

A Scalable Network Resource Allocation Mechanism With Bounded Efficiency Loss

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

Abstract—The design of pricing mechanisms for network resource allocation has two important objectives: 1) a simple and scalable end-to-end implementation and 2) efficiency of the resulting equilibria. Both objectives are met by certain recently proposed mechanisms when users are price takers, but not when users can anticipate the effects of their actions on the resulting prices. In this paper, we partially close this gap, by demonstrating an alternative resource allocation mechanism which is scalable and guarantees a fully efficient allocation when users are price taking. In addition, when links have affine marginal cost, this mechanism has efficiency loss bounded by 1/3 when users are price anticipating. These results are derived by studying *Cournot games*, and in the process we derive the first nontrivial constant factor bounds on efficiency loss in these well-studied economic models.

Index Terms—Computer networks, game theory, resource management.

I. INTRODUCTION

IN THIS PAPER, we consider a *congestion pricing* approach to resource allocation in communication networks; see, e.g., [1] for an overview. Congestion pricing mechanisms in networks must tradeoff two competing goals: simplicity/scalability of the mechanism, and efficiency. In this paper, we define a scalable mechanism as one where each user's strategic decision depends only on end-to-end path information, rather than detailed information on individual link states. Efficiency is defined in terms of the aggregate value of a network allocation. The abstraction we consider assumes that each user has a *utility function*, and each link has a *cost function*; an efficient allocation is one that maximizes *aggregate surplus* (aggregate utility less aggregate cost).

Several previous efforts have considered the design of scalable and efficient market mechanisms, with varying degrees of success depending on the assumptions made about the behavior of network users. Kelly and coworkers have shown that as long as users act as "price takers" (i.e., if users do not anticipate the effect of their strategic decisions on the prices), then a simple proportional bidding mechanism can achieve efficient allocations when link capacities are fixed [2] or elastic [3]. The mech-

anism asks each user to submit a single bid, which conveys the total payment of that user. In the case of a single link, resources are allocated to users in proportion to the bids submitted. The network case is a graceful generalization of this model, where each user is only required to choose a bid in response to the prices on each available *path*, rather than on each *link* of the network.

Since Kelly *et al.* assume that users are price takers, they ignore the possibility of strategic users that are fully rational and anticipate the effect of their decisions on the link prices. A game theoretic formulation of the mechanisms in [2] and [3] for the case of price anticipating users has been considered for a single link by Johari *et al.* [4], [5], and aggregate surplus at a Nash equilibrium was shown to be within a constant factor of the maximum possible aggregate surplus. However, the results in [4] and [5] exhibit an undesirable scaling behavior in the network context: in order to ensure a bounded efficiency loss, the authors required that each user submit *individual bids* to each link in the network, rather than a single bid per path as in the original proposal of [2], [3]. The ramifications of such a mechanism are that each user must be aware of each link he can potentially use, as well as the prices of those links. If, instead, the user submits only a single bid, Hajek and Yang [6] have shown that the efficiency loss can be arbitrarily high when users are price anticipating.

We close the gap in preceding results, by demonstrating that it is possible to design a network resource allocation mechanism that is scalable and has good efficiency properties when users are price taking, as well as price anticipating. We consider a simple market mechanism for link data rate allocation, where users choose the rates along each path they desire from the network, and optimize their payoff only based on the aggregate price of each path available to them (rather than individual link prices). When users are price taking, this mechanism achieves an efficient allocation. We also show that if every link has an affine marginal cost function the efficiency loss is no worse than 1/3 of the maximal aggregate surplus—regardless of the utility functions of the users. This result is of practical importance: while affine approximations to links' marginal cost are still quite rich, it is unlikely that detailed knowledge of users' utilities will be available. Furthermore, the assumption of affine marginal costs is no more restrictive than the affine latency assumption in the seminal work of Roughgarden and Tardos [7].

Mechanisms where market participants choose their desired quantities are known as *Cournot* models [8]. Cournot games are among the best studied economic models for competition between market participants; see, e.g., [9]–[11] for surveys of this rich topic. The results of this paper are among the first constant factor bounds available for Cournot competition models, and

Manuscript received February 15, 2005; revised January 15, 2006. This work was supported in part by the National Science Foundation under a Graduate Research Fellowship and under Grant ECS-0312921.

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

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

Digital Object Identifier 10.1109/JSAC.2006.872880

thus the analysis of this paper constitutes an important contribution to the existing results in economics as well. Additional bounds using similar techniques can be derived for other economic models of interest, in particular for monopoly pricing with general concave demand, and Cournot oligopoly pricing with affine demand; for details, see [12].

Our investigation forms part of a broader body of work that quantifies the efficiency loss in environments where participants may be selfish. Results have been obtained for routing [13], traffic networks [7], [14], and network design problems [15], [16]. We note in particular that Roughgarden and Tardos showed that when nonatomic users selfishly route through a network, and if delays are affine functions of flow on each link of the network, then the average delay is a factor of at most $4/3$ higher than the optimal average delay.

The outline of the remainder of this paper is as follows. In Section II, we describe the operation of the mechanism for a single link, and establish that when users are price taking the resulting allocation is efficient (a consequence of the first fundamental theorem of welfare economics [17]). We then show that at a Nash equilibrium of this game when users are price anticipating and the link has affine marginal cost, the efficiency loss is no more than $1/3$ of the maximal aggregate surplus, regardless of the utility functions of the users; we also show this bound is tight.

We then consider general networks in Section III, where each user chooses the rate at which they wish to send on each available path. It is again the case that price taking users end up maximizing the aggregate surplus. We then establish that as long as all links have affine marginal cost functions, the efficiency loss at a Nash equilibrium is no worse than a factor of $1/3$, matching the result of the single-link case.

II. A SINGLE LINK

In this section, we consider a game where multiple users compete for a single link, and where the strategies of the users represent their desired rates. Although it is easy to check that such games can yield arbitrarily high efficiency loss in general [12], we will establish a bound on efficiency loss in the important special case where the link has affine marginal cost.

Formally, we consider the following model. We assume that N users compete for a single link. We assume that each user n has a utility function U_n , and that the total data rate through the link incurs a cost characterized by a cost function C . We make the following assumptions.

Assumption 1: For each n , and over the domain $x_n \geq 0$, the utility function $U_n(x_n)$ is concave, nondecreasing, and continuously differentiable (where we interpret $U'_n(0)$ as the right directional derivative of U_n at 0).

Assumption 2: There exists a differentiable, convex, nondecreasing function $p(q)$ over $q \geq 0$, with $p(0) \geq 0$ and $p(q) \rightarrow \infty$ as $q \rightarrow \infty$, such that for $q \geq 0$

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

In particular, $C(q)$ is convex and nondecreasing.

We assume that both utility and cost are measured in monetary units, so that an efficient allocation is characterized as an optimal solution of the following optimization problem:

$$\text{maximize} \quad \sum_n U_n(x_n) - C\left(\sum_n x_n\right) \quad (1)$$

$$\text{subject to} \quad x_n \geq 0, \quad n = 1, \dots, N. \quad (2)$$

We refer to the objective function (1) as the *aggregate surplus* [17]. Since $p(q) \rightarrow \infty$ as $q \rightarrow \infty$, while U_n only grows at most linearly, it follows that an optimal solution exists.

We now consider the following pricing scheme for resource allocation. Each user n chooses a desired rate x_n . Given the vector $\mathbf{x} = (x_1, \dots, x_n)$, the link sets a single price $\mu(\mathbf{x}) = p(\sum_n x_n)$. User n then pays $x_n \mu(\mathbf{x})$. We first consider the case where, given a price $\mu > 0$, user n chooses x_n to maximize

$$P_n(x_n; \mu) = U_n(x_n) - \mu x_n. \quad (3)$$

Notice that in the previous expression, each user is acting as a *price taker*; that is, he does not anticipate the effect of a change in his strategy on the resulting price. Since we are using *marginal cost pricing*, i.e., since $\mu(\mathbf{x}) = p(\sum_n x_n)$, we expect that price taking users will maximize aggregate surplus at a competitive equilibrium. This is formalized in the following proposition, a special case of the first fundamental theorem of welfare economics [17]. This result is similar to that proven by Kelly *et al.* in [3] for a proportionally fair pricing scheme, where users choose the total amount they are willing to pay, rather than their total rate.

Proposition 1: Suppose Assumptions 1 and 2 hold. There exists a *competitive equilibrium*, that is, a vector \mathbf{x} and a scalar μ such that $\mu = p(\sum_n x_n)$, and

$$P_n(x_n; \mu) = \max_{\bar{x}_n \geq 0} P_n(\bar{x}_n; \mu), \quad n = 1, \dots, N. \quad (4)$$

Any such vector \mathbf{x} solves (1) and (2). If the functions U_n are strictly concave, such a vector \mathbf{x} is unique.

Proposition 1 shows that with marginal cost pricing, and if the users of the link behave as price takers, there exists a vector of rates \mathbf{x} , where all users have optimally chosen their x_n , with respect to the given price $\mu = p(\sum_n x_n)$; and at this “equilibrium,” the aggregate surplus is maximized.

When the price taking assumption is violated, however, the model changes into a game and the guarantee of Proposition 1 is no longer valid. Consider, then, an alternative model where the users of a single link are price anticipating, rather than price taking, and play a *Cournot game* to acquire a share of the link. We use the notation \mathbf{x}_{-n} to denote the vector of all rates chosen by users other than n ; i.e., $\mathbf{x}_{-n} = (x_1, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_N)$. Then, given \mathbf{x}_{-n} , each user n chooses $x_n \geq 0$ to maximize

$$Q_n(x_n; \mathbf{x}_{-n}) = U_n(x_n) - x_n p\left(\sum_m x_m\right). \quad (5)$$

The payoff function Q_n is similar to the payoff function P_n , except that the user now *anticipates* that the price will be set according to $p(\sum_m x_m)$. A *Nash equilibrium* of the game defined by (Q_1, \dots, Q_N) is a vector $\mathbf{x} \geq 0$ such that for all n

$$Q_n(x_n; \mathbf{x}_{-n}) \geq Q_n(\bar{x}_n; \mathbf{x}_{-n}), \quad \text{for all } \bar{x}_n \geq 0. \quad (6)$$

We first show that a Nash equilibrium exists for this game. The proof is a standard application of Rosen's existence theorem [18], and is omitted; details may be found in [12] and [19].

Proposition 2: Suppose that Assumptions 1 and 2 hold. Then there exists a Nash equilibrium \mathbf{x} for the game defined by (Q_1, \dots, Q_N) .

Because the payoff Q_n is concave in x_n for fixed \mathbf{x}_{-n} , a vector \mathbf{x} is a Nash equilibrium if and only if the following first-order conditions are satisfied for each n , where $q = \sum_m x_m$:

$$U'_n(x_n) = p(q) + x_n p'(q), \quad \text{if } x_n > 0; \quad (7)$$

$$U'_n(0) \leq p(q), \quad \text{if } x_n = 0. \quad (8)$$

We will use these conditions to investigate the efficiency loss when users are price anticipating. Specifically, we are interested in comparing the aggregate surplus achieved at a Nash equilibrium to the aggregate surplus achieved at a competitive equilibrium.

It is not difficult to construct examples in which the efficiency loss is arbitrarily high, under the general assumptions we have made in this section; one such example is presented in [12]. Indeed, the knowledge that Cournot oligopoly models can yield high-efficiency loss is a generally accepted piece of folklore in the economics community; see, e.g., [20]. However, in the next theorem, we consider the important special case where marginal cost is affine. Even though the utility functions of the users may be arbitrary, we show that the efficiency loss of the resulting mechanism is no worse than 1/3 when users are price anticipating.

Theorem 3: Suppose that Assumption 1 holds, and that $p(q) = aq + b$ for some $a > 0$, $b \geq 0$. Suppose also that $U_n(0) \geq 0$ for all n . If \mathbf{x}^S is any solution to (1), (2), and \mathbf{x} is any Nash equilibrium of the game defined by (Q_1, \dots, Q_N) , then

$$\sum_n U_n(x_n) - C\left(\sum_n x_n\right) \geq \frac{2}{3} \left(\sum_n U_n(x_n^S) - C\left(\sum_n x_n^S\right) \right). \quad (9)$$

Furthermore, this bound is tight: for every $a > 0$, $b \geq 0$, and $\delta > 0$, there exists a choice of N and a choice of (linear) utility functions U_n , $n = 1, \dots, N$, such that a Nash equilibrium \mathbf{x} exists with

$$\sum_n U_n(x_n) - C\left(\sum_n x_n\right) \leq \left(\frac{2}{3} + \delta\right) \left(\sum_n U_n(x_n^S) - C\left(\sum_n x_n^S\right) \right). \quad (10)$$

Proof: The proof follows in three steps. We first show that we can assume without loss of generality that $\sum_n U_n(x_n^S) - C(\sum_n x_n^S) > 0$, and $\sum_n U'_n(x_n)x_n - C(\sum_n x_n) > 0$. Next, we show that the worst case occurs when the utility functions of the users are linear (Lemma 4). We then optimize over all games with linear utility functions to determine the worst case efficiency loss.

Define $q = \sum_n x_n$, and $q^S = \sum_n x_n^S$. If $q = 0$, then $x_n = 0$ for all n ; and by (8), we have $U'_n(0) \leq p(0)$ for all n . But this is a sufficient condition for optimality for (1) and (2), so we conclude \mathbf{x} is an optimal solution to (1) and (2). Thus if $q = 0$, the bound trivially holds.

From now on we assume, without loss of generality, that $q > 0$. Now, for any user n with $x_n > 0$, from (7) together with the fact that $p'(q) = a > 0$, we conclude that $U'_n(x_n) > p(q)$. Thus, $\sum_n U'_n(x_n)x_n > qp(q) \geq C(q)$, so that $\sum_n U'_n(x_n)x_n - C(q) > 0$. Furthermore, since U_n is concave, nondecreasing, and nonnegative, we have $U'_n(x_n)x_n \leq U_n(x_n)$. Thus, $\sum_n U_n(x_n) - C(q) > 0$; since \mathbf{x}^S is an optimal solution to (1), (2), we have $\sum_n U_n(x_n^S) - C(q^S) \geq \sum_n U'_n(x_n)x_n - C(q) > 0$.

We will use the next lemma to establish that linear utility functions yield the worst efficiency loss. The proof uses the concavity of the utility functions, and is closely related to the proof of Lemma 4 in [4]. For this reason, we omit the proof; details can be found in [12].

Lemma 4: Suppose that Assumptions 1 and 2 hold. Suppose also that $U_n(0) \geq 0$ for all n . Fix any rate vector $\mathbf{x} \geq 0$, and let \mathbf{x}^S be any optimal solution to (1), (2). Define $\alpha_n = U'_n(x_n)$. If $\sum_n \alpha_n x_n - C(\sum_n x_n) > 0$ and $\sum_n U_n(x_n^S) - C(\sum_n x_n^S) > 0$, then the following inequality holds:

$$\frac{\sum_n U_n(x_n) - C(\sum_n x_n)}{\sum_n U_n(x_n^S) - C(\sum_n x_n^S)} \geq \frac{\sum_n \alpha_n x_n - C(\sum_n x_n)}{\max_{\bar{q} \geq 0} [(\max_n \alpha_n)\bar{q} - C(\bar{q})]}. \quad (11)$$

If we replace the utility function U_n by a new utility function \hat{U}_n for each n , where $\hat{U}_n(\hat{x}_n) = (U'_n(x_n))\hat{x}_n$, then \mathbf{x} continues to be a Nash equilibrium, since the optimality conditions (7) and (8) still hold. Applying Lemma 4, therefore, we see that the ratio of Nash equilibrium aggregate surplus to the maximal aggregate surplus cannot increase if we replace U_n by \hat{U}_n for all n .

Thus, we assume without loss of generality that the utility functions of all users are linear, i.e., $U_n(x_n) = \alpha_n x_n$. Since we have assumed $\sum_n \alpha_n x_n - C(\sum_n x_n) > 0$, we know that $\alpha_n > 0$ for at least one n . Thus, by replacing α_n by $\alpha_n / (\max_m \alpha_m)$, and $C(\cdot)$ by $C(\cdot) / (\max_n \alpha_n)$, we can also assume without loss of generality that $\max_n \alpha_n = 1$. Furthermore, by relabeling if necessary, we can assume that $\alpha_1 = 1$. Note that after rescaling, the new price function p is still affine but may have a different slope.

Since we have restricted attention to settings where $\sum_n \alpha_n x_n - C(\sum_n x_n) > 0$, we must also have $\sum_n x_n > 0$. Thus, from (7) and the fact that $\max_n \alpha_n = 1$, we must have $1 > p(q) = aq + b$; in particular, this implies that $b < 1$.

We start by computing the maximal aggregate surplus under these assumptions. Since the price function is $p(q) = aq + b$, the maximal aggregate surplus is achieved when $p(q^S) = 1$, i.e., when $q^S = (1 - b)/a$; this rate is entirely allocated to user 1. The maximal aggregate surplus is thus

$$\frac{1 - b}{a} - \frac{(1 - b)^2}{2a} - \frac{b(1 - b)}{a} = \frac{(1 - b)^2}{2a}.$$

Since the maximal aggregate surplus is fixed as $(1 - b)^2 / (2a)$, by (7) and (8) the worst case game is identified by solving the following optimization problem (with unknowns $x_1, \dots, x_n, \alpha_1, \dots, \alpha_n, q$):

$$\text{minimize} \quad \sum_{n=1}^N \alpha_n x_n - C(q) \quad (12)$$

$$\text{subject to} \quad \alpha_n = p(q) + x_n p'(q), \quad \text{if } x_n > 0, \\ n = 1, \dots, N \quad (13)$$

$$\alpha_n \leq p(q), \text{ if } x_n = 0, \quad n = 1, \dots, N \quad (14)$$

$$\sum_{n=1}^N x_n = q > 0 \quad (15)$$

$$\alpha_1 = 1; 0 \leq \alpha_n \leq 1, \quad n = 2, \dots, N \quad (16)$$

$$x_n \geq 0, \quad n = 1, \dots, N. \quad (17)$$

The objective function is the aggregate surplus associated with a Nash equilibrium allocation \mathbf{x} . The conditions (13) and (14) are equivalent to the Nash equilibrium conditions established in (7) and (8). The constraint (15) ensures that the total allocation made at the Nash equilibrium is equal to q . The constraints in (16) follow since we have restricted without loss of generality to games where $\alpha_1 = \max_n \alpha_n = 1$. The constraint (17) ensures the rate allocated to each user is nonnegative.

We start by assuming that $q > 0$ is fixed, and optimize only over \mathbf{x} and $\boldsymbol{\alpha}$. In this case, we start by noting that we may assume without loss of generality that $\alpha_n = p(q) + x_n p'(q)$ for all users $n = 2, \dots, N$. Indeed, if $(\boldsymbol{\alpha}, \mathbf{x})$ is a feasible solution and $x_n > 0$ for some $n = 2, \dots, N$, then (13) and (14) imply that $\alpha_n = p(q) + x_n p'(q)$. On the other hand, if $x_n = 0$ for some $n = 2, \dots, N$, we can set $\alpha_n = p(q) = aq + b$; this preserves feasibility, but does not impact the term $\alpha_n x_n$ in the objective function (12). We can, therefore, restrict attention to feasible solutions for which

$$\alpha_n = p(q) + x_n p'(q) = aq + b + ax_n, \quad n = 2, \dots, N. \quad (18)$$

Having done so, observe that the constraint (16), that $\alpha_n \leq 1$, may be written as

$$x_n \leq \frac{1 - aq - b}{a}, \quad n = 2, \dots, N.$$

Finally, the constraint (16) that $\alpha_n \geq 0$ becomes redundant, as it is guaranteed by the fact that $a > 0$, $b \geq 0$, and $q > 0$.

It follows from (16) together with (13) that a candidate solution satisfying (15) can only exist if $x_1 > 0$, in which case we have $1 = p(q) + x_1 p'(q) = aq + b + ax_1$, so that $x_1 = (1 - aq - b)/a$. In particular, we conclude immediately that for a feasible solution to exist, we must have $0 < (1 - aq - b)/a \leq q$. This yields the following reduced optimization problem:

$$\text{minimize } \frac{1 - aq - b}{a} + \sum_{n=2}^N (aq + b + ax_n)x_n - C(q) \quad (19)$$

$$\text{subject to } \sum_{n=2}^N x_n = q - \frac{1 - aq - b}{a} \quad (20)$$

$$x_n \leq \frac{1 - aq - b}{a}, \quad n = 2, \dots, N \quad (21)$$

$$x_n \geq 0, \quad n = 2, \dots, N. \quad (22)$$

The objective function (19) is equivalent to (12) upon substitution for α_n [from (13)] and x_1 [also from (13)]. The constraint (20) is equivalent to the allocation constraint (15); and the constraint (21) ensures $\alpha_n \leq 1$, as required in (16).

For fixed $q > 0$, the resulting problem is symmetric in the rates x_n for $n = 2, \dots, N$. It is clear that a feasible solution exists if and only if

$$\frac{q}{N} \leq \frac{1 - aq - b}{a} \leq q. \quad (23)$$

In this case, the following symmetric solution is feasible:

$$x_n = \frac{q - \frac{(1 - aq - b)}{a}}{N - 1}.$$

Furthermore, since the objective function (19) is strictly convex, this symmetric solution must in fact be optimal. If we substitute in the objective function (19), the resulting optimal value is strictly decreasing as N increases; the worst case occurs as $N \rightarrow \infty$, and the optimal objective value (19) becomes

$$\begin{aligned} & \frac{1 - aq - b}{a} + (aq + b) \left(q - \frac{1 - aq - b}{a} \right) - C(q) \\ & = \frac{1 - b}{a} - q + (aq + b) \left(2q - \frac{1 - b}{a} \right) - \frac{aq^2}{2} - bq. \end{aligned} \quad (24)$$

Furthermore, the feasibility requirements (23) on a , b , and q become $0 < (1 - aq - b)/a \leq q$, or upon rearranging, $(1 - b)/2 \leq aq < 1 - b$.

Until now, we have kept the price function and the total rate q fixed, and found the worst case game. We now optimize over all possible choices of price function p (i.e., over $a > 0$ and $b \geq 0$), as well as over possible Nash equilibrium rates (i.e., over $q > 0$). Recall that the maximal aggregate surplus is $(1 - b)^2/(2a)$. Thus, the worst case ratio is identified by the following optimization problem over q , a , and b :

$$\begin{aligned} & \text{minimize } \frac{2a}{(1 - b)^2} \left(\frac{1 - b}{a} - q + (aq + b) \right) \\ & \quad \times \left(2q - \frac{1 - b}{a} \right) - \frac{aq^2}{2} - bq \\ & \text{subject to } \frac{1 - b}{2} \leq aq \leq 1 - b, \quad a > 0, b \geq 0, q > 0. \end{aligned}$$

If we let $\bar{a} = aq$, then this problem becomes equivalent to the following problem:

$$\begin{aligned} & \text{minimize } \frac{2\bar{a}}{(1 - b)^2} \\ & \quad \times \left(\frac{1 - b}{\bar{a}} - 1 + (\bar{a} + b) \left(2 - \frac{1 - b}{\bar{a}} \right) - \frac{\bar{a}}{2} - b \right) \\ & \text{subject to } \frac{1 - b}{2} \leq \bar{a} \leq 1 - b, \quad \bar{a} > 0, b \geq 0. \end{aligned}$$

By substituting $x = \bar{a}/(1 - b)$ and differentiating, it is straightforward to establish that the minimum value of this optimization problem occurs at any pair \bar{a} and b satisfying the constraints, such that $\bar{a}/(1 - b) = 2/3$. One such pair is given by $\bar{a} = 2/3$, and $b = 0$. At any such solution, the minimum objective value is equal to $2/3$. This establishes (9).

We now show (10), for a fixed price function $p(q) = aq + b$ with $a > 0$ and $b \geq 0$. We choose the utility functions so that

$$U_1(x_1) = \left(\frac{3a}{2} + b \right) x_1;$$

$$U_n(x_n) = \left(a + b + \frac{a}{2(N - 1)} \right) x_n, \quad n = 2, \dots, N.$$

Let $\bar{x} = 1/(2(N - 1))$. It is then straightforward to check that for sufficiently large N , if $x_1 = 1/2$ and $x_n = \bar{x}$ for $n = 2, \dots, N$, the allocation \mathbf{x} is a Nash equilibrium. Furthermore, the maximum aggregate surplus is achieved by choosing q^S so that $3a/2 + b = p(q^S) = aq^S + b$, so $q^S = 3/2$, $x_1^S = q^S =$

$3/2$, and $x_n^S = 0$ for $n = 2, \dots, N$. Thus, the ratio of Nash equilibrium aggregate surplus to maximal aggregate surplus is

$$\frac{\left(\frac{3a}{2} + b\right) \left(\frac{1}{2}\right) + (a + b + a\bar{x}) \left(\frac{1}{2}\right) - \frac{a}{2} - b}{\left(\frac{3a}{2} + b\right) \left(\frac{3}{2}\right) - \left(\frac{a}{2}\right) \left(\frac{3}{2}\right)^2 - b \left(\frac{3}{2}\right)} = \frac{\frac{3}{2} + \bar{x}}{\frac{9}{4}}. \quad (25)$$

Now, as $N \rightarrow \infty$, this ratio approaches $2/3$, as required. \square

We note that by an appropriate choice of utility functions the worst case efficiency loss is always *exactly* $1/3$ for *any* affine price function.

III. GENERAL NETWORKS

In this section, we consider an extension of the single-link model to general networks. Our key contribution is to show that the generalization to a network model is a natural, scalable market mechanism: one where users choose rates over each *path* available, rather than at each *link* in the network. We show that the efficiency loss remains bounded, when marginal cost functions are affine; this is a counterpoint to the previous results on efficiency loss for the proportionally fair network resource allocation mechanism, where efficiency loss is only bounded if users submit *individual* bids to each link of the network [4], [6].

We consider a network consisting of J links, or *resources*, numbered $1, \dots, J$. As before, a set of users numbered $1, \dots, N$, shares this network of resources. We assume a set of paths through the network, numbered $1, \dots, P$. By an abuse of notation, we will use J , N , and P to also denote the *sets* of resources, users, and paths, respectively. Each path $q \in P$ uses a subset of the set of resources J ; if resource j is used by path q , we will denote this by writing $j \in q$. Each user $n \in N$ has a collection of paths available through the network; if path q serves user n , we will denote this by writing $q \in n$. We will assume without loss of generality that paths are uniquely identified with users, so that for each path q there exists a unique user n such that $q \in n$. (There is no loss of generality because if two users share the same path, that is captured in our model by creating two paths which use exactly the same subset of resources.) For notational convenience, we note that the resources required by individual paths are captured by the *path-resource incidence matrix* \mathbf{A} , defined by $A_{jq} = 1$ if $j \in q$, and $A_{jq} = 0$, otherwise. Furthermore, we can capture the relationship between paths and users by the *path-user incidence matrix* \mathbf{H} , defined by $H_{nq} = 1$ if $q \in n$, and $H_{nq} = 0$, otherwise. Note that by our assumption on paths, for each path q we have $H_{nq} = 1$ for exactly one user n .

Let $y_q \geq 0$ denote the rate allocated to path q , and let $d_n = \sum_{q \in n} y_q \geq 0$ denote the rate allocated to user n ; using the matrix \mathbf{H} , we may write the relation between $\mathbf{d} = (d_n, n \in N)$ and $\mathbf{y} = (y_q, q \in P)$ as $\mathbf{Hy} = \mathbf{d}$. Furthermore, if we let f_j denote the total rate on link j , we must have

$$\sum_{q:j \in q} y_q = f_j, \quad j \in J.$$

Using the matrix \mathbf{A} , we may write this constraint as $\mathbf{Ay} = \mathbf{f}$.

We continue to assume that user n receives a utility $U_n(d_n)$ from an allocated rate d_n , and that each link j incurs a cost $C_j(f_j)$ when the total allocated rate at link j is f_j . We will assume that the utility functions satisfy Assumption 1, and each

cost function C_j satisfies Assumption 2; we let p_j denote the marginal cost function associated with link j .

The natural generalization of the problem (1) and (2) to a network context is given by the following optimization problem:

$$\text{maximize} \quad \sum_n U_n(d_n) - \sum_j C_j(f_j) \quad (26)$$

$$\text{subject to} \quad \mathbf{Ay} = \mathbf{f} \quad (27)$$

$$\mathbf{Hy} = \mathbf{d} \quad (28)$$

$$y_q \geq 0, \quad q \in P. \quad (29)$$

We continue to refer to the objective function (26) as the *aggregate surplus*. Since the objective function is continuous and U_n grows at most linearly while C_j grows superlinearly, an optimal solution \mathbf{y} exists. Since the feasible region is convex and the cost functions C_j are each strictly convex, the optimal vector $\mathbf{f} = \mathbf{Ay}$ is uniquely defined (though \mathbf{y} need not be unique). In addition, if the functions U_n are strictly concave, then the optimal vector $\mathbf{d} = \mathbf{Hy}$ is uniquely defined as well. As in the previous development, we will use the optimal solution to (26)–(29) as a benchmark for the outcome of the network game.

We will consider the following network resource allocation mechanism, a natural generalization of the game considered for a single link in the previous section. Each user n chooses a rate y_q for each path $q \in n$; thus the strategy of user n is now a vector $\mathbf{y}_n = (y_q, q \in n)$. The total rate demanded at link j is then $\sum_{q:j \in q} y_q$. We continue to assume that each link chooses a price equal to marginal cost, so that given the composite strategy vector $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_N)$, the price of link j is $\mu_j(\mathbf{y}) = p_j(\sum_{q:j \in q} y_q)$. Each user n then pays a total amount $y_q \sum_{j \in q} \mu_j(\mathbf{y})$ along each path $q \in n$; thus the total payment by user n is $\sum_{q \in n} y_q \sum_{j \in q} \mu_j(\mathbf{y})$. Notice that each user is only choosing rates along each available *path* through the network, rather than at each *link* in the network. This is a much more scalable approach to network resource allocation, because users do not need to be aware of the network topology in making their rate decisions; each user only needs to know the aggregate price of each path available to him.

As in the previous development, if we assume that each user behaves as a price taker, then given a vector of prices $\boldsymbol{\mu} = (\mu_j, j \in J)$, the payoff to user n is

$$P_n(\mathbf{y}_n; \boldsymbol{\mu}) = U_n \left(\sum_{q \in n} y_q \right) - \sum_{q \in n} y_q \sum_{j \in q} \mu_j.$$

Since we are using marginal cost pricing, we again expect price taking users to maximize aggregate surplus at a competitive equilibrium; this is formalized in the following analogue of Proposition 1, which is again a special case of the first fundamental theorem of welfare economics [17].

Proposition 5: Suppose that Assumption 1 holds, and that for each price function p_j and cost function C_j , Assumption 2 holds. There exists a *competitive equilibrium*, that is, a pair of vectors \mathbf{y} and $\boldsymbol{\mu}$ such that $\mu_j = p_j(\sum_{q:j \in q} y_q)$, and

$$P_n(\mathbf{y}_n; \boldsymbol{\mu}) = \max_{\bar{\mathbf{y}}_n \geq 0} P_n(\bar{\mathbf{y}}_n; \boldsymbol{\mu}), \quad n = 1, \dots, N. \quad (30)$$

Furthermore, any such vector \mathbf{y} solves (26)–(29).

However, if each user is price anticipating, rather than price taking, the users may not maximize aggregate surplus. Consider, then, an alternative model where the users are price anticipating, and play a Cournot game to acquire a share of the links of the network. We use the notation \mathbf{y}_{-n} to denote the vector of all rates chosen by users other than n ; i.e., $\mathbf{y}_{-n} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n-1}, \mathbf{y}_{n+1}, \dots, \mathbf{y}_N)$. Then, given \mathbf{y}_{-n} , each user n chooses $\mathbf{y}_n \geq 0$ to maximize

$$Q_n(\mathbf{y}_n; \mathbf{y}_{-n}) = U_n \left(\sum_{q \in n} y_q \right) - \sum_{q \in n} y_q \sum_{j \in q} p_j \left(\sum_{\bar{q}: j \in \bar{q}} y_{\bar{q}} \right). \quad (31)$$

The payoff function Q_n is similar to the payoff function P_n , except that the user now *anticipates* that the price at link j will be set according to $p_j(\sum_{q: j \in q} y_q)$. A Nash equilibrium of the game defined by (Q_1, \dots, Q_N) is a composite vector $\mathbf{y} \geq 0$ such that for all n and all $\bar{\mathbf{y}}_n$, there holds $Q_n(\mathbf{y}_n; \mathbf{y}_{-n}) \geq Q_n(\bar{\mathbf{y}}_n; \mathbf{y}_{-n})$.

As in Proposition 2, it is straightforward to show that a Nash equilibrium exists for this game.

Proposition 6: Suppose that Assumption 1 holds, and that for each price function p_j and cost function C_j Assumption 2 holds. Then there exists a Nash equilibrium \mathbf{y} for the game defined by (Q_1, \dots, Q_N) .

We would now like to investigate the efficiency loss by comparing the aggregate surplus at a Nash equilibrium to the optimal value of (26)–(29). Of course, since a single link is a special case of a general network, the efficiency loss can be arbitrarily high at a Nash equilibrium. However, it is possible to establish a bound on efficiency loss in the special case where all link price functions are affine. The key technique that we use is to show a relationship between the Nash equilibria of the game defined by (Q_1, \dots, Q_N) , and the Nash equilibria of a game where users choose rates independently *at each link*. Once this equivalence is established, we can analyze the latter game using methods similar to those adopted in the network model of [4] and [5]: we reduce the analysis to individual games at each link, and then apply Theorem 3.

Theorem 7: Suppose that Assumption 1 holds, and that for each $j \in J$, $p_j(q_j) = a_j q_j + b_j$ for some $a_j > 0$, $b_j \geq 0$. Suppose also that $U_n(0) \geq 0$ for all n . If \mathbf{y}^S is any solution to (26)–(29), and \mathbf{y} is any Nash equilibrium of the game defined by (Q_1, \dots, Q_N) , then

$$\begin{aligned} & \sum_n U_n \left(\sum_{q \in n} y_q \right) - \sum_j C_j \left(\sum_{q: j \in q} y_q \right) \\ & \geq \frac{2}{3} \left(\sum_n U_n \left(\sum_{q \in n} y_q^S \right) - \sum_j C_j \left(\sum_{q: j \in q} y_q^S \right) \right). \quad (32) \end{aligned}$$

Proof outline: First, we define a new game where each user n chooses a rate d_{jn} demanded *at each link*; the strategy of user n is thus $\mathbf{d}_n = (d_{jn}, j \in J)$. This strategy determines the rate allocation to user n at each link in the network; given this allocation, user n sends at the maximum rate possible using

the paths $q \in n$ which are available. This maximum rate is the optimal value of the following *max-flow* optimization problem:

$$\text{maximize} \quad \sum_{q \in n} y_q \quad (33)$$

$$\text{subject to} \quad \sum_{q \in n: j \in q} y_q \leq d_{jn}, \quad j \in J; \quad (34)$$

$$y_q \geq 0, \quad q \in n. \quad (35)$$

We denote the optimal objective value of this optimization problem by $z_n(\mathbf{d}_n)$, where $\mathbf{d}_n = (d_{jn}, j \in J)$. Finally, the price at each link j is set to $p_j(\sum_n d_{jn})$, and the total payment made by user n is $\sum_j d_{jn} p_j(\sum_n d_{jn})$. Thus, given a composite strategy vector $\mathbf{d} = (\mathbf{d}_1, \dots, \mathbf{d}_N)$, the payoff to user n is

$$T_n(\mathbf{d}_n; \mathbf{d}_{-n}) = U_n(z_n(\mathbf{d}_n)) - \sum_j d_{jn} p_j \left(\sum_m d_{jm} \right).$$

The key property of the game defined by (T_1, \dots, T_N) is that any Nash equilibrium \mathbf{y} of the game defined by (Q_1, \dots, Q_N) is related by a natural transformation to a Nash equilibrium \mathbf{d} of the game defined by (T_1, \dots, T_N) ; and further, the aggregate surplus remains unchanged under this transformation. This is formalized in the following lemma.

Lemma 8: Suppose that Assumption 1 holds, and that each price function p_j and cost function C_j satisfies Assumption 2. Let \mathbf{y} be a Nash equilibrium of the game defined by (Q_1, \dots, Q_N) , and define $d_{jn} = \sum_{q \in n: j \in q} y_q$. Then, \mathbf{d} is a Nash equilibrium of the game defined by (T_1, \dots, T_N) . Furthermore, there holds

$$\begin{aligned} & \sum_n U_n \left(\sum_{q \in n} y_q \right) - \sum_j C_j \left(\sum_{q: j \in q} y_q \right) \\ & = \sum_n U_n(z_n(\mathbf{d}_n)) - \sum_j C_j \left(\sum_n d_{jn} \right). \quad (36) \end{aligned}$$

Proof of Lemma: Suppose that the vector \mathbf{y} is a Nash equilibrium of the game defined by (Q_1, \dots, Q_N) , and define $d_{jn} = \sum_{q \in n: j \in q} y_q$. By definition it follows that:

$$\sum_{q \in n} y_q \sum_{j \in q} p_j \left(\sum_{\bar{q}: j \in \bar{q}} y_{\bar{q}} \right) = \sum_j d_{jn} p_j \left(\sum_s d_{js} \right). \quad (37)$$

We now claim that $U_n(\sum_{q \in n} y_q) = U_n(z_n(\mathbf{d}_n))$. Clearly, \mathbf{y}_n is feasible for (33)–(35), so $\sum_{q \in n} y_q \leq z_n(\mathbf{d}_n)$; and thus $U_n(\sum_{q \in n} y_q) \leq U_n(z_n(\mathbf{d}_n))$. Suppose then that $U_n(\sum_{q \in n} y_q) < U_n(z_n(\mathbf{d}_n))$; this is only possible if $\sum_{q \in n} y_q < z_n(\mathbf{d}_n)$. But in this case if $\bar{\mathbf{y}}_n$ is an optimal solution to (33)–(35) with rate allocation \mathbf{d}_n , then $\bar{\mathbf{y}}_n$ is a profitable deviation for user n in the game defined by (Q_1, \dots, Q_N) : the utility to user n strictly increases, while the payment made by user n does not increase. This is not possible, since \mathbf{y} is a Nash equilibrium of the game defined by (Q_1, \dots, Q_N) . We conclude that $U_n(\sum_{q \in n} y_q) = U_n(z_n(\mathbf{d}_n))$, and thus we have $T_n(\mathbf{d}_n; \mathbf{d}_{-n}) = Q_n(\mathbf{y}_n; \mathbf{y}_{-n})$ for all n .

Now, suppose that \mathbf{d} is not a Nash equilibrium of the game defined by (T_1, \dots, T_N) ; then there exists a user n such that $\bar{\mathbf{d}}_n \neq \mathbf{d}_n$ is a profitable deviation for user n . It is then straightforward to show that if $\bar{\mathbf{y}}_n$ is an optimal solution to (33)–(35), then $\bar{\mathbf{y}}_n$ is a profitable deviation for user n from \mathbf{y}_n in the game defined by (Q_1, \dots, Q_N) , i.e., $Q_n(\bar{\mathbf{y}}_n; \mathbf{y}_{-n}) > Q_n(\mathbf{y}_n; \mathbf{y}_{-n})$. Thus, \mathbf{y} could not have been a Nash equilibrium for the game defined by (Q_1, \dots, Q_N) , a contradiction. So, we conclude \mathbf{d} is a Nash equilibrium for the game defined by (T_1, \dots, T_N) .

Finally, since we have shown that $U_n(\sum_{q \in n} y_q) = U_n(z_n(\mathbf{d}_n))$ for all n , and we have $\sum_{q: j \in q} y_q = \sum_n d_{jn}$ for each j by definition, we conclude that (36) holds. \square

Given Lemma 8, it suffices to focus on the worst case efficiency loss in the game defined by (T_1, \dots, T_N) . This game can be analyzed using techniques similar to those used in [4, Th. 7] and [5, Th. 14], so we only sketch the remainder of the proof here. If \mathbf{d}_n^* is a Nash equilibrium of the game defined by (T_1, \dots, T_N) , we replace the “composite” utility function $U_n(z_n(\mathbf{d}_n))$ with a linear utility function $\alpha_n^\top \mathbf{d}_n$, where α_n is chosen so that \mathbf{d}_n^* is a Nash equilibrium of the new game defined by these linear utility functions. Since the “composite” utility function $\alpha_n^\top \mathbf{d}_n$ for user n is linear in the vector of rate allocations \mathbf{d}_n , the network structure is no longer relevant and the game defined by (T_1, \dots, T_R) decouples into J Cournot games, one at each link j . We finally apply Theorem 3 at each link to arrive at the bound in the theorem. \square

The preceding theorem extends Theorem 3 to general networks, where users may have arbitrary utility functions (subject to Assumption 1) and arbitrary paths available through the network. Observe that the key step is Lemma 8, which relates a Cournot game where users choose rates along each path to a Cournot game where users choose rates at each link. Note that this relationship is only established between Nash equilibria of the two games, but that suffices for the efficiency analysis we perform. Indeed, a direct analysis of the efficiency loss of the game defined (Q_1, \dots, Q_N) would be much more difficult, due to the strong coupling between the players; it is not *a priori* clear that adverse effects such as those discovered by Hajek and Yang [6] will not occur in this context.

Finally, we also note that the preceding theorem essentially uses Theorem 3 as a “black box.” Thus, any other efficiency loss bound that holds for a single-link Cournot game with arbitrary utilities (or even linear utilities, due to Lemma 4) and possibly more general cost functions can be extended to a network Cournot game where users choose rates for each path available.

IV. CONCLUSION

This paper has considered a simple model for network resource allocation: users choose the rate at which they want to send data, and links set prices according to the marginal cost of the total rate allocated. While such a scheme is efficient when all users are price taking, there is a loss of efficiency when users are able to anticipate the effects of their choices on the link prices. We established that this efficiency loss is bounded by 1/3 when links’ marginal costs are affine.

As mentioned in the Section II Introduction, the methods of this paper can be extended to also yield efficiency loss bounds

for a number of related economic environments of interest; we note two such environments here. When a monopolist with general convex cost function faces a concave market demand curve, the “deadweight loss” due to monopoly pricing can be guaranteed to be no worse than 1/3. In addition, when multiple oligopolists with general convex cost functions face an affine market demand curve, the same bound 1/3 on efficiency loss holds. These results are the first published nontrivial constant factor bounds on efficiency loss for Cournot competition; for further details, see [21].

The results of this paper suggest that the dual goals of scalable network mechanism design and robustness to selfish users can indeed be achieved, at least under some specific assumptions on the cost structure of the links. An interesting further direction, therefore, concerns the design of a network resource allocation mechanism that is scalable, and provides efficiency loss bounds *regardless* of the structure of the cost functions.

REFERENCES

- [1] M. Falkner, M. Devetsikiotis, and I. Lambadaris, “An overview of pricing concepts for broadband IP networks,” *IEEE Commun. Surveys*, vol. 3, no. 2, 2000.
- [2] F. P. Kelly, “Charging and rate control for elastic traffic,” *Eur. Trans. Telecommun.*, vol. 8, pp. 33–37, 1997.
- [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] 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.
- [5] R. Johari, S. Mannor, and J. N. Tsitsiklis, “Efficiency loss in a network resource allocation game: The case of elastic supply,” MIT Lab. Inf. Decision Syst., Cambridge, MA, 2004. Publication 2605.
- [6] B. Hajek and S. Yang, “Strategic buyers in a sum-bid game for flat networks,” 2004, submitted for publication.
- [7] T. Roughgarden and É. Tardos, “How bad is selfish routing?,” *J. ACM*, vol. 49, no. 2, pp. 236–259, 2002.
- [8] A. A. Cournot, *Recherches sur les Principes Mathématiques de la Théorie des Richesses*. New York: Macmillan and Co., 1897. Translated by N. T. Bacon; originally published in 1838.
- [9] A. F. Daughety, Ed., *Cournot Oligopoly*. Cambridge, U. K.: Cambridge Univ. Press, 1988.
- [10] J. W. Friedman, *Oligopoly and the Theory of Games*. Amsterdam, The Netherlands: North-Holland, 1977.
- [11] C. Shapiro, “Theories of oligopoly behavior,” in *Handbook of Industrial Organization*, R. Schmalensee and R. D. Willig, Eds. Amsterdam, The Netherlands: Elsevier Science, 1989, vol. 1, pp. 329–414.
- [12] R. Johari, “Efficiency loss in market mechanisms for resource allocation,” Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 2004.
- [13] E. Koutsoupias and C. Papadimitriou, “Worst-case equilibria,” in *Proc. 16th Annu. Symp. Theoretical Aspects of Comput. Sci.*, 1999, pp. 404–413.
- [14] A. S. Schulz and N. Stier Moses, “On the performance of user equilibria in traffic networks,” in *Proc. ACM-SIAM Symp. Discrete Algorithms*, 2003, pp. 86–87.
- [15] E. Anshelevich, A. Dasgupta, É. Tardos, and T. Wexler, “Near-optimal network design with selfish agents,” in *Proc. ACM Symp. Theory Comput.*, 2003, pp. 511–520.
- [16] A. Fabrikant, A. Luthra, E. Maneva, C. Papadimitriou, and S. Shenker, “On a network creation game,” in *Proc. 22nd Annu. ACM Symp. Principles Distrib. Comput.*, 2003, pp. 347–351.
- [17] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*. Oxford, U.K.: Oxford Univ. Press, 1995.
- [18] J. Rosen, “Existence and uniqueness of equilibrium points for concave n -person games,” *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.
- [19] W. Novshek, “On the existence of Cournot equilibrium,” *Review of Economic Studies*, vol. 52, no. 1, pp. 85–98, 1985.
- [20] J. Tirole, *The Theory of Industrial Organization*. Cambridge, MA: MIT Press, 1988.
- [21] R. Johari and J. N. Tsitsiklis, *Efficiency Loss in Cournot Games*. Cambridge, MA: MIT Lab. Inf. Decision Syst., 2005. Publication 2639.



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

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

University, Stanford, CA.



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 from Massachusetts Institute of Technology (MIT), Cambridge, in 1980, 1980, 1981, and 1984, respectively.

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

Dr. Tsitsiklis received 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).