms* (2.3) *or* (2.4),

$$\max_{i\in[n],\hat{z}_i,z_i\in\mathbb{R}^m_{\geq 0}} \frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j}\Big(1-\big(\frac{\theta_i(\hat{z}_i)(\sum_{j\in[m]} z_i^j)}{\theta_i(z_i)(\sum_{j\in[m]} \hat{z}_i^j)}\big)^{\sigma}\Big) \leq k\xi(\sigma)\ .$$

These lemmas, together with Lemma 2.1, prove the theorem as well as Theorem 2.3 below.

□

Note that this bound may not necessarily be tighter in all regimes so our new PoA bound is $\min(\frac{k^\sigma}{1-\xi(\sigma)}, \frac{1}{1-k\xi(\sigma)})$. Though we cannot in closed form determine the region for which it is tighter, we can do so numerically. For example, for affine cost functions with $k=2$, $\frac{k^\sigma}{1-\xi(\sigma)} = \frac{8}{3}$ and $\frac{1}{1-k\xi(\sigma)} = 2$. Later in this section we show via example that the bound in Theorem 2.2 is tight in this case.

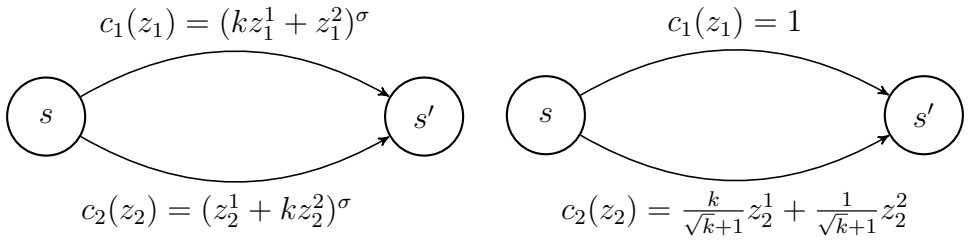
The method used for establishing Theorem 2.2 also gives a bound on the bicriteria, stated in the following theorem.

Theorem 2.3. *Consider a class of nonatomic congestion games with cost functions drawn from $\mathcal{C}_{k,\sigma}$. Let $\hat{z}$ be an equilibrium for this game. If q^* is a social optimum for the same game with $1+k\xi(\sigma)$ times as much flow demand of each type, then*

$$C(\hat{z}) \leq C(q^*)\ .$$

As an example, if road delays are described by polynomials of degree 4, and the maximum asymmetry between the spacing of platooned and nonplatooned vehicles is 3, then the the cost of selfishly routing vehicles will be less than optimally routing $1+3\,\xi(4) \approx 2.61$ times as much vehicle flow of each type.

In the theoretical results above, we only considered the Price of Anarchy, which depends on the worst-case equilibria, in the worst-case over all network topologies and



(a) Network from Example 2.1 with two sided asymmetry. One unit of vehicle type 1 and one unit of vehicle type 2 cross from node s to s'.

(b) Network from Example 2.2 with one sided asymmetry. $\frac{1}{\sqrt{k}}$ units of vehicle type 1 and one unit of vehicle type 2 cross from node s to s'.

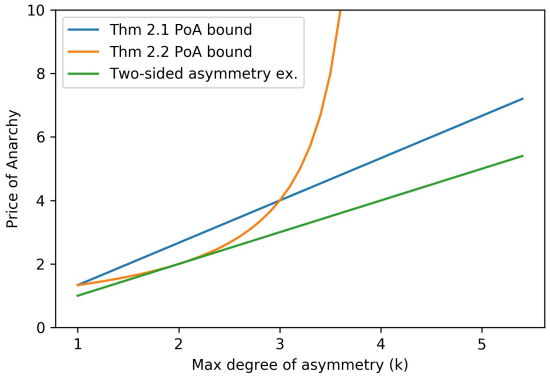
Figure 2.5: Examples showing the tightness of the Price of Anarchy bounds.

cost functions in the considered class of cost functions. In this setting we found that the PoA scales with the maximum degree of asymmetry as k^σ. However, in mixed autonomy, the equilibrium is not necessarily unique, meaning that there may be better equilibria with lower cost, and it may not be a fair assumption to assume that the network will settle on the worst-case equilibrium. It is therefore valuable to understand the worst-case gap between social cost under optimal routing and *best*-case equilibrium, meaning the equilibrium that minimizes the social cost. This quantity is known as the Price of Stability.

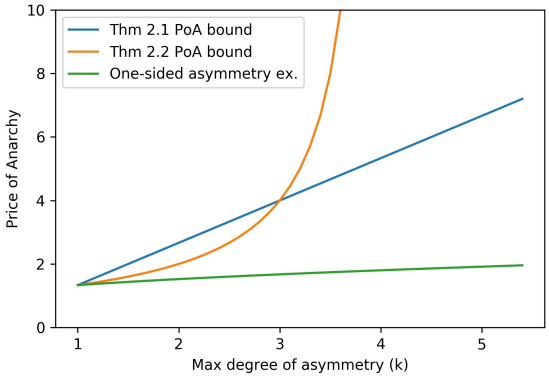
Proposition 2.5. *The Price of Stability for the class of nonatomic congestion games with cost functions drawn from $\mathcal{C}_{k,\sigma}$ is lower bounded by $1 + \frac{k}{2\sqrt{k}+1}$.*

This proposition is proved in Example 2.2. This result shows that even assuming that the most socially optimal equilibrium arises in each network, in the worst case, the ratio between the equilibrium social cost and the optimal social cost can be lower bounded by an expression which scales with $\sqrt{k}/2$.

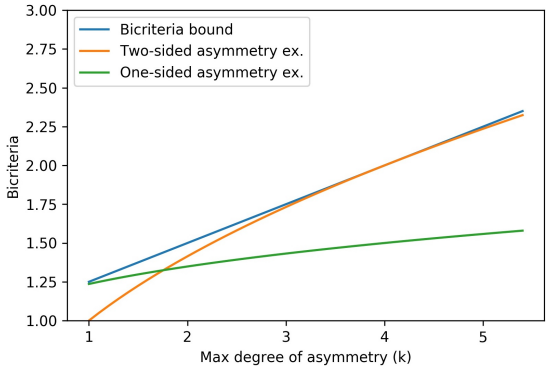
Next we provide examples that give a lower bound on the PoA for this class of networks and serve to illustrate the tightness of the bounds. The examples are shown in Fig. 2.5 and the comparison of the PoA and bicriteria are shown in Fig. 2.6. We discuss



(a) Comparison of the PoA of Example 2.1 with the upper bound, with $\sigma = 1$.



(b) Comparison of the PoA of Example 2.2 with the upper bound, with $\sigma = 1$.



(c) Bicriteria of Examples 2.1 and 2.2 (with $\sigma = 1$), compared with the bicriteria bound.

Figure 2.6: Tightness of theoretical bounds.

notions of one-sided and two-sided asymmetry: a network has one sided asymmetry if $h_i^1 \leq h_i^2 \ \forall i \in [n]$ (human-driven cars always contribute more to road delay than autonomous cars) or $h_i^1 \geq h_i^2 \ \forall i \in [n]$ (human-driven cars always contribute less to road delay than autonomous cars); otherwise the network has two-sided asymmetry. We provide two example networks (Fig. 2.5), one with two-sided asymmetry and one with one-sided asymmetry. We compare the PoA and bicriteria in those networks to the upper bounds established earlier.

Through the first example we prove Corollary 2.1. In this example, we consider two roads, one of which is well-suited for autonomous vehicles (such as a highway) and the other is well-suited for human-driven vehicles (such as an urban road).

Example 2.1. *Consider the network of parallel roads in Fig. 2.5 (a), where one unit of vehicle type* 1 *and one unit of vehicle type* 2 *wish to cross from node* s *to* s'*. The roads have costs* $c_1(z_2) = (kz_1^1 + z_1^2)^\sigma$ *and* $c_2(z_2) = (z_2^1 + kz_2^2)^\sigma$*, where* $k \geq 1$*. In worst-case equilibrium, all flow of type* 1 *are on the top road and all flow of type* 2 *are on the bottom road. In the best case, these routing are reversed. This yields a PoA of* k^σ*. To find the bicriteria, we calculate how much traffic, optimally routed, yields a cost equal to* $2k^\sigma$*, the cost of the original traffic volume at worst-case equilibrium. We find that* $k^{\frac{\sigma}{\sigma+1}}$ *as much traffic, optimally routed, yields this same cost.*

We now analyze the setting in which autonomous vehicles *always* increase the capacity of a road. In this case, the tightness of our bound (which holds for two-sided asymmetry as well) remains open.

Example 2.2. *Consider the network of parallel roads in Fig. 2.5 (b), where* $\frac{1}{\sqrt{k}}$ *unit of vehicle type* 1 *and one unit of vehicle type* 2 *wish to cross from node* s *to* s'*. The roads have costs* $c_1(z_1) = 1$ *and* $c_2(z_1) = \frac{k}{\sqrt{k}+1} z_2^1 + \frac{1}{\sqrt{k}+1} z_2^2$*. At equilibrium, all vehicles take the bottom road; optimally routed, flow of type* 1 *takes the top road and flow of type* 2 *take*

the bottom. This yields a PoA of $1 + \frac{k}{2\sqrt{k}+1}$. *Calculations similar to that in Example 2.1 yield a bicriteria of* $\frac{(-1+\sqrt{1+4\sqrt{k}})(1+\sqrt{k})}{2\sqrt{k}}$.

For affine cost functions, $\sigma = 1$, so $\xi = 1/4$. The PoA bound is then $\min(\frac{4}{4-k}, \frac{4}{3}k)$ and the bicriteria bound is $1 + k/4$. With $\sigma = 1$, the first example has PoA k and bicriteria $\sqrt{k}$, and the second example has PoA of order $\sqrt{k}$ and bicriteria of order $k^{1/4}$. Accordingly, the first example shows that the PoA bound is tight for $k = 2$ and the bicriteria bound is tight for $k = 4$. Note that a realistic range for k is between 1 and 4.

Further, for affine cost functions, the bound in Theorem 2.2 is tighter than that of Theorem 2.1 when the degree of asymmetry is low. However, the bound in Theorem 2.1 scales much better for high degrees of asymmetry. This effect is accentuated for cost functions that have higher order polynomials – the regime for which the bound in Theorem 2.2 is tighter shrinks as the maximum polynomial degree grows.

As stated in Corollary 2.1, our bound is order-optimal with respect to the maximum degree of asymmetry, k. Comparing the bound $(\frac{k^\sigma}{1-\xi(\sigma)})$ with the PoA in Example 2.1 (k^σ) shows that for a fixed σ, the PoA upper bound is within a constant factor of the lower bound, implying that the upper bound is order-optimal in k.

It is also worth noting that under the construction used in Theorem 2.2, the bicriteria is related to the PoA through the quantity $\beta(\mathcal{C})$ [21]. Observe that Example 2.1 provides a bicriteria of 2 for $k = 4$, implying $\beta(\mathcal{C}_4) \geq 1$. Since the PoA is greater than or equal to $\frac{1}{1-\beta(\mathcal{C})}$, this mechanism cannot bound the PoA for $k \geq 4$. This leads us to rely on the mechanism developed for Theorem 2.1 for networks with large asymmetry.

2.4 Conclusion

In this chapter[5] we presented a framework, similar to a congestion game, for considering traffic networks with mixed autonomy. To do so we presented two models for the capacity of roads with mixed autonomy, each corresponding to an assumption about the technological capabilities of autonomous vehicles, and we defined a class of road latency functions that incorporates these capacity models. Using this framework, we developed two methods of bounding the Price of Anarchy and show that these bounds depend on the degree of the polynomial describing latency and the difference in the degree to which platooned and nonplatooned vehicles occupy space on a road. In addition we presented a bound on the bicriteria, another measure of inefficiency due to selfish routing. We presented examples showing these bounds are tight in some cases and recover classical bounds when human-driven and autonomous vehicles affect congestion the same way. Moreover, we showed that our PoA bound is order-optimal with respect to the degree to which vehicle types differently affect latency.

Some directions for future work are as follows. The capacity models presented assume that vehicle types are determined as a result of a Bernoulli process; a more general capacity model could incorporate autonomous vehicles that actively rearrange themselves as to form platoons. Further, autonomous vehicles can affect vehicle flow in ways not limited to platooning. In addition, our proposed latency function considers only the effect of a vehicle on the road upon which it travels; a more general latency function would consider interaction between roads. Finally, the PoA bound is not shown to be tight but is order-optimal in the degree of asymmetry k, and a future work could aim to close this gap. Nonetheless, this chapter presents a framework that can be used in the future for studying traffic networks in mixed autonomy.

[5]This work was supported by California Department of Transportation UC Connect grant number 65A0529, NSF grant 1749357, and UC Office of the President grant LFR-18-548175.

2.5 Appendix

2.5.1 Proof of Lemma 2.1

We prove Lemma 2.1, which offers a useful tool for bounding the PoA. To prove part (a),

$$\begin{aligned}\langle c(z'), z\rangle &= \langle c(z), z\rangle + \langle c(z') - c(z), z\rangle \\ &\leq \langle c(z), z\rangle + \beta(c, z')\langle c(z'), z'\rangle \\ &\leq C(z) + \beta(\mathcal{C})C(z') \end{aligned} \tag{2.18}$$

and replacing z' with $\hat{z}$, by the Variational Inequality, $C(\hat{z}) \leq \langle c(\hat{z}), z\rangle$ for any feasible routing z. Completing the proof requires that $\beta(\mathcal{C}) \leq 1$, then replacing the generic z with z^*.

To prove part (b), element-wise monotonicity implies the feasibility of $(1+\beta(\mathcal{C}))^{-1}q^*$, which routes the same volume of traffic as $\hat{z}$. Using (2.1),

$$\langle c(\hat{z}), \hat{z}\rangle \leq \langle c(\hat{z}), (1+\beta(\mathcal{C}))^{-1}q^*\rangle \ . \tag{2.19}$$

Then,

$$C(\hat{z}) = (1+\beta(\mathcal{C}))\langle c(\hat{z}), \hat{z}\rangle - \beta(\mathcal{C})\langle c(\hat{z}), \hat{z}\rangle \tag{2.20}$$

$$\leq (1+\beta(\mathcal{C}))\langle c(\hat{z}), (1+\beta(\mathcal{C}))^{-1}q^*\rangle - \beta(\mathcal{C})\langle c(\hat{z}), \hat{z}\rangle \tag{2.21}$$

$$\leq C(q^*) \ , \tag{2.22}$$

where (2.21) uses (2.19) and (2.22) uses (2.18). □

2.5.2 Proof of Lemma 2.2

We prove Lemma 2.2, an intermediary step for bounding the PoA. We use Lemma 2.1 and bound $\beta(\mathcal{L}_\sigma)$, where $\mathcal{L}_\sigma$ denotes the set of aggregate cost functions with maximum polynomial degree σ.

$$\begin{aligned}\beta(\mathcal{L}_\sigma) &:= \sup_{c\in\mathcal{C}_{\cdot,\sigma},\hat{z}\in\mathcal{Z}} \max_{f\in\mathbb{R}^n_{\geq 0}} \frac{\langle \ell(\hat{f}(\hat{z});\hat{z},c) - \ell(f;\hat{z},c), f\rangle}{\langle \ell(\hat{f}(\hat{z});\hat{z},c), \hat{f}(\hat{z})\rangle} \\ &\leq \sup_{c\in\mathcal{C}_{\cdot,\sigma},i\in[n],\hat{z}\in\mathcal{Z}} \max_{f_i\in\mathbb{R}_{\geq 0}} \frac{f_i(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i) - \ell(f_i;\hat{z}_i,c_i))}{\hat{f}_i(\hat{z}_i)(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i))},\end{aligned}$$

where the inequality results from all terms in the denominator being nonnegative. Note that since, by definition, $\beta(\mathcal{L}_\sigma) \geq 1$, we know that for this term to be maximized, $f_i \leq \hat{f}_i(\hat{z}_i)$ – otherwise $\beta(\mathcal{L}_\sigma)$ would be negative, since ℓ_i is increasing in f_i. Next we bound this term for each capacity model. We first bound this term when road i uses capacity model 1. Assume, without loss of generality, that the vehicle types are assigned

such that on road i, $h_i^1 \geq h_i^2 \geq \ldots \geq h_i^m > 0$. Plugging in the definition of ℓ_i in this case,

$$\begin{aligned}
\beta(\mathcal{L}_\sigma) &\leq \sup_{c\in\mathcal{C}_{\cdot,\sigma}, i\in[n], \hat{z}\in\mathcal{Z}} \max_{f_i\in\mathbb{R}_{\geq 0}} \frac{f_i(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i) - \ell_i(f_i;\hat{z}_i,c_i))}{\hat{f}_i(\hat{z}_i)(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i))} \\
&= \sup_{\substack{h_i^1\geq h_i^2\geq\ldots\geq h_i^m>0,\\ j'\in[m],\\ \hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0},\\ 1\leq\sigma_i\leq\sigma,\\ t_i\in\mathbb{R}_{\geq 0},\\ \rho_i\in\mathbb{R}_{\geq 0},\\ d_i\in\mathbb{R}_{>0}}} \frac{f_i\frac{\rho_i}{d_i^\sigma}\Big(\big(\sum_{j\in[m]} h_i^j\hat{z}_i^j\big)^{\sigma_i} - \big(\sum_{j\in[j'-1]} h_i^j\hat{z}_i^j + h_i^{j'}(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)\big)^{\sigma_i}\Big)}{(\sum_{j\in[m]}\hat{z}_i^j)t_i\big(1+\frac{\rho_i}{d_i^\sigma}(\sum_{j\in[m]} h_i^j\hat{z}_i^j)^{\sigma_i}\big)} \\
&\leq \sup_{\substack{h_i^1\geq h_i^2\geq\ldots\geq h_i^m>0,\\ j'\in[m],\\ \hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0},\\ 1\leq\sigma_i\leq\sigma}} \frac{f_i\Big(\big(\sum_{j\in[m]} h_i^j\hat{z}_i^j\big)^{\sigma_i} - \big(\sum_{j\in[j'-1]} h_i^j\hat{z}_i^j + h_i^{j'}(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)\big)^{\sigma_i}\Big)}{(\sum_{j\in[m]}\hat{z}_i^j)(\sum_{j\in[m]} h_i^j\hat{z}_i^j)^{\sigma_i}} \qquad (2.23) \\
&= \sup_{\substack{h_i^1\geq h_i^2\geq\ldots\geq h_i^m>0,\\ j'\in[m],\\ \hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0},\\ 1\leq\sigma_i\leq\sigma}} \frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big[1 - \Big(\frac{\sum_{j\in[j'-1]} h_i^j\hat{z}_i^j + h_i^{j'}(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)}{\sum_{j\in[m]} h_i^j\hat{z}_i^j}\Big)^{\sigma_i}\Big] \\
&\leq \sup_{\substack{h_i^1\geq h_i^2\geq\ldots\geq h_i^m>0,\\ j'\in[m],\\ \hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big[1 - \Big(\frac{\sum_{j\in[j'-1]} h_i^j\hat{z}_i^j + h_i^{j'}(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)}{\sum_{j\in[m]} h_i^j\hat{z}_i^j}\Big)^{\sigma}\Big] \qquad (2.24) \\
&\leq \sup_{\substack{h_i^1\geq h_i^2\geq\ldots\geq h_i^{j'}>0,\\ j'\in[m],\\ \hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big[1 - \Big(\frac{\sum_{j\in[j'-1]} h_i^j\hat{z}_i^j + h_i^{j'}(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)}{\sum_{j\in[j'-1]} h_i^j\hat{z}_i^j + h_i^{j'}\sum_{j'}^{m}\hat{z}_i^j}\Big)^{\sigma}\Big] \qquad (2.25) \\
&\leq \sup_{\substack{h_i^{j'}>0,\\ j'\in[m],\\ \hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big[1 - \Big(\frac{h_i^{j'}\sum_{j\in[j'-1]}\hat{z}_i^j + h_i^{j'}(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)}{h_i^{j'}\sum_{j\in[j'-1]}\hat{z}_i^j + h_i^{j'}\sum_{j'}^{m}\hat{z}_i^j}\Big)^{\sigma}\Big] \qquad (2.26) \\
&= \sup_{\substack{\hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big[1 - \Big(\frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big)^{\sigma}\Big] .
\end{aligned}$$

In the above, (2.23) follows from the numerator, denominator, and t_i being nonnegative, (2.24) results from the whole expression being nonnegative and therefore the second term in the bracket being less than or equal to 1, and therefore the bracketed expression is maximized when σ_i is the maximum allowable value. Since we assumed that $h_i^1 \geq h_i^2 \geq \ldots \geq h_i^m > 0$, (2.25) stems from $\sum_{j\in[m]} h_i^j \hat{z}_i^j$ being maximized when the h_i^j terms are larger, and (2.26) follows from the fact that the second term in the brackets is minimized when a term common to the numerator and denominator is minimized due to the structure of the expression.

Then we can show that the final expression above is equal to the following.

$$\begin{aligned}
\sup_{\substack{\hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big[1-\Big(\frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big)^{\sigma}\Big] &= \sup_{\substack{\hat{z}_i\in\mathbb{R}^m_{\geq 0},\\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}-\Big(\frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}\Big)^{\sigma+1} \\
&= \frac{1}{(\sigma+1)^{1/\sigma}}-\frac{1}{(\sigma+1)^{(\sigma+1)/\sigma}} \\
&= \sigma(\sigma+1)^{-\frac{\sigma+1}{\sigma}} \\
&:= \xi(\sigma)\ .
\end{aligned}$$

The logic above follows from the overall expression being concave with respect to $\frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j}$, and therefore maximized when $\frac{f_i}{\sum_{j\in[m]}\hat{z}_i^j} = (\sigma+1)^{-\frac{1}{\sigma}}$. Plugging in this value yields that $\beta(\mathcal{L}_\sigma)$ is upper bounded by $\xi(\sigma)$.

We next show that the expression has the same upper bound when road i uses capacity model 2. We bound the expression for $h_i^1 \leq h_i^2$, but the same method works for when $h_i^1 > h_i^2$. As we established above, we only need to consider the case in which $f_i \leq \hat{f}_i(\hat{z}_i)$. We then bound the expression for two cases: $0 \leq f_i \leq \hat{z}_i^2$ and $\hat{z}_i^2 < f_i \leq \hat{z}_i^1 + \hat{z}_i^2$. For the

first case,

$$\begin{aligned}
\frac{f_i(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i)-\ell_i(f_i;\hat{z}_i,c_i))}{\hat{f}_i(\hat{z}_i)(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i))} &= \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}\Big(1-\frac{1+\rho_i(\frac{h_i^2 f_i}{d_i})^{\sigma_i}}{1+\rho_i(\frac{(\hat{z}_i^1)^2h_i^1+\hat{z}_i^2(\hat{z}_i^2+\hat{z}_i^1)h_i^2}{d_i(\hat{z}_i^1+\hat{z}_i^2)})^{\sigma_i}}\Big) \\
&\le \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}\Big(1-\frac{\frac{h_i^2 f_i}{d_i}}{\frac{(\hat{z}_i^1)^2h_i^1+\hat{z}_i^2(\hat{z}_i^2+\hat{z}_i^1)h_i^2}{d_i(\hat{z}_i^1+\hat{z}_i^2)}})^{\sigma_i}\Big) \qquad (2.27)\\
&= \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}\Big(1-\big(\frac{h_i^2 f_i(\hat{z}_i^1+\hat{z}_i^2)}{(\hat{z}_i^1)^2h_i^1+\hat{z}_i^2(\hat{z}_i^2+\hat{z}_i^1)h_i^2}\big)^{\sigma_i}\Big) \\
&\le \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}\Big(1-\big(\frac{h_i^2 f_i(\hat{z}_i^1+\hat{z}_i^2)}{(\hat{z}_i^1)^2h_i^2+\hat{z}_i^2(\hat{z}_i^2+\hat{z}_i^1)h_i^2}\big)^{\sigma_i}\Big) \qquad (2.28)\\
&= \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}\Big(1-\big(\frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}\big)^{\sigma_i}\Big) \\
&\le \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}\Big(1-\big(\frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}\big)^{\sigma}\Big) \qquad (2.29)\\
&\le \xi(\sigma)\ . \qquad (2.30)
\end{aligned}$$

In the above, (2.27) follows from $\ell_i(f_i;\hat{z}_i,c_i) \le \ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i)$, (2.28) follows from $h_i^1 \le h_i^2$, (2.29) follows from $\sigma_i \le \sigma$ and the expression it applies to being less than or equal to 1, and (2.30) follows from the same logic as the proof for capacity model 1.

For the case that $\hat{z}_i^2 < f_i \leq \hat{z}_i^1 + \hat{z}_i^2$,

$$\begin{aligned}
\frac{f_i(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i)-\ell_i(f_i;\hat{z}_i,c_i))}{\hat{f}_i(\hat{z}_i)(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i))} &= \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}(1-\frac{1+\rho_i(\frac{h_i^2(f_i)^2-(h_i^2-h_i^1)(f_i-\hat{z}_i^1)^2}{d_i f_i})^{\sigma_i}}{1+\rho_i(\frac{(\hat{z}_i^1+\hat{z}_i^2)^2h_i^2-(h_i^2-h_i^1)(\hat{z}_i^1)^2}{d_i(\hat{z}_i^1+\hat{z}_i^2)})^{\sigma_i}}) \\
&\leq \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}(1-(\frac{\frac{h_i^2(f_i)^2-(h_i^2-h_i^1)(f_i-\hat{z}_i^1)^2}{d_i f_i}}{\frac{(\hat{z}_i^1+\hat{z}_i^2)^2h_i^2-(h_i^2-h_i^1)(\hat{z}_i^1)^2}{d_i(\hat{z}_i^1+\hat{z}_i^2)}})^{\sigma_i}) \qquad (2.31) \\
&= \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}(1-((\frac{\hat{z}_i^1+\hat{z}_i^2}{f_i})\frac{(h_i^2(f_i)^2-(h_i^2-h_i^1)(f_i-\hat{z}_i^1)^2)}{(\hat{z}_i^1+\hat{z}_i^2)^2h_i^2-(h_i^2-h_i^1)(\hat{z}_i^1)^2})^{\sigma_i}) \\
&\leq \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}(1-((\frac{\hat{z}_i^1+\hat{z}_i^2}{f_i})\frac{(f_i)^2}{(\hat{z}_i^1+\hat{z}_i^2)^2})^{\sigma_i}) \qquad (2.32) \\
&= \frac{f_i}{\hat{z}_i^1+\hat{z}_i^2}(1-(\frac{f_i}{\hat{z}_i^1+\hat{z}_i^2})^{\sigma_i}) \\
&\leq \xi(\sigma)\,. \qquad (2.33)
\end{aligned}$$

Here, (2.31) follows from $\ell_i(f_i;\hat{z}_i,c_i) \leq \ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i)$, (2.32) follows from $f_i \leq \hat{z}_i^1+\hat{z}_i^2$, and (2.33) follows the same logic as (2.29) and (2.30) above. With all these cases, the lemma is proven. □

2.5.3 Proof of Lemma 2.3

We prove the final intermediary step, bounding the difference in social cost between the optimal routing for our intermediary latency functions and the optimal routing for our original latency function.

First define $w_i : \mathbb{R}^m_{\geq 0} \times \mathcal{C} \to \mathbb{R}^m_{\geq 0}$ such that

$$[w_i(z_i;c_i)]_{j'} = \begin{cases} \sum_{j\in[m]} z_i^j & \text{if } j' = \operatorname{argmin}_{j\in[m]} h_i^j \\ 0 & \text{otherwise} \end{cases}.$$

In words, w_i puts all the flow in z_i as the vehicle type that congests the least according

to cost function c_i. Then considering capacity model 1, for f^* as an optimal routing with respect to $L(f;\hat{z},c)$ and z^* as an optimal routing with respect to C,

$$\begin{aligned}
L(f^*;\hat{z},c) &= \sum_{i\in[n]} f_i^*\ell_i(f_i^*;\hat{z}_i,c_i) \\
&\leq \sum_{i\in[n]}\sum_{j\in[m]} z^{*j}_i\ell_i(\sum_{j\in[m]} z^{*j}_i;\hat{z}_i,c_i) && (2.34)\\
&\leq \sum_{i\in[n]}\sum_{j\in[m]} z^{*j}_i t_i(1+\rho_i(\frac{\max_{j\in[m]} h_i^j \sum_{j\in[m]} z^{*j}_i}{d_i})^\sigma_i && (2.35)\\
&\leq \sum_{i\in[n]}\sum_{j\in[m]} z^{*j}_i k^\sigma t_i(1+\rho_i(\frac{\min_{j\in[m]} h_i^j \sum_{j\in[m]} z^{*j}_i}{d_i})^\sigma_i && (2.36)\\
&= k^\sigma \sum_{i\in[n]}\sum_{j\in[m]} z^{*j}_i c_i(w_i(z_i^*;c_i)) \\
&\leq k^\sigma \sum_{i\in[n]}\sum_{j\in[m]} z^{*j}_i c_i(z_i^*) && (2.37)\\
&= k^\sigma C(z^*)\ . && (2.38)
\end{aligned}$$

In the above, (2.34) stems from f^* as a minimizer of the cost with respect to L, (2.35) follows from the definitions of ℓ_i and c_i, (2.36) follows from the definition of k and σ, and (2.37) from the construction of w_i.

Thus, $L(f^*;\hat{z},c) \leq k^\sigma C(z^*)$, proving the lemma. □

2.5.4 Proof of Lemma 2.4

This Lemma and the following one together prove Theorem 2.2. Using (2.11),

$$\begin{aligned}
\beta(c,\hat{z}) &= \max_{z\in\mathbb{R}^{mn}_{\geq 0}} \frac{\sum_{i\in[n]} t_i\rho_i[(\frac{\sum_{j\in[m]}\hat{z}_i^j}{\theta_i(\hat{z}_i)})^{\sigma_i} - (\frac{\sum_{j\in[m]}z_i^j}{\theta_i(z_i)})^{\sigma_i}](\sum_{j\in[m]} z_i^j)}{\sum_{i\in[n]} t_i[1+\rho_i(\frac{\sum_{j\in[m]}\hat{z}_i^j}{\theta_i(\hat{z}_i)})^{\sigma_i}](\sum_{j\in[m]}\hat{z}_i^j)} \\
&\leq \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \frac{\rho_i[(\frac{\sum_{j\in[m]}\hat{z}_i^j}{\theta_i(\hat{z}_i)})^{\sigma_i} - (\frac{\sum_{j\in[m]}z_i^j}{\theta_i(z_i)})^{\sigma_i}](\sum_{j\in[m]} z_i^j)}{[1+\rho_i(\frac{\sum_{j\in[m]}\hat{z}_i^j}{\theta_i(\hat{z}_i)})^{\sigma_i}](\sum_{j\in[m]}\hat{z}_i^j)} \qquad (2.39) \\
&\leq \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \frac{[(\frac{\sum_{j\in[m]}\hat{z}_i^j}{\theta_i(\hat{z}_i)})^{\sigma_i} - (\frac{\sum_{j\in[m]}z_i^j}{\theta_i(z_i)})^{\sigma_i}](\sum_{j\in[m]} z_i^j)}{(\frac{\sum_{j\in[m]}\hat{z}_i^j}{\theta_i(\hat{z}_i)})^{\sigma_i}(\sum_{j\in[m]}\hat{z}_i^j)} \\
&= \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]}\hat{z}_i^j}(1-(\frac{\theta_i(\hat{z}_i)(\sum_{j\in[m]} z_i^j)}{\theta_i(z_i)(\sum_{j\in[m]}\hat{z}_i^j)})^{\sigma_i}) \\
&\leq \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]}\hat{z}_i^j}(1-(\frac{\theta_i(\hat{z}_i)(\sum_{j\in[m]} z_i^j)}{\theta_i(z_i)(\sum_{j\in[m]}\hat{z}_i^j)})^{\sigma}) \ , \qquad (2.40)
\end{aligned}$$

where the terms of the denominator being nonnegative imply (2.39), since a term in the summation in the numerator that is negative does not need to be accounted for in the upper bound. Then, $\beta(c,q)\geq 0$ implies (2.40), allowing us to consider only the maximum allowable degree of polynomial. □

2.5.5 Proof of Lemma 2.5

For capacity model 1, without loss of generality, let us again assume that on road i that $h_i^1 \geq h_i^2 \geq \ldots \geq h_i^m > 0$. Then,

$$\begin{aligned}
\beta(c, \hat{z}) &\leq \max_{i\in[n], z_i \in \mathbb{R}^m_{\geq 0}} \frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j}\left(1 - \left(\frac{\theta_i(\hat{z}_i)(\sum_{j\in[m]} z_i^j)}{\theta_i(z_i)(\sum_{j\in[m]} \hat{z}_i^j)}\right)^\sigma\right) \\
&= \max_{i\in[n], z_i \in \mathbb{R}^m_{\geq 0}} \frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j}\left(1 - \left(\frac{\sum_{j\in[m]} h_i^j z_i^j}{\sum_{j\in[m]} h_i^j \hat{z}_i^j}\right)^\sigma\right) \\
&\leq \max_{i\in[n], z_i \in \mathbb{R}^m_{\geq 0}} \frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j}\left(1 - \left(\frac{h_i^m \sum_{j\in[m]} z_i^j}{h_i^1 \sum_{j\in[m]} \hat{z}_i^j}\right)^\sigma\right) \\
&\leq \max_{i\in[n], z_i \in \mathbb{R}^m_{\geq 0}} \frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j}\left(1 - \frac{1}{k^\sigma}\left(\frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j}\right)^\sigma\right) . \qquad (2.41)
\end{aligned}$$

In (2.41) we use the Definition 2.3 of the maximum degree of asymmetry. We now note that the above expression is concave with respect to $\frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j}$, and the expression is maximized when $\frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j} = k(\sigma+1)^{-1/\sigma}$. With this, we find

$$\beta(\mathcal{C}_{k,\sigma}) \leq k\sigma(\sigma+1)^{-\frac{\sigma+1}{\sigma}} = k\xi(\sigma) , \qquad (2.42)$$

proving the lemma for capacity model 1.

For capacity model 2, let us assume that $h_i^1 \geq h_i^2$, though the same steps follow if the

opposite is true, with only a mild alteration.

$$
\begin{aligned}
\beta(c,\hat{z}) &\leq \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \frac{\sum_{j\in[m]} z_i^j}{\sum_{j\in[m]} \hat{z}_i^j}\left(1-\left(\frac{\theta_i(\hat{z}_i)(\sum_{j\in[m]} z_i^j)}{\theta_i(z_i)(\sum_{j\in[m]} \hat{z}_i^j)}\right)^\sigma\right) \\
&= \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \left(\frac{z_i^1+z_i^2}{\hat{z}_i^1+\hat{z}_i^2}\right)\left(1-\left(\frac{\hat{z}_i^1+\hat{z}_i^2}{z_i^1+z_i^2}\right)\frac{(z_i^1)^2h_i^1+((z_i^1+z_i^2)^2-(z_i^1)^2)h_i^2}{(\hat{z}_i^1)^2h_i^1+((\hat{z}_i^1+\hat{z}_i^2)^2-(\hat{z}_i^1)^2)h_i^2}\right)^\sigma\right) \\
&\leq \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \left(\frac{z_i^1+z_i^2}{\hat{z}_i^1+\hat{z}_i^2}\right)\left(1-\left(\frac{\hat{z}_i^1+\hat{z}_i^2}{z_i^1+z_i^2}\right)\frac{(z_i^1)^2h_i^2+((z_i^1+z_i^2)^2-(z_i^1)^2)h_i^2}{(\hat{z}_i^1)^2h_i^1+((\hat{z}_i^1+\hat{z}_i^2)^2-(\hat{z}_i^1)^2)h_i^1}\right)^\sigma\right) \qquad (2.43) \\
&\leq \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \left(\frac{z_i^1+z_i^2}{\hat{z}_i^1+\hat{z}_i^2}\right)\left(1-\frac{1}{k^\sigma}\left(\frac{\hat{z}_i^1+\hat{z}_i^2}{z_i^1+z_i^2}\right)\frac{(z_i^1)^2+((z_i^1+z_i^2)^2-(z_i^1)^2)}{(\hat{z}_i^1)^2+((\hat{z}_i^1+\hat{z}_i^2)^2-(\hat{z}_i^1)^2)}\right)^\sigma\right) \qquad (2.44) \\
&= \max_{i\in[n],z_i\in\mathbb{R}^m_{\geq 0}} \left(\frac{z_i^1+z_i^2}{\hat{z}_i^1+\hat{z}_i^2}\right)\left(1-\frac{1}{k^\sigma}\left(\frac{z_i^1+z_i^2}{\hat{z}_i^1+\hat{z}_i^2}\right)^\sigma\right) \\
&\leq k\xi(\sigma)\ . \qquad (2.45)
\end{aligned}
$$

Above, (2.43) follows from the assumption that $h_i^1 \geq h_i^2$, (2.44) follows from the definition of the maximum degree of asymmetry k, and the same logic as applied in the case of capacity model 1 results in (2.45).

Together, this shows that regardless of capacity model, $\beta(\mathcal{C}_{k,\sigma}) \leq k\xi(\sigma)$. □

Chapter 3

Mixed Autonomous Traffic Networks: Design

In this chapter[1] we investigate toll design for traffic networks with mixed autonomy. In the preceding chapter we established a model for traffic networks with mixed autonomy and bounded the inefficiency due to selfish routing in the absence of tolling. In this chapter we investigate how much tolling can help mitigate this inefficiency. Specifically, for the case of affine latency functions, we provide differentiated tolls which completely eliminate inefficiency due to selfish routing. We show that anonymous cannot significantly improve inefficiency, and nonetheless we provide anonymous tolls which improve the performance guarantees from what they are without tolling. Finally, in the setting of parallel networks, we provide tolls that are ϵ-differentiated, meaning that each vehicle type pays almost the same amount on a road, and we show that these tolls completely

[1]This chapter is in part adapted, with permission, from the manuscripts "Optimal tolling for heterogeneous traffic networks with mixed autonomy", previously published by IEEE Conference on Decision and Control, © 2019 IEEE, "Optimal tolling for multitype mixed autonomous traffic networks" previously published by IEEE Control Systems Letters, © 2020 IEEE, and "Anonymous Tolling for Traffic Networks with Mixed Autonomy" [75–77]. These are joint work with Ramtin Pedarsani, and the first is joint work with Sam Coogan.

eliminate inefficiency due to selfish routing.

3.1 Introduction

Autonomous vehicles have been presented as a way to greatly decrease road congestion in popular culture and academic works [31, 78]. However, recent work shows that in the realm of *mixed autonomy*, where some vehicles are autonomous and some are human-driven, if vehicle users choose routes to minimize their own travel times, switching some vehicles from human-driven to autonomous can paradoxically *worsen* congestion [41]. Moreover, as shown in the preceding chapter, the Price of Anarchy, or maximum gap between the total delay under selfish routing as compared to the minimum total delay possible, is much larger in the case of mixed autonomy than when there is only one vehicle type.

When vehicles consist of a single type, well-known tolling schemes can eliminate inefficiency [23, 24, 79] and the *social cost*, or total delay experienced by all users, can be reduced to the minimum that it would be with all vehicles routed by a social planner. However, in the general mixed autonomy setting, optimal tolls are unknown, absent very restrictive assumptions [76, 80]. In our work, we provide optimal tolls for a broad class of settings, namely general networks with multiple source-destination pairs and with multiple vehicle types, where road latency is an affine function of vehicle flows. We do so via a simple, and to the best of our knowledge, novel, application of the Variational Inequality. With the tolls we provide, network inefficiency due to selfish routing is eliminated.

The optimal tolls we derive are *differentiated*, meaning that the different vehicle types must experience different tolls. However, it can be difficult to administer these differentiated tolls, due to logistics and concerns about privacy and fairness [81, 82]. In

the case of mixed autonomy, the tolls levied on the users of human-driven vehicles are likely to be greater than those levied against the users of autonomous vehicles, though users of autonomous vehicle may be wealthier. To address this, we investigate *anonymous* tolls, which are identical for all vehicle types on a given road. We provide a function to calculate anonymous tolls and prove that these tolls can decrease inefficiency, and provide upper and lower bounds on the efficiency improvement possible with anonymous tolls. Finally, we analyze the efficacy of variable marginal cost tolling in our setting.

To summarize, our contributions are as follows:

- We provide optimal differentiated tolls for general networks with mixed autonomy and affine latency functions, with optimality provided via a direct and clear proof,
- we provide anonymous tolls with an upper bound on the induced inefficiency,
- we establish lower bounds on the worst-case induced inefficiency with anonymous tolls for comparison,
- in the setting of parallel networks, we introduce tolls with only an infinitesimally small differentiation between vehicle types which eliminate inefficiency due to selfish routing, and
- we show that variable marginal cost tolling can yield equilibria which have social cost that are arbitrarily worse than that of the socially optimal routing.

3.1.1 Related Works

Classic works in congestion games determine optimal tolls for networks with a single vehicle type as far back as Pigou [79], and canonized in Beckmann *et al.* [23]. Dafermos generalizes this to the multiclass setting, in which different vehicle types affect congestion differently [24]. However, this work requires the social cost of flow on each road to be strictly convex, which is generally not the case in the mixed autonomy setting.

In the mixed-autonomy setting with two vehicle types and polynomial latency functions, previous work provides optimal tolls when the asymmetry between the congestion effects of the two vehicle types is constant across all roads in the network [80]. This work also establishes that when tolls are restricted to be nondifferentiated between vehicle types, tolling may not be able to achieve the socially optimal routing.

These works are based on the foundation of congestion games, which have long been used to analyze optimal routing [13], understand the properties of user (selfish) equilibria [51,83], and bound the *Price of Anarchy*, the maximum gap in social cost between optimal routing and the worst-case user equilibrium in a class of games [19, 21]. In Chapter 2, and the works on which it is based [44], this gap is bounded in the setting of mixed autonomy, and the maximum gap is shown to increase with the ratio between how much the different vehicle types can affect congestion on any road. Similarly, [41] describes and bounds the counterintuitively negative effect which can be caused by converting human-driven vehicles to less congesting autonomous vehicles, and [84] shows a similar effect with autonomous vehicles traveling to park in urban centers.

The effects of mixed autonomy on traffic congestion has been studied in a number of other contexts. This is studied in jointly routing and rebalancing autonomous mobility-on-demand services in the presence of other transportation modes [85] as well as autonomous carpooling [86]. Other works study selectively using route recommendations to improve the efficiency of equilibria [87, 88].

3.2 Model

We consider the setting proposed in Chapter 2, using the first capacity model (2.3) and a latency function with polynomial degree of 1 ($\sigma = 1$), leading to affine latency functions. For the sake of readability, and to introduce new notation, we describe our

full model in this section.

Consider a network described by a graph with n edges, where edges correspond to roads. We consider nonatomic flow, meaning each driver controls an infinitesimally small fraction of vehicle flow. Let m denote the number of vehicle types, where the latency on each edge is a function of the flow of each vehicle type on that edge. We generally use i to index roads and j to index vehicle types. For integer x, let $[x]$ denote $\{1, 2, \ldots, x\}$, and let $\mathbf{1}_n$ to denote the n-dimensional column vector of ones. Let $\langle v, w \rangle$ denote the inner product between the vectors v and w.

Consider inelastic flow demand, with demand described by $\{s_r, s'_r, \beta_r, j_r\}_{r \in [R]}$, where for each *commodity* r in the set of commodities $[R]$, s_r is that commodity's source node, s'_r is the destination node, β_r is the flow demand, and j_r is the vehicle type of the commodity. Thus, each commodity describes the flow demand of a specific vehicle type between a specific source and destination node.

We use z_i^j to describe the flow of vehicle of type i on road j. Then, for road i, z_i is a column vector describing the flow of all vehicle types on road i:

$$z_i := \begin{bmatrix} z_i^1 & z_i^2 & \ldots & z_i^m \end{bmatrix}^T .$$

Let z describe the flow over the entire network:

$$z = \begin{bmatrix} z_1^T & z_2^T & \ldots & z_n^T \end{bmatrix}^T .$$

We can alternately describe the flow through a graph in terms of the flow on each route. Let $\mathcal{P}_r$ denote the set of simple paths available to commodity r. Let $f_p^j \geq 0$ denote the flow of vehicle type j on path p. Then a routing f is feasible if $\sum_{p \in \mathcal{P}_r} f_p^{j_r} = \beta_r$ for commodities $r \in [R]$. Since there is a one-to-one correspondence between path flows and

edge flows, a characterization of feasible routings can be applied to edge routings z as well. Thus, a feasible routing z refers to a routing which satisfies the commodity flow demands and graph constraints. We use $\mathcal{Z}$ to denote the set of feasible routings. We use $\mathcal{I}_p$ to denote the set of edges in path p.

As derived in [75] and [76] (based on the Bureau of Public Roads delay model in conjunction with a capacity model derived for mixed autonomy [42, 73]), we consider multitype mixed autonomous traffic as having an affine latency function, where the scaling effect of each vehicle type on congestion depends on the nominal headway that the vehicle type maintains. Accordingly, each edge i has latency, or cost, function $c_i : \mathbb{R}^m_{\geq 0} \to \mathbb{R}^m_{\geq 0}$

$$c_i(z_i) = A_i z_i + b_i \mathbf{1}_n$$

where $b_i \in \mathbb{R}_{\geq 0}$ denotes the free-flow latency on road i and $A_i \in \mathbb{R}^{m \times m}_{\geq 0}$ describes the linear increase in latency with flow. We consider all vehicle types to experience the same delay from traversing a road, so we can write $A_i = \mathbf{1}_n \otimes a_i$, for some $a_i \in \mathbb{R}^m_{\geq 0}$. To ease our later derivations, we define

$$\begin{aligned}
A &:= \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_n \end{bmatrix}, \\
b &:= \begin{bmatrix} b_1 \mathbf{1}_n^T & b_2 \mathbf{1}_n^T & \dots & b_n \mathbf{1}_n^T \end{bmatrix}^T, \text{ and} \\
c(z) &:= \begin{bmatrix} c_1(z_1)^T & c_2(z_2)^T & \dots & c_n(z_n)^T \end{bmatrix} = Az + b,
\end{aligned}$$

and we use $\mathcal{C}$ to refer to the set of cost functions c which are of this form.

We consider tolling, where vehicle types may experience different tolls. Accordingly,

we denote the cost experienced by users of type j on road i under toll τ_j as $c_i^{\tau_j} \in \mathbb{R}^m_{\geq 0} \to \mathbb{R}^m_{\geq 0}$, where for some toll τ_i^j,

$$c_i^{\tau_j}(z_i) = c_i(z_i) + \tau_i^j \ .$$

We also define the following quantities:

$$\begin{aligned}
c_i^{\tau}(z_i) &:= \begin{bmatrix} c_i^{\tau_1}(z_i)^T & c_i^{\tau_2}(z_i)^T & \dots & c_i^{\tau_m}(z_i)^T \end{bmatrix} \ , \\
c^{\tau}(z) &:= \begin{bmatrix} c_1^{\tau}(z_1)^T & c_2^{\tau}(z_2)^T & \dots & c_n^{\tau}(z_n)^T \end{bmatrix} \ , \text{ and} \\
\tau &:= \begin{bmatrix} \tau_1^T & \tau_2^T & \dots & \tau_n^T \end{bmatrix}^T \ .
\end{aligned}$$

When discussing the efficiency of a routing, we will judge it by its *social cost*, or aggregate latency, which we denote by $C(z)$, where $C :\in \mathbb{R}^{mn}_{\geq 0} \to \mathbb{R}_{\geq 0}$:

$$C(z) := \sum_{i \in [n]} \langle c_i(z_i), z_i \rangle = \langle c(z), z \rangle \ . \tag{3.1}$$

Note that tolls do not deduct from the social cost, as we consider the tolls to be circulated back into public funds.

We are concerned with the social cost of user equilibria in this setting. This means routings for which no user type would wish to change their routing, where users experience the cost of a route as a sum of the latency and the toll for their user type. Specifically, we consider Wardrop Equilibrium, which can be specified in terms of path routing.

Definition 3.1. *A flow is at* Wardrop Equilibrium *if no user can decrease their cost by switching routes. Mathematically, a flow f is at Wardrop Equilibrium relative to toll τ if*

$$\forall r \in [R],\ \forall p, p' \in \mathcal{P}_r,\ f_p^{j_r} > 0 \implies \sum_{i \in \mathcal{I}_p} c_i^{\tau_{j_r}}(z_i) \leq \sum_{i \in \mathcal{I}_{p'}} c_i^{\tau_{j_r}}(z_i) \ .$$

In words, for each commodity, if a path has positive flow, then no other path available to that commodity can have lower cost, where the cost is relative to the user type of the commodity.

Another useful way of characterizing equilibria is using the *Variational Inequality* [16, 51]. Applied to our setting, if feasible routing z is a Wardrop Equilibrium, then for any other feasible routing $z' \in \mathcal{Z}$,

$$\langle c^\tau(z), z - z' \rangle \leq 0 \ .$$

In contrast with equilibrium routing, we define a socially optimal routing z^* as a routing in the set of feasible routing $\mathcal{Z}$ which minimizes the cost function $C(z)$. We define the set of minimizers of $C(z)$ by $\mathcal{Z}^*$

3.3 Differentiated Tolls

We are now prepared to present optimal tolls, which completely eliminate inefficiency in routing in our setting. Specifically, when these differentiated tolls are applied, the social cost of any routing at equilibrium will equal the optimal social cost of any feasible routing.

Theorem 3.1. *Let z^* denote a socially optimal routing for a network with social cost C. Apply tolls $\tau_i = A_i^T z_i^*$. Any resulting Wardrop Equilibrium z' will have optimal social cost, meaning*

$$C(z') = C(z^*) \ .$$

Proof. Let $\hat{z}$ denote a routing at Wardrop Equilibrium, and let z^* denote a socially

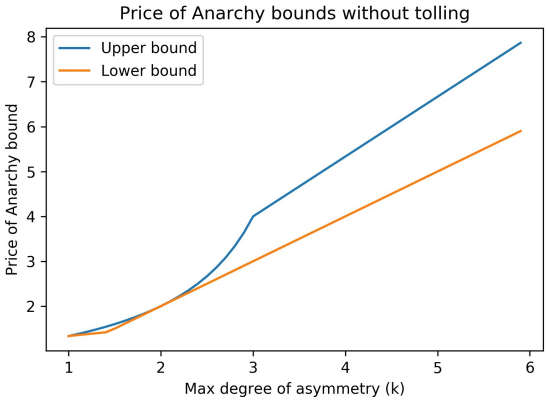
optimal routing.

$$\begin{aligned}
C(\hat{z}) &= \langle c(\hat{z}), \hat{z} \rangle = \langle c^{\tau}(\hat{z}), \hat{z} \rangle - \tau^T \hat{z} \\
&\leq \langle c^{\tau}(\hat{z}), z^* \rangle - \tau^T \hat{z} \\
&= \langle c(\hat{z}), z^* \rangle - \tau^T (\hat{z} - z^*) \\
&= \langle c(z^*), z^* \rangle + \langle c(\hat{z}) - c(z^*), z^* \rangle - \tau^T (\hat{z} - z^*) \\
&= C(z^*) + z^{*T}(A\hat{z} + b - Az^* - b) - \tau^T (\hat{z} - z^*) \\
&= C(z^*) + z^{*T} A(\hat{z} - z^*) - \tau^T (\hat{z} - z^*) \\
&= C(z^*) + (z^{*T} A - \tau^T)(\hat{z} - z^*) \\
&= C(z^*)
\end{aligned}$$

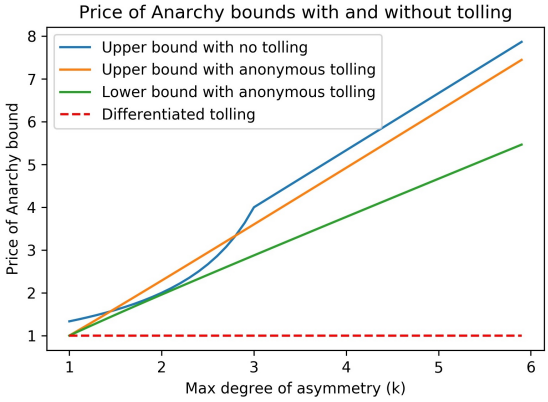
The inequality on the second line results from the Variational Inequality and the final equality stems from the structure of the tolls. By the definition of optimal routing, $C(\hat{z}) \geq C(z^*)$, so we find that $C(\hat{z}) = C(z^*)$. □

This proof is straightforward, and results from a novel application of Variational Inequality, similar to that of [21], but applied to the cost function under the application of tolls. Thus we have optimal tolls for multitype mixed autonomy in general networks with affine latency functions. As mentioned in the theorem, the tolls depend on an optimal flow pattern, which must be found to calculate the tolls.

Remark 3.1. *These optimal tolls are the same as the edge tolls in [24], though in that work the proof of optimality is limited to when the social cost is a convex function, which in our setting it is not. The optimality of these same tolls has been proven in the mixed autonomy setting when there are two vehicle types [80], but it is restricted to the case in which the asymmetry in the congestion effects of the vehicle types is constant across roads,* i.e. *weighted congestion games, though they prove this for polynomial latency functions.*



(a) Bounds on the Price of Anarchy in Theorem 2.1, compared against PoA lower bounds from Chapter 2.



(b) Bounds on inefficiency with anonymous tolls, compared with untolled PoA upper bound and tolled PoA lower bound. The PoA with the proposed differentiated tolling is 1.

Figure 3.1: Bounds on the Price of Anarchy, with and without tolling

Since they consider only two vehicle types, this assumption means that $a_i^1/a_i^2 = k$ *for some constant* k *for all roads* i *in the set of roads* $[n]$*. Our theoretical results greatly expands their results by extending it to multiple types beyond two vehicle types and removing this assumption about the congestion effects of the vehicle types, thereby expanding the model from that of a weighted congestion game.*

3.4 Anonymous Tolls

As mentioned in the introduction, implementing differentiated tolls may be logistically or politically prohibitive. Accordingly, in this section we present anonymous tolls which improve the bound on worst-case equilibrium behavior. In order to contextualize these results, we first review a special case of results from the preceding chapter which bound worst-case equilibrium behavior in the mixed autonomy setting without tolling. This will contextualize the results on anonymous tolling results to follow.

3.4.1 Price of Anarchy Bounds

The Price of Anarchy (PoA) is a metric which measures, for a class of congestion games and cost functions, the maximum ratio between the social cost under the worst-case equilibrium to the socially optimal cost. This gives a worst-case bound on how much worse-off the social cost can be when users act selfishly. The PoA can be established with or without tolls – for example, tolls that bring the equilibrium social cost equal to the socially optimal cost, such as the differentiated tolls in the previous section, have a PoA of 1. We now present a special case of the tolls presented in the preceding chapter.

Definition 3.2. *The* maximum degree of asymmetry*, denoted k, is the maximum ratio of congestion effects due to two vehicle types. In our setting,*

$$k := \max_{i\in[n],j,j'\in[m]} a_i^j/a_i^{j'} \ .$$

Note that by definition, $k \geq 1$. We now present a theorem which extends the affine case of Theorems 1 and 2 in [44].

Corollary 3.1 (Theorem 2.1 and 2.2)**.** *For any feasible equilibrium routing $\hat{z}$ and optimal routing z^* in an untolled network with maximum degree of asymmetry k,*

$$C(\hat{z}) \leq \Lambda(k)C(z^*) \ ,$$

where

$$\Lambda(k) = \begin{cases} 4/(4-k) & k \leq 3 \\ 4k/3 & k > 3 \end{cases} .$$

The bounds above are not proven to be tight in all cases, but in Chapter 2 we provide two examples to bound from below worst-case PoA, one of which reaches a PoA of k,

and one of which reaches $1 + k/(2\sqrt{k} + 1)$. Figure 3.1 (a) compares these lower bounds against the upper bounds from Corollary 3.1.

3.4.2 Anonymous Tolling

In this section we establish anonymous tolls which, when applied, improve the previously established PoA bounds; we defer the proof of our theoretical results to the appendix. We will then provide lower bounds on the worst-case PoA with this anonymous routing via examples.

Theorem 3.2. *Find an optimal routing z^*. Then on each road i, levy the following identical toll for all vehicle types:*

$$\tau_i^j = \min_{j' \in [m]} a_i^{j'} \sum_{j \in [m]} z_i^{*j} \ .$$

Then, for any Wardrop Equilibrium $\hat{z}$,

$$C(\hat{z}) \leq \frac{4k}{3k+1} k C(z^*) \ .$$

This anonymous toll lowers the upper bound on PoA from that of the class of congestion games without tolling for $k < 1.45$ and $k > 2.8$. However, the PoA bound, while tight for maximum degree of asymmetry $k = 1$ and $k = 2$, is not tight for all k, as shown in Fig. 3.1 (a). Therefore, lowering the upper bound is not guaranteed to improve worst-case performance. Nevertheless, the bound with anonymous tolling is lower than without tolling and the subsequent examples show that this tolling can improve the PoA in certain cases.

We next investigate the tightness of this bound. The first example will prove the following proposition, which presents a subclass of anonymous tolls with a specific structure

which makes it easier to compute. Specifically, we consider tolling schemes in which the toll on a road is a function of the vehicle flows of each type on that road, the optimal vehicle flows, and the network parameters, but treating these vehicle types interchangeably. We formulate this mathematically in the following proposition.

Proposition 3.1. *Consider the class of tolls in which the anonymous tolls on each road i is a function of the form:*

$$\tau_i(\{(z_i^j, z_i^{*j}, a_i^j, b_i)\}_{j\in[m]}) \ .$$

The Price of Anarchy with these tolls applied is at least k.

The lower bound on the PoA in the proposition is proven in the following example.

Example 3.1. *Consider the network in Figure 2.5 (a), with flow demand of* 1 *for each vehicle type. The socially optimal routing has the flow of type* 1 *on the bottom road and type* 2 *on the top road, for a social cost of* 2*. The worst-case equilibrium has this routing reversed, for a social cost of* $2k$*. This yields a PoA of* k*.*

Due to the symmetry of the optimal routing, the anonymous tolls of Theorem 3.2 are equal on the two roads, with a worst-case equilibrium routing again having a social cost of $2k$*, yielding a PoA of* k*. Thus, in this network, the anonymous tolls provided do not improve the worst-case performance. Further, due to the symmetry in this example, the two roads have identical inputs to the tolling function for any tolling function of the form proposed in Proposition 3.1. Thus the two roads must have identical tolls, again yielding a PoA of* k*. This proves the proposition.*

This class of tolls in Proposition 3.1 (of which the tolls in Theorem 3.2 are a member) has the advantage of depending only on the optimal routing, and not on calculating a specific equilibrium routing, which is often much more computationally difficult to find.

However, if the tolls are not restricted to the above form, the social cost can be reduced slightly to $(7k+3)/4 - 1/(k+1)$*, as shown in [76]. The resulting Price of Anarchy in this case is used as our lower bound on PoA for anonymous tolling. We compare the untolled PoA upper bound to the anonymous PoA upper bound and this lower bound in Figure 3.1 (b).*

We next present an example in which our proposed anonymous tolling *does* improve the Price of Anarchy in a network from what it is without tolling.

Example 3.2. *Consider the network shown in Figure 2.5 (b), with vehicle flow demand of* $1/\sqrt{k}$ *units of type* 1 *and* 1 *unit of type* 2*. The optimal routing has the flow of type* 1 *entirely on the top road and type* 2 *on the bottom road, for a cost of* $1/\sqrt{k} + 1/(\sqrt{k}+1)$*. The worst-case equilibrium has all flow on the bottom road, for a social cost of* $1 + 1/\sqrt{k}$*, yielding a PoA which scales with* $\sqrt{k}$ *[44].*

With the tolling from Theorem 3.2, both vehicle types experience a toll on road 1 *of* 0 *and a toll on road* 2 *of* $1/(\sqrt{k}+1)$*. The resulting worst-case equilibrium has all flow of vehicle type* 1 *on the lower road and vehicle type* 2 *on the upper road, for a social cost of* $1 + 1/(\sqrt{k}+1)$*. This an improvement from the untolled worst-case equilibrium cost of* $1 + 1/\sqrt{k}$*.*

The worst-case equilibrium social cost could be further improved with a toll on road 2 *of* $\tau_2 = 1/2$*, with the resulting worst-case equilibrium involving flow of vehicle type* 1 *split between the two roads and all vehicle flow of type* 2 *on the top road for a social cost of* $1 + (3\sqrt{k} - 1)/(4k)$*.*

3.5 ϵ-Differentiated Tolls

In this section we consider a setting of a parallel network of n roads, where roads have affine latency functions. We present theoretical results with tolls that are almost

anonymous and only differentiated with ϵ-differentiation, where ϵ is arbitrarily small, which completely eliminate inefficiency due to selfish routing. Before presenting our theoretical results, we first introduce some necessary notation.

For a specific routing z, we use $\mathcal{N}_j^z$ to denote the set of roads with positive flow of vehicle type j:

$$\mathcal{N}_j^z = \{i : z_i^j > 0 \wedge i \in [n]\} \ .$$

Similarly, we use $\mathcal{M}_i^z$ to denote the set of vehicle types with positive flow on road i for the routing z:

$$\mathcal{M}_i^z = \{j : z_i^j > 0 \wedge j \in [m]\} \ .$$

The theoretical results established in the next section conceptualize vehicle flow on roads in the form of a graph, where for each specific routing z, a graph can be constructed. We construct a bipartite graph $G = (U, V, E)$ where one set of nodes is the set of roads ($U = [n]$) and one set of nodes is the set of vehicle types ($V = [m]$). The set of edges connect vehicle types to roads on which they have positive flow, *i.e.*

$$E = \{(i, j) : i \in [n] \wedge j \in \mathcal{M}_i^z\} \ , \tag{3.2}$$

or equivalently, $E = \{(i, j) : z_i^j > 0 \wedge i \in [n] \wedge j \in [m]\}$. In other words, for a routing z, there is an edge between the nodes denoting road i and vehicle type j if there is positive flow of type j on road i. We illustrate this in Figure 3.2.

We make the following assumption regarding the latency functions.

Assumptions 3.1. *The road latency functions are affine and the latency of each road is strictly increasing with the flow of each vehicle type on that road,* i.e. $a_i^j > 0$ *for all* $i \in [n]$, $j \in [m]$.

We rephrase the Wardrop Equilibrium condition for the case of parallel networks.

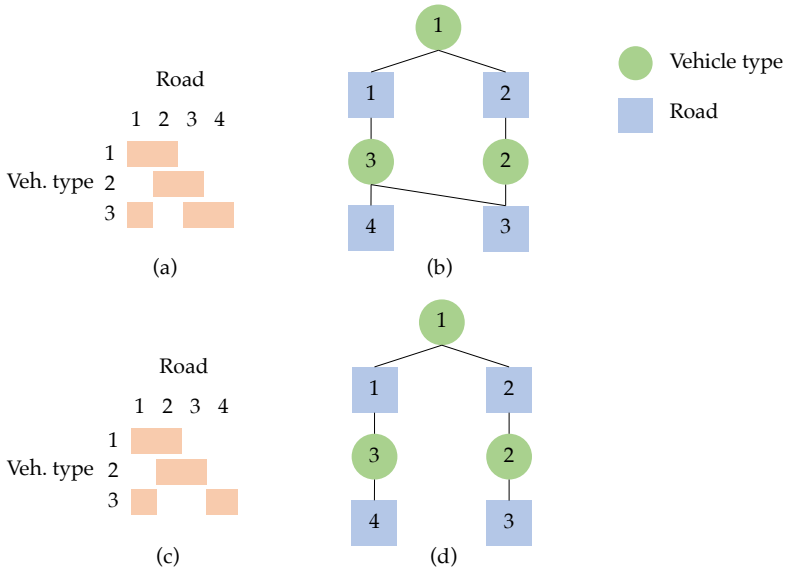


Figure 3.2: Example routings for a network with four roads and three vehicle types. (a) vehicle type 1 has positive flow on roads 1 and 2, type 2 has positive flow on roads 2 and 3, and type 3 has positive flow on roads 1, 3, and 4. (b) shows the corresponding bipartite graph. (c) shows a similar routing but type 3 has zero flow on road 3, and (d) shows its corresponding bipartite graph.

Definition 3.3. *A flow z is a* Wardrop Equilibrium *if $z_i^j > 0$ implies $c_i^j(z) \leq c_{i'}^j(z)$ for all $i, i' \in [n]$, $j \in [m]$.*

Since in this section we consider a network of parallel roads, a flow is at Wardrop Equilibrium if no user can decrease their cost by switching roads.

In this section we establish tolls which ensure that the social cost is minimized in any resulting equilibrium. We do this in two theorems: the first establishes properties about a routing in the set of routings which minimize the social cost, and the second provides optimal tolls which are constructed based on the routing which is proved to exist in the first theorem.

Theorem 3.3. *Consider the setting of multitype congestion games with affine cost functions on parallel networks. There exists a routing in the set of routings minimizing social cost, $z \in \mathcal{Z}^*$, such that $G(z)$ is acyclic, where $G = (U, V, E)$ is constructed as in* (3.2), i.e.

where nodes are the roads and vehicle types, and edges exist between road i and vehicle type j when $z_i^j > 0$.

Corollary 3.2. *For the routing $z \in \mathcal{Z}^*$, such that $G(z)$ is acyclic, provided by Theorem 3.3, no two vehicle types share more than one common road.*

Theorem 3.4. *Consider the setting of multitype congestion games with affine cost functions on parallel networks. Consider any optimal routing $z^* \in \mathcal{Z}^*$ that has an associated acyclic graph. The existence of such a routing is provided in Theorem 3.3. Assume the latency on a road is increasing in the flow of each vehicle type on that road. Then levy the following tolls $\tau(z^*)$:*

$$\tau_i^j({z^*}_i) = \begin{cases} \mu - c_i({z^*}_i) & \text{if } i \in \mathcal{N}_j^{z^*} \\ \mu - c_i({z^*}_i) + \epsilon & \text{if } i \in \cup_{j' \in [m]} \mathcal{N}_{j'}^{z^*} \setminus \mathcal{N}_j^{z^*} \\ P & \text{otherwise.} \end{cases} \tag{3.3}$$

Then for every μ, and every $\epsilon > 0$, and sufficiently large P, the only Nash equilibrium that exists is the flow z^*.

In words, we find an optimal routing with the acyclic property guaranteed by Thm 3.3. Denote this routing z^*. Then, based on this routing, we assign a toll for each vehicle type on each road. If z^* has no vehicle flow on a specific road, we set a very high toll for all vehicle types, ensuring that in equilibrium, vehicle flow on this road will be zero for all vehicle types. If a vehicle type has positive flow on road i under z^*, it is given the toll $\mu - c_i(z_i)$. If it does not have positive flow under z^* but another vehicle type does, it is given the toll $\mu - c_i(z_i) + \epsilon$, where ϵ is any positive number. This guarantees the optimal routing at equilibrium.

We see that optimal routing can be guaranteed with an infinitesimal difference be-

tween the tolls of each vehicle type on a road. We can even consider the setting in which users are given random gifts to help encourage them to follow the route recommendations and use these to ensure optimal routing without having a different toll for the vehicle types. We now prove the theorem and defer the proofs of the supporting lemmas to the appendix.

Proof. We prove this theorem by breaking down a candidate equilibrium $\tilde{z}$ into cases for which roads the vehicle types have positive flow on, relative to the roads they have positive flow on in z^*, the optimal routing on which the tolls are based. By analyzing each case, we show that $\tilde{z} = z^*$.

We first present a proposition suggesting that for sufficiently large P, no equilibrium flow will exist on roads with toll P. We then provide a lemma stating that in any equilibrium in which all vehicle types use the roads they were on in the routing on which the tolls were based, z^*, then the equilibrium flow will be equal to z^*. Finally, we examine the case in which $\tilde{z}$ has at least one vehicle type with positive flow on some road or roads which did not have positive flow of the vehicle type in z^* (though not on a road with toll P). To reiterate, our three cases are as follows:

1. $\tilde{z}$ has some vehicle type with positive flow on a road which did not have positive vehicle flow of any type in z^*,
2. $\tilde{z}$ is such that each vehicle type has positive flow on the same roads as that vehicle type did in z^*, or
3. $\tilde{z}$ is such that at least one vehicle type has positive flow on a road which, in z^*, did not have positive flow of that vehicle type but did have positive flow of another vehicle type.

The first case is handled in a proposition, the second in a lemma, and the third in the subsequent development.

Proposition 3.2. *For sufficiently large P, users in equilibrium will not use a road with toll P. Formally, for equilibrium flow f experiencing tolls $\tau(z^*)$,*

$$f_i^j = 0 \; \forall j \in [m] \wedge i \in [n] \setminus \cup_{j' \in [m]} \mathcal{N}_{j'}^{z^*} \; .$$

This means that no equilibrium will have positive flow on roads that did not have positive flow in z^*.

Lemma 3.1. *Consider the setting of multitype congestion games with affine cost functions on parallel networks. Consider any routing z^* that has an associated acyclic graph; the existence of such a routing is provided in Theorem 3.3. Assume that in routing z^*, every road has positive flow of some vehicle type. Then levy the following tolls $\tau(z^*)$:*

$$\tau_i^j(z^*) = \mu - \ell_i(z^*) \tag{3.4}$$

Consider a new equilibrium routing $\tilde{z}$ with flow demand of each type no greater than the demand of that type in z^*, meaning

$$(\forall j \in [m])[\sum_{i \in [n]} {z^*}_i^j \leq \sum_{i \in [n]} \tilde{z}_i^j] \; . \tag{3.5}$$

Moreover, assume that each vehicle type has positive flow in $\tilde{z}$ only on roads that it has positive flow on in z^*, meaning

$$(\forall j \in [m])[\mathcal{N}_j^{\tilde{z}} \subseteq \mathcal{N}_j^{z^*}] \; . \tag{3.6}$$

Then under Assumption 3.1, for every μ,

1. *Equilibrium routing $\tilde{z}$ is unique, meaning for any other equilibrium $\tilde{z}'$,*

$$(\forall i \in [n] \wedge \forall j \in [m])[\tilde{z}_i'^j = \tilde{z}_i^j] \ .$$

2. *If $\tilde{z}$ has the same flow demand as z^*, then $\tilde{z} = z^*$. Formally,*

$$(\forall j \in [m])[\sum_{i\in[n]} \tilde{z}_i^j = \sum_{i\in[n]} z^{*j}_i] \implies \tilde{z} = z^* \ .$$

This lemma, which is proved in the appendix, handles Case 2 above. We will also use the second statement in proving Case 3, which will be done in the remainder of this proof.

Let us use $\mathcal{I}$ to denote the set of roads with positive flow in z^*:

$$\mathcal{I} = \{i \in [n] \mid (\exists j \in [m])[z_i^{*j} > 0]\} \ . \tag{3.7}$$

Next we define the notation $c_i^{\bar{\tau}}(z_i)$, which we call the *standard cost* on a road, to denote the minimum cost on a road to the vehicle types which are on that road – if it has vehicles which are on it in z^* this road will have their tolled cost, otherwise it will have the large cost of P.

$$c_i^{\bar{\tau}}(z_i) := \begin{cases} c_i(z_i) + \mu - c_i(z_i^*) & \text{if } i \in \mathcal{I} \\ P & \text{otherwise.} \end{cases} \tag{3.8}$$

Proposition 3.3. *By the definition of Wardrop Equilibrium (Definition 3.3) and the structure of the tolls, all pairs of used roads will have standard cost within ϵ of each*

other. More specifically,

$$\forall i, i' \in [n], j \in [m] \text{ s.t. } \hat{z}_i^j > 0,\ c_i^{\bar{\tau}}(\hat{z}_i) \leq c_{i'}^{\bar{\tau}}(\hat{z}_{i'}) + \epsilon\ . \tag{3.9}$$

Proof. The reason for this is as follows. For any road i with positive flow of vehicle type j, if there is another road with $c_{i'}^{\bar{\tau}}(\hat{z}_{i'}) + \epsilon < c_i^{\bar{\tau}}(\hat{z}_i)$ then $c_{i'}^{\tau_j}(\hat{z}_{i'}) < c_i^{\tau_j}(\hat{z}_i)$, violating Wardrop Equilibrium conditions. □

So far we have established that in a Wardrop Equilibrium, all used roads will have standard cost within ϵ of each other, and no roads that were not used in z^* will be used in this equilibrium routing. We will next investigate the standard costs on the roads as it relates the vehicle types on each road.

We now define some additional notation: let $\underline{c}$ denote the minimum standard cost of all roads which were used in z^*, and $\overline{c}$ denote the maximum standard of all used roads which were used in z^*. Mathematically,

$$\begin{aligned} \underline{c} &= \min_{i \in \mathcal{I}} c_i^{\bar{\tau}}(z_i) \\ \overline{c} &= \max_{i \in \mathcal{I}} c_i^{\bar{\tau}}(z_i) \end{aligned} \tag{3.10}$$

Since by Lemma 3.1 if no vehicle types in $\hat{z}$ are on roads they were not on in z^*, meaning for all $j \in [m]$, $\mathcal{N}_j^{\hat{z}} \subseteq \mathcal{N}_j^{z^*}$, then $\hat{z} = z^*$. Accordingly, we restrict our attention to the situation in which some road has a vehicle type in $\hat{z}$ that it did not have in z^*, meaning

$$\exists j \in [m], i \in [n] \text{ s.t. } \hat{z}_i^j > 0 \wedge z_i^{*j} = 0 \tag{3.11}$$

Next we make a statement about the roads containing vehicle types in $\hat{z}$ that were not there in z^*.

Lemma 3.2. *Any road that has a vehicle type on it in $\hat{z}$ but not in z^* will have standard cost $\underline{c}$ under routing $\hat{z}$, and all roads that have this specific vehicle type in z^* have standard cost $\overline{c}$ under routing $\hat{z}$. Mathematically,*

$$(\forall j \in [m] \wedge i \in [n])\big[\hat{z}_i^j > 0 \wedge z_i^{*j} = 0 \implies c_i^{\overline{\tau}}(\hat{z}_i) = \underline{c}\big] \text{ , and}$$
$$(\forall j \in [m] \wedge i \in [n])\big[\hat{z}_i^j > 0 \wedge z_i^{*j} = 0 \implies (\forall i' \in \mathcal{N}_j^{z^*})[c_{i'}^{\overline{\tau}}(\hat{z}_{i'}) = \underline{c} + \epsilon]\big] \ .$$

Moreover,

$$\overline{c} = \underline{c} + \epsilon \ . \tag{3.12}$$

Since this logic applies to all roads which have positive flow of a vehicle type in $\hat{z}$ that did not have it in z^*, we can construct a set of roads which contains all roads with positive flow in $\hat{z}$ that did not have positive flow of those types in z^*. This set will contain these roads and possibly other roads as well. We partition $\mathcal{I}$, the roads which can have positive flow in an equilibrium, into three partitions: $\mathcal{I}_1$, $\mathcal{I}_2$, and $\mathcal{I}_3$. We define two first two iteratively, and we define the third to be the roads not included in the first two.

Definition 3.4. *Let $\mathcal{I}_1$ denote roads which have positive flow of vehicle types in $\hat{z}$ that were not there in z^*, and iteratively add in all roads which share common vehicle types (relative to z^*) with those roads. We define $\mathcal{I}_2$ iteratively as well, starting with roads which had positive flow in z^* of those vehicle types now in $\mathcal{I}_1$, and adding all roads which share common vehicle types (relative to z^*) with those roads. We define $\mathcal{I}_3$ as all remaining roads. To formulate this mathematically, first define a helper function which takes in a set of roads and outputs the union of that set and all roads that share vehicle types with it in z^*:*

$$\Phi(A) = A \cup \{i \in [n] \mid (\exists j \in [m], i' \in A)[z_{i'}^{*j} > 0 \wedge z_i^{*j} > 0]\} \ . \tag{3.13}$$

Then,

$$\begin{aligned}
\mathcal{I}_1 &= \Phi^n(\{i \in \mathcal{I} \mid (\exists j \in [m])[\hat{z}_i^j > 0 \wedge z_i^{*j} = 0]\}) \\
\mathcal{I}_2 &= \Phi^n(\{i' \in \mathcal{I} \mid (\exists j \in [m], i \in [n])[\hat{z}_i^j > 0 \wedge z_i^{*j} = 0 \wedge i' \in \mathcal{N}_j^{z^*}]\}) \\
\mathcal{I}_3 &= \mathcal{I} \setminus (\mathcal{I}_1 \cup \mathcal{I}_2) \, .
\end{aligned} \tag{3.14}$$

The construction of $\mathcal{I}_1$ and $\mathcal{I}_2$ to have roads with continuously overlapping vehicles types with respect to z^*, implying that all its roads have the same standard cost. Using this and Lemma 3.2 we can state the following:

$$\begin{aligned}
&\forall i \in \mathcal{I}_1 : \; c_i^{\overline{\tau}}(\hat{z}_i) = \underline{c} \\
&\forall i \in \mathcal{I}_2 : \; c_i^{\overline{\tau}}(\hat{z}_i) = \overline{c} = \underline{c} + \epsilon \\
&\forall i \in \mathcal{I}_3 : \; \underline{c} \leq c_i^{\overline{\tau}}(\hat{z}_i) \leq \overline{c} \, .
\end{aligned} \tag{3.15}$$

If (3.15) and (3.14) share any roads, then $\underline{c} = \underline{c} + \epsilon$, yielding a contradiction and proving the theorem. Otherwise, these three sets are partitions of $\mathcal{I}$ and do not share any roads in common with each other. Moreover, these roads partition the vehicle types in z^*, meaning that in z^*, there are no vehicle types which have positive flow on more than one of the sets $\mathcal{I}_1$, $\mathcal{I}_2$, or $\mathcal{I}_3$.

Moreover, as an implication of the definition, all roads in $\mathcal{I}_3$ have $\hat{z}_i^j = z_i^{*j}$. Accordingly, by the definition of the tolls,

$$\forall i \in \mathcal{I}_3 : \; c_i^{\overline{\tau}}(\hat{z}_i) = \mu \, . \tag{3.16}$$

We next relate $\underline{c}$ and $\overline{c}$ to μ. First we deal with $\mathcal{I}_1$, which has $c_i^{\overline{\tau}}(\hat{z}_i) = \underline{c}$.

Lemma 3.3. *For any road i in partition $\mathcal{I}_1$, constructed as in Definition 3.4, the standard*

cost on i will be at least as great in $\hat{z}$ as it is in z^*. Mathematically,

$$(\forall i \in \mathcal{I}_1)[c_i^{\overline{\tau}}(\hat{z}_i) \geq c_i^{\overline{\tau}}(z_i^*)] \tag{3.17}$$

We now deal with the roads in $\mathcal{I}_2$.

Lemma 3.4. *For any road i in partition $\mathcal{I}_2$, constructed as in Definition 3.4, the standard cost on i will not be greater in $\hat{z}$ as it is in z^*. Mathematically,*

$$(\forall i \in \mathcal{I}_2)[c_i^{\overline{\tau}}(\hat{z}_i) \leq c_i^{\overline{\tau}}(z_i^*)] \tag{3.18}$$

However, there is a contradiction between Lemmas 3.3 and 3.4 (combined with the fact that in $c_i^{\overline{\tau}}(z_i^*) = \mu$ for all roads i in $\mathcal{I}$) with (3.15), proving Case 3. With all three cases handled, the theorem is proved.

□

The theorem above provides tolls which ensure that any equilibrium will minimize the social cost of routing. Importantly, the tolls for each vehicle type differ by at most ϵ, which must be greater than zero, but can be arbitrarily small. A system designer can consider different incentive methods to create this small differentiation. Some possible methods include actually differentiated tolls, randomly providing gifts to vehicles of some types to differentiate the tolls on expectation, or even pro-social public messaging. This small differentiation addresses many of the problems with non-anonymous tolls, including fairness and privacy.

3.6 Variable Marginal Cost Tolling

The tolls provided in this chapter have been *fixed cost* tolls, meaning that the tolls are calculated based on the network structure and overall flow demand, but the toll on a link does not vary with the flow on that link. However, previous works have shown that fixed tolls are not *strongly robust* to mischaracterizations of latency functions in a network, meaning that if the information used to calculate the tolls is not accurate, the tolls do not necessarily incentivize optimal behavior [89]. Moreover, the tolls provided depend on knowledge of the entire network and the flow demand of all vehicle types. A class of tolls which alleviates this latter requirement is *variable marginal cost tolls*, in which the toll provided to vehicles of type j on road i is

$$\tau_i^j(z_i) = (\sum_{j\in[m]} z_i^j)\frac{\partial}{\partial z_i^j} c_i(z_i) \ .$$

Assuming, as we do, that all vehicle types experience latency identically, if the social cost function is convex, then variable marginal cost tolling ensure that any resulting equilibrium has the social cost of the socially optimal routing[2] [90]. However, in our setting the assumption of convex social cost function does not hold. It is therefore worth investigating whether marginal cost tolls are optimal in the setting of mixed autonomy; the following proposition shows that they are not.

Proposition 3.4. *Consider the setting of multitype congestion games with affine cost functions. The Price of Anarchy, with variable marginal cost tolls applied, is unbounded.*

Proof. We prove this proposition by example. Consider the network shown in Figure 3.3. Consider 1 unit flow demand of vehicle type 1 and ζ units flow demand of vehicles type 2.

[2]This is because the tolls align the Wardrop Equilibrium conditions and the conditions for first-order optimality of the social cost.

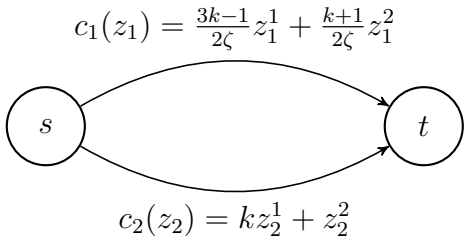


Figure 3.3: Examples proving Proposition 3.4. Consider 1 unit flow demand of vehicle type 1 and ζ units flow demand of vehicles type 2.

The form of the marginal cost tolls make it such that the Wardrop Equilibrium conditions are satisfied when first-order optimality conditions are satisfied. We will therefore lower bound the Price of Anarchy by taking the ratio of the social cost at a saddle point of the social cost function to the social cost at another feasible routing. Since the socially optimal routing will not have greater social cost than this latter term, this will be a lower bound on the PoA.

With the parameters chosen above, a saddle point of the social cost function (and therefore a Wardrop Equilibrium with tolls applied) exists at $z_1^1 = 0$, $z_1^2 = \zeta$, $z_2^1 = 1$, $z_2^2 = 0$. Let us use $\hat{z}$ to denote this routing. Let us use $\tilde{z}$ to denote another routing, $z_1^1 = 1$, $z_1^2 = 0$, $z_2^1 = 0$, $z_2^2 = \zeta$. Then, for any socially optimal routing z^*,

$$\frac{C(\hat{z})}{C(z^*)} \geq \frac{C(\hat{z})}{C(\tilde{z})} = \frac{(k+1)\zeta^2 + 2k\zeta}{2\zeta^3 + 3k - 1} .$$

If we choose $\zeta = \sqrt{k}$, we find

$$\frac{C(\hat{z})}{C(z^*)} \geq \frac{k}{2\sqrt{k} - 1} ,$$

meaning that we have a lower bound on the Price of Anarchy which is unbounded with increasing k. □

This proposition shows that marginal cost tolls do not necessarily induce optimal equilibria in the setting of mixed autonomy. Moreover, we showed that a lower bound

on the Price of Anarchy with variable marginal cost tolls applied scales with $\sqrt{k}$, where k is the maximum degree of asymmetry.

3.7 Conclusion

In this chapter[3] we explored tolling for multitype mixed autonomy in general networks with affine latency functions, where different vehicle types affect congestion differently. We showed via a clear proof mechanism that our provided differentiated tolls yield socially optimal routing. We provided anonymous tolls with lower bounds on the resulting PoA and upper bound the PoA of *any* anonymous toll scheme, and compared these bounds to the PoA of untolled networks. After establishing the fundamental limitations of anonymous tolls, we provide tolls which are ϵ-differentiated, meaning they have only an infinitesimal difference for the different vehicle types, and showed that these tolls are optimal on parallel networks.

There are a number of important directions for continuing this line of work. There is a gap between the upper bound performance guarantees of our anonymous tolls and the lower bounds that are established via example. It is worthwhile to explore different subclasses of anonymous tolls to understand which induce better worst-case performance versus best-case equilibrium performance. Finally, there is a possibility of extending these ϵ-differentiated tolls to more general settings, or if not, characterizing exactly how much differentiation between tolling types is needed to ensure optimality. Continuing this direction holds promise for understanding the tradeoffs in choosing tolling granularity in the presence of many different vehicle types, and choosing a toll method that can be efficiently and fairly implemented, while preserving traffic network performance efficiency.

[3]This work was supported by NSF grants CCF-1755808, 1736582, and ECCS 1952920, and UC Office of the President grant LFR-18-548175

3.8 Appendix

3.8.1 Proof of Theorem 3.2

To prove this theorem, we use the aggregate cost function provided in Proposition 2.4, then prove a Price of Anarchy bound on this aggregate cost function with tolls applied, and then bound the cost of the optimal routing with respect to the aggregate cost function as compared to the original cost function using Lemma 2.3.

Accordingly, we first define $\hat{f}$, $v_i^{j'}$, ℓ, and L as in eqs. (2.12) to (2.14) and (2.17), and by Prop. 2.4, $C(\hat{z}) = L(\hat{f}(\hat{z}); \hat{z}, c)$, and $\hat{f}(\hat{z})$ is a Wardrop Equilibrium for $L(\cdot; \hat{z}, c)$. In an abuse of notation, we redefine $\tau \in \mathbb{R}^n_{\geq 0}$ as a column vector with elements $\tau_i \in \mathbb{R}_{\geq 0}$. We then define

$$\ell_i^{\tau_i}(f_i; \hat{z}_i, c_i) := \ell_i(f_i; \hat{z}_i, c_i) + \tau_i \text{ and}$$
$$[\ell^\tau(f; \hat{z}, c)]_i := \ell_i^{\tau_i}(f_i; \hat{z}_i, c_i) \ .$$

Then,

$$\begin{aligned} L(\hat{f}(\hat{z}); \hat{z}, c) &= \langle \ell(\hat{f}(\hat{z}); \hat{z}, c), \hat{f}(\hat{z}) \rangle \\ &= \langle \ell^\tau(\hat{f}(\hat{z}); \hat{z}, c), \hat{f}(\hat{z}) \rangle - \tau^T \hat{f}(\hat{z}) \\ &\leq \langle \ell^\tau(\hat{f}(\hat{z}); \hat{z}, c), f^* \rangle - \tau^T \hat{f}(\hat{z}) \\ &= \langle \ell(\hat{f}(\hat{z}); \hat{z}), f^* \rangle - \tau^T (\hat{f}(\hat{z}) - f^*) \\ &\leq L(f^*; \hat{z}, c) + \gamma(\mathcal{C}_k) L(\hat{f}(\hat{z}); \hat{z}, c) \ , \end{aligned}$$

where

$$
\begin{aligned}
\gamma(\mathcal{C}_k) = & \sup_{\substack{c\in\mathcal{C}_k \\ \hat{z}\in\mathbb{R}^{nm}_{\geq 0} \\ f\in\mathcal{F}}} \frac{\langle \ell(\hat{f}(\hat{z});\hat{z},c)-\ell(f;\hat{z},c),f\rangle - \tau^T(\hat{f}-f)}{\langle \ell(\hat{f}(\hat{z});\hat{z},c),\hat{f}\rangle} \\
\leq & \max_{\substack{c_i\in\mathcal{C}_k \\ \hat{z}_i\in\mathbb{R}^m_{\geq 0} \\ f_i\in\mathbb{R}_{\geq 0}}} \frac{(\ell_i(\hat{f}_i(\hat{z}_i);\hat{z}_i,c_i)-\ell_i(f_i;\hat{z}_i,c_i))f_i - \tau_i(\hat{f}_i-f_i)}{\ell_i(\hat{f}_i(\hat{z}_i),\hat{z}_i,c_i)\hat{f}_i} \\
\leq & \max_{\substack{a_i\in\mathcal{A}_k \\ \hat{z}_i\in\mathbb{R}^m_{\geq 0} \\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i\big(\sum_j^m a_i^j\hat{z}_i^j - a_i^{j'}(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j) - a_i^m(\sum_{j\in[m]}\hat{z}_i^j - f_i)\big)}{(\sum_{j\in[m]}a_i^j\hat{z}_i^j)(\sum_{j\in[m]}\hat{z}_i^j)} \\
= & \max_{\substack{a_i\in\mathcal{A}_k \\ \hat{z}_i\in\mathbb{R}^m_{\geq 0} \\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i\big(\sum_j^m (a_i^j-a_i^m)\hat{z}_i^j - (a_i^{j'}-a_i^m)(f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)\big)}{(\sum_{j\in[m]}a_i^j\hat{z}_i^j)(\sum_{j\in[m]}\hat{z}_i^j)} \\
\leq & \max_{\substack{a_i\in\mathcal{A}_k \\ \hat{z}_i\in\mathbb{R}^m_{\geq 0} \\ f_i\in\mathbb{R}_{\geq 0}}} \frac{f_i(a_i^{j'}-a_i^m)\big(\sum_j^m\hat{z}_i^j - (f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)\big)}{a_i^{j'}(\sum_{j\in[m]}\hat{z}_i^j)(\sum_{j\in[m]}\hat{z}_i^j)} \\
\leq & \max_{\hat{z}_i\in\mathbb{R}^m_{\geq 0}, f_i\in\mathbb{R}_{\geq 0}} \frac{(k-1)f_i\big(\sum_j^m\hat{z}_i^j - (f_i - \sum_{j\in[j'-1]}\hat{z}_i^j)\big)}{k(\sum_{j\in[m]}\hat{z}_i^j)(\sum_{j\in[m]}\hat{z}_i^j)} \\
= & \max_{\hat{z}_i\in\mathbb{R}^m_{\geq 0}, f_i\in\mathbb{R}_{\geq 0}} \frac{(k-1)f_i\big(\sum_{j\in[m]}\hat{z}_i^j - f_i\big)}{k(\sum_{j\in[m]}\hat{z}_i^j)(\sum_{j\in[m]}\hat{z}_i^j)}
\end{aligned}
$$

where the first inequality stems from the terms in the denominator being nonnegative. As before, without loss of generality we choose the vehicle indices such that on road i, $a_i^1 \geq a_i^2 \geq \ldots \geq a_i^m > 0$. In the above, we denote this set of possible parameters a_i^j with maximum degree of asymmetry k as $\mathcal{A}_k$. The second inequality stems from the constant term in the latency, b_i, being nonnegative and appearing only in the denominator. The third inequality stems from the fact that a_i^j is monotonically nonincreasing in the index j, and the final inequality is from the definition of the maximum degree of inequality, k.

We can then maximize with respect to f_i by setting $f_i = 1/2 \sum_{j \in [m]} \hat{z}_i^j$. Then,

$$\gamma(\mathcal{C}_k) \leq \frac{(k-1)}{4k} \text{ , so}$$
$$L(\hat{f}(\hat{z}); \hat{z}, c) \leq 4k/(3k-1) L(f^*; \hat{z}, c) \text{ .}$$

Then by Lemma 2.3, $L(f^*; \hat{z}, c) \leq kC(z^*)$, where z^* is an optimal routing with respect to the latency function c. Together this proves the theorem. □

3.8.2 Proof of Theorem 3.3

We prove this theorem constructively. We show that if there exists a routing in the set $\mathcal{Z}^*$ that has a cyclic bipartite graph, we can break each cycle without altering the cost. We do this by showing that for a cyclic routing z', there exists a feasible direction d and that moving in the direction d will not alter the cost and will eventually break the cycle.

The second-order partial derivatives of (3.1) are as follows

$$\frac{\partial^2 C}{\partial z_i^j \partial z_{i'}^{j'}}(z) = \begin{cases} 0 & \text{if } i \neq i' \\ a_i^j + a_i^{j'} & \text{otherwise .} \end{cases}$$

Since we define $z = \left[z_1^1, z_1^2, \ldots, z_1^m, z_2^1, z_2^2, \ldots, z_n^m \right]^T$, the Hessian matrix is therefore block-diagonal in the following form:

$$H = \begin{bmatrix} H_1 & 0 & \ldots & 0 \\ 0 & H_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & H_n \end{bmatrix} ,$$

where the block corresponding to road i is

$$H_i = \begin{bmatrix} 2a_i^1 & a_i^1 + a_i^2 & \dots & a_i^1 + a_i^m \\ a_i^1 + a_i^2 & 2a_i^2 & \dots & a_i^2 + a_i^m \\ \vdots & \vdots & \ddots & \vdots \\ a_i^1 + a_i^m & a_i^2 + a_i^m & \dots & 2a_i^m \end{bmatrix}.$$

Consider a routing z' which minimizes social cost and has a corresponding *cyclic* graph. We will alter this routing and break the cycle while maintaining the same social cost.

Since the graph is bipartite, any cycle will have the same number of nodes of each type. For some cycle in z', let us use $r \in \{2, 3, \dots, \min(m, n)\}$ to denote the number of roads (and therefore vehicle types as well) in a simple cycle. The vehicles and roads are indexed arbitrarily, so let us consider the cycle to be comprised of the first r roads (roads $\{1, 2, \dots, r\}$ and the first r vehicle types (vehicle types $\{1, 2, \dots, r\}$). Let road 1 be shared between vehicle types 1 and 2, road 2 be shared between vehicle types 2 and 3, and so on, until road r which is shared between types r and 1. The remaining roads and vehicle types are indexed arbitrarily.

Based on this feasible routing z', we construct another feasible routing $\tilde{z}$. Consider flow $z' + \alpha d$, where $\alpha \in \mathbb{R}_{\geq 0}$ and d is a direction vector as follows. As with z, let $d = \begin{bmatrix} d_1, d_2, \dots, d_n \end{bmatrix}^T$, where d_i corresponds to the flow change on road i, specifically $d_i = \begin{bmatrix} d_i^1, d_i^2, \dots, d_i^m \end{bmatrix}^T$. We choose $d_1 = \begin{bmatrix} -1, 0, \dots, 0, 1, 0, \dots 0 \end{bmatrix}^T$, where $d_1^r = 1$. We choose $d_2 = \begin{bmatrix} 1, -1, 0, \dots, 0 \end{bmatrix}^T$, and, for $i = \{3, \dots, r\}$, d_i is equal to d_{i-1} circularly shifted downward. For $i \notin [r]$, $d_i^j = 0$ for all $j \in [m]$.

The direction defined above corresponds to shifting some flow of type 1 from road 1 to road 2, some flow of type 2 from road 2 to road 3, and so on, ending in some flow of

type r shifting from road r to road 1. With this defined direction, the flow vector $z' + \alpha d$ is feasible for some range $\alpha \in [0, \overline{\alpha}]$, which we show as follows, using the fact the flow demand is inelastic.

When starting from a feasible flow, moving in the direction d satisfies conservation of flow for inelastic flow demand, as $\sum_{i \in [n]} d_i^j = 0$ for all $j \in [m]$. Moreover, by the definition of flow z', $d_i^j < 0$ only when $z_i'^j > 0$, meaning that in the direction d, flow of a certain vehicle type on a road is decreased only if there exists positive flow of that vehicle type already. As such, we find the maximum feasible range to be $\overline{\alpha} = \min_{k \in [r]} z_k^k$, since at this point the nonnegativity constraint becomes active. The reason it is the minimum of z_k^k, with $k \in [r]$, is because in direction d, some of vehicle type k is shifted off road k. Since at $\alpha = \overline{\alpha}$ the nonnegativity constraint becomes active, at the flow $\tilde{z} = z' + \overline{\alpha} d$, the cycle on roads $[r]$ has been broken. Note that no new cycles have been induced, as d_i^j can only be greater than zero when $z_i'^j$ is greater than zero. Thus, no new edges are added to the graph by moving in direction d.

We now investigate the social cost at $\tilde{z}$. Since the objective function is quadratic (and therefore analytic), by a trivial application of Taylor's inequality we can express $C(\tilde{z})$ as a second-order Taylor expansion around z' as follows.

$$C(\tilde{z}) = C(z') + \langle \nabla C(z'), \overline{\alpha} d \rangle + \frac{1}{2} \overline{\alpha} d^T H(z') \overline{\alpha} d \ ,$$

where $H(z')$ is the Hessian of C evaluated at z'. First-order optimality conditions imply that $\langle \nabla C(z'), \overline{\alpha} d \rangle = 0$, since d only has nonzero elements where inequalities for feasible flow are not tight; thus for z' to be optimal, the derivative of $C(z')$ in the direction of d must be zero. We now investigate the latter term, ignoring the scalar $\frac{\overline{\alpha}^2}{2}$. Since H is

block-diagonal,

$$d^T H d = \sum_{i \in [n]} d_i^T H_i d_i \, .$$

Let us inspect $d_i^T H_i d_i$. For any specific road i', $d_{i'}$ has one entry that is 1 and one that is -1. Let us assign p and p' such that $d_{i'}^p = 1$ and $d_{i'}^{p'} = -1$. Then,

$$\begin{aligned} d_{i'}^T H_{i'} d_{i'} &= \sum_{j' \in [m]} d_{i'}^{j'} \sum_{j \in [m]} (a_{i'}^{j'} + a_{i'}^{j}) d_{i'}^{j} \\ &= \sum_{j' \in [m]} d_{i'}^{j'} [(a_{i'}^{j'} + a_{i'}^{p}) - (a_{i'}^{j'} + a_{i'}^{p'})] \\ &= \sum_{j' \in [m]} d_{i'}^{j'} (a_{i'}^{p} - a_{i'}^{p'}) = (a_{i'}^{p} - a_{i'}^{p'}) - (a_{i'}^{p} - a_{i'}^{p'}) = 0 \, . \end{aligned}$$

Thus, $d^T H d = 0$ as well, so

$$C(\tilde{z}) = C(z') \, .$$

We have thus constructed a flow $\tilde{z}$ which has the same social cost as a socially optimal flow z', which has broken a cycle in the graph representing z' without introducing a new cycle. Since the number of roads and vehicle types is finite, this process can be repeated until we arrive at a flow which has no cycles and has a social cost which optimizes (3.1). This proves the theorem statement. □

3.8.3 Proof of Lemma 3.1

We use the following property of the lemma statement.

Property 3.1. *On each road, all users experience the same toll under the tolling structure*

in the lemma statement. Accordingly,

$$(\forall i \in [n] \wedge j, j' \in [m] \wedge \forall z \in \mathcal{Z})[c_i^j(z) = c_i^{j'}(z)] \; .$$

We now prove the theorem by contradiction. Assume there exists two distinct flows $\tilde{z} \neq \hat{z}$ which are both at equilibrium under tolls $\tau(z^*)$. Since $\tilde{z} \neq \hat{z}$,

$$\exists i \in [n] \wedge j \in [m] \text{ s.t. } \tilde{z}_i^j > \hat{z}_i^j \; . \tag{3.19}$$

From the lemma statement, $i \in \mathcal{N}_j^{z^*}$. Then, resulting from (3.19), either

(i) $\exists j' \in \mathcal{M}_i^{z^*}$ s.t. $\tilde{z}_i^{j'} < \hat{z}_i^{j'}$, or

(ii) $c_i^j(\tilde{z}) > c_i^j(\hat{z})$ (from Assumption 3.1).

If (i), then it must be the case that

(iii) $\exists i' \in \mathcal{N}_{j'}^{z^*}$ s.t. $\tilde{z}_{i'}^{j'} > \hat{z}_{i'}^{j'}$,

due to the demand being inelastic, which implies, from Definition 3.3,

$$c_i^{j'}(\tilde{z}) \geq c_{i'}^{j'}(\tilde{z}) \; . \tag{3.20}$$

Then again, as a result of (iii) either (i) or (ii) must be the case, where i is replaced by i' and j is replaced by j'. This process continues until we reach (ii). This terminal point must exist as the routing's bipartite graph is acyclic. Say the termination point is on vehicle type $j^{(k)}$ on road $i^{(p)}$. Then,

$$c_{i^{(p)}}^{j^{(k)}}(\tilde{z}) > c_{i^{(p)}}^{j^{(k)}}(\hat{z}) \geq c_i^j(\hat{z}) \; , \tag{3.21}$$

where the second relationship results from Property 3.1 and Definition 3.3. Further, due

to Definition 3.3 and Property 3.1,

$$c_{i^{(p)}}^{j^{(k)}}(\tilde{z}) \leq c_{i^{(p-1)}}^{j^{(k)}}(\tilde{z}) = c_{i^{(p-1)}}^{j^{(k-1)}}(\tilde{z}) \leq \ldots \leq c_i^{j'}(\tilde{z}) = c_i^j(\tilde{z}) \ . \tag{3.22}$$

We follow a similar logic down another branch. Reusing the indices j' and i' to a new use, we consider the other result of (3.19). Since the flow demand is inelastic,

$$\exists i' \in \mathcal{N}_j^{z^*} \text{ s.t. } \tilde{z}_{i'}^j < \hat{z}_{i'}^j \ . \tag{3.23}$$

Next, consider the results of (3.23). Similarly to above, either

(iv) $\exists j' \in \mathcal{M}_{i'}^{z^*}$ s.t. $\tilde{z}_{i'}^{j'} > \hat{z}_{i'}^{j'}$ (where $j' \in \mathcal{M}_{i'}^{z^*}$ due to the lemma statement), or

(v) $c_{i'}^{j'}(\tilde{z}) < c_{i'}^{j'}(\hat{z})$ (from Assumption 3.1).

If (iv), then it must be the case that

(vi) $\exists i'' \in \mathcal{N}_{j'}^{z^*}$ s.t. $\tilde{z}_{i''}^{j'} < \hat{z}_{i''}^{j'}$,

due to the demand being inelastic, which implies, from Assumption 3.1,

$$c_{i''}^{j'}(\tilde{z}) \geq c_{i'}^{j'}(\tilde{z}) \ . \tag{3.24}$$

Similarly to above, as a result of (vi), either (iv) or (v) must be the case, where i' is replaced by i'' and j is replaced by j'. This process continues until we reach (vi). Say the termination point is on vehicle type $j^{(r)}$ on road $i^{(s)}$. Then,

$$c_{i^{(s)}}^{j^{(r)}}(\tilde{z}) < c_{i^{(s)}}^{j^{(r)}}(\hat{z}) \leq c_i^j(\hat{z}) \ , \tag{3.25}$$

where the second relationship results from Property 3.1 and Definition 3.3. Further, due

to Definition 3.3 and Property 3.1,

$$c_{i^{(s)}}^{j^{(r)}}(\tilde{z}) \geq c_{i^{(s-1)}}^{j^{(r)}}(\tilde{z}) = c_{i^{(s-1)}}^{j^{(r-1)}}(\tilde{z}) \geq \ldots \geq c_i^{j'}(\tilde{z}) = c_i^j(\tilde{z}) \; . \tag{3.26}$$

We then string together (3.21) and (3.25) to find

$$c_{i^{(p)}}^{j^{(k)}}(\tilde{z}) > c_{i^{(s)}}^{j^{(r)}}(\tilde{z}) \; .$$

We find a contradiction by combining (3.22) and (3.26): $c_{i^{(p)}}^{j^{(k)}}(\tilde{z}) \leq c_{i^{(s)}}^{j^{(r)}}(\tilde{z})$.

This contradiction proves that any resulting equilibrium for a flow demand is unique, which is the first lemma statement. To prove the second statement, we note that the tolls are the path tolls of [24], which, for a flow demand equal to that of the demand used to generate z^*, the flow for determining the tolls, guarantees the existence of an equilibrium with flow equal to z^*. This existence, with the uniqueness proven in the previous statement, proves the second lemma statement. □

3.8.4 Proof of Lemma 3.2

The reasoning for this statement is as follows. Due to the fact that j was not on i in z^*, and from the structure of the tolls, we know that $c_i^{\tau_j}(\hat{z}_i) = c_i^{\overline{\tau}}(\hat{z}_i) + \epsilon$. By definition of Wardrop Inequality, due to $\hat{z}_i^j > 0$, we know $c_i^{\tau_j}(\hat{z}_i) \leq c_{i'}^{\tau_j}(\hat{z}_{i'})$ for all other roads i' in $[n]$. Then, there is some road i' such that $c_{i'}^{\tau_j}(\hat{z}_{i'}) = c_{i'}^{\overline{\tau}}(\hat{z}_{i'})$, (since there is at least one road i' with positive flow of type j in z^*). From Wardrop Equilibrium we find $c_i^{\tau_j}(\hat{z}_i) \leq c_{i'}^{\tau_j}(\hat{z}_{i'})$. As we've stated, $c_i^{\tau_j}(\hat{z}_i) = c_i^{\overline{\tau}}(\hat{z}_i) + \epsilon$, and since $i' \in \mathcal{N}_j^{z^*}$, we have $c_{i'}^{\tau_j}(\hat{z}_{i'}) = c_{i'}^{\overline{\tau}}(\hat{z}_{i'})$. This gives us

$$c_i^{\overline{\tau}}(\hat{z}_i) + \epsilon = c_{i'}^{\overline{\tau}}(\hat{z}_{i'}) \; . \tag{3.27}$$

From Proposition 3.3 and the definitions of $\underline{c}$ and $\overline{c}$ in (3.10), we can state that

$$\begin{aligned} c_i^{\overline{\tau}}(\hat{z}_i) &= \underline{c} \quad \text{and} \\ c_{i'}^{\overline{\tau}}(\hat{z}_{i'}) &= \overline{c} \ . \end{aligned} \tag{3.28}$$

□

3.8.5 Proof of Lemma 3.3

By definition of $\mathcal{I}_1$,

$$\sum_{i \in \mathcal{I}_1} \hat{z}_i^j = \sum_{i \in \mathcal{I}_1} z_i^{*j} \quad \forall i' \in \mathcal{I}_1, \ \forall j \in \mathcal{M}_{i'}^{z^*} \ . \tag{3.29}$$

In words, all vehicle types on $\mathcal{I}_1$ in z^* will be on those roads as well in $\hat{z}$, in addition to some vehicle flow that was not there in z^*. Then, we can show by contradiction that the resulting standard cost $c_i^{\overline{\tau}}$ for the roads in $\mathcal{I}_1$ will be greater than or equal to μ as follows. From (3.15), we know that the standard cost will be the same on all roads in $\mathcal{I}_1$. By definition of the toll structure, the standard cost of the roads in $\mathcal{I}_1$ with routing z^* will be μ. In $\hat{z}$ there is not less flow of any type to route on $\mathcal{I}_1$, and cost functions are nondecreasing in the vehicle flows. If $c_i^{\overline{\tau}}(\hat{z}_i) = c_i(\hat{z}_i) + \mu - c_i(z_i^*) < \mu$ for all roads in $\mathcal{I}_1$, then z^* would not be an optimal routing, since this new routing can decrease the social cost from what it was in z^*. By this contradiction,

$$c_i^{\overline{\tau}}(\hat{z}_i) \geq \mu \quad \forall i \in \mathcal{I}_1 \ . \tag{3.30}$$

□

3.8.6 Proof of Lemma 3.4

By definition of $\mathcal{I}_2$,

$$\sum_{i \in \mathcal{I}_2} \hat{z}_i^j \leq \sum_{i \in \mathcal{I}_2} z_i^{*j} \quad \forall j \in [m] \,, \tag{3.31}$$

and each road has only the vehicle flow types that it did in z^*, since $\mathcal{I}_1 \cap \mathcal{I}_2 = \emptyset$ (from their different values of $c_i^{\overline{\tau}}$ of their roads – otherwise we arrive at a contradiction earlier on). From Lemma 3.1, each road having the vehicle flow types that it did in z^* implies the uniqueness of equilibrium flow. Since the cost functions are nondecreasing in the vehicle flows, the equilibrium flow $\hat{z}$ has standard cost less than or equal to that of the routing in z^*, which is μ. Accordingly,

$$c_i^{\overline{\tau}}(\hat{z}_i) \leq \mu \quad \forall i \in \mathcal{I}_2 \,. \tag{3.32}$$

□

Chapter 4

Incentivizing Efficient Equilibria

In this chapter[1], we consider a setting in which, instead of a consumer product, autonomous vehicles are offered as a ride-hailing service. We consider a road network shared between human drivers, which choose their routes to minimize latency, and autonomous vehicles, which offer their riders choices between different routes, each with an associated price. In this case, a system designer can set the route prices so as to incentivize the users to choose routes, that in the presence of selfishly optimizing human drivers, lead to a desired network state. In order to facilitate this, we model how humans choose between various options with different prices and latencies, and leverage an *active learning* framework to learn the preferences of a population. Through a user study, we empirically show this method to be an efficient and accurate way to learn preferences. We then formulate an optimization to select prices which will yield the best network equilibrium. Moreover, we theoretically characterize the range of outcomes possible in

[1]This chapter is adapted, with permission, from the manuscripts "Altruistic autonomy: Beating congestion on shared roads", previously published by International Workshop on the Algorithmic Foundations of Robotics, © 2018 Springer, "The green choice: Learning and influencing human decisions on shared roads", previously published by IEEE Conference on Decision and Control, © 2019 IEEE, and "Incentivizing Efficient Equilibria in Traffic Networks with Mixed Autonomy", previously published by IEEE Transactions on Control of Network Systems, © 2021 IEEE, and is joint work with Erdem Bıyık, Dorsa Sadigh, and Ramtin Pedarsani [91–93].

this setting and epirically show that our algorithm performs well compared to the range of outcomes.

4.1 Introduction

Road congestion is a major and growing source of inefficiency, costing drivers in the United States billions of dollars in wasted time and fuel [3]. As established in Chapter 2, autonomy alone will not suffice in alleviating road congestion, especially in the mixed-autonomy setting. Accordingly, in Chapter 3 we analyzed tolling to mitigate the inefficiency due to selfish routing, and established worst-case guarantees. Implicit in this model was the assumption that autonomous vehicles will be available as consumer products, and therefore tolling will equivalently affect the choices of users of human-driven and autonomous vehicles. This vision of the future, with autonomous vehicles as consumer products, is not the only possible outcome.

There are many challenges faced in creating autonomous vehicles (AVs) as a consumer product. Prohibitive factors include that for AVs as consumer products, the sensors required for most autonomous technology is currently prohibitively expensive and require specialized servicing, the AVs will need to be able to operate in any location and under any weather condition, and the infrastructure, including maps and servicing, must be available in all places [94]. Due to these limitations, there are currently companies which, instead of offering AVs as consumer products, offer AVs as a ride-hailing service [95, 96].

In this setting, a social planner wishing to reduce congestion will no longer have equal influence over autonomous and human-driven vehicles, and will instead need to rely on the autonomous vehicles to influence the behavior of the human-driven vehicles. In the setting of a congestion game, like the one used in Chapters 2 and 3, this yields a Stackelberg Game – a game in which a planner chooses the actions of some fraction of

users and the remaining users respond with an equilibrium strategy. However, it has been definitively shown that in general, in Stackelberg Games one cannot improve worst-case guarantees beyond those that exist without any controlled flow [28]. Accordingly, in this chapter we move towards a more realistic model and provide an algorithm for determining how to leverage pricing control over the autonomous ride hailing service to improve network performance. Instead of providing worst-case guarantees for our algorithm, in this framework we will theoretically determine the range of possible outcomes on the network in order to gauge the efficacy of our pricing algorithm.

Specifically, we consider a setting where a benevolent ride hailing service or social planner sets prices for each route that an autonomous vehicle user may take. The users of the autonomous service will choose their routes, or choose not to travel, based on their time-money tradeoff. The remaining human drivers will choose routes that minimize their latency. The role of the social planner is then to choose prices to optimize some social objective. Though we no longer consider a congestion game, our model is an indirect Stackelberg game – a game in which a social planner controls some fraction of the population's actions, where the remainder of the population responds selfishly. However, in our model the planner only controls its portion of the vehicle flow indirectly, via pricing.

To effectively set prices, the social planner needs a model for a) the flow of vehicles on a road, which depends on how many autonomous and human-driven vehicles are on the road, and b) how people make decisions between routes with various prices and latencies. We model the latter based on multinomial logits [97] and model human drivers as selfish agents who reach a *Nash Equilibrium* (which we called a Wardrop Equilibrium in the setting of the previous chapters), a configuration in which no one can decrease their travel time by switching paths [17].

Moreover, we use active preference-based learning [98, 99], via a series of queries, to

understand the preferences of autonomous service users. This enables the planner to predict how autonomous service users will react to a set of options with only a relatively low number of training queries. We experimentally verify this method accurately predicts human choice, and that our planning algorithm can use this to improve network performance. Furthermore, we provide a theoretical framework which establishes benchmarks for the performance of the algorithm and we show that these benchmarks can be calculated in polynomial time. Finally, we show the efficacy of our pricing scheme in the context of these benchmarks.

Our contributions in this work are as follows:

- We develop an active preference-based algorithm to learn how different people value time and money in choosing their transportation option. This enables learning a model for people's routing choices in a data-efficient manner.
- We formulate and solve an optimization for ride hailing service to minimize congestion and maximize the road network utilization while constraining a minimum profit for the service supplier.
- We provide theoretical benchmarks for understanding the performance of the above and prove that we can calculate these benchmarks in polynomial time.
- We validate the learning algorithm and the planning optimization formulation via simulations and user studies. Our results show carefully designed pricing schemes can significantly improve traffic throughput.

4.1.1 Related Works

Previous works have shown the potential benefits that autonomous vehicles can have for traffic networks by increasing road capacity through platooning [31, 32], damping shockwaves of slowing vehicles [35, 36], managing merges [37], decongesting freeways in

the event of accidents [39], and balancing a supply of vehicles [100]. Relatedly, many works analyze and bound the inefficiency that can arise from network users choosing their routes selfishly [19], including when autonomous vehicles are introduced [72].

As we use financial incentives to influence this behavior, our formulation is related to work in tolling, some of which consider users with different sensitivities to tolls [25, 26]. [101] considers a congestion game framework and derives tolls which drive users to choose socially optimal strategies for a broad class of user strategy update dynamics. In contrast to many of these works, we consider an empirically validated probabilistic model for human choice [102] which incorporates differences in price sensitivity, and we model vehicle flow on shared roads based on the foundational Fundamental Diagram of Traffic (FDT) [34]. Our FDT-based model shares fundamental similarities with the one developed in [33]. Some works consider heterogeneous and stochastic user utilities (*e.g.* [103], [104]). However, the approaches taken in these works cannot incorporate a FDT-based model for road congestion, in which latency is no longer an increasing function of vehicle flow on a road. Also relatedly, [29] considers a Stackelberg game (meaning a planner can route a portion of the vehicle flow) on parallel roads with flow dictated by the FDT when there is a single vehicle type. Hence, a major novelty of our work lies in that we consider a mixed-autonomy network where autonomous service users have different preferences.

To understand human choice, there has recently been much effort on learning human reward functions which are assumed to be sufficient to model preferences. Inverse reinforcement learning [105–107] and preference-based learning [98, 99, 108–110] are the most popular choices. In this paper, we employ preference-based learning, a natural fit to our problem. We actively synthesize queries – a non-trivial generalization and extension of [98] – for data-efficiency and better usability.

4.2 Problem Setting and Objective

4.2.1 Vehicle Flow Model

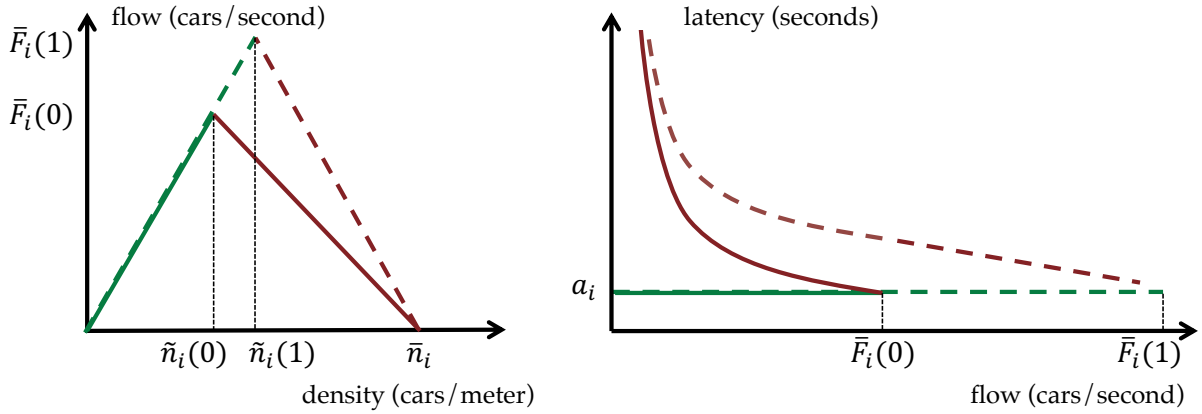
We assume every road p has a maximum flow. This occurs when traffic is in *free-flow* – when all vehicles travel at the nominal road speed $\bar{v}_p$.

Definition 4.1. *The* free-flow latency *of a road p, denoted t_p, is the time it takes vehicles to traverse the road in free-flow. With road length denoted d_p, the free-flow latency is $t_p := d_p/\bar{v}_p$.*

Traffic Density vs Traffic Flow: Adding more cars to a road that is already at maximum flow makes the traffic switch from free-flow to a congested regime, which decreases the vehicle flow. In the extreme case, at a certain density $\bar{\phi}_p$, cars are bumper-to-bumper and vehicle flow stops. The solid lines in Fig. 4.1(a) – Fundamental Diagram of Traffic [34] – illustrates this phenomenon, where flow increases linearly with respect to density until it hits the critical density. The slope corresponds to the free-flow velocity $\bar{v}_p$ on road i. After the critical point, flow decreases linearly until it is zero at the maximum density. In this figure, $\bar{F}$ is the maximum flow and $\tilde{n}$ denotes the critical density, where the argument is the fraction of vehicles that are autonomous. We will formally define this notation below.

Traffic Flow vs Road Latency: The relationship between vehicle flow and road latency reflects the same free-flow/congested divide. As mentioned above, roads in free-flow have constant latency. In the congested regime, however, latency *increases* as vehicle flow decreases, since a high density of vehicles is required to achieve a low traffic flow. This is represented in Fig. 4.1(b) with the solid lines.

Mixed-Autonomy Roads. We assume that on mixed-autonomy roads, the autonomous vehicles can coordinate with one another and potentially form platoons to help with the



(a) The Fundamental Diagram of Traffic for roads with all human-driven (solid) and all autonomous (dashed) vehicles. In the latter, congestion begins at a higher vehicle density as autonomous vehicles require a shorter headway when following other vehicles.

(b) The relationship between vehicle flow and latency also changes in the presence of autonomous vehicles. Free-flow speed remains the same but maximum flow on a road increases.

Figure 4.1: Vehicle flow model for mixed autonomy.

efficiency of the road network. We now extend the traffic model above to mixed-autonomy settings as shown with dashed lines in Fig. 4.1. We define the autonomy level of a road p as the fraction of autonomous vehicle flow on that road: $\alpha_p := \frac{f_p^{\mathrm{a}}}{f_p^{\mathrm{a}}+f_p^{\mathrm{h}}}$ where f_p^{a} and f_p^{h} represent the autonomous and human-driven vehicle flow, respectively. Assuming that neither the nominal velocity $\bar{v}_p$ nor the maximum density $\bar{\phi}_p$ changes, the critical density at which traffic becomes congested will now shift and increase with *autonomy level* α_p, as platooned autonomous vehicles require a shorter headway than human drivers.

To formalize the relationship between autonomy level and critical density on road i, we assume the space occupied by autonomous vehicles and humans at nominal velocity is h_p^{a} and h_p^{h}, respectively, with $h_p^{\mathrm{a}} \leq h_p^{\mathrm{h}}$. This inequality reflects the assumption that autonomous vehicles can maintain a short headway, regardless of the type of vehicle they are following. Then, the critical density is

$$\tilde{n}_p(\alpha_p) := \frac{b_p}{\alpha_p h_p^{\mathrm{a}} + (1-\alpha_p) h_p^{\mathrm{h}}} , \tag{4.1}$$

where b_p is the number of lanes on that road. Here the denominator represents the average length from one car's rear bumper to the preceding car's rear bumper when all cars follow the vehicle in front of them with nominal headway. Note that critical density is expressed here as a function of the autonomy level α_p of the road. Since flow increases linearly with density until hitting the critical point, the maximum flow can also be expressed as a function of autonomy level: $\bar{F}_p(\alpha_p) = \bar{v}_p \tilde{n}_p(\alpha_p)$.

The flow on a road, $f_p = f_p^{\mathrm{h}} + f_p^{\mathrm{a}}$ is a function of the density (ϕ_p^{h} and ϕ_p^{a}, respectively) of each vehicle type as follows.

$$f_p(\phi_p^{\mathrm{h}}, \phi_p^{\mathrm{a}}) := \begin{cases} \bar{v}_p \cdot (\phi_p^{\mathrm{h}} + \phi_p^{\mathrm{a}}), & \text{if } \phi_p^{\mathrm{h}} + \phi_p^{\mathrm{a}} \leq \tilde{n}_p(\alpha_p) \\ \frac{\bar{v}_p \cdot \tilde{n}_p(\alpha_p) \cdot (\bar{\phi}_p - (\phi_p^{\mathrm{h}} + \phi_p^{\mathrm{a}}))}{\bar{\phi}_p - \tilde{n}_p(\alpha_p)}, & \text{if } \tilde{n}_p(\alpha_p) \leq \phi_p^{\mathrm{h}} + \phi_p^{\mathrm{a}} \leq \bar{\phi}_p \\ 0, & \text{otherwise} \quad . \end{cases} \tag{4.2}$$

We can then write the latency as a function of vehicle flow as well as a binary argument s_p, which indicates whether the road is congested [1, 29]:

$$\ell_p(f_p^{\mathrm{h}}, f_p^{\mathrm{a}}, s_p) = \begin{cases} \frac{d_p}{\bar{v}_p} & \text{if } s_p = 0 \\ d_p \left(\frac{\bar{\phi}_p}{f_p^{\mathrm{h}} + f_p^{\mathrm{a}}} + \frac{\tilde{n}_p(\alpha_p) - \bar{\phi}_p}{\bar{v}_p \cdot \tilde{n}_p(\alpha_p)} \right) & \text{if } s_p = 1 \ . \end{cases} \tag{4.3}$$

Fig. 4.1(b) illustrates the effect of mixed autonomy on latency.

4.2.2 Network Model

Assumptions 4.1. *We consider a network of n parallel roads. We assume that no two roads have the same free-flow latency. We order the indices such that* $t_1 < t_2 < \ldots < t_n$.

The role of the assumption above is explained in the Appendix. We use $[k]$ to denote the set of the first k roads; accordingly, $\mathcal{P}$ denotes the set of all roads.

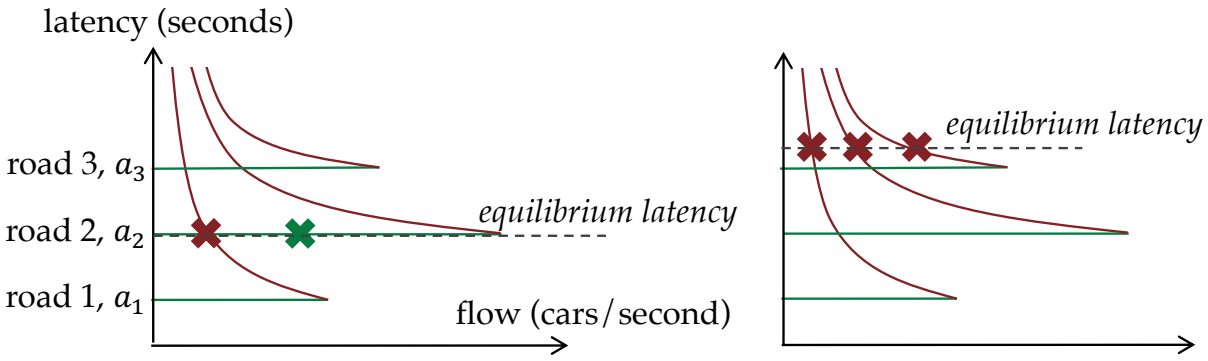
We describe the network state by $(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s})$, where $\boldsymbol{f}^{\mathrm{h}}$, $\boldsymbol{f}^{\mathrm{a}} \in \mathbb{R}^n_{\geq 0}$ and $\boldsymbol{s} \in \{0,1\}^n$. A *feasible routing* is one for which $f^{\mathrm{h}}_p + f^{\mathrm{a}}_p \leq \bar{F}_p(\alpha_p)$ for all roads, and the flow of each vehicle type on the roads sum to $\bar{\lambda}^{\mathrm{h}}$ and $\bar{\lambda}^{\mathrm{a}}$ in the case of inelastic demand, respectively, which denote the total vehicle flow demands. In the case where the demand for autonomous vehicles is elastic, i.e., Section 4.4, this constraint is relaxed such that the total autonomous flow does not exceed the maximum demand. We are interested in finding a routing, i.e. allocation of vehicles into the roads, that minimizes the total latency experienced by all vehicles while maximizing the total flow of the roads in the case of elastic demand. Further, we constrain this optimization based on *total demand*, *selfishness* or *flexibility* of the vehicles. While selfish drivers always take the quickest road possible, flexible vehicles accept relatively longer latencies. We will formalize these terms in Section 4.3.

4.2.3 Human Driver Choice Model

We assume human drivers are selfish, i.e. their only consideration is minimizing their own commute time. This leads to a *Nash Equilibrium* [17], which, on parallel roads, means that if one road has positive flow on it, all other roads must have higher or equal latency. Formally,

$$f^{\mathrm{h}}_p > 0 \implies \ell_p(f^{\mathrm{h}}_p, f^{\mathrm{a}}_p, s_p) \leq \ell_{p'}(f^{\mathrm{h}}_{p'}, f^{\mathrm{a}}_{p'}, s_{p'}) \; \forall p, p' \in \mathcal{P} \; .$$

This implies that all selfish users experience the same latency. It is therefore useful to consider the flow-latency diagrams of roads when studying which equilibria exist – by fixing the latency on the y-axis, one can reason about which roads must be congested to achieve that equilibrium. As shown in Fig. 4.2, equilibria may have one road in free-flow and rest congested, or all may be congested [1, 29].



(a) Equilibrium with one road in free-flow.

(b) Equilibrium with all roads congested.

Figure 4.2: Some possible equilibria of a three-road network with fixed flow demand. Green and red lines denote the free-flow and congested regimes, respectively. An equilibrium has an associated *equilibrium latency* experienced by all selfish users. By considering a given equilibrium latency, we can reason about which roads must be congested at that equilibrium as well how much flow is on each road.

4.2.4 System Objective

In this paper, we are interested in developing a pricing scheme for autonomous vehicles that improves the state of the traffic by alleviating the adverse effects of selfishness. First, we develop two baselines: the first is the case in which all users, including autonomous vehicles, are selfish and the network reaches a *Nash Equilibrium*. In this case we wish to efficiently calculate the Nash Equilibrium that minimizes overall travel latency. Our goal is to achieve lower travel latencies than this equilibrium with the same amount of flow. In the second baseline, we assume we have limited direct control over the routing of autonomous users – more specifically, we can route autonomous vehicles as we wish as long as the latency they experience is within some range of the quickest route available. A specific case of this baseline with full flexibility of the autonomous vehicles serves as a lower bound to our method.

After developing the two baselines in Section 4.3, we present our pricing method that provides the same benefits as flexible behavior through financial incentives in Section 4.4.

For our pricing scheme, we describe the various facets of the problem. We assume the demand of human drivers is fixed, and the demand of people using the autonomous service is elastic – *if prices and latencies are high, some people may choose not to use the autonomous mobility service.* Our goal is then to simultaneously maximize the number of autonomous service users that can use the road, and minimize the average latency experienced by all the people using the roads.

4.3 Performance Benchmarks

Throughout this section, we assume an inelastic demand, i.e., we will route all autonomous and human-driven flow demand, $\bar{\lambda}^{\mathrm{a}}$ and $\bar{\lambda}^{\mathrm{h}}$, into the network. As the demand is inelastic, we are only interested in finding a routing that minimizes the total latency experienced by all vehicles, $C(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}) = \sum_{p\in\mathcal{P}}(f_p^{\mathrm{h}} + f_p^{\mathrm{a}})\ell_p(f_p^{\mathrm{h}}, f_p^{\mathrm{a}}, s_p)$, while satisfying the demand, i.e. $\sum_{p\in\mathcal{P}} f_p^{\mathrm{h}} = \bar{\lambda}^{\mathrm{h}}$ and $\sum_{p\in\mathcal{P}} f_p^{\mathrm{a}} = \bar{\lambda}^{\mathrm{a}}$.

We now make precise the aforementioned notions of selfishness and flexibility. We develop properties of the resulting equilibria, and using those, provide polynomial-time algorithms for computing the benchmark flows.

Selfishness. Human drivers are often thought of as selfish, meaning they will not take a route with long latency if a quicker route is available to them. If all drivers are selfish this leads to a *Nash Equilibrium*, in which no driver can achieve a lower travel time by unilaterally switching routes [17]. This means that all selfish users with the same origin and destination experience the same travel time.

Definition 4.2. *The* **longest equilibrium road** *is the road with maximum free-flow latency which has latency equal to the latency experienced by selfish users. Let* m_{EQ} *denote the index of this road. We use* $NE(\bar{\lambda}^h, \bar{\lambda}^a, m_{\mathrm{EQ}})$ *to denote the set of Nash Equilibria with* longest equilibrium road *having index* m_{EQ}.

Definition 4.3. *The* **longest used road** *is the road with maximum free-flow latency that has positive vehicle flow of any type on it. We use* m_{ALL} *to denote the index of this road; if all vehicles in a network are selfish then* $m_{\text{EQ}} = m_{\text{ALL}}$.

The following lemma will help with the subsequent theoretical results; we defer its proof to the appendix.

Lemma 4.1. *If the set of Nash Equilibria contains a routing with positive flow only on roads* $[m]$*, then there exists a routing in the set of Nash Equilibria with positive flow only on roads* $[m']$ *where* $m' \leq m$*, and road* m' *is in free-flow.*

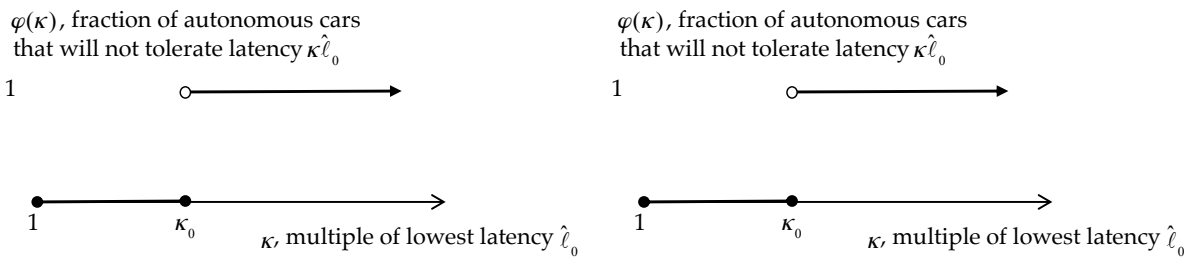
We define the set of ***Best-case Nash Equilibria (BNE)*** as the set of feasible routings in equilibrium that minimize the total latency for flow demand $(\bar{\lambda}^{\text{h}}, \bar{\lambda}^{\text{a}})$, denoted $\text{BNE}(\bar{\lambda}^{\text{h}}, \bar{\lambda}^{\text{a}})$. The following theorem (proof deferred to the appendix) provides properties of the set of BNE for mixed-autonomy roads (for roads with a single vehicle type, see [29]).

Theorem 4.1. *There exists a road index* m^*_{EQ} *such that all routings in the set of BNE have the below properties. Further, this index* m^*_{EQ} *is the minimum index such that a feasible routing can satisfy the properties:*

1. *road* m^*_{EQ} *is in free-flow,*
2. *roads with index less than* m^*_{EQ} *are congested with latency* $a_{m^*_{\text{EQ}}}$*, and*
3. *all roads with index greater than* m^*_{EQ} *have zero flow.*

As the same latency level can be achieved by varying the autonomy levels of the roads, BNE is not necessarily unique.

Flexibility. We also wish to find a lower bound for the social cost when some users are willing (or incentivized) to take longer routes. To that end, we use the term *flexibility profile* to refer to the distribution of the degree to which autonomous users are willing to endure longer routes. For computational reasons, we consider flexibility profiles with a finite number of flexibility levels.



(a) Users will tolerate latency of up to κ_0 times that of the quickest route.

(b) Users have multiple flexibility levels.

Figure 4.3: Flexibility profiles. A fraction $\varphi(\kappa)$ of autonomous users will not accept latency greater than κ times that of the quickest available route.

Formally, we define $\varphi : \mathbb{R}_{\geq 0} \to [0, 1]$ to represent the flexibility profile as a non-decreasing function of a latency value that is mapped to $[0, 1]$. A volume of $\varphi(\kappa)\bar{\lambda}^{\mathrm{a}}$ autonomous flow will reject a route incurring latency κ times the minimum route latency available, which we denote $\ell_{m_{\mathrm{EQ}}}$. If autonomous users have a uniform flexibility level as in Fig. 4.3(a), we call them κ_0-*flexible* users, where κ_0 is the maximum multiple of the minimum latency that autonomous users will accept. Users may have differing flexibility levels, as in Fig. 4.3(b). We use K to denote the set of flexibility levels, with cardinality $|K|$. Accordingly, a feasible routing $(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s})$ is in the set of ***Flexible Nash Equilibria (FNE)*** if

1. all routes with human traffic have latency $\ell_{m_{\mathrm{EQ}}} \leq \ell_p(f_p^{\mathrm{h}}, f_p^{\mathrm{a}}, s_p)\ \forall p \in \mathcal{P}$ and
2. for any $\ell \geq 0$, a volume of at least $\varphi(\ell/\ell_{m_{\mathrm{EQ}}})\bar{\lambda}^{\mathrm{a}}$ autonomous traffic experiences a latency less than or equal to ℓ. Note that it is sufficient to check this condition for $\ell = \ell_p(f_p^{\mathrm{h}}, f_p^{\mathrm{a}}, s_p)$ for all p.

We denote the set of routings at Flexible Nash Equilibria with demand $(\bar{\lambda}^{\mathrm{h}}, \bar{\lambda}^{\mathrm{a}})$, equilibrium latency $\ell_{m_{\mathrm{EQ}}}$, and flexibility profile φ as $\mathrm{FNE}(\bar{\lambda}^{\mathrm{h}}, \bar{\lambda}^{\mathrm{a}}, \ell_{m_{\mathrm{EQ}}}, \varphi)$. The set of ***Best-case Flexible Nash Equilibria (BFNE)*** is the subset of FNE with routings that minimize total latency. Note that as in Theorem 4.1, we use m^*_{EQ} to denote the road with longest free-flow latency that contains selfish vehicle flow in the best routing within

the considered set of equilibria. We defer the proof of the following to the appendix.

Theorem 4.2. *For any given routing in the set of BFNE, there exist a longest equilibrium road m^*_{EQ} and a longest used road m^*_{ALL} with $m^*_{\mathrm{EQ}} \leq m^*_{\mathrm{ALL}}$, such that:*

1. *roads with index less than m^*_{EQ} are congested,*
2. *roads with index greater than m^*_{EQ} are in free-flow,*
3. *roads with index greater than m^*_{EQ} and less than m^*_{ALL} have maximum flow.*

Remark 4.1. *Note that, unlike in BNE, road m^*_{EQ} will not necessarily be in free-flow and m^*_{EQ} is not necessarily the minimum index such that all selfish traffic can be feasibly routed at Nash Equilibrium [1]. Further, different elements of the set BFNE can have different indices for longest equilibrium and longest used road.*

Finding the Best-case Nash Equilibria. In general, the Nash Equilibrium constraint is a difficult combinatorial constraint. Theorem 4.1 however states that we can characterize the congestion profile of the roads by finding the minimum free-flow road such that Nash Equilibrium can be feasibly achieved. This is formalized as follows: find the minimum m_{EQ} such that $\mathrm{NE}(\bar{\lambda}^{\mathrm{h}}, \bar{\lambda}^{\mathrm{a}}, m_{\mathrm{EQ}})$ is nonempty:

$$m^*_{\mathrm{EQ}} = \operatorname*{argmin}_{m_{\mathrm{EQ}} \in \mathcal{P}} a_{m_{\mathrm{EQ}}} \text{ s.t. } \mathrm{NE}(\bar{\lambda}^{\mathrm{h}}, \bar{\lambda}^{\mathrm{a}}, m_{\mathrm{EQ}}) \neq \emptyset \text{ , then}$$
$$\mathrm{BNE}(\bar{\lambda}^{\mathrm{h}}, \bar{\lambda}^{\mathrm{a}}) \subseteq \mathrm{NE}(\bar{\lambda}^{\mathrm{h}}, \bar{\lambda}^{\mathrm{a}}, m^*_{\mathrm{EQ}}). \tag{4.4}$$

Theorem 4.3. (4.4) *can be solved in $O(N^4)$ time.*

We defer the proof to the appendix.

Finding the Best-case Flexible Nash Equilibria. To find an element of the BFNE,

we need to solve:

$$\underset{\substack{m_{\mathrm{EQ}}\in\mathcal{P},\ \hat{\ell}_0\in[a_{m_{\mathrm{EQ}}},a_{m_{\mathrm{EQ}}+1}),\\ (\boldsymbol{f}^{\mathrm{h}},\boldsymbol{f}^{\mathrm{a}},\boldsymbol{s})\in\mathrm{FNE}(\bar{\lambda}^{\mathrm{h}},\bar{\lambda}^{\mathrm{a}},\hat{\ell}_0,\varphi)}}{\operatorname{argmin}} C(\boldsymbol{f}^{\mathrm{h}},\boldsymbol{f}^{\mathrm{a}},\boldsymbol{s})\ . \tag{4.5}$$

As demonstrated in [1], the longest equilibrium road is no longer the road with lowest free-flow latency such that the routing is feasible, as was the case in BNE. Further, road m_{EQ} may not be in free-flow in the set of BFNE. However, we do know that for a fixed m_{EQ}, the latency on road m_{EQ} which minimizes cost, subject to feasibility constraints, is one of a finite number of options.

Theorem 4.4. *Finding a solution to* (4.5) *is equivalent to finding a routing in the set of BFNE, if any exist. Further,* (4.5) *can be solved in* $O(|K|N^5)$ *time, where* $|K|$ *is the number of flexibility levels of autonomous vehicle users.*

Proof. First, Definition 4.2 implies that the latency on the longest equilibrium road m_{EQ}, which we denote $\hat{\ell}_0$, must be less than that of road $m^*_{\mathrm{EQ}}+1$. This, with the definition of BFNE imply that (4.5) solves for an element of the BFNE. Now, note that for a given $\hat{\ell}_0$, the optimal routing will maximize the autonomous flow on roads $[m_{\mathrm{EQ}}]$. We show that this can be computed in $O(N^3)$ time, and the optimal allocation of the remaining autonomous flow can be computed in $O(N)$ time. Next, we note that subject to feasibility, the social cost of a routing decreases monotonically with $\hat{\ell}_0$. In light of this, we show that there are a maximum of $k|N|$ critical points to check for feasibility when searching for the optimal $\hat{\ell}_0$.

We now show that given $\hat{\ell}_0$, the latency on road m_{EQ}, computing the optimal flow on roads $[m_{\mathrm{EQ}}]$ can be done in $O(N^3)$ time. For a given $\hat{\ell}_0$, the optimal routing will fit as much autonomous flow as possible on roads $[m_{\mathrm{EQ}}]$. Let $\boldsymbol{f}^{\mathrm{h,EQ}}$ and $\boldsymbol{f}^{\mathrm{a,EQ}}$ denote the elements of the regular and autonomous routings $\boldsymbol{f}^{\mathrm{h}}$ and $\boldsymbol{f}^{\mathrm{a}}$ that correspond to flows on

the roads $[m_{\mathrm{EQ}}]$, as with $\boldsymbol{s}^{\mathrm{EQ}}$. Then, solve

$$\begin{aligned}
\underset{\boldsymbol{f}^{\mathrm{h,EQ}},\boldsymbol{f}^{\mathrm{a,EQ}}\in\mathbb{R}_{\geq 0}^{|[m_{\mathrm{EQ}}]|},s_{m_{\mathrm{EQ}}}\in\{0,1\}}{\operatorname{argmax}} & \sum_{p\in[m_{\mathrm{EQ}}]} f_i^{\mathrm{a}} \\
\text{s.t.} \quad & \sum_{p\in[m_{\mathrm{EQ}}]} f_p^{\mathrm{h}} = \bar{\lambda}^{\mathrm{h}},\ \sum_{p\in[m_{\mathrm{EQ}}]} f_p^{\mathrm{a}} \leq \bar{\lambda}^{\mathrm{a}}, \\
\forall p \in [m_{\mathrm{EQ}}]\ & f_p^{\mathrm{h}} \geq 0, f_p^{\mathrm{a}} \geq 0, \\
\forall i \in [m_{\mathrm{EQ}}-1]\ & \ell_p(f_p^{\mathrm{h}}, f_p^{\mathrm{a}}, 1) = \hat{\ell}_0, \\
& \ell_{m_{\mathrm{EQ}}}(f_{m_{\mathrm{EQ}}}^{\mathrm{h}}, f_{m_{\mathrm{EQ}}}^{\mathrm{a}}, s_{m_{\mathrm{EQ}}}) = \hat{\ell}_0.
\end{aligned} \tag{4.6}$$

Using similar reasoning as in Theorem 4.3, this can be formulated as a linear program and therefore can be solved with a computational complexity $O(N^3)$ [111].

Having computed the optimal routing on roads $[m_{\mathrm{EQ}}]$, we now consider an optimization which computes the resulting optimal index of the longest used road in order to fit the autonomous flow, and as a result, the optimal routing of the remaining autonomous flow. This follows from Theorem 4.2:

$$\begin{aligned}
\underset{j\in\mathcal{P}\setminus[m_{\mathrm{EQ}}-1]}{\operatorname{argmin}} \quad & j \\
\text{s.t.} \quad & \sum_{p\in[m_{\mathrm{EQ}}]} f_p^{\mathrm{a}} + \sum_{p\in[j]\setminus[m_{\mathrm{EQ}}]} \bar{F}_p(1) \geq \bar{\lambda}^{\mathrm{a}}\ ,
\end{aligned} \tag{4.7}$$

which requires computations of order $O(N)$.

We temporarily restrict our attention to the case in which autonomous users have a

uniform autonomy level. We wish to optimize over the following decision variables:

$m_{\mathrm{EQ}} \in \mathcal{P}$	longest equilibrium road?
$m_{\mathrm{ALL}} \in [m] \backslash [m_{\mathrm{EQ}} - 1]$	longest used road?
$\hat{\ell}_0 \in [t_{m_{\mathrm{EQ}}}, t_{m_{\mathrm{EQ}}+1})$	equilibrium latency?
$\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}} \in \mathbb{R}^n_{\geq 0}, \boldsymbol{s} \in \{0,1\}^n$	actual routing?

The objective function to be minimized is aggregate latency, which, using the Theorem 4.2, can be formulated as follows:

$$\hat{\ell}_0 \sum_{p \in [m_{\mathrm{EQ}}]} (f_p^{\mathrm{h}} + f_p^{\mathrm{a}}) + \sum_{p \in [m_{\mathrm{ALL}}-1] \backslash [m_{\mathrm{EQ}}]} a_p \bar{F}_p(1) + a_{m_{\mathrm{ALL}}} \Big(\bar{\lambda}^{\mathrm{a}} - \sum_{p \in [m_{\mathrm{ALL}}-1]} f_p^{\mathrm{a}} \Big)$$

$$\text{s.t. } (\boldsymbol{f}^{\mathrm{h,EQ}}, \boldsymbol{f}^{\mathrm{a,EQ}}, \boldsymbol{s}^{\mathrm{EQ}}) \in (4.6) \tag{4.8}$$

$$m_{\mathrm{ALL}} = (4.7) \tag{4.9}$$

$$\frac{1}{\bar{\lambda}^{\mathrm{a}}} \Big(\sum_{p \in [m_{\mathrm{EQ}}]} f_p^{\mathrm{a}} + \sum_{p \in [j] \backslash [m_{\mathrm{EQ}}]} \bar{F}_p(1) \Big) \geq \varphi(\frac{a_i}{\hat{\ell}_0}) \; \forall j \in \mathcal{P} \backslash [m_{\mathrm{EQ}} - 1] \tag{4.10}$$

For a given $\hat{\ell}_0$, the optimal routing maximizes the autonomous flow on roads $[m_{\mathrm{EQ}}]$, yielding (4.8). The remaining autonomous flow is routed as in (4.9). Finally, (4.10) ensures that no one is more flexible than they wish.

To solve this, recall that in the case of uniform flexibility,

$$\varphi(\kappa) = \begin{cases} 0 & 0 \leq \kappa \leq \kappa_0 \\ 1 & \kappa > \kappa_0 \; . \end{cases}$$

Further, the volume of autonomous flow that can fit on roads $[m]$ increases with decreasing ℓ_m. Because of this, we can restrict our search of $\hat{\ell}_0$ to critical points of the

function φ:

$$\hat{\ell}_0 \in \{t_{m_{\mathrm{EQ}}}\} \cup \{\frac{t_p}{\kappa_0} : p \in [m_{\mathrm{ALL}}] \backslash [m_{\mathrm{EQ}}], t_{m_{\mathrm{EQ}}} < \frac{t_p}{\kappa_0} < t_{m_{\mathrm{EQ}}+1}\},$$

which is a set with maximum cardinality N. Therefore, we can find the BFNE via the following algorithm:

1. Enumerate through all possible values of m_{EQ} (N possibilities).
2. For each possible m_{EQ}, enumerate through all possible values of $\hat{\ell}_0$ (N possibilities).
3. For each combination of m_{EQ} and $\hat{\ell}_0$, find the optimal routing on roads $[m_{\mathrm{EQ}}]$ via (4.6) (order N^3), and find m_{ALL} and the optimal routing of autonomous vehicles on the remaining roads via (4.7) (order N). As these are sequential, this step requires computations of order $O(N^3)$.

All together, this requires computations of order $O(N^5)$.

Now consider that autonomous vehicles have nonuniform flexibility levels. Then,

$$\varphi(\kappa) = \begin{cases} 0 & 0 \leq \kappa \leq \kappa_0 \\ \varphi_0 & \kappa_0 < \kappa \leq \kappa_1 \\ \varphi_1 & \kappa_1 < \kappa \leq \kappa_2 \\ \dots & \\ 1 & \kappa > \kappa_{|K|-1} \ . \end{cases}$$

We therefore must search over $\hat{\ell}_0$ in the following set:

$$\hat{\ell}_0 \in \{t_{m_{\mathrm{EQ}}}\} \cup \{\frac{t_i}{\kappa_j} : p \in [m_{\mathrm{ALL}}] \backslash [m_{\mathrm{EQ}}], j \in K, t_{m_{\mathrm{EQ}}} < \frac{t_p}{\kappa_j} < t_{m_{\mathrm{EQ}}+1}\},$$

which has maximum cardinality $|K|N$, bringing the total computation complexity to $O(|K|N^5)$. □

4.4 Incentivizing Flexibility

We next present a pricing mechanism to attain the same benefits of flexible behavior through financial incentives in the case of elastic demands. We first explain the human choice model that formalizes how autonomous service users choose between a variety of price and latency pairs depending on their value of time, encoded by their *reward functions*. We then formulate a mathematical optimization problem that aims to find an optimal trade-off between high road usage and low average travel time. We then explain the data-efficient algorithm that we propose for learning human reward functions.

4.4.1 Autonomous Service User Choice Model

How users choose between a variety of price and latency pairs depends on their valuation of time and money. Without knowing this choice model we cannot plan vehicle flows and will not be able to ensure the resulting configuration matches our vehicle flow models for the roads. Also, since different populations may have different valuations, we need to learn this tradeoff for our population so we can estimate how many people will choose which option.

To untangle these constraints, we describe how autonomous service users make routing decisions (the model for human drivers is provided in Sec. 4.2.3). Though human drivers are motivated directly only by latency, autonomous service users experience cost in both latency and the price of the ride. We model the users as having some underlying reward function, which is parameterized by their time/money tradeoff, as well as the desirability of the option of traveling by some other means such as walking or public transit. We assume that a strictly dominated option in terms of latency and price is completely

undesirable. We formally define this set

$$D = \{p \in \mathcal{P} \mid (p_p > p_{p'} \wedge \ell_p \geq \ell_{p'}) \vee (p_p \geq p_{p'} \wedge \ell_p > \ell_{p'}) \text{ for some } p' \in \mathcal{P}\} ,$$

where p_p is the monetary cost of route p. We also define the set of undominated roads, $\bar{D} = \mathcal{P} \backslash D$. We model the reward function of user j for choosing road p as follows, where we assume choosing road 0 denotes declining the service:

$$r_j(\boldsymbol{\ell}, \mathbf{p}, p) = \begin{cases} -\omega_{j_1}\ell_p - \omega_{j_2}p_p & \text{if } p \in \bar{D}, \\ -\infty & \text{if } p \in D, \\ -\zeta_j \ell^w & \text{if } p{=}0 \text{ (user } j \text{ declines the service)} , \end{cases}$$

with the following nonnegative parameters: $\boldsymbol{\ell}$ denotes the vector of road latencies, $\boldsymbol{p}$ denotes prices, $\boldsymbol{\omega}_j = \begin{bmatrix} \omega_{j_1} & \omega_{j_2} \end{bmatrix}^\top$ characterizes the users' time/money tradeoff and ζ_j specifies their willingness to use an alternative option with latency ℓ^w, which could be walking, biking, or public transportation.

We do not assume users are simple reward maximizers. Rather, we adopt the multinomial logit model [97] for the probability with which users choose each option:

$$P(\text{user } j \text{ chooses route } p) \propto \exp\left(r_j(\boldsymbol{\ell}, \boldsymbol{p}, p)\right) \tag{4.11}$$

for all $p \in \mathcal{P} \cup \{0\}$.

In order to determine the optimal pricing, we want to know how many users will choose each route as a function of the route prices and latencies. To this end, we define $q_i(\boldsymbol{\ell}, \boldsymbol{p})$ as the expected fraction of autonomous service users that will choose route p. If

the parameter distribution for autonomous service users is $g(\boldsymbol{\omega}, \zeta)$, then

$$q_i(\boldsymbol{\ell}, \boldsymbol{p}) = \int_0^\infty \int_0^\infty \int_0^\infty g(\boldsymbol{\omega}, \zeta) P(p \mid \boldsymbol{\omega}, \zeta) d\omega_1 d\omega_2 d\zeta \ .$$

where $P(i \mid \boldsymbol{\omega}, \zeta)$ is the probability that a user with reward function parameters $(\boldsymbol{\omega}, \zeta)$ will choose road p. This expression relate prices and latency to human choices, enabling us to determine the prices that will maximize the social objective. This will be important in constraining the optimization to only consider latency/price options that correspond to the desired vehicle flows.

4.4.2 Solution Method

Problem Formulation. We now formulate the problem where we have an indirect control of the autonomous cars' routing through pricing, and the demand of autonomous service is elastic. Because of this elasticity, we cannot just minimize the average latency, which would result in extremely high prices to keep autonomous service users off the network. Hence, we consider an objective that is a combination of maximizing road usage and minimizing average travel time.[2] We parameterize this tradeoff with parameter $\theta \geq 0$ in the cost function

$$J(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}) = \frac{\sum_{p \in \mathcal{P}} (f_p^{\mathrm{h}} + f_p^{\mathrm{a}}) \ell_p(f_p^{\mathrm{h}}, f_p^{\mathrm{a}}, s_p)}{\sum_{p \in \mathcal{P}} (f_p^{\mathrm{h}} + f_p^{\mathrm{a}})} - \theta \sum_{p \in \mathcal{P}} (f_p^{\mathrm{h}} + f_p^{\mathrm{a}}) \ .$$

Our control variables in this optimization are the *latency* on each road and the *price* offered to the users for traveling on each route. However, we cannot arbitrarily choose prices and latencies – we need to respect:

[2]Some other works, e.g. [49], use the *social welfare* objective. While the two objectives have many similarities, we cannot directly adopt it here as we have heterogeneous users who have different price-latency valuations.

1. *the characteristics of the roads*, in terms of how the flow demand for a road corresponds to the latency on that road (Section 4.2.1), and
2. *how people make decisions*, making sure that the number of people who choose each option corresponds to the latency of the roads described in the options (Section 4.4.1).

Moreover, we want to be fair; so we must offer the same pricing and routing options to all autonomous service users.

Given $\mathbf{q}(\boldsymbol{\ell}, \boldsymbol{p})$, the cost function $J(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s})$, inelastic flow demand of human drivers $\bar{\lambda}^{\mathrm{h}}$, and elastic demand of autonomous users $\bar{\lambda}^{\mathrm{a}}$, we are ready to formulate the planning optimization. The most straightforward way is to optimize jointly over $\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{\ell}$ and $\boldsymbol{p}$. However, $\boldsymbol{\ell}$ is fully defined by $\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}$ and $\boldsymbol{s}$. Hence, instead of $\boldsymbol{\ell}$, we use $\boldsymbol{s}$, which will help us solve this nonconvex optimization. The problem is then formulated as:

$$\min_{\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{p} \in \mathbb{R}^n_{\geq 0}, k \in \mathcal{P}, s_{k:N} \in \{0,1\}^{N-k+1}} J(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}) \tag{4.12}$$

$$\text{s.t.} \sum_{p \in \mathcal{P}} f_p^{\mathrm{h}} = \bar{\lambda}^{\mathrm{h}} \tag{4.13}$$

$$f_p^{\mathrm{a}} = \bar{\lambda}^{\mathrm{a}} q_i(\boldsymbol{\ell}(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}), \boldsymbol{p}), \forall p \in (\mathcal{P} \setminus [k]) \cup D \tag{4.14}$$

$$\sum_{i \in [k] \setminus D} f_p^{\mathrm{a}} = \bar{\lambda}^{\mathrm{a}} \sum_{i \in [k] \setminus D} q_i(\boldsymbol{\ell}(\boldsymbol{f}^{\mathrm{h}}, \boldsymbol{f}^{\mathrm{a}}, \boldsymbol{s}), \boldsymbol{p}) \tag{4.15}$$

$$a_k \leq \ell_k(f_k^{\mathrm{h}}, f_k^{\mathrm{a}}, s_k) \leq a_{k+1} \tag{4.16}$$

$$f_p^{\mathrm{h}} = 0, \forall p \in \mathcal{P} \setminus [k] \tag{4.17}$$

$$\ell_p(f_p^{\mathrm{h}}, f_p^{\mathrm{a}}, 1) = \ell_k(f_k^{\mathrm{h}}, f_k^{\mathrm{a}}, s_k), \forall p \in [k-1] \tag{4.18}$$

$$f_p^{\mathrm{h}} + f_p^{\mathrm{a}} \leq \bar{F}\left(f_p^{\mathrm{a}} / (f_p^{\mathrm{a}} + f_p^{\mathrm{h}})\right), \forall p \in \mathcal{P} \tag{4.19}$$

$$\sum_{p \in \mathcal{P}} \left(f_p^{\mathrm{a}} p_p - f_p^{\mathrm{a}} d_p c\right) \geq \bar{P} \tag{4.20}$$

with $s_{1:k-1} = 1$ due to selfishness of human-driven vehicles. Here, k is the longest

road with human-driven vehicle flow, $\bar{P}$ is the minimum profit per unit time for the autonomous service provider and c is the constant fuel cost per unit length. We can describe the constraints as follows.

4.13. The human-driven vehicle flow demand is fulfilled.

4.14. Autonomous flow will be distributed into the roads in $(\mathcal{P} \setminus [k]) \cup D$ based on the choice model described in the preceding section.

4.15. Total autonomous flow in $[k] \setminus D$ will satisfy the user choices, but can be distributed arbitrarily as the roads have the same latency and price.

4.16. The "longest equilibrium road" has latency on the given interval of free flow latency.

4.17. Human-driven cars are selfish, i.e. no human-driven car will experience higher latency than the road k.

4.18. The congested roads have the same latency as the "longest equilibrium road".

4.19. The maximum capacities of the roads are respected.

4.20. The minimum profit per unit time is satisfied.

We can further improve the search space by relying on the heuristic that the roads that are not used by the human-driven vehicles will be in free-flow, i.e. $s_{k+1:N} = 0$. While we do not have a proof for the conditions that lead to this, we also note constructing counterexamples seems to require extremely careful design, which suggests the heuristic holds in general. Furthermore, the following theorem shows we could also set $s_k = 0$ under an additional assumption.

Theorem 4.5. *Assume $\omega_2 > 0$ for all users. Then there exists a free-flow road k in the optimal solution to the problem such that $\ell_p = \ell_k$ for $\forall p \in [k]$, and $f_p^h = 0$ for $\forall p \in \mathcal{P} \setminus [k]$ as long as the optimization is feasible.*

We defer the proof to the appendix.

Generalizations. We assumed all autonomous cars are controlled by a centralized social

planner. To extend our framework to scenarios where this is not the case and the social planner has the control over a fraction of autonomous cars, we can simply do the following modifications: The optimization will also be over $\boldsymbol{f}^{\mathrm{b}} \in \mathbb{R}^n_{\geq 0}$, which will now represent the autonomous flow that does *not* belong to the planner. Add the corresponding constraint of (4.13) for f^{b}_p. Similar to (4.17), $f^{\mathrm{b}}_p = 0$ for $p \in \mathcal{P} \setminus [k]$ due to selfishness. Also, replace f^{a}'s in (4.16), (4.18) and (4.19) with $f^{\mathrm{a}} + f^{\mathrm{b}}$ with appropriate subscripts. These simple modifications enable a more general use.

Solving the Optimization. After learning the distribution $g(\boldsymbol{\omega}, \zeta)$ (described below), we first take M samples $(\bar{\boldsymbol{\omega}}, \bar{\zeta}) \sim g(\boldsymbol{\omega}, \zeta)$. Using these samples, we approximate the expected fraction of autonomous users that will choose route p as:

$$q_i(\boldsymbol{\ell}, \boldsymbol{p}) \doteq \frac{1}{M} \sum_{\bar{\boldsymbol{\omega}}, \bar{\zeta}} P(p \mid \bar{\boldsymbol{\omega}}, \bar{\zeta})$$

where $\doteq$ denotes asymptotic equality as $M \to \infty$.

We then locally solve the nonconvex planning optimization using interior point algorithms [112] with 100 random initial points for each run to get closer to global optimum.

Data-Efficient Learning of Human Reward Functions. While the routes in a specific network can be fully modeled with the physical properties and the speed limits, user's decision models (parameterized by $(\boldsymbol{\omega}, \zeta)$) must be learned in a data-driven way. The parameters might be different among the users. While a business executive might prefer paying extra to reach their company quickly, a graduate student may decide to go to the lab a little later in order to save a few dollars. Therefore, we have to learn personalized parameters $\boldsymbol{\omega}$ and ζ.

We learn the parameters from users' previous choices, which is known as *preference-based learning*. If user j chooses from a variety of options, the user's choice gives us a noisy estimate of which road $p \in \mathcal{P} \cup \{0\}$ maximizes their reward function $r_j(\boldsymbol{\ell}, \boldsymbol{p}, p)$.

We could start from either uniform priors or priors informed by domain knowledge, then sample from the distribution $g(\boldsymbol{\omega}, \zeta)$.

However, a major drawback of doing so is how quickly we learn the user preferences. Preference-based learning suffers from the small amount of information that each query carries. For example, if we show 4 options to a user (including the option to decline the service), then the maximum information we can get from that query is only 2 bits. To tackle this issue, previous works pose the query synthesis problem as an optimization and maximize a measure of the expected change in the learned distribution after the queries [98, 99].

While those works focus on pairwise queries, in this case we expect to pose several route options to the users and therefore need more general query types. By using these general queries which offer a variety of routes with varied latency and price, we can consider various ways of using this learning framework to learn the human preferences.

- We could do a user study on a few people to learn a good prior distribution.
- We could use an exploration/exploitation strategy if we are allowed to break the fairness constraint a few times for some small portion of the users. This could be made through user-specific promotions; for each user we may either choose to use the learned model or to offer special rates that would help us profile the user better.
- We could do an initial profiling study for each new user.

To implement any of these options, we formulate the following active learning optimization. First, we discuss the general preference-based learning framework. Given the data from previous choices of user j, which we denote as $\mathcal{D}_j$, we formalize the probability of $(\boldsymbol{\omega}_j, \zeta_j)$ being true parameters for that user as follows:

$$P(\boldsymbol{\omega}_j, \zeta_j \mid \mathcal{D}_j) \propto P(\boldsymbol{\omega}_j, \zeta_j) \prod_m P(\mathcal{D}_{j_m} \mid \boldsymbol{\omega}_j, \zeta_j)$$

where $\mathcal{D}_{j_m}$ denotes the road user j chose in their m^{th} choice (with $\mathcal{D}_{j_m} = 0$ meaning that the user declined the service and preferred the alternative option). The relation is due to the assumption that the users' choices are conditionally independent from each other given the reward function parameters.

The second term comes from the human choice model. For the prior, we can use a uniform distribution over nonnegative parameters. The prior may be crucial especially when we do not have enough data for a new user. In such settings, we incorporate domain knowledge to start with a better prior.

We then use this unnormalized $P(\boldsymbol{\omega}_j, \zeta_j \mid \mathcal{D}_j)$ to obtain the samples of $(\boldsymbol{\omega}_j, \zeta_j)$ using Metropolis-Hastings algorithm. Doing this for each user, which can be easily parallelized, we directly obtain the samples $(\bar{\boldsymbol{\omega}}, \bar{\zeta}) \sim g(\boldsymbol{\omega}, \zeta)$.

Next we formulate the active learning framework, which is needed so that it will not take an excessive number of queries to learn human preferences. For this, we want to maximize the expectation of the difference between the prior and the unnormalized posterior:

$$\begin{aligned}\text{query}_m^* &= \underset{\text{query}_m}{\text{argmax}}\, \mathbb{E}_{\mathcal{D}_{j_m}}\Big[P(\boldsymbol{\omega}_j, \zeta_j|\mathcal{D}_{j\,1:m-1}) - P(\boldsymbol{\omega}_j, \zeta_j|\mathcal{D}_{j\,1:m-1})P(\mathcal{D}_{j_m}|\boldsymbol{\omega}_j, \zeta_j, \mathcal{D}_{j\,1:m-1})\Big] \\ &= \underset{\text{query}_m}{\text{argmin}}\, \mathbb{E}_{\mathcal{D}_{j_m}}\left[P(\mathcal{D}_{j_m}|\boldsymbol{\omega}_j, \zeta_j, \mathcal{D}_{j\,1:m-1})\right] \ .\end{aligned}$$

As we will use the sampled $(\boldsymbol{\omega}_j, \zeta_j)$ to compute the probabilities of road choices, we can write the optimization as:

$$\text{query}_m^* \doteq \underset{\text{query}_m}{\text{argmin}}\, \mathbb{E}_{\mathcal{D}_{j_m}}\left[\sum_{\bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j} P(\mathcal{D}_{j_m}|\bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j, \mathcal{D}_{j\,1:m-1})\right] \ ,$$

where we have M samples denoted as $(\bar{\boldsymbol{\omega}}_j, \bar{\zeta}_j)$, and the term $1/M$ is canceled. Using the

law of total probability,

$$P(\mathcal{D}_{j\,m}|\mathcal{D}_{j\,1:m-1}) \doteq \frac{1}{M}\sum_{\bar{\boldsymbol{\omega}}_j,\bar{\zeta}_j} P(\mathcal{D}_{j\,m}|\bar{\boldsymbol{\omega}}_j,\bar{\zeta}_j,\mathcal{D}_{j\,1:m-1})\ ,$$

which leads to the following optimization for finding query_m^*:

$$\begin{aligned}\underset{\text{query}_m}{\text{argmin}}\sum_{\mathcal{D}_{j\,m}} P(\mathcal{D}_{j\,m}|\mathcal{D}_{j\,1:m-1})\sum_{\bar{\boldsymbol{\omega}}_j,\bar{\zeta}_j} P(\mathcal{D}_{j\,m}|\bar{\boldsymbol{\omega}}_j,\bar{\zeta}_j,\mathcal{D}_{j\,1:m-1})\\ \doteq \underset{\text{query}_m}{\text{argmin}}\sum_{\mathcal{D}_{j\,m}}\left(\sum_{\bar{\boldsymbol{\omega}}_j,\bar{\zeta}_j} P(\mathcal{D}_{j\,m}|\bar{\boldsymbol{\omega}}_j,\bar{\zeta}_j)\right)^2\ .\end{aligned}$$

We can easily compute this objective value for any given query_m. This optimization is nonconvex due to the human choice model. As in previous works, we assume local optima is good enough [98, 99]. We then use a Quasi-Newton method (L-BFGS [113]) to find the local optima, and we repeat this for 1000 times starting from random initial points.

4.5 User Experiments

To validate our framework, we conducted different simulations and a user study approved by Stanford University's Research Compliance Office.

Hypotheses. We test three hypotheses that together suggest our framework successfully reduces traffic congestion through pricing, after it learns the humans' choice models:

H1: Our active learning algorithm can learn the autonomous service user preferences in a data-efficient way.

H2: Our planning optimization reduces the overall latency by creating flexible behavior through pricing.

H3: When used by humans, the overall framework works well and is advantageous

over inflexible algorithms.

Implementation Details. In the planning optimization, we used the heuristic $s_{k+1:N} = 0$. We assumed the only alternative for autonomous service users is walking. We set $c = 6 \times 10^{-5}$ USD/meter, $\mathbf{b} = 1$. We assumed the regular and autonomous cars keep a 2-second and 1-second headway distance with the leading car, respectively. The length of the cars is 5 meters, and the minimum gap between two cars is 2 meters.

Experiments and Analyses. To validate **H1**, we simulated 5 autonomous service users with different preferences. We tested our active learning framework by asking two sets of 200 queries, each of which consisted of 4 road options, similar to Fig. 4.6, and a walking option. The queries were generated actively in the first set and randomly in the second. After each query, we recorded the sample $(\bar{\boldsymbol{\omega}}, \bar{\zeta})$ which has the highest likelihood as our estimates.

Fig. 4.4 shows how the estimates evolved within active learning setting for one of the users. All values are overestimated initially. Intuitively, this is because getting noiseless responses has higher likelihood in the beginning. As we query more, accepting some of the responses as noisy maximizes the likelihood. Therefore, the values start decreasing.

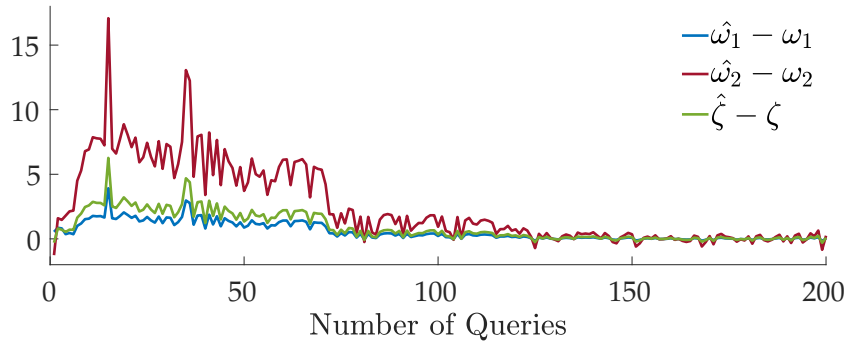


Figure 4.4: The errors of the reward function estimates are shown with varying number of queries. $\hat{\omega}_1$, $\hat{\omega}_2$ and $\hat{\zeta}$ represent the estimates.

Another important observation is that the estimates of the parameters increase and decrease together even in the early iterations. This suggests we are able to learn the

ratio between the parameters, e.g. ω_1/ω_2 very quickly. To check this claim, we used the following error metric:

$$e_{x,y} = \|x/y - \hat{x}/\hat{y}\|_1 \ ,$$

where $x, y \in \{\omega_1, \omega_2, \zeta\}$ and $\hat{x}, \hat{y}$ represent the corresponding estimates. Fig. 4.5 shows how this error decreases with increasing number of queries. It also shows how active querying enables data-efficient learning compared to the random querying baseline. We are able to learn the relationship between parameters even under 20 queries with active learning. All these results strongly support **H1**.

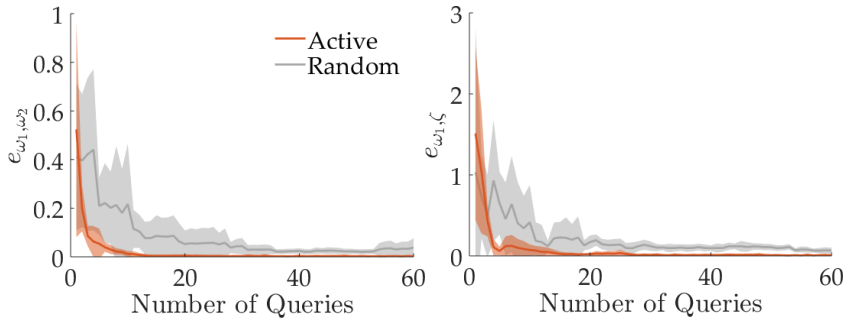


Figure 4.5: The error metric is averaged over 5 different reward functions.

The fact that we are able to learn the ratios implies we can estimate which road the user is most likely to choose. We will only be unsure about how noisy the user is if the parameter estimates did not converge yet. Therefore, we can still use the estimates for our planning optimization even when we have small number of queries.

To validate **H2**, we use the road network from [1] and the equilibria benchmarks we developed: NE, BNE and BFNE with full flexibility. Here, we give the average latencies for the 4-road network from that study which we visualize in Fig. 4.6, and where $\bar{\lambda}^{\mathrm{h}} = 0.4, \bar{\lambda}^{\mathrm{a}} = 1.2$ cars per second: (NE: 400.00 sec, BNE: 125.66 sec, BFNE (full flexibility): 102.85 sec).

We then assumed we perfectly learned the preferences of the 5 simulated users. We

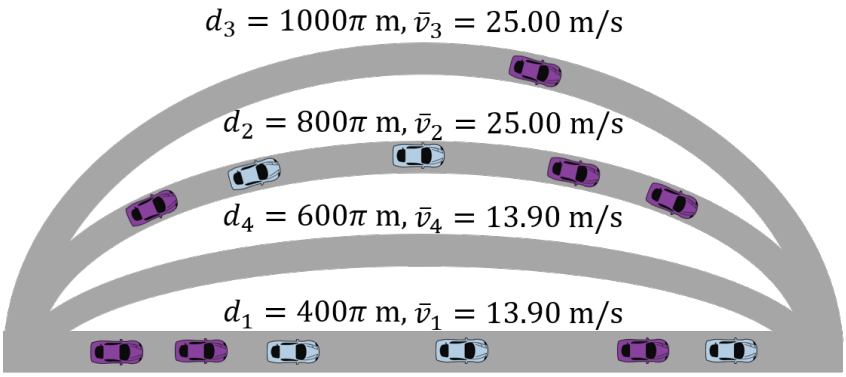


Figure 4.6: The 4-road network from [1]. The roads are not to the scale and ordered with respect to the free-flow latencies.

ran the planning optimization with $\bar{P} = 0$ and 3 different θ to show the trade-off. The results are summarized in Table 4.1.

Table 4.1: Results of Routing Simulation

θ	Avg. Latency (seconds)	Flow (cars/second)
1	90.41	0.4412
20	97.03	1.2746
10^6	111.28	1.5964

It can be seen we can adjust the trade-off between average latency and the served flow by tuning θ. Also, given the human preferences, even when we served (almost) all of autonomous demand, our framework outperforms BNE. This shows its effectiveness on creating flexibility and supports **H2**.

For **H3**, we recruited 21 subjects (9 female, 12 male) with an age range from 19 to 60. In the first phase of the experiment, each participant was asked 40 actively synthesized queries (4 roads + 1 walking option). We then used their responses to get the maximum likelihood sample $(\bar{\boldsymbol{\omega}}, \bar{\zeta})$. Afterwards, we designed 5 different road networks each with 4 different roads and an additional route where people may choose to walk. The 5 different networks cover a range of different road lengths from 1.8 kilometers to 78 kilometers. For each network, we also set different θ, $\bar{\lambda}^{\mathrm{h}}$, $\bar{\lambda}^{\mathrm{a}}$, and $\bar{P}$. By assuming all autonomous flow

is in the service of these 21 subjects, or of the groups that match with their preferences, we locally solved the planning optimization to get the pricing scheme for each traffic network. We refer to these results as *anticipated* values.

In the second phase of the user study, we presented the route-price pairs and the walking option to the same 21 subjects. For each of the 5 networks, they picked the option they would prefer. Using these responses, we allocated the autonomous flow into the roads. However, it is technically possible that more users select a road than its maximum flow. To handle such cases, we assumed extra flow is transferred to the roads with smaller latencies without making the users pay more. If that is not feasible, extra flow is transferred to the slower roads, without any discount. While these break the fairness constraint, it rarely happens and affects only a very small portion of the flow. After autonomous flows are allocated, human driven cars selfishly chose their routes in a way to minimize the overall average latency. We refer to the results of this allocation as *actual* values.

Table 4.2 compares the anticipated and the actual values. We report latencies in seconds, flows in cars per second, and profit is a rate value with the unit USD per second. In order to show how our framework incentivizes flexible behavior, we also added other benchmarks: two where the same flow as actual flow is routed (BNE1 and BFNE1) and two where all flow demand is routed, i.e. no walking option (BNE2 and BFNE2). While BNE and BFNE assume completely inelastic demand, using them as benchmarks under both the actual flow and the all flow demand gives us insights about the success of our framework, because the allocation of vehicles under the actual flow is comparable to the elastic demand case.

It can be seen that there is generally an alignment between the anticipated and actual values. While the mismatch may be further reduced by doing more queries or having more users, the difference with BNE1 is significant. In all cases, our framework achieved

to incentivize flexibility, which yielded lower average latencies compared to BNE1. Especially in Case 3 and Case 5, our framework approximately halved the average latency compared to the best Nash equilibria. In fact, our framework achieved a latency that is close to the lower bound set by BFNE1. We visualize these in Fig. 4.7. Furthermore, our framework successfully reduced flow demand when satisfying the full demand under selfishness is impossible (Case 1).

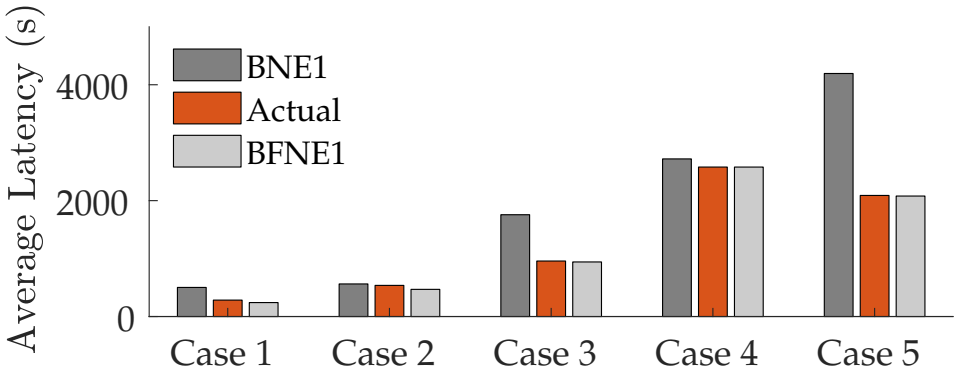


Figure 4.7: The comparison of the actual results and BNE1, both of which allocate the same amount of flow.

One caveat is the small amount of actual flow in Case 1, which also caused an important profit loss. This is because the roads are relatively shorter, and most users preferred walking over paying for an autonomous car. Our framework could not predict this, because the learned reward functions failed to accurately estimate the probabilities.

Table 4.2: Results of Real-User Experiments

	Anticipated			**Actual**	**Pricing Scheme**		**BNE1**	**BFNE1**	**Full**	**BNE2**	**BFNE2**
	Flow	Av. Lat.	Profit	Flow	Av. Lat.	Profit	Av. Lat.	Av. Lat.	Flow	Av. Lat.	Av. Lat.
C1	1.662	338.10	11.11	1.043	283.30	3.85	501.97	240.80	1.90	Infeas.	421.45
C2	1.559	510.97	4.50	1.505	537.30	3.95	562.20	468.79	1.60	562.20	474.35
C3	1.268	921.06	4.59	1.300	957.71	3.99	1756.89	941.81	1.30	1756.89	941.81
C4	1.554	2576.36	40.01	1.600	2579.52	49.33	2720.00	2579.33	1.60	2720.00	2579.33
C5	1.238	2083.42	48.48	1.205	2089.53	46.32	4194.26	2079.15	1.30	4194.26	2183.56

Table 4.3: Comparing the anticipated and actual performance of five cases, C1-C5, in terms of flow, average latency, and profit. For a fixed flow, BFNE serves as a lower bound on the latency. Achieving lower latency than BNE means we successfully incentivize flexible behavior.

4.6 Conclusion

In this chapter[3] we addressed the efficiency of traffic networks with mixed autonomy when autonomous vehicles are used for a ride-hailing service, and vehicle flow on roads is dictated by a model based on the Fundamental Diagram of Traffic. We developed a method of pricing rides with autonomous vehicles such that when a population chooses from these route and price options, and the human drivers choose the quickest routes available to them, the objective of decreasing travel latency and increasing road usage is achieved. To do so, we modeled how people choose between different route options with varying prices and latencies. Moreover, we developed a method for actively learning the parameters that describe the preferences of a population of users. We developed theoretical results which we use to gauge the performance of our algorithm, and conducted a user study showing that our method of parameterizing and actively learning the preferences of a human population is effective.

A wide horizon for further research remains. One could relax the assumption that the reward functions are linear to improve the prediction accuracy [114] and optimize for information gain in the active learning scheme, which can yield better data efficiency [115]. Another direction is to consider an elastic demand for human drivers, as well as a variety of nondriving options including walking, biking, or taking the bus. In that case, ζ will have a multimodal distribution; we then need to learn the mixture. Future works could consider optimizing for objectives aside from a linear combination of minimizing delay and maximizing the number of people traveling [116].

More broadly, one can look at more general network topologies – if route pricing is computed only with local information on each edge, pricing can make congestion worse

[3]Toyota Research Institute ("TRI") provided funds to assist the authors with their research but this article solely reflects the opinions and conclusions of its authors and not TRI or any other Toyota entity. This work was also supported by NSF grant CCF-1755808, NSF ECCS-1952920, FLI grant RFP2-000, and UC Office of the President grant LFR-18-548175.

when users have different price sensitivities [117]. One can also expand the model to include the role of information in decision making, as well as biases such as risk aversion. Through these future directions we can ensure the efficient operation of transportation networks that include autonomous vehicle-based ride hailing services.

4.7 Appendix

4.7.1 Proof of Lemma 4.1.

We begin by noting that for any given network and feasible flow demand, a *continuum* of equilibria exist which satisfy the flow demand, where different equilibria have positive flow on different sets of roads. This fact stems from the two regimes (free-flow and congested) that exist, as well as the two different vehicle types. We note that for a congested road, a lower density yields a higher vehicle flow and lower latency. Accordingly, via the relationship with density, the latency function decreases smoothly with an increase in the flow of regular or autonomous cars. Note this does not mean that transferring flow to a road in a dynamic setting decreases its latency, rather that the function dictating the relationship between latency and flow in the congested regime is smoothly decreasing (in accordance with each quantity's relationship with the corresponding vehicle density).

With this in mind, we consider the continuum of equilibria that exist. Consider a Nash Equilibrium with positive flow on roads $[m]$. Assume road m is congested, as otherwise the lemma would be satisfied with $m' = m$. Roads $[m-1]$ must be congested so as to have the same latency as road m, as it is an equilibrium. We constructively find a different equilibrium as follows. Consider another configuration serving the same flow demand, where the flow on road m serves less flow and roads $[m-1]$ serve more flow as follows. Note that in each configuration serving the flow demand, vehicle flow

is conserved but vehicle density is not. Accordingly, road m has higher latency (from higher density) and roads $[m-1]$ have lower latency (from lower density). Since the flow-latency relationship in (4.3) has latency as a monotonically decreasing function with flow of each vehicle type, we can consider a new configuration where each road in $[m-1]$ has equal latency (which is less than the first equilibrium configuration) while serving more flow. Accordingly, we can construct one of the two following configurations.

1. the latencies on roads $[m-1]$ are reduced to a_m, the free-flow latency on road m, or
2. road m serves no flow and the latencies on roads $[m-1]$ are greater than a_m.

In the first case, the lemma is satisfied with $m' = m$. In the second case, we can consider the same logic again, again considering a new configuration with less flow on road $m-1$ and more flow on roads $[m-2]$. This continues until either we achieve case 1 above or until we are reduced to a single road. If that occurs, traffic can be routed in free-flow on that road, since any feasible flow in the congested regime is less than the maximum free-flow on a road. □

4.7.2 Proof of Theorem 4.1.

The definition of Nash Equilibrium and the fact that latency on a road is always equal to or greater than its free-flow latency together imply that at Nash Equilibrium, if road m has positive flow then all roads with free flow latency less than a_m have positive flow as well. These also imply that if a road m in free-flow has positive flow, roads with greater free-flow latency will have zero flow. Further, we use Lemma 4.1 to show that all routings in the set of BNE will have one road in free-flow. Assume for the purposes of contradiction that we have a routing in the set of BNE in which only roads $[m]$ have positive flow and all are congested, with equilibrium latency $\hat{\ell}_0$. The total cost is then

$\hat{\ell}_0(\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}})$, where $\hat{\ell}_0 > a_m$. By Lemma 4.1, another routing exists in the set of NE which uses roads $[m']$, where $m' \leq m$ and road m' is in free flow. The cost of this equilibrium is $a_{m'}(\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}}) \leq t_m(\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}}) < \hat{\ell}_0(\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}})$ contradicting our premise.

So far we have proved the numbered claims. We prove the remaining claim by contradiction. Assume there are two routings in f, $f' \in \mathrm{BNE}(\bar{\lambda}^{\mathrm{h}}, \bar{\lambda}^{\mathrm{a}})$ which have different free-flow roads, m and m' respectively. Assumption 4.1 implies $t_m \neq t_{m'}$; let $t_m < t_{m'}$. Since all selfish users experience the same latency, the total latency of routing f is $(\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}})t_m < (\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}})t_{m'}$, which is the total latency of routing f'. However by the definition of BNE, the total latency of the two routings are equal, yielding a contradiction. □

4.7.3 Proof of Theorem 4.2.

The first property directly follows from Theorem 4.1, as the regular vehicles have to be at a Nash equilibrium due to selfishness. To prove the second property, we note that for roads that have higher latencies than road m^*_{EQ}, a Nash equilibrium is not necessary due to flexibility. As $\ell_i(f_i^{\mathrm{h}}, f_i^{\mathrm{a}}, s_i)$ is a non-increasing continuous function of f_i^{a} and decreasing for $s_i = 1$, roads that have higher latencies than road m^*_{EQ} will always be in free-flow.

Now we assume some of the roads with indices greater than m^*_{EQ} and less than m^*_{ALL} are in free-flow, but not at maximum flow. Then we could simply transfer some flow from the road m^*_{ALL} to those roads and have lower overall costs. This is a contradiction, completing the proof for the third property. □

4.7.4 Proof of Theorem 4.3.

Since road m_{EQ}, the longest equilibrium road, is in free-flow by Theorem 4.1 and m^*_{EQ} is the minimum feasible m_{EQ}, our solution is restricted to the set of BNE. Accordingly,

we can restrict our optimization to routing in which the longest equilibrium road is in free-flow. This allows us to write an optimization equivalent to checking the feasibility of a routing with the desired congestion profile:

$$\begin{aligned}
&\max_{\boldsymbol{f}^{\mathrm{h}},\boldsymbol{f}^{\mathrm{a}}\in\mathbb{R}^N_{\geq 0}} 1 \\
&\text{s.t.} \sum_{p\in[m^*_{\mathrm{EQ}}]} f^{\mathrm{h}}_p = \bar{\lambda}^{\mathrm{h}}, \quad \sum_{i\in[m^*_{\mathrm{EQ}}]} f^{\mathrm{a}}_p = \bar{\lambda}^{\mathrm{a}} \\
&\qquad \ell_p(f^{\mathrm{h}}_p, f^{\mathrm{a}}_p, 1) = t_{m^*_{\mathrm{EQ}}} \;\forall p \in [m^*_{\mathrm{EQ}} - 1] \\
&\qquad f^{\mathrm{h}}_{m^*_{\mathrm{EQ}}} + f^{\mathrm{a}}_{m^*_{\mathrm{EQ}}} \leq \bar{F}_{m^*_{\mathrm{EQ}}}(f^{\mathrm{h}}_{m^*_{\mathrm{EQ}}}, f^{\mathrm{a}}_{m^*_{\mathrm{EQ}}})
\end{aligned}$$

The constraints can be shown to be affine in the decision variables. As it is a linear program, it can be solved in $O(N^3)$ time [111]. Finding m^*_{EQ} requires a search in $O(N)$ time. □

4.7.5 Proof of Theorem 4.5.

We first prove, similar to Luce's choice axiom [118], changing the latency or the price of some roads does not alter the autonomous flow ratio between the other options, including the alternative option. For this, we look at $\mathbb{E}_{\boldsymbol{\omega},\zeta}\left[\frac{P(p_1|\boldsymbol{\omega},\zeta,\boldsymbol{\ell},\boldsymbol{p})}{P(p_2|\boldsymbol{\omega},\zeta,\boldsymbol{\ell},\boldsymbol{p})}\right]$, where p_2 is an undominated option, $P(p \mid \boldsymbol{\omega},\zeta,\boldsymbol{\ell},\boldsymbol{p})$ is the probability of choosing option $p \in [N]\cup\{0\}$ under the given reward parameters, latency and price. We note

$$\begin{aligned}
&P(p \mid \boldsymbol{\omega},\zeta,\boldsymbol{\ell},\boldsymbol{p}) = \frac{\exp(r(\boldsymbol{\ell},\boldsymbol{p},p;\boldsymbol{\omega},\zeta))}{\sum_{p'=0}^{N}\exp(r(\boldsymbol{\ell},\boldsymbol{p},p';\boldsymbol{\omega},\zeta))}\,, \quad \text{so} \\
&\mathbb{E}_{\boldsymbol{\omega},\zeta}\left[\frac{P(p_1 \mid \boldsymbol{\omega},\zeta,\boldsymbol{\ell},\boldsymbol{p})}{P(p_2 \mid \boldsymbol{\omega},\zeta,\boldsymbol{\ell},\boldsymbol{p})}\right] = \mathbb{E}_{\boldsymbol{\omega},\zeta}\left[\frac{\exp(r(\boldsymbol{\ell},\boldsymbol{p},p_1;\boldsymbol{\omega},\zeta))}{\exp(r(\boldsymbol{\ell},\boldsymbol{p},p_2;\boldsymbol{\omega},\zeta))}\right]\,.
\end{aligned}$$

As the reward of an option depends only on that option's price and latency, this proves the first statement above.

Equipped with this result, we now prove Theorem 4.5. Assume the optimal solution is such that all human-driven flow is in congested roads. Let k be the index of the highest free-flow latency road with nonzero human-driven flow. We first show there exists an equally optimal solution with no dominated roads in $[k]$. For this, we simply set the prices of roads $[k]$ such that they are all equal and the total autonomous demand in $[k]$ in the original and the new solution is the same. Since the new solution has no dominated roads in $[k]$, their new price has to be at least as high as the undominated roads in $[k]$ of the original solution, which implies the profit constraint is still satisfied. As the total demand served and the overall latency values are the same between these two solutions by the first statement above, the two solutions are equally good. In the remaining of the proof, we refer to the new solution as the "optimal solution" for clarity, and show there exists a better solution, leading to a contradiction.

Denote the optimal solution with $(\boldsymbol{\ell}^*, \boldsymbol{p}^*, \boldsymbol{f}^{\mathrm{h}^*}, \boldsymbol{f}^{\mathrm{a}*})$, and its dominated options with D^*, noting $[k] \cap D^* = \emptyset$. Let the ratio of autonomous service users in road i to the autonomous service users who decline the service be β_i, i.e., $\beta_i = \frac{f^{\mathrm{a}*}_p}{\lambda^{\mathrm{a}} - \sum_{p' \in [N]} f^{\mathrm{a}*}_{p'}}$.

Let $k' = \mathrm{argmax}_p\, t_p$ subject to $p \leq k$ and $\left(\sum_{p' \in [k]} f^{\mathrm{h}*}_{p'}, \sum_{p' \in [k]} f^{\mathrm{a}*}_{p'}\right)$ can be allocated into $[k']$ when all roads in $[k']$ have latency equal to $t_{k'}$. Existence of such a k' is guaranteed by Lemma 4.1.

Using k', we propose an alternative solution $(\boldsymbol{\ell}',\boldsymbol{p}',\boldsymbol{f}^{\mathrm{h}'},\boldsymbol{f}^{\mathrm{a}\prime})$:

$$\ell'_p = \begin{cases} \ell^*_p & \text{if } p \in [N] \setminus [k] \\ a_i & \text{if } p \in [k] \setminus [k'] \\ t_{k'} & \text{if } p \in [k'] \end{cases},$$

$$p'_p = \begin{cases} p^*_p & \text{if } p \in [N] \setminus ([k] \cup D^*) \\ p^*_p + \epsilon & \text{if } p \in [k] \cup D^* \end{cases}.$$

where $\epsilon \geq 0$ is such that the ratio of autonomous service users in $[k']$ to the autonomous service users who decline the service in the alternative solution is equal to $\sum_{p\in[k]} \beta_p$. Since $\ell'_p < \ell^*_p$ for $\forall p \in [k']$, $\omega_2 > 0$ for all users, and $[k] \cap D^* = \emptyset$, the existence of such an ϵ is guaranteed.

By the first statement, we know the ratios (β_p's) in the alternative solution will be equal to the optimal solution for roads in $[N] \setminus ([k] \cup D^*)$ as their latencies and prices are the same. Similarly, roads in D^* are also dominated in the alternative solution, because no road has higher latency and prices went up maximally (by ϵ) for those roads. This means the ratios are equal between two solutions for all roads in $[N] \setminus [k]$. Noting roads in $[k] \setminus [k']$ are dominated in the alternative solution, as they have the same price as the roads in $[k']$ but higher latency, we write the total autonomous demand:

$$\bar{\lambda}^{\mathrm{a}} = \left(\bar{\lambda}^{\mathrm{a}} - \sum_{p\in[N]} f^{\mathrm{a}\prime}_p \right) (1 + \beta_1 + \beta_2 + \cdots + \beta_N)$$

by the construction of ϵ. As the total autonomous demand is the same between two solutions, we get $\bar{\lambda}^{\mathrm{a}} - \sum_{p\in[N]} f^{\mathrm{a}\prime}_p = \bar{\lambda}^{\mathrm{a}} - \sum_{p\in[N]} f^{\mathrm{a}*}_p$ meaning the flow of users who decline the service is the same. This implies $f^{\mathrm{a}\prime}_p = f^{\mathrm{a}*}_p$ for $\forall p \in [N]/[k]$, and $\sum_{p\in[k]} f^{\mathrm{a}\prime}_p =$

$\sum_{p\in[k]} f^{\mathrm{a}*}_p$. As the selection of k' ensures there is enough room for selfish vehicles in $[k']$ in the alternative solution, we have $\sum_{p\in[k']} f^{\mathrm{h}\prime}_p = \sum_{p\in[k]} f^{\mathrm{h}*}_p$.

Finally, the alternative solution satisfies the profit constraint because the price and the autonomous flow in $[N]\setminus[k]$ are the same as the optimal solution, and the remaining autonomous flow pays higher price in the alternative solution.

Overall, the alternative solution is feasible, serves the same amount of flow, but has lower latency overall, $(\boldsymbol{\ell}^*, \boldsymbol{p}^*, \boldsymbol{f}^{\mathrm{h}^*}, \boldsymbol{f}^{\mathrm{a}*})$ cannot be the optimal solution. □

4.7.6 Role of Assumption 4.1.

Without this assumption, m^*_{EQ} in Theorem 4.1 would instead represent a set of roads, all of which would be in free-flow. Theorem 4.2 would change similarly to Theorem 4.1. Theorem 4.3 is not altered, the constraint associated with m^*_{EQ} would instead apply to all roads within the set. Finally, Theorem 4.4 would remain the same as well. What would potentially change this theorem is if the set of roads in m^*_{EQ} could have different congestion levels in the computed BFNE. To see why this is not the case, note that the flexibility profile is with respect to the quickest route available to users (*i.e.* the least congested road in the set m^*_{EQ}). Accordingly, having another road in the set be more congested would not help satisfy any constraint related to the flexibility profile and would also not serve more flow than if it had the same latency as the minimum latency road in m^*_{EQ}. Hence, all roads in m^*_{EQ} will have the same latency, so the computational complexity remains the same. Theorem 4.5 would still hold: there exists a free-flow road used by human drivers in the optimal solution.

Chapter 5

Learning to Dynamically Influence Human Routing

The previous chapters of this dissertation analyzed how to influence the equilibria of traffic networks with mixed autonomy. In Chapter 3, the tolling schemes proposed either guaranteed uniqueness of equilibria, or bounded the worst-case performance of equilibria. When analyzing the case of autonomous vehicles serving as a ride hailing service in Chapter 4, the equilibria found were no longer unique. Further, though we found price controls which would maintain a good equilibrium, the control scheme was not designed to *guide* traffic conditions from a bad equilibrium to a good one. Accordingly, in this chapter[1] we analyze the dynamics of a traffic network in mixed autonomy. Due to the complexity of the formulation, we use Reinforcement Learning to learn a policy which dynamically controls the autonomous vehicle routing to influence human routing in such a way that road congestion is alleviated. We characterize the optimal equilibria in parallel networks and show that the learned policy brings the network to the best equilibrium,

[1]This chapter is adapted, with permission, from "Learning How to Dynamically Route Autonomous Vehicles on Shared Roads", previously published in Transportation Research Part C, © Elsevier 2021, and is joint work with Erdem Bıyık, Dorsa Sadigh, and Ramtin Pedarsani [119].

while model predictive control and greedy optimization methods fail to do so, and result in unboundedly large delays. We also show the efficacy of our method on large, general networks.

5.1 Introduction

There are a few mechanisms by which road network inefficiency and congestion emerge. Previous chapters have focused on road network inefficiency due to route choice, and specifically, have analyzed and controlled the equilibria of networks. However, in realistic models of congestion, the congestion on a road is a function not just of the incoming flows of vehicles, but also of the *current state* of network – specifically, the density of vehicles in each portion of the network. Because of this, choosing an equilibrium is not as simple as choosing an input routing to the network, as the same input can result in different equilibria, depending on the network state.

If a system controller had control over the input routing of vehicles entering a network starting from an empty state, it could directly control the state of the network. There are a few issues with this. Even if this unlikely situation was to happen, there could still be system shocks, such as fluctuations in demand or traffic collisions, which suddenly change the state of the network. Because of this, if one wishes to decrease congestion by controlling or influencing routing, it is important to develop a method which is robust to these effects, and which can start from any feasible network state.

In this chapter we consider a similar setting to that of Chapter 4, in which some users of a network are users of an autonomous ride-hailing service, and some users are human drivers. We consider that a system designer can directly control the routing of the autonomous service users and that human drivers chooser their routes selfishly in response. In contrast with equilibrium analysis of Chapter 4, in the current chapter we

consider the *dynamic* setting. Our goal is to design a controller which can route the autonomous vehicles in a way which leverages the selfish routing response of the human drivers to *decongest* the network, and move the network from a 'bad' equilibrium to a 'good' one.

Our key idea is that by controlling the routing of autonomous vehicles we can change the delay associated with traversing each road, thereby indirectly influencing peoples' routing choices. By influencing people to use more "socially advantageous" routes, we can eliminate long queues and significantly reduce traffic jams on roads.

The model for *mixed-autonomy* traffic, meaning traffic with both human-driven and autonomous vehicles, is complex, involving very large and continuous state space and continuous action space. Having human drivers dynamically respond to the choices of the autonomous vehicles further complicates the matter, making a dynamic programming-based approach and other classical methods infeasible. Because of this, we use model-free deep reinforcement learning (RL) to learn a policy without requiring access to the dynamics of the transportation network. Specifically, we show it is possible to learn a policy via proximal policy optimization (PPO) [120] that mitigates traffic congestion by managing routing of autonomous cars given the network state.

To understand the performance of the learned policy, we investigate the *equilibrium* behavior of the network. In the preceding chapters we showed that there is a wide spectrum of equilibria in traffic networks, meaning situations in which everyone is taking the quickest route immediately available to them, and these equilibria can have greatly varying average user delay. In this chapter we establish efficient ways to compute equilibria in the network and compare the best equilibrium (in terms of latency) with the RL policy, which works regardless of whether equilibrium conditions hold or not. We show that the learned policy reaches the 'desirable' equilibria that have low travel times when starting with varying traffic patterns, and can recover network functionality after a disturbance

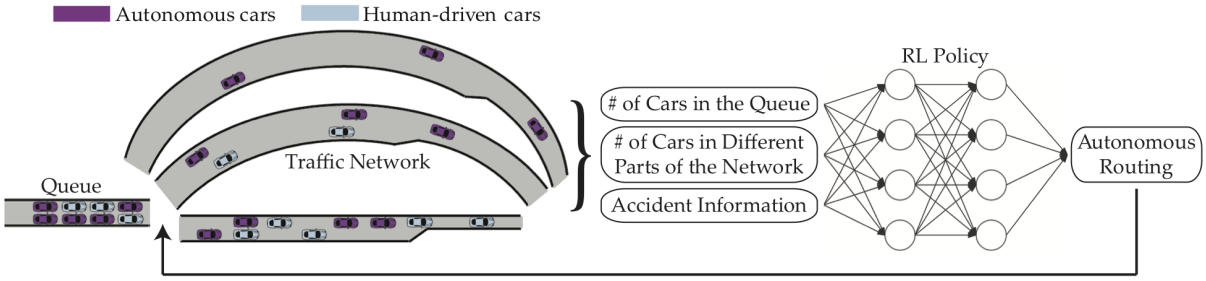


Figure 5.1: The schematic diagram of our framework. Our deep RL agent processes the state of the traffic and outputs a control policy for autonomous cars' routing.

such as a traffic accident. To summarize, our contributions are as follows:

- *Theoretical analysis:* We characterize equilibria in the network and derive a polynomial-time computation for finding optimal equilibria of parallel networks.
- *Finding a control policy via deep RL:* We employ deep RL methods to learn a routing policy for autonomous cars that effectively saves the traffic network from unboundedly large delays. We show via simulation that the RL policy is able to bring our network to the best possible equilibrium when starting from a congested state or after a network disturbance on parallel networks. We further show that an MPC-based approach and a greedy optimization method fail to do so, and thus is outperformed by the RL-based method in general networks.

We portray our framework in the schematic diagram Fig. 5.1.

5.1.1 Related Works

Many works seek to understand how much traffic network latency could be improved if vehicle routing was controlled by a central planner, including works on *congestion games* [14,19,53,72]. Some study how indirectly influencing peoples' routing choices by providing them network state information affects network performance [10,11]. *Stackelberg Routing*, in which only some of the vehicles are controlled, is another way to influence routing

[27, 121]; some works incorporate the *dynamics* of human routing choices [122]. While providing useful techniques for analysis, the congestion game framework does not reflect a fundamental empirical understanding about vehicle flow on roads, namely that roads with low vehicle density have a roughly constant latency, and roads with high density see latency *increase* as flow *decreases*.

Works on CTM [34, 123] capture this phenomenon, including works that characterize equilibria on roads described with CTM [5]. Notably, some consider equilibria of parallel-path Stackelberg Games, including with mixed autonomy [29,91]. However, their analyses are limited to steady-state and do not capture the dynamics. [124] considers a Fundamental Diagram of Traffic-based model for slowly varying traffic. They formulate this as a Stackelberg Game and design routing information for users to minimize overall latency and bound the resulting inefficiency in a simple network. However, they only consider a single-vehicle type, not a mixed-autonomy setting.

Previous works have shown the negative congestion effects of ride-hailing services [4]. There is a wide range of works on how to decrease congestion, including congestion pricing [7], variable speed limits [8] and highway ramp metering [5].

Some works look at the low-level control of autonomous cars, specifically controlling acceleration to smooth flow and ease congestion at bottlenecks [36, 62, 63]; [125] provides a benchmark for gauging the performance of these techniques. Other works learn ramp metering policies [6], localize congestion [39], and model lane-change behavior with a neural network [126]. One study surveys the way in which autonomous vehicles provide the opportunity for better congestion management [127].

In addition to these learning-based methods, there has also been an effort to use RL for route selection [128] and driver choice modeling in traffic assignment problem [129–132]. Again using RL, [133] shows reward shaping mechanisms could be utilized to reach better equilibria. More recently, [134] develops a hierarchical approach to optimize tolling and

signal control in the high-level whereas a multi-agent RL method models the drivers in the lower level. Although these works show the effectiveness and potential of RL methods in transportation, to the best of our knowledge, these methods have not been used in a routing game with mixed-autonomy traffic where a central planner aims to reduce congestion by indirectly influecing humans' routing via the routing of autonomous vehicles.

Without any reinforcement learning component, some works provide macroscopic models of roads shared between human-driven and autonomous cars. [65] models highway bottlenecks in the presence of platoons of autonomous vehicles mixed in with human-driven vehicles. The authors relate their model to a CTM type-model similar to the model presented below, though it is specific to a single highway. [135] describes a microscopic model to determine the effect of autonomy on throughput, yielding fundamental diagrams. The fundamental relationship between autonomy level and critical density in our model mirrors that of [33], which develops a CTM model for mixed autonomy traffic.

Some works solve the dynamic traffic assignment problem for networks with a CTM-based flow model, including some which decompose the optimization to enable optimizing flow on large networks [136]. In contrast, our works studies the setting in which some flow demand is controlled to optimize the system performance, and some flow demand updates according to a selfish update rule. This precludes the use of such decomposition techniques, since the optimization can no longer be formulated as a linear program. Because of this, we use RL to solve for a routing policy in our setting.

5.2 Vehicle flow dynamics: modeling roads

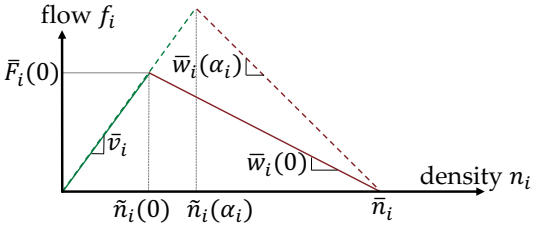
In this section we describe dynamics governing how vehicle flow travels on a road. We extend the CTM, a widely used model that discretizes roads into *cells*, each with uniform

density [34,123], for mixed-autonomy traffic. In CTM, each road segment has a maximum flow that can traverse it. The key idea of our extension is that since autonomous vehicles can keep a shorter headway (distance to the car in front of it), the greater the fraction of autonomous vehicles on a road, the greater the maximum flow that the road can serve [91]. Accordingly, our extension of CTM lies in the dependence of cell parameters on the *autonomy level*, or the fraction of autonomous vehicles, in each cell.

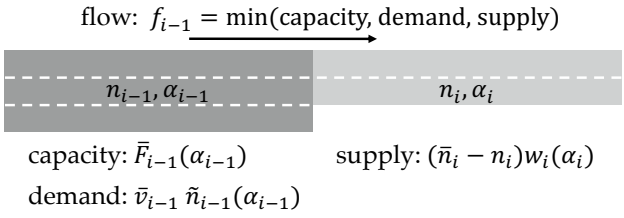
We use our capacity model in conjunction with Daganzo's CTM formulation in [34, 137], the combination of which we describe in the following. We consider a network of roads with a single origin and destination for all vehicles in the network. The origin and destination are connected by the set of simple paths $\mathcal{P}$. Each path is composed of a number of cells, and we denote the set of cells composing path p by $\mathcal{I}_p$. We generally use i and p as indices for cells and paths, respectively.

In the CTM, every cell has a *critical density*, and when the density of a cell exceeds the critical density, that cell is *congested*. We model the critical density as being dependent on the autonomy level. This is because autonomous vehicles maintain a different nominal headway than human-driven vehicles; in other words, autonomous vehicles may require more space in front of them due to prediction error, or less space, as they may react faster than human drivers. Accordingly, we use the model in [42] to model the capacity of a cell.

Using this model, each cell i has a free-flow velocity, $\bar{v}_i$, as well as a nominal headway for vehicles traveling at the free-flow velocity — h_i^{h} cells/vehicle for human-driven vehicles and h_i^{a} for autonomous vehicles. The capacity of the cell then varies with the autonomy level, denoted $\alpha_i \in [0, 1]$. We use b_i to denote the number of lanes in a cell. We model vehicles as slowing down when the headway experienced decreases below the nominal headway required and accordingly model the critical density as follows, as in [32,42,72,91]:



(a) Fundamental diagram of traffic governing vehicle flow in each cell of the Cell Transmission Model. The solid line corresponds to a cell with only human-driven vehicles; the dashed line represents a cell with both vehicle types at autonomy level α_i.



(b) The flow from one cell to another is a function of the density ϕ and autonomy level α in each cell.

Figure 5.2: Vehicle flow model. Green and red respectively represent a cell in free-flow and congestion, and we suppress the notation for path p.

$$\tilde{n}_i(\alpha_i) := b_i/(\alpha_i h_i^{\mathrm{a}} + (1-\alpha_i)h_i^{\mathrm{h}}) \; . \tag{5.1}$$

Each cell also has a vehicle density, $\phi_i = \phi_i^{\mathrm{h}} + \phi_i^{\mathrm{a}}$, where ϕ_i^{h} and ϕ_i^{a} are, respectively, the number of human-driven and autonomous vehicles. Thus, $\alpha_i = \phi_i^{\mathrm{a}}/(\phi_i^{\mathrm{h}} + \phi_i^{\mathrm{a}})$. As the cells are very large compared to the vehicles, we consider these quantities to be continuous variables. As mentioned above, CTM has two regimes for vehicle flow: free-flow, when cell density is less than the critical density, and congestion, when cell density is greater than the critical density but less than the *jam density* $\bar{\phi}_i$, the density at which flow stops completely.

Three factors limit the flow from one cell to another. One is the *capacity*, or maximum flow out of a cell, which is the flow of vehicles that traverse the cell at the critical density:

$$\bar{F}_i(\alpha_i) := \bar{v}_i \tilde{n}_i(\alpha_i) \; . \tag{5.2}$$

The flow out of a cell is limited by the sending function of that cell, which is the minimum of the capacity of the cell and the *demand* of vehicles in the cell: $S_i(\alpha_i(k)) =$

$\min(\bar{F}_i(\alpha_i), \bar{v}_i\phi_i(k))$. The flow entering a cell is limited by that cell's receiving function, which is the minimum of its capacity and its *supply* of vehicles: $R_i(\alpha_i(k)) = \min(\bar{F}_i(\alpha_i), (\bar{\phi}_i - \phi_i)w_i(\alpha_i))$, where w_i is the *shockwave speed*, the speed at which slowing waves of traffic propagate upstream: $w_i(\alpha_i) := \bar{v}_i\tilde{n}_i(\alpha_i)/(\bar{\phi}_i - \tilde{n}_i(\alpha_i))$. In the following, we use $f_i(k)$ to denote the flow out of cell i at time k and $y_i(k)$ to denote the flow into cell i. We use the standard superscripts for human-driven and autonomous flow, with the relationships $f_i^{\mathrm{h}}(k) + f_i^{\mathrm{a}}(k) = f_i(k)$ and $y_i^{\mathrm{h}}(k) + y_i^{\mathrm{a}}(k) = y_i(k)$. Accordingly,

$$\begin{aligned} \phi_i^{\mathrm{h}}(k+1) &= \phi_i^{\mathrm{h}}(k) + y_i^{\mathrm{h}}(k) - f_i^{\mathrm{h}}(k) \ , \\ \phi_i^{\mathrm{a}}(k+1) &= \phi_i^{\mathrm{a}}(k) + y_i^{\mathrm{a}}(k) - f_i^{\mathrm{h}}(k) \ . \end{aligned} \tag{5.3}$$

Since some cells might be a part of more than one path, we also track the paths of the human-driven and autonomous vehicles in each cell. We use $\mu_i^{\mathrm{h}}(p,k)$ and $\mu_i^{\mathrm{a}}(p,k)$ to denote the fraction of human-driven and autonomous vehicles, respectively, in cell i at time k that are taking path p. If cell i is not on path p, let $\mu_i^{\mathrm{h}}(p,k) = \mu_i^{\mathrm{a}}(p,k) = 0$.

Extending the development in [138], we formulate a calculation of the flow of mixed autonomous vehicles through general junctions. We define O as the set of intersections, or junctions, in the network. We use $\Xi(o)$ to denote the set of turning movements through intersection o, with a turning movement denoted by a tuple, such as $[i,o,j] \in \Xi(o)$, where i denotes the incoming cell, and j denotes the outgoing cell. As before, we consider all cells to have one direction of travel. For intersection o we define a set of conflict points $C(o)$, and $\Xi(c)$ denotes the set of turning movements through the intersection which pass through conflict point c, where $c \in C(o)$. These routes may have different priority levels, so for each $[i,o,j] \in \Xi(c)$ we define $\beta_{ioj}^c > 0$ as the priority of turning movement $[i,o,j]$ through conflict point c. Each conflict point has some supply R_c, which we assume is independent of the level of autonomy of the vehicles passing through

it. The relative priority of the turning movements will determine the relative flow of each turning movement through the conflict point. In a slight abuse of notation, we use $f_{ioj}(k)$ to denote the *total* flow of vehicles through turning movement $[i, o, j]$ at time k; we use $f^{\mathrm{h}}_{ioj}(k)$ and $f^{\mathrm{a}}_{ioj}(k)$ to denote the flow of human and autonomous vehicles, respectively, through the turning movement. We use $\Gamma(o)$ and $\Gamma^{-1}(o)$ to denote the set of cells exiting and entering junction o, respectively.

We then calculate the flows at each time step according to **procedure** FLOW CALCULATION, on the subsequent page.

An interpretation of this algorithm is as follows. The set A denotes the set of turning movements with flows that can yet be increased, and each turning movement is assigned a rate at which its flow increases. As sending and receiving limits are reached, turning movements are removed from A until there are no more turning movements left to increase.

In more concrete terms, first calculate the fraction of vehicles in each incoming cell which are headed to each outgoing cell. Then initialize all flows to 0 and initialized the unused sending and receiving capacity for each cell and conflict point. We then find relative rates of flow increase, δ_{ioj}, for the turning movements. In the loop, we calculate the similar rates of flow increases for the receiving cells and conflict points based on the rates previously found. Then, the flows are increased by the established rates until either a sending limit, a cell receiving limit, or conflict point capacity is reached. Any turning movement that has reached its sending limit is removed from the set of turning movements with further flow increases, A. Similarly, any turning movement that exits from a cell which has reached its receiving limit is removed from A, and the same with turning movements through conflict points which have reached their capacity. The loop repeats until A is empty.

Having calculated the flow through the intersection, the states of each cell is updated

1: **procedure** FLOW CALCULATION(Intersection o)

2:

$$\forall [i,o,j] \in \Xi(o),\ p^{\mathrm{h}}_{ioj} \leftarrow \sum_{p \in \mathcal{P}: i,j \in \mathcal{I}_p} \mu^{\mathrm{h}}_i(p,k)$$

$$p^{\mathrm{a}}_{ioj} \leftarrow \sum_{p \in \mathcal{P}: i,j \in \mathcal{I}_p} \mu^{\mathrm{a}}_i(p,k)$$

$$p_{ioj} \leftarrow \frac{n^{\mathrm{h}}_i(k) p^{\mathrm{h}}_{ioj} + n^{\mathrm{a}}_i(k) p^{\mathrm{a}}_{ioj}}{n^{\mathrm{h}}_i(k) + n^{\mathrm{a}}_i(k)}$$

3:

$$\forall [i,o,j] \in \Xi(o),\ f_{ioj} \leftarrow 0$$

$$f^{\mathrm{h}}_{ioj} \leftarrow 0$$

$$f^{\mathrm{a}}_{ioj} \leftarrow 0$$

$$\tilde{S}_{ioj} \leftarrow S_i(\alpha_i(k)) p_{ioj}$$

$$\forall (o,j) \in \Gamma(o),\ \tilde{R}_{oj} \leftarrow R_j(\alpha_j(k))$$

$$\forall c \in C(o),\ \tilde{R}_c \leftarrow R_c$$

4: For all $[i,o,j] \in \Xi(o)$, set δ_{ioj} such that

$$\forall [i,o,j] \in \Xi(o),$$

$$\forall [i',o,j'] \in \Xi(o),\ \frac{\delta_{ioj}}{\delta_{i'oj'}} = \frac{p_{ioj}}{p_{i'oj'}}\ ,$$

where i can equal i' and j can equal j', and

$$\forall c \in C(o);\ \forall [i,o,j] \in \Xi(c)$$

$$\forall [i',o,j'] \in \Xi(c),\ \frac{\delta_{ioj}}{\delta_{i'oj'}} = \frac{\beta^c_{ioj} p_{ioj}}{\beta^c_{i'oj'} p_{i'oj'}}\ ,$$

where i can equal i' and j can equal j'.

5: $A \leftarrow \Xi(o)$

6: **while** $A \neq \emptyset$ **do**

7:
$$\forall (o,j) \in \Gamma(o),\ \delta_{oj} \leftarrow \sum_{[i,o,j] \in A} p_{ioj}$$
$$\forall c \in C(o),\ \delta_c \leftarrow \sum_{[i,o,j] \in \Xi(c) \cap A} p_{ioj}$$

8:
$$\theta = \min\{ \min_{[i,o,j] \in A} \frac{\tilde{S}_{ioj}}{\delta_{ioj}}, \min_{(o,j) \in \Gamma(o), \delta_{oj} > 0} \frac{\tilde{R}_{oj}}{\delta_{oj}}, \min_{c \in C(o): \delta_c > 0} \frac{\tilde{R}_c}{\delta_c} \}$$

9:
$$\begin{aligned} \forall [i,o,j] \in A,\ f_{ioj} &\leftarrow f_{ioj} + \theta \delta_{ioj} \\ f^{\mathrm{h}}_{ioj} &\leftarrow f^{\mathrm{h}}_{ioj} + \theta \delta_{ioj} (1 - \alpha_i(k)) \\ f^{\mathrm{a}}_{ioj} &\leftarrow f^{\mathrm{a}}_{ioj} + \theta \delta_{ioj} \alpha_i(k) \\ \tilde{S}_{ioj} &\leftarrow \tilde{S}_{ioj} - \theta \delta_{ioj} \\ \forall (oj) \in \Gamma(o),\ \tilde{R}_{oj} &\leftarrow \tilde{R}_{oj} - \theta \delta_{oj} \\ \forall c \in C(o),\ \tilde{R}_c &\leftarrow \tilde{R}_c - \theta \delta_{oj} \end{aligned}$$

10:
$$\begin{aligned} A &\leftarrow A \setminus \{[i,o,j] \in A : \tilde{S}_{ioj} = 0\} \\ A &\leftarrow A \setminus \{[i,o,j'] \in A : \tilde{R}_{oj} = 0 \wedge p_{ioj'} > 0\} \\ A &\leftarrow A \setminus \{[i,o,j] \in c : \tilde{R}_c = 0\ \forall c \in C(o)\} \end{aligned}$$

11: **end while**

12: **return** f_{ioj}, f^{h}_{ioj}, f^{a}_{ioj}, $\forall [i,o,j] \in \Xi(o)$

13: $a = 5$

14: **end procedure**

as follows. We compute the incoming flows for the outgoing cells as follows:

$$\begin{aligned} \forall (o,j) \in \Gamma(o),\, y_j^{\mathrm{h}}(k) &= \sum_{[i,o,j]\in\Xi(o)} f_{ioj}^{\mathrm{h}} \\ y_j^{\mathrm{a}}(k) &= \sum_{[i,o,j]\in\Xi(o)} f_{ioj}^{\mathrm{a}} \\ y_j(k) &= y_j^{\mathrm{h}}(k) + y_j^{\mathrm{a}}(k) \end{aligned} \tag{5.4}$$

To calculate the outgoing flows of the incoming cells,

$$\begin{aligned} \forall (i,o) \in \Gamma^{-1}(o),\, f_i^{\mathrm{h}}(k) &= \sum_{[i,o,j]\in\Xi(o)} f_{ioj}^{\mathrm{h}} \\ f_i^{\mathrm{a}}(k) &= \sum_{[i,o,j]\in\Xi(o)} f_{ioj}^{\mathrm{a}} \\ f_i(k) &= f_i^{\mathrm{h}}(k) + f_i^{\mathrm{a}}(k)\ , \end{aligned} \tag{5.5}$$

where $\Gamma^{-1}(o)$ denotes the set of cells going into intersection o. (5.3) updates the human-driven and autonomous vehicle densities of each cell at the next time step. To update the fraction of vehicles in the outgoing cells on each path,

$$\begin{aligned} &\forall (o,j) \in \Gamma(o), \\ &\mu_j^{\mathrm{h}}(p,k+1) = \frac{\sum_{[i',o,j]\in\Xi(o)} f_i^{\mathrm{h}}(k)\mu_i^{\mathrm{h}}(p,k) + \mu_j^{\mathrm{h}}(p,k)(n_j^{\mathrm{h}}(k) - f_j^{\mathrm{h}}(k))}{n_j^{\mathrm{h}}(k+1)}\ , \\ &\mu_j^{\mathrm{a}}(p,k+1) = \frac{\sum_{[i',o,j]\in\Xi(o)} f_i^{\mathrm{a}}(k)\mu_i^{\mathrm{a}}(p,k) + \mu_j^{\mathrm{a}}(p,k)(n_j^{\mathrm{a}}(k) - f_j^{\mathrm{a}}(k))}{n_j^{\mathrm{a}}(k+1)}\ . \end{aligned}$$

Accidents. To evaluate the performance of the developed RL policy in reacting to disturbances, we consider stochastic accidents occuring in the network, each of which causes one lane to be closed. We let accidents occur in any cell at any time with equal

probability as long as the jam density does not decrease below the current density of the cell. Each accident is cleared out after some number of time steps, drawn from a Poisson distribution. If $\bar{b}_i$ lanes of cell i are closed due to accidents, then the jam density and the critical density for the cell reduce to $(b_i - \bar{b}_i)/b_i$ of their original values. Thus, accidents introduce time-dependency to these variables.

5.3 Network dynamics: routing for humans and autonomous vehicles

As mentioned above, we consider a network with a set of possible paths $\mathcal{P}$. We use λ^{h} and λ^{a} to denote the human-driven and autonomous vehicle demands, respectively. We model all vehicles entering the network as entering a *queue*, a single cell with infinite capacity. We use 0 for the index of this cell. The routing choices of autonomous vehicles leaving the queue is determined by the central controller, and the routing choices of human-driven vehicles leaving it are determined from the latencies associated with each path, detailed below.

5.3.1 Human choice dynamics

In general, people wish to minimize the amount of time spent traveling. However, people do not change routing choices instantaneously in response to new information; rather they have some inertia and only change strategies sporadically. Moreover, we assume people only account for current conditions and do not strategize based on predictions of the future [139]. Accordingly, we use an *evolutionary dynamic* to describe how a population of users choose their routes.[2] Specifically, we model the human driver

[2]Alternately, one could model individual users as learning agents, posing it as a Multi-Agent Reinforcement Learning problem. However, we consider large networks with too many human agents for this

population as following Hedge Dynamics, also called Log-linear Learning [140–142].

Let $(\mu_0^{\mathrm{h}}(p,k))_{p\in\mathcal{P}}$ represent the initial routing of human-driven vehicles at time k; accordingly, $\sum_{p\in\mathcal{P}}\mu_0^{\mathrm{h}}(p,k)=1$ for all k. Humans will update their routes based on their estimates of how long it will take to traverse each path. However, it is not always possible to predict travel time accurately on general networks, since vehicles entering later on a different path may influence the travel time of vehicles entering earlier. Because of this, we consider that humans have an estimate $\hat{\ell}_p(k)$ of the true latency $\ell_p(k)$. With these estimates, the routing vector is updated as follows.

$$\mu_0^{\mathrm{h}}(p,k+1)=\frac{\mu_0^{\mathrm{h}}(p,k)\exp(-\eta^{\mathrm{h}}(k)\hat{\ell}_p(k))}{\sum_{p'\in\mathcal{P}}\mu_0^{\mathrm{h}}(p',k)\exp(-\eta^{\mathrm{h}}(k)\hat{\ell}_{p'}(k))}\,. \tag{5.6}$$

The ratio of the volume of vehicles using a path at successive time steps is inversely proportional to the exponential of the delay experienced by users of that path. The learning rate $\eta^{\mathrm{h}}(k)$ may be decreasing or constant. Krichene *et al.* introduce this model in the context of humans' routing choices and simulate a congestion game with Amazon Mechanical Turk users to show the model accurately predicts human behavior [143]. We note that though we use this specific model for human choice for our simulations, the control method described later does not require this specific choice of human choice model. Our theoretical analysis similarly is not restricted to this choice of dynamics and works for any human choice model in which all fixed points of the dynamics satisfy human selfishness.

5.3.2 Autonomous vehicle control policy

We assume that we have control over the routing of autonomous vehicles. We justify this by envisioning a future in which autonomous vehicles are offered as a service rather

to be feasible.

than a consumer product. We then assume that a city can coordinate with the owner of an autonomous fleet to decrease congestion in the city. Moreover, unlike traditional tolling, coordination between autonomous vehicles and city infrastructure allows for fast-changing and geographically finely quantized tolls, enabling routing control to be achieved through incentives [92,116]. The initial routing of autonomous vehicles is then our control parameter by which we influence the state of traffic on the network. Consistent with the previous notation, we denote the initial autonomous routing as $(\mu_0^{\mathrm{a}}(p,k))_{p\in\mathcal{P}} \in \mathbb{R}_{\geq 0}^{|\mathcal{P}|}$, where $\sum_{p\in\mathcal{P}} \mu_0^{\mathrm{a}}(p,k) = 1$.

We assume the existence of a central controller, or social planner, which dictates μ_0^{a} by processing the state of the network. At each time step, we let the controller observe:

- the number of human-driven and autonomous vehicles in each cell and in the queue,
- binary states for each lane that indicates whether the lane is closed due to an accident or not.

We use deep RL to arrive at a policy for the social planner to control the autonomous vehicle routing, μ_0^{a}. Since the state space is very large and both state and action spaces are continuous, a dynamic programming-based approach is infeasible. For instance, even if we discretized the spaces, say with 10 quantization levels, and did not have accidents, we would have 10^{82} possible states and 10 actions for a moderate-size network with only 2 paths and 40 cells in total.

We wish to minimize the total latency experienced by users, which is equal to summing over time the number of users in the system at each time step. Accordingly, the stage cost is:

$$J(k) = \sum_{i\in\mathcal{I}} \phi_i(k)\,. \tag{5.7}$$

Due to their high performance in continuous control tasks [120,144], we employ policy gradient methods to learn a policy that produces μ_0^{a} given the observations. Specifically,

we use state-of-the-art PPO with an objective function augmented by adding an entropy bonus for sufficient exploration [120, 145]. We build a deep neural network, and train it using Adam optimizer [146]. An overview of the PPO method and the set of parameters we use are presented in the appendix (Sec. 5.7.3 and Sec. 5.7.4). Each episode has a fixed number of time steps.

In order to evaluate the performance of our control policy, we use three criteria. The first is the *throughput* of the network – we wish to have a policy that can serve any feasible demand, thereby stabilizing the queue. The second is the *average delay* experienced by users of the network, which we measure by counting the number of vehicles in the system. The third is the *convergence* to some steady state; we wish to avoid wild oscillations in congestion. To contextualize the performance of our control policy in this framework, we first establish the performance of *equilibria* of the network.

5.4 Equilibrium analysis

In this section, we examine the possible *equilibria* of our dynamical system, which characterize the possible steady state behaviors of the system. A network with a given demand can have a variety of equilibria with varying average user delay. If our control achieves overall delay equal to that of the best possible equilibrium, it is a successful policy. Section 5.5 shows empirically that our learned policy can achieve the best equilibrium in a variety of settings.

In this section, we first formulate an optimization which solves for the most efficient equilibrium, which is computationally hard. Motivated by this, we restrict the class of networks considered and prove theoretical properties of this restricted class. Using these properties, we formulate a new optimization formulation to solve for the most efficient equilibrium and prove it is solvable in polynomial time.

5.4.1 Equilibrium Formulation

We define two notions of equilibrium: one related to the vehicle flow dynamics, and one related to human choice dynamics.[3]

Definition 5.1 (Path Equilibrium)**.** *We define a* path equilibrium *for path p as a set of cell densities $(\phi_i^h(k), \phi_i^a(k))_{i\in\mathcal{I}_p}$ that, for a given constant flow entering the first cell on the path, $y_i^h(k)$ and $y_i^a(k)$, the cell densities are constant.*

Definition 5.2 (Network Equilibrium)**.** *We define a* network equilibrium *as a set of cell densities $(\phi_i^h(k), \phi_i^a(k))_{i\in\mathcal{I}}$ and human vehicle routing $(\mu_0^h(p,k))_{p\in\mathcal{P}}$, such that for a given constant entering flow $y_0^h(k)$ and $y_0^a(k)$ and a given constant autonomous vehicle routing $(\mu_0^a(p,k))_{p\in\mathcal{P}}$, the human vehicle routing, subject to the dynamics in* (5.6)*, is constant.*

We are interested in satisfying both notions of equilibrium – both the path equilibrium, which deals with the vehicle flow dynamics, and the network equilibrium, which deals with the human choice dynamics. Accordingly, the pair can be considered a Stackelberg Equilibrium for a leader controlling the autonomous vehicles who wishes to maximize the social utility in the presence of selfish human demand. We formulate the following optimization to solve for the most efficient equilibrium (satisfying both notions of equilibrium defined above), *i.e.* the equilibrium which minimizes the total travel time of all users of the network. We drop all time indices since we consider quantities that are constant over time.

$$\min_{(\phi_i^{\mathrm{h}},\phi_i^{\mathrm{a}},f_i^{\mathrm{h}},f_i^{\mathrm{a}},y_i^{\mathrm{h}},y_i^{\mathrm{a}},\mu_i^{\mathrm{h}}(p),\mu_i^{\mathrm{a}}(p),\ell_p)_{i\in\mathcal{I},p\in\mathcal{P}}} \sum_{i\in\mathcal{I}} \phi_i$$

[3]These define equilibria in the sense of dynamical systems, and do not strictly correspond to game-theoretic notions of equilibria. Under Assumption 5.3 below, the set of equilibria for the dynamics of human choice will correspond to the set of Nash Equilibria where the payoff is the path latency.

$$\text{s.t. } \forall o \in O : \textbf{procedure } \textsc{Flow Calculation}(\text{Intersection } o)$$

$$(5.4), (5.5)$$

$$\forall i \in \mathcal{I} : y_i^{\mathrm{h}} = f_i^{\mathrm{h}}, \quad y_i^{\mathrm{a}} = f_i^{\mathrm{a}}, \quad \alpha_i = \phi_i^{\mathrm{a}}/(\phi_i^{\mathrm{h}} + \phi_i^{\mathrm{a}})$$

$$\sum_{i' \in \mathcal{U}_i} f_{i'}^{\mathrm{h}} \mu_{i'}^{\mathrm{h}}(p) = f_i^{\mathrm{h}} \mu_i^{\mathrm{h}}(p), \ \sum_{i' \in \mathcal{U}_i} f_{i'}^{\mathrm{a}} \mu_{i'}^{\mathrm{a}}(p) = f_i^{\mathrm{a}} \mu_i^{\mathrm{a}}(p)$$

$$\ell_p = \sum_{i \in p} (\phi_i^{\mathrm{h}} + \phi_i^{\mathrm{a}})/(f_i^{\mathrm{h}} + f_i^{\mathrm{a}})$$

$$\forall p, p' \in \mathcal{P} : \mu_0^{\mathrm{h}}(p)(\ell_p - \ell_{p'}) \leq 0$$

While this formulation solves for the most efficient equilibrium of any traffic network, it is computationally difficult, especially due to the final constraint. Due to this, we introduce a restricted class of networks that we consider for the remainder of this section, which allows us to compute equilibria in polynomial time with respect to the number of paths.

Definition 5.3 (Bottleneck)**.** *We define a* bottleneck *as a regular junction at which the number of lanes decreases, decreasing the capacity of the cells.*

Assumptions 5.1. *We consider a parallel network in which leaving the first cell, vehicles choose a path and paths do not share cells, meaning that each cell is identified with only one path, aside from the downstream-most cell which has infinite capacity. We further consider that all cells in the path have the same model parameters, except for a bottleneck after the m_p^n upstream-most cells.*

In other words, we consider a parallel network where each path is composed of identical cells except for a single junction with a decrease in the number of lanes. Fig. 5.1 shows an example of such a network. For ease of analysis, we first establish properties of Path Equilibria, then Network Equilibria.

5.4.2 Path equilibrium

As mentioned above, we restrict our considered class of paths to those with a single bottleneck, meaning one point on the path at which cell capacity drops. Formally, we consider each path p to have m_p^{n} cells, each with b_p^{n} lanes, followed by m_p^{b} cells downstream, each with b_p^{b} lanes, where $b_p^{\mathrm{b}} < b_p^{\mathrm{n}}$. We define $r_p := b_p^{\mathrm{b}}/b_p^{\mathrm{n}} \in (0, 1)$.

In a slight abuse of notation, we use the subscript p for parameters that are constant over a path under Assumption 5.1, and the superscript n for cells before the bottleneck and b for the bottleneck and cells downstream of it. We now present a theoretical result that completely analytically characterizes the path latencies that can occur at equilibrium.

Theorem 5.1. *Under Assumption 5.1 a path p with flow dynamics described in Section 5.2 that is at Path Equilibrium will have the same autonomy level in all cells. Denote this autonomy level α_p. If the vehicle flow demand is strictly less than the minimum cell capacity, the path will have no congested cells. Otherwise, the path will have one of the following latencies, where $\gamma_p \in \{0, 1, 2, \ldots, m_p^n\}$:*

$$\ell_p = \frac{|\mathcal{I}_p|}{\bar{v}_p} + \gamma_p \frac{(1 - r_p)\bar{\phi}_p^n(\alpha_p h_p^a + (1 - \alpha_p)h_p^h)}{r_p \bar{v}_p b_p^n} \,.$$

Proof. The proof is composed of three lemmas. We first establish a property of path equilibria that allows us to treat the vehicle flow as if it were composed of a single car type. With this, we use the CTM to characterize possible equilibria on a path. We then derive the delay associated with each congested cell. Combining the latter two lemmas yields the theorem.

Lemma 5.1. *A path in equilibrium with nonzero incoming flow has the same autonomy level in all cells of the path, which is equal to the autonomy level of the vehicle flow onto*

the path. Formally, a path p with demand $(\bar{\lambda}_p^h, \bar{\lambda}_p^a)$ in equilibrium has, for all cells i in $\mathcal{I}_p$,

$$\alpha_i = \bar{\lambda}_p^a / (\bar{\lambda}_p^h + \bar{\lambda}_p^a) \ .$$

We defer the proof of the lemma to the appendix. With this, our path equilibria analysis simplifies to that of single-typed traffic, with the autonomy level treated as a variable parameter. The next lemma, similarly to Theorem 4.1 of [5], completely characterizes the congestion patterns that can occur in cell equilibria. For this lemma, we consider the cell indices in a path to be increasing, where the cell immediately downstream from a cell i has index $i+1$.

Lemma 5.2. *Under Assumption 5.1, if the demand on a path is less than the minimum capacity of its cells, they will be uncongested at path equilibrium. Otherwise, a path with demand equal to the minimum cell capacity will have m_p^n possible path equilibria, corresponding to one of the following sets of congested cells, where j is the index of the m_p^nth cell:*

$$\big\{\emptyset, \{j\}, \{j-1, j\}, \ldots, \{j - m_p^n + 1, \ldots, j-2, j-1, j\}\big\} \ .$$

Proof. As mentioned above, this lemma relates closely to Theorem 4.1 of [5]. However, we cannot directly apply that theorem due to differing assumptions; namely they assume $\bar{F}_{i+1} = (\bar{\phi}_i - \tilde{n}_i) w_i$ for all i. We therefore offer a similar proof, tailored to our assumptions.

For ease of notation, we drop all path subscripts p as well as the cell index for the free-flow velocity parameter $\bar{v}$. In light of Lemma 5.1, we also suppress the autonomy level arguments to capacity $\bar{F}_i$ and critical density $\tilde{n}_i$. The flow equation then becomes $f_i = \min(\bar{v}\phi_i,\ (\bar{\phi}_{i+1} - \phi_{i+1}) w_{i+1},\ \bar{F}_i, \bar{F}_{i+1})$.

We begin by proving that if the vehicle flow demand is strictly less than the minimum capacity, *i.e.* the bottleneck capacity, then the only equilibrium has no congested cells.

Let us use j' to denote the index of the final cell in the path. Under Assumption 5.1 there is no supply limit to the flow exiting a path, so $f_{j'} = \min(\bar{v}\phi_{j'}, \bar{F}_{j'})$. Since $f_0 = f_{j'} < \bar{F}_{j'}$, $f_0 = f_{j'} = \bar{v}\phi_{j'}$. The definition of capacity, $\bar{F}_i = \bar{v}\tilde{n}_i$, then implies that $\phi_{j'} < \tilde{n}_{j'}$, meaning that cell j' is uncongested, so $\bar{v}\phi_{j'} < (\bar{\phi}_{j'} - \phi_{j'})w_{j'}$.

This is the base case for a proof by induction. Consider cell i that is uncongested (*i.e.* $\phi_i < \tilde{n}_i$). Since by assumption all cells have flow strictly less than the cell's capacity, $f_i = \bar{v}\phi_i < \bar{F}_i$. Then consider the flow entering cell i: $f_{i-1} = \min(\bar{v}\phi_{i-1}, (\bar{\phi}_i - \phi_i)w_i, \bar{F}_{i-1}) = f_i < \bar{F}_i < (\bar{\phi}_i - \phi_i)w_i$.

The fact that $\bar{F}_i \leq \bar{F}_{i-1}$ then implies that $f_{i-1} = \bar{v}\phi_{i-1}$, so cell $i-1$ is uncongested, proving the lemma's first statement.

The second statement assumes the flow on the path is equal to the minimum capacity. The cells in the bottleneck segment all have the same capacity, which we denote $\bar{F}^{\mathrm{b}}$; this capacity is less than the capacity of the cells in the nonbottleneck segment. This means all bottleneck cells will be operating at capacity (and therefore have vehicle density equal to their critical density); flow on the path is therefore equal to $\bar{F}^{\mathrm{b}}$.

We now turn to the nonbottleneck segment. We first note that if a nonbottleneck cell is uncongested then the preceeding cell must be uncongested as well, using the same reasoning as that proving the first statement above. Next, consider the flow out of the downstream-most cell of the nonbottleneck segment: $f_j = \min(\bar{v}\phi_j, (\bar{\phi}_{j+1} - \phi_{j+1})w_{j+1}, \bar{F}_j) = \bar{F}^{\mathrm{b}} < \bar{F}_j$, so $f_j = \min(\bar{v}\phi_j, (\bar{\phi}_{j+1} - \phi_{j+1})w_{j+1})$. Cell j can be uncongested, in which case the cell density is such that $\bar{v}\phi_j = \bar{F}^{\mathrm{b}}$, or the cell can be congested, in which case the second term dominates. Then, if nonbottleneck cell i is congested, the flow into it is $f_{i-1} = \min(\bar{v}\phi_{i-1}, (\bar{\phi}_i - \phi_i)w_i)$. Again, to achieve this flow, cell $i-1$ can be either congested or uncongested. As shown above, if uncongested, then all upstream cells must be uncongested as well, yielding the second statement in the lemma. □

We use these properties to find a closed-form expression for the latency incurred by traveling through a bottleneck cell, which when combined with Lemma 5.2, completes the proof.

Lemma 5.3. *The latency incurred by traveling through a congested cell is as follows.*

$$\frac{1}{\bar{v}_p} + \frac{(1-r_p)\bar{\phi}_p^n(\alpha_p h_p^a + (1-\alpha_p)h_p^h)}{r_p \bar{v}_p b_p^n} \,.$$

Proof. Recall that we assume paths have a uniform free-flow velocity across all cells in a path, where path p has free-flow velocity $\bar{v}_p$. We define $[m_p^{\mathrm{n}}]$ as the set of cells before the bottleneck, which have b_p^{n} lanes. The remaining cells, with indices in the set $\mathcal{I}_p \setminus [m_p^{\mathrm{n}}]$, have b_p^{b} lanes. Further recall the definition $r_p = b_p^{\mathrm{b}}/b_p^{\mathrm{n}}$. Let $\bar{F}_p^{\mathrm{n}}(\alpha_p)$ denote the capacity of the cells before the bottleneck of path p with autonomy level α_p and let $\bar{F}_p^{\mathrm{b}}(\alpha_p)$ be the same for the bottleneck cell. Note that $\bar{F}_p^{\mathrm{b}}(\alpha_p) = r_p \bar{F}_p^{\mathrm{n}}(\alpha_p)$. Similarly, let $w_p^{\mathrm{n}}(\alpha_p)$ and $w_p^{\mathrm{b}}(\alpha_p)$ denote the shockwave speed for prebottleneck cells and bottleneck cell, respectively, on path p with autonomy level α_p, as with jam densities $\bar{\phi}_p^{\mathrm{n}}$ and $\bar{\phi}_p^{\mathrm{b}}$ and critical densities $\tilde{n}_p^{\mathrm{n}}(\alpha_p)$ and $\tilde{n}_p^{\mathrm{b}}(\alpha_p)$.

Lemma 5.2 establishes all possible combinations of congested cells that a path at equilibrium can experience. We now investigate how much delay each configuration induces on the path, parameterized by the autonomy level of the path. By Lemma 5.2 and the definitions of r and capacity (5.2),

$$f_p = \bar{F}_p^{\mathrm{b}} = r_p \bar{F}_p^{\mathrm{n}}(\alpha_p) = r_p w_p^{\mathrm{n}}(\alpha_p)(\bar{\phi}_p^{\mathrm{n}} - \tilde{n}_p^{\mathrm{n}}(\alpha_p)) \,. \tag{5.8}$$

Let $n_p^{\mathrm{c}}(\alpha_p)$ denote the vehicle density in a congested cell on path p, which we know must occur upstream of the bottleneck (Lemma 5.2). Then, the flow entering a congested cell before the bottleneck is $f_p = w_p^{\mathrm{n}}(\alpha_p)(\bar{\phi}_p^{\mathrm{n}} - n_p^{\mathrm{c}}(\alpha_p))$. Equating this with (5.8), we find

$n_p^{\mathrm{c}}(\alpha_p) = (1 - r_p)\bar{\phi}_p^{\mathrm{n}} + r_p\tilde{n}_p^{\mathrm{n}}(\alpha_p)$.

To use this to find the latency incurred by traveling through a congested cell, we divide the density by the flow, as follows.

$$\begin{aligned}\frac{n_p^{\mathrm{c}}(\alpha_p)}{f_p} &= \frac{n_p^{\mathrm{c}}(\alpha_p)}{\bar{F}_p^{\mathrm{b}}(\alpha_p)} = \frac{(1 - r_p)\bar{\phi}_p^{\mathrm{n}} + r_p\tilde{n}_p^{\mathrm{n}}(\alpha_p)}{r_p\bar{v}_p\tilde{n}_p^{\mathrm{n}}(\alpha_p)} \\ &= \frac{1}{\bar{v}_p} + \frac{(1 - r_p)\bar{\phi}_p^{\mathrm{n}}(\alpha_p h_p^{\mathrm{a}} + (1 - \alpha_p)h_p^{\mathrm{h}})}{r_p\bar{v}_p b_p^{\mathrm{n}}}\,. \qquad \square\end{aligned}$$

Together, the lemmas prove the theorem. □

The two terms above are the free-flow delay and the per-cell latency due to congestion, respectively. Theorem 5.1 allows us to calculate the possible latencies of a path as a function of its autonomy level α_p. Since in a network equilibrium all used paths have the same latency, we can calculate network equilibria more efficiently than comprehensively searching over all possible routings. However, equilibria may not exist, even with a fine time discretization – in equilibrium the path latencies must be equal, but by Theorem 5.1, road latency is a function of the integer γ_p. To avoid this artifact, when analyzing network equilibria we consider the cells to be small enough that we can consider the continuous variable $\gamma_p \in [0, m_p^{\mathrm{n}}]$.

5.4.3 Network equilibrium

We define the best equilibrium to be the equilibrium that serves a given flow demand with minimum latency. We are now ready to establish properties of network equilibria, as well as how to compute the best equilibria. We use the following two assumptions in our analysis of network equilibrium.

Assumptions 5.2. *No two paths have the same free-flow latency.*

Assumptions 5.3. *The initial choice distribution has positive human-driven and autonomous vehicle flow on each path.*

Note that the Assumption 5.2 is not strictly necessary but is useful for easing analysis. A similar analysis could be performed in its absence. We justify Assumption 5.3 by noting that humans are not entirely rational and that our choice model does not capture all reasons a person may wish to choose a route, and some small fraction of people will choose routes that seem less advantageous at first glance.

Theorem 5.2. *Under Assumptions 5.1 and 5.2, a routing that minimizes total latency when all users (both human drivers and autonomous users) are selfish can be computed in $O(|\mathcal{P}|^3 \log |\mathcal{P}|)$ time. A routing that minimizes total latency when human drivers are selfish and autonomous users are controlled can also be computed in $O(|\mathcal{P}|^3 \log |\mathcal{P}|)$ time.*

Proof. To establish properties of network equilibria, we introduce some notation. We use $t_p = |\mathcal{I}_p|/\bar{v}_p$ to denote the free-flow latency of path p. We also use $\mathcal{P}_{\leq t_p} = \{p' \in \mathcal{P} : t_{p'} \leq t_p\}$, which denotes the set of paths with free-flow latency less than or equal to that of path p. We similarly define the expression with other comparators, *e.g.* $\mathcal{P}_{<t_p}$ or $\mathcal{P}_{>t_p}$.

This proposition follows from Definition 5.2 and Assumption 5.3. The next lemma follows, with proof deferred to the appendix.

Proposition 5.1. *In a network equilibrium,*

1. *All paths with selfish drivers have the same latency, and*
2. *All paths without selfish drivers have equal or greater latency.*

Lemma 5.4. *If the set of equilibria contains a routing with positive flow only on paths $\mathcal{P}_{\leq t_p}$, then there exists a routing in the set of equilibria in which path p is in free-flow.*

Lemma 5.5. *Under Assumption 5.2, if some users are selfish and some users are not selfish, then the best equilibrium will have the following properties:*

1. *the path with largest free-flow latency used by selfish users will be in free-flow,*
2. *all paths with lower free-flow latency will be congested,*
3. *paths with greater free-flow latency may have nonselfish users, and*
4. *paths used with larger free-flow latency that have nonselfish users on them will be at capacity, except perhaps the path with largest free-flow latency used by nonselfish users.*

Proof. Consider a network with some selfish and some non-selfish (controlled) users. Let p denote the path with the longest free-flow latency that contains selfish users. For the purpose of contradiction, let this path contain congested cells, and let this be the best equilibrium. Fix the nonselfish flow on all roads with longer free-flow latency than p. By Lemma 5.4, there exists an equilibrium for the selfish users in which p is in free-flow. This results in less latency for the users on path p, and no selfish user will have greater delay (Proposition 5.1). This contradicts the premise, proving the first property.

The second property follows directly from Proposition 5.1 and Assumption 5.2. The third property follows from the definition of nonselfish users, which can take a path with a larger latency than other available paths. The best equilibrium minimized total latency. If there was a road with nonselfish users that was not at capacity, while another path with higher latency has positive flow, this would not be the best equilibrium, since a more efficient routing would shift flow from the higher latency path to the lower latency one. This yields the final property. □

Using these properties, we prove Theorem 5.2. We first consider the setting in which all users are selfish. We use $\hat{\ell}_p(\alpha_p)$ to denote the per-cell latency due to congestion, *i.e.* $\hat{\ell}_p(\alpha_p) = \frac{(1-r_p)\bar{\phi}_p^{\mathrm{n}}(\alpha_p h_p^{\mathrm{a}} + (1-\alpha_p)h_p^{\mathrm{h}})}{r_p \bar{v}_p b_p^{\mathrm{n}}}$. Lemma 5.5 implies that for a given demand, all equilibria in the set of most efficient equilibria for that demand have one path that is in free-flow. We can then formulate the search for a best equilibrium as an optimization. We are

helped by the fact that the best equilibria will use the minimum number of feasible paths, since all users experience the same delay. Then, for each candidate free-flow path (denote with index p'), check feasibility of only using paths $\mathcal{P}_{\leq t_{p'}}$, and choose a routing that minimizes $|\mathcal{P}_{\leq t_{p'}}|$, *i.e.* the number of roads used. The reason for minimizing the number of used roads is that all users are experiencing the same latency (Proposition 5.1) and in the best equilibrium, the road with flow on it that has longest free-flow latency will be in free-flow (Lemma 5.5). The feasibility can be checked as follows, with an optimization that utilizes Lemma 5.5.

$$\begin{aligned}
\underset{(f_p^{\mathrm{h}}, f_p^{\mathrm{a}})_{p\in\mathcal{P}_{\leq t_{p'}}},\, \gamma\in\prod_{p\in\mathcal{P}_{<t_{p'}}}[0,m_p^{\mathrm{n}}]}{\operatorname{argmin}} \quad & 1 \\
\text{s.t.} \quad & \sum_{p\in\mathcal{P}_{\leq t_{p'}}} f_p^{\mathrm{h}} = \bar{\lambda}^{\mathrm{h}},\ \sum_{p\in\mathcal{P}_{\leq t_{p'}}} f_p^{\mathrm{a}} = \bar{\lambda}^{\mathrm{a}} \\
& f_{p'}^{\mathrm{h}} + f_{p'}^{\mathrm{a}} \leq \bar{F}_{p'}\left(\frac{f_{p'}^{\mathrm{a}}}{f_{p'}^{\mathrm{h}} + f_{p'}^{\mathrm{a}}}\right) \\
\forall p \in \mathcal{P}_{<t_{p'}} : \ & \gamma_p \hat{\ell}_p\left(\frac{f_p^{\mathrm{a}}}{f_p^{\mathrm{h}} + f_p^{\mathrm{a}}}\right) = t_{p'} - t_p \\
& f_p^{\mathrm{h}} + f_p^{\mathrm{a}} = \bar{F}_p\left(\frac{f_p^{\mathrm{a}}}{f_p^{\mathrm{h}} + f_p^{\mathrm{a}}}\right) .
\end{aligned}$$

The last constraint yields an affine relationship between f_p^{h} and f_p^{a} for paths $\mathcal{P}_{<t_{p'}}$. Solving for f_p^{h} and plugging into the first constraint yields an affine relationship between γ_p and f_p^{a}. This way, the optimization can be converted to a linear program, and we must solve $\log|\mathcal{P}|$ linear programs to search the minimum feasible p'.

This formulation assumes that all vehicles are selfish. If instead we consider selfish human drivers and fully controlled autonomous users, we can construct a similar optimization to find the best equilibrium. For each choice of free-flow path p', we minimize the total latency of the autonomous vehicles not on free-flow paths. We then choose the routing corresponding to free-flow path p' that minimizes total latency (which may not

necessarily minimize the number of paths used by human drivers). For each candidate free-flow path p' we solve the following optimization.

$$\begin{aligned}
&\underset{(f_p^{\mathrm{h}},f_p^{\mathrm{a}})_{p\in\mathcal{P}},\,\gamma\in\prod_{p\in\mathcal{P}_{<t_{p'}}}[0,m_p^{\mathrm{n}}]}{\operatorname{argmin}} \sum_{p\in\mathcal{P}_{>t_{p'}}} f_p^{\mathrm{a}} t_p \\
&\qquad\qquad \text{s.t.} \sum_{p\in\mathcal{P}_{\leq t_{p'}}} f_p^{\mathrm{h}} = \bar{\lambda}^{\mathrm{h}}, \sum_{p\in\mathcal{P}} f_p^{\mathrm{a}} = \bar{\lambda}^{\mathrm{a}} \\
&\qquad\qquad\qquad f_{p'}^{\mathrm{h}} + f_{p'}^{\mathrm{a}} \leq \bar{F}_{p'}\left(\frac{f_{p'}^{\mathrm{a}}}{f_{p'}^{\mathrm{h}} + f_{p'}^{\mathrm{a}}}\right) \\
&\forall p \in \mathcal{P}_{<t_{p'}} : \gamma_p \hat{\ell}_p(\alpha_p) = t_{p'} - t_p \\
&\qquad\qquad\qquad f_p^{\mathrm{h}} + f_p^{\mathrm{a}} = \bar{F}_p\left(\frac{f_p^{\mathrm{a}}}{f_p^{\mathrm{h}} + f_p^{\mathrm{a}}}\right) \\
&\forall p \in \mathcal{P}_{>t_{p'}} : f_p^{\mathrm{a}} \leq \bar{F}^{\mathrm{b}}(1)
\end{aligned}$$

This can be reformulated as a linear program by the same mechanism. Again, we solve $\log|\mathcal{P}|$ linear programs and choose the one corresponding to the minimum feasible p'. □

Using these properties to compute optimal equilibria, we establish a framework for understanding the performance of our learned control policy. If the policy can reach the best equilibrium latency starting from arbitrary path conditions we view the policy as successful. We use this baseline to evaluate our experimental results in the following section.

A question then arises: if we have computed the best possible equilibria, why do we not directly implement that control? This approach is not fruitful, since the theoretical analysis of best equilibria gives the control policy only in the steady state. In practice, the network can start in any state, including worse equilibria, from which good equilibria will not emerge when autonomous vehicles unilaterally use their routing in the best equilibrium. Besides, our equilibrium analysis is limited to parallel networks and extending

it to more general networks would yield a nonconvex optimization problem. A dynamic policy which depends on the current traffic state is therefore needed to guide the network to the best equilibrium. As shown in the following section, the policy learned via deep reinforcement learning achieves this guidance and reaches the best equilibrium in a variety of settings.

5.5 Experiments and Results

In all of the experiments, we adopt the following parameters. All vehicles are 4 meters long. Human drivers keep a 2 second headway distance, whereas autonomous cars can keep 1 second. Each time step corresponds to 1 minute of real-life simulation. Each episode during deep RL training covers 5 hours of real-life simulation (300 time steps). In test time, we simulate 6 hours of real-life (360 time steps) to ensure the RL policy did not learn to minimize the latency in the first 300 time steps and leave excess vehicles in the network at the end. We divide paths into the cells such that it takes 1 time step to traverse each cell in free-flow. We initialize $\phi_i(0) \sim \text{unif}(0, 1.2\tilde{n}_i)$ for all $i \in \mathcal{I}_p$ for all $p \in \mathcal{P}$. We set the standard deviations of the zero-mean Gaussian demand noise to be $\bar{\lambda}^{\text{h}}/10$ and $\bar{\lambda}^{\text{a}}/10$ for human-driven and autonomous vehicles, respectively.

Our overall control scheme can be seen in Fig. 5.1. As the learning model, we build a two-hidden-layer neural network, with each layer having 256 nodes. We train an RL agent for each configuration that we will describe later on in simulated traffic networks that are based on the mixed-autonomy traffic model and the dynamics that we described in Sections 5.2 and 5.3. All trainings simulate 40 million time steps.[4] Depending on whether we evaluate our RL-based approach with (or without) the accidents, we enable (or disable) accidents at the training phase. However, we note that the number of possible accident

[4]Other hyperparameter values we use for PPO are in the Appendix.

configurations in the network is far more than the expected number of accidents during all training episodes. Hence, successfully handling accidents requires good generalization performance. Similar to accidents, the demand distributions match between training and test environments.

We compare our method with two baselines: first, a *selfish* routing scheme, where all cars are selfish and use the human choice dynamics presented in Sec. 5.3.1, and second, a model predictive control (MPC) based controller which can perfectly simulate the network other than the uncertainty due to accidents and noisy demand. It plans for the receding horizon of 4 minutes and re-plans after every 1 minute to minimize the number of cars in the network using a Quasi-Newton method (L-BFGS [113]). To increase robustness against the uncertainty, it samples 12 different simulations of the network and takes the average. We note that this MPC can only be useful in small networks where some cars can enter the network and reach the destination within the MPC horizon of 4 minutes. While increasing the horizon may help MPC operate in larger networks, it causes a huge computational burden. In fact, even though we parallelized the controller over 12 Intel® Xeon® Gold 6244 CPUs (3.60 GHz), it took the controller 32 seconds on average to decide the routing of autonomous vehicles for the next 1 minute, which clearly indicates a practical problem. In all experiments, we set $\eta^{\mathrm{h}}(k)$ (and $\eta^{\mathrm{a}}(k)$ for the selfish baseline) to be 0.5 for all k.

5.5.1 General Class of Networks

We first start by considering a small network of 9 cells and 7 junctions (1 regular junction, 3 merges and 3 diverges) as shown in Fig. 5.3, where the priority levels of cells at merges are equal to their numbers of lanes. We set the autonomy level of the demand $\bar{\alpha} = 0.6$ and the total demand $\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}} = 2.60$ cars per second. We set the probability

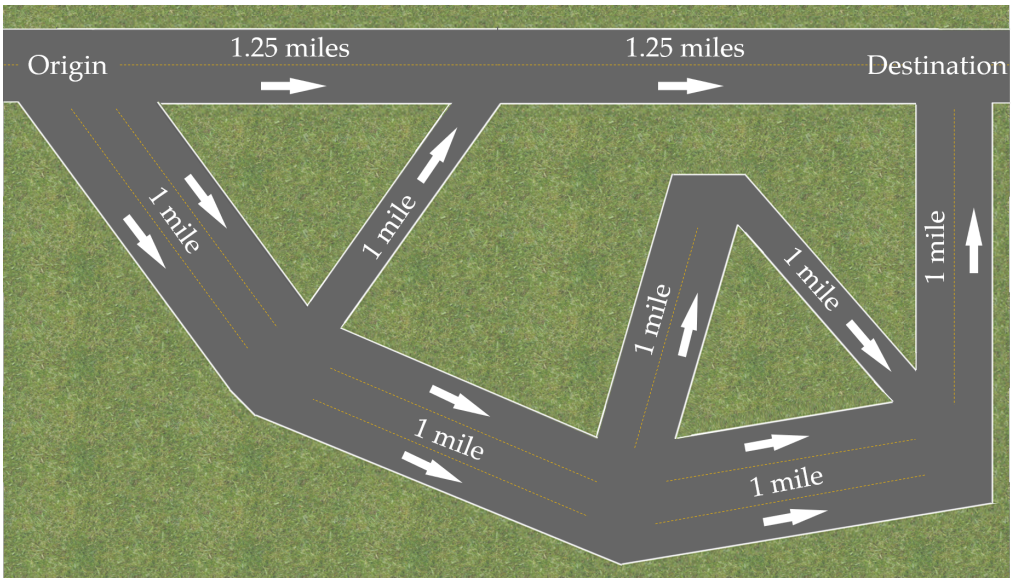


Figure 5.3: The small general class network used for experiments.

of accidents such that the expected frequency of accidents is 1 per 100 minutes, and clearing out an accident takes 30 minutes on average [147]. For human choice dynamics, we assume humans' latency estimates are based on the current states of each cell, i.e., they estimate the latencies as if the network is in steady-state.

Fig. 5.4 shows the number of cars in the network over time (mean $\pm$ standard error over 100 simulations). While MPC controller improves over the selfish routing, they both suffer from linearly growing queues. On the other hand, RL controller stabilizes the queue and keeps the network uncongested.

Next, we consider a larger network shown in Fig. 5.5 as a graph where the numbers noted on the links denote the number of cells in that link in one direction. Each cell, excluding queues which has infinite capacity, has 2 lanes. This is a quantized version of the OW network due to Ortúzar and Willumsen [2], and is widely used in the transportation literature [129, 133, 148, 149]. This is a larger network with 4 origin-destination pairs, 102 cells (and 2 queues) and 41 junctions (28 junctions with only one incoming and one outgoing cell, and 13 more general junctions). We set the total demand to be $\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}} =$

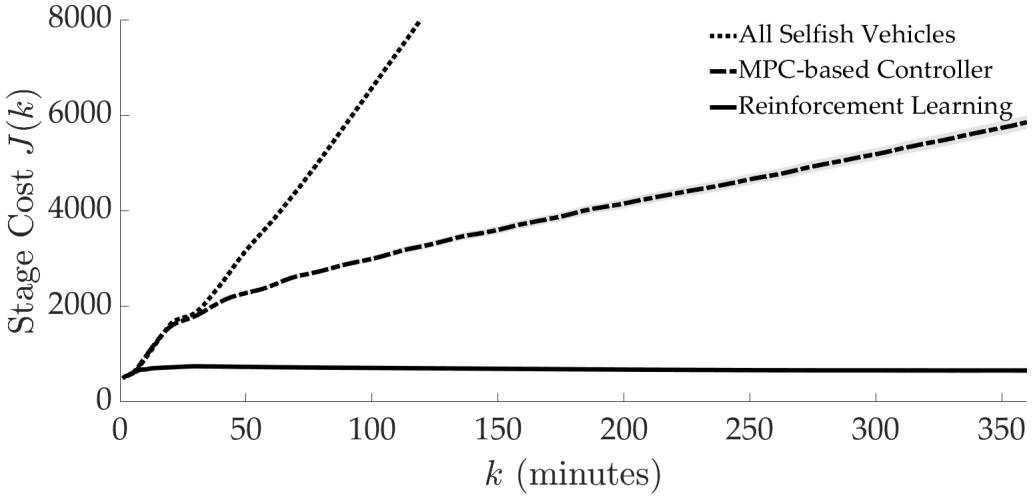


Figure 5.4: Time vs. number of cars under selfish, MPC and RL routing on the small general class network.

3.46 cars per second, distributed equally to the 4 origin-destination pairs in expectation. As there are 1752 possible different simple paths that vehicles could be taking, our action space is 1752 dimensional. While such an optimization is still possible with powerful computation resources, it might be unnecessary because an optimal solution is unlikely to utilize the paths that traverse too many cells. We therefore restrict our action space to the 10 shortest paths (with respect to the free-flow latencies) between each origin and destination, and so adopt a 40-dimensional action space. We keep the other experiment parameters the same as the small network experiment above.

Due to the network size and the computation cost to simulate the OW network, the MPC-based controller does not produce any useful results in a reasonable time as explained before. We instead implemented the greedy optimization method of [122] as a baseline. Specifically we used a genetic algorithm for the optimization with a constraint on the run time of one minute, as it is an online algorithm. It is important to note that RL policy makes a routing decision within a millisecond during test time. We compare the RL controller with this greedy method and the selfish routing.

Fig. 5.6 shows the number of cars in the network over time (mean ± standard error

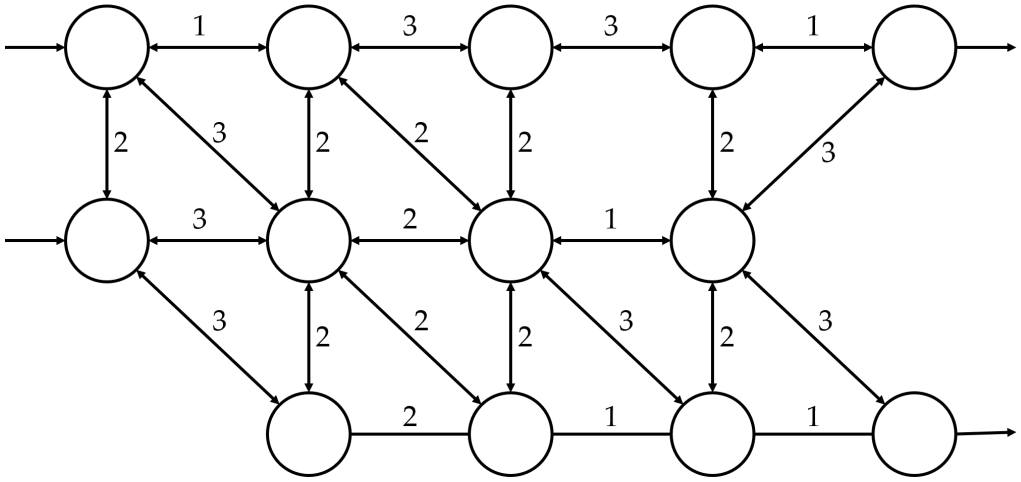


Figure 5.5: OW network (adapted from [2]) used for experiments.

over 100 simulations). Again, the selfish routing and the greedy optimization method of [122] suffer from linearly growing queues, while RL controller is able to stabilize the queues and keeps the network uncongested even though the network may start from a congested state. Furthermore, we check whether the reduced action space is really sufficient. We observe that, over 100 episodes, 98.92% of the autonomous vehicles were routed to the paths that are faster than the fastest path that is not in the action space.

To analyze the performance RL controller in comparison with the optimal equilibrium, we now move to parallel networks.

5.5.2 Parallel Networks

We consider a parallel network from downtown Los Angeles to the San Fernando Valley with 3 paths. The highway numbers and the approximated parameter tuples (length, number of lanes, speed limit) are:

1. 110N (5 miles, 3 lanes, 60 mph); 101N (10 miles, 3 lanes for 5 miles then 2 lanes, 60 mph)
2. 10E (5 miles, 4 lanes, 75 mph); 5N (10 miles, 4 lanes, 75 mph); 134W (5 miles, 3

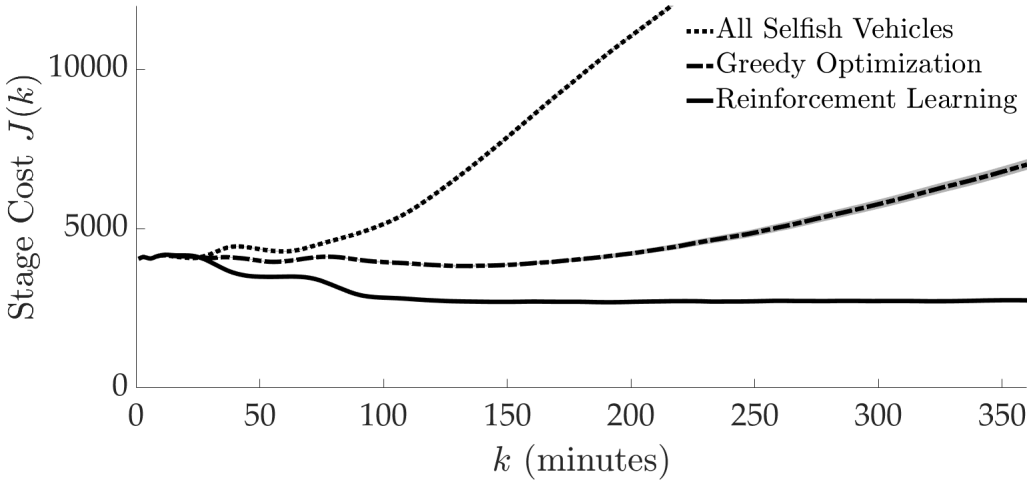


Figure 5.6: Time vs. number of cars under selfish, greedy and RL routing on OW network.

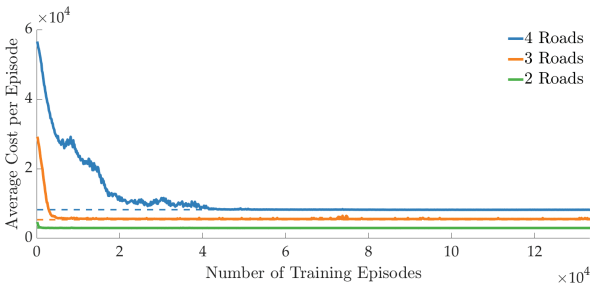
lanes, 75 mph)

3. 10W; 405N (both 10 miles, 4 lanes, 75 mph); 101S (5 miles, 3 lanes, 75 mph)

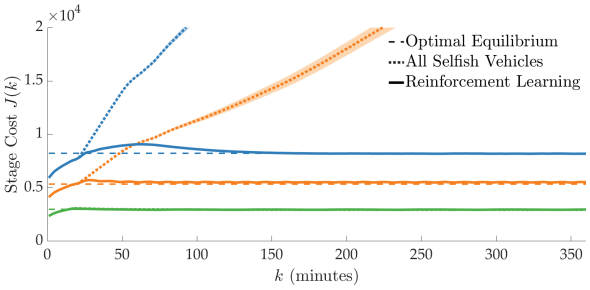
As the cells are now not shared between the paths, we employ better latency estimates for human choice dynamics: we compute them as the actual latencies that would occur if there were no accidents and no more demand into the network.

We perform 3 sets of experiments. In the first two, we disable accidents and analyze the effects of varying the number of paths and autonomy. As the shortest path has 15 cells, we exclude MPC-based controller from our analysis as it is computationally prohibitive to adopt a receding horizon longer than 15 minutes.

Varying number of paths. We first vary the number of paths $|\mathcal{P}| \in \{2, 3, 4\}$ by duplicating, or removing, the third path. We set the autonomy level of the demand $\bar{\alpha} = 0.6$, and $\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}}$ to be 95% of the maximum capacity under this autonomy level. We plot learning curves in Fig. 5.7 (a). It can be seen that even with $|\mathcal{P}| = 4$ when observation space is 144-dimensional, the agent successfully learns routing within 40 million time steps. With randomized initial states, the agents learn routing policies that achieve nearly as good as optimal equilibrium for all $|\mathcal{P}| \in \{2, 3, 4\}$. In Fig. 5.7 (b), we



(a) Average number of cars in the system per episode during RL training.



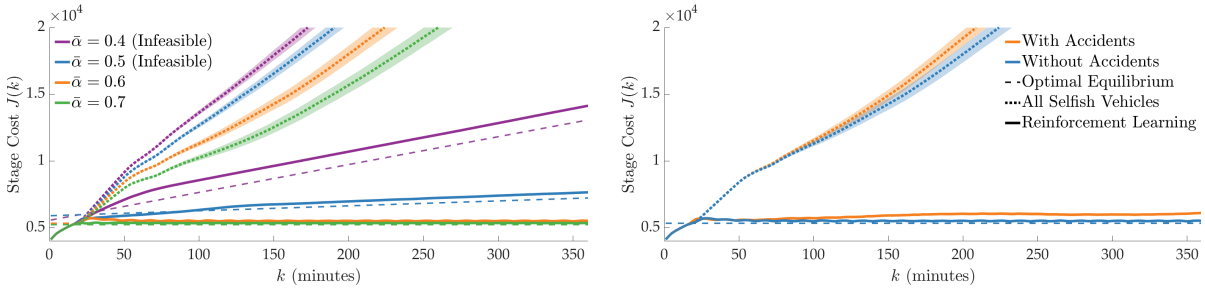
(b) Time vs. number of cars in the system for the comparison of selfish and RL routing in parallel networks.

Figure 5.7: Varying number of paths.

plot the number of cars (mean ± standard error over 100 simulations) in the system over time. While selfish routing causes congestion by creating linearly growing queues when $|\mathcal{P}| > 2$, RL policies successfully stabilize queues and even reach car numbers of optimal equilibria.

Varying autonomy. We take $|\mathcal{P}| = 3$ and vary the autonomy of demand $\bar{\alpha} \in \{0.4, 0.5, 0.6, 0.7\}$ without changing the total demand $\bar{\lambda}^{\mathrm{h}} + \bar{\lambda}^{\mathrm{a}}$. Note the demands are infeasible when $\bar{\alpha} \in \{0.4, 0.5\}$. In Fig. 5.8 (a), we plot the number of cars (mean ± standard error over 100 simulations) in the system over time. The result is similar to the previous experiment when the demand is feasible. With infeasible demand, RL agent keeps a queue that is only marginally longer than the queue that optimal equilibrium would create. On the other hand, selfish routing grows the queue with much faster rates. These experiments show RL policy successfully handles random initializations.

Accidents. In the third set, we fix $|\mathcal{P}| = 3$ and $\bar{\alpha} = 0.6$ for the same total average demand and enable accidents. As before, the expected frequency of accidents is 1 per 100 minutes, and clearing out an accident takes 30 minutes on average. Fig. 5.8 (b) shows the RL policy successfully handles accidents, indicating a good generalization performance by the RL controller. To give a clearer picture, we provide the space-time diagrams and



(a) Varying autonomy.

(b) Varying the presence of accidents and noise in the demand.

Figure 5.8: Varying autonomy and perturbations.

the detailed information about the system states of a sample run in Figs. 5.9 and 5.10, respectively. Fig. 5.9 shows that selfish routing causes congestion by not utilizing the third route, whereas RL can avoid congestion and handle accidents. Fig. 5.10 shows the number of cars in each cell as well as the queue lengths over time. The small oscillations, which occur even after the effect of the accidents disappear (between third and fourth hours), are due to noisy demand and the discretization of cells. With selfish routing, the vehicles use the longest path only when there is an accident in another path (around first and third hours) or the other two paths are congested (third and fifth hours). In contrast, RL makes good use of the network and leads to altruistic behavior. It also handles the accidents by effectively altering the routing of autonomous cars (around third hour, autonomous cars start using the first route until the accident in the third route is cleared). Hence, it manages to stabilize the queue and prevent congestion. We provide video visualizations of this run at `https://youtu.be/XwdSJuUb09o`.

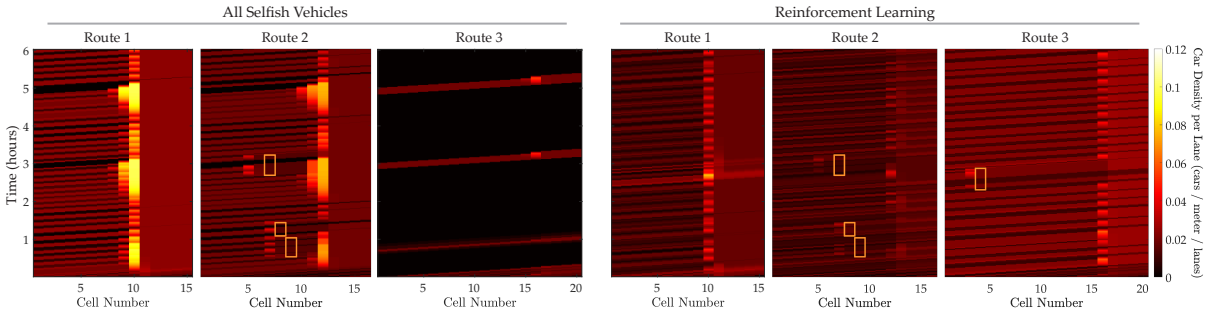


Figure 5.9: Space-time diagrams on a parallel traffic network with accidents and noisy demand. Orange rectangles represent accidents.

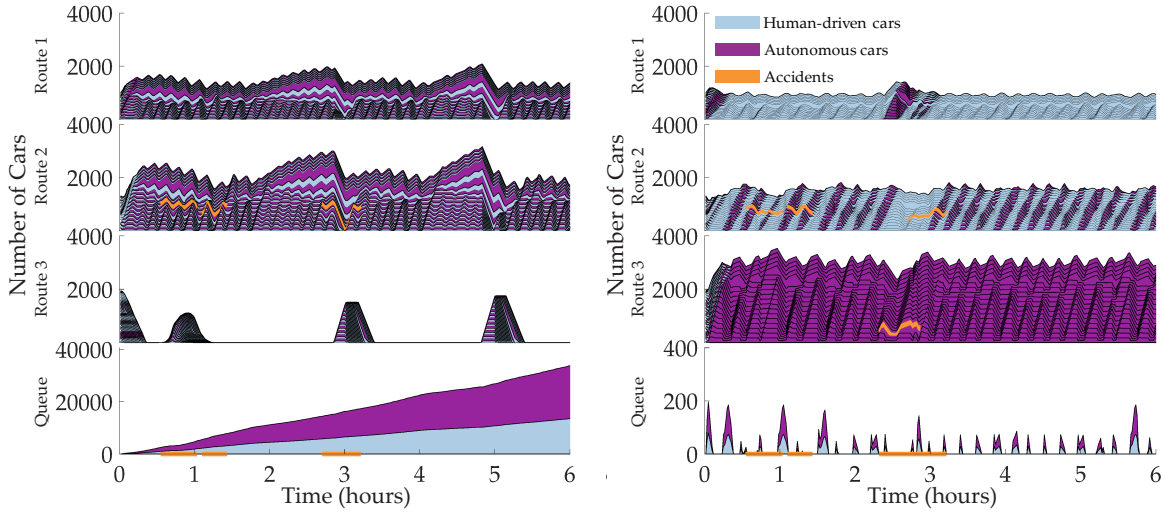
5.6 Conclusion

In this chapter[5] we presented a framework for understanding a dynamic traffic network shared between selfish human drivers and controllable autonomous cars. We showed, using deep RL, we can find a policy to minimize the average travel time experienced by users of the network. We developed theoretical results to describe and calculate the best equilibria that can exist and empirically showed that our policy reaches the best possible equilibrium performance in parallel networks. Further, we provided case studies showing how the training period scales with the number of paths, and we showed our control policy is empirically robust to accidents and stochastic demand.

Limitations. We used the number of cars in each cell as predictive features for RL training. Although this makes the state space dimensionality grow only linearly with the number of cells, it may not be scalable to much larger traffic networks. Moreover, the action space grows linearly with the number of source-destination pairs, also impacting the scalability of the algorithm.

Future work. This work opens up many future directions for research, including using multiagent reinforcement learning to model autonomous vehicles with competitive

[5]This work was supported by NSF grant 1953032 and Toyota. Toyota Research Institute (TRI) provided funds to assist the authors with their research but this article solely reflects the opinions and conclusions of its authors and not TRI or any other Toyota entity



(a) Selfish routing.

(b) RL routing.

Figure 5.10: The network under perturbations due to accidents and noisy demand. For each path and time step, from bottom to top, the stacked color segments show the number of cars in the cells from origin to the destination. Congestion occurs only upstream to the bottlenecks.

goals and/or en route decision making ability and improving how the training time scales with the complexity of the network. Another interesting future work is to investigate how an RL policy can be deployed and the simulation imperfections (including the dependency on the simulated human choice dynamics) can be alleviated by collecting online data using sensors from the real traffic network.

5.7 Appendix

5.7.1 Summary of notation

See Table 5.1.

Table 5.1: Summary of Notation

p	Path index	unitless
$\mathcal{P}$	Set of paths in the network	set of paths
i	Cell index	unitless
$\mathcal{I}$	set of cells in the network	set of cells
$\mathcal{I}_p$	set of cells in path p	set of cells
$\mathcal{U}_i$	set of cells upstream of cell i	set of cells
$\bar{v}_i$	Free-flow velocity of cell i	cells/time step
b_i	Number of lanes of cell i	unitless
h_i^{h} (h_i^{a})	Nominal vehicle headway on cell i	cells/vehicle
ϕ_i^{h} (ϕ_i^{a})	Density of vehicles on cell i	vehicles/cell
ϕ_i	Total vehicle density on cell i	vehicles/cell
f_i^{h} (f_i^{a})	Flow of vehicles from cell i	vehicles/time step
y_i^{h} (y_i^{a})	Hum. (aut.) veh flow into cell i	vehicles/time step
α_i	Autonomy level of cell i	unitless
$\tilde{n}_i(\alpha)$	Critical density of cell i, at aut. α	vehicles/cell
$\bar{\phi}_i$	Jam (maximum) density of cell i	vehicles/cell
$\bar{F}_i(\alpha)$	Capacity of cell i, at aut. α	vehicles/time step
$w_i(\alpha)$	Shockwave speed of cell i, at aut. α	cells/time step
k	Time index	unitless
$\ell_p(k)$	Latency of path p if starting at time k	time steps
$q_i(k)$	Priority for cell i at a merge at time k	unitless
$\mu_i^{\mathrm{h}}(p,k)$ (μ^{a})	Frac. of hum. (aut.) vehs in i on p at k	unitless
$\beta_i^{\mathrm{h}}(i',k)$ (β^{a})	Frac. of hum. (aut.) vehs $i \to i'$ at k	unitless
$J(k)$	Stage cost at time k	vehicles
m_p^{b} (m_p^{n})	# of (non)bottleneck cells on path p	cells
b^{b} (b^{n})	# of lanes in (non)bottleneck cells on p	unitless
r_p	$:= b^{\mathrm{b}}/b^{\mathrm{n}}$	unitless
γ_p	Number of congested cells on path p	cells

5.7.2 Proofs for Section 5.4.3

Proof of Lemma 5.1. By definition, at equilibrium, the number of vehicles in each cell i in $\mathcal{I}_p$, $\phi_i^{\mathrm{a}}(k)$ and $\phi_i^{\mathrm{a}}(k)$ is constant for all times k. Since by definition the incoming flow is also constant, by the definition of the sending and receiving functions, constant cell densities implies constant flows. By (5.3), a constant density also implies that the incoming and outgoing flow in each cell are equal. This means that all cells will have the same incoming flow as the first cell. Further, we know that since the density of

autonomous vehicles is constant over time, incoming and outgoing autonomy levels are equal. Accordingly, if cell i' is the cell immediately upstream of i, then $\alpha_{i'}(k)f_{i'}(k) = \alpha_i(k)f_i(k)$. Since we also have $f_{i'}(k) = f_i(k)$, this implies that $\alpha_{i'}(k) = \alpha_i(k)$. Therefore the autonomy level of all cells is the same. Let us denote this uniform autonomy level α_p. Let the index of the first cell in the path be 0. Then, $\bar{\lambda}_p^{\mathrm{h}} + \bar{\lambda}_p^{\mathrm{a}} = f_0$ and $\bar{\lambda}_p^{\mathrm{a}} = \alpha_p f_0$. Combining these two expressions, we find $\alpha_p = \bar{\lambda}_p^{\mathrm{a}}/(\bar{\lambda}_p^{\mathrm{h}} + \bar{\lambda}_p^{\mathrm{a}})$. □

Proof of Lemma 5.4. Under Assumption 5.2, no two paths have the same free-flow latency. With Proposition 5.1, this implies that if an equilibrium has a used path with no congestion, it must be the used path with greatest free-flow latency, as otherwise all used paths would not have the same latency. Therefore, if an equilibrium routing with positive flow on paths $[p]$ has a path in free-flow, it must be path p. Otherwise, we can construct an equilibrium with the same demand that has path p in free-flow. Recall that the latency on paths in equilibrium is increasing with the length of the congested portion of the path, $\gamma_{p'}$, and $\gamma_{p'} = 0$ corresponds to an uncongested path. If all paths are congested, we consider decreasing the length of congestion on all paths simulatenously, at rates which keep the path latencies equal. This continues until path p becomes completely uncongested. This construction proves the lemma. □

5.7.3 Overview of Proximal Policy Optimization (PPO)

In this section, we give a brief overview of the PPO method [120] we used for training our deep reinforcement learning model. We first start with formalizing the problem. We then introduce the policy gradients and the details of PPO. To keep the notation consistent with the reinforcement learning literature, we abuse the notation for some variables. Hence, this section of the appendix is written in a standalone way, and the variables should not be confused with the notation introduced in the main paper (e.g. f

is going to denote the transition distribution of the system as introduced below, instead of flow values as in the main paper).

Problem Setting. We consider a sequential decision making problem in a Markov decision process (MDP) represented by a tuple $(\mathcal{S}, \mathcal{A}, f, T, r, \gamma)$, where $\mathcal{S}$ is the set of states. $\mathcal{A}$ denotes the set of actions, and the system transitions with respect to the transition distribution $f : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$. For example, if $f(s, a, s') = p$, this means taking action $a \in \mathcal{A}$ at state $s \in \mathcal{S}$ transitions the system into state s' with probability p. Next, T denotes the horizon of the system, i.e., the process gets completed after T time steps. The reward function $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ maps state-actions to reward values. The decision maker is then trying to maximize the cumulative reward over T time steps by only observing the observations (not states). Finally γ is a discount factor that sets how much priority we give to optimizing earlier rewards in the system.

Let us now describe how we formulate a transportation network with the CTM model as an MDP in this paper. The state of the network is fully defined by the following information:

- Location of each vehicle (which cell or queue it is in),
- Type of each vehicle (human-driven or autonomous),
- Accident information (where and when it happened), and
- Planned path of each vehicle (which cells it is going to traverse).

In our model, we assumed the first three items in the above list are available as observations. While this breaks the Markov assumption, deep RL techniques often perform well in partially observable MDPs, too. So our deep RL policy is trying to make its decisions based only on those first three observations, and the non-observability of the planned paths increases the stochasticity of the problem. The action set of the decision maker is defined by the set of available routing paths of autonomous vehicles. The transition distribution follows the dynamics of CTM, human choice dynamics, as well as the acci-

dents which also introduce stochasticity into the system. Finally, as a reward function, one can think of using the negative of the number of cars in the system as a proxy to negative of overall latency in the network.

Policy Gradients. To solve this problem using deep neural networks, we model the decision-maker agent with a stochastic policy π_θ parameterized with θ (e.g. weights of the neural network), such that $\pi_\theta(a \mid s)$ gives the probability of taking action a when observing state s. The goal of the agent is to maximize the expected cumulative discounted reward:

$$J(\theta) = \mathbb{E}_{\tau\sim\pi_\theta}\left[\sum_{t=0}^{T-1}\gamma^t r(s_t, a_t)\right]$$

where τ denotes a trajectory $(s_0, a_0, \ldots, s_{T-1}, a_{T-1}, s_T)$ in the system. The discount factor is to improve robustness and to reduce susceptibility against high variance. We can equivalently write this objective as:

$$J(\theta) = \int_\Xi \pi_\theta(\tau) r(\tau) d\tau$$

where Ξ is the set of all possible trajectories, $\pi_\theta(\tau)$ is the probability of trajectory τ under policy π_θ, and $r(\tau)$ is the cumulative discounted reward of trajectory τ. The idea in policy gradients is to take gradient steps to maximize this quantity by optimizing θ:

$$\begin{aligned}
\nabla_\theta J(\theta) &= \nabla_\theta \int_\Xi \pi_\theta(\tau) r(\tau) d\tau \\
&= \int_\Xi \nabla_\theta \pi_\theta(\tau) \frac{\pi_\theta(\tau)}{\pi_\theta(\tau)} r(\tau) d\tau \\
&= \int_\Xi \pi_\theta(\tau) r(\tau) \nabla_\theta \log \pi_\theta(\tau) d\tau \\
&= \mathbb{E}_{\tau\sim\pi_\theta}\left[r(\tau)\nabla_\theta \log \pi_\theta(\tau)\right]
\end{aligned}$$

which we can efficiently approximate by sampling trajectories using the policy.

Unfortunately, this vanilla policy gradient method is not robust against variance (due to stochasticity in the environment and trajectory sampling) and suffers from data-inefficiency. In recent years, several works have developed alternative ways to approximate the gradients. One such idea is based on using baselines to reduce variance:

$$\nabla_\theta J(\theta) = \mathbb{E}_{\tau\sim\pi_\theta}\left[\sum_{t=0}^{T-1} \nabla_\theta \log \pi_\theta(a_t^\tau \mid s_t^\tau)\hat{A}_t^\tau\right]$$

where $\hat{A}$ is called the estimated advantage function, which is usually defined as $G_t^\tau - V(s_t^\tau)$, where G_t^τ is the cumulative discounted reward of the trajectory τ after (and including) time step t, and $V(s_t^\tau)$ is some baseline that quantifies the value of state s_t^τ. This new equation for $\nabla_\theta J(\theta)$ holds due to the Markov assumption and that the baseline is independent from the policy parameter θ.

Having presented the policy gradients and the use of baselines for variance reduction, we are now ready to give an overview of PPO.

Proximal Policy Optimization (PPO). PPO further improves the robustness and data-efficiency of policy gradient methods by using a surrogate objective that prevents the policy from being updated with large deviations. Instead of the usual objective $\mathbb{E}_{\tau\sim\pi_\theta}\left[\log \pi_\theta(a_t^\tau \mid s_t^\tau)\hat{A}_t^\tau\right]$, PPO uses the following objective:

$$J_1(\theta) = \mathbb{E}_{\tau\sim\pi_\theta}\left[\min(g_t^\tau(\theta)\hat{A}_t^\tau, \text{clip}(g_t^\tau(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t^\tau)\right]$$

where

$$g_t^\tau(\theta) = \frac{\pi_\theta(a_t^\tau \mid s_t^\tau)}{\pi_{\theta_{\text{old}}}(a_t^\tau \mid s_t^\tau)} \text{ and } \text{clip}(x, \epsilon_1, \epsilon_2) = \begin{cases} \epsilon_1 & x < \epsilon_1, \\ x & \epsilon_1 \le x \le \epsilon_2, \\ \epsilon_2 & \text{otherwise.} \end{cases}$$

In addition to $J_1(\theta)$, PPO uses two more objective functions and converts the prob-

lem into a multi-objective optimization problem. The first additional objective is for the baseline $V(s_t^\tau)$. Specifically, PPO learns a parameterized value function V_ϕ in a supervised way to minimize $(V_\phi(s_t^\tau) - V_t^{\text{target}})^2$ where V_t^{target} is calculated using the sampled trajectories as a sum of discounted rewards after (and including) time step t. It should be noted that this does not make $G_t^\tau - V_\phi(s_t^\tau) = 0$, because $V_\phi(s_t^\tau)$ is an estimate of the true value function and is updated after the computation of the estimated advantage. Therefore,

$$J_2(\phi) = -\mathbb{E}_{\tau\sim\pi_\theta}\left[V_\phi(s_t^\tau) - V_t^{\text{target}}\right] .$$

Finally, PPO uses an entropy bonus (inspired by [145]) to ensure sufficient exploration:

$$J_3(\theta) = \mathbb{E}_{\tau\sim\pi_\theta} H(\pi_\theta(\cdot \mid s_t^\tau)) ,$$

where H is information entropy. At the end, PPO tries to solve:

$$\text{maximize}_{\theta,\phi} \quad J_1(\theta) + J_2(\phi) + cJ_3(\theta)$$

where c is the coefficient for the entropy term.

5.7.4 Experiment details

In implementation, we used $J(k) - J(k-1)$ as a proxy cost for time step k, where $J(0) = 0$.

Below are the set of hyperparameters we used for PPO. We refer to Section 5.7.3 and [120] for the definitions of PPO-specific parameters. While this set yields good results as we presented in the paper, a careful tuning may improve the performance.

- Number of Time Steps: 40 million
- Number of Actors: 32 (32 CPUs in parallel)

- Time Steps per Episode During Training: 300
- Time Steps per Actor Batch: 1200
- ϵ for Clipping in the Surrogate Objective: 0.2
- Optimization Step Size (OSS): 3×10^{-4}
- Annealing for ϵ (Clipping) and OSS: Linear (down to 0)
- Entropy Coefficient: 0.005
- Number of Optimization Epochs: 5
- Optimization Batch Size: 64
- γ for Advantage Estimation: 0.99
- λ for Advantage Estimation: 0.95
- ϵ for Adam Optimization: 10^{-5}

Finally, we report the training times (for 40 million time steps) and the number of time steps of empirical convergence (in terms of reward value) for each RL policy in Table 5.2. In test time, RL policies produce an action in under 1 ms.

Table 5.2: Training and Convergence Times

Policy	**Training Time**	**Time Step of Convergence**
Simple General Network	10.0 hours	26.3 million
OW Network	253.1 hours	31.0 million
$\lvert\mathcal{P}\rvert = 2$	22.2 hours	0.7 million
$\lvert\mathcal{P}\rvert = 3$	38.9 hours	10.0 million
$\lvert\mathcal{P}\rvert = 3$, w/ accidents	40.5 hours	22.0 million
$\lvert\mathcal{P}\rvert = 3$, $\bar{\alpha} = 0.4$	50.6 hours	25.5 million
$\lvert\mathcal{P}\rvert = 3$, $\bar{\alpha} = 0.5$	43.1 hours	19.3 million
$\lvert\mathcal{P}\rvert = 3$, $\bar{\alpha} = 0.7$	38.6 hours	6.6 million
$\lvert\mathcal{P}\rvert = 4$	101.4 hours	23.3 million

Chapter 6

Platoon Formation by Influencing Individual Driving Actions

Previous chapters have built on a capacity model for mixed autonomy which assumes that autonomous vehicles can keep a short headway (*i.e.* platoon) behind any vehicle, whether human-driven or autonomous. In this chapter[1] we relax that assumption and provide a method by which autonomous vehicles can leverage their influence over human drivers so as to reorder themselves to form platoons, thereby reaching, or approaching, the capacity in the aformentioned capacity function.

6.1 Introduction

In recent years we have seen many advances in the field of autonomous driving. The emergence of autonomous cars on public roads has become a reality. However, it is becoming very clear that autonomous cars cannot immediately avoid interactions with

[1]This chapter is adapted, with permission, from "Maximizing road capacity using cars that influence people", previously published in IEEE Conference on Decision and Control, © 2018 IEEE, and is joint work with Kabir Chandrasekher, Ramtin Pedarsani, and Dorsa Sadigh [150].

humans [151]. Despite incorporating the highest levels of autonomy, today's self-driving cars have to drive on roads shared by human-driven vehicles, bicyclists, or even pedestrians. This emphasizes the urgency and importance of understanding the interactions between humans and autonomy on shared roads.

The work in this area has mainly diverged into two distinct approaches:

1. Local and distributed techniques to enable safe and effective interaction between the vehicles [152–159]; and
2. Network-level control that focuses on studying global properties such as optimal routing in the network or optimal throughput [19, 31, 42, 63, 67].

While these techniques are quite valuable at solely studying local or network level properties, it turns out these two paradigms can actively influence each other. For instance as shown in Fig. 6.1, if all autonomous cars (purple cars) on a road take interactive local actions that enforce human-driven cars (blue cars) to open up some space, the autonomous cars can then leverage the opportunity to form a platoon (Fig. 6.1 (b)).

The mobility benefits of platooning autonomous or semi-autonomous cars have been largely studied in the literature for both freeway networks [30, 59, 160–162] and urban networks [31, 46]. Most of these works have assumed that all vehicles are equipped with some level of autonomy. Another line of work has focused on designing efficient scheduling policies at intersections by leveraging the autonomy of vehicles and their communication capabilities [163–168]. Our approach is significantly different from the existing work in the literature.

> *Our key insight is that we need to intelligently design local level interactions based on the knowledge of the desirable high level objectives for the network of autonomous and human-driven cars.*

Therefore, we attempt to attack the problem using a unified approach. We do so by

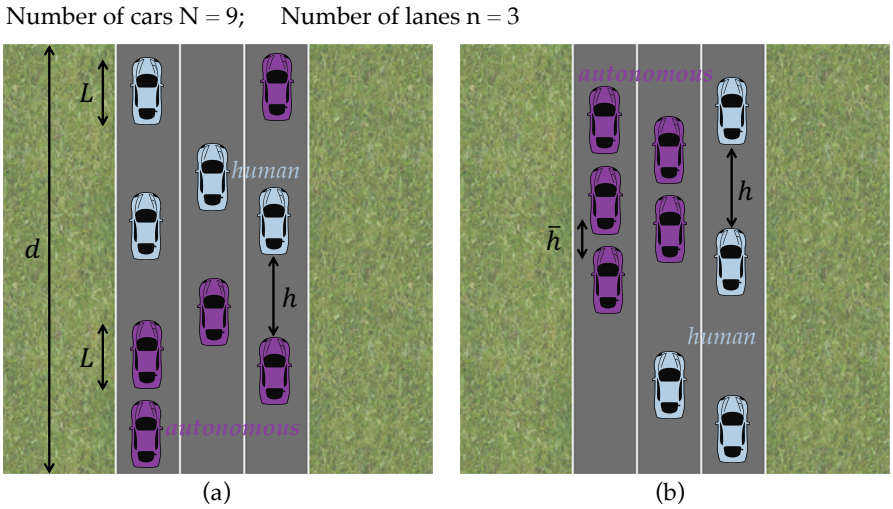


Figure 6.1: Road shared by autonomous (purple) and human-driven (blue) cars.

extending the mathematical framework outlined in Chapter 2 to include the notion of optimal lane allocation and ordering within it. We then provide theoretical results on the structure of the optimal solution as well as the price of having no control. This optimal allocation and ordering can practically be achieved using state-of-the-art interaction-aware controllers [153]. Our contributions in this chapter are summarized as follows:

- Leveraging local interactions between autonomous and human driven cars to modify the road configuration in open loop to an optimal configuration.
- Precisely characterizing the optimal solution of the network allocation problem and characterizing the price of no control.
- Implementing the controller for a mixed-autonomy setting via extensive simulation showing tight agreement with our theoretical results.

6.2 Running Example

In this section, we introduce a running example that describes an instance of the problem of our interest. Imagine a road shown in Fig. 6.1 shared by autonomous and

human driven cars. We let the purple cars represent autonomous cars, and the blue cars represent human-driven cars.

The vehicles, autonomous or human-driven, probabilistically arrive at any of the three lanes 1, 2, 3 (where lane 1 is the leftmost lane). As we will discuss throughout the paper, the vehicles keep a headway between each other, and this headway is much smaller if two autonomous cars are following each other, i.e., $\bar{h} < h$. This reduced headway also corresponds to the term *platooning*, which intuitively means that autonomous cars, unlike human-driven cars, are capable of coordinating with each other and driving close to each other to save space, energy, and time.

Our goal in this work is to start with a road configuration as shown in Fig. 6.1 (a), and plan for the autonomous cars to navigate intelligently for the goal of converging to a configuration similar to Fig. 6.1 (b).

We theoretically construct the most efficient road configuration based on an optimal vehicle configuration problem and study its properties. In addition, to achieve this optimal configuration, we leverage local interaction-aware controllers for autonomous cars. As shown in previous work [153], the autonomous vehicle can indirectly *affect* a human driven car to take desirable actions that result in the ideal road configuration. In constructing these configurations, we assume that the heavy, local optimizations are solved in a distributed fashion (separately by each vehicle), and that a central controller exists which functions only to send and receive simple messages to coordinate.

6.3 Model

Our goal in this paper is to leverage the low level interactions between autonomous and human-driven cars to *indirectly* order the vehicles in a fashion that is desirable from the road network's perspective. To achieve this goal, we plan to answer two questions:

1. Assuming autonomous and human-driven vehicles can reorder themselves from any configuration to another (for example from Fig. 6.1 (a) to Fig. 6.1 (b)), what would be the optimal vehicle ordering for the road?
2. How can we leverage the capability of autonomous cars to enforce this optimal ordering?

To study the optimal ordering algorithm, we first discuss road *capacity* as a desirable property for the road network. Road capacity has been used as a common measure that needs to be optimized for a desirable traffic network [32, 42, 43, 46][2].

To model road capacity, we consider how many vehicles can be packed onto a road at nominal speed. To this end, each car has a certain headway in front of it depending on the type of the car, as well as the type of the car it follows. In this model motivated by previous work in capacity modeling [32, 42, 43, 46], we assume that human-driven cars follow all vehicles at a distance of h and autonomous cars follow other autonomous cars by a distance of $\bar{h}$ (typically $\bar{h} < h$) and follow human-driven cars with a distance of h (as shown in Fig. 6.1 (a)). Note that these quantities will vary by road as they depend on the nominal speed on that road.

We consider the capacity on a multi-lane road to be the sum of capacities of each lane. To find lane capacity, we first find the average headway, which will be a function of the *level of autonomy* of the lane, defined as $\alpha_i = \frac{y_i}{x_i+y_i} \in [0, 1]$, where x_i and y_i are respectively the volume of human-driven and autonomous vehicles in lane i. For instance, the level of autonomy for the leftmost lane in Fig. 6.1 (a) is $\alpha_1 = \frac{2}{4}$.

Let h_j be the headway in front of vehicle j, and let H_N be the total expected headway for N vehicles on road i. Then, we see that $H_N = \sum_{j=1}^{N-1} \mathbb{E}[h_j] = (N-1)(\alpha_i^2\bar{h}+(1-\alpha_i^2)h)$.

If each vehicle has length L, then as the number of vehicles N increases, the average

[2]Though some works show that increasing road capacity can lead to worse traffic equilibria [72,169], a system designer can selectively use this capability and choose which roads on which to increase capacity to avoid these negative effects and improve equilibrium performance.

space taken by a vehicle approaches $\alpha_i^2\bar{h} + (1-\alpha_i^2)h + L$. The capacity is the number of vehicles that can be fit on a lane of length d, so

$$c(\alpha_i) = \frac{d}{\alpha_i^2\bar{h} + (1-\alpha_i^2)h + L} = \frac{d}{k_1 - \alpha_i^2 k_2}\ , \tag{6.1}$$

where we define two constants $k_1 \triangleq L + h$ and $k_2 \triangleq h - \bar{h}$, and $k_1 > k_2$. Note that throughout the paper we consider same network parameters, $d, h, \bar{h}, L$ for all the lanes.

Upper Bound on Capacity. If we do not assume that the vehicles will arrive from some Bernoulli process, and allow some arbitrary process to determine the ordering of the vehicles, the capacity model described above will no longer be accurate. Though the specific form of the capacity function will depend on the ordering, we define a capacity model that serves as an upper bound on the capacity for any ordering. This upper bound corresponds to the situation in which all autonomous cars are ordered optimally, meaning they are all adjacent and can form one long platoon.

In this case, each autonomous car has a short headway $\bar{h}$ and each human-driven vehicle has headway h, yielding average headway $\alpha_i\bar{h} + (1-\alpha_i)h$. This results in the following capacity function:

$$c^{UB}(\alpha_i) = \frac{d}{\alpha_i\bar{h} + (1-\alpha_i)\bar{h} + L} = \frac{d}{k_1 - \alpha_i k_2}\ . \tag{6.2}$$

Since for all i lanes $\alpha_i \in [0,1]$, it is clear that $c^{UB}(\alpha_i) \geq c(\alpha_i)$. Further, c^{UB} serves as a tight upper bound with regards to vehicle ordering with a large number of vehicles.

6.3.1 Optimal Lane Assignment

We now pose optimization problems for assigning vehicles lanes in which to travel when operating at capacity, under the capacity models in (6.1) and (6.2). We also provide theoretical results pertaining to the optimal lane assignment and attendant total road capacities.

Optimal Lane Assignment with Bernoulli Arrivals. Assuming that the autonomy level for each lane α_i is given by a system operator, we now describe the high-level optimal lane assignment problem that maximizes the total capacity of the road. Here we consider a multi-lane road to have a total capacity that is the sum of the capacity of each lane. All lane parameters will be the same, and the only difference between capacities in the lanes is due to different lane levels of autonomy α_i.

Let $\bar{\alpha}$ denote the overall autonomy level on the road that is the fraction of autonomous cars to all cars in the network. Further, let $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1, \ldots, \alpha_n \end{bmatrix}^T$. Then, we define

$$C(\boldsymbol{\alpha}) \triangleq \sum_{i=1}^{n} c(\alpha_i) \ , \tag{6.3}$$

$$G(\boldsymbol{\alpha}) \triangleq \sum_{i=1}^{n} (\alpha_i - \bar{\alpha}) c(\alpha_i) \ . \tag{6.4}$$

We consider the capacity of a road to be the sum of the lane capacities, so the social planner's optimization problem is as follows:

$$\boldsymbol{\alpha}^* = \arg\max_{\boldsymbol{\alpha}} C(\boldsymbol{\alpha}) \tag{6.5}$$

$$\text{s.t. } G(\boldsymbol{\alpha}) = 0 \ , \tag{6.6}$$

where $\alpha_i \in [0, 1]$ for all lanes and (6.6) constrains the solution to have overall autonomy level equal to $\bar{\alpha}$, the autonomy level of the traffic feeding the road. To see this, observe

that $\sum_{i=1}^{n} \alpha_i c(\alpha_i)$ is the sum of autonomous cars on all lanes and has to be equal to $\bar{\alpha} \sum_{i=1}^{n} c(\alpha_i)$, which implies (6.6). Moreover, note that all the lanes have the same length, i.e. $d_i = d$; thus, we implicitly assume that lane utilization is at capacity.

We now present our optimal lane assignment theorem:

Theorem 6.1. *Consider an optimization problem of the form* (6.5)*. Any solution will have at most one lane with mixed autonomy,* i.e. *at most one lane with* $\alpha_i \in (0, 1)$*.*

See Appendix 6.7.1 for the proof of this theorem.

Remark 6.1. *This theorem formalizes the intuition that mixing autonomous and regular cars only decreases capacity, as it lessens the likelihood of adjacent autonomous vehicles, the condition necessary for platooning.*

Remark 6.2. *Note that a corollary of the proof of this theorem is that* $G(\boldsymbol{\alpha})$ *is increasing with an increase of any element of* $\boldsymbol{\alpha}$*. This quality of the constraint function will be used in proving later theorems.*

Now that we have information that narrows down the set of possible solutions, we can derive a closed-form expression for optimal lane assignment, as expressed in the following theorem.

Theorem 6.2. *Let* $\boldsymbol{\alpha}^* = \left[\alpha_1^*, \ \ldots, \ \alpha_n^*\right]^T$ *be an optimum of* (6.5)*, with autonomy levels ordered in decreasing order. Let* m *denote the last lane with full autonomy,* i.e. $m = \max i$ *s.t.* $\alpha_i^* = 1$*, where* $m = 0$ *implies that there are no lanes with full autonomy. Then,*

$$m = \lfloor \frac{\bar{\alpha} n (k_1 - k_2)}{k_1 - \bar{\alpha} k_2} \rfloor \ .$$

See Appendix 6.7.2 for a sketch of the proof of this theorem[3].

[3] A full proof can be found in the extended version [170]

Remark 6.3. *One can also derive a closed-form expression for α^*_{m+1} by solving a quadratic equation, but for the sake of brevity we omit this result.*

After characterizing the optimal solution to (6.5), we now describe the lane assignment that minimizes total capacity. This will be used to characterize the cost of declining to optimally assign vehicles to lanes, which is developed in Section 6.3.3.

Proposition 6.1. *Let $\boldsymbol{\alpha}_*$ denote the worst-case lane assignment, corresponding to the minimum sum of capacities:*

$$\boldsymbol{\alpha}_* = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}}\, C(\boldsymbol{\alpha}) \ s.t.\ (6.6)\ .$$

Then, $\boldsymbol{\alpha}_* = \mathbf{1}_n^T \bar{\alpha}$ and $C(\boldsymbol{\alpha}_*) = nc(\bar{\alpha})$, where $\mathbf{1}_n$ is the row vector of ones of length n.

See Appendix 6.7.3 for a sketch of the proof. Intuitively, the worst-case lane assignment is when all lanes have autonomy level equal to the overall autonomy level. This means that any perturbation from uniform autonomy level allows benefits from platooning.

Upper Bound on Capacity Under Optimal Lane Assignment. Using the upper bound capacity in (6.2), we now formulate an optimization problem to find the theoretical maximum capacity on a road with n lanes, with autonomy level $\bar{\alpha}$.

For reasons of readability, we use $\boldsymbol{\beta}$ when discussing autonomy levels that are solutions to the upper bound capacity maximization. As before, let $C^{UB}(\boldsymbol{\beta}) \triangleq \sum_{i=1}^n c^{UB}(\beta_i)$ and $G^{UB}(\boldsymbol{\beta}) \triangleq \sum_{i=1}^n (\beta_i - \bar{\alpha}) c^{UB}(\beta_i)$. The optimization problem is then:

$$\max_{\boldsymbol{\beta}} C^{UB}(\boldsymbol{\beta}) \tag{6.7}$$

$$\text{s.t.} \quad G^{UB}(\boldsymbol{\beta}) = 0 \ , \tag{6.8}$$

$$\beta_i \in [0,1] \ \forall \ i$$

We next present a lemma showing that the total capacity when using the upper bound on capacity is invariant to lane assignment.

Proposition 6.2. *Any feasible solution to* (6.7) *has cost* $nc^{UB}(\bar{\alpha})$.

This proposition follows from the ordering-invariant nature of the capacity upper bound. One can also construct a proof similar to that of Proposition 6.1.

Remark 6.4. *Given that all feasible lane assignments have the same total capacity, there are two notable lane assignment to consider. One such lane assignment is the uniform lane assignment,* $\boldsymbol{\beta}^* = \left[\bar{\alpha}, \ \ldots, \ \bar{\alpha}\right]^T$. *Another such lane assignment is the one in which there is a maximum of one mixed lane, such as the one developed in Theorem 6.2 for the original capacity function. In this case, there will be same number of purely autonomous lanes, which is apparent from the proof of the theorem. However, the level of autonomy in the mixed lane will differ from that in the solution to* (6.5). *Again, for brevity we do not state the closed-form solution for this autonomy level, but it can be found by manipulating the constraint* (6.8).

6.3.2 Local Interaction on Shared Roads

Now that we have discussed the optimal lane assignment problem, we would like to address the second question of our interest, i.e., *how can we leverage the capability of autonomous cars to enforce this optimal lane assignment?*

The solution of the optimal lane assignment problem provides an optimal level of autonomy $\boldsymbol{\alpha}^*$ based on the model of road capacity as we have discussed so far. Since the cars are driving on shared roads, we would like to leverage the power of autonomous cars to enforce this level of autonomy by allowing them to navigate intelligently on the road.

The autonomous cars in each lane can take actions that *affect* the behavior of the human-driven cars. For instance, it can result in the human changing lanes, slowing down, or speeding up. These local actions can in fact create a reordering of the vehicles. Our goal in this paper is to intelligently create such reorderings in order to get closer to the optimal autonomy level $\boldsymbol{\alpha}^*$. In this section, we discuss how such local reorderings can be initiated.

We model the local interactions between one autonomous car and one human-driven car as a dynamical system: $x^{t+1} = f(x^t, u_A^t, u_H^t)$, where the states of the world x^t evolve based on the actions of the human-driven car u_H^t and the autonomous car u_R^t. Our goal in these local interactions is to design a controller for the autonomous car, i.e., $u_A^*(u_H^*)$, that not only achieves reaching its destination but also can affect the actions of the human-driven car. In other words, it can target for specific reactions from the human-driven car. We follow the work of Sadigh et al. [153] to design such controllers by formulating the problem as an optimization:

$$u_A^* = \arg\max_{u_A} R_A(x, u_A, u_H^*(x, u_A)) \ . \tag{6.9}$$

Here R_A denotes the autonomous car's reward function, which if optimized outputs a sequence of actions for the autonomous car u_A^* that avoids collisions and achieves the goals of the car safely and efficiently. In addition, it can also implicitly *manipulate* the actions of the human-driven car.

The optimization in 6.9 depends on a model of the human driving behavior that

outputs u_H^*. In this work, we model the humans as agents who optimize their own reward function:

$$u_H^* = \arg\max_{u_H} R_H(x, u_A^*, u_H) \ . \tag{6.10}$$

Similarly, this reward function encodes the goals and objectives of the human-driven car such as collision avoidance or keeping velocity and heading. The reward R_H is usually learned through learning from demonstration techniques such as inverse reinforcement learning (IRL) [107, 171, 172], where trajectories of human driving gets collected in an offline setting and R_H is then computed based on the collected trajectories. In our work, we use the reward function trained and learned in [153].

Through this interaction between an autonomous car and a human-driven car 6.9, we can intelligently design reward functions for the autonomous car R_A, which then lead to actions from the autonomous car u_A^* that help him navigate for better efficiency of the road network and affect the actions of the human-driven car u_H^*.

6.3.3 Cost of lack of planning

We have now discussed how one can solve the optimal lane assignment problem, and how using local interactions can affect the actions of the human-driven car. Before presenting our simulation results, we first would like to present theoretical limits on how much performance can be improved by optimally assigning lanes to cars and optimally controlling them through local interactions. This will allow us to gauge the efficacy of our control scheme, as well as judge when it is worthwhile to attempt more difficult vehicle manipulation by examining the magnitude of potential gains that can be achieved by these more complicated maneuvers. To this end, we introduce the following two quantities:

Definition 6.1. *The* price of negligence*, denoted* Λ*, is the maximum ratio between road*

capacity at optimal vehicle arrangement on a road to capacity at the worst-case vehicle lane assignment. This is formalized as follows:

$$\Lambda = \max_{\bar{\alpha},n,L,h,\bar{h}} \frac{C^{UB}(\boldsymbol{\beta}^*)}{C(\boldsymbol{\alpha}_*)}$$
$$\textit{s.t. feasible network parameters}$$

Remark 6.5. *Note that worst-case lane assignment in the denominator above maintains the Bernoulli assumption,* i.e. *not considering worst-case car ordering. Under worst-case ordering, autonomous and human-driven cars would be interleaved, unlike the assumption leading to the capacity model in* (6.1).

Definition 6.2. *The* price of no control*, denoted* Γ*, is the maximum ratio between road capacity at optimal vehicle arrangement on a road and road capacity at best-case vehicle lane assignment without low-level vehicle arrangement. This is formalized as*

$$\Gamma = \max_{\bar{\alpha},n,L,h,\bar{h}} \frac{C^{UB}(\boldsymbol{\beta}^*)}{C(\boldsymbol{\alpha}^*)}$$
$$\textit{s.t. feasible network parameters}$$

With this in mind, we present our bounds on these quantities.

Theorem 6.3. *The* price of negligence *is bounded by*

$$\Lambda \leq 2\frac{L+h-\sqrt{(L+h)(L+\bar{h})}}{h-\bar{h}} \leq 2\,,$$

and the price of no control *is bounded by*

$$\Gamma \leq \frac{2n(L+h)}{(2n-1)(L+h)+\sqrt{(L+h)(L+\bar{h})}} \leq \frac{2n}{2n-1}\,.$$

The proof sketch of this theorem is found in Appendix 6.7.4.

Remark 6.6. *The derived upper bounds on* Λ *and* Γ *are achieved by setting vehicle length* L *and short headway* $\bar{h}$ *to 0. Intuitively, the gain due to platooning increases as the space taken up by a platoon decreases.*

Remark 6.7. *Note that* $\Gamma \leq \Lambda$*, with equality when* $n = 1$*. This is because when there is only one lane, there is no gain from optimal lane assignment.*

The implications of the upper bounds on price of negligence and price of no control are perhaps surprising. This means that if vehicle type is decided by a Bernoulli process, optimally ordering the vehicles will result in at most a factor of 2 increase in capacity. In a more realistic scenario, consider $L = 4$, $h = 30$, $\bar{h} = 11$. Then, $\Lambda \leq 1.202$, meaning that the maximum possible improvement is approximately 20%.

Note that a 20% increase in capacity may yield far more than a 20% improvement in road latency. For example, a fourth-order polynomial is commonly used in traffic literature. Consider, as in [74, Ch. 13], road latency of the form $\ell(x) = t^0[1 + \rho(\frac{x}{c'})^\sigma]$, where x is the volume of traffic on the road, c' is the "practical capacity", t^0 is the free-flow travel time, and ρ and σ are typically 0.15 and 4, respectively. Using this model, under a given traffic flow, a 20% increase in capacity leads to a 50% reduction in latency due to congestion. In high congestion, this results in a large decrease in total latency.

6.4 Achieving Optimal Routing: Mid-level Optimization

In this section we describe our algorithm for approaching the capacity increases yielded from the optimal lane assignment and optimal vehicle ordering. We describe

our algorithm and discuss the theoretical upper bound for performance at each stage. Below is a summary of the policy:

- Phase 0: Cars are initialized in an intermixed configuration and ordering in all lanes are determined by a Bernoulli process with parameter $\bar{\alpha}$.
- Phase 1: Autonomous vehicles follow optimal lane assignment.
- Phase 2: Autonomous vehicles *affect* human drivers in local interactions to follow optimal lane assignment.
- Phase 3: Autonomous vehicles start platooning in the mixed lane, if one exists.

Note that if we were to terminate the policy at phase 2, we would want the lane assignment to follow the solution to the optimization in (6.5). However, if we terminate the policy at phase 3, the optimal lane assignment is the solution to (6.7). The number of purely autonomous lanes in the solutions are the same (see Remark 6.4), but the autonomy level in the mixed lanes will differ. We choose the optimistic lane assignment, which assumes that we can optimally rearrange the vehicles.

6.4.1 Phase 1: Autonomous cars following optimal lane assignment

In this stage, pictured in Fig. 6.2 (Phase 1), autonomous vehicles switch lanes to satisfy their optimal lane assignment determined by $\boldsymbol{\alpha}^*$, the ordered solution to (6.7). The human-driven vehicles though are not assumed to follow the optimal lane assignment by the end of this stage.

To accomplish this, each autonomous vehicle can be directed to switch to a specific lane, if there is a strong centralized control with close sensing and communication with the vehicles. This can be accomplished in a decentralized manner as well, if the autonomous vehicles are given a vector of probabilities $\boldsymbol{q}^* = [q_1^* \ q_2^* \ ... \ q_n^*]^T$ corresponding to $\boldsymbol{\alpha}^*$,

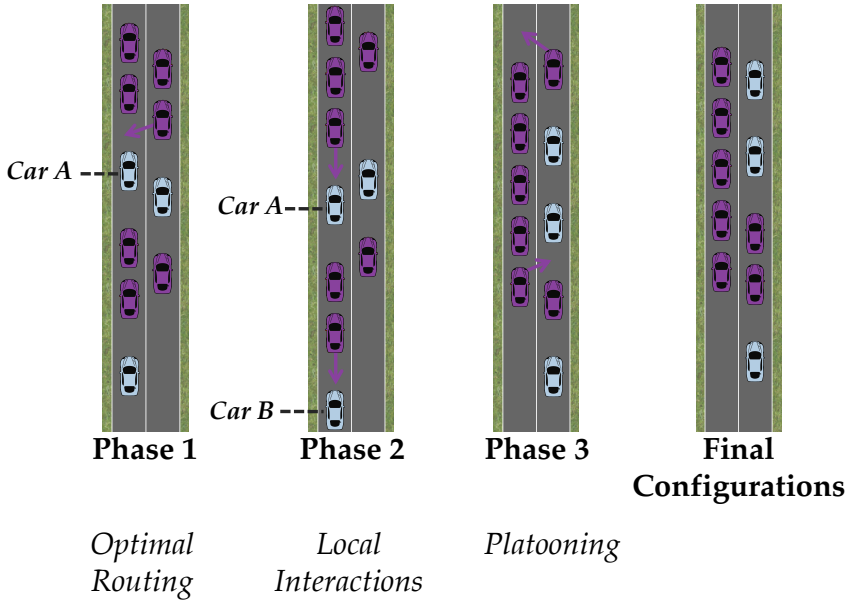


Figure 6.2: Phase 1: Autonomous vehicles follow optimal lane assignment. Here, the acting autonomous car pairs with human car A. Phase 2: Autonomous vehicles influence humans to follow optimal lane assignment. Here, the acting autonomous cars pair with the human cars A & B. Phase 3: Autonomous vehicles platoon in the mixed lane.

where q_i^* is the probability that an autonomous vehicle is optimally assigned to lane i. The vehicles then sample from a generalized Bernoulli distribution with parameter $\boldsymbol{q}^*$ to determine which lane to inhabit.

However the desired lane is selected, in the process of switching lanes, each autonomous vehicle will determine its action by performing the nested optimization in (6.9) while including behavior of the human-driven vehicles in their path. This pairing allows the autonomous car to model the human's responses, allowing it to merge even when there is not a large space for it, by relying on trailing human-driven vehicles slowing down in response to the autonomous car. The result of this phase is shown for a simulated example in Fig. 6.2.

6.4.2 Phase 2: Autonomous cars locally interacting with humans

This phase takes place only if the solution to (6.5) yields at least one lane with all autonomous vehicles, *i.e.* $\alpha_1 = 1$. Otherwise, all influence over human cars occurs in the mixed lane, which takes place during Phase 3.

If there are some lanes that under optimal lane assignment have only autonomous vehicles, in this stage the autonomous vehicles in those lanes influence the human-driven vehicles to enforce this solution, as shown in Fig. 6.2 (Phase 2). Lane-by-lane, starting with lane 1, autonomous cars influence human drivers to leave their lanes. This is done by having each autonomous car check if there is a human-driven vehicle behind it; if so, it pairs with that vehicle and with the goal to influence them to merge rightward. In a decentralized manner, each autonomous vehicle in an autonomous lane monitors its surroundings, and if there are no human-driven vehicles to the left of it, it checks for a human-driven vehicle behind it. If there is one, it influences it to move to the lane on the right.

By the end of this phase, lanes designated as purely autonomous will be so, with all autonomous vehicles in these lane forming long platoons. We expect the total capacity at the end of this phase to compare to the solution of (6.5).

6.4.3 Phase 3: Autonomous cars platooning

Though the theoretical maximum capacity under the Bernoulli assumption has been achieved at this point, if there is a lane with mixed autonomy, the capacity can be further improved by platooning the vehicles in that lane. If the mixed lane is the only lane with autonomous vehicles, i.e., $\forall i,\ \alpha_i^* < 1$, then autonomous vehicles can pair with human-driven cars behind them, leading them out of the lane, until they reach another

autonomous car and platoon with it.

However, if there is a lane with full autonomy, this scheme can be more easily achieved through a series of swaps, as seen in Fig. 6.2 (Phase 3). If there is a lone autonomous vehicle, it can merge into the adjacent full-autonomy lane, and another car can exit the lane to join a platoon in the mixed lane. The exiting vehicle can do this by pairing with the human-driven vehicle behind the vehicle it wants to platoon with, so that it will not be too cautious in switching lanes. Alternatively, the vehicle to be platooned with can slow down, and the merging vehicle move in front of it.

By the end of this phase, total capacity will be lower bounded by the optimal capacity in (6.5), and upper bounded by the optimum of (6.7).

6.5 Experimental Results

In this section, we discuss our autonomous driving simulation framework and compare the results of the simulation to the theoretical bounds established in Section 6.3.1.

6.5.1 Experimental Setup

We simulate a two-lane road with 20 vehicles, at varying levels of autonomy. To determine vehicle arrangement at the start of the simulation, we create a population with desired number of vehicles of each type, shuffle it, then randomly place the vehicles into lanes, ensuring that each lane has the same number of vehicles. This begins phase 0, in which cars are intermixed as the result of a Bernoulli process. Human-driven cars are simulated as agents optimizing a reward function, where reward functions are taken from those used in [153].

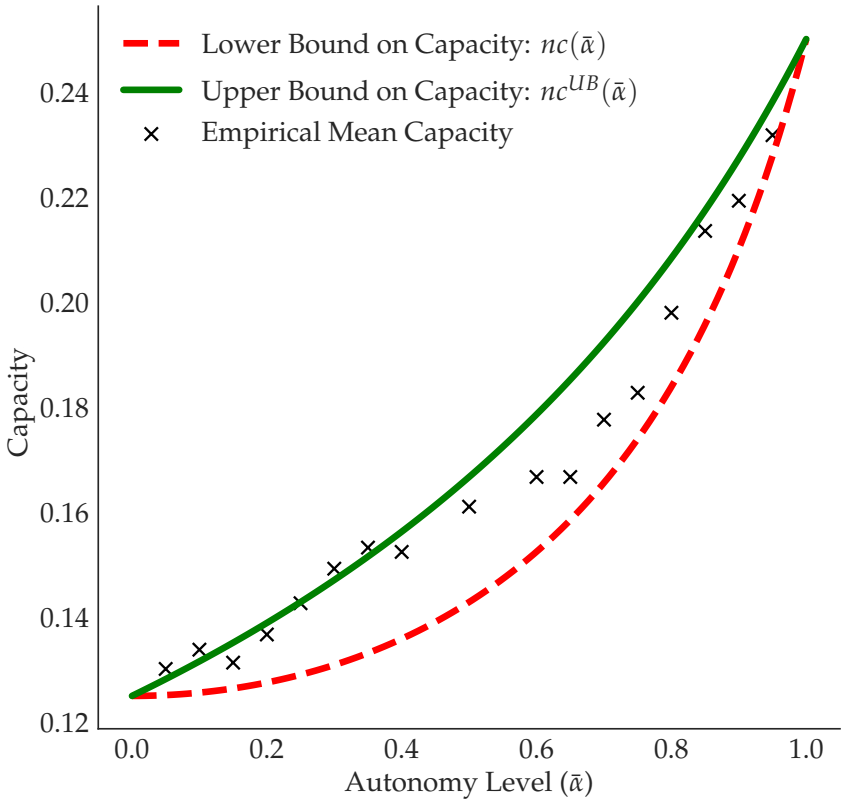


Figure 6.3: Capacity with Local Interactions. The achieved capacity at various autonomy levels plotted against the achievable capacity with optimal local interactions (green) and achievable capacity without any local interaction.

6.5.2 Driving Simulator

We use a simple point-mass model for the dynamics of each vehicle, where the state of the system is: $x = [x \;\; y \;\; \theta \;\; v]^\top$. x, y represent the coordinates of the vehicle on the road, θ is its heading, and v is its velocity. We assume two control inputs $u = [u_1 \;\; u_2]$ each representing the steering angle and the acceleration of the vehicle. We then represent the dynamics of each vehicle as the following, where γ is the friction coefficient:

$$[\dot{x} \;\; \dot{y} \;\; \dot{\theta} \;\; \dot{v}] = [v \cdot \cos(\theta) \quad v \cdot \sin(\theta) \quad v \cdot u_1 \quad u_2 - \alpha \cdot v] \; .$$

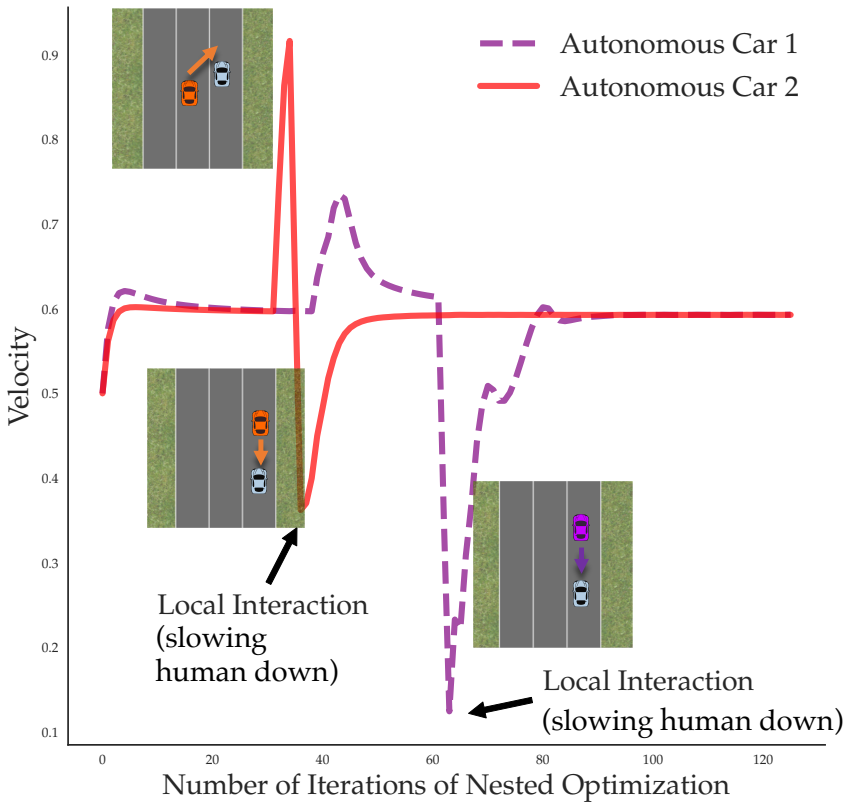


Figure 6.4: Timing of Local Interactions. The variation in velocity over 125 iterations of the mid-level optimization for 2 of the autonomous cars in the network.

6.5.3 Simulation Results

Recall the situation of Fig. 6.1 and note that our goal is to start in a configuration similar to the left side and end in a situation such as the right. Noting our theoretical upper and lower bounds on the effects of lane assignment and ordering, we show in Fig. 6.3 the results of the mid-level optimization and compare it to the earlier derived upper and lower bounds. Constrained by computation, we simulate 200 iterations of the mid-level optimization using a setup with 20 cars at varying autonomy levels. As expected, each run lies between the two curves. We note that the few data points which exceed the upper bound are due to the asymptotic nature of the bound and the fact that we are using only 20 cars in the network. Note also that at lower levels of autonomy, the

data points are closer to the upper bound, whereas at higher levels, the data points are closer to the lower bound. This is due to the fact that the nested optimization is done one car at a time and with only one other car in the network. We illustrate the timing of the local interactions in Fig. 6.4, noting that when the velocity of the autonomous cars drops dramatically, it indicates a movement intended to influence a human car.

6.5.4 Implementation Details

In our implementation of the nested optimization, we used the software package Theano [173,174] to symbolically compute all Jacobians and Hessians. Theano optimizes the computation graph into efficient C code, which is crucial for real-time applications. In our implementation, each step of our optimization is solved with a horizon length $N = 5$. We run the large-scale simulations with 20 cars in the network on a cluster using 4 CPU's with a maximum utilization of 36 GB RAM between them. We make the code publicly available here: https://github.com/Stanford-HRI/MultilaneInteractions.

6.6 Discussion and Conclusion

In this chapter[4] we introduced a procedure to connect network-level control of traffic networks with low-level local interactions amongst vehicles. To this end, we provided theoretical guarantees on the potential of re-ordering vehicles as well as allocating proportions of vehicles optimally. Finally, we implemented an algorithm to achieve the optimal ordering.

Limitations. The nested optimization is extremely computationally heavy, and is

[4]Some of the computing for this project was performed on the Sherlock cluster. We would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results. This work was supported in part by NSF grant CCF-1755808 and the UC Office of the President grant LFR-18-548175.

natural to be run in a distributed setting in which each car runs its own nested optimization. It would be interesting to implement a fully distributed version of the algorithm and compare the results to the current implementation. Additionally, we emphasize that the simulations were not run with real humans, but rather with agents whose reward functions were learned from humans. Each human is, however, different and this could have large impacts on the performance of the algorithm. This may be addressed in future work.

Conclusion. We demonstrated the connection between network-level and vehicle-level control in transportation networks and designed a unifying scheme utilizing both to achieve optimal efficiency in the network.

6.7 Appendix

6.7.1 Proof of Theorem 6.1

Assume, for the purpose of contradiction, that there exists a solution $\boldsymbol{\alpha}^*$ to (6.5) with components $\alpha_i^*,\ \alpha_j^* \in (0,1)$, where $i \neq j$. Assume without loss of generality that $\alpha_i \leq \alpha_j$. If $\alpha_i = \alpha_j$, we can construct a new solution with same total capacity in which $\alpha_i \neq \alpha_j$. This is because $\frac{\partial}{\partial \alpha_i} C(\boldsymbol{\alpha}) = \frac{\partial}{\partial \alpha_j} C(\boldsymbol{\alpha})$ and $\frac{\partial}{\partial \alpha_i} G(\boldsymbol{\alpha}) = \frac{\partial}{\partial \alpha_j} G(\boldsymbol{\alpha})$, so the autonomy level on road i can be decreased by some infinitesimal value ϵ and autonomy level on road j increased by ϵ, thereby maintaining the same capacity sum and satisfying the constraint.

Now that we have established an optimal solution with $\alpha_i^* < \alpha_j^*$ with $\alpha_i^*,\ \alpha_j^* \in (0,1)$, we explore a further perturbation of our solution. We construct a new lane assignment $\hat{\boldsymbol{\alpha}}$ with $\hat{\alpha}_i = \alpha_i^* - \epsilon_1$ and $\hat{\alpha}_j = \alpha_j^* + \epsilon_2$, where $\epsilon_1,\ \epsilon_2 > 0$ with these unequal infinitesimal perturbations designed such that the constraint remains satisfied. Note that this can be accomplished with nonnegative perturbations as $\frac{\partial}{\partial \alpha_i} G(\boldsymbol{\alpha}) = \frac{k_2 \alpha_i^2 - 2k_2 \alpha_i \bar{\alpha} + k_1}{(k_1 - k_2 \alpha_i^2)^2} > 0$, which is

positive since $k_1 > k_2$.

Since we choose the relative scaling of ϵ_1 and ϵ_2 based on their contribution to the constraint function, we can find the relative change in the objective due to the perturbations by the expression

$$\frac{\frac{\partial}{\partial\alpha_j}C(\boldsymbol{\alpha})}{\frac{\partial}{\partial\alpha_j}G(\boldsymbol{\alpha})} - \frac{\frac{\partial}{\partial\alpha_i}C(\boldsymbol{\alpha})}{\frac{\partial}{\partial\alpha_i}G(\boldsymbol{\alpha})} = \frac{2k_2\alpha_j}{k_2\alpha_j^2 - 2k_2\alpha_j\bar{\alpha} + k_1} - \frac{2k_2\alpha_i}{k_2\alpha_i^2 - 2k_2\alpha_i\bar{\alpha} + k_1} > 0$$

since $\frac{2k_2\alpha_i}{k_2\alpha_i^2 - 2k_2\alpha_i\bar{\alpha} + k_1} > 0$ and

$$\frac{\partial}{\partial\alpha_i}\frac{2k_2\alpha_i}{k_2\alpha_i^2 - 2k_2\alpha_i\bar{\alpha} + k_1} = \frac{2k_2(k_1 - \alpha^2 k_2)}{(k_2\alpha_i^2 - 2k_2\alpha_i\bar{\alpha} + k_1)^2} > 0\ ,$$

as $k_1 > k_2 > 0$. Therefore, the addition of ϵ_2 to α_j increases the objective more than the subtraction of ϵ_1 from α_i decreases it. Then, $C(\hat{\boldsymbol{\alpha}}) > C(\boldsymbol{\alpha}^*)$, so $\boldsymbol{\alpha}^*$ is not an optimum, proving the theorem. □

6.7.2 Proof of Theorem 6.2

Assume that $\bar{\alpha} < 1$, so $m < n$. Let $g(\alpha_i) = (\alpha_i - \bar{\alpha})c(\alpha_i)$. Then using (6.6),

$$\begin{aligned} 0 = G(\boldsymbol{\alpha}^*) &= \sum_{i=1}^{n} g(\alpha_i^*) = \sum_{i=1}^{m} g(1) + g(\alpha_{m+1}^*) + \sum_{i=m+2}^{n} g(0) \\ &= mg(1) + g(\alpha_{m+1}^*) + (n - m - 1)g(0)\ . \end{aligned}$$

We can solve for m, giving

$$m = \frac{(k_1 - k_2)(n\bar{\alpha}(k_1 - \alpha_{m+1}^*)k_2 - \alpha_{m+1}^*(k_1 - \alpha_{m+1}^*\bar{\alpha}k_2))}{(k_1 - \alpha_{m+1}^{*2})(k_1 - \bar{\alpha}k_2)}\ .$$

One can show that $\frac{\partial m}{\partial\alpha_{m+1}^*} < 0$. Note that by definition, $\alpha_{m+1}^* \in [0, 1)$ (as m represents

the index of the first lane such that $\alpha_i \neq 1$), so we can bound m using these values:

$$\frac{\bar{\alpha}n(k_1 - k_2)}{k_1 - \bar{\alpha}k_2} - 1 < m \leq \frac{\bar{\alpha}n(k_1 - k_2)}{k_1 - \bar{\alpha}k_2} .$$

Therefore, $m = \lfloor \frac{\bar{\alpha}n(k_1-k_2)}{k_1-\bar{\alpha}k_2} \rfloor$. Note that in the case excluded at the start of the proof, when $\bar{\alpha} = 1$, this expression yields $m = n$, which is correct. Therefore, this expression is true for $\bar{\alpha} \in [0, 1]$. □

6.7.3 Proof of Proposition 6.1

First note that given feasible lane assignment vector $\boldsymbol{\alpha}$, if there exists one element $\alpha_i > \bar{\alpha}$, then there must also exist element $\alpha_j < \bar{\alpha}$. This follows from the fact that $G(\nVdash_n^T \bar{\alpha}) = 0$ and $\frac{\partial}{\partial \alpha_i} G(\boldsymbol{\alpha}) > 0$.

Now we can prove the proposition by recursion. The base case is $\boldsymbol{\alpha} = \nVdash_n^T \bar{\alpha}$. For any other feasible lane assignment, there are at most n lane autonomy levels α_i not equal to $\bar{\alpha}$. Pick any pair of lanes i and j such that $\alpha_i > \bar{\alpha} > \alpha_j$ (which, when not in the base case, are guaranteed to exist by the fact above). Then, using the reverse of the mechanism in the proof of Theorem 6.1, we keep the constraint satisfied while decreasing α_i and increasing α_j, while monotonically decreasing the road capacity. This continues until either $\alpha_i = \bar{\alpha}$, $\alpha_j = \bar{\alpha}$, or both. Now we have a lane assignment vector with lower road capacity than the original, with at most $n - 1$ lane autonomy levels not equal to $\bar{\alpha}$, proving the proposition. □

6.7.4 Proof of Theorem 6.3

To prove the first statement, note that due to Proposition 6.2, the cost for any feasible lane assignment for the upper bound capacity function is equivalent to $nc^{UB}(\bar{\alpha})$. For given

network parameters k_1 and $k_2 h$, we find our price of negligence:

$$\Lambda \leq \max_{\bar{\alpha}\in[0,1]} \frac{nc^{UB}(\bar{\alpha})}{nc(\bar{\alpha})} = \max_{\bar{\alpha}\in[0,1]} \frac{k_1 - \bar{\alpha}^2 k_2}{k_1 - \bar{\alpha} k_2} \, .$$

This term is concave with respect to $\bar{\alpha}$ for $\bar{\alpha} \in [0,1]$, with maximum at $\bar{\alpha} = \frac{k_1 - \sqrt{k_1^2 - k_1 k_2}}{k_2}$. Plugging this in,

$$\begin{aligned} \Lambda &\leq 2\frac{k_1 - \sqrt{k_1^2 - k_1 k_2}}{k_2} = 2\frac{L + h - \sqrt{(L+h)(L+\bar{h})}}{h - \bar{h}} \\ &\leq 2. \end{aligned} \tag{6.11}$$

To observe (6.11), note that $k_1 - 2k_2 \leq \sqrt{(k_1 - k_2)k_1}$.

For the second statement, first note that if $\bar{\alpha} = 1$, the price of no control would be 1. Since $\Gamma \geq 1$, let us exclude that case. Now, assuming $\bar{\alpha} < 1$ we apply Theorem 6.1 and Proposition 6.2:

$$\begin{aligned} \Gamma &\leq \frac{nc^{UB}(\bar{\alpha})}{\sum_{i=1}^{m} c(1) + c(\alpha^*_{m+1}) + \sum_{i=m+2}^{n} c(0)} \\ &\leq \frac{nc^{UB}(\bar{\alpha})}{c(\alpha^*_1) + (n-1)c(0)} \, , \end{aligned} \tag{6.12}$$

where (6.12) follows from the fact that $\Gamma \geq 1$ and $c(1) > c(\alpha^*_{m+1}) \geq c(0)$ so the denominator is minimized when $m = 0$ (meaning there are no purely autonomous lanes).

We can then solve $G(\boldsymbol{\alpha}^*) = 0$ to find an expression for $\bar{\alpha}$ as a function of α^*_1, yielding

$$\Gamma \leq \frac{n(k_1 - \alpha^*_1 k_2)}{n(k_1 - \alpha^*_1 k_2) - \alpha^*_1 k_2 (1 - \alpha^*_1)} \, .$$

We then consider bounding $1/\Gamma$, which we show to be convex.

$$\frac{\partial^2}{\partial \alpha_1^{*2}}(1/\Gamma) = \frac{2k_2(k_1^2 + 3k_1k_2(-1+\alpha_1^*)\alpha - k_2^2\alpha_1^{*3})}{n(k_1 - \alpha_1^{*2}k_2)^3}$$

The inner term on the numerator is convex for $\alpha_1^* \in [0,1]$, with second derivative $6k_2(k_1 - \alpha_1^* k_2)$. Its minimum over that interval occurs at $\alpha_1^* = \frac{k_1 - \sqrt{k_1^2 - k_1k_2}}{k_2}$, yielding $\frac{\partial^2}{\partial \alpha_1^{*2}}(1/\Gamma) > 0$. Solving $\frac{\partial}{\partial \alpha_1^*}(1/\Gamma) = 0$, we find that the minimum of this outer equation also occurs at $\alpha_1^* = \frac{k_1 - \sqrt{k_1^2 - k_1k_2}}{k_2}$. Using this,

$$\begin{aligned}
\Gamma &\leq \frac{4k_1n^2 - 2n(k_1 + \sqrt{k_1^2 - k_1k_2})}{4k_1n(n-1) + k_2} \\
&= \frac{2n(L+h)}{(2n-1)(L+h) + \sqrt{(L+\bar{h})(L+\bar{h})}} \\
&\leq \frac{2n}{2n-1}\,.
\end{aligned}$$

As k_2 approaches k_1, the above inequality becomes tight. □

Chapter 7

Conclusion

In this dissertation we examined a number of mechanisms by which a system designer can decrease congestion in the setting of mixed autonomy. We provided a congestion game-like model for traffic networks in mixed autonomy and provided worst-case bounds on the inefficiency due to selfish routing. We provided optimal tolls, analyzed the potential benefits of anonymous tolling, and provided anonymous tolls with a performance guarantee. We also studied the setting in which autonomous vehicles are used for ride-hailing services, and provided a method for pricing these services to incentivize good traffic equilibria. We extended this to the dynamic setting, using Reinforcement Learning to learn routing for autonomous vehicles which influences human drivers to choose routing which decongests the network. Finally we bridged the gap between two capacity models for mixed autonomy, and provided a method by which autonomous vehicles can rearrange themselves to form platoons to maximize road capacity in the presence of human-driven vehicles. In all these settings we provided theoretical results to contextualize and serve as a benchmark for the performance of our algorithms.

7.1 Future Directions

There are many promising directions to continue research in this field. We organize this by specifying directions specific to each chapter, followed by more general comments.

Chapter 2. In this chapter we considered a model for road capacity under mixed autonomy which considers the effects of platooning. This can be extended to include other effects of mixed autonomy, as well as including other models for the effects of platooning – *e.g.* generalizing the second capacity model to more than two vehicle types. Moreover, while the Price of Anarchy upper bound is order-optimal with respect to the maximum degree of asymmetry, it is not necessarily so with respect to the maximum polynomial degree, warranting further investigation. We lower bound the Price of Stability as scaling with the square root of the maximum degree of asymmetry; it would be valuable to upper bound this quantity as well, and compare it to the Price of Anarchy, which scales linearly with the maximum degree of asymmetry.

Chapter 3. Our tolling results all consider an affine latency model, so a natural extension is to consider polynomial latency functions. This direction is valuable, though it requires sacrificing the main tools used to derive our results, namely the usage of the Variational Inequality in the proofs of Theorem 3.1 and 3.2, as well as Theorem 3.3, which is used to prove Theorem 3.4 about almost anonymous tolls. However, such an extension would be a valuable addition. It would also be valuable to analyze the robustness of the considered tolls to mischaracterizations of flow demand or network parameters. One approach to improve robustness is a variable marginal cost toll, in which the toll on a road is depends on the current flow of each vehicle type on that road. We show that such a toll does not guarantee that all resulting equilibria will coincide with the socially optimal routing, and that the gap between one resulting equilibrium and the socially optimal routing can be unbounded. However, further characterizing this gap would be

a valuable contribution. Specifically, we provide a lower bound on the Price of Anarchy with respect to the maximum degree of asymmetry; it would be valuable to provide an upper bound as well and compare this bound to the Price of Anarchy bound on an untolled network.

Another possible extension is to consider a population in which different users have different sensitivities to tolls. Another direction is to consider a middle ground between fully differentiated tolls and completely anonymous tolls, such as grouping vehicle classes together which have similar effects on road congestion, therefore requiring a smaller number of distinct tolls than the number of vehicle classes. In such a scheme, instead of Price of Anarchy scaling with the maximum degree of asymmetry, it is possible that it will scale with the maximum degree of asymmetry within each grouping of vehicle types. Such an approach would allow for a more flexible implementation with correspondingly flexible performance guarantees.

Chapter 4. When estimating and leveraging the sensitivity of users to features such as cost and latency, it is important to ensure that the estimates accurately reflect true user preferences, including changing preferences. Since the active learning to learn preferences may depend on stated preferences instead of revealed preferences (in the event that these answers are found when users complete a survey rather than actually choosing options they will follow through with), it is necessary to adjust the estimates to account for a possible resulting distortion. However, revealed preferences will become available as users use the service. Accordingly, a promising direction of future research is to adjust the preferences learned from the active learning phase with the revealed preferences found when the user is using the service (and the system designer no longer has complete control over the queries posed to the user). This method can also be used to track changing user preferences. Another direction for future work is to extend the work to more general networks.

We use Fundamental Diagram of Traffic-based model to relate vehicle flow and density. Kerner has argued that a three-phase model more accurately reflects empirical observations of vehicle flow [175]. However, in this context, a FDT-based model is useful in that we solve for the *best* equilibrium, and implicitly rely on methods such as those in Chapter 5 to dynamically move to the best available equilibrium. We can consider the FDT model used to be an optimistic version of a model which includes hysteresis. Thus, if we can relate the static equilibrium to the fluctuations which emerge in the three-phase model, and show that a static equilibrium is an optimistic estimate of the average flow under the fluctuations, pricing according to the model derived using this FDT-based model is useful even in a network with flow which follows the three-phase model. This future work will further justify the use of such equilibrium-based methods.

Chapter 5. As mentioned above, there are many models which relate vehicle flow, density, and velocity. In this chapter we use a Cell Transmission Model-based model, extended to incorporate the effects of mixed autonomy. We then use deep Reinforcement Learning (RL) to learn a controller for the autonomous flow to minimize total travel time. By using a CTM-based approach, we are able to establish benchmarks to showcase the efficacy of our method. Since we have proven the concept of using RL for controlling one vehicle type to influence the choices of the other vehicle type, the advantage of using RL is that our method can easily be extended to other models for vehicle flow, such as Kerner's three-phase model. Another important direction for future research, though we provide a method which works for a relatively large state space, is to truly address the scalability issue and allow for larger networks with more source-destination pairs.

Chapter 6. In this chapter we provided a low-level scheme for increasing capacity on roads through platooning. However, in some cases, increasing the capacity on a road can yield *worse* equilibrium behavior. Accordingly, one feasible direction is having vehicles which adjust their platooning behavior, and therefore effect on capacity, based on the

road they are on in order to improve a global goal of decreasing congestion. In another direction, our scheme only allows for computing *pairwise* interactions; future work can extend this to groups of vehicles. This work can also be improved by using a more decentralized method.

On a broader scale, there are other directions which this dissertation did not consider. It is possible that the free-flow latency of a road depends on the autonomy level as well; this can be considered in future models. In this dissertation we focused primarily on decreasing congestion by influencing route choice. Another important direction is to consider influencing the *timing* of trip departures to decrease congestion. We mentioned above that improving capacity via platooning may worsen equilibrium routing. This can be incorporated in an important extension of this work, namely fusing the different approaches on different scales presented in this work. One can consider a scheme which includes pricing, dynamic routing, and platoon formation control. The goal would then be to use these tools in conjunction to improve the performance of a traffic network. This task, though challenging, would allow planners to use the emergence of autonomous vehicles as a tool by which to improve the efficiency of traffic networks for all users.

Bibliography

[1] E. Bıyık, D. A. Lazar, R. Pedarsani, and D. Sadigh, *Altruistic autonomy: Beating congestion on shared roads*, in *Workshop on the Algorithmic Foundations of Robotics*, 2018.

[2] J. de Dios Ortúzar and L. G. Willumsen, *Modelling transport*. John wiley & sons, 2011.

[3] D. Schrank, B. Eisele, T. Lomax, and J. Bak, *Urban mobility scorecard*, 2015.

[4] A. Henao, *Impacts of Ridesourcing-Lyft and Uber-on Transportation Including VMT, Mode Replacement, Parking, and Travel Behavior*. University of Colorado at Denver, 2017.

[5] G. Gomes, R. Horowitz, A. A. Kurzhanskiy, P. Varaiya, and J. Kwon, *Behavior of the cell transmission model and effectiveness of ramp metering*, *Transportation Research Part C: Emerging Technologies* (2008).

[6] F. Belletti, D. Haziza, G. Gomes, and A. M. Bayen, *Expert level control of ramp metering based on multi-task deep reinforcement learning*, *IEEE Transactions on Intelligent Transportation Systems* (2018).

[7] W. Hu, *Over $10 to drive in manhattan? what we know about the congestion pricing plan*, *The New York Times* (2019).

[8] X.-Y. Lu, P. Varaiya, R. Horowitz, D. Su, and S. E. Shladover, *Novel freeway traffic control with variable speed limit and coordinated ramp metering*, *Transportation Research Record* (2011).

[9] S. Coogan, E. Kim, G. Gomes, M. Arcak, and P. Varaiya, *Offset optimization in signalized traffic networks via semidefinite relaxation*, *Transportation Research Part B: Methodological* **100** (2017) 82–92.

[10] J. Lazarus, J. Ugirumurera, S. Hinardi, M. Zhao, F. Shyu, Y. Wang, S. Yao, and A. M. Bayen, *A decision support system for evaluating the impacts of routing applications on urban mobility*, in *21st International Conference on Intelligent Transportation Systems*, IEEE, 2018.

[11] M. Wu, S. Amin, and A. E. Ozdaglar, *Value of information systems in routing games*, *arXiv preprint arXiv:1808.10590* (2018).

[12] R. Johari, S. Mannor, and J. N. Tsitsiklis, *Efficiency loss in a network resource allocation game: the case of elastic supply*, *IEEE Transactions on Automatic Control* **50** (2005), no. 11 1712–1724.

[13] S. C. Dafermos and F. T. Sparrow, *The traffic assignment problem for a general network*, *Journal of Research of the National Bureau of Standards B* (1969).

[14] D. W. Hearn, S. Lawphongpanich, and S. Nguyen, *Convex programming formulations of the asymmetric traffic assignment problem*, *Transportation Research Part B: Methodological* (1984).

[15] M. Florian, *A traffic equilibrium model of travel by car and public transit modes*, *Transportation Science* (1977).

[16] M. J. Smith, *The existence, uniqueness and stability of traffic equilibria*, *Transportation Research Part B: Methodological* (1979).

[17] S. Dafermos, *Traffic equilibrium and variational inequalities*, *Transportation science* (1980).

[18] C. Papadimitriou, *Algorithms, games, and the internet*, in *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pp. 749–753, ACM, 2001.

[19] T. Roughgarden and É. Tardos, *How bad is selfish routing?*, *Journal of the ACM (JACM)* (2002).

[20] T. Roughgarden, *The price of anarchy is independent of the network topology*, *Journal of Computer and System Sciences* (2003).

[21] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses, *A geometric approach to the price of anarchy in nonatomic congestion games*, *Games Econ. Behavior* (2008).

[22] G. Perakis, *The "price of anarchy" under nonlinear and asymmetric costs*, *Mathematics of Operations Research* (2007).

[23] M. Beckmann, C. B. McGuire, and C. B. Winsten, *Studies in the economics of transportation*, tech. rep., 1956.

[24] S. C. Dafermos, *Toll patterns for multiclass-user transportation networks*, *Transportation science* **7** (1973), no. 3 211–223.

[25] L. Fleischer, K. Jain, and M. Mahdian, *Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games*, in *Symposium on Foundations of Computer Science*, IEEE, 2004.

[26] P. N. Brown and J. R. Marden, *The robustness of marginal-cost taxes in affine congestion games*, *IEEE Transactions on Automatic Control* (2017).

[27] T. Roughgarden, *Stackelberg scheduling strategies*, *SIAM Journal on Computing* (2004).

[28] V. Bonifaci, T. Harks, and G. Schäfer, *Stackelberg routing in arbitrary networks*, *Mathematics of Operations Research* **35** (2010), no. 2 330–346.

[29] W. Krichene, J. D. Reilly, S. Amin, and A. M. Bayen, *Stackelberg routing on parallel transportation networks*, *Handbook of Dynamic Game Theory* (2018).

[30] S. E. Shladover, *Longitudinal control of automated guideway transit vehicles within platoons*, *Journal of Dynamic Systems, Measurement, and Control* (1978).

[31] J. Lioris, R. Pedarsani, F. Y. Tascikaraoglu, and P. Varaiya, *Platoons of connected vehicles can double throughput in urban roads*, *Transportation Research Part C: Emerging Technologies* (2017).

[32] A. Askari, D. A. Farias, A. A. Kurzhanskiy, and P. Varaiya, *Effect of adaptive and cooperative adaptive cruise control on throughput of signalized arterials*, in *IEEE Intelligent Vehicles Symposium*, 2017.

[33] M. W. Levin and S. D. Boyles, *A multiclass cell transmission model for shared human and autonomous vehicle roads*, *Transportation Research Part C: Emerging Technologies* **62** (2016) 103–116.

[34] C. F. Daganzo, *The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory*, *Transportation Research Part B: Methodological* (1994).

[35] R. E. Stern, S. Cui, M. L. Delle Monache, R. Bhadani, M. Bunting, M. Churchill, N. Hamilton, H. Pohlmann, F. Wu, B. Piccoli, *et. al.*, *Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments*, *Transportation Research Part C: Emerging Technologies* (2018).

[36] C. Wu, A. M. Bayen, and A. Mehta, *Stabilizing traffic with autonomous vehicles*, in *International Conference on Robotics and Automation*, 2018.

[37] L. Jin, M. Čičič, S. Amin, and K. H. Johansson, *Modeling the impact of vehicle platooning on highway congestion: A fluid queuing approach*, in *Conference on Hybrid Systems: Computation and Control*, ACM, 2018.

[38] M. Čičić, L. Jin, and K. H. Johansson, *Coordinating vehicle platoons for highway bottleneck decongestion and throughput improvement*, *arXiv preprint arXiv:1907.13049* (2019).

[39] S. Sivaranjani, Y.-S. Wang, V. Gupta, and K. Savla, *Localization of disturbances in transportation systems*, in *IEEE Conference on Decision and Control (CDC)*, 2015.

[40] G. Gunter, D. Gloudemans, R. E. Stern, S. McQuade, R. Bhadani, M. Bunting, M. L. Delle Monache, R. Lysecky, B. Seibold, J. Sprinkle, *et. al.*, *Are commercially implemented adaptive cruise control systems string stable?*, *IEEE Transactions on Intelligent Transportation Systems* (2020).

[41] N. Mehr and R. Horowitz, *How will the presence of autonomous vehicles affect the equilibrium state of traffic networks?*, *arXiv preprint arXiv:1901.05168* (2019).

[42] D. A. Lazar, S. Coogan, and R. Pedarsani, *Capacity modeling and routing for traffic networks with mixed autonomy*, in *IEEE Conference on Decision and Control (CDC)*, 2017.

[43] D. A. Lazar, S. Coogan, and R. Pedarsani, *The price of anarchy for transportation networks with mixed autonomy*, in *IEEE American Control Conference (ACC)*, 2018.

[44] D. A. Lazar, S. Coogan, and R. Pedarsani, *Routing for traffic networks with mixed autonomy*, *IEEE Transactions on Automatic Control* (2020).

[45] *Beyond Traffic 2045: Trends and Choices*. US Department of Transportation, 2015.

[46] A. Askari, D. A. Farias, A. A. Kurzhanskiy, and P. Varaiya, *Measuring impact of adaptive and cooperative adaptive cruise control on throughput of signalized intersections*, *arXiv preprint arXiv:1611.08973* (2016).

[47] T. Roughgarden, *Selfish routing and the price of anarchy*, vol. 174. MIT press Cambridge, 2005.

[48] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses, *Selfish routing in capacitated networks*, *Mathematics of Operations Research* (2004).

[49] C. K. Chau and K. M. Sim, *The price of anarchy for non-atomic congestion games with symmetric cost maps and elastic demands*, *Operations Research Letters* (2003).

[50] J. R. Correa and N. E. Stier-Moses, *Wardrop equilibria*, *Wiley encyclopedia of operations research and management science* (2011).

[51] A. De Palma and Y. Nesterov, *Optimization formulations and static equilibrium in congested transportation networks*, tech. rep., 1998.

[52] M. Florian and D. Hearn, *Network equilibrium models and algorithms*, *Handbooks in Operations Research and Management Science* (1995).

[53] S. C. Dafermos, *The traffic assignment problem for multiclass-user transportation networks*, *Transportation science* (1972).

[54] F. Farokhi, W. Krichene, M. Alexandre, and K. H. Johansson, *When do potential functions exist in heterogeneous routing games?*, 2014.

[55] E. Koutsoupias and C. Papadimitriou, *Worst-case equilibria*, in *Annual Symposium on Theoretical Aspects of Computer Science*, Springer, 1999.

[56] K. Bhawalkar, M. Gairing, and T. Roughgarden, *Weighted congestion games: the price of anarchy, universal worst-case examples, and tightness*, *ACM Transactions on Economics and Computation (TEAC)* **2** (2014), no. 4 1–23.

[57] S. Kapoor and J. Shin, *Price of anarchy in networks with heterogeneous latency functions*, *Mathematics of Operations Research* **45** (2020), no. 2 755–773.

[58] S. Darbha and K. Rajagopal, *Intelligent cruise control systems and traffic flow stability*, *Transportation Research Part C: Emerging Technologies* **7** (1999), no. 6 329–352.

[59] J. Yi and R. Horowitz, *Macroscopic traffic flow propagation stability for adaptive cruise controlled vehicles*, *Transportation Research Part C: Emerging Technologies* (2006).

[60] R. Pueboobpaphan and B. van Arem, *Driver and vehicle characteristics and platoon and traffic flow stability: Understanding the relationship for design and assessment of cooperative adaptive cruise control*, *Transportation Research Record: Journal of the Transportation Research Board* (2010).

[61] G. Orosz, *Connected cruise control: modelling, delay effects, and nonlinear behaviour*, *Vehicle System Dynamics* (2016).

[62] S. Cui, B. Seibold, R. Stern, and D. B. Work, *Stabilizing traffic flow via a single autonomous vehicle: Possibilities and limitations*, in *IEEE Intelligent Vehicles Symposium*, 2017.

[63] C. Wu, A. Kreidieh, E. Vinitsky, and A. M. Bayen, *Emergent behaviors in mixed-autonomy traffic*, in *Conference on Robot Learning*, 2017.

[64] M. Motie and K. Savla, *Throughput analysis of a horizontal traffic queue under safe car following models*, in *IEEE Conference on Decision and Control (CDC)*, 2016.

[65] L. Jin, M. Čičić, S. Amin, and K. H. Johansson, *Modeling impact of vehicle platooning on highway congestion: A fluid queuing approach*, in *International Conference on Hybrid Systems: Computation and Control*, ACM, 2018.

[66] K.-Y. Liang, J. Mårtensson, and K. H. Johansson, *Heavy-duty vehicle platoon formation for fuel efficiency*, *IEEE Transactions on Intelligent Transportation Systems* (2016).

[67] A. Adler, D. Miculescu, and S. Karaman, *Optimal policies for platooning and ride sharing in autonomy-enabled transportation*, in *Workshop on Algorithmic Foundations of Robotics (WAFR)*, 2016.

[68] V. Turri, B. Besselink, and K. H. Johansson, *Cooperative look-ahead control for fuel-efficient and safe heavy-duty vehicle platooning*, *IEEE Transactions on Control Systems Technology* (2017).

[69] S. van de Hoef, K. H. Johansson, and D. V. Dimarogonas, *Fuel-efficient en route formation of truck platoons*, *IEEE Transactions on Intelligent Transportation Systems* (2018).

[70] F. Rossi, R. Zhang, Y. Hindy, and M. Pavone, *Routing autonomous vehicles in congested transportation networks: Structural properties and coordination algorithms*, *Autonomous Robots* (2018).

[71] D. Chen, S. Ahn, M. Chitturi, and D. A. Noyce, *Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles*, *Transportation research part B: methodological* (2017).

[72] N. Mehr and R. Horowitz, *Can the presence of autonomous vehicles worsen the equilibrium state of traffic networks?*, in *IEEE Conference on Decision and Control (CDC)*, 2018.

[73] *Bureau of public roads traffic assignment manual*, *US Department of Commerce* (1964).

[74] Y. Sheffi, *Urban transportation networks*, vol. 6. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[75] D. A. Lazar, S. Coogan, and R. Pedarsani, *Optimal tolling for heterogeneous traffic networks with mixed autonomy*, in *58th Conference on Decision and Control (CDC)*, pp. 4103–4108, IEEE, 2019.

[76] D. A. Lazar and R. Pedarsani, *Optimal tolling for multitype mixed autonomous traffic networks*, *IEEE Control Systems Letters* (2020).

[77] D. A. Lazar and R. Pedarsani, *Anonymous tolling for traffic networks with mixed autonomy*, *Submitted, IEEE Conference on Decision and Control (CDC)* (2021).

[78] Y. Noguchi, *Self-Driving Cars Could Ease Our Commutes, But That'll Take A While*. National Public Radio, 2017.

[79] A. C. Pigou, *The economics of welfare*, *McMillan&Co., London* (1920).

[80] N. Mehr and R. Horowitz, *Pricing traffic networks with mixed vehicle autonomy*, in *American Control Conference (ACC)*, 2019.

[81] M. U. Iqbal and S. Lim, *Designing tolling technologies with privacy in mind: A user perspective*, *Transportation Research Part C: Emerging Technologies* **61** (2008), no. 2 1–25.

[82] L. Rößger, J. Schade, and T. Tretvik, *Motivational factors influencing behavioural responses to charging measures in freight operator sector*, .

[83] J. Wardrop, *Some theoretical aspects of road traffic research*, in *Inst. Civil Engineers Proc. London, UK*, 1900.

[84] P. N. Brown, *A tragedy of autonomy. self-driving cars and urban congestion externalities*, in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 981–986, IEEE, 2019.

[85] S. Wollenstein-Betech, M. Salazar, A. Houshmand, M. Pavone, I. C. Paschalidis, and C. G. Cassandras, *Routing and rebalancing intermodal autonomous mobility-on-demand systems in mixed traffic*, *IEEE Transactions on Intelligent Transportation Systems* (2021).

[86] S. Amin, P. Jaillet, and M. Wu, *Efficient carpooling and toll pricing for autonomous transportation*, *arXiv preprint arXiv:2102.09132* (2021).

[87] Y. Zhu and K. Savla, *On routing drivers through persuasion in the long run*, in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 4091–4096, IEEE, 2019.

[88] M. Wu, S. Amin, and A. E. Ozdaglar, *Value of information in bayesian routing games*, *Operations Research* **69** (2021), no. 1 148–163.

[89] P. N. Brown and J. R. Marden, *Studies on robust social influence mechanisms: Incentives for efficient network routing in uncertain settings*, *IEEE Control Systems Magazine* **37** (2017), no. 1 98–115.

[90] W. H. Sandholm, *Negative externalities and evolutionary implementation*, *The Review of Economic Studies* **72** (2005), no. 3 885–915.

[91] E. Bıyık, D. A. Lazar, R. Pedarsani, and D. Sadigh, *Altruistic autonomy: Beating congestion on shared roads*, in *Workshop on the Algorithmic Foundations of Robotics*, 2018.

[92] E. Bıyık, D. A. Lazar, D. Sadigh, and R. Pedarsani, *The green choice: Learning and influencing human decisions on shared roads*, in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 347–354, IEEE, 2019.

[93] E. Bıyık, D. A. Lazar, R. Pedarsani, and D. Sadigh, *Incentivizing efficient equilibria in traffic networks with mixed autonomy*, *Accepted, IEEE Transactions on Control of Network Systems* (2021).

[94] E. Frazzoli, *notonomy's vision for autonomous driving*, *MIT 6.S094: Deep Learning for Self-Driving Cars, https://www.youtube.com/watch?v=dWSbItd0HEA* (2018).

[95] K. Korosec, *Aptiv's self-driving cars have given 100,000 paid rides on the lyft app*, *Tech Chrunch, https://techcrunch.com/2020/02/11/aptivs-self-driving-cars-have-given-100000-paid-rides-on-the-lyft-app/.*

[96] I. Boudway, *Waymo's self-driving future looks real now that the hype is fading*, *Bloomberg Businessweek, https://www.bloomberg.com/news/articles/2021-01-21/waymo-self-driving-taxis-are-coming-to-more-u-s-cities.*

[97] M. E. Ben-Akiva, S. R. Lerman, and S. R. Lerman, *Discrete choice analysis: theory and application to travel demand*, vol. 9. MIT press, 1985.

[98] D. Sadigh, A. D. Dragan, S. Sastry, and S. A. Seshia, *Active preference-based learning of reward functions*, in *Robotics: Science and Systems (RSS)*, 2017.

[99] E. Biyik and D. Sadigh, *Batch active preference-based learning of reward functions*, in *Conference on Robot Learning*, pp. 519–528, 2018.

[100] M. Salazar, M. Tsao, I. Aguiar, M. Schiffer, and M. Pavone, *A congestion-aware routing scheme for autonomous mobility-on-demand systems*, in *European Control Conference (ECC)*, 2019.

[101] W. H. Sandholm, *Evolutionary implementation and congestion pricing*, *The Review of Economic Studies* **69** (2002), no. 3 667–689.

[102] N. D. Daw, J. P. O'doherty, P. Dayan, B. Seymour, and R. J. Dolan, *Cortical substrates for exploratory decisions in humans*, *Nature* (2006).

[103] P. Koster, E. Verhoef, S. Shepherd, and D. Watling, *Preference heterogeneity and congestion pricing: The two route case revisited*, *Transportation Research Part B: Methodological* **117** (2018) 137–157.

[104] A. Sumalee and W. Xu, *First-best marginal cost toll for a traffic network with stochastic demand*, *Transportation Research Part B: Methodological* **45** (2011), no. 1 41–59.

[105] A. Y. Ng, S. J. Russell, *et. al.*, *Algorithms for inverse reinforcement learning.*, in *International Conference on Machine Learning (ICML)*, 2000.

[106] P. Abbeel and A. Y. Ng, *Exploration and apprenticeship learning in reinforcement learning*, in *International Conference on Machine Learning (ICML)*, ACM, 2005.

[107] B. D. Ziebart, A. L. Maas, J. A. Bagnell, and A. K. Dey, *Maximum entropy inverse reinforcement learning.*, in *AAAI*, 2008.

[108] W. Cheng, J. Fürnkranz, E. Hüllermeier, and S.-H. Park, *Preference-based policy iteration: Leveraging preference learning for reinforcement learning*, in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 312–327, Springer, 2011.

[109] N. Kallus and M. Udell, *Revealed preference at scale: Learning personalized preferences from assortment choices*, in *Proceedings of the 2016 ACM Conference on Economics and Computation*, pp. 821–837, 2016.

[110] M. Zadimoghaddam and A. Roth, *Efficiently learning from revealed preference*, in *International Workshop on Internet and Network Economics*, pp. 114–127, Springer, 2012.

[111] C. C. Gonzaga, *Path-following methods for linear programming*, *SIAM review* (1992).

[112] R. A. Waltz, J. L. Morales, J. Nocedal, and D. Orban, *An interior algorithm for nonlinear optimization that combines line search and trust region steps*, *Mathematical programming* (2006).

[113] G. Andrew and J. Gao, *Scalable training of l 1-regularized log-linear models*, in *International Conference on Machine Learning (ICML)*, ACM, 2007.

[114] E. Biyik, N. Huynh, M. J. Kochenderfer, and D. Sadigh, *Active preference-based gaussian process regression for reward learning*, in *Proceedings of Robotics: Science and Systems (RSS)*, July, 2020.

[115] E. Biyik, M. Palan, N. C. Landolfi, D. P. Losey, and D. Sadigh, *Asking easy questions: A user-friendly approach to active reward learning*, in *Conference on Robot Learning (CoRL)*, 2019.

[116] M. Beliaev, E. Biyik, D. A. Lazar, W. Z. Wang, D. Sadigh, and R. Pedarsani, *Incentivizing routing choices for safe and efficient transportation in the face of the covid-19 pandemic*, in *12th ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*, May, 2021.

[117] P. N. Brown and J. R. Marden, *Fundamental limits of locally-computed incentives in network routing*, in *IEEE American Control Conference (ACC)*, 2017.

[118] R. D. Luce, *Individual choice behavior: A theoretical analysis*. Courier Corporation, 2012.

[119] D. A. Lazar, E. Bıyık, D. Sadigh, and R. Pedarsani, *Learning how to dynamically route autonomous vehicles on shared roads*, *Accepted, IEEE Transactions on Control of Network Systems* (2021).

[120] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, *Proximal policy optimization algorithms*, *Preprint, arXiv:1707.06347* (2017).

[121] C. Swamy, *The effectiveness of stackelberg strategies and tolls for network congestion games*, *ACM Transactions on Algorithms (TALG)* (2012).

[122] W. Krichene, M. S. Castillo, and A. Bayen, *On social optimal routing under selfish learning*, *IEEE Transactions on Control of Network Systems* (2018).

[123] A. Muralidharan, G. Dervisoglu, and R. Horowitz, *Freeway traffic flow simulation using the link node cell transmission model*, in *IEEE American Control Conference*, 2009.

[124] A. Aswani and C. Tomlin, *Game-theoretic routing of gps-assisted vehicles for energy efficiency*, in *IEEE American Control Conference*, 2011.

[125] E. Vinitsky, A. Kreidieh, L. Le Flem, N. Kheterpal, K. Jang, F. Wu, R. Liaw, E. Liang, and A. M. Bayen, *Benchmarks for reinforcement learning in mixed-autonomy traffic*, in *Conference on Robot Learning*, pp. 399–409, 2018.

[126] M. A. Wright, S. F. Ehlers, and R. Horowitz, *Neural-attention-based deep learning architectures for modeling traffic dynamics on lane graphs*, *arXiv preprint arXiv:1904.08831* (2019).

[127] X. Di and R. Shi, *A survey on autonomous vehicle control in the era of mixed-autonomy: From physics-based to ai-guided driving policy learning*, *arXiv preprint arXiv:2007.05156* (2020).

[128] C. Mao and Z. Shen, *A reinforcement learning framework for the adaptive routing problem in stochastic time-dependent network*, *Transportation Research Part C: Emerging Technologies* **93** (2018) 179–197.

[129] A. L. Bazzan and R. Grunitzki, *A multiagent reinforcement learning approach to en-route trip building*, in *2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 5288–5295, IEEE, 2016.

[130] B. Zhou, Q. Song, Z. Zhao, and T. Liu, *A reinforcement learning scheme for the equilibrium of the in-vehicle route choice problem based on congestion game*, *Applied Mathematics and Computation* **371** (2020) 124895.

[131] G. d. O. Ramos, A. L. Bazzan, and B. C. da Silva, *Analysing the impact of travel information for minimising the regret of route choice*, *Transportation Research Part C: Emerging Technologies* **88** (2018) 257–271.

[132] F. Stefanello, B. C. da Silva, and A. L. Bazzan, *Using topological statistics to bias and accelerate route choice: preliminary findings in synthetic and real-world road networks*, in *ATT@ IJCAI*, 2016.

[133] R. Grunitzki, G. de Oliveira Ramos, and A. L. C. Bazzan, *Individual versus difference rewards on reinforcement learning for route choice*, in *2014 Brazilian Conference on Intelligent Systems*, pp. 253–258, IEEE, 2014.

[134] Z. Shou and X. Di, *Multi-agent reinforcement learning for dynamic routing games: A unified paradigm*, *arXiv preprint arXiv:2011.10915* (2020).

[135] H. S. Mahmassani, *50th anniversary invited article—autonomous vehicles and connected vehicle systems: Flow and operations considerations*, *Transportation Science* (2016).

[136] M. Mehrabipour, L. Hajibabai, and A. Hajbabaie, *A decomposition scheme for parallelization of system optimal dynamic traffic assignment on urban networks with multiple origins and destinations*, *Computer-Aided Civil and Infrastructure Engineering* **34** (2019), no. 10 915–931.

[137] C. F. Daganzo, *The cell transmission model, part ii: network traffic*, *Transportation Research Part B: Methodological* (1995).

[138] S. D. Boyles, N. E. Lownes, and A. Unnikrishnan, *Transportation Network Analysis*, vol. 1. 0.85 ed., 2020.

[139] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT press, 2010.

[140] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*. Cambridge university press, 2006.

[141] J. R. Marden and J. S. Shamma, *Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation*, *Games and Economic Behavior* (2012).

[142] L. E. Blume, *The statistical mechanics of strategic interaction*, *Games and economic behavior* (1993).

[143] W. Krichene, M. C. Bourguiba, K. Tlam, and A. Bayen, *On learning how players learn: Estimation of learning dynamics in the routing game*, *Transactions on Cyber-Physical Systems* (2018).

[144] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, *Trust region policy optimization*, in *International Conference on Machine Learning*, 2015.

[145] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, *Asynchronous methods for deep reinforcement learning*, in *International conference on machine learning*, 2016.

[146] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, *Preprint, arXiv:1412.6980* (2014).

[147] H. TranStar, *2017 annual report*, report, Houston TranStar, 2018. Accessed: January 28, 2019.

[148] G. d. O. Ramos and A. L. C. Bazzan, *Towards the user equilibrium in traffic assignment using grasp with path relinking*, in *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pp. 473–480, 2015.

[149] A. L. Bazzan, D. Cagara, and B. Scheuermann, *An evolutionary approach to traffic assignment*, in *2014 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS)*, pp. 43–50, IEEE, 2014.

[150] D. A. Lazar, K. Chandrasekher, R. Pedarsani, and D. Sadigh, *Maximizing road capacity using cars that influence people*, in *IEEE Conference on Decision and Control (CDC)*, 2018.

[151] *Self-driving uber car kills arizona pedestrian*, 2018.

[152] D. Sadigh, S. S. Sastry, S. A. Seshia, and A. Dragan, *Information gathering actions over human internal state*, in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),*, pp. 66–73, IEEE, 2016.

[153] D. Sadigh, S. Sastry, S. A. Seshia, and A. D. Dragan, *Planning for autonomous cars that leverage effects on human actions.*, in *Robotics: Science and Systems*, 2016.

[154] D. Sadigh, *Safe and Interactive Autonomy: Control, Learning, and Verification.* PhD thesis, EECS Department, University of California, Berkeley, Aug, 2017.

[155] A. Gray, Y. Gao, J. K. Hedrick, and F. Borrelli, *Robust predictive control for semi-autonomous vehicles with an uncertain driver model*, in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, pp. 208–213, IEEE, 2013.

[156] V. Raman, A. Donzé, D. Sadigh, R. M. Murray, and S. A. Seshia, *Reactive synthesis from signal temporal logic specifications*, in *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*, pp. 239–248, ACM, 2015.

[157] M. P. Vitus and C. J. Tomlin, *A probabilistic approach to planning and control in autonomous urban driving*, in *2013 IEEE 52nd Annual Conference on Decision and Control (CDC)*, pp. 2459–2464.

[158] C. Hermes, C. Wohler, K. Schenk, and F. Kummert, *Long-term vehicle motion prediction*, in *2009 IEEE Intelligent Vehicles Symposium*, pp. 652–657, 2009.

[159] R. Vasudevan, V. Shia, Y. Gao, R. Cervera-Navarro, R. Bajcsy, and F. Borrelli, *Safe semi-autonomous control with enhanced driver modeling*, in *American Control Conference (ACC), 2012*, pp. 2896–2903, IEEE, 2012.

[160] J. Vander Werf, S. Shladover, M. Miller, and N. Kourjanskaia, *Effects of adaptive cruise control systems on highway traffic flow capacity*, *Transportation Research Record: Journal of the Transportation Research Board* (2002), no. 1800 78–84.

[161] B. Van Arem, C. J. Van Driel, and R. Visser, *The impact of cooperative adaptive cruise control on traffic-flow characteristics*, *IEEE Transactions on Intelligent Transportation Systems* **7** (2006), no. 4 429–436.

[162] G. M. Arnaout and S. Bowling, *A progressive deployment strategy for cooperative adaptive cruise control to improve traffic dynamics*, *International Journal of Automation and Computing* **11** (2014), no. 1 10–18.

[163] J. Lee and B. Park, *Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment*, *IEEE Transactions on Intelligent Transportation Systems* **13** (2012), no. 1 81–90.

[164] E. Dallal, A. Colombo, D. Del Vecchio, and S. Lafortune, *Supervisory control for collision avoidance in vehicular networks with imperfect measurements*, in *2013 IEEE 52nd Annual Conference on Decision and Control (CDC)*, pp. 6298–6303, IEEE, 2013.

[165] D. Miculescu and S. Karaman, *Polling-systems-based control of high-performance provably-safe autonomous intersections*, in *53rd IEEE Conference on Decision and Control*, pp. 1417–1423, IEEE, 2014.

[166] A. Colombo and D. Del Vecchio, *Least restrictive supervisors for intersection collision avoidance: A scheduling approach*, *IEEE Transactions on Automatic Control* **60** (2015), no. 6 1515–1527.

[167] P. Tallapragada and J. Cortés, *Coordinated intersection traffic management*, *IFAC-PapersOnLine* **48** (2015), no. 22 233–239.

[168] Y. J. Zhang, A. A. Malikopoulos, and C. G. Cassandras, *Optimal control and coordination of connected and automated vehicles at urban traffic intersections*, in *2016 American Control Conference (ACC)*, pp. 6227–6232, July, 2016.

[169] D. Braess, *Über ein paradoxon aus der verkehrsplanung*, *Unternehmensforschung* **12** (1968), no. 1 258–268.

[170] D. A. Lazar, K. Chandrasekher, R. Pedarsani, and D. Sadigh, *Maximizing road capacity using cars that influence people*, in *Conference on Decision and Control (CDC)*, pp. 1801–1808, IEEE, 2018. `https://web.stanford.edu/~kabirc/files/CapacityInfluenceLazChaPedSad18.pdf`.

[171] P. Abbeel and A. Y. Ng, *Apprenticeship learning via inverse reinforcement learning*, in *International Conference on Machine learning (ICML)*, ACM, 2004.

[172] S. Levine and V. Koltun, *Continuous inverse optimal control with locally optimal examples*, *arXiv preprint arXiv:1206.4617* (2012).

[173] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, and Y. Bengio, *Theano: a CPU and GPU math expression compiler*, in *Proceedings of the Python for Scientific Computing Conference (SciPy)*, June, 2010. Oral Presentation.

[174] F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. J. Goodfellow, A. Bergeron, N. Bouchard, and Y. Bengio, "Theano: new features and speed improvements." Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop, 2012.

[175] B. S. Kerner, *Criticism of generally accepted fundamentals and methodologies of traffic and transportation theory: A brief review*, *Physica A: Statistical Mechanics and its Applications* **392** (2013), no. 21 5261–5282.