# Efficiency Loss in a Network Resource Allocation Game: The Case of Elastic Supply

Ramesh Johari, *Member, IEEE*, Shie Mannor, *Member, IEEE*, and John N. Tsitsiklis, *Fellow, IEEE*

***Abstract*—We consider a resource allocation problem where individual users wish to send data across a network to maximize their utility, and a cost is incurred at each link that depends on the total rate sent through the link. It is known that as long as users do not anticipate the effect of their actions on prices, a simple proportional pricing mechanism can maximize the sum of users' utilities minus the cost (called aggregate surplus). Continuing previous efforts to quantify the effects of selfish behavior in network pricing mechanisms, we consider the possibility that users anticipate the effect of their actions on link prices. Under the assumption that the links' marginal cost functions are convex, we establish existence of a Nash equilibrium. We show that the aggregate surplus at a Nash equilibrium is no worse than a factor of $4\sqrt{2}-5$ times the optimal aggregate surplus; thus, the efficiency loss when users are selfish is no more than approximately 34%.**

***Index Terms*—Congestion pricing, network resource allocation.**

THE current Internet is used by a widely heterogeneous population of users; not only are different types of traffic sharing the same network, but different end users place different values on their perceived network performance. This has led to a surge of interest in *congestion pricing*, where the network is treated as a market, and prices are set to mediate demand and supply of network resources; see, e.g., [1] and [2].

We investigate a specific price mechanism considered by Kelly *et al.* in [3] (motivated by the proposal made in [4]). For simplicity, let us first consider the special case of a single link; in this case the mechanism works as follows. Each user submits a bid, or total *willingness-to-pay*, to the link manager. This represents the total amount the user expects to pay. The link manager then chooses a total rate and price such that the product of price and rate is equal to the sum of the bids, and the price is equal to marginal cost; note, in particular, that the supply of the link is *elastic*, i.e., it is not fixed in advance. Finally, each user receives a fraction of the allocated rate in proportion to their bid. It is shown in [3] that if users do not anticipate the effect of their bid on the price, such a scheme maximizes the sum of users' utilities minus the cost of the total allocated rate, known as the *aggregate surplus* (see [5, Ch. 10]).

The pricing mechanism of [3] takes as input the bids of the users, and produces as output the price of the link, and the resulting rate allocation to the users. Kelly *et al.* [3] continue on to discuss distributed algorithms for implementation of this market-clearing process: given the bids of the users, the authors present two algorithms which converge to the market-clearing price and rate allocation. Indeed, much of the interest in this market mechanism stems from its desirable properties as a decentralized system, including both stability and scalability. For details, we refer the reader to [6]–[9].

One important interpretation of the price given to users in the algorithms of [3] is that it can provide early notification of congestion. Building on the explicit congestion notification (ECN) proposal [10], this interpretation suggests that the network might charge users proactively, in hopes of avoiding congestion later. From an implementation standpoint, such a shift implies that rather than a hard capacity constraint (i.e., a link is overloaded when the rate through it exceeds the capacity of the link), the link has an elastic capacity (i.e., the link gradually begins to signal a buildup of congestion before the link's true capacity is actually met). Many proposals have been made for "active queue management" (AQM) to achieve good performance with ECN; see, e.g., [11]–[14]. This issue is of secondary importance to our discussion, as we do not concern ourselves with the specific interpretation of the cost function at the link. (An insightful discussion of the relationship between active queue management and the cost function of the link may be found in [15].)

In this paper, we investigate the robustness of the market mechanism of [3] when users attempt to manipulate the market. Formally, we consider a model where users anticipate the effects of their actions on the link prices. This makes the model a game, and we ask two fundamental questions. First, does a Nash equilibrium exist for this game? And second, how inefficient is such an equilibrium relative to the maximal aggregate surplus? We show that Nash equilibria exist, and that the efficiency loss is no more than a factor $6-4\sqrt{2}$ of the maximal aggregate surplus (approximately 34%) when users are price anticipating.

Such an investigation forms part of a broader body of work on quantifying efficiency loss in environments where participants are selfish. Results have been obtained for routing [16]–[18], traffic networks [19], [20] and network design problems [21], [22]. Our work is most closely related to that of [23], where the

Manuscript received June 14, 2004; revised June 24, 2005. Recommended by Associate Editor R. S. Srikant. This work was supported in part by the National Science Foundation under a Graduate Research Fellowship and Grant ECS-0312921, and by the Defense Advanced Research Projects Agency under the Next Generation Internet Initiative. A preliminary version of this paper appeared in the *Proceedings of the 43rd IEEE Conference on Decision and Control*, Bahamas, December 2004.

R. Johari is with the Department of Management Science and Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: ramesh.johari@stanford.edu).

S. Mannor is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada (e-mail: shie@ece.mcgill.ca).

J. N. Tsitsiklis is with the Department of Electrical Engineering and Computer Science, the Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: jnt@mit.edu).

Digital Object Identifier 10.1109/TAC.2005.858687

0018-9286/$20.00 © 2005 IEEE

same market mechanism as in this paper was considered for the case where the supply of a link is fixed, or *inelastic*; this was the mechanism first presented in [4]. Johari and Tsitsiklis show the efficiency loss when users are price anticipating is no worse than 25% [23].

The outline of the remainder of the paper is as follows. We consider a single link in isolation; extensions to general networks are discussed in [24]. In Section I, we describe the market mechanism for a single link, and recapitulate the results of Kelly *et al.* [3]. In Section II, we describe a game where users are price anticipating, and establish the existence of a Nash equilibrium. We also establish necessary and sufficient conditions for a strategy vector to be a Nash equilibrium. These conditions are used in Section III to prove the main result of the paper for a single link: That when users are price anticipating, the efficiency loss—that is, the loss in aggregate surplus relative to the maximum—is no more than 34%.

In Section IV, we compare the settings of inelastic and elastic supply. In particular, we consider a limit of cost functions which approach a hard capacity constraint. We show that if these cost functions are monomials and we let the exponent tend to infinity, then the efficiency loss approaches 25%, which is consistent with the result of [23]. Some conclusions are offered in Section V.

## I. Background

Suppose $R$ users share a single communication link. Let $d_r \geq 0$ denote the rate allocated to user $r$. We assume that user $r$ receives a *utility* equal to $U_r(d_r)$ if the allocated rate is $d_r$. In addition, we let $f = \sum_r d_r$ denote the total rate allocated at the link, and let $C(f)$ denote the cost incurred at the link when the total allocated rate is $f \geq 0$. We will assume that both $U_r$ and $C$ are measured in the same monetary units. A natural interpretation is that $U_r(d_r)$ is the monetary value to user $r$ of a rate allocation $d_r$, and $C(f)$ is a monetary cost for congestion at the link when the total allocated rate is $f$.

We make the following assumptions regarding $U_r$ and $C$.

*Assumption 1:* For each $r$, over the domain $d_r \geq 0$ the utility function $U_r(d_r)$ is concave, strictly increasing, and continuously differentiable, and the right directional derivative at 0, denoted $U_r'(0)$, is finite.

*Assumption 2:* There exists a continuous, convex, strictly increasing function $p(f)$ over $f \geq 0$ with $p(0) = 0$, such that for $f \geq 0$:

$$C(f) = \int_0^f p(z)dz.$$

Thus, $C(f)$ is strictly convex and strictly increasing.

Concavity in Assumption 1 corresponds to *elastic* traffic, as defined by Shenker [25]; such traffic includes file transfers such as FTP connections and peer-to-peer connections. Note that Assumption 2 does not require the price function $p$ to be differentiable. Indeed, assuming smoothness of $p$ would simplify some of the technical arguments in the paper. However, we later require the use of nondifferentiable price functions in our proof of Theorem 8.

Given complete knowledge and centralized control of the system, a natural problem for the network manager to try to solve is the following [4]:

SYSTEM :

$$\text{maximize} \quad \sum_r U_r(d_r) - C\left(\sum_r d_r\right) \tag{1}$$

$$\text{subject to} \quad d_r \geq 0, \qquad r = 1, \ldots, R. \tag{2}$$

Since the objective function (1) is continuous, and $U_r$ increases at most linearly while $C$ increases superlinearly, an optimal solution $\mathbf{d}^S = (d_1^S, \ldots, d_R^S)$ exists for (1), (2); since the feasible region is convex and $C$ is strictly convex, if the functions $U_r$ are strictly concave, then the optimal solution is unique. We refer to the objective function (1) as the *aggregate surplus*; this is the net monetary benefit to the economy consisting of the users and the single link [5]. For convenience, we define a function surplus$(\mathbf{d})$ which gives the aggregate surplus at an allocation $\mathbf{d}$

$$\text{surplus}(\mathbf{d}) \triangleq \sum_r U_r(d_r) - C\left(\sum_r d_r\right). \tag{3}$$

Due to the decentralized nature of the system, the resource manager may not have an exact specification of the utility functions [4]. As a result, we consider the following pricing scheme for rate allocation. Each user $r$ makes a payment (also called a *bid*) of $w_r$ to the resource manager. Given the vector $\mathbf{w} = (w_1, \ldots, w_r)$, the resource manager chooses a rate allocation $\mathbf{d}(\mathbf{w}) = (d_1(\mathbf{w}), \ldots, d_R(\mathbf{w}))$. We assume the manager treats all users alike—in other words, the network manager does not *price differentiate*. Thus the network manager sets a single price $\mu(\mathbf{w})$; we assume that $\mu(\mathbf{w}) = 0$ if $w_r = 0$ for all $r$, and $\mu(\mathbf{w}) > 0$ otherwise. All users are then charged the same price $\mu(\mathbf{w})$, leading to

$$d_r(\mathbf{w}) = \begin{cases} 0, & \text{if } w_r = 0 \\ \frac{w_r}{\mu(\mathbf{w})}, & \text{if } w_r > 0. \end{cases}$$

Associated with this choice of price is an aggregate rate function $f(\mathbf{w})$, defined by

$$f(\mathbf{w}) = \sum_r d_r(\mathbf{w}) = \begin{cases} 0, & \text{if } \sum_r w_r = 0 \\ \frac{\sum_r w_r}{\mu(\mathbf{w})}, & \text{if } \sum_r w_r > 0. \end{cases} \tag{4}$$

We will assume that $w_r$ is measured in the same monetary units as both $U_r$ and $C$. In this case, given a price $\mu > 0$, user $r$ acts to maximize the following payoff function over $w_r \geq 0$:

$$P_r(w_r; \mu) = U_r\left(\frac{w_r}{\mu}\right) - w_r. \tag{5}$$

The first term represents the utility to user $r$ of receiving a rate allocation equal to $w_r/\mu$; the second term is the payment $w_r$ made to the manager. Observe that since utility is measured in monetary units, the payoff is *quasilinear* in money, a typical assumption in modeling market mechanisms [5].

Notice that as formulated before, the payoff function $P_r$ assumes that user $r$ acts as a *price taker*; that is, user $r$ does not *anticipate* the effect of his choice of $w_r$ on the price $\mu$ and, hence,

on his resulting rate allocation $d_r(\mathbf{w})$. Informally, we expect that in such a situation the aggregate surplus will be maximized if the network manager sets a price equal to marginal cost, i.e., if the price function satisfies

$$\mu(\mathbf{w}) = p\left(f(\mathbf{w})\right). \tag{6}$$

According to the following proposition a joint solution to (4) and (6) can always be found; the proof is straightforward, and details may be found in [24]. This proposition is then used to show that when users optimize (5) and the price is set to satisfy (6), aggregate surplus is maximized.

*Proposition 1:* Suppose Assumption 2 holds. Given any vector of bids $\mathbf{w} \geq 0$, there exists a unique pair $(\mu(\mathbf{w}), f(\mathbf{w})) \geq 0$ satisfying (4) and (6), and in this case $f(\mathbf{w})$ is the unique solution $f$ to:

$$\sum_r w_r = fp(f). \tag{7}$$

Furthermore, $f(\cdot)$ has the following properties: 1) $f(\mathbf{0}) = 0$; 2) $f(\mathbf{w})$ is continuous for $\mathbf{w} \geq 0$; 3) $f(\mathbf{w})$ is a strictly increasing and strictly concave function of $\sum_r w_r$; and 4) $f(\mathbf{w}) \to \infty$ as $\sum_r w_r \to \infty$.

Observe that we can view (7) as a market-clearing process. Given the total revenue $\sum_r w_r$ from the users, the link manager chooses an aggregate rate $f(\mathbf{w})$ so that the revenue is exactly equal to the aggregate charge $f(\mathbf{w})p(f(\mathbf{w}))$. Due to Assumption 2, this market-clearing aggregate rate is uniquely determined. Kelly *et al.* present two algorithms in [3] which amount to dynamic processes of market-clearing; as a result, a key motivation for the mechanism we study in this paper is that it represents the equilibrium behavior of the algorithms in [3]. Kelly *et al.* show in [3] that when users are nonanticipating, and the network sets the price $\mu(\mathbf{w})$ according to (4) and (6), the resulting allocation solves SYSTEM. This is formalized in the following theorem, adapted from [3].

*Theorem 2 (Kelly et al., [3]):* Suppose Assumptions 1 and 2 hold. For any $\mathbf{w} \geq 0$, let $(\mu(\mathbf{w}), f(\mathbf{w}))$ be the unique solution to (4) and (6). Then there exists a vector $\mathbf{w}$ such that $\mu(\mathbf{w}) > 0$, and

$$P_r\left(w_r; \mu(\mathbf{w})\right) = \max_{\overline{w}_r \geq 0} P_r\left(\overline{w}_r; \mu(\mathbf{w})\right), \qquad r = 1, \ldots, R. \tag{8}$$

For any such vector $\mathbf{w}$, the vector $\mathbf{d}(\mathbf{w}) = \mathbf{w}/\mu(\mathbf{w})$ solves $SYSTEM$. If the functions $U_r$ are strictly concave, such a vector $\mathbf{w}$ is unique.

Theorem 2 shows that with an appropriate choice of price function (as determined by (4) and (6)), and under the assumption that the users behave as price takers, there exists a bid vector $\mathbf{w}$ where all users have optimally chosen their bids $w_r$, with respect to the given price $\mu(\mathbf{w})$; and the aggregate surplus is maximized at this "equilibrium." However, when the price taking assumption is violated, the model changes into a game and the guarantee of Theorem 2 is no longer valid. We investigate this game in the following section.

## II. Single Link Game

We now consider an alternative model where the users of a single link are price anticipating, rather than price taking, and play a game to acquire a share of the link. Throughout the remainder of this section and the next, we will assume that the link manager sets the price $\mu(\mathbf{w})$ according to the unique choice prescribed by Proposition 1, as follows.

*Assumption 3:* For any $\mathbf{w} \geq 0$, the aggregate rate $f(\mathbf{w})$ is the solution to (7): $\sum_r w_r = f(\mathbf{w})p(f(\mathbf{w}))$. Furthermore, for each $r$, $d_r(\mathbf{w})$ is given by

$$d_r(\mathbf{w}) = \begin{cases} 0, & \text{if } w_r = 0 \\ \frac{w_r}{p(f(\mathbf{w}))}, & \text{if } w_r > 0. \end{cases} \tag{9}$$

Note that we have $f(\mathbf{w}) > 0$ and $p(f(\mathbf{w})) > 0$ if $\sum_r w_r > 0$ and, hence, $d_r$ is always well defined.

We adopt the notation $\mathbf{w}_{-r}$ to denote the vector of all bids by users other than $r$, i.e., $\mathbf{w}_{-r} = (w_1, w_2, \ldots, w_{r-1}, w_{r+1}, \ldots, w_R)$. Given $\mathbf{w}_{-r}$, each user $r$ chooses $w_r \geq 0$ to maximize

$$Q_r(w_r; \mathbf{w}_{-r}) \triangleq U_r\left(d_r(\mathbf{w})\right) - w_r \tag{10}$$

over nonnegative $w_r$. The payoff function $Q_r$ is similar to the payoff function $P_r$, except that the user now anticipates that the network will set the price according to Assumption 3, as captured by the allocated rate $d_r(\mathbf{w})$. A *Nash equilibrium* of the game defined by $(Q_1, \ldots, Q_R)$ is a vector $\mathbf{w} \geq 0$ such that for all $r$

$$Q_r(w_r; \mathbf{w}_{-r}) \geq Q_r(\overline{w}_r; \mathbf{w}_{-r}) \quad \text{for all } \overline{w}_r \geq 0. \tag{11}$$

In the next section, we show that a Nash equilibrium always exists, and give necessary and sufficient conditions for a vector $\mathbf{w}$ to be a Nash equilibrium. In Section II-B, we outline a class of price functions for which the Nash equilibrium is unique.

### A. Existence of Nash Equilibrium

In this section, we establish that a Nash equilibrium exists for the game defined by $(Q_1, \ldots, Q_R)$. We start by establishing certain properties of $d_r(\mathbf{w})$ in the following proposition.

*Proposition 3:* Suppose that Assumptions 1–3 hold. Then: 1) $d_r(\mathbf{w})$ is a continuous function of $\mathbf{w}$; and 2) for any $\mathbf{w}_{-r} \geq 0$, $d_r(\mathbf{w})$ is strictly increasing and concave in $w_r \geq 0$, and $d_r(\mathbf{w}) \to \infty$ as $w_r \to \infty$.

*Proof:* We first show 1): that $d_r(\mathbf{w})$ is a continuous function of $\mathbf{w}$. Recall from Proposition 1 that $f(\mathbf{w})$ is a continuous function of $\mathbf{w}$, and $f(\mathbf{0}) = 0$. Now at any vector $\mathbf{w}$ such that $\sum_s w_s > 0$, we have $p(f(\mathbf{w})) > 0$, so $d_r(\mathbf{w}) = w_r/p(f(\mathbf{w}))$; thus continuity of $d_r$ at $\mathbf{w}$ follows by continuity of $f$ and $p$. Suppose instead that $\mathbf{w} = \mathbf{0}$, and consider a sequence $\mathbf{w}^n$ such that $\mathbf{w}^n \to \mathbf{0}$ as $n \to \infty$. Then, $\sum_r d_r(\mathbf{w}^n) = f(\mathbf{w}^n) \to 0$ as $n \to \infty$, from parts 1) and 2) of Proposition 1; since $d_r(\mathbf{w}^n) \geq 0$ for all $n$, we must have $d_r(\mathbf{w}^n) \to 0 = d_r(\mathbf{0})$ as $n \to \infty$, as required.

We now show 2): that $d_r(\mathbf{w})$ is concave and strictly increasing in $w_r \geq 0$, with $d_r(\mathbf{w}) \to \infty$ as $w_r \to \infty$. From Assumption 3, we can rewrite the definition of $d_r(\mathbf{w})$ as

$$d_r(\mathbf{w}) = \begin{cases} 0, & \text{if } w_r = 0 \\ \frac{w_r}{\sum_s w_s} f(\mathbf{w}), & \text{if } w_r > 0. \end{cases} \quad (12)$$

From this expression and Proposition 1, it follows that $d_r(\mathbf{w})$ is strictly increasing in $w_r$. To show $d_r(\mathbf{w}) \to \infty$ as $w_r \to \infty$, we only need that $f(\mathbf{w}) \to \infty$ as $w_r \to \infty$, a fact that was shown in Proposition 1.

It remains to be shown that for fixed $\mathbf{w}_{-r}$, $d_r$ is a concave function of $w_r \geq 0$. Since we have already shown that $d_r$ is continuous, we may assume without loss of generality that $w_r > 0$. We will only consider the case where $p$ is twice differentiable; the extension to general $p$ uses a simple limiting argument, and details can be found in [24]. When $p$ is twice differentiable, it follows from (7) that $f$ is twice differentiable in $w_r$. Since $w_r > 0$, we can differentiate (12) twice to find

$$\frac{\partial^2 d_r(\mathbf{w})}{\partial w_r^2} = -\frac{2\sum_{s\neq r} w_s}{\left(\sum_s w_s\right)^3} f(\mathbf{w}) + \frac{2\sum_{s\neq r} w_s}{\left(\sum_s w_s\right)^2} \cdot \frac{\partial f(\mathbf{w})}{\partial w_r} + \frac{w_r}{\sum_s w_s} \cdot \frac{\partial^2 f(\mathbf{w})}{\partial w_r^2}.$$

From Proposition 1, $f$ is a strictly concave function of $\sum_s w_s$; thus the last term in the sum above is nonpositive. To show that $d_r$ is concave in $w_r$, therefore, it suffices to show that the sum of the first two terms is negative, i.e.,

$$\frac{f(\mathbf{w})}{\sum_s w_s} \geq \frac{\partial f(\mathbf{w})}{\partial w_r}. \quad (13)$$

By differentiating both sides of (7), we find that

$$\frac{\partial f(\mathbf{w})}{\partial w_r} = \frac{1}{p\left(f(\mathbf{w})\right) + f(\mathbf{w}) p'\left(f(\mathbf{w})\right)}.$$

From (7), we have

$$\frac{f(\mathbf{w})}{\sum_s w_s} = \frac{1}{p\left(f(\mathbf{w})\right)}.$$

Substituting these relations, and noting that $f(\mathbf{w})p'(f(\mathbf{w})) \geq 0$ since $p$ is strictly increasing, we conclude that (13) holds, as required. Thus $d_r(\mathbf{w})$ is concave in $w_r$, as long as $p$ is twice differentiable. □

The previous proposition establishes concavity and continuity of $d_r$; this guarantees existence of a Nash equilibrium, as the following proposition shows. The proof is an application of Rosen's existence theorem [26]; details may be found in [24].

*Proposition 4:* Suppose that Assumptions 1–3 hold. Then there exists a Nash equilibrium $\mathbf{w}$ for the game defined by $(Q_1, \ldots, Q_R)$.

In the remainder of this section, we establish necessary and sufficient conditions for a vector $\mathbf{w}$ to be a Nash equilibrium. Because the price function $p$ may not be differentiable, we will use *subgradients* to describe necessary local conditions for a vector $\mathbf{w}$ to be a Nash equilibrium. Since the payoff of user $r$ is concave, these necessary conditions will in fact be sufficient for $\mathbf{w}$ to be a Nash equilibrium.

We begin with some concepts from convex analysis [27], [28]. An *extended real-valued function* is a function $g : \mathbb{R} \to [-\infty, \infty]$; such a function is called *proper* if $g(x) > -\infty$ for all $x$, and $g(x) < \infty$ for at least one $x$. We say that a scalar $\gamma$ is a *subgradient* of an extended real-valued function $g$ at $x$ if for all $\overline{x} \in \mathbb{R}$, we have $g(\overline{x}) \geq g(x) + \gamma(\overline{x} - x)$. The *subdifferential* of $g$ at $x$, denoted $\partial g(x)$, is the set of all subgradients of $g$ at $x$. Finally, given an extended real-valued function $g$, we denote the right directional derivative of $g$ at $x$ by $\partial^+ g(x)/\partial x$ and left directional derivative of $g$ at $x$ by $\partial^- g(x)/\partial x$ (if they exist). If $g$ is convex, then $\partial g(x) = [\partial^- g(x)/\partial x, \partial^+ g(x)/\partial x]$, provided the directional derivatives exist.

For the remainder of this paper, we view any price function $p$ as an extended real-valued convex function, by defining $p(f) = \infty$ for $f < 0$. Our first step is a lemma identifying the directional derivatives of $d_r$ as a function of $w_r$; for notational convenience, we introduce the following definitions of $\varepsilon^+(f)$ and $\varepsilon^-(f)$, for $f > 0$:

$$\varepsilon^+(f) \triangleq \frac{f}{p(f)} \cdot \frac{\partial^+ p(f)}{\partial f} \quad \varepsilon^-(f) \triangleq \frac{f}{p(f)} \cdot \frac{\partial^- p(f)}{\partial f}. \quad (14)$$

Note that under Assumption 2, we have $0 < \varepsilon^-(f) \leq \varepsilon^+(f)$ for $f > 0$.

*Lemma 5:* Suppose Assumptions 1–3 hold. Then, for all $\mathbf{w}$ with $\sum_s w_s > 0$, $d_r(\mathbf{w})$ is directionally differentiable with respect to $w_r$. These directional derivatives are given by

$$\frac{\partial^+ d_r(\mathbf{w})}{\partial w_r} = \frac{1}{p\left(f(\mathbf{w})\right)} \times \left(1 - \frac{d_r(\mathbf{w})}{f(\mathbf{w})} \cdot \frac{\varepsilon^+\left(f(\mathbf{w})\right)}{1 + \varepsilon^+\left(f(\mathbf{w})\right)}\right) \quad (15)$$

$$\frac{\partial^- d_r(\mathbf{w})}{\partial w_r} = \frac{1}{p\left(f(\mathbf{w})\right)} \times \left(1 - \frac{d_r(\mathbf{w})}{f(\mathbf{w})} \cdot \frac{\varepsilon^-\left(f(\mathbf{w})\right)}{1 + \varepsilon^-\left(f(\mathbf{w})\right)}\right). \quad (16)$$

Furthermore, $\partial^+ d_r(\mathbf{w})/\partial w_r > 0$, and if $w_r > 0$ then $\partial^- d_r(\mathbf{w})/\partial w_r > 0$.

*Proof:* Existence of the directional derivatives is obtained because $d_r(\mathbf{w})$ is a concave function of $w_r$ (from Proposition 3). Fix a vector $\mathbf{w}$ of bids, such that $\sum_r w_r > 0$. Since $f$ is an increasing concave function of $w_r$, and the convex function $p$ is directionally differentiable at $f(\mathbf{w})$ (see [28, Th. 23.1]), we can apply the chain rule to compute the right directional derivative of (7) with respect to $w_r$

$$1 = \frac{\partial^+ f(\mathbf{w})}{\partial w_r} \cdot p\left(f(\mathbf{w})\right) + f(\mathbf{w}) \cdot \frac{\partial^+ p\left(f(\mathbf{w})\right)}{\partial f} \cdot \frac{\partial^+ f(\mathbf{w})}{\partial w_r}.$$

Thus, as long as $\sum_r w_r > 0$, $\partial^+ f(\mathbf{w})/\partial w_r$ exists, and is given by

$$\frac{\partial^+ f(\mathbf{w})}{\partial w_r} = \left(p\left(f(\mathbf{w})\right) + f(\mathbf{w}) \cdot \frac{\partial^+ p\left(f(\mathbf{w})\right)}{\partial f}\right)^{-1}.$$

We conclude from (9) that the right directional derivative of $d_r(\mathbf{w})$ with respect to $w_r$ is given by

$$\frac{\partial^+ d_r(\mathbf{w})}{\partial w_r} = \frac{1}{p(f(\mathbf{w}))} - \frac{w_r}{(p(f(\mathbf{w})))^2} \cdot \frac{\partial^+ p(f(\mathbf{w}))}{\partial f} \cdot \frac{\partial^+ f(\mathbf{w})}{\partial w_r}.$$

Simplifying, this reduces to (15). Note that since $d_r(\mathbf{w}) \le f(\mathbf{w})$ and $\varepsilon^+(f(\mathbf{w}))/[1+\varepsilon^+(f(\mathbf{w}))] < 1$, we have $\partial^+ d_r(\mathbf{w})/\partial w_r > 0$. A similar analysis follows for the left directional derivative.□

For notational convenience, we make the following definitions for $f > 0$:

$$\beta^+(f) \triangleq \frac{\varepsilon^+(f)}{1+\varepsilon^+(f)} \quad \beta^-(f) \triangleq \frac{\varepsilon^-(f)}{1+\varepsilon^-(f)}. \tag{17}$$

Under Assumption 2, we have $0 < \beta^-(f) \le \beta^+(f) < 1$ for $f > 0$.

The next proposition is the central result of this section: It provides simple local conditions that are necessary and sufficient for a vector $\mathbf{w}$ to be a Nash equilibrium.

*Proposition 6:* Suppose that Assumptions 1–3 hold. Then, $\mathbf{w}$ is a Nash equilibrium of the game defined by $(Q_1, \ldots, Q_R)$, if and only if $\sum_r w_r > 0$, and with $\mathbf{d} = \mathbf{d}(\mathbf{w})$, $f = f(\mathbf{w})$, the following two conditions hold for all $r$:

$$U_r'(d_r)\left(1 - \beta^+(f) \cdot \frac{d_r}{f}\right) \le p(f) \tag{18}$$

$$U_r'(d_r)\left(1 - \beta^-(f) \cdot \frac{d_r}{f}\right) \ge p(f), \quad \text{if } d_r > 0. \tag{19}$$

Conversely, if $\mathbf{d} \ge 0$ and $f > 0$ satisfy (18), (19), and $\sum_r d_r = f$, then the vector $\mathbf{w} = p(f)\mathbf{d}$ is a Nash equilibrium with $\mathbf{d} = \mathbf{d}(\mathbf{w})$ and $f = f(\mathbf{w})$.

*Proof:* We first show that if $\mathbf{w}$ is a Nash equilibrium, then we must have $\sum_r w_r > 0$. Suppose not; then $w_r = 0$ for all $r$. Fix a user $r$; for $\overline{w}_r > 0$, we have $d_r(\overline{w}_r; \mathbf{w}_{-r})/\overline{w}_r = f(\overline{w}_r; \mathbf{w}_{-r})/\overline{w}_r = 1/p(f(\overline{w}_r; \mathbf{w}_{-r}))$, which approaches infinity as $\overline{w}_r \to 0$. Thus, $\partial^+ d_r(\mathbf{w})/\partial w_r = \infty$, and thus we have

$$\frac{\partial^+ Q_r(w_r; \mathbf{w}_{-r})}{\partial w_r} = U_r'(0) \cdot \frac{\partial^+ d_r(\mathbf{w})}{\partial w_r} - 1 = \infty.$$

In particular, an infinitesimal increase of $w_r$ strictly increases the payoff of user $r$, so $\mathbf{w} = \mathbf{0}$ cannot be a Nash equilibrium. Thus, if $\mathbf{w}$ is a Nash equilibrium, then $\sum_r w_r > 0$.

Now, let $\mathbf{w}$ be a Nash equilibrium. We established in Lemma 5 that $d_r$ is directionally differentiable in $w_r$ for each $r$, as long as $\sum_s w_s > 0$. Thus, from (11), if $\mathbf{w}$ is a Nash equilibrium, then the following two conditions must hold:

$$\frac{\partial^+ Q_r(w_r; \mathbf{w}_{-r})}{\partial w_r} = U_r'(d_r(\mathbf{w})) \cdot \frac{\partial^+ d_r(\mathbf{w})}{\partial w_r} - 1 \le 0$$

$$\frac{\partial^- Q_r(w_r; \mathbf{w}_{-r})}{\partial w_r} = U_r'(d_r(\mathbf{w})) \cdot \frac{\partial^- d_r(\mathbf{w})}{\partial w_r} - 1 \ge 0, \quad \text{if } w_r > 0.$$

We may substitute using Lemma 5 to find that if $\mathbf{w}$ is a Nash equilibrium, then

$$U_r'(d_r(\mathbf{w}))\left(1 - \beta^+(f(\mathbf{w})) \cdot \frac{d_r(\mathbf{w})}{f(\mathbf{w})}\right) \le p(f(\mathbf{w}))$$

$$U_r'(d_r(\mathbf{w}))\left(1 - \beta^-(f(\mathbf{w})) \cdot \frac{d_r(\mathbf{w})}{f(\mathbf{w})}\right) \ge p(f(\mathbf{w})) \quad \text{if } w_r > 0.$$

Since the condition $w_r > 0$ is identical to the condition $d_r(\mathbf{w}) > 0$, this establishes the conditions in the proposition. Conversely, if $\sum_r w_r > 0$ and the preceding two conditions hold, then we may reverse the argument: since the payoff function of user $r$ is a concave function of $w_r$ for each $r$ (from Proposition 3), (18), (19) are sufficient for $\mathbf{w}$ to be a Nash equilibrium.

Finally, suppose that $\mathbf{d}$ and $f > 0$ satisfy (18), (19), with $\sum_r d_r = f$. Then let $w_r = d_r p(f)$. We then have $\sum_r w_r > 0$ (since $f > 0$); and $\sum_r w_r = fp(f)$, so that $f = f(\mathbf{w})$. Since $f > 0$, we have $d_r = w_r/p(f) = w_r/p(f(\mathbf{w}))$, so that $d_r = d_r(\mathbf{w})$. Thus, $\mathbf{w}$ is a Nash equilibrium, as required. □

Note that the preceding proposition identifies a Nash equilibrium entirely in terms of the allocation made; and conversely, if we find a pair $(\mathbf{d}, f)$ which satisfies (18), (19) with $f > 0$ and $\sum_r d_r = f$, then there exists a Nash equilibrium which yields that allocation. In particular, the set of allocations $\mathbf{d}$ which can arise at Nash equilibria coincides with those vectors $\mathbf{d}$ such that $f = \sum_r d_r > 0$, and (18), (19) are satisfied.

### *B. Nondecreasing Elasticity Price Functions: Uniqueness of Nash Equilibrium*

In this section, we demonstrate that for a certain class of differentiable price functions, there exists a *unique* Nash equilibrium of the game defined by $(Q_1, \ldots, Q_R)$. We consider price functions $p$ which satisfy the following additional assumption.

*Assumption 4:* The price function $p$ is differentiable, and exhibits *nondecreasing elasticity*: For $0 < f_1 \le f_2$, there holds

$$\frac{f_1 p'(f_1)}{p(f_1)} \le \frac{f_2 p'(f_2)}{p(f_2)}.$$

To gain some intuition for the concept of nondecreasing elasticity, consider a price function $p$ satisfying Assumption 2. The quantity $fp'(f)/p(f)$ is known as the *elasticity* of a price function $p$ [5]. Note that the elasticity of $p(f)$ is the derivative of $\ln(p(f))$ with respect to $\ln f$. From this viewpoint, we see that nondecreasing elasticity is equivalent to the requirement that $\ln(p(f))$ is a convex function in $\ln f$. (Note that this is not equivalent to the requirement that $p$ is a convex function of $f$.)

Nondecreasing elasticity can also be interpreted by considering the price function as the inverse of the *supply function* $s(\mu) = p^{-1}(\mu)$; the supply function gives the amount of rate the provider is willing to supply at a given price $\mu$ [5]. In this case, nondecreasing elasticity of the price function is equivalent to nonincreasing elasticity of the supply function.

Nondecreasing elasticity captures a wide range of price functions; we give two common examples that follow.

*Example 1 (The M/M/1 Queue):* Consider the cost function $C(f) = af/(s - f)$, where $a > 0$ and $s > 0$ are constants;

then the cost is proportional to the steady-state queue size in an M/M/1 queue with service rate $s$ and arrival rate $f$. (Note that we must view $p$ as an extended real-valued function, with $p(f) = \infty$ for $f > s$; this does not affect any of the analysis of this paper.) It is straightforward to check that, as long as $0 < f < s$, we have

$$\frac{fp'(f)}{p(f)} = \frac{2f}{s-f}$$

which is a strictly increasing function of $f$. Thus, $p$ satisfies Assumption 4.

*Example 2 (M/M/1 Overflow Probability):* Consider the function $p(f) = a(f/s)^B$, where $a > 0$, $s > 0$, and $B \geq 1$ is an integer. Then, the price is set proportional to the probability that an M/M/1 queue exceeds a buffer level $B$, when the service rate is $s$ and the arrival rate is $f$. In this case we have $fp'(f)/p(f) = B$, so that $p$ satisfies Assumption 4.

We now prove the key property of differentiable nondecreasing elasticity price functions in the current development: For such functions, there exists a unique Nash equilibrium of the game defined by $(Q_1, \ldots, Q_R)$.

*Proposition 7:* Suppose Assumptions 1–3 hold. If in addition $p$ is differentiable and exhibits nondecreasing elasticity (Assumption 4 holds), then there exists a unique Nash equilibrium for the game defined by $(Q_1, \ldots, Q_R)$.

*Proof:* We use the expressions (18), (19) to show that the Nash equilibrium is unique under Assumption 4. Observe that in this case, from (17), we may define $\beta(f) = \beta^+(f) = \beta^-(f)$ for $f > 0$, and conclude that $\mathbf{w}$ is a Nash equilibrium if and only if $\sum_s w_s > 0$ and the following optimality conditions hold:

$$U_r'(d_r(\mathbf{w}))\left(1 - \beta(f(\mathbf{w})) \cdot \frac{d_r(\mathbf{w})}{f(\mathbf{w})}\right) = p(f(\mathbf{w})), \quad \text{if } w_r > 0 \tag{20}$$

$$U_r'(0) \leq p(f(\mathbf{w})), \quad \text{if } w_r = 0. \tag{21}$$

Suppose we have two Nash equilibria $\mathbf{w}^1$, $\mathbf{w}^2$, with $0 < \sum_s w_s^1 < \sum_s w_s^2$; then $p(f(\mathbf{w}^1)) < p(f(\mathbf{w}^2))$, and $f(\mathbf{w}^1) < f(\mathbf{w}^2)$. Note that $U_r'(d_r)$ is nonincreasing as $d_r$ increases; and $\beta(f)$ is nondecreasing as $f$ increases (from Assumption 4) and, therefore, $\beta(f(\mathbf{w}^1)) \leq \beta(f(\mathbf{w}^2))$. Furthermore, if $w_r^2 > 0$, then from (20) we have $U_r'(0) > p(f(\mathbf{w}^2))$; thus $U_r'(0) > p(f(\mathbf{w}^1))$, so $w_r^1 > 0$ as well [from (21)].

Now note that the right-hand side of (20) is strictly larger at $\mathbf{w}^1$ than at $\mathbf{w}^2$; thus the left hand side must be strictly larger at $\mathbf{w}^1$ than at $\mathbf{w}^2$ as well. This is only possible if $d_r(\mathbf{w}^1)/f(\mathbf{w}^1) > d_r(\mathbf{w}^2)/f(\mathbf{w}^2)$ for each user $r$, since we have shown in the preceding paragraph that $f(\mathbf{w}^1) < f(\mathbf{w}^2)$; $U_r'(d_r)$ is nonincreasing as $d_r$ increases; and $\beta(f(\mathbf{w}^1)) \leq \beta(f(\mathbf{w}^2))$. Since $f(\mathbf{w}) = \sum_r d_r(\mathbf{w})$, we have

$$1 = \sum_{r:w_r^2>0} \frac{d_r(\mathbf{w}^2)}{f(\mathbf{w}^2)} < \sum_{r:w_r^2>0} \frac{d_r(\mathbf{w}^1)}{f(\mathbf{w}^1)} \leq 1$$

which is a contradiction. Thus, at the two Nash equilibria, we must have $\sum_s w_s^1 = \sum_s w_s^2$, so we can let $f_0 = f(\mathbf{w}^1) = f(\mathbf{w}^2)$, $p_0 = p(f_0)$, and $\beta_0 = \beta(f_0)$. Then, all Nash equilibria $\mathbf{w}$ satisfy

$$U_r'(d_r(\mathbf{w}))\left(1 - \beta_0 \frac{d_r(\mathbf{w})}{f_0}\right) = p_0, \quad \text{if } w_r > 0 \tag{22}$$

$$U_r'(0) \leq p_0, \quad \text{if } w_r = 0. \tag{23}$$

However, now we observe that the left hand side of (22) is strictly decreasing in $d_r(\mathbf{w})$, so given $p_0$, there exists at most one solution $d_r(\mathbf{w})$ to (22). Since $w_r = d_r(\mathbf{w})p_0$, this implies the Nash equilibrium $\mathbf{w}$ must be unique. □

We observe that uniqueness of the Nash equilibrium implies an additional desirable property in the case of symmetric users. If two users share the same utility function, and the price function $p$ is differentiable, we conclude from Proposition 7 that at the unique Nash equilibrium, these users submit exactly the same bid (and, hence, receive exactly the same rate allocation).

We also note that in general, Nash equilibria need not be unique. As an example, consider a case with two users, where $U_r(d_r) = 3d_r$, for $r = 1, 2$. Let the price function be $p(f) = f$ for $0 \leq f \leq 1$, and $p(f) = 2f - 1$ for $f \geq 1$. Then, it is straightforward to verify that any pair $(d_1, d_2)$ such that $1/3 \leq d_i \leq 2/3$ for $i = 1, 2$, and $d_1 + d_2 = 1$, satisfies (18), (19) with $f = 1$; thus, by Proposition 6, there exist Nash equilibria that correspond to each of these $(d_1, d_2)$. Note, however, that the price function $p$ is not differentiable; it is not clear whether, in general, differentiability of $p$ suffices to guarantee a unique Nash equilibrium.

## III. Efficiency Loss: The Single Link Case

We let $\mathbf{d}^S$ denote an optimal solution to SYSTEM, defined in (1), (2), and let $\mathbf{w}$ denote any Nash equilibrium of the game defined by $(Q_1, \ldots, Q_R)$. We now investigate the efficiency loss of this system; that is, how much aggregate surplus is lost because the users attempt to "game" the system? To answer this question, we must compare the aggregate surplus $\sum_r U_r(d_r(\mathbf{w})) - C(\sum_r d_r(\mathbf{w}))$ obtained when the users fully evaluate the effect of their actions on the price, and the aggregate surplus $\sum_r U_r(d_r^S) - C(\sum_r d_r^S)$ obtained by choosing an allocation which maximizes aggregate surplus. The following theorem is the main result of this paper: It states that the efficiency loss is no more than approximately 34%, and that this bound is essentially tight.

*Theorem 8:* Suppose that Assumptions 1–3 hold. Suppose also that $U_r(0) \geq 0$ for all $r$. Let $\mathbf{d}^S$ be any solution to $SYSTEM$, and let $\mathbf{w}$ be any Nash equilibrium of the game defined by $(Q_1, \ldots, Q_R)$. Then, we have the following bound:

$$\text{surplus}(\mathbf{d}(\mathbf{w})) \geq (4\sqrt{2} - 5) \cdot \text{surplus}(\mathbf{d}^S) \tag{24}$$

where surplus$(\cdot)$ is defined in (3). In other words, there is no more than approximately a 34% efficiency loss when users are price anticipating.

Furthermore, this bound is tight: For every $\delta > 0$, there exists a choice of $R$, a choice of (linear) utility functions $U_r$, $r = 1, \ldots, R$, and a (piecewise linear) price function $p$ such that a Nash equilibrium $\mathbf{w}$ and a solution $\mathbf{d}^S$ to SYSTEM exist with

$$\text{surplus}(\mathbf{d}(\mathbf{w})) \leq (4\sqrt{2} - 5 + \delta) \cdot \text{surplus}(\mathbf{d}^S). \tag{25}$$

*Proof:* The proof of (24) consists of a sequence of steps.

1) We show that the worst case ratio occurs when the utility function of each user is linear.
2) We restrict attention to games where the total allocated Nash equilibrium rate is $f = 1$.
3) We compute the worst case choice of linear utility functions, for a fixed price function $p(\cdot)$ and total Nash equilibrium rate $f = 1$.
4) We prove that it suffices to consider a special class of piecewise linear price functions.
5) Combining steps 1)–3), we compute the worst case efficiency loss by minimizing the ratio of Nash equilibrium aggregate surplus to maximal aggregate surplus, over the worst case choice of games with linear utility functions [from step 2)] and our restricted class of piecewise linear price functions [from step 3)].

*Step 1: Show that we may assume without loss of generality that $U_r$ is linear for each user $r$*, i.e., *without loss of generality we may assume $U_r(d_r) = \alpha_r d_r$, where $\alpha_1 = 1$ and $0 < \alpha_r \leq 1$ for $r > 1$*. The proof of this claim is similar to the proof of [23, Lemma 4]. Let $\mathbf{d}^S$ denote any solution to $SYSTEM$, and let $\mathbf{w}$ denote a Nash equilibrium, for an arbitrary collection of utility functions $(U_1, \ldots, U_R)$ satisfying the assumptions of the theorem. We let $\mathbf{d} = \mathbf{d}(\mathbf{w})$ denote the allocation vector at the Nash equilibrium. For each user $r$, we define a new utility function $\overline{U}_r(d_r) = \alpha_r d_r$, where $\alpha_r = U_r'(d_r)$; we know that $\alpha_r > 0$ by Assumption 1. Then, observe that if we replace the utility functions $(U_1, \ldots, U_R)$ with the linear utility functions $(\overline{U}_1, \ldots, \overline{U}_R)$, the vector $\mathbf{w}$ remains a Nash equilibrium; this follows from the necessary and sufficient conditions of Proposition 6.

We first show that $\sum_r \alpha_r d_r - C(f) > 0$. To see this, note from (19) that $\alpha_r > p(f)$ for all $r$ such that $d_r > 0$. Thus, $\alpha_r d_r > d_r p(f)$ for such a user $r$, so $\sum_r \alpha_r d_r > f p(f) \geq C(f)$, by convexity (Assumption 2).

Next, we note that $\sum_r U_r(d_r^S) - C(\sum_r d_r^S) > 0$. This follows since $U_r$ is strictly increasing and nonnegative, while $C'(0) = p(0) = 0$; thus if $\overline{d}_r$ is sufficiently small for all $r$, we will have $\sum_r U_r(\overline{d}_r) - C(\sum_r \overline{d}_r) > 0$, which implies $\sum_r U_r(d_r^S) - C(\sum_r d_r^S) > 0$ (since $\mathbf{d}^S$ is a solution to SYSTEM).

Using concavity, we have for each $r$ that $U_r(d_r^S) \leq U_r(d_r) + \alpha_r(d_r^S - d_r)$. Defining $\Omega_r = U_r(d_r) - \alpha_r d_r$ and expanding the definition of surplus$(\cdot)$, we have

$$\begin{aligned}\frac{\text{surplus}(\mathbf{d})}{\text{surplus}(\mathbf{d}^S)} &= \frac{\sum_r U_r(d_r) - C\left(\sum_r d_r\right)}{\sum_r U_r\left(d_r^S\right) - C\left(\sum_r d_r^S\right)} \\ &\geq \frac{\sum_r \Omega_r + \sum_r \alpha_r d_r - C\left(\sum_r d_r\right)}{\sum_r \Omega_r + \sum_r \alpha_r d_r^S - C\left(\sum_r d_r^S\right)} \\ &\geq \frac{\sum_r \Omega_r + \sum_r \alpha_r d_r - C\left(\sum_r d_r\right)}{\sum_r \Omega_r + \max_{\overline{\mathbf{d}} \geq 0}\left(\sum_r \alpha_r \overline{d}_r - C\left(\sum_r \overline{d}_r\right)\right)}.\end{aligned}$$

(Note that all denominators are positive, since we have shown that $\sum_r U_r(d_r^S) - C(\sum_r d_r^S) > 0$.) Since we assumed $U_r(0) \geq 0$, we have $\Omega_r = U_r(d_r) - U_r'(d_r)d_r \geq 0$ by concavity; and since $0 < \sum_r \alpha_r d_r - C(f) \leq \max_{\overline{\mathbf{d}} \geq 0}(\sum_r \alpha_r \overline{d}_r - C(\sum_r \overline{d}_r))$, we have the inequality

$$\begin{aligned}&\frac{\sum_r U_r(d_r) - C\left(\sum_r d_r\right)}{\sum_r U_r\left(d_r^S\right) - C\left(\sum_r d_r^S\right)} \\ &\qquad \geq \frac{\sum_r \alpha_r d_r - C\left(\sum_r d_r\right)}{\max_{\overline{\mathbf{d}} \geq 0}\left(\sum_r \alpha_r \overline{d}_r - C\left(\sum_r \overline{d}_r\right)\right)}.\end{aligned}$$

Now, observe that the right hand side of the previous expression is the ratio of the Nash equilibrium aggregate surplus to the maximal aggregate surplus, when the utility functions are $(\overline{U}_1, \ldots, \overline{U}_R)$; since this ratio is no larger than the same ratio for the original utility functions $(U_1, \ldots, U_R)$, we can restrict attention to games where the utility function of each user is linear. Finally, by replacing $\alpha_r$ by $\alpha_r/(\max_s \alpha_s)$, and the cost function $C(\cdot)$ by $C(\cdot)/(\max_s \alpha_s)$, we may assume without loss of generality that $\max_r \alpha_r = 1$. Thus, by relabeling the users if necessary, we assume for the remainder of the proof that $U_r(d_r) = \alpha_r d_r$ for all $r$, where $\alpha_1 = 1$ and $0 < \alpha_r \leq 1$ for $r > 1$.

Before continuing, we observe that under these conditions, we have the following relation:

$$\max_{\overline{\mathbf{d}} \geq 0}\left(\sum_r \alpha_r \overline{d}_r - C\left(\sum_r \overline{d}_r\right)\right) = \max_{\overline{f} \geq 0}\left(\overline{f} - C(\overline{f})\right).$$

To see this, note that at any fixed value of $\overline{f} = \sum_r \overline{d}_r$, the left hand side is maximized by allocating the entire rate $\overline{f}$ to user 1. Thus, the ratio of Nash equilibrium aggregate surplus to maximal aggregate surplus becomes

$$\frac{\sum_r \alpha_r d_r - C\left(\sum_r d_r\right)}{\max_{\overline{f} \geq 0}\left(\overline{f} - C(\overline{f})\right)}. \tag{26}$$

Note that the denominator is positive, since $C'(0) = p(0) = 0$; and further, the optimal solution in the denominator occurs at the unique value of $\overline{f} > 0$ such that $p(\overline{f}) = 1$.

*Step 2: Show that we may restrict attention to games where the total allocated rate at the Nash equilibrium is $f = 1$.* Fix a cost function $C$ satisfying Assumption 2. Let $\mathbf{w}$ be a Nash equilibrium, and let $\mathbf{d} = \mathbf{d}(\mathbf{w})$ be the resulting allocation. Let $f = \sum_r d_r$ be the total allocated rate at the Nash equilibrium; note that $f > 0$ by Proposition 6. We now define a new price function $\hat{p}$ according to $\hat{p}(\hat{f}) = p(f \cdot \hat{f})$, and a new cost function $\hat{C}(\hat{f}) = \int_0^{\hat{f}} \hat{p}(z)dz$; note that $\hat{C}(\hat{f}) = C(f \cdot \hat{f})/f$. Then, it is straightforward to check that $\hat{p}$ satisfies Assumption 2. We will use hats to denote the corresponding functions when the price function is $\hat{p}$: $\hat{\beta}^+(\hat{f})$, $\hat{\beta}^-(\hat{f})$, $\hat{d}_r(\mathbf{w})$, $\hat{f}(\mathbf{w})$, etc.

Define $\hat{w}_r = w_r/f$. Then we claim that $\hat{\mathbf{w}}$ is a Nash equilibrium when the price function is $\hat{p}$. First observe that $\sum_r \hat{w}_r = \sum_r w_r/f = p(f) = \hat{p}(1)$; thus $\hat{f}(\hat{\mathbf{w}}) = 1$. Furthermore, $\hat{d}_r(\hat{\mathbf{w}}) = \hat{w}_r/\hat{p}(\hat{f}(\hat{\mathbf{w}})) = \hat{w}_r/\hat{p}(1) = w_r/(fp(f)) = d_r/f$. Finally, note that

$$\frac{\partial^+ \hat{p}(1)}{\partial \hat{f}} = f\frac{\partial^+ p(f)}{\partial f}$$

from which we conclude that $\hat{\beta}^+(1) = \beta^+(f)$, and similarly $\hat{\beta}^-(1) = \beta^-(f)$. Recall that $\mathbf{w}$ is a Nash equilibrium for the price function $p$; thus, if we combine the preceding conclusions and apply Proposition 6, we have that $\hat{\mathbf{w}}$ is a Nash equilibrium when the price function is $\hat{p}$, with total allocated rate $\hat{f} = 1$ and allocation $\hat{\mathbf{d}} = \mathbf{d}/f$.

To complete the proof of this step, we note the following chain of equalities:

$$\frac{\sum_r \alpha_r d_r - C\left(\sum_r d_r\right)}{\max_{\hat{f}\geq 0}\left(\hat{f} - C(\hat{f})\right)} = \frac{\sum_r \alpha_r \hat{d}_r - \hat{C}(1)}{\max_{\hat{f}\geq 0}\left(\frac{\hat{f}}{f} - \frac{C(\hat{f})}{f}\right)} = \frac{\sum_r \alpha_r \hat{d}_r - \hat{C}(1)}{\max_{g\geq 0}\left(g - \hat{C}(g)\right)}$$

where we make the substitution $g = \hat{f}/f$. But now note that the right hand side is the ratio of Nash equilibrium aggregate surplus to maximal aggregate surplus for a game where the total allocated rate at the Nash equilibrium is equal to 1. Consequently, in computing the worst case efficiency loss, we may restrict our attention to games where the Nash equilibrium allocated rate is equal to 1.

*Step 3: For a fixed price function* $p$, *determine the instance of linear utility functions that minimizes Nash equilibrium aggregate surplus, for a fixed Nash equilibrium allocated rate* $f = \sum_r d_r = 1$. Note that fixing the price function $p$ fixes the optimal aggregate surplus; thus minimizing the aggregate surplus at Nash equilibrium also yields the worst case efficiency loss.

We will optimize over the set of all games where users have linear utility functions [satisfying the conditions of step 1)], and where the total Nash equilibrium rate is $f = 1$. We use the necessary and sufficient conditions of Proposition 6. Note that by fixing the price function $p$ and the total rate $f > 0$, the Nash equilibrium price is fixed, $p(1)$, and $\beta^+(1)$ and $\beta^-(1)$ are fixed as well [from the definition (17)]; for notational convenience, we abbreviate $p = p(1)$, $C = C(1)$, $\beta^+ = \beta^+(1)$, and $\beta^- = \beta^-(1)$ for the duration of this step. Since $\alpha_1 = 1$, for a fixed value of $R$ the game with linear utility functions that minimizes aggregate surplus is given by solving the following optimization problem (with unknowns $d_1, \ldots, d_R, \alpha_2, \ldots, \alpha_R$):

$$\begin{aligned}
\text{minimize} \quad & d_1 + \sum_{r=2}^{R} \alpha_r d_r - C && (27)\\
\text{subject to} \quad & \alpha_r(1-\beta^+ d_r) \leq p, \qquad r = 1,\ldots,R && (28)\\
& \alpha_r(1-\beta^- d_r) \geq p, \quad \text{if} \quad d_r > 0 \\
& \qquad\qquad r = 1,\ldots,R && (29)\\
& \sum_{r=1}^{R} d_r = 1 && (30)\\
& 0 < \alpha_r \leq 1, \qquad r = 2,\ldots,R && (31)\\
& d_r \geq 0, \qquad r = 1,\ldots,R. && (32)
\end{aligned}$$

(Note that we have applied Proposition 6: if we solve the preceding problem and find an allocation $\mathbf{d}$ and coefficients $\boldsymbol{\alpha}$, then there exists a Nash equilibrium $\mathbf{w}$ with $\mathbf{d} = \mathbf{d}(\mathbf{w})$.) The objective function is the aggregate surplus given a Nash equilibrium allocation $\mathbf{d}$. Conditions (28) and (29) are equivalent to the Nash equilibrium conditions established in Proposition 6. The constraint (30) ensures that the total allocation made is equal to 1, and the constraint (31) follows from Step 1. The constraint (32) ensures the rate allocated to each user is nonnegative.

We solve this problem through a sequence of reductions. We first show we may assume without loss of generality that the constraint (29) holds with equality for all users $r = 2, \ldots, R$. The resulting problem is symmetric in the users $r = 2, \ldots, R$; we next show that a feasible solution exists if and only if $1 - \beta^+ \leq p < 1$ and $R$ is sufficiently large, and we conclude using a convexity argument that $d_r = (f - d_1)/(R-1)$ at an optimal solution. Finally, we show the worst-case occurs in the limit where $R \to \infty$, and calculate the resulting Nash equilibrium aggregate surplus.

We first show that it suffices to optimize over all $(\boldsymbol{\alpha}, \mathbf{d})$ such that (29) holds with equality for $r = 2, \ldots, R$. Note that if $(\boldsymbol{\alpha}, \mathbf{d})$ is a feasible solution to (27)–(32), then from (29)–(32), and the fact that $0 < \beta^- < 1$, we conclude that $p < 1$. Now, if $d_r > 0$ for some $r = 2, \ldots, R$, but the corresponding constraint in (29) does not hold with equality, we can reduce $\alpha_r$ until the constraint in (29) does hold with equality; by this process, we obtain a smaller value for the objective function (27). On the other hand, if $d_r = 0$ for some $r = 2, \ldots, R$, we can set $\alpha_r = p$; since $p < 1$, this preserves feasibility, but does not impact the term $\alpha_r d_r$ in the objective function (27). Therefore, we can restrict attention to feasible solutions for which

$$\alpha_r = \frac{p}{1 - \beta^- d_r}, \qquad r = 2, \ldots, R. \tag{33}$$

Having done so, observe that the constraint (31) that $\alpha_r \leq 1$ may be written as

$$d_r \leq \frac{1-p}{\beta^-}, \qquad r = 2, \ldots, R.$$

Finally, the constraint (31) that $\alpha_r > 0$ becomes redundant, as it is guaranteed by the fact that $d_r \leq 1$ [from (30)], $\beta^- < 1$ (by definition), and (33).

We now use the preceding observations to simplify the optimization problem (27)–(32) as follows:

$$\begin{aligned}
\text{minimize} \quad & d_1 + p\sum_{r=2}^{R} \frac{d_r}{1-\beta^- d_r} - C && (34)\\
\text{subject to} \quad & 1 - \beta^+ d_1 \leq p \leq 1 - \beta^- d_1 && (35)\\
& \sum_{r=1}^{R} d_r = 1 && (36)\\
& d_r \leq \frac{1-p}{\beta^-}, \qquad r = 2,\ldots,R && (37)\\
& d_r \geq 0, \qquad r = 1,\ldots,R. && (38)
\end{aligned}$$

The objective function (34) equals (27) upon substitution for $\alpha_r$ for $r = 2, \ldots, R$, from (33). We know that $d_1 > 0$ when $p(f) < 1$ [from (28) and (29)]; thus the constraint (35) is equivalent to the constraints (28) and (29) for user 1 with $d_1 > 0$. The constraint (28) for $r > 1$ is redundant and eliminated, since (29)

holds with equality for $r > 1$. The constraint (36) is equivalent to the allocation constraint (30); and the constraint (37) ensures $\alpha_r \leq 1$, as required in (31).

We first note that for a feasible solution to (34)–(38) to exist, we must have $1 - \beta^+ \leq p < 1$. We have already shown that we must have $p < 1$ if a feasible solution exists. Furthermore, from (35) we observe that the smallest feasible value of $d_1$ is $d_1 = (1-p)/\beta^+$. We require $d_1 \leq 1$ from (36) and (38), so we must have $(1-p)/\beta^+ \leq 1$, which yields the restriction that $1 - \beta^+ \leq p$. Thus, there only exist Nash equilibria with total rate 1 and price $p$ if:

$$1 - \beta^+ \leq p < 1. \tag{39}$$

We will assume for the remainder of this step that (39) is satisfied.

We note that if $\overline{\mathbf{d}} = (\overline{d}_1, \ldots, \overline{d}_R)$ is a feasible solution to (34)–(38) with $R$ users, then letting $\overline{d}_{R+1} = 0$, the vector $(\overline{d}_1, \ldots, \overline{d}_{R+1})$ is a feasible solution to (34)–(38) with $R+1$ users, and with the same objective function value (34) as $\overline{\mathbf{d}}$. Thus, the minimal objective function value cannot increase as $R$ increases, so the worst case efficiency loss occurs in the limit where $R \to \infty$.

We now solve (34)–(38) for a fixed feasible value of $d_1$. From the constraints (36), (37), we observe that a feasible solution to (34)–(38) exists if and only if the following condition holds in addition to (39):

$$d_1 + (R-1) \cdot \frac{1-p}{\beta^-} \geq 1. \tag{40}$$

In this case, the following symmetric solution is feasible:

$$d_r = \frac{1-d_1}{R-1}, \qquad r = 2, \ldots, R. \tag{41}$$

Furthermore, since the objective function is strictly convex and symmetric in the variables $d_2, \ldots, d_R$, and the feasible region is convex, the symmetric solution (41) must be optimal.

If we substitute the optimal solution (41) into the objective function (34) and take the limit as $R \to \infty$, then the constraint (40) is vacuously satisfied, and the objective function becomes $d_1 + p(1-d_1) - C$. Since we have shown that $p < 1$, the worst case occurs at the smallest feasible value of $d_1$; from (35), this value is

$$d_1 = \frac{1-p}{\beta^+}. \tag{42}$$

The resulting worst case Nash equilibrium aggregate surplus is

$$p + \frac{(1-p)^2}{\beta^+} - C.$$

To complete the proof of the theorem, we will consider the ratio of this Nash equilibrium aggregate surplus to the maximal aggregate surplus; we denote this ratio by $F(p)$ as a function of the price function $p(\cdot)$

$$F(p) = \frac{p(1) + \frac{(1-p(1))^2}{\beta^+(1)} - C(1)}{\max_{f \geq 0} (f - C(f))}. \tag{43}$$

Note that henceforth, the scalar $p$ used throughout step 3) will be denoted $p(1)$, and we return to denoting the price function by $p$. Thus, $F(p)$ as defined in (43) is a function of the entire price function $p(\cdot)$.

For completeness, we summarize in the following lemma an intermediate tightness result which will be necessary to prove the tightness of the bound in the theorem.

*Lemma 9:* Suppose that assumptions 2) and 3) are satisfied. Then, there exists $R > 0$ and a choice of linear utility functions $U_r(d_r) = \alpha_r d_r$, where $\alpha_1 = \max_s \alpha_s = 1$, with total Nash equilibrium rate 1, if and only if (39) is satisfied, i.e.,

$$1 - \beta^+(1) \leq p(1) < 1. \tag{44}$$

In this case, given $\delta > 0$, there exists $R > 0$ and a collection of $R$ users where user $r$ has utility function $U_r(d_r) = \alpha_r d_r$, such that $\mathbf{d}$ is a Nash equilibrium allocation with $\sum_r d_r = 1$, and

$$\frac{\sum_r \alpha_r d_r - C(1)}{\max_{\mathbf{d} \geq 0} \left(\sum_r \alpha_r d_r - C\left(\sum_r d_r\right)\right)} \leq F(p) + \delta. \tag{45}$$

*Proof of Lemma:* The proof follows from step 3). We have shown that if there exists a Nash equilibrium with total rate 1, then (44) must be satisfied. Conversely, if (44) is satisfied, we proceed as follows: Define $d_1$ according to (42); choose $R$ large enough that (40) is satisfied; define $d_r$ according to (41); and then define $\alpha_r$ according to (33) with $p = p(1)$. Then, it follows that $(\mathbf{d}, \boldsymbol{\alpha})$ is a feasible solution to (27)–(32), which (by Proposition 6) guarantees there exists a Nash equilibrium whose total allocated rate equals 1. The bound in (45) then follows by the proof of step 3). □

The remainder of the proof amounts to minimizing the worst case ratio of Nash equilibrium aggregate surplus to maximal aggregate surplus, over all valid choices of $p$. A valid choice of $p$ is any price function $p$ such that at least one choice of linear utility functions satisfying the conditions of step 1) leads to a Nash equilibrium with total allocated rate 1. By Lemma 9, all such functions $p$ are characterized by the constraint (44). We will minimize $F(p)$, given by (43), over all choices of $p$ satisfying (44).

*Step 4: Show that in minimizing $F(p)$ over $p$ satisfying* (44), *we may restrict attention to functions $p$ satisfying the following conditions*:

$$p(f) = \begin{cases} af, & 0 \leq f \leq 1 \\ a + b(f-1), & f \geq 1 \end{cases} \tag{46}$$

$$0 < a \leq b \tag{47}$$

$$\frac{1}{a+b} \leq 1 < \frac{1}{a}. \tag{48}$$

Observe that $p$ as defined in (46)–(48) is a convex, strictly increasing, piecewise linear function with two parts: an initial segment which increases at slope $a > 0$, and a second segment which increases at slope $b \geq a$. In particular, such a function satisfies Assumption 2. Furthermore, we have $\partial^+ p(1)/\partial f = b$, so that $\varepsilon^+(1) = b/a$. This implies $\beta^+(1) = b/(a+b)$; thus, multiplying through (48) by $a$ yields (44).

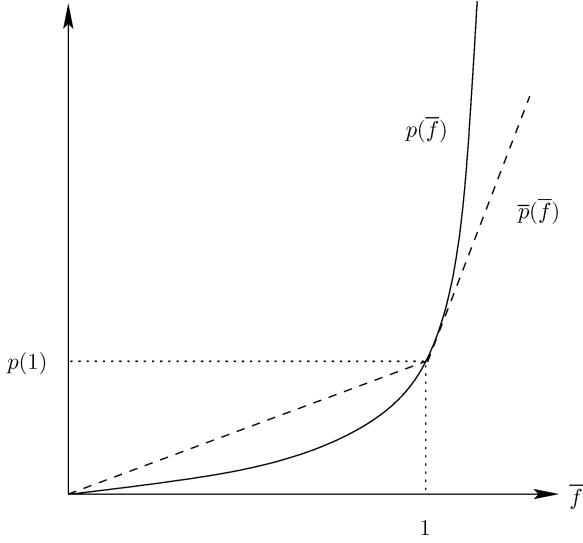



Fig. 1. Proof of Theorem 8, step 4): Given a price function $p$ (solid line) and Nash equilibrium rate 1, a new price function $\overline{p}$ (dashed line) is defined according to (49).

To verify the claim of step 4), we consider any function $p$ such that (44) holds. We define a new price function $\overline{p}$ as follows:

$$\overline{p}(f) = \begin{cases} fp(1), & 0 \le f \le 1 \\ p(1) + \frac{\partial^+ p(1)}{\partial f}(f-1), & f \ge 1. \end{cases} \quad (49)$$

(See Fig. 1 for an illustration.) Let $a = p(1)$, and let $b = \partial^+ p(1)/\partial f$. Then $a > 0$; and since $p(0) = 0$, we have $\partial^+ p(1)/\partial f \ge p(1)$ by convexity of $p$, so that $b \ge a$. Furthermore, since $p(1) < 1$ from (44), we have $1/a > 1$. Finally, we have

$$\frac{1}{a+b} = \frac{1}{p(1)}\left(1 - \beta^+(1)\right) \le 1$$

where the equality follows from the definition of $\beta^+(1)$ and the inequality follows from (44). Thus, $\overline{p}$ satisfies (46)–(48). Observe also that $\overline{p}(1) = p(1)$, and $\partial^+\overline{p}(1)/\partial f = \partial^+ p(1)/\partial f$, and thus $\overline{\beta}^+(1) = \beta^+(1)$.

We now show that $F(\overline{p}) \le F(p)$. As an intermediate step, we define a new price function $\hat{p}(\cdot)$ as follows:

$$\hat{p}(f) = \begin{cases} p(f), & 0 \le f \le 1 \\ \overline{p}(f), & f \ge 1. \end{cases}$$

Of course, $\hat{p}(1) = p(1)$ and $\partial^+\hat{p}(1)/\partial f = \partial^+\overline{p}(1)/\partial f = \partial^+ p(1)/\partial f$, so that (44) is satisfied for $\hat{p}$. Let $\hat{C}(f) = \int_0^f \hat{p}(z)\,dz$ denote the cost function associated with $\hat{p}(\cdot)$. Observe that (by convexity of $p$), we have for all $f$ that $\hat{p}(f) \le p(f)$, so that $\hat{C}(f) \le C(f)$. Thus

$$\max_{f\ge 0}\left(f - \hat{C}(f)\right) \ge \max_{f \ge 0}\left(f - C(f)\right).$$

Furthermore, $\hat{C}(1) = C(1)$ so that $F(\hat{p}) \le F(p)$.

Next, we let $\overline{C}(f) = \int_0^f \overline{p}(z)\,dz$ denote the cost function associated with $\overline{p}(\cdot)$. By convexity of $p$, we know $\overline{p}(1) \ge p(1)$ for $0 \le f \le 1$; thus $\overline{C}(f) \ge C(f)$ in that region. We let $\Delta \triangleq \overline{C}(1) - C(1) \ge 0$. Then, we have the following relationship:

$$\begin{aligned} F(\hat{p}) &= \frac{p(1) + \frac{(1-p(1))^2}{\beta^+(1)} - C(1)}{\max_{f\ge 0}\left(f - \hat{C}(f)\right)} \\ &\ge \frac{p(1) + \frac{(1-p(1))^2}{\beta^+(1)} - (C(1) + \Delta)}{\max_{f\ge 0}\left(f - \left(\hat{C}(f) + \Delta\right)\right)} \\ &= F(\overline{p}). \end{aligned}$$

The last equality follows by observing that since $\hat{p}(1) = p(1) < 1$, the solution to $\max_{f\ge 0}(f - \hat{C}(f))$ occurs at $\hat{f}^S > 1$ where $\hat{p}(\hat{f}^S) = 1$; and at all points $f \ge 1$, we have the relationship $\hat{C}(f) + \Delta = \overline{C}(f)$. Combining the preceding results, we have $F(p) \ge F(\overline{p})$, as required.

*Step 5: The minimum value of $F(p)$ over all $p$ satisfying* (46)–(48) *is* $4\sqrt{2} - 5$. We first show that given $p$ satisfying (46)–(48), $F(p)$ is given by

$$\begin{aligned} F(p) &= \frac{\frac{1}{2}a + \left(1 + \frac{a}{b}\right)(1-a)^2}{1 - \frac{1}{2}a + \frac{1}{2}\frac{(1-a)^2}{b}} \\ &= \frac{ab + 2(a+b)(1-a)^2}{2b - ab + (1-a)^2}. \end{aligned} \quad (50)$$

The numerator results by simplifying the numerator of (43), when $p$ takes the form described by (46)–(48). To arrive at the denominator, we note that the solution to $\max_{f\ge 0}(f - C(f))$ occurs at $f^S$ satisfying $p(f^S) = 1$. Since $a < 1$, we must have $f^S > 1$ and $a + b(f^S - 1) = 1$. Simplifying, we find

$$f^S = 1 + \frac{1-a}{b}. \quad (51)$$

The expression $f^S - C(f^S)$, upon simplification, becomes the denominator of (50), as required.

Fix $a$ and $b$ such that $0 < a \le b$, and $1/(a+b) \le 1 < 1/a$, and define $p$ as in (46). We note here that the constraints $0 < a \le b$ and $1/(a+b) \le 1 < 1/a$ may be equivalently rewritten as $0 < a < 1$, and $\max\{a, 1-a\} \le b$. Define $H(a,b) \triangleq F(p)$; from (50), note that for a fixed $a$, $H(a,b)$ is a ratio of two affine functions of $b$, and thus the minimal value of $H(a,b)$ is achieved either when $b = \max\{a, 1-a\}$ or as $b \to \infty$. Define $H_1(a) = H(a,b)|_{\max\{a,1-a\}}$, and $H_2(a) = \lim_{b\to\infty} H(a,b)$. Then

$$H_1(a) = \begin{cases} H(a,b)|_{b=1-a} = \frac{2-a}{3-2a} \\ \qquad \text{if } 0 < a \le \frac{1}{2} \\ H(a,b)|_{b=a} = a^2 + 4a(1-a)^2 \\ \qquad \text{if } \frac{1}{2} \le a < 1; \end{cases} \quad (52)$$

$$H_2(a) = \lim_{b\to\infty} H(a,b) = \frac{a + 2(1-a)^2}{2-a}. \quad (53)$$

We now minimize $H_1(a)$ and $H_2(a)$ over $0 < a < 1$. Over $0 < a \le 1/2$, the minimum value of $H_1(a)$ is 2/3, achieved as $a \to 0$. Over $1/2 \le a < 1$, the minimum value of $H_1(a)$ is 20/27, achieved at $a = 2/3$. Finally, over $0 < a < 1$, the minimum value of $H_2(a)$ is $4\sqrt{2} - 5$, achieved at $a = 2 - \sqrt{2}$. Since $\min\{2/3, 20/27, 4\sqrt{2} - 5\} = 4\sqrt{2} - 5$, we conclude

that the minimal value of $F(p)$ over all $p$ satisfying (46)–(48) is equal to $4\sqrt{2}-5$. This completes the proof of (24), the lower bound in the theorem.

We now show that this lower bound is tight. Fix $\delta > 0$. The preceding argument shows that the worst case occurs for price functions satisfying (46)–(48), where $a = 2-\sqrt{2}$ and $b \to \infty$. For fixed $b \geq a = 2-\sqrt{2}$, let $p_b$ be the associated price function defined according to (46). Then, we have established that

$$\lim_{b\to\infty} F(p_b) = 4\sqrt{2}-5.$$

From Lemma 9, we know there exists $\gamma_b$ such that $\gamma_b < F(p_b) + \delta/2$, and where $\gamma_b$ is the ratio of Nash equilibrium aggregate surplus to maximal aggregate surplus for some game with price function $p_b$ and total allocated rate 1 at the Nash equilibrium. We thus have

$$\lim_{b\to\infty} \gamma_b = \lim_{b\to\infty} F(p_b) + \frac{\delta}{2} = 4\sqrt{2}-5+\frac{\delta}{2}.$$

Thus, for $b$ sufficiently large, we will have $\gamma_b < 4\sqrt{2}-5+\delta$, establishing (25). □

Theorem 8 shows that in the worst case, aggregate surplus falls by no more than approximately 34% when users are able to anticipate the effects of their actions on the price of the link. Furthermore, this bound is essentially tight. In fact, from the proof of the theorem we see that this ratio is achieved via a sequence of games where the following hold true.

1) The price function $p$ has the form given by (46)–(48), with $a = 2-\sqrt{2}$, $b \to \infty$, and $f = 1$.
2) The number of users becomes large ($R \to \infty$).
3) User 1 has linear utility with $U_1(d_1) = d_1$, and all users have linear utility with $U_r(d_r) = \alpha_r d_r$, where $\alpha_r \approx p(1) = 2-\sqrt{2}$ (for $r > 1$). The last item follows by substituting the solution (41) in (33), and taking the limit as $R \to \infty$. (Note that formally, the limits of $R \to \infty$ and $b \to \infty$ should be taken in the correct order; in particular, in the proof we first have $R \to \infty$, and then $b \to \infty$.)

Note that the price function $p$ used to achieve the worst case efficiency loss is not differentiable. As discussed in Section I, this is the main reason that we allow nondifferentiable price functions in Assumption 2. Indeed, some of the results of Section II-A can be simplified if we restrict attention only to differentiable price functions. Nevertheless, we note that even if we only consider differentiable price functions, the worst case efficiency loss remains approximately 34%. This result can be established by approximating the piecewise linear price functions described in (46)–(48) by differentiable price functions.

It is interesting to note that the worst case is obtained by considering instances where the price function is becoming steeper and steeper at the Nash equilibrium rate 1, since $b \to \infty$. This forces the optimal rate $f^S$ at the solution to SYSTEM to approach the Nash equilibrium rate $f = 1$, as we observe from (51); nevertheless, the shortfall between the Nash equilibrium aggregate surplus and the maximal aggregate surplus approaches 34%.

## IV. Inelastic Supply Versus Elastic Supply

In this section, we briefly compare the model of this paper (allocation of a resource in elastic supply) with the model of [23] (allocation of a resource in inelastic supply). In [23], a model is considered with a single link having exactly $D_{\max}$ units of rate available to allocate among the users. As in the model of this paper, user $r$ submits a bid $w_r$. The link manager then sets a price $\mu = \sum_r w_r/D_{\max}$; and user $r$ receives an allocation $d_r$ given by

$$d_r = \begin{cases} 0, & \text{if } w_r = 0 \\ \frac{w_r}{\mu}, & \text{if } w_r > 0. \end{cases}$$

As in this paper, the payoff to user $r$ is $U_r(d_r) - w_r$. It is shown in [23] that when users are price anticipating and the link supply is inelastic, the efficiency loss is at most 25% of the maximal aggregate utility.

Intuitively, we would like to model a system with an inelastic supply $D_{\max}$ by a cost function which is zero for $0 \leq f < D_{\max}$, and infinite for $f > D_{\max}$. Formally, we show in this section that if the price function is given by $p(f) = af^B$ for $a \geq 0$ and $B \geq 1$, then as $B \to \infty$ the worst case efficiency loss approaches 25%—the same value obtained in [23]. While this does not formally establish the result in [23], the limit is intuitively plausible, because as the exponent $B$ increases, the price function $p$ and associated cost function begin to resemble an inelastic capacity constraint with $D_{\max} = 1$: For $f < 1$, $f^B \to 0$ as $B \to \infty$; and for $f > 1$, $f^B \to \infty$ as $B \to \infty$.

*Theorem 10:* Suppose that Assumptions 1–3 hold. Suppose also that $U_r(0) \geq 0$ for all $r$, and that $p(f) = af^B$ for $a \geq 0$ and $B \geq 1$. Define the function $g(B)$ by

$$g(B) \triangleq \left(\frac{B+1}{2B+1}\right)^{\frac{1}{B}} \left(\frac{(B+1)(3B+2)}{(2B+1)^2}\right). \quad (54)$$

If $\mathbf{d}^S$ is any solution to $SYSTEM$, and $\mathbf{w}$ is any Nash equilibrium of the game defined by $(Q_1, \ldots, Q_R)$, then

$$\text{surplus}\,(\mathbf{d}(\mathbf{w})) \geq g(B) \cdot \text{surplus}(\mathbf{d}^S) \quad (55)$$

where surplus$(\cdot)$ is defined in (3). Furthermore, $g(B)$ is strictly increasing, with $g(B) \to 3/4$ as $B \to \infty$; and the bound (55) is tight: for fixed $B \geq 1$, for every $\delta > 0$, there exists a choice of $R$ and a choice of (linear) utility functions $U_r$, $r = 1, \ldots, R$, such that a Nash equilibrium $\mathbf{w}$ a solution $\mathbf{d}^S$ to SYSTEM exist with

$$\text{surplus}\,(\mathbf{d}(\mathbf{w})) \leq (g(B)+\delta) \cdot \text{surplus}(\mathbf{d}^S). \quad (56)$$

*Proof:* We follow the proof of Theorem 8. Steps 1)–4) follow as in that proof, provided we can show that two scalings of the function $p(\cdot)$ do not affect our result—in step 1), where we replace $p(\cdot)$ by $p(\cdot)/\max_r \alpha_r$, and in step 2), where we replace $p(\cdot)$ by $p(f\cdot)$, where $f$ is the Nash equilibrium rate. Indeed, both these scalings remain valid, since the rescaled price function is still a monomial with the same exponent as $p$, but a different constant coefficient. In particular, we may continue to restrict attention to the special case where $U_r(d_r) = \alpha_r d_r$, with $\max_r \alpha_r = \alpha_1 = 1$, and where the total Nash equilibrium allocated rate is 1.

From steps 1)–4) of the proof of Theorem 8, we must minimize $F(p)$, defined in (43), for all choices of $p$ such that (44) is satisfied, i.e., such that $1-\beta^+(1) \le p(1) < 1$. For $p(f) = af^B$, we have $\beta^+(f) = B/(1+B)$ and, thus, we require

$$\frac{1}{1+B} \le a < 1. \tag{57}$$

Note that at the maximal aggregate surplus, $p(f^S) = a(f^S)^B = 1$ implies that $f^S = a^{-1/B}$. Furthermore, $C(f) = af^{B+1}/(B+1)$ for $f \ge 0$. Thus, $f^S - C(f^S)$ is given by

$$f^S - C(f^S) = \left(\frac{B}{B+1}\right) \cdot \left(\frac{1}{a}\right)^{\frac{1}{B}}.$$

From (43), we conclude that $F(p)$ is given by

$$F(p) = \frac{a + (1-a)^2\left(1+\frac{1}{B}\right) - \frac{a}{(B+1)}}{\left(\frac{B}{B+1}\right)\left(\frac{1}{a}\right)^{\frac{1}{B}}}.$$

We now minimize $F(p)$ over the set of $a$ satisfying (57). We begin by differentiating $F(p)$ with respect to $a$, and setting the derivative to zero; simplifying, this yields the following equation:

$$Ba + \left(1+\frac{1}{B}\right)\left((2B+1)a^2 - 2(B+1)a + 1\right) = 0.$$

This equation is quadratic in $a$, and has two solutions $a_1$ and $a_2$: $a_1 = 1/(B+1)$, and $a_2 = (B+1)/(2B+1)$. Both solutions satisfy (57). Let $p_1(f) = a_1 f^B$, and $p_2(f) = a_2 f^B$. We have

$$F(p_1) = \left(\frac{1}{B+1}\right)^{\frac{1}{B}}\left(\frac{B+2}{B+1}\right)$$

$$F(p_2) = \left(\frac{B+1}{2B+1}\right)^{\frac{1}{B}}\left(\frac{(B+1)(3B+2)}{(2B+1)^2}\right).$$

To minimize $F(p)$ over $a$ satisfying (57), we need also to check the endpoint where $a = 1$. If $p = f^B$, we find $F(p) = 1$; since $F(p_1), F(p_2) \le 1$ from the definition of $F(p)$, the minimum value is achieved at either $p_1$ or $p_2$.

For $B \ge 1$, we define $g_1(B) = F(p_1)$, and $g_2(B) = F(p_2)$. We need the following technical lemma.

*Lemma 11:* The functions $g_1(B)$ and $g_2(B)$ are strictly increasing for $B \ge 1$. Furthermore, $g_1(B) \ge 3/4$ for $B \ge 1$, while $\lim_{B\to\infty} g_2(B) = 3/4$.

*Proof:* We begin by noting that $g_1(1) = 3/4$. Let $\hat{g}_1(B) = \ln(g_1(B))$; it suffices to show that $\hat{g}_1(B)$ is strictly increasing for $B \ge 1$. Differentiating $\hat{g}_1$ yields

$$\hat{g}_1'(B) = \frac{(B+2)\ln(B+1) - 2B}{(B+2)B^2}.$$

It suffices to check that $h_1(B) > 0$, where $h_1(B) = (B+2)\ln(B+1) - 2B$. We have $h_1(1) = 3\ln 2 - 2 > 0$; $h_1'(1) = \ln 2 - 1/2 > 0$; and $h_1''(B) = B/(B+1)^2 > 0$. This implies $h_1(B) > 0$ for all $B \ge 1$, so $g_1$ is strictly increasing for $B \ge 1$.

Next, we consider $g_2(B)$. Note first that $g_2(1) = 20/27$. Furthermore, as $B \to \infty$, $((B+1)/(2B+1))^{1/B} \to 1$, and $(B+1)(3B+2)/(2B+1)^2 \to 3/4$. Thus, $g_2(B) \to 3/4$ as $B \to \infty$.

Finally, let $\hat{g}_2(B) = \ln(g_2(B))$; it suffices to show $\hat{g}_2$ is strictly increasing for $B \ge 1$. Differentiating $\hat{g}_2(B)$ yields

$$\hat{g}_2'(B) = \frac{(3B+2)\ln\left(\frac{2B+1}{B+1}\right) - 2B}{(3B+2)B^2}.$$

As before, it suffices to check that $h_2(B) > 0$, where $h_2(B) = (3B+2)\ln((2B+1)/(B+1)) - 2B$. We have $h_2(1) = 5\ln(3/2) - 2 > 0$; $h_2'(1) = 3\ln(3/2) - 7/6 > 0$; and $h_2''(1) = B/[(B+1)^2(2B+1)^2] > 0$. Thus, $h_2(B) > 0$ for all $B \ge 1$, which implies $g_2$ is strictly increasing for $B \ge 1$. □

From the previous lemma, we conclude that the minimum value of $F(p)$ over $p = af^B$ satisfying (57) is given by $g_2(B)$; this establishes (55). As in Theorem 8, by construction this bound is tight, so (56) holds as well. □

The preceding theorem shows that for a particular sequence of price functions which approach an inelastic supply constraint, the efficiency loss gradually decreases from 7/27 (at $B = 1$) to 1/4 (as $B \to \infty$). In the limit as $B \to \infty$, we recover the same efficiency loss as in the earlier work of [23]. However, while we have demonstrated such a limit holds as long as the price functions are monomials, there remains an open question: If the price functions "converge" (in an appropriate sense) to a fixed capacity constraint, under what conditions does the efficiency loss also converge to 1/4? It is straightforward to check that such a limit cannot always hold. For example, consider price functions $p$ of the form specified in (46)–(48). Using the expression for $F(p)$ given in (50), it is possible to show that by first taking $b \to \infty$, and then taking $a \to 0$, the worst case efficiency loss approaches zero; see (53).

## V. Conclusion

This paper considers a pricing mechanism where the available resources in a network are in elastic supply. For a game where users' strategies are the payments they are willing to make, we showed that the efficiency loss is no more than 34% when users are price anticipating, for the setting of a single link (Theorem 8).

This result can be extended to general networks, using the same approach as in [23]. We consider a standard multicommodity flow model, where each user has multiple paths available, and each path uses a subset of links in the network. The utility to a user depends on the maximum rate at which he can send through the network. We study a game where users submit individual bids to each link in the network; it is straightforward to establish existence of a Nash equilibrium for such a game. Using techniques similar to the results proven in a network context in [23], it can be shown that the efficiency loss is no more than 34% when users are price anticipating, matching the result of Section III. For details of this network extension, the reader is referred to [24] and [29].

Important questions remain regarding an extension of this work to a dynamic context. While our results suggest that ma-

nipulation of the market in a static game setting cannot lead to arbitrarily high efficiency loss, such a result does not necessarily imply users will not be able to manipulate an algorithmic implementation of this mechanism (such as those proposed in [3]). Investigation of this point is an open research topic.

Critical to any investigation of dynamics is the nature of the *information* available to the players of the pricing game. In order to compute an optimal strategic decision users need to know not only the current price level $p(f(\mathbf{w}))$, but also the total allocated rate $f(\mathbf{w})$ and the derivative of the price $p'(f(\mathbf{w}))$ (where we have assumed for simplicity that $p$ is differentiable). We postulate that the overhead of actually collecting such detailed information in a large scale communication network is quite high; in fact, in general users do not have knowledge of either the total allocated rate or the derivative of the price at the resource. This raises an important question of information availability when users respond to price signals: users may not react optimally, so what are the users' conjectures about how their strategies affect the price? Developing more detailed models for the users' response to available price information from the network is a research direction for the future.

## References


[1] B. Briscoe, V. Darlagiannis, O. Heckman, H. Oliver, V. Siris, D. Songhurst, and B. Stiller, "A market managed multiservice Internet (M3I)," *Comput. Commun.*, vol. 26, no. 4, pp. 404–414, 2003.

[2] M. Falkner, M. Devetsikiotis, and I. Lambadaris, "An overview of pricing concepts for broadband IP networks," *IEEE Commun. Surv.*, vol. 3, no. 2, 2000.

[3] F. P. Kelly, A. K. Maulloo, and D. K. Tan, "Rate control for communication networks: shadow prices, proportional fairness, and stability," *J. Oper. Res. Soc.*, vol. 49, pp. 237–252, 1998.

[4] F. P. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, pp. 33–37, 1997.

[5] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*. Oxford, U.K.: Oxford Univ. Press, 1995.

[6] R. Johari and D. K. Tan, "End-to-end congestion control for the Internet: delays and stability," *IEEE/ACM Trans. Networking*, vol. 9, no. 6, pp. 818–832, Dec. 2001.

[7] F. Kelly, "Stability and fairness of end-to-end congestion control," *Eur. J. Control*, vol. 9, pp. 159–176, 2003.

[8] R. Srikant, *The Mathematics of Internet Congestion Control*. Boston, MA: Birkhäuser, 2004.

[9] G. Vinnicombe, "On the stability of networks operating TCP-like congestion control," in *Proc. IFAC World Congr.*, 2002.

[10] K. Ramakrishnan, S. Floyd, and D. Black, "The addition of explicit congestion notification (ECN) to IP," Internet Engineering Task Force, 2001.

[11] S. Athuraliya, S. H. Low, V. H. Li, and Q. Yin, "REM: active queue management," *IEEE Network*, vol. 15, no. 3, pp. 48–53, 2001.

[12] F. P. Kelly, "Models for a self managed Internet," *Phil. Trans.: Math., Phys., Eng. Sci.*, vol. 358, no. 1773, pp. 2335–2348, 2000.

[13] S. Kunniyur and R. Srikant, "Analysis and design of an adaptive virtual queue (AVQ) algorithm for active queue management," in *Proc. ACM SIGCOMM*, 2001, pp. 123–134.

[14] ——, "End-to-end congestion controls schemes: utility functions, random losses, and ECN marks," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 689–702, 2003.

[15] R. J. Gibbens and F. P. Kelly, "Resource pricing and the evolution of congestion control," *Automatica*, vol. 35, pp. 1969–1985, 1999.

[16] A. Czumaj and B. Voecking, "Tight bounds for worst-case equilibria," in *Proc. 13th Annu. ACM-SIAM Symp. Discrete Algorithms*, 2002, pp. 413–420.

[17] E. Koutsoupias and C. Papadimitriou, "Worst-case equilibria," in *Proc. 16th Annu. Symp. Theoretical Aspects of Computer Science*, 1999, pp. 404–413.

[18] M. Mavronicolas and P. Spirakis, "The price of selfish routing," in *Proc. 33rd Annu. ACM Symp. Theory of Computing*, 2001, pp. 510–519.

[19] J. R. Correa, A. S. Schulz, and N. Stier Moses, "Selfish routing in capacitated networks," *Math. Oper. Res.*, 2004.

[20] T. Roughgarden and E. Tardos, "How bad is selfish routing?," *J. ACM*, vol. 49, no. 2, pp. 236–259, 2002.

[21] E. Anshelevich, A. Dasgupta, E. Tardos, and T. Wexler, "Near-optimal network design with selfish agents," in *Proc. 35th Annu. ACM Symp. Theory of Computing*, 2003, pp. 511–520.

[22] A. Fabrikant, A. Luthra, E. Maneva, C. Papadimitriou, and S. Shenker, "On a network creation game," in *Proc. 22nd Annu. ACM Symp. Principles of Distributed Computing*, 2003, pp. 347–351.

[23] R. Johari and J. N. Tsitsiklis, "Efficiency loss in a network resource allocation game," *Math. Oper. Res.* , vol. 29, no. 3, pp. 407–435, 2004.

[24] R. Johari, S. Mannor, and J. N. Tsitsiklis. (2005) Efficiency loss in a network resource allocation game: The case of elastic supply. ArXiV. [Online]http://www.arxiv.org/abs/cs.GT/0 506 054

[25] S. Shenker, "Fundamental design issues for the future Internet," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1176–1188, Sep. 1995.

[26] J. Rosen, "Existence and uniqueness of equilibrium points for concave $n$-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

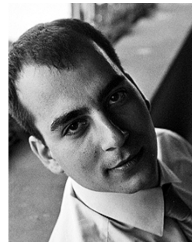
[27] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA: Athena Scientific, 2003.

[28] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton Univ. Press, 1970.

[29] R. Johari, "Efficiency loss in market mechanisms for resource allocation," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 2004.

**Ramesh Johari** (M'05) received the A.B. degree in mathematics from Harvard University, Cambridge, MA, the Certificate of Advanced Study in mathematics from the University of Cambridge, Cambridge, U.K., and the Ph.D. in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1998, 1999, and 2004, respectively.

He is currently an Assistant Professor of Management Science and Engineering, and by courtesy, Electrical Engineering, at Stanford University, Stanford, CA.



**Shie Mannor** (S'00–M'03) received the B.Sc. degree in electrical engineering, the B.A. degree in mathematics (both *summa cum laude*), and the Ph.D. degree in electrical engineering from the Technion—Israel Institute of Technology, Haifa, in 1996, 1996, and 2002, respectively.

During the spring semester of 2002, he was a Lecturer at the Electrical Engineering Department of the Technion . From 2002 to 2004, he was a Postdoctoral Associate at the Massachusetts Institute of Technology, Cambridge. He is currently an Assistant Professor of Electrical and Computer Engineering at McGill University, Montreal, QC, Canada. His research interests include machine learning and pattern recognition, planning and control, multiagent systems, and communications.

Dr. Mannor was a Fulbright Scholar in 2002, and he is currently a Canada Research Chair in Machine Learning.

**John N. Tsitsiklis** (F'99) received the B.S. degree in mathematics, and the B.S., M.S., and Ph.D. degrees in electrical engineering, all from the Massachusetts Institute of Technology (MIT), Cambridge, in 1980, 1980, 1981, and 1984, respectively.

He is currently a Professor of Electrical Engineering and Computer Science, and a Co-Director of the Operations Research Center at MIT. His research interests are in the fields of systems, optimization, communications, control, and operations research. He has coauthored four books and about 100 journal papers.

Dr. Tsitsiklis' awards include an Outstanding Paper Award by the IEEE Control Systems Society, the MIT Edgerton Faculty Achievement Award (1989), the Bodossakis Foundation Prize (1995), and the INFORMS/CSTS prize (1997). He is currently a member of the editorial board for the Springer-Verlag "Lecture Notes in Control and Information Sciences" series, and an Associate Editor of *Mathematics of Operations Research*.