# Studies on
# Mathematical Models of Traffic Equilibria

RHODA PADUA AGDEPPA

# Studies on Mathematical Models of Traffic Equilibria

**RHODA PADUA AGDEPPA**

A DISSERTATION
Submitted in partial fulfillment of
the requirements for the degree of
**DOCTOR OF INFORMATICS**

(Applied Mathematics and Physics)

**Kyoto University**
**Kyoto, Japan**
DECEMBER 2008

*to HONEY*

*... the wind beneath my wings ...*

# Preface

Traffic assignment is the process of allocating a given set of traffic movements to a specific traffic system. The purposes of traffic assignment include the assigning of estimated future trips to the existing traffic system in order to assess the deficiencies of the system, to evaluate the effects of possible improvements in the traffic system and to test possible alternate routes for the traffic system.

One of the most popular principles used in the study of traffic assignment models is the **Wardrop user equilibrium principle** which states that *users of the traffic network will choose the route having the minimum cost between each OD pair, and through this process, the routes that are used will have equal costs; moreover, routes with costs higher than the minimum will have no flow.* Moreover, the traffic equilibrium problem (TEP) is to find a vector pair of route flows and minimum route costs such that the Wardrop user equilibrium principle, together with the nonnegativity condition imposed on the travel demand function, are satisfied.

The TEP has been studied for many decades. Various formulations have been proposed under various assumptions. Earlier TEP formulations made use of assumptions which are found to be unnatural or unrealistic (e.g., that the travel costs are independent of the link flows) in order to obtain TEP models which are easy to analyze. Most of the existing TEP formulations assume that route costs are additive, that is, *the route costs are simply the sum of the link costs for all the links on the route being considered.* Another assumption used in most TEP models is that every traveler has a complete and accurate information about the characteristics of the traffic network and all travelers have the same route cost perception and travel behavior.

However, due to the rapid technological and economic advancement seen all over the world, enormous changes in the road traffic conditions have occurred. Thus the additivity and certainty assumptions on the route costs is no longer appropriate. Moreover, different individuals may have different travel behavior and such travel behavior may be affected by the different time or weather of the day.

Hence, there is indeed a need to include nonadditivity and uncertainty in the route cost function formulation in order to present a more realistic view of the traffic system. Unfortunately, formulating the TEP with nonadditive route costs results in an increased difficulty in the analysis and computation of its equilibrium solution which is usually done by formulating the TEP into the equivalent mixed complementarity problem (MCP). The TEP with additive costs may be formulated as a monotone MCP having a unique solution. However, an MCP derived from the TEP with nonadditive costs does not immediately possess monotonicity unless restrictive assumptions are made or a certain reformulation is introduced. Moreover, the TEP under uncertainty may not have a solution in general.

It is the aim of this dissertation to present a more realistic TEP model which is solvable using existing solution methods. In Chapter 3, we consider the TEP with nonadditive route costs and propose an MCP formulation of this nonadditive TEP model which is solvable. We then apply this MCP reformulation of the nonadditive TEP model to the road pricing problem and show that the resulting reformulation is solvable using an existing solution method. This is done in Chapter 4. Moreover, in Chapter 5, we propose models of the TEP with uncertainty which give a reasonable equilibrium for the TEP. Moreover, we show that these models can be converted into convex programming problems under some reasonable conditions, and hence we can obtain an equilibrium solution of the model using the existing optimization solver.

Kyoto, Japan RHODA PADUA AGDEPPA

December 2008

# Acknowledgments

First and foremost, I would like to express my sincerest gratitude to my supervisors, Prof. Masao Fukushima and Prof. Nobuo Yamashita of Kyoto University, for without their professional guidance and support, this research would not have been realized. I especially thank them for giving me the chance to pursue doctoral studies and do research under their supervision. I would also like to thank them for patiently sharing their expertise with me and for untiringly providing me with helpful inputs during the entire process of this study. Furthermore, I am grateful to them for carefully reading my manuscript and for giving valuable comments and suggestions in order to make this a better one.

I would also like to thank Prof. Hiroshi Nagamochi for his careful review of my dissertation to improve its quality.

I am greatly indebted to the Ministry of Education, Science, Sports and Culture of Japan, for without its financial support through the Monbukagakusho Scholarship, pursuing graduate studies at Kyoto University would not have been possible.

I would also like to thank Dr. Shunsuke Hayashi, Ms. Fumie Yagura, Rocs, Mendee, Theang, and all the past and present members of Fukushima Laboratory for their kindness and support which made my stay at the laboratory and the University easier.

I also thank my KXU and KAPS family for their encouragement, prayers and support. I especially thank my KAPSpamilya *Nicolle*, *Gennie*, *Jay Arre*, *Lenee*, *Eloi*, *Laarni*, *Mike*, *Jong*, *Vince*, *Ninang Madam Casilda* and *Ninong Pastor Dan*, for providing me with food, shelter, laughter and shoulders to cry on during the last phase of my thesis writing.

I would also like to thank all my *beautiful friends*, especially *Grat* and *Maruy*, who have continuously prayed for me and who have always been there for me, even only in cyberspace.

I am also very thankful to ***DocNav***†, who is so lovingly remembered yet so sadly missed, for having encouraged me to pursue my dreams and for being instrumental to my coming to Japan.

I am also very grateful to Mama Jane, Timoy and the whole Padua, Agdeppa and Namoco family, for their continued love, prayers, support and understanding.

I share the success of this dissertation with Concon, my constant source of strength, love and understanding. I thank him for always believing in me and for always reminding me to keep going even when the going gets tough.

Finally, I would like to give back all the glory, honor and praise to the Lord my God, for without Him, all these would not have been possible.

***Odang***

# List of Tables

# List of Figures

# Notations

| | |
|---|---|
| $\mathcal{G} = (\mathcal{A},\mathcal{N})$ | network under consideration |
| $\mathcal{A}$ | the set of links (with cardinality $n_{\mathcal{A}}$) |
| $\mathcal{N}$ | the set of nodes (with cardinality $n_{\mathcal{N}}$) |
| W | the set of origin-destination (OD) pairs w in $\mathcal{G}$ (with cardinality $n_{\mathrm{W}}$) |
| $R_{\mathrm{w}}$ | the set of routes $r$ connecting the OD pair w $\in$ W |
| $R$ | the set of all routes (with cardinality $n_R$) |
| $f_a$ | the flow on link $a \in \mathcal{A}$ |
| $f$ | the vector of link flows |
| $F_r$ | the flow on route $r \in R$ |
| $F$ | the vector of route flows |
| $t_a$ | the travel time on link $a \in \mathcal{A}$ |
| $C_r$ | the cost experienced by a person using route $r$ |
| $u_{\mathrm{w}}$ | the minimal route cost for the OD pair w |
| $u$ | the vector with components $u_{\mathrm{w}}$ |
| $D_{\mathrm{w}}$ | the demand associated with each OD pair w |
| $\Delta = (\delta_{ar})$ | the link-route incidence matrix whose elements are $\delta_{ar}$ where $\delta_{ar}$ is 1 if route $r$ passes through link $a$ and 0 otherwise |
| $\Gamma = (\Gamma_{r\mathrm{w}})$ | the route-OD pair incidence matrix whose elements are $\Gamma_{r\mathrm{w}}$ where $\Gamma_{r\mathrm{w}}$ is 1 if $r \in R_{\mathrm{w}}$ and 0 otherwise |

# Table of Contents

# Chapter 1

# Introduction

The movement of people and goods between different places is one of the major activities that can be observed in a highly urbanized area. Since a society depends upon the mobility provided by transportation networks, an efficient transportation system is therefore necessary for its social and economic growth.

The recent modernization and economic advancement in most countries around the world have caused an increase in the demand for transportation. However, the increase of mobility has brought many serious social and environmental problems such as increased air pollution, increased accident rates and traffic congestion.

The quantitative analysis of existing transport systems and traffic phenomena has been done by traffic planners and researchers in order to address the above-mentioned problems. The real-world transport system is very complex, hence properly formulated traffic models are necessary in order to evaluate and manage such a system.

Traffic assignment is a major component in transportation planning. The basic concepts of traffic assignment began when the demand for transportation enormously increased after World War II [58]. The main purpose of the early traffic assignment problem was to estimate the diversion of traffic from existing roads to new, improved alternative roads in order to minimize the travel cost.

The early traffic assignment studies used assumptions which were highly unrealistic. For example, they assumed that the travel time and cost were independent of link flows when in reality, travel times and costs correspondingly increase with the increase in the number of users.

In 1952, a breakthrough in traffic assignment modeling came with the publication of the paper of Wardrop [70] on two principles of flow distribution in a transportation network, namely, the *user equilibrium principle* and the *system optimum principle.* The user equilibrium principle states that *users of the traffic network will choose the route having the minimum cost between each origin-destination (OD) pair, and through this process, the routes that are used will have equal costs; moreover, routes with costs higher than the minimum will have no flow.* The system optimum principle assumes that *users of the traffic network will choose their routes so as to minimize the total travel time in the transportation system.* These two principles have been the basis of most traffic assignment models.

A traffic equilibrium model aims at predicting flow patterns and travel times which are the results of the network user's choices with regard to routes from their origins to their destinations. The model is based on the behavioral assumption that all travelers compete noncooperatively for the network resources in order to minimize their travel costs [70]. This is the **Wardrop user equilibrium principle**. The traffic flows that satisfy this principle are usually referred to as *user-equilibrium* flows since the routes chosen by the network users are those which are individually perceived to be the shortest route under the prevailing condition. It should be noted that the only situation where the user optimum flows and the system optimum flows are equal is in the ideal condition when no congestion exists [58].

In this study, we focus on the user equilibrium principle. This is because in the real traffic system, route flows are likely closer to a user rather than a system optimum [58].

This chapter is organized as follows. In the next section, we discuss the mathematical models for the traffic equilibrium problem. We also discuss some of the solution methods for the TEP in Section 1.1. Section 1.2 presents the motivations and objectives of the study. Organization and contributions of this study are discussed in Section 1.3.

## 1.1 Mathematical Models for the Traffic Equilibrium Problem

The **traffic equilibrium problem** (TEP) is to find the flow pattern by allocating the OD demands to the network in such a way that no user can reduce his/her travel time by unilaterally changing his/her route.

Various mathematical models for the TEP have been introduced to better understand the problem. In what follows, we present the general TEP model, a bilevel model on the TEP, and the TEP model under uncertainty.

Let $\mathcal{G} = (\mathcal{A},\mathcal{N})$ be a transportation network, where $\mathcal{A}$ is the set of links (with cardinality $n_{\mathcal{A}}$) and $\mathcal{N}$ is the set of nodes (with cardinality $n_{\mathcal{N}}$). Let W be the set of origin-destination (OD) pairs in $\mathcal{G}$ (with cardinality $n_{\mathrm{W}}$). For every OD pair w $\in$ W, there corresponds the set $R_{\mathrm{w}}$ of routes connecting the OD pair w. We denote by $R$ the set of all routes (with cardinality $n_R$), i.e., $R = \bigcup_{\mathrm{w}\in\mathrm{W}} R_{\mathrm{w}}$. All throughout this study, we assume that the network $\mathcal{G}$ is connected, that is, there exists a route between each pair of nodes.

The cost experienced by a person using route $r$ is denoted by $C_r$. In general, route costs can be a function of the entire vector of route flows. The demand associated with each OD pair w, denoted by $D_{\mathrm{w}}$, is a function of the vector of minimum OD travel costs.

Figure 1.1 shows a sample traffic network with two origins and three destinations. The routes between origin $O_2$ to destination $D_3$, $R_{\mathrm{w}}$ for w $= \{2-3\}$, are shown in bold.

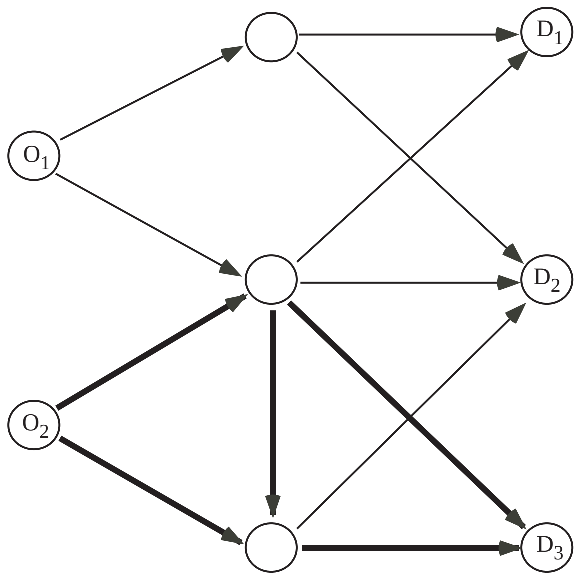



Figure 1.1: A sample traffic network.

### 1.1.1 General TEP Model

The TEP is to find a vector pair $(F,u)$ of route flows and minimum route costs such that conditions

$$0 \leq C_r(F) - u_{\rm w} \perp F_r \geq 0, \ \forall r \in R_{\rm w}, \ {\rm w} \in {\rm W}, \tag{1.1.1}$$

$$\sum_{r \in R_{\rm w}} F_r = D_{\rm w}(u), \ \forall {\rm w} \in {\rm W}, \tag{1.1.2}$$

$$u_{\rm w} \geq 0, \ \forall {\rm w} \in {\rm W}, \tag{1.1.3}$$

where $F \in \Re_+^{n_R}$ is the vector of route flows $F_r$, $u_{\rm w}$ is the minimal route cost for the OD pair w, and $u \in \Re_+^{n_{\rm W}}$ is the vector with components $u_{\rm w}$. The notation "$x \perp y$" means that vectors $x$ and $y$ are orthogonal and thus (1.1.1) implies $(C_r(F) - u_{\rm w})F_r = 0$ for all $r \in R_{\rm w}, {\rm w} \in {\rm W}$ are satisfied.

(1.1.1) is the mathematical representation of the Wardrop equilibrium principle, (1.1.2) means that the (elastic) travel demand must be satisfied, while (1.1.3) indicates that the minimum travel costs must be nonnegative.

Studies on the traffic equilibrium problem assumes that each traveler has complete and accurate information about the available routes and other characteristics of the traffic network. Such a TEP model is called the *deterministic model*. In reality, however, it can be observed that most traffic users have incomplete information of the network. This is known as the *stochastic model*.

The TEP model is also formulated using the fixed demand case or the elastic demand case. In the fixed demand case, $D(u) \equiv D$. In the *elastic demand* case, on the other hand, the TEP is formulated as a problem where the demand is a function of the minimum route costs $u$ between the origin and destination. The elastic TEP model is more realistic in the sense that, in general, a number of route choices are available to a traveler and that economic considerations usually affect the travel decisions of a user.

Another assumption used in the study of the TEP is that the route costs faced by the users in the network are *additive*, that is, the route costs are simply the sum of the link costs for all the links on the route being considered. Mathematically, the additive route cost can be written as

$$C_r(F) = \sum_{a \in \mathcal{A}} \delta_{ar} t_a(f) \text{ for all } r \in R_w, w \in W, \tag{1.1.4}$$

where $\delta_{ar}$ are the elements of the link-route incidence matrix $\Delta$, i.e.,

$$\delta_{ar} = \begin{cases} 1 & \text{if route } r \text{ passes through link } a \\ 0 & \text{otherwise,} \end{cases}$$

and $t_a(f)$ is the travel time on link $a \in \mathcal{A}$, $f_a$ is the flow on link $a \in \mathcal{A}$, $f$ is the vector of link flows $f_a = \sum_{r \in R} \delta_{ar} F_r$, $a \in \mathcal{A}$.

When the route costs are additive, TEP (1.1.1) – (1.1.3) can be reformulated as the monotone variational inequality problem. Hence, we can efficiently obtain an equilibrium of the TEP by using existing solution methods for the VIP.

However, as will be discussed later in Section 1.2, there are many situations where this additivity assumption on the route costs is inappropriate [35]. Hence, recent studies on the TEP have considered the case of *nonadditive route costs*.

### 1.1.2 Bilevel Model on the TEP

Traffic congestion is an important notion in the analysis of traffic equilibrium models. In a real traffic network, it can be observed that as the number of traffic users increases, the average speed on a link tends to decrease, which may lead to traffic congestion. The need for measures to reduce congestion in the urban traffic areas is becoming more serious as more and more people cluster in the cities, as a result of the modernization.

Road pricing is considered one of the effective means to reduce traffic congestion and environmental damage, and it has been introduced in major highways of most countries. Studies on road pricing consider a bilevel model wherein the traffic planner is assigned as "the leader" (upper-level decision maker) while the traffic users are called "the followers" (lower-level decision makers). Here, the leader makes some actions in order to achieve his goal (e.g., collects toll in order to alleviate traffic congestion), while the followers react to the actions of the leader by changing their behaviors (e.g., varying their travel schedules, route choices or travel modes) according to the traffic equilibrium principle.

The bilevel model can be formulated as a mathematical program with equilibrium constraints (MPEC) [52], which is a constrained optimization problem whose constraints are defined by a parametric variational inequality or complementarity system. The road pricing problem (RPP) on the TEP can be formulated as the following MPEC:

$$\begin{array}{rl} \min & \theta(\tau, F) \\ \text{s.t.} & \tau \in \mathbf{T}, \\ & (F, u) \in S(\tau), \end{array} \tag{1.1.5}$$

where $\theta(\tau, F)$ is an objective function (say, maximize revenue), $\mathbf{T}$ is the set of possible tolls imposed on all tollable links in the network (for example, $\mathbf{T} = \{\tau : \tau_a \geq 0, a \in \mathcal{A}\}$) and $S(\tau)$ is the set of solutions to the TEP satisfying conditions (1.1.1) – (1.1.3).

### 1.1.3 The TEP Model under Uncertainty

The concept of user equilibrium is generally associated with each traveler having full and accurate information about travel costs, and all travelers being uniform and rational in their decision-making [58]. That is, it is assumed that each traffic situation that happens is deterministic. In reality, however, traffic conditions are not deterministic but contain uncertainty. For example, traffic conditions vary depending on the weather condition. Travel time on a day with fine weather (when visibility is good) will probably differ from a snowy day when visibility is poor. Hence, it is necessary to include uncertainty in the study of traffic equilibrium models in order to present a more realistic view of the traffic conditions.

The traffic equilibrium problem with uncertainty can be formulated by considering $\Omega$ as the sample space of factors contributing to the uncertainty in the traffic network, say, weather or time of the day. The route cost function $C(F)$ is replaced by $C(F, \omega)$ and the travel demand $D$ is replaced by $D(\omega)$, $\omega \in \Omega$. The TEP with uncertainty can be written as the following stochastic variational inequality problem (SVIP): Find $F$ such that

$$\begin{aligned} 0 \leq C_r(F, \omega) - u_{\mathrm{w}} \perp F_r \geq 0, \ \forall r \in R_w, \ \mathrm{w} \in \mathrm{W}, \\ \sum_{r \in R_{\mathrm{w}}} F_r = D_{\mathrm{w}}(\omega), \ \forall \mathrm{w} \in \mathrm{W}, \\ u_{\mathrm{w}} \geq 0, \ \forall \mathrm{w} \in \mathrm{W} \end{aligned} \tag{1.1.6}$$

for each $\omega \in \Omega$.

It is important to note, however, that (1.1.6) does not have a solution in general. Hence, it is necessary to consider a reasonable solution instead.

### 1.1.4 Solution Methods for the TEP

The properties of an equilibrium solution to the TEP have been studied by considering reformulations of the Wardrop equilibrium conditions. These reformulations include [58]: (i)

Mathematical program (MP) (e.g., [46]); (ii) Fixed-point problem (FPP) (e.g., [6]); (iii) Variational inequality problem (VIP) (e.g., [21]); and (iv) Nonlinear complementarity problem (NCP) (e.g., [1]).

Various approaches for solving the equivalent reformulations of the TEP have been proposed. These solution methods include the generalized Newton's method [44], PATH solver [22], simplicial decomposition method [46], projection and contraction method [14], the smoothing method [16], and the regularization method [25]. Another approach is to reformulate the equivalent VIP reformulation as a minimization problem by the introduction of a merit function [35].

Convergence results for these approaches have been established under the key assumption of monotonicity. A VIP equivalent to the TEP with additive costs may usually be formulated as a monotone VIP [26]. However, a VIP derived from the TEP with nonadditive costs does not immediately possess monotonicity unless restrictive assumptions are made or a certain reformulation is introduced.

Solution methods for the bilevel problem in transportation have also been proposed. The bilevel problem, however, is known to be very complex, hence difficult to solve. Reformulations to the bilevel problems have been introduced, which include (i) transforming the bilevel problem into a (one-level) mixed 0-1 problem (e.g., [7]); and (ii) replacing the lower level problem by its stationarity conditions, provided it is convex, which leads to the MPEC.

Solution methods to the above-mentioned reformulation of the bilevel problem include the branch-and-bound algorithm [8] and trust region method [20]. Other solution methods include smoothing approach [17, 24, 52], penalty approach [49, 63] and implicit programming approach (ImPA) [23, 52]. Details of the ImPA will be presented in Chapter 4.

In the case of the TEP model under uncertainty, the resulting model is not deterministic. Hence, there is a need of transforming such a model into a certain deterministic formulation in order to solve the problem. The deterministic reformulations include (i) expected value method [38]; and (ii) expected residual minimization method [18]. The details are given in Chapter 5.

## 1.2 Motivations and Objectives of the Study

Various researches on the TEP have been done in order to better understand and present a more realistic representation of the real traffic conditions. More realistic traffic models are necessary for an efficient transportation system. However, realistic traffic models are known to be difficult to solve, or cannot be solved at all. Therefore, existing studies on the TEP have used various assumptions in order to present realistic TEP models which are solvable. It is the aim of this study to propose a solvable TEP model that is more realistic than the existing ones.

In the study of the TEP, one of the basic assumptions used is that the route costs are simply the sum of the link costs for all the links on the route being considered [1, 15, 21, 26]. Recall that the additive route cost function $C_r$ can be written as (1.1.4).

There are many situations, however, where this additivity assumption on the route costs is inappropriate. In particular, Gabriel and Bernstein [35] discussed several situations where the route costs are nonadditive:

(i) Nonadditive toll and fare schemes – most existing toll and fare schemes being implemented around the world are nonadditive.

(ii) Nonlinear valuation of travel time – different individuals have different valuations of time, which contributes to the nonadditivity of route costs.

(iii) Transportation policies – the different transportation policies, such as congestion pricing and the collection of emission fees, also add to the nonadditivity of route costs. For example, emissions of hydrocarbons and carbon monoxide are nonlinear functions of travel times.

Hence, it is indeed necessary to consider the nonadditivity of route costs in order to present a more realistic view of the traffic situation.

Various models of TEPs with nonadditive route costs have been proposed in the last decade [9, 35, 50]. However, those traffic models used a route cost function which is too simple or assumptions which are too restrictive in order to obtain solvable TEP models.

In this study, we consider the traffic equilibrium problem with nonadditive route costs. We introduce a route cost function which is a nonadditive disutility function of time (with money converted to time). Our objective is to reformulate the equivalent VIP formulation of the TEP into a monotone Mixed Complementarity Problem (MCP) under appropriate conditions such that the resulting reformulation of the TEP with nonadditive route cost can then be solved using existing solution methods.

Moreover, we consider the road pricing formulation for the TEP with nonadditive costs. This study aims to transform the road pricing formulation into an MPEC model by applying the results obtained from the above-mentioned proposed TEP with nonadditive route costs.

On the other hand, the presence of uncertainty in most conditions affecting traffic, such as different weather conditions, makes it necessary to consider the TEP with uncertainty in order to present a more realistic view of the traffic situation. Unfortunately, the resulting model of the TEP under uncertainty is not deterministic. Hence it is necessary to find deterministic formulations for such problems. Recently, a new approach called the expected residual (ER) method has been proposed to give a reasonable solution to problems of this kind [18, 19, 28, 48].

The ER method regards a minimizer of an expected residual function for the AVIP as a solution of SAVIP. Previous studies on the ER method employed the "min" function or the Fischer-Burmeister (FB) function [18, 19, 28]. Such functions however are nonconvex in general and hence we may not get a global solution.

In this study, we employ the regularized gap function and the D-gap function to define a residual in the ER model and show that our proposed ER models are convex under some conditions and hence a global solution can be obtained using existing solution methods.

## 1.3 Organization and Contributions

In the subsequent chapters, we will present some results related to the study of the traffic equilibrium problem (TEP). Below, we summarize the organization of the rest of the thesis as well as brief descriptions of the main contributions done in this study.

In Chapter 2, we provide an overview of the concepts used in this study. Specifically, we give some important concepts on monotonicity and convexity for VIP and MPEC that are necessary for better understanding of later arguments.

In Chapter 3, we consider the TEP with nonadditive costs. We modify the model presented in [35] by introducing a disutility function. We show that the equivalent VIP reformulation of the TEP can be transformed into a monotone MCP. We then establish the existence and uniqueness result for an equilibrium of this reformulation.

In Chapter 4, we consider the road pricing formulation for the TEP with nonadditive costs. We apply the results in Chapter 3 to transform the road pricing formulation into a mathematical program with equilibrium constraint (MPEC). We then show that this MPEC formulation of the road pricing problem can be reformulated as a mathematical program with strictly monotone MCP which can be solved using existing solution methods.

In Chapter 5, we consider the TEP under uncertainty. We begin first by considering the affine variational inequality problem (AVIP) under uncertainty. We then propose two new ER models, the *ER-R model* which uses the regularized gap function and the *ER-D model* which uses the D-gap function for the stochastic affine variational inequality problem (SAVIP). We establish the convexity of the two ER models and solve the traffic equilibrium problem under uncertainty using the ER-D model.

In Chapter 6, we give a brief summary and conclusion of the main contributions of this dissertation. We also mention some issues for future consideration.

# Chapter 2

# Preliminaries

In this chapter, we provide an overview of the important concepts used in this thesis. Specifically, we give some important concepts and existing results for the variational inequality problem, the mathematical program with equilibrium constraints and their solution methods.

## 2.1 Convexity and Monotonicity

In this section, we define the concepts of convexity and monotonicity. Note that monotonicity plays an important role in the existence and uniqueness of equilibrium of the TEP while convexity guarantees the global optimality of local minimum. We first define the convexity for sets and real-valued functions.

**Definition 2.1.1.** *A set $K \in \Re^n$ is said to be convex if $(1-a)x + ay \in K$ holds for any vectors $x, y \in K$ and scalar $a \in (0,1)$.*

**Definition 2.1.2.** *Let $K \in \Re^n$ be a nonempty and convex set. Then, a function $\rho : K \to \Re^n$ is said to be*

(i) *convex (on K) if* $\rho((1-a)x+ay) \leq (1-a)\rho(x)+a\rho(y)$ *holds for any* $x, y \in K$ *and* $a \in (0,1)$*;*

(ii) *strictly convex (on K) if* $\rho((1-a)x+ay) < (1-a)\rho(x)+a\rho(y)$ *holds for any* $x, y \in K$ *with* $x \neq y$ *and* $a \in (0,1)$*; and*

(iii) *strongly convex (on K) with modulus* $\varepsilon > 0$ *if* $\rho((1-a)x+ay) \leq (1-a)\rho(x)+a\rho(y) - (1-a)a||x-y||^2$ *holds for any* $x, y \in K$ *and* $a \in (0,1)$*.*

It is obvious that any strongly convex function is strictly convex, and any strictly convex function is convex. For example, a linear function is convex but not strictly convex, $\rho(\alpha) = e^{\alpha}$ is strictly convex but not strongly convex, and $\rho(\alpha) = \alpha^2$ is strongly convex.

Convexity plays a crucial role in the field of optimization. In particular, in the nonlinear programming problem

$$\text{minimize} \quad \rho(z) \quad \text{subject to} \quad \kappa_i(z) \leq 0 \ \ (i = 1, \ldots, m),$$

if functions $\kappa_1, \ldots, \kappa_m$ and $\rho$ are convex, then any local minimum of the problem is a global minimum.

Next, we define monotonicity for vector-valued mappings from a subset of $\Re^n$ to $\Re^n$.

**Definition 2.1.3.** *Let* $K \subseteq \Re^n$ *be a nonempty and convex set. A function* $G : \Re^n \to \Re^n$ *is called*

(i) *monotone (on K) if* $(x-y)^T\big(G(x)-G(y)\big) \geq 0, \forall x, y \in K$*;*

(ii) *strictly monotone (on K) if* $(x-y)^T\big(G(x)-G(y)\big) > 0, \forall x, y \in K$ *with* $x \neq y$*; and*

(iii) *strongly monotone (on K) with modulus* $\varepsilon > 0$ *if* $(x-y)^T\big(G(x)-G(y)\big) \geq \varepsilon\|x-y\|^2, \forall x, y \in K$*.*

It is obvious that any strongly monotone function is strictly monotone, and any strictly monotone function is monotone.

It is important to note that monotonicity of the function $G$ in Definition 2.1.3 particularly plays an important role in the existence and uniqueness of an equilibrium of TEP. Most of the existing results for TEP rely on the assumption that the function $G$ involved satisfies certain conditions such as strong or strict monotonicity [26]. Moreover, monotonicity is also important for solution methods for TEP to work efficiently.

The following proposition shows the relation between convexity and monotonicity.

**Proposition 2.1.1.** *Let $K \subseteq \Re^n$ be an open convex set, and $\rho : K \to \Re$ be a continuously differentiable function. Then*

*(i) $\rho$ is convex on $K$ if and only if $\nabla\rho$ is monotone on $K$;*

*(ii) $\rho$ is strictly convex on $K$ if and only if $\nabla\rho$ is strictly monotone on $K$; and*

*(iii) $\rho$ is strongly convex on $K$ with modulus $\varepsilon > 0$ if and only if $\nabla\rho$ is strongly monotone on $K$ with modulus $\varepsilon > 0$.*

It can be seen from the above proposition that there is a close relation between convexity and monotonicity. Moreover, these properties also have much relevance to the positive (semi)definiteness of matrices.

**Proposition 2.1.2.** *Let $K \subseteq \Re^n$ be an open convex set, and $G : K \to \Re^n$ be a continuously differentiable function. Then*

*(i) $G$ is monotone on $K$ if and only if $\nabla G(x)$ is positive semidefinite for any $x \in K$;*

*(ii) $G$ is strictly monotone on $K$ if $\nabla G(x)$ is positive definite for any $x \in K$; and*

*(iii) $G$ is strongly monotone on $K$ if and only if there exists $\varepsilon > 0$ such that*

$$\min_{||e||=1} e^T \nabla G(x) e \geq \varepsilon$$

*for any $x \in K$.*

Note that (ii) above does not hold when "if" is replaced by "if and only if". For example, though a function $G : \Re \to \Re$ defined by $G(\alpha) = \alpha^3$ is monotonically increasing on $\Re$, $\nabla G(\alpha) = 3\alpha^2$ is not positive when $\alpha = 0$.

The above two propositions directly lead to the following corollary which mentions the relation between the convexity of a real-valued function and the positive (semi)definiteness of its Hessian matrix.

**Corollary 2.1.3.** *Let $K \subseteq \Re^n$ be an open convex set, and $\rho : K \to \Re$ be a twice continuously differentiable function. Then*

*(i) $\rho$ is convex on $K$ if and only if $\nabla^2 \rho(x)$ is positive semidefinite for any $x \in K$;*

*(ii) $\rho$ is strictly convex on $K$ if $\nabla^2 \rho(x)$ is positive definite for any $x \in K$; and*

*(iii) $\rho$ is strongly convex on $K$ if and only if there exists $\varepsilon > 0$ such that*

$$\min_{||e||=1} e^T \nabla^2 \rho(x) e \geq \varepsilon$$

*for any $x \in K$.*

Moreover, the following corollary on monotonicity of affine functions and convexity of quadratic functions can be easily shown.

**Corollary 2.1.4.** *Let $M \in \Re^{nxn}$ and $q \in \Re^n$ be a given matrix and a vector, respectively. Let $G : \Re^n \to \Re$ and $\rho : \Re \to \Re$ be defined by $G(x) = Mx + q$ and $\rho(x) = \frac{1}{2} x^T M x + q^T x$, respectively. Then, we have*

*(i) $M$ is positive semidefinite $\Leftrightarrow$ $G$ is monotone $\Leftrightarrow$ $\rho$ is convex; and*

*(ii) $M$ is positive definite $\Leftrightarrow$ $G$ is strongly monotone $\Leftrightarrow$ $\rho$ is strongly convex.*

Corollary 2.1.4 implies that the strong monotonicity is equivalent to the strict monotonicity for affine functions, and that the strong convexity is equivalent to the strict convexity for quadratic functions.

## 2.2 Variational Inequality Problem and Mathematical Program with Equilibrium Constraints

The *variational inequality problem (VIP)* provides a convenient framework for the TEP, especially in the analysis and computation of an equilibrium of the TEP. The VIP [26] is generally stated as follows: Find a vector $x \in K$ such that

$$(y - x)^T G(x) \geq 0, \quad \forall y \in K, \tag{2.2.1}$$

where $K$ is a nonempty closed convex subset of $\Re^n$ and $G : K \to \Re^n$ is a continuous function. This problem is denoted by VIP$(K, G)$. The VIP is a very large class of problems containing systems of equations, convex programming problems, and complementarity problems. In particular, when $K = \Re^n$, (2.2.1) is equivalent to the equation:

$$G(y) = 0.$$

Moreover, the following constrained optimization problem can be reformulated as a VIP:

$$\begin{aligned} &\text{minimize} \quad \rho(y) \\ &\text{subject to} \quad y \in K, \end{aligned} \tag{2.2.2}$$

where the objective function $\rho$ is continuously differentiable ($C^1$) on an open superset of the closed convex set $K \subseteq \Re^n$. Any local minimizer $x$ of (2.2.2) must satisfy

$$(y - x)^T \nabla\rho(x) \geq 0, \quad \forall y \in K,$$

which is the VIP$(K, \nabla\rho)$. A solution of the VIP$(K, \nabla\rho)$ is called a stationary point of (2.2.2). Moreover, if $\rho$ is a convex function, then every stationary point of (2.2.2) is a global minimum of the above optimization problem. Hence, for a convex function $\rho$ and a convex set $K$, VIP$(K, \nabla\rho)$ is equivalent to the optimization problem (2.2.2).

Special cases of the VIP include the Nonlinear Complementarity Problem (NCP) and the Mixed Complementarity Problem (MCP). The NCP is the VIP with $K = \Re^n_+ \equiv \{x \in \Re^n | x_i \geq 0, i = 1, \ldots, n\}$ and the MCP is the VIP with $K = \{x \in \Re^n | a_i \leq x_i \leq b_i$, $i = 1, \ldots, n\}$, where $a_i \in \Re \cup \{-\infty\}$, $b_i \in \Re \cup \{+\infty\}$, $a_i \leq b_i$, $i = 1, \ldots, n$. We denote the NCP with the function $G$ by NCP$(G)$ and the MCP with the function $G$ and the set $K$ by MCP($G$,$K$).

The TEP can be formulated as a VIP. When the travel cost and the demand functions $C_r(F)$ and $D_{\mathrm{w}}(u)$ are nonnegative, and for each OD pair $\mathrm{w} \in \mathrm{W}$,

$$\Big[\sum_{r\in R_{\mathrm{w}}} F_r C_r(F) = 0, F \geq 0\Big] \Longrightarrow [F_r = 0,\ \forall\, r \in R_{\mathrm{w}}], \tag{2.2.3}$$

then the Wardrop equilibrium conditions (1.1.1) – (1.1.3) are equivalent to the NCP($H$) with the function $H$ defined by

$$H(F,u) \equiv \begin{pmatrix} C(F) - \Gamma u \\ \Gamma^T F - D(u) \end{pmatrix}, \tag{2.2.4}$$

where $C(F)$ is the vector of route costs $C_r(F)$, and $\Gamma = (\Gamma_{r\mathrm{w}})$ is the route-OD pair incidence matrix whose entries are given by

$$\Gamma_{r\mathrm{w}} = \begin{cases} 1 & \text{if} \quad r \in R_{\mathrm{w}} \\ 0 & \text{otherwise.} \end{cases}$$

**Remark 2.2.1.** *If the route cost function $C_r$ is positive, then since for each* $\mathrm{w} \in \mathrm{W}$*, $F_r$ is positive for some $r \in R_{\mathrm{w}}$, we have $C_r(F) - u_w = 0$ from (1.1.1), hence, $u_w = C_r(F) > 0$. Thus, the NCP(H) can be rewritten as*

$$0 \leq C(F) - \Gamma u \perp F \geq 0,$$

$$\Gamma^T F - D(u) = 0,$$

*which is the MCP(H,L) with the set $L$ defined by*

$$L = \Re_+^{n_R} \times \Re^{n_{\mathrm{w}}}.$$

Various approaches for solving the MCP have been proposed. Those solution methods include the generalized Newton's method [44], the smoothing method [16] and the regularization method [25]. Another method is to reformulate the VIP as a minimization problem by the introduction of a merit function [35]. Convergence results for these approaches have been established under the key assumption of monotonicity on $H$.

Another important class related to the TEP is the *mathematical program with equilibrium constraints* (MPEC), which has two sets of variables, namely, an upper-level variable $x \in \Re^n$ and a lower-level variable $y \in \Re^m$, and in which some or all of its constraints are defined by a parametric variational inequality or complementarity system with $y$ as its primary variables and $x$ the parameter vector [52]. The MPEC is generally stated as follows:

$$\begin{aligned} \min \quad & \theta(x, y) \\ \text{s.t.} \quad & (x, y) \in Z, \\ & y \text{ solves VIP}(K(x), G(x, \cdot)). \end{aligned} \tag{2.2.5}$$

Here, $Z$ is a subset of $\Re^{n+m}$, $\theta : \Re^{n+m} \to \Re^m$, $K : \Re^n \to 2^{\Re^m}$ are mappings, and $y$ solves $\text{VIP}(K(x), G(x, \cdot))$ if and only if $y \in K(x)$ and

$$(y' - y)^T G(x, y) \geq 0, \quad \text{for all } y' \in K(x).$$

Note that the above variational inequality is what is generally referred to as the *equilibrium constraints* in MPEC.

The MPEC (2.2.5) is a generalization of the bilevel programming problem that is a mathematical program with optimization constraints. Moreover, the MPEC is also closely related to the so-called *Stackelberg game* [65], in which a distinctive player (called the leader) can anticipate the actions (or reactions) of the other players (called the followers) and he then uses such knowledge to select his optimal strategy. Each follower, on the other hand, acts and devices his own strategy based on the particular strategy of the leader and his cost function is dependent on both the leader's and all the other followers' strategies.

Other economic applications of MPEC include oligopolistic market analysis [64], volatility estimation of American pricing option [42] and pricing of electric transmission [41]. MPECs are also used in the study of engineering problems such as problems in elastoplasticity [32] and obstacle problems [57].

Many transportation planning and design problems can be formulated as MPECs. In network design problems, one is concerned with the modification of a transportation infrastructure by adding new links or improving existing ones in order to maximize social welfare and/or minimize design and other costs [12, 15].

Although MPEC plays a very important role in many fields such as engineering design, economic equilibrium and multilevel game, MPEC is very difficult to deal with because, from the geometric point of view, its feasible region is not convex and not connected in general. Moreover, its constraints fail to satisfy a standard constraint qualification such as the linear independence constraint qualification of the Mangasarian-Fromovitz constraint qualification at any feasible point [52]. Hence, existing solution methods for nonlinear programming cannot be applied to MPEC directly.

Many methods have been proposed in order to solve MPECs. Such solution methods include smoothing approach [17, 24, 52], penalty approach [49, 63] and implicit programming approach (ImPA) [23, 52]. The ImPA is known to be useful when the lower level problem has a unique solution for every value of the upper level variable.

## 2.3 Existence and Uniqueness of Solution of the VIP

In the last decades, the VIP has been studied extensively. The monotonicity of $G$ particularly plays an important role in the existence and uniqueness of solutions of VIP. Existence and uniqueness are especially important in the study of the TEP since route flows do exist and are certainly unique in practice.

In what follows, we discuss various attributes of the solution set of VIP (2.2.1). We begin with the following proposition.

**Proposition 2.3.1.** *Let $G : \mathfrak{R}^n \to \mathfrak{R}^n$ be a continuous function, and $K \in \mathfrak{R}^n$ be a nonempty closed convex set. If $K$ is bounded, then* (2.2.1) *has at least one solution.*

Proposition 2.3.1 guarantees the solvability of (2.2.1) under the boundedness assumption on K.

Using Proposition 2.3.1, we can show that the following results guarantee the existence of a solution to the equivalent NCP formulation of the TEP.

**Proposition 2.3.2.** *If each route cost function $C_r$ is nonnegative and continuous for all $r \in R_{\mathrm{w}}$, and each demand function $D_{\mathrm{w}}$ is a nonnegative continuous function bounded above for all $\mathrm{w} \in \mathrm{W}$, then the NCP(H) with the function $H$ defined by* (2.2.4) *has a solution.*

*Proof.* See Theorem 3.17 [58]. □

**Proposition 2.3.3.** *Let $G : \mathfrak{R}^n \to \mathfrak{R}^n$ be a continuous function, and $K \subseteq \mathfrak{R}^n$ be a nonempty closed convex set. Let $S$ be the (possibly empty) solution set of the VIP* (2.2.1).

*(i) If $G$ is monotone on $K$, then $S$ is a closed convex set;*

*(ii) If $G$ is strictly monotone on $K$, then $S$ consists of at most one element; and*

*(iii) If $G$ is strongly monotone on $K$, then $S$ consists of exactly one element.*

Note that in this proposition, only (iii) guarantees the existence of a solution to (2.2.1).

A VIP equivalent to the TEP with additive costs may usually be formulated as a monotone VIP [26]. However, a VIP derived from the TEP with nonadditive costs does not immediately possess monotonicity unless restrictive assumptions are made or a certain reformulation is introduced.

**Proposition 2.3.4.** *Suppose that the demand function $D_{\mathrm{w}}$ is a positive continuous function bounded above for each $\mathrm{w} \in \mathrm{W}$. Suppose further that the travel time function $t_a$ is positive and continuous for each $a \in \mathcal{A}$. If $D$ is positive and $t$ is strictly monotone, then the NCP($H$) has a unique solution.*

*Proof.* See Theorem 3.19 [58]. $\square$

In Chapter 3, we present the TEP with nonadditive route costs and show that under some reasonable assumptions, the proposed TEP with nonadditive route costs can be reformulated so that it can then be solved using existing solution methods.

## 2.4 Solution Methods for VIP and NCP

Various approaches for solving the VIP and MCP have been proposed. These solution methods include the generalized Newton's method [44], the smoothing method [16] and the regularization method [25]. Another method is to reformulate the VIP as a minimization problem or as an equivalent system of equations by the introduction of a merit function [35]. In this section, we introduce reformulation approaches for VIP and NCP.

### 2.4.1 Merit Function for VIP and NCP

A real-valued function which takes 0 at a solution of an equilibrium problem and takes a positive value otherwise is called a *merit function* of the equilibrium problem. We formally define merit functions in the following.

**Definition 2.4.1.** *A merit function for the VIP(K, G) on a (closed) set $K \subseteq X$ is a nonnegative function $\Psi : X \to \Re_+$ such that $x$ solves VIP(K, G) if and only if $x \in X$ and $\Psi(x) = 0$, that is, if and only if the solutions of the VIP(K, G) coincide with the global solutions of the problem*

$$\begin{aligned} &\min\ \Psi(x) \\ &\text{s.t. } x \in X \end{aligned} \tag{2.4.1}$$

*and the optimal objective value of this problem is zero [26].*

Before we introduce a merit function for NCP, we first define NCP functions.

A function $\phi : R^2 \to R$ is called an *NCP function* if it has the property:

$$\phi(a, b) = 0 \ \Leftrightarrow\ a \geq 0, b \geq 0, ab = 0.$$

Two popular NCP functions are

(i) *min function* : $\phi(a, b) = \min(a, b)$;

(ii) *Fischer-Burmeister (FB) function* : $\phi(a, b) = a + b - \sqrt{a^2 + b^2}$ .

Using an NCP function $\phi$, the NCP($G$) can be reformulated as the following system of nonlinear equations:

$$\Phi(x) = \begin{pmatrix} \phi(x_1, G_1(x)) \\ \vdots \\ \phi(x_n, G_n(x)) \end{pmatrix} = 0. \tag{2.4.2}$$

Thus, the NCP (2.2.4) (or MCP) of the TEP can be reformulated as the following system of nonlinear equations:

$$\Phi(F,u) = \begin{pmatrix} \phi\Big(F_1, C_1(F) - u_{\mathrm{w}}\Big) \\ \vdots \\ \phi\Big(F_{n_R}, C_{n_R}(F) - u_{\mathrm{w}}\Big) \\ \Gamma^T F - D(u) \end{pmatrix} = 0. \tag{2.4.3}$$

The nonlinear equations (2.4.2) can be reformulated as the following unconstrained minimization problem:

$$\min \Psi(x), \tag{2.4.4}$$

where $\Psi(x) = ||\Phi(x)||^2$. The value of $\Psi(x)$ is always larger than or equal to 0. Hence, $\Psi(x)$ is a merit function of the NCP($G$).

Next we introduce merit functions for VIP. Popular merit functions for VIP($K, G$) are [26]:

(i) *Regularized gap function* ($X = K$):

$$f_\alpha(x) = \max_{y \in K} \left\{ \langle G(x), x - y \rangle - \frac{1}{2\alpha} \|y - x\|^2 \right\}, \quad \alpha > 0; \tag{2.4.5}$$

and

(ii) *D-gap function* ($X = \Re^n$) :

$$g_\alpha(x) = f_\alpha(x) - f_{1/\alpha}(x), \quad \alpha > 1. \tag{2.4.6}$$

Note that $f_\alpha(x)$ and $g_\alpha(x)$ are also merit functions for NCP.

Different merit functions $\Psi$ result in different optimization problems. Hence, it is necessary to choose appropriate merit functions having as many desirable properties as possible. Important properties of merit functions are as follows [33]:

- NCP($G$) can be reformulated by means of a merit function into a constrained or unconstrained optimization problem. *Unconstrained problems* can be solved by simpler algorithms. *Constrained problems*, on the other hand, may possess desirable properties, for example, they may contain fewer stationary points which do not solve the NCP.

- The merit function $\Psi$ must be continuously differentiable (smooth). *Smooth merit functions* allow the use of well-developed (existing) smooth optimization theory and numerical methods in nonlinear programming.

- Merit functions can be used in the design of numerical algorithms for solving the VIP and MCP [26]. Specifically, one can use an iterative algorithm to minimize the merit function in order to obtain its global minimum. However, merit functions are not convex in general so that obtaining a global minima is not always guaranteed. Often, one can only find a stationary point of the problem. Thus, it is indeed necessary to know when such a stationary point becomes a solution to the VIP/NCP.

In the following, we summarize the properties of some merit functions for VIP and NCP.

**Theorem 2.4.1.** *[26]* ***(Equivalence)***

*(i) Squared FB function: Let*

$$\Psi_{FB}(x) = \frac{1}{2}\sum_{i=1}^{n}(\phi(x_i, G_i(x)))^2,$$

*where $\phi$ is the Fischer-Burmeister function. Then solving the NCP(H) is equivalent to finding a global solution of the following unconstrained minimization problem:*

$$\min_{x\in\Re^n} \Psi_{FB}(x), \tag{2.4.7}$$

*if the NCP has a solution.*

*(ii) Regularized gap function: Let $K \subseteq \Re^n$ be a closed convex set. The regularized gap function $f_\alpha$ is nonnegative on $K$ and $x$ solves VIP $(K, G)$ if and only if $f_\alpha(x) = 0$ and $x \in K$. Hence, VIP$(K, G)$ is equivalent to the constrained optimization problem*

$$\begin{aligned} \min \quad & f_\alpha(x) \\ \text{s.t.} \quad & x \in K. \end{aligned} \tag{2.4.8}$$

(iii) *D-gap function: The D-gap function $g_\alpha$ is nonnegative in $\Re^n$ and $g_\alpha(x) = 0$ if and only if $x$ is a solution of the VIP($K, G$). Hence, solving VIP($K, G$) is equivalent to finding a solution of the unconstrained minimization problem*

$$\begin{array}{ll} \min & g_\alpha(x) \\ \text{s.t.} & x \in \Re^n. \end{array} \tag{2.4.9}$$

**Theorem 2.4.2.** *[26] **(Differentiability)** Suppose that $G$ is continuously differentiable.*

(i) *The squared FB function $\Psi_{FB}$ is continuously differentiable.*

(ii) *The regularized gap function $f_\alpha$ is continuously differentiable.*

(iii) *The D-gap function $g_\alpha$ is continuously differentiable.*

In reformulating (2.4.2) into an unconstrained minimization problem, one may use the FB function which is known to have some advantageous properties over other NCP functions, namely, the FB-function is semismooth and the merit function defined by the squared FB-function is smooth and differentiable [27].

**Theorem 2.4.3.** ***(Stationarity)** [26]*

(i) *Squared FB function: Suppose that $G$ is a monotone function. Then $x^*$ is a stationary point of $\Psi_{FB}$ if and only if $x^*$ is a solution of NCP($G$).*

(ii) *Regularized gap function: Suppose that $G$ is a strongly monotone function. Then $x^*$ is a stationary point of* (2.4.8) *if and only if $x^*$ solves VIP($K, G$).*

(iii) *D-gap function: Let $G$ be a strongly monotone function. Then $x^*$ is a stationary point of* (2.4.9) *if and only if $x^*$ is a solution of VIP($K, G$).*

The following results for the convexity of the regularized gap function (2.4.5) and the D-gap function (2.4.6) for LCP have been established by Peng [61].

**Theorem 2.4.4.** ***(Theorem 2.1, [61])*** *Let $G(x) = Mx + q$. If $M$ is positive definite, then the regularized gap function $f_\alpha$ is convex for any sufficiently large $\alpha > 0$.*

**Theorem 2.4.5.** ***(Theorem 3.1, [61])*** *Let $G(x) = Mx + q$. If $M$ is positive definite, then the D-gap function $g_\alpha$ is convex if $\alpha > \overline{\alpha}$, where $\overline{\alpha}$ is given by*

$$\overline{\alpha} = \max_{\|x\|=1} \frac{1 + x^T M^T M x}{2x^T M x} .$$

The above results guarantee the convexity of (2.4.5) and (2.4.6) under the conditions that $G$ is affine, $\nabla G$ is positive definite and the parameter $\alpha$ is sufficiently large.

In this study, we extend the above results to the affine variational inequality problem (AVIP).

### 2.4.2 The Generalized Newton Method for NCP

We now present the Generalized Newton Method (GNM) to solve the NCP($G$) proposed by Jiang [43]. In this method, the Fischer-Burmeister (FB) function $\phi$, which is known to be nonsmooth but convex, is used to reformulate the NCP($G$) into a system of nonsmooth equations $\Phi(x) = 0$ given by (2.4.2). This is further equivalent to finding a global solution of the minimization problem

$$\min_{x \in \Re^n} \Psi_{FB}(x), \tag{2.4.10}$$

where

$$\Psi_{FB}(x) \;=\; \frac{1}{2}||\Phi(x)||^2 = \frac{1}{2}\sum_{i=1}^{n}(\phi(x_i, G_i(x)))^2 \tag{2.4.11}$$

and $\phi$ is the Fischer-Burmeister function, if the NCP($G$) has a solution.

In what follows, we present some concepts necessary for the understanding of the generalized Newton method.

Let $\mathcal{H} : \Re^n \to \Re^n$ be locally Lipschitz on $\Re^n$. Then *Clarke's generalized Jacobian* of $\mathcal{H}$ at $x$, denoted by $\partial\mathcal{H}(x)$, can be defined as

$$\partial\mathcal{H}(x) := co\{\lim_{x^k \to x} \nabla\mathcal{H}(x^k) \;:\; \mathcal{H} \text{ is differentiable at } x^k \in \Re^n\},$$

where $co$ stands for the convex hull. Denote $y \to^d x$ if $y \to x$, $y \neq x$ and $(y - x)/||y - x|| \to d/||d||$ for some $d \in \Re^n \backslash \{0\}$.

A function $\mathcal{H} : \Re^n \to \Re^n$ is said to *semismooth* at $x \in \Re^n$ if $\mathcal{H}$ is Lipschitz continuous on an open neighborhood of $x$ and the following limit exists

$$\lim_{V \in \partial\mathcal{H}(y),\, y \to^d x} Vd.$$

A necessary condition for $\mathcal{H}$ to be semismooth at $x$ is that $\mathcal{H}$ is directionally differentiable at $x$ [62]. Moreover, the semismoothness of $\mathcal{H}$ at $x$ implies that for any $d \in \Re^n$,

$$\lim_{||d|| \to 0} \frac{\mathcal{H}(x+d) - \mathcal{H}(x) - \mathcal{H}'(x,d)}{||d||} = 0, \tag{2.4.12}$$

and

$$\lim_{V \in \partial\mathcal{H}(x+d)\, ||d|| \to 0} \frac{Vd - \mathcal{H}'(x,d)}{||d||} = 0, \tag{2.4.13}$$

where $\mathcal{H}'(x,d)$ denotes the *directional derivative* of $\mathcal{H}$ at $x$ along the direction $d$.

Since $\phi$ is Lipschitz and $G$ is smooth, it follows that $\Phi$ is locally Lipschitz on $\Re^n$. However, $\Phi$ is not continuously differentiable at a given point $x$ in general except that some strong assumptions are imposed such as the condition that $x_i^2 + G(x)_i^2 > 0$ for $i = 1, \ldots, n$.

Let $N(x) = \{i |\, x_i^2 + G_i(x)^2 = 0\}$ and let $x \in \mathfrak{F} = \{x \in \Re^n : N(x) = \emptyset\}$. Then

$$\nabla\Phi(x) = \mathrm{diag}(\gamma_i(x))\nabla G(x) + \mathrm{diag}(\mu_i(x)),$$

where $\mathrm{diag}(\varsigma_i)$ denotes the diagonal matrix with diagonal elements $\varsigma_1, \varsigma_2, \ldots, \varsigma_n$; and

$$\gamma_i(x) = \frac{G_i(x)}{\sqrt{(G_i(x))^2 + x_i^2}} - 1,$$

$$\mu_i(x) = \frac{x_i}{\sqrt{(G_i(x))^2 + x_i^2}} - 1.$$

For $x \notin \mathfrak{F}$, it follows from the definition of Clarke's generalized Jacobian that any $V \in \partial\Phi(x)$ can be written as

$$V = \mathrm{diag}(\gamma_i)\nabla G(x) + \mathrm{diag}(\mu_i),$$

where $(\gamma_i + 1)^2 + (\mu_i + 1)^2 \leq 1$, $i = 1, 2, \ldots, n$.

We now present the *generalized Newton method of Jiang and Qi [44]* for the general nonlinear equations $G(x) = 0$ :

**Step 1.** Choose an initial point $x^0 \in \Re^n$ and let $k = 0$.

**Step 2.** Choose $V_k \in \partial\Phi(x^k)$ and solve the following Newton equations for the direction $d^k \in \Re^n$:

$$\Phi(x^k) + V_k d^k = 0. \tag{2.4.14}$$

**Step 3.** Set $x^{k+1} = x^k + d^k$. If $x^{k+1}$ solves $\Phi(x) = 0$, stop. Otherwise, let $k := k+1$ and go to Step 2.

Note that (2.4.14) is not necessarily solvable if $V_k$ is singular. Moreover, the generalized Newton method may not converge globally. To remedy such drawbacks of the method, the GNM can be combined with the Levenberg-Marquardt and line search methods using merit functions. This *generalized Newton method (GNM) with line search* proposed by Jiang [43] and used in our numerical experiments in Chapter 3 is described as follows:

**Step 1.** (Initialization) Choose an initial starting point $x^0 \in \Re^n$, two scalars $\wp$, $\varrho \in (0,1)$, and let $k := 0$.

**Step 2.** (Search direction) Choose $V_k \in \partial\Phi(x^k)$ and solve the following generalized Newton equation:

$$V_k^T \Phi(x^k) + (V_k^T V_k + \upsilon_k I)d = 0, \tag{2.4.15}$$

with $I$ the identity matrix in $\Re^{n\times n}$. If $d = 0$ is a solution of the generalized Newton equation, the algorithm terminates. Otherwise, let $d^k$ be the solution of the above equation and go to Step 3.

**Step 3.** (Line search) Let $\lambda_k = \varrho^{i_k}$, where $i_k$ is the smallest nonnegative integer $i$ such that

$$\Psi(x^k + (\varrho)^i d^k) - \Psi(x^k) \leq \wp(\varrho)^i \nabla\Psi(x^k)^T d^k.$$

**Step 4.** (Update) Choose $\upsilon_{k+1} > 0$. Let $x^{k+1} := x^k + \lambda_k d^k$ and $k := k+1$. Go to Step 1.

The following results were established in [43].

**Theorem 2.4.6.** *[43] Suppose that $G$ is monotone. Let $x^*$ be an accumulation point of $\{x^k\}$ generated by the GNM with line search. Let $\upsilon_k = \min\{\Psi(x^k), ||\nabla\Psi(x^k)||\}$. Then $x^*$ is a solution of the NCP. Moreover, $\{x^k\}$ converges to $x^*$ superlinearly if $\wp \in (0, \frac{1}{2})$ and if each element of $\partial\Phi(x^*)$ is nonsingular.*

## 2.5 The Solution Method for the MPEC Model of the Road Pricing Problem

In this study, we consider the road pricing problem (RPP) which is formulated as a mathematical program with equilibrium constraints (MPEC) of the form

$$\begin{array}{ll} \min & \theta(\tau, F) \\ \text{s.t.} & \tau \in \mathbf{T}, \\ & (F, u) \in S(\tau), \end{array} \tag{2.5.1}$$

where $\mathbf{T}$ is the set of possible tolls imposed on all tollable arcs in the network and $S(\tau)$ is the solution set of MCP

$$\begin{array}{l} 0 \leq \widetilde{C}(\tau, F) - \Gamma u \perp F \geq 0, \\ \Gamma^T F - D(u) = 0, \end{array} \tag{2.5.2}$$

where $\widetilde{C}(\tau, F)$ is the vector of nonadditive route costs with a given toll $\tau$.

If $S(\tau)$ is a singleton for all $\tau \in \mathbf{T}$, then MPEC (2.5.1) can be written as an *implicit optimization problem* in the upper-level variable $\tau$ alone:

$$\begin{array}{ll} \min & \widetilde{\theta}(\tau) \\ \text{s.t} & \tau \in \mathbf{T}, \end{array} \tag{2.5.3}$$

where $\widetilde{\theta}(\tau) \equiv \theta(\tau, y(\tau))$, $y \equiv (F, v)$, $y(\tau) \equiv (F(\tau), v(\tau))$.

The ImP approach is considered to be restrictive since it requires unnecessarily strong assumptions such as strong monotonicity for the approach to be applicable to certain problems. However, the ImPA assumes that for $\tau$ sufficiently close to $\bar{\tau}$, $y(\tau)$ is locally single-valued and has a certain first-order directional smoothness property. Moreover, with the existence of such an implicit function, the implicit optimization problem can then be solved using existing optimization solvers [52].

## 2.6 Expected Residual Approach for the VIP with Uncertainty

The stochastic variational inequality problem (SVIP) is to find a vector $x \in K$ such that

$$(y-x)^T G(x,\omega) \geq 0,\ \forall y \in K, \tag{2.6.1}$$

where $K \subseteq \Re^n$ is a nonempty closed convex set, $G : \Re^n \times \Omega \to \Re^n$ is a vector-valued function and $(\Omega, P)$ is a probability space with $\Omega \subseteq \Re^n$.

In general, there is no $x$ satisfying the above SVIP for all $\omega \in \Omega$. One approach is to consider the following deterministic reformulation of (2.6.1):

$$(y-x)^T E[G(x,\omega)] \geq 0,\ \forall y \in K, \tag{2.6.2}$$

where $E[G(x,\omega)]$ is the expectation function of $G(x,\omega)$. This is known as the *expected value* (EV) model.

Another existing approach considers the following deterministic formulation which is to find a vector $x \in R^n_+$ that minimizes the expected total residual defined by an NCP function, that is,

$$\min_{x\in R^n_+} E\,[\,||\Phi(x,\omega)||\,^2] \tag{2.6.3}$$

where $E[||\Phi(x,\omega)||^2]$ is the expectation function of $||\Phi(x,\omega)||^2$.

The problem (2.6.3) is referred to as the *expected residual* (ER) model. The idea of the ER method is to minimize the expected residual of the equilibrium for each event $\omega$ by using an NCP function as introduced in Section 2.4. Here, $\Phi(x^*,\omega) = 0$ means that $x^*$ is an equilibrium solution for the event $\omega \in \Omega$. Thus, $||\Phi(x,\omega)||$ can be regarded as the "distance" from a vector $x$ to the equilibrium solution for $\omega \in \Omega$.

Previous studies on the stochastic complementarity problem made use of an NCP function, such as the "min" function and Fischer-Burmeister (FB) function, to formulate ER models [18, 19, 28, 48]. In the deterministic case, all NCP functions are equivalent in the sense that they can reformulate any complementarity problem as a system of nonlinear equations having the same solution set [18]. In the stochastic case, the situation is different, however.

The ER model with the "min" function has been studied in [18, 19, 28] for the stochastic linear complementarity problem (SLCP). In particular, it is shown that, for a class of SLCPs, if the EV model has a bounded solution set, then the ER model also has a bounded solution set, but the converse is not true in general. Moreover, if $M(\omega)$ is a stochastic $R_0$ matrix, then the ER model has a bounded solution set. Recall that a stochastic matrix $M(\cdot)$ is called a stochastic $R_0$ matrix [28] if $x \geq 0,\ M(\omega)x \geq 0,\ x^T M(\omega)x = 0$ *a.e.* imply that $x = 0$. The studies [18, 19, 28], however, do not consider the convexity of the ER model.

In Chapter 5, we establish the convexity of the ER model and apply it to the traffic equilibrium problem under uncertainty.

# Chapter 3

# The Traffic Equilibrium Problem with Nonadditive Costs and Its Monotone Mixed Complementarity Problem Formulation

## 3.1 Introduction

In the study of the traffic equilibrium problem (TEP), the researchers have presented various formulations in which many different assumptions are made to represent the "real" traffic conditions [1, 15, 21, 26]. One of the standard assumptions used is that *the route costs faced by the users in the network are* ***additive***. That is, *the route costs are simply the sum of the link costs for all the arcs on the route being considered.*

There are many situations, however, where this additivity assumption on the route costs is inappropriate. Gabriel and Bernstein [35] discussed some of the situations where nonadditive route costs occur. They claimed that almost all toll and fare schemes being implemented around the world are nonadditive. For example, the different pricing policies such as congestion pricing and the collection of emission fees add to the nonadditivity of travel costs. Moreover, different individuals have different valuations of time, which contributes to the nonadditivity of route costs.

Although nonadditivity is important in presenting a more realistic view of the traffic situation, it causes a difficulty in the analysis and computation of an equilibrium, which are usually done by formulating the TEP as the variational inequality problem (VIP).

In the last decades, the VIP has been studied extensively. Monotonicity particularly plays an important role in the existence and uniqueness of solutions of VIP. Moreover, the monotonicity is also important for solution methods for VIP to work efficiently. Most of the existing results for the VIP rely on the assumption that the function involved satisfies certain conditions such as strong or strict monotonicity [26].

A VIP equivalent to the TEP with additive costs may usually be formulated as a monotone VIP [26]. However, a VIP derived from the TEP with nonadditive costs does not immediately possess monotonicity unless restrictive assumptions are made or a certain reformulation is introduced.

Lo and Chen [50] considered a special case of the TEP with nonadditive cost functions. Specifically, they introduced a route-specific cost structure, where the route cost is assumed to be the sum of the travel time and an additional charge which is route-specific (a specific travel cost, possibly in the form of toll, is added only to a particular route in the network). This additional cost is only incurred by travelers on that route. They showed that the equivalent NCP becomes monotone. However, they reported that other users of the network (not necessarily using this route) are affected by this added route cost only when they share a common link with the route with the added cost. Moreover, the route cost function they considered was very simple, hence not so realistic. In order to solve the TEP, they converted the NCP formulation into an equivalent optimization problem by using a merit function.

Gabriel and Bernstein [35] proposed a more general route cost function. They also used some assumptions on the route costs in order to ensure monotonicity of their formulation. However, as will be shown in Section 3.2, those assumptions imply that the cost function is an affine function of time. In their work, they proposed a merit function approach to solve the NCP formulation of the TEP with nonadditive costs. Their method was based on transforming the NCP first into a problem of finding a zero of a system of nonsmooth equations. The problem can be solved by using an existing method when the NCP is monotone.

In this chapter, we modify the model presented by Gabriel and Bernstein [35] by introducing a disutility function. We show that the equivalent VIP can be transformed into a monotone MCP, and then give the existence and uniqueness results for the proposed model.

This chapter is organized as follows. In the next section, we discuss the nonadditive travel costs for the TEP. The proposed TEP and its monotone MCP reformulation are presented in Section 3.3. We also establish the existence and uniqueness results in this section. Computational results for TEPs with different disutility functions and various networks to compare our reformulation to the original VIP formulation are given in Section 3.4. We give a brief conclusion in Section 3.5.

## 3.2 Nonadditive Travel Costs

Previous studies on the TEP focused on the assumption that *the cost on route $r$ is simply the sum of the costs on each link $a$ comprising the route $r$.* Although the additivity assumption is convenient, there are various situations in which the route costs in the network are no longer additive. A particular case of a nonadditive route cost model considers both time and money in the formulation. Moreover, different individuals normally have different values for time. Hence, the additivity assumption is no longer appropriate for such a case. A detailed discussion on various situations where route costs are nonadditive can be found in [35].

Gabriel and Bernstein [35] and Larsson, et al. [46] presented two different formulations of the nonadditive route cost functions:

1. Gabriel and Bernstein [35]:
$$C_r(F) = \varphi_r\big(\sum_{a\in A}\delta_{ar}t_a(f)\big) + \eta_1\sum_{a\in A}\delta_{ar}t_a(f) + \Lambda_r(F),\ \forall r\in R_{\mathrm{w}}, \mathrm{w}\in \mathrm{W}, \tag{3.2.1}$$
where $\eta_1 > 0$ is the time-based operating costs factor (e.g., gasoline consumption), $\varphi_r$ is a function which converts time into money, and $\Lambda_r(F)$ is the route-specific financial costs (e.g., tolls) which are allowed to vary in cost according to route flows.

2. Larsson et al. [46]:
$$C_r(F) = \sum_{a\in A}\delta_{ar}t_a(f) + \phi_r(m_r),\ \forall r\in R_{\mathrm{w}}, \mathrm{w}\in \mathrm{W}, \tag{3.2.2}$$
where $m_r$ is the monetary outlay (e.g., route-specific financial cost which is allowed to vary according to route) and the function $\phi_r$ converts money into time.

In Gabriel and Bernstein [35], the route cost function is based on money ("money-based"), while in the formulation of Larsson et al. [46], the route cost is expressed in terms of time ("time-based"). There has been no clear explanation as to which formulation is better, or as to why the route costs should be represented as such. It has been noted by Bernstein and Wynter [10], however, that even if one chooses $\phi_r = \varphi_r^{-1}$ in (3.2.2), this will not make the two formulations equivalent.

We point out that, although the route cost function in Gabriel and Bernstein [35] is a general form of the route cost function, the assumptions they used to establish its monotonicity are somewhat restrictive. They assumed that there exists a function $\ell : \mathfrak{R}^{n_R} \to \mathfrak{R}$ such that $\varphi_r'(\xi_r) = \ell(\xi) \geq 0$, for all $r = 1, \ldots, n_R$, where $\xi$ is the vector of route travel times, i.e., $\xi = \Delta^T t(\Delta F)$. This assumption implies that $\ell(\xi)$ is a constant independent of $\xi$ and hence the function $\varphi_r$ must be affine. To see this, consider $\xi$ and $\overline{\xi}$ such that $\overline{\xi}_r = \xi_r$ for all $r$ except for some $\overline{r}$, and $\overline{\xi}_{\overline{r}} = \xi_{\overline{r}} + \delta$. Then $\ell(\xi) = \varphi_{\overline{r}}'(\xi_{\overline{r}}) = \ell(\overline{\xi})$. This holds for all $\delta$ and for any $\overline{r}$. Therefore, $\ell(\xi)$ must be constant, and hence, $\varphi_r(\xi)$ is affine.

In our proposed model, we will present a route cost function that can deal with both linear and nonlinear cases by introducing a particular disutility function, and show its monotonicity.

## 3.3 TEP with Disutility Functions and Its Monotone MCP Reformulation

In this section, a new formulation of the TEP with nonadditive costs that can be reformulated as a monotone MCP is proposed. The existence and uniqueness result for an equilibrium of this reformulation is then established.

### 3.3.1 TEP Model with Disutility Function

We consider a special case of the "time function" given in the form

$$T_r(F) = \sum_{a \in \mathcal{A}} \delta_{ar} t_a(f) + g_r(\Lambda_r), \ \forall r \in R_{\mathrm{w}}, \mathrm{w} \in \mathrm{W}, \tag{3.3.1}$$

where $\Lambda_r$ is the route toll (assumed to be fixed) and $g_r$ is a function that converts money into time. Next, we introduce a disutility function $U_{\mathrm{w}}$ for each OD pair $\mathrm{w} \in \mathrm{W}$.

We propose the following new route cost function:

$$C_r(F) = U_{\mathrm{w}}\big(T_r(F)\big) = U_{\mathrm{w}}\Big(\sum_{a\in\mathcal{A}} \delta_{ar} t_a(f) + g_r(\Lambda_r)\Big),\ \forall r \in R_{\mathrm{w}}, \mathrm{w} \in \mathrm{W}. \tag{3.3.2}$$

Note that when each disutility function $U_{\mathrm{w}}$ is the identity function, the route cost function (3.3.2) reduces to the route cost function (3.2.2) proposed by Larsson et al. (2002). Also, when $g_r(\Lambda_r)$ and $\Lambda_r(F)$ are absent, (3.3.2) becomes equivalent to (3.2.1) by letting $U_{\mathrm{w}}(\chi_r) = \varphi_r(\omega_r) + \eta_1\omega_r$. The model (3.2.1) may describe more realistic situations than (3.3.2). However, in (3.2.1), $\varphi_r$ must be affine to ensure the monotonicity of the equivalent MCP as pointed out in Section 3.2. We stress that the disutility function $U_{\mathrm{w}}$ includes both linear and nonlinear cases. Moreover, formulation (3.3.2) can be used to deal with the multimodal TEP where different modes (such as trucks, cars, etc.) use different disutility functions.

In what follows, we make use of (3.3.2) in order to obtain a monotone MCP reformulation of the TEP.

### 3.3.2 A Monotone MCP Reformulation

In this subsection, we present a monotone MCP equivalent to the TEP with (3.3.2). In the succeeding discussions, we assume that the functions $D_{\mathrm{w}}$, $U_{\mathrm{w}}$ and $C_r$ are continuous.

We also assume the following conditions for our purpose.

**Assumption 3.3.1.** *For all* $\mathrm{w} \in \mathrm{W}$*, the demand function* $D_{\mathrm{w}}$ *is always positive,* $U_{\mathrm{w}} : [0,\infty) \to [0,\infty)$ *is a strictly increasing function such that* $U_{\mathrm{w}}(0) = 0$ *and* $\lim_{v\to\infty} U_{\mathrm{w}}(v) = \infty$*. Also, for each* $r$*,* $T_r(F) > 0$ *for all* $F \geq 0$*, and* $g_r(\Lambda_r)$ *in (3.3.1) is nonnegative.*

Assumption 3.3.1 holds in general, since most network users would prefer the shortest travel time, and hence the disutility function is strictly increasing.

Note that Assumption 3.3.1 implies that $C_r$ defined by (3.3.2) is positive and thus we can reformulate the TEP with (3.3.2) as the following MCP($H$,$L$):

$$U_{\mathrm{w}}(T_r(F)) - u_{\mathrm{w}} \geq 0,\ F_r \geq 0,\ \big(U_{\mathrm{w}}(T_r(F)) - u_{\mathrm{w}}\big)F_r = 0,\ \forall r \in R_{\mathrm{w}}, \mathrm{w} \in \mathrm{W},$$
$$\sum_{r \in R_{\mathrm{w}}} F_r = D_{\mathrm{w}}(u),\ \forall \mathrm{w} \in \mathrm{W},$$

where

$$H(F,u) \equiv \begin{pmatrix} U(T(F)) - \Gamma u \\ \Gamma^T F - D(u) \end{pmatrix},$$

$U(T(F)) = \big(\ldots, U_{\mathrm{w}}(T_r(F)), \ldots\big)^T$, and $L = R_+^{n_R} \times R^{n_{\mathrm{W}}}$ .

**Remark 3.3.1.** *Note that from* (1.1.1) – (1.1.3)*,* $u_{\mathrm{w}}$ *above is the minimal route cost for the OD pair* $\mathrm{w} \in \mathrm{W}$*. On the other hand,* $v_{\mathrm{w}}$ *in the next proposition is the minimal time based on the "time function"* (3.3.1)*.*

However, the above MCP formulation is not monotone in general. In what follows, we reformulate MCP($H$,$L$) into an MCP with cost functions $T_r(F)$. We then show that this reformulation is monotone under appropriate conditions.

**Theorem 3.3.1.** *Suppose that Assumption 3.3.1 holds. Then MCP(H,L) is equivalent to MCP($\widetilde{H}$,L) with*

$$\widetilde{H}(F,v) \equiv \begin{pmatrix} T(F) - \Gamma v \\ \Gamma^T F - D(U(v)) \end{pmatrix} \tag{3.3.3}$$

*and* $U(v) = \big(\ldots, U_{\mathrm{w}}(v_{\mathrm{w}}), \ldots\big)^T$.

*Proof.* First we show that MCP($H$,$L$) implies MCP($\widetilde{H}$,$L$). Let $(F^*, u)$ be a solution of MCP($H$,$L$). By Assumption 3.3.1, for each $\mathrm{w} \in \mathrm{W}$ there exists a unique $v_{\mathrm{w}} \geq 0$ such that $U_{\mathrm{w}}(v_{\mathrm{w}}) = u_{\mathrm{w}}$. If $F_r^* > 0$, then $U_{\mathrm{w}}(T_r(F^*)) = u_{\mathrm{w}} = U_{\mathrm{w}}(v_{\mathrm{w}})$. Thus, $T_r(F^*) = v_{\mathrm{w}}$, and hence $\big(T_r(F^*) - v_{\mathrm{w}}\big)F_r^* = 0$. If $F_r^* = 0$, then $U_{\mathrm{w}}(T_r(F^*)) \geq u_{\mathrm{w}} = U_{\mathrm{w}}(v_{\mathrm{w}})$. Since $U_{\mathrm{w}}$ is strictly increasing, we have $T_r(F^*) \geq v_{\mathrm{w}}$ and $\big(T_r(F^*) - v_{\mathrm{w}}\big)F_r^* = 0$. Moreover, $\sum_{r \in R_{\mathrm{w}}} F_r - D_{\mathrm{w}}(u) = \sum_{r \in R_{\mathrm{w}}} F_r - D_{\mathrm{w}}\big(U(v)\big) = 0$. Therefore, $(F^*, v)$ is a solution of MCP($\widetilde{H}$,$L$).

To show that MCP($\widetilde{H}$,$L$) implies MCP($H$,$L$), let $(F^*, v)$ be a solution of MCP($\widetilde{H}$,$L$). Then, since, $F_r^* \geq 0$ and $\sum_{r \in R_{\mathrm{w}}} F_r^* = D_{\mathrm{w}}\big(U(v)\big)$, we can find, for each $\mathrm{w} \in \mathrm{W}$, a route $j_{\mathrm{w}} \in R_{\mathrm{w}}$ such that $F_{j_{\mathrm{w}}}^* > 0$. For such $j_{\mathrm{w}} \in R_{\mathrm{w}}$, $T_{j_{\mathrm{w}}}(F^*) = v_{\mathrm{w}}$. Since a route cost function is assumed to be always positive, we have $T_{j_{\mathrm{w}}}(F^*) > 0$ and $v_{\mathrm{w}} > 0$.

Let $u_{\mathrm{w}} = U_{\mathrm{w}}(v_{\mathrm{w}})$. Since $U_{\mathrm{w}}(v_{\mathrm{w}}) > 0$, we have $u_{\mathrm{w}} > 0, \forall \mathrm{w} \in \mathrm{W}$. To complete the proof, we need to show that $U_{\mathrm{w}}(T_r(F^*)) - u_{\mathrm{w}} \geq 0$ and $\big(U_{\mathrm{w}}(T_r(F^*)) - u_{\mathrm{w}}\big)F_r^* = 0$ for all $r \in R_{\mathrm{w}}$.

Now, suppose $F_r^* > 0$. Then $T_r(F^*) = v_{\mathrm{w}}$. This implies that $U_{\mathrm{w}}(T_r(F^*)) = U_{\mathrm{w}}(v_{\mathrm{w}}) = u_{\mathrm{w}}$. Hence, $U_{\mathrm{w}}(T_r(F^*)) = u_{\mathrm{w}}$ and $\big(U_{\mathrm{w}}(T_r(F^*)) - u_{\mathrm{w}}\big)F_r^* = 0$. If $F_r^* = 0$, then $T_r(F^*) \geq v_{\mathrm{w}}$ and $U_{\mathrm{w}}(T_r(F^*)) \geq U_{\mathrm{w}}(v_{\mathrm{w}})$. Thus, $U_{\mathrm{w}}(T_r(F^*)) - u_{\mathrm{w}} \geq 0$ and $\big(U_{\mathrm{w}}(T_r(F^*)) - u_{\mathrm{w}}\big)F_r^* = 0$. Consequently, $(F^*, u)$ is a solution of MCP($H$,$L$). □

Having shown that MCP($H$,$L$) is equivalent to MCP($\widetilde{H}$,$L$), in the succeeding discussions we focus our attention to MCP($\widetilde{H}$,$L$). Note that MCP($H$,$L$) is not monotone in general. However, we can show that under the following additional assumption the MCP($\widetilde{H}$,$L$) becomes monotone.

**Assumption 3.3.2.** *There exists a nonincreasing function $d_w : \Re \to \Re$ such that $D_{\mathrm{w}}(u) = d_{\mathrm{w}}(u_{\mathrm{w}})$ for each $\mathrm{w} \in \mathrm{W}$. Moreover, $t(f) = (\ldots, t_a(f), \ldots)^T$ is monotone on $f$.*

Assumption 3.3.2 means that $D_{\mathrm{w}}$ is a nonincreasing function of $u_{\mathrm{w}}$ only for each $\mathrm{w} \in \mathrm{W}$.

**Theorem 3.3.2.** *Suppose that Assumptions 3.3.1 and 3.3.2 hold. Then MCP($\widetilde{H}$,L) is monotone.*

*Proof.* From Assumption 3.3.2 and the fact that $T(F) = \Delta t(\Delta F) + (\ldots, g_r(\Lambda_r), \ldots)^T$, $T$ is monotone. Also, since $d_{\mathrm{w}}$ is a nonincreasing function and $D_{\mathrm{w}}(u) = d_{\mathrm{w}}(u_{\mathrm{w}})$ for each $\mathrm{w} \in \mathrm{W}$ from Assumption 3.3.2, it follows that $-D\big(U(v)\big)$ is monotone. For any $(F_1, v_1)^T$,

$(F_2, v_2)^T \in \Re^{n_R} \times \Re^{n_W}$, we have

$$
\begin{aligned}
&\left(\widetilde{H}(F_1, v_1) - \widetilde{H}(F_2, v_2)\right)^T \left( \begin{pmatrix} F_1 \\ v_1 \end{pmatrix} - \begin{pmatrix} F_2 \\ v_2 \end{pmatrix} \right) \\
= \;& \Big(T(F_1) - T(F_2)\Big)^T (F_1 - F_2) - \Big(\Gamma v_1 - \Gamma v_2\Big)^T (F_1 - F_2) \\
& + \Big(\Gamma^T (F_1 - F_2)\Big)^T (v_1 - v_2) - \Big(D(U(v)) - D(U(v))\Big)^T (v_1 - v_2) \\
= \;& \Big(T(F_1) - T(F_2)\Big)^T (F_1 - F_2) - \Big(D(U(v)) - D(U(v))\Big)^T (v_1 - v_2) \geq 0,
\end{aligned}
$$

where the last inequality follows from the monotonicity of $T(F)$ and $-D(U(v))$. Hence MCP($\widetilde{H}$,$L$) is monotone. $\square$

Using a similar argument, we can show the following result.

**Corollary 3.3.3.** *If $D_{\rm w}(u)$ is constant for each ${\rm w} \in {\rm W}$, then MCP($\widetilde{H}$,$L$) is monotone.*

### 3.3.3 Existence and Uniqueness Results

In this subsection, we present some existence and uniqueness results for our proposed model (3.3.2).

The first result ensures that MCP($\widetilde{H}$,$L$) has a solution, i.e., our model has an equilibrium. To prove it, we make use of a result by Facchinei and Pang (2003).

**Assumption 3.3.3.** *The function $C_r$ defined by (3.3.2) is nonnegative, and $D_{\rm w}$ is bounded above on the set $\{u \in \Re^{n_{\rm W}} | u > 0\}$.*

**Theorem 3.3.4.** *Suppose Assumption 3.3.3 holds. Then MCP($\widetilde{H}$,$L$) has a nonempty bounded solution set. Moreover, if Assumptions 3.3.1 and 3.3.2 hold, the set of solutions is convex.*

*Proof.* Since $C_r$ is nonnegative and $D_w$ is bounded above by Assumption 3.3.3, it follows from Proposition 2.2.14 in [26] that MCP($\widetilde{H}$,$L$) has a solution. Next we show that the solution set is bounded. Let $S$ be its solution set and let $S_F = \{F | (F, v) \in S\}$ and $S_v = \{v | (F, v) \in S\}$.

Since $\sum_{r\in R_{\rm w}} F_r = D_{\rm w}(U(v))$, $\forall {\rm w} \in {\rm W}$ for all $(F, v) \in S$ and, by assumption, the demand function $D$ is bounded, it follows that $S_F$ is bounded. Moreover, we note that $T_r(F) \geq v_{\rm w} \geq 0$, $\forall r \in R_{\rm w}, {\rm w} \in {\rm W}$, from Assumption 3.3.1 and the definitions of MCP($\widetilde{H}$,$L$). Hence, $S_v$ is bounded since $S_F$ is bounded. Thus, the solution set $S$ of MCP($\widetilde{H}$,$L$) is bounded.

Suppose that Assumptions 3.3.1 and 3.3.2 hold. Since MCP($\widetilde{H}$,$L$) is monotone by Theorem 3.3.2, it follows that the set of solutions is convex [26]. □

Next we show that the set of solutions of MCP($\widetilde{H}$,$L$) is a singleton under the following assumption together with Assumption 3.3.1.

**Assumption 3.3.4.** *There exists a strictly decreasing function $d_{\rm w} : \mathfrak{R} \to \mathfrak{R}$ such that $D_{\rm w}(u) = d_{\rm w}(u_{\rm w})$ for each ${\rm w} \in W$. Moreover, $T(F) = (\ldots, T_r(F), \ldots)^T$ is a strictly monotone function.*

**Remark 3.3.2.** *Assumption 3.3.4 on $T$ holds when $\nabla t(f)$ is positive definite for all $f$, where $t(f) = (\ldots, t_a(f), \ldots)^T$, and the rank of the link-route incidence matrix $\Delta$ is $n_R$. This is because $T(F) = \Delta t(\Delta F) + (\ldots, g_r(\Lambda_r), \ldots)^T$ so that $\nabla_F T(F) = \Delta \nabla t(\Delta F)\Delta^T$ is positive definite.*

Under Assumption 3.3.4, both $T$ and $-D(U(\cdot))$ are strictly monotone.

**Theorem 3.3.5.** *Suppose that Assumptions 3.3.1, 3.3.4 and 3.3.3 hold. Then MCP($\widetilde{H}$,L) has a unique solution.*

*Proof.* It follows from Theorem 3.3.4 that MCP($\widetilde{H}$,$L$) has a solution. To show that this solution is unique, let $x_1 = (F_1^T, v_1^T)^T$ and $x_2 = (F_2^T, v_2^T)^T$ be two solutions of MCP($\widetilde{H}$,$L$). Since $x_1$, $x_2$, $\widetilde{H}(x_1)$ and $\widetilde{H}(x_2)$ are nonnegative, from the complementarity conditions $x_1^T\widetilde{H}(x_1) = 0$ and $x_2^T\widetilde{H}(x_2) = 0$, we have

$$(x_1 - x_2)^T\big(\widetilde{H}(x_1) - \widetilde{H}(x_2)\big) \leq 0.$$

From the definition (3.3.3) of $\widetilde{H}$ and $x$, the above inequality can be rewritten as

$$(F_1 - F_2)^T \big(T(F_1) - \Gamma v_1 - T(F_2) + \Gamma v_2\big) + (v_1 - v_2)^T \big(\Gamma^T F_1 - D(U(v_1)) - \Gamma^T F_2 + D(U(v_2))\big) \leq 0,$$

which implies that

$$(F_1 - F_2)^T \big(T(F_1) - T(F_2)\big) + (v_1 - v_2)^T \big( - D(U(v_1)) + D(U(v_2))\big) \leq 0. \qquad (3.3.4)$$

Since $T$ and $-D_w(U(\cdot))$ are strictly monotone from Assumption 3.3.4, the inequality (3.3.4) implies that $F_1 = F_2$ and $v_1 = v_2$. Therefore, the solution set is a singleton. □

## 3.4 Numerical Results

Figure 3.1: The 7-link Network A.

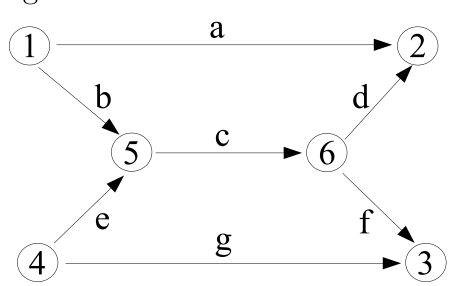
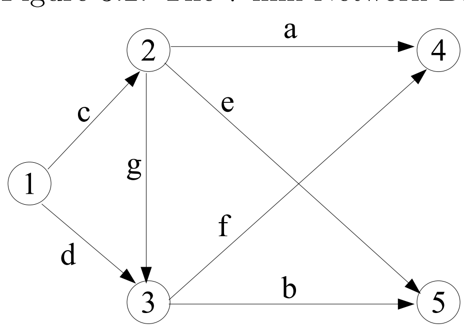



Figure 3.2: The 7-link Network B.




In this section, we present our computational results. Under Assumption 3.3.1, we can easily verify that NCP($H$) and NCP($\widetilde{H}$) are equivalent to MCP($H$,$L$) and MCP($\widetilde{H}$,$L$),

Figure 3.3: The 11-link Network.

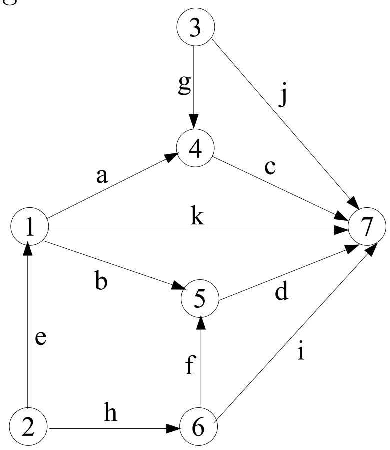



Table 3.1: Network routes and OD pairs.

| Network | OD pair | Route |
|---|---|---|
| 7-link A | 1-2 | 1 = {a}, 2 = {b,c,d} |
| | 1-3 | 3 = {b,c,f} |
| | 4-2 | 4 = {c,d,e} |
| | 4-3 | 5 = {c,e,f}, 6 = {g} |
| 7-link B | 1-4 | 1 = {c,f,g}, 2 = {a,c}, 3 = {d,f} |
| | 1-5 | 4 = {b,c,g}, 5 = {c,e}, 6 = {b,d} |
| 11-link | 1-7 | 1 = {a,c}, 2 = {k}, 3 = {b,d} |
| | 2-7 | 4 = {h,i}, 5 = {b,d,e}, 6 = {e,k}, 7 = {a,c,e} |
| | 3-7 | 8 = {j}, 9 = {c,g} |
| | 6-7 | 10 = {i}, 11 = {d,f} |

respectively. In our numerical experiments, we try to obtain an equilibrium solution of the TEP with (3.3.2) by solving NCP($H$) and NCP($\widetilde{H}$) instead of MCP($H$,$L$) and MCP($\widetilde{H}$,$L$). To solve NCPs, we use the Generalized Newton Method (GNM) of Jiang [43]. (See Chapter 2 for the detail).

Table 3.2: Coefficients of the demand function of the 7-link Network A for the 2-mode case.

| Coefficients of the demand function | OD pair | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | MODE A | | | | MODE B | | | |
| | 1-2 | 1-3 | 4-2 | 4-3 | 1-2 | 1-3 | 4-2 | 4-3 |
| $b_{\mathrm{w}}^1$ | 400 | 400 | 400 | 400 | 400 | 400 | 400 | 400 |
| $b_{\mathrm{w}}^2$ | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |

Table 3.3: Coefficients of the demand function for the single-mode case.

| Network | Coefficients of the demand function | |
|---|---|---|
| | $b_{\mathrm{w}}^1$ | $b_{\mathrm{w}}^2$ |
| 7-link A | 600 | 0.04 |
| | 500 | 0.03 |
| | 500 | 0.05 |
| | 400 | 0.05 |
| 7-link B | 200 | 0.2 |
| | 220 | 0.2 |
| 11-link | 600 | 0.04 |
| | 500 | 0.03 |
| | 500 | 0.05 |
| | 400 | 0.05 |

The main reason for choosing the GNM is due to the fact that our proposed model satisfies monotonicity properties, and the GNM has nice convergence properties under these conditions. There are other methods such as the PATH solver (Cao and Ferris [13], Dirkse and Ferris [22]) that can be used to solve the NCPs however.

Table 3.4: Coefficients of the link cost function of the 7-link Network A for the 2-mode case.

| Coefficients of the link cost function | Link | | | | | | |
|---|---|---|---|---|---|---|---|
| | a | b | c | d | e | f | g |
| $c_{\mathrm{w}}^1$ | 60 | 10 | 5 | 8 | 12 | 5 | 70 |
| $c_{\mathrm{w}}^2$ | 200 | 300 | 700 | 300 | 300 | 300 | 200 |

The numerical experiments consist of two parts. In the first part we check the validity of our model by comparing it with the traditional model with additive costs. Here, we use a network with two transportation modes. Our model uses different disutility functions for the different transportation modes.

In the second part of the experiments, we aim to find a solution for the two NCP formulations, namely, NCP($H$) and NCP($\widetilde{H}$), in order to compare the two formulations.

The coding was done in Matlab 6.5. In our experiments, we used three different sample networks (Figures 3.1, 3.2 and 3.3). The network shown in Figure 3.1 is taken from [15], the one shown in Figure 3.2 is taken from [67] and the one in Figure 3.3 is taken from [68].

The routes and OD pairs are given in Table 3.1. The demand function used is

$$D_{\mathrm{w}}(u_{\mathrm{w}}) = -b_{\mathrm{w}}^1(exp(-b_{\mathrm{w}}^2 u_{\mathrm{w}})),$$

where $b_{\mathrm{w}}^1$ and $b_{\mathrm{w}}^2$ are given in Table 3.2 (for the 2-mode case) and Table 3.3 (for the single-mode case). The link cost function used is

$$t_a(f) = c_{\mathrm{w}}^1(1.0 + 0.15(f/c_{\mathrm{w}}^2)^4),$$

where $c_{\mathrm{w}}^1$ and $c_{\mathrm{w}}^2$ are given in Table 3.4 (for the 2-mode case) and Table 3.5 (for the single-mode case).

Table 3.5: Coefficients of the link cost function for the single-mode case.

| Network | Link | Coefficients of the Link Cost Function | |
|---|---|---|---|
| | | $c_{\mathrm{w}}^1$ | $c_{\mathrm{w}}^2$ |
| 7-link A | a | 60 | 50 |
| | b | 10 | 2 |
| | c | 5 | 6 |
| | d | 8 | 2 |
| | e | 12 | 13 |
| | f | 5 | 6 |
| | g | 70 | 60 |
| 7-link B | a | 6 | 15 |
| | b | 4 | 15 |
| | c | 3 | 30 |
| | d | 5 | 30 |
| | e | 6 | 15 |
| | f | 4 | 15 |
| | g | 1 | 15 |
| 11-link | a | 6 | 200 |
| | b | 5 | 200 |
| | c | 6 | 200 |
| | d | 7 | 200 |
| | e | 6 | 100 |
| | f | 1 | 100 |
| | g | 5 | 150 |
| | h | 10 | 150 |
| | i | 11 | 200 |
| | j | 11 | 200 |
| | k | 15 | 200 |

Table 3.6: Route flows of 7-link Network A for the 2-mode case.

| Route Flow | MODEL | |
|---|---|---|
| | Traditional | Proposed |
| $F_1^A$ | 0.0000 | 0.0000 |
| $F_2^A$ | 75.8216 | 76.8721 |
| $F_3^A$ | 101.9756 | 103.2007 |
| $F_4^A$ | 144.9559 | 146.1842 |
| $F_5^A$ | 104.7306 | 105.5160 |
| $F_6^A$ | 0.0000 | 0.0000 |
| $F_1^B$ | 0.0000 | 0.0000 |
| $F_2^B$ | 75.8216 | 72.8016 |
| $F_3^B$ | 101.9756 | 99.4810 |
| $F_4^B$ | 144.9559 | 143.2517 |
| $F_5^B$ | 104.7306 | 101.8342 |
| $F_6^B$ | 0.0000 | 0.0000 |

### 3.4.1 Comparison of the Proposed Model and the Traditional Model

We have tested the validity of our proposed formulation. In this experiment, we compare our proposed model to that of the traditional model on the 7-link Network A with two transportation modes.

In this experiment, we suppose that both modes have the same OD pairs (Table 3.1), set of routes (Table 3.1) and demand functions (Table 3.2). For the traditional model, we use the same route cost functions for both modes A and B, that is, $C_r(F) = T_r(F)$. On the other hand, for the proposed model we use different route cost functions for mode A and for

Table 3.7: Residuals when $U_{\mathrm{w}}(T_r(F)) = (T_r(F))^2$.

| NETWORK | initial point of GNM | RESIDUAL | |
|---|---|---|---|
| | | NCP($H$) | NCP($\widetilde{H}$) |
| 7-link A | (0,0,0,0,0,0,0,0,0,0) | 7.047E-003 | 1.1479E-005 |
| | (1,1,1,1,1,1,1,1,1,1) | 6.645E-003 | 1.7912E-005 |
| | (10,10,10,10,10,10,10,10,10,10) | 3.499E-003 | 2.5162E-005 |
| 7-link B | (0,0,0,0,0,0,0,0) | 2.1822E-009 | 8.0737E-006 |
| | (1,1,1,1,1,1,1,1) | 1.9485E-009 | 2.9617E-008 |
| | (10,10,10,10,10,10,10,10) | 1.7703E-009 | 2.9615E-008 |
| 11-link | (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0) | 6.5077E-009 | 1.4218E-006 |
| | (1,1,1,1,1,1,1,1,1,1,1,1,1,1,1) | 9.8253E-007 | 2.9473E-009 |
| | (10,10,10,10,10,10,10,10,10,10,10,10,10,10,10) | 7.5163E-007 | 6.5297E-007 |

mode B, that is, $C_r^A(F) = T_r(F)$ and $C_r^B(F) = T_r(F) + 0.001(T_r(F))^2$, respectively.

The results are shown in Table 3.6. In the table, $F_r^A, r = 1, \ldots, 6$, stand for the route flows corresponding to mode A, and $F_r^B, r = 1, \ldots, 6$, stand for the route flows corresponding to mode B. The results show that, compared to the route flows for the traditional TEP model, there is a significant difference in the route flows of the two modes for our proposed model. As expected, the routes with lower travel costs (i.e., lower disutility function values) have higher route flows (in the case of mode A), while routes with higher disutility function values have lesser flows (in the case of mode B).

### 3.4.2 Comparison of NCP($H$) and NCP($\widetilde{H}$) Formulations

We have also compared the NCP formulations of the TEP, namely, NCP($H$) and NCP($\widetilde{H}$). The networks are tested using nonlinear link cost functions, an elastic demand function and

Table 3.8: Residuals when $U_{\mathrm{w}}(T_r(F)) = T_r(F) + 0.01(T_r(F))^2$.

| NETWORK | initial point of GNM | RESIDUAL | |
|---|---|---|---|
| | | NCP($H$) | NCP($\widetilde{H}$) |
| 7-link A | (0,0,0,0,0,0,0,0,0,0) | 3.9814E+004 | 6.1748E-013 |
| | (1,1,1,1,1,1,1,1,1,1) | 3.8881E+004 | 4.3999E-014 |
| | (10,10,10,10,10,10,10,10,10,10) | 1.2173E-008 | 1.7146E-011 |
| 7-link B | (0,0,0,0,0,0,0,0) | 8.7901E-006 | 5.0343E-009 |
| | (1,1,1,1,1,1,1,1) | 9.8626E-007 | 4.2881E-010 |
| | (10,10,10,10,10,10,10,10) | 9.1695E-007 | 1.1765E-012 |
| 11-link | (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0) | 3.6525E-006 | 1.5524E-006 |
| | (1,1,1,1,1,1,1,1,1,1,1,1,1,1,1) | 3.6775E-006 | 1.0126E-008 |
| | (10,10,10,10,10,10,10,10,10,10,10,10,10,10,10) | 2.0176E-006 | 1.6949E-008 |

various disutility functions. Here we introduce two disutility functions, namely,

(i) $U_{\mathrm{w}}(T_r(F)) = (T_r(F))^2$; and

(ii) $U_{\mathrm{w}}(T_r(F)) = T_r(F) + 0.01(T_r(F))^2$

for the route cost functions on each network.

The computational results are shown in Tables 3.7 and 3.8. In these tables, "NETWORK" stands for the sample network used, the columns NCP($H$) and NCP($\widetilde{H}$) under "RESIDUAL" respectively show the values of the residuals for the two NCP formulation. The residual is defined as $r(x) = |x^T H(x)| + \sum_{i=1}^{n_R+n_W} \min\{0, x_i\} + \sum_{i=1}^{n_R+n_W} \min\{0, H_i(x)\}$ and it is computed in order to evaluate the quality of the solutions. Therefore, the residuals should be as small as possible; a value very close to zero is ideal.

We have also tested our proposed reformulation for the case where there are two different

Table 3.9: Residuals for the 7-link Network A for the case when there are 2 modes of transportation using the routes in the network.

| Disutility Function | initial point of GNM | RESIDUAL | |
|---|---|---|---|
| | | NCP($H$) | NCP($\widetilde{H}$) |
| $U_{\mathrm{w}}(T_r(F)) = (T_r(F))^2$ | (0,0,0,0,0,0,0,0,0,0) | 2.3817E-002 | 2.2658E-006 |
| | (1,1,1,1,1,1,1,1,1,1) | 8.8063E-004 | 3.2301E-005 |
| | (10,10,10,10,10,10,10,10,10,10) | 4.2729E-002 | 3.3028E-005 |
| $U_{\mathrm{w}}(T_r(F)) = T_r(F) + 0.01(T_r(F))^2$ | (0,0,0,0,0,0,0,0,0,0) | 1.9057E+004 | 7.9972E-009 |
| | (1,1,1,1,1,1,1,1,1,1) | 1.8794E+004 | 7.9962E-009 |
| | (10,10,10,10,10,10,10,10,10,10) | 9.1208E-010 | 7.9978E-009 |

transportation modes in the network. In this example, we use Network A (Figure 3.1). The results for this case are shown in Table 3.9.

In both cases, the results reveal that our proposed reformulation NCP($\widetilde{H}$) successfully yields an equilibrium of the original TEP as evident by the computed residual for each formulation (see for example, in Table 3.8 for the 7-link A Network and Table 3.9 for $U_{\mathrm{w}}(T_r(F)) = T_r(F) + 0.01(T_r(F))^2$ ). However, we have a difficulty in obtaining an equilibrium of the TEP by solving NCP($H$) as it lacks the monotonicity.

## 3.5 Conclusions

In this chapter, we have formulated the TEP with nonadditive route costs by introducing a disutility function, then presented its monotone MCP reformulation. For this reformulation, we have established the existence and uniqueness of the equilibrium of the proposed model. Moreover, we have shown through numerical experiments that our new MCP reformulation is useful in identifying an equilibrium of the TEP. We note that, aside from GNM, there are other methods (e.g. PATH solver) available for solving NCPs.

# Chapter 4

# An MPEC Model of the Road Pricing Problem with Nonadditive Route Costs

## 4.1 Introduction

In recent years, the modern economic growth has caused traffic congestion problems around the world. Means to solve such a problem have become the focus of attention of most traffic planners and researchers. Road pricing is considered one of the effective means and it has been introduced in major highways of most countries [12, 53, 55, 60, 66]. In most cases, tolls are imposed in order to reduce network congestion since it causes an increase in the transportation cost of some links of the network and thus may result in the reduction of transportation demand for certain routes in the network, or force road users to change their travel routes or travel schedules. Another reason for collecting tolls is to cover for the maintenance costs of the network or to compensate for the social and environmental damage (such as pollution) that may be brought about by the use of the road network [29, 53, 68].

Studies on road pricing consider a bilevel model wherein the traffic planner is assigned as "the leader" (upper-level decision maker) while the traffic users are called "the followers" (lower-level decision makers). Here, the leader makes some actions in order to achieve his goal (e.g., collects toll in order to alleviate traffic congestion), while the followers react to the actions of the leader by changing their behaviors (e.g., varying their travel schedules, route choices or travel modes) according to the traffic equilibrium principle. The bilevel model can be formulated as a mathematical program with equilibrium constraints (MPEC) [52], which is a constrained optimization problem whose constraints are defined by a parametric variational inequality or complementarity system.

MPEC has been studied extensively in the last decade. MPECs however are known to be quite difficult to handle due to its complexities [52]. Such difficulty in handling MPECs arises from the fact that its feasible region is in general nonconvex and nonsmooth. Many methods have already been proposed in order to solve MPECs. Such solution methods include smoothing approach [17, 24, 52], penalty approach [49, 63] and implicit programming approach (ImPA) [23, 52]. The ImPA is known to be useful when the lower level problem has a unique solution for every upper level variable. In this chapter, we apply the ImPA to our bilevel optimization problem.

Previous studies on road pricing usually assumed the additive route costs [12, 60, 67], that is, *the route costs are simply the sum of the link costs for all the arcs on the route being considered*. Although the additivity assumption is convenient, there are various situations in which the route costs in the network are no longer additive. Some of these situations are presented in Section 1.2 as discussed by Gabriel and Bernstein in [35].

In the study of road pricing on the traffic equilibrium problem (TEP), it is indeed important to consider nonadditivity in order to present a more realistic view of the traffic situation. However, it causes a difficulty in the analysis and computation of an equilibrium, since they are usually done by formulating the TEP into the equivalent mixed complementarity problem (MCP). The TEP with additive costs may be formulated as a monotone MCP having a unique solution [26]. However, an MCP derived from the TEP with nonadditive costs does not immediately possess monotonicity unless restrictive assumptions are made or a certain reformulation is introduced.

In this chapter, we consider the road pricing formulation for the TEP with nonadditive costs. We make use of the results in Chapter 3 in order to transform the road pricing formulation into an MPEC. Since the proposed MPEC model has a strictly monotone lower level constraints as will be shown later in this Chapter, we employ the ImPA for solving it.

This chapter is organized as follows. In the next section, the MPEC model of the TEP and its reformulations are presented. We then introduce an implicit programming formulation for the MPEC and show its properties in Section 4.3. Numerical examples are given in Section 4.4. We give a brief conclusion in Section 4.5.

## 4.2 Mathematical Program with Monotone MCP Constraints Model of the Road Pricing Problem with Nonadditive Route Costs

In this section, we consider a road pricing model on the TEP with nonadditive route costs. We show that this model can be reformulated as a mathematical problem with monotone MCP constraints.

Let $\widetilde{C}(\tau, F)$ be the vector of nonadditive route costs with a given toll $\tau \in \Re^{n_{\mathcal{A}}}$. Then the route flow must satisfy the following MCP:

$$\begin{aligned} &0 \le \widetilde{C}(\tau, F) - \Gamma u \perp F \ge 0, \\ &\Gamma^T F - D(u) = 0. \end{aligned} \tag{4.2.1}$$

Let $S(\tau)$ be the solution set of MCP (4.2.1).

We consider the situation where the traffic planner aims to minimize an objective function $\theta(\tau, F)$ by choosing an optimal toll $\tau$. Various types of objective function can be used. For example,

1. to maximize the total revenue, we set

$$\theta(\tau, F) = -\sum_{a \in \mathcal{A}} \tau_a f_a, \tag{4.2.2}$$

   where $f_a$ is the flow on link $a \in \mathcal{A}$ which is determined by $F \in \Re^{n_R}$;

2. to minimize the total travel cost, we set

$$\theta(\tau, F) = \sum_{\mathrm{w} \in \mathrm{W}} \sum_{r \in R_{\mathrm{w}}} F_r \widetilde{C}_r(\tau, F). \tag{4.2.3}$$

Then the road pricing problem is formulated as the following MPEC:

$$\begin{aligned} \min \quad & \theta(\tau, F) \\ \text{s.t.} \quad & \tau \in \mathbf{T}, \\ & (F, u) \in S(\tau), \end{aligned} \tag{4.2.4}$$

where $\mathbf{T}$ is the set of possible tolls imposed on all tollable arcs in the network (for example, $\mathbf{T} = \{\tau : \tau_a \geq 0, a \in \mathcal{A}\}$).

When MCP (4.2.1) is monotone, a number of methods, such as ImPA, are available to solve MPEC (4.2.4). However, when the route cost function $C_r$ is nonadditive (which usually happens in the real world), MCP (4.2.1) is not necessarily monotone and hence such methods may not be applied directly. Below, we consider the nonadditive route cost function $C_r$ introduced in Chapter 3, and show that MCP (4.2.1) can be reformulated as a monotone MCP.

First, we consider a special case of the "time function" given in the form

$$T_r(\tau, F) = \sum_{a \in \mathcal{A}} \delta_{ar} t_a(\Delta F) + g_r(\tau), \ \forall r \in R_{\mathrm{w}}, \mathrm{w} \in \mathrm{W}. \tag{4.2.5}$$

By introducing a disutility function $U_{\mathrm{w}}$ for each OD pair $\mathrm{w} \in \mathrm{W}$, we define the following nonadditive route cost function

$$C_r(\tau, F) = U_{\mathrm{w}}\big(T_r(\tau, F)\big) = U_{\mathrm{w}}\Big(\sum_{a \in \mathcal{A}} \delta_{ar} t_a(\Delta F) + g_r(\tau)\Big), \ \forall r \in R_{\mathrm{w}}, \mathrm{w} \in \mathrm{W}. \tag{4.2.6}$$

This route cost function can deal with both the linear and nonlinear cases. For example, one may consider a nonlinear disutility function of the form $U_{\mathrm{w}}(t) = t + 0.01t^2$.

Using the route cost function (4.2.6), we reformulate the MCP (4.2.1). To this end, we make the following assumption.

**Assumption 4.2.1.** *The functions $D_{\mathrm{w}}$, $U_{\mathrm{w}}$ and $C_r$ are continuous. Also, for each $\mathrm{w} \in \mathrm{W}$, the demand function $D_{\mathrm{w}}$ is positive, and $U_{\mathrm{w}} : [0, \infty) \to [0, \infty)$ is a strictly increasing function such that $U_{\mathrm{w}}(0) = 0$ and $\lim_{v \to \infty} U_{\mathrm{w}}(v) = \infty$. Moreover, for each $r$, $T_r(\tau, F) > 0$ for all $F \geq 0$, and $g_r(\tau)$ in (4.2.5) is nonnegative for all $\tau \in \boldsymbol{T}$.*

The above assumption means that the disutility function is strictly increasing. It further implies that $C_r$ defined by (4.2.6) is positive. We then rewrite MCP (4.2.1) with the route cost function (4.2.6) as the following MCP:

$$\begin{aligned} &0 \leq\ U(T(\tau, F)) - \Gamma u \perp F\ \geq 0, \\ &\Gamma^T F - D(u) = 0, \end{aligned} \tag{4.2.7}$$

where $U(T(\tau, F)) = (\ldots, U_{\mathrm{w}}(T_r(\tau, F)), \ldots)^T$. Note, however, that MCP (4.2.7) is not monotone in general.

By the equivalence between (4.2.1) and (4.2.7), we can rewrite MPEC (4.2.4) as

$$\begin{aligned} \min \quad & \theta(\tau, F) \\ \text{s.t.} \quad & \tau \in \mathbf{T} \\ & (U(T(\tau, F)) - \Gamma u)^T F = 0 \\ & \Gamma^T F - D(u) = 0 \\ & U(T(\tau, F)) - \Gamma u \geq 0 \\ & F \geq 0. \end{aligned} \tag{4.2.8}$$

We may solve the above problem using a general optimization method, such as Sequential Quadratic Programming (SQP), or some existing solution methods for MPEC [52]. However, some efficient methods, such as the ImPA, cannot be applied since MCP (4.2.7), which is involved in the constraints of (4.2.8), is not monotone in general.

In Chapter 3, it was shown that under Assumption 3.3.1, MCP (4.2.7) is equivalent to the following MCP:

$$\begin{aligned} &0 \leq\ T(\tau, F) - \Gamma v \perp F\ \geq 0, \\ &\Gamma^T F - D(U(v)) = 0, \end{aligned} \tag{4.2.9}$$

where $U(v) = (\ldots, U_{\mathrm{w}}(v_{\mathrm{w}}), \ldots)^T$ and $T(\tau, F) = (\ldots, T_r(\tau, F), \ldots)^T$. The function $T_r(\tau, F)$ is given by (4.2.5). Moreover, we showed in Theorem 3.3.2 that MCP (4.2.9) is monotone under the following additional assumption.

**Assumption 4.2.2.** *There exist a nonincreasing function* $d_{\mathrm{w}} : \Re \to \Re$ *and a strictly increasing function* $\bar{t}_a : \Re \to \Re$ *such that* $D_{\mathrm{w}}(u) = d_{\mathrm{w}}(u_{\mathrm{w}})$ *for each* $\mathrm{w} \in \mathrm{W}$ *and* $t_a(f) = \bar{t}_a(f_a)$ *for each* $a \in \mathcal{A}$*.*

Thus, we obtain the following optimization problem with monotone MCP constraints equivalent to MPEC(4.2.8):

$$\begin{array}{ll} \min & \theta(\tau, F) \\ \text{s.t.} & \tau \in \mathbf{T} \\ & (T(\tau, F) - \Gamma v)^T F = 0 \\ & \Gamma^T F - D(U(v)) = 0 \\ & T(\tau, F) - \Gamma v \geq 0 \\ & F \geq 0. \end{array} \tag{4.2.10}$$

## 4.3 An Implicit Programming Approach (ImPA)

In this section we present an implicit programming approach (ImPA) for solving MPEC (4.2.10) given in the previous section. The ImPA has already been studied in the literature [52]. However, its applicability is somewhat limited since the uniqueness of a solution of the lower level problem is required. We will show that the lower level problem of MPEC (4.2.10) satisfies one of the required monotonicity properties, and hence the ImPA can be applied to our problem.

First, we give sufficient conditions for the uniqueness of a solution of MCP (4.2.9).

**Assumption 4.3.1.** *There exist a strictly decreasing function $d_{\mathrm{w}} : \Re \to \Re$ such that $D_{\mathrm{w}}(u) = d_{\mathrm{w}}(u_{\mathrm{w}})$ for each $\mathrm{w} \in \mathrm{W}$. Moreover, $T(\tau, F) = (\ldots, T_r(\tau, F), \ldots)^T$ is a strictly monotone function with respect to $F$.*

**Remark 4.3.1.** *Assumption 4.3.1 holds when $\nabla t(\Delta F)$ is positive definite for all $F$, where $t(\Delta F) = (\ldots, t_a(\Delta F), \ldots)^T$, and the rank of the link-route incidence matrix $\Delta$ is $n_R$. This is because $T(\tau, F) = \Delta t(\Delta F) + (\ldots, g_r(\tau), \ldots)^T$ so that $\nabla_F T(\tau, F) = \Delta \nabla t(\Delta F) \Delta^T$ is positive definite.*

**Remark 4.3.2.** *When $T(\tau, F)$ is merely monotone, we may consider a regularized function $\widetilde{T}(\tau, F) := T(\tau, F) + \varepsilon F$ as a "time function" (4.2.5), where $\varepsilon$ is a positive constant. The* *function $\widetilde{T}(\tau, F)$ is strictly monotone. Moreover, since a route flow $F$ is bounded, a term $\varepsilon F$ is vanishingly small, provided $\varepsilon$ is sufficiently small. Moreover, we showed that MCP has a unique solution if the function $C_r$ defined by (4.2.6) is nonnegative, and $D_w$ is bounded above on the set $\Re_{++}^{n_W} = \{u \in \Re^{n_W} : u > 0\}$.*

It follows from Theorem 3.3.5 that for each $\tau \in \mathbf{T}$, the solution set $\widetilde{S}(\tau)$ of MCP (4.2.9) is a singleton. Let $(F(\tau), v(\tau))$ be the unique element of $\widetilde{S}(\tau)$.

In the succeeding discussions, we let $y \equiv (F, v)$, $y(\tau) \equiv (F(\tau), v(\tau))$, and $\widetilde{\theta}(\tau) \equiv \theta(\tau, y(\tau))$. Then MPEC (4.2.10) can be written as the following optimization problem in the upper-level variable $\tau$ alone:

$$\begin{aligned} \min \quad & \widetilde{\theta}(\tau) \\ \text{s.t} \quad & \tau \in \mathbf{T}. \end{aligned} \tag{4.3.1}$$

This problem can be solved using existing optimization solvers. Such an optimization method usually requires the differentiability of the objective function, particularly near the solution. From now on, our aim is to investigate the differentiability of the function $y(\tau)$. To begin with, consider the reformulation on MCP (4.2.9) using the *min* function.

Let $H_{min}$ be defined by

$$H_{min}(\tau, y) = \begin{pmatrix} \min\left\{F_1, \widetilde{H}_1(\tau, y)\right\} \\ \vdots \\ \min\left\{F_{n_R}, \widetilde{H}_{n_R}(\tau, y)\right\} \\ \Gamma^T F - D(U(v)) \end{pmatrix}, \tag{4.3.2}$$

where

$$\widetilde{H}(\tau, y) \equiv T(\tau, F) - \Gamma v. \tag{4.3.3}$$

Then MCP (4.2.9) is equivalent to $H_{min}(\tau, y) = 0$.

Note that the function $H_{min}$ is not differentiable at a point $(\tau, y)$ where $F_i = \widetilde{H}_i(\tau, y)$ for some $i \in \{1, \ldots, n_R\}$. Hence, $y(\tau)$ is not differentiable at such a point. For this reason, we make use of the following strict complementarity assumption.

**Assumption 4.3.2.** *For each $i = 1, \ldots, n_R$, $F_i(\overline{\tau}) > 0$ holds when $\widetilde{H}_i(\overline{\tau}, y(\overline{\tau})) = 0$ at a solution $\overline{\tau}$ of MPEC (4.3.1).*

Under Assumption 4.3.2, $y(\overline{\tau})$ is a strictly complementary solution of MCP (4.2.9). Hence, we can define the following index sets associated with the solution $y(\overline{\tau})$:

$$A = \{i \in \{1, \ldots, n_R\} : (F(\overline{\tau}))_i > 0 = \widetilde{H}_i(\overline{\tau}, y(\overline{\tau}))\}, \tag{4.3.4}$$

$$B = \{i \in \{1, \ldots, n_R\} : (F(\overline{\tau}))_i = 0 < \widetilde{H}_i(\overline{\tau}, y(\overline{\tau}))\}, \tag{4.3.5}$$

which form a partitioning of $\{1, \ldots, n_R\}$. When $y$ is close to $y(\overline{\tau})$, $H_{min}(\overline{\tau}, y) = 0$ is locally reduced to

$$\widehat{H}(\overline{\tau}, y) \equiv \begin{pmatrix} (H_{min}(\overline{\tau}, y))_A \\ (H_{min}(\overline{\tau}, y))_B \\ \Gamma^T F - D(U(v)) \end{pmatrix} = \begin{pmatrix} F_A \\ \widetilde{H}_B(\overline{\tau}, y) \\ \Gamma^T F - D(U(v)) \end{pmatrix} = 0.$$

Now we show that the Jacobian of $\widehat{H}$ with respect to $y$ is nonsingular at $(\overline{\tau}, y(\overline{\tau}))$. We use the following additional assumption.

**Assumption 4.3.3.** $\nabla_{F_B} T_B(\tau, F)$ *is a positive definite matrix on some open sphere with center* $(\overline{\tau}, F(\overline{\tau}))$.

We show that

$$\begin{aligned} 0 &\neq \det \nabla_y \widehat{H}(\overline{\tau}, y(\overline{\tau}))^T \\ &= \det \begin{pmatrix} I & 0 & 0 \\ \nabla_{F_A} \widetilde{H}_B(\overline{\tau}, y(\overline{\tau}))^T & \nabla_{F_B} \widetilde{H}_B(\overline{\tau}, y(\overline{\tau}))^T & -\Gamma_B^T \\ \Gamma_A & \Gamma_B & -\nabla D(U(v))^T \nabla U(v)^T \end{pmatrix}, \end{aligned}$$

where $\Gamma_A$ and $\Gamma_B$ are submatrices of the route-OD pair incidence matrix $\Gamma$ corresponding to the index sets A and B, respectively. Let $M = \nabla_{F_B} \widetilde{H}_B(\overline{\tau}, y(\overline{\tau}))^T$ and $N = -\nabla D(U(v))^T \nabla U(v)^T$. Then

$$\det \nabla_y \widehat{H}(\overline{\tau}, y(\overline{\tau}))^T = \det \begin{pmatrix} M & -\Gamma_B^T \\ \Gamma_B & N \end{pmatrix}.$$

**Claim.** $\begin{pmatrix} M & -\Gamma_B^T \\ \Gamma_B & N \end{pmatrix}$ is nonsingular.

*Proof.* First, recall that from Assumption 3.3.1, $v$ is always nonnegative. Let vectors $p$ and $q$ satisfy

$$\begin{pmatrix} M & -\Gamma_B^T \\ \Gamma_B & N \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

i.e.,

$$Mp - \Gamma_B^T q = 0, \tag{4.3.6}$$

$$\Gamma_B p + Nq = 0. \tag{4.3.7}$$

From (4.3.7) we have $\Gamma_B p = -Nq$ so that

$$0 = p^T Mp - p^T \Gamma_B^T q = p^T Mp + q^T N^T q.$$

From Assumptions 3.3.1 and 3.3.2, $N$ is positive semidefinite. Hence,

$$0 = p^T Mp + q^T N^T q \geq p^T Mp.$$

Since $\nabla_{F_B} T_B(\tau, F)$ is positive definite by Assumption 4.3.3, it follows that $M = \nabla_{F_B} \widetilde{H}_B(\overline{\tau}, y(\overline{\tau}))^T$ is positive definite. Thus, $p = 0$.

Moreover, from the definition of the route-OD pair incidence matrix $\Gamma$, $\Gamma_B$ is a full rank matrix. Thus, by (4.3.6), we also have $q = 0$. It then follows that $\begin{pmatrix} M & -\Gamma_B^T \\ \Gamma_B & N \end{pmatrix}$ is nonsingular. $\square$

Based on the preceding discussion, we can apply the Implicit Function Theorem to obtain the following result.

**Theorem 4.3.1.** *Suppose that Assumptions 4.3.2 and 4.3.3 hold. Then the function $y(\tau)$ is differentiable near $\overline{\tau}$. Moreover, $\widetilde{\theta}$ is differentiable near $\overline{\tau}$ whenever $\theta$ is differentiable near $(\overline{\tau}, y(\overline{\tau}))$.*

## 4.4 Numerical Results

In this section, we present our computational results. In the numerical experiments, we solve MPEC (4.3.1) and MPEC (4.2.8). We call the former the implicit programming approach (ImPA) and the latter the direct optimization approach (DOA). In MPEC (4.3.1), the decision variable consists only of the toll $\tau$. In MPEC (4.2.8), on the other hand, the decision variables include the toll $\tau$, route flow $F$ and the minimum route cost $u$ associated with each OD pair. Both problems (4.3.1) and (4.2.8) were solved using the solver **fmincon** in the Optimization Toolbox of Matlab, which is based on the sequential programming method (SQP). We employ the generalized Newton method (GNM) of Jiang [43] to obtain the equilibrium solution $(F(\tau), v(\tau))$ of MCP (4.2.9) for a given $\tau$.

We used the sample networks shown in Chapter 3 (Figures 3.1 − 3.3) in our experiments. The routes and OD pairs in these networks are given in Table 3.1. The demand function used is

$$D_{\mathrm{w}}(u_{\mathrm{w}}) = -b^1_{\mathrm{w}}(\exp(-b^2_{\mathrm{w}} u_{\mathrm{w}})),$$

where $b^1_{\mathrm{w}}$ and $b^2_{\mathrm{w}}$ are given in Table 4.1. The link cost function used is

$$t_a(f_a) = c^1_a(1.0 + 0.15(f_a/c^2_a)^4),$$

where $c^1_{\mathrm{w}}$ and $c^2_{\mathrm{w}}$ are given in Table 4.2.

We have solved the two problems, i.e., (4.3.1) and (4.2.8), using two objective functions, namely, maximizing the total revenue (4.2.2) and minimizing the total travel cost (4.2.3), subject to $\tau_a \geq 0$, $\forall a \in \mathcal{A}$, with either of the following disutility functions:

$$U_{\mathrm{w}}(t) = t + 0.01t^2 \tag{4.4.1}$$

$$U_{\mathrm{w}}(t) = t^{1.1} \tag{4.4.2}$$

for the three sample networks. In the case of minimizing the total travel cost, we impose the additional constraints that $\tau_a \leq 5$, $\forall a \in \mathcal{A}$. The initial toll values used in the experiments for the two methods were randomly generated as follows. First we choose an initial toll value $\tau^0_i$ from $[0, 100]$ for the problem of maximizing the total revenue, and from $[0, 5]$ for the problem of minimizing the total travel cost. Then, using the generalized Newton method, we compute the corresponding $(F(\tau^0), u(\tau^0))$ and assign them as the initial route flow and initial route cost for DOA.

We used the **fmincon** solver to solve each problem with 10 randomly generated initial points. For all runs, the solver was terminated with the message *"OPTIMIZATION TERMINATED"* and exitflag values 1, 4 or 5 were obtained. The corresponding meanings of these exitflag values are as follows:

**1** : The first order optimality conditions are satisfied to the specified tolerance.

Table 4.1: Coefficients of the demand function.

| Network | Coefficients of the demand function | | |
|---|---|---|---|
| | O-D pair | $b_{\mathrm{w}}^1$ | $b_{\mathrm{w}}^2$ |
| 7-link A | 1-2 | 600 | 0.04 |
| | 1-3 | 500 | 0.03 |
| | 4-2 | 500 | 0.05 |
| | 4-3 | 400 | 0.05 |
| 7-link B | 1-4 | 200 | 0.2 |
| | 1-5 | 220 | 0.2 |
| 11-link | 1-7 | 600 | 0.04 |
| | 2-7 | 500 | 0.03 |
| | 3-7 | 500 | 0.05 |
| | 6-7 | 400 | 0.05 |

Table 4.2: Coefficients of the link cost function (lcf).

| Network | Link | Coefficients of the lcf | |
|---|---|---|---|
| | | $c_a^1$ | $c_a^2$ |
| 7-link A | a | 60 | 50 |
| | b | 10 | 2 |
| | c | 5 | 6 |
| | d | 8 | 2 |
| | e | 12 | 13 |
| | f | 5 | 6 |
| | g | 70 | 60 |
| 7-link B | a | 6 | 15 |
| | b | 4 | 15 |
| | c | 3 | 30 |
| | d | 5 | 30 |
| | e | 6 | 15 |
| | f | 4 | 15 |
| | g | 1 | 15 |
| 11-link | a | 6 | 200 |
| | b | 5 | 200 |
| | c | 6 | 200 |
| | d | 7 | 200 |
| | e | 6 | 100 |
| | f | 1 | 100 |
| | g | 5 | 150 |
| | h | 10 | 150 |
| | i | 11 | 200 |
| | j | 11 | 200 |
| | k | 15 | 200 |

**4** : The magnitude of a search direction is smaller than the specified tolerance and the constraint violation is less than options.TolCon.

**5** : The magnitude of a directional derivative is less than the specified tolerance and the constraint violation is less than options.TolCon.

Here "options.TolCon" is the *termination tolerance on the constraint violation*, whose default value is set at $10^{-7}$. The respective exitflag values of fmincon are shown in Table 4.3. In the table, the values under the column "Exitflag Values" correspond to the actual exitflag values obtained during the experiment, and the values shown under the column "Frequency" stand for the number of times each exitflag value was obtained both for the maximization and minimization problems with their respective disutility functions (4.4.1) and (4.4.2).

Table 4.3: Exitflag values of **fmincon**.

| Network | Exitflag values | Frequency | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Maximizing (4.2.2) | | | | Minimizing (4.2.3) | | | |
| | | (4.4.1) | | (4.4.2) | | (4.4.1) | | (4.4.2) | |
| | | DOA | ImP | DOA | ImP | DOA | ImP | DOA | ImP |
| 7-link A | 1 | 0 | 0 | 0 | 0 | 10 | 10 | 10 | 8 |
| | 4 | 3 | 0 | 4 | 0 | 0 | 0 | 0 | 0 |
| | 5 | 7 | 10 | 6 | 10 | 0 | 0 | 0 | 2 |
| 7-link B | 1 | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 9 |
| | 4 | 3 | 0 | 4 | 0 | 9 | 0 | 9 | 0 |
| | 5 | 7 | 10 | 6 | 10 | 1 | 0 | 1 | 1 |
| 11-link | 1 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 |
| | 4 | 8 | 0 | 9 | 0 | 6 | 0 | 6 | 0 |
| | 5 | 2 | 10 | 1 | 6 | 4 | 10 | 4 | 10 |

The computational results of 10 trials with different initial points for each problem are shown in Tables 4.4 and 4.5. In these tables, "Network" stands for the sample network used in the experiment, the columns "DOA" and "ImP" under "FVAL" show the mean, maximum and minimum objective function values of MPEC (4.2.8) and MPEC (4.3.1) obtained by the DOA and ImPA, respectively.

Table 4.4: Mean, maximum and minimum values of the objective function when the objective is to maximize the total revenue.

| Network | Disutility Function | | FVAL | | |
|---|---|---|---|---|---|
| | | | mean | maximum | minimum |
| 7-link A | (4.4.1) | DOA | 245.10 | 420.52 | 42.33 |
| | | ImP | 415.66 | 420.52 | 371.85 |
| | (4.4.2) | DOA | 394.13 | 531.84 | 253.40 |
| | | ImP | 531.90 | 531.90 | 531.90 |
| 7-link B | (4.4.1) | DOA | 21.96 | 78.77 | 0.03 |
| | | ImP | 85.31 | 93.57 | 83.86 |
| | (4.4.2) | DOA | 13.35 | 51.06 | 0.03 |
| | | ImP | 58.22 | 65.30 | 57.24 |
| 11-link | (4.4.1) | DOA | 2643.62 | 6099.8 | 42.74 |
| | | ImP | 6206.37 | 6521.5 | 5963.2 |
| | (4.4.2) | DOA | 3056.65 | 5519.4 | 302.32 |
| | | ImP | 5547.15 | 5832.3 | 5411.2 |

Table 4.4 shows the case of maximizing the total revenue. In all experiments considered, the mean objective function values obtained by ImPA were relatively greater than the corresponding values obtained by DOA. It has been observed that DOA is highly influenced by the choice of an initial point. In particular, for the case of the 7-link Network A with disutility function (4.4.2), ImPA always converged to the same limit point, while DOA converged to 10 different limit points for all 10 different initial points.

For the problem of minimizing the total travel cost, it can be seen from Table 4.5 that

Table 4.5: Mean, maximum and minimum values of the objective function when the objective is to minimize the total cost.

| Network | Disutility Function | | FVAL | | |
|---|---|---|---|---|---|
| | | | mean | maximum | minimum |
| 7-link A | (4.4.1) | DOA | 1153.9 | 1153.9 | 1153.9 |
| | | ImP | 1153.9 | 1153.9 | 1153.9 |
| | (4.4.2) | DOA | 1466.8 | 1466.8 | 1466.8 |
| | | ImP | 1466.8 | 1466.8 | 1466.8 |
| 7-link B | (4.4.1) | DOA | 99.38 | 176.77 | 55.70 |
| | | ImP | 41.08 | 41.08 | 41.08 |
| | (4.4.2) | DOA | 59.97 | 133.85 | 25.15 |
| | | ImP | 25.73 | 50.09 | 23.02 |
| 11-link | (4.4.1) | DOA | 9741.31 | 9816 | 9714.3 |
| | | ImP | 9733.85 | 9774.8 | 9716.3 |
| | (4.4.2) | DOA | 9724.57 | 9753.2 | 9714.3 |
| | | ImP | 8897.62 | 9026.2 | 8570.6 |

the two approaches give similar results for the 7-link A Network. However for the 7-link B and 11-link Networks, the mean objective function values obtained by ImPA were relatively smaller than the ones obtained by DOA. Similarly, as in the case of maximizing the total revenue, it has been observed that DOA is highly influenced by the choice of an initial point. In particular, for the 7-link Network B with disutility function (4.4.1), ImPA always converged to the same limit point, while DOA converged to 10 different limit points for all 10 initial points used in the experiment.

The numerical results reveal that ImPA generally performs better than DOA. They particularly show that the solutions obtained by DOA vary depending on the chosen initial points, while those of ImPA are stable. For the exitflag values in Table 4.3, we do not see any significant difference between ImPA and DOA. However, in general, solving problem (4.3.1) with ImPA spent more CPU time than solving problem (4.2.8) with DOA. This is because we used the generalized Newton method with fixed initial points in our codes. This problem may be remedied by using methods other than GNM in order to improve the efficiency of the proposed solution method.

## 4.5 Conclusions

In this chapter, we have introduced a road pricing model of the TEP with nonadditive route costs. Our approach makes use of the disutility function introduced in Chapter 3 for the lower level problems. We have shown that the resulting MPEC model can be reformulated as a mathematical program with strictly monotone MCP, hence, the ImPA is appropriate for solving the proposed model. Moreover, we have shown through numerical experiments that for solving our proposed model, the ImPA works better than the direct optimization approach in terms of stability of solutions.

# Chapter 5

# Convex Expected Residual Models for the Traffic Equilibrium Problem under Uncertainty

## 5.1 Introduction

In the study of the traffic equilibrium problem, most models assume that traffic users are given complete and accurate information about the actual traffic situation. However, in reality, the traffic situation is generally affected by various factors such as changes in weather conditions (e.g., different rainfall or snowfall intensities), different disturbances on the road (e.g., traffic accident or road improvements/construction) and changes in travel demand (e.g., morning or evening rush hour or special celebrations). Such variations in the travel time or travel demand from day to day imply that uncertainties do occur in the actual traffic system, hence the assumption that each driver knows completely the traffic conditions all over the network is rarely true. In fact, traffic users will have to make their route choices without exactly knowing the actual traffic conditions.

As an example, we consider weather. On a day with fine weather, the visibility is good and the road surface condition is good. On a rainy day, on the other hand, travel visibility is usually poor and the road user will have to travel at a reduced speed due to slippery road surface. Correspondingly, the travel cost on a fine day will probably differ from the travel cost on a rainy day. Moreover, travelers may even have to choose to postpone their trips scheduled on a rainy day to another day with good weather.

In this chapter, we consider the situation when the actual traffic situation is uncertain and the travel cost function or the demand function changes correspondingly with the situation. That is, we consider the traffic equilibrium problem (TEP) under uncertainty. The TEP under uncertainty can be reformulated as the stochastic variational inequality problem (SVIP) as will be shown in Section 5.2.

The SVIP, however, has no solution in general. Thus it is necessary to define a reasonable solution of the SVIP. In this chapter, we regard a global solution of the expected residual (ER) model for the SVIP as a reasonable solution of the stochastic affine variational inequality problem (SAVIP). We propose ER models based on the regularized gap function and the D-gap function for the VIP. In particular, we establish convexity of both the regularized gap function and the D-gap function and show that the resulting ER models with the proposed residual functions are convex. Thus, a reasonable solution of the TEP under uncertainty can be obtained by solving a convex programming problem.

This chapter is organized as follows. In the next section, we introduce the traffic equilibrium problem (TEP) under uncertainty and reformulate it as SVIP. We then discuss the expected residual model for SAVIP in Section 5.3. In Section 5.4, we introduce the regularized gap function and the D-gap function for the affine variational inequality problem (AVIP), and establish the convexity results of these functions. The proposed ER models for SAVIP are then presented in Section 5.5. We also establish the convexity results for these models in this section. Computational results for the TEP under uncertainty with the proposed ER models are given in Section 5.6. We give a brief conclusion in Section 5.7.

## 5.2 Traffic Equilibrium Problem under Uncertainty

Let $\Omega$ denote the sample space of factors contributing to the uncertainty in the traffic network, such as weather and accidents. For each event $\omega \in \Omega$, we assign an occurrence probability $p$. Let

| | |
|---|---|
| $u_{\mathrm{w}}$ : | minimal route cost for OD pair $\mathrm{w} \in \mathrm{W}$, |
| $u$ : | vector with components $u_{\mathrm{w}}$, |
| $D_{\mathrm{w}}(\omega)$ : | travel demand under uncertainty for OD pair $\mathrm{w} \in \mathrm{W}$, |
| $D(\omega)$ : | vector with components $D_{\mathrm{w}}(\omega)$, |
| $C(F, \omega)$ : | vector of route cost functions $C_r(F, \omega)$. |

In what follows, we are concerned with the special case where the travel demands do not depend on the route costs, which is the case of fixed travel demands.

Let $S(\omega) = \{F \in \Re^{n_R} | F \geq 0, \Gamma^T F = D(\omega)\}$. The traffic equilibrium problem (TEP) under uncertainty can be written as the following stochastic variational inequality problem (SVIP): Find $F^* \in S(\omega)$ such that

$$\langle C(F^*, \omega), F - F^* \rangle \geq 0, \ \forall F \in S(\omega). \tag{5.2.1}$$

The route cost function $C_r$ is defined by

$$C_r(F, \omega) = \sum_{a \in \mathcal{A}} \delta_{ar} t_a(\Delta F, \omega), \tag{5.2.2}$$

where $\Delta = (\delta_{ar})$ is the link-route incidence matrix and $t_a(\Delta F, \omega)$ is the travel time with uncertainty on link $a$. Note that if $t_a(\cdot, \omega)$ is affine, then (5.2.1) becomes a stochastic affine variational inequality problem (SAVIP). Details of the SAVIP will be discussed in the next section.

Note also that the TEP under uncertainty (5.2.1) – (5.2.2) may also be written as the following stochastic mixed complementarity problem (SMCP):

$$\begin{aligned} & 0 \leq F \perp C(F, \omega) - \Gamma u \geq 0, \\ & \Gamma^T F - D(\omega) = 0, \end{aligned} \tag{5.2.3}$$

where $x \perp y$ means vector $x$ and $y$ are perpendicular to each other. However, we observe that

$$M(\omega) = \begin{pmatrix} \nabla_F C(F, \omega) & -\Gamma \\ \Gamma^T & 0 \end{pmatrix}$$

is not a positive definite matrix, even if $C$ is affine with respect to $F$ and $\nabla_F C(F, \omega)$ is positive definite.

## 5.3 Expected Residual Models for the Stochastic Affine Variational Inequality Problem (SAVIP)

The affine variational inequality problem (AVIP) is to find $x \in S$ such that

$$\langle Mx + q, y - x \rangle \geq 0, \ \forall y \in S,$$

where $S = \{y \in \Re^n \mid Ay = b, y \geq 0\}$ with $A \in \Re^{m\times n}$ and $b \in \Re^m$, and $M \in \Re^{n\times n}$, $q \in \Re^n$. The AVIP is a wide class of problems which includes the quadratic programming problem and the linear complementarity problem.

A stochastic version of the AVIP is the stochastic affine variational inequality problem (SAVIP) which is to find $x \in S(\omega)$ such that

$$\langle M(\omega)x + q(\omega), y - x \rangle \geq 0, \ \forall y \in S(\omega), \tag{5.3.1}$$

where $S(\omega) = \{y \in \Re^n \mid A(\omega)y = b(\omega), y \geq 0\}$ with $A : \Omega \to \Re^{m\times n}$ and $b : \Omega \to \Re^m$, $M : \Omega \to \Re^{n\times n}$, $q : \Omega \to \Re^n$ and $(\Omega, P)$ is a probability space with $\Omega \subseteq \Re^l$. When $S(\omega) \equiv \Re^n_+$, the problem is reduced to the stochastic linear complementarity problem (SLCP) [18].

There is no vector $x$ satisfying (5.3.1) for all $\omega \in \Omega$ in general. We may consider two approaches in order to get a reasonable solution of SAVIP. One is the ***expected value*** (EV) method which formulates the problem as follows: Let $\bar{M} = E[M(\omega)]$, $\bar{q} = E[q(\omega)]$, $\bar{A} = E[A(\omega)]$ and $\bar{b} = E[b(\omega)]$, where $E$ denotes the expectation. The EV formulation is to find a vector $x \in \bar{S} = \{x | \bar{A}x = \bar{b}, \ x \geq 0\}$ such that

$$\langle \bar{M}x + \bar{q}, y - x \rangle \geq 0, \quad \forall y \in \bar{S}.$$

Another approach is the ***expected residual*** (ER) method which makes use of a residual function for AVIP. The ER method solves the following optimization problem:

$$\begin{array}{ll} \min & E[r(x,\omega)] \\ \text{s.t.} & x \in X, \end{array}$$

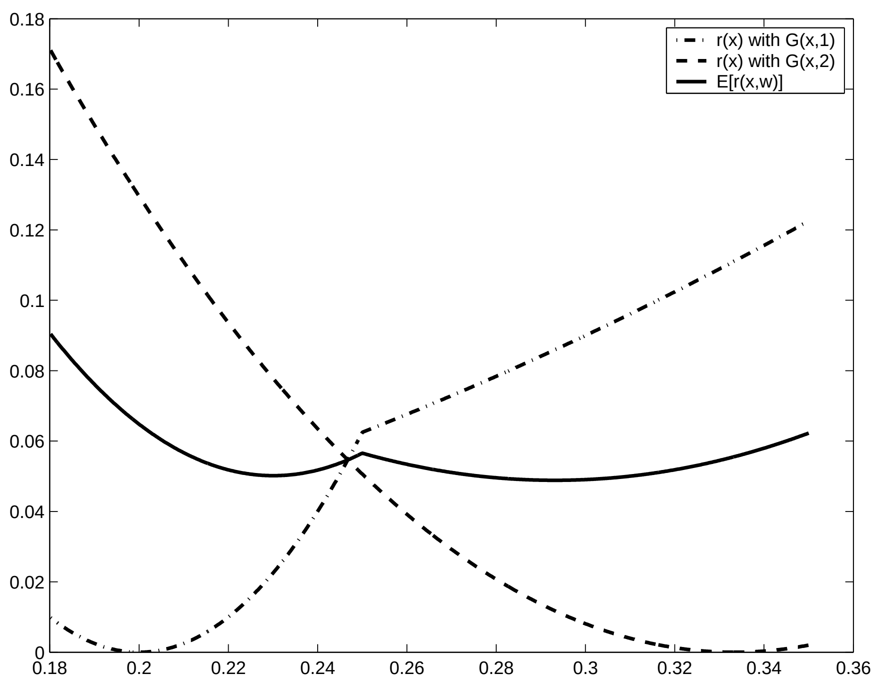
where $r(\cdot, \omega) : \Re^n \to \Re_+$ is a residual function for the variational inequality problem.

For the stochastic complementarity problem, the previous studies [18, 19, 28, 48] made use of an NCP function to formulate ER models. The ER model with the "min" function has been studied in [18, 19, 28] for the stochastic linear complementarity problem (SLCP). In particular, it is shown that, for a class of SLCPs, if the EV model has a bounded solution set, then the ER model also has a bounded solution set, but the converse is not true in general. Moreover, based on their theoretical and numerical results, Chen and Fukushima [18] and Chen, et al. [19] pointed out that a solution of the ER model is more reasonable than that of the EV model. Thus we consider the ER model in this chapter.

We can expect to obtain a solution of the ER model by using existing solution methods. However, there is no guarantee that such a solution is a global optimal solution of the ER model. The following example shows the nonconvexity of the ER model with the natural residual function.

**Example 5.3.1.** *Consider the SLCP with* $G(x,\omega) = \begin{cases} 5x-1 & \text{if } \omega = 1 \\ 2.7x - 0.9 & \text{if } \omega = 2 \end{cases}$ *and* $\Omega = \{1,2\}$, $p(1) = p(2) = \frac{1}{2}$, *and* $r(x,\omega) = \min\ (x, G(x,\omega))^2$. *Then the expected residual function* $E[r(x,\omega)]$ *is not convex as shown in Figure 5.1.*

Figure 5.1: The ER function with the natural residual for Example 5.3.1.




Recently, Luo and Lin [51] consider the stochastic variational inequality problem (SVIP) which is to find a vector $x \in S \subseteq \mathfrak{R}^n$ such that

$$\langle G(x,\omega), y - x\rangle \geq 0, \ \forall y \in S, \tag{5.3.2}$$

where $G : \mathfrak{R}^n \times \Omega \to \mathfrak{R}^n$. The SVIP (5.3.2) is a generalization of the SLCP studied in [18, 19, 28]. In [51], the authors consider the ER model with the regularized gap function

$$g(x,\omega) = \max_{y\in S}\langle G(x,\omega), x - y\rangle - \frac{\alpha}{2}\|x - y\|^2,$$

where $G$ is an affine function, that is, $G(x,\omega) = M(\omega)x + q(\omega)$. They establish the differentiability of this regularized gap function and the objective function $E[g(x,\omega)]$ of the ER model. They also establish the conditions for the level boundedness of $E[g(x,\omega)]$. They then propose a quasi-Monte Carlo method to solve the ER model for the SVIP by means of sequential approximations of $E[g(x,\omega)]$. The convergence properties of such an approximation method have also been established. However, they do not consider the convexity of the ER model.

## 5.4 Convexity of the Regularized Gap Function and D-gap Function for the AVIP

In this section, we show that the regularized gap function and the D-gap function are convex when $M$ is positive definite. The results are extensions of [61] where these functions are shown to be convex for the deterministic LCP.

Let $G(x) = Mx+q$ and $S = \{x \in \mathfrak{R}^n | Ax = b, x \geq 0\}$. Then the regularized gap function $f_\alpha : \mathfrak{R}^n \to \mathfrak{R}$ and the D-gap function $g_\alpha : \mathfrak{R}^n \to \mathfrak{R}_+$ for the AVIP are defined, respectively, by

$$f_\alpha(x) = \max_{y\in S}\left\{\langle G(x), x - y\rangle - \frac{1}{2\alpha}\|y - x\|^2\right\} \tag{5.4.1}$$

and

$$g_\alpha(x) = f_\alpha(x) - f_{1/\alpha}(x),$$

where $\alpha > 1$ is a positive constant.

In what follows we show the main results of this section which are natural extensions of [61, Theorems 2.1 and 3.1].

**Theorem 5.4.1.** *Suppose that $S$ is nonempty and $M$ is positive definite. Then the following statements hold.*

**(a)** *The regularized gap function $f_\alpha$ is convex for all $\alpha \geq \frac{1}{\beta_{\min}}$, where $\beta_{\min} > 0$ is the minimum eigenvalue of $M + M^T$. Moreover, if $\alpha \geq \frac{1}{\beta_{\min}}(1+\beta)$ with a positive constant $\beta$, then $f_\alpha$ is strongly convex with modulus $\beta$.*

**(b)** *The D-gap function $g_\alpha$ is convex for all $\alpha \geq \bar{\alpha}$, where $\bar{\alpha}$ is given by*

$$\bar{\alpha} = \max_{\|x\|=1} \frac{1 + x^T M^T M x}{2x^T M x} > 0.$$

*Moreover, $g_\alpha$ is strongly convex with modulus $\beta > 0$ for all $\alpha \geq \bar{\alpha} + \beta$.*

**Proof.** (a) Suppose that $\alpha \geq \frac{1}{\beta_{\min}}(1+\beta)$ with a nonnegative constant $\beta$. Then $v^T(\alpha(M + M^T) - I)v \geq \beta\|v\|^2$ for all $v \in \Re^n$. It then follows that the maximand in (5.4.1) is convex in $x$ for any $y$, and hence $f_\alpha$ is convex. Moreover, if $\beta > 0$, then the maximand is strongly convex with modulus $\beta$ for every $y$, and hence $f_\alpha$ is also strongly convex with modulus $\beta$.

(b) First notice that $-f_{1/\alpha}(x)$ is the optimum value of the following convex quadratic programming problem:

$$\begin{array}{ll} \min_y & -\langle G(x), x - y\rangle + \frac{\alpha}{2}\|y - x\|^2 \\ \text{s.t.} & Ay = b \\ & y \geq 0. \end{array} \tag{5.4.2}$$

Then by direct calculation, the Lagrangian dual problem of (5.4.2) is formulated as

$$\begin{array}{ll} \max_{(\lambda,\mu)} & h(x, \lambda, \mu) \\ \text{s.t.} & \lambda \in \Re^m \\ & \mu \geq 0, \end{array}$$

where $\lambda \in \Re^m$ and $\mu \in \Re^n$ are the Lagrange multipliers of (5.4.2), and $h(x, \lambda, \mu)$ is given by

$$\begin{aligned} h(x, \lambda, \mu) \;=\; & -\frac{1}{2\alpha}\|A^T\lambda - \mu\|^2 - \langle b, \lambda\rangle - \frac{1}{\alpha}\langle A^T\lambda - \mu, Mx + q - \alpha x\rangle \\ & -\frac{1}{2\alpha}\left\langle x, \left(M^T M - \alpha(M + M^T) + \alpha^2 I\right)x\right\rangle - \frac{1}{\alpha}\langle q, (M - \alpha I)x\rangle \\ & -\frac{1}{2\alpha}\|q\|^2 + \frac{\alpha}{2}\|x\|^2 - \langle x, Mx\rangle - \langle q, x\rangle. \end{aligned}$$

By the duality theorem, we have $-f_{1/\alpha}(x) = \max\{h(x,\lambda,\mu) | \lambda \in \Re^m, \mu \geq 0\}$. Hence, the D-gap function is written as

$$\begin{aligned} g_\alpha(x) &= \max_{y \in S}\{\langle G(x), x-y\rangle - \frac{1}{2\alpha}\|y-x\|^2\} + \max_{\lambda \in \Re^m, \mu \geq 0} h(x,\lambda,\mu) \\ &= \max_{y \in S, \lambda \in \Re^m, \mu \geq 0} \left\{\langle G(x), x-y\rangle - \frac{1}{2\alpha}\|y-x\|^2 + h(x,\lambda,\mu)\right\} \\ &= \max_{y \in S, \lambda \in \Re^m, \mu \geq 0} p(x,y,\lambda,\mu), \end{aligned}$$

where

$$p(x,y,\lambda,\mu) = \langle G(x), x-y\rangle - \frac{1}{2\alpha}\|y-x\|^2 + h(x,\lambda,\mu).$$

Next we show that $p(\cdot,y,\lambda,\mu)$ is convex for every fixed $(y,\lambda,\mu)$. Note that

$$\begin{aligned} \nabla_x^2 p(x,y,\lambda,\mu) &= M + M^T - \frac{1}{\alpha}I - \frac{1}{\alpha}M^TM + M + M^T - \alpha I + \alpha I - M - M^T \\ &= M + M^T - \frac{M^TM + I}{\alpha}. \end{aligned}$$

Then we can deduce that $M + M^T - \frac{M^TM+I}{\alpha}$ is positive semidefinite for any $\alpha \geq \bar{\alpha}$ in a way similar to the proof of [61, Theorem 3.1]. Therefore $p(\cdot,y,\lambda,\mu)$ is convex for all $(y,\lambda,\mu)$, and hence $g_\alpha$ is convex. □

**Remark 5.4.1.** *In [61], the linear complementarity problem (LCP) with the general cone is considered. We can extend Theorem 5.4.1 to the AVIP with the general cone by assuming Slater's constraint qualification.*

## 5.5 ER Models for SAVIP and Their Convexity

We formulate two ER models using the regularized gap function and the D-gap function. Let $G(x,\omega) = M(\omega)x + q(\omega)$. Then the regularized gap function and the D-gap function with random variable $\omega \in \Omega$ are defined by

$$f_\alpha(x,\omega) = \max_{y \in S(\omega)} \left\{\langle G(x,\omega), x-y\rangle - \frac{1}{2\alpha}\|y-x\|^2\right\}$$

and

$$g_\alpha(x,\omega) = f_\alpha(x,\omega) - f_{1/\alpha}(x,\omega).$$

Using these functions, we formulate the following two ER models:

$$\begin{array}{lll} \text{ER-R} & \min & E[f_\alpha(x,\omega) + \tau\|A(\omega)x - b(\omega)\|] \\ & \text{s.t.} & x \geq 0. \end{array}$$

$$\begin{array}{lll} \text{ER-D} & \min & E[g_\alpha(x,\omega)] \\ & \text{s.t.} & x \in \Re^n. \end{array}$$

The parameter $\tau > 0$ in ER-R is used for controlling the balance between the residual and the feasibility.

Let $\theta_\alpha^R(x)$ and $\theta_\alpha^D(x)$ be the objective functions of ER-R and ER-D, respectively, i.e.,

$$\begin{aligned} \theta_\alpha^R(x) &= E[f_\alpha(x,\omega) + \tau\|A(\omega)x - b(\omega)\|], \\ \theta_\alpha^D(x) &= E[g_\alpha(x,\omega)]. \end{aligned}$$

Now we investigate the conditions under which $\theta_\alpha^R(x)$ and $\theta_\alpha^D(x)$ are convex.

We call $M(\omega)$ uniformly positive definite with modulus $\beta_0$ if there exists a positive constant $\beta_0$ such that

$$\inf_{\omega\in\Omega,\|x\|=1} x^T M(\omega)x \geq \beta_0.$$

**Theorem 5.5.1.** *Suppose that $M(\omega)$ is uniformly positive definite with modulus $\beta_0$. Suppose also that $S(\omega)$ is nonempty for all $\omega \in \Omega$. Then the following statements hold.*

**(a)** *$\theta_\alpha^R$ is convex for all $\alpha \geq \frac{1}{2\beta_0}$ and strongly convex with modulus $\beta > 0$ for all $\alpha \geq \frac{1}{2\beta_0}(1+\beta)$.*

**(b)** *Suppose that $M(\omega)$ is bounded on $\Omega$. Then $\theta_\alpha^D$ is convex for all $\alpha \geq \bar{\alpha}$, where $\bar{\alpha}$ is given by*

$$\bar{\alpha} = \sup_{w\in\Omega,\|x\|=1} \frac{1 + x^T M(\omega)^T M(\omega)x}{2x^T M(\omega)x}.$$

*Moreover, $\theta_\alpha^D$ is strongly convex with modulus $\beta > 0$ for all $\alpha \geq \bar{\alpha} + \beta$.*

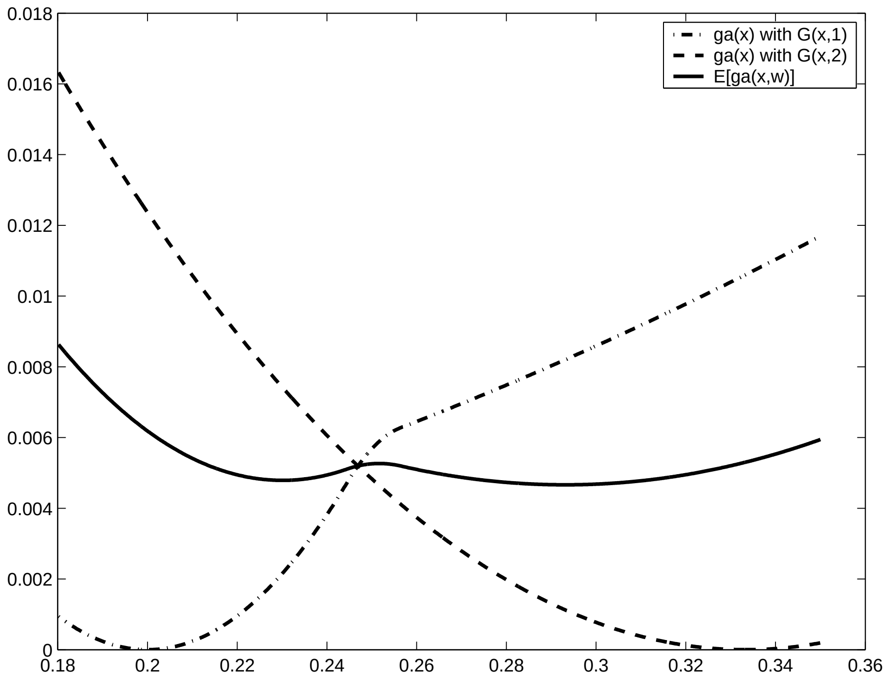
**Proof.** Since the sum of (strongly) convex functions is (strongly) convex, the statements (a) and (b) follow from Theorem 5.4.1. □

The theorem indicates that both ER-R and ER-D are convex programming problems, and hence we can obtain a global optimal solution using existing solution methods. These methods include the quasi-Newton methods and the interior point methods [56, 11].

We show the effects of $\alpha$ on $E[g_\alpha(x,\omega)]$ in the following example.

**Example 5.5.1.** *Let* $\Omega = \{1,2\}$ *and* $p(1) = p(2) = \frac{1}{2}$. *Let* $G(x,\omega) = \begin{cases} 5x-1 & \text{if } \omega = 1 \\ 2.7x - 0.9 & \text{if } \omega = 2, \end{cases}$ *where* $S(\omega) = \{x \in \Re \mid x \geq 0\}$. *Figures 5.2 and 5.3 show that* $E[g_\alpha(x,\omega)]$ *becomes convex when* $\alpha$ *is large.*

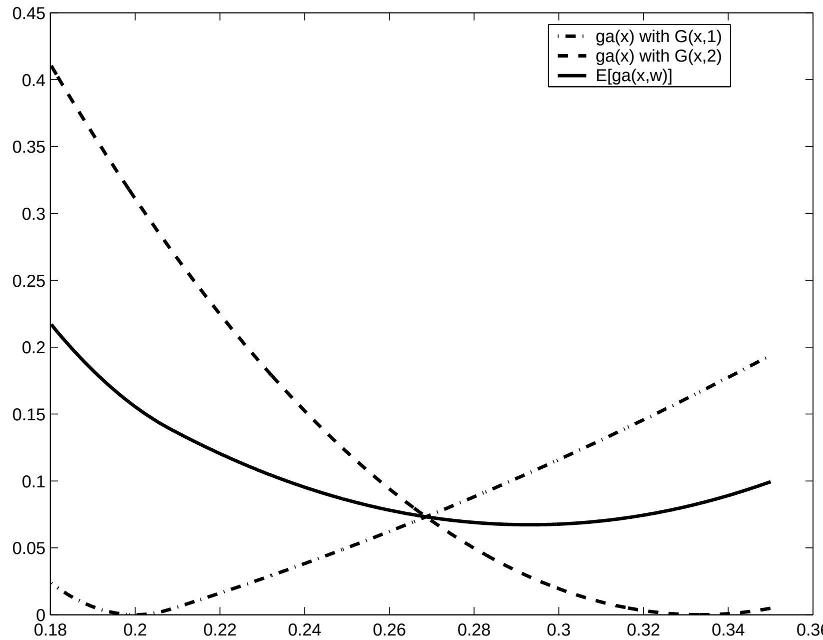
Figure 5.2: The ER-D function for Example 5.5.1 when $\alpha = 1.1$.




**Theorem 5.5.2.** *Suppose that* $M(\omega)$ *is uniformly positive definite. Suppose also that* $S(\omega)$ *is nonempty for all* $\omega \in \Omega$. *Then there exists a solution of ER-D. Moreover, if* $\alpha \geq \frac{1}{2\beta_0}(1+\beta)$, *then there exists a solution of ER-R.*

**Proof.** Since $M(\omega)$ is uniformly positive definite, $g_\alpha(x,\omega)$ is coercive for all $\omega \in \Omega$, see Proposition 10.3.9 in [26]. Therefore, $\theta^D_\alpha$ is also coercive. Hence, ER-D has a solution.

Figure 5.3: The ER-D function for Example 5.5.1 when $\alpha = 5$.




Moreover, if $\alpha \geq \frac{1}{2\beta_0}(1+\beta)$, it follows from Theorem 5.5.1 (a) that $\theta_\alpha^R$ is strongly convex. Thus, ER-R has a solution. $\square$

**Remark 5.5.1.** *When the travel time $t_a(\cdot, \omega)$ is an affine function for each $a$ and any $\omega$, the SVIP* (5.2.1) – (5.2.2) *becomes the SAVIP. In* (5.2.1)*, $M(\omega) = \nabla_F C(F, \omega)$ is positive definite under some conditions such as that $t_a$ is an increasing function of the link flows. Hence, by Theorems 5.5.1 and 5.5.2, the convexity of the proposed ER models (ER-R and ER-D) guarantees that we can obtain a global solution of the ER model for the AVIP formulation of the TEP with uncertainty.*

*Note however that we cannot apply Theorems 5.5.1 and 5.5.2 to the SMCP* (5.2.3) *since, as mentioned in Section 5.2, $M(\omega)$ is not positive definite.*

In the following, we give a particular example to illustrate the meaning of the solutions obtained by the ER-R and ER-D models.

Consider the case where there are two events, $\omega_1$ and $\omega_2$ that can happen, say, $\omega_1 = fine\ day$ and $\omega_2 = rainy\ day$. The TEP without uncertainty only considers the case when the traffic users know the exact weather of the day. That is, either $\omega_1$ happens with probability 1 or $\omega_2$ happens with probability 1. However, in reality, nobody can exactly predict the weather, and the available weather information such as the weather forecast cannot be trusted completely. The TEP with uncertainty considers the case when the occurrence probability of $\omega_1$ is, say, 0.6 and that of $\omega_2$ is, say, 0.4. The solution obtained by the ER model is regarded as the traffic flow pattern that satisfies the equilibrium condition on average.

## 5.6 Numerical Results

In this section, we present our computational results. In the numerical experiments, we solve the TEP under uncertainty (5.2.1) – (5.2.2) using the ER-D model proposed in Section 3. We solve the problem using the solver **fminunc** in the Optimization Toolbox of Matlab. We employ the quadratic programming solver **quadprog** of Matlab to compute $g_\alpha(x,\omega)$ for each $\omega \in \Omega$. The TEP under uncertainty is solved using different values of the parameter $\alpha$ to find its influence on the solution obtained. Moreover, we consider the case where $D(\omega)$ is fixed for all $\omega \in \Omega$ and the case where there is also uncertainty in $D(\omega)$. We also solve the MCP formulation (5.2.3) of the TEP under uncertainty using the ER method with Fischer-Burmeister (FB) function and compare the solutions obtained with those of the ER-D method.

The sample network shown in Figure 5.4 is used in our experiment. The attributes of this sample network are given in Table 5.1. We use the linear link cost function given by $t_a(f,\omega) = M(\omega)f + k(\omega)$, where $f = KF$ is the vector of link flows $f_a$, $k_i(\omega)$ represents the free travel cost of link $i$ and $M_{ij}(\omega)$ represents the magnitude of the effect of flows on link $j$ to the link cost of link $i$. The corresponding values of $H(\omega)$ and $k(\omega)$ are as follows:

$$M(\omega) = \begin{pmatrix} 22 & 0 & 2 & 2 & 4 & 1 & 2 & 0 & 4 & 5 \\ 0 & 15 & 0 & 0 & 1 & 2 & 0 & 3 & 5 & 3 \\ 2 & 0 & 14 & 0 & 2 & 0 & 1 & 3 & 2 & 3 \\ 2 & 0 & 0 & 16+50\omega & 0 & 2 & 3 & 1 & 2 & 4 \\ 4 & 1 & 2 & 0 & 12 & 0 & 2 & 2 & 0 & 0 \\ 1 & 2 & 0 & 2 & 0 & 10 & 0 & 0 & 1 & 2 \\ 2 & 0 & 1 & 3 & 2 & 0 & 11 & 0 & 0 & 0 \\ 0 & 3 & 3 & 1 & 2 & 0 & 0 & 14 & 0 & 1 \\ 4 & 5 & 2 & 2 & 0 & 1 & 0 & 0 & 16+50\omega & 0 \\ 5 & 3 & 3 & 4 & 0 & 2 & 0 & 1 & 0 & 20 \end{pmatrix}$$

Figure 5.4: A 10-link traffic network.

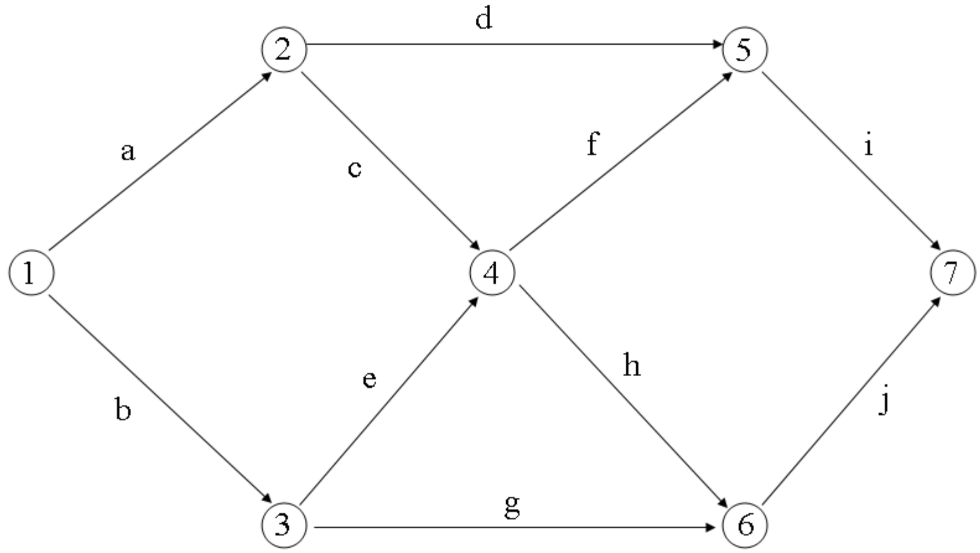



Table 5.1: OD pair, routes and links of the 10-link network.

| O-D pair | Routes | Links |
|---|---|---|
| 1-7 | 1 | {a,d,i} |
| | 2 | {a,c,f,i} |
| | 3 | {a,c,h,j} |
| | 4 | {b,e,f,i} |
| | 5 | {b,e,h,j} |
| | 6 | {b,g,j} |

and $k(\omega) = [50, 30, 40, 40 + 60\omega, 30, 50, 20, 60, 40 + 40\omega, 70]^T$.

Note that in this sample network, only the costs of links $d = (2, 5)$ and $i = (5, 7)$ depend on the random variable $\omega$. We assume that $\omega$ is uniformly distributed in the interval $[\frac{1}{2} - \delta, \frac{1}{2} + \delta]$. Hence, the expectation of $\omega$ is $\frac{1}{2}$ and its variance is $\frac{\delta^2}{3}$.

In our experiments, we choose $L$ samples of $\omega$ from the interval $[\frac{1}{2} - \delta, \frac{1}{2} + \delta]$ to approximate the actual continuous distribution. Hence, the occurrence probability of each event $\omega_i$ is $p_i = \frac{1}{L}$. We set $L = 21$. The values of $\alpha$ have been arbitrarily chosen in our numerical experiments.

### 5.6.1 Comparison of Link Flows for Different Values of $\alpha$

In this experiment, we look at the influence of $\alpha$. Here we set $\delta = 0.1$ and assume the fixed demand $D(\omega) = 200$ for all $\omega \in \Omega$. We present the results for various values of $\alpha$ in Tables 5.2 and 5.3. Table 5.2 shows that some route flows obtained are negative for small values of $\alpha$. However, as the value of $\alpha$ becomes large, the route flows obtained become all positive. In Table 5.3, it can be seen that as the value of $\alpha$ increases, the link flows on links $d = (2, 5)$ and $i = (5, 7)$ increase correspondingly. It is also interesting to observe that as the value of $\alpha$ becomes larger, the corresponding route flows obtained get closer to satisfying the demand, as shown in Figure 5.5. Moreover, the increase in the total route flow is small when $\alpha$ is large, so the effect of $\alpha$ to a solution is small for large $\alpha$.

Table 5.2: Route flows for different values of $\alpha$ when $D(\omega) = 200$.

| $\alpha$ | Route Flows |
|---|---|
| 3.3 | (35.85, 19.85, 8.47, 17.72, -2.28, 148.07 ) |
| 5 | (35.76, 19.81, 8.46, 17.68, -2.28, 147.77) |
| 18 | (29.39, 21.20, 2.45, 8.57, 3.40, 125.25) |
| 50 | (29.90, 23.63, 0.39, 6.58, 5.58, 127.24) |
| 100 | (30.38, 24.24, 0.13, 6.43, 5.93, 129.13) |
| 1000 | (31.00, 18.67, 6.14, 12.56, 0.06, 131.53) |

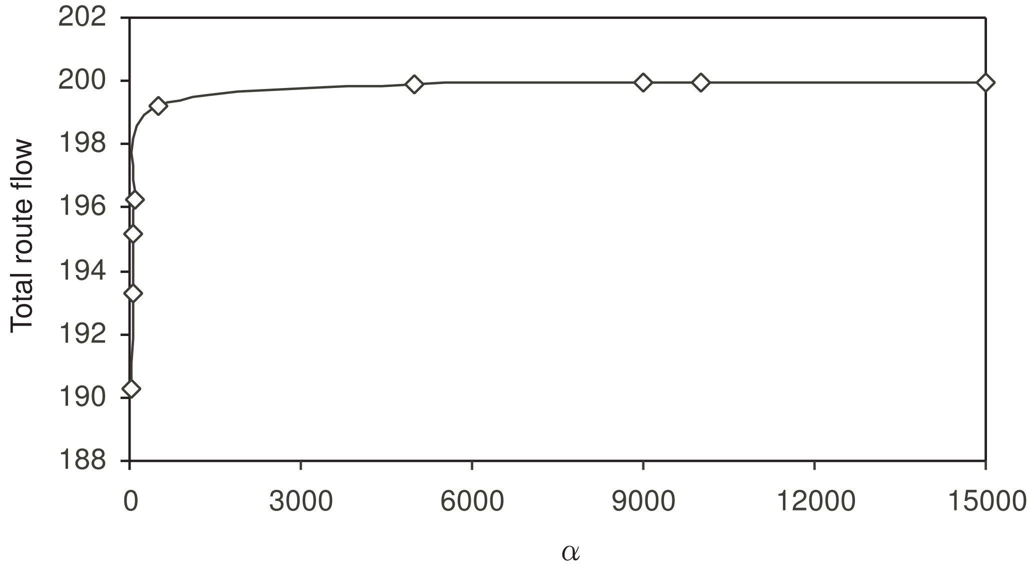
### 5.6.2 The Case of Travel Demand with Uncertainty

In our experiments, we also consider the case where the travel demand is subject to uncertainty. Here, we set $\delta = 0.1$ and assume that the travel demand is given by $D(\omega) = 500\omega - 100$.

Table 5.3: Link flows for different values of $\alpha$ when $D(\omega) = 200$.

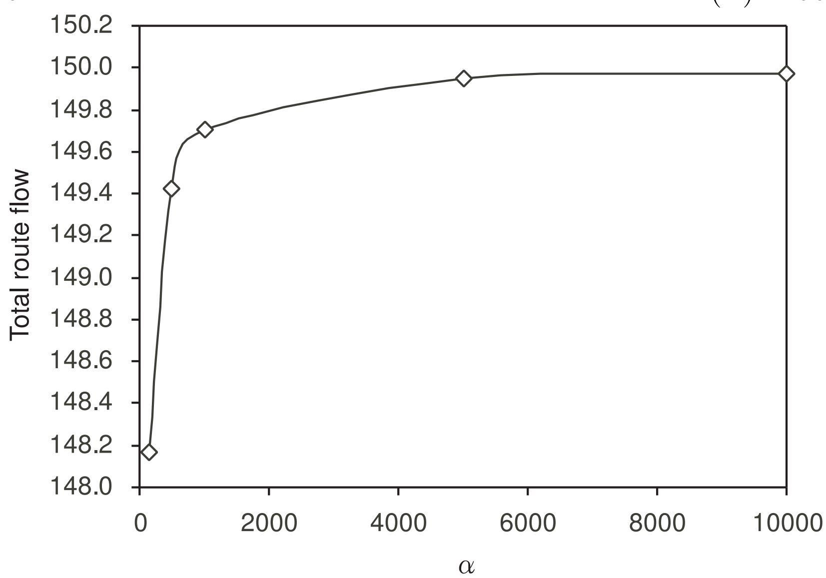
| $\alpha$ | Link Flows |
|---|---|
| 18 | (53.04, 137.22, 23.65, 29.39, 11.97, 29.77, 125.25, 5.85, 59.16, 131.10) |
| 50 | (53.91, 139.40, 24.01, 29.90, 12.16, 30.21, 127.24, 5.96, 60.11, 133.20) |
| 75 | (54.45, 140.73, 24.24, 30.21, 12.27, 30.50, 128.44, 6.03, 60.70, 134.47) |
| 100 | (54.75, 141.49, 24.37, 30.38, 12.36, 30.66, 129.13, 6.07, 61.04, 135.20) |
| 500 | (55.59, 143.59, 24.72, 30.87, 12.56, 31.11, 131.03, 6.17, 61.98, 137.20) |
| 5000 | (55.80, 144.11, 24.81, 30.99, 12.61, 31.22, 131.50, 6.20, 62.22, 137.70) |
| 9000 | (55.82, 144.14, 24.82, 31.00, 12.61, 31.23, 131.52, 6.20, 62.23, 137.72) |
| 10000 | (55.82, 144.14, 24.82, 31.00, 12.61, 31.23, 131.53, 6.20, 62.23, 137.73) |

Figure 5.5: Total route flow for different values of $\alpha$ when $D(\omega) = 200$.




The results are shown in Table 5.4.

Figure 5.6: Total route flow for different values of $\alpha$ when $D(\omega) = 500\omega - 100$.




It can be seen from Table 5.4 that some route flows obtained are negative when $\alpha$ is small. It is also observed that, similar to the case where the demand is fixed as $D(\omega) = 200$, the total route flow increases as the value of $\alpha$ increases. It can be seen from Figure 5.6 that as $\alpha$ becomes large, the total route flow approaches 150. Moreover, the increase in the total

Table 5.4: Route flows for different values of $\alpha$ when $D(\omega) = 500\omega - 100$.

| $\alpha$ | Route Flows |
|---|---|
| 10 | (25.6609, 14.1777, 5.9815, 12.7606, -2.0497, 107.0979) |
| 50 | (22.4198, 19.0718, -1.2081, 3.6273, 5.1809, 96.2403) |
| 145 | (22.8787, 18.2127, -0.0121, 4.9095, 4.0867, 98.0357) |
| 150 | (22.8883, 18.1652, 0.0425, 4.9659, 4.0342, 98.0733) |
| 500 | (23.0951, 17.1067, 1.2529, 6.2162, 2.8694, 98.8816) |
| 1000 | ( 23.1423, 15.8479, 2.5464, 7.5188, 1.5863, 99.0661) |
| 5000 | (23.1808, 15.9142, 2.5084, 7.4884, 1.6328, 99.2165) |
| 10000 | ( 23.1857, 15.9310, 2.4952, 7.4760 1.6471, 99.2355) |

route flow is small when $\alpha$ is large. Hence, the effect of $\alpha$ to a solution is small when $\alpha$ is large.

**Remark 5.6.1.** *Note that even if ER-D is not convex, i.e., when $\alpha$ is small, we may still obtain a reasonable solution. However, it can be seen from Table 5.2 and Table 5.4 that the solutions tend to be infeasible when $\alpha$ is small.*

### 5.6.3 Comparison of ER-D Model with Another ER Model

In this experiment, we also compare the ER-D model proposed in this chapter with another ER model which is based on the MCP formulation (5.2.3) and uses the Fischer-Burmeister (FB) function. This ER model is referred to as ER-FB and is defined as follows:

$$\begin{aligned} \min \quad & E[\Psi(F, u, \omega)] \\ \text{s.t.} \quad & \Gamma^T F - \bar{D} = 0, \end{aligned} \tag{5.6.1}$$

where $\Psi(F, u, \omega) = \|\Phi(F, u, \omega)\|^2$ with

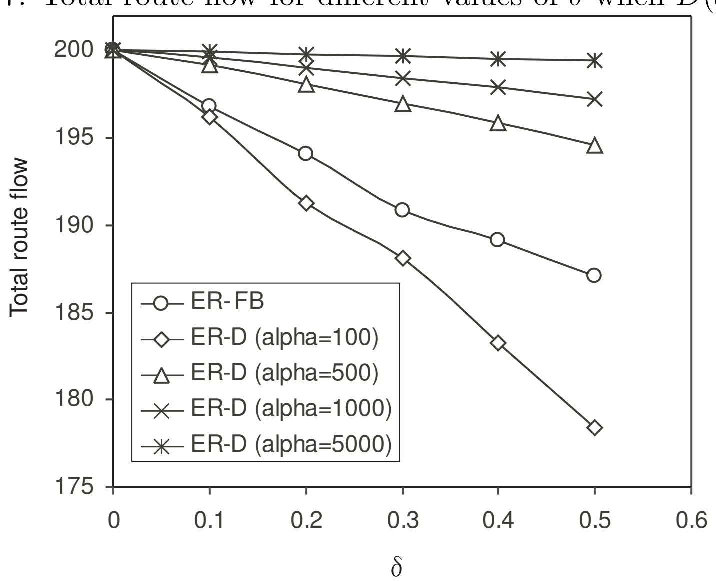
$$\Phi(F, u, \omega) = \begin{pmatrix} \phi(F_1, (C(F, \omega) - \Gamma u)_1) \\ \vdots \\ \phi(F_{n_R}, (C(F, \omega) - \Gamma u)_{n_R}) \end{pmatrix},$$

and the travel demand is assumed to be fixed at $\bar{D} = 200$.

Here, we consider the effect of $\delta$, which defines the interval $\Omega = [\frac{1}{2} - \delta, \frac{1}{2} + \delta]$, on the feasibility of the solutions obtained by the two ER methods. Link flows obtained for different values of $\alpha$ and different values of $\delta$ are shown in Table 5.5. It can be seen from the table that ER-FB and ER-D obtained the same solution when $\delta$ is very small. However, the results vary when $\delta$ and $\alpha$ become larger.

Moreover, as shown in Figure 5.7, as $\delta$ increases, that is, as the variance of $\omega$ becomes larger, the obtained route flows tend to violate the demand condition. More specifically, the bigger the value of $\delta$, the smaller the total route flow. However, as seen from the figure, the decrease in the total route flow for the ER-FB model is more significant than the ER-D models with large $\alpha$. Thus the solutions obtained by the ER-D models with larger values of $\alpha$ are more stable than the solutions obtained by the ER-FB model when the variance of random variable $\omega$ becomes large.

Figure 5.7: Total route flow for different values of $\delta$ when $D(\omega) = 200$.




**Remark 5.6.2.** *Note that when we consider ER-FB for the SAVIP, we need to convert*

*SAVIP into SMCP. Hence, the ER-FB model may lose some properties of the SAVIP. The difference between the solutions obtained by the ER-D model and the ER-FB model can be seen in the numerical results above, where the solution of ER-FB tends to violate the conditions more than the solution of ER-D.*

## 5.7 Conclusion

In this chapter, we have proposed two new ER models, the *ER-R model* which uses the regularized gap function and the *ER-D model* which uses the D-gap function for the stochastic affine variational inequality problem (SAVIP). Sufficient conditions for the models to be convex have been established. One of the ER models proposed in this chapter, the ER-D model, is then applied to the traffic equilibrium problem under uncertainty. In the numerical experiment, we compare the ER-D model with the MCP-based ER model with the Fischer-Burmeister function (ER-FB).

The numerical results show that, when the demand $D(\omega)$ is fixed ($D(\omega) = 200$), the proposed ER-D model with large $\alpha$ can obtain more reasonable solutions since the obtained route flows tend to satisfy the demand condition. Moreover, the demand condition is not greatly affected by the increase in the variance of $\omega$, that is, in the change in $\delta$, as compared to the ER-FB model.

In this study, the values of $\alpha$ used in the numerical experiments were only chosen arbitrarily. Determining a suitable value of $\alpha$ (and $\bar{\alpha}$) based on Theorem 5.5.1 and investigating its effect on the feasibility of the solutions would be an interesting topic to consider in the future.

Table 5.5: Link flows for ER-D and ER-FB for different values of $\delta$ when $D(\omega) = 200$.

| $\delta$ | Link Flows | | | | |
|---|---|---|---|---|---|
| | ER-D ($\alpha = 100$) | ER-D ($\alpha = 500$) | ER-D ($\alpha = 1000$) | ER-D ($\alpha = 5000$) | ER-FB |
| 0.0001 | 54.85 | 54.85 | 54.85 | 54.85 | 54.85 |
| | 145.15 | 145.15 | 145.15 | 145.15 | 145.15 |
| | 26.60 | 26.60 | 26.60 | 26.60 | 26.60 |
| | 28.25 | 28.25 | 28.25 | 28.25 | 28.25 |
| | 11.78 | 11.78 | 11.78 | 11.78 | 11.78 |
| | 31.59 | 31.59 | 31.59 | 31.59 | 31.59 |
| | 133.37 | 133.37 | 133.37 | 133.37 | 133.37 |
| | 6.79 | 6.79 | 6.79 | 6.79 | 6.79 |
| | 59.85 | 59.85 | 59.85 | 59.85 | 59.85 |
| | 140.15 | 140.15 | 140.15 | 140.15 | 140.15 |
| 0.01 | 54.75 | 55.60 | 55.71 | 55.80 | 54.61 |
| | 141.49 | 143.59 | 143.88 | 144.11 | 142.20 |
| | 24.37 | 24.72 | 24.77 | 24.81 | 24.02 |
| | 30.38 | 30.87 | 30.94 | 30.99 | 30.59 |
| | 12.36 | 12.56 | 12.59 | 12.61 | 15.50 |
| | 30.66 | 31.11 | 31.18 | 31.22 | 32.05 |
| | 129.14 | 131.03 | 131.29 | 131.50 | 126.68 |
| | 6.07 | 6.17 | 6.19 | 6.20 | 7.50 |
| | 61.05 | 61.98 | 62.11 | 62.22 | 62.64 |
| | 135.20 | 137.20 | 137.48 | 137.70 | 134.20 |
| 0.02 | 53.67 | 55.61 | 55.88 | 56.11 | 55.66 |
| | 137.62 | 142.47 | 143.15 | 143.70 | 138.40 |
| | 22.79 | 23.61 | 23.72 | 23.81 | 21.60 |
| | 30.88 | 32.01 | 32.17 | 32.30 | 34.07 |
| | 12.07 | 12.54 | 12.61 | 12.66 | 16.40 |
| | 28.82 | 29.82 | 29.96 | 30.08 | 33.16 |
| | 125.54 | 129.93 | 130.54 | 131.04 | 121.97 |
| | 6.05 | 6.32 | 6.36 | 6.40 | 4.80 |
| | 59.70 | 61.83 | 62.13 | 62.37 | 67.29 |
| | 131.59 | 136.26 | 136.90 | 137.43 | 126.80 |

# Chapter 6

# Summary and Conclusions

In this dissertation, the mathematical models of the traffic equilibrium problems have been studied. Nonadditive route cost functions as well as uncertainty have been considered in order to present a more realistic view of the traffic conditions. Specifically, the main results of this study are detailed in Chapters 3 – 5.

In Chapter 3, we have formulated the TEP with nonadditive route costs by introducing a disutility function (which is a function of the travel costs between OD pairs) , then presented its monotone MCP reformulation. For this reformulation, we have established the existence and uniqueness of the equilibrium of the proposed model. Moreover, we have shown through numerical experiments that our new MCP reformulation is useful in identifying an equilibrium of the TEP.

In Chapter 4, we have considered a road pricing model of the TEP with nonadditive route costs. Our approach makes use of the disutility function introduced in Chapter 3 for the lower level problems. We have shown that the resulting MPEC model of the road pricing problem can be reformulated as a mathematical program with strictly monotone MCP constraints, hence, the ImPA is appropriate for solving the proposed model. Moreover, we have shown through numerical experiments that for solving our proposed model, the ImPA works better than the direct optimization approach in terms of stability of solutions.

In Chapter 5, we consider the TEP under uncertainty. To solve this problem we have first proposed two new ER models, the *ER-R model* which uses the regularized gap function and the *ER-D model* which uses the D-gap function for the stochastic affine variational inequality problem (SAVIP). Sufficient conditions for the models to be convex have been established. One of the ER models proposed in this paper, the ER-D model, is then applied to the traffic equilibrium problem under uncertainty. In the numerical experiment, we compare the ER-D model with the MCP-based ER model with the Fischer-Burmeister function (ER-FB). The numerical results show that the proposed ER-D model can obtain better solutions when the model parameter $\alpha$ is large since the obtained route flows tend to satisfy the demand condition. Moreover, the demand condition is not greatly affected by the increase in the variance of the random variables as compared to the ER-FB model.

As stated above, we have made significant contributions in the study of the TEP. However, there still remain topics that need further improvement and issues that remain unresolved. We point out some of these in the following.

In Chapter 3, we restricted ourselves to the disutility function which is only a function of the travel costs between OD pairs. It is also important to consider the case when this disutility function is a function of the travel costs between each route. Moreover, the case of asymmetric demand function is also an important topic for future study. Finding an efficient nonadditive shortest path algorithm for the purpose of column generation, which is necessary when solving the nonadditive TEP for real-sized networks, is another important topic to consider.

In Chapter 4, we have only considered a simple traffic network. Extending our model to deal with the case of the more general multimodal or multiclass networks will be another important topic to explore.

Finally, in Chapter 5, we have only arbitrarily chosen the value of $\alpha$ in our experiments. Determining how large the value of $\alpha$ should be and investigating its effect on the feasibility of the solutions would be an interesting topic to consider in the future.

# Bibliography


[1] Aashtiani, H.Z. and Magnanti, T.L., 1981. Equilibria on a congested transportation network. *SIAM Journal on Algebraic and Discrete Methods* **2**, 213 – 226.

[2] Agdeppa, R.P., Yamashita, N. and Fukushima, M., 2007. The traffic equilibrium problem with nonadditive costs and its monotone mixed complementarity problem formulation. *Transportation Research B* **41** (8), 862 – 874.

[3] Agdeppa, R.P., Yamashita, N. and Fukushima, M., 2008. An implicit programming approach for the road pricing problem with nonadditive route costs. *Journal of Industrial and Management Optimization* **4** (1), 183 – 197.

[4] Agdeppa, R.P., Yamashita, N. and Fukushima, M., Convex expected residual models for stochastic affine variational inequality problems and its application to the traffic equilibrium problem. To appear in *Pacific Journal of Optimization.*

[5] Altman, E. and Wynter, L., 2004. Equilibrium, games and pricing in transportation and telecommunication networks. *Networks and Spatial Economics* **4**, 7 – 21.

[6] Asmuth, R., Eaves, B.C. and Peterson, E.L., 1979. Computing economic equilibria on affine networks with Lemke's Algorithm. *Mathematics of Operations Research* **4**, 207 – 214.

[7] Audet, C., Hansen, P., Jaumard, B. and Savard, G., 1997. Links between linear bilevel and mixed 0-1 programming problems. *Journal of Optimization Theory and Applications* **93**, 273 – 300.

[8] Bard, J.F. and Moore, J.T., 1990. A branch-and-bound algorithm for the bilevel programming problem. *SIAM Journal of Scientific and Statistical Computing* **11** (2), 281 – 292.

[9] Bernstein, D. and Gabriel, S. A., 1997. Solving the nonadditive traffic equilibrium problem. In: Pardalos, P.M., Hearn, D.W., Hager, W.W., (Eds.), *Lecture Notes in Economics and Mathematical Systems: Proceedings of the Network Optimization Conference*. Springer, New York, 72 – 102.

[10] Bernstein, D. and Wynter, L., 2000. Issues of uniqueness and convexity in non-additive bi-criteria equilibrium models. In: Bielli, M. and Carotenuto, P., (Eds.), *Proceedings of the Rome Jubilee 2000 Conference*. Rome, Italy.

[11] Bertsekas, D.P., 1999. *Nonlinear Programming 2nd Ed.*, Athena Scientific.

[12] Brotcorne, L., Labbé, M., Marcotte P. and Savard, G., 2001. A bilevel model for toll optimization on a multicommodity transportation network. *Transportation Science* **35** 345 – 358.

[13] Cao, M. and Ferris, M.C., 1996. A pivotal method for affine variational inequalities. *Mathematics of Operations Research* **21**, 44 – 64.

[14] Chen, A., Lo, H.K. and Yang, H., 2001. A self-adaptive projection and contraction algorithm for the traffic assignment problem with path-specific costs. *European Journal of Operational Research* **135**, 27 – 41.

[15] Chen, M., Bernstein, D. H., Chien, S. I. J. and Mouskos, K., 1999. Simplified formulation of the toll design problem. *Transportation Research Record 1667*, 88 – 95.

[16] Chen, X., 2000. Smoothing methods for complementarity problems and their applications: A survey. *Journal of the Operations Research Society of Japan* **43**, 32 – 47.

[17] X. Chen, and Fukushima, M., 2004. A smoothing method for a mathematical program with P-matrix linear complementarity constraints. *Computational Optimization and Applications* **27**, 223 – 246.

[18] Chen, X. and Fukushima, M., 2005. Expected residual minimization method for stochastic linear complementarity problems. *Mathematics of Operations Research* **30**, 1022 – 1038.

[19] Chen, X., Zhang, C. and Fukushima, M., 2009. Robust solution of monotone stochastic linear complementarity problems. *Mathematical Programming B* **117**, 51 – 80.

[20] Colson, B., Marcotte, P. and Savard, G., 2005. A trust-region method for nonlinear bilevel programming: Algorithm and computational Experience. *Computational Optimization and Applications* **30**, 211 – 227.

[21] Dafermos, S., 1980. Traffic equilibrium and variational inequalities. *Transportation Science* **14** (1), 42 – 54.

[22] Dirkse, S.P. and Ferris, M.C., 1995. The PATH Solver: A non-monotone stabilization scheme for mixed complementarity problems. *Optimization Methods and Software* **5**, 123 – 156.

[23] Dussault, J.P., Marcotte, P., Roch, S. and Savard, G., 2006. A smoothing heuristic for a bilevel pricing problem. *European Journal of Operational Research* **174**, 1396 – 1413.

[24] Facchinei, F., Jiang, H. and Qi, L., 1999. A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming* **85**, 107 – 134.

[25] Facchinei, F. and Kanzow, C., 1999. Beyond monotonicity in regularization methods for complementarity problems. *SIAM Journal of Control and Optimization* **37**, 1150 – 1161.

[26] Facchinei, F. and Pang, J.S., 2003. *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer-Verlag, New York.

[27] Facchinei, F. and Soares, J., 1997. A new merit function for nonlinear complementarity problems and related algorithm. *SIAM Journal on Optimization* **7** (1), 225 – 247.

[28] Fang, H., Chen, X. and Fukushima, M., 2007. Stochastic $R_0$ matrix linear complementarity problems. *SIAM Journal on Optimization* **18**, 482 – 506.

[29] Ferrari, P., 2002. Road network toll pricing and social welfare. *Transportation Research Part B* **36**, 471 – 483.

[30] Ferris, M.C. and Kanzow, C., 1998. Complementarity and related problems: A survey. *Mathematical Programming Technical Report*, 98 – 117.

[31] Ferris, M., Meeraus, A., and Rutherford, T., 1999. Computing Wardropian equilibrium in a complementarity framework. *Optimization Methods and Software* **10**, 669 – 685.

[32] Ferris, M. and Tin-Loi, F., 1999. On the solution of a minimum weight elastoplastic problem involving displacement and complementarity constraints. *Computer methods in Applied Mechanics and Engineering* **174**, 107 – 120.

[33] Fischer, A. and Jiang, H., 2000. Merit functions for complementarity and related problems: A survey. *Computational Optimization and Applications* **17**, 159 – 182.

[34] Florian, M., 1989. Mathematical Programming Applications in National, Regional and Urban Planning. *Mathematical Programmng*, 58 – 71.

[35] Gabriel, S. A. and Bernstein, D., 1997. The traffic equilibrium problem with nonadditive path costs. *Transportation Science* **31** (4), 337 – 348.

[36] Gabriel, S. A. and Bernstein, D., 2000. Nonadditive shortest paths: Subproblems in multi-agent competitive network model. *Computational and Mathematical Organization Theory* **6** (1), 29 – 45.

[37] Geiger, C. and Kanzow, C., 1996. On the resolution of monotone complementarity problems. *Computational Optimization and Applications* **5**, 155 – 173.

[38] Gürkan, G., Özge, A.Y., and Robinson, S.M., 1999. Sample-path solution of stochastic variational inequalities. *Mathematical Programming* **84**, 313 – 333.

[39] Han, D. and Lo, H.K., 2004. Solving non-additive traffic assignment problems: A descent method for co-coercive varational inequalities. *European Journal of Operational Research* **159**, 529 – 544.

[40] Hearn, D. and Ramana, M., 1998. Solving congestion toll pricing models. In: Marcotte, P. and Nguyen, S., (Eds.), *Equilibrium and Advanced Transportation Modeling.* Kluwer Academic Publishers, 109 – 124.

[41] Hobbs, B.F. and Kelly, K.A., 1992. Using game theory to analyze electric transmission pricing policies in the United States. *European Journal of Operational Research* **56**, 154 – 171.

[42] Huang, J. and Pang, J.S., 2000. A mathematical programming with equilibrium constraints approach to inverse pricing of American options: The case of an implied volatility surface. *The Journal of Computational Finance* **4**, 21 – 56.

[43] Jiang, H., 1999. Global convergence analysis of the generalized Newton and Gauss-Newton methods of the Fischer-Burmeister equation for the complementarity problem. *Mathematics of Operations Research* **24** (3), 529 – 543.

[44] Jiang, H. and Qi, L., 1997. A new nonsmooth equations approach to the nonlinear complementarity problems. *SIAM Journal of Control and Optimization* **35** (1), 178 – 193.

[45] Labbé, M., Marcotte, P. and Savard, G., 1998. A bilevel model of taxation and its application to optimal highway pricing. *Management Science* **44** (12), 1608 – 1622.

[46] Larsson, T., Lindberg, P. O., Lundgren, J., Patriksson, M. and Rydergren, C., 2002. On traffic equilibrium models with a nonlinear time/money relation. In: Patriksson, M. and Labbé M., (Eds.), *Transportation Planning – State of the Art*. Kluwer Academic Publishers, Boston, MA., 19 – 31.

[47] Larsson, T. and Patriksson, M., 1994. A class of gap functions for variational inequalities. *Mathematical Programming* **64**, 53 – 79.

[48] Lin, G. H., Chen, X. and Fukushima, M., New restricted NCP functions and their applications to stochastic NCP and stochastic MPEC. To appear in *Optimization*.

[49] Lin, G.H. and Fukushima, M., 2003. Some exact penalty results for nonlinear programs and their applications to mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications* **118**, 67 – 80.

[50] Lo, H.K. and Chen, A., 2000. Traffic equilibrium problem with route-specific costs: Formulation and algorithms. *Transportation Research Part B* **34**, 493 – 513.

[51] Luo, M.J. and Lin, G.H., Expected residual minimization method for stochastic variational inequality problems. To appear in *Journal of Optimization, Theory and Application.*

[52] Luo, Z.Q., Pang, J.S. and Ralph, D., 1996. *Mathematical Programs with Equilibrium Constraints.* Cambridge University Press, Cambridge, United Kingdom.

[53] Migdalas, A., 1995. Bilevel programming in traffic planning: Models, methods and challenge. *Journal of Global Optimization* **7**, 381 – 405.

[54] Moré, J.J., 1996. Global methods for nonlinear complementarity problems. *Mathematics of Operations Research* **21** (3), 589 – 614.

[55] Morrison, S.A., 1986. A survey of road pricing. *Transportation Research A* **20**, 87 – 97.

[56] Nocedal, J. and Wright, S.J., 1999. *Numerical Optimization*, Springer.

[57] Outrata, J., Kocvara, M. and Zowe, J., 1998. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints.* Kluwer Academic Publishers, Dordrecht.

[58] Patriksson, M., 1994. *The Traffic Assignment Problem: Models and Methods.* VSP BV, Utrecht, The Netherlands.

[59] Patriksson, M., 2004. Algorithms for computing traffic equilibria. *Networks and Spatial Economics* **4**, 23 – 38.

[60] Patriksson, M. and Rockafellar, R.T., 2002. A mathematical model and descent algorithm for bilevel traffic management. *Transportation Science* **36**, 271 – 291.

[61] Peng, J., 1997. Convexity of the implicit Lagrangian. *Journal of Optimization Theory and Applications* **92**, 331 – 341.

[62] Qi, L. and Sun, J., 1993. A nonsmooth version of Newton's method. *Mathematical Programming* **58**, 353 – 368.

[63] Scholtes, C. and Stöhr, M., 1999. Exact penalization of mathematical programs with complementarity constraints. *SIAM Journal on Control and Optimization* **37**, 617 – 652.

[64] Sherali, H.D., Soyster, A.L., and Murphy, F.H., 1983. Stackelberg-Nash-Cournot equilibria: Characteristics and computations. *Operations research* **31**, 253 – 276.

[65] Stackelberg, H.V., 1952. *The Theory of Market Economy.* Oxford University Press.

[66] Yan, H. and Lam, W.H.K., 1996. Optimal road tolls under conditions of queueing and congestion. *Transportation Research Part A* **30**, 319 – 332.

[67] Yang, H., 1997. Sensitivity analysis for the elastic-demand network equilibrium problem with applications. *Transportation Research Part B* **31** (1), 55 – 70.

[68] Yang, H. and Bell, M.G.H., 1997. Traffic restraint, road pricing and network equilibrium. *Transportation Research Part B* **31** (4), 303 – 314.

[69] Yang, H., Bell, M.G.H. and Meng, Q., 1997. Modeling the capacity and level of service of urban transportation networks. *Transportation Research Part B* **34**, 255 – 275.

[70] Wardrop, J. G., 1952. Some theoretical aspects of road traffic research, *Proceedings of the Institution of Civil Engineers Part II*, 325 - 378.