

This is a postprint version of the following document:

Caballero, P., Banchs, A., Veciana, G. y Costa-Pérez, X. (2017). Network Slicing Games: Enabling Customization in Multi-Tenant Mobile Networks. In *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*.

DOI: <https://doi.org/10.1109/INFOCOM.2017.8057046>

© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

# Network Slicing Games: Enabling Customization in Multi-Tenant Mobile Networks

Pablo Caballero\*    Albert Banchs<sup>†</sup>    Gustavo de Veciana\*    Xavier Costa-Pérez<sup>‡</sup>

\*The University of Texas at Austin, Austin, TX. Email: pablo.caballero@utexas.edu, Gustavo@ece.utexas.edu

<sup>†</sup>University Carlos III of Madrid and IMDEA Networks Institute, Madrid, Spain. Email: banchs@it.uc3m.es

<sup>‡</sup>NEC Laboratories Europe, Heidelberg, Germany. Email: xavier.costa@neclab.eu

**Abstract**—Network slicing to enable resource sharing among multiple tenants –network operators and/or services– is considered a key functionality for next generation mobile networks. This paper provides an analysis of a well-known model for resource sharing, the ‘share-constrained proportional allocation’ mechanism, to realize network slicing. This mechanism enables tenants to reap the performance benefits of sharing, while retaining the ability to customize their own users’ allocation. This results in a *network slicing game* in which each tenant reacts to the user allocations of the other tenants so as to maximize its own utility. We show that, under appropriate conditions, the game associated with such strategic behavior converges to a Nash equilibrium. At the Nash equilibrium, a tenant always achieves the same, or better, performance than under a static partitioning of resources, hence providing the same level of protection as such static partitioning. We further analyze the efficiency and fairness of the resulting allocations, providing tight bounds for the price of anarchy and envy-freeness. Our analysis and extensive simulation results confirm that the mechanism provides a comprehensive practical solution to realize network slicing. Our theoretical results also fill a gap in the literature regarding the analysis of this resource allocation model under strategic players.

## I. INTRODUCTION

There is consensus among the relevant industry and standardization communities [1], [2] that a key element in 5G mobile networks will be *network slicing*. The idea is to allow the mobile infrastructure to be “sliced” into logical networks, which are operated by different entities and may be tailored to support specific services. This provides a basis for efficient infrastructure sharing among diverse entities, ranging from classical or virtual mobile network operators to new players that simply view connectivity as a service. Such new players could be, for instance, Over-The-Top (OTT) service providers which use a *network slice* to ensure satisfactory service to their customers (e.g., Amazon Kindle’s support for downloading content or a pay TV channel including a premium subscription). In the literature, the term *tenant* is often used to refer to the owner of a network slice.

A network slice is a collection of resources and functions that are orchestrated to support a specific service. This includes software modules running at different locations as well as the nodes’ computational resources, and communication resources in the backhaul and radio network. The intention is to only provide what is necessary for the service, avoiding unnecessary overheads and complexity. Thus, network slices enable tenants to compete with each other using the same physical

infrastructure, but customizing their slices and network operation according to their market segment’s characteristics and requirements. For instance, slices can be geared at supporting various IoT or M2M applications, such as the connectivity required to realize ‘intelligent’ vehicular systems.

A key problem underlying network slicing is enabling efficient sharing of mobile network resources. One of the approaches considered in 3GPP suggests that resources could be statically partitioned based on fixed ‘network shares’ [3]. However, given that slices’ loads may be spatially inhomogeneous and time varying, it is desirable to allow resource allocations to be ‘elastic’, e.g., dependent on the slices’ loads at different base stations. At the same time, tenants should be protected from one another, and retain the ability to autonomously manage their slice’s resources, in order to better customize allocations to their customers. To that end, it is desirable to adopt resource allocation models in which tenants can communicate their preferences to the infrastructure (say by dynamically subdividing their network share amongst their customers) and then have base stations’ resources allocated according to their preferences (i.e., proportionally to the customers’ shares).

Under such a dynamic resource allocation model, a tenant might exhibit strategic behavior, by adjusting its preferences depending on perceived congestion at resources, so as to maximize its own utility. Such behavior could in turn have adverse effects on the network; for instance, the overall efficiency may be harmed, or one may see instability in slice requests. The focus of this paper is on (i) the analysis and performance of this simple resource allocation model, and (ii) the validation of its feasibility as a means to enable tenants to customize resource allocation within their slice while protecting them from one another.

*Related work:* The resource allocation mechanism informally described above corresponds to a Fisher market, which is a standard framework in economics. In such markets, buyers (in our case slices) have fixed budgets (in our case network shares) and (according to their preferences) bid for resources within their budget, which are then allocated to buyers proportionally to their bids. Analysis of the Fisher market shows that, as long as buyers are price-taking (i.e., they do not anticipate the impact of their bids on the price – in our case, the impact of the slices’ preferences on the overall congestion), the Nash equilibrium is socially optimal,

and distributed algorithms can be easily devised to reach it [4]. This assumption may be reasonable for markets where the impact of a single buyer on a resource's price is negligible, but does not apply to our case where a relatively small number of active tenants might be sharing resources.

There is a substantial literature on Fisher markets with strategic buyers, which, as will be studied in this paper, anticipate the impact of their bids [5]. The analysis, so far, has been limited to the case of buyers with *linear utility* functions of the allocated resources, which can lead to extremely unfair allocations. While such utility functions may be suitable for goods, they are not an appropriate model for tenants wishing to customize allocations amongst their customers. This paper includes a comprehensive analysis for a wide set of slice utility functions, including the convergence of best response dynamics and other results which to our knowledge are new.

A related resource allocation model often considered in the networking field is the so-called 'Kelly's mechanism' [6]; this mechanism allocates resources to players proportionally to their bids and, assuming that they are price-taking, converges to a social optimum. Follow-up work has considered price-anticipating players in this setting; for example, [7] analyze efficiency losses, while [8] devise a scalar-parametrized modification that is once again socially optimal for price-anticipating players. However, in Kelly's mechanism players respond to their payoff (given by the utility minus cost) whereas in our model tenants' behavior is only driven by their utilities (since they have a fixed budget: the network share). Consequently, results on the analysis of Kelly's mechanism are not applicable to our setting.

In the context of the existing resource allocation models described above, this work covers the following gap in the literature: the analysis of *budget-constrained resource allocation* under *price-anticipating users* with *nonlinear utilities*. The reader is referred to the extended version [9] for a table illustrating this gap in the context of existing literature.

From a more practical angle, multi-tenant sharing has been studied from different points of view, including planning, economics, coverage, performance, etc. [10], [11]. This paper focuses specifically on the design of algorithms for resource sharing among tenants, which has been previously addressed by [12]–[15]. The work of [15] considers sharing via a bid-based auction, which may incur substantial overhead and complexity; in contrast, our approach relies on fixed (pre-negotiated) network shares. The works of [12]–[14] also fix a network share per slice, but consider approaches where the infrastructure makes centralized decisions on the resources allocated to each tenant's customers; hence, these approaches do not enable tenants to make their own decisions on how to allocate resources to their customers.

Network slicing has emerged as a desirable feature for 5G [1]. 3GPP has started work on defining requirements for network slicing [2], whereas the Next Generation Mobile Network (NGMN) alliance has identified network sharing among slices (the focus of this paper) as a key issue [16]. In spite of these efforts, most of the work so far has addressed

architectural aspects with only a limited focus on resource allocation algorithms [17], [18]. To the best of our knowledge, this is the first work investigating how to enable tenants to customize their allocations in a dynamic slicing model.

*Key contributions:* The rest of the paper is organized as follows. After introducing our system model (Section II), we show that with the resource sharing model under study, each slice has the ability to achieve the same or better utility than under static resource slicing irrespective of how the other slices behave, which confirms that this model effectively protects slices from one another (Section III-A). Next we show that if tenants exhibit strategic behavior (i.e, optimize their utilities), then (i) a Nash equilibrium exists under mild conditions; and (ii) the system converges to such an equilibrium when tenants sequentially take their best response (Sections III-B and III-C). The resulting efficiency and fairness among tenants are then studied, providing: (i) a tight bound on the Price of Anarchy of the system, and (ii) a bound on the Envy-freeness (Section IV). Our results are validated via simulation, confirming that the approach provides substantial gains, protects network slices from each other, operates close to optimal performance and is effectively envy-free (Section V).

## II. SYSTEM MODEL

We consider a wireless network consisting of a set of resources  $\mathcal{B}$  (the base stations or sectors) shared by a set of network slices  $\mathcal{O}$  (the tenants). At a given point in time, the network supports a set of users  $\mathcal{U}$  (the customers or devices), which can be subdivided into subsets  $\mathcal{U}_b$  (the users at base station  $b$ ),  $\mathcal{U}^o$  (the users of slice  $o$ ) and  $\mathcal{U}_b^o$  (their intersection). We further assume that a user  $u \in \mathcal{U}$  has a mean peak capacity  $c_u$  depending on the choice of modulation and coding at the base station it is associated with. For any user  $u$ , we let  $b(u)$  denote the base station it is currently associated with.

### A. Resource allocation model

As indicated in the introduction, we focus on a well established resource sharing model known in economics as a Fisher market. Hereafter, we will refer to this model as the 'Share-Constrained Proportional Allocation' (SCPA) mechanism.

In our setting, each slice  $o$  is allocated a network share  $s_o$  (corresponding to its budget) such that  $\sum_{o \in \mathcal{O}} s_o = 1$ . The slice is at liberty in turn to distribute its share amongst its users, assigning them weights (corresponding to the bids):  $w_u$  for  $u \in \mathcal{U}^o$ , such that  $\sum_{u \in \mathcal{U}^o} w_u = s_o$ . We let  $\mathbf{w}^o = (w_u : u \in \mathcal{U}^o)$  be the weights of slice  $o$ ,  $\mathbf{w} = (w_u : u \in \mathcal{U})$  those of all slices and  $\mathbf{w}^o = (w_u : u \in \mathcal{U} \setminus \mathcal{U}^o)$  the weights of all users excluding those of slice  $o$ .

We shall assume users are allocated a fraction of resources at their base station proportionally to their weights  $w_u$ . Thus the rate of user  $u$  is given by

$$r_u(\mathbf{w}) = \frac{w_u}{\sum_{v \in \mathcal{U}_{b(u)}} w_v} c_u = \frac{w_u}{l_{b(u)}(\mathbf{w})} c_u$$

where  $l_b(\mathbf{w}) = \sum_{u \in \mathcal{U}_b} w_u$  denotes the overall load at  $b$  (recall that  $c_u$  is the achievable rate if the user had the entire base station to itself).

To implement the above resource allocation, a slice needs to communicate the weights of its users  $\mathbf{w}^o$  to the infrastructure. When selecting its weights, we assume that the slice is aware of the overall load at each base station (indeed, a slice could infer these by varying its users' weights and observing the resulting resource allocations).<sup>1</sup>

In the case where a slice  $o$  is the only one with users at a given base station  $b$ , we shall assume that the slice's users are allocated the entire capacity at that base station independent of their weights. Thus such a slice would set  $w_u = 0$  for these users, allowing them to receive all the resources of this base station without consuming any share. In order to avoid dealing with this special case, and without loss of generality, we will make the following assumption for the rest of the paper.

**Assumption 1.** (*Competition at all resources*) We assume that all resources have active users from at least two slices.

### B. Network Slice Utility and Service Differentiation

Network slices may support services and customers of different types and needs. Alternatively, competing slices with similar customer types may wish to differentiate the service they provide. To that end, we assume each network slice has a *private* utility that reflects the benefit obtained by the slice from a given allocation and is given by

$$U^o(\mathbf{w}) = \sum_{u \in \mathcal{U}^o} \phi_u f_u(r_u(\mathbf{w})),$$

where  $\phi_u$  is the relative priority of user  $u$ , with  $\phi_u \geq 0$  and  $\sum_{u \in \mathcal{U}^o} \phi_u = 1$ , and  $f_u(\cdot)$  is a (concave) utility function associated with the user. In the sequel, we will often focus on the following well-known class of utility functions [19].

**Definition 1.** A network slice  $o$  has a homogenous  $\alpha_o$ -fair utility if for all  $u \in \mathcal{U}^o$  we have that

$$f_u(r_u) = \begin{cases} \frac{(r_u)^{1-\alpha_o}}{(1-\alpha_o)}, & \alpha_o \neq 1 \\ \log(r_u), & \alpha_o = 1. \end{cases}$$

Thus, in our setting, a slice is free to choose different fairness criteria in allocating resources across its users, by selecting the appropriate  $\alpha_o$  parameter. Note that  $\alpha_o = 1$  corresponds to the widely accepted proportional fairness criterion, while  $\alpha_o = 2$  corresponds to potential delay fairness,  $\alpha_o \rightarrow \infty$  to max-min fairness and  $\alpha_o = 0$  to linear sum utility.

A slice can also 'strategically' optimize the weight allocation of its users to maximize its own utility. We will consider such strategic behavior of weight allocations in Section III.

### C. Baseline allocations

Next we introduce two natural resource allocation comparative baselines: socially optimal allocations and static slicing.

<sup>1</sup>It is worth noting that, with the SCPA mechanism under study, the weights of a given tenant are not disclosed to the others, which only see the overall load at each base station.

a) *Socially Optimal Allocations (SO)*: If slices were to share their utility functions with a centralized authority, one could in principle consider a socially optimal allocation of weights and resources. These would be given by the maximizer to the *overall network utility*  $U(\mathbf{w})$  given by (see [14]):

$$\begin{aligned} \max_{\mathbf{w} \geq 0} \quad & U(\mathbf{w}) := \sum_{o \in \mathcal{O}} s_o U^o(\mathbf{w}) \\ \text{s.t.} \quad & r_u(\mathbf{w}) = \frac{w_u}{l_{b(u)}(\mathbf{w})} c_u, \quad \forall u \in \mathcal{U} \\ & \sum_{u \in \mathcal{U}^o} w_u = s_o, \quad \forall o \in \mathcal{O}. \end{aligned}$$

Note that (as in [14]) we have weighted the slices' utilities to reflect their shares (thus prioritizing those with higher shares). We shall denote the resulting optimal weight and resource allocations under the socially optimal allocations by  $\mathbf{w}^*$  and  $\mathbf{r}^* = (r_u^* : u \in \mathcal{U})$ , respectively.

b) *Static Slicing (SS)*: By static slicing (also known as static splitting [20]) we refer to a complete partitioning of resources based on the network shares  $s_o, o \in \mathcal{O}$ . In this setting, each slice  $o$  receives a fixed fraction  $s_o$  of each resource and can unilaterally optimize its weight allocation as follows:

$$\begin{aligned} \max_{\mathbf{w}^o \geq 0} \quad & U^o(\mathbf{w}^o) = \sum_{u \in \mathcal{U}^o} \phi_u f_u(r_u(\mathbf{w}^o)) \\ \text{s.t.} \quad & r_u(\mathbf{w}^o) = \frac{w_u}{\sum_{v \in \mathcal{U}_{b(u)}^o} w_v} s_o c_u \quad \forall u \in \mathcal{U}^o \\ & \sum_{u \in \mathcal{U}^o} w_u = s_o, \end{aligned}$$

where we have abused notation to indicate that, in this case,  $U^o$  and  $r_u$  depend only on  $\mathbf{w}^o$ . We shall denote the resulting optimal weight and resource allocations under static slicing for all slices by  $\mathbf{w}^{ss}$  and  $\mathbf{r}^{ss} = (r_u^{ss} : u \in \mathcal{U})$  respectively, where

$$r_u^{ss} = \frac{w_u^{ss}}{\sum_{v \in \mathcal{U}_{b(u)}^o} w_v^{ss}} s_o c_u \quad \forall u \in \mathcal{U}^o, \forall o \in \mathcal{O}. \quad (1)$$

## III. STRATEGIC BEHAVIOR AND NASH EQUILIBRIUM

Under the SCPA resource allocation model, it is reasonable to assume that a player (network slice) would choose to adjust its weights so as to optimize its utility (and thus the service delivered to its customers). Since the resources allocated to a user depend on the weight allocations of the other slices, such behavior would be predicated on the aggregate weight of the other slices at each resource. From the point of view of slice  $o$ , the overall load at resource  $b$  can be decomposed as

$$l_b(\mathbf{w}) = a_b^o(\mathbf{w}^o) + d_b^o(\mathbf{w}^o)$$

where

$$a_b^o(\mathbf{w}^o) = \sum_{o' \in \mathcal{O} \setminus \{o\}} \sum_{u \in \mathcal{U}_{b'}^{o'}} w_u \quad \text{and} \quad d_b^o(\mathbf{w}^o) = \sum_{u \in \mathcal{U}_b^o} w_u,$$

correspond to the aggregate weight of the other slices and that of slice  $o$ , respectively. As indicated in Section II, we assume  $\mathbf{a}^o(\mathbf{w}^o) = (a_b^o(\mathbf{w}^o) : b \in \mathcal{B})$  are readily available to slice  $o$ .

### A. Gain over Static Slicing

We first analyze if strategic behavior on the part of network slices may result in allocations that are worse than those under static slicing. Note that static slicing provides complete isolation among slices but potentially poor utilization. A critical question is whether dynamic sharing, which achieves a higher resource utilization, also provides the same level of protection. This is confirmed by the following result.

**Lemma 1.** *Consider slice  $o$  and any feasible weight allocation  $\mathbf{w}^o$  for other slices satisfying the network share constraints. Then, there exists a weight allocation  $\mathbf{w}^o$  for slice  $o$ , possibly dependent on  $\mathbf{w}^o$ , such that the resulting weight allocation  $\mathbf{w}$  satisfies  $r_u(\mathbf{w}) \geq r_u^{ss}$  for all  $u \in \mathcal{U}_o$ .*

This lemma is easily shown by choosing  $\mathbf{w}^o$  such that

$$w_u = \frac{w_u^{ss}}{\sum_{u \in \mathcal{U}_o} w_u^{ss}} \frac{a_{b(u)}^o(\mathbf{w}^o)}{\sum_{b' \in \mathcal{B}_o} a_{b'}^o(\mathbf{w}^o)} s_o, \quad \forall u \in \mathcal{U}_o$$

where  $\mathcal{B}^o$  is the set of base stations where slice  $o$  has users. The intuitive interpretation for this choice is that by distributing its weights proportionally to the load at each base station, slice  $o$  can achieve the same resource allocation as static slicing at each base station. Further, by redistributing these allocations amongst its user in the same manner as static slicing, it achieves at least as much rate per user.

It follows immediately from this lemma that under the SCPA resource allocation model, if all slices exhibit strategic behavior attempting to maximize their utilities, they necessarily achieve a higher utility than under static slicing.

**Theorem 1.** *If the game where each network slice maximizes its utility has a Nash equilibrium, then each slice achieves a higher utility than under static slicing.*

Note this result does not require slices to have homogenous or concave utilities, just that they be increasing in the users' rate allocations.

### B. Existence of Nash Equilibrium

Next we study whether there exists a Nash equilibrium (NE) under which no slice can benefit by unilaterally changing its weight allocation. To that end, we first characterize the best response of a slice.

Given the weights of the other slices,  $\mathbf{w}^o$ , the best response of slice  $o$  is the unique maximizer  $\mathbf{w}^o$  of its utility, i.e.,

$$\begin{aligned} \max_{\mathbf{w}^o \geq 0} \quad & \sum_{u \in \mathcal{U}_o} \phi_u f_u \left( \frac{w'_u c_u}{a_{b(u)}^o(\mathbf{w}^o) + d_{b(u)}^o(\mathbf{w}^o)} \right) \\ \text{s.t.} \quad & \sum_{u \in \mathcal{U}_o} w'_u = s_o. \end{aligned}$$

The following lemma characterizes the best response for a network slice with homogenous  $\alpha_o$ -fair utility (see [5] for the best response when  $\alpha_o = 0$ ).

**Lemma 2.** *Suppose slice  $o$  has a homogeneous  $\alpha_o$ -fair utility (with  $\alpha_o > 0$ ). Given the weights of the other slices  $\mathbf{w}^o >$*

*0, slice  $o$ 's best response  $\mathbf{w}^o$  is the unique solution to the following nonlinear set of equations:*

$$w_u = \frac{\beta_u \frac{(a_{b(u)}^o(\mathbf{w}^o))^{\frac{1}{\alpha_o}}}{(a_{b(u)}^o(\mathbf{w}^o) + d_{b(u)}^o(\mathbf{w}^o))^{\frac{2}{\alpha_o} - 1}}}{\sum_{v \in \mathcal{U}_o} \beta_v \frac{(a_{b(v)}^o(\mathbf{w}^o))^{\frac{1}{\alpha_o}}}{(a_{b(v)}^o(\mathbf{w}^o) + d_{b(v)}^o(\mathbf{w}^o))^{\frac{2}{\alpha_o} - 1}}} s_o, \quad \forall u \in \mathcal{U}_o, \quad (2)$$

where  $\beta_u := (\phi_u)^{\frac{1}{\alpha_o}} (c_u)^{\frac{1}{\alpha_o} - 1}$ .

Note that slice  $o$  need only know  $\mathbf{a}^o(\mathbf{w}^o)$  to compute its best response. Building on this characterization, we will study the game in which all slices choose to allocate their weights based on their best response. The following theorem proves that this game admits a Nash equilibrium, i.e., there is a weight allocation  $\mathbf{w}$  such that no slice can improve its utility by modifying its weights unilaterally.<sup>2</sup>

**Theorem 2.** *Suppose all slices have homogenous  $\alpha_o$ -fair utilities (with possibly different  $\alpha_o > 0$ ). Then, there exists a (not necessarily unique) Nash equilibrium satisfying (2) for each slice.*

The proof of this result is technical and been relegated to [9]. The argument proceeds as follows. We consider a perturbed game where an additional slice assigns a weight  $\varepsilon$  at each base station. For this perturbed game, we have concave utilities and compact strategy spaces such that the result of [21] gives existence of a Nash equilibrium.<sup>3</sup> We then consider a sequence of such equilibria as  $\varepsilon \rightarrow 0$ . By compactness of the strategy space it must have a converging subsequence. One can further show that the weight allocations for the perturbed equilibria have uniform positive lower bounds, so the limit of the converging subsequence also has positive weights. Note that (as it can be seen from Lemma 2) slice  $o$ 's best response in the perturbed game  $\mathbf{w}^o$  is a continuous differentiable function of  $\mathbf{w}^o$  as long as  $\mathbf{w}^o > 0$ . It then follows by continuity that the limit of the converging subsequence is a Nash equilibrium.

### C. Convergence of Best Response Dynamics

Below we will consider a best response game wherein slices realize their best responses in rounds; specifically, they update their weights ( $\mathbf{w}^o$ ) sequentially, one at a time and in the same fixed order, in response to the other slices' weights ( $\mathbf{a}^o$ ).

**Theorem 3.** *If slices have homogeneous  $\alpha_o$ -fair utilities, possibly with different  $\alpha_o \in [1, 2]$  for  $o \in \mathcal{O}$ , then the best response game converges to a Nash equilibrium.*

<sup>2</sup>The existence of a NE had already been proven by [4] for the case  $\alpha_o = 0 \forall o$ . Here we extend this result to any combination of  $\alpha_o$  values.

<sup>3</sup>In particular, [21] shows by applying the Kakutani fixed point theorem that there exists a solution to the equations defining a Nash equilibrium. Note that in the case of users with log utilities, the function is not defined for a weight of 0 and hence the conditions of [21] are not satisfied; however, a careful reading of the proof of [21] shows the result still applies.

Note that the value of  $\alpha_o$  impacts a slice's best response and consequently the game dynamics. As seen in Lemma 2, the best response weights are proportional to:

$$w_u \propto g(a_b^o, d_b^o) := \frac{(a_b^o)^{\frac{1}{\alpha_o}}}{(a_b^o + d_b^o)^{\frac{2}{\alpha_o} - 1}},$$

where we have suppressed the dependency of  $a_b^o$  on  $\mathbf{w}^o$  and  $d_b^o$  on  $\mathbf{w}^o$ . The function  $g(\cdot, \cdot)$  has different properties depending on  $\alpha_o$  which are shown in Table II. The regime where  $1 \leq \alpha_o \leq 2$ , considered in Theorem 3, is of particular interest since it includes proportional ( $\alpha_o = 1$ ) and potential delay ( $\alpha_o = 2$ ) fairness. It is known that convergence is not ensured when  $\alpha_o = 0$  for all slices (see [5]); for other regimes, we resort to the simulations results provided in [9], which suggest convergence for any  $\alpha_o > 0$ .

	$\alpha_o = 0$	$0 < \alpha_o < 1$	$1 \leq \alpha_o \leq 2$	$2 < \alpha_o < \infty$
$g$ w.r.t. $d_b^o$	linear	convex	convex	concave
$g$ w.r.t. $a_b^o$	linear	convex	concave	concave
NE existence	✓ [5]	✓ Theorem 2 for heterogeneous $\alpha_o$		
convergence	× [5]	✓ simulations	✓ Theorem 3	✓ simulations

TABLE I: Impact of  $\alpha_o$  on slice's Best Responses.

Perhaps surprisingly, the above result is quite challenging to show. The key challenge lies in the “price-anticipating” aspect of the best response, in which players anticipate the impact of their own allocation.<sup>4</sup> The rest of this section is a sketch of the proof for this result.

We shall denote time as slotted  $\{0, 1, \dots, t, \dots\}$  and assume a single slice makes an update each time slot. Without loss of generality, we will index slices  $\{1, 2, \dots, |\mathcal{O}|\} = \mathcal{O}$  according to their updating order in a round. We let  $\mathbf{w}(t) = (\mathbf{w}^o(t) : o \in \mathcal{O})$  be the weights of all slices at the end the time slot  $t$  update, where  $\mathbf{w}^o(t) = (w_u(t) : u \in \mathcal{U}^o)$ . Suppose that slices have arbitrary positive initial weight vectors at time zero denoted  $\mathbf{w}(0) = (\mathbf{w}^1(0), \mathbf{w}^2(0), \dots, \mathbf{w}^{|\mathcal{O}|}(0))$ . Consequently, slice 1 will update its weights at time slots:  $\{1, |\mathcal{O}|+1, \dots, r \cdot |\mathcal{O}|+1\}$ , corresponding to rounds  $\{0, 1, \dots, r, \dots\}$ .

We will further define  $\Delta \mathbf{w}^o(t+1) = (\Delta w_u(t+1) : u \in \mathcal{U}^o)$ , where  $\Delta \mathbf{w}^o(t+1) = (\Delta w_u(t+1) : u \in \mathcal{U}^o)$  such that,

$$w_u(t+1) = w_u(t)(1 + \Delta w_u(t+1)), \quad \forall o \in \mathcal{O}, u \in \mathcal{U}^o$$

where  $1 + \Delta w_u(t+1)$  captures the relative change in slice  $o$ 's weight update at time slot  $t+1$ . Furthermore, to capture the overall changes in slices weights at the end of each round, we shall define  $\boldsymbol{\omega}(0) = \mathbf{w}(0)$ ,  $\boldsymbol{\omega}(r) = (\boldsymbol{\omega}^o(r) : o \in \mathcal{O})$  where  $\boldsymbol{\omega}^o(r) = \mathbf{w}^o(r \cdot |\mathcal{O}| + 1)$  and  $\Delta \boldsymbol{\omega}^o(r)$  such that  $\Delta \boldsymbol{\omega}(r) = (\Delta \boldsymbol{\omega}_u^o(r) : u \in \mathcal{U}^o)$ . For all  $o \in \mathcal{O}$ , we define

$$\bar{\Delta} \boldsymbol{\omega}^o(r) := \max_{u \in \mathcal{U}^o} \Delta \boldsymbol{\omega}_u^o(r), \quad \underline{\Delta} \boldsymbol{\omega}^o(r) := \min_{u \in \mathcal{U}^o} \Delta \boldsymbol{\omega}_u^o(r).$$

<sup>4</sup>Indeed, as mentioned in the introduction, there are very few results in the literature on the convergence of price-anticipating best response dynamics.

The key step in our convergence proof is the following lemma – see the appendix for a sketch and [9] for the detailed proof.

**Lemma 3.** *If the game has not converged to a Nash equilibrium, i.e.  $\Delta \boldsymbol{\omega}(r) \neq \mathbf{0}$  for  $r > 1$ , then:*

$$\max_{o \in \mathcal{O}} \left( 1 + \bar{\Delta} \boldsymbol{\omega}^o(r+1), \frac{1}{1 + \underline{\Delta} \boldsymbol{\omega}^o(r+1)} \right) < \max_{o \in \mathcal{O}} \left( 1 + \bar{\Delta} \boldsymbol{\omega}^o(r), \frac{1}{1 + \underline{\Delta} \boldsymbol{\omega}^o(r)} \right).$$

The above lemma suggests that when slices have not reached an equilibrium, then in the next round

$$\max_{o \in \mathcal{O}} \left( 1 + \bar{\Delta} \boldsymbol{\omega}^o(r+1), \frac{1}{1 + \underline{\Delta} \boldsymbol{\omega}^o(r+1)} \right)$$

will decrease. This in turn suggests that maximum and minimum components of the vector of relative changes,  $1 + \Delta \boldsymbol{\omega}(r)$ , are getting closer to 1.

With this result in hand, one can show the existence of a Lyapunov function guaranteeing convergence of the best response game, thus completing the proof of Theorem 3 – see [9] for the detailed proof.

#### IV. PERFORMANCE BOUNDS ANALYSIS

In this section we analyze the performance of the Nash equilibrium in terms of two standard metrics for efficiency and fairness: (i) the *price of anarchy*, which gives the loss in overall utility resulting from slices' strategic behavior, and (ii) *envy-freeness*, which captures the degree to which a slice would prefer another slice's allocations across the network resources. We will focus on the case where slice utilities are 1-fair homogeneous i.e.,  $U^o(\mathbf{w}) = \sum_{u \in \mathcal{U}^o} \phi_u \log(r_u(\mathbf{w})) \forall o \in \mathcal{O}$  – a widely accepted case leading to the well-known proportionally fair allocations.

##### A. Efficiency: Price of Anarchy

According to [22], for slices with 1-fair homogenous utilities, the socially optimal allocation of resources  $\mathbf{w}^*$  is such that  $w_u^* = \phi_u s_o$ ,  $\forall u \in \mathcal{U}^o$  and  $\forall o \in \mathcal{O}$ . The following theorem bounds the difference between the *overall network utility* resulting from such socially optimal allocation,  $U(\mathbf{w}^*)$ , and that obtained at a Nash equilibrium of the SCPA resource allocation mechanism,  $U(\mathbf{w})$  – a sketch of the proof is provided in the Appendix.

**Theorem 4.** *If all slices have 1-fair homogenous utilities, then the Price of Anarchy (PoA) associated with a Nash equilibrium  $\mathbf{w}$  satisfies*

$$\text{PoA} := U(\mathbf{w}^*) - U(\mathbf{w}) \leq \log(e).$$

Furthermore, there exists a game instance for which this bound is tight.

Note that, with 1-fair utilities, if we increase the capacity of all resources by a factor  $\Delta c$ , we have a utility increase of  $\log(\Delta c)$ . Thus, the performance improvement achieved by

the socially optimal allocation over SCPA is (in the upper bound) equivalent to having a capacity  $e$  times larger, i.e., almost the triple capacity. While there are some (pathological) cases in which such a bound can be achieved, our simulation results show that for practical scenarios the actual performance difference between the two allocations is much smaller, confirming that (for  $\alpha_o = 1$ ) the flexibility gained with the SCPA mechanism comes at a very small price in performance.

### B. Fairness: Envy-freeness

Next we consider a Nash equilibrium  $\mathbf{w}$  and analyze whether a slice, say  $o$ , with utility  $U^o(\mathbf{w})$ , might have a better utility if it were to exchange its resources with those of another slice, say  $o'$ . To that end, we denote by  $\tilde{\mathbf{w}}$  the resulting weight allocation when the users of slices  $o$  and  $o'$  exchange their allocated resources. It is easy to see that  $\tilde{\mathbf{w}}^o$  is such that

$$\tilde{w}_u^o = \frac{\phi_u}{\sum_{v \in \mathcal{U}_b^o} \phi_v} d_b^{o'}(\mathbf{w}) \text{ for all } b \in \mathcal{B} \text{ and all } u \in \mathcal{U}_b^o, \quad (3)$$

i.e., slice  $o$  takes the aggregate weight of  $o'$  at base station  $b$  under the Nash equilibrium,  $d_b^{o'}(\mathbf{w})$ , and allocates it proportionally to its user priorities. Clearly,  $\tilde{\mathbf{w}}^{o'}$  is defined similarly and the remaining slices weights remain unchanged under  $\tilde{\mathbf{w}}$ .

We define the envy of slice  $o$  for  $o'$ 's resources under the Nash equilibrium  $\mathbf{w}$  by

$$E^{o,o'} := U^o(\tilde{\mathbf{w}}) - U^o(\mathbf{w}).$$

Note that envy is a “directed” concept, i.e., it is defined from slice  $o$ 's point of view. When  $E^{o,o'} \leq 0$ , we say slice  $o$  is not envious. The following theorem provides a bound on  $E^{o,o'}$  – see the Appendix for a sketch of the proof.

**Theorem 5.** *Consider a slice  $o$  with 1-fair homogeneous utilities and the remaining slices  $\mathcal{O} \setminus \{o\}$  with arbitrary slice utilities. Consider a slice  $o'$  such that  $s_o = s_{o'}$ . Let  $\mathbf{w}$  denote a Nash equilibrium and  $\tilde{\mathbf{w}}$  denote the resulting weights when  $o$  and  $o'$  exchange their resources. Then, the envy of slice  $o$  for  $o'$  satisfies*

$$E^{o,o'} = U^o(\tilde{\mathbf{w}}) - U^o(\mathbf{w}) \leq 0.060.$$

Furthermore, there is a game instance where  $0.041 \leq E_{o,o'}$ .

Given that, if one increases the rates of all users by a factor  $\Delta r$  this yields a utility increase of  $\log(\Delta r)$ , one can interpret this result as saying that, by exchanging resources with  $o'$ , slice  $o$  may see a gain equivalent to increasing the rate of all its users by a factor between 4.1% and 6.1% (given by the lower and upper bounds of the above theorem). This is quite low and, moreover, simulation results show that in practical settings there is actually (almost) never any envy, confirming that our system is (practically) envy-free.

## V. PERFORMANCE EVALUATION

Next, we evaluate the performance of the SCPA resource allocation mechanism via simulation. The mobile network scenario considered is based on the IMT-A evaluation guidelines for dense ‘small cell’ deployments [23], which consider

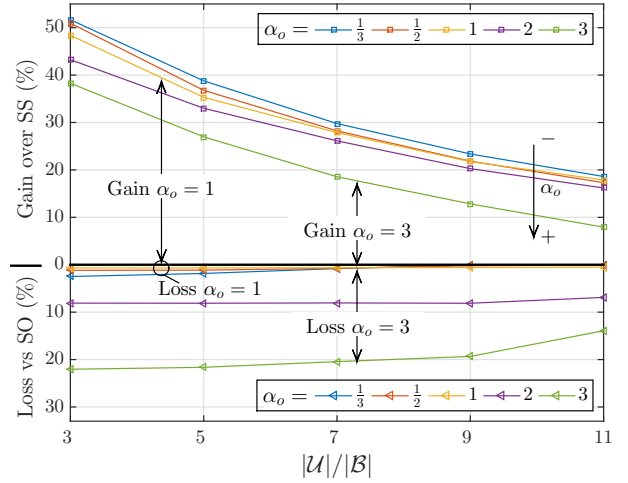


Fig. 1: Average Gain over Static slicing and Loss against Social optimum for different scenarios.

base stations with an intersite distance of 200 meters in a hexagonal cell layout with 3 sector antennas.<sup>5</sup> The network size  $|\mathcal{B}|$  is 57 sectors and users move according to the Random Waypoint Model (RWP).<sup>6</sup> Users’ Signal Interference to Noise Ratio ( $\text{SINR}_u$ ) is computed based on physical layer network model specified in [23] (which includes path loss, shadowing, fast fading and antenna gain) and user association follows the strongest signal policy. The achievable rate for users,  $c_u$ , are determined based on the thresholds reported in [24]. For all our simulation results, we obtained 95% confidence intervals with relative errors below 1% (not shown in the figures).

### A. Overall performance

Throughout the paper we have used *static slicing* and the *socially optimal* resource allocations as our baselines. In order to confirm our analytical results and gain additional insights, we have evaluated the performance of the SCPA mechanism versus these two baselines via simulation. As an intuitive metric for comparison, we have used the extra capacity required by these baseline schemes to deliver the same performance as SCPA:

- (i) *Gain over SS*: additional resources required by *static slicing* to provide the same utility as SCPA (in %).
- (ii) *Loss versus SO*: additional resources required by SCPA to provide the same utility as the *socially optimal* allocation (in %); note that this metric is closely related to the *Price of Anarchy* analyzed in Section IV-A.

The results shown in Figure 1 are for different user densities ( $|\mathcal{U}|/|\mathcal{B}|$ ) and different slice utilities ( $\alpha_o$  parameter). As expected, the SCPA mechanism always has a gain over static slicing and a loss over the social optimal. However, for  $\alpha_o = 1$  the loss is well below the bound given in Section IV-A. We

<sup>5</sup>Note that, in this setting, users associate with sectors rather than the base stations we used in the mechanism description and analysis.

<sup>6</sup>In the extended version, additional simulation results are given for different mobility models [9].

further observe that performance is particularly good as long as  $\alpha_o$  does not exceed 1 (Gain over SS up to 50% and Loss over SO below 5%), and it degrades mildly as  $\alpha_o$  increases.

### B. Fairness

In addition to overall performance, it is of interest to evaluate the fairness of the resulting allocations. While in Section IV-B we derived analytically a bound on the envy, we have further explored this via simulation by evaluating up to  $10^7$  randomly generated scenarios, with parameters drawn uniformly in the ranges:  $|\mathcal{O}| \in [2, 12]$ ,  $|\mathcal{B}| \in [10, 90]$ ,  $|\mathcal{U}|/|\mathcal{B}| \in [3, 15]$ ,  $\alpha_o \in [0.01, 30]$  and  $\phi$  vectors in the simplex. Our results show that  $E^{o,o'} < 0$  holds for *all* the cases explored, confirming that in practice the system is envy-free.

### C. Protection against other slices

One of the main objectives of our proposed framework is to enable slices to customize their resource allocations. This can be done by adjusting (i) the user priorities  $\phi_u$ , and (ii) the parameter  $\alpha_o$ , which regulates the desired level of fairness among the slice's users. In order to evaluate the impact that these settings have amongst slices, we simulated a scenario with three slices: Slice 1 has  $\alpha_1 = 1$ , Slice 2 has  $\alpha_2 = 4$ , and Slice 3 has  $\alpha_3$  with varying values. For simplicity, we set the priorities  $\phi_u$  equal for all users.

Figure 2 shows the rate distributions of the 3 slices. We observe that the choice of  $\alpha_3$  is effective in adjusting the level of user fairness for Slice 3; indeed, as  $\alpha_3$  grows, the rate distribution becomes more homogeneous. Such customization at Slice 3 has a higher impact on Slice 1 than on Slice 2. This is the case because, as  $\alpha_2$  is quite large, the distribution of Slice 2's rates remains homogeneous, making the slice fairly insensitive to the choices of the other slices. As can be seen in the subplots, the utilities of Slices 1 and 2 are not only larger than the utility of static slicing, but remain fairly insensitive to  $\alpha_3$ , showing that in both cases we have a good level of protection between slices.

## VI. CONCLUSIONS

In this paper we have analyzed a 'share-constrained proportional allocation' framework for network slicing. The framework allows slices to customize the resource allocation to their users, leading to a *network slicing game* in which each slice reacts to the settings of the others. Our main conclusion is that the framework provides an *effective and implementable* scheme for dynamically sharing resources across slices. Indeed, this scheme involves simple operations at base stations and incurs a limited signaling between the slices and the infrastructure. Our results confirm system stability (*best response dynamics* converge), substantial gains over *static slicing*, and fairness of the allocations (*envy-freeness*). Moreover, as long as the majority of the slices do not choose  $\alpha_o$  values larger than 1 (i.e., they do not all demand very homogeneous rate distributions), the overall performance is close to optimal (*price of anarchy* is very small). Thus, in this case the flexibility provided by this framework comes at no

cost. If a substantial number of slices choose higher  $\alpha_o$ 's, then we pay a (small) price for enabling slice customization.

## ACKNOWLEDGEMENTS

The work of P. Caballero and G. de Veciana has been supported by NSF Award CNS1232283, and the work of A. Banchs and X. Costa-Pérez by the H2020-ICT-2014-2 project 5G NORMA under Grant Agreement No. 671584.

## REFERENCES

- [1] NGMN Alliance, "Description of Network Slicing Concept," NGMN 5G P1, Jan. 2016.
- [2] 3GPP, "Study on Architecture for Next Generation System," TR 23.799, v0.5.0, May 2016.
- [3] —, "Study on Radio Access Network (RAN) sharing enhancements," TR 22.852, v13.1.0, Sep. 2014.
- [4] L. Zhang, "Proportional response dynamics in the Fisher market," *Theoretical Computer Science*, vol. 412, no. 24, pp. 2691–2698, May 2011.
- [5] M. Feldman, K. Lai, and L. Zhang, "The Proportional-Share Allocation Market for Computational Resources," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 8, pp. 1075–1088, Aug. 2009.
- [6] F. P. Kelly, A. K. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness, and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, Mar. 1998.
- [7] R. Johari and J. N. Tsitsiklis, "Efficiency Loss in a Network Resource Allocation Game," *Mathematics of Operations Research*, vol. 29, no. 3, pp. 407–435, Aug. 2004.
- [8] S. Yang and B. Hajek, "VCG-Kelly Mechanisms for Allocation of Divisible Goods: Adapting VCG Mechanisms to One-Dimensional Signals," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 6, pp. 1237–1243, Aug. 2007.
- [9] Extended version of the paper "Network Slicing Games: Enabling Customization in Multi-Tenant Mobile Networks", available at <https://arxiv.org/abs/1612.08446>.
- [10] P. D. Francesco *et al.*, "A Sharing- and Competition-Aware Framework for Cellular Network Evolution Planning," *IEEE Transactions on Cognitive Communications and Networking*, vol. 1, no. 2, Jun. 2015.
- [11] B. Leng, P. Mansourifard, and B. Krishnamachari, "Microeconomic Analysis of Base-station Sharing in Green Cellular Networks," in *Proc. of IEEE INFOCOM*, Apr. 2014.
- [12] A. Gudipati, L. E. Li, and S. Katti, "RadioVisor: A Slicing Plane for Radio Access Networks," in *Proc. of HotSDN*, Aug. 2014.
- [13] I. Malanchini, S. Valentin, and O. Aydin, "Generalized resource sharing for multiple operators in cellular wireless networks," in *Proc. of IWCMC*, Aug. 2014.
- [14] R. Mahindra *et al.*, "Radio Access Network sharing in cellular networks," in *Proc. of IEEE ICNP*, Oct. 2013.
- [15] S. Rathinakumar and M. K. Marina, "GAVEL: strategy-proof ascending bid auction for dynamic licensed shared access," in *Proc. of ACM MobiHoc*, Jul. 2016.
- [16] NGMN Alliance, "5G White Paper," Feb. 2015.
- [17] X. Zhou, R. Li, T. Chen, and H. Zhang, "Network slicing as a service: enabling enterprises' own software-defined cellular networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 146–153, Jul. 2016.
- [18] K. Samdanis, X. Costa-Perez, and V. Sciancalepore, "From Network Sharing to Multi-tenancy: The 5G Network Slice Broker," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 32–39, Jul. 2016.
- [19] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on Networking (ToN)*, vol. 8, no. 5, pp. 556–567, Oct. 2000.
- [20] X. Costa-Perez, J. Swetina, T. Guo, R. Mahindra, and S. Rangarajan, "Radio access network virtualization for future mobile carrier networks," *IEEE Communications Magazine*, vol. 51, no. 7, Jul. 2013.
- [21] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, Jul. 1965.
- [22] L. Li, M. Pal, and Y. R. Yang, "Proportional fairness in multi-rate wireless LANs," in *Proc. of IEEE INFOCOM*, Apr. 2008.
- [23] ITU-R, "Report ITU-R M.2135-1, Guidelines for evaluation of radio interface technologies for IMT-Advanced," Technical Report, Dec. 2009.
- [24] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," TS 36.213, v12.5.0, Rel. 12, Mar. 2015.



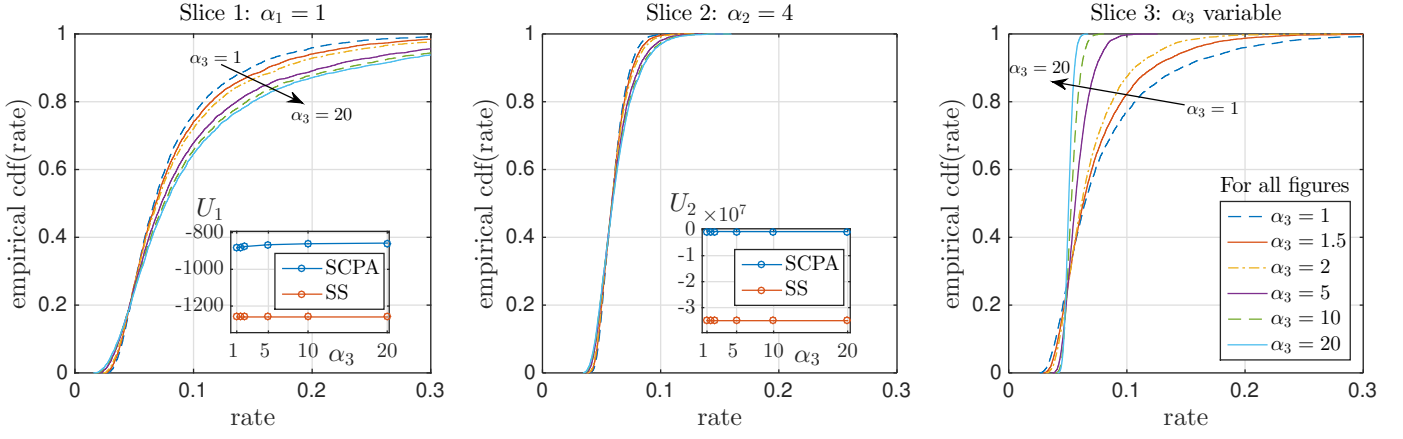


Fig. 2: Impact of  $\alpha_3$  decision on the slice rate distributions.

#### APPENDIX

In the following we provide sketches of the proofs of some of the most relevant theoretical results of this paper. The complete proofs of all the results can be found in the extended version of the paper [9].

##### Sketch of the proof of Lemma 3

Let  $\Delta a_b^o(t)$  be the relative change of  $a_b^o(t)$  since the previous round, i.e.,  $a_b^o(t) = a_b^o(t-1)(1 + \Delta a_b^o(t))$ , and let  $\underline{\Delta a}^o(t) = \min_b a_b^o(t)$  and  $\overline{\Delta a}^o(t) = \max_b a_b^o(t)$ . To prove the lemma we will start by showing the following intermediate result:

$$\max \left( 1 + \overline{\Delta w}^o(t+1), \frac{1}{1 + \underline{\Delta w}^o(t+1)} \right) < \max \left( 1 + \overline{\Delta a}^o(t), \frac{1}{1 + \underline{\Delta a}^o(t)} \right), \quad (4)$$

We will first show that

$$1 + \overline{\Delta w}^o(t+1) < \max \left( 1 + \overline{\Delta a}^o(t), \frac{1}{1 + \underline{\Delta a}^o(t)} \right). \quad (5)$$

Let  $u, u' \in \mathcal{U}^o$  be, respectively, the user for which  $\Delta w_u(t+1)$  takes the largest value and the one for which it takes the smallest value. Then,

$$\frac{w_u(t+1)}{w_{u'}(t+1)} = \frac{w_u(t)(1 + \overline{\Delta w}^o(t+1))}{w_{u'}(t)(1 + \underline{\Delta w}^o(t+1))} \quad (6)$$

From Lemma 2 we have

$$\frac{w_u(t+1)}{w_{u'}(t+1)} = \frac{\frac{\beta_u(a_b^o(t-1)|\mathcal{O}|+1)(1 + \Delta a_b^o(t))}{(a_b^o(t-1)|\mathcal{O}|+1)(1 + \Delta a_b^o(t)) + d_u w_u(t)(1 + \overline{\Delta w}^o(t+1))}^{\frac{1}{\alpha_o}}}{\frac{\beta_{u'}(a_{b'}^o(t-1)|\mathcal{O}|+1)(1 + \Delta a_{b'}^o(t))}{(a_{b'}^o(t-1)|\mathcal{O}|+1)(1 + \Delta a_{b'}^o(t)) + d_{u'} w_{u'}(t)(1 + \underline{\Delta w}^o(t+1))}^{\frac{1}{\alpha_o}}}$$

where  $b$  and  $b'$  are the stations of user  $u$  and  $u'$  respectively.

By operating on the above equation, we can obtain the following inequality for  $1 \leq \alpha_o \leq 2$ :

$$\frac{w_u(t+1)}{w_{u'}(t+1)} \leq \frac{\frac{\beta_u(a_b^o(t-1)|\mathcal{O}|+1)^{\frac{1}{\alpha_o}}}{(a_b^o(t-1)|\mathcal{O}|+1 + d_u w_u(t))^{\frac{2}{\alpha_o}-1}}}{\frac{\beta_{u'}(a_{b'}^o(t-1)|\mathcal{O}|+1)^{\frac{1}{\alpha_o}}}{(a_{b'}^o(t-1)|\mathcal{O}|+1 + d_{u'} w_{u'}(t))^{\frac{2}{\alpha_o}-1}}} \frac{(1 + \overline{\Delta a}^o(t))^{\frac{1}{\alpha_o}}}{(1 + \underline{\Delta a}^o(t))^{\frac{1}{\alpha_o}}}$$

$$= \frac{w_u(t)}{w_{u'}(t)} \frac{(1 + \overline{\Delta a}^o(t))^{\frac{1}{\alpha_o}}}{(1 + \underline{\Delta a}^o(t))^{\frac{1}{\alpha_o}}}. \quad (7)$$

Combining (20) with the above yields

$$\frac{w_u(t)(1 + \overline{\Delta w}^o(t+1))}{w_{u'}(t)(1 + \underline{\Delta w}^o(t+1))} < \frac{w_u(t)}{w_{u'}(t)} \frac{(1 + \overline{\Delta a}^o(t))^{\frac{1}{\alpha_o}}}{(1 + \underline{\Delta a}^o(t))^{\frac{1}{\alpha_o}}}.$$

From  $x^a y^b \leq \max(x, y)$  for  $x, y \geq 1$  and  $a + b = 1$

$$\frac{(1 + \overline{\Delta a}^o(t+1))^{\frac{1}{\alpha_o}}}{(1 + \underline{\Delta a}^o(t))^{\frac{1}{\alpha_o}}} \leq \max \left( 1 + \overline{\Delta a}^o(t), \frac{1}{1 + \underline{\Delta a}^o(t)} \right).$$

From the above two equations we have

$$\frac{1 + \overline{\Delta w}^o(t+1)}{1 + \underline{\Delta w}^o(t+1)} < \max \left( 1 + \overline{\Delta a}^o(t), \frac{1}{1 + \underline{\Delta a}^o(t)} \right). \quad (8)$$

We can now prove (18) by contradiction. Suppose that

$$1 + \overline{\Delta w}^o(t+1) \geq \max \left( 1 + \overline{\Delta a}^o(t), \frac{1}{1 + \underline{\Delta a}^o(t)} \right).$$

and combining this with (??) yields  $\frac{1}{1 + \underline{\Delta w}^o(t+1)} < 1$ , which contradicts the fact that  $\underline{\Delta w}^o(t+1)$  needs to be necessarily smaller than 0 unless we have already converged.<sup>7</sup> Thus, this proves (18).

To prove (17), we also need to show the equation below. This can be proven based on a similar argument to the one we have used for (18) (see [9] for the details).

$$\frac{1}{1 + \underline{\Delta w}^o(t+1)} < \max \left( 1 + \overline{\Delta a}^o(t), \frac{1}{1 + \underline{\Delta a}^o(t)} \right).$$

Once we have shown (17), we proceed as follows. From the fact that  $a_b^o(t) = \sum_{o' \in \mathcal{O} \setminus \{o\}} \sum_{u \in \mathcal{U}_{b'}^o} w_u(t)$ , it can be seen that

$$(1 + \overline{\Delta a}^o(t)) \leq \max_{t' \in \{t, |\mathcal{O}|+2, \dots, t\}} (1 + \overline{\Delta w}^o(t')(t')) \text{ and } (1 + \underline{\Delta a}^o(t)) \geq \min_{t' \in \{t, |\mathcal{O}|+2, \dots, t\}} (1 + \underline{\Delta w}^o(t')(t')).$$

<sup>7</sup>The case in which we have converged,  $\underline{\Delta w}^o(t+1) = 0$ , is detailed in the extended version of the proof [9].

Combining the above with (17) yields

$$\max \left( 1 + \bar{\Delta} w^o(t+1), \frac{1}{1 + \underline{\Delta} w^o(t+1)} \right) < \max_{t' \in \{t, |O|+2, \dots, t\}} \left( 1 + \bar{\Delta} w^{o(t')}(t'), \frac{1}{1 + \underline{\Delta} w^{o(t')}(t')} \right)$$

Finally, the lemma is proven by applying the above expression recursively.  $\square$

#### Sketch of the proof of Theorem 4

Since in the NE each slice maximizes its utility, we have

$$\sum_{u \in \mathcal{U}^o} \phi_u \log \left( \frac{w_u}{l_{b(u)}(\mathbf{w})} \right) \geq \sum_{u \in \mathcal{U}^o} \phi_u \log \left( \frac{w_u^*}{d_{b(u)}^o(\mathbf{w}^*) + a_{b(u)}^o(\mathbf{w})} \right)$$

Given that  $d_{b(u)}^o(\mathbf{w}^*) + a_{b(u)}^o(\mathbf{w}) \leq l_{b(u)}(\mathbf{w}) + l_{b(u)}(\mathbf{w}^*)$ ,

$$\sum_{u \in \mathcal{U}^o} \phi_u \log \left( \frac{w_u}{l_{b(u)}(\mathbf{w})} \right) \geq \sum_{u \in \mathcal{U}^o} \phi_u \log \left( \frac{w_u^*}{l_{b(u)}(\mathbf{w}) + l_{b(u)}(\mathbf{w}^*)} \right)$$

From the above it follows that

$$\begin{aligned} & \sum_{u \in \mathcal{U}^o} \phi_u \log(r_u(\mathbf{w}^*)) - \sum_{u \in \mathcal{U}^o} \phi_u \log(r_u(\mathbf{w})) \\ & \leq \sum_{u \in \mathcal{U}^o} \phi_u \log \left( \frac{w_u^* c_u}{l_{b(u)}(\mathbf{w}^*)} \right) - \sum_{u \in \mathcal{U}^o} \phi_u \log \left( \frac{w_u^* c_u}{l_{b(u)}(\mathbf{w}) + l_{b(u)}(\mathbf{w}^*)} \right) \\ & = \sum_{u \in \mathcal{U}^o} \phi_u \log \left( \frac{l_{b(u)}(\mathbf{w}^*)}{l_{b(u)}(\mathbf{w}) + l_{b(u)}(\mathbf{w}^*)} \right) \end{aligned}$$

Summing the above over all slices weighted by the corresponding shares yields

$$U(\mathbf{w}^*) - U(\mathbf{w}) \leq \sum_{u \in \mathcal{U}} \phi_u s_o \log \left( \frac{l_{b(u)}(\mathbf{w}^*)}{l_{b(u)}(\mathbf{w}) + l_{b(u)}(\mathbf{w}^*)} \right)$$

Given  $w_u^* = \phi_u s_o$ , we have

$$\begin{aligned} U(\mathbf{w}^*) - U(\mathbf{w}) & \leq \sum_{b \in \mathcal{B}} \log \left( \frac{l_b(\mathbf{w}^*)}{l_b(\mathbf{w}) + l_b(\mathbf{w}^*)} \right)^{\sum_{u \in \mathcal{U}_b} w_u^*} \\ & = \sum_{b \in \mathcal{B}} \sum_{u \in \mathcal{U}_b} w_u \log \left( \frac{l_b(\mathbf{w}^*)/l_b(\mathbf{w})}{1 + l_b(\mathbf{w}^*)/l_b(\mathbf{w})} \right)^{\frac{l_b(\mathbf{w}^*)}{l_b(\mathbf{w})}} \end{aligned}$$

and, given that  $(x/(1+x))^x > 1/e$  for  $x \geq 0$ , this yields

$$U(\mathbf{w}^*) - U(\mathbf{w}) \leq \sum_{b \in \mathcal{B}} \sum_{u \in \mathcal{U}_b} w_u \log(e) = \log(e)$$

To show that the bound is tight, see the game instance given in [9] for which it holds  $U(\mathbf{w}^*) - U(\mathbf{w}) = \log(e)$ .  $\square$

#### Sketch of the proof of Theorem 5

In order to bound the envy  $U^o(\tilde{\mathbf{w}}) - U^o(\mathbf{w})$  at the NE, we will construct a weight allocation  $\mathbf{m}$  that satisfies  $U^o(\mathbf{m}) \leq U^o(\mathbf{w})$  and  $U^o(\tilde{\mathbf{m}}) \geq U^o(\tilde{\mathbf{w}})$  – where  $\tilde{\mathbf{w}}$  and  $\tilde{\mathbf{m}}$  are the allocations resulting from exchanging the resources of slices  $o$  and  $o'$  in  $\mathbf{w}$  and  $\mathbf{m}$ , respectively. It then follows that  $U^o(\tilde{\mathbf{m}}) - U^o(\mathbf{m})$  is an upper bound on the envy.

Specifically, the weight allocation  $\mathbf{m}$  will be chosen such that: (i) for all slices different from  $o$ , the weights remain the same as in the NE, i.e.,  $\mathbf{m}^o = \mathbf{w}^o$ ; and (ii) the weights of slice  $o$  are chosen so as to maximize  $U^o(\mathbf{m})$  subject

to  $d_b^o(\mathbf{m}^o) = \sum_{u \in \mathcal{U}_b^o} m_u \leq a_b^o(\mathbf{m}^o) \forall b \in \mathcal{B}$  and slice  $o$ 's share constraint. Note that with this weight allocation we have  $a_{b(u)}^o(\mathbf{m}^o) = a_{b(u)}^o(\mathbf{w}^o)$  – for readability purposes, we will use just  $a_{b(u)}^o$ . Note also that the weights that slice  $o$  would have with the resources of  $o'$  remain the same, i.e.,  $\tilde{\mathbf{m}}^o = \tilde{\mathbf{w}}^o$ .

By following a similar argument to that of Lemma 2, it can be seen that the above leads to the weights  $m_u$  for  $u \in \mathcal{U}^o$  solving the set of equations below, which have a feasible solution as long as  $s_o < \sum_{u \in \mathcal{U}^o} a_{b(u)}^o(\mathbf{m}^o)$  (see [9] for the case when this does not hold).

$$m_u = \begin{cases} a_{b(u)}^o \frac{\phi_u}{\sum_{v \in \mathcal{U}_{b(u)}^o} \phi_v}, & a_{b(u)}^o = d_{b(u)}^o(\mathbf{m}^o) \\ \frac{\phi_u \frac{a_{b(u)}^o}{a_{b(u)}^o + d_{b(u)}^o(\mathbf{m}^o)}}{\sum_{v \in \mathcal{U}^o} \phi_v \frac{a_{b(v)}^o}{a_{b(v)}^o + d_{b(v)}^o(\mathbf{m}^o)}} s'_o, & a_{b(u)}^o > d_{b(u)}^o(\mathbf{m}^o) \end{cases}$$

where  $\hat{\mathcal{U}}^o$  is the set of users of slice  $o$  for which  $a_{b(u)}^o > d_{b(u)}^o(\mathbf{m}^o)$  and  $s'_o = s_o - \sum_{u \in \mathcal{U}^o \setminus \hat{\mathcal{U}}^o} m_u$ .

Note that  $U^o(\mathbf{m}) \leq U^o(\mathbf{w})$  is a direct consequence of the fact that  $\mathbf{w}$  is a NE and  $\mathbf{m}^o = \mathbf{w}^o$ . The inequality  $U^o(\tilde{\mathbf{m}}) \geq U^o(\tilde{\mathbf{w}})$  is proven in the extended version [9] by showing that we can move from  $\tilde{\mathbf{w}}$  to  $\tilde{\mathbf{m}}$  through intermediate steps such that in each of them  $U^o(\tilde{\mathbf{w}})$  increases.

To find an upper bound on  $U^o(\tilde{\mathbf{m}}) - U^o(\mathbf{m})$ , recall that  $U^o(\tilde{\mathbf{m}}) = \sum_{u \in \mathcal{U}^o} \phi_u \log(\frac{\tilde{m}_u c_u}{l_b(\tilde{\mathbf{m}})})$  and  $U^o(\mathbf{m}) = \sum_{u \in \mathcal{U}^o} \phi_u \log(\frac{m_u c_u}{l_b(\mathbf{m})})$ . Given that  $l_b(\tilde{\mathbf{m}}) = l_b(\mathbf{m})$  and  $\tilde{m}_u = m_u$  for  $u \notin \hat{\mathcal{U}}^o$ , this yields

$$U^o(\tilde{\mathbf{m}}) - U^o(\mathbf{m}) = \sum_{u \in \hat{\mathcal{U}}^o} \phi_u \log(\tilde{m}_u) - \sum_{u \in \hat{\mathcal{U}}^o} \phi_u \log(m_u).$$

Since  $\sum_{u \in \hat{\mathcal{U}}^o} \log(\tilde{m}_u)$  subject to  $\sum_{u \in \hat{\mathcal{U}}^o} \tilde{m}_u = s'_o$  takes a maximum at  $\tilde{m}_u = \hat{\phi}_u s'_o$  (where  $\hat{\phi}_u = \phi_u / \sum_{v \in \hat{\mathcal{U}}^o} \phi_v$ ),

$$\begin{aligned} U^o(\tilde{\mathbf{m}}) - U^o(\mathbf{m}) & \leq \sum_{u \in \hat{\mathcal{U}}^o} \phi_u \log(\hat{\phi}_u s'_o) - \sum_{u \in \hat{\mathcal{U}}^o} \phi_u \log(m_u) \\ & \leq \sum_{u \in \hat{\mathcal{U}}^o} \hat{\phi}_u \log(\hat{\phi}_u s'_o) - \sum_{u \in \hat{\mathcal{U}}^o} \hat{\phi}_u \log(m_u) \quad (9) \end{aligned}$$

In order to bound the term  $\sum_{u \in \hat{\mathcal{U}}^o} \hat{\phi}_u \log(m_u)$  above, we look for a bound on  $\frac{m_u}{m_v}$ . Given that  $a_b^o \geq d_b^o(\mathbf{m}^o)$  holds for all  $b$ , we have for  $u, v \in \hat{\mathcal{U}}^o$ :

$$\frac{m_u}{m_v} = \frac{\phi_u \frac{a_{b(u)}^o}{a_{b(u)}^o + d_{b(u)}^o(\mathbf{m}^o)}}{\phi_v \frac{a_{b(v)}^o}{a_{b(v)}^o + d_{b(v)}^o(\mathbf{m}^o)}} > \frac{\phi_u \frac{a_{b(u)}^o}{a_{b(u)}^o + a_{b(u)}^o}}{\phi_v \frac{a_{b(v)}^o}{a_{b(v)}^o}} = \frac{1}{2} \frac{\phi_u}{\phi_v}$$

It can be seen that  $\sum_{u \in \hat{\mathcal{U}}^o} \hat{\phi}_u \log(m_u)$  subject to  $\frac{m_u}{m_v} \geq \frac{1}{2} \frac{\phi_u}{\phi_v}$  is maximized when the  $\frac{m_u}{\phi_u}$  of all users but one is equal to the lower bound given by the constraint, which yields  $\frac{m_u}{\phi_u} = \frac{1}{2} \frac{m_v}{\phi_v}, \forall u \neq v$ . Substituting these  $\mathbf{m}$  values in (9) and simplifying the resulting expression, we obtain  $U^o(\tilde{\mathbf{m}}) - U^o(\mathbf{m}) \leq \log(1 + \hat{\phi}) - \hat{\phi} \log(2)$ . The  $\hat{\phi}$  value that maximizes this expression is  $\hat{\phi} = \frac{1}{\log 2} - 1$ . Substituting this value in the expression gives  $U^o(\tilde{\mathbf{m}}) - U^o(\mathbf{m}) \leq 0.060$ . To show that the bound is tight, see [9] for a game instance that satisfies  $U^o(\tilde{\mathbf{m}}) - U^o(\mathbf{m}) = 0.041$ .  $\square$