Design, Analysis and Simulation of the Progressive Second Price Auction for Network Bandwidth Sharing*

CU/CTR/TR 487-98-21, `Revision: 2.10`

Aurel A. Lazar and Nemo Semret†
Dept. of Electrical Engineering, Columbia University
New York, NY, 10027-6699, USA.
{aurel, nemo}@ctr.columbia.edu

April 1, 1998

Abstract

We present the Progressive Second Price auction, a new decentralized mechanism for allocation of variable-size shares of a resource among multiple users. Unlike most mechanisms in the economics litterature, PSP is designed with a very small message space, making it suitable for real-time market pricing of communication bandwidth. Under elastic demand, the PSP auction is incentive compatible and stable, in that it has a "truthful" ϵ-Nash equilibrium where all players bid at prices equal to their marginal valuation of the resource. PSP is efficient in that the equilibrium allocation maximizes total user value.

The equilibrium holds when PSP is applied by independent resource controllers on each link of a network with arbitrary topology, with users having arbitrary but fixed routes. In the network case, the distributed mechanism has a further incentive compatibility in that submitting the same bid at all links along the route is an optimal strategy for each user, regardless of other players' actions.

Using a prototype implementation of the auction game on the Internet, we investigate how convergence times scale with the number of bidders, as well as the trade-off between economic efficiency and signalling load.

---

*Parts of this work were presented at the 8^{th} *International Symposium on Dynamic Games and Applications*, Maastricht, The Netherlands, July 1998.

†Corresponding author.

Finally, we provide a rate-distortion theoretic basis for valuation of bandwidth, which leads naturally to the elastic demand model that is assumed in the analysis of the mechanism.

Keywords: resource allocation, auctions, game theory, mechanism design, network pricing.

1 Introduction

Communication networks are characterized by what economists call externalities. The value a user gets from the network depends on the other users. The positive externalities are that a communication network is more valuable if more people are connected. The negative externalities are that resources are shared by users who – because of distance, population size, or selfishness – cannot or will not coordinate their actions sufficiently to achieve the most desirable allocation of resources. The recognition of this reality in many aspects of networks and distributed computations has lead in recent years to the emergence of game theoretic approaches in their analysis and design [21, 9, 23, 28, 15, 16].

Prices, whether they relate to "real money" in a public network or "funny money" (based on quotas) in a private system, play a key role as allocation control signals. In the former case, this role is of course intimately tied to another, which is to allow a network provider to remain in business [7].

The telephone system and the current Internet represent two extremes of the relationship between resource allocation and pricing. The resources allocated to a telephone call are fixed, and usage prices are based on the predictability of the total demand at any given time. On the Internet, the current practice of pricing by the maximum capacity of the user's connection (flat-rate pricing) decouples the allocation (actual use) of resources from the prices. In the emerging multiservice networks (ATM, Next-Generation Internet), neither of these approaches are viable. The former because of the wide and rapidly evolving range of applications (including some which adapt to resource availability) will make demand more difficult to predict. And the latter because, once the flat fee is paid, there are no incentives to limit usage since increasing consumption benefits the user individually, whereas limiting it to sustainable levels brings benefits which are shared by all. This makes it vulnerable to the well-known "tragedy of the commons". With flat rates alone, the tendency is toward increasing congestion which chases away high-intensity users, or increasing rates which exclude low-intensity users [6], in both cases leading to decreased network revenue.

Thus there is a need to develop new approaches to pricing of network resources. Among the requirements are: sensitivity to the range of resource requirements (either through a sufficiently rich range of traffic classes which are priced differently, or by allowing users to explicitly quantify resource requirements); prices must be dynamically responsive to unpredictable demand (market based system); perhaps most importantly the pricing architecture should constrain as little as possible the efficiency trade-offs of the policies.

The fundamental issue in designing pricing policies is trade-offs between engineering efficiency and economic efficiency. This trade-off, which is more or less constrained by the underlying network technology, has many dimensions, including:

- how much measurement (from usage to capacity pricing),
- the granularity of differently priced service offerings (e.g. number of traffic classes),
- the level of resource aggregation – both in time and in space – at which pricing is done (per packet/cell or per connection, at the edge of the network or at each hop), and
- the information requirement (how much *a priori* knowledge of user behavior and preferences is required/assumed by the network in computing prices).

An approach which achieves economic efficiency is the smart-market approach of [18], wherein each packet contains a bid, and if it is served, pays a clearing price given by the highest bid among packets which are denied service (dropped). This approach is incentive compatible in that the optimal strategy for a (selfish) user is to set the bid price in each packet equal to the true valuation. Each node in the network becomes an efficient market, but the engineering cost (sorting packets by bid price, per-packet and per-hop accounting) could be significant if line speeds are high relative to the processing power in the router. In [14], users are charged according to a combination of declared and measured characteristics of traffic. By taking an equivalent bandwidth model of resource utilization, and assuming appropriate traffic models, a menu of pricing plans indexed by the declared traffic can be offered which encourages users to make truthful declarations (e.g. of the mean rate), and also encourages the users' characterization efforts to be directed where they are most relevant to the network resource allocation.

As the pricing is relative, [14] does not aim to address the problem of determining the actual monetary values of the market price (that users would be willing to pay). Another pricing scheme which incorporates multiplexing gain is formulated in [12]. These and a number of other schemes are summarized in [11], in a comprehensive view of the connection establishment process, which identifies the user-network negotiation as the key "missing link" in network engineering/economic research. In terms of our taxonomy of the previous paragraph, this is part of the information requirement trade-off. Indeed, in the absence of formal mechanisms to deal with the information problem, complex and (at least intuitively) undesirable things happen. For example, some providers offer expensive "front of the book" rates to uninformed customers, and lower "back of the book" rates to informed customers who may be about to defect to another carrier (see [7] and also the recent wars between AT&T and MCI in consumer long-distance service in the United States). In [29], it is argued that architectural considerations such as where charges are assessed should take precedence over the pursuit of optimal efficiency, and spatial aggregation, i.e. edge pricing, is proposed as a useful paradigm.

In this paper, we propose a new auction mechanism which accommodates various dimensions of the engineering-economics trade-off. The mechanism applies to a generic arbitrarily divisible and additive resource model (which may be equivalent bandwidth, peak rate, contract regions, etc., at any level of aggregation.) It does not assume any specific mapping of resource allocation to quality of service. Rather, users are defined as having an explicit monetary valuation of quantities of resource, which the network doesn't or can't know a priori. Thus, in terms of our trade-off taxonomy, this mechanism aims for unlimited granularity, flexibility in the level of aggregation and minimal information requirement.

In the most likely auction scenaria, the bidders would be automated agents requesting bulk bandwidth for aggregate flows like virtual paths, virtual private networks, or edge capacity[2].

We begin by formally presenting the design of our Progressive Second Price auction mechanism for sharing a single arbitrarily divisible resource in Section 2. In Section 3, after describing our model of user preferences and the elastic demand assumption, we prove that PSP has the desired properties of incentive compatibility, stability, and efficiency. The section concludes with a simulation results on the convergence properties, and the efficiency trade-offs.

2 Design of an Auction for a Divisible Resource

2.1 Message Process

Following [31], it is useful to expose the design in terms of its two aspects: realization, where a message process that enables a certain allocation objective is defined; and Nash implementation, where allocation rules are designed with incentives which drive the players to an equilibrium where the (designer's) desired allocation is achieved.

In this section we define the message process. Here we make the fundamental choice which will constrain the subsequent aspects of the design. Our first concern here is with engineering. For the sake of scalability in a network setting, we shall aim for a process where a) the exchanged messages are as small as possible, while still conveying enough information to allow resource allocation and pricing to be performed without any a-priori knowledge of demand (market research, etc.); and b) the amount of computation at the center is minimized.

Given a quantity Q of a resource, and a set of players $\mathcal{I} = \{1, \dots, I\}$, an auction is a mechanism consisting of: 1) players submitting bids, i.e. declaring their desired share of the total resource and a price they are willing to pay for it, and 2) the auctioneer allocating shares of the resource to the players based on their bids.

Player i's bid is $s_i = (q_i, p_i) \in \mathcal{S}_i = [0, Q] \times [0, \infty)$, meaning he would like a quantity q_i at a *unit* price p_i. A bid profile is $s = (s_1, \dots, s_I)$. Following standard game theoretic notation, let $s_{-i} \equiv (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_I)$, i.e. the bid profile of player i's opponents, obtained from s by deleting s_i. When we wish to emphasize a dependence on a particular player's bid s_i, we will write the profile s as $(s_i; s_{-i})$.

The allocation is done by an **allocation rule** A,

$$\begin{array}{rrcl} A: & \mathcal{S} & \longrightarrow & \mathcal{S} \\ & s = (q, p) & \longmapsto & A(s) = (a(s), c(s)), \end{array}$$

where $\mathcal{S} = \prod_{i \in \mathcal{I}} \mathcal{S}_i$.

The i-th row of $A(s)$, $A_i(s) = (a_i(s), c_i(s))$, is the allocation to player i: she gets a quantity $a_i(s)$ for which she is charged $c_i(s)$. Note that p is a price per unit and c is a total cost.

An allocation rule A is **feasible** if $\forall s$,

$$\sum_{i \in \mathcal{I}} a_i(s) \leq Q$$

and $\forall i \in \mathcal{I}$,

$$\begin{aligned} a_i(s) &\leq q_i, \\ c_i(s) &\leq p_i q_i. \end{aligned}$$

Remark a: The above formulation is a generalization of what is usually meant by an auction. The latter is the special case where $a_w(s) = Q$ for some winner $w \in \mathcal{I}$ and $a_i(s) = 0, \forall i \neq w$, i.e. the sale of a single indivisible object to one buyer, for which the theory is well developed [20, 22]. In our approach, allocations are for arbitrary shares of the total available quantity of resource. Equivalently, one could slice the resource into many small units, each of which is auctioned as an indivisible object. But in a practical implementation of auctions for sharing a resource, a process of bidding for each individual unit would result in a tremendous signaling overhead. More importantly, since the users would be bidding on a discrete grid of quantities, analytical predictions of outcomes could be misleading since they could be sensitive to the particular choice of grid[1].

Remark b: Most of the mechanism design literature in Economics makes use of the following "Revelation Principle":

> Given any feasible[2] auction mechanism, there exists an equivalent[3] feasible direct revelation mechanism which gives to the seller and all bidders the same expected utilities as the given mechanism. ([22], Lemma 1)

In this sharing context, a direct revelation mechanism would be one where each user message consists of the user's type, which is the valuation[4] of the resource over the whole range of their possible demands, i.e. a function $\theta_i : [0, Q] \rightarrow [0, \infty)$, and the budget (see Section 3.1). A consequence of revelation principle is that the mechanism designer can restrict her attention to direct revelation mechanisms, find the best mechanism in terms of her (economic) efficiency objectives, and then – if necessary – transform it into an equivalent mechanism in the desired message space. This is convenient

[1] For a more detailed discussion of this point, see [4] p. 34, and references therein.

[2] Myerson's feasible auctions satisfy a number of requirements, notably an *incentive compatibility* constraint.

[3] By equivalent, it is meant that, at some equilibrium, all players get the same utility. There may be other, possibly ill-behaved, equilibria.

[4] The valuation of a given amount of resource is how much the user is willing to pay for that quantity. The inverse of the valuation is the user's demand function, giving a desired quantity for each price.

because one can exclude the infinitely many mechanisms with larger message spaces, without fear of missing any better designs.

In our sharing problem, the user's type is infinite-dimensional, but for engineering reasons we choose a message space that is 2-dimensional. Thus, a given message can come from many possible types, so there is no single way to do the transformation from the direct revelation mechanism to the desired one. Thus, unlike most of the mechanism design literature, we will take a direct approach, where we posit an allocation rule for our desired message space, and then show that it has an equilibrium, and that the design objective is met at equilibrium[5]. This is equivalent to guessing the right direct-revelation-to-desired-mechanism transformation and building it into the allocation rule from the start.

2.2 Allocation Rule

Define, for $y \geq 0$

$$\underline{Q}_i(y, s_{-i}) = \left[Q - \sum_{p_k \geq y, k \neq i} q_k \right]^+ . \tag{1}$$

and

$$Q_i(y, s_{-i}) = \lim_{\eta \searrow y} \underline{Q}_i(\eta, s_{-i}) = \left[Q - \sum_{p_k > y, k \neq i} q_k \right]^+ .$$

The "progressive second price" (PSP) allocation rule is defined as follows:

$$\begin{aligned} a_i(s) &= q_i \wedge \underline{Q}_i(p_i, s_{-i}), & (2) \\ c_i(s) &= \sum_{j \neq i} p_j \left[a_j(0; s_{-i}) - a_j(s_i; s_{-i}) \right], & (3) \end{aligned}$$

where $\wedge$ means taking the minimum.

Remark a: For a fixed opponent profile s_{-i}, $Q_i(p_i, s_{-i})$ represents the maximum available quantity at a bid price of p_i. **The intuition behind PSP** is an exclusion-compensation principle: player i pays for his allocation so as to exactly cover the "social opportunity cost" which is given by the declared willingness to pay (bids) of the users who are excluded by i's presence

[5]Our aim is not to show that we don't need message spaces that are larger than in the direct revelation mechanisms, but rather that we *can* use this smaller message space and still achieve our objective.

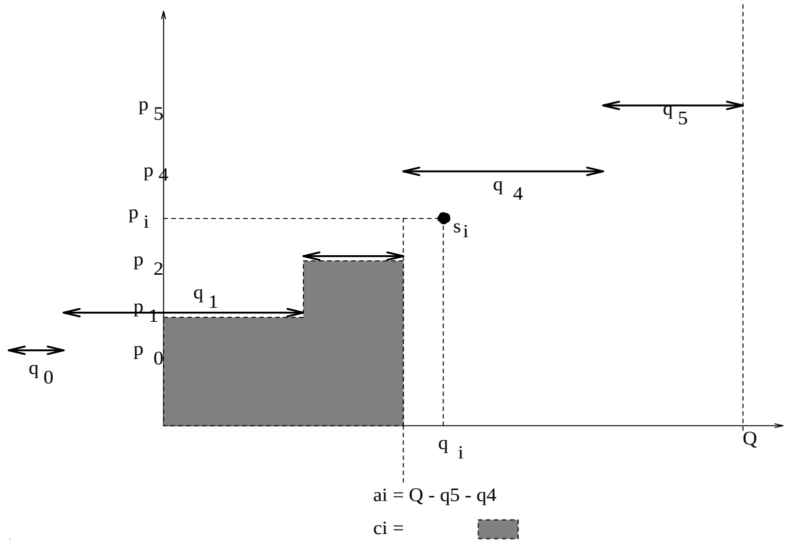

Figure 1: Exclusion-compensation principle: the intuition behind the PSP rule

(see Figure 1), and thus also compensates the seller for the maximum lost potential revenue. Note that this amounts to implicitly assuming that the bid price accurately reflects the marginal valuation θ_i' on the range $[a_i, q_i]$. In other words, by this rule the auctioneer is saying to the player: "if you bid (q_i, p_i), I take it to mean that in the vicinity of q_i, θ_i can be approximated by a line of slope p_i." This is the (built-in) transformation from the direct-revelation mechanism to the desired message process discussed in the second remark at the end of Section 2.1.

The charge c_i increases with a_i in a manner similar to the income tax in a progressive tax system. For a fixed opponent profile s_{-i}, imagine player i is increasing q_i, starting from 0. The first few units that player i gets will be taken away from the lowest clearing opponent (i.e. $m = \arg\min_j\{p_j : a_j > 0\}$), and player i will pay a price (marginal cost) p_m per unit. When a_m reaches 0, the subsequent units that player i gets will cost him $p_{m'} > p_m$, where m' is the new lowest clearing player, the one just above m. The PSP rule is the natural generalization of second-price auctions (or Vickrey auctions). In a Vickrey auction of a *single non-divisible object*, each player submits a sealed bid, and the object is sold to the highest bidder at the bid price of the second highest bidder, which is what happens here if $q_i = Q, \forall i$. This is widely known to have many desirable properties [30, 22, 4], the most important of which is that it has an equilibrium profile where all players bid their true valuation. As we will presently show, this property is preserved by the PSP rule in the more general case of sharing an arbitrarily divisible resource, and this leads to stability (Nash equilibrium). The PSP rule is analogous to Clarke-Groves mechanisms [3, 8, 19] in the direct-revelation case.

Remark b: When two players bid at exactly the same price, and they are asking for more than is available at that price, (2) punishes both of them. For example, if $Q = 100$ and $s_1 = (4, 60)$ and $s_2 = (4, 70)$, the allocations would be $a_1 = 60 \wedge (100 - 70) = 30$, and $a_2 = 70 \wedge (100 - 60) = 40$. Since the bid prices are equal, there is no "right" way to decide who to give the remaining capacity to. One could divide it equally, or proportionally to their requests, etc. For the subsequent analysis, it turns out it is simpler to not give it to either one (of course, it will be allocated to the lower bidders if there are any). This is just a technicality since by deciding this, we ensure that it will never happen (at equilibrium), since the users will always prefer to change their prices and/or reduce their quantity.

Considering the computational complexity of PSP, a straightforward implementation would at worst, sort the bids in time $I \log I$, perform (2) in linear time, and (3) can be done in time I^2. Thus, the complexity of computing the allocations is $O(I^2)$.

3 Analysis of the Progressive Second Price Auction

3.1 User Preferences

Since the allocation rule A is given by design, the only analytical assumptions we make is on the form of the players' preferences.

Player i's preferences are given by his utility function

$$\begin{array}{rccc} u_i : & \mathcal{S} & \longrightarrow & (-\infty, \infty) \\ & s & \longmapsto & u_i(s). \end{array}$$

Player i has a **valuation** of the resource $\theta_i(a_i(s)) \geq 0$, which is the total value to her of her allocation. Thus, for a bid profile of s, under allocation rule A, player i getting an allocation $A_i(s)$ has the quasi-linear utility

$$u_i(s) = \theta_i(a_i(s)) - c_i(s) \tag{4}$$

which is simply the value of what she gets minus the cost.

In addition, the player can be constrained by a **budget** $b_i \in [0, \infty]$, so the bid s_i must lie in the set

$$S_i(s_{-i}) = \{s_i \in \mathcal{S}_i : c_i(s_i; s_{-i}) \le b_i\}. \tag{5}$$

In the proofs of the following section, we will assume that users have elastic demand, that is:

Assumption 1 *For any $i \in \mathcal{I}$,*

- $\theta_i(0) = 0$,
- θ_i *is differentiable,*
- $\theta_i' \ge 0$*, non-increasing and continuous*
- $\exists \gamma_i > 0$*,* $\forall z \ge 0$*,* $\theta_i'(z) > 0 \Rightarrow \forall \eta < z, \theta_i'(z) \le \theta_i'(\eta) - \gamma_i(z - \eta)$.

The last item says that as long as the valuation is strictly increasing, it must also be strictly concave (with minimum curvature γ_i). However, it is allowed to "flatten" beyond a certain amount of resource.

Valuations of this form have wide applicability. In addition to the obvious economic justification of diminishing returns (the value of additional units of capacity is diminishes), such valuations can be justified by information theoretic fundamentals as well. For examples, see Section 3.4 and Appendix A.

3.2 Equilibrium of PSP

The auction game is given by $(Q, u_1, \ldots, u_I, A)$, that is, by specifying the resource, the players, and a feasible allocation rule. We analyze it as a strategic game of complete information.

Define the set of best replies to a profile s_{-i} of opponents bids: $S_i^*(s_{-i}) = \{s_i \in S_i(s_{-i}) : u_i(s_i; s_{-i}) \ge u_i(s_i'; s_{-i}), \forall s_i' \in S_i(s_{-i})\}$. Let $S^*(s) = \prod_i S_i^*(s_{-i})$. A Nash equilibrium is a fixed point of the point-to-set mapping S^*, i.e. a profile $s \in S^*(s)$. Such a point is what is most accepted as a consistent prediction of the actual outcome of a game, and has been repeatedly confirmed by experiments, as well as a wide range of theoretical approaches. Indeed, in a dynamic game, where players recompute the best response to the current strategy profile of their opponents, this iteration can only converge to a

Nash equilibrium (if it converges at all). In addition, an important trend in modern game theory is the development of learning models, and there too, it has been shown that Nash equilibria result also from rational learning through repeated play among the same players [13].

A more general (and hence weaker) notion of stability is the existence of an ϵ-Nash equilibrium. Let the ϵ-best replies be $S_i^\epsilon(s_{-i}) = \{s_i \in S_i(s_{-i}) : u_i(s_i; s_{-i}) \geq u_i(s_i'; s_{-i}) - \epsilon, \forall s_i' \in S_i(s_{-i})\}$. An ϵ-Nash equilibrium is a fixed point of S^ϵ.

In a dynamic auction game, $\epsilon > 0$ can be interpreted as a *bid fee* paid by a bidder each time they submit a bid. Thus, the user will send a best reply bid as long as it improves her current utility by ϵ, and the game can only end at an ϵ-Nash equilibrium.

Define

$$P_i(z, s_{-i}) = \inf \{y \geq 0 : Q_i(y, s_{-i}) \geq z\}. \tag{6}$$

Thus, for fixed s_{-i}, $\forall y, z \geq 0$,

$$z \leq Q_i(y, s_{-i}) \Rightarrow y \geq P_i(z, s_{-i}) \tag{7}$$

and[6]

$$y > P_i(z, s_{-i}) \Rightarrow z \leq Q_i(y, s_{-i}). \tag{8}$$

The graph of $P_i(., s_{-i})$ is the "staircase" shown in Figure 1, and that of $Q_i(., s_{-i})$ is obtained by flipping it 90 degrees.

It is readily apparent that

$$c_i(s) = \int_0^{a_i(s)} P_i(z, s_{-i})\, dz. \tag{9}$$

The key property of PSP is that, for a given opponent profile, a player cannot do much better than simply tell the truth, which in this setting means bidding at a price equal to the marginal valuation, i.e. set $p_i = \theta_i'(q_i)$. By doing so, she can always get within $\epsilon > 0$ of the best utility.

Let $\mathcal{T}_i = \{s_i \in \mathcal{S}_i : p_i = \theta_i'(q_i)\}$, the (unconstrained) set of player i's truthful bids, and $\mathcal{T} = \prod_i \mathcal{T}_i$.

Proposition 1 *(*Incentive compatibility*) Under Assumption 1, $\forall i \in \mathcal{I}$, $\forall s_{-i} \in \mathcal{S}_{-i}$, such that $Q_i(0, s_{-i}) = 0$, for any $\epsilon > 0$, there exists a truthful ϵ-best reply $t_i(s_{-i}) \in \mathcal{T}_i \cap S_i^\epsilon(s_{-i})$.*

[6] Actually, since $Q_i(., s_{-i})$ is upper-semi-continuous (jumps up), we have $z \leq Q_i(y, s_{-i}) \Leftrightarrow y \geq P_i(z, s_{-i})$

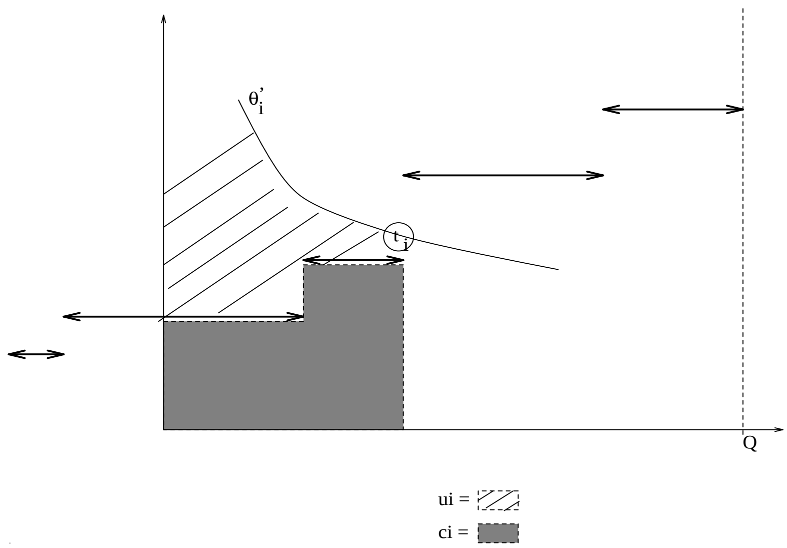

Figure 2: Truthful ϵ-best reply

In particular, let

$$G_i(s_{-i}) = \left\{ z \in [0, Q] : z \leq Q_i(\theta_i'(z), s_{-i}) \text{ and } \int_0^z P_i(\eta, s_{-i})\, d\eta \leq b_i \right\}.$$

Then with $v_i = [\sup G_i(s_{-i}) - \epsilon/\theta_i'(0)]^+$ *and* $w_i = \theta_i'(v_i)$, $t_i = (v_i, w_i) \in \mathcal{T}_i \cap S_i^\epsilon(s_{-i})$.

The truthful best reply can be found in a straightforward manner, as illustrated in Figure 2.

Proof: Fix $s_{-i} \in \mathcal{S}_{-i}$. Let $z_i = \sup G_i(s_{-i})$ and $y_i = \theta_i'(z_i)$.

By definition of z_i, $\exists \{z(n)\} \subset G_i(s_{-i})$ such that $\lim_n z(n) = z_i$. Hence $b_i \geq \lim_n \int_0^{z(n)} P_i(\eta, s_{-i})\, d\eta = \int_0^{z_i} P_i(\eta, s_{-i})\, d\eta \geq c_i(t_i; s_{-i})$, where the equality comes from the boundedness of P_i and the Lebesgue dominated convergence theorem, and the second inequality from (9) and (2). Thus $t_i \in \mathcal{T} \cap S_i(s_{-i})$.

Next we show that $t_i \in S_i^\epsilon(s_{-i})$. First, $z_i = \lim_n z(n) \leq \lim_n Q_i(\theta_i'(z(n)), s_{-i}) \leq Q_i(\lim_n \theta_i'(z(n)), s_{-i})$, where the inequalities follow respectively from $z(n) \in G_i(s_{-i})$, and the upper semi-continuity of $Q_i(., s_{-i})$. Now by the continuity of θ_i', $Q_i(\lim_n \theta_i'(z(n)), s_{-i}) = Q_i(\theta_i'(z_i), s_{-i}) = Q_i(y_i, s_{-i})$, hence

$$z_i \leq Q_i(y_i, s_{-i}). \tag{10}$$

Now, we claim that $a_i(t_i; s_{-i}) = v_i$. Indeed, if $z_i = 0$ then $v_i = 0$ and $a_i(t_i; s_{-i}) = 0$. If $z_i > 0$, then by (10), $Q_i(y_i, s_{-i}) > 0$ and since by hypothesis $Q_i(0, s_{-i}) = 0$, we have $\theta_i'(z_i) = y_i > 0$. Also, $z_i > 0$ implies $v_i < z_i$. Therefore, by Assumption 1, we have $w_i = \theta_i'(v_i) > \theta_i'(z_i) = y_i$. Hence, since $\underline{Q}_i(., s_{-i})$ is non-decreasing, $\underline{Q}_i(w_i, s_{-i}) \geq \lim_{\eta \searrow y} \underline{Q}_i(\eta, s_{-i}) = Q_i(y_i, s_{-i}) \geq z_i > v_i$. Thus, by (2),

$$a_i(t_i; s_{-i}) = v_i \tag{11}$$

Now $\forall s_i \in S_i(s_{-i})$,

$$\begin{aligned}
& u_i(t_i; s_{-i}) - u_i(s) \\
&= \theta_i(a_i(t_i; s_{-i})) - \theta_i(a_i(s)) - c_i(t_i; s_{-i}) + c_i(s) \\
&= \int_{a_i(t_i;s_{-i})}^{a_i(s)} [P_i(z, s_{-i}) - \theta_i'(z)]\, dz. \\
&= \int_{z_i}^{a_i(s)} [P_i(z, s_{-i}) - \theta_i'(z)]\, dz + \int_{v_i}^{z_i} [P_i(z, s_{-i}) - \theta_i'(z)]\, dz \\
&\geq \int_{z_i}^{a_i(s)} [P_i(z, s_{-i}) - \theta_i'(z)]\, dz - \epsilon
\end{aligned} \tag{12}$$

where the inequality follows from $(z_i - v_i) \leq \epsilon/\theta_i'(0)$ and the fact that θ_i' is non-increasing. Thus, it suffices to show that the integral is ≥ 0.

If $z_i < a_i(s)$, take any $z \in (z_i, a_i(s)]$. By the definition of z_i, $z \notin G_i(s_{-i})$. Now $s_i \in S_i(s_{-i})$ implies $b_i \geq c_i(s) = \int_0^{a_i(s)} P_i(\eta, s_{-i})\, d\eta \geq \int_0^z P_i(\eta, s_{-i})\, d\eta$. Therefore, we must have $z > Q_i(\theta_i'(z))$, which by (8), implies $\theta_i'(z) \leq P_i(z)$ and the integrand in (12) is ≥ 0 as desired.

Suppose $z_i \geq a_i(s_i)$. Since θ_i' is non-increasing, $Q_i(., s_{-i})$ is non-decreasing and $P_i(., s_{-i}) \geq 0$, any point to the left of z_i is in the set $G_i(s_{-i})$, $\forall z < z_i, z \in G_i(s_{-i})$, hence $z \leq Q_i(\theta_i'(z), s_{-i})$ which by (7), implies $\theta_i'(z) \geq P_i(z, s_{-i})$, so the integrand in (12) is ≤ 0 as desired. □

Figure 3 shows the utility function of player 4, $u_4(s_4)$, in a PSP auction with $I = 5$ players, with s_{-4} fixed, and a valuation $\theta_4(q) = 10q$. The plateaus correspond to the points where $q_4 \geq Q_4(p_4, s) = \left[Q - \sum_{\{j: p_j > p_4\}} a_j(s)\right]^+$, and $a_4(s)$ can no longer be increased at that bid price – see (2). At bid prices $p_4 > p_5$, the utility decreases when $a_4 > Q - q_5$, because after that point, each additional unit of resource is taken away from player 5, and thus costs p_5, which is more than θ_4' its value to player i. Thus, each additional unit starts bringing negative utility. This is what discourages users from bidding above their valuation. Proposition 1 is illustrated by the fact that for any given quantity q_4, the utility u_4 is maximized on the plane $p_4 = \theta_4' \equiv 10$.

Remark: When the players have linear valuations and no budget constraint ($b_i = \infty$), PSP becomes identical to a second-price auction for a

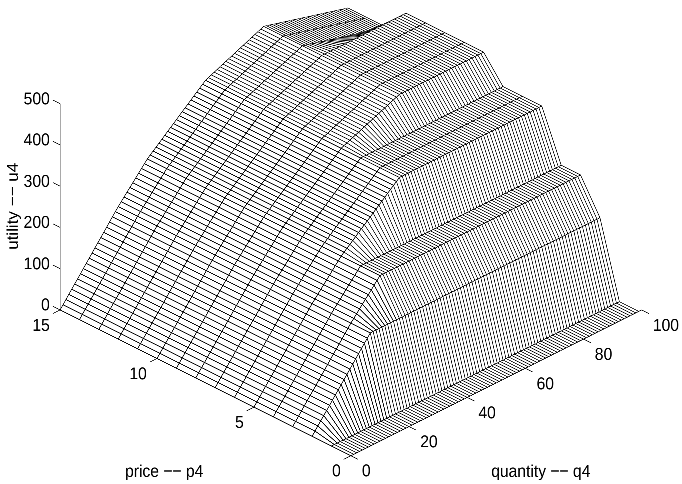

Figure 3: Utility $u_4(s_4)$ for $s_1 = (100, 1)$, $s_2 = (10, 2)$, $s_3 = (20, 4)$, $s_5 = (20, 7)$, $s_6 = (30, 12)$

non-divisible object. Then the existence of a Nash equilibrium follows directly from incentive compatibility.

Were the message process such that players declared a price and a budget (rather than desired quantity), it may have been possible to design an allocation rule A such that they are inclined to reveal their true budget, thus obtaining incentive compatibility in both dimensions, and hence equilibrium. But such a rule A would likely not have a simple closed form like (2) and (3). In essence, the computational load of translating budgets into shares would be centralized at the auctioneer, thus making the system less scalable to large numbers of users. On the other hand, decentralization has a cost too, which is the signaling overhead resulting from players possibly adjusting bids based on opponent bids in the iterated game. Our design is based on the premise that the latter approach is the more scalable of the two (indeed that was the reason for choosing a small message space).

The next property is that the truthful best reply is continuous in opponent profiles (this can be seen in Figure 2: as the "staircase" is varied smoothly, the point of intersection with θ_i' moves smoothly, provided θ_i' is not flat – which is given by the last time item in Assumption 1). To prove that, we will need the following:

Lemma 1 *$\forall s, s' \in \mathcal{S}, \forall y, z \geq 0, \forall \delta > 0$, if $||s_{-i} - s'_{-i}|| < \delta$ then*

$$Q_i(y+\delta, s_{-i}) + \delta\sqrt{I} \geq Q_i(y, s'_{-i}) \geq Q_i(y-\delta, s_{-i}) - \delta\sqrt{I}, \tag{13}$$

and

$$P_i(z+\delta\sqrt{I}, s_{-i}) + \delta \geq P_i(z, s'_{-i}) \geq P_i(z-\delta\sqrt{I}, s_{-i}) - \delta. \tag{14}$$

Proof: First, $||s_{-i} - s'_{-i}|| < \delta$ implies $\sum_k |q'_k - q_k| < \delta\sqrt{I}$, and $p_k + \delta > p'_k > p_k - \delta$. Thus, $\sum_k q_k 1_{\{p_k+\delta>y\}} + \delta\sqrt{I} \geq \sum_k q'_k 1_{\{p'_k>y\}} \geq \sum_k q_k 1_{\{p_k-\delta>y\}} - \delta\sqrt{I}$. Then, using (1) and the identity $(a+b)^+ \leq (a)^+ + (b)^+$, the first result follows.

For any $y < P_i(z, s'_{-i})$, by (7), we have $z > Q_i(y, s'_{-i}) \geq Q_i(y-\delta, s_{-i}) - \delta\sqrt{I}$, which by (8), implies $y - \delta \leq P_i(z+\delta\sqrt{I}, s_{-i})$. Letting $y \nearrow P_i(z, s'_{-i})$, we get $P_i(z, s'_{-i}) \leq P_i(z+\delta\sqrt{I}, s_{-i}) + \delta$.

For any $y > P_i(z, s'_{-i})$, by (8), we have $z \leq Q_i(y, s'_{-i}) \leq Q_i(y+\delta, s_{-i}) + \delta\sqrt{I}$, which by (7) implies $y + \delta \geq P_i(z - \delta\sqrt{I})$. Letting $y \searrow P_i(z, s'_{-i})$, we get $P_i(z, s'_{-i}) \geq P_i(z-\delta\sqrt{I}) - \delta$. $\square$

Lemma 2 *(Continuity of best reply) Under Assumption 1, $\forall i \in \mathcal{I}$, the ϵ-best reply t_i given in Proposition 1 is continuous in s_{-i} on any subset $V_i(\underline{P}, \overline{P}) = \{s_{-i} \in \mathcal{S}_i : \forall z > 0, \overline{P} \geq P_i(z, s_{-i}) \geq \underline{P}\}$, with $\infty > \overline{P} \geq \underline{P} > 0$.*

Proof: Let $z_i = \sup G_i(s_{-i})$. We will show z_i is continuous, and the continuity of $v_i = [z_i - \epsilon/\theta'_i(0)]^+$ and $w_i = \theta'_i(v_i)$ follow immediately (recall that by Assumption 1, θ'_i is continuous).

Suppose there is a discontinuity at some s_{-i}. Then, $\exists \epsilon_1 > 0$, such that $\forall \delta > 0$, $\exists s'_{-i} \in V_i(\underline{P}, \overline{P})$ with $||s_{-i} - s'_{-i}|| < \delta$ and $||z_i - z'_i|| \geq \epsilon_1$, where $z'_i = \sup G_i(s'_{-i})$.

Suppose $z_i + \epsilon_1 \leq z'_i = \sup G_i(s'_{-i})$ (the case $z'_i + \epsilon_1 \leq z_i$ is handled identically, with s_{-i} and s'_{-i} interchanged.). Consider the definition of G_i; since θ'_i is decreasing and $Q_i(., s_{-i})$ is non-decreasing and $P_i(., s_{-i}) \geq 0$, any point to the left of z'_i is in the set $G_i(s'_{-i})$, therefore

$$z_i + \epsilon_1 \in G_i(s'_{-i}). \tag{15}$$

Thus, $z_i + \epsilon_1 \leq Q_i(\theta'_i(z_i+\epsilon_1), s'_{-i}) = \left[Q - \sum_k q'_k 1_{\{p'_k > \theta'_i(z_i+\epsilon_1)\}}\right]^+$. Therefore,

$$z_i + \epsilon_1 \leq Q_i(\theta'_i(z_i+\epsilon_1) + \delta, s_{-i}) + \delta\sqrt{I},$$

using Lemma 1.

Also, by (7) $z_i + \epsilon_1 \leq Q_i(\theta'_i(z_i+\epsilon_1), s'_{-i}) \Rightarrow \theta'_i(z_i+\epsilon_1) > P(z_i+\epsilon_1, s'_{-i})$. Now since $s'_{-i} \in V_i(\underline{P}, \overline{P})$, this last expression is $\geq \underline{P} > 0$, hence $\theta'_i(z_i+\epsilon_1) > 0$. Then using Assumption 1, $\theta'_i(z_i+\epsilon_2) \geq \theta'_i(z_i+\epsilon_1) + \gamma_i(\epsilon_1 - \epsilon_2) > \theta'_i(z_i+\epsilon_1) + \delta$, for

$\delta < \delta_1 = \frac{\epsilon_1}{\sqrt{I}} \wedge \epsilon_1 \gamma_i$, and $0 < \epsilon_2 < (\epsilon_1 - \delta_1 \sqrt{I}) \wedge (\epsilon_1 - \delta_1/\gamma_i)$. Therefore, since $Q_i(., s_{-i})$ is non-decreasing,

$$\begin{aligned} z_i + \epsilon_2 &\leq Q_i(\theta_i'(z_i + \epsilon_2), s_{-i}) + \delta\sqrt{I} - \epsilon_1 + \epsilon_2 \\ &< Q_i(\theta_i'(z_i + \epsilon_2), s_{-i}). \end{aligned} \tag{16}$$

Now (15) also implies that

$$\begin{aligned} b_i &\geq \int_0^{z_i+\epsilon_1} P_i(\eta, s'_{-i})\, d\eta \\ &\geq \int_0^{z_i+\epsilon_3} P_i(\eta, s_{-i})\, d\eta + \int_0^{z_i+\epsilon_3} \left[P_i(\eta, s'_{-i}) - P_i(\eta, s_{-i})\right]\, d\eta + (\epsilon_1 - \epsilon_3)\underline{P}, \end{aligned}$$

and this holds $\forall \epsilon_3 < \epsilon_1$. Now, using Lemma 1,

$$\begin{aligned} &\int_0^{z_i+\epsilon_3} \left[P_i(\eta, s'_{-i}) - P_i(\eta, s_{-i})\right]\, d\eta \\ &\quad\geq -\delta Q + \int_0^{z_i+\epsilon_3} P_i(\eta - \delta\sqrt{I}, s_{-i})\, d\eta - \int_0^{z_i+\epsilon_3} P_i(\eta, s_{-i})\, d\eta \\ &\quad\geq -\delta Q - \int_{z_i+\epsilon_3-\delta\sqrt{I}}^{z_i+\epsilon_3} P_i(\eta, s_{-i})\, d\eta \\ &\quad\geq -(Q + \overline{P})\delta\sqrt{I}. \end{aligned}$$

Let $\delta_2 = \frac{\epsilon_1 \underline{P}}{(Q+\overline{P})\sqrt{I}}$, and ϵ_3 such that $0 < \epsilon_3 < \epsilon_1 - (Q + \overline{P})\delta_2\sqrt{I}/\underline{P}$. Then

$$b_i \geq \int_0^{z_i+\epsilon_3} P_i(\eta, s_{-i})\, d\eta, \tag{17}$$

for $\delta < \delta_2$.

Now choosing $\delta < \delta_1 \wedge \delta_2$, (16) and (17) imply that $G_i(s_{-i}) \ni (z_i + \epsilon_3) \wedge (z_i + \epsilon_2) > z_i = \sup G_i(s_{-i})$, a contradiction. □

We introduce one additional player, player 0, whose valuation is $\theta_0(z) = p_0 z$, and whose bid can therefore be fixed at $s_0 = (q_0, p_0) = (Q, p_0)$. Player 0 can be viewed as the auctioneer, and $p_0 > 0$ as a "reserve price" at which the seller is willing to "buy" all of the resource from himself. From (1), the presence of the bid $s_0 = (Q, p_0)$ implies $\forall i \in \mathcal{I}, Q_i(y, s_{-i}) = 0, \forall y < p_0$. In particular, setting $y = 0$, the condition of Proposition 1 holds. Thus, we can restrict our attention to truthful strategies only, and still have feasible best replies. This forms a "truthful" game embedded within the larger auction game, where the strategy space is $\mathcal{T} \subset \mathcal{S}$, the feasible sets are $\mathcal{T}_i \cap S_i(s_{-i})$, and the best replies are $R_i^\epsilon(s) = \mathcal{T}_i \cap S_i^\epsilon(s)$. A fixed point of R^ϵ in $\mathcal{T}$ is a fixed point of S^ϵ in $\mathcal{S}$. Thus an equilibrium of the embedded game is an equilibrium of the whole game.

Proposition 2 *(Nash equilibrium) In the auction game with the PSP rule given by (2) and (3) and a reserve price $p_0 > 0$, and players described by (4) and (5), if Assumption 1 holds, then for any $\epsilon > 0$, there exists a truthful ϵ-Nash equilibrium $s^* \in \mathcal{T}$.*

Proof: $\forall s \in \mathcal{T}$, $\forall i \in \mathcal{I}$, $\forall z > 0$, we have $z > 0 = Q_i(p_0/2, s_{-i})$, which by (8) implies $P_i(z, s_{-i}) \geq p_0/2 = \underline{P}$. Let $\overline{P} = \max_{k \in \mathcal{I} \cup \{0\}} \theta'_k(0)$. Then, the conditions of Lemma 2 are satisfied and $t = (v, w)$ **is continuous in** s **on** $\mathcal{T}$. By Assumption 1, θ'_i is continuous therefore $v(q,p) = v(q, \theta'(q))$ (as defined in Proposition 1), can be viewed as a continuous mapping of $[0, Q]^I$ onto itself. By Brouwer's fixed-point theorem (see for example [10]), any continuous mapping of a convex compact set into itself has at least one fixed point, i.e. $\exists q^* = v(q^*) \in [0, Q]^I$. Now with $s^* = (q^*, \theta'(q^*))$, we have $s^* = t(s^*) \in \mathcal{T}$. $\square$

3.3 Efficiency

The objective in designing the auction is that, at equilbrium, resources always go to those who value them most. Indeed, the PSP mechanism does have that property. This can be loosely argued as follows: for each player, the marginal valuation is never greater than the bid price of any opponent who is getting a non-zero allocation. Thus, whenever there is a player j whose marginal valuation is less than player i's and j is getting a non-zero allocation, i can take some away from j, paying a price less than i's marginal valuation, i.e. increasing u_i, but also increasing the total value, since i's marginal value is greater. Thus at equilibrium, i.e. when no one can unilaterally increase their utility, the total value is maximized. Formally, consider $a \in \arg\max_{\mathcal{A}} \sum_i \theta_i(a_i)$. The Karush-Kuhn-Tucker optimality conditions are that there exists a Lagrange multiplier λ such that $\theta'_i(a_i) = \lambda$, if $a_i > 0$, and $\theta'_i(a_i) \leq \lambda$, otherwise.

Assumption 2 *For any $i \in \mathcal{I}$, $b_i = \infty$, and θ'_i satisfies*[7]

$$\theta'_i(z) - \theta'_i(z') > -\kappa(z - z'),$$

whenever $z > z' \geq 0$.

[7] If θ'_i is differentiable, the condition is $0 \geq \theta''_i > -\kappa$.

Given any $\epsilon > 0$, for any ϵ-Nash-equilbrium $s^* \in \mathcal{T}$, let $a^* \equiv a(s^*)$, and let $\underline{a}^* \equiv \min_{i \in \mathcal{I} \cup \{0\}, a_i > 0} a_i^*$, the smallest non-zero allocation. The following is the "ϵ version" of the Karush-Kuhn-Tucker conditions.

Lemma 3 *Suppose Assumptions 1 and 2 hold. If for some j, $a_j^* > \sqrt{\epsilon/\kappa}$, then $\forall i \in \mathcal{I} \cup \{0\}$,*

$$\theta_i'(a_i^*) < \theta_j'(a_j^*) + 2\sqrt{\epsilon\kappa}.$$

An immediate corollary is that if $\underline{a}^* > \sqrt{\epsilon/\kappa}$ then

$$\lambda^* - 2\sqrt{\epsilon\kappa} < \theta_i'(a_i^*) < \lambda^* + 2\sqrt{\epsilon\kappa}$$

if $a_i^* > \sqrt{\epsilon/\kappa}$ and

$$\theta_i'(a_i^*) < \lambda^* + 2\sqrt{\epsilon\kappa},$$

if $a_i^* = 0$, for some $\lambda^* \geq 0$.

Proof: Suppose $\theta_i'(a_i^*) \geq \theta_j'(a_j^*) + 2\sqrt{\epsilon\kappa}$. Since $\theta_j'(a_j^*) \geq \theta_j'(q_j^*) \geq p_j^*$, we have $\theta_i'(a_i^*) \geq p_j^* + 2\sqrt{\epsilon\kappa}$.

Since $a_j^* > 0$, if player i bids at a price above p_j^*, he can take all of player j's allocation, without losing anything of his own, i.e. $a_i^* + a_j^* \leq Q_i(p_j^*, s_{-i}^*)$. By (7), this implies

$$p_j^* \geq P_i(a_i^* + a_j^*, s_{-i}^*).$$

Let $q_i = (a_i^* + \sqrt{\epsilon/\kappa})$ and $s_i = (q_i, \theta_i'(q_i))$. Then

$$\begin{aligned}
u_i(s_i; s_{-i}^*) - u_i(s^*) &= \int_{a_i^*}^{a_i^* + \sqrt{\epsilon/\kappa}} \theta_i'(z) - P_i(z, s_{-i}^*)\, dz \\
&\geq \left[\theta_i'(a_i^* + \sqrt{\epsilon/\kappa}) - p_j^*\right] \sqrt{\epsilon/\kappa} \\
&\geq \left[\theta_i'(a_i^*) - \kappa\sqrt{\epsilon/\kappa} - p_j^*\right] \sqrt{\epsilon/\kappa} \\
&\geq \left[2\sqrt{\epsilon\kappa} - \kappa\sqrt{\epsilon/\kappa}\right] \sqrt{\epsilon/\kappa} \\
&= \epsilon
\end{aligned}$$

which contradicts the fact that s^* is an ϵ-Nash equilibrium. □

Proposition 3 *(Efficiency) Suppose Assumptions 1 and 2 hold. If $\underline{a}^* > \sqrt{\epsilon/\kappa}$, then*

$$\max_{\mathcal{A}} \sum_i \theta_i(a_i) - \sum_i \theta_i(a_i^*) = O(\sqrt{\epsilon\kappa}),$$

where $\mathcal{A} = \{a \in [0, Q]^{I+1} : \sum_i a_i \leq Q\}$.

Proof: (of Proposition 3) Let $\mathcal{I}^+ = \{k : a_k > a_k^*\}$ and $\mathcal{I}^- = \{k : a_k < a_k^*\}$. For $i \in \mathcal{I}^+$, we have $\theta_i'(a_i^*) \leq \lambda^* + 2\sqrt{\epsilon\kappa}$. For $i \in \mathcal{I}^-$, we have $a_i^* > a_i \geq 0$, therefore by the lemma, $\theta_i'(a_i^*) > \lambda^* - 2\sqrt{\epsilon\kappa}$. Therefore,

$$\begin{aligned}\sum_{\mathcal{I}} \theta_i(a_i) - \theta_i(a_i^*) &\leq \sum_{\mathcal{I}^+} \theta_i'(a_i^*)(a_i - a_i^*) - \sum_{\mathcal{I}^-} \theta_i'(a_i^*)(a_i^* - a_i) \\ &\leq (\lambda^* + 2\sqrt{\epsilon\kappa})\Delta - (\lambda^* - 2\sqrt{\epsilon\kappa})\Delta,\end{aligned}$$

where $\Delta = \sum_{\mathcal{I}^+}(a_i - a_i^*) = \sum_{\mathcal{I}^-}(a_i^* - a_i)$. Since $\Delta \leq Q$ the result follows, with the bound $4Q\sqrt{\epsilon\kappa}$. □

The condition $b_i = \infty$ is sufficient, but not necessary to achieve efficient outcomes. In fact with any budget profile, efficiency can be achieved if the users cooperate. For example, if they all choose a bid quantity close to what they can actually obtain (which they do if they use the strategy given by Proposition 1), then the price paid would be p_0 per unit for all the allocations, and if p_0 or the shares a_i^* are not too large, then budget constraints are irrelevant and a^* is efficient. More generally, efficiency is attained if the budgets are not too far out of line with the valuations, i.e. there are no players with very high demand and very low budget.

Proposition 3 provides a key to understanding the basic trade-off between engineering and economic efficiency. The smaller ϵ, the closer we get to the value-optimal allocations. But in a dynamic game, where players iteratively adjust their bids to the opponent profile, a player will bid as long as he can gain at least ϵ utility (since that is the cost of the bid), thus a smaller ϵ makes the iteration take longer to converge, i.e. entails more signaling.

3.4 Convergence

An issue of obvious concern is how the convergence time scales with the number of bidders. We consider this experimentally using the software described in Appendix B.

In all our simulations we let $Q = 100$. For each user, the valuation is strictly increasing and concave up to a maximum corresponding to a physical line capacity, and flat beyond that. Since, as shown by Proposition 3, only the second derivative of the valuation is needed to measure the efficiency of the PSP auction, a second order (parabolic) model is deemed sufficient. Thus we use valuations of the form:

$$\theta_i(z) = -\kappa_i(z \wedge \overline{q}_i)^2/2 + \kappa_i\overline{q_i}(z \wedge \overline{q}_i),$$

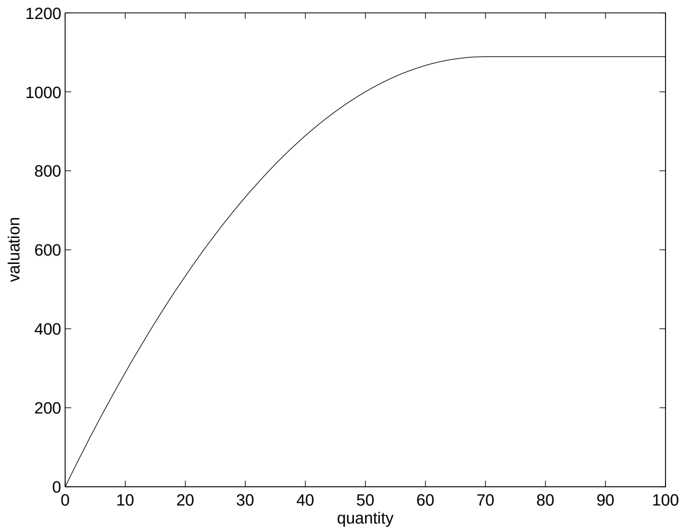

Figure 4: Parabolic valuation with $\kappa = 0.5$ and $\overline{q} = 70$

where $\overline{q}_i$ is the line rate, and $\kappa_i > 0$ (see Figure 4).

We generate our user population with independent random variables $\{\theta_i'(0)\}_{\mathcal{I}}$ (corresponding to the maximum unit price the user would pay) uniformly distributed in $[10, 20]$, and $\kappa_i = \theta_i'(0)/\overline{q}_i$, and $\overline{q}_i$ uniformly distributed on $[50, 100]$. All players have a budget $b_i = 100$. The bid fee is fixed at $\epsilon = 5$. Each user has a bidding agent which can submit at most one bid per second (see Algorithm 1 in Appendix B).

With this set-up, the results are shown in Figures 5-6. Simulations were run for 11 population sizes ranging from 2 to 96 player. Each point is simulated 10 times with new random valuations for all players. The overall mean is 11.9 bids per player. From Figure 5, the number of bids seems to grow as the square of the number of players.

The actual time to converge, shown in Figure 6, grows more slowly, since the computation of bids is done in parallel by all the players. In fact, for small numbers of players, the time decreases. This can best be explained as follows. Suppose there are only two players, with similar valuations. They will both start by asking for their maximum quantity, at their marginal valuation (which at their maximum quantity is near zero). Then as each sees the other's bid, each will reduce the quantity and increase the price a little bit. And they go on taking turns, gradually raising the market price

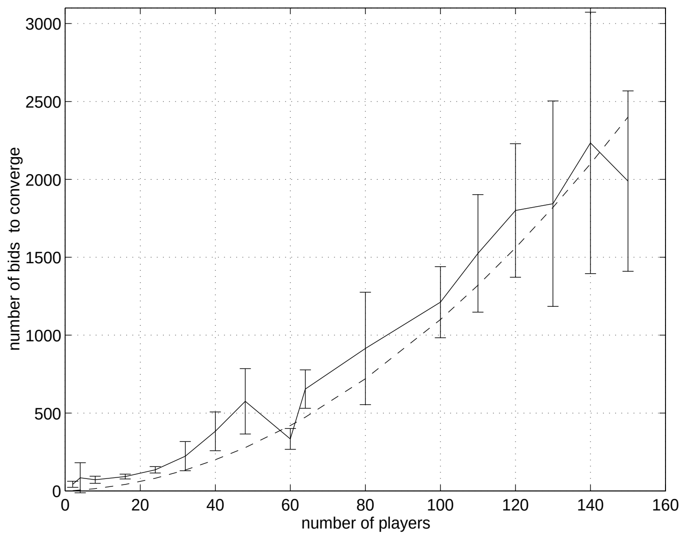

Figure 5: Mean (+/- std. dev.) number of bids – solid line. The dashed line is $I + I^2/10$.

until they reach an equilibrium. However if there are 10 players, in between two bids by the same player, the 9 others will already have bid up the price, so he will jump to higher price than if there was only one opponent. Thus the equilibrium market price will be reached more quickly. For large populations, this effect becomes small compared to the sheer volume of bids, and the convergence time starts to grow.

The trade-off between signaling and economic efficiency discussed in light of Proposition 3 is illustrated by Figures 7-8. Increasing the bid fee speeds up convergence, at a cost of lost efficiency. A resource manager should select a bid fee which optimally balances the two in a given context. Figure 8 also illustrates the validity of the lower bound given by Proposition 3.

4 Decentralized PSP Auctioning of Networked Resources

4.1 Formulation

In this section, we extend the formulation of Section 2 to the network case, where there is a set of resources $\mathcal{L} = \{1, \ldots, L\}$, of which the quantities are

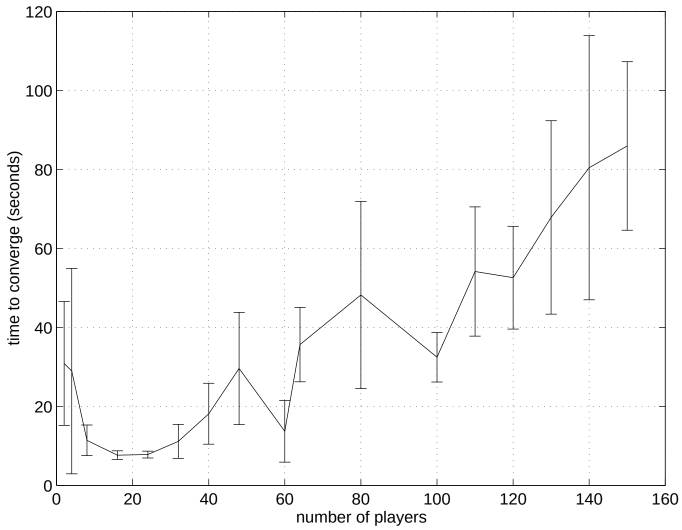

Figure 6: Mean (+/- std. dev.) convergence time in seconds (for a 1 second bid interval).

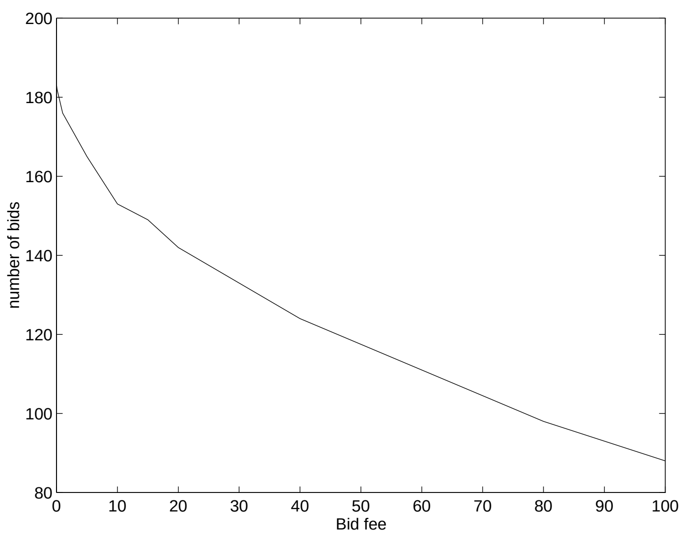

Figure 7: Number of bids to converge vs ϵ

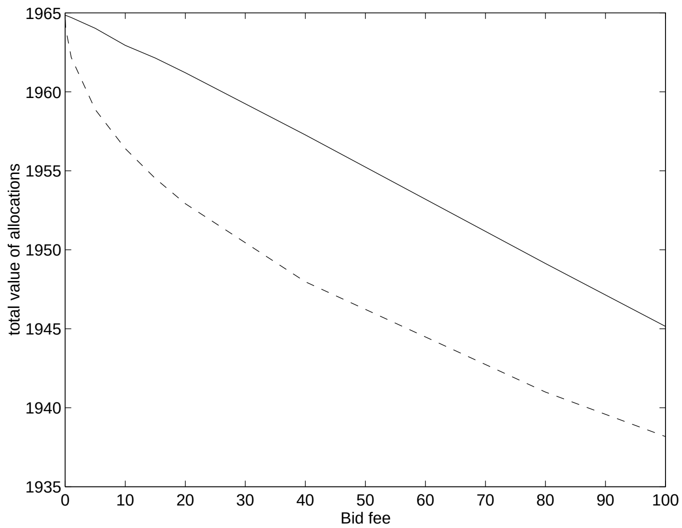

Figure 8: $\sum_i \theta_i(a_i^*)$ – solid line, and $\max_a \sum_i \theta_i(a_i) - 4(\epsilon\kappa)^{-1/2}$ – dashed line, vs ϵ.

$Q^1, \ldots, Q^L$, and as before, a set of players $\mathcal{I} = \{1, \ldots, I\}$.

A basic goal is that the mechanism be decentralized in that the allocations at any node depend only on local information: the resources available at that node and the bids for that node only. This makes the mechanism applicable to cases where the various resources being auctioned may be owned by different entities. Each player is responsible for coordinating (or not) her bids at the different nodes on her route in such a way that maximizes her utility.

Let $\mathcal{Q}^l = [0, Q^l]$, and $\mathcal{Q} = \prod_{l \in \mathcal{L}} \mathcal{Q}^l$. Player i's bid is now $s_i = (s_i^1, \ldots, s_i^L) \in \mathcal{S}_i = \prod_{l \in \mathcal{L}} \mathcal{S}_i^l$, where $s_i^l = (q_i^l, p_i^l) \in \mathcal{S}_i^l = \mathcal{Q}^l \times [0, \infty)$ is the bid for resource $l \in \mathcal{L}$. At each node $l \in \mathcal{L}$, we have an allocation rule A^l, as defined in Section 2, which maps a profile $s^l \in \mathcal{S}^l = \prod_{i \in \mathcal{I}} \mathcal{S}_i^l$ to an allocation $A^l(s^l) = (a^l(s^l), c^l(s^l))$.

In keeping with our motivation, which comes from communication networks, player i's type now includes a *route*[8], $r_i \subset \mathcal{L}$, as well as a valuation and budget as in Section 3. We will assume that players only care about the end-to-end "thickness" of their allocated "pipe", which is given by the

[8] Our analysis will not require that r_i form a continuous path, or any specific type of subgraph – "route" means any arbitrary subset of nodes.

thinnest allocation, and the total charge. Thus,

$$u_i(s) = \theta_i(\min_{l \in r_i} a_i^l(s)) - c_i(s), \tag{18}$$

where

$$c_i(s) = \sum_{l \in \mathcal{L}} c_i^l(s).$$

Remark: In reality, routing itself is a competitive game, and the decentralized nature of the auction makes it possible for players to make the route part of their strategy and thus vary it in response to other users' actions. In our analysis however, we assume players have obtained a (fixed) route before entering the auction game. In the broader context, the auction game may be nested within a larger game which includes routing.

4.2 Equilibrium of Networked PSP Auctions

Assume that the allocation at each node $l \in \mathcal{L}$ is performed by a PSP rule, i.e. A^l satisfies (2) and (3).

The key property in the analysis of the network case is that, given a fixed opponent profile, a player cannot do better than place *consistent* bids, i.e, the same bid at all the nodes on her path and bid zero on all nodes not in her path.

For each $i \in \mathcal{I}$, we define

$$\begin{array}{rrcl} x_i : & \mathcal{S} & \longrightarrow & \mathcal{S}_i \\ & s & \longmapsto & x_i(s) = (z_i, y_i), \text{ where} \\ & & & z_i^l = 1_{r_i}(l) \min_{m \in r_i} a_i^m(s), \\ & & & y_i^l = 1_{r_i}(l) \max_{m \in r_i} p_i^m, \ 1 \leq l \leq L. \end{array}$$

Lemma 4 *$\forall s \in \mathcal{S}$, and $i \in \mathcal{I}$,*

$$u_i(x_i(s); s_{-i}) \geq u_i(s).$$

Moreover ,

$$s_i \in S(s_{-i}) \Rightarrow x_i(s) \in S(s_{-i}).$$

Proof: First, we prove that, leaving bid prices unchanged, there is no loss of utility for a player who reduces the bid quantities to z_i, i.e.

$$u_i((z_i, p_i); s_{-i}) \geq u_i(s).$$

For any $l \in r_i$, $z_i^l \leq a_i^l(s) \leq Q_i^l(p_i^l, s^l)$, therefore $a_i^l((z_i, p_i); s_{-i}) = z_i^l \wedge Q_i^l(p_i^l, s_{-i}^l) = z_i^l = min_{m \in r_i} a_i^m(s)$. Thus,

$$
\begin{aligned}
& u_i((z_i, p_i); s_{-i}) - u_i(s) \\
&= \theta_i(min_{m \in r_i} a_i^m((z_i, p_i); s_{-i})) - \theta_i(min_{m \in r_i} a_i^m(s)) - c_i((z_i, p_i); s_{-i}) + c_i(s) \\
&= 0 - c_i((z_i, p_i); s_{-i}) + c_i(s) \\
&= \sum_{l \in \mathcal{L}} \int_{a_i^l((z_i, p_i); s_{-i})}^{a_i^l(s)} P_i^l(z, s_{-i})\, dz \\
&\geq 0,
\end{aligned}
$$

since $P_i^l \geq 0$ and $a_i^l((z_i, p_i); s_{-i}) \leq a_i^l(s)$.

Second, for any $l \in r_i$, $y_i^l \geq p_i^l$, hence $Q_i^l(y_i^l, s^l) \geq Q_i^l(p_i^l, s^l) \geq z_i^l$. Thus, $a_i^l((z_i^l, y_i^l); s_{-i}^l) = z_i^l = a_i^l((z_i, p_i); s_{-i})$. Now since $z_i^l = 0$ for $l \notin r_i$, we have

$$
\begin{aligned}
c_i((z_i, y_i); s_{-i}) &= \sum_{l \in \mathcal{L}} \int_0^{a_i^l((z_i, y_i); s_{-i})} P_i^l(z, s_{-i})\, dz \\
&= \sum_{l \in r_i} \int_0^{a_i^l((z_i, y_i); s_{-i})} P_i^l(z, s_{-i})\, dz \\
&= \sum_{l \in r_i} \int_0^{a_i^l((z_i, p_i); s_{-i})} P_i^l(z, s_{-i})\, dz \\
&= c_i((z_i, p_i); s_{-i}),
\end{aligned}
$$

hence

$$u_i((z_i, y_i); s_{-i}) = u_i((z_i, p_i); s_{-i}),$$

which completes the proof of the first statement.

Now $s_i \in S(s_{-i}) \Rightarrow b_i \geq c_i(s) \geq c_i((z_i, y_i); s_{-i}) = c_i((z_i, p_i); s_{-i}) \Rightarrow (z_i, p_i) \in S(s_{-i})$. □

Thanks to Lemma 4, we can restrict our attention to consistent strategies only, and still have feasible best replies[9]. This forms a "consistent" embedded game with feasible sets replaced by the consistent strategy set

$$\tilde{S}_i(s_{-i}) = \{x_i(s) : s_i \in S_i(s_{-i})\}.$$

Just as in Proposition 2, an equilibrium of the embedded game is an equilibrium of the whole game.

[9] If $\theta_i' > 0$, an even stronger statement holds: a bid can be a best reply only if it results in the same quantity allocation at all the nodes in the route.

Define for $\forall y, z \geq 0$, $\forall s \in \mathcal{S}$, $\forall i \in \mathcal{I}$,

$$\begin{aligned}
\tilde{P}_i(z, s_{-i}) &= \sum_{l \in r_i} P_i^l(z, s_{-i}^l), \\
\tilde{Q}_i(y, s_{-i}) &= \sup\{z \in \bigcap_{l \in r_i} \mathcal{Q}^l : \tilde{P}_i(z, s_{-i}) < y\}, \\
\underline{\tilde{Q}}_i(y, s_{-i}) &= \min_{l \in r_i} \underline{Q}_i^l(y, s_{-i}^l), \\
\tilde{a}_i(s) &= q_i^1 \wedge \underline{\tilde{Q}}_i(p_i^1, s_{-i}), \\
\tilde{c}_i(s) &= \int_0^{\tilde{a}_i(s)} \tilde{P}_i(z, s_{-i})\, dz.
\end{aligned}$$

Lemma 5 *$\forall s_{-i} \in \mathcal{S}_{-i}, \forall s_i \in \tilde{S}(s_{-i}), \forall l \in r_i$, $\tilde{a}_i(s) = \min_{l \in r_i} a_i^l(s)$ and $\tilde{c}_i(s) = c_i(s)$.*

Proof: Follows trivially from the definitions and the fact that $s_i \in \tilde{S}_i(s_{-i}) \Rightarrow q_i^l = a_i^l(s) = q_i^1, \forall l \in r_i$. □

Then, from (18), on the feasible set $s_i \in \tilde{S}_i(s_{-i})$,

$$u_i(s) = \theta_i(\tilde{a}_i(s)) - \tilde{c}_i(s).$$

Now we see that, within the feasible sets, the embedded game is identical to the single node game, with all elements being replaced with the ˜ version.

Proposition 4 *(Network Nash equilibrium) In the network auction game with the PSP rule applied independently at each node, reserve prices $p_0^l > 0, \forall l \in \mathcal{L}$, and players described by (18) and (5), if Assumption 1 holds, then for any $\epsilon > 0$, there exists a consistent and truthful ϵ-Nash equilibrium $s^* \in \mathcal{T}$.*

Proof: The proof consists of checking that all the steps leading up to Proposition 2 are valid in the network case.

Proposition 1 relies only on a) the relationship (7)-(8) between P and Q, which also holds for $\tilde{P}$ and $\tilde{Q}$, and b) the fact that $w_i > y_i \Rightarrow \underline{Q}_i(w_i, s_{-i}) \geq Q_i(y_i, s_{-i})$. We want property b) to hold for the ˜ game. Since it holds at each node, i.e. $\underline{Q}_i^l(w_i, s_{-i}^l) \geq Q_i^l(y_i, s_{-i}^l)$, it suffices to show that $\tilde{Q}_i(y_i, s_{-i}) \leq Q_i^l(y_i, s_{-i}^l), \forall l \in r_i$. But

$$\begin{aligned}
z \leq \tilde{Q}_i(y_i, s_{-i}) &\Rightarrow \sum_{l \in r_i} P_i^l(z, s_{-i}) \leq y_i \\
&\Rightarrow P_i^l(z, s_{-i}) < y_i, \forall l \in r_i \\
&\Rightarrow z \leq Q_i^l(y, s_{-i}), \forall l \in r_i
\end{aligned}$$

where the first line follows from (7) (applied tp $\tilde{Q}, \tilde{P}$), and the third from (8) applied to Q^l, P^l. In the second line, the strict inequality holds assuming that $P_i^l(., s_{-i}^l) \geq \underline{P}^l > 0$ for at least two $l \in r_i$. Putting a reserve price $p_0^l > 0$ at each node ensures that. Now in the particular case $z = \tilde{Q}_i(y_i, s_{-i})$, we get the desired property, and the analogue of Proposition 1 holds in the network case. Note that, like in Proposition 1, we get $v_i = \tilde{a}_i(t_i, s_{-i}) = \min_{l \in r_i} a_i^l(t_i, s_{-i})$. But $a_i^l(t_i, s_{-i}) \leq q_i^l = v_i, \forall l$, therefore $t_i \in \tilde{S}_i(s_{-i})$.

For Lemma 1, by summing over l, it can be trivially shown that (14) holds. From that, just as was done in the proof of Lemma 1 to get (14) from (13), using (7) and (8), we can do the reverse and get (13) from (14).

Finally, Lemma 2 depends only on Lemma 1 and (7)-(8), so it holds. □

5 Conclusion

> Auctions are one of oldest surviving classes of economic institutions [...] As impressive as the historical longevity is the remarkable range of situations in which they are currently used. [20]

We proposed the progressive second price auction, a new auction which generalizes key properties of traditional single non-divisible object auctions to the case where an arbitrarily divisible resource is to be shared. We have shown that our auction rule, assuming an elastic-demand model of user preferences, constitutes a stable and efficient allocation and pricing mechanism. Even though we are motivated by problems of bandwidth and buffer space reservation in a communication network, the auction was formulated in a manner which is generic enough for use in a wide range of situations.

In a companion paper [17], we show that the incentive compatibility and equilibrium results can be generalized to a setting where multiple resources are auctioned ny independent resource controllers, with users bidding on arbitrary but fixed routes/topologies.

In our current work, we are considering the case of stochastically arriving players bidding for advance reservations (i.e. resources for a given period of time). Another interesting direction of future work is learning strategies, and evolutionary behavior which can emerge from repeated inter-action between the same players.

6 Acknowledgements

The authors would like to thank Wang Ke for his valuable help in the simulations.

A Information-theoretic basis for the valuation

In general, valuations are simply assumed to be given as external factors. Indeed, the fundamental assumption in any market theory is that buyers know what the goods are worth to them. The "elastic demand" or "diminishing returns" nature of Assumption 1 is fully justified from a purely economic standpoint for virtually all resources in everyday life.

In the case of variable bandwidth, we can go even further by better quantifying what the goods are. For a user sending video, say, how much value is lost when the channel capacity goes from 1.5 to 1.2 Mbps? Ultimately, the value lies not in the amount of raw bandwidth but in the information that is successfully sent. Our goal in this section is to give a brief description of how Information Theory can be used for a bottom-up construction of bandwidth valuations – based on the fundamental thing the user cares about which is communication of information – and that such valuations will generally be of the type in Assumption 1.

Any information source has a function $D(.)$, such that when compressed to a rate R, the signal has a distortion of at least $D(R)$[1]. The distortion is the least possible expected "distance" between the original and compressed signals, where the minimization is over all possible coding/decoding schemes. In this context, we make the distance measure the monetary cost of the error. This cost can be chosen, for example, to be proportional to some common measures like the mean squared error, the Hamming distance (probability of error), the maximum error, etc., or heuristic measures based on experiments with human perception. Given that modern source-coding techniques can, given a distortion measure, achieve distortions close to the theoretical lower bound[5], it is not unreasonable to use the rate-distortion curve as an indication of the value of the bandwidth share.

Let $D_i(.)$ be the distortion-rate function of $\{X_i(t)\}$, a stochastic process modeling the source of information associated with user i. In the most likely auction scenaria, sources would be aggregates of many application streams for which bulk capacity is being purchased, e.g. Virtual Paths, Virtual Private Networks, or edge capacity[2]. X_i is encoded as Y_i which has a rate of R bits per second.

Shannon's channel coding theorem[27] states that Y_i can be received without errors if and only if the channel has a capacity $C > R$. In our auction context, user i has capacity (bandwidth allocation) $C = a_i$, and thus has to suffer a distortion of at least $D_i(a_i)$. The value of the bandwidth is then

$$\theta_i(a_i) = \overline{\theta_i} - D_i(a_i), \tag{19}$$

where $\overline{\theta_i}$ is the value of the full information.

The relevant properties of the distortion-rate functions are:

- when the rate is greater than the entropy of the source, the distortion is zero, and
- for many common source models and cost functions, the distortion-rate function is convex, and has a continuous derivative.

It is easy to see that, with these properties, (19) satisfies Assumption 1.

Example 1: Let $\{X\}$ be a Bernoulli source, taking two values with probabilities p and $1-p$. Without loss of generality, let $p \in [0, 1/2]$. In this case, since the source is i.i.d, one can define the distortion on a per symbol basis. Using a Hamming cost function

$$d(X, Y) = 1_{\{X \neq Y\}},$$

i.e. assuming it costs one unit of money every time one bit is wrong, we have the distortion $D = Ed(X, Y) = P(X \neq Y)$. the rate-distortion function is

$$R(D) = [H(p) - H(D)]^{+},$$

where $H(x) = -x\log(x) - (1-x)\log(1-x)$, and the distortion-rate function is the inverse function. It can be easily seen that $D(R)$ is strictly convex and decreasing for $0 \leq R < H(p)$, and $D(R) = 0$ for $R \geq H(p)$. The continuity of D' on $0 \leq R < H(p)$ and $R > H(p)$ is obvious. At the critical point $(R = H(p), D = 0)$, we have

$$\begin{aligned}
\lim_{R \nearrow H(p)} D'(R) &= \lim_{D \searrow 0} 1/R'(D) \\
&= \lim_{D \searrow 0} 1/\log(D/1-D) \\
&= 0 \\
&= \lim_{R \searrow H(p)} D'(R).
\end{aligned}$$

Thus continuity of D' holds throughout, and Assumption 1 is valid for the valuation of the form (19) for this source – see Figure 9.

Example 2: Let $\{X\}$ be a Gaussian source with Markovian time-dependency, i.e a covariance matrix $\Phi = \left(\sigma^2 r^{|i-j|}\right)_{i,j}$, $r \in [0, 1)$. Suppose we

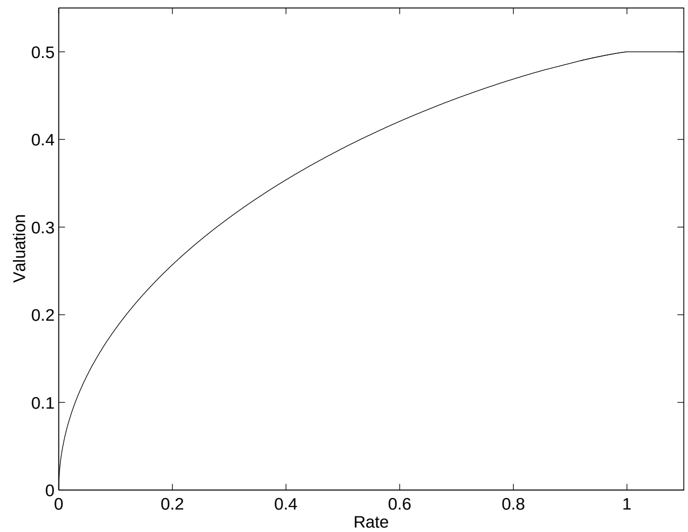

Figure 9: Distortion-rate based valuation for a Bernoulli $p = 1/2$ source

use the squared error cost, i.e. it costs one unit of money for one unit of energy in the error signal. Then, we have for low distortions $D \leq (1-r)/(1+r)$, $R(D) = \frac{1}{2} \log \frac{1-r^2}{D}$, or

$$D(R) = (1 - r^2)\sigma^2 2^{-2R},$$

and Assumption 1 clearly holds for (19). In the i.i.d. case ($r = 0$), the formula holds for all R.

As the source models get more complex, it rapidly becomes impossible to give closed-form expressions for either $R(D)$ or $D(R)$. Often parametric forms are available, and the functions can be evaluated numerically. Fortunately, the convexity property extends to a wide class of models, including for example auto-regressive sources, even when the generating sequence is non-Gaussian (see [1] for a full treatment of $R(D)$, including the above cases).

It can happen, e.g. for some video source models, that the R-D curve, which gives the best (R,D) pairs achievable by any coder/decoder, is not convex. But, for tractability, practical codecs are usually optimized on a convex hull of the space of possible (R,D) pairs[24, 25]. Thus, even when the theoretical D(R) curve is not convex, the actual distortion achieved in real-life systems almost always varies in a convex manner with the available bandwidth.

B Simulation software and bidding algorithm

A prototype software agent based implementation of the auction game, called TREX, has been developed and extensively used since December 1995. Much of the intuition behind the mechanism design and the analysis in this work came from experiments done on this inter-active distributed auction game on the World Wide Web, using the Java programming language [26]. The game can be played in real-time by any number of players from anywhere on the Internet (see Figure 10).

Each user plays in the dynamic auction game using the following:

Algorithm 1 **1** *Let* $s_i = 0$, *and* $\hat{s}_{-i} = \emptyset$. *Start an independent thread which receives updates of* $\hat{s}_{-i}$.

2 *Compute the truthful* ϵ*-best-reply of Proposition 1,* $t_i \in \mathcal{T}_i \cap S_i^{\epsilon}(\hat{s}_{-i})$.

3 *If* $u_i(t_i; \hat{s}_{-i}) > u_i(s_i; \hat{s}_{-i}) + \epsilon$, *then send the bid* $s_i = t_i$.

4 *Sleep for 1 second.*

5 *Go to* **2**.

No assumption is made on the order of the turns. Players join the game at different times, and depending on the execution context of the client program, the sleep time of 1 second is more or less approximate. This, along with communication delays which make the times at which bids arrive at the server and updates at the clients essentially random times, make the distributed game completely asynchronous.

Algorithm 1 can be described as selfish and short-sighted. Selfish because it will submit a new bid if and only if it can improve it's own utility (by more than the fee for the bid). Thus, the game can only converge to an ϵ-Nash equilibrium, if it converges at all. And short-sighted because it does not take the extensive form of the game into account, i.e. does not use strategies which may result in a temporary loss but a better utility in the long run.

References

[1] T. Berger. *Rate Distortion Theory: A Mathematical Basis for Data Compression.* Prentice-Hall, 1971.

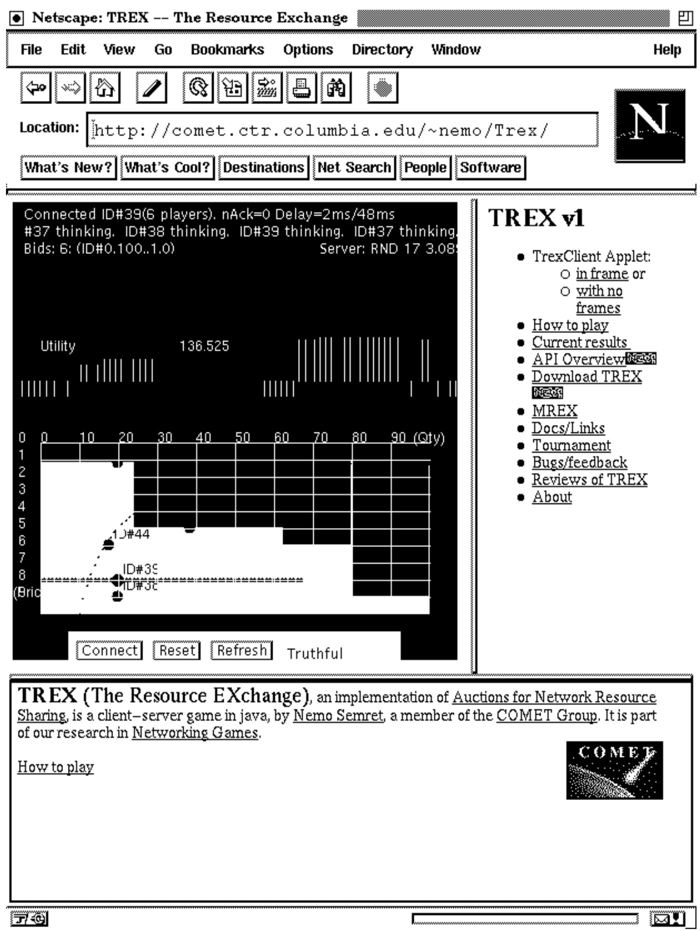

Figure 10: Screen shot of bidding agent in TREX – the prototype on-line PSP auction

[2] D. Clark. Internet cost allocation and pricing. In L. W. McKnight and J. P. Bailey, editors, *Internet Economics*. MIT Press, 1997.

[3] E. H. Clarke. Multipart pricing of public goods. *Public Choice*, 8:17–33, 1971.

[4] D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991.

[5] A. Gersho and R. M Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1992.

[6] J. Gong and S. Marble. Pricing common resources under stochastic demand. preprint – Bellcore, April 1997.

[7] J. Gong and P. Sriganesh. An economic analysis of network architectures. *IEEE Network*, pages 18–21, March/April 1996.

[8] T. Groves. Incentives in teams. *Econometrica*, 41(3):617–631, July 1973.

[9] M. T. T. Hsiao and A. A. Lazar. Optimal decentralized flow control of markovian queueing networks with multiple controllers. *Performance Evaluation*, 13(3), Dec 1991.

[10] Hurewicz and Wallman. *Dimension Theory*. Princeton University Press, 1948.

[11] H. Jiang and S. Jordan. Connection establishment in high speed networks. *IEEE J. Select. Areas Commun.*, 13(7):1150–1161, 1995. http://www.ece.nwu.edu/~scott/conn_pricing.htm.

[12] H. Jiang and S. Jordan. The role of price in the connection establishment process. *European Trans. Telecommunications*, 6(4):421–429, 1995. http://www.ece.nwu.edu/~scott/conn_pricing.html.

[13] E. Kalai and E. Lehrer. Rational learning leads to Nash equilibrium. *Econometrica*, 61(5):1019–1045, September 1993.

[14] F. P. Kelly. Charging and accounting for bursty connections. In L. W. McKnight and J. P. Bailey, editors, *Internet Economics*. MIT Press, 1997.

[15] Y. A. Korilis and A. A. Lazar. On the existence of equilibria in noncooperative optimal flow control. *J. ACM*, 42(3), May 1995.

[16] Y. A. Korilis, A. A. Lazar, and A. Orda. Architecting noncooperative networks. *IEEE J. Select. Areas Commun.*, 13(7), September 1995.

[17] A. A. Lazar and N. Semret. The progressive second price auction mechanism for network resource sharing. In 8^{th} *International Symposium on Dynamic Games and Applications*, Maastricht, The Netherlands, July 1998.

[18] J. K. MacKie-Mason and H. R. Varian. Pricing the internet. In B. Kahin and J. Keller, editors, *Public Access to the Internet.* Prentice Hall, 1994. ftp://ftp.econ.lsa.umich.edu/pub/Papers.

[19] A. Mas-Colell and M. D. Whinston. *Microeconomic Theory.* Oxford University Press, 1995.

[20] P. R. Milgrom. *Advances in Economic Theory: Fifth World Congress*, chapter Auction Theory, pages 1–31. Number 12 in Econometric Society Monographs. Cambridge University Press, 1987.

[21] M. S. Miller and K. E. Drexler. Markets and computation: Agoric open systems. In Bernardo Huberman, editor, *The Ecology of Computation.* Elsevier Science Publishers/North-Holland, 1988.

[22] R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, February 1981.

[23] A. Orda, R. Rom, and N. Shimkin. Competitive routing in multiuser communication networks. *IEEE/ACM Trans. Networking*, 1(5), October 1993.

[24] G. Pacifici, G. Karlsson, M. Garrett, and N. Ohta, editors. *Real-Time Video Services in Multimedia Networks*, volume 15 of *IEEE J. Select. Areas Commun.*, Aug. 1997.

[25] G. M. Schuster and A. K. Katsaggelos. *Rate-Distortion Based Video Compression.* Kluwer Academic Publishers, 1997.

[26] N. Semret. TREX-The Resource EXchange, 1995-96. http://comet.columbia.edu/˜nemo/Trex.

[27] C. E. Shannon. A mathemetical theory of communication. *Bell Syst. Tech. J.*, 27, 1948.

[28] S. Shenker. Making greed work in networks: A game-theoretic analysis of switch service disciplines. In *Proc. ACM SIGCOMM*, 1994.

[29] S. Shenker, D. Clark, D. Estrin, and S. Herzog. Pricing in computer networks: Reshaping the research agenda. *ACM Comput. Commun. Review*, 26(2):19–43, 1996.

[30] W. Vickrey. Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance*, 16, 1961.

[31] S. R. Williams. Realization and Nash implementation: Two aspects of mechanism design. *Econometrica*, 54(1):139–151, January 1986.