Modeling Travel Time Uncertainty in Traffic Networks

by

Daizhuo Chen

B.S. Physics, Peking University (2008)

Submitted to the School of Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Computation for Design and Optimization

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2010

Author . . . .
School of Engineering
August 5, 2010

Certified by . . . .
Georgia Perakis
William F. Pounds Professor of Operations Research
Thesis Supervisor

Accepted by . . . .
Karen Willcox
Associate Professor of Aeronautics and Astronautics
Codirector, Computation for Design and Optimization Program

Modeling Travel Time Uncertainty in Traffic Networks

by

Daizhuo Chen

Submitted to the School of Engineering
on August 5, 2010, in partial fulfillment of the
requirements for the degree of
Master of Science in Computation for Design and Optimization

Abstract

Uncertainty in travel time is one of the key factors that could allow us to understand and manage congestion in transportation networks. Models that incorporate uncertainty in travel time need to specify two mechanisms: the mechanism through which travel time uncertainty is generated and the mechanism through which travel time uncertainty influences users' behavior. Existing traffic equilibrium models are not sufficient in capturing these two mechanisms in an integrated way.

This thesis proposes a new stochastic traffic equilibrium model that incorporates travel time uncertainty in an integrated manner. We focus on how uncertainty in travel time induces uncertainty in the traffic flow and vice versa. Travelers independently make probabilistic path choice decisions, inducing stochastic traffic flows in the network, which in turn result in uncertain travel times. Our model, based on the distribution of the travel time, uses the mean-variance approach in order to evaluate travelers' travel times and subsequently induce a stochastic traffic equilibrium flow pattern.

In this thesis, we also examine when the new model we present has a solution as well as when the solution is unique. We discuss algorithms for solving this new model, and compare the model with existing traffic equilibrium models in the literature. We find that existing models tend to overestimate traffic flows on links with high travel time variance-to-mean ratios.

To benchmark the various traffic network equilibrium models in the literature relative to the model we introduce, we investigate the total system cost, namely the total travel time in the network, for all these models. We prove three bounds that allow us to compare the system cost for the new model relative to existing models. We discuss the tightness of these bounds but also test them through numerical experimentation on test networks.

Thesis Supervisor: Georgia Perakis
Title: William F. Pounds Professor of Operations Research

Acknowledgments

I am very fortunate, happy and honored to have had the opportunity to work with my advisors, Professor Georgia Perakis and Professor Retsef Levi, over the past year. They have taught me how to do research. Their guidance, patience, and invaluable insights are key in preparing me to achieve my goals.

I would like express my heartfelt gratitude to Georgia for her expert guidance, patience, and optimism. I have benefited greatly from her invaluable knowledge and piercing insights. She is also my technical writing teacher. I would like to thank her for the great amount of time and effort she put to supervise my work. Without her pushing me doing the best, this thesis wouldn't have been possible.

I am deeply grateful to Retsef. He has been my role model for his rigorous thinking and enthusiasm in research. His creative ideas and insightful comments are always a great source of inspiration.

Thanks to the Singapore-MIT Alliance for the graduate scholarship. It brings me such a wonderful opportunity. Thanks to Laura Koller for her sincere help in the Computation for Design and Optimization (CDO) program.

Thanks to all my friends who have made my life at MIT a joyful experience. I would like to thank Thai Dung Nguyen and Gonzalo Romero for many cheerful and inspiring discussions on our research. Thanks to Gil Kwak and Wei Wang for their taking the initiative of building a cordial friendship. Thanks to Bonnie Vierthaler, Daniel Enking, John Cumbers, Shibo Song, and Chunbo Lou for their warmth and hospitality. Cheers to my three flatmates, Di Yang, Linjiang Cai, and Renyu Hu, and many visitors of our dorm throughout the past year, for all the delicious food together we have made and tasted. Thanks to many old and new friends of mine at Tufts University. Thanks to my friends at the Operations Research Center, from whom I have learned so much. And a special thanks to all my CDO fellow students. I enjoyed so much being one of us.

I am also grateful to many friends who are outside Cambridge. Thanks to Hantao Yin, Yifan Yang, and Li An for many wonderful academic discussions we had. Thanks to Xiaogang Liu for his most sincere help that allows me to focus on my research during difficult times. Thanks to many others who have helped me in one way or another. I would like to thank Wuyan for her cheerful support, patience, and trust throughout the past year.

Last but not least, I am deeply indebted to my family. Their unconditional love, support, and care allow me to move forward with my full potentials.

Contents

Chapter 1

Introduction

One of the most serious problems in the contemporary transportation system is traffic congestion. In the year of 2007 alone, it is estimated that congestion is responsible for 4.2 billion hours of travel delay and causes travelers to consume 2.8 billion gallons of extra fuel that in turn cost the U.S. economy $87 billion (Schrank and Lomax, 2009). Despite the fact that considerable efforts have been devoted to alleviate congestion (for example, by increasing road capacity, enhancing the infrastructure planning, utilizing Advanced Traveler Information System, and finally, through congestion pricing), so far relatively little attention has been directed to addressing travel time uncertainty.

Travel time uncertainty is often a direct consequence of traffic congestion. In the presence of congestion, the unpredictability of travel time causes additional frustration to travelers. It is common for travelers to dislike travel time uncertainty just as much as traffic congestion itself (Liu et al., 2004), and consequently to adjust their behavior to cope with the unpredictable travel time that they will experience (for a review, see Noland and Polak, 2002). Furthermore, data shows that the utility loss due to travel time uncertainty is of the same order of magnitude as the total travel costs (see, for example, Hallenbeck et al. 2003, Skabardonis et al. 2003, and Marchal and de Palma 2008). This observation strongly suggests that it is important to model and understand the implications of travel time uncertainty on transportation systems.

This thesis considers travel time uncertainty in traffic networks. It is well-known that vehicles running on roads experience uncertain travel times, and traffic flows on roads fluctuate from day to day. Roads constitute links in the traffic network of highways, major and minor roads, and junctions. Traffic flow theory predicts that traffic flows influence travel times, and transportation network models often reasonably assume that, the larger the traffic flow on a link is, the longer the time it takes to traverse that link. This thesis specifically focuses on how to model and handle the uncertainty of travel times resulting from the uncertainty in traffic flows and in turn the influence that travel time uncertainty has on travelers' path choice decisions.

Although it is not the intent of this thesis to model all sources of travel time uncertainty (e.g., accidents, weather conditions), the goal of this thesis is to consider the sources and influences of travel time uncertainty in an integrated manner. Previous models on this topic appear to make little effort to consider travel time uncertainty in an integrated way. This thesis attempts to make the missing connection.

1.1 Motivations

This research has been motivated by two main issues. First, at a general level, recent developments in the study of transportation networks have shown that travel time uncertainty is an important factor that influences traffic networks (see, for example, Liu et al. 2004, Brilon et al. 2005, and Marchal and de Palma 2008). Despite this realization, the understanding of travel time uncertainty in the literature is still very limited. Second and more specifically, most existing research has focused on either the sources or the influences of travel time uncertainty. As a result, our goal is to develop an integrated model that incorporates both aspects.

1.1.1 Travel time uncertainty

Researchers have built various mathematical models of traffic networks since the 1920s[1]. These models serve various purposes that range from infrastructure planning, network performance evaluation, and vehicle fleet operations management, to congestion pricing.

Traditionally most traffic network models assume that traffic network statistics (for example, traffic flows and travel times) are deterministic, largely because deterministic models are theoretically and computationally tractable and produce useful predictions for the purpose of study. Nevertheless, in reality traffic networks are subject to random fluctuations that are not foreseeable neither by the modeler nor by the traveler. These random fluctuations may result, for example, from the random waiting time at road junctions, traffic accidents, bad weather, or the fluctuations in the number of total travelers. The random fluctuations lead to travel time uncertainty in the traffic network. Deterministic models do not capture this uncertainty, and hence they are insufficient in describing the impacts of these uncertainties on the traffic network.

However, uncertainty plays an important role in traffic networks. Intuitively, uncertain travel times often cause frustration to travelers, and as a result travelers adjust their travel behavior in response to travel time uncertainty. For example, travelers catching a train or plane may be very concerned about the uncertainty of the travel time they might experience. As a result, they react by adjusting their

[1] Pigou (1920) considered a two-node, two-link transportation network. His idea was further developed by Knight (1924).

departure times or path choices.

Transportation studies have also confirmed the important role that travel time uncertainty plays in traffic networks. Liu et al. (2004) found that travelers value a reduction in travel time variability more highly than a corresponding reduction in the overall travel time for that journey. Marchal and de Palma (2008) concluded that the utility loss due to uncertainty is of the same order of magnitude as the total travel costs. Brilon et al. (2005) commented that deterministic traffic network models are insufficient for traffic network performance assessment because they overemphasize small differences in travel times, but in the meantime, the significant influences that travel time uncertainty will cause are not adequately represented. Travel time uncertainty has also gained considerable attention from public authorities in recent years. See, for example, the reports of Cambridge Systematics et al. (2000) and of the U.S. Federal Highway Administration (2007).

Accounting for the influences of travel time uncertainty in traffic networks calls for an explicit approach to modeling travel time uncertainty. The incorporation of travel time uncertainty will not only improve traffic network modeling, but also enhance our general understanding of uncertainty and its influence on human decisions.

1.1.2 Integration of sources and influences

There are two crucial steps in order to incorporate travel time uncertainty in traffic networks: 1) identifying the sources of travel time uncertainty, so that travel time uncertainty can be properly represented in a traffic network, and 2) determining the influences of travel time uncertainty on travelers' decisions.

Existing models have identified different sources of travel time uncertainty and adopted different approaches to representing travel time uncertainty in traffic networks (see Section 1.2.1). Many approaches have also been used to model the influence of travel time uncertainty on travelers' behavior (see Section 1.2.2; for a review, see also Noland and Polak, 2002). Despite considerable progress in both aspects, we are not aware of models that incorporate both the sources and the influences of travel time uncertainty and treat them in an integrated manner.

1.2 Literature Review

As stated in the previous section, this thesis contributes to enhancing the literature of existing traffic network models in their ability to capture the interactions between network travel time uncertainty and travelers' behavior. This section contains some related background material.

1.2.1 What causes travel time uncertainty

Conventional traffic network models are concerned with predictions of link flows and travel times on a traffic network, under the assumptions that 1) the link travel times are dependent on the link traffic flows, and 2) the travelers' path choice decisions are dependent on path travel times. The deterministic user equilibrium (DUE, also called Wardrop equilibrium) and the stochastic user equilibrium (SUE) are by far the most well studied traffic equilibrium models. The DUE model was first proposed by Wardrop (1952); it assumes that travelers only travel on paths with minimum path travel times. Daganzo and Sheffi (1977) introduced the SUE model, which assumes that travelers have random perception errors of their travel times, and therefore choose paths that minimize their perceived travel times. For both the DUE and SUE models, the word "user" emphasizes the noncooperative nature of the individual travelers. These models are contrasted with the system optimum (SO) model, which assumes that flow patterns are optimized by a central planner in order to minimize the aggregate travel times of all travelers. The monographs of Sheffi

(1984), Ben-Akiva and Lerman (1985) and Cascetta (2001) extensively investigate the theoretical and practical aspects of these models. For a review of the DUE model, see also Florian and Hearn (1995), Nagurney (1999), and Correa and Stier-Moses (2010).

An important feature of the DUE and SUE models is that traffic flows and (actual) travel times are assumed to be deterministic.[2] As a result, both models are referred to as deterministic network models.

However, it is well-known that in reality both traffic flows and travel times are stochastic and vary from day to day. Traffic flows are stochastic because travelers make unpredictable travel decisions. This unpredictability comes from both the modeler's lack of understanding of the travelers' exact decision mechanisms and the modeler's inability to capture all the factors that influence travelers' decisions. The travel time uncertainty also results from two sources. On the demand side, uncertain traffic flows cause uncertain travel times (because the link travel times are dependent on the link traffic flows). On the supply side, stochastic events, for example, waiting times at traffic signals, traffic accidents, and unexpected closure of roads, will result in uncertain travel times.

Previous traffic network models have developed multiple approaches to modeling travel time uncertainty, as presented in the following.

The first approach focuses on the travel time uncertainty resulting from uncertain traffic flows. In this approach, despite the assumption that the link travel times deterministically depend on the link traffic flows, the uncertain traffic flows lead to uncertain travel times. Models that adopt this approach include Watling (2002) and Clark and Watling (2005).

[2]The word "stochastic" in the name of the SUE model emphasizes the model's assumption of stochastic perception errors. However, the SUE model interprets the probability that a traveler chooses a certain path as the proportion of travelers of the same trip choosing that path. Hence this model results in deterministic predictions of traffic flows and (actual) travel times.

The second approach focuses on the travel time uncertainty resulting from the uncertain link capacities due to unexpected events. Models that adopted this approach generally assume that unexpected events such as traffic accidents and bad weather will temporarily reduce the link capacity in a stochastic manner. As a result, despite the assumption that the traffic flows remain deterministic, the link travel times are stochastic. Modelers can adjust the various parameters including the probability, severity (that is, how much capacity is affected), and duration of unexpected events to represent networks of different uncertainty levels. This approach is adopted, for example, by Noland et al. (1998), Brilon et al. (2005), and Marchal and de Palma (2008).

The third approach is to explicitly specify the distribution of link travel times for all links in the network. In this approach, the travel time on each link is composed of a deterministic component and a stochastic component. While the deterministic component depends on the link traffic flow, the stochastic component is generally assumed to be a random variable independent of the link traffic flow. This approach is adopted by, for example, Mirchandani and Soroush (1987), Ordonez and Stier-Moses (2007), and Connors and Sumalee (2009). This approach is particularly suitable for modeling uncertain travel times as a result of random waiting time at traffic signals.

Finally, travel time uncertainty can also be represented using a microscopic simulation model. This last approach differs in that it models the movement of individual cars in real time. Studies that follow this approach include Schadschneider (2000).

The different approaches described above are not mutually exclusive. In fact it could be argued that each of the first three approaches explains a component of the travel time uncertainty experienced by travelers. It remains an open research problem to develop a traffic model that unifies all these approaches.

1.2.2 How travelers respond to travel time uncertainty

An essential component of traffic network models is traffic demand modeling. The idea is to study travelers' behavior, namely, how many travelers will be traveling from one place to another at a given time (the trip generation problem) and what paths these travelers will choose (the path choice problem). For simplicity and clarity, in this thesis, fixed trip demands are assumed, i.e., we do not model demand stochasticity. We are mainly concerned with the path choice problem under the influence of travel time uncertainty.

The travelers' behavior is governed by the (dis)utility associated with each of the available paths. In the transportation literature, travel (dis)utility is also referred to as travel cost. These two terms will be used interchangeably in this thesis. In general, path choice models involve two steps. In the first step, the travel costs of all available paths are evaluated. In the second step, travelers make path choice decisions that minimize their path travel costs.

The concept of the travel cost has evolved over time and remains an area of active academic research. Conventional traffic network models assume that the (deterministic) path travel time is the only component of the travel cost, as is most obviously reflected in Wardrop's first principle (see Wardrop, 1952): "The journey times in all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route." Since then modelers have included other factors (for example, the cost of fuel, wear and tear, and toll or congestion charges) into the travel cost. McFadden (1974) introduced the concept of random utility theory in traffic demand modeling. This theory further justifies the use of demographic data as travel cost factors. It interprets the travel cost as a variable determined by observable economic factors, which include the travel time, monetary costs and travelers' social economic status, and unobservable factors, such as road quality or scenery, the need to make an extra mid-path stop, etc. The aggregate effect of the unobservable factors on the travel cost is modeled as a random variable that follows a known distribution. Recently, Zeid (2009) has enriched the traffic demand models by incorporating travelers' subjective well-being as a component of the travel cost.

For simplicity, we assume that travel costs are decided by travel times alone. Travel time uncertainty generally has a negative influence on travel costs. There are many different ways of incorporating travel time uncertainty into travel costs. Next we will discuss these approaches.

The first approach measures the travel cost as the sum of the mean travel time and the variance or the standard deviation of travel time multiplied by a trade-off factor. This approach is adopted by a number of empirical studies. For example, Jackson and Jucker (1982) studied travelers' route-to-work choices using the mean-variance approach. Senna (1994) and Nam et al. (2005) estimated the value of travel time variability using the mean-variance approach. These empirical studies developed the methodologies of measuring the trade-off factor of a population through techniques such as stated preferences. Nikolova et al. (2006) studied the optimum routing problem in stochastic traffic networks. In a special case of their model the travel cost is a linear combination of the mean and variance of path travel time. However, these models do not address how the individual travelers' routing decisions influence the traffic network.

The second approach utilizes the expected utility (EU) theory. Under this theory, the traveler uses a continuous utility function to transform travel times into travel (dis)utilities. The expected value of the travel (dis)utility is then taken as the travel cost. For example, when the utility function is linear, the travel cost equals the expected travel time (multiplied by a constant). When the utility function assigns very large (dis)utility values for long travel times, travelers become risk averse. This approach is adopted by Mirchandani and Soroush (1987) and Senna (1994).

The third approach utilizes the principles of prospect theory. Due to the drawbacks of the EU theory (for example, the EU theory has limited representation of the behavior process and violates perception rationality; for a review, see Tversky and Kahneman, 1992), various non-expected utility theories have also been developed as alternatives to the expected utility theory. One of the most popular among them is prospect theory (see Kahneman and Tversky, 1979). This theory, in addition to modeling travel times using a utility function, uses a probability weighting function that underweights large and overweights small probabilities for travel cost evaluation. Avineri and Prashker (2005) and Connors and Sumalee (2009) recently applied this prospect theory approach to traffic network modeling.

The fourth approach is the percentile approach. This approach typically measures the travel cost as the 90th or 95th percentile of travel time (see the report of the U.S. Department of Transportation, 2007). Because percentile measures are easy to communicate, this approach has received a lot of attention among transportation authorities recently. However, although the percentile approach is often used by transportation planners as a measure of network reliability, it is less understood whether this approach provides a good measure of the travel cost that governs travelers' behavior. Moreover, the calculation of travel time percentiles usually involves convolutions of probability density functions and may pose computational challenges to traffic network modelers.

Lastly, the fifth approach is the robust optimization approach adopted by Ordonez and Stier-Moses (2007). This approach assumes that all users are risk-averse and minimize their worst case travel time. As a result, in this sense users make "robust" decisions. The robust optimization approach was initially considered by Soyster (1973). However, Soyster's approach was deemed to be too conservative at his time (see Ben-Tal and Nemirovski, 2000). A significant step forward for developing a theory for robust optimization was taken independently by Ben-Tal and Nemirovski (1998) and El Ghaoui et al. (1998). In traffic network modeling, instead of specifying a distribution function for the travel time as a random variable, the robust approach defines an uncertainty set for the variations of travel times. Travelers measure their travel costs as the worst case travel times, given that the realization of travel times comes from the uncertainty set. Bertsimas and Sim (2003) developed a robust optimization network flow model. This distribution-free model provides another conceptually and computationally attractive alternative to the previously mentioned approaches. Nevertheless, due to its considering a worst case approach, some may criticize it as too conservative.

1.2.3 Stochastic traffic equilibrium models

In this last part of the literature review section we review four existing traffic equilibrium models that incorporate travel time uncertainty. Because these models represent travel times as random variables, we refer to them as *stochastic network equilibrium models.* These models contrast with the deterministic traffic equilibrium models (see the DUE and the SUE model) which assume deterministic travel times.

Mirchandani and Soroush (1987) developed one of the earliest stochastic traffic network equilibrium models. To represent travel time stochasticity, they exogenously specify the distributions of link travel times. Users measure their travel costs using the expected utility approach. Although users have heterogeneous perception errors for travel costs, the resulting traffic flows are deterministic (similar to the SUE model, this model interprets the probability of a randomly picked traveler choosing a certain path as the proportion of travelers choosing that path).

Watling (2002) considered another stochastic traffic network equilibrium model and named it the generalized stochastic user equilibrium (GSUE). The GSUE model extends the SUE model by assuming that users make probabilistic path choice decisions according to the SUE path choice probabilities. In other words, unlike the SUE model, the GSUE model does not interpret the path choice probabilities as path choice proportions. As a result, the traffic flows in the GSUE model are stochastic, inducing stochastic travel times. However, the GSUE model assumes that the path travel cost is the expected path travel time, and thus the model is insufficient in modeling the influences of travel time uncertainty on travelers' behavior. Because the traffic flows in the GSUE model are stochastic, we call the GSUE model a generalized stochastic traffic equilibrium model.

Ordonez and Stier-Moses (2007) developed a robust user equilibrium (RUE) model. They assume that travelers measure travel costs as the worst case travel time, and travelers further choose paths that minimize the travel costs (no perception error is present; similar to the assumption of the DUE model). The uncertainty set of travel time variations is exogenously specified and does not depend on traffic flows. The resulting equilibrium traffic flows are deterministic. Nevertheless, this model may be too conservative.

Connors and Sumalee (2009) considered a stochastic traffic network equilibrium model that utilized prospect theory. They exogenously specify the link travel time distributions (similar to Mirchandani and Soroush, 1987), and transform the travel time distributions into travel costs using the cumulative prospect theory. Users have no perception errors, and they choose paths that minimize their travel costs. The resulting equilibrium traffic flows are deterministic.

Among the existing stochastic network equilibrium models, only the GSUE model has assumed stochastic traffic flows, yet it fails to address the problem of how uncertain travel times influence travelers' decisions. The other three traffic equilibrium models require an exogenous specification of travel time distributions, and they cannot incorporate how travelers' decisions under uncertain travel times influence the travel time uncertainty.

To address the deficiencies of the existing stochastic network models, this thesis develops a new stochastic network model in which the travel time uncertainty is induced by traffic flow uncertainty[3] and traffic flow uncertainty is influenced by travel time uncertainty. In this sense the new model incorporates travel time uncertainty in an integrated manner.

1.3 Contributions

The study of travel time uncertainty in traffic networks is recently gaining momentum. Despite the growing literature that examines the influences of travel time uncertainty on travelers' behavior, to the best of our knowledge, no studies provide an integrated model that is concerned with both how travel time uncertainty is generated and how travelers react to it. Existing models are not sufficient in capturing these two aspects in an integrated way and hence cannot reveal the interactions between the two.

This thesis has two main contributions. The first main contribution is the development and testing of a new stochastic traffic network model that accounts for travel time uncertainty in an integrated manner. First, we developed a new stochastic user equilibrium model, presented the conditions under which this model has a unique solution, and discussed algorithms that solve this model. Second, we tested this model on test networks and compared it to previous models existing in the lit-

[3]It is possible to enrich the present model by incorporating other sources of travel time stochasticity.

erature. Lastly, we studied a related day-to-day dynamic traffic network model in order to study the stability of our new model.

The main finding of this new model is that deterministic traffic network models tend to overestimate traffic flows on more congested links and underestimate traffic flows on less congested links.

The second main contribution is that we benchmarked different traffic network equilibrium models in terms of the total system cost. The system cost is defined as the total travel time of all users of the network. We proved three new bounds for the system cost that improved on existing bounds. We studied the tightness of these bounds and presented numerical results of the system cost on test networks.

Finally, this work also contributes in terms of developing theoretically the mean-variance approach in order to model the influence of travel time uncertainty on travelers' behavior. This approach has been used in the literature before, but primarily in empirical studies. To the best of our knowledge, it has not received any attention in traffic equilibrium modeling from a theoretical perspective.

1.4 Thesis Outline

Chapter 2 introduces the traffic user equilibrium model, the assumptions we made, and previously studied traffic equilibrium models. In this chapter we distinguish the deterministic network equilibrium model and the stochastic network equilibrium model. Chapter 3 develops a new model, the Truly Stochastic User Equilibrium (TSUE) model. In this chapter we study the existence and uniqueness of its equilibrium solution, propose two solution algorithms, and present numerical results on the Sioux Falls test network. In Chapter 4, we benchmark various traffic network equilibrium models in terms of the system cost. We establish theoretical bounds for the system cost. Chapter 5 considers a day-to-day dynamic model and how this model converges to the newly proposed equilibrium. Chapter 6 concludes the thesis.

Chapter 2

Traffic Network User Equilibrium Models

This chapter introduces various traffic network user equilibrium models that exist in the literature. These models can be separated into two groups. The first group consists of *deterministic network equilibrium models.* These models assume that the traffic network is deterministic; that is, in the state of equilibrium the traffic flows and travel times are represented by a single point in the state space. The second group consists of *stochastic network equilibrium models.* These models introduce stochasticity in the underlying traffic network, so that travel times (and, for some models, traffic flows) are stochastic. Stochastic network models must specify both how the stochasticity is generated and how the uncertain travel times will influence travelers' behavior.

In general the traffic network consists of two components, traffic supply and traffic demand. An equilibrium is reached when these two components match each other. This relationship is plotted in Figure 2.5. Figures like this will repeat multiple times in this thesis. They will be explained in detail in this chapter.

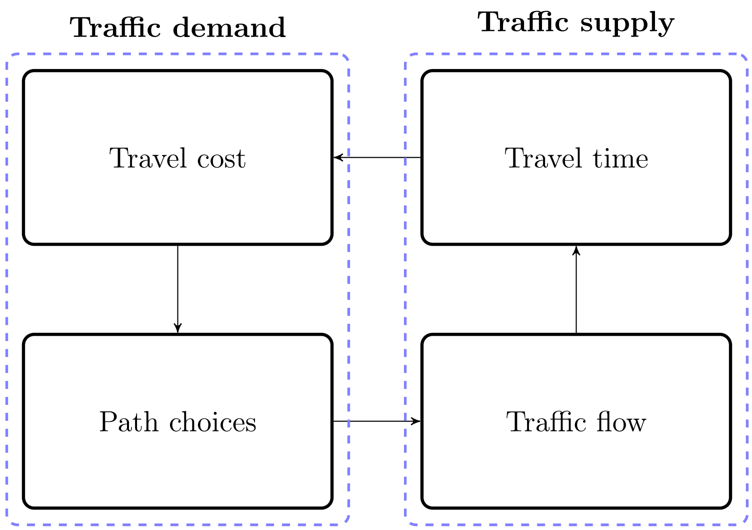


Figure 2.1: Diagram of traffic network user equilibrium models.

This chapter is organized as follows. Section 2.1 defines the traffic network, states the model assumptions, and introduces the concept of user equilibrium. Section 2.2 focuses on deterministic network models. Two models are introduced, namely, the deterministic user equilibrium (DUE, also called Wardrop equilibrium) and the stochastic user equilibrium (SUE). Section 2.3 focuses on the stochastic network models. The robust user equilibrium (RUE) and the cumulative prospect user equilibrium (CPUE) are introduced.

2.1 The Traffic Network

The transportation system at its full scale is an extremely complex system. It is highly dynamic and stochastic, as it operates and evolves in both spacial and temporal dimensions and involves complicated human decisions. It is well-known that attempts to manage traffic networks often lead to unintended consequences, which poses great challenges to modelers.[1]

The traffic network model provides a simplified view of the real transportation system, and it is introduced for the purpose of studying a specific aspect of the transportation system. Traffic network equilibrium models aim to predict the long-term traffic flows and travel times in the network (and any network attributes that can be derived from traffic flows and travel times). The main feature of these models is to balance traffic demand and traffic supply. The mathematical representation, along with the necessary simplifying assumptions, of the traffic network studied in this thesis will be described below.

2.1.1 Representation of the traffic network

The spacial aspect of the transportation system is modeled as a directed graph, which consists of a set of nodes and a set of directed links. Users (travelers) move from node to node through links, and their movements induce traffic flows on links. Each user is associated with an origin node and a destination node, and he travels on one path that connects the origin node to the destination node. A pair of origin and destination nodes is called an OD pair (or a trip). Let N be the set of nodes, L be set of links, and K be the set of all OD pairs. For any OD pair $k \in K$, let R^k be the set of available paths (also called routes) that connects the origin to the destination nodes, and q^k be the travel demand of OD pair k (that is, the number of users traveling on that OD pair in a unit of time). Let R be the set of all paths, thus $R = \cup_{k \in K} R^k$. The link-path relationship is characterized by the link-path incidence matrix $\mathbf{\Delta} \in \{0,1\}^{|L| \times |R|}$, with $\Delta_{lr} = 1$ if link l is included in path r, and $\Delta_{lr} = 0$ otherwise. Let $\mathbf{\Gamma} \in \{0,1\}^{|K| \times |R|}$ be the OD-pair-path incidence matrix, with $\Gamma_{kr} = 1$ if path r connects OD pair k, and $\Gamma_{kr} = 0$ otherwise. Here $|\cdot|$ represents

[1] An interesting example that reveals the complicated feedback loops related to road congestion is discussed in Section 5.6 of Sterman and Sterman (2000).

the cardinality (i.e., the number of elements) of a set.[2]

Users of the same OD pair are assumed to use the same mode of transportation (for example, a car), and share similar characteristics (for example, the RUE model assumes that these users share the same approach towards uncertainty and, as we will see later, the same level of "uncertainty budget"). This assumption does not sacrifice the generality of our model, because modelers can always divide users into homogeneous groups at the expense of increasing the number of OD pairs.

To be specific, in a very simple traffic network model, the links represent roads, and the nodes represent road junctions. However, it should be noted that the transportation system can be modeled at different levels of details. The level of details depends on the purpose of the study.[3]

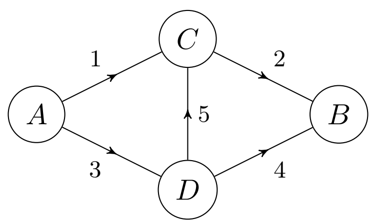


Figure 2.2: A simple traffic network

As an example, for the network in Figure 2.2

- node set $N = \{A,\ B,\ C,\ D\}$
- link set $L = \{1,\ 2,\ 3,\ 4,\ 5\}$
- set of OD pairs $K = \{(A, B)\}$

[2]As a convention in this thesis, we shall always use symbols in bold fonts to indicate vectors and matrices. For the convenience of readers, we have complied a summary of notations in the appendix of this thesis.

[3]For example, in the simple case, a 4-way junction may be modeled as a single node in the traffic network; yet in more complicated applications, it can be modeled as multiple nodes with each lane modeled separately. Readers may refer to Chapter 2 of the monograph of Cascetta (2001) for a complete presentation on this aspect.

- travel demand $q^{(A,D)} = 1000$ veh/hour
- path set $R^{(A,D)} = \{(1,3),\ (2,4),\ (1,5,4)\}$
- link-path incidence matrix

$$\mathbf{\Delta} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

- OD-pair-path incidence matrix

$$\mathbf{\Gamma} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

Now we turn to the temporal aspect of traffic network models. Because traffic networks evolve with time, it is of great importance to identify the time horizon analyzed in a traffic network model. Traffic networks usually exhibit cyclic behavior in the temporal dimension. For example, there is the daily cycle, the weekly cycle, and the annual cycle. The traffic network models in this chapter (and in this thesis) study short time periods (within one cycle) repeated over a long time horizon (covering multiple cycles). The short time period is also called the reference time period. Its typical length varies from a few minutes to a few hours in different studies. For example, it is typical for these models to study the morning rush hour from 7am to 8am on week days, spanning the horizon of a year. Here the morning rush hour is the short time period, and a year is the long time horizon. Such particular temporal settings allow the following approximation that tremendously simplifies our model:

The *intra-period stationarity* approximation. This approximation assumes a stationary state of the network within the short time period, namely the network statistics (traffic flows, link travel times) do not change with time during the short time period.[4] Commonly used by transportation planners, this approximation leads to a tractable traffic network model that provides useful predictions.

A relaxation of the above approximation leads to the development of the (intra-period) dynamic traffic assignment (DTA) models, where traffic flows vary with time. As can be imagined, these models are in general much less tractable compared to the intra-period stationary models, both theoretically and computationally. Nevertheless, the DTA models are of growing research interest over the past years. Interested readers may refer to, for example, the review article of Peeta and Ziliaskopoulos (2001) or the monograph of Friesz (2010).

Depending on whether the stationary traffic flows in the short time period remain the same for different periods, there are *inter-period static* models (stationary flows persist across periods) and *inter-period dynamic* models (stationary flows do not persist across periods; also referred to as day-to-day dynamic models).

Under the intra-period stationarity approximation and the inter-period static assumption, the state of the traffic network reduces to a timeless fixed point state (see Figure 2.3). There are two obvious advantages of studying a fixed point state. On one hand, powerful analytical procedures exist to find the fixed point state, as will be demonstrated later in this thesis. On the other hand, the fixed point model does not need to specify a potentially complicated dynamic mechanism in which the traffic network evolves over time. In applications, it is often implicitly assumed that the traffic network will converge to the fixed point state over the long run. However, whether or why this assumption is true is still an open research question.

[4] As a result of stationary traffic flows, the flow conservation law requires that for any two points on the same link, traffic flows are equal.

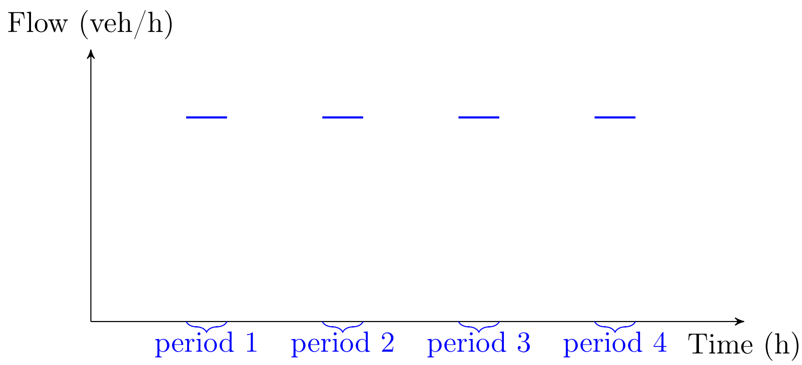


Figure 2.3: For a fixed point state, traffic flows are stationary within a period, and persist across periods

2.1.2 Traffic supply and traffic demand

In traffic network models, travelers' behavior is governed by travel costs. These are quantities that measure the dissatisfaction level of traveling on different paths. On one hand, travel costs influence users' path choice decisions (i.e., which path to travel on), and hence the traffic flows. On the other hand, travel costs are dependent of the traffic flows (for example, large traffic flow may cause congestion and hence increase the level of dissatisfaction). While the former relationship is referred to as traffic demand, the latter is referred as traffic supply.

Traffic supply

Consider a stationary short time period as described previously, for any link $l \in L$, let the traffic flow be f_l, and the travel time on this link be τ_l. Similarly, for any path $r \in R$, let h_r and μ_r be the path flow and path travel time, respectively. Let $\boldsymbol{f}$, $\boldsymbol{h}$ be the respective link and path flow vectors, $\boldsymbol{\tau}$ and $\boldsymbol{\mu}$ be the respective link and path travel time vectors, then $\boldsymbol{f}, \boldsymbol{\tau} \in \mathbb{R}^{|L|}$, and $\boldsymbol{h}, \boldsymbol{\mu} \in \mathbb{R}^{|R|}$. The link and path flows satisfy the following relationship $\boldsymbol{f} = \boldsymbol{\Delta h}$; the link and path costs satisfy $\boldsymbol{\mu} = \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{\tau}$.

It is reasonable to assume that link travel times depend on link flows. Due to congestion, when the link flow increases, the link travel time is likely to increase as well. Such relationship is assumed to be deterministic[5] and represented by the link flow-time function $t_l = t_l(\boldsymbol{f})$, $\forall l \in L$.[6]

In vector form the flow-time function can be written as $\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{f})$. We define the equivalent travel time function in the path space: $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{h}) = \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{t}(\boldsymbol{\Delta h})$.

The link flow-time function is called *separable* if $t_l = t_l(f_l)$, $\forall l \in L$. This implies that the travel time on link l only depends on the flow on link l. The link flow-time function is called *symmetric* if it is differentiable and $\frac{\partial t_l}{\partial f_{l'}} = \frac{\partial t_{l'}}{\partial f_l}$, $\forall l, l' \in L$. Obviously separable link flow-time functions are also symmetric.

Two specific types of link flow-time functions will be used in this thesis. The first is the affine flow-time function, namely,

$$t_l(\boldsymbol{f}) = a_l f_l + b_l. \tag{2.1}$$

This functional form is often used for its theoretical simplicity. Here a_l and b_l are link specific parameters. The second is the BPR link-travel-time function recommended by the U.S. Bureau of Public Roads (see U.S. Department of Commerce, 1964), namely,

$$t_l(\boldsymbol{f}) = \alpha_l \left(1 + \beta_l \left(\frac{f_l}{\gamma_l}\right)^4\right). \tag{2.2}$$

This functional form is widely used in empirical traffic network modeling. Here α_l, β_l, and γ_l are link specific parameters.

[5]In fact, the traffic flow theory studies the flow-time relationship and concludes that when link flow approaches its maximum level, the link travel time becomes highly stochastic. Interested readers may refer to Chapter 2 of Cascetta (2001) for more details. However, the traffic flow theory and its results will not be the subject of this thesis. The deterministic flow-time relationship is a good approximation for traffic networks and has been proved effective in applications.

[6]Some stochastic network models assume that such a relationship is not deterministic and that link flow-time functions give rise to the mean travel times.

Traffic demand

In many applications it suffices to assume that the travel cost is solely determined by the travel time. If travel time is deterministic, the obvious choice is $\boldsymbol{c} = \boldsymbol{\mu} = \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{t}(\boldsymbol{\Delta h})$. Building upon this, other components of travel costs can be included in at least three manners:

1. For quantifiable travel costs such as tolls or delay penalties. The resulting travel costs are called generalized travel costs.

2. Travel costs that are not observable to the modeler or that are observable but difficult to quantify, can be incorporated into the model as random "perception errors". The perception errors follow a modeler specified distribution. The stochastic user equilibrium model follows this approach.

3. In the case of stochastic network models, the travel times are random and the travel time uncertainty can be translated into travel costs through certain risk measures. These risk measures will be discussed in detail in Chapter 3.

There are two commonly adopted approaches for the traffic demand model. The first approach imposes users only traveling on paths of minimum cost. The deterministic user equilibrium follows this approach. The second approach assumes users have perception errors of the path costs, and hence minimize the perceived travel cost. The stochastic user equilibrium follows this approach.

Both approaches will be analyzed in greater detail later in this chapter.

2.1.3 User equilibrium

In a traffic network, user equilibrium is achieved when the inputs and outputs of the traffic supply model and the traffic demand model match each other, namely the flows and the costs in the two models are equal. This definition is formally stated in

Bell and Iida (1997) as follows: under equilibrium, demand (flow) is equal to supply (flow) at the prevailing costs to the users. Historically user equilibrium is defined as the state in which the traffic flow is such that there is no incentive for path switching. These two definitions are equivalent since “no incentive for path switching” implies that the traffic demand flow at prevailing costs coincides with the existing flow, and vice versa.

One key feature of the user equilibrium is that the traffic network reaches a static state across different time periods. On the other hand, if we assume that both travel demand and travel supply remain (approximately) static across different time periods, it is generally true that the traffic network will also be in an equilibrium state.

Equilibrium models are widely used because they are powerful and simple. These models are effectively timeless, and hence avoid the need of specifying and calibrating a dynamic mechanism according to which the traffic network evolves over time.

2.2 Deterministic Network Equilibrium Models

Deterministic network equilibrium models are models in which the traffic flows and travel times (costs) are deterministic. This type of models include the deterministic user equilibrium (DUE) model and the stochastic user equilibrium (SUE) model.

First proposed by Wardrop (1952), the DUE model assumes travelers only choose paths that minimize their travel costs, and is perhaps the most studied traffic network model in the literature. The SUE model was first studied by Daganzo and Sheffi (1977) as an extension to the DUE model. It assumes that travelers minimize their perceived travel costs. The perceived travel cost deviates from the actual travel cost by a stochastic perception error term. In this chapter, we shall investigate the exact meaning of these perception errors.

2.2.1 The deterministic user equilibrium (DUE) model

One of the most well studied traffic network equilibrium models is the deterministic user equilibrium (DUE) model. It assumes that users minimize their travel cost, which in the simple case is just the travel time. This implies, for a given OD pair, any path with a positive flow has minimum cost. For any OD pair k, let π^k be the minimum path cost. Let S_h be the set of feasible path flows, that is, $S_h = \left\{\boldsymbol{h} \in \mathbb{R}^{|R|} \geqslant 0 : \boldsymbol{\Gamma h} = \boldsymbol{q}\right\}$.

This gives rise to the following formulation for the DUE model.

DUE-FIX. Find a path flow vector $\boldsymbol{h}^{\mathrm{DUE}} \in S_h$, such that there exists a vector $\boldsymbol{\pi} \in \mathbb{R}^{|K|\times 1}$ of minimum path costs for all OD pairs, so that the following relationship holds

$$\begin{aligned} \boldsymbol{c}^{\mathrm{DUE}} &= \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{t}(\boldsymbol{\Delta h}^{\mathrm{DUE}}) \\ c_r^{\mathrm{DUE}} &= \pi^k, \quad \text{if } h_r^{\mathrm{DUE}} > 0, \quad \forall k \in K,\, r \in R^k \\ c_r^{\mathrm{DUE}} &\geqslant \pi^k, \quad \text{if } h_r^{\mathrm{DUE}} = 0, \quad \forall k \in K,\, r \in R^k \end{aligned} \tag{2.3}$$

We refer to the above formulation as the fixed point (FIX) formulation of the DUE model. For the convenience of the theoretical and computational study of the DUE model, a number of equivalent formulations have been developed in the literature. Two particular formulations, namely the variational inequality (VI) formulation and the optimization (OPT) formulation will be of particular interest because they will be used later in this thesis. We first state the VI formulation. It can be formulated in either the link flow space or the path flow space. Let S_f be the set of feasible link flows, that is, $S_f = \{\boldsymbol{f} = \boldsymbol{\Delta h} : \boldsymbol{h} \in S_h\}$.

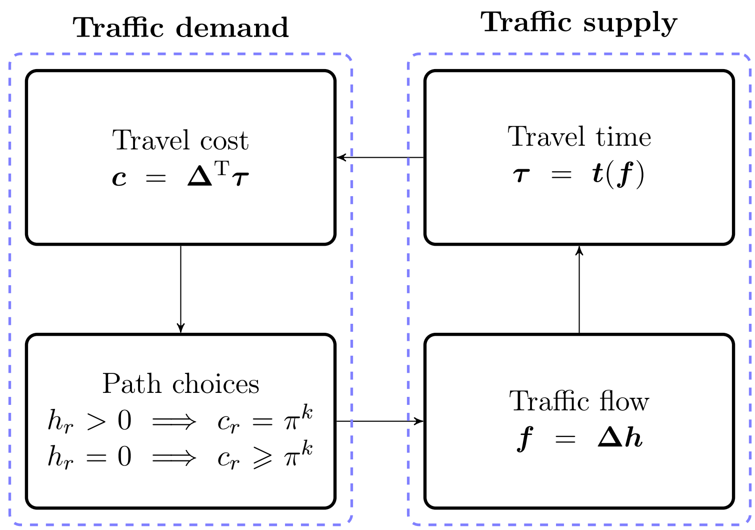


Figure 2.4: Diagram of the deterministic user equilibrium model.

DUE-VI-PATH. Find $\boldsymbol{h}^{\mathrm{DUE}} \in S_h$, such that

$$\begin{aligned} \boldsymbol{c}^{\mathrm{DUE}} &= \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{t}(\boldsymbol{\Delta}\boldsymbol{h}^{\mathrm{DUE}}) \\ \left(\boldsymbol{c}^{\mathrm{DUE}}\right)^{\mathrm{T}}\left(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{DUE}}\right) &\geqslant 0, \quad \forall \boldsymbol{h} \in S_h \end{aligned} \tag{2.4}$$

DUE-VI-LINK. Find $\boldsymbol{f}^{\mathrm{DUE}} \in S_f$, such that

$$\boldsymbol{t}(\boldsymbol{f}^{\mathrm{DUE}})^{\mathrm{T}}\left(\boldsymbol{f}-\boldsymbol{f}^{\mathrm{DUE}}\right) \geqslant 0, \quad \forall \boldsymbol{f} \in S_f \tag{2.5}$$

When function $\boldsymbol{t}(\cdot)$ is smooth and has a symmetric Jacobian matrix everywhere, an optimization formulation exists. In this case, the vector field defined by $\boldsymbol{t}(\cdot)$ is the gradient of a scalar field (refer to Lemma 2.1). Namely, there exists a scalar function $y(\cdot)$ such that $\boldsymbol{t}(\boldsymbol{f}) = \boldsymbol{\nabla} y(\boldsymbol{f})$, for any $\boldsymbol{f} \in S_f$. This function can be explicitly defined by a line integral $y(\boldsymbol{f}) \equiv \int_{\boldsymbol{0}}^{\boldsymbol{f}} \boldsymbol{t}(\boldsymbol{\omega}) \cdot \mathrm{d}\boldsymbol{\omega}$.[7] The optimization formulation is stated below.

[7] A line integral can be converted to a standard integral by expressing the integrand as a function of a single (scalar) parameter, and then performing the integration along the curve defined by the parameter. If the integrand function is smooth and has s symmetric Jacobian matrix, the integral

DUE-OPT. Find the minimizer $\boldsymbol{f}^{\mathrm{DUE}}$ to the following minimization problem

$$\min_{\boldsymbol{f} \in S_f} \; y(\boldsymbol{f}) \tag{2.6}$$

The following lemma formally states that existence of the scalar function $y(\cdot)$.

Lemma 2.1. *If function $\boldsymbol{t}(\cdot)$ is smooth and has a symmetric Jacobian matrix everywhere, then there exists a smooth scalar function $y(\cdot)$ such that $\boldsymbol{t}(\boldsymbol{f}) = \boldsymbol{\nabla} y(\boldsymbol{f})$ for any $\boldsymbol{f} \in S_f$.*

Remark. This lemma can be proved using Poincare's lemma (refer to Chapter 10 of Rudin, 1986).

We then proceed to state the equivalences of the various formulations.

Theorem 2.2. *The following DUE formulations are equivalent.*

- *DUE-FIX: the original formulation* (2.3)
- *DUE-VI-PATH: the VI formulation in the path flow space* (2.4)
- *DUE-VI-LINK: the VI formulation in the link flow space* (2.5)

Furthermore, if the flow-time function $\boldsymbol{t}(\cdot)$ is smooth and has a symmetric and positive semidefinite Jacobian matrix everywhere, then the above formulations are also equivalent to the following.

- *DUE-OPT: the optimization formulation* (2.6).

Remark. Refer to Cascetta (2001) for a proof of this theorem.

only depends on the starting and ending points of the integration and is independent of the path.

2.2.2 The stochastic user equilibrium (SUE) model

While the DUE model allocates travel demands to the set of paths with the least travel costs, the SUE model assumes that due to variations in travelers' perceptions of travel costs, trip makers distribute themselves along competitive paths. In SUE, users' perceived travel cost $\tilde{c}_r$ deviates from the "actual" travel cost c_r by some random perception error ε_r, namely,

$$\tilde{c}_r = c_r^{\mathrm{SUE}} + \varepsilon_r,$$

with the "actual" travel cost generated by the flow-time function $c^{\mathrm{SUE}} = \mu(h^{\mathrm{SUE}}) = \Delta^{\mathrm{T}} t(\Delta h^{\mathrm{SUE}})$. The perception error ε_r is identically and independently distributed across the traveler population and across different routes.

The SUE literature has provided several interpretations of the exact meaning of the perception errors:

- The perception error may correspond to the deviation of the realization of the random individual travel time from its mean value (see Daganzo and Sheffi 1977). Despite the intra-period stationarity approximation, this interpretation acknowledges the randomness of individual travel time and model them as perception error.

- The perception error may correspond to the imprecise measurement of the travel time by users (see Sheffi and Powell 1982). This is the deviation of perceived travel time from the actual travel time.

- The perception error may correspond to factors unobservable to modelers but that contribute to the travel cost (see Sheffi 1984, Bell and Iida 1997, and de Palma and Picard 2005). This interpretation acknowledges that individual travelers may place different weights on such attributes as the number of stops, variability in travel time, scenery and road quality. Since these factors are unobservable to modelers, it may be captured by through a random perception error.

In what follows we assume the multinomial logit model, in which the perception error ε_r is assumed to follow the Gumble distribution with parameter θ. Under the multinomial logit model, the probability that path r has the least perceived travel time compared to all available paths for OD pair k is $p_r = \frac{\exp(-c_r^{\text{SUE}}/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}^{\text{SUE}}/\theta)}$. The SUE model translates these probabilities as the proportions of users of OD pair k traveling on path r.

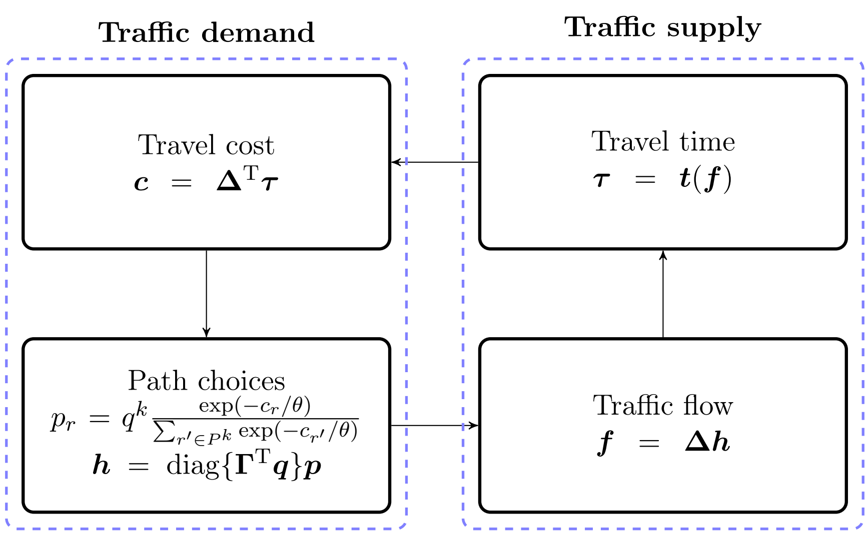


Figure 2.5: Diagram of the (multinomial logit) stochastic user equilibrium model.

The SUE equilibrium can then be formulated as below

SUE-FIX. Find a path flow vector $\boldsymbol{h}^{\text{SUE}} \in S_h$, such that the following relationship holds

$$\begin{aligned} \boldsymbol{c}^{\mathrm{SUE}} &= \boldsymbol{\Delta}^{\mathrm{T}} \boldsymbol{t}(\boldsymbol{\Delta} \boldsymbol{h}^{\mathrm{SUE}}), \\ h_r^{\mathrm{SUE}} &= q^k \frac{\exp(-c_r^{\mathrm{SUE}}/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}^{\mathrm{SUE}}/\theta)}, \quad \forall k \in K, \, r \in R^k. \end{aligned} \tag{2.7}$$

The above formulation has an equivalent VI and optimization formulation, as stated below

SUE-VI-PATH. Find $\boldsymbol{h}^{\mathrm{SUE}} \in S_h$, such that

$$\begin{aligned} \boldsymbol{c}^{\mathrm{SUE}} &= \boldsymbol{\Delta}^{\mathrm{T}} \boldsymbol{t}(\boldsymbol{\Delta} \boldsymbol{h}^{\mathrm{SUE}}), \\ \left(\boldsymbol{c}^{\mathrm{SUE}} + \theta \log \frac{\boldsymbol{h}^{\mathrm{SUE}}}{\boldsymbol{q}} \right)^{\mathrm{T}} &\left(\boldsymbol{h} - \boldsymbol{h}^{\mathrm{SUE}} \right) \geqslant 0, \quad \forall \boldsymbol{h} \in S_h. \end{aligned} \tag{2.8}$$

SUE-OPT-PATH. Find the minimizer of the following optimization problem

$$\min_{\boldsymbol{h} \in S_h} \; y(\Delta \boldsymbol{h}) + \theta g(\boldsymbol{h}), \tag{2.9}$$

In this setting $y(\cdot)$ is defined as $y(\boldsymbol{f}) \equiv \int_{\mathbf{0}}^{\boldsymbol{f}} \boldsymbol{t}(\boldsymbol{\omega}) \cdot \mathrm{d}\boldsymbol{\omega}$, and $g(\boldsymbol{h}) \equiv \left(\log \frac{\boldsymbol{h}}{\boldsymbol{q}} \right)^{\mathrm{T}} \boldsymbol{h}$. Note that we can prove the inequality in the VI formulation (2.8) always holds as an equality. See the proof of Theorem 3.1.

Theorem 2.3. *The following SUE formulations are equivalent.*

- *SUE-FIX: the fixed point formulation* (2.7)
- *SUE-VI-PATH: the VI formulation in the path flow space* (2.8)

Furthermore, if function $\boldsymbol{t}(\cdot)$ *is smooth and has a symmetric and positive semidefinite Jacobian matrix everywhere, then the above formulations are also equivalent to the following.*

- *SUE-OPT-PATH: the optimization formulation in the path flow space* (2.9)

Remark. The proof is similar to that in the DUE case. We shall omit the proof.

2.3 Stochastic Network Equilibrium Models

In this section we provide a unified framework for stochastic network equilibrium models. We will discuss two specific stochastic network equilibrium models, namely the robust user equilibrium (RUE) and the cumulative prospect user equilibrium (CPUE).

All stochastic network models must specify two fundamental mechanisms. One fundamental mechanism defines how travel time uncertainty is generated; we refer to this mechanism as uncertainty generator. The other fundamental mechanism defines how travel time uncertainty influences users decision; we refer to this mechanism as uncertainty responder. Different stochastic network models have specified different combinations of these two mechanisms.

However, as we will discuss below, among existing stochastic network equilibrium models, some have a sophisticated uncertainty generator, others have sophisticated uncertainty responder. The lack of balance between these two mechanisms motivates our development of a new stochastic network user equilibrium model in the next chapter.

2.3.1 The robust user equilibrium (RUE) model

Ordonez and Stier-Moses (2007) developed a stochastic network user equilibrium model that assumes users adopt a robust optimization approach in making their route choice decisions. We refer to this model as the robust user equilibrium (RUE). RUE uses a relatively simple uncertainty generator. It assumes that while the traffic flow is still deterministic, the actual link travel times have random deviations from the link travel times predicted by the flow-time functions:

$$\boldsymbol{T}(\boldsymbol{f}) = \boldsymbol{t}(\boldsymbol{f}) + \boldsymbol{\varepsilon}.$$

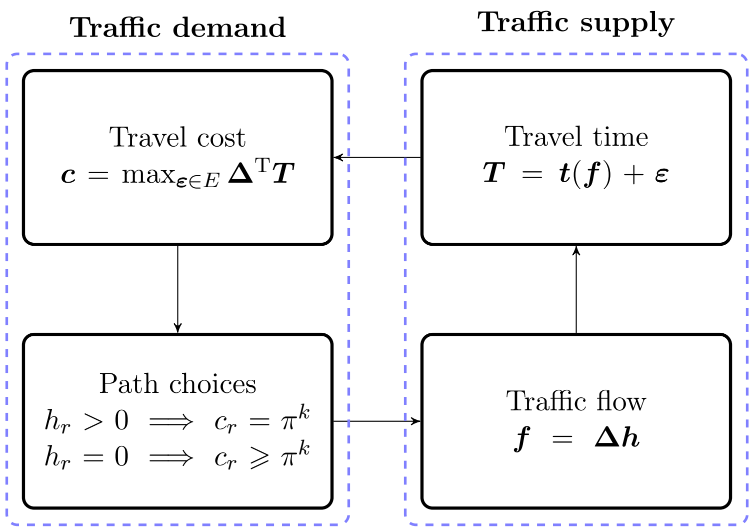


Figure 2.6: Diagram of the robust user equilibrium model.

In this case, $\boldsymbol{\varepsilon}$ is the vector of random deviations. It is assumed that these deviations are independent of the traffic flow. Consistent with other robust optimization models, RUE does not specify a distribution function for $\boldsymbol{\varepsilon}$. Instead RUE specifies an uncertainty set for $\boldsymbol{\varepsilon}$, denoted by E.

RUE adopts a robust approach in defining its uncertainty responser. It assumes that all users are risk-averse and hence only travel on paths with minimum worst case travel time. This assumption implies that the path cost is the worst case travel time, namely

$$c_r^{\mathrm{RUE}} = \max_{\boldsymbol{\varepsilon} \in E} \boldsymbol{\Delta}_r^{\mathrm{T}} \boldsymbol{T} = \max_{\boldsymbol{\varepsilon} \in E} \boldsymbol{\Delta}_r^{\mathrm{T}} \left(\boldsymbol{t}(\boldsymbol{\Delta h}) + \varepsilon \right).$$

Users then choose the paths that minimize this cost.

2.3.2 The cumulative prospect user equilibrium (CPUE) model

Connors and Sumalee (2009) considered a stochastic network equilibrium model that features a sophisticated approach of modeling users' perception of travel time uncertainty. Their model utilizes the cumulative prospect theory (CPT), which

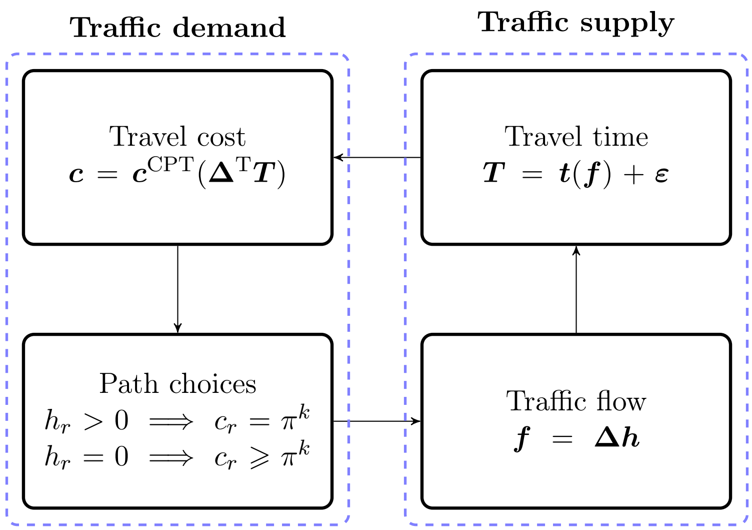


Figure 2.7: Diagram of the cumulative prospect user equilibrium model.

provides an empirically well-supported paradigm for choices made under uncertainty. In their model, the path travel cost[8] is evaluated from the actual outcomes (namely, travel times) and their probabilities of occurring via two transformations

- a value function $g(\cdot)$ that describes the payoff level experienced by users for each possible travel time, and
- a probability weighting function $w(\cdot)$ that transforms the probability into the perceived likelihood of occurrence.

In particular, let M be the random variable of path travel time, then

$$g(M) = \begin{cases} (M - m_0)^{\alpha} & M \geqslant m_0 \\ -\lambda(m_0 - M)^{\beta} & M < m_0 \end{cases}$$

$$w(p) = \exp\left(-\left(-\log p\right)^{\gamma}\right)$$

[8]The term actually used by Connor and Sumalee is perceived value. Path travel cost is the general term this thesis adopts to provide a consistent framework for various models.

where the parameters satisfy $0 < \alpha, \beta, \gamma \leqslant 1$, and m_0 is a reference travel time (as reflected in function g, depending on whether M is larger or smaller compared to the reference travel time, users value the outcome differently). If M is a continuous random variable, let $F_M(\cdot)$ be its cumulative density function (CDF) for M. The path travel cost is (we denote it as c^{CPT})[9]

$$c = c^{\mathrm{CPT}}(M) = \int_{-\infty}^{m_0} g(m) \frac{\mathrm{d}}{\mathrm{d}m} \left(w \left(F_M(m) \right) \right) \mathrm{d}m + \int_{m_0}^{\infty} g(m) \frac{\mathrm{d}}{\mathrm{d}m} \left(-w \left(1 - F_M(m) \right) \right) \mathrm{d}m$$

While on the demand side, Connors and Sumalee (2009) adopted a sophisticated way in modeling users' perception and response to uncertain travel time, on the supply side they assumed a much simpler model, namely, that 1) the flow is deterministic and 2) the stochasticity of travel time is modeled by a flow-independent error term

$$\boldsymbol{M} = \boldsymbol{\Delta}^{\mathrm{T}} \left(\boldsymbol{t} \left(\boldsymbol{\Delta f} \right) + \boldsymbol{\varepsilon} \right).$$

$\boldsymbol{\varepsilon}$ is independent on different links.

[9] Note that $w(\cdot)$ applies on the cumulative probability weight instead of probability of individual events. Hence this approach is named *cumulative* prospect theory.

Chapter 3

Mean-Variance Approach to Stochastic Network Equilibrium Model

This section develops a new stochastic user equilibrium model. This model assumes stochasticity, in the sense that travelers make probabilistic path choices from day to day. The probabilistic path choice decisions induce stochastic traffic flows in the network, which result in uncertain travel times. The travel time uncertainty in turn encourages travelers to reduce their probabilities of choosing the more uncertain paths.

This model is unique in the sense that travel time uncertainty is captured both as a source of network stochasticity (through its impact on travelers' path choice decisions) and as a result of network stochasticity (through the path choice model that generates stochastic traffic flows). This new model contrasts with most previously developed models in assuming that both traffic flows and travel times are stochastic. It is on this ground that we name this new model the "Truly Stochastic

User Equilibrium" (TSUE) model.

This chapter is organized as below. Section 3.1 formulates this model. Section 3.2 introduces a number of alternative approaches to modeling the travel cost in stochastic network. Section 3.3 investigates the existence and uniqueness of the equilibrium solution. Section 3.4 proposes a solution algorithm. Section 3.5 presents the numerical results, followed by discussion.

3.1 Model Formulations

This section introduces the new model. The model involves four key steps (approximations):

1. We assume that travelers make probabilistic path choice decisions. Travelers associate with each of their available paths a probability. Everyday, each traveler *independently* chooses between alternative paths according to these probabilities. These probabilities are assumed to be the *same* for travelers of the same OD pair. We refer to this as the *assumption of probabilistic path choice decisions.* As a result, the path flows are random variables that follow the multinomial distribution.

2. We approximate the travel time function with its first order expansion in the neighborhood of the expected flow. From this approximation and the distribution of path flows, we obtain the mean and variance of travel times.

3. In the equilibrium state, the traveler's travel cost is determined by the mean-variance approach. That is, the travel cost associated with a path is the sum of its mean travel time and the variance of its travel time multiplied by a trade-off factor $\lambda \geqslant 0$. This trade-off factor characterizes how much cost the travelers

associate with the variance of travel times, relative to the mean travel times. It is assumed that this factor is fixed across the traveler population.

4. Travelers have random perception errors of their travel cost. The perceived travel cost differs from the actual travel cost by a random perception error term. This error term is assumed to be *independently and identically distributed* from day to day and across the traveler population. In this thesis we assume that this perception error follows the *Gumbel distribution* with dispersion level θ, resulting to the multinomial logit model.

We discuss these steps in detail in this section. In Section 3.2, we will discuss some alternative approaches to the third step.

3.1.1 Probabilistic path choice decisions

To motivate the new model, we first recall the SUE model, which assumes that travelers have perception errors of their path travel costs. The perception errors are assumed to be independent from traveler to traveler, and follow a known distribution across the population. Based on these assumptions, the SUE model derives from the actual path travel costs the probability of one randomly selected traveler choosing a certain path. This probability on the individual traveler level is then interpreted aggregately as the proportion of travelers choosing that path. Despite the stochasticity involved in deriving the path choice probabilities, the interpretation of probabilities as proportions renders the SUE model into a deterministic network model with deterministic traffic flows and travel times.

However, in reality, it is well known that traffic flows fluctuate from day to day, hence it is more appropriate to represent traffic flows as random variables. Noticing this, we consider an alternative interpretation of the probabilities in the SUE model. Namely, we will avoid characterizing the path choice behavior in an aggregate level through a deterministic approach, and instead we will characterize the path choice behavior at an individual level using a stochastic approach, where each traveler independently and at random chooses from available paths using a probability distribution. Under this assumption the traffic flows are random variables. We assume that the traffic demand is fixed at each OD pair, then traffic flows follow the multinomial distribution. As a result, the actual path traffic flows on a certain day are realizations from this multinomial distribution.

In the equilibrium state, we assume that travelers of the same OD pair share the common set of path choice probabilities. We will explain why we make this assumption when we discuss the random utility model in Section 3.1.4.

Next we introduce the notations we use. Let $\boldsymbol{p} \in [0,1]^{|R|}$ be the vector of path choice probabilities and S_p be the set of feasible path choice probabilities[1]. Written in explicit form, we set $S_p = \left\{\boldsymbol{p} \in [0,1]^{|R|} : \boldsymbol{\Gamma p} = \boldsymbol{1}\right\}$, where $\boldsymbol{\Gamma}$ is the trip-path incidence matrix, $\boldsymbol{1}$ is a column vector of all 1's, and $\boldsymbol{\Gamma p} = \boldsymbol{1}$ states that the path choice probabilities sum up to 1 for each OD pair. Let $\boldsymbol{H}$ be the vector of path traffic flows; in the TSUE model, $\boldsymbol{H}$ is a multivariate random variable of dimension $|R|$, that we define next.

For an OD pair $k \in K$, let $\boldsymbol{p}^k$ and $\boldsymbol{H}^k$ be the vectors of path choice probabilities and path traffic flows, respectively, of paths available for k. Since $|R^k|$ paths are available for OD pair k, both $\boldsymbol{p}^k$ and $\boldsymbol{H}^k$ have $|R^k|$ elements. By the path choice probability assumption, $\boldsymbol{H}^k$ follows the multinomial distribution. Assuming that the total demand for OD pair k is q^k, we have that

$$\boldsymbol{H}^k \sim \text{Multinomial}(q^k, \boldsymbol{p}^k). \tag{3.1}$$

[1]In our notation, letter p always stands for probability, and letter r always stands for paths (routes).

The above relationship holds for any OD pair $k \in K$. To simplify our notation, we write the random variable in vector form

$$\boldsymbol{H} \sim \text{Multinomial}(\boldsymbol{q}, \boldsymbol{p}).$$

Let $\boldsymbol{h} = \mathbb{E}[\boldsymbol{H}]$ be the vector of the expected path flows, then

$$\boldsymbol{h}^k = \mathbb{E}[\boldsymbol{H}^k] = q^k \boldsymbol{p}^k. \tag{3.2}$$

Notice that this relationship coincides with that in the SUE model. However, contrary to the TSUE model, the SUE model interprets $\boldsymbol{h}$ as the flow that is deterministic, rather than the expected flow as in the TSUE model.

Let $\boldsymbol{F}$ be the vector of link traffic flows and $\boldsymbol{f} = \mathbb{E}[\boldsymbol{F}]$ be the vector of expected link flows. Then

$$\boldsymbol{F} = \boldsymbol{\Delta} \boldsymbol{H}, \tag{3.3}$$

$$\boldsymbol{f} = \boldsymbol{\Delta} \boldsymbol{h}. \tag{3.4}$$

3.1.2 First-order approximation of the travel time function

To incorporate the travel time uncertainty into the travel cost, we need to evaluate the mean and variance of the path travel time. Let $\boldsymbol{M}$ be the vector of path travel times. We assume that the travel time is deterministically determined by the traffic flow, i.e., $\boldsymbol{M} = \boldsymbol{\Delta}^{\mathrm{T}} \boldsymbol{t}(\boldsymbol{F})$, where $\boldsymbol{t}$ is the link travel time function.

In practice, both $\mathbb{E}[\boldsymbol{M}]$ and $\mathrm{Var}(\boldsymbol{M})$ may be difficult to evaluate for general nonlinear travel time functions. To overcome this difficulty, we propose the following approximation scheme. We write the Taylor expansion of the link travel time function in the neighborhood of the expected flow $\boldsymbol{f}$, namely

$$\boldsymbol{t}(\boldsymbol{F}) = \boldsymbol{t}\left(\boldsymbol{f}\right) + \boldsymbol{J}_t\left(\boldsymbol{f}\right)\left(\boldsymbol{F} - \boldsymbol{f}\right) + O\left(\left(\boldsymbol{F} - \boldsymbol{f}\right)^2\right),$$

where $\boldsymbol{J}_t\left(\boldsymbol{f}\right)$ is the Jacobian matrix of the travel time function $\boldsymbol{t}$ evaluated at $\boldsymbol{f}$. We consider the first order approximation of the travel time function

$$\boldsymbol{t}(\boldsymbol{F}) \approx \hat{\boldsymbol{t}}(\boldsymbol{F}) = \boldsymbol{t}\left(\boldsymbol{f}\right) + \boldsymbol{J}_t\left(\boldsymbol{f}\right)\left(\boldsymbol{F} - \boldsymbol{f}\right).$$

Let $\hat{\boldsymbol{M}} = \boldsymbol{\Delta}^{\mathrm{T}}\hat{\boldsymbol{t}}(\boldsymbol{F})$ be the corresponding approximate path travel time. Using Eq. (3.3) and Eq. (3.4), we have that

$$\hat{\boldsymbol{M}} = \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{t}\left(\boldsymbol{\Delta}\boldsymbol{h}\right) + \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{J}_t\left(\boldsymbol{f}\right)\boldsymbol{\Delta}\left(\boldsymbol{H} - \boldsymbol{h}\right). \tag{3.5}$$

The expectation of $\hat{\boldsymbol{M}}$ is

$$\mathbb{E}[\hat{\boldsymbol{M}}] = \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{t}\left(\boldsymbol{\Delta}\boldsymbol{h}\right). \tag{3.6}$$

Let $\mathrm{Var}(\hat{\boldsymbol{M}})$ be the vector of the variances of $\hat{\boldsymbol{M}}$. Then

$$\mathrm{Var}(\hat{\boldsymbol{M}}) = \mathrm{Diag}\left\{\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{J}_t\left(\boldsymbol{\Delta}\boldsymbol{h}\right)\boldsymbol{\Delta}\mathrm{Cov}\left(\boldsymbol{H}\right)\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{J}_t\left(\boldsymbol{\Delta}\boldsymbol{h}\right)\boldsymbol{\Delta}\right\}.$$

In this equation, the $\mathrm{Diag}\left\{\cdot\right\}$ operator extracts the main diagonal of a square matrix. For example, applying the $\mathrm{Diag}\left\{\cdot\right\}$ operator on the identity matrix $\boldsymbol{I}$ gives the vector of all 1's $\mathbf{1}$, i.e., $\mathrm{Diag}\left\{\boldsymbol{I}\right\} = \mathbf{1}$). $\mathrm{Cov}(\boldsymbol{H})$ gives the covariance matrix of the multivariate random variable $\boldsymbol{H}$. For two paths r, $r' \in R$, the (r, r') element of

Cov $(\boldsymbol{H})$ is

$$(\text{Cov}(\boldsymbol{H}))_{r,r'} = \begin{cases} q^k(p_r - p_r^2) & \text{if } r = r' \in R^k, \\ -q^k p_r p_{r'} & \text{if } r \neq r', \text{ but both } r \text{and } r' \text{are in } R^k, \\ 0 & \text{otherwise.} \end{cases}$$

We may rewrite the above result in a more compact form. In fact, Cov $(\boldsymbol{H})$ is a block diagonal matrix with $|K|$ blocks on this diagonal. For $k \in K$, its kth block is

$$\begin{aligned} \text{Cov}(\boldsymbol{H}^k) &= q^k \left(\text{diag}\left\{ \boldsymbol{p}^k \right\} - \boldsymbol{p}^k (\boldsymbol{p}^k)^{\text{T}} \right) \\ &= \text{diag}\left\{ \boldsymbol{h}^k \right\} - \frac{1}{q^k} \boldsymbol{h}^k (\boldsymbol{h}^k)^{\text{T}}. \end{aligned} \tag{3.7}$$

In this equation, the diag $\{\cdot\}$ operator turns a vector into a diagonal matrix. It is the inverse of the Diag $\{\cdot\}$ operator, e.g., diag $\{\mathbf{1}\} = \boldsymbol{I}$. Let $\boldsymbol{\Delta}^k$ be the link-path incidence matrix restricted to OD pair $k \in K$ (with dimension $|L| \times |R^k|$), then we have

$$\begin{aligned} \text{Var}(\hat{\boldsymbol{M}}) &= \text{Diag}\left\{ \boldsymbol{\Delta}^{\text{T}} \boldsymbol{J}_t (\boldsymbol{\Delta h}) \sum_{k \in K} \left(\boldsymbol{\Delta}^k \text{Cov}\left(\boldsymbol{H}^k \right) (\boldsymbol{\Delta}^k)^{\text{T}} \right) \boldsymbol{J}_t (\boldsymbol{\Delta h}) \boldsymbol{\Delta} \right\} \\ &= \text{Diag}\Bigg\{ \boldsymbol{\Delta}^{\text{T}} \boldsymbol{J}_t (\boldsymbol{\Delta h}) \\ &\quad \sum_{k \in K} \left(\boldsymbol{\Delta}^k \left(\text{diag}\left\{ \boldsymbol{h}^k \right\} - \frac{1}{q^k} \boldsymbol{h}^k (\boldsymbol{h}^k)^{\text{T}} \right) (\boldsymbol{\Delta}^k)^{\text{T}} \right) \boldsymbol{J}_t (\boldsymbol{\Delta h}) \boldsymbol{\Delta} \Bigg\}. \end{aligned} \tag{3.8}$$

3.1.3 Mean-variance approach to the travel cost

Essential to our stochastic network model is the mechanism in which travelers determine their travel cost in the presence of uncertain travel times. Consider any path $r \in R$. The travel time on this path is M_r. We assume that its path travel cost c_r is measured as the mean travel time plus a constant multiplied with the variance of the travel time, namely

$$c_r = \mathbb{E}[M_r] + \lambda \mathrm{Var}(M_r), \tag{3.9}$$

where the multiplier λ is called the mean-variance trade-off factor. Multiplier λ represents the travelers' risk aversion level; the greater its value, the more risk averse are the travelers. It is important to note that λ is not a dimensionless parameter. If the travel time is measured in minutes, λ has the unit of 1/minute.

This above mean-variance approach was frequently used in empirical studies of travelers' path choice behavior under travel time uncertainty (including Jackson and Jucker 1982, Senna 1994, and Nam et al. 2005). These studies also allow modelers to calibrate the value of λ. Furthermore, Small (1999) showed that the variance of travel time is a significant and positive attribute in the travel cost, which implies that $\lambda > 0$. Another empirical study by de Palma and Picard (2005) suggests that there are two different groups of travelers with respect to their preferences of travel time uncertainty: the majority of travelers are risk-averse, others are risk neutral or risk seeking. For the present thesis, the value of λ is assumed to be nonnegative and uniform across the population. Relaxing the assumption that λ is fixed across the population is an open field of exploration. For example, one can assume that λ follows a distribution that models traveler heterogeneity.

There are two main reasons to justify our choice of the mean-variance approach.

1. Although empirical traffic network studies have frequently used the travel time variance as a variable to explain travelers' path choice decisions (see, for example, Jackson and Jucker, 1982), no traffic equilibrium model appears to have incorporated the travel time variance as a factor in its path choice model; and

2. From a computational point of view, the evaluation of the travel time variance is relatively easy (see the following discussion).

We will list some alternative approaches to modeling travel costs in Section 3.2.

Using the first-order approximation introduced in Section 3.1.2, the traveler's path choice is characterized by the travel cost

$$\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{h}) = \mathbb{E}[\hat{\boldsymbol{M}}] + \lambda \mathrm{Var}(\hat{\boldsymbol{M}}), \tag{3.10}$$

where $\lambda \geqslant 0$ is a constant parameter. The expectation and variance of $\hat{\boldsymbol{M}}$ can be obtained through Eq. (3.6) and Eq. (3.8), respectively. From these two equations we know that the path travel cost $\boldsymbol{c}$ is a function of the expected path flow $\boldsymbol{h}$. We have that

$$\begin{aligned} \boldsymbol{c}(\boldsymbol{h}) = \boldsymbol{\Delta}^{\mathrm{T}} \boldsymbol{t}\left(\boldsymbol{\Delta h}\right) + \lambda \mathrm{Diag}\Bigg\{ & \boldsymbol{\Delta}^{\mathrm{T}} \boldsymbol{J}_t\left(\boldsymbol{\Delta h}\right) \\ & \sum_{k \in K} \left(\boldsymbol{\Delta}^k \left(\mathrm{diag}\left\{ \boldsymbol{h}^k \right\} - \frac{1}{q^k} \boldsymbol{h}^k (\boldsymbol{h}^k)^{\mathrm{T}} \right) (\boldsymbol{\Delta}^k)^{\mathrm{T}} \right) \boldsymbol{J}_t\left(\boldsymbol{\Delta h}\right) \boldsymbol{\Delta} \Bigg\}. \end{aligned} \tag{3.11}$$

When $\lambda = 0$, the TSUE travel cost is $\boldsymbol{c} = \boldsymbol{\Delta}^{\mathrm{T}} \boldsymbol{t}\left(\boldsymbol{\Delta h}\right)$. In this special case the TSUE model reduces to the SUE model.

3.1.4 The random perception errors of the travel cost

We assume that the travelers have perception errors of their travel cost $\boldsymbol{c}$. The perceived travel cost is

$$\bar{\boldsymbol{c}} = \boldsymbol{c} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon}$ is the vector of random perception errors. In this thesis we assume that $\boldsymbol{\varepsilon}$ is independent and identically distributed for all paths and all travelers, and it follows the Gumbel distribution with dispersion level θ. Travelers then choose paths that minimize their perceived travel cost. This process results in the multinomial logit model. According to this model, the probability that a traveler traveling on OD pair

$k \in K$ chooses path $r \in R^k$ is

$$p_r = \frac{\exp(-c_r/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}/\theta)}. \tag{3.12}$$

We assume that travelers have common path choice probabilities in the equilibrium state. This is because: 1) travelers use the same mechanism to determine their path choice probability (that is, Eq. (3.12)), and 2) in the long run, travelers will have enough samples of the travel times that allow them to evaluate the mean and variance of travel times, hence they will experience the same travel cost $\boldsymbol{c}$. In order to study how traveler update their cost based their actual travel experience, we need to specify a *learning mechanism* in which travelers update their path choice probabilities. A specific learning mechanism will be considered in Chapter 5 when we study a related inter-period dynamic model. However, we do assume that travelers are aware of all the paths available to them and the travel costs of these paths. Ramming (2002) examined the relationship between network knowledge and path choice behavior. Nevertheless, we will not cover this topic in the thesis.

In the equilibrium state, the path choice probability $\boldsymbol{p}$ does not change over time, and as a result, the (actual) travel cost $\boldsymbol{c}$ is also stationary. We further assume that the perception errors are independent not only across the traveler population, but also from day to day. This assumption ensures that for the same traveler the perceived travel cost is random from day to day. Hence the traveler makes probabilistic path choice decisions everyday.

The TSUE model seeks to find the equilibrium path choice probabilities. We propose two equivalent formulations

TSUE-FIX. Find $\boldsymbol{h}^{\text{TSUE}} \in S_h$, such that

$$\begin{aligned} \boldsymbol{c}^{\text{TSUE}} &= \boldsymbol{c}(\boldsymbol{h}^{\text{TSUE}}), \\ h_r^{\text{TSUE}} &= q^k \frac{\exp(-c_r^{\text{TSUE}}/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}^{\text{TSUE}}/\theta)}, \quad \forall k \in K, r \in R^k. \end{aligned} \tag{3.13}$$

TSUE-VI. Find $\boldsymbol{h}^{\text{TSUE}} \in S_h$, such that

$$\begin{aligned} \boldsymbol{c}^{\text{TSUE}} &= \boldsymbol{c}(\boldsymbol{h}^{\text{TSUE}}), \\ \left(\boldsymbol{c}^{\text{TSUE}} + \theta \log \frac{\boldsymbol{h}^{\text{TSUE}}}{\boldsymbol{q}}\right)^{\text{T}} &\left(\boldsymbol{h} - \boldsymbol{h}^{\text{TSUE}}\right) \geqslant 0, \quad \forall \boldsymbol{h} \in S_h. \end{aligned} \tag{3.14}$$

The first formulation is a fixed point formulation. The second formulation is a variational inequality formulation. We note that the inequality in the TSUE-VI formulation always holds as an equality (as shown in the proof). The next theorem states the equivalence of these two formulations.

Theorem 3.1. *The TSUE-FIX and TSUE-VI formulations for the TSUE model are equivalent.*

Proof. We first show that TSUE-FIX implies TSUE-VI. Suppose $\boldsymbol{h}^{\text{TSUE}}$ satisfies the TSUE-FIX formulation. For any OD pair $k \in K$, let $\pi^k = -\theta \log \left(\sum_{r' \in R^k} \exp(-c_{r'}^{\text{TSUE}}/\theta)\right)$. Then, according to Eq. (3.13), we have that

$$c_r^{\text{TSUE}} + \theta \log \frac{h_r^{\text{TSUE}}}{q^k} = \pi^k \quad \forall k \in K, r \in R^k.$$

For any $\boldsymbol{h} \in S_h$, we have

$$\begin{aligned}
&\left(\boldsymbol{c}^{\mathrm{TSUE}} + \theta \log \frac{\boldsymbol{h}^{\mathrm{TSUE}}}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\boldsymbol{h} - \boldsymbol{h}^{\mathrm{TSUE}}\right) \\
=& \sum_{k \in K} \sum_{r \in R^k} \pi^k \left(h_r - h_r^{\mathrm{DUE}}\right) \\
=& \sum_{k \in K} \pi^k \left(\sum_{r \in R^k} h_r - \sum_{r \in R^k} h_r^{\mathrm{DUE}}\right) \\
=& \sum_{k \in K} \pi^k \left(q^k - q^k\right) = 0.
\end{aligned}$$

Hence $\boldsymbol{h}^{\mathrm{TSUE}}$ satisfies the TSUE-VI formulation.

Next we show that TSUE-VI implies TSUE-FIX. Suppose $\boldsymbol{h}^{\mathrm{TSUE}}$ satisfies the TSUE-VI formulation Eq. (3.14). Notice that for any OD pair $k \in K$, we have

$$\sum_{r \in R^k} h_r^{\mathrm{TSUE}} = q^k = \sum_{r \in R^k} q^k \frac{\exp(-c_r^{\mathrm{TSUE}}/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}^{\mathrm{TSUE}}/\theta)}.$$

If, by contradiction, the TSUE-FIX formulation doesn't hold, then there must exist some OD pair k and paths $r, s \in R^k$, such that $h_r^{\mathrm{TSUE}} > q^k \frac{\exp(-c_r^{\mathrm{TSUE}}/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}^{\mathrm{TSUE}}/\theta)}$ and $h_s^{\mathrm{TSUE}} < q^k \frac{\exp(-c_s^{\mathrm{TSUE}}/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}^{\mathrm{TSUE}}/\theta)}$. Let $\pi^k = -\theta \log\left(\sum_{r' \in R^k} \exp(-c_{r'}^{\mathrm{TSUE}}/\theta)\right)$. Then we have that

$$c_r^{\mathrm{TSUE}} + \theta \log \frac{h_r^{\mathrm{TSUE}}}{q^k} > \pi^k > c_s^{\mathrm{TSUE}} + \theta \log \frac{h_s^{\mathrm{TSUE}}}{q^k}. \tag{3.15}$$

Consider the path flow vector defined in the following

$$h_{r'} = \begin{cases} h_{r'}^{\mathrm{TSUE}} & r' \neq r,\, r' \neq s, \\ 0 & r' = r, \\ h_r^{\mathrm{TSUE}} + h_s^{\mathrm{TSUE}} & r' = s. \end{cases}$$

This is clearly a feasible path flow, i.e., $\boldsymbol{h} \in S_h$. But

$$\begin{aligned}
&\left(\boldsymbol{c}^{\text{TSUE}} + \theta \log \frac{\boldsymbol{h}^{\text{TSUE}}}{\boldsymbol{q}}\right)^{\text{T}} \left(\boldsymbol{h} - \boldsymbol{h}^{\text{TSUE}}\right) \\
=& - \left(c_r^{\text{TSUE}} + \theta \log \frac{h_r^{\text{TSUE}}}{q^k}\right) h_r^{\text{TSUE}} + \left(c_s^{\text{TSUE}} + \theta \log \frac{h_s^{\text{TSUE}}}{q^k}\right) h_r^{\text{TSUE}} \\
<& 0.
\end{aligned}$$

The last line follows from Eq. (3.15) and the fact that $h_r^{\text{TSUE}} > 0$. The above inequality contradicts with the TSUE-VI formulation. Hence we proved that TSUE-VI implies TSUE-FIX. □

It is easy to see that when $\lambda = 0$, the TSUE-VI formulation reduces to the SUE-VI formulation (Eq. (2.8)).

We further note that if we set $\lambda = 0$, and substitute the first-order approximation of the travel time function in the TSUE model with a second-order approximation, we obtain the GSUE model developed by Watling (2002). In the GSUE model, the travel time function is approximated by, for any link $l \in L$,

$$t_l(\boldsymbol{F}) \approx \hat{\hat{t}}_l(\boldsymbol{F}) = t_l(\boldsymbol{f}) + \left(\boldsymbol{\nabla} t_l(\boldsymbol{f})\right)^{\text{T}} \left(\boldsymbol{F} - \boldsymbol{f}\right) + \frac{1}{2} \left(\boldsymbol{F} - \boldsymbol{f}\right)^{\text{T}} \boldsymbol{\nabla}^2 t_l(\boldsymbol{f}) \left(\boldsymbol{F} - \boldsymbol{f}\right),$$

where $\boldsymbol{\nabla} t_l(\boldsymbol{f})$ and $\boldsymbol{\nabla}^2 t_l(\boldsymbol{f})$ respectively give the gradient of Hessian matrix of function t_l at $\boldsymbol{f}$. Hence, the expected link travel time is

$$\mathbb{E}[\hat{\hat{t}}_l(\boldsymbol{F})] = t_l(\boldsymbol{f}) + \frac{1}{2} \left\|\boldsymbol{\nabla}^2 t_l(\boldsymbol{f}), \text{Cov}(\boldsymbol{F})\right\|,$$

where the scalar product of two $m \times n$ matrices is defined as $\|\boldsymbol{X}, \boldsymbol{Y}\| = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij}$, and $\text{Cov}(\boldsymbol{F})$ is the covariance matrix of the link flows. We have $\text{Cov}(\boldsymbol{F}) = \boldsymbol{\Delta} \text{Cov}(\boldsymbol{H}) \boldsymbol{\Delta}^{\text{T}}$.

The (actual) travel cost of the GSUE model is

$$\begin{aligned} \boldsymbol{c}^{\text{GSUE}} &= \mathbb{E}[\boldsymbol{\Delta}^{\text{T}}\hat{\hat{\boldsymbol{t}}}(\boldsymbol{F})] \\ &= \boldsymbol{\Delta}^{\text{T}}\boldsymbol{t}(\boldsymbol{f}) + \frac{1}{2}\boldsymbol{\Delta}^{\text{T}} \left\| \boldsymbol{\nabla}^2\boldsymbol{t}(\boldsymbol{f}), \boldsymbol{\Delta}\text{Cov}(\boldsymbol{H})\boldsymbol{\Delta}^{\text{T}} \right\|. \end{aligned} \tag{3.16}$$

For networks with affine travel time functions, the GSUE model reduces to the SUE model.

We provide the TSUE travel cost function for comparison

$$\begin{aligned} \boldsymbol{c}^{\text{TSUE}} &= \mathbb{E}[\boldsymbol{\Delta}^{\text{T}}\hat{\boldsymbol{t}}(\boldsymbol{F})] + \lambda\text{Var}(\boldsymbol{\Delta}^{\text{T}}\hat{\boldsymbol{t}}(\boldsymbol{F})) \\ &= \boldsymbol{\Delta}^{\text{T}}\boldsymbol{t}(\boldsymbol{f}) + \lambda\text{Diag}\left\{\boldsymbol{\Delta}^{\text{T}}\boldsymbol{J}_t\left(\boldsymbol{f}\right)\boldsymbol{\Delta}\text{Cov}\left(\boldsymbol{H}\right)\boldsymbol{\Delta}^{\text{T}}\boldsymbol{J}_t\left(\boldsymbol{f}\right)\boldsymbol{\Delta}\right\}. \end{aligned} \tag{3.17}$$

Both the TSUE and GSUE travel cost functions involve the covariance matrix of link flows ($\boldsymbol{\Delta}\text{Cov}\left(\boldsymbol{H}\right)\boldsymbol{\Delta}^{\text{T}}$). However, while the TSUE model relates to the first-order derivatives of the travel time function, the GSUE model relates to the second-order derivatives of the travel time function. The TSUE model incorporates the travel time uncertainty in the travel cost, and hence this model can capture the influence of travel time uncertainty on travelers' behavior. This is not accomplished by the GSUE model.

3.2 Alternative Approaches to Modeling Travel Costs

This section discusses alternative approaches to modeling travel costs that have been previously used in transportation studies. These approaches share the same purpose of turning the stochastic travel times into travel costs. Similar to the fact that uncertainty in investment returns cause financial risk to investors, uncertain travel times are also a form of risk for travelers. Risk can be modeled using operators called risk measures. Given any distribution of travel times, a risk measure returns a deterministic scalar travel cost, which represents the risk level experienced by the traveler. To be specific, let $\mathcal{R}(\cdot)$ be a risk measure. Then for any path $r \in R$, we have that

$$c_r = \mathcal{R}(M_r), \tag{3.18}$$

where $c_r \in \mathbb{R}$ is the travel cost for path r.

Previous studies of stochastic traffic networks models have adopted different approaches to choosing risk measures, and hence these studies used different forms of operator $\mathcal{R}$ in Eq. (3.18). We list some of these approaches below. We restrict our attention to traffic network modeling. For a general purpose review of optimization under uncertainty, see Rockafellar (2007); for using random utility models in the study of decision making under uncertainty, see de Palma et al. (2008). The following list attempts to be comprehensive but not necessarily exhaustive.

Approach 1: Guessing a point

This approach identifies a single point in the travel time distribution as an estimate of the travel cost. This is a commonly used approach in practice that essentially avoids the issue of travel time uncertainty. Namely, let μ_r be the identified travel time, we have that

$$c_r = \mu_r.$$

We point out that the SUE model can be interpreted as using this approach. By interpreting the path choice probabilities as path choice proportions and obtaining deterministic traffic flows, the SUE model essentially identifies a single point in the traffic flow space, which by Eq. (3.2) is $\boldsymbol{h} = \mathbb{E}[\boldsymbol{H}]$. Meanwhile, the SUE model also identifies a single point in the travel time distribution, namely the travel time decided by the identified traffic flow, which is $\boldsymbol{\mu} = \boldsymbol{\Delta}^{\mathrm{T}}(\boldsymbol{\Delta h})$. The SUE model proceeds to take the travel time $\boldsymbol{\mu}$ to be the same as the travel cost, that is, $\boldsymbol{c} = \boldsymbol{\mu}$. Therefore, the SUE model essentially obtains the travel costs by a single point estimation.

Approach 2: Relying on expectation

In this case, the travel cost is the expected travel time, namely

$$c_r = \mathbb{E}[M_r]. \tag{3.19}$$

This approach corresponds to a special case of the expected utility approach that we will introduce below. Although Eq. (3.19) is relatively easy to evaluate, this approach effectively assumes that travelers are "risk-neutral". As a result, it fails to explicitly account for the impact of travel time uncertainty on the travel cost.

Travel costs defined as expected travel times will be distinguished from the travel costs in the SUE model. In the SUE model, $\boldsymbol{c} = \boldsymbol{\mu} = \boldsymbol{\Delta}^{\mathrm{T}}(\boldsymbol{\Delta h}) = \boldsymbol{\Delta}^{\mathrm{T}}(\boldsymbol{\Delta}\mathbb{E}[\boldsymbol{H}])$. But Eq. (3.19) implies that $\boldsymbol{c} = \mathbb{E}[\boldsymbol{M}] = \mathbb{E}[\boldsymbol{\Delta}^{\mathrm{T}}(\boldsymbol{\Delta H})]$. These costs coincide if the flow-time function is affine, but in general they are different.

This approach was adopted in the GSUE model studied by Watling (2002).

Approach 3: Expected utility analysis

Expected utility (EU) theory is perhaps the most widely used theory to study choices under uncertainty. It proposes a utility function that transforms the uncertain outcomes (here the travel times) into utilities and define the cost as the expected utility. Let $u(\cdot)$ be the utility function, then we have that

$$c_r = \mathbb{E}[u(M_r)]. \tag{3.20}$$

Mirchandani and Soroush (1987) used this approach for stochastic network modeling. Barberis and Thaler (2003) reviewed the drawbacks of the expected utility approach. A few types of commonly used utility functions would result in the following risk measures

- Linear utility function. In this case $c_r = \mathbb{E}[M_r]$;
- Quadratic utility function. In this case $c_r = \mathbb{E}[M_r + bM_r^2]$, where b is a constant parameter; and
- Exponential utility function. In this case $c_r = \mathbb{E}[\exp(bM_r)]$, where b is a constant parameter.

Approach 4: Cumulative prospect theory analysis

Tversky and Kahneman (1992) proposed a new framework for choice modeling. In this framework, the outcomes of uncertain events are transformed by a utility function (as in the case of expected utility), meanwhile the probability weights of these outcomes are also transformed by a probability weighting function to represent travelers' perceptions of these probabilities. In a recent work, Connors and Sumalee (2009) proposed a stochastic network model that adopts this approach. Interested readers should refer to their work for the details of this approach.

Approach 5: Worst-case analysis

The robust approach uses a fundamentally different idea to characterize travel time uncertainty. Instead of assuming a probabilistic distribution of travel times, this approach considers only the worst case travel time and hence is best suited to capture the behavior of risk averse travelers. We include this approach here for completeness.

3.3 Existence and Uniqueness of the Equilibrium Solution

In this section we discuss the existence and uniqueness of the TSUE equilibrium solution.

3.3.1 Existence of the equilibrium solution

Existence of the equilibrium solution to the TSUE model is stated in the following theorem. To prove this theorem, we apply Brouwer's fixed point theorem to the TSUE-FIX formulation.

Theorem 3.2. *If the flow-time function $\boldsymbol{t}(\cdot)$ is continuously differentiable, then the TSUE model has at least one solution in S_h.*

Proof. Consider the following function $\boldsymbol{U}$ that maps from S_h to S_h, with

$$U_r(\boldsymbol{h}) = q^k \frac{\exp(-c_r/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}/\theta)}, \quad \forall k \in K,\ r \in R^k, \tag{3.21}$$

where $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{h})$ as defined in Eq. (3.11).

By the TSUE-FIX formulation, finding the solutions to the TSUE model is equivalent to finding the fixed points of $\boldsymbol{U}$. We now claim that $\boldsymbol{U}$ is a continuous function of $\boldsymbol{h}$. To prove this, it suffices to show that $\boldsymbol{c}$ is a continuous function of $\boldsymbol{h}$. Since $\boldsymbol{t}(\cdot)$ is continuously differentiable, both $\boldsymbol{t}(\cdot)$ and $\boldsymbol{J}_t(\cdot)$ are also continuous functions. By Eq. (3.11) we know that $\boldsymbol{c}$ is a continuous functions of $\boldsymbol{h}$.

Thus $\boldsymbol{U}$ is a continuous function on a convex compact set S_h. By Brouwer's fixed point theorem, it must have at least one fixed point. □

3.3.2 Uniqueness of the equilibrium solution

In general, the TSUE model can have multiple equilibria. Recall the proof for the uniqueness of the SUE equilibrium solution (see, for example, Cascetta, 2001). The literature requires that the travel cost function is monotone.

Definition 3.3. The cost function $\boldsymbol{c}(\cdot)$ is *monotone* if for any $\boldsymbol{h}, \tilde{\boldsymbol{h}} \in S_h$, it satisfy the following condition

$$\left(\boldsymbol{c}(\boldsymbol{h}) - \boldsymbol{c}(\tilde{\boldsymbol{h}})\right)^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}}\right) \geqslant 0.$$

If the inequality holds as a strict inequality for all $\boldsymbol{h} \neq \tilde{\boldsymbol{h}}$, then the cost function $\boldsymbol{c}(\cdot)$ is called *strictly monotone*.

Under this condition, the VI formulation of the SUE model implies that this model must have a unique solution (for a proof, see Lemma 3.5).

For the TSUE model, the cost function has two components. One component is the expected travel time (this is the same as in the SUE model). The other component is the variance of the travel time multiplied by λ. While under mild assumptions we can easily prove that the first term is monotone, the second term can break the monotonicity assumption, because the variance of the travel time is in general not a monotonic function of the path choice probabilities. Fortunately, when λ is relatively small, the non-monotonicity of the variance term is dominated by the monotonicity of the expectation term. As a result, the travel cost remains a monotone function, and the uniqueness of the equilibrium solution still holds. In what follows we prove this statement formally.

In the following discussion, we first present a few lemmas before we prove the uniqueness theorem.

Lemma 3.4. *If $\boldsymbol{h} \neq \tilde{\boldsymbol{h}} \in S_h$, with $\boldsymbol{h}, \tilde{\boldsymbol{h}} > 0$, then the following inequality holds*

$$\left(\log \frac{\boldsymbol{h}}{\boldsymbol{q}} - \log \frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}}\right) > 0.$$

Proof. For any OD pair $k \in K$, consider a path $r \in R^k$, since $\log p_r$ is a strictly increasing function of p_r, we have that

$$(\log p_r - \log \tilde{p}_r)(p_r - \tilde{p}_r) > 0.$$

Multiplying by q^k, we have that

$$\left(\log \frac{h_r}{q^k} - \log \frac{\tilde{h}_r}{q^k}\right)\left(h_r - \tilde{h}_r\right) \geqslant 0.$$

By summing up the above inequalities over all paths, we obtain the result in the lemma. □

Remark. In Chapter 4 we will prove a stronger version of this lemma. See Corollary 4.5.

Lemma 3.5. *Consider a traffic network model. If the path travel cost function $\boldsymbol{c}(\boldsymbol{h})$ is monotone, that is,*

$$\left(\boldsymbol{c}(\boldsymbol{h}) - \boldsymbol{c}(\tilde{\boldsymbol{h}})\right)^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}}\right) \geqslant 0, \quad \forall \boldsymbol{h}, \tilde{\boldsymbol{h}} \in S_h,$$

then the TSUE model has at most one solution.

Proof. Suppose, by contradiction, that the TSUE model has two different solutions

$\boldsymbol{h} \neq \tilde{\boldsymbol{h}}$. By its VI formulation, we have that

$$\left(\boldsymbol{c}(\boldsymbol{h}) + \theta \log \frac{\boldsymbol{h}}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\tilde{\boldsymbol{h}} - \boldsymbol{h}\right) = 0,$$
$$\left(\boldsymbol{c}(\tilde{\boldsymbol{h}}) + \theta \log \frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}}\right) = 0.$$

Summing up these two inequalities, we have that

$$\left(\boldsymbol{c}(\boldsymbol{h}) - \boldsymbol{c}(\tilde{\boldsymbol{h}})\right)^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}}\right) + \theta \left(\log \frac{\boldsymbol{h}}{\boldsymbol{q}} - \log \frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}}\right) = 0.$$

By assumption, the first term in the left hand side of the above equation is non-negative. By Lemma 3.4, the second term is positive, hence

$$\left(\boldsymbol{c}(\boldsymbol{h}) - \boldsymbol{c}(\tilde{\boldsymbol{h}})\right)^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}}\right) + \theta \left(\log \frac{\boldsymbol{h}}{\boldsymbol{q}} - \log \frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}}\right) > 0,$$

which is a contradiction. □

Lemma 3.6. *Let* $\boldsymbol{x} \in \mathbb{R}^n$ *be a vector. Let* $|\boldsymbol{x}|$ *represent the Euclidean vector norm of* $\boldsymbol{x}$*. Then the following relationship holds*

$$|\boldsymbol{x}| \leqslant \sum_{i=1}^{n} |x_i| \leqslant \sqrt{n}\, |\boldsymbol{x}| \, .$$

Proof. To prove the stated inequality, we only need to prove the following

$$\sum_{i=1}^{n} x_i^2 \leqslant \left(\sum_{i=1}^{n} |x_i|\right)^2 \leqslant n \sum_{i=1}^{n} x_i^2.$$

The left hand side follows easily. The right hand side follows from Cauchy's inequality. □

Theorem 3.7. *Consider a traffic network with*

- *an affine travel time function* $\boldsymbol{t}(\boldsymbol{f}) = \boldsymbol{G}\boldsymbol{f} + \boldsymbol{b}$*, where* $\boldsymbol{G}$ *is a positive definite matrix,*

- *all the elements of matrix* $\boldsymbol{G}$ *are non-negative, and*

- *the mean-variance trade-off factor* $0 < \lambda \leqslant \frac{1}{2\sqrt{|R|}} \frac{\nu_{\min}(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta})}{\nu_{\max}(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta})^2}$*, where* $|R|$ *is the number of paths,* $\boldsymbol{\Delta}$ *is the link-path incidence matrix, and* $\nu_{\min}(\cdot)$ *and* $\nu_{\max}(\cdot)$ *respectively give the smallest and largest eigenvalues of a matrix.*

Then the TSUE model has at most one solution.

Remark. This is the main theorem proving the uniqueness of the solution. Briefly, it states that when λ is sufficiently small, the TSUE solution is unique. In fact, the upper bound on λ given in this theorem is more restrictive than actually necessary. Computationally, even if we choose values greater than this upper bound for λ, the TSUE model can still have a unique solution.

Proof. By Lemma 3.5, to prove this theorem, we only need to prove that $\boldsymbol{c}(\boldsymbol{h})$ is a monotone function of $\boldsymbol{h}$. Using Eq. (3.11), this means that for any $\boldsymbol{h}, \tilde{\boldsymbol{h}} \in S_h$, the following expression is nonnegative.

$$\begin{aligned}
&\Big(\boldsymbol{c}(\boldsymbol{h}) - \boldsymbol{c}(\tilde{\boldsymbol{h}})\Big)^{\mathrm{T}} (\boldsymbol{h} - \tilde{\boldsymbol{h}}) \\
=&\Big(\boldsymbol{\Delta}^{\mathrm{T}} (\boldsymbol{G}\boldsymbol{\Delta}\boldsymbol{h} + \boldsymbol{b}) - \boldsymbol{\Delta}^{\mathrm{T}}(\boldsymbol{G}\boldsymbol{\Delta}\tilde{\boldsymbol{h}} + \boldsymbol{b})\Big)^{\mathrm{T}} (\boldsymbol{h} - \tilde{\boldsymbol{h}}) \\
&+ \lambda \mathrm{Diag}\big\{\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G} \sum_{k\in K} \Big(\boldsymbol{\Delta}^k\big(\mathrm{Cov}(\boldsymbol{H}^k) - \mathrm{Cov}(\tilde{\boldsymbol{H}}^k)\big)(\boldsymbol{\Delta}^k)^{\mathrm{T}}\Big)\boldsymbol{G}\boldsymbol{\Delta}\big\}^{\mathrm{T}}(\boldsymbol{h} - \tilde{\boldsymbol{h}}).
\end{aligned} \tag{3.22}$$

We consider the two terms in the above expression separately. For the first term in Eq. (3.22), we have that

$$\begin{aligned}
&\left(\boldsymbol{\Delta}^{\mathrm{T}}\left(\boldsymbol{G}\boldsymbol{\Delta}\boldsymbol{h}+\boldsymbol{b}\right)-\boldsymbol{\Delta}^{\mathrm{T}}(\boldsymbol{G}\boldsymbol{\Delta}\tilde{\boldsymbol{h}}+\boldsymbol{b})\right)^{\mathrm{T}}(\boldsymbol{h}-\tilde{\boldsymbol{h}})\\
=&(\boldsymbol{h}-\tilde{\boldsymbol{h}})^{\mathrm{T}}\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}(\boldsymbol{h}-\tilde{\boldsymbol{h}})\\
\geqslant&\nu_{\min}\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right)\left|\boldsymbol{h}-\tilde{\boldsymbol{h}}\right|^{2},
\end{aligned}\tag{3.23}$$

where $|\cdot|$ represents Euclidean vector norm and $\nu_{\min}(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta})$ gives the smallest eigenvalue of $\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}$. Note that $\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}$ is a symmetric and positive definite matrix, hence all its eigenvalues are real and positive.

For the second term in Eq. (3.22), consider OD pair $k \in K$. For $s, s' \in R^k$, the (s, s') component of the matrix $\left(\mathrm{Cov}(\boldsymbol{H}^k) - \mathrm{Cov}(\tilde{\boldsymbol{H}}^k)\right)$ is

$$\begin{aligned}
&\left(\mathrm{Cov}(\boldsymbol{H}^k)-\mathrm{Cov}(\tilde{\boldsymbol{H}}^k)\right)_{s,s'}\\
&=q^k\left(\mathrm{diag}\{\boldsymbol{p}^k-\tilde{\boldsymbol{p}}^k\}-\boldsymbol{p}^k(\boldsymbol{p}^k)^{\mathrm{T}}+\tilde{\boldsymbol{p}}^k(\tilde{\boldsymbol{p}}^k)^{\mathrm{T}}\right)_{s,s'}\\
&=q^k\left((p_s-\tilde{p}_s)\delta_{s,s'}-p_sp_{s'}+\tilde{p}_s\tilde{p}_{s'}\right)\\
&=q^k\left((p_s-\tilde{p}_s)\delta_{s,s'}-p_sp_{s'}+\tilde{p}_sp_{s'}-\tilde{p}_sp_{s'}+\tilde{p}_s\tilde{p}_{s'}\right)\\
&=q^k\left((p_s-\tilde{p}_s)(\delta_{s,s'}-p_{s'})+\tilde{p}_s(\tilde{p}_{s'}-p_{s'})\right),
\end{aligned}$$

where $\delta_{s,s'}$ is the Kronecker delta function. Since $|\delta_{s,s'} - p_{s'}| \leqslant 1$ and $|\tilde{p}_s| \leqslant 1$, we have that

$$\begin{aligned}
&\left|\left(\mathrm{Cov}(\boldsymbol{H}^k)-\mathrm{Cov}(\tilde{\boldsymbol{H}}^k)\right)_{s,s'}\right|\\
=&q^k\left|(p_s-\tilde{p}_s)(\delta_{s,s'}-p_{s'})+\tilde{p}_s(\tilde{p}_{s'}-p_{s'})\right|\\
\leqslant&q^k\left|(p_s-\tilde{p}_s)(\delta_{s,s'}-p_{s'})\right|+q^k\left|\tilde{p}_s(\tilde{p}_{s'}-p_{s'})\right|\\
\leqslant&\left|h_s-\tilde{h}_s\right|+\left|\tilde{h}_{s'}-h_{s'}\right|.
\end{aligned}\tag{3.24}$$

To simplify our notations, let $\bar{\boldsymbol{G}} = \boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}$. We have that

$$
\begin{aligned}
&\left(\operatorname{Diag}\left\{\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\sum_{k\in K}\Big(\boldsymbol{\Delta}^{k}\big(\operatorname{Cov}(\boldsymbol{H}^{k})-\operatorname{Cov}(\tilde{\boldsymbol{H}}^{k})\big)(\boldsymbol{\Delta}^{k})^{\mathrm{T}}\Big)\boldsymbol{G}\boldsymbol{\Delta}\right\}\right)_{r}\\
=&\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\sum_{k\in K}\Big(\boldsymbol{\Delta}^{k}\big(\operatorname{Cov}(\boldsymbol{H}^{k})-\operatorname{Cov}(\tilde{\boldsymbol{H}}^{k})\big)(\boldsymbol{\Delta}^{k})^{\mathrm{T}}\Big)\boldsymbol{G}\boldsymbol{\Delta}\right)_{r,r}\\
=&\sum_{k\in K}\sum_{s,s'\in R^{k}}\bar{G}_{r,s}\bar{G}_{r,s'}\left(\operatorname{Cov}(\boldsymbol{H}^{k})-\operatorname{Cov}(\tilde{\boldsymbol{H}}^{k})\right)_{s,s'}.
\end{aligned}
$$

Because the elements of both matrices $\boldsymbol{\Delta}$ and $\boldsymbol{G}$ are nonnegative, the elements of matrix $\bar{\boldsymbol{G}}$ must be nonnegative. Taking the norm and applying Eq. (4.18), we have that

$$
\begin{aligned}
&\left|\left(\operatorname{Diag}\left\{\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\sum_{k\in K}\Big(\boldsymbol{\Delta}^{k}\big(\operatorname{Cov}(\boldsymbol{H}^{k})-\operatorname{Cov}(\tilde{\boldsymbol{H}}^{k})\big)(\boldsymbol{\Delta}^{k})^{\mathrm{T}}\Big)\boldsymbol{G}\boldsymbol{\Delta}\right\}\right)_{r}\right|\\
\leqslant&\sum_{k\in K}\sum_{s,s'\in R^{k}}\bar{G}_{r,s}\bar{G}_{r,s'}\left|\left(\operatorname{Cov}(\boldsymbol{H}^{k})-\operatorname{Cov}(\tilde{\boldsymbol{H}}^{k})\right)_{s,s'}\right|\\
\leqslant&\sum_{k\in K}\sum_{s,s'\in R^{k}}\bar{G}_{r,s}\bar{G}_{r,s'}\left(\left|h_{s}-\tilde{h}_{s}\right|+\left|h_{s'}-\tilde{h}_{s'}\right|\right)\\
=&2\sum_{s,s'\in R}\bar{G}_{r,s}\bar{G}_{r,s'}\left|h_{s}-\tilde{h}_{s}\right|.
\end{aligned}
$$

Applying Lemma 3.6, we have that

$$
\begin{aligned}
&\left|\operatorname{Diag}\left\{\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\sum_{k\in K}\Big(\boldsymbol{\Delta}^{k}\big(\operatorname{Cov}(\boldsymbol{H}^{k})-\operatorname{Cov}(\tilde{\boldsymbol{H}}^{k})\big)(\boldsymbol{\Delta}^{k})^{\mathrm{T}}\Big)\boldsymbol{G}\boldsymbol{\Delta}\right\}\right|\\
\leqslant&2\sum_{r\in R}\sum_{s,s'\in R}\bar{G}_{r,s}\bar{G}_{r,s'}\left|h_{s}-\tilde{h}_{s}\right|\\
=&2\boldsymbol{1}\bar{\boldsymbol{G}}\bar{\boldsymbol{G}}\boldsymbol{d}\\
\leqslant&2\left|\boldsymbol{1}\right|\left\|\bar{\boldsymbol{G}}\bar{\boldsymbol{G}}\right\|\left|\boldsymbol{d}\right|\\
=&2\sqrt{|R|}\nu_{\max}\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right)^{2}\left|\boldsymbol{h}-\tilde{\boldsymbol{h}}\right|,
\end{aligned}
$$

where $\mathbf{1}$ is the column vector of all 1's (with dimension $|R| \times 1$), $\boldsymbol{d}$ is the vector with its sth component being $\left|h_s - \tilde{h}_s\right|$, and $\|\cdot\|$ denotes the Euclidean matrix norm. The last line follows from the facts that (1) $|\mathbf{1}| = \sqrt{|R|}$, $|\boldsymbol{d}| = \left|\boldsymbol{h} - \tilde{\boldsymbol{h}}\right|$, and (2)$\left\|\bar{\boldsymbol{G}}\bar{\boldsymbol{G}}\right\| = \left\|\bar{\boldsymbol{G}}\right\|^2 = \left\|\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right\|^2 = \nu_{\max}\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right)^2$.

Now return to the second term in Eq. (3.22). We have that

$$\begin{aligned}
&\operatorname{Diag}\big\{\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\sum_{k\in K}\Big(\boldsymbol{\Delta}^k\big(\operatorname{Cov}(\boldsymbol{H}^k) - \operatorname{Cov}(\tilde{\boldsymbol{H}}^k)\big)(\boldsymbol{\Delta}^k)^{\mathrm{T}}\Big)\boldsymbol{G}\boldsymbol{\Delta}\big\}^{\mathrm{T}}(\boldsymbol{h} - \tilde{\boldsymbol{h}}) \\
\geqslant &-\left|\operatorname{Diag}\left\{\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\sum_{k\in K}\Big(\boldsymbol{\Delta}^k\big(\operatorname{Cov}(\boldsymbol{H}^k) - \operatorname{Cov}(\tilde{\boldsymbol{H}}^k)\big)(\boldsymbol{\Delta}^k)^{\mathrm{T}}\Big)\boldsymbol{G}\boldsymbol{\Delta}\right\}\right|\left|\boldsymbol{h} - \tilde{\boldsymbol{h}}\right| \\
\geqslant &-2\sqrt{|R|}\nu_{\max}\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right)^2\left|\boldsymbol{h} - \tilde{\boldsymbol{h}}\right|^2.
\end{aligned} \tag{3.25}$$

Plugging Eq. (3.23) and (3.25) to Eq. (3.22), we have that

$$\begin{aligned}
&\left(\boldsymbol{c}(\boldsymbol{h}) - \boldsymbol{c}(\tilde{\boldsymbol{h}})\right)^{\mathrm{T}}(\boldsymbol{h} - \tilde{\boldsymbol{h}}) \\
\geqslant &\left(\nu_{\min}\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right) - 2\lambda\sqrt{|R|}\nu_{\max}\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right)^2\right)\left|\boldsymbol{h} - \tilde{\boldsymbol{h}}\right|^2.
\end{aligned}$$

Therefore, when $\lambda \leqslant \frac{1}{2\sqrt{|R|}}\frac{\nu_{\min}\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right)}{\nu_{\max}\left(\boldsymbol{\Delta}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\Delta}\right)^2}$, the above expression is nonnegative. By Lemma 3.5, the TSUE model has at most one solution. □

3.4 Solution Algorithms

In the previous sections, we have formulated the TSUE model both as a fixed point problem (Eq. (3.13)) and as a variational inequality problem (Eq. (3.14)). In Theorem 3.7, we provided conditions under which the TSUE travel cost $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{h})$, defined in Eq. (3.11), is a monotone function. Under this condition, and by Lemma 3.4, function $\left(\boldsymbol{c}(\boldsymbol{h}) + \theta \log \frac{\boldsymbol{h}}{\boldsymbol{q}}\right)$ in the variational inequality formulation is strictly monotone. Many algorithms already exisit for solving variational inequality problems with strictly monotone (or strongly monotone) functions (see, for example, the projection and relaxation method discussed in Nagurney, 1999). We can solve the TSUE model by applying these algorithms.

However, the conditions developed in Theorem 3.7 are very restrictive. In general cases we can not ensure the strict monotonicity of function $\left(c(h) + \theta \log \frac{\boldsymbol{h}}{\boldsymbol{q}}\right)$, and hence we are not guaranteed that the traditional algorithms in the literature will converge to a solution. This motivates us to develop special algorithms to solve the TSUE model in order to find the equilibrium solution for general networks.

In this section we describe two algorithms for solving the TSUE model. The first algorithm (using the PATH solver) only works for small-scale networks ($|R| < 50$). The second algorithm (a modified Frank-Wolfe algorithm) works well for medium and large scale networks ($|R| \geqslant 50$). Nevertheless, we would like to point out that for small-scale networks, the first algorithm is faster.

3.4.1 An algorithm solving the mixed complementarity formulation

To apply this algorithm, we first reformulate the problem as a set of nonlinear equations. We interpret these equations as a mixed complementarity problem, and then apply a standard solver (PATH) to solve the mixed complementarity problem.

The fixed point formulation of the TSUE model (Eq. (3.13)) implies that, for any $k \in K$, $r \in R^k$, the following is true

$$c_r^{\text{TSUE}} + \theta \log \frac{h_r^{\text{TSUE}}}{q^k} = -\theta \log \left(\sum_{r' \in \mathcal{P}^k} \exp \left(-\frac{c_{r'}^{\text{TSUE}}}{\theta} \right) \right).$$

We define the right hand side as a new variable $\boldsymbol{\pi}$, i.e., for any $k \in K$, define the component π^k of vector $\boldsymbol{\pi}$ as

$$\pi^k = -\theta \log \left(\sum_{r' \in \mathcal{P}^k} \exp \left(-\frac{c_{r'}^{\text{TSUE}}}{\theta} \right) \right).$$

Then we have that

$$c^{\mathrm{TSUE}} + \theta \log \frac{\boldsymbol{h}^{\mathrm{TSUE}}}{\boldsymbol{q}} - \boldsymbol{\Gamma}^{\mathrm{T}} \boldsymbol{\pi} = 0.$$

With this new variable, the TSUE model can be formulated as a set of nonlinear equations, that is, find $\boldsymbol{h} \in \mathbb{R}_+^{|R|}$ and $\boldsymbol{\pi} \in \mathbb{R}^{|K|}$, such that

$$\begin{pmatrix} \boldsymbol{c}(\boldsymbol{h}) + \theta \log \frac{\boldsymbol{h}}{\boldsymbol{q}} - \boldsymbol{\Gamma}^{\mathrm{T}} \boldsymbol{\pi} \\ \boldsymbol{\Gamma h} - \boldsymbol{q} \end{pmatrix} = \mathbf{0}. \tag{3.26}$$

Here the presence of the term $\log \frac{\boldsymbol{h}}{\boldsymbol{q}}$ requires that $\boldsymbol{h} > 0$. This requirement, together with equation $\boldsymbol{\Gamma h} - \boldsymbol{q} = \mathbf{0}$, ensures that $\boldsymbol{h} \in S_h$. The above formulation can be solved using standard methods described in text books (see, for example, Chapter 11 of Nocedal and Wright, 2006).

Eq. (3.26) can also be viewed as a mixed complementarity problem (MCP). In this thesis, we use the PATH solver (see Ferris and Munson, 1999) to solve the TSUE model.

3.4.2 The modified Frank-Wolfe algorithm

The second algorithm is a modification of the Frank-Wolfe algorithm that is commonly used to solve traffic equilibrium models (see, for example, Bell and Iida, 1997). The Frank-Wolfe algorithm and its modifications are frequently used to solve the SUE model (see, for example, Damberg et al. 1996, Maher 1998, and Bekhor and Toledo 2005).

The proposed algorithm for the TSUE model is as follows.

1. (initialization)

   (a) Set $p_r' \leftarrow \frac{1}{|R^k|}$, for all $k \in K$, $r \in R^k$.

   (b) Set $n \leftarrow 1$.

2. (auxiliary link flows)

   (a) Set $\boldsymbol{c}' \leftarrow \boldsymbol{c}(\boldsymbol{p}')$.

   (b) Set $p_r^* \leftarrow \frac{\exp(-c_r/\theta)}{\sum_{r' \in R^k} \exp(-c_{r'}/\theta)}$, for $k \in K$, $r \in R^k$.

3. (approximate line search)

   (a) Set $\boldsymbol{p}' \leftarrow$ the minimizer of $|U(\boldsymbol{p}) - \boldsymbol{p}|$ where $\boldsymbol{p} = \boldsymbol{p}' + \alpha(\boldsymbol{p}^* - \boldsymbol{p}')$, with $0 \leqslant \alpha \leqslant \frac{1}{n}$. This is done by an approximate line search.

   (b) Set $n \leftarrow n + 1$.

   (c) If the convergence criterion is satisfied (i.e., $|U(\boldsymbol{p}) - \boldsymbol{p}| < 10^{-4}$), exit the algorithm. Otherwise, go to Step 2.

In the above algorithm, function $U(\boldsymbol{p})$ is defined in Eq. (3.21).

The approximate line search in Step 3(a) is performed using a heuristic algorithm. Many standard line search algorithms exist in the literature (see, for example, Nocedal and Wright, 2006 and Gill et al., 1982). However, these algorithms typically work under the conditions that (1) the objective function is convex and/or (2) the moving direction is a descent direction. Neither of these two conditions necessarily hold in our case. Hence we develop the following heuristic line search algorithm.

The purpose of this heuristic is to decide an optimal step length that minimizes the objective defined in Step 3(a). For this purpose, we sequentially try the following step lengths: $\alpha = \frac{1}{n}$, $\alpha = \frac{1}{2n}$, $\alpha = \frac{1}{4n}$, $\alpha = \frac{1}{8n}$, etc. Let $y(\alpha) = |\boldsymbol{U}(\boldsymbol{p}' + \alpha(\boldsymbol{p}^* - \boldsymbol{p}')) - (\boldsymbol{p}' + \alpha(\boldsymbol{p}^* - \boldsymbol{p}'))|$ be the objective function. Ideally we arrive at some step length $\alpha = \alpha_0$ such that $y(\alpha_0) \leqslant y(0)$ and $y(\alpha_0) \leqslant y(2\alpha_0)$. This ensures that function y has a local minimum in $[0, 2\alpha_0]$. Then we make a quadratic approximation of function y using its value at $\alpha = 0$, $\alpha = \alpha_0$, and $\alpha = 2\alpha_0$, and set the step length $\alpha = \alpha^*$ to be the minimizer of this quadratic approximation (α^* must satisfy $0 \leqslant \alpha^* \leqslant 2\alpha_0$ because $y(\alpha_0) \leqslant y(0)$ and $y(\alpha_0) \leqslant y(2\alpha_0)$). If we cannot find a step length α_0 that satisfies the above requirements after trying for 10 different step lengths, we will set $\alpha = \frac{1}{n}$. The line search algorithm is as follows

1. (initialization)

   (a) Set $maxi \leftarrow 10$.

   (b) Set $i \leftarrow 1$.

   (c) Set $\alpha \leftarrow \frac{1}{n}$.

   (d) Set $obj(0) \leftarrow |U(\boldsymbol{p}') - \boldsymbol{p}'|$

2. (try a new point)

   (a) Set $\boldsymbol{p} \leftarrow \boldsymbol{p}' + \alpha(\boldsymbol{p}^* - \boldsymbol{p}')$.

   (b) Set $obj(i) \leftarrow |U(\boldsymbol{p}) - \boldsymbol{p}|$.

   (c) If $i \geqslant 2$ and $obj(i) \geqslant obj(i-1)$, then set $\alpha = \frac{1}{n}$ and go to Step 4.

   (d) If $i = maxi$, then set $\alpha = \frac{1}{n}$ and go to Step 4.

   (e) If $i \geqslant 2$ and $obj(i) \leqslant obj(i-1)$ and $obj(i) \leqslant obj(0)$, then let $\alpha \leftarrow \frac{\alpha}{2} \frac{obj(i-1)+3obj(0)-4obj(i)}{obj(i-1)+obj(0)-2obj(i)}$ and go to Step 3.

   (f) Set $i \leftarrow i + 1$.

   (g) Set $\alpha \leftarrow \alpha/2$.

   (h) Go to Step 2.

3. (moving the point)

   (a) Set $\boldsymbol{p} \leftarrow \boldsymbol{p}' + \alpha(\boldsymbol{p}^* - \boldsymbol{p}')$.

In the numerical experiment section we will consider a large-scale network with 26,400 paths, and as a result, we will use the second algorithm. We would also like to point out that the first algorithm fails to converge for some network instances with $|R| > 50$.

3.5 Numerical Results on Test Networks

In this section we present the numerical results of our new model on test networks. We first investigate the convergence speed of our proposed algorithm. Next we study the equilibrium solution of the new TSUE model, compare it with the SUE equilibrium solution, and state the implications of the results on traffic network modeling.

3.5.1 Convergence speed of solution algorithm

In this subsection we investigate how fast our proposed algorithm converges to the equilibrium solution. For this purpose we solve the TSUE model for the Sioux Falls network (see Figure B.3 in the appendix). This network is commonly used by transportation modelers to test algorithm performances. It has 24 nodes, 76 links, 528 OD pairs, and 1,632,820 simple paths (a simple path is a path that has no repeating nodes). Most of these paths are not used by travelers. In practice, for each OD pair, we select the 50 shortest paths in terms of the free-flow travel time. In total 26,400 paths are selected. We restrict our attention to these paths. The network assumes BPR type link travel time functions (see Eq. (2.2)).

Using the proposed algorithm, we solve the TSUE model for the Sioux Falls network. The TSUE model has two parameters, namely, θ, the perception dispersion level, and λ, the mean-variance trade-off factor. For the numerical study in this chapter we fix $\theta = 1$. We will investigate the influence of θ on traffic equilibrium models in the next chapter. For the Sioux Falls network we pick $\lambda = 0$, 100, and 200. We will explain how we pick these values in the following subsection.

Recall that the TSUE model has a fixed point formulation $\boldsymbol{U}(\boldsymbol{p}) = \boldsymbol{p}$ (See Eq. (3.21)). The convergence speed of an algorithm can be measured in terms of $|\boldsymbol{U}(\boldsymbol{p}) - \boldsymbol{p}|$, that is, the Euclidean norm of $(\boldsymbol{U}(\boldsymbol{p}) - \boldsymbol{p})$. The convergence of the proposed algorithm is shown in Figure 3.1 and Table 3.1. In the figure and table, the label TSUE(0) stands for the TSUE model with $\lambda = 0$ (i.e., the SUE model), and the label TSUE(100) stands for the TSUE model with $\lambda = 100$, etc. The average rate of convergence is defined as $\left(|\boldsymbol{U}(\boldsymbol{p}^n) - \boldsymbol{p}^n| \,/\, \left|\boldsymbol{U}(\boldsymbol{p}^0) - \boldsymbol{p}^0\right|\right)^{\frac{1}{n}}$, where n is the number of iterations, $\boldsymbol{p}^0$ is the initial value for the path choice probability, and $\boldsymbol{p}^n$ is the path choice probability at Step n.

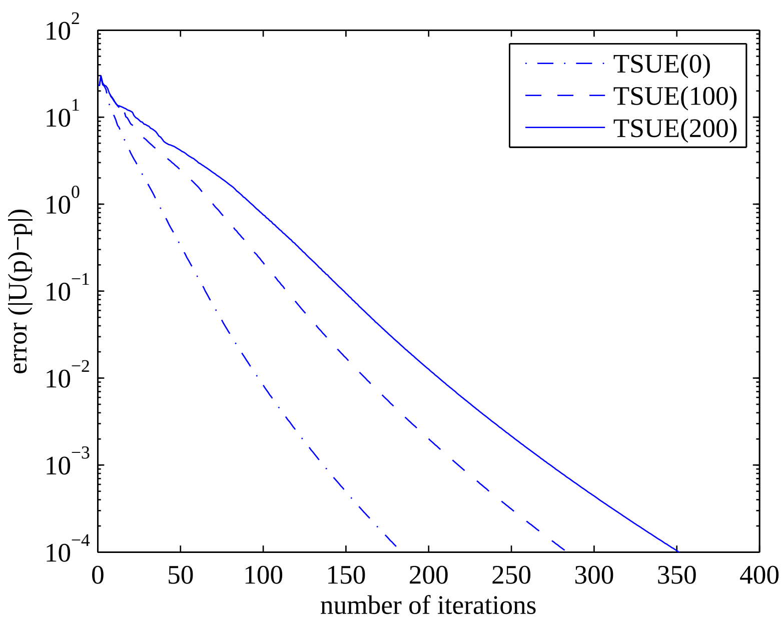


Figure 3.1: Convergence speed of the proposed TSUE algorithm.

From the numerical results we observe that

- For the Sioux Falls network, the TSUE algorithm converges in a few hundred

Model	Number of iterations	Run time (sec)	Ave. rate of convergence
TSUE(0)	184	83.34	0.935
TSUE(100)	284	131.84	0.958
TSUE(200)	352	167.37	0.966

Table 3.1: Number of iterations and run time of the proposed TSUE algorithm.

iterations. On a personal computer with a 3.00 GHz dual core CPU, the total computation time is less than three minutes (see Table 3.1). This indicates that the algorithm can be applied to large scale networks.

- This algorithm converges approximately linearly to the TSUE equilibrium (see Figure 3.1; note that the error is plotted in log scale). This meets our expectation that Frank-Wolfe type algorithms in general converge linearly or sublinearly (see, for example, Bertsekas, 1995).

- As λ increases, the convergence speed becomes slower (see Table 3.1). This indicates that the extra variance term in the travel cost makes the convergence slower. However, even for very large values of λ, the algorithm still converges, although it uses more iterations. For example, when $\lambda = 10,000$, the algorithm converges after 3000 iterations (by convergence, we mean $|\boldsymbol{U}(\boldsymbol{p}) - \boldsymbol{p}| < 10^{-4}$).

3.5.2 Numerical results on the TSUE equilibrium

To study the TSUE model we introduced, we need to examine the relative magnitude of the mean travel time and the variance of the travel time. As noted before, when $\lambda = 0$, the TSUE model reduces to the SUE model, that is, the expected TSUE equilibrium flows are equal to the SUE equilibrium flows, and the TSUE path choice probabilities are equal to the SUE path choice probabilities. (Note that the TSUE and SUE interpret these probabilities differently. See our discussion in Section 3.1.1.) With these path choice probabilities we can evaluate $\mathbb{E}[M_r]$ (i.e., the expected path travel time for any path $r \in R$) and $\text{Var}(M_r)$ (i.e., the variance of the path travel time for any path $r \in R$).

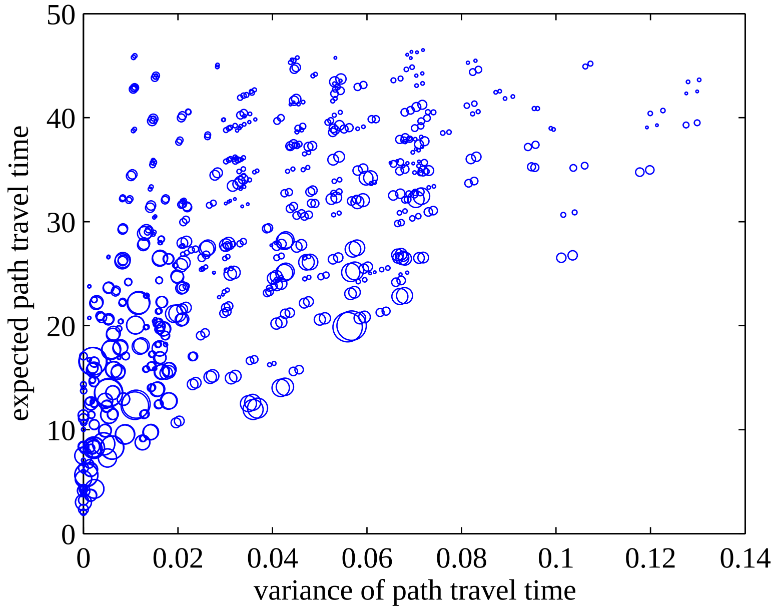
In general, we call $\frac{\text{Var}(M_r)}{\mathbb{E}[M_r]}$ the variance-to-mean (VM) ratio for path r. For any link $l \in L$, let T_l be the random variable for the travel time on link l. We call $\frac{\text{Var}(T_l)}{\mathbb{E}[T_l]}$ the variance-to-mean (VM) ratio for link l. The VM ratio measures the level of travel time uncertainty experienced on a path or a link. We further define the median variance-to-mean (MVM) ratio for the network as the median value of the VM ratio of all the paths, weighted by the their expected path flows (that is, $\mathbb{E}[\boldsymbol{H}]$). The MVM ratio measures the level of travel time uncertainty experienced by a typical traveler. It has the following interpretation: on average, 50% of the travelers travel on a path r that has a VM ratio less than the MVM ratio.

The MVM ratio is helpful when picking the value of λ for the TSUE model. As a rule of thumb, for the numerical tests in this chapter and next chapter, we will pick values of λ to be roughly 0.1 or 0.2 divided by the MVM ratio. This implies that for a typical traveler, the travel cost attributed to the travel time uncertainty is roughly 10% or 20% of that attributed to the mean travel time.

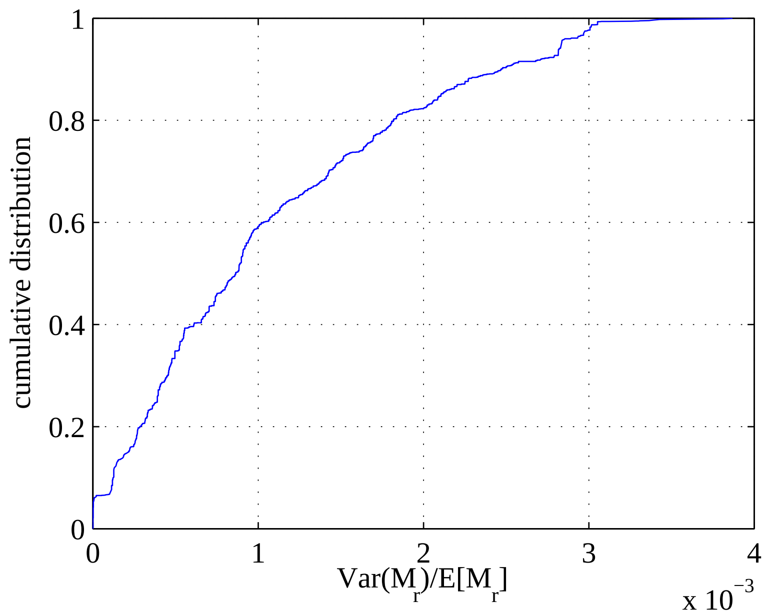
The TSUE equilibrium with $\lambda = 0$

We present results related to the VM ratio in Figures 3.2a and 3.2b.

Figure 3.2a shows the mean path travel time relative to the variance of the path travel time. In this figure, each path $r \in R$ is represented by a circle. The coordinates of the circle's center are the mean and variance of path r's travel time. The area of the circle is proportional to the expected flow on path r, i.e., $\mathbb{E}[H_r]$. (The figure shows 740 paths, i.e., 740 circles, out of all 26,400 paths. Other paths have very small expected flows, hence they are not shown in the figure.) This figure implies that the mean path travel time is positively correlated with the variance of the path



(a) Mean vs. variance of the path travel time



(b) Cumulative distribution of the path VM ratio

Figure 3.2: Mean and variance of the path travel time for the TSUE equilibrium with $\lambda = 0$.

travel time. The correlation coefficient is 0.66.

Figure 3.2b shows the cumulative distribution of the path VM ratio, weighted by the expected path flow. This figure implies that, on average, 50% of the travelers choose a path with VM ratio less than 0.86×10^{-3}, 80% of the travelers choose a path with VM ratio between 0.12×10^{-3} and 2.5×10^{-3}, and almost no travelers choose a path with VM ratio greater than 3.1×10^{-3}. We provide the quantiles of the path VM ratio in Table 3.2.

Quantile	0.01	0.10	0.25	0.50	0.75	0.90	0.99
Path VM ratio (10^{-3})	2.1e-5	0.12	0.39	0.86	1.6	2.5	3.1

Table 3.2: Quantiles of the path VM ratio for the TSUE(0) equilibrium solution.

The above result has important implications on choosing the value of λ for the TSUE model. The TSUE travel cost has two components. The first component is $\mathbb{E}[M_r]$, and the second component is $\lambda \mathrm{Var}(M_r)$. Consider the following example, if $\lambda = 100$, then on average, for 50% of the travelers, the travel cost attributed to the travel time uncertainty is no more than 10% of that attributed to the mean travel time. For 80% of the travelers, the travel cost attributed to the travel time uncertainty is between 1% to 25% of that attributed to the mean travel time.

The TSUE equilibrium with $\lambda > 0$

We are now in a position to solve the TSUE model. For the purpose of presentation we will choose two values of λ, that is, $\lambda = 100$ and $\lambda = 200$. We are particularly interested in two questions: 1) how different are the TSUE solutions from the SUE solution and GSUE solutions, and 2) what can we learn from these differences.

To show the difference of the TSUE solutions from the SUE and GSUE solutions, we investigate the mean travel time in both the path space and the flow space. We measure the relative change of the mean travel time and the mean traffic flow in the TSUE and GSUE solutions compared to the SUE solution (i.e., the TSUE equilibrium with $\lambda = 0$). Figures 3.3a and 3.3b present these results.

Figure 3.3a plots the cumulative distribution of the relative change of the mean path travel time experienced by users. This figure implies that, for example, in the case of $\lambda = 100$, on average around 20% of the travelers will experience at least 5% less mean travel time compared to the predicted travel time of the SUE model.

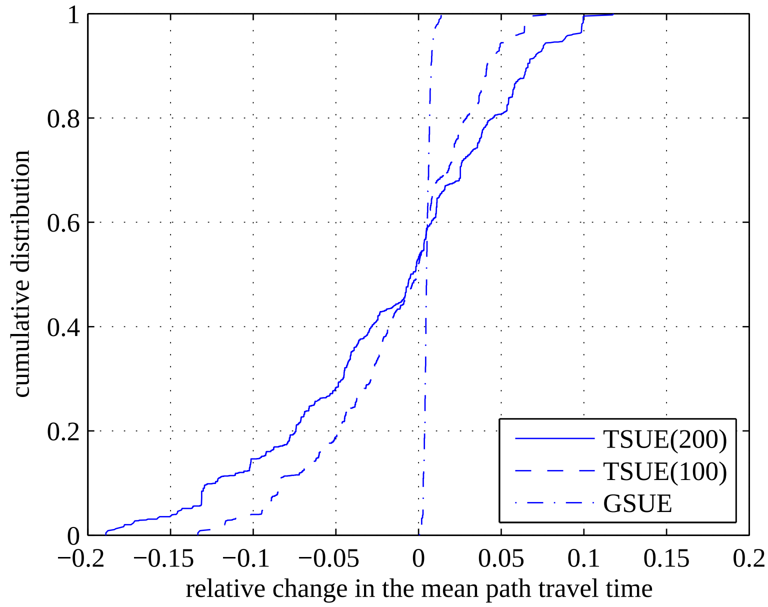
Figure 3.3b plots the cumulative distribution of the relative change of the mean link traffic flow. For example, in the case of $\lambda = 200$, the TSUE model predicts a longer mean travel time on about 50% of the links. The gap is up to 17% of the mean travel time predicted by the SUE model. On the other hand, the TSUE model also predicts about 9% less mean travel time on some link. (See Table 3.3.)

	Rel. change in the mean link flow		
Quantile	GSUE	TSUE(100)	TSUE(200)
0.01	-0.0038	-0.0605	-0.0976
0.10	-0.0025	-0.0466	-0.0734
0.25	-0.0010	-0.0162	-0.0277
0.50	0.0005	0.0029	0.0042
0.75	0.0008	0.0213	0.0340
0.90	0.0014	0.0413	0.0746
0.99	0.0019	0.1086	0.1777

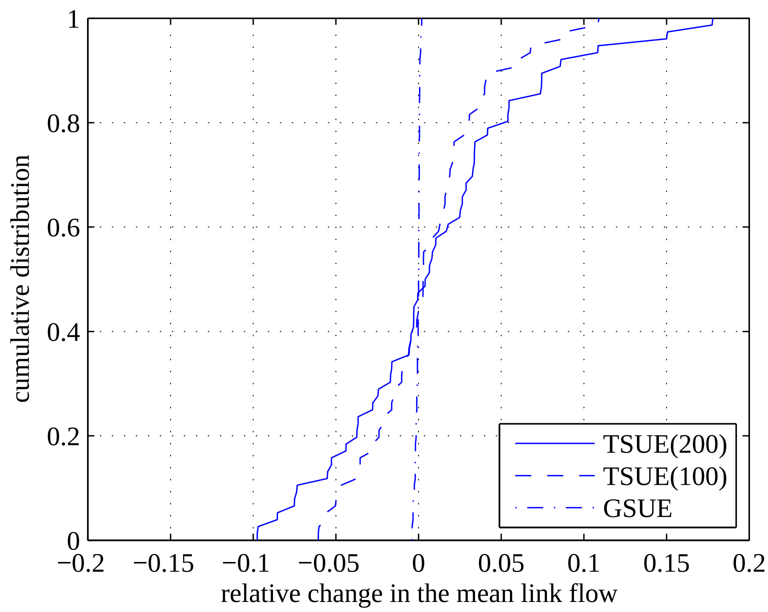
Table 3.3: Quantiles of the relative change in the mean link traffic flow.

We also conclude from Figures 3.3a and 3.3b and Table 3.3 that the mean travel time and traffic flow predicted by the GSUE model is close to that of the SUE model. This shows that accounting for the second-order term in the travel time function does not significantly change the equilibrium solution. This justifies our use of the first order approximation in the TSUE model.

We further investigate the changes in the predicted link flows of the TSUE model. To understand how travel time uncertainty influences the equilibrium link flows, we plot the relative change of expected link flows against the link travel time uncertainty



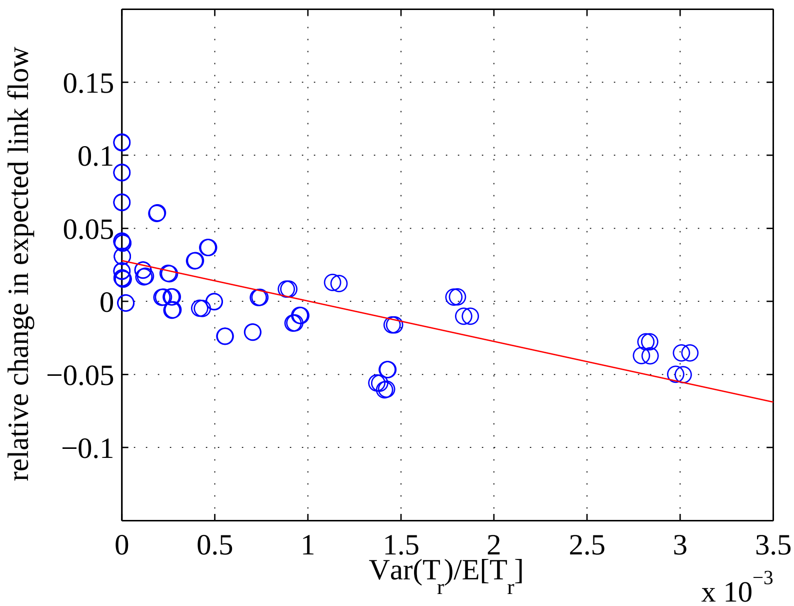
(a) The mean path travel time.



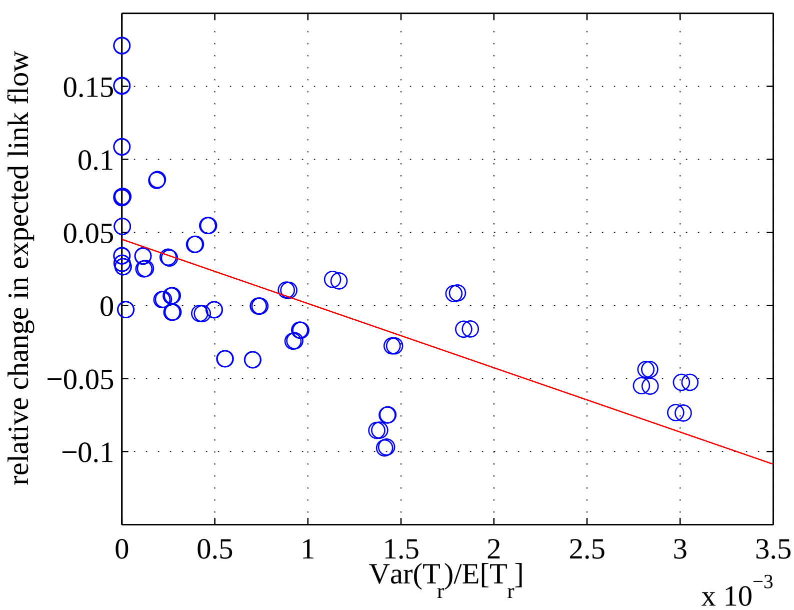
(b) The mean link traffic flow.

Figure 3.3: Relative change in mean travel time and traffic flow of the TSUE and GSUE equilibrium solutions, compared to the SUE equilibrium solution.

level. For link $l \in L$, we measure the link travel time uncertainty level using its VM ratio, i.e., $\frac{\mathrm{Var}(T_l)}{\mathbb{E}[T_l]}$. The results are shown in Figure 3.4a (the case of $\lambda = 100$) and 3.4b (the case of $\lambda = 200$). Both figures indicate that the link VM ratio is negatively correlated with the link flow. When λ increases, the influences become more obvious. This is a strong evidence that the SUE model underestimates flows on links with high VM ratio and underestimates flows on links with low VM ratio.



(a) TSUE(100)



(b) TSUE(200)

Figure 3.4: The SUE model under-estimates flows on links with high VM ratio and over-estimate flows on links with low VM ratio.

Chapter 4

Benchmarking the Traffic Network Equilibrium Models with Respect to the System Cost

Given a traffic network with fixed traffic demands[1], we can build different traffic equilibrium models (see Chapter 2). Each of these equilibria predicts a total system cost (for example, the total traverse time) for the traffic system. It is important to study the total system cost and how it relates for each model, for the following reasons.

First, this study allows us to understand different traffic network equilibrium models using a common measure. This contributes to our understanding of each of the individual models and their relationship.

Second, since all traffic network equilibrium models are simplifications of reality, by evaluating the system cost with different models, we achieve a better understanding of the system cost and its possible range in a real traffic network.

[1]In this chapter, by traffic demands we refer to the OD pair flow.

Closely related to the concept of the system cost is the system optimum (SO) model. The SO model assumes that a central planner is in full control of the flows in the network and optimizes the flows in order to minimize the system cost. This assumption contrasts with the assumption of a selfish and non-cooperative routing behavior in the traffic equilibrium models. The system cost is a measure of the overall performance of the traffic network. For example, the lower the total travel cost is, the more efficient is the traffic system. Hence the gap between the system cost in equilibrium and the optimal system cost measures the loss of system efficiency due to the selfish routing behavior of network users. Our study will also address this issue.

In this chapter we consider the following models: the DUE model, the SUE model, the TSUE model, and the SO model. This chapter is organized as follows. Section 4.1 introduces notations and defines the system cost for traffic network models. Section 4.2 establishes upper bounds for the TSUE system cost. Section 4.3 illustrates the tightness and quality of the bounds.

4.1 The System Cost

Throughout this chapter we restrict our attention to traffic networks with affine, symmetric and positive semidefinite travel time functions. That is, the path travel time is

$$\boldsymbol{\mu}(\boldsymbol{h}) = \boldsymbol{G}\boldsymbol{h} + \boldsymbol{b},$$

where $\boldsymbol{G}$ is a symmetric and positive semidefinite matrix and $\boldsymbol{b}$ is a constant vector (the vector of free-flow path travel times). To keep our notations simple, all discussions in this section will be conducted in the path space.

In the following, we will define the system cost for the following models: the

DUE and SUE models, the SO model, and the TSUE model.

4.1.1 The system cost for the DUE and SUE models

For deterministic models, the system cost, defined as the aggregate travel time experienced by all users, is $z(\boldsymbol{h}) = \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{h} = (\boldsymbol{G}\boldsymbol{h} + \boldsymbol{b})^{\mathrm{T}}\boldsymbol{h}$. In particular, let $\boldsymbol{h}^{\mathrm{DUE}}$ and $\boldsymbol{h}^{\mathrm{SUE}}(\theta)$ be the vectors of equilibrium path flows for the DUE and SUE models, respectively, then

$$z^{\mathrm{DUE}} = \left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}} + \boldsymbol{b}\right)^{\mathrm{T}} \boldsymbol{h}^{\mathrm{DUE}},$$
$$z^{\mathrm{SUE}}(\theta) = \left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{SUE}}(\theta) + \boldsymbol{b}\right)^{\mathrm{T}} \boldsymbol{h}^{\mathrm{SUE}}(\theta).$$

Note that the SUE model is parameterized by the perception dispersion level θ, whose range is $(0, \infty)$. We consider two extreme cases of the SUE model:

Case 1: $\theta \to 0$. Zero perception dispersion level implies that users have accurate perceptions of their travel times. Hence the SUE model reduces to the DUE model. We define $\boldsymbol{h}^{\mathrm{SUE}}(0) = \lim_{\theta\to 0^+} \boldsymbol{h}^{\mathrm{SUE}}(\theta)$ and $z^{\mathrm{SUE}}(0) = \lim_{\theta\to 0^+} z^{\mathrm{SUE}}(0)$. Then $\boldsymbol{h}^{\mathrm{SUE}}(0) = \boldsymbol{h}^{\mathrm{DUE}}$ and $z^{\mathrm{SUE}}(0) = z^{\mathrm{DUE}}$. This fact is formally stated in Proposition 4.1 below.

Proposition 4.1. *Consider a traffic network with a continuous travel time function* $\boldsymbol{t}(\cdot)$*. If there exists a positive number* $\delta > 0$*, such that for all* $\theta \in (0, \delta)$*, the SUE model has a unique solution and* $\boldsymbol{h}^{\mathrm{SUE}}(0) = \lim_{\theta\to 0^+} \boldsymbol{h}^{\mathrm{SUE}}(\theta)$ *exists, then* $\boldsymbol{h}^{\mathrm{SUE}}(0)$ *is a solution to the DUE model.*

Proof. By the SUE-VI-PATH formulation (see Eq. (2.8)), for all $\theta \in (0, \infty)$, the following equation holds for all $\boldsymbol{h} \in S_h$

$$\left(\boldsymbol{c}\left(\boldsymbol{h}^{\mathrm{SUE}}(\theta)\right) + \theta \log \frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\boldsymbol{h} - \boldsymbol{h}^{\mathrm{SUE}}(\theta)\right) = 0.$$

That is,

$$
\begin{aligned}
&\boldsymbol{c}\left(\boldsymbol{h}^{\mathrm{SUE}}(\theta)\right)^{\mathrm{T}}\left(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{SUE}}(\theta)\right)\\
=&\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}^{\mathrm{SUE}}(\theta)-\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}.
\end{aligned}
$$

Since $\boldsymbol{t}(\cdot)$ is continuous, $\boldsymbol{c}(\cdot)$ is continuous as well. Taking θ to 0 from above, we have that

$$
\begin{aligned}
&\boldsymbol{c}\left(\boldsymbol{h}^{\mathrm{SUE}}(0)\right)^{\mathrm{T}}\left(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{SUE}}(0)\right)\\
=&\lim_{\theta\to 0^+}\left(\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}^{\mathrm{SUE}}(\theta)-\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}\right).
\end{aligned}
$$

We claim that $\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}^{\mathrm{SUE}}(\theta)$ is bounded for all $\theta\in(0,\delta)$. To prove this, we only need to show that for all $k\in K$, $r\in R^k$, $\left(\log\frac{h_r^{\mathrm{SUE}}(\theta)}{q^k}\right)h_r^{\mathrm{SUE}}(\theta)$ is bounded.[2] By taking derivatives, we can show that $\left(\log\frac{x}{q^k}\right)x$ is a convex function of x for $x>0$, and it takes its minimum value at $x=\frac{q^k}{e}$. Hence $\left(\log\frac{h_r^{\mathrm{SUE}}(\theta)}{q^k}\right)h_r^{\mathrm{SUE}}(\theta)\geqslant -\frac{q^k}{e}$. Furthermore we notice that $h_r^{\mathrm{SUE}}(\theta)\leqslant q^k$. Thus $\left(\log\frac{h_r^{\mathrm{SUE}}(\theta)}{q^k}\right)h_r^{\mathrm{SUE}}(\theta)\leqslant 0$. Therefore, $\left(\log\frac{h_r^{\mathrm{SUE}}(\theta)}{q^k}\right)h_r^{\mathrm{SUE}}(\theta)$ is bounded. As a result,

$$
\lim_{\theta\to 0^+}\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}^{\mathrm{SUE}}(\theta)=0,
$$

which means that

$$
\boldsymbol{c}\left(\boldsymbol{h}^{\mathrm{SUE}}(0)\right)^{\mathrm{T}}\left(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{SUE}}(0)\right)=-\lim_{\theta\to 0^+}\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}.
$$

Since $\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}$ is always non-positive, we have that $\lim_{\theta\to 0^+}\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}\leqslant$

[2] In fact, we can show that $0\geqslant\left(\log\frac{\boldsymbol{h}^{\mathrm{SUE}}(\theta)}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}^{\mathrm{SUE}}(\theta)\geqslant\left(\log\frac{\boldsymbol{h}^{\mathrm{INF}}}{\boldsymbol{q}}\right)^{\mathrm{T}}\boldsymbol{h}^{\mathrm{INF}}$, where $\boldsymbol{h}^{\mathrm{INF}}$ is defined in Eq. (4.1).

0. This implies that

$$\boldsymbol{c}\left(\boldsymbol{h}^{\mathrm{SUE}}(0)\right)^{\mathrm{T}}\left(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{SUE}}(0)\right) \geqslant 0, \quad \forall \boldsymbol{h} \in S_h.$$

This coincides with the DUE-VI-PATH formulation, namely, Eq. (2.4). Hence $\boldsymbol{h}^{\mathrm{SUE}}(0)$ is a solution to the DUE model. □

Case 2: $\theta \to \infty$. In this case travelers are assumed to have very large perception errors. According to the random utility model, travelers should have equal probabilities of choosing any of the available paths. Let

$$\boldsymbol{h}^{\mathrm{INF}} = \lim_{\theta \to \infty} \boldsymbol{h}^{\mathrm{SUE}}(\infty). \tag{4.1}$$

By the SUE-FIX formulation (see Eq. (2.7)) and the fact that the cost function $\boldsymbol{c}(\cdot)$ is bounded, we have $h_r^{\mathrm{INF}} = \frac{q^k}{|R^k|}$ for any $k \in K$, $r \in R^k$. Let $z^{\mathrm{INF}} = \lim_{\theta \to \infty} z^{\mathrm{SUE}}(\theta)$, then

$$z^{\mathrm{INF}} = \left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}} + \boldsymbol{b}\right)^{\mathrm{T}} \boldsymbol{h}^{\mathrm{INF}}.$$

4.1.2 The system cost for the SO model

Next we define the system cost for the system optimum model. It has the following optimization formulation.

SO-OPT. Find the minimizer $\boldsymbol{h}^{\mathrm{SO}} \in S_h$ of the following minimization problem

$$\min_{\boldsymbol{h} \in S_h} \; (\boldsymbol{G}\boldsymbol{h} + \boldsymbol{b})^{\mathrm{T}} \boldsymbol{h}. \tag{4.2}$$

Let $\boldsymbol{h}^{\mathrm{SO}}$ be the vector of the SO path flows. Then the SO system cost is

$$z^{\mathrm{SO}} = \left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{SO}} + \boldsymbol{b}\right)^{\mathrm{T}} \boldsymbol{h}^{\mathrm{SO}}.$$

By this optimization formulation, we have that $z^{\text{DUE}} \geqslant z^{\text{SO}}$, $z^{\text{INF}} \geqslant z^{\text{SO}}$, and $z^{\text{SUE}}(\theta) \geqslant z^{\text{SO}}$, for all $\theta \in [0, \infty)$.

4.1.3 The system cost for the TSUE model

Lastly, we define the system cost for the TSUE model. We consider two different definitions of the TSUE system cost.

In the first definition, we define the TSUE system cost as the system cost of the expected flow, namely

$$z^{\text{TSUE}}(\theta, \lambda) = \left(\boldsymbol{G}\boldsymbol{h}^{\text{TSUE}}(\theta, \lambda) + \boldsymbol{b}\right)^{\text{T}} \boldsymbol{h}^{\text{TSUE}}(\theta, \lambda). \tag{4.3}$$

In this equation, $\boldsymbol{h}^{\text{TSUE}}(\theta, \lambda)$ is the expected equilibrium path flow of the TSUE model. In the special case of $\lambda = 0$, since $\boldsymbol{h}^{\text{TSUE}}(\theta, 0) = \boldsymbol{h}^{\text{SUE}}(\theta)$, it also follows that $z^{\text{TSUE}}(\theta, 0) = z^{\text{SUE}}(\theta)$. This definition is a natural extension to the definitions of the DUE, SUE and SO system costs. It keeps the functional form of the previous definitions and hence allows us to compare these system costs in a consistent manner.

Alternatively, we can define the TSUE system cost as the expected system cost. Unlike the deterministic models such as the DUE and SUE models, the TSUE model assumes that traffic flows and travel times are random variables. As a result the aggregate travel cost experienced by all users is a random variable. Let $\boldsymbol{H}^{\text{TSUE}}(\theta, \lambda)$ be the vector of equilibrium path flows of the TSUE equilibrium. Then the TSUE system cost is defined as

$$\tilde{z}^{\text{TSUE}}(\theta, \lambda) = \mathbb{E}\left[\left(\boldsymbol{G}\boldsymbol{H}^{\text{TSUE}}(\theta, \lambda) + \boldsymbol{b}\right)^{\text{T}} \boldsymbol{H}^{\text{TSUE}}(\theta, \lambda)\right].$$

Here θ is the perception dispersion level, and λ is the mean-variance trade-off factor. Let $\boldsymbol{h}^{\text{TSUE}}(\theta, \lambda) = \mathbb{E}[\boldsymbol{H}^{\text{TSUE}}(\theta, \lambda)]$. When $\lambda = 0$, the TSUE model reduces to the

SUE model. As a result, the expected flows in the TSUE equilibrium are equal to the flows in the SUE equilibrium, i.e.,

$$\boldsymbol{h}^{\mathrm{TSUE}}(\theta, 0) = \boldsymbol{h}^{\mathrm{SUE}}(\theta). \tag{4.4}$$

However, the deterministic system cost for the SUE model is not necessarily equal to the expected system cost for the TSUE model.

For simplicity, in the following discussion we will omit the parameters θ and λ in the expression. For example, we will write $\boldsymbol{H}^{\mathrm{TSUE}}$ instead of $\boldsymbol{H}^{\mathrm{TSUE}}(\theta, \lambda)$. For arbitrary $\theta \in (0, \infty)$, $\lambda \in [0, \infty)$, we have

$$\begin{aligned}
\tilde{z}^{\mathrm{TSUE}} &= \mathbb{E}\left[\left(\boldsymbol{H}^{\mathrm{TSUE}}\right)^{\mathrm{T}} \boldsymbol{G}\boldsymbol{H}^{\mathrm{TSUE}}\right] + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}} \\
&= \sum_{r,r' \in R} G_{r,r'} \mathbb{E}\left[H_r^{\mathrm{TSUE}} H_{r'}^{\mathrm{TSUE}}\right] + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}} \\
&= \sum_{r,r' \in R} G_{r,r'} \left(h_r^{\mathrm{TSUE}} h_{r'}^{\mathrm{TSUE}} + \mathrm{Cov}\left(H_r^{\mathrm{TSUE}}, H_{r'}^{\mathrm{TSUE}}\right)\right) + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}} \\
&= \left(\boldsymbol{h}^{\mathrm{TSUE}}\right)^{\mathrm{T}} \boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} + \sum_{r,r' \in R} G_{r,r'} \mathrm{Cov}\left(H_r^{\mathrm{TSUE}}, H_{r'}^{\mathrm{TSUE}}\right) + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}} \\
&= \left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} + \boldsymbol{b}\right)^{\mathrm{T}} \boldsymbol{h}^{\mathrm{TSUE}} + \boldsymbol{e}^{\mathrm{T}} \left(\boldsymbol{G} \circ \mathrm{Cov}\left(\boldsymbol{H}^{\mathrm{TSUE}}\right)\right) \boldsymbol{e},
\end{aligned}$$

where $\boldsymbol{e}$ is a column vector of all 1's, $\mathrm{Cov}\left(\boldsymbol{H}^{\mathrm{TSUE}}\right)$ is the covariance matrix of $\boldsymbol{H}^{\mathrm{TSUE}}$, and $\boldsymbol{A} \circ \boldsymbol{B}$ denotes the Hadamard product of two matrices of the same dimension, with $(\boldsymbol{A} \circ \boldsymbol{B})_{i,j} = A_{i,j} B_{i,j}$.

The above result suggests that, when $\lambda = 0$, by Eq. (4.4), we have that

$$\tilde{z}^{\mathrm{TSUE}}(\theta, 0) = z^{\mathrm{SUE}}(\theta) + \boldsymbol{e}^{\mathrm{T}} \left(\boldsymbol{G} \circ \mathrm{Cov}\left(\boldsymbol{H}^{\mathrm{TSUE}}(\theta, 0)\right)\right) \boldsymbol{e}. \tag{4.5}$$

In general, we have the following relationship between these two definitions.

$$\tilde{z}^{\mathrm{TSUE}}(\theta, \lambda) = z^{\mathrm{TSUE}}(\theta, \lambda) + \boldsymbol{e}^{\mathrm{T}} \left(\boldsymbol{G} \circ \mathrm{Cov} \left(\boldsymbol{H}^{\mathrm{TSUE}}(\theta, \lambda) \right) \right) \boldsymbol{e}. \tag{4.6}$$

4.2 Bounds for the System Cost

In this section we will develop three bounds for the TSUE system cost. These bounds establish the relationship between the TSUE system cost and the system costs of other models. Section 4.2.1 briefly reviews previous studies on the system cost. Section 4.2.2 introduces three upper bounds for the TSUE system cost. In the special case of $\lambda = 0$, these bounds reduces to upper bounds for the SUE system cost. We present a number of lemmas necessary to prove these bounds in Section 4.2.3. The proofs are shown in Section 4.2.4. We discuss the tightness of these bounds in Section 4.2.5.

4.2.1 Literature review

The system optimum model was initially formulated by Wardrop (1952). Wardrop's second principle states: "at equilibrium the average journey time is minimum". Our formulation of the system optimum model follows this principle.

Maher et al. (2005) formulated the Stochastic System Optimum (SSO) model. The system cost for the SSO model, in addition to accounting for the total travel times, also includes travelers' perception errors. However, because we focus on the total travel time in this chapter, we will not address the SSO model.

Roughgarden and Tardos (2002) were perhaps the first to study the loss of efficiency between the user equilibrium and system optimum in transportation networks. They provided an upper bound for the ratio of the DUE system cost over the SO system cost. They called the ratio price of anarchy, following Koutsoupias and Papadimitriou (1999), who initially coined the term in the context of a load balancing game. The price of anarchy is a measure of the efficiency loss due to the selfish routing behavior of travelers. Roughgarden and Tardos (2002) proved that for networks with affine and separable travel time functions, the DUE system cost is at most $\frac{4}{3}$ of the system optimal cost. That is,

$$z^{\mathrm{DUE}} \leqslant \frac{4}{3} z^{\mathrm{SO}}. \tag{4.7}$$

Roughgarden (2002) subsequently generalized the bound for special classes of separable and nonlinear travel time functions (such as polynomials). Correa et al. (2004) studied the "price of anarchy" for capacitated networks. See also, among others, Chau and Sim (2003) and Correa et al. (2008).

Perakis (2007) extended the above results to networks with nonlinear and asymmetric travel time functions (with positive semidefinite Jacobian matrix). She proved that

$$\frac{z^{\mathrm{DUE}}}{z^{\mathrm{SO}}} \leqslant \begin{cases} \frac{4}{4-c^2\mathcal{A}} & \text{if } c^2 \leqslant \frac{2}{\mathcal{A}}, \\ c^2\mathcal{A}^2 - 2(\mathcal{A}-1) & \text{if } c^2 > \frac{2}{\mathcal{A}}, \end{cases} \tag{4.8}$$

where c measures the degree of "asymmetry" of the travel time function $\boldsymbol{t}(\cdot)$ (with $c = 1$ if the Jacobian matrix of $\boldsymbol{t}(\cdot)$ is symmetric and $c \geqslant 1$ otherwise), and $\mathcal{A}$ measures the degree of "nonlinearity" of $\boldsymbol{t}(\cdot)$ (with $\mathcal{A} = 1$ if $\boldsymbol{t}(\cdot)$ is affine and $\mathcal{A} \geqslant 1$ otherwise). This bound reduces to Eq. (4.7) for affine and symmetric link travel time functions. Furthermore, it was shown that this bound is tight.

Guo and Yang (2005) were perhaps the first to develop a bound for the SUE system cost. They considered networks with separable and non-decreasing travel time functions. Let $t_l(\cdot)$ be the travel time function on link l. Following Roughgarden

(2002), for each link, they defined

$$\gamma_l = \sup_{f,f' \geqslant 0} \frac{(t_l(f) - t_l(f'))f'}{t_l(f)f}.$$

For the whole network, they defined

$$\gamma = \max_{l \in L} \gamma_l.$$

For example, for networks with separable affine travel time functions, we have that $\gamma = \frac{1}{4}$. For networks with more general travel time functions, we have that $\frac{1}{4} \leqslant \gamma < 1$. Guo and Yang showed that

$$z^{\text{SUE}}(\theta) \leqslant \frac{1}{1-\gamma} \left(z^{\text{SO}} + \theta \mathcal{B} \right), \tag{4.9}$$

where B is a network specific constant, defined as

$$\mathcal{B} = \sum_{k \in K} q^k \mathcal{K}(|R^k|).$$

Here $|R^k|$ is the number of paths available to travelers of OD pair k, and $\mathcal{K}(\cdot)$ is a function defined as follows. For any positive integer n, $\mathcal{K}(n)$ is a positive real number that solves the following equation

$$\mathcal{K} e^{\mathcal{K}+1} = n - 1.$$

It is easy to show that this equation has a unique solution. The numerical values of this function for some values of $|R^k|$ are shown in Table 4.1.

The bound in Eq. (4.9) grows linearly as a function of θ. For networks with affine travel time functions, it reduces to the bound in Eq. (4.7) when $\theta \to 0$. This

$\lvert R^k \rvert$	1	10	10^2	10^3	10^4	10^5
$\mathcal{K}(\lvert R^k \rvert)$	0	1.10	2.63	4.42	6.36	8.39

Table 4.1: Numerical values of $\mathcal{K}(|R^k|)$.

bound grows to infinity when $\theta \to \infty$.

We note that the performance of this bound is unsatisfactory for both extreme cases. When $\theta \to 0$, we know $z^{\text{SUE}}(0) \leqslant z^{\text{DUE}}$ in addition to $z^{\text{SUE}}(0) \leqslant \frac{4}{3} z^{\text{SO}}$. The first bound is stronger than the second bound because $z^{\text{DUE}} \leqslant \frac{4}{3} z^{\text{SO}}$. When $\theta \to \infty$, we know that $z^{\text{SUE}}(\infty) \leqslant z^{\text{INF}}$, i.e., the bound is finite. These observations motivate us to develop new and stronger bounds for the TSUE system cost in the following subsection.

4.2.2 Bounds for the TSUE system cost

We develop three upper bounds for $z^{\text{TSUE}}(\theta, \lambda)$ defined in Eq. (4.3). These bounds benchmark $z^{\text{TSUE}}(\theta, \lambda)$ against the total travel times for the various models, that is, z^{DUE}, z^{SO} and z^{INF}, respectively. These bounds are as follows (refer to Theorem 4.8, 4.9 and 4.10 for the formal statements of the conditions under which these bounds hold)

$$z^{\text{TSUE}}(\theta, \lambda) \leqslant z^{\text{DUE}} + \frac{\theta B + \lambda J}{2 + 2\theta A} + \sqrt{\frac{\theta B + \lambda J}{2 + 2\theta A}} \sqrt{\frac{\theta B + \lambda J}{2 + 2\theta A} + 2 z^{\text{DUE}}}, \tag{4.10}$$

$$z^{\text{TSUE}}(\theta, \lambda) \leqslant \frac{4 + 4\theta A}{3 + 4\theta A} \left(z^{\text{SO}} + \theta B + \lambda J \right), \tag{4.11}$$

$$z^{\text{TSUE}}(\theta, \lambda) \leqslant \frac{4 + 8\theta A}{3 + 8\theta A} (z^{\text{INF}} + \lambda J) \leqslant \frac{4}{3} (z^{\text{INF}} + \lambda J). \tag{4.12}$$

Here A, B, and J are three network dependent constants. To be specific, A depends on the magnitude of traffic demands and the travel time function, with $A = \frac{1}{2 q_{\max} \nu_{\max}(\boldsymbol{G})}$, where the $q_{\max} = \max_{k \in K} q^k$ denotes the largest demand flow among all OD pairs, and $\nu_{\max}(\boldsymbol{G})$ denotes the largest eigenvalue of matrix $\boldsymbol{G}$. B depends on the traffic demands as well as the number of paths in the network, with $B = \sum_{k\in K} q^k \log |R^k|$, where $|R^k|$ is the number of available paths for OD pair k. J is a constant defined as $J = 4\sqrt{|R|}\nu_{\max}(\boldsymbol{G})^2 \left(\sum_k q^k\right)^{3/2}$. These three constants can be easily calculated for a given network without solving any of the user equilibrium or system optimum models.

To understand the meaning of A, consider a network with a single OD pair and n parallel links. Suppose the link travel time functions are $t_i(f_i) = a_i f_i + b_i$, for $i = 1, \cdots, n$. Let q be the demand flow. Then $\nu_{\max}(\boldsymbol{G}) = \max_i\{a_i\}$, and $q_{\max}\nu_{\max}(\boldsymbol{G}) = q \max_i\{a_i\} = \max_i\{(a_i q + b_i) - b_i\}$. Therefore, for this network, $q_{\max}\nu_{\max}(\boldsymbol{G})$ represents the maximum difference between the travel time when a link is free of flow and the travel time when the link is fully loaded. Roughly speaking this quantity measures the magnitude of potential variation in the link travel time. Hence A is inversely proportional to the level of potential variation in the link travel time.

Dependence on λ

When $\lambda = 0$, these bounds reduces to upper bounds for the SUE system cost. For a fixed value of θ, all these bounds increase linearly as λ increases.

Dependence on θ

To investigate how these bounds depend on the value of θ, we focus on the case of $\lambda = 0$.

The first bound is tight when $\theta = 0$; in this case the bound reduces to $z^{\mathrm{SUE}}(0) \leqslant z^{\mathrm{DUE}}$. The second bound implies that $z^{\mathrm{SUE}}(0) \leqslant \frac{4}{3} z^{\mathrm{SO}}$. Since $z^{\mathrm{SUE}}(0) = z^{\mathrm{DUE}}$, we have $z^{\mathrm{DUE}} \leqslant \frac{4}{3} z^{\mathrm{SO}}$. This reproduces the bound for the DUE model (see Roughgarden and Tardos, 2002). The third bound is tight when $\theta \to \infty$; under this limit the bound reduces to $z^{\text{SUE}}(\infty) \leqslant z^{\text{INF}}$. Furthermore, the third bound implies that $z^{\text{SUE}}(\theta) \leqslant \frac{4}{3} z^{\text{INF}}$ for any dispersion level θ.

A summary of the bounds for various values of θ are provided in Table 4.2.

	The first bound	The second bound	The third bound
Eq. Number	Eq. (4.10)	Eq. (4.11)	Eq. (4.12)
$\theta = 0$	z^{DUE}	$\frac{4}{3} z^{\text{SO}}$	$\frac{4}{3} z^{\text{INF}}$
$\theta = \infty$	∞	∞	z^{INF}
$\theta \in (0, \infty)$	increasing	increasing	decreasing

Table 4.2: Bounds for various values of θ. The first and second bound are always increasing functions of θ. The third bound is always a decreasing function of θ.

In general, the first bound is the most restrictive among all three when θ is small, and the third upper bound is the most restrictive when θ is large. The second upper bound may be the most restrictive when θ lies in the middle range.

4.2.3 Useful inequalities

In this subsection, we prove a number of useful inequalities (Lemma 4.2, 4.3, 4.4 and 4.7 and Corollary 4.5 and 4.6) necessary to prove the upper bounds (see Theorem 4.8, 4.9 and 4.10).

Lemma 4.2. *If $\boldsymbol{G}$ is a positive semidefinite matrix, then for any feasible path flow vectors $\boldsymbol{h}$ and $\tilde{\boldsymbol{h}}$ in S_h, and any real number $a > 0$, the following inequality holds*

$$a\boldsymbol{h}^{\text{T}}\boldsymbol{G}\boldsymbol{h} + \frac{1}{4a}\tilde{\boldsymbol{h}}^{\text{T}}\boldsymbol{G}\tilde{\boldsymbol{h}} \geqslant \boldsymbol{h}^{\text{T}}\boldsymbol{G}\tilde{\boldsymbol{h}}. \tag{4.13}$$

Proof. Since $\boldsymbol{G}$ is positive semidefinite, the following inequality must hold

$$\left(\sqrt{a}\boldsymbol{h} - \frac{1}{2\sqrt{a}}\tilde{\boldsymbol{h}}\right)^{\text{T}} \boldsymbol{G} \left(\sqrt{a}\boldsymbol{h} - \frac{1}{2\sqrt{a}}\tilde{\boldsymbol{h}}\right) \geqslant 0.$$

By expanding the above expression we obtain (4.13). □

Lemma 4.3. *Consider the INF solution* $\boldsymbol{h}^{\mathrm{INF}}$ *(defined as* $h_r^{\mathrm{INF}} = \frac{q^k}{|R^k|}$ *for any* $k \in K$*,* $r \in R^k$*). For any feasible path flow vector* $\boldsymbol{h} \in S_h$*, the following relationship holds*

$$(\log \frac{\boldsymbol{h}^{\mathrm{INF}}}{\boldsymbol{q}})^{\mathrm{T}} \left(\boldsymbol{h} - \boldsymbol{h}^{\mathrm{INF}}\right) = 0. \tag{4.14}$$

Proof. For any OD pair $k \in K$ and any path $r \in R^k$ for the OD pair, we have $\frac{h_r^{\mathrm{INF}}}{q^k} = \frac{1}{|R^k|}$, hence

$$\sum_{r \in R^k} \frac{h_r^{\mathrm{INF}}}{q^k} \left(h_r - h_r^{\mathrm{INF}}\right) = \frac{1}{|R^k|} \sum_{r \in R^k} \left(h_r - h_r^{\mathrm{INF}}\right) = \frac{1}{|R^k|} \left(q^k - q^k\right) = 0.$$

Eq. (4.14) is obtained by summing up the above result for all OD pairs. □

Remark. It is possible to show that $\boldsymbol{h}^{\mathrm{INF}}$ is the only path flow vector that ensures the equation holds for all feasible path flow vectors. But since this result is related to our topic we will not prove it.

Lemma 4.4. *If* $\boldsymbol{G}$ *is a non-zero, symmetric and positive semidefinite matrix, then for any feasible path flow vectors* $\boldsymbol{h}$ *and* $\tilde{\boldsymbol{h}}$ *in* S_h*, such that all components of* $\tilde{\boldsymbol{h}}$ *are positive, the following inequality holds*

$$(\log \frac{\boldsymbol{h}}{\boldsymbol{q}} - \log \frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}})^{\mathrm{T}} \boldsymbol{h} \geqslant A(\boldsymbol{h} - \tilde{\boldsymbol{h}})^{\mathrm{T}} \boldsymbol{G} (\boldsymbol{h} - \tilde{\boldsymbol{h}}). \tag{4.15}$$

Here A *is a constant defined by* $A = \frac{1}{2 q_{\max} \nu_{\max}(\boldsymbol{G})}$*, where* $q_{\max} = \max_{k \in K} q^k$ *represents the largest demand of all OD pairs, and* $\nu_{\max}(\boldsymbol{G})$ *represents the largest eigenvalue of* $\boldsymbol{G}$*.*

Remark. The assumption that $\boldsymbol{G}$ is symmetric and positive semidefinite implies that $\boldsymbol{G}$ must have real and nonnegative eigenvalues. The assumption that $\boldsymbol{G}$ is not a zero matrix ensures that $\nu_{\max}(\boldsymbol{G}) > 0$. The assumption that all components of $\tilde{\boldsymbol{h}}$ are positive, namely $\tilde{\boldsymbol{h}}$ is in the interior of S_h ensures that the term $(\log \frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}})^{\mathrm{T}}\boldsymbol{h}$ is well-defined.

Proof. For any OD pair $k \in K$, consider the following scalar function $f: S_p^k \to \mathbb{R}$, defined as $f(\boldsymbol{p}^k) = (\log \boldsymbol{p}^k)^{\mathrm{T}}\boldsymbol{p}^k$. The domain of this function S_p^k is the set of vectors of feasible path flow probabilities, namely $S_p^k = \left\{\boldsymbol{p}^k \in [0,1]^{|R^k|} : \sum_{r \in R^k} p_r = 1\right\}$. The gradient and Hessian matrix of this function are respectively

$$\boldsymbol{\nabla} f = \log \boldsymbol{p}^k + \mathbf{1},$$

$$(\boldsymbol{H}_f)_{rs} = \begin{cases} 1/p_r, & \text{if } r = s, \\ 0, & \text{otherwise.} \end{cases}$$

Note that in the above expressions, $\mathbf{1}$ is a $|R^k|$-dimensional vector of all 1's. The Hessian matrix $\boldsymbol{H}_f$ is a diagonal matrix, and its rth diagonal component is $\frac{1}{p_r} \geqslant 1$. Hence for any two vectors $\boldsymbol{p}^k$ and $\tilde{\boldsymbol{p}}^k$ in S_p^k, by Taylor Expansion, there exists a path flow vector $\bar{\boldsymbol{p}}^k = \tilde{\boldsymbol{p}}^k + \lambda(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k)$, with $\lambda \in [0,1]$, such that

$$\begin{aligned}(\log \boldsymbol{p}^k)^{\mathrm{T}}\boldsymbol{p}^k - (\log \tilde{\boldsymbol{p}}^k)^{\mathrm{T}}\tilde{\boldsymbol{p}}^k &= f(\boldsymbol{p}^k) - f(\tilde{\boldsymbol{p}}^k) \\ &= \nabla f(\tilde{\boldsymbol{p}}^k)^{\mathrm{T}}(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k) + \frac{1}{2}(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k)^{\mathrm{T}}\boldsymbol{H}_f(\bar{\boldsymbol{p}}^k)(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k) \\ &\geqslant (\log \tilde{\boldsymbol{p}}^k + \mathbf{1})^{\mathrm{T}}(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k) + \frac{1}{2}(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k)^{\mathrm{T}}(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k) \\ &= (\log \tilde{\boldsymbol{p}}^k)^{\mathrm{T}}(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k) + \frac{1}{2}(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k)^{\mathrm{T}}(\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k).\end{aligned}$$

The third line follows from the fact that all diagonal elements of $\boldsymbol{H}_f(\bar{\boldsymbol{p}}^k)$ are at least 1. Note that the same term $-(\log \tilde{\boldsymbol{p}}^k)^{\mathrm{T}}\tilde{\boldsymbol{p}}^k$ appears on both sides of the inequality, so it cancels out. Hence

$$(\log \boldsymbol{p}^k - \log \tilde{\boldsymbol{p}}^k)^{\mathrm{T}} \boldsymbol{p}^k \geqslant \frac{1}{2} (\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k)^{\mathrm{T}} (\boldsymbol{p}^k - \tilde{\boldsymbol{p}}^k).$$

Let $\boldsymbol{p}^k = \frac{\boldsymbol{h}^k}{q^k}$ and $\tilde{\boldsymbol{p}}^k = \frac{\tilde{\boldsymbol{h}}^k}{q^k}$, we have that

$$\begin{aligned}(\log \frac{\boldsymbol{h}^k}{q^k} - \log \frac{\tilde{\boldsymbol{h}}^k}{q^k})^{\mathrm{T}} \boldsymbol{h}^k &\geqslant \frac{q^k}{2} (\frac{\boldsymbol{h}^k}{q^k} - \frac{\tilde{\boldsymbol{h}}^k}{q^k})^{\mathrm{T}} (\frac{\boldsymbol{h}^k}{q^k} - \frac{\tilde{\boldsymbol{h}}^k}{q^k}) \\ &\geqslant \frac{1}{2q^k} (\boldsymbol{h}^k - \tilde{\boldsymbol{h}}^k)^{\mathrm{T}} (\boldsymbol{h}^k - \tilde{\boldsymbol{h}}^k) \\ &\geqslant \frac{1}{2q_{\max}} (\boldsymbol{h}^k - \tilde{\boldsymbol{h}}^k)^{\mathrm{T}} (\boldsymbol{h}^k - \tilde{\boldsymbol{h}}^k).\end{aligned}$$

Summing up this inequality over all $k \in K$, we obtain that

$$\begin{aligned}(\log \frac{\boldsymbol{h}}{\boldsymbol{q}} - \log \frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}})^{\mathrm{T}} \boldsymbol{h} &\geqslant \frac{1}{2q_{\max}} (\boldsymbol{h} - \tilde{\boldsymbol{h}})^{\mathrm{T}} (\boldsymbol{h} - \tilde{\boldsymbol{h}}) \\ &\geqslant \frac{1}{2q_{\max} \nu_{\max}(\boldsymbol{G})} (\boldsymbol{h} - \tilde{\boldsymbol{h}})^{\mathrm{T}} \boldsymbol{G} (\boldsymbol{h} - \tilde{\boldsymbol{h}}).\end{aligned}$$

In the last inequality we have used the fact that $(\boldsymbol{h} - \tilde{\boldsymbol{h}})^{\mathrm{T}} \boldsymbol{G} (\boldsymbol{h} - \tilde{\boldsymbol{h}}) \leqslant \nu_{\max}(\boldsymbol{G}) (\boldsymbol{h} - \tilde{\boldsymbol{h}})^{\mathrm{T}} (\boldsymbol{h} - \tilde{\boldsymbol{h}})$. □

Corollary 4.5. *If $\boldsymbol{G}$ is a non-zero, symmetric and positive semidefinite matrix, then for any feasible path flow vectors $\boldsymbol{h}$ and $\tilde{\boldsymbol{h}}$ in S_h, such that all components of both $\boldsymbol{h}$ and $\tilde{\boldsymbol{h}}$ are positive, the following inequality holds*

$$(\log \frac{\boldsymbol{h}}{\boldsymbol{q}} - \log \frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}})^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}} \right) \geqslant 2A (\boldsymbol{h} - \tilde{\boldsymbol{h}})^{\mathrm{T}} \boldsymbol{G} (\boldsymbol{h} - \tilde{\boldsymbol{h}}), \tag{4.16}$$

where A is defined in Lemma 4.4.

Proof. By Lemma 4.4, we have that

$$(\log\frac{\boldsymbol{h}}{\boldsymbol{q}} - \log\frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}})^{\mathrm{T}}\boldsymbol{h} \geqslant A(\boldsymbol{h}-\tilde{\boldsymbol{h}})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{h}-\tilde{\boldsymbol{h}}),$$

$$(\log\frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}} - \log\frac{\boldsymbol{h}}{\boldsymbol{q}})^{\mathrm{T}}\tilde{\boldsymbol{h}} \geqslant A(\boldsymbol{h}-\tilde{\boldsymbol{h}})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{h}-\tilde{\boldsymbol{h}}).$$

Eq. (4.17) is obtained by summing up the above two inequalities. □

Corollary 4.6. *If $\boldsymbol{G}$ is a non-zero, symmetric and positive semidefinite matrix, then for any feasible path flow vectors $\boldsymbol{h}$ and $\tilde{\boldsymbol{h}}$ in S_h, such that all components of both $\boldsymbol{h}$ and $\tilde{\boldsymbol{h}}$ are positive, the following inequality holds*

$$(\log\frac{\boldsymbol{h}}{\boldsymbol{q}})^{\mathrm{T}}\left(\boldsymbol{h}-\tilde{\boldsymbol{h}}\right) \geqslant A(\boldsymbol{h}-\tilde{\boldsymbol{h}})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{h}-\tilde{\boldsymbol{h}}) + A(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{INF}}) - B, \quad (4.17)$$

where A is defined in Lemma 4.4, and $B = -(\log\frac{\boldsymbol{h}^{\mathrm{INF}}}{\boldsymbol{q}})^{\mathrm{T}}\boldsymbol{h}^{\mathrm{INF}} = \sum_{k\in K} q^k \log|R^k|$ is a constant that depends on the traffic demand flows and the number of paths available for each OD pair.

Proof. By Lemma 4.4 and Lemma 4.3, we have that

$$(\log\frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}} - \log\frac{\boldsymbol{h}}{\boldsymbol{q}})^{\mathrm{T}}\tilde{\boldsymbol{h}} \geqslant Az^{\mathrm{DUE}} + \frac{\theta B+\lambda J}{2+2\theta A} + \sqrt{\frac{\theta B+\lambda J}{2+2\theta A}}\sqrt{\frac{\theta B+\lambda J}{2+2\theta A}+2z^{\mathrm{DUE}}}(\boldsymbol{h}-\tilde{\boldsymbol{h}})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{h}-\tilde{\boldsymbol{h}}),$$

$$(\log\frac{\boldsymbol{h}}{\boldsymbol{q}} - \log\frac{\boldsymbol{h}^{\mathrm{INF}}}{\boldsymbol{q}})^{\mathrm{T}}\boldsymbol{h} \geqslant A(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{h}-\boldsymbol{h}^{\mathrm{INF}}),$$

$$(\log\frac{\boldsymbol{h}^{\mathrm{INF}}}{\boldsymbol{q}})^{\mathrm{T}}\boldsymbol{h} = (\log\frac{\boldsymbol{h}^{\mathrm{INF}}}{\boldsymbol{q}})^{\mathrm{T}}\boldsymbol{h}^{\mathrm{INF}}\,(=-B)\,.$$

Furthermore, since all components in $\log\frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}}$ are nonpositive, we have that

$$0 \geqslant (\log\frac{\tilde{\boldsymbol{h}}}{\boldsymbol{q}})^{\mathrm{T}}\tilde{\boldsymbol{h}}.$$

Eq. (4.17) is obtained by summing up the above four equations. □

Lemma 4.7. *If* $\boldsymbol{G}$ *is a non-zero, symmetric and positive semidefinite matrix, with only non-negative elements, then for any feasible path flow vectors* $\boldsymbol{h}$ *and* $\tilde{\boldsymbol{h}}$ *in* S_h*, the following inequality holds*

$$(\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H})\boldsymbol{G}\})^{\mathrm{T}}\left(\boldsymbol{h}-\tilde{\boldsymbol{h}}\right) \geqslant -J$$

where $\mathrm{Cov}(\boldsymbol{H})$ *is a function of* $\boldsymbol{h}$ *defined in Eq.* (3.7)*, and* J *is a constant defined as* $J = 4\sqrt{|R|}\nu_{\max}(\boldsymbol{G})^2\left(\sum_k q^k\right)^{3/2}$*, and* $\nu_{max}(\boldsymbol{G})$ *gives the largest eigenvalue of matrix* $\boldsymbol{G}$*.*

Proof. Consider OD pair $k \in K$. For $s, s' \in R^k$, the (s, s') component of matrix $\left(\mathrm{Cov}(\boldsymbol{H}^k)\right)$ is

$$\left(\mathrm{Cov}(\boldsymbol{H}^k)\right)_{s,s'} = q^k p_s(\delta_{s,s'} - p_{s'}),$$

where $\delta_{s,s'}$ is the Kronecker delta function. Since $|\delta_{s,s'} - p_{s'}| \leqslant 1$, we have that

$$\left|\left(\mathrm{Cov}(\boldsymbol{H}^k)\right)_{s,s'}\right| \leqslant q^k p_s = h_s. \tag{4.18}$$

Let $\boldsymbol{G} = \boldsymbol{\Delta}^{\mathrm{T}}\bar{\boldsymbol{G}}\boldsymbol{\Delta}$. We have that

$$\begin{aligned}
&(\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H})\boldsymbol{G}\})_r \\
=&\left(\mathrm{Diag}\left\{\boldsymbol{\Delta}^{\mathrm{T}}\bar{\boldsymbol{G}}\sum_{k\in K}\left(\boldsymbol{\Delta}^k\mathrm{Cov}(\boldsymbol{H}^k)(\boldsymbol{\Delta}^k)^{\mathrm{T}}\right)\bar{\boldsymbol{G}}\boldsymbol{\Delta}\right\}\right)_r \\
=&\left(\boldsymbol{\Delta}^{\mathrm{T}}\bar{\boldsymbol{G}}\sum_{k\in K}\left(\boldsymbol{\Delta}^k\mathrm{Cov}(\boldsymbol{H}^k)(\boldsymbol{\Delta}^k)^{\mathrm{T}}\right)\bar{\boldsymbol{G}}\boldsymbol{\Delta}\right)_{r,r} \\
=&\sum_{k\in K}\sum_{s,s'\in R^k} G_{r,s}G_{r,s'}\left(\mathrm{Cov}(\boldsymbol{H}^k)\right)_{s,s'}.
\end{aligned}$$

Since the elements of matrix $\boldsymbol{G}$ are nonnegative, taking the norm and applying

Eq. (4.18), we have that

$$
\begin{aligned}
& |(\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H})\boldsymbol{G}\})_r| \\
\leqslant & \sum_{k \in K} \sum_{s,s' \in R^k} G_{r,s} G_{r,s'} \left| \left(\mathrm{Cov}(\boldsymbol{H}^k) \right)_{s,s'} \right| \\
\leqslant & \sum_{k \in K} \sum_{s,s' \in R^k} G_{r,s} G_{r,s'} h_s \\
= & 2 \sum_{s,s' \in R} G_{r,s} G_{r,s'} h_s .
\end{aligned}
$$

Applying Lemma 3.6, we have that

$$
\begin{aligned}
& |\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H})\boldsymbol{G}\}| \\
\leqslant & 2 \sum_{r \in R} \sum_{s,s' \in R} G_{r,s} G_{r,s'} h_s \\
= & 2 \mathbf{1} \boldsymbol{G}\boldsymbol{G}\boldsymbol{h} \\
\leqslant & 2 \left| \mathbf{1} \right| \left\| \boldsymbol{G}\boldsymbol{G} \right\| \left| \boldsymbol{h} \right| \\
= & 2 \sqrt{|R| \sum_k q^k} \nu_{\max} \left(\boldsymbol{G} \right)^2 ,
\end{aligned}
\tag{4.19}
$$

where $\mathbf{1}$ is the column vector of all 1's (with dimension $|R| \times 1$) and $\|\cdot\|$ denotes the Euclidean matrix norm. The last line follows from the facts that (1) $|\mathbf{1}| = \sqrt{|R|}$, (2) $|\boldsymbol{h}| \leqslant \sum_k q^k$, and (3) $\|\boldsymbol{G}\boldsymbol{G}\| = \|\boldsymbol{G}\|^2 = \nu_{\max}(\boldsymbol{G})^2$. Hence we have that

$$
\begin{aligned}
& (\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H})\boldsymbol{G}\})^{\mathrm{T}} \left(\boldsymbol{h} - \tilde{\boldsymbol{h}} \right) \\
\geqslant & - |\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H})\boldsymbol{G}\}| \left| \boldsymbol{h} - \tilde{\boldsymbol{h}} \right| \\
\geqslant & - 4 \sqrt{|R|} \nu_{\max} \left(\boldsymbol{G} \right)^2 (\sum_k q^k)^{3/2} .
\end{aligned}
\tag{4.20}
$$

In the last step we have used the fact that $\left| \boldsymbol{h} - \tilde{\boldsymbol{h}} \right| \leqslant \sum_r |h_r - \tilde{h}_r| \leqslant \sum_r (h_r + \tilde{h}_r) =$

$2\sum_k q^k$. □

Remark. The lower bound proved in this lemma is a loose bound, largely because we have used a series of loose inequalities in Eq. (4.19) and (4.20). Nevertheless this result allows us to prove the upper bounds for the TSUE system cost in the following subsection.

4.2.4 Proof of the bounds

Next we prove three uppers bounds for the TSUE system cost. It should be noted that we can obtain a single upper bound by taking the minimum of all three bounds. We present the three bounds separately for the sake of clarity.

Theorem 4.8. *If $\boldsymbol{G}$ is a non-zero, symmetric and positive semidefinite matrix, with only non-negative elements, if $\boldsymbol{b} \geqslant 0$, if θ satisfies $2(\theta B + \lambda J)(1 + 2\theta A) \leqslant z^{\mathrm{DUE}}$, then the system cost for the TSUE model $z^{\mathrm{TSUE}}(\theta, \lambda)$ has the following upper bound*

$$z^{\mathrm{TSUE}}(\theta, \lambda) \leqslant z^{\mathrm{DUE}} + \frac{\theta B + \lambda J}{2 + 2\theta A} + \sqrt{\frac{\theta B + \lambda J}{2 + 2\theta A}}\sqrt{\frac{\theta B + \lambda J}{2 + 2\theta A} + 2z^{\mathrm{DUE}}}, \tag{4.21}$$

where $\theta > 0$ is the dispersion level, $\lambda \geqslant 0$ is the mean-variance trade-off factor, $A = \frac{1}{2q_{\max}^2 \nu_{\max}(\boldsymbol{G})}$, $B = \sum_{k \in K} q^k \log |R^k|$, and $J = 4\sqrt{|R|}\nu_{\max}(\boldsymbol{G})^2 \left(\sum_k q^k\right)^{3/2}$.

Proof. Using the VI formulation of the TSUE model, namely Eq. (3.14), we have that

$$\left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} + \boldsymbol{b} + \lambda \mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H})\boldsymbol{G}\} + \theta \log \frac{\boldsymbol{h}^{\mathrm{DUE}}}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\boldsymbol{h}^{\mathrm{DUE}} - \boldsymbol{h}^{\mathrm{TSUE}}\right) = 0.$$

That is,

$$
\begin{aligned}
&\left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}+\boldsymbol{b}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{DUE}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)\\
=&\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{TSUE}}}{\boldsymbol{q}}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{TSUE}}-\boldsymbol{h}^{\mathrm{DUE}}\right)\\
&+\lambda(\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H})\boldsymbol{G}\})^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{TSUE}}-\boldsymbol{h}^{\mathrm{DUE}}\right)\\
\geqslant&-\theta B-\lambda J+\theta A\left(\boldsymbol{h}^{\mathrm{DUE}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)^{\mathrm{T}}\boldsymbol{G}\left(\boldsymbol{h}^{\mathrm{DUE}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)\\
&+\theta A\left(\boldsymbol{h}^{\mathrm{INF}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)^{\mathrm{T}}\boldsymbol{G}\left(\boldsymbol{h}^{\mathrm{INF}}-\boldsymbol{h}^{\mathrm{TSUE}}\right),
\end{aligned}
\tag{4.22}
$$

where the last line follows from Corollary 4.6 and Lemma 4.7.

Using the VI formulation of the DUE model, namely Eq. (2.4), we have that

$$
\left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}+\boldsymbol{b}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{DUE}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)\geqslant 0. \tag{4.23}
$$

Furthermore, Lemma 4.2 implies that

$$
a_1(\boldsymbol{h}^{\mathrm{DUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}}+\frac{1}{4a_1}(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}\geqslant(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}}, \tag{4.24}
$$

$$
a_2(\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}+\frac{1}{4a_2}(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}\geqslant(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}, \tag{4.25}
$$

where $a_1 > 0$ and $a_2 > 0$.

Multiply Eq. (4.23), (4.24) and (4.25) respectively by nonnegative constants C_1, C_2 and C_3, and add to Eq. (4.22). We have that

$$
\begin{aligned}
&(\frac{C_3}{4a_2}-\theta A)(\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}+\theta B+\lambda J\\
&+(-C_3+2\theta A)(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}+(1+C_1-C_2+2\theta A)(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}}\\
&+(-1+\frac{C_2}{4a_1}+a_2C_3-2\theta A)(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}-(1-C_1)\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}}\\
&+(-C_1+a_1C_2-\theta A)\,(\boldsymbol{h}^{\mathrm{DUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}}+(1-C_1)\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{DUE}}\geqslant 0.
\end{aligned}
\tag{4.26}
$$

We select

$$\begin{cases} C_2 = 1 + C_1 + 2\theta A, \\ C_3 = 2\theta A, \\ a_1 = \frac{1+\theta A}{1+C_1+2\theta A}, \\ a_2 = \frac{1}{2}. \end{cases}$$

Then the following equations hold.

$$-C_3 + 2\theta A = 0,$$
$$1 + C_1 - C_2 + 2\theta A = 0,$$
$$\frac{C_3}{4a_2} - \theta A = 0,$$
$$-C_1 + a_1 C_2 - \theta A = 1 - C_1.$$

The above equations suggest that, in Eq. (4.26), the three terms, namely, $(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}$, $(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}}$, and $(\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}$, are eliminated. Furthermore, the coefficient of these two terms $(\boldsymbol{h}^{\mathrm{DUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}}$ and $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{DUE}}$ are equalized. We have that

$$\begin{aligned} &(1 - C_1)(4 + 4\theta A)\left((\boldsymbol{h}^{\mathrm{DUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{DUE}}\right) + (4 + 4\theta A)(\theta B + \lambda J) \\ \geqslant &(1 - C_1)(3 + C_1 + 4\theta A)(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} + (1 - C_1)(4 + 4\theta A)\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}}. \end{aligned}$$

When $C_1 < 1$, we have $(1 - C_1)(4 + 4\theta A) > (1 - C_1)(3 + C_1 + 4\theta A)$, hence

$$\begin{aligned} &(1 - C_1)(4 + 4\theta A)\left((\boldsymbol{h}^{\mathrm{DUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{DUE}} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{DUE}}\right) + (4 + 4\theta A)(\theta B + \lambda J) \\ \geqslant &(1 - C_1)(3 + C_1 + 4\theta A)\left((\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}}\right). \end{aligned}$$

That is,

$$z^{\mathrm{TSUE}}(\theta) \leqslant \frac{4 + 4\theta A}{3 + C_1 + 4\theta A}\left(z^{\mathrm{DUE}} + \frac{\theta B + \lambda J}{1 - C_1}\right). \tag{4.27}$$

This upper bound holds for any value of $C_1 \in [0, 1)$. By minimizing this upper bound over C_1, we set

$$C_1 = 1 + \frac{\theta B + \lambda J}{z^{\mathrm{DUE}}} - \frac{\theta B + \lambda J}{z^{\mathrm{DUE}}} \sqrt{1 + \frac{4 z^{\mathrm{DUE}}}{\theta B + \lambda J} + \frac{4\theta A z^{\mathrm{DUE}}}{\theta B + \lambda J}}.$$

It is obvious that $C_1 < 1$. When $2(\theta B + \lambda J)(1 + 2\theta A) \leqslant z^{\mathrm{DUE}}$, we have $C_1 \geqslant 0$. Substituting the value of C_1 back into Eq. (4.27), we obtain Eq. (4.21). □

Theorem 4.9. *If $\boldsymbol{G}$ is a non-zero, symmetric and positive semidefinite matrix, with only non-negative elements, if $\boldsymbol{b} \geqslant 0$, then the system cost for the TSUE model $z^{\mathrm{TSUE}}(\theta, \lambda)$ has the following upper bound*

$$z^{\mathrm{TSUE}}(\theta, \lambda) \leqslant \frac{4 + 4\theta A}{3 + 4\theta A} \left(z^{\mathrm{SO}} + \theta B + \lambda J \right), \tag{4.28}$$

where $\theta > 0$ is the dispersion level, $\lambda \geqslant 0$ is the mean-variance trade-off factor, $A = \frac{1}{2 q_{\max}^2 \nu_{\max}(\boldsymbol{G})}$, $B = \sum_{k \in K} q^k \log |R^k|$, and $J = 4\sqrt{|R|} \nu_{\max}(\boldsymbol{G})^2 (\sum_k q^k)^{3/2}$.

Proof. Using Lemma (4.2), the following inequality holds

$$a(\boldsymbol{h}^{\mathrm{SO}})^{\mathrm{T}} \boldsymbol{G} \boldsymbol{h}^{\mathrm{SO}} + \frac{1}{4a} (\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}} \boldsymbol{G} \boldsymbol{h}^{\mathrm{TSUE}} \geqslant (\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}} \boldsymbol{G} \boldsymbol{h}^{\mathrm{SO}}, \tag{4.29}$$

where $a > 0$. The VI formulation of the TSUE model (see Eq. (3.14)) implies that

$$\left(\boldsymbol{G} \boldsymbol{h}^{\mathrm{TSUE}} + \boldsymbol{b} + \lambda \mathrm{Diag}\{\boldsymbol{G} \mathrm{Cov}(\boldsymbol{H}^{\mathrm{TSUE}}) \boldsymbol{G}\} + \theta \log \frac{\boldsymbol{h}^{\mathrm{TSUE}}}{\boldsymbol{q}} \right)^{\mathrm{T}} \left(\boldsymbol{h}^{\mathrm{SO}} - \boldsymbol{h}^{\mathrm{TSUE}} \right) = 0.$$

That is,

$$
\begin{aligned}
&\left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}+\boldsymbol{b}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{SO}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)\\
=&\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{TSUE}}}{\boldsymbol{q}}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{TSUE}}-\boldsymbol{h}^{\mathrm{SO}}\right)\\
&+\lambda(\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H}^{\mathrm{TSUE}})\boldsymbol{G}\})^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{TSUE}}-\boldsymbol{h}^{\mathrm{SO}}\right)\\
\geqslant&-\theta B-\lambda J+\theta A\left(\boldsymbol{h}^{\mathrm{SO}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)^{\mathrm{T}}\boldsymbol{G}\left(\boldsymbol{h}^{\mathrm{SO}}-\boldsymbol{h}^{\mathrm{TSUE}}\right).
\end{aligned}
\tag{4.30}
$$

The third line follows from Corollary 4.6 and Lemma 4.7.

Multiplying Eq. (4.29) by a constant $C \geqslant 0$ and adding to Eq. (4.30), we have that

$$
\begin{aligned}
&(aC-\theta A)(\boldsymbol{h}^{\mathrm{SO}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{SO}}+(-1+\frac{C}{4a}-\theta A)(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}\\
&+(1-C+2\theta A)(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{SO}}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{SO}}-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}}+\theta B+\lambda J\geqslant 0.
\end{aligned}
\tag{4.31}
$$

We select

$$
\begin{cases}
C = 1+2\theta A,\\
a = \frac{1+\theta A}{1+2\theta A}.
\end{cases}
$$

Then the following equations hold.

$$
1-C+2\theta A = 0,
$$

$$
aC-\theta A = 1.
$$

The above equations suggest that, in Eq. (4.31), the term $(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{SO}}$ is eliminated. Furthermore, the coefficient of these two terms $(\boldsymbol{h}^{\mathrm{SO}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{SO}}$ and $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{SO}}$ are equalized. We have that

$$
\begin{aligned}
&(\boldsymbol{h}^{\mathrm{SO}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{SO}} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{SO}} + \theta B + \lambda J \\
\geqslant & \frac{3+4\theta A}{4+4\theta A}(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}} \\
\geqslant & \frac{3+4\theta A}{4+4\theta A}\left((\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}}\right).
\end{aligned}
$$

The third line follows from the facts that $1 \geqslant \frac{3+4\theta A}{4+4\theta A}$ and $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}} \geqslant 0$. The above equation is $z^{\mathrm{TSUE}}(\theta,\lambda) \leqslant \frac{4+4\theta A}{3+4\theta A}(z^{\mathrm{SO}} + \theta B + \lambda J)$. □

Theorem 4.10. *If $\boldsymbol{G}$ is a non-zero, symmetric and positive semidefinite matrix, with only non-negative elements, if $\boldsymbol{b} \geqslant 0$, then the system cost for the TSUE model $z^{\mathrm{TSUE}}(\theta,\lambda)$ has the following upper bound*

$$
z^{\mathrm{TSUE}}(\theta,\lambda) \leqslant \frac{4+8\theta A}{3+8\theta A}(z^{\mathrm{INF}} + \lambda J) \leqslant \frac{4}{3}(z^{\mathrm{INF}} + \lambda J), \tag{4.32}
$$

where $\theta > 0$ is the dispersion level, $\lambda \geqslant 0$ is the mean-variance trade-off factor, $A = \frac{1}{2q_{\max}^2 \nu_{\max}(\boldsymbol{G})}$, $B = \sum_{k\in K} q^k \log|R^k|$, and $J = 4\sqrt{|R|}\nu_{\max}(\boldsymbol{G})^2 \left(\sum_k q^k\right)^{3/2}$.

Proof. Lemma 4.2 implies that

$$
a(\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}} + \frac{1}{4a}(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} \geqslant (\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}, \tag{4.33}
$$

where $a > 0$ is a positive constant. The VI formulation of the TSUE model (see Eq. (3.14)) implies that

$$
\left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}} + \boldsymbol{b} + \lambda \mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H}^{\mathrm{TSUE}})\boldsymbol{G}\} + \theta \log \frac{\boldsymbol{h}^{\mathrm{SUE}}}{\boldsymbol{q}}\right)^{\mathrm{T}} \left(\boldsymbol{h}^{\mathrm{INF}} - \boldsymbol{h}^{\mathrm{SUE}}\right) = 0.
$$

That is,

$$
\begin{aligned}
&\left(\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}+\boldsymbol{b}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{INF}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)\\
=&\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{TSUE}}}{\boldsymbol{q}}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{SUE}}-\boldsymbol{h}^{\mathrm{INF}}\right)\\
&\quad+\lambda(\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H}^{\mathrm{TSUE}})\boldsymbol{G}\})^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{TSUE}}-\boldsymbol{h}^{\mathrm{INF}}\right)\\
=&\theta\left(\log\frac{\boldsymbol{h}^{\mathrm{TSUE}}}{\boldsymbol{q}}-\log\frac{\boldsymbol{h}^{\mathrm{INF}}}{\boldsymbol{q}}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{TSUE}}-\boldsymbol{h}^{\mathrm{INF}}\right)\\
&\quad+\lambda(\mathrm{Diag}\{\boldsymbol{G}\mathrm{Cov}(\boldsymbol{H}^{\mathrm{TSUE}})\boldsymbol{G}\})^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{TSUE}}-\boldsymbol{h}^{\mathrm{INF}}\right)\\
\geqslant&2\theta A\left(\boldsymbol{h}^{\mathrm{INF}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)^{\mathrm{T}}\boldsymbol{G}\left(\boldsymbol{h}^{\mathrm{INF}}-\boldsymbol{h}^{\mathrm{TSUE}}\right)-\lambda J.
\end{aligned}
\tag{4.34}
$$

In the third line we used the fact that $\left(\log\frac{\boldsymbol{h}^{\mathrm{INF}}}{\boldsymbol{q}}\right)^{\mathrm{T}}\left(\boldsymbol{h}^{\mathrm{TSUE}}-\boldsymbol{h}^{\mathrm{INF}}\right)=0$ (see Lemma 4.3). The last line follows from Corollary 4.6 and Lemma 4.7.

Multiplying (4.33) by a constant $C\geqslant 0$ and adding to (4.34) implies that

$$
\begin{aligned}
&(aC-2\theta A)(\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}+(-1+\frac{C}{4a}-2\theta A)(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}\\
&\quad+(1-C+4\theta A)(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{INF}}-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}}-\lambda J\geqslant 0.
\end{aligned}
\tag{4.35}
$$

We select

$$
\begin{cases}
C=1+4\theta A,\\
a=\frac{1+2\theta A}{1+4\theta A}.
\end{cases}
$$

Then the following equations hold.

$$
\begin{aligned}
1-C+4\theta A&=0,\\
aC-2\theta A&=1.
\end{aligned}
$$

The above equations suggest that, in Eq.(4.35), the term $(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}$ is eliminated. Furthermore, the coefficient of these two terms $(\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}$ and $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{INF}}$

are equalized. We have that

$$\begin{aligned}
&(\boldsymbol{h}^{\mathrm{INF}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{INF}}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{INF}}+\lambda J\\
\geqslant&\frac{3+8\theta A}{4+8\theta A}(\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}}\\
\geqslant&\frac{3+8\theta A}{4+8\theta A}\left((\boldsymbol{h}^{\mathrm{TSUE}})^{\mathrm{T}}\boldsymbol{G}\boldsymbol{h}^{\mathrm{TSUE}}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}}\right).
\end{aligned}$$

The third line follows from the facts that $1 \geqslant \frac{3+8\theta A}{4+8\theta A}$ and $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{TSUE}} \geqslant 0$. The above equation is $z^{\mathrm{TSUE}}(\theta,\lambda) \leqslant \frac{4+8\theta A}{3+8\theta A}(z^{\mathrm{INF}}+\lambda J)$. □

4.2.5 Tightness of the bounds

In this subsection we discuss the tightness of the three upper bounds for the TSUE system cost. When $\lambda > 0$, as noted in the remark that follows Lemma 4.7, these upper bounds can be loose. Hence we instead focus on the case of $\lambda = 0$, namely we will study the tightness of the upper bounds for the SUE system cost.

In general, an upper bound for the SUE system cost can be written in the following form

$$z^{\mathrm{SUE}}(\theta) \leqslant F^{\mathrm{BD}}(z^{\mathrm{DUE}}, z^{\mathrm{SO}}, z^{\mathrm{INF}}, A, B, \theta), \tag{4.36}$$

where $F^{\mathrm{BD}}(\cdot)$ is a function describing the corresponding bound. For $\theta_0 \in [0,\infty) \cup \{+\infty\}$, we say F^{BD} is *tight* at $\theta = \theta_0$ if and only if there exists at least one network instance for which the equality sign in Eq. (4.36) holds for $\theta = \theta_0$. We say F^{BD} is *exact* at $\theta = \theta_0$ if and only if for any arbitrary network, Eq. (4.36) holds as an equality for $\theta = \theta_0$.

It is easy to see that an exact bound is always tight. Nevertheless, the reverse is not always true. For example, the bound on the DUE system cost developed by Roughgarden and Tardos (2002), i.e., $z^{\mathrm{DUE}} \leqslant \frac{4}{3}z^{\mathrm{SO}}$, is tight but not exact. It is tight because we can find a network instance (discussed below) for which $z^{\mathrm{DUE}} = \frac{4}{3}z^{\mathrm{SO}}$.

However the equality does not hold for any arbitrary networks.

Recall that the three upper bounds we proved are (when $\lambda = 0$)

$$z^{\mathrm{SUE}}(\theta) \leqslant z^{\mathrm{DUE}} + \frac{\theta B}{2+2\theta A} + \sqrt{\frac{\theta B}{2+2\theta A}}\sqrt{\frac{\theta B}{2+2\theta A} + 2z^{\mathrm{DUE}}}, \tag{4.37}$$

$$z^{\mathrm{SUE}}(\theta) \leqslant \frac{4+4\theta A}{3+4\theta A}\left(z^{\mathrm{SO}} + \theta B\right), \tag{4.38}$$

$$z^{\mathrm{SUE}}(\theta) \leqslant \frac{4+8\theta A}{3+8\theta A} z^{\mathrm{INF}} \left(\leqslant \frac{4}{3} z^{\mathrm{INF}}\right). \tag{4.39}$$

Using the above definitions of tightness and exactness, we have the following results.

The first bound is exact at $\theta = 0$.

When $\theta = 0$, the first bound becomes $z^{\mathrm{SUE}}(0) \leqslant z^{\mathrm{DUE}}$. Nevertheless we know that $z^{\mathrm{SUE}}(0) = z^{\mathrm{DUE}}$. Hence this bound is exact.

The second bound is tight at $\theta = 0$.

When $\theta = 0$, the second bound becomes $z^{\mathrm{SUE}}(0) \leqslant \frac{4}{3} z^{\mathrm{SO}}$. Consider the following network with a single OD pair and two parallel links, as shown in Figure B.1a. The demand flow is 100 (we omit the units), and the travel time functions on each link are

$$t_1 = 0.1 f_1,$$

$$t_2 = 10.$$

Solving this network, we have that $\boldsymbol{h}^{\mathrm{SUE}}(0) = (100,\, 0)^{\mathrm{T}}$ and $\boldsymbol{h}^{\mathrm{SO}} = (50,\, 50)^{\mathrm{T}}$. Hence $z^{\mathrm{SUE}}(0) = 1000$ and $z^{\mathrm{SO}} = 750$. For this network $z^{\mathrm{SUE}}(0) = \frac{4}{3} z^{\mathrm{SO}}$. This shows that the second bound is tight.

The third bound is tight at $\theta = 0$, and exact at $\theta = \infty$.

When $\theta = 0$, the third bound becomes $z^{\text{SUE}}(0) \leqslant \frac{4}{3} z^{\text{INF}}$. To show that this bound is tight, consider the previous network. We have that $\boldsymbol{h}^{\text{INF}} = (50, 50)^{\text{T}}$ and $z^{\text{INF}} = 750$. That is, $z^{\text{SUE}}(0) = \frac{4}{3} z^{\text{INF}}$. Hence we showed that the third bound is tight at $\theta = 0$.

When $\theta \to \infty$, the third bound becomes $z^{\text{SUE}}(\infty) \leqslant z^{\text{INF}}$. Since we know $z^{\text{SUE}}(\infty) = z^{\text{INF}}$, this bound is exact.

4.3 Numerical Results on Test Networks

In this section we present the numerical results on test networks.

4.3.1 The test networks

We will examine the DUE, SUE, SO, and TSUE models for seven networks. The properties of these networks are summarized in Table 4.3. Refer to Appendix B for the detailed descriptions of these networks.

We will refer to a network as a small-scale network if it has less than 50 paths. We will refer to a network as a medium-scale network if it has 50 to 1000 paths. We will refer to a network as a large-scale network if it has more than 1000 paths.

Among the seven networks, the first four are small-scale networks. The first network is the network we used to show the tightness of our bounds. The first three networks all have a single OD pair. The fourth network has 5 OD pairs.

The fifth and sixth networks are medium-size networks. They share the same graph (nodes and links) but differ in terms of the number of OD pairs and demand flows. These networks are labeled as Grid-1 and Grid-2 because their nodes and links form a 3×3 grid.

The last network is the modified Sioux Falls network. The link travel time functions of the original Sioux Falls network are BPR-type functions, i.e., $t_l(f_l) = \alpha_l \left(1 + \beta_l \left(\frac{f_l}{\gamma_l}\right)^4\right)$, for all $l \in L$. For the purpose of study in this chapter, we transform these functions into affine functions. The new link travel time functions are defined as $\tilde{t}_l(f_l) = \frac{\alpha_l \beta_l}{\gamma_l} f_l + \alpha_l$, for all $l \in L$. This transformation allows us to evaluate the bounds proved in the previous section. We label this transformed network as Sioux Falls*. (We have also done this transformation for the Grid-1 and Grid-2 networks.)

Table 4.3 also contains the a summary of measures: $\frac{1}{A}$, B, J, $\frac{z^{\text{DUE}}}{z^{\text{SO}}}$, and $\frac{z^{\text{INF}}}{z^{\text{SO}}}$, for these networks. In general, as the scale of the network increases, A becomes smaller, and B and J becomes larger. The value of J grows very quickly as the scale of the network increases, largely because the term $(\sum_k q^k)^{3/2}$ in the expression of J grows very quickly. This observation suggests that our bound in terms of λ may be loose (see also the remark that follows Lemma 4.7).

Network	$\lvert N\rvert$	$\lvert L\rvert$	$\lvert K\rvert$	$\lvert R\rvert$	$\frac{1}{A}$	B	J	$\frac{z^{\text{DUE}}}{z^{\text{SO}}}$	$\frac{z^{\text{INF}}}{z^{\text{SO}}}$
Small-1	2	2	1	2	20.00	69.31	28.28	1.3333	1.0000
Small-2	2	2	1	2	40.00	69.31	113.1	1.0340	1.0014
Small-3	4	5	1	3	114.6	109.9	1138	1.0196	1.2941
Small-4	8	10	5	14	1689	427.7	4.269e6	1.0073	1.1405
Grid-1*	9	24	7	69	331.7	1194	2.339e5	1.0142	1.7967
Grid-2*	9	24	72	644	1382	6946	6.819e8	1.0028	5.3097
Sioux Falls*	24	76	528	26400	8.099e4	1.411e6	2.180e15	1.0031	4.7104

Table 4.3: Summary of test networks.

4.3.2 Results on the SUE system cost

The results on the SUE system cost and its upper bounds are shown in Figure 4.1. We plot the DUE, SUE, and INF system costs, together with the three bounds for the SUE system cost. All values are normalized against the SO system cost. We plot the SUE system cost and its upper bounds vs. different values of θ that range from 10^{-2} to 10^3. Note that in reality the value of θ should be at the level of 10^{-1} to 10^1, since the link travel times for these networks are typically between 1 to 10. However, we intentionally consider a very large range for the value of θ in order to study the performance of the bounds in extreme conditions. The horizontal axis (for θ) is plotted in log scale.

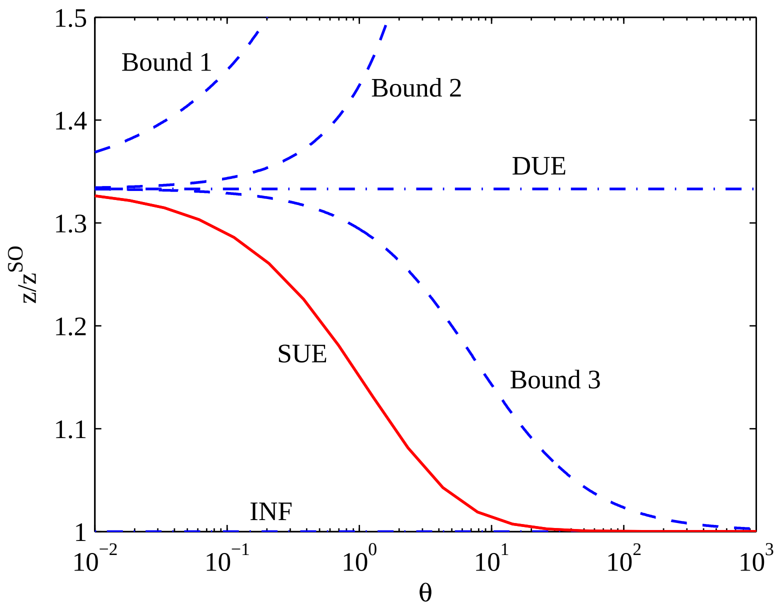
The DUE and INF system costs are not dependent on θ, hence they are plotted as a horizontal line.

We make the following observations:

- The SUE system cost as a function of θ can be increasing (see the plots for Small-4, Grid-2*, and Sioux Falls*), decreasing (see the plots for Small-1 and Small-2), or not monotone (see the plot for Small-3 and Grid-1*).

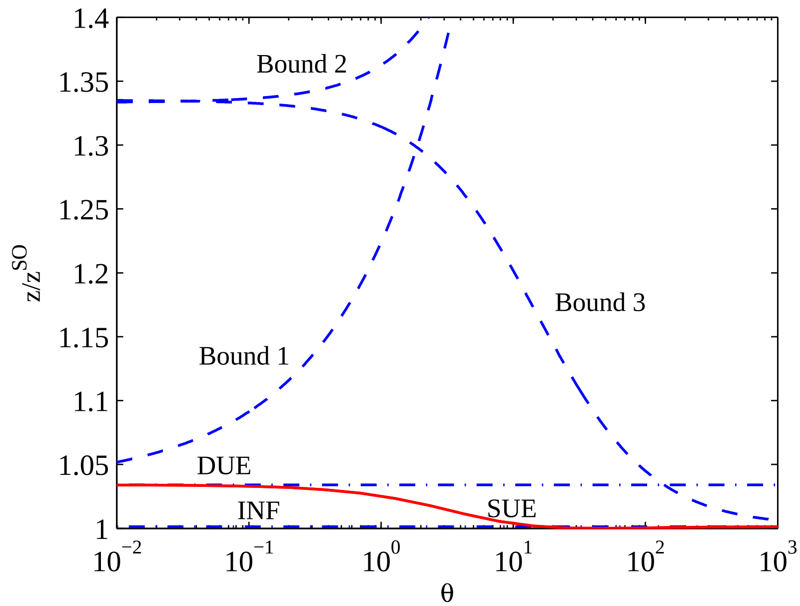
- When θ is small, the SUE system cost is close to the DUE system cost. When θ is large, the SUE system cost is close to the INF system cost. The transition typically happens when θ is between 10^0 and 10^2.

- In general, when θ is close to 0, the first bound is smaller than the second bound (the only exception is the Small-1 network). As θ increases, the first bound increases much faster than the second bound does. When θ passes a critical value (from the plots this value is typically between 10^0 and 10^2), the first bound is greater than the second bound.

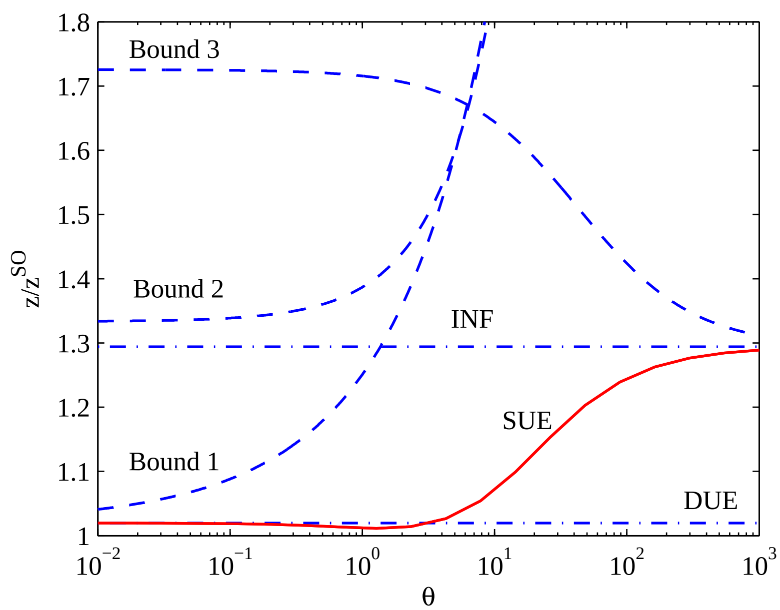


(a) Small-1

Figure 4.1: Numerical results on the SUE system cost and its bounds.

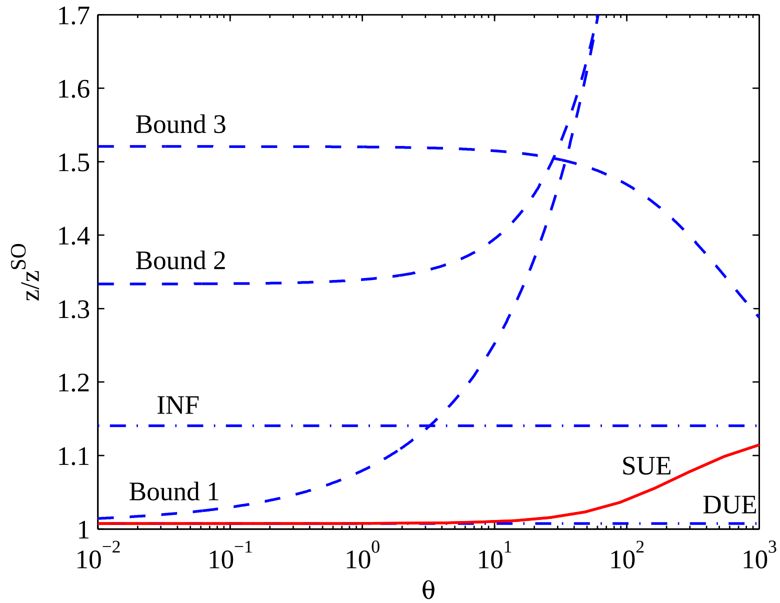


(b) Small-2

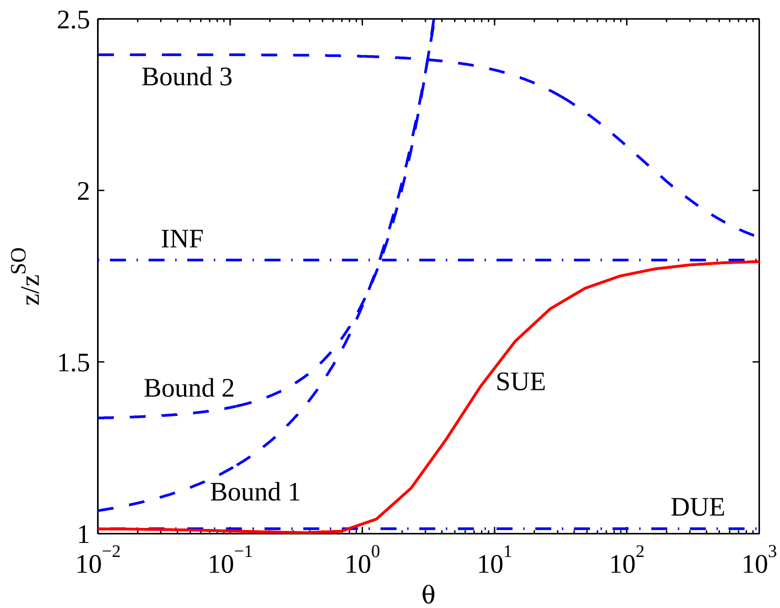


(c) Small-3

Figure 4.1: Numerical results on the SUE system cost and its bounds.(continued)

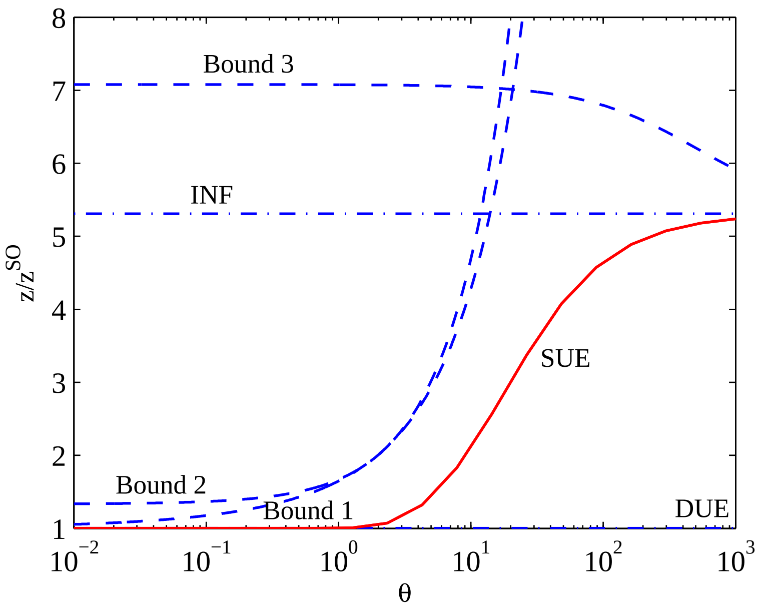


(d) Small-4

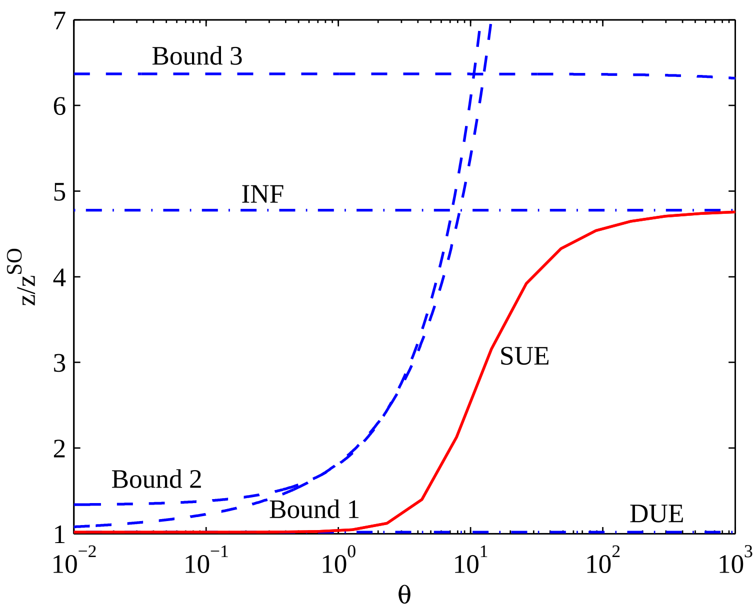


(e) Grid-1*

Figure 4.1: Numerical results on the SUE system cost and its bounds. (continued)



(f) Grid-2*



(g) Sioux Falls*

Figure 4.1: Numerical results on the SUE system cost and its bounds. (continued)

4.3.3 Results on the TSUE system cost

As discussed before (see, for example, the remark following Lemma 4.7), the three upper bounds on the TSUE system cost are very loose when $\lambda > 0$. Hence we will not show these bounds in the following discussion. Instead, we focus on how the TSUE system cost is different from the SUE system cost. (We have already discussed how the SUE system cost is different from the DUE, SO, and INF system costs in the previous subsection.)

In order to numerically solve the TSUE model, we must decide what values of λ to choose. Our choices are shown in Table 4.4. For each network, we present the median variance-to-mean (MVM) ratio defined in Section 3.5.2. The MVM ratio measures the ratio of the variance of travel time over the mean travel time for a typical network user.

As a rule of thumb, for the numerical tests in this section, we will pick values of λ to be roughly 0.1 or 0.2 divided by the MVM ratio. This implies that for a typical traveler, the travel cost attributed to the travel time uncertainty is roughly 10% or 20% of that attributed to the mean travel time. The values of λ are shown in Table 4.4.

Network ID	MVM ratio	λ_1	λ_2
Grid-1*	0.0028	36	72
Grid-2*	0.0020	50	100
Sioux Falls*	4.83e-6	2.1e4	4.2e4

Table 4.4: Choice of the values of λ for the TSUE model.

The results on the TSUE system cost are shown in Figure 4.2 and Table 4.5. We plot the TSUE system cost as a function of θ. The range of θ is chosen to be from 1 to 10. As commented before, this is a "realistic" range for the value of θ. In Figure 4.2, the label TSUE(0) stands for the TSUE system cost when $\lambda = 0$ (i.e., the SUE system cost). Similarly, the labels TSUE(λ_1) and TSUE(λ_2) respectively stand for the TSUE system costs when $\lambda = \lambda_1$ and $\lambda = \lambda_2$.

	Grid-1*		Grid-2*		Sioux-Falls*	
	$\theta = 1$	$\theta = 10$	$\theta = 1$	$\theta = 10$	$\theta = 1$	$\theta = 10$
$z^{\text{TSUE}}(0)/z^{\text{SO}}$	1.0218	1.4859	1.0035	2.1018	1.0336	2.5363
$z^{\text{TSUE}}(\lambda_1)/z^{\text{SO}}$	1.0159	1.4395	1.0023	1.6861	1.0307	1.7431
$z^{\text{TSUE}}(\lambda_2)/z^{\text{SO}}$	1.0124	1.4012	1.0026	1.4995	1.0329	1.5080
$\tilde{z}^{\text{TSUE}}(0)/z^{\text{SO}}$	1.0241	1.4905	1.0043	2.1048	1.0336	2.5363
$\tilde{z}^{\text{TSUE}}(\lambda_1)/z^{\text{SO}}$	1.0180	1.4439	1.0030	1.6886	1.0307	1.7431
$\tilde{z}^{\text{TSUE}}(\lambda_2)/z^{\text{SO}}$	1.0144	1.4055	1.0033	1.5016	1.0329	1.5080

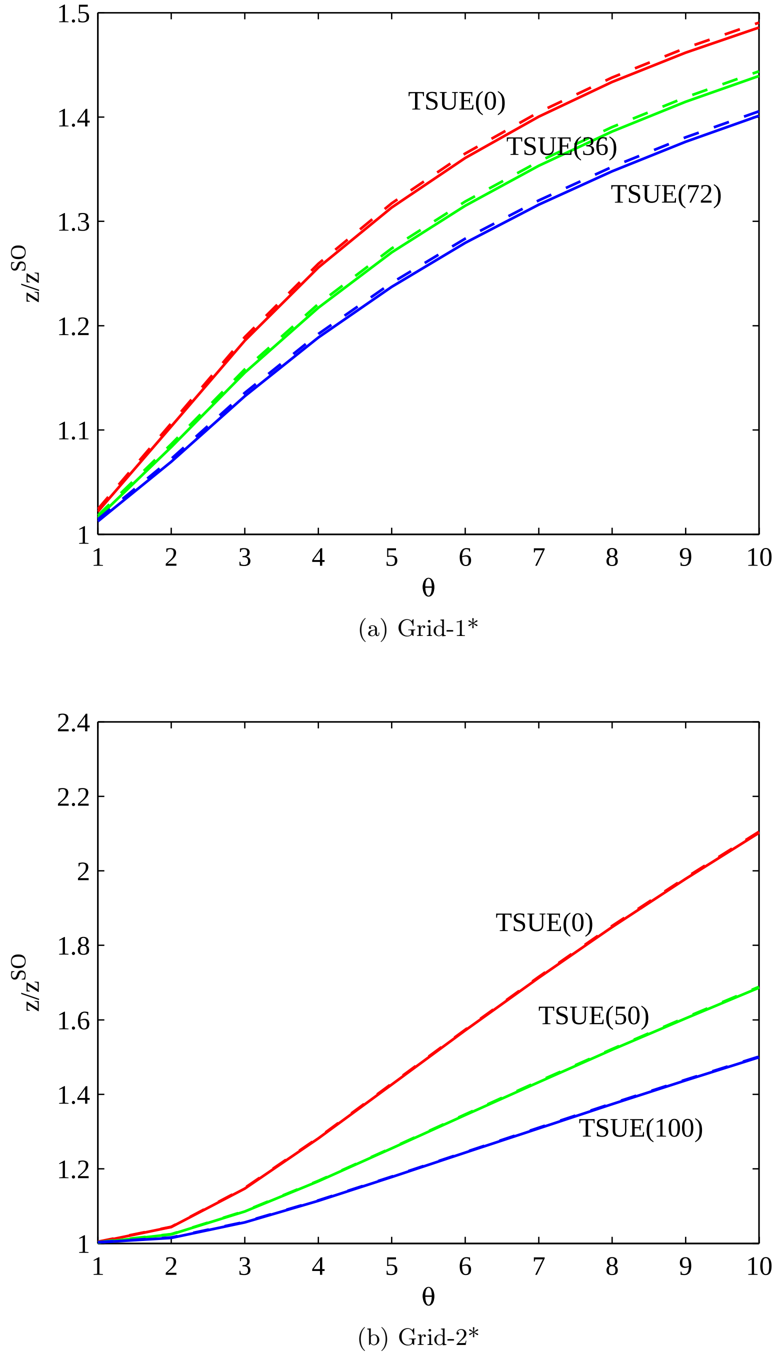
Table 4.5: The TSUE system cost for different values of λ and θ.

Recall that we have two definitions for the TSUE system cost. The first definition, $z^{\text{TSUE}}(\theta, \lambda)$, is the system cost of the expected path flow. The second definition, $\tilde{z}^{\text{TSUE}}(\theta, \lambda)$, is the expected system cost. In Figure 4.2, we plot $z^{\text{TSUE}}(\theta, \lambda)$ in solid lines and $\tilde{z}^{\text{TSUE}}(\theta, \lambda)$ in dashed lines.

From these results we have the following observations:

- For our choice of the values of λ, the difference between $z^{\text{TSUE}}(\theta, \lambda)$ and $\tilde{z}^{\text{TSUE}}(\theta, \lambda)$ is relatively small compared to the difference between $z^{\text{TSUE}}(\theta, \lambda)$ and $z^{\text{TSUE}}(\theta, 0)$. This is perhaps because these networks have small VM ratios (i.e., $\frac{\text{Var}(M_r)}{\mathbb{E}[M_r]}$), namely, the variance of travel times is relatively small compared to the mean travel times. As a result, the difference between $z^{\text{TSUE}}(\theta, \lambda)$ and $\tilde{z}^{\text{TSUE}}(\theta, \lambda)$, which is the extra term $\boldsymbol{e}^{\text{T}} \left(\boldsymbol{G} \circ \text{Cov} \left(\boldsymbol{H}^{\text{TSUE}}(\theta, \lambda) \right) \right) \boldsymbol{e}$ in Eq. (4.6), is small.

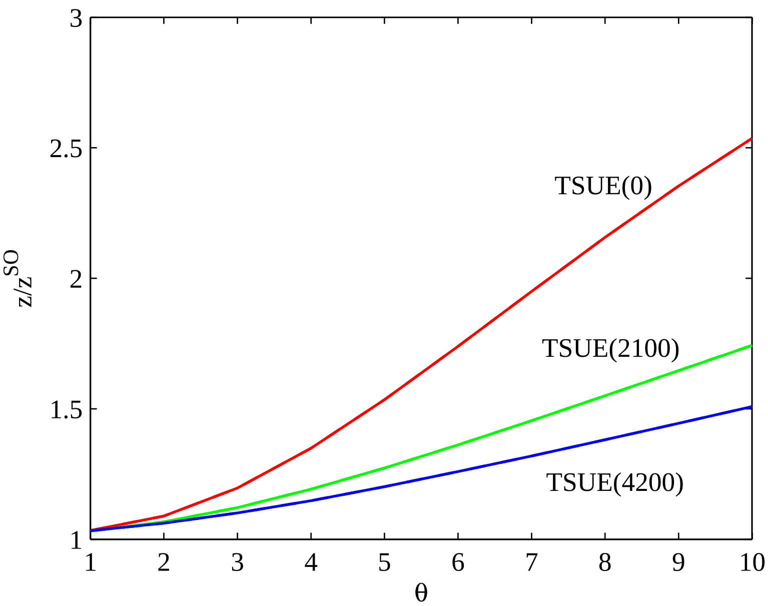
- For these networks, when λ increase, the TSUE system cost tends to decrease (see Table 4.5). This indicates that, when the travel time variance is a part of the travel cost, users' selfish behavior of minimizing their own travel cost seem to reduce the total system travel time. We are yet to test this effect on larger networks.



(a) Grid-1*

(b) Grid-2*

Figure 4.2: Numerical results on the TSUE system cost.



(c) Sioux Falls*

Figure 4.2: Numerical results on the TSUE system cost. (continued)

Chapter 5

An Inter-Period Dynamic Traffic Network Model

In this chapter we develop an inter-period dynamic learning traffic network model. Contrary to inter-period stationary models (namely, the equilibrium models discussed in the previous chapters), inter-period dynamic models assume that users update their travel decisions in each period to account for their most recent traveling experience. Hence travelers' decisions and traffic flows evolve over time.

The main purposes of studying inter-period dynamic models include the following

- Studying whether traffic models converge to their corresponding equilibria. As indicated in previous chapters, early researches of traffic network models take for granted that traffic networks converge to their equilibria in the long run. However this assumption is not well justified. Indeed, some inter-period dynamic models, including Cascetta (1989) and Nakayama et al. (1999), suggest that traffic models may not converge to traffic equilibria.

- Modeling the short-term inter-period evolution of a traffic network in the presence of a sudden change in the network. For example, this sudden change may be the opening of a new highway, the closing of a bridge, or the temporary increase of traffic demands due to a major event such as the FIFA World Cup. Inter-period dynamic models are necessary for studying the evolution of traffic network over time.

- Utilizing finer traffic network information to provide more accurate predictions of traffic networks (compared to the equilibrium models). Traditional traffic network performance data usually only includes estimation of average performance over a quarter or a year, and the data is not available on a daily basis. At this large time scale it may be reasonable to argue that the average traffic flows are decided by the balance between traffic supply and traffic demand, and the equilibrium models attempt to capture this balance. However, modern traffic network performance data is collected at a much finer time scale. Typically the data is available on a daily scale, an hourly scale, or even on a real-time basis. Traffic flows on these finer time scales are more susceptible to specific events and learning and updating mechanism of users, and dynamic traffic network models can outperform the equilibrium models by utilizing finer traffic network information.

However, inter-period dynamic models are to be distinguished from intra-period dynamic models. While the former assumes constant traffic flows within a time period, the latter considers traffic flows as time-dependent variables within a time period. These two thus differ in the major factors they intend to capture in the traffic networks. With a simplifying within period stationarity assumption, the inter-period dynamic models mainly capture the day-to-day learning and updating behavior of users. On the contrary, the intra-period dynamic models attempt to more closely examine the structure of time-dependent traffic network flows and the real-time interactions between travelers, traffic networks and traveler information systems.

In this chapter we focus on a simple network with one OD pair and two parallel links. In Section 5.1 we introduce the notations and assumptions of the model. The dynamic evolution of this model is presented in Section 5.2.

5.1 Model Formulation

Consider a network with single OD pair and two parallel links. The traffic demand is $q = 100$ veh/period. The travel time on each link is decided by its link flow-time function. For this model we consider the BPR function

$$\tau_i = \alpha_i(1 + \beta_i(\frac{f_i}{\gamma_i})^4), \tag{5.1}$$

where f_i is the link flow, α_i is the free flow travel time, β_i is the congestion factor and γ_i measures the link capacity. In the model, the parameters are given below

i	α_i	β_i	γ_i
1	1.0	1.0	60
2	1.0	1.0	40

Table 5.1: Flow-time function parameters

The link choice behavior of each user is characterized by a link choice probability. Users can have non-uniform link choice probabilities, and they update these probabilities over different time periods. In time period t, the probability that user u chooses link i is $P^t_{u,i}$. (We use capitalized letters to indicate random variables.) We impose that

$$P^t_{u,1} + P^t_{u,2} = 1.$$

Namely each user's link choice probabilities for the two links sum up to 1. We assume that users make their link choices independently, and their link choice decisions in period t are also independent from their previous decisions.

In time period t, let $F_{u,i}^t$ represent the link choice actually made by user h, defined as

$$F_{u,i}^t = \begin{cases} 1 & \text{if user } u \text{ chooses link } i \text{ in time period } t, \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, let F_i^t the traffic flow on link i. Then $F_i^t = \sum_u F_{u,i}^t$. It follows from our assumption that $F_{u,i}^t$ and F_i^t are random variables that the travel time on link i is also a random variable, denoted as $T_i^t(F_i^t) = \alpha_i(1 + \beta_i(\frac{F_i^t}{\gamma_i})^4)$.

We assume that users believe that the traffic network is inter-period stationary, namely, that their experienced travel time T_i^t is sampled from an unknown but time-period-invariant travel time distribution associated with link i. This assumption is common in the study of learning in games called fictitious play (see Brown 1951).

Under this assumption of stationary travel time distribution, users can estimate the mean and variance (denoted as $\mu_{u,i}^t$ and $v_{u,i}^t$) of the travel time distribution based their respective experiences in the first t time periods, i.e.,

$$\mu_{u,i}^t = \sum_{s=1}^{t} T_{u,i}^s F_{u,i}^s / \sum_{s=1}^{t} F_{u,i}^s,$$

$$v_{u,i}^t = \sum_{s=1}^{t} \left(T_{u,i}^s - \mu_{u,i}^s\right)^2 F_{u,i}^s / \sum_{s=1}^{t} F_{u,i}^s.$$

In this equation, the term $\sum_{s=1}^t F_{u,i}^s$ gives the number of periods that user u has chosen link i in the first t time periods. It should be noted that our model assumes users only know the travel time they actually experiences. The user doesn't know the travel time on the link that he doesn't travel on.

We assume that the users adopt the mean-variance approach in Chapter 3 to evaluate the link travel cost. Furthermore, assume that they have perception errors of the travel cost. As a result, the link travel costs are transformed into link choice probabilities through the following MNL model

$$P_{u,i}^{(t+1)} = \frac{\exp(-(\mu_{u,i}^t + \lambda v_{u,i}^t)/\theta)}{\exp(-(\mu_{u,1}^t + \lambda v_{u,1}^t)/\theta) + \exp(-(\mu_{u,2}^t + \lambda v_{u,2}^t)/\theta)},$$

where λ is the mean-variance trade-off factor, and θ is the dispersion level of the perception error.

5.2 Results of Numerical Simulations

In the numerical simulations we performed, we set $\theta = 0.5$ and $\lambda = 2$. The model starts on time period 1 with randomly chosen link choice probabilities $P_{u,1}^1$ and $P_{u,2}^1$. The probabilities are drawn from a uniform distribution on $[0.1, 0.9]$. We discard the very large and very small probabilities in order to make sure that users have reasonable chances of sampling on both links. We also assume users only start to update his link choice probabilities when both links have been sampled.

We run simulations with the above setting. The results are presented in Figure 5.1 and Figure 5.2.

Figure 5.1 plots the evolution of the link choice probability for link 1 ($P_{u,1}^t$) over time (t). For a given day, we plot the follow 4 values: the maximum link choice probability ($\max_u P_{u,1}^t$), the medium link choice probability ($\text{median}_u P_{u,1}^t$), the minimum link choice probability ($\min_u P_{u,1}^t$), and the realized link flow (F_1^t).

We observe that, over a long time, the link choice probabilities converge to a constant for all users. This justifies the assumption of uniform path choice probability across the traveler population in the TSUE model. The speed of convergence can be measured in terms of how fast the difference between the maximum and minimum link choice probabilities, i.e., $\left(\max_u P_{u,1}^t - \min_u P_{u,1}^t\right)$, converges to 0. As shown in Figure 5.2, the convergence speed is approximately $t^{-0.5}$.

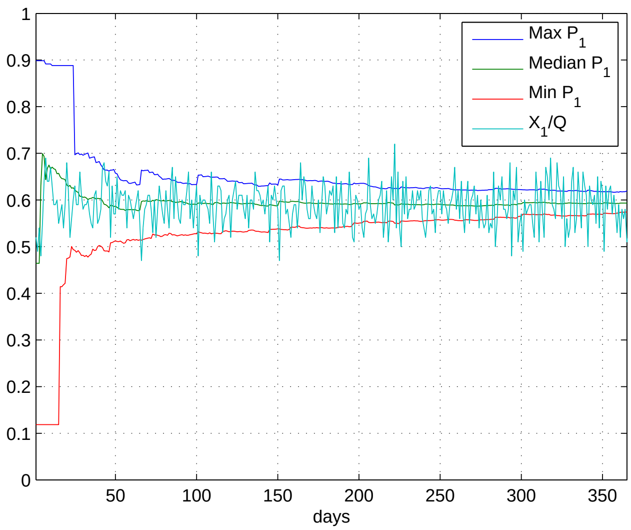


Figure 5.1: Evolution of the link choice probability over time. This figure shows the maximum, median and minimum link choice probabilities of all users on link 1. It also shows the actual link flow on link 1.

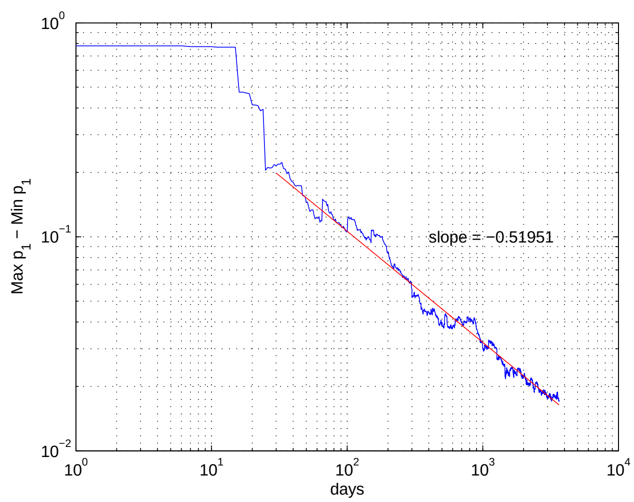


Figure 5.2: The convergence of link choice probabilities over time.

Chapter 6

Conclusions

This thesis focuses on modeling travel time uncertainty in traffic network models. Travel time uncertainty plays an important role in the transportation system. Accounting for travel time uncertainty in traffic network models is a field of active academic research.

Models that incorporate travel time uncertainty must specify two fundamental mechanisms: the mechanism through which travel time uncertainty is generated and the mechanism through which travel time uncertainty influences users' behavior. Previous studies have adopted different approaches to defining these two mechanisms; most existing stochastic traffic equilibrium models feature a sophisticated approach to one of the two mechanisms, but nevertheless adopt a simplistic view on the other mechanism. This thesis proposed a new stochastic traffic equilibrium model, namely the Truly Stochastic User Equilibrium (TSUE), that combines both mechanisms. We showed that there exists a solution to this new model and the conditions under which the solution is unique. We compared this new model with the existing traffic equilibrium models in the literature. Through numerical results on test networks we observed that the conventional stochastic user equilibrium model tends to overestimate traffic flows on the links with larger travel time variances.

To benchmark the various traffic network equilibrium models, we investigate the total system cost, namely the total travel time of all users, of these models. We proved upper bounds of the system cost for the TSUE model and presented numerical results on test networks. As a special case, we also derived upper bounds for the SUE system cost. These bounds improved the previous bounds obtained by researchers. Through numerical study, we discovered that the total system cost of the network is reduced when travelers minimize the variance of their travel time in addition to the mean travel time.

This thesis contributes to the traffic network modeling literature and our general understanding of modeling uncertainty in large and complicated systems.

Appendix A

Summary of Notations

This appendix provides a summary of the notations used in this thesis.

In general, we use lower case letters for deterministic variables, and upper case letters for random variables. For example, while f refers to the deterministic traffic flow, as used in the DUE model, F refers to the random variable of traffic flow, as used in the TSUE model.

We use bold fonts to indicate vector and matrices. Many equations used in this thesis can be written in either a vector form or componentwise. This thesis frequently uses the vector form to simplify notations. As a special rule, when the logarithm function applies to a vector, it means applying the function component-wise and the output variable is a vector of the same size as the input variable. In particular, $\boldsymbol{y} = \log \frac{\boldsymbol{h}}{\boldsymbol{q}}$ should be understood as $y_r = \log \frac{h_r}{q^k}, \forall k \in K, r \in R^k$.

Letter p stands for probability. Letter r stands for paths (routes).

For a set X, $|X|$ represents the cardinality of the set.

Notations for the traffic network

A set of all nodes, with cardinality $|A|$.

L set of all links (arcs), with cardinality $|L|$.

K set of all OD pairs (trips, movements), with cardinality $|K|$.

R set of all paths (routes), with cardinality $|R|$.

R^k set of all paths available to OD pair $k \in K$, with cardinality $|R^k|$.

$\mathbf{\Delta}$ link-path incidence matrix, with dimension $|L| \times |R|$

$\mathbf{\Delta}^k$ link-path incidence matrix restricted to only paths for OD pair $k \in K$, with dimension $|L| \times |R^k|$

$\mathbf{\Gamma}$ OD-pair-path incidence matrix, with dimension $|K| \times |R|$

Notations for traffic demand and traffic supply

$\boldsymbol{q}$ vector of travel demand flows, with dimension $|K| \times 1$

$\boldsymbol{f}$ vector of deterministic link flows, with dimension $|L| \times 1$

$\boldsymbol{F}$ vector of random link flows, with dimension $|L| \times 1$

$\boldsymbol{t}$ vector of deterministic link travel times, with dimension $|L| \times 1$

$\boldsymbol{T}$ vector of random link travel times, with dimension $|L| \times 1$

$\boldsymbol{\mu}$ vector of deterministic path travel times, with dimension $|R| \times 1$

$\boldsymbol{M}$ vector of random path travel times, with dimension $|R| \times 1$

$\boldsymbol{c}$ vector of deterministic path costs, with dimension $|R| \times 1$

$\boldsymbol{p}$ vector of path choice probabilities, with dimension $|R| \times 1$

$\boldsymbol{h}$ vector of deterministic path flows, with dimension $|R| \times 1$

$\boldsymbol{H}$	vector of random path flows, with dimension $\lvert R\rvert \times 1$
S_h	set of feasible path flow vectors, $S_h = \left\{\boldsymbol{h} \in \mathbb{R}^{\lvert R\rvert} \geqslant 0 : \boldsymbol{\Gamma h} = \boldsymbol{q}\right\}$
S_f	set of feasible link flow vectors, $S_f = \{\boldsymbol{f} = \boldsymbol{\Delta h} : \boldsymbol{h} \in S_h\}$
S_p	set of feasible path choice probability vectors, $S_p = \left\{\boldsymbol{p} \in \mathbb{R}^{\lvert R\rvert} \geqslant 0 : \boldsymbol{\Gamma p} = \boldsymbol{1}^{\lvert K\rvert}\right\}$
θ	dispersion level parameter, used in the multinomial logit path choice model
λ	mean-variance trade-off factor, used in the TSUE cost function
$\boldsymbol{\mu}^k, \boldsymbol{c}^k, \boldsymbol{p}^k, \boldsymbol{h}^k$	These variables respectively represent the vectors of deterministic path travel times, path costs, path choice probabilities, and path flows, but restricted to a single OD pair $k \in K$
$\boldsymbol{M}^k, \boldsymbol{H}^k$	These variables respectively represent the vectors of random path travel times and random path flows, but restricted to a single OD pair $k \in K$

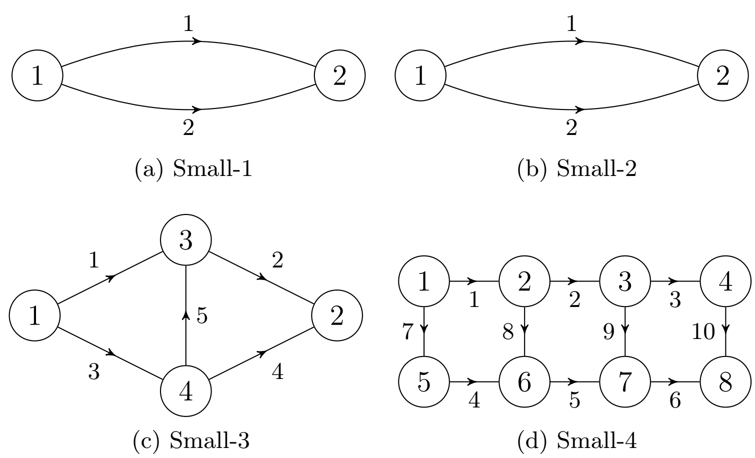
	Link space		Path space	
	Deterministic	Random	Deterministic	Random
Traffic flow	$\boldsymbol{f}$	$\boldsymbol{F}$	$\boldsymbol{h}$	$\boldsymbol{H}$
Travel time	$\boldsymbol{t}$	$\boldsymbol{T}$	$\boldsymbol{\mu}$	$\boldsymbol{M}$
Travel cost			$\boldsymbol{c}$	
Choice probability			$\boldsymbol{p}$	

Table A.1: Summary of variables

Appendix B

Illustrations of Test Networks

B.1 Small-scale networks



(a) Small-1

(b) Small-2

(c) Small-3

(d) Small-4

Figure B.1: Illustration of the small-scale test networks.

Link	1	2
a	0.1	0.0
b	0	10

(a) Small-1

Link	1	2
a	0.20	0.12
b	10	20

(b) Small-2

Link	1	2	3	4	5
a	0.1	0.1	0.2	0.1	0.1
b	5	5	4	16	10

(c) Small-3

Link	1	2	3	4	5	6	7	8	9	10
a	0.7	0.2	0.6	0.3	0.0	0.8	0.7	0.1	0.8	0.7
b	35	20	10	15	40	10	10	35	20	10

(d) Small-4

Table B.1: Link flow-time function parameters for small-scale networks.

No.	1
Origin	1
Destination	2
Flow	100

(a) Small-1

No.	1
Origin	1
Destination	2
Flow	100

(b) Small-2

No.	1
Origin	1
Destination	2
Flow	100

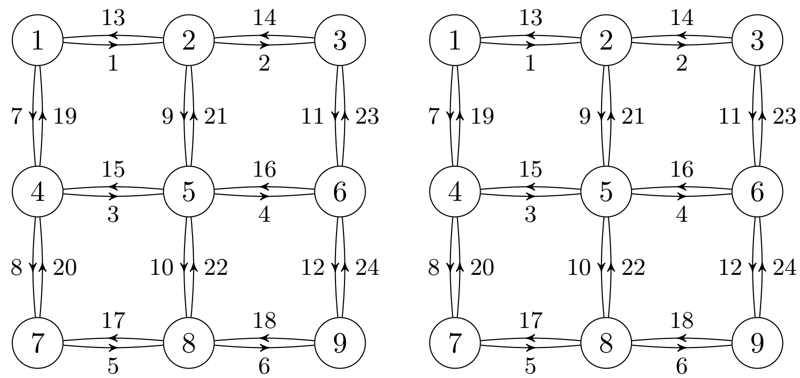
(c) Small-3

No.	1	2	3	4	5
Origin	1	1	1	2	3
Destination	6	7	8	8	8
Flow	50	100	100	50	100

(d) Small-4

Table B.2: OD pair flows for small-scale networks.

B.2 Medium-scale networks



(a) Grid-1 (b) Grid-2

Figure B.2: Illustration of the medium-scale test networks.

Link	1	2	3	4	5	6	7	8	9	10	11	12
α	2.0	2.0	1.0	1.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
β	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
γ	100	100	100	100	100	100	100	100	100	100	100	100
Link	13	14	15	16	17	18	19	20	21	22	23	24
α	2.0	2.0	1.0	1.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
β	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
γ	100	100	100	100	100	100	100	100	100	100	100	100

(a) Grid-1

Link	1	2	3	4	5	6	7	8	9	10	11	12
α	2.0	2.0	1.0	1.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
β	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
γ	100	100	100	100	100	100	100	100	100	100	100	100
Link	13	14	15	16	17	18	19	20	21	22	23	24
α	2.0	2.0	1.0	1.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
β	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
γ	100	100	100	100	100	100	100	100	100	100	100	100

(b) Grid-2

Table B.3: Link flow-time function parameters for medium-scale networks.

No.	1	2	3	4	5	6	7
Origin	1	1	2	3	4	6	6
Destination	5	9	8	7	9	7	8
Flow	20	150	80	100	50	50	60

(a) Grid-1

Destination	1	2	3	4	5	6	7	8	9
Origin	OD pair flow								
1		50	40	70	80	70	60	20	20
2	30		70	50	10	60	50	30	40
3	70	80		10	70	20	20	30	70
4	40	30	80		70	50	10	80	10
5	40	70	10	20		40	20	40	70
6	80	10	70	60	60		10	10	40
7	10	20	20	20	40	50		80	70
8	70	80	50	30	10	50	40		40
9	30	40	60	70	50	20	50	50	

(b) Grid-2

Table B.4: OD pair flows for medium-scale networks.

B.3 Large-scale networks

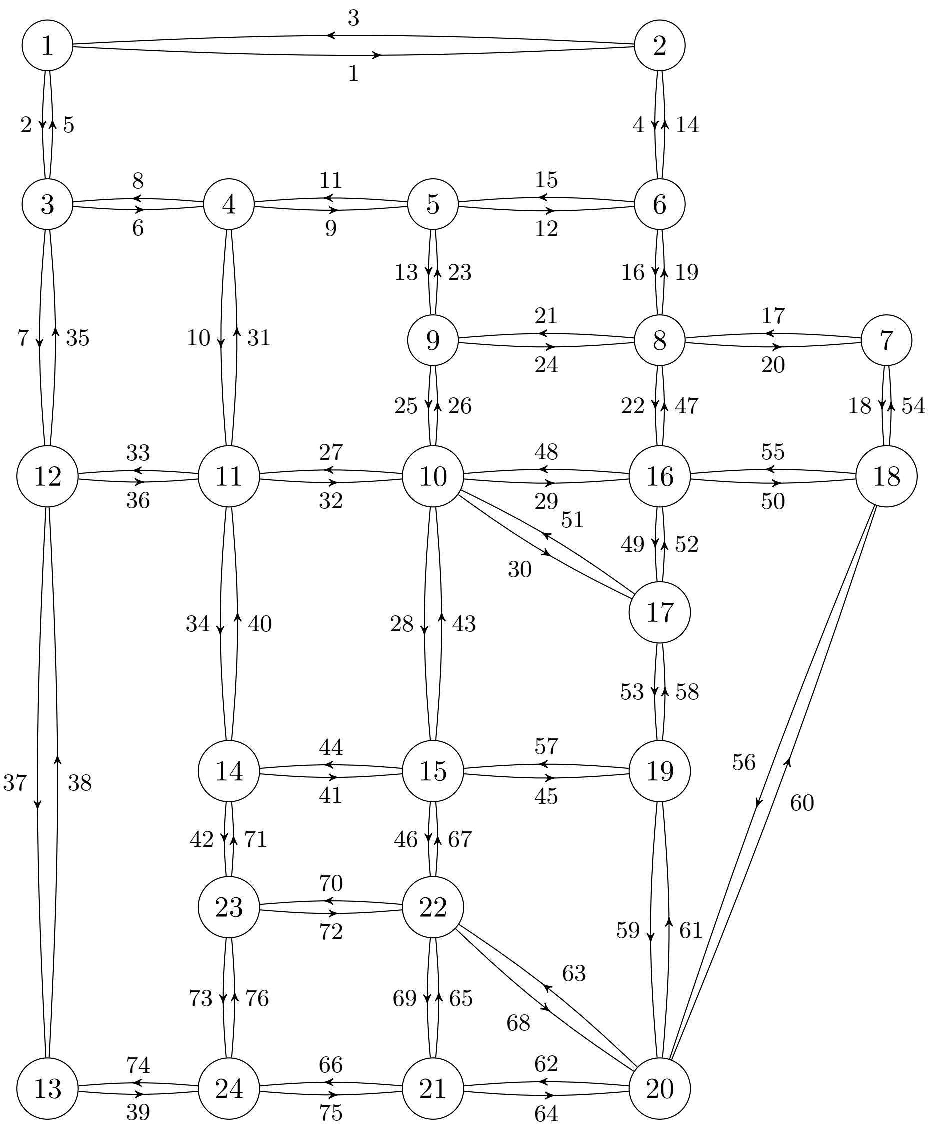


Figure B.3: Illustration of the Sioux Falls network.

Link	1	2	3	4	5	6	7	8
α	6.00	4.00	6.00	5.00	4.00	4.00	4.00	4.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	25900	23403	25900	4958	23403	17111	23403	17111
Link	9	10	11	12	13	14	15	16
α	2.00	6.00	2.00	4.00	5.00	5.00	4.00	2.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	17783	4909	17783	4948	10000	4958	4948	4899
Link	17	18	19	20	21	22	23	24
α	3.00	2.00	2.00	3.00	10.00	5.00	5.00	10.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	7842	23403	4899	7842	5050	5046	10000	5050
Link	25	26	27	28	29	30	31	32
α	3.00	3.00	5.00	6.00	4.00	8.00	6.00	5.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	13916	13916	10000	13512	4855	4994	4909	10000
Link	33	34	35	36	37	38	39	40
α	6.00	4.00	4.00	6.00	3.00	3.00	4.00	4.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	4909	4877	23403	4909	25900	25900	5091	4877
Link	41	42	43	44	45	46	47	48
α	5.00	4.00	6.00	5.00	3.00	3.00	5.00	4.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	5128	4925	13512	5128	14565	9599	5046	4855
Link	49	50	51	52	53	54	55	56
α	2.00	3.00	8.00	2.00	2.00	2.00	3.00	4.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	5230	19680	4994	5230	4824	23403	19680	23403
Link	57	58	59	60	61	62	63	64
α	3.00	2.00	4.00	4.00	4.00	6.00	5.00	6.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	14565	4824	5003	23403	5003	5060	5076	5060
Link	65	66	67	68	69	70	71	72
α	2.00	3.00	3.00	5.00	2.00	4.00	4.00	4.00
β	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
γ	5230	4885	9599	5076	5230	5000	4925	5000
Link	73	74	75	76				
α	2.00	4.00	3.00	2.00				
β	0.15	0.15	0.15	0.15				
γ	5079	5091	4885	5079				

Table B.5: Link flow-time function parameters for the Sioux Falls network.

Destination	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Origin	OD pair flow/100																							
1		1	1	5	2	3	5	8	5	13	5	2	5	3	5	5	4	1	3	3	1	4	3	1
2	1		1	2	1	4	2	4	2	6	2	1	3	1	1	4	2		1	1		1		
3	1	1		2	1	3	1	2	1	3	3	2	1	1	1	2	1					1	1	
4	5	2	2		5	4	4	7	7	12	14	6	6	5	5	8	5	1	2	3	2	4	5	2
5	2	1	1	5		2	2	5	8	10	5	2	2	1	2	5	2		1	1	1	2	1	
6	3	4	3	4	2		4	8	4	8	4	2	2	1	2	9	5	1	2	3	1	2	1	1
7	5	2	1	4	2	4		10	6	19	5	7	4	2	5	14	10	2	4	5	2	5	2	1
8	8	4	2	7	5	8	10		8	16	8	6	6	4	6	22	14	3	7	9	4	5	3	2
9	5	2	1	7	8	4	6	8		28	14	6	6	6	9	14	9	2	4	6	3	7	5	2
10	13	6	3	12	10	8	19	16	28		40	20	19	21	40	44	39	7	18	25	12	26	18	8
11	5	2	3	15	5	4	5	8	14	39		14	10	16	14	14	10	1	4	6	4	11	13	6
12	2	1	2	6	2	2	7	6	6	20	14		13	7	7	7	6	2	3	4	3	7	7	5
13	5	3	1	6	2	2	4	6	6	19	10	13		6	7	6	5	1	3	6	6	13	8	8
14	3	1	1	5	1	1	2	4	6	21	16	7	6		13	7	7	1	3	5	4	12	11	4
15	5	1	1	5	2	2	5	6	10	40	14	7	7	13		12	15	2	8	11	8	26	10	4
16	5	4	2	8	5	9	14	22	14	44	14	7	6	7	12		28	5	13	16	6	12	5	3
17	4	2	1	5	2	5	10	14	9	39	10	6	5	7	15	28		6	17	17	6	17	6	3
18	1			1		1	2	3	2	7	2	2	1	1	2	5	6		3	4	1	3	1	
19	3	1		2	1	2	4	7	4	18	4	3	3	3	8	13	17	3		12	4	12	3	1
20	3	1		3	1	3	5	9	6	25	6	5	6	5	11	16	17	4	12		12	24	7	4
21	1			2	1	1	2	4	3	12	4	3	6	4	8	6	6	1	4	12		18	7	5
22	4	1	1	4	2	2	5	5	7	26	11	7	13	12	26	12	17	3	12	24	18		21	11
23	3		1	5	1	1	2	3	5	18	13	7	8	11	10	5	6	1	3	7	7	21		7
24	1			2		1	1	2	2	8	6	5	7	4	4	3	3		1	4	5	11	7	

Table B.6: OD pair flows for the Sioux Falls network.

Bibliography

Avineri, E., and Prashker, J. N. Sensitivity to travel time variability: Travelers' learning perspective. *Transportation Research Part C: Emerging Technologies*, 13 (2):157–183, Apr. 2005.

Barberis, N., and Thaler, R. A survey of behavioral finance. In *Handbook of the Economics of Finance*, volume 1(2), pages 1053–1128. Elsevier, 2003.

Bekhor, S., and Toledo, T. Investigating path-based solution algorithms to the stochastic user equilibrium problem. *Transportation Research Part B: Methodological*, 39(3):279–295, Mar. 2005.

Bell, M. G. H., and Iida, Y. *Transportation Network Analysis.* Wiley, 1997.

Ben-Akiva, M. E., and Lerman, S. R. *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press, 1985.

Ben-Tal, A., and Nemirovski, A. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, Nov. 1998.

Ben-Tal, A., and Nemirovski, A. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88(3):411–424, 2000.

Bertsekas, D. P. *Nonlinear Programming.* Athena Scientific, 1995.

Bertsimas, D., and Sim, M. Robust discrete optimization and network flows. *Mathematical Programming*, 98(1):49–71, 2003.

Brilon, W., Geistefeldt, J., and Regler, M. Reliability of freeway traffic flow: a stochastic concept of capacity. In *Transportation and traffic theory: flow, dynamics and human interaction: proceedings of the 16th International Symposium on Transportation and Traffic Theory, University of Maryland, College Park, Maryland, 19-21 July 2005*, page 125, 2005.

Brown, G. W. Iterative solution of games by fictitious play. In *Activity Analysis of Production and Allocation*, volume 13, page 374–376. New York: Wiley, 1951.

Cambridge Systematics, National Research Council, American Association of State Highway and Transportation Officials, and National Cooperative Highway Research Program. *A guidebook for performance-based transportation planning.* Transportation Research Board, 2000.

Cascetta, E. A stochastic process approach to the analysis of temporal dynamics in transportation networks. *Transportation Research Part B: Methodological*, 23(1): 1–17, Feb. 1989.

Cascetta, E. *Transportation Systems Engineering.* Springer, 2001.

Chau, C. K., and Sim, K. M. The price of anarchy for non-atomic congestion games with symmetric cost maps and elastic demands. *Operations Research Letters*, 31 (5):327–334, Sept. 2003.

Clark, S., and Watling, D. Modelling network travel time reliability under stochastic demand. *Transportation Research Part B: Methodological*, 39(2):119–140, 2005.

Connors, R. D., and Sumalee, A. A network equilibrium model with travellers' perception of stochastic travel times. *Transportation Research Part B: Methodological*, 43(6):614–624, July 2009.

Correa, J. R., Schulz, A. S., and Stier-Moses, N. E. Selfish routing in capacitated networks. *Mathematics of Operations Research*, 29(4):961–976, Nov. 2004.

Correa, J. R., Schulz, A. S., and Stier-Moses, N. E. A geometric approach to the price of anarchy in nonatomic congestion games. *Games and Economic Behavior*, 64(2):457–469, Nov. 2008.

Correa, J. R., and Stier-Moses, N. E. Wardrop equilibria. In *Encyclopedia of Operations Research and Management Science (to be published)*. 2010.

Daganzo, C. F., and Sheffi, Y. On stochastic models of traffic assignment. *Transportation Science*, 11(3):253, 1977.

Damberg, O., Lundgren, J. T., and Patriksson, M. An algorithm for the stochastic user equilibrium problem. *Transportation Research Part B: Methodological*, 30(2): 115–131, Apr. 1996.

de Palma, A., Ben-Akiva, M., Brownstone, D., Holt, C., Magnac, T., McFadden, D., Moffatt, P., Picard, N., Train, K., Wakker, P., and Walker, J. Risk, uncertainty and discrete choice models. *Marketing Letters*, 19(3):269–285, Dec. 2008.

de Palma, A., and Picard, N. Route choice decision under travel time uncertainty. *Transportation Research Part A: Policy and Practice*, 39(4):295–324, May 2005.

El Ghaoui, L., Oustry, F., and Lebret, H. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9:33–52, 1998.

Ferris, M. C., and Munson, T. S. Interfaces to PATH 3.0: Design, implementation and usage. *Computational Optimization and Applications*, 12(1):207–227, Jan. 1999.

Florian, M., and Hearn, D. Network equilibrium models and algorithms. In *Network Routing*, volume 8 of *Handbooks in Operations Research and Management Science*, page 485–550. Elsevier, 1995.

Friesz, T. L. *Dynamic Optimization and Differential Games*. Springer, 2010.

Gill, P. E., Murray, W., and Wright, M. H. *Practical Optimization*. Academic Press, Feb. 1982.

Guo, X. L., and Yang, H. The price of anarchy of stochastic user equilibrium in traffic networks. In *Proceedings of the 10th International Conference of Hong Kong Society for Transportation Studies (HKSTS). Hong Kong: HKSTS Ltd*, page 63–72, 2005.

Hallenbeck, M. E., Ishimaru, J., and Nee, J. *Measurement of Recurring Versus Non-Recurring Congestion*. Washington State Transportation Center (TRAC), 2003.

Jackson, W. B., and Jucker, J. V. An empirical study of travel time variability and travel choice behavior. *Transportation Science*, 16(4):460–475, Nov. 1982.

Kahneman, D., and Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, Mar. 1979.

Knight, F. H. Some fallacies in the interpretation of social cost. *The Quarterly Journal of Economics*, 38(4):582–606, Aug. 1924.

Koutsoupias, E., and Papadimitriou, C. Worst-Case equilibria. In *Proceedings of* *the 16th Annual Symposium on Theoretical Aspects of Computer Science Trier, Germany, March 4–6, 1999*, pages 404–413. 1999.

Liu, H. X., Recker, W., and Chen, A. Travel time reliability. In *Assessing the Benefits and Costs of ITS*, pages 241–261. 2004.

Maher, M. Algorithms for logit-based stochastic user equilibrium assignment. *Transportation Research Part B: Methodological*, 32(8):539–549, Nov. 1998.

Maher, M., Stewart, K., and Rosa, A. Stochastic social optimum traffic assignment. *Transportation Research Part B: Methodological*, 39(8):753–767, Sept. 2005.

Marchal, F., and de Palma, A. Measurement of uncertainty costs with dynamic traffic simulations. *Transportation Research Record: Journal of the Transportation Research Board*, 2085(-1):67–75, Dec. 2008.

McFadden, D. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 8:105–142, 1974.

Mirchandani, P., and Soroush, H. Generalized traffic equilibrium with probabilistic travel times and perceptions. *Transportation Science*, 21(3):133–152, Aug. 1987.

Nagurney, A. *Network Economics: A Variational Inequality Approach.* Springer, 1999.

Nakayama, S., Kitamura, R., and Fujii, S. Drivers' learning and network behavior: Dynamic analysis of the Driver-Network system as a complex system. *Transportation Research Record: Journal of the Transportation Research Board*, 1676(-1): 30–36, Jan. 1999.

Nam, D., Park, D., and Khamkongkhun, A. Estimation of value of travel time reliability. *Journal of Advanced Transportation*, 39(1):39–61, 2005.

Nikolova, E., Brand, M., and Karger, D. R. Optimal route planning under uncertainty. In *Proceedings of International Conference on Automated Planning and Scheduling*, 2006.

Nocedal, J., and Wright, S. J. *Numerical Optimization*. 2006.

Noland, R. B., and Polak, J. W. Travel time variability: a review of theoretical and empirical issues. *Transport Reviews: A Transnational Transdisciplinary Journal*, 22(1):39, 2002.

Noland, R. B., Small, K. A., Koskenoja, P. M., and Chu, X. Simulating travel reliability. *Regional Science and Urban Economics*, 28(5):535–564, Sept. 1998.

Ordonez, F., and Stier-Moses, N. Robust wardrop equilibrium. In *Network Control and Optimization*, pages 247–256. 2007.

Peeta, S., and Ziliaskopoulos, A. Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics*, 1(3):233–265, 2001.

Perakis, G. The "price of anarchy" under nonlinear and asymmetric costs. *Mathematics of Operations Research*, 32(3):614–628, Aug. 2007.

Pigou, A. C. *The Economics of Welfare*. 1920.

Ramming, M. S. *Network Knowledge and Route Choice*. Thesis, Massachusetts Institute of Technology, 2002.

Rockafellar, T. Coherent approaches to risk in optimization under uncertainty. *Tutorials in Operations Research: Or Tools and Applications: Glimpses of Future Technologies*, 2007.

Roughgarden, T. *Selfish Routing*. Thesis, Cornell University, 2002.

Roughgarden, T., and Tardos, Éva. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002.

Rudin, W. *Principles of Mathematical Analysis*. McGraw-Hill, 1986.

Schadschneider, A. Statistical physics of traffic flow. *Physica A: Statistical Mechanics and its Applications*, 285(1-2):101–120, Sept. 2000.

Schrank, D., and Lomax, T. *2009 Urban Mobility Report*. Texas Transportation Institute, The Texas A&M University System, 2009.

Senna, L. A. D. S. The influence of travel time variability on the value of time. *Transportation*, 21(2):203–228, May 1994.

Sheffi, Y. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, 1984.

Sheffi, Y., and Powell, W. B. An algorithm for the equilibrium assignment problem with random link times. *Networks*, 12(2):191–207, 1982.

Skabardonis, A., Varaiya, P., and Petty, K. Measuring recurrent and nonrecurrent traffic congestion. *Transportation Research Record: Journal of the Transportation Research Board*, 1856(-1):118–124, Jan. 2003.

Small, K. A. *Valuation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation*. Transportation Research Board, Jan. 1999.

Soyster, A. L. Convex programming with Set-Inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157, Oct. 1973.

Sterman, J., and Sterman, J. D. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill/Irwin, Feb. 2000.

Tversky, A., and Kahneman, D. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, Oct. 1992.

U.S. Department of Commerce. *Traffic Assignment Manual for Application with a Large, High Speed Computer.* U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division, Washington, 1964.

U.S. Department of Transportation. *Making It There on Time, All the Time.* U.S. Department of Transportation, Federal Highway Administration, Brochure, 2007.

U.S. Federal Highway Administration. *Making It There on Time, All the Time.* U.S. Department of Transportation, Federal Highway Administration, 2007.

Wardrop, J. G. Some theoretical aspects of road research. *Institute of Civil Engineers Proceedings: Engineering Divisions*, 1(3):325–362, 1952.

Watling, D. A second order stochastic network equilibrium model, I: theoretical foundation. *Transportation Science*, 36(2):149–166, May 2002.

Zeid, M. A. *Measuring and Modeling Activity and Travel Well-Being.* Thesis, Massachusetts Institute of Technology, 2009.