Non-Cooperative Cost Sharing Games via Subsidies*

Niv Buchbinder † Liane Lewin-Eytan‡ Joseph (Seffi) Naor§ Ariel Orda¶

Abstract

We consider a cost sharing system where users are selfish and act according to their own interest. There is a set of facilities and each facility provides services to a subset of the users. Each user is interested in purchasing a service, and will buy it from the facility offering it at the lowest cost. The overall system performance is defined to be the total cost of the facilities chosen by the users. A central authority can encourage the purchase of services by offering subsidies that reduce their price, in order to improve the system performance. The subsidies are financed by taxes collected from the users.

Specifically, we investigate a non-cooperative game, where users join the system, and act according to their *best response*. We model the system as an instance of a set cover game, where each element is interested in selecting a cover minimizing its payment. The subsidies are updated dynamically, following the selfish moves of the elements and the taxes collected due to their payments. Our objective is to design a *dynamic* subsidy mechanism that improves on the overall system performance while collecting as taxes only a small fraction of the sum of the payments of the users. The performance of such a subsidy mechanism is thus defined by two different quality parameters: (i) the *price of anarchy*, defined as the ratio between the cost of the Nash equilibrium obtained and the cost of an optimal solution; and (ii) the *taxation ratio*, defined as the fraction of payments collected as taxes from the users.

We investigate two different models: (i) an *integral* model in which each element is covered by a single set; and (ii) a *fractional* model in which an element can be fractionally covered by several sets. Let f denote the maximum number of sets that an element can belong to. For the fractional model, we provide a subsidy mechanism such that, for any $\epsilon \leq 1$, the price of anarchy is $O(\frac{\log f}{\epsilon})$ and the taxation ratio is ϵ. For the integral model, we provide a subsidy mechanism such that, for any $\epsilon \leq 1$, the price of anarchy is $O(\frac{\log f \log\left(\frac{n}{\epsilon}\right)}{\epsilon})$ and the taxation ratio is ϵ, where n is the number of elements.

1 Introduction

Individual self-interest is the basis for the modern market system in which a consumer acts in its self-interest when buying goods at lowest prices. A government, or any other central authority,

*A preliminary version of this paper appeared in the *Proceedings of the First International Symposium on Algorithmic Game Theory (SAGT '08)*.

†Computer Science Department, Technion, Haifa 32000, Israel. E-mail: nivb@cs.technion.ac.il.

‡Department of Electrical Engineering, Technion, Haifa 32000, Israel.
E-mail: liane@tx.technion.ac.il.

§Computer Science Department, Technion, Haifa 32000, Israel. E-mail: naor@cs.technion.ac.il. Research supported in part by ISF Grant 1366/07 and US-Israel BSF Grant 2002276.

¶Department of Electrical Engineering, Technion, Haifa 32000, Israel.
E-mail: ariel@ee.technion.ac.il.

can influence natural market forces in several ways, such as taxation or regulation. In cases where a government wishes to support and encourage the production of a good that is regarded as being in the public interest, it gives out an assistance called a *subsidy* (also called negative taxation). Subsidies are thus a way to influence the state of the market in a world of independent self-interested consumers.

An example where government supervision can be very effective is an urban passenger transportation system. An employee commuting to work in a city usually has many transportation options. He can use a private car, join a car-pool, or use public transportation, e.g., bus or train. The common choices as to how to travel to work have significant environmental impacts and a major influence on road traffic congestion. It is thus a governmental interest to reduce the number of single occupancy vehicles on the road and encourage people to use public transport when commuting to and from work.

Letting the invisible hand of the free market take its course can sometimes be devastating. Consider, for example, a setting in which a new building is being built. Each new resident can either purchase a private car, or initiate the use of some public transport at a much higher cost. As no bus line is available at the new residence when it is established, the cheapest way for each new resident to commute is to buy his own car, and then no public transport will ever be established. Thus, in this case, it is the role of a central authority to develop public transport by offering subsidies. After public transportation means are established, it is likely that residents will switch from private to public transport, since the latter is cheaper.

Central authorities have limited budgets. Therefore, subsidies are financed by taxes collected from the users. The taxes collected by a central authority should only be a bounded fraction of the total payments made by the users. In the sequel, we develop a formal model of a cost sharing system with selfish non-cooperative users, and introduce a *dynamic* subsidy mechanism that improves on the overall system performance.

1.1 Our Model

We investigate a system where facilities provide services to users. Each user is interested in purchasing a service which is typically provided by only a subset of the facilities. Users naturally buy the service from the facility offering it at the lowest cost. A central authority can encourage the purchase of services by offering subsidies that reduce their price. We investigate settings where users share services and thereby also share their cost. The notion of *social welfare* or *social cost* corresponds to the overall system performance, and is defined to be the total cost of the facilities chosen by the users. Back to the public transportation example, each transportation option corresponds to a different facility having a different cost. The cost of each facility is essentially the cost of operating the type of transport it represents. The cost of a facility that provides service to several users is shared amongst them, and can be subsidized by the central authority in order to shift market share of facilities and users to it, e.g., from cars to public transport.

The Set Cover Setting. We model the system as an instance of the set cover problem. Let $N = \{1, 2, \ldots, n\}$ be a ground set of n elements (the users), and let $\mathbb{S}$ be a family of subsets of N, $|\mathbb{S}| = m$ (the facilities). Each facility $s \in \mathbb{S}$ thus consists of the users to whom it can provide service. A *cover* of $N' \subseteq N$ is a collection of sets such that their union contains N'. In our

public transportation example, a cover is a choice of transport types allowing all users belonging to N' to get to work. Each subset $s \in \mathbb{S}$ has a non-negative *cost* c_s associated with it. The *social cost* of a collection of sets $\mathbb{T}$ is defined to be the total cost of the sets belonging to $\mathbb{T}$.

In a feasible cover, each user is assigned to one of the sets in the cover containing it. Users sharing the same set also share its cost. We consider an egalitarian cost sharing mechanism, which evenly splits the cost of a set among its users. This cost sharing mechanism has an intuitive appeal, and satisfies essential properties such as *cross monotonicity* (the cost share of a user for using a set cannot increase when additional users join the system) and *budget balance* (the sum of the payments of the users receiving service from a set is equal to its cost).

The Non-Cooperative Game. We consider a set cover game with selfish non-cooperative players (also called users, or elements). Each player is interested in selecting its cover in a way that minimizes its payment. Thus, the strategies of the players in the game correspond to their different possible covers, that is, the different sets that can provide service to the players. Each player independently chooses a strategy minimizing its payment, i.e., its *best response*. The best response of a player in the set cover game is thus defined as the set(s) that can provide service to the player at minimum cost (with respect to the *current* state of the system). The mutual influence of the players is determined by the egalitarian cost sharing mechanism.

We focus on a dynamic setting, where players follow the natural game course induced by *best-response dynamics*. Each player, in his turn, chooses a cover that minimizes his cost. In this paper, we take an approach that does not rely on starting the game in a specific starting configuration. There are many situations in which not all players might be available at the same time. We thus explore a natural setting where users join the game starting from an empty configuration. Upon arrival, a user chooses a cover selfishly. As a result, players that have joined the game previously may change their strategy later on by choosing a cover of lower cost. The central authority is allowed to increase the subsidies of the sets in every step of the game in order to improve the social welfare of the final cover. We assume that the game is controlled by an adversarial scheduler that decides which user plays in each step. The order by which the users play is not known beforehand (as it is chosen adversarially) as well as the set of elements (users) $N' \subseteq N$ that actually participates in the game. (Note that N' may be a strict subset of N in general.) However, we assume that the set cover instance, i.e., N and $\mathbb{S}$, is known in advance.

The natural game course continues until Nash equilibrium is reached. A Nash equilibrium of the set cover game corresponds to a choice of covers for all users in N', where no user can unilaterally reduce its payment by choosing a different cover. We note that the set cover game is a special case of the well known class of *congestion games* [15]. Rosenthal [15] showed that a potential function can be defined for each congestion game with the property that it decreases in case a player makes a move that improves his cost, thus establishing convergence to Nash equilibrium through best response dynamics.

Subsidies & Taxes. The Nash equilibrium of the set cover game is not unique and the greedy nature of the users could lead to very inefficient Nash equilibrium points, even when initializing the game from an empty configuration. We use subsidies in order to guarantee that best-response dynamics will not converge to such bad equilibria. The following example is instructive as to why subsidies are needed for minimizing the cost of the final solution when considering an arbitrary set system. Consider n users where each user can be covered by a unit-cost "private"

set containing only himself. There is also a set containing all the users that costs $\sqrt{n}$. The users appear one by one and the best response of each user is to pick the private set covering him. Once a user picks his private set, he will have no incentive to change his strategy. How will the set covering all the users (the social optimal solution) be chosen without subsidies? Clearly, a worse case could be achieved in case the cost of the set containing all users is $1 + \epsilon$. However, in this case, an optimal solution could be trivially achieved by giving a negligible subsidy (higher than ϵ) to this set.

During the course of the game, each subset $s \in \mathbb{S}$ is associated with a subsidy value (possibly equal to zero, in case no subsidy is offered to the set). The *effective cost* of a set s, denoted by $\hat{c}_s$, is defined to be c_s minus the subsidy associated with s. Thus, following the egalitarian cost sharing mechanism used in our setting, if n_s users use set s, then each user pays $\hat{c}_s/n_s$ for this set. We note that as the subsidies can only lower the cost of the users, the potential function of the set cover game decreases in the presence of a subsidy mechanism as well, and convergence to Nash equilibrium is still guaranteed.

Ideally, we would like our subsidy mechanism to spend on subsidies only a bounded fraction of the revenue. However, in a non-cooperative game setting, the users' payments are dynamic, and can vary significantly during the game course due to strategy changes. Consider, for example, a set s shared by many users, who decide to leave it at some point of the game in order to join other subsidized sets. In this case, the revenue that was accrued from the use of s is now reduced to zero. In order to cope with such dynamic scenarios, we propose a natural framework where subsidies are offered via *taxes*. A tax is a non-refundable sum paid to the central authority only in case a user purchases a new set. It is equal to a fixed fraction of the effective cost of the purchased set. The taxes collected by a central authority are equal to a fraction of the payments made by the users that open new sets, and later on offers the revenue from the taxes as subsidies. The total amount of subsidies offered should always be bounded by the amount of taxes collected.

Quality Parameters. The performance of a subsidy mechanism is a function of two quality parameters:

- **The price of anarchy:** The ratio between the social cost of a Nash equilibrium solution (that is, the sum of the subsidies and the payments of the users) and the social cost of an optimal solution.
- **The taxation ratio:** The fraction of the payments collected as taxes from the users.

There is a trade-off between the taxation ratio and the price of anarchy achieved by our subsidy mechanism. The higher the fraction of payments collected as taxes and spent on subsidies is, the lower the cost of the final solution is. The taxation ratio is determined by a parameter $\epsilon \leq 1$ which is given as input to the subsidy mechanism. Denoting by P the total payments of the user, the objective is to achieve the best price of anarchy while collecting taxes (and spending on subsidies) at most ϵP.

Compare the set cover game to the multicast game [3, 7] in which users (terminals) connect to a source by making a routing decision that minimizes their payment. Chuzhoy et al [7] analyze the price of anarchy of a Nash equilibrium resulting from the best-response dynamics of a game course in which the players first join the game sequentially beginning from an *empty*

configuration. Their setting is thus a special case of our model in which there is no central authority intervention. It is shown in [6] that the price of anarchy of this setting is $O(\log^3 n)$. In the multicast setting, unlike the set cover game, an initial empty configuration coupled with best response dynamics does guarantee a low price of anarchy with no need to offer subsidies.

1.2 Results and Techniques.

We consider two different models: (i) an *integral* model in which sets can only be fully bought (i.e., integrally) and each element is covered by a single set; and (ii) a *fractional* model in which a fraction of a set can be bought and each user can be covered by several sets (provided that their fractions add up to 1). Note that the fraction of coverage an element gets from a set cannot be greater than the fraction associated with the set. The subsidies, similarly to the choices of the sets, can be given either integrally or fractionally, depending on the model. In the fractional model, the central authority is allowed to subsidize only a fraction of a set.

The importance of the fractional model is two-fold. First, a fractional solution turns out to be an intermediate step towards obtaining an integral solution. Second, it is interesting in its own right as it captures many practical "fractional" scenarios. In the urban transport system example, a fractional solution can correspond to the case where a user uses different transport options during the week. Then, subsidizing a fraction of a set can be interpreted as subsidizing a transportation mean only during part of the day, or part of the week. Thus, a fraction in this example can be interpreted as "rate".

Let f denote the maximum *frequency* of an element, that is, the maximum number of sets that an element can belong to. For the fractional model, we prove the following theorem:

Theorem 1 (**Fractional Cover**). *There exists a subsidy policy such that, for any $\epsilon \leq 1$, the price of anarchy is $O(\frac{\log f}{\epsilon})$ and the taxation ratio is ϵ.*

Theorem 1 provides a trade-off between the taxation ratio and the price of anarchy: as the central authority collects a higher fraction of the payments as taxes (later on invested in subsidies), the cost of the final solution decreases.

For the integral model, we obtain the following slightly inferior bound.

Theorem 2 (**Integral Cover**). *There exists a subsidy policy such that, for any $\epsilon \leq 1$, the price of anarchy is $O(\frac{\log f \log(\frac{n}{\epsilon})}{\epsilon})$ and the taxation ratio is ϵ.*

In order to design mechanisms for both the fractional and integral models, we draw on ideas from [1, 4]. In [1] Alon et al. considered an online version of the set cover problem, where elements arrive one by one and need to be covered upon arrival. The goal of [1] is to design an online algorithm achieving the best possible competitive ratio with respect to the optimal solution, i.e., optimal social welfare. In [1], an $O(\log m \log n)$-competitive algorithm is presented for this online setting.

In our work, we take into consideration not only the overall system performance, but also the selfish nature of the users who play according to their best response. As the goal of the users is to minimize the payment for their cover, they may change their strategy after joining the system until Nash equilibrium is reached. We thus go beyond the online version of the set cover problem considered in [1], and analyze its non-cooperative game extension. The model investigated in [1] can be seen as a special case of ours, where the central authority pays for the full cost of the

cover and users pay nothing. (Also, users join one by one without reaching equilibrium.) Thus, using the algorithm of [1], a central authority will not be able to finance the subsidies from taxes. Bounding the taxation ratio while maintaining a low price of anarchy requires a new algorithm and a different analysis which is achieved via the primal-dual approach of [4].

We note that for both the integral and fractional models our results are tight, since our subsidy mechanism applies also to the special case of an online setting where users join the system one by one and act according to their *best response* by choosing a cover of minimum cost. In case ϵ is a fixed constant, the resulting price of anarchy almost matches the lower bound of $\Omega(\frac{\log m \log n}{\log\log m + \log\log n})$ shown in [1] for the online setting.

In the fractional model, our subsidy mechanism keeps a bounded taxation ratio by investing money in subsidies only when a user purchases a new set (or a fraction thereof), and pays a tax. The idea is therefore to invest in subsidies in each iteration only a small fraction of the payment of a user (corresponding to the tax paid). The total cost of the taxes is bounded by maintaining an (almost) feasible dual solution during the execution of the algorithm, which also allows us to bound the price of anarchy of the solution. As in [1], we maintain a fractional primal solution, however, in our case, it is not always feasible. Rather, the feasibility of the cover is obtained by the best response of the users joining the system. The primal solution we maintain corresponds to the subsidies given by the central authority.

Developing an integral subsidy mechanism requires several more ideas. As opposed to the fractional case, it is no longer possible to offer in each iteration a fraction of the user's payment. Instead, the algorithm keeps a bounded taxation ratio by giving an integral subsidy only after accumulating the taxes paid by the users over several iterations. In [1], Alon et al. obtained an integral solution for their online setting by maintaining at each iteration a fractional feasible solution and using a potential function that determines which of the sets should be chosen to the integral cover. We design a new potential function and note that the potential function defined in [1] cannot satisfy our needs, as it would lead to a high taxation ratio. The analysis we perform is more delicate and bounds both the price of anarchy and the taxation ratio of the algorithm.

Perspective on Other Approaches that Improve on the Social Welfare. The issue of improving on the overall system performance even in the face of selfish behavior has been considered extensively in the game theory literature, and designing mechanisms to improve the coordination of selfish agents is a well known idea. A central topic in game theory is the notion of mechanism design (see e.g. [14]) in which the rules of a game are designed to achieve a specific outcome. This is done by setting up a structure where players are paid (or penalized), and thus each player has an incentive to behave as the designer intends. Planning such a mechanism is based on an assumption that the players have private information known only to them and which affects their decisions.

Coordination mechanisms [8, 13] is a game theoretic concept that improves on the performance of systems with independent selfish and non-colluding agents by redesigning the system, i.e., by selecting policies and rules of the game (for example, adding delays and priorities to a congestion game [8, 9]). Another approach for improving on the overall system performance and reducing the price of anarchy is to impose economic incentives upon users in the form of tolls [5, 10, 11, 17]. In such systems, the performance of a user is determined by a monetary payment to a central authority for the use of particular resources. A different model used in order to improve on the social welfare assumes that the central authority can impose particular

strategies on some fraction of the self-optimizing users [12, 16]. This is called a *Stackelberg strategy*.

An important aspect of both mechanism design and coordination mechanisms is that the designer must design the system once and for all. The same applies to settings in which tolls are used, and which assume global knowledge of the system (in particular, the set of users, or commodities, is known beforehand). The tolls are thus computed off-line, prior to the course of the game. In contrast, in our setting, the policy of the central authority is dynamic, changes over time, and is determined by the state of the system.

Extensions. Our fractional subsidy algorithm can be generalized for the game extensions of the wide range of online graph and network optimization problems considered in [2] and which concern connectivity and cut problems in graphs. In a general online connectivity problem, there is a communication network known to the algorithm in advance, where each edge in the network has a nonnegative cost. The connectivity demands, specifying subsets of vertices to be connected, arrive online. The notion of *social welfare* of a subgraph G is defined to be the total cost of the edges belonging to G. Thus, an optimal solution with respect to the overall system performance consists of a minimum cost subgraph satisfying all connectivity demands. The algorithm presented in [2] satisfies each new demand, so as to achieve the best possible competitive ratio.

In the non-cooperative game version of these problems, a user corresponds to a connectivity demand, and is thus interested in choosing a minimum cost subgraph satisfying its own demand. The central authority is allowed to subsidize the costs of some of the edges by collecting taxes, in order to improve on the overall system performance. The game extensions of this range of problems belong to the class of congestion games [15], and thus their natural game course induced by best-response dynamics converges to a Nash equilibrium. Our subsidy algorithm achieves a taxation ratio of ϵ, while maintaining a price of anarchy of $O(\frac{\log m}{\epsilon})$, where m is the number of edges in the graph. Examples of problems belonging to this class are fractional versions of Steiner trees, generalized Steiner trees, and the group Steiner problem. It remains an open question whether an integral solution can be obtained for this set of problems as well.

2 Formal Definitions

In this section we formally describe our model. Let $N = \{1, 2, \ldots, n\}$ be a ground set of n elements (the users), and let $\mathbb{S}$ be a family of subsets of N, $|\mathbb{S}| = m$ (the facilities). Each $s \in \mathbb{S}$ has a non-negative *cost* c_s associated with it. Let f be the maximum frequency of an element, i.e., the maximum number of sets that can contain an element. A *cover* is a collection of sets such that their union is N. The *cost* of a cover is the sum of the costs of the sets that are included in the cover. A *fractional cover* is an assignment of weights, w_s, to each $s \in \mathbb{S}$, such that the total weight of the sets that contain each element is at least 1. The cost of a fractional cover is $\sum_{s \in \mathbb{S}} w_s c_s$. A linear programming formulation of the minimum fractional set cover problem appears in Figure 1. We have a variable w_s for each set $s \in \mathbb{S}$ indicating the fraction of set s that is taken to the cover. For each element, we demand that the sum of the fractions of the sets that contain the element is at least 1. In the dual program (see also Figure 1) we have a variable y_e corresponding to each of the elements. We require that the total sum of variables that correspond

Primal		Dual	
Minimize:	$\sum_{s\in\mathbb{S}} c_s w_s$	Maximize:	$\sum_{e\in N} y_e$
Subject to:		Subject to:	
$\forall e \in N$:	$\sum_{s\|e\in s} w_s \geq 1$	$\forall s \in \mathbb{S}$:	$\sum_{e\in s} y_e \leq c_s$
$\forall s \in \mathbb{S} \| w_s \geq 0$		$\forall e \in N \| y_e \geq 0$	

Figure 1: A primal-dual pair for the set-cover problem.

to elements that belong to a set s is at most the cost of the set. The integral set cover problem corresponds to the special case where $w_s \in \{0, 1\}$.

The Set Cover Game: Cost Shares & Subsidies Structures. We turn to define the set cover game more precisely. As our subsidy mechanism works under fairly general assumptions, not all the definitions here are needed for the algorithms and analysis in the next sections. Rather, any setting in which the subsidies offered are fully financed from taxes, is sufficient. We provide here a precise and natural definition of the game for completeness.

For simplicity, we assume taxes are collected only when a user purchases a new set (or a fraction thereof). This can happen either when a user joins the system or when it changes its strategy. The cost c_s of a new set s that has not yet been opened (purchased) is called its *opening cost*. When a new set is opened by a user, a fraction equal to ϵ of its opening cost is collected as tax. The *operating* cost of a facility is defined to be its opening cost minus the payment collected as tax. The operating cost of a facility that provides service to several users is shared amongst them. The tax paid by a user is non-refundable, while the remaining part of the payment (the operating cost) is a variable amount which may decrease when additional users join the same set and share its cost. We note that our mechanism can support other settings where both taxes and operating costs are shared by the users, as long as taxes are non-refundable, and can thus cover the subsidies. Each set is associated with a subsidy value (possibly zero, in case no subsidy is offered to the set). The subsidy can be applied either to the opening cost, in case the set has not been opened yet, or to the operating cost in case it is used by at least one user (see Figure 2). In the latter case, subsidies are given to a set that has already been purchased so as to lower its cost and encourage more users to join it. The *effective opening cost* $\hat{c}_s$ of a set s, is defined to be its opening cost minus the subsidy associated with s. The *effective operating cost* of s is defined similarly with respect to the operating cost of the set.

In the *integral* model, sets are taken integrally and each element is covered by a single set. The effective opening cost $\hat{c}_s$ that a user will have to pay for purchasing a new set s that is not subsidized, is composed of a non-refundable tax of $\epsilon \cdot \hat{c}_s$, and a variable payment of $(1-\epsilon)\hat{c}_s$ that is equal to the effective operating cost of the set. In case a user joins a set s that is not subsidized, and shared by other users, its payment is equal to $(1-\epsilon)\hat{c}_s/n_s$, where n_s is the number of users sharing s.

In the *fractional* model, each set s is associated with a fraction x_s which is fully subsidized (that is, its effective cost equals zero). The cost of any other fraction of this set, that is, a fraction λ that is not subsidized, is equal to $\lambda \cdot c_s$. Each element can be covered by several fractions of different sets adding up to 1. Denote the fraction of set s used by user i by $\lambda_{s,i}$ and the number of users using set s by n_s. Assume without loss of generality that $\lambda_{s,1} \leq \lambda_{s,2} \leq \cdots \leq \lambda_{s,n_s} \leq 1$. Define $\lambda_{s,0} = x_s$. The cost of each fraction of s is as follows: the interval

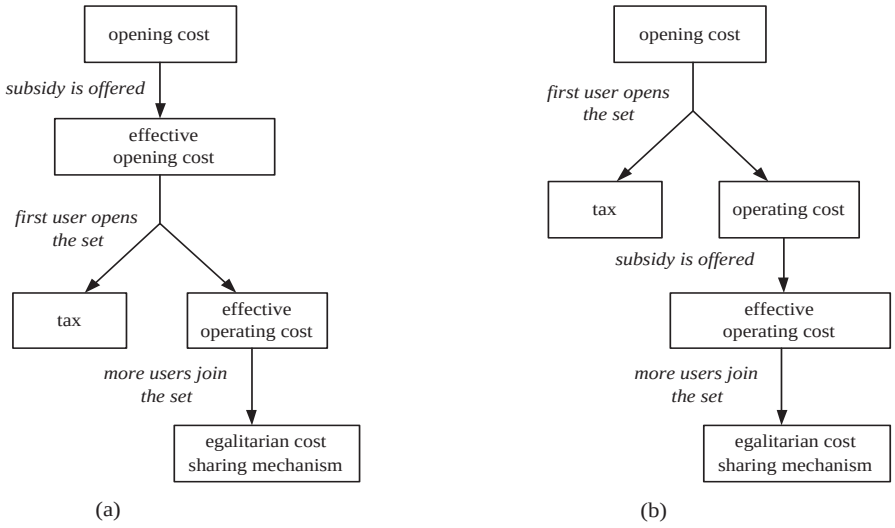

Figure 2: (a) A subsidy is offered to a set that has not been opened yet. The first user joining the set pays its effective opening cost, consisting of a tax and the effective operating cost. In case more users join the set, they share its effective operating cost. (b) A user opens a new set that is not subsidized, and pays its opening cost, consisting of a tax and the operating cost. The subsidy offered lowers the operating cost, changing it to effective operating cost. In case more users join the set, they share its effective operating cost.

$[\lambda_{s,j-1}, \lambda_{s,j}]$ is shared by $(n_s - j + 1)$ users, where the variable payment of each user equals $(\lambda_{s,j} - \lambda_{s,j-1}) \cdot (1 - \epsilon) \cdot c_s/(n_s - j + 1)$. The first user who opened the interval $[\lambda_{s,j-1}, \lambda_{s,j}]$ will also pay a non-refundable tax equal to $(\lambda_{s,j} - \lambda_{s,j-1}) \cdot \epsilon \cdot c_s$.

Let us consider the following example, where there is a set s with price c_s. The opened fraction of s is equal to $2/3$. Assume that 1/3 of s is fully subsidized, and element 1 is covered by 1/2 of s, and element 2 is covered by 2/3 of s. Assume also that element 1 was the first to open the interval $[1/3, 1/2]$ of s. Then, the tax paid by element 1 for opening s equals $\epsilon \cdot c_s \cdot (1/2 - 1/3)$. The variable payment of element 1 equals $(1/2 - 1/3) \cdot (1 - \epsilon)c_s/2$, as (1/2-1/3) of its cover is shared with element 2. The tax paid by element 2 for using s equals $\epsilon \cdot c_s \cdot (2/3 - 1/2))$, as he is the first (and only) user that uses the interval $[1/2, 2/3]$ of s. The variable payment of element 2 equals $(1/2 - 1/3) \cdot (1 - \epsilon)c_s/2 + (2/3 - 1/2) \cdot (1 - \epsilon)c_s$, as (1/2-1/3) of its cover is shared with element 1, and (2/3-1/2) of its cover is not shared with any other element. Note that the elements do not need to contribute any payments for the first 1/3 fraction of s as it is fully subsidized.

Following is a graphical example of the fractional model, where an element chooses a single set, and some fraction of this set is fully subsidized.

Nash Equilibrium Existence & Convergence. For both the fractional and integral models, the set cover game always converges to a Nash equilibrium. This property is established by means of a global potential function Φ on the strategy space. We denote by $\mathcal{T}$ the strategy profile consisting of the integral cover choices of all players, and by $\tilde{\mathcal{T}}$ the family of sets that have already been opened. Note that $\tilde{\mathcal{T}}$ may include sets that have been opened, and later "deserted", following strategy changes performed by users. The potential function $\Phi(\mathcal{T}, \tilde{\mathcal{T}})$ defined for our integral set cover game is the following:

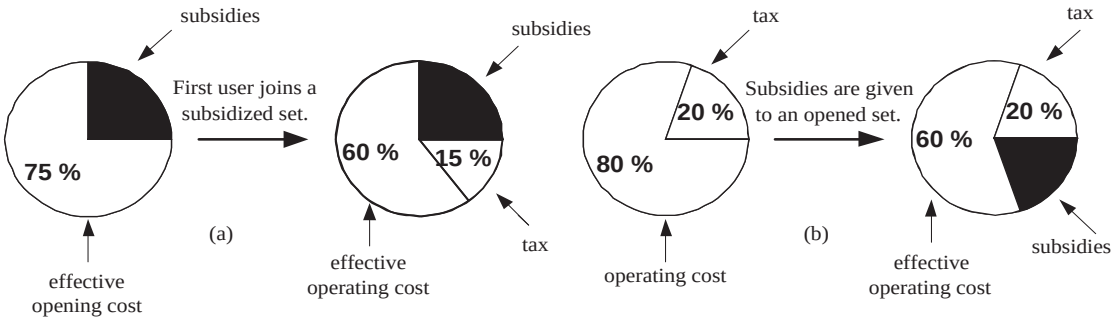

Figure 3: (a) Set s has a fully subsidized fraction of 1/4. The opening cost of the remaining fraction of s is thus equal to $3/4c_s$. Assuming the tax is 20%, the first user joining this fraction pays a tax of $3/20c_s$, and an operating cost of $3/5c_s$. A similar scenario is represented by Figure 2a. (b) Set s' is not subsidized, used by a single user. Assuming the tax percentage is 20%, the user pays a tax of $1/5c_{s'}$ and an operating cost of $4/5c_{s'}$. Now, in case a value of $1/5c_{s'}$ is given as subsidies, the operating cost of the set is reduced to $3/5c_{s'}$. A similar scenario is represented by Figure 2b.

$$\Phi(\mathcal{T}, \tilde{\mathcal{T}}) = \sum_{s \in \tilde{\mathcal{T}}} \epsilon \cdot \hat{c}_s + \sum_{s \in \mathcal{T}} \left(\sum_{j=1}^{n_s} \frac{(1-\epsilon)\hat{c}_s}{j} \right).$$

The potential of the fractional model follows directly, as each fraction λ_s of set s can be considered as a different set, with opening cost $\lambda_s \cdot c_s$ and operating cost $\lambda_s \cdot (1-\epsilon)c_s$.

Theorem 3. *A Nash equilibrium exists for every instance of the set cover game. Moreover, this game always converges to Nash equilibrium via best response dynamics.*

The proof of theorem 3 appears in Appendix A.

3 The Fractional Model

In this section we design a fractional subsidy algorithm that is executed by the central authority. The algorithm receives as input a parameter $\epsilon \leq 1$, and generates a solution with price of anarchy $O(\frac{\log f}{\epsilon})$, and taxation ratio ϵ. The subsidy algorithm runs in iterations, where each iteration corresponds to a new set fraction purchased by some user. This can be the case either when a new user joins the system, or when an existing user changes its strategy. In each such case, the user pays as tax an ϵ fraction of its payment. The amount of subsidies given in each iteration is bounded by the amount of collected taxes, thus allowing the subsidies to be fully financed from taxes. The central authority does not determine which sets are chosen to the cover by the users. The only guarantee is that the users act according to their best response by choosing a fractional cover of minimum cost. Each iteration of the algorithm solely consists of an update of the subsidies. Consider an element (user) e that either joins the system or changes its strategy. There are four different types of set fractions that can be chosen by e.

1. Fractions that are fully subsidized. These fractions have zero cost.

2. Fractions that are not subsidized, yet are used by other users. A user joining such fractions does not have to pay any tax. The operating cost of such a fraction is evenly split between its users.

3. Fractions that are not subsidized and are **not** used by other users. A user choosing these fractions will have to pay their full cost (tax and a full operating cost).

4. Fractions that have been previously opened, but are currently **not** used by any user (users left them following strategy changes). A user joining such a fraction does not have to pay any tax (as a tax was paid when opening it for the first time), but has to pay its full operating cost.

The minimum cost cover chosen by element e consists of the lowest cost combination of fractions of sets adding up to one, while taking into account the charging associated with the four types of set fractions. In case the element chooses a second-type fraction, its payment will lower the payments of other elements using the fraction, but will not have any effect on the cost of the solution. In addition, we do not consider in our analysis the gain from fractions that are "deserted" following strategy changes performed by users (that is, fractions that have been opened, but left by all their users). As we do not reduce the cost of our solution when such fractions are unused, we do not take them into account when an element reuses them by choosing a fourth-type fraction. Moreover, as both fractions of type three and four are chosen from the minimum cost feasible set that covers e, any user that already opened a new fraction of a set in the past will always prefer to return to this fourth-type fraction (that requires no tax payment), before opening new third-type fractions (that do require tax payment).

Thus, following the best response of user e, the total cost of the solution increases only due to third-type fractions. Note that a third-type fraction is chosen by e in order to "complete" its cover, after choosing all possible first, second, and fourth-type fractions of lower cost. Let ρ be the third-type fraction chosen by e. User e chooses the fraction ρ from the minimal cost set that covers it. Let $c_{\min}$ be the opening cost of this minimal set. We refer to ρ as the *greedy choice*, or *greedy cover* of the user, and to $\rho \cdot c_{\min}$ as its *greedy cost*. Let x_s be the fraction of set s that is subsidized by the central authority. Initially, $x_s = 0$ for all sets, and the dual variables $y_e = 0$ for all elements. The algorithm that updates the subsidies offered by the central authority is the following:

Fractional Subsidy Algorithm (with input ϵ):
When user e purchases a new (third-type) set fraction:

1. $y_e \leftarrow y_e + \epsilon \cdot \rho \cdot c_{\min}$
2. For each set s that contains e:

$$x_s \leftarrow x_s \cdot \left(1 + \frac{\epsilon \cdot \rho \cdot c_{\min}}{2c_s}\right) + \frac{\epsilon \cdot \rho \cdot c_{\min}}{f \cdot 2c_s}$$

The variables y_e are the variables of the dual linear program of the fractional set cover problem (Figure 1). These variables are used to maintain an (almost) feasible dual solution. The cost of the dual solution allows us to bound both the price of anarchy and the taxation ratio of the algorithm. Note that the value of the primal variables w_s, indicating the fraction of set s that is taken to the cover (Figure 1), is determined both by the third-type fractions chosen by the user, and by the subsidized fractions x_s.

Let Δx_s^i be the change of x_s in the ith iteration (i.e., the additional subsidy given to set s). The amount of subsidies given in the ith iteration is $\sum_{s \in \mathbb{S}} \Delta x_s^i c_s$. We show that this amount is

bounded by $\epsilon \cdot \rho \cdot c_{\min}$, which is the tax paid by the user. To do so, we establish a relationship between the fractional greedy cost G, the fractional subsidy cost F and the total profit D of the dual solution we produce. Note that in each iteration the amount of taxes collected is exactly the change in the dual cost ($\epsilon \cdot \rho \cdot c_{\min}$). Let ΔG_i, ΔF_i and ΔD_i be the change of the fractional greedy cost, the fractional subsidy cost, and the dual cost, respectively, in the ith iteration.

Lemma 4. *In each iteration i, $\Delta G_i = \frac{1}{\epsilon}\Delta D_i$, and $\Delta F_i \leq \Delta D_i$. Thus, $\Delta F_i / \Delta G_i \leq \epsilon$.*

Proof. In each iteration the greedy cost $\Delta G_i = \rho \cdot c_{\min}$, and $\Delta D_i = \epsilon \cdot \rho \cdot c_{\min}$. Thus, $\Delta G_i = \frac{1}{\epsilon} \cdot \Delta D_i$. In case $\sum_{s|e\in s} x_s > 1$, element e is covered by fully subsidized set fractions and thus $\rho = 0$. Thus, we get that in each iteration, the subsidy cost, ΔF_i, is:

$$\sum_{s|e\in s} c_s \frac{\epsilon \cdot \rho \cdot c_{\min}}{2c_s}\left(x_s + \frac{1}{f}\right) \leq \epsilon \cdot \rho \cdot c_{\min} \leq \Delta D_i.$$

□

As ΔD_i equals the amount of new taxes collected in the ith iteration, ΔF_i is the subsidy cost in the ith iteration, and ΔG_i is the opening cost of the new set fraction purchased in the same iteration, the next corollary follows directly.

Corollary 5. *The taxation ratio of the fractional subsidy algorithm is ϵ. Moreover, the subsidy cost is bounded by the amount of taxes collected.*

Lemma 6. *The dual solution D produced is feasible up to factor of $O(\log f)$.*

Proof. To prove the lemma, consider the dual constraint of a set s, and consider the fraction of s subsidized by the central authority. We prove by induction, that for all sets s, x_s is at least: $x_s \geq \frac{1}{f}\left(2^{\frac{1}{2c_s}\sum_{e\in s} y_e} - 1\right)$.

Initially, this inequality holds trivially. Consider an iteration in which we increase y_e by Δy_e and also increase the value of x_s. Let x_s and x'_s be the values in the beginning and at the end of the iteration, respectively. Similarly, let y_e and y'_e be the values in the beginning and at the end of the iteration, respectively (where $y'_e = y_e + \Delta y_e$). Then:

$$\begin{aligned}
x'_s &= x_s \cdot \left(1 + \frac{\epsilon \cdot \rho \cdot c_{\min}}{2c_s}\right) + \frac{\epsilon \cdot \rho \cdot c_{\min}}{f \cdot 2c_s} \\
&= x_s \cdot \left(1 + \frac{\Delta y_e}{2c_s}\right) + \frac{\Delta y_e}{f \cdot 2c_s} \\
&\geq \frac{1}{f}\left(2^{1/(2c_s)\sum_{e\in s} y_e} - 1\right) \cdot \left(1 + \frac{\Delta y_e}{2c_s}\right) + \frac{\Delta y_e}{f \cdot 2c_s} \\
&= \frac{1}{f}\left(2^{1/(2c_s)\sum_{e\in s} y_e} \cdot \left(1 + \frac{\Delta y_e}{2c_s}\right) - 1\right) \cdot \\
&\geq \frac{1}{f}\left(2^{1/(2c_s)\sum_{e\in s} y_e} \cdot 2^{\Delta y_e/(2c_s)} - 1\right) = \frac{1}{f}\left(2^{1/(2c_s)\sum_{e\in s} y'_e} - 1\right).
\end{aligned}$$

The first inequality follows by the induction hypothesis. The second inequality follows since $2^y \leq 1 + y$ for $y \leq 1$ (note that as $\epsilon, \rho \leq 1$, and $c_{\min} \leq c_s$, it holds that $\frac{\epsilon \cdot \rho \cdot c_{\min}}{c_s} \leq 1$). Finally, it is easy to observe that x_s is at most 3 (in the beginning of the iteration, $x_s < 1$, as s is not

fully subsidized, and its value is increased up to 3 following the current iteration). Thus we get that $\frac{1}{f}\left(2^{1/(2c_s)\sum_{e\in s} y_e} - 1\right) \leq 3$. Therefore, the sum of the variables y_e in the dual constraint corresponding to s is at most $2c_s \cdot \log_2(1+3f) = O(\log f) \cdot c_s$. □

Theorem 7. *The price of anarchy of the final solution is $O(\frac{\log f}{\epsilon})$, and the taxation ratio is ϵ.*

Proof. The taxation ratio follows from Corollary 5. Let D' be a feasible dual solution obtained from D by dividing it by $O(\log f)$. The total cost of the solution is bounded by the sum of the subsidies given by the central authority and the total greedy cost. By Lemmas 4 and 6 we get that the total cost of the solution is then at most:

$$G + F = O\left((1+\frac{1}{\epsilon})\log f\right) D' = O\left(\frac{\log f}{\epsilon}\text{OPT}\right).$$

□

4 The Integral Model

In this section we show how to obtain a subsidy algorithm for the integral version of the problem, which requires several more ideas and a careful analysis. The algorithm receives as input a parameter $\epsilon \leq 1$, and generates a solution with price of anarchy $O\left(\frac{1}{\epsilon}\log f \log\left(\frac{n}{\epsilon}\right)\right)$, and taxation ratio ϵ.

Let OPT be the cost of an optimal integral solution. We design a subsidy algorithm that computes a solution with the properties stated above, given the value of OPT. Note that we can assume (using doubling) that the value of OPT is known up to a factor of 2. The complete subsidy algorithm runs in phases, as follows. We start by guessing $\alpha = \min_{s\in\mathbb{S}} c_s$. If it turns out that the total cost of the solution exceeds $\Theta(\alpha\frac{\log f \log(\frac{n}{\epsilon})}{\epsilon})$, we update the value of α by doubling it, and start a new phase by restarting the algorithm from the current event. Since the success of our algorithm is guaranteed whenever $\alpha \geq$ OPT, then it holds in the last phase that $\alpha \leq 2$OPT. Therefore, the total cost of the solution is the sum of a geometric sequence which is at most twice the bound on the cost of the last phase of our algorithm. Moreover, this does not influence the taxation ratio, that is ϵ in each such phase separately. Note that as we guess the value of the optimum solution, we can ignore all sets with cost greater than α, since these sets cannot belong to an optimal solution (and we thus assume that $\alpha \geq \max_{s\in\mathbb{S}} c_s$).

The algorithm maintains a variable $x_s \geq 0$ for each $s \in \mathbb{S}$, and updates it as in the fractional case. Unlike the fractional case, these variables do not denote (fractional) subsidies. Rather, the value of x_s is used in order to determine whether the set s should be (fully) subsidized. Let $x_j = \sum_{s\in\mathbb{S}_j} x_s$ for each element $j \in N$, where $\mathbb{S}_j$ denotes the collection of sets containing element j. We define $\mathbb{C}$ to be the family of sets in $\mathbb{S}$ that are chosen to the cover, either by the greedy choices of the users, or by the central authority, and define $\tilde{\mathbb{C}} \subseteq \mathbb{C}$ to be the family of sets that are (fully) subsidized. We denote by C and $\tilde{C}$ the set of all elements covered by the members of $\mathbb{C}$ and $\tilde{\mathbb{C}}$, respectively. The following potential function is used throughout the algorithm:

$$\Phi(\epsilon) = \sum_{j\notin\tilde{C}} \exp\left((x_j - 1)\cdot\ln\left(\frac{e\cdot n}{\epsilon}\right)\right) + \exp\left(\frac{1}{2\alpha}\sum_{s\in\mathbb{S}}\left[c_s\cdot\mathcal{I}_{\tilde{\mathbb{C}}}(s) - \frac{3}{2}x_s c_s\cdot\ln\left(\frac{e\cdot n}{\epsilon}\right)\right] - \epsilon\right).$$

The function $\mathcal{I}_{\tilde{\mathbb{C}}}$ above is the characteristic function of $\tilde{\mathbb{C}}$, that is, $\mathcal{I}_{\tilde{\mathbb{C}}}(s) = 1$ if $s \in \tilde{\mathbb{C}}$, and $\mathcal{I}_{\tilde{\mathbb{C}}}(s) = 0$ otherwise. The potential function is used to determine whether a set s should be subsidized. More specifically, after increasing the value x_s, the set s is added to the cover $\tilde{\mathbb{C}}$ (that is, s is subsidized), only if as a result the potential function decreases. Throughout the analysis of the algorithm, the first term of the potential function ensures that whenever the fraction assigned to an element j is at least 1 (that is, $x_j \geq 1$), then j is covered by a fully subsidized set. The second term is used to both bound the cost of the subsidized sets and the total cost of the solution. This part of the potential function was carefully adjusted, so that both the total cost of the solution and the subsidy ratio are maintained. Several constants that are used later on by the algorithm were also chosen carefully so that the subsidy ratio is exactly ϵ.

Consider a user e that either joins the system, or performs a best-response move. In either case, e chooses an integral cover of minimum cost, i.e., it chooses a min cost set covering it. If this set is either subsidized, or used by users that joined the system previously, then the total solution cost does not increase and the subsidy algorithm does nothing. In addition, we do not consider in our analysis the gain from sets that are "deserted" following strategy changes performed by users. As we do not reduce the cost of our solution when such sets are unused, we do not take them into account when an element reuses them. Moreover, any user that already opened a set in the past will prefer to return to this "deserted" set (requiring no tax payment), instead of opening a new one (requiring tax payment). In case the user chooses a new set, that is, a set that is neither subsidized, "deserted", nor used by other users, we implement the following subsidy algorithm:

Integral Subsidy Algorithm (with input ϵ):
Let $\epsilon'' = \frac{1}{16}\epsilon$ and let $\epsilon' = \frac{\epsilon}{3\ln\left(\frac{e \cdot n}{\epsilon''}\right)}$.
Let s' be the new set chosen by the user and let $c_{\min}$ be the cost of the set:

1. $y_e \leftarrow y_e + \epsilon' \cdot c_{\min}$
2. For each set s that contains e:
   (a) $x_s \leftarrow x_s \cdot \left(1 + \frac{\epsilon' c_{\min}}{2c_s}\right) + \frac{\epsilon' \cdot c_{\min}}{f \cdot 2c_s}$.
   (b) Subsidize the full cost of set s (add it to $\tilde{\mathbb{C}}$) if by doing so the value of the potential function $\Phi(\epsilon'')$ is at most its value before the increment of x_s.

The algorithm updates the variables x_s each time a user purchases a new set. In each such iteration the tax collected from the user is $\epsilon c_{\min}$. We will show that the total subsidy given by the algorithm is at most the amount of tax that was collected until that time. The analysis of our algorithm's performance is based on the following lemma.

Lemma 8. *For any value $\epsilon \leq 1$, $\Phi(\epsilon)$ satisfies the following properties:*

1. *At start $\Phi(\epsilon) \leq 1$, and at any time during the execution of the algorithm $\Phi(\epsilon) > 0$.*
2. *Each time the fraction x_s of a set s is increased by the algorithm, then either adding it to $\tilde{\mathbb{C}}$, or not adding it, does not increase the value of $\Phi(\epsilon)$.*

Proof. In order to simplify notation, we use in the proof Φ instead of $\Phi(\epsilon)$. We prove the two parts of the lemma:

Proof of (1): At start,

$$\Phi = n \cdot \exp\left(-\ln\left(\frac{e \cdot n}{\epsilon}\right)\right) + \exp(-\epsilon) = n \cdot \frac{\epsilon}{e \cdot n} + \exp(-\epsilon) = \frac{\epsilon}{e} + \exp(-\epsilon) \leq 1.$$

The last equality follows since $\epsilon \leq 1$. It is easy to verify that each term in the potential function is always positive, and therefore the potential function is always positive.

Proof of (2): The proof is probabilistic. We prove that, adding s to $\tilde{\mathbb{C}}$ with probability p and not adding it with probability $(1-p)$, decreases the expected cost of the potential function Φ. Therefore, at least one of these options does not increase the value of Φ. Let δ_s be the value by which the fraction of set s is incremented. We choose to include s in $\tilde{\mathbb{C}}$ with probability $p = 1 - \exp(-\delta_s \ln\left(\frac{e \cdot n}{\epsilon}\right))$. By linearity of expectation, we may consider each term of the potential function separately.

We consider the contribution of some element $j \notin \tilde{\mathbb{C}}$ to the first term of the potential function. If $j \notin s$, then the first term remains as is. Otherwise, the expected value of its contribution is:

$$\begin{aligned}
&(1-p) \cdot \exp\left((x_j + \delta_s - 1) \cdot \ln\left(\frac{e \cdot n}{\epsilon}\right)\right) + p \cdot 0 \\
&= \exp\left(-\delta_s \ln\left(\frac{e \cdot n}{\epsilon}\right)\right) \cdot \exp\left((x_j + \delta_s - 1) \cdot \ln\left(\frac{e \cdot n}{\epsilon}\right)\right) = \exp\left((x_j - 1) \cdot \ln\left(\frac{e \cdot n}{\epsilon}\right)\right).
\end{aligned}$$

It remains to bound the expected value of the second term of the potential function. Let $\tilde{\mathbb{C}}$ be the sets that are fully subsidized before the current iteration, where the value of x_s is increased. The set s is not in $\tilde{\mathbb{C}}$, as otherwise the current element would already be covered by a fully subsidized set and the current iteration would not have been initiated. Let T be the value of the second term of the potential function before increasing the fraction x_s:

$$T = \exp\left(\frac{1}{2\alpha} \sum_{s \in \mathbb{S}} \left[c_s \mathcal{I}_{\tilde{\mathbb{C}}}(s) - \frac{3}{2} x_s c_s \cdot \ln\left(\frac{e \cdot n}{\epsilon}\right)\right] - \epsilon\right).$$

Let $\tilde{\mathbb{C}}'$ be the sets that are fully subsidized after the current iteration. Therefore, $\tilde{\mathbb{C}}' = \tilde{\mathbb{C}} \cup \{s\}$ with probability p, and $\tilde{\mathbb{C}}' = \tilde{\mathbb{C}}$ with probability $(1-p)$. Let T' denote the value of the second term with respect to the cover $\tilde{\mathbb{C}}'$, and $\mathbb{E}[\cdot]$ denote the expectation value. Therefore,

$$\mathbb{E}[T'] = T \cdot \exp\left(-\frac{1}{2\alpha} \cdot \frac{3}{2} \delta_s c_s \cdot \ln\left(\frac{e \cdot n}{\epsilon}\right)\right) \cdot \mathbb{E}\left[\exp\left(\frac{1}{2\alpha} \cdot c_s \mathcal{I}_{\tilde{\mathbb{C}}'}(s)\right)\right]. \tag{1}$$

We would like to bound the term $\mathbb{E}\left[\exp\left(\frac{1}{2\alpha} c_s \mathcal{I}_{\tilde{\mathbb{C}}'}(s)\right)\right]$. As $\mathcal{I}_{\tilde{\mathbb{C}}'}(s) = 1$ with probability p and $\mathcal{I}_{\tilde{\mathbb{C}}'}(s) = 0$ with probability $(1-p)$, we get that

$$
\begin{aligned}
& \mathbb{E}\left[\exp\left(\frac{1}{2\alpha}\cdot c_s \mathcal{I}_{\tilde{\mathbb{C}}'}(s)\right)\right] \\
= \;& \exp\left(-\delta_s \ln\left(\frac{e\cdot n}{\epsilon}\right)\right) + \left(1-\exp\left(-\delta_s \ln\left(\frac{e\cdot n}{\epsilon}\right)\right)\right)\cdot \exp\left(\frac{c_s}{2\alpha}\right) \\
\leq \;& 1-\delta_s \ln\left(\frac{e\cdot n}{\epsilon}\right) + \delta_s \ln\left(\frac{e\cdot n}{\epsilon}\right)\exp\left(\frac{c_s}{2\alpha}\right) \qquad (2)\\
= \;& 1+\delta_s \ln\left(\frac{e\cdot n}{\epsilon}\right)\cdot\left(\exp\left(\frac{c_s}{2\alpha}\right)-1\right) \qquad (3)\\
\leq \;& 1+\delta_s \ln\left(\frac{e\cdot n}{\epsilon}\right)\cdot\frac{3c_s}{4\alpha} \leq \exp\left(\frac{3\delta_s c_s \cdot \ln\left(\frac{e\cdot n}{\epsilon}\right)}{4\alpha}\right). \qquad (4)
\end{aligned}
$$

Here, (2) follows since for all $y \geq 0$ and $z \geq 1$, $e^{-y}+(1-e^{-y})\cdot z \leq 1-y+y\cdot z$, (4) follows since $e^y - 1 \leq 3y/2$ for all $0 \leq y \leq 1/2$ (recall that $\alpha \geq \max_{s\in\mathbb{S}} c_s$); and since $1+y \leq e^y$ for all $y \geq 0$. Plugging the above term in (1), we conclude that the expected value of the second term of the potential function after the current iteration is at most

$$
\mathbb{E}[T'] \leq T\cdot \exp\left(-\frac{1}{2\alpha}\cdot\frac{3}{2}\delta_s c_s \cdot \ln\left(\frac{e\cdot n}{\epsilon}\right)\right)\cdot\exp\left(\frac{3\delta_s c_s\cdot\ln\left(\frac{e\cdot n}{\epsilon}\right)}{4\alpha}\right) = T,
$$

completing the proof of the second part of Lemma 8. □

By Lemma 8 the algorithm is well defined throughout the execution of the algorithm, and it follows that $\Phi(\epsilon)$ is monotonically non-increasing. Using Lemma 8, we now prove our main Theorem:.

Theorem 9. *For any $\epsilon \leq 1$, the price of anarchy of the solution is $O\left(\frac{\log f \log\left(\frac{n}{\epsilon}\right)}{\epsilon}\right)$, and the taxation ratio is ϵ.*

Proof. Let F be the fractional subsidy cost, that is $F = \sum_{s\in\mathbb{S}} x_s c_s$. Let I be the integral subsidy cost, $\sum_{s\in\tilde{\mathbb{C}}} c_s$. Let G be the total cost of the greedy integral choices of the users. Finally, let D be the total profit of the dual solution produced by the algorithm. The total cost of the solution is the sum of the cost of the greedy choices and the cost of the subsidies (that is, $(G+I)$).

By the first part of Lemma 8, the value of the potential function stays at most 1 during the execution of the subsidy algorithm. Therefore, if during the execution, $x_j \geq 1$ for some user j, then $j \in \tilde{C}$, since otherwise the contribution of the term $\exp\left(\ln\left(\frac{e\cdot n}{\epsilon''}\right)[x_j - 1]\right)$ itself would be at least 1. That is, if a new user j arrives, and $x_j \geq 1$, then j is covered by a fully subsidized set, and its greedy cost is zero. Due to this property, Lemmas 4 and 6 hold for the integral model as well, with respect to ϵ'. Thus, we get that

- In each iteration i, $\Delta G_i = \frac{1}{\epsilon'}\cdot\Delta D_i$, and $\Delta F_i \leq \Delta D_i$. Hence, we get that $F \leq D$ and that $G = D/\epsilon'$.
- The dual solution D is feasible up to factor of $O(\log f)$.

Next, we prove that the price of anarchy achieved by our algorithm is $O(\frac{\log f \log\left(\frac{n}{\epsilon}\right)}{\epsilon})$. Our subsidy algorithm runs with the function $\Phi(\epsilon'')$. By the second term of the potential function, we get that:

$$\exp\left(\frac{1}{2\alpha}\sum_{s\in\mathbb{S}}\left[c_s\mathcal{I}_{\tilde{\mathbb{C}}}(s) - \frac{3}{2}x_s c_s \cdot \ln\left(\frac{e\cdot n}{\epsilon''}\right)\right] - \epsilon''\right) \leq 1.$$

Therefore, it holds that:

$$\begin{aligned} I &= \sum_{s\in\mathbb{S}} c_s\mathcal{I}_{\tilde{\mathbb{C}}}(s) \leq \sum_{s\in\mathbb{S}}\frac{3}{2}x_s c_s \cdot \ln\left(\frac{e\cdot n}{\epsilon''}\right) + 2\alpha\epsilon'' && (5)\\ &\leq \sum_{s\in\mathbb{S}}\frac{3}{2}x_s c_s \cdot \ln\left(\frac{e\cdot n}{\epsilon''}\right) + 4\epsilon''\cdot \text{OPT} = \frac{3}{2}F\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + 4\epsilon''\cdot\text{OPT}. && (6)\end{aligned}$$

Thus, we get that the total solution cost is at most:

$$\begin{aligned} G + I &\leq \frac{D}{\epsilon'} + \frac{3}{2}D\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + 4\epsilon''\cdot\text{OPT} && (7)\\ &= \frac{1}{\epsilon}\cdot 3D\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + \frac{3}{2}D\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + 4\epsilon''\cdot\text{OPT} && (8)\\ &\leq \frac{1}{\epsilon}O(\log f)\cdot\text{OPT}\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + O(\log f)\cdot\text{OPT}\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + 4\epsilon''\cdot\text{OPT} && (9)\\ &= O\left(\frac{\log f\log\left(\frac{n}{\epsilon}\right)}{\epsilon}\right)\cdot\text{OPT}.\end{aligned}$$

Inequality (7) follows since $F \leq D$ and G is added to both sides of (6). Inequality (8) follows by plugging the value $\epsilon' = \frac{\epsilon}{3\ln\left(\frac{e\cdot n}{\epsilon''}\right)}$ and (9) follows by since the dual solution D is feasible up to factor of $O(\log f)$.

We now prove that the ratio between the sum of the subsidies offered, I, and the total payments of the users, G, along the game course, is at most ϵ. Thus, by collecting as taxes a fraction equal to ϵ of the payment made by the user in each iteration where a new set is purchased, the central authority is able to finance the subsidies by the taxes collected from the users. We thus get that the taxation ratio is ϵ. Going back to Inequality (6), we get that:

$$\begin{aligned} I &\leq \frac{3}{2}F\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + 4\epsilon''\cdot\text{OPT} \leq \frac{3}{2}D\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + 4\epsilon''\cdot\text{OPT} && (10)\\ &\leq \frac{3}{2}G\epsilon'\cdot\ln\left(\frac{e\cdot n}{\epsilon''}\right) + 4\epsilon''\cdot(G+I) = \frac{3}{4}\epsilon\cdot G + \frac{1}{4}\epsilon\cdot I. && (11)\end{aligned}$$

Inequality (10) follows as $F \leq D$. Inequality (11) follows since $G = \frac{D}{\epsilon'}$, the fact that $\text{OPT} \leq G + I$ (since G and I together form a feasible integral solution), and setting $\epsilon'' = \frac{\epsilon}{16}$. Simplifying the last inequality, we get that for any value $\epsilon \leq 1$:

$$\frac{I}{G} \leq \frac{\frac{3}{4}\epsilon}{1-\frac{1}{4}\epsilon} \leq \epsilon.$$

□

5 Conclusions

We considered a non-cooperative set cover game in which each user acts selfishly and chooses a cover that minimizes his cost. We focused on a dynamic setting, where players follow the natural game course induced by *best-response dynamics*. We took an approach that does not rely on any prior knowledge of the set of users that actually participates in the game, and thus explored a natural setting in which users join the game starting from an empty configuration.

We designed a dynamic subsidy mechanism, where a central authority creates incentives for users to purchase services by offering subsidies that reduce their price, in order to improve on the system performance. The subsidies are financed by taxes collected from the users on purchased sets. We addressed both an integral model in which sets can only be fully bought and each element is covered by a single set, and a fractional model in which a fraction of a set can be bought and each user can be covered by several sets. For both models we proposed a mechanism that achieves our main goal which is keeping the price of anarchy as low as possible, while collecting as taxes (and spending in subsidies) only a small fraction of the sum of the payments of the users.

To the best of our knowledge, this is the first work to suggest a *dynamic* online subsidy mechanism, where there are no game rules or system rules to be defined beforehand. Rather, the policy of the central authority changes over time, and is only determined by the current state of the system. As being the first such suggested setting, we believe our work leaves room for further research. One main direction is to apply similar mechanisms for other problems, for example, cost-sharing problems in settings with congestion. Another direction for future work is the time convergence to a Nash equilibrium, which remains to be explored.

References

[1] N. Alon, B. Awerbuch, Y. Azar, N. Buchbinder and J. Naor. The online set cover problem. *Proc. of the 35th Annual ACM Symposium on the Theory of Computation*, pp. 100–105, 2003. To appear, SIAM J. Comput.

[2] N. Alon, B. Awerbuch, Y. Azar, N. Buchbinder, and J. Naor. A general approach to online network optimization problems. *ACM Transactions on Alg.*, Vol. 2, pp. 640–660, 2006.

[3] E. Anshelevich, A. Dasgupta, J. Kleinberg, É. Tardos, T. Wexler and T. Roughgarden. The price of stability for network design with fair cost allocation. *Proc. of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pp. 295–304, 2004.

[4] N. Buchbinder and J. Naor. Online primal-dual algorithms for covering and packing problems. *Proc. of the 13th Annual European Symposium on Algorithms*, pp. 689–701, 2005.

[5] I. Caragiannis, C. Kaklamanis and P. Kanellopoulos. Taxes for linear atomic congestion games. *Proc. of the 14th conference on Annual European Symposium*, pp. 184–195, 2006.

[6] M. Charikar, H. Karloff, C. Mathieu, M. Saks and J. Naor. Online multicast with egalitarian cost sharing, *Proc. of the twentieth Annual Symposium on Parallelism in Algorithms and Architectures*, pp. 70–76, 2008.

[7] C. Chekuri, J. Chuzhoy, L. Lewin-Eytan, J. Naor and A. Orda, Non-Cooperative multicast and facility location games", IEEE Journal on Selected Areas in Communication (Special Issue on Non-Cooperative Behavior in Networking), Vol. 25 (2007), pp. 1193-1206.

[8] G. Christodoulou, E. Koutsoupias and A. Nanavati. Coordination mechanisms. *Proc. of the 31st Inter. Colloq. on Automata, Languages and Programming*, pp. 345–357, 2004.

[9] M. Chrobak and E. Koutsoupias. Coordination mechanisms for congestion games. *ACM SIGACT News*, Vol. 35(4), pp. 58–71, 2004.

[10] R. Cole, Y. Dodis, and T. Roughgarden. How much can taxes help selfish routing? *Proc. of the 4th ACM Conference on Electronic Commerce*, pp. 98-107, 2003.

[11] L.K.Fleischer, K. Jain and M. Mahdian. Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. *Proc. of the 45th Annual Symposium of the Foundations of Computer Science*, pp. 277–285, 2004.

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

[13] N. Immorlica, L. Li, V. Mirrokni and A. Schulz. Coordination mechanisms for selfish scheduling. *Proc. of the 1st Inter. Workshop on Internet and Network Econonomics*, pp. 55-69, 2005.

[14] N. Nisan and A. Ronen. Algorithmic mechanism design. *Proc. of the 31th annual ACM Symp. on the Theory of Computation - STOC'99*, pp. 129–140, 1999.

[15] R.W. Rosenthal. A class of games possessing pure strategy Nash equilibria. *International Journal of Game Theory*, Vol. 2, pp. 65–67, 1973.

[16] T. Roughgarden. Stackelberg scheduling strategies. *Proc. of the 33rd annual ACM Symp. on the Theory of Computation - STOC'01*, pp. 104–113, 2001.

[17] C. Swamy. The effectiveness of Stackelberg strategies and tolls for network congestion games. *Proc. of the 18th annual ACM-SIAM Symp. on Discrete Algorithms*, pp. 1133–1142, 2007.

A Proof of Theorem 3

Theorem 3. *A Nash equilibrium exists for the set cover game. Moreover, this game always converges to Nash equilibrium via best response dynamics.*

Proof. We prove the theorem for the integral model. The proof can be easily extended for the fractional case, as done for the fractional model presented in [7].

We prove that the function Φ is an exact potential for the set cover game. That is, the decrease in the value of the potential function following a move performed by a player is equal to the decrease in the payoff of the respective player. The strategy space of the set cover game consists of all possible feasible covers, and is thus finite. Consequently, as Φ admits a minimal

value over the strategy space, the game possesses a (pure-strategy) equilibrium. Moreover, there is a one-to-one correspondence between Nash equilibrium solutions and the solutions defining a local minimum of the potential function, and convergence of best-response dynamics to Nash equilibrium is thus guaranteed.

Consider a strategy profile $\mathcal{T}$ where s_i is the set chosen by player i. We denote by n_s the number of elements using set s according to $\mathcal{T}$. In case i is the first player to purchase s_i, it pays a tax equal to $\epsilon \cdot c_{s_i}$. Note that as this tax is non-refundable, it cannot be reduced by any change of strategy. The variable cost share of player i is equal to $\frac{(1-\epsilon)\hat{c}_{s_i}}{n_{s_i}}$. Now, assume that this variable sum can be reduced in case i performs a best-response move and switches from set s_i to set s_i^*, resulting in a new strategy profile $\mathcal{T}^*$. We denote by n_s^* the number of elements using set s according to $\mathcal{T}^*$. We compare the change in i's payoff and in the value of the potential function in the two following possible cases:

1. Assume set s_i^* was opened previously, and thus i does not have to pay any tax for using it. Thus, the family of sets $\tilde{\mathcal{T}}^*$ is identical to $\tilde{\mathcal{T}}$. As i performs a best-response move, its payoff is reduced, and thus

$$\frac{(1-\epsilon)\hat{c}_{s_i^*}}{n^*_{s_i^*}} < \frac{(1-\epsilon)\hat{c}_{s_i}}{n_{s_i}}.$$

As $n^*_{s_i^*} = n_{s_i^*} + 1$, it holds that

$$\frac{(1-\epsilon)\hat{c}_{s_i^*}}{n_{s_i^*}+1} < \frac{(1-\epsilon)\hat{c}_{s_i}}{n_{s_i}}. \tag{1}$$

We turn to compare the potential value of $(\mathcal{T}, \tilde{\mathcal{T}})$ and $(\mathcal{T}^*, \tilde{\mathcal{T}}^*)$:

$$\begin{aligned}
\Phi(\mathcal{T}^*, \tilde{\mathcal{T}}^*) &= \sum_{s\in\tilde{\mathcal{T}}^*} \epsilon\cdot\hat{c}_s + \sum_{s\in\mathcal{T}^*}\left(\sum_{j=1}^{n_s^*}\frac{(1-\epsilon)\hat{c}_s}{j}\right) && (2)\\
&= \sum_{s\in\tilde{\mathcal{T}}} \epsilon\cdot\hat{c}_s + \sum_{s\in\mathcal{T}}\left(\sum_{j=1}^{n_s}\frac{(1-\epsilon)\hat{c}_s}{j}\right) + \frac{(1-\epsilon)\hat{c}_{s_i^*}}{n_{s_i^*}+1} - \frac{(1-\epsilon)\hat{c}_{s_i}}{n_{s_i}} && (3)\\
&< \sum_{s\in\tilde{\mathcal{T}}} \epsilon\cdot\hat{c}_s + \sum_{s\in\mathcal{T}}\left(\sum_{j=1}^{n_s}\frac{(1-\epsilon)\hat{c}_s}{j}\right) = \Phi(\mathcal{T}, \tilde{\mathcal{T}}). && (4)
\end{aligned}$$

Equality (3) follows as the taxes paid in $\tilde{\mathcal{T}}$ and $\tilde{\mathcal{T}}^*$ are equal, and due to the change in the number of users of s_i and s_i^* in $\mathcal{T}$ and $\mathcal{T}^*$. Inequality (4) follows from Inequality (1).

2. Assume that player i is the first player that purchases s_i^*. Thus, the family of sets $\tilde{\mathcal{T}}^*$ includes the sets in $\tilde{\mathcal{T}}$, plus the additional set s_i^*. In that case, i has to pay both a tax and a variable sum (the effective operating cost) for using s_i^*. As i performs a best-response move, its payoff is reduced, and thus

$$\epsilon \cdot \hat{c}_{s_i^*} + (1-\epsilon)\hat{c}_{s_i^*} < \frac{(1-\epsilon)\hat{c}_{s_i}}{n_{s_i}}. \tag{5}$$

We turn to compare the potential value of $(\mathcal{T}, \tilde{\mathcal{T}})$ and $(\mathcal{T}^*, \tilde{\mathcal{T}}^*)$:

$$\begin{aligned}
\Phi(\mathcal{T}^*, \tilde{\mathcal{T}}^*) &= \sum_{s \in \tilde{\mathcal{T}}^*} \epsilon \cdot \hat{c}_s + \sum_{s \in \mathcal{T}^*} \left(\sum_{j=1}^{n_s^*} \frac{(1-\epsilon)\hat{c}_s}{j} \right) && (6) \\
&= \sum_{s \in \tilde{\mathcal{T}}} \epsilon \cdot \hat{c}_s + \epsilon \cdot \hat{c}_{s_i^*} + \sum_{s \in \mathcal{T}} \left(\sum_{j=1}^{n_s} \frac{(1-\epsilon)\hat{c}_s}{j} \right) + (1-\epsilon)\hat{c}_{s_i^*} - \frac{(1-\epsilon)\hat{c}_{s_i}}{n_{s_i}} && (7) \\
&< \sum_{s \in \tilde{\mathcal{T}}} \epsilon \cdot \hat{c}_s + \sum_{s \in \mathcal{T}} \left(\sum_{j=1}^{n_s} \frac{(1-\epsilon)\hat{c}_s}{j} \right) = \Phi(\mathcal{T}, \tilde{\mathcal{T}}). && (8)
\end{aligned}$$

Equality (7) follows as in $\tilde{\mathcal{T}}^*$ there is an additional tax of $\epsilon \cdot \hat{c}_{s_i^*}$ compared to the taxes paid in $\tilde{\mathcal{T}}$, and due to the change in the number of users of s_i and s_i^* in $\mathcal{T}$ and $\mathcal{T}^*$. Inequality (8) follows from Inequality (5).

Note that in both cases, the difference between the potential function $\Phi(\mathcal{T}^*, \tilde{\mathcal{T}}^*)$ and $\Phi(\mathcal{T}, \tilde{\mathcal{T}})$ is equal to the difference between player i's payoff with respect to the strategy profiles $\mathcal{T}^*$ and $\mathcal{T}$.

Note that in the second case, following the purchase of a new set s_i^*, this set might be fully subsidized, changing its effective cost to zero. The change in the payoff of player i remains equal to the change in the value of the potential function in this case as well. □