December 2003 **LIDS REPORT WP-1696**

Flow Control, Routing, and Performance from Service Provider Viewpoint [1]

by

Daron Acemoglu [2] and Asuman Ozdaglar [3]

Abstract

We consider a game theoretic framework to analyze traffic in a congested network, where a profit-maximizing monopolist sets prices for different routes. Each link in the network is associated with a flow-dependent latency function which specifies the time needed to traverse the link given its congestion. Users have utility functions defined over the amount of data flow transmitted, the delays they incur in transmission, and the expenditure they make for using the bandwidth. Given the prices of the links, each user chooses the amount of flow to send and the routes to maximize the utility he receives. We define an equilibrium of user choices given the prices, show its existence and essential uniqueness, and characterize how this equilibrium changes in response to changes in prices. We then define a monopoly equilibrium (ME) as the equilibrium prices set by the monopolist and the corresponding user equilibrium, and characterize this equilibrium.

We also study the performance of the ME relative to the user equilibrium at zero prices and the social optimum, which would result from the choice of a network planner with full information and full control over the flow and routing choices of users. Although equilibria for a given price vector or without prices are typically inefficient relative to the social optimum, we show that the ME achieves full efficiency for the routing problem (i.e., where each user has a fixed amount of data to transmit).

Keywords: Flow control, routing, generalized Wardrop equilibrium, monopoly pricing, externalities, efficiency.

[1] Corresponding Author: Asuman Ozdaglar, asuman@mit.edu
[2] Dept. of Economics, M.I.T., Cambridge, Mass., 02142.
[3] Dept. of Electrical Engineering and Computer Science, M.I.T., Cambridge, Mass., 02139.

1. INTRODUCTION

A fundamental problem in communication and data networks is the management of congestion, both to ensure timely transmission of information and to prevent loss of data in transmission. The standard approach is to use optimization methods to achieve the best potential network performance by adjusting the input flow rates of users and routing the resulting traffic. However, in many scenarios, it is impossible or impractical to regulate the traffic in such a centralized manner. Moreover, this approach requires considerable knowledge about the needs (preferences) of all the users in the network, an increasingly unrealistic assumption in the major networks of today such as the Internet.

The recognition of this problem has motivated a recent theoretical literature to consider the selfish flow choice and routing behavior of users in the absence of central planning (see, among others, [ORS93], [Kel97], [KMT98], [LoL99], [ABS02], [BaS02], [RoT02], [JoT03]). In these models, individuals choose their input flow rates and the routes to optimize their own objective, and are assumed to form conjectures about the behavior of other users consistent with the game theoretic notion of Nash Equilibrium or the notion of Wardrop Equilibrium first introduced in the analysis of congestion in transport networks (see [War52]). Not surprisingly, the resulting allocation differs markedly from the full-information social optimum, which would be chosen by a network manager with full information. For example, a recent paper by Roughgarden and Tardos [RoT03] studies a model where agents decide the routing of a given flow of information and finds that generally the performance of selfish routing can be much worse than the full-information optimum, but they also provide a bound on the performance gap in a specific case (when the latency functions are linear).

With a few notable exceptions (e.g., [BaS02]), however, this literature considers only situations in which users face no monetary costs of sending information. In contrast, most networks in practice are for-profit entities that charge prices for transmission of information. In this paper, we construct a model to analyze traffic in a congested network where there are prices associated with different routes and a profit-maximizing monopolist setting prices. Our objective is twofold. First, we develop a tractable framework to analyze both the sensitivity of the total flow of information and the routing choices to prices, and the determination of the equilibrium prices. Second, we use this model to study the performance gap between the monopolized network and the full-information social optimum.

Our model consists of a given network, a large number of users, and a profit-maximizing monopolist. To emphasize the main ideas, we consider a network with parallel links accessed by a number of heterogeneous users. Each link in the network is associated with a flow-dependent latency function specifying the time needed to traverse this link given the total flow through this link. Users have potentially different utility functions defined over the amount of data flow transmitted, the delays they incur in transmission and the expenditure they make for using the bandwidth. The monopolist sets prices per unit bandwidth for each link, and each user pays a price proportional to the amount of bandwidth he uses over the links. Given prices of the links, each user chooses the amount of flow to send and the routes. Formally, this corresponds to a two-stage game, where the monopolist is the Stackelberg leader setting prices anticipating the subsequent behavior of users. Each price vector defines a different subgame, and users play this (sub)game taking prices as given. It is important to note that, even though users are price-takers, they have to anticipate the amount of congestion on all routes, which they do according to the notion of Nash Equilibrium (NE) or Generalized Wardrop Equilibrium (GWE— we refer to this as "generalized" because each Wardrop Equilibrium is now conditional on a given vector of prices). In the NE, users take the effect of their own flow on congestion into account, while in the GWE, they ignore their own effect. Therefore, in terms of the general equilibrium analysis in the economics literature, the GWE also corresponds to a generalized competitive equilibrium where agents are both price takers and "latency takers" (i.e., they take latency on each route as given).[1] We define a Monopoly Equilibrium (ME) as the equilibrium prices set by the monopolist and the corresponding GWE or NE.

We first establish a number of important results on the behavior of such a network. In particular, under some relatively natural and general assumptions, we show that:

1. Given a vector of prices, there always exists a GWE and in every GWE, there is a unique distribution of flow rates among users and traffic load on the links of the network.

2. In every GWE, the flow rate of each user is continuous and non-increasing in the price vector.

3. In every GWE, the traffic load on each link is continuous in the price vector, non-increasing in its own price and non-decreasing in other links' prices.

4. There always exists a pure strategy NE, and the link loads at an NE converges to the link

[1] See, for example, Debreu [Deb59], or Arrow and Hahn [ArH71] on general competitive analysis. The major difference between the analysis here and the standard models of competitive analysis is the presence of congestion externalities. See the second footnote on the next page for a definition of externality.

loads at a GWE as the number of users approaches infinity.

5. There always exists a pure strategy ME. We further study properties and provide a characterization of equilibrium prices.

These results are not only useful for our subsequent analysis, but also provide a framework for future studies of network equilibria with flow control and routing in the presence of for-profit service providers.

We next study the performance of the ME relative to the full-information social optimum and the GWE without any prices (i.e., at zero prices). There has been a lot of interest in monopoly distortions among economists, where the focus is generally the price distortion created by monopoly versus its potential innovation benefits.[1] In our setting, the issue is rather different. We show that the GWE without prices always generates too much flow for each user relative to the social optimum. This is because each user creates an externality on other users by making the routes that he or she uses more congested.[2] Monopoly pricing may improve the performance of the system because the monopolist internalizes this externality. In particular, a key insight of our analysis is that the monopolist realizes that a higher price for a particular route may not reduce the attraction of this route to users by much, because, with any reduction in traffic, there will be a corresponding decrease in congestion. We show that for the routing problem (i.e., where each user has a predetermined amount of data that he or she wants to transmit), the ME achieves the full-information social optimum despite the selfish behavior of both users and the monopolist.

Although the ME with pure routing achieves the same load as the social optimum for each

[1] Monopoly distortions refer to the fact that because of monopoly pricing consumers reduce their demands below the social-welfare maximizing ("optimal") allocation. Monopoly distortions are the focus of the bulk of economic research on monopoly. Economists also emphasize that monopoly sometimes emerges because of technological reasons (the so-called "natural monopoly"), in particular when economies of scale and fixed costs imply that only one firm can cover its costs in the market. Balancing the pricing distortions of monopoly may be the potential gains in innovation under monopoly, first proposed by Joseph Schumpeter [Sch75]. See, for example, Tirole [Tir90].

[2] An externality arises when an individual's actions affects the utility of others through non-market means. For example, in our framework when latency functions are not constant, by using a particular link, an individual creates a negative externality on all other users transmitting data through that link. The negative externality implies that individuals tend to transmit too much data and create too much congestion.

link, it provides different levels of utilities to users than the choice of a fully-informed network planner who would implement the same allocation without charging users.[1] In fact, in the ME, users pay a considerable amount to the monopolist (in fact, all of their consumer surplus is taken by the monopolist.[2]) In a companion paper, we study a similar model with multiple providers controlling various links, and show that, under certain circumstances, the corresponding Oligopoly Equilibrium also achieves the social optimum, but in addition, transfers the entire consumer surplus to the users.

Our paper is related to the burgeoning literature on network pricing and selfish routing mentioned above. Much of this work does not consider price-setting by a service provider. Our model differs from these papers by incorporating simultaneously flow control, routing, and service provision. In this regard, it is most closely related to [BaS02], which studies the asymptotic behavior of equilibrium with specific utility and latency functions. Our work is therefore a generalization of this model to more general utility and latency functions, and contains a full characterization of the monopoly equilibrium (which is made difficult by the fact that the monopolist problem is not convex). There has been considerable research to design strategies and economic incentives to cope with the inefficiency created by the selfish behavior of the agents (e.g., taxing, artificial Stackelberg games, see for example [KLO97], [CDR03]). But, except for special scenarios, these can be viewed as outside interventions in the system. In this paper, we analyze the performance of the ME and show that it may improve performance since it internalizes congestion externalities. Hence, we show that having a service provider in the system automatically creates a natural market mechanism and the right incentives to price the effects of congestion. To the best of our knowledge, this is a new insight in the literature.

The paper is organized as follows. In Section 2, we introduce our model in detail. For a given price vector, we define a GWE among the users of the network and explore the properties

---

[1] Note that the standard definition of efficiency in economics does not make any reference to the utilities of the users, but is in terms of the allocation, here flow rates and link loads; see [Deb59] or [MWG95].

[2] Consumer surplus refers to the net utility that the individual obtains over and above what he would have obtained had he not participated in this market. In other words, since his utility without participation is 0, the consumer surplus is the net utility level that he achieves by choosing his optimal flow rates and routing at a given vector of prices. Note that in the pure routing case, the monopolist captures all the consumer surplus, i.e., all consumers obtain zero utility from the network (see Section 6).

of this equilibrium. In Section 3, we study the relation of the GWE to the Nash Equilibrium, and explain why the focus on the GWE in this context is natural. In Section 4, we characterize the Monopoly Equilibrium. In Section 5, we compare the ME to the social optimum. In Section 6, we present a model for the routing problem under monopoly pricing and analyze the performacne of the monopolized system. We present our conclusions and summarize future directions in Section 7.

2. MODEL

We consider a network with I parallel links accessed by J users. Let $\mathcal{I} = \{1, \ldots, I\}$ denote the set of links and $\mathcal{J} = \{1, \ldots, J\}$ denote the set of users. Let x_j^i denote the flow of user j on link i and $x_j = [x_j^1, \ldots, x_j^I]'$ denote the vector of flows of user j. We assume that user j receives a utility of $u_j(\Gamma_j)$, where

$$\Gamma_j = \sum_{i=1}^{I} x_j^i$$

denotes the total flow rate of user j on all links. Throughout the paper, we will refer to Γ_j as the flow rate of user j. Each link in the network has a flow-dependent latency function $l^i(\gamma^i)$, where

$$\gamma^i = \sum_{j=1}^{J} x_j^i \tag{2.1}$$

denotes the total flow (link load) on link i. This latency function specifies the delays in transmission as a function of the link load. We denote the price per unit bandwidth of link i by p^i, and we will later allow a monopolist service provider to control prices.

An important aspect of our model is our assumption that each user is "small" in the sense that when he switches his flows from one path to another, there is no considerable change in the link latencies. This is known as the "Wardrop's Principle", due to a paper by Wardrop [War52]. Wardrop's Principle is used extensively in problems where users have fixed rates to transmit and only their routing choices are subject to optimization; see [RoT02], [HaM85], [ScS03], [CSS03]. Here, we extend this idea to the problem of combined flow control and routing. In addition, we establish below that, similar to [HaM85], the equilibria that we characterize are limits of sequences of Nash Equilibria obtained from sequences of games as the number of users goes to infinity.

We denote the vector of total flows on the links [cf. Eq. (2.1)] by $\gamma = [\gamma^1, \ldots, \gamma^I]$, and the vector of prices of the links by $p = [p^1, \ldots, p^I]$. Given γ and p, each user chooses $x_j \geq 0$ to maximize its payoff function $v_j(x_j; \gamma, p)$, given by

$$v_j(x_j; \gamma, p) = u_j\Big(\sum_{i=1}^{I} x_j^i\Big) - \sum_{i=1}^{I} l^i(\gamma^i) x_j^i - \sum_{i=1}^{I} p^i x_j^i. \tag{2.2}$$

The fact that user utility is additively-separable between total flow, Γ_j, and delay on link i, $l^i(\gamma^i)$, is a useful simplification in line with the rest of the literature. The important feature is that the utility of each user (or alternately the total flow allocated to each user) depends on the flows of all the users.

Implicit in our notation is the fact that each user acts as a "price taker" and a "link load taker" (the utility function is conditioned on the price vector, p, and the link load vector, γ). This is a natural assumption when there are many users, and we show in Section 3 that there is no loss of generality in this assumption as the number of users becomes large.

An equilibrium of this game is defined as follows.

Definition 2.1: Let $x = [x'_1, \ldots, x'_J]$ denote the vector of flows of all the users in the network. For a given price vector $p \geq 0$, a flow vector x is an equilibrium of the game where the payoff functions of the users are given by Eq. (2.2) if

$$x_j \in \arg \max_{0 \leq y_j^i \leq C^i, \ \forall \ i} v_j(y_j; \gamma, p), \qquad \forall \ j \in \mathcal{J},$$

where C^i denotes the capacity of link i, and

$$\gamma^i = \sum_{j=1}^{J} x_j^i, \qquad \forall \ i \in \mathcal{I}.$$

We will refer to this flow vector as the *Generalized Wardrop Equilibrium* (GWE) of this game. Note that the GWE is a function of the price vector p. We will write it as $x(p)$ to make the dependence on price explicit, whenever there is a danger of confusion. Note also that whenever we are talking about a GWE, we will use the notation Γ_j to denote the flow rate of user j, and γ^i to denote the load of link i.

Assumption 2.1: Assume that for each j, the utility function $u_j : [0, \infty) \mapsto [0, \infty)$ satisfies the following conditions:

(a) u_j is strictly concave, nondecreasing, and continuously differentiable.

(b) $0 < u'_j(0) < \infty$.

Also assume that for each i, the link latency function $l^i : [0, C^i) \mapsto [0, \infty)$ satisfies the following conditions:

(a') l^i is continuous and strictly increasing.

(b') $l^i(0) = 0$.

(c') $l^i(x) \to \infty$ as $x \to C^i$.

Assumption 2.1 is natural in this context. It specifies that users derive greater utility from transmitting more data. Concavity of the utility function implies that we are considering elastic traffic (i.e., traditional data applications like file transfer and e-mail, which are tolerant of delays, see [She95]), for which transmitting more and more data has diminishing returns for users. This assumption enables us to obtain a characterization of the equilibrium. In Section 6, we will also consider some non-concave utility functions. Assumption 2.1(a') on the latency function ensures that more data transmission creates greater congestion, which is the essence of the problem understudy here. Assumption 2.1(b') imposes that the cost of transmitting information without any congestion is equal to 0 on all links.[1] Finally, Assumption 2.1(c') serves to capture the capacity constraints on the links within the latency functions.[2] In view of this assumption and Assumption 2.1(b), it can be seen that at any GWE, $x(p)$, we have $\sum_{j\in\mathcal{J}} x_j^i(p) < C^i$. Therefore, the capacity constraint in the definition of GWE could be neglected.

In the next proposition, we show that under Assumption 2.1, a GWE always exists. Although the existence of an equilibrium (GWE) is a useful starting point, in many general equilibrium environments there can be multiple equilibria, making comparisons between equilibria and optima and the analysis of sensitivity of the equilibrium to underlying variables difficult. Therefore in the next proposition, we also establish the essential uniqueness of the equilibrium. By essential uniqueness, we mean that, even though the x's may not be unique, the flow rate of each user and the load of each link are uniquely defined. The proof uses a classical technique that was used for showing related uniqueness results for the Wardrop equilibrium (see [BMW56], [Kel91]).

Proposition 2.1: (Existence-Essential Uniqueness) Let Assumption 2.1 hold. For a given $p \geq 0$, consider the J-player game where the payoff functions of the users are given by Eq. (2.2). Then, there exists a GWE, and the flow rates and the link loads at any GWE are unique. In particular, let x and $\overline{x}$ be two generalized Wardrop equilibria of this game. We define, for all $i \in \mathcal{I}$,

$$\gamma^i = \sum_{j\in\mathcal{J}} x_j^i, \qquad \overline{\gamma}^i = \sum_{j\in\mathcal{J}} \overline{x}_j^i,$$

[1] This assumption simplifies the analysis. It could be relaxed at the expense of added notation.

[2] An alternative assumption could be that, for each j, there exists a nonzero scalar B_j such that $u_j'(B_j) = 0$. This assumption would guarantee that no individual has an infinite demand or infinite willingness to pay for data transmission.

and for all $j \in \mathcal{J}$,

$$\Gamma_j = \sum_{i \in \mathcal{I}} x_j^i, \qquad \overline{\Gamma}_j = \sum_{i \in \mathcal{I}} \overline{x}_j^i.$$

We have

$$\gamma^i = \overline{\gamma}^i, \qquad \forall\ i \in \mathcal{I},$$

$$\Gamma_j = \overline{\Gamma}_j, \qquad \forall\ j \in \mathcal{J}.$$

Proof: Given any $p \geq 0$, consider the following optimization problem

$$\begin{aligned} \text{maximize} \ & \sum_{j \in \mathcal{J}} u_j(\Gamma_j) - \sum_{i \in \mathcal{I}} \int_0^{\gamma^i} l^i(z)dz - \sum_{i \in \mathcal{I}} p^i \gamma^i \\ \text{subject to} \ & \Gamma_j = \sum_{i \in \mathcal{I}} x_j^i, \qquad \forall j \\ & \gamma^i = \sum_{j \in \mathcal{J}} x_j^i, \qquad \forall i \\ & x_j^i \geq 0, \qquad \forall\ i,\ j. \end{aligned} \tag{2.3}$$

It can be seen by checking the first order necessary and sufficient conditions that the optimal solutions of this problem correspond exactly to the set of GWE at the price vector p. In view of Assumption 2.1, the objective function of problem (2.3) is continuous and the feasible set is compact. Hence, this problem has an optimal solution, showing the existence of a GWE. Note further that, since u_j is strictly concave for all j and l^i is strictly increasing for all i, it follows that the objective function of problem (2.3) is a strictly concave function of Γ and γ. Hence, there exists a unique optimum for the flow rate vector Γ and the link load vector γ, showing the uniqueness of the flow rates and the link loads at a GWE. **Q.E.D.**

Sensitivity to Prices

The essential uniqueness of the equilibrium enables us to examine the continuity and monotonicity properties of the link loads and flow rates of users as functions of the link prices. We will see that the key equilibrium notions, the flow rates and the link loads, are both continuous and monotonic functions of prices.

Proposition 2.2: (Continuity) Let Assumption 2.1 hold. For a given $p \geq 0$, let $x(p)$ be a GWE of the J-player game where the payoff functions of the users are given by Eq. (2.2). Denote the corresponding link loads and flow rates by

$$\gamma^i(p) = \sum_{j=1}^{J} x_j^i(p), \qquad \forall\ i \in \mathcal{I},$$

$$\Gamma_j(p) = \sum_{i=1}^{I} x_j^i(p), \qquad \forall\, j \in \mathcal{J}.$$

For all i and j, $\gamma^i(p)$ and $\Gamma_j(p)$ are continuous at all $p \geq 0$.

Proof: Let $\{p_k\}$ be some nonnegative sequence that converges to some $p \geq 0$. We show that $\gamma^i(p_k)$ converges to $\gamma^i(p)$ for all i, and $\Gamma_j(p_k)$ converges to $\Gamma_j(p)$ for all j. By Proposition 2.1, for each p_k, there exists a GWE of this game, denoted by $x(p_k)$, which satisfies,

$$\gamma^i(p_k) = \sum_{j=1}^{J} x_j^i(p_k), \qquad \Gamma_j(p_k) = \sum_{i=1}^{I} x_j^i(p_k).$$

The sequence $\{x(p_k)\}$ lies in the compact set since $x_j^i(p_k) \in [0, C^i]$ for all k (cf. Assumption 2.1). Therefore, it has a limit point, denoted by $\tilde{x}$. We assume without loss of generality that the sequence $\{x(p_k)\}$ converges to $\tilde{x}$, which implies that

$$\gamma^i(p_k) \to \tilde{\gamma}^i = \sum_{j=1}^{J} \tilde{x}_j^i,$$

$$\Gamma_j(p_k) \to \tilde{\Gamma}_j = \sum_{i=1}^{I} \tilde{x}_j^i.$$

We next show that $\tilde{x}$ is a GWE of this game with the prices given by p, which by Proposition 2.1, guarantees that $\gamma^i(p_k) \to \gamma^i(p)$ and $\Gamma_j(p_k) \to \Gamma_j(p)$. Assume, to arrive at a contradiction, that $\tilde{x}$ is not a GWE. This implies that for some $j \in \mathcal{J}$,

$$\tilde{x}_j \notin \arg\max_{y_j \geq 0} \left\{ u_j\Big(\sum_{i=1}^{I} y_j^i\Big) - \sum_{i=1}^{I} l^i\Big(\sum_{j=1}^{J} \tilde{x}_j^i\Big) y_j^i - \sum_{i=1}^{I} p^i y_j^i \right\}.$$

Hence, there exists some $\epsilon > 0$ and some $y_j \geq 0$ such that

$$u_j\Big(\sum_{i=1}^{I} y_j^i\Big) - \sum_{i=1}^{I} l^i(\tilde{\gamma}^i) y_j^i - \sum_{i=1}^{I} p^i y_j^i \ \geq u_j\Big(\sum_{i=1}^{I} \tilde{x}_j^i\Big) - \sum_{i=1}^{I} l^i(\tilde{\gamma}^i)\tilde{x}_j^i - \sum_{i=1}^{I} p^i \tilde{x}_j^i + 3\epsilon.$$

Since u_j is continuous (which follows by concavity of u_j, see [BNO03]), l^i is continuous (cf. Assumption 2.1), and $p_k \to p$, $x(p_k) \to \tilde{x}$, we have, for k sufficiently large,

$$\begin{aligned} u_j\Big(\sum_{i=1}^{I} y_j^i\Big) - \sum_{i=1}^{I} l^i\big(\gamma^i(p_k)\big) y_j^i - \sum_{i=1}^{I} p_k^i y_j^i &\geq u_j\Big(\sum_{i=1}^{I} y_j^i\Big) - \sum_{i=1}^{I} l^i(\tilde{\gamma}^i) y_j^i - \sum_{i=1}^{I} p^i y_j^i - \epsilon \\ &\geq u_j\Big(\sum_{i=1}^{I} \tilde{x}_j^i\Big) - \sum_{i=1}^{I} l^i(\tilde{\gamma}^i)\tilde{x}_j^i - \sum_{i=1}^{I} p^i \tilde{x}_j^i + 2\epsilon \\ &\geq u_j\Big(\sum_{i=1}^{I} x_j^i(p_k)\Big) - \sum_{i=1}^{I} l^i\big(\gamma^i(p_k)\big) x_j^i(p_k) \\ &\qquad - \sum_{i=1}^{I} p_k^i x_j^i(p_k) + \epsilon, \end{aligned}$$

contradicting the fact that $x(p_k)$ is a GWE for each p_k. This shows that $\tilde{x}$ is a GWE for the game with the prices given by p, and completes the proof. **Q.E.D.**

The next proposition establishes the monotonicity of flow rates in prices, and captures the intuitive notion that when prices are higher, users choose lower(no higher) flow rates. Although this result is intuitive, it can be seen through simple examples that concavity of utility functions is not strong enough to guarantee it, and we need strict concavity of the utility functions, as imposed in Assumption 2.1, for the result to hold.

Proposition 2.3: (Monotonicity of Flow Rates) Let Assumption 2.1 hold. Let p and $\tilde{p}$ be two price vectors such that $p \geq \tilde{p}$. Then, we have

$$\Gamma_j(p) \leq \Gamma_j(\tilde{p}), \qquad \forall\ j \in \mathcal{J},$$

where $\Gamma_j(p)$ and $\Gamma_j(\tilde{p})$ are the flow rates of user j at a GWE given prices p and $\tilde{p}$, respectively.

Proof: Let us partition the set of users into two sets R and S as

$$R = \big\{ r \in \mathcal{J} \mid \Gamma_r(p) > \Gamma_r(\tilde{p}) \big\},$$

$$S = \big\{ s \in \mathcal{J} \mid \Gamma_s(p) \leq \Gamma_s(\tilde{p}) \big\}.$$

We show that the set R is empty. Assume to arrive at a contradiction that the set R is nonempty. Define a subset of links as

$$I_{act} = \big\{ i \in \mathcal{I} \mid x_j^i(p) > 0, \text{ for some } j \in R \big\}.$$

We show that, for all $i \in I_{act}$, we have

$$\gamma^i(p) < \gamma^i(\tilde{p}), \tag{2.4}$$

and

$$x_s^i(\tilde{p}) = 0, \qquad \forall\ s \in S. \tag{2.5}$$

Let $i \in I_{act}$. This implies that $x_j^i(p) > 0$ for some $j \in R$. Using the first order optimality conditions, we obtain

$$u_j'\big(\Gamma_j(p)\big) - l^i\big(\gamma^i(p)\big) - p^i \geq u_j'\big(\Gamma_j(\tilde{p})\big) - l^i\big(\gamma^i(\tilde{p})\big) - \tilde{p}^i.$$

Since $j \in R$ and u_j is strictly concave, we have $u_j'\big(\Gamma_j(p)\big) < u_j'\big(\Gamma_j(\tilde{p})\big)$, which implies that

$$l^i\big(\gamma^i(p)\big) + p^i < l^i\big(\gamma^i(\tilde{p})\big) + \tilde{p}^i, \tag{2.6}$$

which together with $p \geq \tilde{p}$ yields

$$\gamma^i(p) < \gamma^i(\tilde{p}),$$

hence proving claim (2.4). To show (2.5), suppose to arrive at a contradiction, that for some $i \in I_{act}$, $x_s^i(\tilde{p}) > 0$ for some $s \in S$. This implies by the first order optimality conditions that

$$u_s'\big(\Gamma_s(\tilde{p})\big) - l^i\big(\gamma^i(\tilde{p})\big) - \tilde{p}^i \geq u_s'\big(\Gamma_s(p)\big) - l^i\big(\gamma^i(p)\big) - p^i.$$

Since $s \in S$ and u_j is concave, we have $u_s'\big(\Gamma_s(\tilde{p})\big) \leq u_s'\big(\Gamma_s(p)\big)$, which together with Eq. (2.6) implies that

$$\begin{aligned} l^i\big(\gamma^i(\tilde{p})\big) + \tilde{p}^i &\leq l^i\big(\gamma^i(p)\big) + p^i \\ &< l^i\big(\gamma^i(\tilde{p})\big) + \tilde{p}^i, \end{aligned}$$

thus yielding a contradiction and showing (2.5).

We next use Eqs. (2.4) and (2.5) to obtain

$$\sum_{i \in I_{act}} \gamma_i(\tilde{p}) = \sum_{i \in I_{act}} \sum_{j=1}^{J} x_j^i(\tilde{p}) = \sum_{i \in I_{act}} \sum_{j \in R} x_j^i(\tilde{p}) \leq \sum_{j \in R} \sum_{i=1}^{I} x_j^i(\tilde{p}) = \sum_{j \in R} \Gamma_j(\tilde{p}),$$

where the second equality follows from Eq. (2.5). We also have

$$\sum_{i \in I_{act}} \gamma_i(p) \geq \sum_{i \in I_{act}} \sum_{j \in R} x_j^i(p) = \sum_{j \in R} \sum_{i=1}^{I} x_j^i(p) = \sum_{j \in R} \Gamma_j(p),$$

where the second equality follows from definition of set R. The preceding sets of equations together with the definition of set R imply that

$$\sum_{i \in I_{act}} \gamma^i(p) > \sum_{i \in I_{act}} \gamma^i(\tilde{p}).$$

Summing Eq. (2.4) over all $i \in I_{act}$ yields a contradiction, thus proving that the set R is empty. **Q.E.D.**

The next proposition shows the monotonicity of link loads in prices. It establishes that link loads are monotonically non-increasing in their own prices and monotonically non-decreasing in other links' prices. Both of these are intuitive; the first implies that a higher price reduces the traffic on the link, while the second implies that, from the point of view of users, different links are substitutes—a higher price for one increases the traffic on the others.

Proposition 2.4: (Monotonicity of Link Loads) Let Assumption 2.1 hold. For some scalar $\epsilon > 0$ and some $t = 1, \dots, I$, let e_t denote the vector whose t^{th} component is ϵ and all the remaining components are equal to 0. Then, we have

$$\gamma^t(p + e_t) \leq \gamma^t(p), \tag{2.7}$$

$$\gamma^s(p + e_t) \geq \gamma^s(p), \qquad \forall\ s \neq t. \tag{2.8}$$

Proof: Let us partition the set of links into two sets R and S in the following way:

$$R = \big\{ r \in \mathcal{I} \mid \gamma^r(p + e_t) > \gamma^r(p) \big\},$$

$$S = \big\{ s \in \mathcal{I} \mid \gamma^s(p + e_t) \leq \gamma^s(p) \big\}.$$

Assume to arrive at a contradiction that $\gamma^t(p + e_t) > \gamma^t(p)$. This implies that $t \in R$. Define a subset of users as

$$\mathcal{J}_\epsilon = \big\{ j \mid x_j^r(p + e_t) > 0 \text{ for some } r \in R \big\}.$$

We show that, for all $j \in \mathcal{J}_\epsilon$, we have

$$\Gamma_j(p) > \Gamma_j(p + e_t), \tag{2.9}$$

and

$$x_j^s(p) = 0, \qquad \forall\ s \in S. \tag{2.10}$$

Since $x(p)$ and $x(p + e_t)$ are GWE's for the games with prices given by p and $p + e_t$, respectively, we have, by Definition 2.1, that for all $j = 1, \dots, J$,

$$x_j(p) \in \arg\max_{y_j \geq 0} \left\{ u_j\big(\Gamma_j(p)\big) - \sum_{i=1}^{I} l^i\big(\gamma^i(p)\big) y_j^i - \sum_{i=1}^{I} p^i y_j^i \right\},$$

$$x_j(p + e_t) \in \arg\max_{y_j \geq 0} \left\{ u_j\big(\Gamma_j(p + e_t)\big) - \sum_{i=1}^{I} l^i\big(\gamma^i(p + e_t)\big) y_j^i - \sum_{i \neq t} p^i y_j^i - (p^t + \epsilon) y_j^t \right\}.$$

These imply by the first order necessary optimality conditions that for all j and i,

$$\begin{aligned} u_j'\big(\Gamma_j(p)\big) - l^i\big(\gamma^i(p)\big) - p^i &\leq 0, \qquad \text{if } x_j^i(p) \geq 0, \\ &= 0, \qquad \text{if } x_j^i(p) > 0, \end{aligned} \tag{2.11}$$

and for all j and $i \neq t$,

$$u_j'\big(\Gamma_j(p+e_t)\big) - l^i\big(\gamma^i(p+e_t)\big) - p^i \leq 0, \qquad \text{if } x_j^i(p+e_t) \geq 0,$$
$$= 0, \qquad \text{if } x_j^i(p+e_t) > 0. \tag{2.12}$$

By definition, for any $j \in \mathcal{J}_\epsilon$, we have $x_j^r(p+e_t) > 0$ for some $r \in R$. This implies by Eqs. (2.11) and (2.12) that[1]

$$u_j'\big(\Gamma_j(p)\big) - l^r\big(\gamma^r(p)\big) - p^r \leq u_j'\big(\Gamma_j(p+e_t)\big) - l^r\big(\gamma^r(p+e_t)\big) - p^r. \tag{2.13}$$

Since $r \in R$ and the latency function l^r is strictly increasing, the preceding relation implies

$$u_j'\big(\Gamma_j(p)\big) < u_j'\big(\Gamma_j(p+e_t)\big),$$

from which, using the concavity of the utility function u_j, we obtain

$$\Gamma_j(p) > \Gamma_j(p+e_t),$$

thus proving claim (2.9). To show claim (2.10), suppose, to arrive at a contradiction, that for any $j \in \mathcal{J}_\epsilon$, we have $x_j^s(p) > 0$ for some $s \in S$. Since $t \notin S$ by assumption, this implies that $s \neq t$ and we obtain from Eqs. (2.11) and (2.12) that

$$u_j'\big(\Gamma_j(p+e_t)\big) - l^s\big(\gamma^s(p+e_t)\big) - p^s \leq u_j'\big(\Gamma_j(p)\big) - l^s\big(\gamma^s(p)\big) - p^s.$$

Using the preceding relation and the facts that the latency functions l^s are stricly increasing, and the utility functions u_j are concave, we obtain

$$\Gamma_j(p) \leq \Gamma_j(p+e_t),$$

contradicting Eq. (2.9), and showing that for any $j \in \mathcal{J}_\epsilon$, we have $x_j^s(p) = 0$ for all $s \in S$.

We next show that Eqs. (2.9) and (2.10) yield a contradiction. Note that we have

$$\sum_{j\in\mathcal{J}_\epsilon} \Gamma_j(p) = \sum_{j\in\mathcal{J}_\epsilon}\sum_{r=1}^{I} x_j^r(p) = \sum_{j\in\mathcal{J}_\epsilon}\sum_{r\in R} x_j^r(p) = \sum_{r\in R}\sum_{j\in\mathcal{J}_\epsilon} x_j^r(p)$$
$$\leq \sum_{r\in R}\sum_{j=1}^{J} x_j^r(p) = \sum_{r\in R}\gamma^r(p),$$

---

[1] Note that r could be equal to t, in which case this equation becomes

$$u_j'\big(\Gamma_j(p)\big) - l^t\big(\gamma^t(p)\big) - p^t \leq u_j'\big(\Gamma_j(p+e_t)\big) - l^t\big(\gamma^t(p+e_t)\big) - p^t - \epsilon$$
$$\leq u_j'\big(\Gamma_j(p+e_t)\big) - l^t\big(\gamma^t(p+e_t)\big) - p^t,$$

and thus the same subsequent analysis applies.

where the second equality follows from Eq. (2.10). We also have

$$\begin{aligned}
\sum_{j\in\mathcal{J}_\epsilon}\Gamma_j(p+e_t) = \sum_{j\in\mathcal{J}_\epsilon}\sum_{r=1}^{I} x_j^r(p+e_t) &= \sum_{r=1}^{I}\sum_{j\in\mathcal{J}_\epsilon} x_j^r(p+e_t) \\
&= \sum_{r\in R}\sum_{j\in\mathcal{J}_\epsilon} x_j^r(p+e_t) + \sum_{r\notin R}\sum_{j\in\mathcal{J}_\epsilon} x_j^r(p+e_t) \\
&\geq \sum_{r\in R}\sum_{j\in\mathcal{J}_\epsilon} x_j^r(p+e_t) \\
&= \sum_{r\in R}\sum_{j=1}^{J} x_j^r(p+e_t) \\
&= \sum_{r\in R}\gamma^r(p+e_t),
\end{aligned}$$

where the second equality from the end follows from the definition of set $\mathcal{J}_\epsilon$ The two preceding sets of equations together with the definition of set R imply that

$$\sum_{j\in\mathcal{J}_\epsilon}\Gamma_j(p+e_t) > \sum_{j\in\mathcal{J}_\epsilon}\Gamma_j(p).$$

Summing Eq. (2.9) over all $j\in\mathcal{J}_\epsilon$ yields a contradiction, thus showing that $t\notin R$, proving Eq. (2.7).

Next we show Eq. (2.8). Let $s\neq t$. If $\gamma^s(p)=0$, then we are done. Assume that $\gamma^s(p)>0$. This implies that there exists some k such that $x_k^s(p)>0$, from which, using the first order conditions, we get

$$u_k'\big(\Gamma_k(p)\big) - l^s\big(\gamma^s(p)\big) - p^s \geq u_k'\big(\Gamma_k(p+e_t)\big) - l^s\big(\gamma^s(p+e_t)\big) - p^s.$$

Combining the preceding with the fact that $\Gamma_k(p)\geq\Gamma_k(p+e_t)$ [cf. Proposition 2.3], we obtain

$$\gamma^s(p+e_t)\geq\gamma^s(p),$$

thus proving the desired claim. **Q.E.D.**

The above propositions provided some general results about the continuity and monotonicity of GWE's. We now establish two lemmas which enable a sharper characterization of the GWE. This characterization will be useful in the analysis of subsequent sections.

Lemma 2.1: Let Assumption 2.1 hold. For a given $p\geq 0$, let γ^i be the load of link i at a GWE. Then, for all i with $\gamma^i>0$, we have

$$p^i + l^i(\gamma^i) = \min_{m\in\mathcal{I}}\big\{p^m + l^m(\gamma^m)\big\}.$$

Proof: Since $\gamma^i > 0$, there exists some j such that $x_j^i > 0$ and satisfies

$$u_j'(\Gamma_j) - l^i(\gamma^i) - p^i = 0.$$

We also have

$$u_j'(\Gamma_j) - l^m(\gamma^m) - p^m \leq 0, \qquad \forall\, m,$$

[cf. Eq. (2.11)]. Combining the preceding two relations, we obtain

$$p^i + l^i(\gamma^i) \leq p^m + l^m(\gamma^m), \qquad \forall\, m,$$

showing the result. **Q.E.D.**

Let us define the effective cost to users of link i as $p^i + l^i(\gamma^i)$, which is the monetary cost plus delay cost of using the link. The lemma states that the effective cost of all links with positive flow must be equal. This is very intuitive; since the links are perfect substitutes from the viewpoint of users, they must all have the same effective cost.

Lemma 2.2: Let Assumption 2.1 hold. For a given $p \geq 0$, let x be a GWE. Define the sets $\overline{\mathcal{I}} = \{i \mid \gamma^i > 0\}$ and $\overline{\mathcal{J}} = \{j \mid \Gamma_j > 0\}$.

(a) Then for all $i \in \overline{\mathcal{I}}$ and $j \in \overline{\mathcal{J}}$, we have

$$u_j'(\Gamma_j) - l^i(\gamma^i) - p^i = 0.$$

(b) There exists a GWE $\tilde{x}$ at this price vector that satisfies $\tilde{x}_j^i > 0$ for all $i \in \overline{\mathcal{I}}$ and $j \in \overline{\mathcal{J}}$.

Proof:

(a) Let $i \in \overline{\mathcal{I}}$ and $j \in \overline{\mathcal{J}}$. Since $\Gamma_j > 0$ by assumption, it follows that there exists some link s such that $x_j^s > 0$, which implies by the first order conditions that

$$u_j'(\Gamma_j) - l^s(\gamma^s) - p^s = 0.$$

Since $\gamma^s > 0$ and $\gamma^i > 0$, we have by Lemma 2.1 that

$$l^i(\gamma^i) + p^i = l^s(\gamma^s) + p^s.$$

Substituting this in the previous equation yields the desired result.

(b) Consider the vector $\tilde{x} = [\tilde{x}_j^i]$ generated in the following way: Set $\tilde{x}_j^i = 0$ if $i \notin \overline{\mathcal{I}}$ or $j \notin \overline{\mathcal{J}}$. Let $i \in \overline{\mathcal{I}}$ and $j \in \overline{\mathcal{J}}$. If $x_j^i > 0$, then set $\tilde{x}_j^i = x_j^i$. Assume $x_j^i = 0$. Then, since $i \in \overline{\mathcal{I}}$ and $j \in \overline{\mathcal{J}}$, there exists some $k \neq j$ and $s \neq i$ such that $x_k^i > 0$ and $x_j^s > 0$. We set

$$\tilde{x}_j^i = \epsilon, \qquad \tilde{x}_j^s = x_j^s - \epsilon,$$

$$\tilde{x}_k^i = x_k^i - \epsilon, \qquad \tilde{x}_k^s = x_k^s + \epsilon,$$

where $\epsilon > 0$ is small enough such that all of the above terms are positive. It can be seen that the link loads and the input flow rates of users are kept constant by this transformation (to see this, add vertically to see the change in link loads and horizontally to see the change in flow rates). Also, since $k \in \overline{\mathcal{J}}$ and $s \in \overline{\mathcal{I}}$, it follows from part (a) that

$$u_k'(\Gamma_k) - l^s(\gamma^s) - p^s = 0.$$

This implies that $\tilde{x}$ satisfies the first order necessary and sufficient optimality conditions, and therefore is a GWE at the price vector p. **Q.E.D.**

The second part of this lemma exploits the fact that the allocation of the unique GWE flow rates and link loads across individuals is indeterminate, and shows that starting from any GWE, we can always construct an alternative GWE, with the same individual flow rates and link loads, in which all individual flows to all links with minimum effective cost are positive.

3. CONVERGENCE OF NASH EQUILIBRIA TO GENERALIZED WARDROP EQUILIBRIUM

The purpose of this section is to show that the link loads at a GWE are limits of the link loads at the Nash Equilibria of a sequence of games, where users recognize their impact on congestion.

The key difference between the GWE and the Nash Equilibrium is that in the latter, each user is no longer a "link-load" taker, but instead anticipates the effect of its actions on the congestion on the links. Therefore, to define the Nash equilibrium, we use a different notation for the payoff functions of the users. We let x_{-j} denote the vector of flows of all users except user j. Given (x_j, x_{-j}) and a price vector p, we denote the payoff function of user j by

$$V_j(x_j, x_{-j}; p) = u_j\Big(\sum_{i=1}^{I} x_j^i\Big) - \sum_{i=1}^{I} l^i\Big(\sum_{k\neq j} x_k^i + x_j^i\Big)x_j^i - \sum_{i=1}^{I} p^i x_j^i. \tag{3.1}$$

Nash equilibrium of this game is then defined as follows.

Definition 3.2: (Nash Equilibrium) For a given price vector $p \geq 0$, a flow vector $\tilde{x}$ is a *Nash equilibrium* (NE) of the game where the payoff functions are given by (3.1) if

$$\tilde{x}_j \in \arg\max_{x_j \geq 0} V_j(x_j, \tilde{x}_{-j}; p), \qquad \forall\, j \in \mathcal{J}.$$

Let Assumption 2.1 hold. Further assume that the latency function l^i is convex for all i. It can be seen, using fixed point arguments, that there exists an NE under these assumptions (cf. [Ros65]).

In problems where users only make routing decisions, the relation between the NE and the Wardrop equilibrium has been studied in [HaM85]. The result that we show next illustrates that a similar relation exists between NE and GWE in our model. Similar to [HaM85], we use a replication strategy idea, which was pioneered by Debreu and Scarf [DeS63]. For this purpose, let us define $G(n)$ to be the game where we have J classes of users, denoted by $N_1, \ldots, N_j$, of n identical users each, with the payoff function of user h for some $h \in N_j$ given by

$$V_h^{(n)}(x_j, x_{-j}; p) = u_j\Big(\sum_{i=1}^{I} x_h^i\Big) - \sum_{i=1}^{I} l^{(n)i}\left(\sum_{k \neq h} x_k^i + x_h^i\right) x_h^i - \sum_{i=1}^{I} p^i x_h^i,$$

where $l^{(n)i}$ is the latency function of link i in game $G(n)$.

It is important to change the latency functions with replication, since, otherwise, the demand for resources in the network will increase as n increases, but available resources would remain constant. The most natural way of doing this is to also increase the capacity of each link with n. To do this, we define

$$l^{(n)i}(\gamma^i) = l^i\left(\frac{\gamma^i}{n}\right), \tag{3.2}$$

which ensures that network resources grow at the same rate as the demand for these resources [cf. Assumption 2.1(c')].

Given these definitions, we can establish the following result.

Proposition 3.5: (Convergence) Let Assumption 2.1 hold. For a given $p \geq 0$, let γ^i denote the total load on link i corresponding to the GWE of the J-player game, where the payoff functions of the users are given by Eq. (2.2). There exists an NE of the game $G(n)$, $\tilde{x}(n)$, for which the corresponding link load $\tilde{\gamma}^i(n) = \sum_{j=1}^{J} \sum_{h \in N_j} \tilde{x}_h^i(n)$ satisfies

$$\lim_{n \to \infty} \frac{\tilde{\gamma}^i(n)}{n} = \gamma^i, \qquad \forall\, i \in \mathcal{I}.$$

Proof: Since all players in class N_j are identical, there exists a Nash equilibrium for which all players of N_j have the same strategy. We can represent this Nash equilibrium as

$$\tilde{x}_h(n) = \frac{\tilde{x}_j(n)}{n}, \qquad \forall\, h \in N_j, \forall\, j,$$

where $\tilde{x}_j(n)$ is the I-dimensional total flow vector of class j on all links. In view of Assumption 2.1, it can be seen that each component of $\tilde{x}_h(n)$ lies in the compact set $[0, C^i]$, and therefore the sequence $\{\tilde{x}_h(n)\}$ is bounded. Assume without loss of generality that $\tilde{x}_h(n) \to \overline{x}_h$ for each h (and also, by implication, $\tilde{\Gamma}_h(n) = \sum_{i\in\mathcal{I}} \tilde{x}_h^i(n) \to \overline{\Gamma}_h$). Since $\tilde{x}(n)$ is a Nash Equilibrium, we have by the necessary optimality conditions, for all i, all j, and all $h \in N_j$,

$$\Big(u_j'\big(\tilde{\Gamma}_h(n)\big) - l^{(n)i}\big(\tilde{\gamma}^i(n)\big) - (l^{(n)i})'\big(\tilde{\gamma}^i(n)\big)\tilde{x}_h^i(n) - p^i\Big)\big(y_h^i - \tilde{x}_h^i(n)\big) \le 0, \qquad \forall\, y_h^i \ge 0,$$

or equivalently,

$$\left(u_j'\big(\tilde{\Gamma}_h(n)\big) - l^i\left(\frac{\tilde{\gamma}^i(n)}{n}\right) - (l^i)'\left(\frac{\tilde{\gamma}^i(n)}{n}\right)\tilde{x}_h^i(n) - p^i\right)\big(y_h^i - \tilde{x}_h^i(n)\big) \le 0, \qquad \forall\, y_h^i \ge 0,$$

[cf. Eq. (3.2)]. The sequence $\frac{\tilde{\gamma}^i(n)}{n} = \frac{1}{n}\sum_j \sum_{h\in N_j} \tilde{x}_h^i(n)$, is bounded, and satisfies $\frac{\tilde{\gamma}^i(n)}{n} \to \overline{\gamma}^i = \sum_j \overline{x}_{h_j}^i$ for some $h_j \in N_j$ and for all i. Taking the limit in the preceding relation, we obtain

$$\big(u_j'(\overline{\Gamma}_h) - l^i(\overline{\gamma}^i) - p^i\big)\big(y_h^i - \overline{x}_h^i\big) \le 0, \qquad \forall\, y_h^i \ge 0.$$

In view of the concavity assumptions (cf. Assumption 2.1), this implies, by the necessary and sufficient optimality conditions for a GWE, that the vector $x = [x_1, \ldots, x_J]$, where $x_j = \overline{x}_h$ for some $h \in N_j$, is a GWE at the price vector p. By Proposition 2.1, the corresponding link loads γ^i are uniquely defined, showing that $\overline{\gamma}^i = \gamma^i$. **Q.E.D.**

4. PRICE DETERMINATION BY PROFIT MAXIMIZATION

In Section 2, we characterized the flow rates and the link loads at a GWE for a given price vector. In this section, we show the existence of profit-maximizing prices for the monopolist service provider and provide a characterization of these prices. The monopolist sets the prices as the optimal solution of the problem, [1]

$$\begin{aligned} &\text{maximize } \sum_{i\in\mathcal{I}} p^i\gamma^i(p) \\ &\text{subject to } \; p \ge 0, \end{aligned} \tag{4.1}$$

[1] This formulation ignores the costs incurred by the monopolist in data transmission. It is straightforward to add a constant marginal cost of transmission, without affecting the results. We do not do so here to reduce notation.

where $\gamma^i(p)$ is the load of link i at a GWE given price vector p; i.e., if $x(p)$ is a GWE, as given in Definition 2.1, then

$$\gamma^i(p) = \sum_{j \in \mathcal{J}} x_j^i(p).$$

We start by establishing the existence of profit-maximizing prices.

Proposition (Existence): Let Assumption 2.1 hold. Then problem (4.1) has an optimal solution, denoted by p^*.

Proof: By Proposition 2.2, we have that $\gamma^i(p)$ is a continuous function of p. Moreover, using Assumption 2.1(b), we can see that the upper level sets of Problem (4.1) are compact. Then it follows that problem (4.1) has an optimal solution. **Q.E.D.**

We will refer to p^* as the *monopoly equilibrium price* of the two-stage game, where the monopolist service provider (Stackelberg leader) sets prices and each price defines a subgame among the users. Let $x(p^*)$ be a GWE given price vector p^*. We will refer to $\big(p^*, x(p^*)\big)$, or more simply $\big(p^*, x^*\big)$, as the *monopoly equilibrium* (ME) of the overall game.

In the following, we show various characteristics of the equilibrium prices.

Proposition 4.6: Let Assumption 2.1 hold. Let p be the vector of monopoly equilibrium prices. Then

$$p^i > 0, \qquad \forall\ i \in \mathcal{I}.$$

Proof: Assume, to arrive at a contradiction, that $p^s = 0$ for some s. We first show that this implies that $\gamma^s > 0$. Assume the contrary, i.e., $\gamma^s = 0$. This implies by the first order optimality conditions and Assumption 2.1 (u_j is nondecreasing and $l^i(0) = 0$), that

$$u_j'(\Gamma_j) = 0, \qquad \forall\ j. \tag{4.2}$$

But this implies that $\gamma^i = 0$ for all i. [If $\gamma^i > 0$ for some i, we would have $-l^i(\gamma^i) - p^i < 0$ for all j, implying that $x_j^i = 0$ for all j , and we get a contradiction.] But, since $\sum_i \gamma^i = \sum_j \Gamma_j$, this implies that $\Gamma_j = 0$ for all j, which is a contradiction by Eq. (4.2) [cf. Assumption 2.1(b)], and shows that we must have $\gamma^s > 0$.

We next consider the price vector $\tilde{p} = p + e_s$, where e_s is a vector whose components are all equal to 0, except the s^{th} component, which is equal to some $\epsilon > 0$. By Proposition 2.4, we have

$$\gamma^s(\tilde{p}) \leq \gamma^s(p),$$

$$\gamma^i(\tilde{p}) \geq \gamma^i(p), \qquad \forall\ i \neq s.$$

We choose ϵ small enough such that $\gamma^s(\tilde{p}) > 0$. [This can be done given the continuity of γ^i in p, cf. Proposition 2.2]. This shows that

$$\sum_{i \in \mathcal{I}} p^i \gamma^i(p) < \sum_{i \in \mathcal{I}} \tilde{p}^i \gamma^i(\tilde{p}),$$

thus contradicting the fact that (p, x) is an ME, and showing that $p^i > 0$ for all i. **Q.E.D.**

This proposition shows that the monopolist will charge positive prices for all links. This result is intuitive. A zero price for link i would hurt the monopolist in two ways; first, directly by not generating any revenues from link i, and second, indirectly, by reducing traffic and profits on other links, in view of our result in Proposition 2.4.

Proposition 4.7: Let Assumption 2.1 hold. Let (p, x) be an ME. Then, the corresponding link loads satisfy

$$\gamma^i > 0, \qquad \forall\ i \in \mathcal{I}.$$

Proof: Define $\overline{\mathcal{I}} = \{i \mid \gamma^i > 0\}$. Assume, to arrive at a contradiction, that $\gamma^s = 0$ for some s. Using Lemma 2.1 and Assumption 2.1(b'), this implies that

$$K = p^i + l^i(\gamma^i) \leq p^s + l^s(0) = p^s, \qquad \forall\ i \in \overline{\mathcal{I}},$$

for some $K > 0$. Since $l^i(\gamma^i) > 0$ for all $i \in \overline{\mathcal{I}}$, we can choose some $\epsilon > 0$ such that

$$p^i < p^s - \epsilon < p^i + l^i(\gamma^i), \qquad \forall\ i \in \overline{\mathcal{I}}. \tag{4.3}$$

We next consider the price vector $\tilde{p} = p - e_s$, where e_s is a vector whose components are all equal to 0, except the s^th component, which is equal to ϵ. By Proposition 2.4, we have

$$\gamma^s(\tilde{p}) \geq \gamma^s(p),$$

$$\gamma^i(\tilde{p}) \leq \gamma^i(p), \qquad \forall\ i \neq s. \tag{4.4}$$

This shows that if $\gamma^i(p) = 0$ for some $i \neq s$, then $\gamma^i(\tilde{p}) = 0$. By Proposition 2.3, we also have

$$\Gamma_j(\tilde{p}) \geq \Gamma_j(p), \qquad \forall\ j, \tag{4.5}$$

which, after summing over all j, yields

$$\sum_{m \in \mathcal{I}} \gamma^m(\tilde{p}) \geq \sum_{m \in \mathcal{I}} \gamma^m(p). \tag{4.6}$$

There are two cases to consider:

Case 1: $\gamma^i(\tilde{p}) = \gamma^i(p)$, for all $i \in \overline{\mathcal{I}}$. In this case, we have that $\gamma^s(\tilde{p}) > 0$. To see this, note that, if $\gamma^s(\tilde{p}) = 0$, we would have by Lemma 2.1 that

$$p^s - \epsilon \geq p^i + l^i\big(\gamma^i(\tilde{p})\big), \qquad \forall\, i \in \overline{\mathcal{I}},$$

which contradicts Eq. (4.3). But this implies that

$$\sum_{m\in\mathcal{I}} p^m \gamma^m(p) < \sum_{m\in\mathcal{I}} \tilde{p}^m \gamma^m(\tilde{p})$$

thus contradicting the fact that (p, x) is an ME.

Case 2: If $\gamma^i(\tilde{p}) < \gamma^i(p)$ for some $i \in \overline{\mathcal{I}}$, then it follows from Eq. (4.6) that $\gamma^s(\tilde{p}) > 0$. Then the change in the profit can be written as

$$\begin{aligned}
\sum_{m\in\mathcal{I}} \tilde{p}^m \gamma^m(\tilde{p}) - \sum_{m\in\mathcal{I}} p^m \gamma^m(p) &= (p^s - \epsilon)\gamma^s(\tilde{p}) + \sum_{m\neq s} p^m \big(\gamma^m(\tilde{p}) - \gamma^m(p)\big) \\
&> (p^s - \epsilon)\gamma^s(\tilde{p}) + (p^s - \epsilon) \sum_{m\neq s} \big(\gamma^m(\tilde{p}) - \gamma^m(p)\big) \\
&= (p^s - \epsilon) \sum_j \big(\Gamma_j(\tilde{p}) - \Gamma_j(p)\big) \\
&\geq 0,
\end{aligned}$$

where the strict inequality follows from Eqs. (4.3) and (4.4), and the last inequality follows from Eq. (4.5). But this again contradicts the fact that (p, x) is an ME.

This shows that for all i, we must have $\gamma^i > 0$, and completes the proof. **Q.E.D.**

This proposition establishes that the monopolist would not choose to have zero link load on any link. Intuitively, allowing some positive traffic on a link would provide positive profits on that link. The proof is somewhat more complicated than this observation, however, because it also creates the indirect effect of reducing the loads on other links, thus reducing the rest of the monopolist's profits. The proof amounts to showing that this indirect effect is always dominated by the direct effect.

Note also that Assumption 2.1(b'), which imposes that $l^i(0) = 0$ for all i is important for this result. If some link i had $l^i(0) > 0$, our previous results remain unchanged, but it would be possible for the monopolist to have zero link load on some link.

We are now ready to provide the characterization of equilibrium prices. Before doing so, however, it is convenient to establish a lemma which will be useful in deriving this characterization.

Lemma 4.3: Let Assumption 2.1 hold. Let (p, x) be an ME, and $\overline{\mathcal{J}} = \{j \in \mathcal{J} \mid \Gamma_j > 0\}$. Assume without loss of generality that $1 \in \overline{\mathcal{J}}$. Then $\left(p, \gamma, [\Gamma_j]_{j \in \overline{\mathcal{J}}}\right)$ is an optimal solution of the following problem,

$$\begin{aligned}
\text{maximize} \ & \sum_{i \in \mathcal{I}} p^i \gamma^i \\
\text{subject to} \ & u_1'(\Gamma_1) - l^i(\gamma^i) - p^i = 0, \qquad \forall\, i \in \mathcal{I}, \\
& u_j'(\Gamma_j) - l^1(\gamma^1) - p^1 = 0, \qquad \forall\, j \in \overline{\mathcal{J}} - \{1\}, \\
& \sum_{i \in \mathcal{I}} \gamma^i = \sum_{j \in \overline{\mathcal{J}}} \Gamma_j, \\
& \gamma^i \geq 0,\ \Gamma_j \geq 0, \qquad \forall\, i \in \mathcal{I},\ j \in \overline{\mathcal{J}}.
\end{aligned} \tag{4.7}$$

Proof: Assume to arrive at a contradiction that there exists some vector $\left(\overline{p}, \overline{\gamma}, [\overline{\Gamma}_j]_{j \in \overline{\mathcal{J}}}\right)$ feasible for problem (4.7) such that

$$\sum_{i \in \mathcal{I}} \overline{p}^i \overline{\gamma}^i > \sum_{i \in \mathcal{I}} p^i \gamma^i. \tag{4.8}$$

We claim that this implies

$$\overline{\gamma}^i \leq \gamma^i(\overline{p}), \qquad \forall\, i \in \mathcal{I}, \tag{4.9}$$

where $\gamma^i(\overline{p})$ is the load of link i at a GWE given $\overline{p}$. Suppose the preceding relation does not hold for some s, i.e., $\overline{\gamma}^s > \gamma^s(\overline{p})$. Since $\left(\overline{p}, \overline{\gamma}, [\overline{\Gamma}_j]_{j \in \overline{\mathcal{J}}}\right)$ is a feasible solution for problem (4.7), it can be seen that we have

$$u_k'(\overline{\Gamma}_k) - l^s(\overline{\gamma}^s) - \overline{p}^s = 0, \qquad \forall\, k \in \overline{\mathcal{J}}.$$

We also have using the first order optimality conditions for a GWE that

$$u_k'\big(\Gamma_k(\overline{p})\big) - l^s\big(\gamma^s(\overline{p})\big) - \overline{p}^s \leq 0, \qquad \forall\, k \in \mathcal{J},$$

where $\Gamma_k(\overline{p})$ is the flow rate of user k at a GWE given $\overline{p}$. Since $\overline{\gamma}^s > \gamma^s(\overline{p})$, we obtain

$$\Gamma_k(\overline{p}) > \overline{\Gamma}_k, \qquad \forall\, k \in \overline{\mathcal{J}}. \tag{4.10}$$

Moreover, we have from the feasibility of $\left(\overline{p}, \overline{\gamma}, [\overline{\Gamma}_j]_{j \in \overline{\mathcal{J}}}\right)$ that

$$l^i(\overline{\gamma}^i) + \overline{p}^i = l^s(\overline{\gamma}^s) + \overline{p}^s, \qquad \forall\, i \in \mathcal{I}.$$

It also follows from Lemma 2.1 that for all i with $\gamma^i(\overline{p}) > 0$, we have

$$l^i\big(\gamma^i(\overline{p})\big) + \overline{p}^i \leq l^s\big(\gamma^s(\overline{p})\big) + \overline{p}^s.$$

Since $\overline{\gamma}^s > \gamma^s(\overline{p})$, the preceding relations imply that

$$\gamma^i(\overline{p}) < \overline{\gamma}^i, \qquad \forall\ i \text{ with } \gamma^i(\overline{p}) > 0,$$

and therefore that

$$\gamma^i(\overline{p}) \le \overline{\gamma}^i, \qquad \forall\ i \in \mathcal{I}.$$

Summing both sides of this relation over all $i \in \mathcal{I}$, we obtain

$$\sum_{k\in\overline{\mathcal{J}}} \overline{\Gamma}_k = \sum_{i\in\mathcal{I}} \overline{\gamma}^i \ge \sum_{i\in\mathcal{I}} \gamma^i(\overline{p}) = \sum_{k\in\mathcal{J}} \Gamma_k(\overline{p}) \ge \sum_{k\in\overline{\mathcal{J}}} \Gamma_k(\overline{p}) > \sum_{k\in\overline{\mathcal{J}}} \overline{\Gamma}_k,$$

which is a contradiction [the last strict inequality uses Eq. (4.10)]. This implies that Eq. (4.9) holds for all $i \in \mathcal{I}$. Combining Eqs. (4.8) and (4.9), we obtain

$$\sum_{i\in\mathcal{I}} \overline{p}^i \gamma^i(\overline{p}) > \sum_{i\in\mathcal{I}} p^i \gamma^i.$$

But this contradicts the fact that (p, x) is an ME, hence proves the desired result. **Q.E.D.**

Using this lemma, we can prove the following proposition.

Proposition 4.8 (Equilibrium Price Characterization): Let Assumption 2.1 hold. Assume further that u_j is twice continuously differentiable for each j, and l^i is continuously differentiable for each i. Let (p, x) be an ME, and $\overline{\mathcal{J}} = \{j \mid \Gamma_j > 0\}$. Then, for all $i \in \mathcal{I}$, we have

$$p^i = (l^i)'(\gamma^i)\gamma^i + \frac{\sum_{m\in\mathcal{I}} \gamma^m}{-\sum_{j\in\overline{\mathcal{J}}} \frac{1}{u_j''(\Gamma_j)}}.$$

Proof: By Lemma 4.3, it follows that $\left(p, \gamma, [\Gamma_j]_{j\in\overline{\mathcal{J}}}\right)$ is an optimal solution of problem (4.7). To simplify the notation, let us denote $\Gamma = [\Gamma_j]_{j\in\overline{\mathcal{J}}}$. One can check that, in view of the assumptions that u_j is strictly concave for all j and l^i is strictly increasing for all i, the constraint gradients of this problem are linearly independent at (p, γ, Γ), i.e., (p, γ, Γ) is a *regular point* (for the proof, see the Appendix). Hence, problem (4.7) admits Lagrange multipliers (see [BNO03], Chapter 5). The Lagrangian function $L(p, \gamma, \Gamma)$ for this problem is obtained by assigning the multipliers λ^i, $i \in \mathcal{I}$, to the first set of equality constraints in problem (4.7), the multipliers μ_j , $j \in \overline{\mathcal{J}} - \{1\}$, to the second set of equality constraints, and μ_1 to the last equality constraint; i.e.,

$$\begin{aligned} L(p, x, \lambda) = {} & \sum_{i\in\mathcal{I}} p^i\gamma^i + \sum_{i\in\mathcal{I}} \lambda^i \big[u_1'(\Gamma_1) - l^i(\gamma^i) - p^i\big] \\ & + \sum_{j\in\overline{\mathcal{J}}-\{1\}} \mu_j \big[u_j'(\Gamma_j) - l^1(\gamma^1) - p^1\big] + \mu_1 \Big[\sum_{i\in\mathcal{I}} \gamma^i - \sum_{j\in\overline{\mathcal{J}}} \Gamma_j\Big]. \end{aligned}$$

Using the first order necessary optimality conditions at the optimal solution (p, γ, Γ) for problem (4.7), together with Proposition 4.6 ($p_i > 0$ for all i), Proposition 4.7 ($\gamma^i > 0$ for all i), and the fact that $\Gamma_j > 0$ for all $j \in \overline{\mathcal{J}}$, we obtain

$$\gamma^1 = \lambda^1 + \sum_{j \in \overline{\mathcal{J}} - \{1\}} \mu_j, \tag{4.11}$$

$$\gamma^i = \lambda^i, \qquad \forall\, i \neq 1, \tag{4.12}$$

and

$$p^1 - \lambda^1 (l^1)'(\gamma^1) - \Big(\sum_{j \in \overline{\mathcal{J}} - \{1\}} \mu_j \Big) (l^1)'(\gamma^1) + \mu_1 = 0,$$

$$p^i - \lambda^i (l^i)'(\gamma^i) + \mu_1 = 0, \qquad \forall\, i \neq 1.$$

Using Eqs. (4.11) and (4.12), the preceding two relations can be rewritten as

$$p^i - (l^i)'(\gamma^i)\gamma^i + \mu_1 = 0, \qquad \forall\, i \in \mathcal{I}. \tag{4.13}$$

Taking the partial derivatives of $L(p, \gamma, \Gamma)$ with respect to Γ_j also yields

$$u_1''(\Gamma_1) \sum_{i \in \mathcal{I}} \lambda^i - \mu_1 = 0,$$

$$\mu_j u_j''(\Gamma_j) - \mu_1 = 0, \qquad \forall\, j \in \overline{\mathcal{J}} - \{1\}.$$

Adding Eqs. (4.11) and (4.12) over all i, and using the preceding, we obtain

$$\sum_{m \in \mathcal{I}} \gamma^m = \mu_1 \sum_{j \in \overline{\mathcal{J}}} \frac{1}{u_j''(\Gamma_j)}.$$

Substituting the previous relation for μ_1 in (4.13) yields the desired result. **Q.E.D.**

This proposition is the central result of this section. An important implication is that the profit-maximizing price for the monopolist service provider consists of two markups (above the marginal cost of the monopolist, which is equal to zero here):[1] the first is $(l^i)'(\gamma^i)\gamma^i$. We will see in the next section that this term essentially internalizes the congestion externality ignored by the users, while the second is a further monopoly markup over this. It is important to emphasize the intuition for the term $(l^i)'(\gamma^i)\gamma^i$: the monopolist realizes that individual users' willingness to pay is reduced by congestion, and charges a greater price to each user in order to reduce congestion. This term is the reason why monopoly pricing may improve the allocation of resources.

The second effect is a further markup, similar to the monopoly markup in the standard economic models. We will see below that this term causes further reduction in the flow rates, and typically creates a distortion relative to the social optimum.

[1] The markup is defined as the difference between price and marginal cost.

5. PERFORMANCE COMPARISON WITH SOCIAL OPTIMUM

A network planner with full information and centralized control of the system allocates the resources as the optimal solution of the following problem:

$$\begin{aligned} \text{maximize} \quad & \sum_{j \in \mathcal{J}} u_j(\Gamma_j) - \sum_{i \in \mathcal{I}} l^i(\gamma^i)\gamma^i \\ \text{subject to} \quad & \Gamma_j = \sum_{i \in \mathcal{I}} x_j^i, \qquad \forall\, j \in \mathcal{J} \\ & \gamma^i = \sum_{j \in \mathcal{J}} x_j^i, \qquad \forall\, i \in \mathcal{I} \\ & x_j^i \geq 0, \qquad \forall\, i \in \mathcal{I},\ j \in \mathcal{J}. \end{aligned} \tag{5.1}$$

We call this problem the *social problem* and the optimal solution of this problem (which exists under Assumption 2.1) the *social optimum.* Note that the social problem is defined in terms of link load and flow allocations, and the prices do not appear in the problem. This is consistent with both studies in general equilibrium analysis (see [Deb59], [ArH71]) and also our objective in this framework; i.e, the goal of the network manager is to maximize the social surplus of the system. The payments are merely transfers between different agents in the system.

Assuming that the objective function of the social problem is concave, an equivalent characterization of the social optimum, denoted by $\tilde{x}$ is given by the first order conditions,

$$u_j'(\tilde{\Gamma}_j) - l^i(\tilde{\gamma}^i) - (l^i)'(\tilde{\gamma}^i)\tilde{\gamma}^i \begin{cases} \leq 0 & \text{if } \tilde{x}_j^i \geq 0, \\ = 0 & \text{if } \tilde{x}_j^i = 0. \end{cases}$$

It can be seen from these conditions that if each user is charged $(l^i)'(\tilde{\gamma}^i)\tilde{\gamma}^i$, the allocation resulting from the corresponding GWE will be the same as the allocation of the social optimum. This amount is called the *marginal congestion cost* or the *Pigovian tax.* In a decentralized implementation, this price is charged to the users to internalize the congestion externalities that they ignore.

From the ME price characterization given in Proposition 4.8, we see that the ME price is the sum of the marginal congestion cost and a markup. Hence, even if the monopolist is a selfish agent maximizing profits, he internalizes the congestion externalities. The reason is that the monopolist wants to charge higher prices when the demand is more inelastic, i.e., the change in demand as a function of the changes in price is low. The demand inelasticity in this problem is related to the marginal congestion cost. The monopolist realizes that if it charges a high price for a particular route, thus reducing traffic on that route, this may actually not reduce the attraction of this route to users by much, because there will be a corresponding decrease in congestion.

In this section, we explore the relationship between the flow rates resulting from the ME, the social optimum, and the GWE at zero prices, with the goal of illustrating the effect of congestion externalities. We also illustrate with an example why monopoly pricing may improve the performance relative to the zero price case.

Proposition 5.9: Let Assumption 2.1 hold. Assume further that u_j is twice continuously differentiable for each j, l^i is continuously differentiable for each i, and the function $\tilde{l}^i$, defined by $\tilde{l}^i(x) = l^i(x)x$ for all $x \geq 0$, is convex for each i. Let $\overline{x}$ denote the GWE at 0 prices, $\tilde{x}$ denote the social optimum, and x denote the ME. For all $j \in \mathcal{J}$, we have,

$$\Gamma_j \leq \tilde{\Gamma}_j \leq \overline{\Gamma}_j.$$

Proof: See the Appendix.

This result shows that when the prices are equal to 0, the users generate too much flow relative to the social optimum since they ignore the congestion externalities. Monopoly pricing improves this behavior; however it may introduce distortion relative to the social optimum due to flow control decisions of the users.

The next example compares the performance of monopoly pricing to the performance of GWE at 0 prices, and illustrates the effects of different utility and latency functions on the performance.

Example 5.1

Consider and network with J identical users and a single link. Assume that the utility function of the users $u(x)$ and the latency function of the link $l(x)$ are given by

$$u(x) = x^a, \qquad \text{for some } 0 < a \leq 1,$$

$$l(x) = \frac{1}{J^b} x^b, \qquad \text{for some } b \geq 0.$$

By symmetry, it can be seen that at any GWE with any price, the flow rate of each user is the same. Let x_G denote the flow rate of a single user at the GWE with prices set equal to 0. Similarly let x_M denote the flow rate of a single user at the ME. Using the optimality conditions, we compute x_G and x_M as

$$x_G = a^{\frac{1}{1+b-a}},$$

$$x_M = \left(\frac{a^2}{1+b}\right)^{\frac{1}{1+b-a}}.$$

Let U_G and U_M denote the total utility that the system gets at the GWE and ME, respectively, i.e.,

$$\begin{aligned} U_G &= Ju(x_G) - l(Jx_G)Jx_G \\ &= Ja^{\frac{a}{1+b-a}}(1-a), \end{aligned}$$

$$U_M = Ju(x_M) - l(Jx_M)Jx_M$$
$$= J\left(\frac{a^2}{1+b}\right)^{\frac{a}{1+b-a}}\left(1 - \frac{a^2}{1+b}\right).$$

We denote the ratio of U_M to U_G by E,

$$E = \frac{U_M}{U_G} = \left(\frac{a}{1+b}\right)^{\frac{a}{1+b-a}} \frac{1-\frac{a^2}{1+b}}{1-a},$$

i.e., E is a measure of performance of the ME to the performance of the GWE at 0 prices. Figure 5.1 illustrates the values of E as a function of a for different values of b. As can be seen from the figure, at fixed values of a, increasing b improves the performance of the ME, i.e., the more convex the latency function of the link, the better performance we have under monopoly pricing. Moreover, at a fixed value of b, ME performs better at higher values of a as expected, i.e., the less concave the utility function is, the better performance we have in the ME.

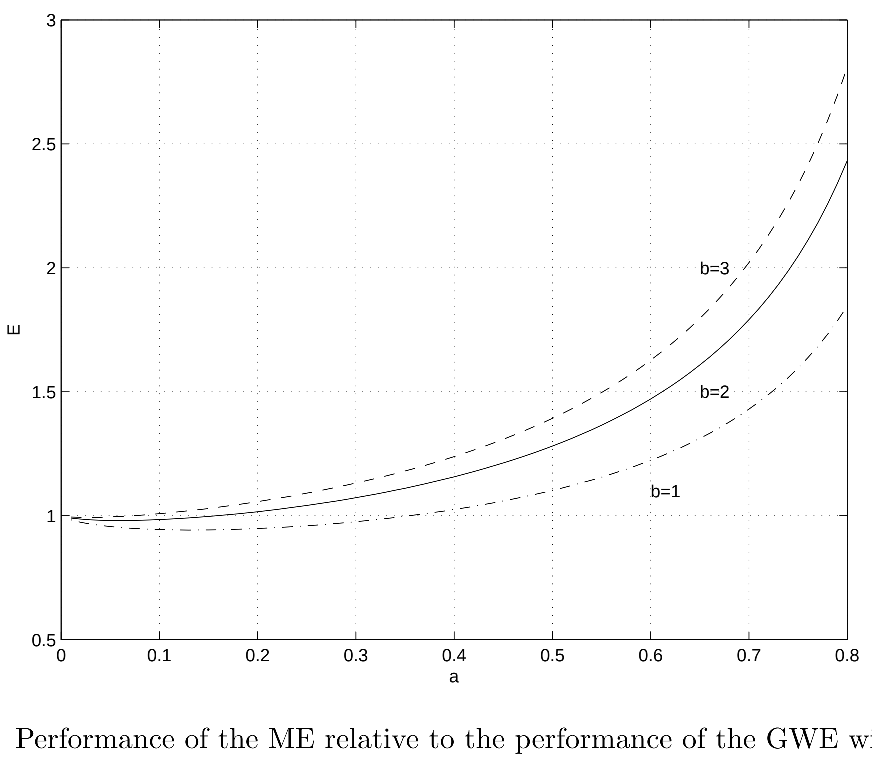

Figure 5.1: Performance of the ME relative to the performance of the GWE with 0 prices.

This example illustrates an important point in the performance of ME. As our discussion of pricing illustrated, the monopolist internalizes the congestion externality in its pricing decision. This congestion externality is the reason why decentralized routing is inefficient. The more pronounced this externality, the more useful is monopoly pricing. The example illustrates this by showing that the performance of monopoly pricing improves relative to GWE with 0 prices as

b increases, thus the latency function becomes more convex and the externalities become more important.

We illustrated in this section that monopoly pricing introduces a natural market mechanism to price congestion externalities. However, as can be seen from Proposition 4.8, monopoly pricing also introduces a distortion related to the markup term. It can be inferred from this price characterization that as the utility functions converge to linear functions, the markup term vanishes and the ME achieves the full information social optimum. In the next section, we analyze other utility functions for which the ME results in the socially optimal allocation.

6. ROUTING WITH PARTICIPATION CONTROL

In this section, we analyze traffic in a parallel-link network accessed by users that have a fixed amount of traffic to send, i.e., user j has t_j units of traffic. This problem is the standard routing problem. However, in a system with the service provider setting the prices, the users should have the option of not sending any data if the prices are set too high. Otherwise, the service provider would charge infinite prices on the links. We refer to this problem as the *routing problem with participation control.* It can be modelled by using the following utility function u_j for user j,

$$u_j(x) = \begin{cases} 0 & \text{if } 0 \leq x < t_j, \\ t_j & \text{if } x \geq t_j, \end{cases} \tag{6.1}$$

together with binary participation decision variables z_j; i.e., $z_j = 1$ if user j chooses to send his t_j units of traffic and $z_j = 0$ if user j chooses not to send any traffic.

An equilibrium of this problem can be defined as follows.

Definition 6.3: For a given price vector $p \geq 0$, a vector $(x^*, z^*) = (x_j^*, z_j^*)_{j\in\mathcal{J}}$, where $x_j^* = \left[(x_j^1)^*, \ldots, (x_j^I)^*\right]$ is a nonnegative vector and $z_j^* \in \{0,1\}$, is a GWE of the routing problem with participation control, if for all $j \in \mathcal{J}$,

$$(x_j^*, z_j^*) \in \arg \max_{x_j \geq 0,\ z_j \in \{0,1\}} \left\{ u_j(\Gamma_j z_j) - \sum_{i\in\mathcal{I}} \big(l^i(\gamma^i) + p^i\big) x_j^i \right\}, \tag{6.2}$$

where u_j is given by Eq. (6.1), and

$$\gamma^i = \sum_{j\in\mathcal{J}} x_j^{*i}, \qquad \forall\ i \in \mathcal{I}.$$

Using Eq. (6.1), we can rewrite condition (6.2) equivalently as

$$(x_j^*, z_j^*) \in \arg \max_{\substack{x_j \geq 0,\ z_j \in \{0,1\} \\ \sum_{i\in\mathcal{I}} x_j^i = t_j, \text{ if } z_j = 1}} \left\{ z_j t_j - \sum_{i\in\mathcal{I}} \big(l^i(\gamma^i) + p^i\big) x_j^i \right\}. \tag{6.3}$$

Note that, due to non-concavity of the utility function [see Eq. (6.1)], best response correspondences are not convex-valued, therefore one cannot use fixed point arguments to show the existence of a GWE for a given p. In the following, we show that at the price set by the monopolist, there exists a GWE, which also achieves the social optimum.

Using the linear structure of problem (6.3), we obtain the following characterization of a GWE (proof ommitted).

Lemma 6.4: For a given price vector $p \geq 0$, a vector $(x_j, z_j)_{j\in\mathcal{J}}$ is a GWE if and only if it satisfies the following conditions:

(1) $\sum_{j\in\mathcal{J}} x_j^i = \gamma^i$ for all $i \in \mathcal{I}$.

(2) If $z_j = 1$, $\sum_{i\in\mathcal{I}} x_j^i = t_j$.

(3) If $z_j = 0$, $x_j^i = 0$ for all $i \in \mathcal{I}$.
Define the set $\overline{\mathcal{I}} = \Big\{ i \mid l^i(\gamma^i) + p^i \leq \min\big\{1, \min_{m\in\mathcal{I}}\{l^m(\gamma^m) + p^m\}\big\}\Big\}$.

(4) If $i \notin \overline{\mathcal{I}}$, then $\gamma^i = 0$.

(5) If $\min_{m\in\mathcal{I}}\{l^m(\gamma^m) + p^m\} < 1$, then $z_j = 1$ for all $j \in \mathcal{J}$.

We next show the following lemma related to the uniqueness of link loads at a GWE of the routing problem.

Lemma 6.5 Assume that l^i is a continuous strictly increasing function. For a given $p \geq 0$, the resulting link loads, $\gamma = [\gamma^1, \ldots, \gamma^I]$, at any GWE are unique.

Proof: Let γ and $\tilde{\gamma}$ be two distinct link load vectors at the GWE (x, z) and $(\tilde{x}, \tilde{z})$, respectively. By Lemma 6.4 and condition (4), we have

$$p^i + l^i(\gamma^i) \begin{cases} = K & \text{if } i \in \overline{\mathcal{I}}, \\ > K & \text{if } i \notin \overline{\mathcal{I}}. \end{cases} \tag{6.4}$$

for some constant K, where $\overline{\mathcal{I}}$ is the set of links as defined in Lemma 6.4. We also have

$$p^i + l^i(\tilde{\gamma}^i) \begin{cases} = \tilde{K} & \text{if } i \in \tilde{\mathcal{I}}, \\ > \tilde{K} & \text{if } i \notin \tilde{\mathcal{I}}. \end{cases} \tag{6.5}$$

for some constant $\tilde{K}$, where $\tilde{\mathcal{I}}$ is the set of links as defined in Lemma 6.4, with γ^i replaced by $\tilde{\gamma}^i$.

Assume without loss of generality that there exists some s such that $\tilde{\gamma}^s > \gamma^s$. This implies that

$$\tilde{K} = p^s + l^s(\tilde{\gamma}^s) > p^s + l^s(\gamma^s) \geq K,$$

from which we get $\tilde{K} > K$. This implies that $\tilde{\gamma}^i \geq \gamma^i$ for all $i \in \mathcal{I}$. To see this, assume that $\tilde{\gamma}^i < \gamma^i$ for some i. We see from (6.4) that

$$p^i = K - l^i(\gamma^i),$$

and from (6.5) and $\tilde{K} > K$ that

$$p^i \geq \tilde{K} - l^i(\tilde{\gamma}^i) > K - l^i(\gamma^i),$$

yielding a contradiction. Hence, we have

$$\sum_{i \in \mathcal{I}} \tilde{\gamma}^i > \sum_{i \in \mathcal{I}} \gamma^i. \tag{6.6}$$

Since $\tilde{K} > K$, we also have $\tilde{z}_j \leq z_j$ for all $j \in \mathcal{J}$, from which we get

$$\sum_{i \in \mathcal{I}} \tilde{\gamma}^i = \sum_{j \in \mathcal{J}} \tilde{z}_j t_j \leq \sum_{j \in \mathcal{J}} z_j t_j = \sum_{i \in \mathcal{I}} \gamma^i,$$

contradicting Eq. (6.6), and showing that the link loads at a GWE for a given p are unique. **Q.E.D.**

Consider next the monopoly problem for the routing problem with participation constraint,

$$\begin{aligned} \text{maximize} \quad & \sum_{i \in \mathcal{I}} p^i \gamma^i \\ \text{subject to} \quad & \gamma^i = \sum_{j \in \mathcal{J}} x_j^i, \qquad \forall\, i \in \mathcal{I}, \\ & (x, z) \in G(p),\; p \geq 0, \end{aligned} \tag{6.7}$$

where $G(p)$ denotes the set of (x, z) that satisfies conditions (1)-(5) stated in Lemma 6.4. We call the optimal solution of problem (6.7) $\big(p, (x, z)\big)$ as the *monopoly equilibrium* (ME) and p as the *monopoly equilibrium price*.

We next show the following ME price characterization.

Lemma 6.6: Let $\big(p, (x, z)\big)$ be an ME. Let $\gamma^i = \sum_{j \in \mathcal{J}} x_j^i$. Then for all i with $\gamma^i > 0$, we have

$$p^i = 1 - l^i(\gamma^i).$$

Proof: Since $\gamma^i > 0$, it follows by condition (4) in Lemma 6.4 that $i \in \overline{\mathcal{I}}$, which implies that

$$p^i + l^i(\gamma^i) \leq 1.$$

We next show that $p^i + l^i(\gamma^i) = 1$. Assume, to arrive at a contradiction, that $p^i + l^i(\gamma^i) < 1$. This implies that

$$p^k + l^k(\gamma^k) < \min\{1, \min_{m \notin \overline{\mathcal{I}}} p^m\}, \qquad \forall\ k \in \overline{\mathcal{I}}.$$

Hence, there exists some $\epsilon > 0$ such that

$$p^k + \epsilon + l^k(\gamma^k) < \min\{1, \min_{m \notin \overline{\mathcal{I}}} p^m\}, \qquad \forall\ k \in \overline{\mathcal{I}}.$$

Hence, (x, z) satisfies conditions (1)-(5) at the price vector $\tilde{p} = p + e_i$, where e_i is an I-dimensional vector, whose i^th component is ϵ if $i \in \overline{\mathcal{I}}$, and 0 otherwise. But this shows that $(x, z) \in G(\tilde{p})$, i.e., $\big(\tilde{p}, (x, z)\big)$ is feasible for problem (6.7), and has strictly higher objective function value. This contradicts the fact the $\big(p, (x, z)\big)$ is an ME, showing that $p^i + l^i(\gamma^i) = 1$. **Q.E.D.**

Using the preceding proposition, we can rewrite problem (6.7) as

$$\begin{aligned}
\text{maximize}_{\ x_j^i \geq 0,\ z_j \in \{0,1\}} \ & \sum_{i \in \mathcal{I}} \big(1 - l^i(\gamma^i)\big)\gamma^i \\
\text{subject to}\ \ & \gamma^i = \sum_{j \in \mathcal{J}} x_j^i, \qquad \forall\ i \in \mathcal{I}, \\
& \sum_{i \in \mathcal{I}} x_j^i = t_j,\ \text{if } z_j = 1, \\
& x_j^i = 0,\ \forall\ i \in \mathcal{I},\ \text{if } z_j = 0,
\end{aligned}$$

or,

$$\begin{aligned}
\text{maximize}_{\ x_j^i \geq 0,\ z_j \in \{0,1\}} \ & \sum_{j \in \mathcal{J}} \left\{ \sum_{i \in \mathcal{I}} x_j^i - \sum_{i \in \mathcal{I}} l^i(\gamma^i) x_j^i \right\} \\
\text{subject to}\ \ & \gamma^i = \sum_{j \in \mathcal{J}} x_j^i, \qquad \forall\ i \in \mathcal{I}, \\
& \sum_{i \in \mathcal{I}} x_j^i = t_j,\ \text{if } z_j = 1, \\
& x_j^i = 0,\ \forall\ i \in \mathcal{I},\ \text{if } z_j = 0,
\end{aligned}$$

which can equivalently be written as

$$\begin{aligned}
\text{maximize}_{\ x_j^i \geq 0,\ z_j \in \{0,1\}} \ & \sum_{j \in \mathcal{J}} \left\{ z_j t_j - \sum_{i \in \mathcal{I}} l^i(\gamma^i) x_j^i \right\} \\
\text{subject to}\ \ & \gamma^i = \sum_{j \in \mathcal{J}} x_j^i, \qquad \forall\ i \in \mathcal{I}, \qquad (6.8) \\
& \sum_{i \in \mathcal{I}} x_j^i = t_j,\ \text{if } z_j = 1.
\end{aligned}$$

This problem has an optimal solution; i.e., at each $z \in \{0,1\}^I$, there exists an optimal solution x (since the objective function is continuous and the constraint set is compact.) This shows that there exists an ME of the routing problem with participation control.

Moreover, note that the social problem for the routing problem with participation control is the following optimization problem that maximizes the aggregate social surplus,

$$\begin{aligned} \text{maximize}_{\ x_j^i \geq 0,\ z_j \in \{0,1\}} & \sum_{j \in \mathcal{J}} \left\{ z_j t_j - \sum_{i \in \mathcal{I}} l^i(\gamma^i) x_j^i \right\} \\ \text{subject to} \ \ & \gamma^i = \sum_{j \in \mathcal{J}} x_j^i, \qquad \forall\, i \in \mathcal{I}, \\ & \sum_{i \in \mathcal{I}} x_j^i = t_j, \ \text{if } z_j = 1, \end{aligned}$$

[cf. Eq. (6.3)]. This problem is identical to problem (6.8). Hence, we establish the following result.

Proposition 6.10: A vector $(\hat{x}, \hat{z})$ is a social optimum if and only if there exists a price vector $\tilde{p}$ such that $\big(p, (x, z)\big)$ is an ME.

7. CONCLUSIONS AND EXTENSIONS

A central objective of data network analysis is to characterize and implement relatively efficient flows of data in large networks. Although much of the literature approaches this problem as a network optimization problem, the practice is often different. First, most networks are decentralized: individual users often have control over their flow and routing decisions, while network planners typically lack detailed knowledge about the needs and preferences of users. Second, most networks are controlled by for-profit entities, whose objectives are not efficiency or user welfare, but profit maximization. While a recent literature has investigated implications of the decentralized nature of modern data networks, particularly emphasizing the potential inefficiencies that result from this feature, there has been no systematic attempt to incorporate the second feature into models of data networks.

This paper is a first attempt towards a systematic analysis of decentralized data networks with many users that control their own flow and routing decisions, and service provider(s) charging prices for bandwidth and data transmission to maximize their own profits. Such an analysis first necessitates a unified approach to flow control and routing in the presence of prices, and a characterization of the response of link loads to changes in prices. After developing such a unified approach, this paper provides a characterization of prices, flow rates and link loads in the presence of a profit-maximizing monopolist service provider.

The most important feature of the monopoly equilibrium is that, despite its profit-maximizing objective, the monopolist may induce an allocation that is socially optimal. In particular, in the absence of the monopoly pricing, flow and routing decisions are inefficient, because users ignore the congestion externality that they create on others by their data transmission. The monopolist recognizes that, in the presence of the congestion externality, higher prices do not diminish the attractiveness of its product by as much, because the reduction in data transmission reduces congestion. Consequently, monopoly price for each link consists of two terms; the first exactly internalizing the congestion externality, and the second a further markup to increase profits. We show that in some important special cases, such as a model with only routing decisions, the monopoly equilibrium achieves the social optimum that the network planner with full information and complete control over flow and routing decisions would have implemented.

This research opens the way for a number of further analyses. In ongoing research, we are considering the following problems:

(1) The results are presented for parallel-link networks. With a different formulation, they can be generalized to other network topologies.

(2) Although the monopolist service provider may achieve the social optimum in some cases, it does so by capturing the entire consumer surplus. A potentially more realistic situation is with oligopolistic service providers competing for users. In this case, a similar efficiency result can be obtained under more general circumstances, and importantly, with the entire consumer surplus potentially accruing to users.

(3) When the monopolist creates distortions relative to the social optimum, these are partly caused by the fact that the monopolist is forced to charge the same price to all users. In practice, monopolists can charge different prices to different users, for example in the form of two-part tariffs, which combine an entry fee and bandwidth prices. However, not knowing the exact preferences of each user, the monopolist has to make two-part tariffs or other non-linear pricing schemes incentive compatible. This leads to a different type of optimization problem, which can be approached with the tools of mechanism design analysis.

(4) The service-provider viewpoints developed in this paper also applies to other aspects of resource allocation in communication networks, for example power control in wireless systems. In this case, the monopolist will choose between different allocation methods, recognizing the effect that these will have on the willingness of users to pay and to participate in its network.

All of these problems are useful for applied network analysis and the design of appropriate network regulation. They also require the development of new techniques and approaches, which we are currently undertaking.

8. APPENDIX

Proof of Proposition 4.8: We prove that the constraint gradients of problem (4.7) are linearly independent at (p,γ,Γ), where (p,x) is the ME, $\gamma^i = \sum_{j\in\overline{\mathcal{J}}} x_j^i$, for each $i\in\mathcal{I}$, and $\Gamma_j = \sum_{i\in\mathcal{I}} x_j^i$, for each $j\in\overline{\mathcal{J}}$. Let $\overline{J} = |\overline{\mathcal{J}}|$ and assume without loss of generality that $\overline{\mathcal{J}} = \{1,\ldots,\overline{J}\}$. Denote the set of $I+\overline{J}$ equality constraints of problem (4.7) by

$$f(p,\gamma,\Gamma) = 0.$$

We show that $\nabla_{(\gamma,\Gamma)} f(p,\gamma,\Gamma)$ is nonsingular under Assumption 2.1. If it were not, there would be some nonzero $I+\overline{J}$-dimensional vector $(y',t')'$ that belongs to the nullspace of $\nabla_{(\gamma,\Gamma)} f(p,\gamma,\Gamma)$, i.e.,

$$\nabla_{\gamma,\Gamma} f(p,\gamma,\Gamma)(y',t')' = 0.$$

Multiplying the preceding out, we obtain the following set of equations,

$$t_{\overline{J}} = (l^1)'(\gamma^1)\left(\sum_{k=1}^{\overline{J}-1} t_k + y_1\right),$$

$$t_{\overline{J}} = (l^m)'(\gamma^m) y_m, \qquad \forall\, m = 2,\ldots,I,$$

$$t_{\overline{J}} = u_1''(\Gamma_1)\sum_{m=1}^{I} y_m,$$

$$t_{\overline{J}} = u_k''(\Gamma_k) t_{k-1}, \qquad \forall\, k = 2\ldots,\overline{J}.$$

Combining the previous relations, we obtain

$$t_{\overline{J}}\left(\sum_{j\in\overline{\mathcal{J}}} \frac{1}{u_j''(\Gamma_j)} - \sum_{i=1}^{I} \frac{1}{(l_i)'(\gamma^i)}\right) = 0.$$

Since $u_j''(\Gamma_j) < 0$ for all j and $(l^i)'(\gamma^i) > 0$ for all i, the term in the paranthesis is negative, implying that $t_{\overline{J}} = 0$. From this, we get that $(y',t')' = 0$, hence proving that $\nabla_{(\gamma,\Gamma)} f(p,\gamma,\Gamma)$ is nonsingular.

Proof of Proposition 5.9: First we show that $\tilde{\Gamma}_j \le \overline{\Gamma}_j$ for all $j\in\mathcal{J}$. Partition the set of users into two sets R and S as

$$R = \{r\in\mathcal{J} \mid \tilde{\Gamma}_r > \overline{\Gamma}_r\},$$

$$S = \{s \in \mathcal{J} \mid \tilde{\Gamma}_s \leq \overline{\Gamma}_s\}.$$

We show that the set R is empty. Assume to arrive at a contradiction that the set R is nonempty. Define a subset of links as

$$I_{act} = \{i \in \mathcal{I} \mid \tilde{x}^i_j > 0, \text{ for some } j \in R\}.$$

We show that, for all $i \in I_{act}$, we have

$$\tilde{\gamma}^i < \overline{\gamma}^i, \tag{8.1}$$

and

$$\overline{x}^i_s = 0, \qquad \forall\, s \in S. \tag{8.2}$$

Let $i \in I_{act}$. This implies that $\tilde{x}^i_j > 0$ for some $j \in R$. The first order optimality conditions for the social optimum and for the GWE at 0 prices imply that

$$\begin{aligned} u'_j(\tilde{\Gamma}_j) - l^i(\tilde{\gamma}^i) - (l^i)'(\tilde{\gamma}^i)\tilde{\gamma}^i &\leq 0, \qquad \text{if } \tilde{x}^i_j \geq 0, \\ &= 0, \qquad \text{if } \tilde{x}^i_j > 0, \end{aligned} \tag{8.3}$$

and

$$\begin{aligned} u'_j(\overline{\Gamma}_j) - l^i(\overline{\gamma}^i) &\leq 0, \qquad \text{if } \overline{x}^i_j \geq 0, \\ &= 0, \qquad \text{if } \overline{x}^i_j > 0. \end{aligned} \tag{8.4}$$

Using the preceding, we obtain

$$u'_j(\tilde{\Gamma}_j) - l^i(\tilde{\gamma}^i) - (l^i)'(\tilde{\gamma}^i)\tilde{\gamma}^i \geq u'_j(\overline{\Gamma}_j) - l^i(\overline{\gamma}^i).$$

Since $j \in R$ and u_j is strictly concave, we have $u'_j(\tilde{\Gamma}_j) < u'_j(\overline{\Gamma}_j)$, which implies that

$$l^i(\tilde{\gamma}^i) + (l^i)'(\tilde{\gamma}^i)\tilde{\gamma}^i < l^i(\overline{\gamma}^i). \tag{8.5}$$

Since l^i is strictly increasing, it follows that

$$\tilde{\gamma}^i < \overline{\gamma}^i,$$

hence proving claim (8.1). To show (8.2), suppose to arrive at a contradiction, that $\overline{x}^i_s > 0$ for some $s \in S$. This implies by the first order optimality conditions [cf. Eqs. (8.3) and (8.4)] that

$$u'_s(\overline{\Gamma}_s) - l^i(\overline{\gamma}^i) \geq u'_s(\tilde{\Gamma}_s) - l^i(\tilde{\gamma}^i) - (l^i)'(\tilde{\gamma}^i)\tilde{\gamma}^i.$$

Since $s \in S$ and u_j is concave, we have $u'_s(\overline{\Gamma}_s) \leq u'_s(\tilde{\Gamma}_s)$. Combining this with the preceding equation and Eq. (8.5), we obtain

$$\begin{aligned} l^i(\overline{\gamma}^i) &\leq l^i(\tilde{\gamma}^i) + (l^i)'(\tilde{\gamma}^i)\tilde{\gamma}^i \\ &< l^i(\overline{\gamma}^i) \end{aligned},$$

thus yielding a contradiction and showing (8.2).

We next use Eqs. (8.1) and (8.2) to obtain

$$\sum_{i \in I_{act}} \overline{\gamma}_i = \sum_{i \in I_{act}} \sum_{j=1}^{J} \overline{x}_j^i = \sum_{i \in I_{act}} \sum_{j \in R} \overline{x}_j^i \leq \sum_{j \in R} \sum_{i=1}^{I} \overline{x}_j^i = \sum_{j \in R} \overline{\Gamma}_j.$$

We also have

$$\sum_{i \in I_{act}} \tilde{\gamma}_i \geq \sum_{i \in I_{act}} \sum_{j \in R} \tilde{x}_j^i = \sum_{j \in R} \sum_{i=1}^{I} \tilde{x}_j^i = \sum_{j \in R} \tilde{\Gamma}_j.$$

The preceding sets of equations together with the definition of set R imply that

$$\sum_{i \in I_{act}} \tilde{\gamma}^i \geq \sum_{i \in I_{act}} \overline{\gamma}^i.$$

Summing Eq. (8.1) over all $i \in I_{act}$ yields a contradiction, thus proving that the set R is empty.

We next show that $\Gamma_j \leq \tilde{\Gamma}_j$ for all $j \in \mathcal{J}$. Partition the set of users into two sets R and S as

$$R = \{r \in \mathcal{J} \mid \Gamma_r > \tilde{\Gamma}_r\},$$

$$S = \{s \in \mathcal{J} \mid \Gamma_s \leq \tilde{\Gamma}_s\}.$$

We show that the set R is empty. Assume to arrive at a contradiction that the set R is nonempty. Define a subset of links as

$$I_{act} = \{i \mid x_j^i > 0, \text{ for some } j \in R\}.$$

We show that, for all $i \in I_{act}$, we have

$$\gamma^i < \tilde{\gamma}^i, \tag{8.6}$$

and

$$\tilde{x}_s^i = 0, \qquad \forall\, s \in S. \tag{8.7}$$

Let $i \in I_{act}$. This implies that $x_j^i > 0$ for some $j \in R$. The first order optimality conditions for the ME imply that

$$\begin{aligned} u'_j(\Gamma_j) - l^i(\gamma^i) - (l^i)'(\gamma^i)\gamma^i + \frac{\sum_{m=1}^{I} \gamma^m}{\sum_{\{j \in \overline{J}\}} \frac{1}{u''_j(\Gamma_j)}} &\leq 0, \qquad \text{if } x_j^i \geq 0, \\ &= 0, \qquad \text{if } x_j^i > 0, \end{aligned} \tag{8.8}$$

[cf. Proposition 4.8]. Together with Eq. (8.3), we obtain

$$u'_j(\Gamma_j) - \bar{l}^i(\gamma^i) + \frac{\sum_{m=1}^{I} \gamma^m}{\sum_{\{j \in \bar{J}\}} \frac{1}{u''_j(\Gamma_j)}} \geq u'_j(\tilde{\Gamma}_j) - \bar{l}^i(\tilde{\gamma}^i),$$

where

$$\begin{aligned} \bar{l}^i(x) &= l^i(x) + (l^i)'(x)x, \\ &= \big(l^i(x)x\big)'. \end{aligned}$$

Since $j \in R$ and u_j is strictly concave, we have $u'_j(\Gamma_j) < u'_j(\tilde{\Gamma}_j)$, which implies that

$$\bar{l}^i(\gamma^i) - \frac{\sum_{m=1}^{I} \gamma^m}{\sum_{\{j \in \bar{J}\}} \frac{1}{u''_j(\Gamma_j)}} < \bar{l}^i(\tilde{\gamma}^i). \tag{8.9}$$

Since $\bar{l}^i$ is nondecreasing [in view of the convexity of $l^i(x)x$] and u_k is strictly concave for all k, this implies that

$$\gamma^i < \tilde{\gamma}^i,$$

hence proving claim (8.6). To show (8.7), assume to arrive at a contradiction, that $\tilde{x}_s^i > 0$ for some $s \in S$. This implies by the first order optimality conditions [cf. Eqs. (8.3) and (8.8)] that

$$u'_s(\tilde{\Gamma}_s) - \bar{l}^i(\tilde{\gamma}^i) \geq u'_s(\Gamma_s) - \bar{l}^i(\gamma^i) + \frac{\sum_{m=1}^{I} \gamma^m}{\sum_{\{j \in \bar{J}\}} \frac{1}{u''_j(\Gamma_j)}}.$$

Since $s \in S$ and u_j is concave, we have

$$u'_s(\Gamma_s) \geq u'_s(\tilde{\Gamma}_s),$$

which together with Eq. (8.9) implies that

$$\begin{aligned} \bar{l}^i(\tilde{\gamma}^i) &\leq \bar{l}^i(\gamma^i) - \frac{\sum_{m=1}^{I} \gamma^m}{\sum_{\{j \in \bar{J}\}} \frac{1}{u''_j(\Gamma_j)}}, \\ &< \bar{l}^i(\tilde{\gamma}^i) \end{aligned}$$

thus yielding a contradiction and showing (8.7).

The rest is a similar argument as before and proves that the set R is empty. **Q.E.D.**

9. ACKNOWLEDGEMENTS

The authors would like to thank seminar participants at MIT Operations Research Center, University of Illinois at Urbana-Champaign, and Caltech, and in particular, Tamer Basar, Ramesh Johari, Tim Roughgarden, and Ivan Werning for useful comments and helpful discussions.

REFERENCES

[ABS02] Altman E., Basar T., and Srikant R., "Nash Equilibria for combined flow control and routing in networks: asymptotic behavior for a large number of users," *IEEE Transactions on Automatic Control*, 2002.

[ArH71] Arrow, K. and Hahn, F., General Competitive Analysis, Holden-Day, San Francisco, 1971.

[BaS02] Basar, T. and Srikant R., "Revenue-maximizing pricing and capacity expansion in a many-users regime," *IEEE INFOCOM*, 2002.

[BMW56] Beckmann, M., Mcguire, C. B., and Winsten, C. B., Studies in the Economics of Transportation. Yale University Press, 1956.

[BNO03] Bertsekas, D. P., Nedic, A., and Ozdaglar, A. E., Convex Analysis and Optimization, Athena Scientific, Belmont, MA, 2003.

[CDR03] Cole, R., Dodis, Y., and Roughgarden, T., "How much can taxes help selfish routing?," *Proceedings of the 35th Annual ACM Symposium on the Theory of Computing*, 2003.

[CSS03] Correa J. R., Schulz, A. S., and N. Stier Moses, "Selfish Routing in Capacitated Networks," MIT Sloan School of Management, Working paper.

[DeS63] Debreu, G. and Scarf, H., "A limit theorem on the core of an economy," Int. Econ. Rev. vol. 4, pp. 235-246, 1963.

[Deb59] Debreu, G., Theory of Value, Wiley, New York, 1959.

[HaM85] Haurie, A. and Marcotte, P., "On the relationship between Nash-Cournot and Wardrop Equilibria," *Networks*, 15, 295-308, 1985.

[JoT03] Johari, R. and Tsitsiklis, J., "Network resource Allocation and a congestion game," Working paper, 2003.

[KLO97] Korilis, Y. A., Lazar, A. A., and Orda, A.,"Achieving network optima using Stackelberg routing strategies," *IEEE/ACM Transactions on Networking*, vol. 5(1, pp. 161-173, 1997.)

[KMT98] Kelly, F. P., Maulloo A. K., and Tan D. K.,“Rate control for communication networks: shadow prices, proportional fairness, and stability,” *Journal of the Operational Research Society*, vol. 49, pp. 237-252, 1998.

[Kel91] Kelly, F. P., “ Network routing,” *Phil. Trans. R. Soc. Lond.*, vol. 337, pp. 343-367, 1991.

[Kel97] Kelly, F. P., “Charging and rate control for elastic traffic,” *European Transactions on Telecommunications*, vol. 8, pp. 33-37, 1997.

[LoL99] Low, S. and Lapsley, D. E., “Optimization flow control, I: Basic algorithm and convergence,” *IEEE/ACM Transactions on Networking*, vol. 7(6), pp. 861-874, 1999.

[MWG95] Mascolell, A., Whinston, M. D., and Green, J. R., Microeconomic Theory, Oxford, United Kingdom: Oxford University Press, 1995.

[ORS93] Orda, A., Rom, R., and Shimkin, N., “Competitive routing in multi-user communication networks,” *IEEE/ACM Transactions on Networking*, vol. 1, pp. 510-521, 1993.

[RoT02] Roughgarden, T. and Tardos, E., “ How bad is selfish routing?” *Journal of the ACM*, vol. 49, no. 2, pp 236-259, 2002.

[Ros65] Rosen, J. B., “Existence and uniqueness of equilibrium points for concave N-person games,” *Econometrica* 33, 3, pp. 520-534, 1965.

[ScS03] Schulz, A. S. and Stier Moses, N., “On the performance of user equilibria in traffic networks,” MIT Sloan School of Management, Working paper 4274-02, 2002.

[Sch75] Schumpeter, J., Capitalism, Socialism, and Democracy, Harper and Row, N.Y., 1975.

[SCEH96] Shenker, S., Clark, D., Estrin, D., and Herzog, S., “Pricing in computer networks: reshaping the research agenda,” *Telecommunications Policy*, vol. 20, no. 3, pp. 183-201, 1996.

[She95] Shenker, S., “Fundamental design issues for the future Internet,” *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1176-1188, 1995.

[Tir90] Tirole, J., The Theory of Industrial Organization, MIT Press, Cambridge, 1990.

[War52] Wardrop, J. G., “Some theoretical aspects of road traffic research,” *Proceedings of the Institute of Civil Engineers*, Pt. II, vol. 1, pp. 325-378.