

# Optimal Auctions with Positive Network Externalities

Nima Haghpanah  
Northwestern University  
Evanston, IL  
nima.haghpanah@gmail.com

Nicole Immorlica<sup>\*</sup>  
Northwestern University  
Evanston, IL  
nicimm@gmail.com

Vahab Mirrokni  
Google Research  
New York, NY  
mirrokni@gmail.com

Kamesh Munagala<sup>†</sup>  
Duke University  
Durham, NC  
kamesh@cs.duke.edu

## ABSTRACT

We consider the problem of designing auctions in social networks for goods that exhibit *single-parameter submodular network externalities* in which a bidder's value for an outcome is a fixed private type times a known submodular function of the allocation of his friends. Externalities pose many issues that are hard to address with traditional techniques; our work shows how to resolve these issues in a specific setting of particular interest. We operate in a Bayesian environment and so assume private values are drawn according to known distributions. We prove that the optimal auction is APX-hard. Thus we instead design auctions whose revenue approximates that of the optimal auction. Our main result considers *step-function externalities* in which a bidder's value for an outcome is either zero, or equal to his private type if at least one friend has the good. For these settings, we provide a  $\frac{e}{e+1}$ -approximation. We also give a 0.25-approximation auction for general single-parameter submodular network externalities, and discuss optimizing over a class of simple pricing strategies.

## Categories and Subject Descriptors

H.4 [Computer Applications]: Social and Behavioral Sciences—*Economics*

## General Terms

Algorithms, Economics

<sup>\*</sup>Supported by Microsoft New Faculty Fellowship, and by NSF grant SMA-1019169.

<sup>†</sup>Supported by an Alfred P. Sloan Research Fellowship, and by NSF via CAREER award CCF-0745761 and grant CCF-1008065.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

EC'11, June 5–9, 2011, San Jose, California, USA.

Copyright 2011 ACM 978-1-4503-0261-6/11/06 ...\$10.00.

## Keywords

Auctions, Social Networks, Externalities

## 1. INTRODUCTION

Many goods have higher value when used in conjunction with others. A classic example of this phenomenon is the telephone, which clearly has positive value for a consumer only if he or she has people to call. Telephones, and other goods with similar stories, are called *networked goods* and said to exhibit *network effects* or *network externalities*. Modern technology has given birth to a new generation of networked goods. Internet services like email, instant messaging, and online social networks are used primarily to connect with friends and, as such, have strong network externalities. But even more significantly, these services, particularly online social networks, provide platforms upon which developers can generate new applications – applications with very strong networking components. It is now possible to play poker in *Texas HoldEm*, visit cafes in *Cafe World*, and even be a virtual farmer in the immensely popular *FarmVille* with friends in online social networks like Facebook. Such applications are more fun when used with friends, and many such applications even explicitly reward players with many friends. The unique feature of such modern networked goods is that the underlying social network is *explicit*. This enables application distributors to use the network structures to market and sell these goods.

In this paper, we leverage explicit network structure to design mechanisms for selling networked goods. We primarily focus on goods, such as applications like *FarmVille* in online social networks, that are available in unlimited supply or, more precisely, can be produced at zero marginal cost. The network externalities of the good are implied by the private valuations of the social network users. In the most general case, users have a private value for each possible allocation of the good to a subset of users. This allows for arbitrary externalities, enabling say John Doe to value the good only if Kim Kardashian owns it despite having no direct relationship to her. While this makes sense for some goods, like fashion, many network goods like telephones or social network applications have value to a user only if users in his or her immediate neighborhood also own the good. The main focus of the paper is on special case of this sort of direct externality, which we call *step function externalities*: that is, we suppose a user's value for the good is zero unless

at least one of his or her neighbors or friends in the social network is also allocated the good.

We study *auction mechanisms*, or mechanisms that solicit bids from agents indicating their private value for various allocations, and then determine an allocation and prices in a way that maximizes expected revenue. As is common in economics, we work in a Bayesian setting where, while the realization of the private value is known only to the agent, it is drawn according to a commonly known distribution. Most literature on auction design assumes that agents value allocations solely based on the bundle of goods they receive, i.e., they are indifferent for the allocation of the other players. This is clearly violated in settings with externalities. Unfortunately, externalities significantly complicate auction design for the following reasons:

1. The efficient representation of values is no longer a trivial task, since in the most general case each bidder might need to report a value for each subset of allocated bidders.
2. More dimensions make satisfying incentive constraints harder (multi-parameter mechanism design is not well understood).
3. The space of feasible allocations might be more complex, which can make finding the optimal allocation a computationally hard problem.
4. Furthermore, the complexity of the feasible allocation space can easily cause the setting to violate *downward-closure*, i.e., not every subset of a feasible allocation is necessarily feasible. Thus the few known results for multi-parameter mechanism design can not be adopted generically.

We circumvent the first two issues by assuming a special structure on the players' values, namely that valuations satisfy step function externalities as defined above. Thus our problem is a single-parameter one, and so the representation and incentive constraints straight-forward. Revenue maximization is also well understood for single-parameter settings. The seminal paper by Myerson [15] fully characterizes mechanisms that maximize revenue. By this characterization, the expected revenue of any auction is equal to the expected *virtual value* of the allocated agents, where the virtual value of an agent is a function of the valuation and its distribution and may be negative. In our setting, this characterization converts the optimal allocation problem to a combinatorial optimization problem, which is to maximize the sum of the virtual values over all *feasible* subsets. For step-function externalities, the feasibility constraint requires that all allocated agents have a neighbor who is also allocated. Graph-theoretically, this equates to finding, in a vertex-weighted graph with possibly negative vertex weights, a maximum-weight subset of vertices whose induced subgraph has no singleton components.

Although the optimal auction is easy to define, the third and fourth issues of auction design with externalities remain in our setting. We observe via reduction to set-buying that approximating the optimization problem within even a linear factor on *every* sampling of the values is NP-hard. On the other hand, we only need to find algorithms that perform well *in expectation* rather than in worst case. Even on average, we prove that our problem remains APX-hard.

However, we are able to design constant approximations for several versions of the problem.

We first note that there's a simple (1/2)-approximation for our problem. The algorithm divides the graph into two subsets of vertices, such that each vertex in each set has a neighbor in the other. This can be done, for example, by constructing a spanning tree of the graph and then taking a bipartite partitioning of it. The allocation strategy is to then pick the set with better expected revenue, extract revenue from that set, and allocate to all users in the other set in order to maintain feasibility. This very simple algorithm does not use the structure of the social network in any deep way, and is therefore unable to give better approximations in even very simple social networks consisting of a single edge. In order to leverage knowledge of the network structure, we construct a linear program (LP) whose variables represent the allocation and whose constraints use the network structure to characterize feasibility. Our main result shows how to use this LP formulation to give a  $\frac{e}{e+1} \approx 0.73$ -approximation to the optimal revenue for any distribution of values. The interesting observation is that, although the gap of this LP is linear in the number of agents for a particular instantiation of values, we manage to prove a constant approximation through a novel average-case analysis of the rounding technique. We also show that the expected integrality gap of our LP is 0.828, and thereby bound the approximation ratio of any LP-based mechanism. In the full version, we also give a simple greedy-based  $\frac{e}{e+1}$ -approximation for our problem.

We extend our setting to *single-parameter submodular externalities* in which a bidder's value for an outcome is his private value times a known function of the set of players who receive the item. For such settings we study a class of strategies called *influence and exploit* in which some bidders (the *influencers*) are given the good for free and the remainder (the *exploited*) are offered an optimal price conditioned on the set of influencers. We show that the revenue is a submodular function of the set of influencers and hence can use recent submodular function maximization results [7, 8] to design an influence-and-exploit strategy whose revenue is within a 0.41-factor of the optimal influence and exploit strategy. We also show that a randomization over influence and exploit strategies gives a 0.25-approximation to the optimal expected revenue by further submodularity arguments.

**Related Work.** Various settings with positive, negative, or mixed externalities have been studied in economics as well as computer science literature. Rohlfs [16] discusses positive externalities in the telephone industry in which a person's value for a telephone increases as more friends use it. A well-studied scenario with negative externalities is the allocation of ad slots in which, a company's valuation for being listed as one of the *sponsored search* results decreases if their competitor is also listed [1, 4, 9, 10, 13, 14]. Finally, the valuation might have mixed externalities, as in the sale of nuclear weapons [12], in which countries prefer their allies rather than their foes to win the auction. Our work can be viewed as another in this line of literature, which addresses the difficulties of externalities in a specific setting of practical import by making application-specific assumptions.

Our work considers *auctions* with externalities. In contrast, some prior work considers instead posted price mechanisms [3, 2, 5, 11]. Particularly relevant to our work is that of Hartline, Mirrokni, and Sundararajan [11]. They consider the problem of finding a revenue-maximizing sequence of

prices that are offered sequentially to myopic buyers. They observed that simple influence and exploit strategies have revenue within a constant factor of the revenue of any equilibrium of any pricing sequence. They are reminiscent of our auctions which subsidize certain subsets of agents, and also our influence and exploit strategies for general single-parameter submodular externalities. However, unlike Hartline, Mirrokni, and Sundararajan [11], we provide approximation results with regards to the optimal auction revenue, which can be larger than the upper bound in their paper.

## 2. PRELIMINARIES

We consider a society of  $n$  bidders located on the vertices in a social network  $G(V, E)$ , where the undirected edges model *friendship*. We assume for ease of exposition that the social network is connected. There is a limited supply  $k$  of a homogeneous good. Unless otherwise specified, we assume  $k \geq n$  so that the supply is essentially unlimited (equivalently, the good can be reproduced at zero marginal cost).

An outcome  $o \in \Omega = \{0, 1\}^n$  is a distribution of goods among bidders, where  $o_i = 1$  if bidder  $i$  receives a copy of the good and 0 otherwise. Bidder  $i$ 's *type*  $v_i : \Omega \rightarrow \mathbb{R}^+ \cup \{0\}$  maps outcomes to non-negative real numbers, where  $v_i(o)$  represents his value for outcome  $o$  and is positive only if he receives a copy of the good (i.e.,  $o_i = 1$ ). We study Bayesian mechanism design, in which one assumes that each type  $v_i(\cdot)$  is drawn independently from a commonly-known distribution  $F_i$ . Let  $F = F_1 \times \dots \times F_n$  be the product distribution of  $F_i$  for all  $i$ ;  $v$  be the vector of types, called the *type profile*;  $v_{-i}$  be the vector of types of agents other than  $i$ , and  $F_{-i}$  the distribution of  $v_{-i}$ . Throughout the paper, our algorithms assume access to expectations defined with respect to the distribution  $F$ . We assume these can be computed to within sufficient accuracy via sampling.

A (direct) mechanism is specified by two functions  $\chi : \mathbb{R}^{n2^n} \rightarrow \Omega$  and  $\rho : \mathbb{R}^{n2^n} \rightarrow \mathbb{R}^n$  in which  $\rho_i(v)$  is the *payment* of agent  $i$  given the reported type profile  $v$ . The utility of an agent for outcome  $o$  and price  $p$  is his value for the outcome minus the price he pays,  $v_i(o) - p$ . We say that a mechanism  $(\chi, \rho)$  is Bayesian incentive compatible (BIC) if reporting the true type maximizes any player  $i$ 's expected utility assuming that other players also report their true types, that is for every agent  $i$  and types  $v_i$  and  $v'_i$ ,

$$\begin{aligned} & \mathbf{E}_{v_{-i} \sim F_{-i}} [v_i(\chi(v)) - \rho_i(\chi(v))] \\ & \geq \mathbf{E}_{v_{-i} \sim F_{-i}} [v_i(\chi(v'_i, v_{-i})) - \rho_i(\chi(v'_i, v_{-i}))]. \end{aligned}$$

Note that this is an interim notion, i.e., the agent's choose the strategy that gives them the highest expected utility after observing their own private value. Similarly, we assume an interim notion of *individual rationality*, i.e., each agent's expected utility conditioned on their private value should be non-negative.

We consider *single-parameter* settings. In these settings, agents values are a function of just one private parameter, called their *type*. As types are represented by a single parameter,  $v_i$ , the Bayesian assumption reduces to assuming that  $v_i$  is drawn independently from a distribution  $F_i$  over the non-negative reals, henceforth referred to as the type distribution of bidder  $i$ . We assume type distributions are *regular* and hence the corresponding *virtual values* are non-decreasing (see Subsection 2.1 for definitions).

In the following subsection, we discuss optimal auction design for single-parameter settings. We encourage the reader familiar with these subjects to skip to Subsection 2.3 where we define the problem studied in this paper.

### 2.1 Optimal Auction Characterization

In his seminal paper [15], Myerson characterized the revenue of the optimal (i.e., revenue-maximizing) auction in terms of the *virtual values* of the agents. We first define virtual values and then discuss the characterization result.

**DEFINITION 1.** *Suppose type  $v_i$  is drawn independently from a distribution  $F_i(v) = \Pr[v_i \leq v]$  and let  $f_i(v)$  be the density function of the distribution (note that if the distribution is discrete then  $f_i(v)$  is simply  $\Pr[v_i = v]$ ). Then the virtual value function  $\phi_i(v)$  is  $v - \frac{1 - F_i(v)}{f_i(v)}$ .*

Note that virtual values may be negative. However, they are non-negative in expectation, a fact which enables many of our results.

**FACT 1.** *For any distribution  $F$  and value  $v$ , the virtual value  $\phi(v)$  is non-negative.*

We will further assume that the distributions we study are *regular*, meaning that the corresponding virtual value function is non-decreasing in the support of  $F$ .

For a mechanism  $(\chi, \rho)$ , let  $x_i(v) = v_i(\chi(v))/v_i$  if  $v_i > 0$ , and zero otherwise. In Myerson's characterization, it is the function  $x$  that is relevant for determining the revenue of the mechanism, and hence in a slight abuse of terminology we will refer to  $x$  as the *allocation function* even though there may be bidders  $i$  with  $x_i(v) = 0$  that receive copies of the good (however they do not value the copy of the good because of the externalities). Accordingly define  $x_i(v_i) = \mathbf{E}_{v_{-i} \sim F_{-i}} [x_i(v_i, v_{-i})]$  to be agent  $i$ 's expected allocation for type  $v_i$ , where the expectation is over the types of other players.

In the single-parameter setting with regular distributions, Myerson showed that for any monotone increasing rule  $x$ , there is a unique corresponding payment rule  $\rho$  such that the resulting mechanism  $(\chi, \rho)$  is BIC (where  $\chi$  is any function that induces allocation function  $x$  and is not necessarily unique). The expected revenue of the mechanism is equal to its expected virtual value,  $\sum \mathbf{E}_{v_i \sim F_i} [x_i(v_i) \phi_i(v_i)]$ . Furthermore, if  $x$  is not monotone increasing, then there is no payment rule that makes the corresponding mechanism BIC. Restricting attention to BIC mechanisms is without loss of generality due to the revelation principle, and so to maximize revenue, one simply needs to find a rule  $\chi$  satisfying all exogenous constraints (e.g., limited supply) whose corresponding feasible allocation function  $x$  is monotone and maximizes expected virtual value. We can therefore analyze the revenue of any monotone mechanism without explicitly defining the prices.

### 2.2 Submodular Externalities

In this setting, we assume that each player  $i$  is assigned *local influence function*  $g_i : 2^V \rightarrow \mathbb{R}$  which is common knowledge. Following the previous literature on network influence, we assume that this local influence function is submodular for each player, i.e.,  $g_i(S \cup \{j\}) - g_i(S) \geq g_i(S' \cup \{j\}) - g_i(S')$ , for all  $S \supseteq S'$ , and  $j \notin S$ . Submodularity models the

diminishing return properly and makes sense in this context. Without loss of generality assume  $g_i$  is normalized such that  $g_i(V) = 1$ . Given this function and  $i$ 's type  $v_i$ , define  $S(o) = \{j : o_j = 1\}$  to be the set of players that are given the item in an outcome  $o$ . Then the value of  $i$  for  $o$  is defined to be  $v_i(o) = v_i \cdot g_i(S(o))$ .

Given any mechanism  $(\chi, \rho)$ , define the *allocation function*  $x_i(v) = g_i(S(\chi(v)))$ . Using this definition we can write  $v_i(\chi(v)) = v_i x_i(v)$ . Now we can invoke Myerson's characterization and write the expected revenue of the mechanism to be  $\sum_i E[x_i(v) \phi_i(v_i)]$ .

We consider two special cases of submodular externalities: *concave externalities* and *step-function externalities*. In concave externalities, for each player  $i$  and subset  $S$ ,  $g_i(S) = \mathcal{G}(|S \cap N_i|)$  for some concave function  $\mathcal{G}(\cdot)$  if  $i \in S$ . That is, the valuation of each bidder  $i$  depends on the number of his neighbors who have the good but not their identity, and also the local influence function is the same among all players. Step-function externalities are detailed in the next section.

### 2.3 Step-Function Externalities

Step-function externalities are a special case of submodular externalities in which the value of the influence function is 0 if the set of neighbors who receive the item is empty, and 1 otherwise. Let  $N(i)$  be the neighborhood of  $i$  in  $G$ ,  $N(i) = \{j : (i, j) \in E\}$ . Formally, bidder  $i$ 's value for an outcome is:

$$v_i(o) = \begin{cases} v_i & : o_i = 1, |j \in N(i), o_j = 1| \geq 1 \\ 0 & : \text{otherwise} \end{cases}$$

We say that a bidder  $i$  is *satisfied* by an allocation if  $v_i(o) = v_i$ . This models applications, like bridge tournaments, that require just one friend to be of value.

In this setting, for any mechanism  $(\chi, \rho)$ , we have  $x_i(v) = 1$  if outcome  $\chi(v)$  satisfies bidder  $i$  and zero otherwise. Note that  $\chi$  and  $x$  may differ on certain bidders  $i$ . In particular, it can be that outcome  $o = \chi(v)$  gives a good to bidder  $i$  (i.e.,  $o_i = 1$ ) and yet does not satisfy the bidder (i.e.,  $o \notin O_i$ ) so that  $x_i(v) = 0$ . As a result, allocation functions  $x$  must satisfy the condition that  $x_i(v) = 1$  only if for at least one (more generally,  $s$ ) neighbor  $j \in N(i)$  of  $i$ , we also have  $x_j(v) = 1$ . This means that in the subgraph induced by the allocated agents, every vertex must have degree at least 1 (more generally,  $s$ ). Call such a subset of agents *feasible*. By the Myerson characterization discussed above, the optimal auction is thus specified by an allocation function that, given a type profile, allocates to a feasible subset of agents with maximum sum of virtual values (note this rule is necessarily monotone).

Graph-theoretically, the problem of finding an optimal allocation function equates to finding a subset of vertices of maximum weight whose induced subgraph has no isolated vertices. Unfortunately, we show in Section 3 that this problem is more general than the *set buying* problem, and therefore hard to approximate within a linear factor on every sampling of values. We also show that the problem of maximizing the *expected* revenue (over randomness of values), is APX-hard.

As the problem is NP-hard to solve optimally, we instead design a polynomial-time monotone allocation function whose expected revenue (as defined by the sum of virtual values) is close to the optimal expected revenue  $OPT$ ,

where the expectations are over the type distributions. We say an auction is an  $\alpha$ -*approximation* if its expected revenue is at least  $\alpha \times OPT$ .

## 3. HARDNESS

By Myerson's characterization of optimal allocations, the problem of finding an optimal allocation function equates to finding a subset of vertices of maximum weight whose induced subgraph has no isolated vertices. Unfortunately, since virtual values and hence vertex weights might be negative, this problem is more general than the set buying problem (see, e.g., Feige et. al. [6]). We prove this formally in lemma3. We next show that the problem of maximizing the expected, over the randomness of values, revenue is APX-hard, and therefore does not admit a PTAS unless  $P=NP$ , which justifies the search for constant factor approximations to the problem in later sections. The reduction is from a special case of set buying, which we call the *prize collecting set cover problem* (PCSCP).

**DEFINITION 2.** *A set buying instance is specified by a set of elements  $U$  and a collection  $\mathcal{F}$  of subsets of  $U$ . There is a non-negative cost  $c(S)$  associated with each set  $S \in \mathcal{F}$ , and a non-negative value  $v(u)$  associated with each element  $u \in U$ . The problem to pick some subsets  $\mathcal{S} \subseteq \mathcal{F}$  to maximize the value of the elements covered by those sets minus the total cost of those sets, that is  $\sum_{u \in \text{Span}(\mathcal{S})} v(u) - \sum_{S \in \mathcal{S}} c(S)$ , where  $\text{Span}(\mathcal{S}) = \cup_{S \in \mathcal{S}} S$ .*

**THEOREM 2** (FEIGE ET AL [6]). *It is NP-hard to approximate the set-buying problem to within a linear factor.*

**LEMMA 3.** *The optimal auction with step-function externalities is NP-hard to approximate to within a linear factor on every instantiation of values.*

**PROOF.** For any instance  $\mathcal{I} = (U, \mathcal{S})$  of the set-buying problem we construct a bipartite graph  $G_{\mathcal{I}} = ((L, R), E)$  with a vertex  $l_u \in L$  for each  $u \in U$  and a vertex  $r_S \in R$  for each  $S \in \mathcal{F}$ . We introduce an edge  $(l_u, r_S) \in E$  for any element  $u$  and set  $S$  such that  $u \in S$ .

Consider an instance  $\mathcal{I}$  and social network defined by the corresponding bipartite graph  $G_{\mathcal{I}}$ . Let the type of bidder  $l_u$  be  $v(u)$  with probability 1 and the type of bidder  $r_S$  be  $-c(S)$  with probability 1. Note these point mass distributions are in fact regular, and furthermore a subset of bidders with maximum sum of virtual values directly corresponds to an optimal solution to the set-buying instance. The first claim follows from the inapproximability of set buying.  $\square$

The PCSCP is a special case of set buying problem in which all set and all elements have equal costs and values, respectively. We first give an approximation preserving reduction from PCSCP to our problem, and then prove that PCSCP is APX-hard, which using the approximation preserving reduction implies the APX-hardness of our problem.

**DEFINITION 3.** *In the prize collecting set cover problem, there is a collection of  $n$  sets  $\{S_1, S_2, \dots, S_n\}$  over a universe  $U$ . For a collection  $C$  of sets, let  $Q_C = \cup_{i \in C} S_i$ . The goal is to find a collection  $C^*$  that maximizes  $\alpha |Q_{C^*}| + n - |C^*|$  for some  $\alpha > 0$ .*

Note that this is equivalent to maximizing the value of covered elements plus the cost of unused sets. While this is

equivalent, in optimality, to maximizing the value of covered elements minus the cost of the used sets, the two problems differ in approximability. The PCSCP is easier to approximate: there is a  $e/(e+1)$ -approximation for it. On the other hand, set-buying is not approximable to within a linear factor. We will show an approximation-preserving reduction from PCSCP to our problem and then give an  $e/(e+1)$ -approximation for our problem.

LEMMA 4. *There is an approximation preserving reduction from the prize collecting set cover problem to our problem.*

PROOF. Given an instance of the prize collecting set cover problem, where the sets are denoted  $\{S_1, S_2, \dots, S_n\}$  and the elements are denoted  $e_1, e_2, \dots, e_m$ , we construct a graph where there is a vertex for each set and each element, and an edge between  $S_i$  and  $e_j$  if  $e_j \in S_i$ . For each element  $e_j$ , the value is deterministic  $\alpha$ . Let  $L \gg mn\alpha$ . For each set  $S_i$ , the valuation follows distribution Bernoulli( $L, 1/L$ ), so that the virtual valuation is  $-1$  w.p.  $1 - 1/L$  and  $(L - 1)$  w.p.  $1/L$ .

To compute the revenue, we let  $L \rightarrow \infty$ . There are two events:

1. If at least one set has positive virtual valuation (which happens w.p. approximately  $n/L \rightarrow 0$ ), the solution chooses all such sets to obtain expected revenue  $n$ . The expected revenue from the remaining nodes (sets with negative virtual valuation and elements) is at most  $\alpha m/L \ll 1$ . Therefore, the optimal solution has contribution  $n$  from this event as  $L \rightarrow \infty$ , and this solution is trivial to compute.
2. If no set has positive virtual valuation (which happens w.p.  $1 - n/L \rightarrow 1$ ), the value of the solution is precisely  $\alpha|Q_{C^*}| - |C^*|$ , and this is the contribution from this event.

Therefore, the value of the optimal revenue solution is  $\alpha|Q_{C^*}| + n - |C^*|$  as  $L \rightarrow \infty$ , and this completes the reduction.  $\square$

THEOREM 5. *The prize collecting set cover problem (PCSCP) is APX-complete.*

PROOF. We start with a 4-regular graph. On such a graph with  $n = 152k$  nodes, for any  $\epsilon > 0$ , it is NP-HARD to decide if there is an independent set of size at least  $(74 - \epsilon)k$  or at most  $(73 + \epsilon)k$ .

Given this graph, construct the following prize collecting set cover instance: There is a set for every vertex, and an element for every edge.  $\alpha$  is set to  $1/3$ . It is easy to check that any feasible solution will correspond to choosing an independent set of vertices, and the incident edges. The goal becomes to maximize  $\frac{1}{3}(3n + I)$ , where  $I$  is the size of an independent set. Therefore, the optimal solution has value  $(530 - \epsilon)k/3$ , and it is NP-HARD to find a solution of value  $(529 + \epsilon)k/3$ , so that the approximation ratio is at least  $\frac{530}{529} \approx 1.002$ .  $\square$

COROLLARY 6. *The problem of maximizing the expected revenue is APX-hard.*

## 4. STEP-FUNCTION EXTERNALITIES

Although the optimal auction is NP-hard to compute and NP-hard to approximate on *every* instantiation of values, it is in fact easy to approximate on average. The following very simple allocation function has expected revenue within a factor  $1/2$  of the optimal expected revenue. In Appendix A, we show that this can be generalized to a  $(1/4)$ -approximation for the limited-supply setting.

Divide vertices into two sets  $S_0$  and  $S_1$  such that each vertex  $i \in S_0$  (respectively  $S_1$ ) has a neighbor in the opposing set  $S_1$  (respectively  $S_0$ ). Note that this can be done efficiently, e.g. by computing a spanning tree of  $G$  and considering an arbitrary 2-coloring of it. Suppose  $S_0$  has higher expected positive virtual value, that is,

$$\sum_{i \in S_0} E[\max(\phi_i(v_i), 0)] \geq \sum_{i \in S_1} E[\max(\phi_i(v_i), 0)].$$

Let  $S_0^+ = \{i \in S : \phi(v_i) \geq 0\}$  be the bidders with positive virtual value in set  $S_0$  for a particular instantiation of values. Then give a copy of the good to every bidder  $i \in S_0^+ \cup S_1$ .<sup>1</sup>

To see that this is a  $(1/2)$ -approximation, note that the expected optimum revenue is at most  $\sum_i E[\max(\phi_i(v_i), 0)]$  since at best an auction can extract  $\phi_i(v_i)$  from all bidders  $i$  with positive virtual value. The above strategy gets expected revenue  $\sum_{i \in S_0^+} E[\max(\phi_i(v_i), 0)]$  from the bidders in  $S_0^+$ , which is at least half the optimum expected revenue by linearity of expectation and our choice of  $S_0$ . Although a copy of the good is always given to any  $i \in S_1$ ,  $i$  is not always satisfied. In particular, for  $i$  to be satisfied it needs to have a neighbor in  $S_0^+ \cup S_1$ . But note that  $i$ 's expected allocation is independent of its value, i.e., we have  $x_i(v_i) = x_i$  for some constant  $x_i$ . As a result, the revenue from  $i$  is  $E[x_i \phi_i(v_i)] = x_i E[\phi_i(v_i)]$ . Thus, since the expected virtual value of any bidder is non-negative (see Fact 1), the expected revenue of bidders in  $S_1$  is non-negative.

Further note that this analysis is tight, as shown by the simple example of a single edge whose endpoints have value 1 with probability  $p$  and 0 with probability  $1 - p$  for some  $0 < p < 1$ . Then the virtual value is 1 with probability  $p$  and  $\frac{-p}{1-p}$  with probability  $1 - p$ . Consider the strategy in which we allocate both nodes when at least one of them has positive value. The value that we get is  $2p^2 + 2p(1-p)(1 + \frac{-p}{1-p}) = 2p^2 + 2p(1-2p) = 2p - 2p^2$ . The  $(1/2)$ -approximation algorithm described above gets expected value  $p$ . The ratio of the algorithm's solution to the given strategy goes to  $1/2$  as  $p \rightarrow 0$ .

There are two reasons why this simple algorithm can not guarantee better than a 0.5-approximation of the optimum solution. First, the current upper bound that we are using is quite loose. Second, the current algorithm gives the good to a bidder solely based on that bidder's value, without considering the values of its neighbors. We can generalize this algorithm to a broader class of *threshold strategies* in which,

<sup>1</sup>A plausibly more natural strategy would be to select set  $S_0$  based on the instantiation of the virtual values rather than their expectation. Such a strategy is still monotone, but bounding its revenue is more complex since the virtual value of  $S_1$  conditional on the selection is now less than zero.

there exists a threshold assigned to each player and he receives the item when his value passes that threshold. Note that the simple algorithm falls in that class since each player is either assigned the threshold of zero, or  $\phi^{-1}(0)$ . We show in the full version that no threshold strategy can have better than 0.5 approximation.

In order to improve this approximation ratio, we need to leverage detailed knowledge of the graph structure. In the remainder of this section, we present a linear-programming-based approach that gets a 0.73-approximation for general distributions. In the full version, we also provide a simple greedy-based algorithm that gets this same approximation ratio.

## 4.1 Linear Program

As discussed above, the main hurdle in the analysis of the simple auction schemes was the loose upper bound on the optimal expected revenue. In this section, we significantly improve the upper bound using linear programs whose constraints characterize the feasible allocation rules. We then use these LPs to bound the expected revenue of an LP-based auction. As a result, we are able to get a better approximation ratio and also generalize the setting to non-identical distributions.

Recall that for each profile of types  $v$  with virtual valuation functions  $\{\phi_i(\cdot)\}$ , the optimal revenue is equal to the maximum sum of virtual values among feasible allocations. In step-function externalities, an allocation  $x(\cdot)$  is feasible if each vertex  $i$  with  $x_i(v) = 1$  had a neighbor  $j$  with  $x_j(v) = 1$ . Hence we can write the following LP relaxation of the optimum revenue:

$$\begin{aligned} \max_x \quad & \sum_i x_i(v) \phi_i(v_i) \\ \text{s.t.} \quad & x_i(v) \leq \sum_{j \in N(i)} x_j(v) \quad \forall i \\ & 0 \leq x_i(v) \leq 1 \quad \forall i. \end{aligned} \quad (1)$$

Each instantiation of types induces one such LP. As discussed in Section 3, given the instantiation of types, our problem is more general than the set-buying problem studied in Feige et. al. [6] whose LP-relaxation is shown to have linear gap. Hence the LP value might seem like a very loose upper bound. However, recall that we only require our auction to have close-to-optimal revenue *on average*. In other words, we need a rounding scheme whose expected value, over the distribution of LPs induced by the type distributions, is close to the expected value of the LPs. Thus we can perform poorly on hard instances so long as we do well on average, and so LPs with linear worst-case integrality gaps might still be useful in designing an LP-based auction with constant approximation ratios.

## 4.2 Auction Scheme

We present an LP-based auction whose expected value is within a constant factor of the expected LP value. Note the LP gets positive contributions from two types of nodes: those with positive virtual value who also have a neighbor with positive virtual value, and those with positive virtual value whose neighbors all have negative virtual value. Our auction first estimates the relative contributions of these two types and then tailors its strategy accordingly. In the extremes, where one type contributes most of the revenue, a simple deterministic scheme has a good approximation.

When the contributions are more-or-less equal, we use a randomized rounding of the LP solution.

To define the auction, we first introduce some notation to capture the LP contribution from the types discussed above. For an instantiation of values  $v$ , let  $x_i^*(v)$  be the optimum LP value of variable  $x_i^*(v)$ . Recall that the LP for any instantiation  $v$  bounds the optimal revenue for that instantiation and hence the optimal expected revenue is at most  $E_v[\max_x \sum_i x_i(v) \phi_i(v_i)] = E_v[\sum_i x_i^*(v) \phi_i(v_i)]$ . Fix a player  $i$  and define the following events:

- $P_i^+$  is the event that  $\phi_i(v_i) \geq 0$  and there exists  $j \in N(i)$  such that  $\phi_j(v_j) \geq 0$ .
- $P_i^-$  is the event that  $\phi_i(v_i) \geq 0$  and all neighbors of  $i$  have negative virtual value.
- $N_i$  is the event that  $\phi_i(v_i) < 0$ .

Observe that the expected revenue of the optimum allocation from agent  $i$  can be written as

$$\begin{aligned} E_v[x_i^*(v) \phi_i(v_i)] &= E_v[x_i^*(v) \phi_i(v_i) | P_i^+] Pr(P_i^+) \\ &\quad + E_v[x_i^*(v) \phi_i(v_i) | N_i] Pr(N_i) \\ &\quad + E_v[x_i^*(v) \phi_i(v_i) | P_i^-] Pr(P_i^-) \end{aligned}$$

Define

$$A_i^* = E_v[x_i^*(v) \phi_i(v_i) | P_i^-] Pr(P_i^-),$$

$$B_i^* = E_v[x_i^*(v) \phi_i(v_i) | P_i^+] Pr(P_i^+),$$

and

$$C_i^* = E_v[x_i^*(v) \phi_i(v_i) | N_i] Pr(N_i)$$

(note  $C_i^*$  is negative). The auction works as follows.

**LP-Based Auction.** Compute  $A^* = \sum_i A_i^*$ ,  $B^* = \sum_i B_i^*$ , and  $C^* = \sum_i C_i^*$ , where the expectation is computed via sampling, and let  $R = (1 - 1/e)A^* + B^* + C^*$ . Solve LP 1 for the instantiation of types  $v$  and let  $x^*$  be an optimal solution. If

- $A^* \geq \max(B^*, R)$ , allocate to all the nodes  $i$  for which  $\phi_i(v_i) \geq 0$  as well as all nodes  $i$  for which  $N_i$  happens *and* the LP value  $x_i^*(v) > 0$ ,
- else if  $B^* \geq \max(A^*, R)$ , allocate to all the nodes for which  $P_i^+$  happens.

Otherwise

1. For each  $i$  with  $\phi_i(v_i) < 0$ , give  $i$  a copy of the good with probability  $x_i^*(v)$ .
2. For each  $i$  with  $\phi_i(v_i) \geq 0$ , give  $i$  a copy of the good if it has a neighbor  $j$  that either
  - (a) has non-negative virtual value  $\phi_j(j) \geq 0$ , or
  - (b) has negative virtual value  $\phi_j(j) < 0$  and received a copy of the good in step 1.

## 4.3 Analysis

The proof of the approximation guarantee requires two technical lemmas that appear at the end of this section. The first lemma shows that the contribution of a node  $i$  when  $P_i^+$  happens outweighs his contribution when  $N$  happens (for

reasonable allocation rules). The second lemma provides the key argument for the analysis of the randomized rounding step: the LP constraints corresponding to an agent  $i$  with positive virtual value must be tight in an optimal solution  $x^*(v)$ . Namely,  $x_i^*(v) = \min(1, \sum_{j \in N(i)} x_j^*(v))$ . Hence to round and get constant contribution from these agents, we can round the nodes with negative virtual value with probability equal to their LP values and then round nodes with positive virtual value to one if some neighbor was rounded to one. To bound the expected allocation of such an agent  $i$  in the rounding, we note that in the worst-case all neighbors of  $i$  have negative virtual value. However, even in this case,  $i$  is allocated so long as at least one  $j \in N(i)$  receives the good. This happens with probability  $x_j^*$  for neighbor  $j$  and so the allocation probability of  $i$  from the rounding scheme is at least  $1 - \prod_{j \in N(i)} (1 - x_j^*)$ . This is within a  $(1 - 1/e)$  fraction of  $x_i^*$ .

**THEOREM 7.** *The LP-based auction is a  $e/(e+1) \approx 0.73$  approximation.*

**PROOF.** For an instantiation of values  $v$ , let  $x_i(v)$  be the expected allocation of  $i$  (over the randomization of the auction). Correspondingly, define  $A_i, B_i, C_i$  and  $A, B, C$  for the auction's allocation function as we did for the LP. There are three cases.

If  $B^* \geq \max(A^*, R)$ , then  $x_i(v) \in \{0, 1\}$  and is 1 only for those  $i$  for which  $P_i^+$  happens. In the LP,  $x_i^*(v)$  also equals 1 for all such  $i$  and hence the revenue of the auction  $B = B^*$ .

If  $A^* \geq \max(B^*, R)$ , then  $x_i(v) = 1$  for all  $i$  for which  $\phi_i(v_i) \geq 0$  so  $A \geq A^*$ . Lemma 8 further shows that  $B + C \geq 0$  and so the total revenue of the auction is at least  $A \geq A^*$ .

Finally, if  $R \geq \max(A^*, B^*)$ , then Lemma 9 shows  $A \geq (1 - 1/e)A^*$ . Furthermore, from the construction of  $x$  we see that conditioned on  $P^+$  and  $N$ ,  $x$  and  $x^*$  are equal so  $B + C = B^* + C^*$ . Therefore the total revenue of the auction is at least  $(1 - 1/e)A^* + B^* + C^*$ .

The optimal expected revenue is at most  $A^* + B^* + C^*$ , and so the approximation ratio of the auction is at least

$$\min \frac{\max(A^*, B^*, (1 - 1/e)A^* + B^* + C^*)}{A^* + B^* + C^*}.$$

For computing the above minimum, normalize  $A^* = 1$  and suppose  $B^* = x$  and  $B^* + C^* = rx$  for  $0 \leq r \leq 1$  (such  $r$  exists since  $B^* + C^* \geq 0$  by Lemma 8 and  $C^* \leq 0$ ). Thus we want to compute the minimum of  $\max(1, x, 1 - 1/e + xr)/(1 + xr)$  where  $0 \leq r \leq 1$ . We can do a case analysis on the maximum:

1.  $xr \leq 1/e$ . Then, we are minimizing  $\max(1, x)/(1 + xr)$ . We can set  $xr = 1/e$ , so that the lowest possible value is  $e/(e+1)$ .
2.  $xr \geq 1/e$  and  $x(1-r) \leq 1 - 1/e$ . Then we have  $(1 - 1/e + xr)/(1 + xr)$ . Setting  $xr = 1/e$  implies  $e/(e+1)$ .
3.  $x \geq 1$  and  $x(1-r) \geq 1 - 1/e$ . Then we have  $x/(1 + xr)$ . But  $xr \leq x + 1/e - 1$ , so that we are minimizing  $x/(x + 1/e)$  for  $x \geq 1$ , so that we again have  $e/(e+1)$ .

Thus the approximation ratio is  $e/(e+1) \approx 0.73$ .  $\square$

**LEMMA 8.** *For any monotone non-decreasing allocation function  $x$  that allocates to nodes  $i$  with  $\phi_i(v_i) < 0$  only if there is a neighbor  $j$  with  $\phi_j(v_j) \geq 0$ , and corresponding  $B, C$ , we have  $B - C \geq 0$ .*

**PROOF.** We prove the inequality for each node  $i$  separately. Let  $N(i)$  be the neighborhood of  $i$  and note that:

$$\begin{aligned} B_i - C_i &= E_{v_i}[x_i(v_i)\phi_i(v_i)|P^+]Pr(P^+) \\ &\quad + E_{v_i}[x_i(v_i)\phi_i(v_i)|N]Pr(N) \\ &= E_{v_i}[x_i(v_i)\phi_i(v_i)|P^+]Pr(P^+) \\ &\quad + (E_{v_i}[x_i(v_i)\phi_i(v_i)|N, \exists j \in N(i), \phi_j(v_j) \geq 0] \\ &\quad \cdot Pr(\exists j \in N(i), \phi_j(v_j) \geq 0)Pr(N) \\ &\quad + E_{v_i}[x_i(v_i)\phi_i(v_i)|N, \forall j \in N(i), \phi_j(v_j) < 0] \\ &\quad \cdot Pr(\forall j \in N(i), \phi_j(v_j) < 0))Pr(N). \end{aligned}$$

But by assumption conditioned on  $N$  and the event  $[\forall j \in N(i), \phi_j(v_j) < 0]$ ,  $x_i^*(v_i) = 0$ , and therefore, letting  $E$  be the event  $[\exists j \in N(i), \phi_j(v_j) \geq 0]$ , we have

$$\begin{aligned} B - C &= E_{v_i}[x_i(v_i)\phi_i(v_i)|P^+]Pr(\phi_i(v_i) \geq 0)Pr(E) \\ &\quad + E_{v_i}[x_i(v_i)\phi_i(v_i)|N, E]Pr(N)Pr(E) \\ &= (E_{v_i}[x_i^*(v_i)\phi_i(v_i)|P^+, E]Pr(\phi_i(v_i) \geq 0) \\ &\quad + E_{v_i}[x_i^*(v_i)\phi_i(v_i)|N, E]Pr(N))Pr(E) \\ &= E_{v_i}[x_i^*(v_i)\phi_i(v_i)|E]Pr(E) \\ &\geq 0, \end{aligned}$$

where the second equality follows because the event  $P_i^+$  implies event  $E$  and the last inequality follows because  $x^*(v_i)$  is a monotone non-decreasing function of  $v_i$  as  $\phi(\cdot)$  is regular and also that  $E_{v_i}[\phi_i(v_i)] = 0$  (see Fact 1).  $\square$

**LEMMA 9.** *If  $R \geq \max(A^*, B^*)$ , then  $A \geq (1 - 1/e)A^*$ .*

**PROOF.** First note that conditioned on event  $P^-$ , we have  $x_i^*(v) = \min(1, \sum_{j \in N(i)} x_j^*(v))$ , and  $x_i(v) = 1 - \prod_{j \in N(i)} (1 - x_j^*(v))$ . Let  $y = \sum_{j \in N(i)} x_j^*(v)$  and  $d = |N(i)|$ . Fixing the value of  $\sum_{j \in N(i)} x_j^*(v)$ , the minimum of  $1 - \prod_{j \in N(i)} (1 - x_j^*(v))$  happens when all the variables are equal, in which case we have  $x_i(v) = 1 - (1 - \frac{y}{d})^d \geq 1 - \frac{1}{e^y}$ . Thus when  $y \leq 1$ , we have  $x_i^*(v) = y$  and so  $\frac{x_i(v)}{x_i^*(v)}$  is at least  $\frac{1 - e^{-y}}{y}$ , whose minimum value is equal to  $1 - \frac{1}{e}$ . When  $y \geq 1$ , we have  $x_i^*(v) = 1$  and so  $\frac{x_i(v)}{x_i^*(v)}$  is at least  $1 - e^{-y}$ , whose minimum value is again  $1 - \frac{1}{e}$ . Therefore we have

$$\begin{aligned} A_i &= E_{v_i}[x_i(v_i)\phi_i(v_i)|P^-]Pr(P^-) \\ &\geq (1 - \frac{1}{e})E_{v_i}[x_i^*(v_i)\phi_i(v_i)|P^-]Pr(P^-) \\ &= (1 - \frac{1}{e})B_i^* \end{aligned}$$

Summing over  $i$  yields the result.  $\square$

## 4.4 Integrality Gap

We now prove that the above LP has integrality gap at most 0.828. This means that we can not use the LP solutions as an upper bound in order to get approximation guarantees better than 0.828. We show the gap by proving the gap on the analogous LP for the PCSCP, which using the reduction in lemma 3 implies the gap on the original LP.

**THEOREM 10.** *The above LP has integrality gap at most 0.828.*

**PROOF.** We construct an LP gap instance for the prize collecting set cover problem. In our instance, the input is a graph; the sets are vertices and the elements are edges,

so that each edge is present in the sets corresponding to its incident vertices. For an  $n$ -vertex graph, the goal is to choose a subset  $X$  of vertices to maximize  $\alpha|E(X)| + n - |X|$ , where  $E(X)$  is the subset of edges incident to some vertex in  $X$ .

The LP has a variable  $x_e$  for each edge, which is 1 if the edge is selected in the event that all vertices in the graph have negative virtual valuation. Similarly,  $y_v$  is the variable denoting whether vertex  $v$  is selected in the same event. The LP can be reformulated as:

$$\begin{aligned} \text{Maximize} \quad & n - \sum_v y_v + \alpha \sum_e x_e \\ x_e \leq & y_u + y_v \quad \forall e = (u, v) \in E \\ x_e, y_u \in & [0, 1] \quad \forall e \in E, u \in V \end{aligned}$$

Consider a complete graph on  $n$  vertices, for large  $n$ . By appropriate scaling, let us rewrite the objective as  $|E(X)| + \alpha n(n - |X|)$ . Set all  $y_v = 1$  and  $x_e = 1/2$ . For this fractional solution, the objective is approximately  $n^2(1 + \alpha)/2$ . Suppose the optimal integer solution chooses  $k$  vertices and all incident edges. Its value is approximately  $nk - k^2/2 + \alpha n(n - k)$ . Optimizing over  $k$ , we obtain  $k = n(1 - \alpha)$ , so that the optimal value is  $n^2(1 + \alpha^2)/2$ . The ratio is therefore  $(1 + \alpha^2)/(1 + \alpha)$ , so that  $\alpha = \sqrt{2} - 1$ . This yields a ratio of  $2(\sqrt{2} - 1) = 0.828$ .  $\square$

## 5. SUBMODULAR EXTERNALITIES

In order to design an approximately optimal mechanism for the more general problem with submodular externalities, we identify a set of mechanisms, called *influence-and-exploit* mechanisms. In the following, we first show that a simple random-sampling mechanism which belongs to this category of mechanisms achieves a 0.25-approximate mechanism for this problem. Then, we focus on optimizing over these mechanisms and design improved approximation algorithms for this problem. We start by defining influence and exploit-mechanisms:

**DEFINITION 4.** *Let  $S$  be a set of users, called the influence set. For a fixed price  $p$  and any set of players  $S$ , define the Influence-and-Exploit Mechanism  $IE(S)$  as follows. Give the item to any  $i \in V \setminus S$  regardless of its value. Give the item to any  $i \in S$  if his value is more than the threshold  $p$ .*

### 5.1 Constant Approximation

First, we observe that a simple IE mechanism gives a 0.25-approximation to the optimal revenue for the setting of single-parameter submodular externalities. Consider the following algorithm:

- Let  $S$  be a random subset of bidders where each  $i \in S$  is chosen independently with probability  $\frac{1}{2}$ .
- *Influence:* Give the good to all  $i \in V \setminus S$  regardless of the value.
- *Exploit:* Give the good to a bidder  $j \in S$  if  $v_j \geq p_j(S)$ , where  $p_j(S) = \phi_j^{-1}(0)$  is the inverse virtual value of zero for the distribution  $F_{j,S}$ , where  $F_{j,S}(p) = Pr_{v_j \in F_j}(v_j g_j(S) \leq p)$ .

In order to prove the approximation guarantee, we make use of the following lemma.

**LEMMA 11** ([7]). *For a ground set  $V$ , Let  $f : 2^V \rightarrow \mathbb{R}$  be a monotone submodular set function. Form set  $S$  by picking elements  $i \in V$  independently at random with some fixed probability  $p$ . Then*

$$\mathbf{E}[f(S)] \geq p\mathbf{E}[f(V)]$$

Now define the *revenue function*  $R_i(S) = \max_p p(1 - F_{i,S}(p))$ . We first prove that  $\sum_i R_i(V)$  is an upper bound on the revenue of any mechanism.

**LEMMA 12.** *The expected revenue of any truthful mechanism is at most  $\sum_i R_i(V)$ .*

**PROOF.** Recall that we normalized  $g_i(V) = 1$ . As a result,  $F_{i,V}(p) = F_i(p)$ . So by definition  $R_i(V) = \max_p p(1 - F(p))$ . Consider any mechanism with allocation function  $x_i(v_i) \leq 1$ . By Myerson's characterization, the expected revenue of the mechanism is

$$\begin{aligned} \sum_i E_{v_i} [x_i(v_i)\phi_i(v_i)] &\leq \sum_i E[\max(0, \phi_i(v_i))] \\ &= \int_{\phi^{-1}(0)}^{\infty} \phi_i(x)f(x)dx \\ &= \int_{\phi^{-1}(0)}^{\infty} (xf(x) - (1 - F(x)))dx \\ &= -x(1 - F(x))\Big|_{\phi^{-1}(0)}^{\infty} \\ &= p_i(S)(1 - F_{i,S}(p_i(S))) \\ &= R_i(V). \end{aligned}$$

$\square$

**LEMMA 13.** *If the revenue function is submodular for all agents, then the above mechanism is a 4-approximation of the optimal mechanism.*

**PROOF.** Consider any agent  $i$ . With probability 1/2, it is chosen to be in  $S$ . Fixing the set  $S$ , the expected revenue we get from  $i$  is  $R_i(S) = p_i(S)(1 - F_{i,S}(p_i(S)))$ . Now note that each agent is independently sampled, so over the random choices of the mechanism, and by submodularity of  $R_i(S)$ , the expected revenue from  $i$  (conditioned on being in  $S$ ) is at least  $R_i(V)/2$ . Since we get this revenue with probability 1/2, the expected revenue from  $i$  is at least  $R_i(V)/4$ . This gives a 4-approximation.  $\square$

Similar to [11], we may simply assume that the revenue function  $R_i$  is monotone and submodular for each bidder, and indeed our result holds for any settings that induce monotone submodular revenue functions. Interestingly, for the single-parameter submodular setting, the submodularity of the revenue function follows from the submodularity of the local influence function.

**LEMMA 14.** *The revenue function is submodular for the single-parameter submodular externality setting and the concave externality setting.*

**PROOF.** Consider a player  $i$  with distribution  $F_i$  over  $v_i$ . Then

$$\begin{aligned} R_i(S) &= \max_p p(1 - F_{i,S}(p)) \\ &= \max_p p(1 - F_i(p/g(S))) \\ &= g(S) \max_{p'} p'(1 - F(p')) \end{aligned}$$



where  $p' = p/g(S)$ . Submodularity of  $R_i(\cdot)$  then follows directly from submodularity of  $g(\cdot)$ .  $\square$

Applying the above two lemmas, we conclude that the following:

**THEOREM 15.** *There exists a  $\frac{1}{4}$ -approximate IE mechanism to the optimum revenue in the single-parameter submodular externality model, and thus in the concave externality model.*

## 5.2 Optimizing over IE Mechanisms

Now that we proved that IE mechanisms achieve constant-factor approximation to the optimal revenue, it would be interesting to optimize among IE mechanisms. To do so we need to find a set  $V \setminus S$  of initial (influential) bidders to get the item regardless of their value, and then exploit the remaining bidders  $S$  by setting optimal thresholds as above. Let  $x_S(v)$  be the outcome of this strategy, that is,  $x_{i,S}(v) = 1$  if the item is given to  $i$  for the profile of types  $v$  in  $\text{IE}(S)$ . Let  $\Phi(S)$  be the expected revenue of  $\text{IE}(S)$ . Our goal is to find a subset  $S$  of bidders that maximizes  $\Phi(S)$ . We do so by arguing that  $\Phi(S)$  is a (not necessarily monotone) submodular function and then using submodular function maximization results. We present the results in this section with regard to concave externalities in order to keep notation simple; the results extend easily to the more general submodular externalities. We first characterize the expected revenue of any IE strategy.

**LEMMA 16.** *Let  $X_{i,S}(v) = |\{j \in N(i) : \chi_j(v) = 1\}|$ . Then the expected revenue of any IE strategy,  $\text{IE}(S)$ , for each  $i \in S$  is equal to  $p(1 - F_i(p))E_v[h(X_{i,S}(v))]$  where  $h(\cdot)$  is the concave function defining the externality (i.e.,  $g_i(o) = o_i \cdot h(|\{j \in N(i) : o_j = 1\}|)$ ).*

**PROOF.** Consider  $\text{IE}(S)$  with allocation function  $x$  and outcome function  $\chi$ . By Myerson's characterization, we can write the expected revenue of  $i$  in  $\text{IE}(S)$  as

$$\begin{aligned} E_v[x_i(v)\phi_i(v_i)] &= E_v[g_i(\chi(v))\phi_i(v)] \\ &= E_v[\chi_i(v)h(X_{i,S}(v))\phi_i(v)]. \end{aligned}$$

Note in any IE strategy,  $\chi_i(v)$  and  $\chi_j(v)$  are independent random variables (when  $v$  is drawn from  $F$ ) for any  $i \neq j$ . Thus  $\chi_i(v)$  is also independent from  $X_{i,S}(v)$ . So we can write the revenue of  $i$  as  $E[h(X_{i,S}(v))]E[\chi_i(v)\phi_i(v_i)]$ . Since we set  $\chi_i(v_i) = 1$  when  $\phi(v_i) \geq 0$ ,  $E[\chi_i(v)\phi(v_i)]$  is equal to the optimum revenue from distribution  $F_i$ , which is equal to  $p(1 - F(p))$ .  $\square$

We next prove the key structural property of the revenue function  $\Phi(S)$  for IE mechanisms, namely that it is submodular.

**LEMMA 17.** *The set function  $\Phi$  is a non-negative submodular function of  $S$ .*

**PROOF.** First note that each agent  $i \in V \setminus S$  contributes 0 to the revenue, and each  $i \in S$  contributes

$$p(1 - F(p))E_v[g(X_{i,S}(v))],$$

where  $X_{i,S}(v)$  is a random variable denoting the number of  $i$ 's neighbors that are given the item to at profile  $v$ , that is  $X_{i,S}(v) = \sum_{j \in N(i)} x_{i,S}(v)$ . For all  $i, S$ , let  $F_{i,S}$  be the distribution (with density function  $f_{i,S}$ ) of  $X_{i,S}(v)$  when  $v$

is drawn from the joint distribution of types. We show submodularity of  $\Phi(\cdot)$  by proving submodularity of all  $\Phi_i(\cdot)$  for all  $i$ , that is  $\Phi_i(S \cup \{j\}) - \Phi_i(S) \leq \Phi_i(S' \cup \{j\}) - \Phi_i(S')$ , for all  $S' \subseteq S$  and all  $i$  and  $j$ . Submodularity of  $\Phi(\cdot)$  follows from submodularity of  $\Phi_i(\cdot)$ 's, since  $\Phi(S) = \sum_{i \notin S} \Phi_i(S)$ , and First note that if  $i$  is not a neighbor of  $j$ , then we have  $0 = \Phi_i(S \cup \{j\}) - \Phi_i(S) \leq \Phi_i(S' \cup \{j\}) - \Phi_i(S') = 0$ . Now assume that  $i$  is a neighbor of  $j$ . Define  $\Phi = \Phi_i(S \cup \{j\}) - \Phi_i(S)$ . Now we have

$$\begin{aligned} \Phi &= p(1 - F(p))(E_v[g(X_{i,S \cup \{j\}}(v))] - E_v[g(X_{i,S}(v))]) \\ &= p(1 - F(p))(E_{k \sim F_{i,S \cup \{j\}}}[g(k)] - E_{k \sim F_{i,S}}[g(k)]) \\ &= p(1 - F(p)) \sum_k g(k)(f_{i,S \cup \{j\}}(k) - f_{i,S}(k)) \end{aligned}$$

Note that we can write  $f_{i,S \cup \{j\}}(k) = F(p)f_{i,S}(k+1) + (1 - F(p))f_{i,S}(k)$ . As a result,

$$\begin{aligned} \Phi &= p(1 - F(p))F(p) \sum_k g(k)(f_{i,S}(k+1) - f_{i,S}(k)) \\ &= p(1 - F(p))F(p) \sum_k f_{i,S}(k)(g(k-1) - g(k)) \end{aligned}$$

Now recall that  $g$  is a concave function of  $k$ . As a result,  $G(k) = g(k-1) - g(k)$  is a non-decreasing function of  $k$ . Therefore,

$$\begin{aligned} &\Phi_i(S' \cup \{j\}) - \Phi_i(S') - (\Phi_i(S \cup \{j\}) - \Phi_i(S)) \\ &= p(1 - F(p))F(p) \sum_k G(k)(f_{i,S'}(k) - f_{i,S}(k)) \\ &= p(1 - F(p))F(p) \sum_k G(k)(F_{i,S'}(k) - \\ &\quad F_{i,S'}(k-1) - (F_{i,S}(k) - F_{i,S}(k-1))) \\ &= p(1 - F(p))F(p) \\ &\quad \cdot \sum_k (F_{i,S'}(k) - F_{i,S}(k))(G(k) - G(k+1)) \end{aligned}$$

Note that for  $S' \supset S$ ,  $F_{i,S'}(k) \geq F_{i,S}(k)$ , and therefore  $F_{i,S'}(k) - F_{i,S}(k) \geq 0$  for all  $k$ . Also, Since  $G$  is a non-decreasing function,  $G(k) - G(k+1) \leq 0$ .

It only remains to consider the revenue function of  $j$  when we add  $j$  to sets. If For  $S$  such that  $j \notin S$ , we have

$$\Phi_j(S \cup \{j\}) - \Phi_j(S) = p(1 - F(p))E_v[g(X_{j,S}(v))]$$

Again, note that for  $S' \supset S$ ,  $F_{j,S}(k) \leq F_{j,S'}(k)$ , therefore

$$\begin{aligned} E_{F_{j,S'}}[g(k)] &- E_{F_{j,S}}[g(k)] \\ &= \sum_k g(k)(f_{j,S'}(k) - f_{j,S}(k)) \\ &= \sum_k (F_{j,S'}(k) - F_{j,S}(k)) \\ &\quad \cdot (g(k) - g(k+1)) \leq 0 \end{aligned}$$

$\square$

Function  $\Phi(\cdot)$  as described above is non-negative and submodular, but not necessarily monotone. In order to obtain a constant-factor approximation for maximizing over IE mechanisms, we can simply apply non-monotone submodular maximization algorithms for this problem [7, 8]. For example, the following simple local search algorithm gives a 0.33-approximation to this problem[7]: (i) Let  $S = \{v | v = \text{argmax}_{i \in V} (\Phi(\{i\}))\}$ , and (ii) at each step either add or

remove a bidder  $i$  from  $S$  if this adding or removing increases the value of  $\Phi(S)$  be a  $1 + \frac{\epsilon}{n}$  factor, (ii) After reaching a local optimal  $L$ , output the better of  $L$  and  $\bar{L}$ . The above simple algorithm achieves 0.33-approximation for the problem of maximizing over IE strategies. One can apply a recently-developed randomized local search algorithm [8] to achieve a 0.41-approximation for this problem.

## 6. REFERENCES

- [1] G. Aggarwal, J. Feldman, S. Muthukrishnan, and M. Pál. Sponsored search auctions with markovian users. In *WINE*, pages 621–628, 2008.
- [2] H. Akhlaghpour, M. Ghodsi, N. Haghpanah, V. S. Mirrokni, H. Mahini, and A. Nikzad. Optimal iterative pricing over social networks (extended abstract). In *WINE*, pages 415–423, 2010.
- [3] N. Anari, S. Ehsani, M. Ghodsi, N. Haghpanah, N. Immorlica, H. Mahini, and V. S. Mirrokni. Equilibrium pricing with positive externalities (extended abstract). In *WINE*, pages 424–431, 2010.
- [4] S. Athey and G. Ellison. Position auctions with consumer search. Forthcoming, *Quarterly Journal of Economics*, 2010.
- [5] O. Candogan, K. Bimpikis, and A. E. Ozdaglar. Optimal pricing in the presence of local network effects. In *WINE*, pages 118–132, 2010.
- [6] U. Feige, N. Immorlica, V. S. Mirrokni, and H. Nazerzadeh. Pass approximation. In *APPROX-RANDOM*, pages 111–124, 2009.
- [7] U. Feige, V. S. Mirrokni, and J. Vondrak. Maximizing non-monotone submodular functions. In *FOCS '07*, pages 461–471, Washington, DC, USA, 2007. IEEE Computer Society.
- [8] S. O. Gharan and J. Vondrak. Submodular maximization by simulated annealing. In *SODA*, 2011.
- [9] I. Giotis and A. R. Karlin. On the equilibria and efficiency of the gsp mechanism in keyword auctions with externalities. In *WINE*, pages 629–638, 2008.
- [10] R. Gomes, N. Immorlica, and V. Markakis. Externalities in keyword auctions: an empirical and theoretical assessment. In *WINE'09*, 2009.
- [11] J. D. Hartline, V. S. Mirrokni, and M. Sundararajan. Optimal marketing strategies over social networks. In *WWW*, pages 189–198, 2008.
- [12] P. Jehiel, B. Moldovanu, and E. Stacchetti. How (not) to sell nuclear weapons. *American Economic Review*, 86(4):814–829, 1996.
- [13] P. Jeziorski and I. Segal. What makes them click: Empirical analysis of consumer demand for internet search advertising. working paper, 2009.
- [14] D. Kempe and M. Mahdian. A cascade model for externalities in sponsored search. In *WINE*, pages 585–596, 2008.
- [15] R. Myerson. Optimal auction design. *Mathematics of Operations Research*, pages 58–73, 1981.
- [16] J. Rohlfs. A theory of interdependent demand for a communications service. *The Bell Journal of Economics and Management Science*, 5(1):16–37, 1974.

## APPENDIX

### A. LIMITED-SUPPLY SETTING

**THEOREM 18.** *There is a (1/4)-approximate auction for the limited-supply setting.*

**PROOF.** Suppose that the auctioneer has  $k$  copies of the good. Compute a spanning tree of the social network and color the nodes red and blue such that each red node has a blue neighbor in the spanning tree (and vice versa). Pick a color uniformly at random and name the nodes of this color  $S_1$  and nodes of the other color  $S_2$ . Allocate to the  $k/2$  highest positive virtual values in  $S_1$ , and their neighbors in  $S_2$  to ensure feasibility. We now compute the expected virtual value of this allocation for the red nodes. We condition on the event  $E$  that the red nodes were chosen to be set  $S_1$ .

$$\begin{aligned} E_v\left[\sum_{i \text{ red}} x_i(v)\phi_i(v_i)\right] &= E_v\left[\sum_{i \text{ red}} x_i(v)\phi_i(v_i)|E\right] \Pr[E] \\ &\quad + E_v\left[\sum_{i \text{ red}} x_i(v)\phi_i(v_i)|\bar{E}\right] \Pr[\bar{E}] \\ &\geq \left(\frac{1}{2}\right) E_v\left[\sum_{i \text{ red}} x_i(v)\phi_i(v_i)|E\right] \\ &= \left(\frac{1}{2}\right) E_v\left[\max_{S \subseteq \text{red}: |S| \leq k/2} \sum_{i \in S} \phi_i(v_i)\right] \end{aligned}$$

where the second step follows since the expected allocation of any red node  $i$  is independent of its value conditioned on  $\bar{E}$ . Therefore by fact 1 each such vertex contributes a non-negative amount to the revenue. The third step follows since conditioned on  $E$  we picked the best set of size at most  $k/2$  from the red nodes.

Now to prove the approximation guarantee first define  $X = E[\max_{S: |S| \leq k} \sum_{i \in S} \phi(v_i)]$ , and note this is an upper bound on the optimum revenue, since in the best case we can allocate the highest (positive)  $k$  virtual values. But we know that for any sampling of the values,

$$\begin{aligned} \max_{S: |S| \leq k} \sum_{i \in S} \phi(v_i) &\leq \max_{S \subseteq \text{red}: |S| \leq k} \sum_{i \in S} \phi(v_i) \\ &\quad + \max_{S \subseteq \text{blue}: |S| \leq k} \sum_{i \in S} \phi(v_i) \\ &\leq 2 \cdot \left( \max_{S \subseteq \text{red}: |S| \leq k/2} \sum_{i \in S} \phi(v_i) \right) \\ &\quad + \max_{S \subseteq \text{blue}: |S| \leq k/2} \sum_{i \in S} \phi(v_i), \end{aligned}$$

and therefore,

$$\begin{aligned} X &= E\left[\max_{S: |S| \leq k} \sum_{i \in S} \phi(v_i)\right] \\ &\leq 2(E\left[\max_{S \subseteq S_1: |S| \leq k/2} \sum_{i \in S} \phi(v_i)\right] \\ &\quad + E\left[\max_{S \subseteq S_2: |S| \leq k/2} \sum_{i \in S} \phi(v_i)\right]), \end{aligned}$$

by linearity of expectation. Recalling that the expected value of our allocation is at least

$$E\left[\max_{S \subseteq S_1: |S| \leq k/2} \left(\sum_{i \in S} \phi_i(v_i)\right)\right],$$

and noting that we picked each of the two sets with probability 1/2 to be  $S_1$ , we conclude that the expected revenue of our allocation, is at least 1/4 of the upper bound.  $\square$