

# Arbitrators in Overlapping Coalition Formation Games

Yair Zick

Division of Mathematical Sciences  
School of Physical and Mathematical Sciences  
Nanyang Technological University, Singapore  
yair0001@ntu.edu.sg

Edith Elkind

Division of Mathematical Sciences  
School of Physical and Mathematical Sciences  
Nanyang Technological University, Singapore  
eelkind@ntu.edu.sg

## ABSTRACT

Overlapping Coalition Formation (OCF) games [3, 4] are cooperative games where the players can simultaneously participate in several coalitions. Capturing the notion of stability in OCF games is a difficult task: a player may deviate by abandoning some, but not all of the coalitions he is involved in, and the crucial question is whether he then gets to keep his payoff from the unaffected coalitions. In [4] the authors introduce three stability concepts for OCF games—the *conservative*, *refined*, and *optimistic* core—that are based on different answers to this question. In this paper, we propose a unified framework for the study of stability in the OCF setting, which encompasses the concepts considered in [4] as well as a wide variety of alternative stability concepts. Our approach is based on the notion of an *arbitrator*, which can be thought of as an external party that determines payoff to deviators. We give a complete characterization of outcomes that are stable under arbitration. In particular, our results provide a criterion for the outcome to be in the refined or optimistic core, thus complementing the results in [4] for the conservative core, and answering questions left open in [4]. We also introduce a notion of the nucleolus for arbitrated OCF games, and argue that it is non-empty. Finally, we extend the definition of the Shapley value [12] to the OCF setting, and provide an axiomatic characterization for it.

## Categories and Subject Descriptors

I.2.11 [Distributed Artificial Intelligence]: Multiagent Systems

## General Terms

Theory

## Keywords

Overlapping Coalition Formation, Core, Nucleolus, Shapley Value

## 1. INTRODUCTION

Cooperation among agents plays a crucial role in the functioning of multi-agent systems. Therefore, developing a better understanding of coalition formation processes is an important research agenda in the multiagent community, and a lot of recent research effort has been spent on the design and analysis of cooperation mechanisms

**Cite as:** Arbitrators in Overlapping Coalition Formation Games, Yair Zick and Edith Elkind, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 55-62.  
Copyright © 2011, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

for realistic multi-agent environments [15, 10, 5, 14, 8]. In such environments, agents are often selfish, and therefore need to be given incentives to act together and share the benefits of cooperation in a fair manner. Cooperative game theory [9, 18] provides the theoretical underpinnings for the study of such settings. Traditionally, it models a multiagent system as a (transferable-utility) game. Such a game can be described by its *characteristic function*, which for every set of agents specifies the profit that these agents can attain by working together. The agents are expected to split into teams, i.e., form a *coalition structure*; the profits of each team are then distributed among its members.

Remarkably, the traditional model assumes that each agent participates in exactly one coalition. However, this is often not the case in real-life settings, where agents form multiple coalitions on the fly in order to perform a specific task and only devote part of their attention and resources to each such coalition. Indeed, Shehory and Kraus in their seminal paper [14] already mention that agents can benefit from forming overlapping coalitions, and propose algorithms for iterative formation of an overlapping coalition structure for their setting. This line of work has been continued by Dang et al. [5], where the authors consider overlapping coalition formation in sensor networks. However, these papers assume that agents are fully cooperative, and will always form the socially optimal (overlapping) coalition structure. While this assumption is appropriate for the specific scenarios considered in these papers, in general, agents may want to maximize their own welfare, and a fully expressive model for overlapping coalition formation should take incentive issues into account.

Recently, Chalkiadakis et al. [3, 4] addressed this problem by proposing a game-theoretic model for overlapping coalition formation. In their model, each agent is endowed with a certain amount of resources, which he is free to distribute across multiple coalitions. The value of such (partial) coalition is determined both by the identities of agents that participate in it and the amount of resources that they contribute. Chalkiadakis et al. [3, 4] focus on the study of stability in their model. Compared to the non-overlapping setting, the stability of an overlapping coalition structure is a delicate issue: if an agent is participating in several projects at once and decides to withdraw all or some of her contributions from one of them, can she expect to continue to receive the payoff from the coalitions that were not harmed by the deviation? In [4], the authors propose three different stability concepts—the conservative core, the refined core, and the optimistic core—that correspond to three possible ways of answering this question. Briefly, under the conservative core the deviators do not expect to get any payoffs from their coalitions with non-deviators. In contrast, in the refined core they continue to get payoffs from coalitions not affected by the deviation. Finally, in the optimistic core the deviators may get

some payoffs from an affected coalition, as long as they continue to contribute to it, and the members of that coalition were able to regroup and focus on a different task so that each non-deviator still gets as much profit as before from that coalition.

While the three concepts of the core proposed in [4] all correspond to reasonable reactions to deviation, this list is by no means exhaustive. For instance, a player may want to punish the deviators and refuse to cooperate with them altogether as soon as they lower the value of one of the coalitions he is involved in, even if other coalitions between that player and the deviator remain unaffected. Alternatively, the players may form a social network, and stop collaborating with a deviator if his behavior harmed one of their friends. Yet another possibility is that a central authority imposes a fine on each of the deviators, making the deviation costly.

In this paper, we propose a stability concept that captures all of the scenarios considered above. Our approach is based on the notion of an *arbitrator*: a function that takes the description of a deviation as an argument and returns the payoff that the deviators receive from each coalition. Different arbitrators correspond to different sets of stable outcomes, or *arbitrated cores*. We show that the three core concepts proposed in [4] can be viewed as special cases of our model. Further, paper [4] characterizes the set of outcomes that belong to the conservative core. We extend this characterization to all arbitrated cores. In particular, this allows us to characterize the outcomes in the refined core and the optimistic core, thus answering the open question proposed in [4].

Now, while the core is an attractive stability concept, it is known that some games have an empty core. This is true even in the overlapping model for most realistic arbitrators. Thus, it is desirable to have a solution concept that identifies “the most stable” (overlapping) coalitions, yet is guaranteed to be non-empty. In the non-overlapping setting, this role is fulfilled by the nucleolus [11]. Motivated by this intuition, we introduce a concept of nucleolus for the OCF setting, and demonstrate that it is always non-empty. However, in contrast to the traditional model, we show that in OCF games the nucleolus may contain more than one outcome.

Finally, we extend the definition of the Shapley value [12] to the OCF setting. Just as in the classic case, we present a set of natural axioms, and demonstrate that our variant of the OCF Shapley value is characterized by these axioms.

The rest of this paper is organized as follows. After presenting the necessary background material in Section 2, we introduce the notion of arbitration and arbitrated core in Section 3, and present our characterization of the outcome in the arbitrated core in Section 4. Section 5 focuses on the nucleolus for OCF games, and Section 6 describes our extension of the Shapley value to the OCF setting. Section 7 presents our conclusions and suggests directions for future work.

## 2. PRELIMINARIES

We begin by describing our notation and the formal model of OCF games. Our definitions mostly follow those in [4]. We, however, describe deviation in a manner more conducive to our analysis.

**Notation** Throughout the paper, we write  $N = \{1, \dots, n\}$ . Given a vector  $\mathbf{x} = (x^1, \dots, x^n) \in \mathbb{R}^n$  and a set  $S \subseteq N$ , we write  $\mathbf{x}(S) = \sum_{i \in S} x^i$ ,  $\mathbf{x}|_S$  equals  $\mathbf{x}$  on the  $S$  coordinates and is 0 otherwise, and  $e^S$  is the indicator vector of  $S$ .

**Classic TU Cooperative Games** A *transferable utility (TU) cooperative game* is defined by a set of players  $N$  and a *characteristic function*  $u : 2^N \rightarrow \mathbb{R}$  with  $u(\emptyset) = 0$ . The set of *feasible payoffs* for a game  $G = (N, u)$  is the set of all vectors  $\mathbf{x} \in \mathbb{R}^n$  such that  $\mathbf{x}(N) = u(N)$ . A *solution concept* is a function that assigns

every TU game  $G = (N, u)$  a set of feasible payoff vectors. A solution concept that assigns  $G$  a single point is called a *value*. For a detailed discussion of solution concepts and their axiomatization see [9], Section 2.3, pp. 19–25.

**OCF Games** Let  $N = \{1 \dots n\}$  be a set of agents. A *partial coalition* of players in  $S \subseteq N$  is a vector  $\mathbf{c} \in [0, 1]^n$ , where  $c^i = 0$  for all  $i \notin S$ . That is, each player in  $S$  may contribute a fraction of their resources to  $\mathbf{c}$ . In what follows, we will omit the word “partial”, and refer to vectors in  $[0, 1]^n$  as *coalitions*.

**DEFINITION 2.1.** An OCF game  $G = (N, v)$  is given by a set of players  $N$  and a characteristic function  $v : [0, 1]^n \rightarrow \mathbb{R}$  assigning a real value to each partial coalition; we require  $v((0)^n) = 0$ .

A *coalition structure* over  $N$  is a  $n \times k$  matrix  $CS = (\mathbf{c}_1, \dots, \mathbf{c}_k)$ , where  $k$  is the number of coalitions. We require that for all  $i \in N$  it holds that  $\sum_{j=1}^k c_j^i \leq 1$ . This means that  $CS$  is a valid division of players’ resources. Coalition structures over subsets of  $N$  are defined in a similar manner. Throughout the paper, we will assume that  $v$  is monotone; thus, we can assume that all players would want to invest all their resources, i.e.  $\sum_{j=1}^k c_j^i = 1$ ; such coalition structures are called *efficient*. We denote the set of all possible coalition structures over  $S$  as  $CS_S$ . We also overload notation and define  $v(CS) = \sum_{j=1}^k v(\mathbf{c}_j)$ . Two coalition structures  $CS = (\mathbf{c}_1, \dots, \mathbf{c}_k)$  and  $CS' = (\mathbf{d}_1, \dots, \mathbf{d}_k)$  are *equivalent* if there is some permutation  $\sigma$  such that for all  $1 \leq j \leq k$ ,  $\mathbf{c}_j = \mathbf{d}_{\sigma(j)}$ .

Similarly to [4], we would sometimes like to limit the maximum number of coalitions players can form; indeed, oftentimes an agent who gives less than a certain fraction of her resources to a coalition can no longer contribute to a coalition. If the number of coalitions is limited by  $U \in \mathbb{N}$  we say that the game  $G$  is *U-finite*.

For any  $CS \in CS_S$ , we call the vector  $w(CS) = \sum_{j=1}^k \mathbf{c}_j$  the *weight vector* of  $CS$ . Note that  $w(CS) \in [0, 1]^n$ , and if  $CS$  is efficient then  $w(CS)$  is the indicator vector for the set  $S$ .

For each  $S \subseteq N$  we denote by  $v^*(S)$  the maximum value achievable by  $S$ :  $v^*(S) = \sup\{v(CS) \mid CS \in CS_S\}$ . Note that  $(N, v^*)$  can be viewed as a classic TU cooperative game; we will refer to  $(N, v^*)$  as the *crisp analogue* of  $G$ . We extend the function  $v^*$  to partial coalitions by setting  $v^*(\mathbf{c}) = \sup\{v(CS) \mid w(CS) \leq \mathbf{c}\}$ ;  $v^*$  is the *superadditive cover* of  $v$ . Note that  $v^*(e^J) = v^*(J)$ . We remark that we borrow the term “crisp” from Aubin, who introduced the concept of fuzzy games (and their crisp analogues) in [1]. Like OCF games, fuzzy games are also defined by functions from  $[0, 1]^n$  to  $\mathbb{R}$ . However, they are based on very different intuition, and, in particular, employ a very different notion of stability. We refer the reader to [4] for a detailed discussion of the differences between OCF games and fuzzy games.

It is often useful to think of the agents as using their resources to complete a given set of tasks. Such games are described in [4] and are called *Threshold Task Games (TTGs)*. A TTG comprises of a finite list of tasks,  $T = \{t_1, \dots, t_k\}$ , each  $t_i$  requires some weight  $w(t_i) \geq 0$  for its completion, and gives a certain payoff  $p(t_i) \geq 0$ . Each player  $i$  has some weight  $w_i \geq 0$  that he may allocate to the completion of any task. The worth of a coalition is

$$v(\mathbf{c}) = \max\{p(t_i) : w(t_i) \leq \sum_{i=1}^n c^i w_i\}.$$

We say that a function  $v$  has *Efficient Coalition Structure (ECS) property* if for any  $J \subseteq N$  and any  $\mathbf{w}_J \leq e^J$  there exists a coalition structure  $CS_J \in CS_J$  such that  $v^*(\mathbf{w}_J) = v(CS_J)$ . All U-finite continuous functions have the ECS property, since the set of all coalition structures with weight less than  $\mathbf{w}_J$  is compact (due to U-finiteness);  $v$  is continuous and hence achieves a maximum over

a compact set. TTGs also have the ECS property, as well as games with superadditive valuations.

We now define how agents share their payoffs. The *set of payoff vectors*  $X^*(G)$  of a game  $G = (N, v)$  is the set of all feasible payoff vectors for its crisp analogue:

$$X^*(G) = \{\mathbf{x} \in \mathbb{R}_+^n \mid \sum_{i=1}^n x^i \leq v^*(N)\}.$$

An *OCF solution concept* is a function that assigns every OCF game a subset of its payoff vectors. Note that the definition of a payoff vector allows transfers between different partial coalitions. However, usually we want to divide payoffs in a way that respects the coalition structure. The next definition paves the way for this.

**DEFINITION 2.2.** *An imputation for a coalition structure  $CS = (\mathbf{c}_1, \dots, \mathbf{c}_k) \in \mathcal{CS}_N$  is a  $n \times k$  matrix  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_k) \in M_{n \times k}(\mathbb{R}_+)$  that satisfies:*

- *Individual Rationality:*  $\sum_{j=1}^k x_j^i \geq v^*(\{i\})$  for all  $i \in N$ .
- *Payoff Distribution:* for all  $1 \leq j \leq k$  we have  $\sum_{i=1}^n x_j^i \leq v(\mathbf{c}_j)$ , and if  $c_j^i = 0$  then  $x_j^i = 0$ .

An imputation is a way for members of each partial coalition to divide profits among themselves; observe that inter-coalitional transfers are not allowed. We call a tuple  $(CS, \mathbf{x})$  of a coalition structure and an imputation a *feasible outcome*, let  $I(CS)$  denote the set of all matrices  $\mathbf{x}$  such that  $(CS, \mathbf{x})$  is a feasible outcome, and let  $\mathcal{F}(S)$  denote the set of all feasible outcomes over  $S$ .

Let  $(CS, \mathbf{x}) \in \mathcal{F}(N)$  be a feasible outcome. We define the *payoff* to an agent  $i \in N$  as  $p_i(CS, \mathbf{x}) = \sum_{j=1}^k x_j^i$ . This is the total payoff of  $i$  from all coalitions in  $CS$ . Similarly, the total payoff to a set  $J$  is  $p_J(CS, \mathbf{x}) = \sum_{i \in J} p_i(CS, \mathbf{x})$ . Note that the vector  $(p_1(CS, \mathbf{x}), \dots, p_n(CS, \mathbf{x}))$  is a payoff vector for  $G$ , since  $p_N(CS, \mathbf{x}) = \sum_{i=1}^n p_i(CS, \mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^k x_j^i = \sum_{j=1}^k \sum_{i=1}^n x_j^i \leq \sum_{j=1}^k v(\mathbf{c}_j) = v(CS) \leq v^*(N)$ .

The *support* of a coalition  $\mathbf{c} \in [0, 1]^n$  is the set of all players who devote their resources to  $\mathbf{c}$ . They are "interested parties" that may be hurt by any change to  $\mathbf{c}$ ; we write  $\text{supp}(\mathbf{c}) = \{i \in N \mid c^i > 0\}$ .

Given a coalition structure  $CS = (\mathbf{c}_1, \dots, \mathbf{c}_k)$  and some  $M \subseteq \{1, \dots, k\}$ , the coalitions whose indices are in  $M$  form a coalition structure; this coalition structure is denoted  $R(CS, M)$ .

Given a set  $J \subseteq N$ , we denote  $K_J = \{j \in \{1, \dots, k\} \mid \text{supp}(\mathbf{c}_j) \subseteq J\}$ .

**DEFINITION 2.3.** *The coalition structure  $CS$  reduced to  $J$  is defined as*

$$CS|_J = R(CS, K_J).$$

*These are all coalitions that are supported only by members of  $J$ . We let  $\overline{CS|_J}$  denote the complement of  $CS|_J$  in  $CS$ ;  $\overline{CS|_J}$  is the coalition structure consisting of all coalitions in  $CS$  that have non- $J$  members in their support.*

**DEFINITION 2.4.** *A coalition structure  $CS'$  is a deviation of  $J$  from  $CS$  if:*

- (1)  $\overline{CS|_J} = (\mathbf{c}_1 \dots \mathbf{c}_m)$ ,  $\overline{CS'|_J} = (\mathbf{d}_1 \dots \mathbf{d}_m)$  and there is some permutation of  $m$  elements,  $\sigma$ , such that for all  $1 \leq l \leq m$ :

$$\forall i \notin J : d_l^i = c_{\sigma(l)}^i \text{ and } \forall i \in J : d_l^i \leq c_{\sigma(l)}^i$$

- (2)  $w(CS'|_J) = w(CS|_J) + \sum_{l=1}^m (\mathbf{c}_{\sigma(l)} - \mathbf{d}_l)$ .

Note that if we assume that  $CS'$  is efficient, then condition (1) implies condition (2). The deviation  $CS'$  describes how a set  $J$  retracts resources from some coalitions and uses them in order to maximize its own welfare. Given a deviation  $CS'$  of  $J$  from  $CS$ , we define  $v^*(CS, J, CS')$  to be  $v^*(w(CS'|_J))$ ; if  $J$  withdraws all of its resources from  $\overline{CS|_J}$ , then the total weight available to  $J$  is  $e^J$ , and  $v^*(CS, J, CS') = v^*(J)$ . For brevity, given some  $\mathbf{c}_l$  in  $\overline{CS|_J}$  that  $J$  deviated from, we refer to the coalition after the deviation as  $dev_{CS'}(\mathbf{c}_l)$ .

### 3. THE ARBITRATION FUNCTION

To discuss stability in OCF games, we need to describe how agents react if some  $J$  deviates from  $(CS, \mathbf{x})$ . Paper [4] presents three different alternatives for such a reaction. The *conservative deviation*, or c-deviation, completely denies payoffs to  $J$ , even from coalitions that  $J$  did not affect. A relaxation of this approach leads to the notion of a *refined deviation*, or r-deviation. Under this deviation rule, deviators receive their share of the profit from all coalitions that were unaffected by the deviation, i.e. if no member of the deviating subset  $J$  changed his contribution to a coalition  $\mathbf{c}$ , then  $J$ 's payoff from  $\mathbf{c}$  is the same as before the deviation. Under the *optimistic deviation*, or o-deviation, the players in  $J$  receive their share of the profits from any coalition in which all non-deviators can still earn the same payoff as before the deviation. Paper [4] defines three notions of the core that correspond to the deviations.

One could easily think of many other reactions to deviation;  $J$  receives only half of its original payoffs in all coalitions outside of  $J$ 's support,  $J$  receives payoff only from those agents who are not worse off after the deviation, and many others. Note that all such rules can be thought of as a payoff function that is given the original coalition structure and  $J$ 's deviation, and then decides on an appropriate payoff to  $J$ . We call this function an *arbitration function* or an *arbitrator*. This function decides how much  $J$  gets from each coalition, given the nature of its deviation.

#### 3.1 Arbitration Functions

Suppose that we are given an outcome  $(CS, \mathbf{x}) \in \mathcal{F}(N)$ , a set of agents  $J \subseteq N$  and a deviation  $CS'$  of  $J$  from  $CS$ . Set  $\overline{CS|_J} = (\mathbf{c}_1 \dots \mathbf{c}_m)$  and  $\overline{CS'|_J} = (\mathbf{d}_1 \dots \mathbf{d}_m)$ .

**DEFINITION 3.1.** *The arbitration function is a mapping that assigns a real value to each coalition in  $\overline{CS|_J}$ .*

$$A(CS, \mathbf{x}, J, CS') = (\phi_l(CS, \mathbf{x}, J, CS'))_{l=1}^m$$

where  $\phi_l$  is a function that determines how much the coalition  $\mathbf{c}_l$  is willing to give  $J$  given its deviation. We require  $\phi_l$  to satisfy the following constraints:

- (1)  $\phi_l$  is non-negative.
- (2)  $\phi_l(CS, \mathbf{x}, J, CS')$  is only distributable between  $J \cap \text{supp}(\mathbf{c}_l)$ .
- (3) If  $v(dev_{CS'}(\mathbf{c}_l)) - p_{N \setminus J}(\mathbf{c}_l, \mathbf{x}) \geq 0$ , then  $\phi_l(CS, \mathbf{x}, J, CS') \leq v(dev_{CS'}(\mathbf{c}_l)) - p_{N \setminus J}(\mathbf{c}_l, \mathbf{x})$ , otherwise  $\phi_l(CS, \mathbf{x}, CS') = 0$ .
- (4)  $\phi_l$  is deviation-monotone: If  $K \subseteq J$  withdraws less resources from  $\mathbf{c}_l$ , then its payoff from  $\phi_l$  is higher.

Condition (2) means that members of  $J$  that were not entitled to payoffs from  $\mathbf{c}_l$  before the deviation are not entitled to payoffs after deviating; condition (3) states that a deviating set  $J$  must ensure that all members of  $N \setminus J$  in the support of  $\mathbf{c}_l$  are paid what they received under  $\mathbf{x}$  if it hopes to receive any profit from  $\mathbf{c}_l$ ; condition

(4) guarantees that the arbitrator behaves in a reasonable manner: agents know that their payoff from  $c_l$  is inversely proportional to the degree to which they hurt  $c_l$ . We illustrate the logic behind the concept of an arbitrator in the following example.

EXAMPLE 3.2. Consider a two-player TTG where there are three tasks:  $t_1, t_2, t_3$  with  $w(t_1) = 5, w(t_2) = 3, w(t_3) = 2$  and  $p(t_1) = 10, p(t_2) = 7, p(t_3) = 4$ . We have two agents: Alice, who has weight 4, and Bob, who has weight 1. The total weight of Alice and Bob is 5; they can complete  $t_2$  and  $t_3$  together and earn a total of 11. However, for the sake of our discussion, suppose that Alice and Bob agree on completing  $t_1$  using all of their weight, and dividing the payoff between them so that Alice receives 6 and Bob receives 4. The corresponding outcome is:  $((\frac{1}{1}), (\frac{6}{4}))$ . Consider the following deviation by Alice; she withdraws 2 units of weight in order to complete  $t_3$  on her own. This means that she still contributes 1 unit of weight to working with Bob, and they can still complete  $t_2$ , earning 7. According to condition (3) Bob must receive at least what he did before the deviation. Therefore, the most that Alice can expect to get from  $t_2$  is 3, under any arbitrator.

REMARK 3.3. We can allow the arbitrator more flexibility by permitting it to collect fines/make additional payments to the deviating subset. This can be captured by adding a *freely distributable value*, or a function  $\Psi(CS, \mathbf{x}, J, CS')$  to the arbitrator, where  $\Psi$  is also deviation-monotone. This model is more general as  $\Psi$  can be distributed among *all* members of  $J$  in any way they wish, while  $\phi_l$  may only be distributed among the members of  $J \cap \text{supp}(c_l)$ .  $\Psi$  can be, for example, a constant arbitration fee that must be paid by a deviating set. This fee is not related to any specific coalition, so the cost can be distributed between the agents in  $J$  in any way they see fit. All proofs in our paper go through for this more general model.

DEFINITION 3.4. The arbitration value of  $\mathcal{A}$  is

$$v^*(CS, J, CS') + \sum_{l=1}^m \phi_l(CS, \mathbf{x}, J, CS')$$

and denoted  $\text{val}(\mathcal{A}, CS, \mathbf{x}, J, CS')$ .

The arbitration value is the total payoff to a deviating set  $J$  given its deviation. The payoff is comprised of the most that  $J$  can make on its own plus the total payoff  $J$  receives from the coalitions it formed with non- $J$  members; the greater the arbitration value, the higher the incentive to deviate.

### 3.2 The Arbitrated Core

In the spirit of the definition given in [3], we now define a profitable deviation of a subset in an arbitrated game.

DEFINITION 3.5. Let  $\mathcal{A}$  be an arbitrator over  $G$ . An  $\mathcal{A}$ -profitable deviation of  $J$  from an outcome  $(CS, \mathbf{x})$  is an outcome  $(CS', \mathbf{y})$ , where

1.  $CS'$  is a deviation of  $J$  from  $CS$ .
2. For all  $c_l$  in  $\overline{CS|_J}$ ,  $p_J(\{dev_{CS'}(c_l)\}, \mathbf{y}) \leq \phi_l(CS, \mathbf{x}, CS')$ .
3. For all  $c_l$  in  $\overline{CS|_J}$ , if  $\phi_l(CS, \mathbf{x}, CS') > 0$ , then for all  $i \in N \setminus J$ , the payoff to  $i$  from the coalition  $dev_{CS'}(c_l)$  is equal to her payoff under  $c_l$ .
4.  $v(CS'|_J) = v^*(CS, J, CS')$  and  $\mathbf{y}$  reduced to the coalitions in  $CS'|_J$  is an imputation over  $J$ .
5. For any  $j \in J$ ,  $p_j(CS', \mathbf{y}) > p_j(CS, \mathbf{x})$ .

Condition 3 implies that a coalition  $c_l$  agrees to pay a deviating  $J$  only if each non- $J$  member gets the same payoff it received under  $\mathbf{x}$ .

DEFINITION 3.6. The  $\mathcal{A}$ -core of  $G = (N, v)$  is the set of all feasible outcomes in  $\mathcal{F}(N)$  that no subset of agents has an  $\mathcal{A}$ -profitable deviation from. The  $\mathcal{A}$ -core is denoted  $\mathcal{C}(\mathcal{A}, G)$ .

EXAMPLE 3.7. The arbitration function for the c-core is  $\phi_l \equiv 0$ . The arbitration function for the r-core is  $p_J(c_l, \mathbf{x})$  if  $c_l$  is the same after the deviation and is 0 otherwise. The arbitration function for the o-core is  $\max\{0, v(dev_{CS'}(c_l)) - p_{N \setminus J}(c_l, \mathbf{x})\}$ . We can define other forms of arbitrators. Set

$$N' = \{i \in N : i \notin J \text{ and } \exists c_l, i \in \text{supp}(c_l), dev_{CS'}(c_l) \neq c_l\}.$$

$N'$  is the set of all non-members of  $J$  who were hurt in some way by  $J$ 's deviation. One could naturally assume that players in  $N'$  would not like to pay members of  $J$  anymore in any coalition. In this case, the arbitration function would be  $p_J(c_l, \mathbf{x})$  if  $\text{supp}(c_l) \cap N' = \emptyset$  and 0 otherwise. We denote the core that corresponds to this arbitrator the *sensitive core*.

Note that if  $J$  can  $\mathcal{A}_1$ -profitably deviate using some deviation  $CS'$  and  $\mathcal{A}_1(CS, \mathbf{x}, J, CS') \leq \mathcal{A}_2(CS, \mathbf{x}, J, CS')$  (coordinate-wise), then  $J$  can  $\mathcal{A}_2$ -profitably deviate from  $(CS, \mathbf{x})$ . This implies that if for all outcomes  $(CS, \mathbf{x})$  and all deviations  $CS'$  of any  $J \subseteq N$  we have  $\mathcal{A}_1(CS, \mathbf{x}, J, CS') \leq \mathcal{A}_2(CS, \mathbf{x}, J, CS')$ , then  $\mathcal{C}(\mathcal{A}_2, G) \subseteq \mathcal{C}(\mathcal{A}_1, G)$ . Particularly we have

$$\text{o-core} \subseteq \text{r-core} \subseteq \text{sensitive-core} \subseteq \text{c-core}$$

This is a generalization of the result shown in [4].

## 4. CHARACTERIZATION OF THE ARBITRATED CORE

We now give a general characterization of the core under some arbitration function  $\mathcal{A}$ . Our proof method is similar to the proof of the characterization result given in [4].

THEOREM 4.1. If  $G = (N, v)$  has the ECS property, then an outcome  $(CS, \mathbf{x}) \in \mathcal{F}(N)$  is in  $\mathcal{C}(\mathcal{A}, G)$  if and only if for any  $J \subseteq N$  and deviation  $CS'$  we have  $p_J(CS, \mathbf{x}) \geq \text{val}(\mathcal{A}, CS, \mathbf{x}, J, CS')$ .

Simply put, an outcome is stable if and only if for any coalition  $J$  and any deviation proposed by  $J$ , the payoff that the members of  $J$  can obtain under  $\mathcal{A}$  does not exceed their current payoff.

PROOF. Suppose first that for every  $J \subseteq N$  and every deviation  $CS'$  of  $J$  from  $CS$  we have  $p_J(CS, \mathbf{x}) \geq \text{val}(\mathcal{A}, CS, \mathbf{x}, J, CS')$ . Therefore, for all  $\mathbf{y} \in I(CS')$ ,

$$p_J(CS', \mathbf{y}) \leq \text{val}(\mathcal{A}, CS, \mathbf{x}, J, CS') \leq p_J(CS, \mathbf{x}).$$

Hence, there exists a player  $j \in J$  that does not strictly benefit from  $(CS', \mathbf{y})$ , and  $J$  cannot  $\mathcal{A}$ -profitably deviate from  $(CS, \mathbf{x})$ .

Conversely, suppose that for some nonempty  $J \subseteq N$  there exists a deviation  $CS'$  such that  $p_J(CS, \mathbf{x}) < \text{val}(\mathcal{A}, CS, \mathbf{x}, J, CS')$ . We show that  $(CS, \mathbf{x})$  is not in the  $\mathcal{A}$ -core of  $G$ . Let  $\overline{CS|_J} = (c_1 \dots c_m)$ . For all  $j \in J$  let  $p_j = p_j(CS, \mathbf{x})$  and for all  $c_l$  in  $\overline{CS|_J}$ , let  $r_l = \phi_l(CS, \mathbf{x}, J, CS')$ . As  $v$  has the ECS property,  $v^*(CS, J, CS') = v(CS_M)$  for some coalition structure  $CS_M \in CS_J$ .

$\text{val}(\mathcal{A}, CS, \mathbf{x}, J, CS') = v(CS_M) + \sum_{l=1}^m r_l$ , which is strictly greater than  $p_J(CS, \mathbf{x})$ ; while  $J$  can strictly gain by deviating, it is possible that the members of  $J$  cannot divide the payoffs from the deviation in a manner that strictly benefits all of them. We now

show that there is a subset of  $J$  that can profitably deviate. Recall that  $r_l$  may only be distributed among  $J_l = \text{supp}(\mathbf{c}_l) \cap J$ . Given  $r_l$ , we define the set of all viable payoff divisions of  $r_l$  among the members of  $J_l$  as

$$\Delta_l = \{\rho \in \mathbb{R}_+^n \mid \sum_{i=1}^n \rho^i = r_l \text{ and } \rho^i = 0 \text{ for all } i \notin J_l\}.$$

Note that  $\Delta_l$  is compact. Given  $(CS, \mathbf{x})$ , we define its *total loss function*

$$TL_{(CS, \mathbf{x})} : I(CS_M) \times \prod_{l=1}^m \Delta_l \rightarrow \mathbb{R}.$$

Given an imputation  $\mathbf{y} \in I(CS_M)$  and  $(\rho_l)_{l=1}^m \in \prod_{l=1}^m \Delta_l$ , we define the total payoff to player  $j \in J$  as

$$q_j = p_j(CS_M, \mathbf{y}) + \sum_{l=1}^m \rho_l^j.$$

The total loss of a payoff division is

$$TL_{(CS, \mathbf{x})}(\mathbf{y}, (\rho_l)_{l=1}^m) = \sum_{j \in J \mid p_j > q_j} p_j - q_j.$$

$TL_{(CS, \mathbf{x})}$  is a continuous, real valued function over a compact set, so there is some payoff division  $(\mathbf{y}, (\rho_l)_{l=1}^m)$  that minimizes  $TL_{(CS, \mathbf{x})}$ . Given a loss-minimizing payoff division  $(\mathbf{y}, (\rho_l)_{l=1}^m)$ , we construct a directed graph  $\Gamma = (V, E)$  where  $V = J$ , and there is a directed edge from  $i \in J$  to  $j \in J$  if and only if  $i$  can legally transfer payments to  $j$ ; this can happen if and only if both  $i$  and  $j$  are in the support of some coalition  $\mathbf{c}$  and  $i$  receives a positive payoff from  $\mathbf{c}$ . We color the vertices of the graph as follows: a vertex  $j$  is green if  $p_j < q_j$ , white if  $p_j = q_j$ , and red if  $p_j > q_j$ . Since  $\sum_{j \in J} q_j > p_J(CS, \mathbf{x})$ , the graph has at least one green vertex. If all vertices are green, then  $p_j < q_j$  for all  $j \in J$ , i.e.,  $CS'$  is a profitable deviation for all players in  $J$ , and we are done. We now assume that there is at least one non-green vertex.

Note that if  $g \in J$  is green, then if there is an edge from  $g$  to some  $j$ , then  $g$  can transfer a small amount of payoff  $0 < \delta < q_g - p_g$  to  $j$ . If  $\delta$  is small enough, then the resulting outcome is still a viable imputation. Similarly,  $j$  can legally transfer the same amount to any vertex that  $j$  is connected to. Therefore, if there is a path from a green vertex  $g$  to some  $j$ , then  $g$  can transfer a small amount  $\delta$  to  $j$  while remaining green. Following [4], we observe that since we chose a payoff distribution that minimizes  $TL_{(CS, \mathbf{x})}$ , if there is a path from a green vertex  $g$  to some vertex  $i$ , then  $i$  is not red. Thus, we can assume w.l.o.g. that if a vertex  $i$  is not green, then there is no path from a green vertex to  $i$ . Let  $G_J$  be the set of all green vertices; we claim that  $G_J$  can  $\mathcal{A}$ -profitably deviate. Indeed, note that making  $J \setminus G_J$  return to their original contributions according to  $CS$  will not negatively affect  $G_J$  (here we use the fact that  $\mathcal{A}$  is deviation-monotone); if  $G_J$  decides to deviate from  $CS$ , without having  $J \setminus G_J$  deviate as well, its payoffs cannot decrease. Therefore,  $G_J$  can  $\mathcal{A}$ -profitably deviate from  $CS$ , and we are done.  $\square$

## 4.1 The Refined and Optimistic Cores

Theorem 4.1 immediately implies the characterization of the conservative core given in [4]. Let  $\mathcal{A}^c$  denote the conservative arbitrator; under  $\mathcal{A}^c$  we have  $\phi_l(CS, \mathbf{x}, J, CS') \equiv 0$ , so any deviating set should not leave any of its resources in any coalition, but rather devote all of its resources to maximize  $v^*(CS, J, CS')$ . Therefore,

$$\sup_{CS'} \{val(\mathcal{A}^c, CS, \mathbf{x}, J, CS')\} = \sup_{CS'} \{v^*(CS, J, CS')\} = v^*(J).$$

Our characterization result indeed shows that  $(CS, \mathbf{x})$  is c-stable if and only if for all  $J \subseteq N$ ,  $p_J(CS, \mathbf{x}) \geq v^*(J)$ .

Theorem 4.1 also gives an intuitive characterization of the refined and optimistic cores; under the refined arbitrator, denoted  $\mathcal{A}^r$ , a deviating subset  $J$  can expect payoff only from coalitions it did not change. Thus, if  $J$  decides to deviate from a coalition, it should withdraw all of its resources from that coalition. The arbitration value of  $\mathcal{A}^r$  is  $v^*(CS, J, CS') + p_J(U, \mathbf{x})$ , where  $U$  is a matrix whose columns are the coalitions unchanged by  $J$ 's deviation. Consequently, an outcome  $(CS, \mathbf{x})$  is in the r-core if and only if for all  $J \subseteq N$  and any deviation of  $J$ ,  $CS'$ , we have  $p_J(CS, \mathbf{x}) \geq v^*(CS, J, CS') + p_J(U, \mathbf{x})$ . To conclude, an outcome  $(CS, \mathbf{x})$  is in the r-core if and only if for any  $J \subseteq N$  and any matrix  $Q$  whose column vectors are coalitions in  $\overline{CS|_J}$  we have

$$p_J(CS|_J, \mathbf{x}) + p_J(Q, \mathbf{x}) \geq v^*(w(CS|_J) + w(Q)|_J).$$

Note that if  $Q = \overline{CS|_J}$ , we get the c-core condition.

Under the optimistic arbitrator, denoted  $\mathcal{A}^o$ ,  $J$  can expect payoff from a coalition  $\mathbf{c}$  if all non- $J$  members of  $\mathbf{c}$  get the payoff they received under  $\mathbf{x}$ . Thus, the payoff available to  $J$  is  $v(\text{dev}_{CS'}(\mathbf{c}_l)) - p_{N \setminus J}(\mathbf{c}_l, \mathbf{x})$ . Let us denote by  $P(CS)$  the coalitions from which  $J$  can expect payoff if it makes the deviation  $CS'$ , and by  $P(CS')$  the same coalitions after  $J$ 's deviation. We define  $N(CS)$  to be the coalitions that will not pay  $J$ . By Theorem 4.1 we have that  $(CS, \mathbf{x}) \in \mathcal{C}(\mathcal{A}^o, G)$  if and only if

$$p_J(CS, \mathbf{x}) \geq v^*(CS, J, CS') + v(P(CS')) - p_{N \setminus J}(P(CS), \mathbf{x}).$$

Since  $p_N(P(CS), \mathbf{x}) = v(P(CS))$ , it follows that  $(CS, \mathbf{x}) \in \mathcal{C}(\mathcal{A}^o, G)$  if and only if

$$p_J(N(CS), \mathbf{x}) \geq v^*(CS, J, CS') + v(P(CS')) - v(P(CS)).$$

Note that if we only consider deviations where  $J$  withdraws all of its resources from  $N(CS)$  and does not change its contribution to  $P(CS)$ , then we get the characterization of the r-core. Also note that checking if an outcome is in the r-core or c-core can be done by considering a finite number of deviations, but for o-core this is not the case.

## 5. THE NUCLEOLUS OF AN ARBITRATED OCF GAME

Although the core of a game is a useful solution concept, it may be empty in some cases; it is desirable to have a solution concept that is more robust, and, in particular, is guaranteed to be non-empty for all (reasonable) OCF games. In the non-overlapping setting, this role is fulfilled by the nucleolus [11]. We extend the notion of nucleolus to OCF games, and show that it exhibits many of the desirable properties of its non-OCF counterpart.

Let  $\mathcal{A}^*(CS, \mathbf{x}, J)$  denote the most that  $J \subseteq N$  can receive from an arbitrator given an outcome  $(CS, \mathbf{x})$ . In this section, we only consider OCF games for which  $\mathcal{A}^*(CS, \mathbf{x}, J)$  is a well-defined real value for any  $CS, \mathbf{x}$  and  $J$ .

**DEFINITION 5.1.** *Given an outcome  $(CS, \mathbf{x})$ , the excess of  $J \subseteq N$  is defined as  $e(CS, \mathbf{x}, J) = \mathcal{A}^*(CS, \mathbf{x}, J) - p_J(CS, \mathbf{x})$ .*

The excess is a measure of a subset's "unhappiness" with a given outcome: the lower the excess, the happier the subset. Note also that Theorem 4.1 states that  $(CS, \mathbf{x}) \in \mathcal{C}(\mathcal{A}, G)$  if and only if  $e(CS, \mathbf{x}, J) \leq 0$  for all  $J \subseteq N$ . Given an outcome  $(CS, \mathbf{x})$ , we define its *excess vector* as

$$\theta(CS, \mathbf{x}) = (e(CS, \mathbf{x}, S_1), e(CS, \mathbf{x}, S_2), \dots, e(CS, \mathbf{x}, S_{2^n})),$$

where  $e(CS, \mathbf{x}, S_1) \geq \dots \geq e(CS, \mathbf{x}, S_{2^n})$ . We write  $(CS, \mathbf{x}) \preceq_L (CS', \mathbf{y})$  if  $\theta(CS, \mathbf{x})$  is lexicographically smaller than  $\theta(CS', \mathbf{y})$ .

We point out that Definition 5.1 coincides with the definition of excess for classic TU cooperative games; given a classic TU game  $(N, u)$ , the most that the set  $J$  can get is simply  $u(J)$ , and the excess is defined as the difference between  $u(J)$  and the payoff to  $J$ . This analogous definition gives rise to an analogous definition of an arbitrated nucleolus.

Given an OCF game  $G = (N, v)$  arbitrated by  $\mathcal{A}$ , the arbitrated nucleolus of  $G$ , denoted  $\mathcal{N}(\mathcal{A}, G)$ , is the set of all outcomes in  $\mathcal{F}(N)$  that are minimal with respect to  $\preceq_L$ . Observe that just like in the non-overlapping case, if  $\mathcal{C}(\mathcal{A}, G) \neq \emptyset$ , then  $\mathcal{N}(\mathcal{A}, G) \subseteq \mathcal{C}(\mathcal{A}, G)$ .

## 5.1 Non-Emptiness of the Nucleolus

Unlike the arbitrated core, the nucleolus is never empty as long as  $\mathcal{A}^*(CS, \mathbf{x}, J)$  is continuous with respect to  $(CS, \mathbf{x})$ . In fact, it suffices that the excess of a set be achievable by using some deviation from a given outcome. This is true for the arbitrators defined above, assuming that  $v$  has the ECS property.

**THEOREM 5.2.** *If  $v$  has the ECS property and  $\mathcal{A}^*(CS, \mathbf{x}, J)$  is continuous w.r.t  $(CS, \mathbf{x})$ , then  $\mathcal{N}(\mathcal{A}, G) \neq \emptyset$*

**PROOF.** First, we would like to note that the excess vector is comprised of continuous functions over  $\mathcal{F}(N)$ . Indeed, observe that for any outcome  $(CS, \mathbf{x})$  and any  $k = 1, \dots, 2^n$  we have

$$\theta_k(CS, \mathbf{x}) = \max_{S_1, \dots, S_k \subseteq N} \{\min\{e(CS, \mathbf{x}, S_1), \dots, e(CS, \mathbf{x}, S_k)\}\},$$

where all  $S_1, \dots, S_k$  are different subsets of  $N$ . Since  $\mathcal{A}^*(CS, \mathbf{x}, J)$  is continuous, so is the excess. Thus,  $\theta_k$  is obtained by combining continuous functions using a finite number of min and max operations, and therefore it is continuous as well.

Set  $X_1 = \{(CS, \mathbf{x}) = \operatorname{argmin}_{(CS', \mathbf{y}) \in \mathcal{F}(N)} \{\theta_1(CS', \mathbf{y})\}\}$ , and for every  $k = 2, \dots, 2^n$ , let

$$X_k = \{(CS, \mathbf{x}) = \operatorname{argmin}_{(CS', \mathbf{y}) \in X_{k-1}} \{\theta_k(CS', \mathbf{y})\}\}.$$

$X_{2^n} \subseteq \mathcal{N}(\mathcal{A}, G)$ , since if  $(CS, \mathbf{x}) \in X_k$  then  $\theta_k(CS, \mathbf{x}) \leq \theta_k(CS', \mathbf{y})$  for every  $k = 1, \dots, 2^n$ . Thus, it remains to show that  $X_{2^n}$  is non-empty. Now, the set  $\mathcal{F}(N)$  is compact and non-empty. From elementary calculus, we know that if  $C \subseteq \mathbb{R}^m$  is a non-empty compact set, and  $f : C \rightarrow \mathbb{R}$  is a continuous function, then the set  $X = \{x \in C \mid f(x) = \min_{y \in C} \{f(y)\}\}$  is a non-empty compact set. Hence,  $X_1$  is compact and non-empty, and inductively so is  $X_{2^n}$ . Consequently,  $\mathcal{N}(\mathcal{A}, G) \neq \emptyset$ .  $\square$

## 5.2 Properties of the Nucleolus

The nucleolus in the non-overlapping setting exhibits some attractive properties. For example, in the non-overlapping setting, the nucleolus is a single point [9, 18]. In the arbitrated OCF setting, however, the nucleolus may have a richer structure.

**EXAMPLE 5.3.** Consider the following TTG:  $N = \{1, 2\}$ . Both players have weight of 1 and there is one task  $t$  with  $w(t) = 2, p(t) = 20$ . Assume that  $G$  is arbitrated by the refined arbitrator. The  $r$ -core of the game is not empty and the only coalition structure that is in the  $r$ -core is  $CS = (\frac{1}{1})$ . Let us consider a payoff distribution where player 1 gets  $10 - \varepsilon$  and player 2 gets  $10 + \varepsilon$  where  $0 < \varepsilon < 10$ . The maximum value that can be provided to player 1 under the refined arbitrator is if he offers  $CS$  as his objection; any other deviation will leave him with nothing. Indeed,  $\mathcal{A}^*(CS, \mathbf{x}, \{1\}) = p_1(CS, \mathbf{x}) = 10 - \varepsilon$ , thus his excess is 0. One can verify that all nucleolus outcomes have excess of 0 for all sets

in this game. However, if the same game is arbitrated by the conservative arbitrator, then the excess of player 1 is  $0 - (10 - \varepsilon) = \varepsilon - 10$ , which will make him sensitive to the fact that he is being cheated.

Example 5.3 demonstrates that outcomes in  $\mathcal{N}(\mathcal{A}, G)$  need not be unique, nor distribute payoff among players in the same manner. However, it turns out that if  $\mathcal{A}^*(CS, \mathbf{x}, J)$  is convex as a function of  $\mathbf{x}$  when  $J$  and  $CS$  are fixed, then for any two outcomes in the nucleolus that have the same coalition structure, each subset of players has the same excess under both of these outcomes. First, we need the following technical lemma:

**LEMMA 5.4.** *If  $(CS, \mathbf{x}), (CS, \mathbf{y}) \in \mathcal{N}(\mathcal{A}, G)$  and  $\mathbf{z} = \frac{\mathbf{x} + \mathbf{y}}{2}$ , then  $(CS, \mathbf{z}) \in \mathcal{N}(\mathcal{A}, G)$ .*

**PROOF.** Suppose that  $(CS, \mathbf{x}), (CS, \mathbf{y}) \in \mathcal{N}(\mathcal{A}, G)$ . Both outcomes must have the same excess vector, i.e.  $\theta(CS, \mathbf{x}) = \theta(CS, \mathbf{y})$ . Set  $\mathbf{z} = \frac{\mathbf{x} + \mathbf{y}}{2}$ . Since  $I(CS)$  is convex,  $\mathbf{z} \in I(CS)$ . Consider  $\theta(CS, \mathbf{z})$ . Denote

$$\begin{aligned} \theta(CS, \mathbf{x}) &= (e(CS, \mathbf{x}, J_1), \dots, e(CS, \mathbf{x}, J_{2^n})), \\ \theta(CS, \mathbf{y}) &= (e(CS, \mathbf{y}, K_1), \dots, e(CS, \mathbf{y}, K_{2^n})), \\ \theta(CS, \mathbf{z}) &= (e(CS, \mathbf{z}, L_1), \dots, e(CS, \mathbf{z}, L_{2^n})). \end{aligned}$$

Given a deviation  $CS'$  of  $J$  from  $CS$ , we have

$$\mathcal{A}^*(CS, \mathbf{z}, J) \leq \frac{\mathcal{A}^*(CS, \mathbf{x}, J) + \mathcal{A}^*(CS, \mathbf{y}, J)}{2},$$

since  $\mathcal{A}^*$  is convex. Since the payoffs to  $J$  are linear in the imputations, we conclude that

$$e(CS, \mathbf{z}, J) \leq \frac{1}{2}e(CS, \mathbf{x}, J) + \frac{1}{2}e(CS, \mathbf{y}, J).$$

Denote  $e(CS, \mathbf{x}, J_l) = e(CS, \mathbf{y}, K_l) = V_l$ . We get

$$e(CS, \mathbf{z}, L_1) \leq \frac{e(CS, \mathbf{x}, L_1) + e(CS, \mathbf{y}, L_1)}{2} \leq V_1.$$

If at any point the inequality is strict,  $e(CS, \mathbf{z}, L_1) < V_1$  and  $\theta(CS, \mathbf{z})$  is strictly smaller lexicographically than  $\theta(CS, \mathbf{x})$ , a contradiction. We similarly conclude that  $e(CS, \mathbf{z}, L_k) = V_k$  for all  $k = 1, \dots, 2^n$ . Therefore  $\theta(CS, \mathbf{z}) = \theta(CS, \mathbf{x}) = \theta(CS, \mathbf{y})$ , and  $(CS, \mathbf{z}) \in \mathcal{N}(\mathcal{A}, G)$ .  $\square$

**THEOREM 5.5.** *Let  $G = (N, v)$  be a game arbitrated by some convex arbitrator  $\mathcal{A}$ . If  $(CS, \mathbf{x}), (CS, \mathbf{y}) \in \mathcal{N}(\mathcal{A}, G)$  then for any  $J \subseteq N$  we have  $e(CS, \mathbf{x}, J) = e(CS, \mathbf{y}, J)$ .*

**PROOF.** The proof scheme is somewhat similar to the proof that the nucleolus for non-OCF games is unique [18, 9]. Let  $(CS, \mathbf{x})$  and  $(CS, \mathbf{y})$  be in  $\mathcal{N}(\mathcal{A}, G)$ . Set  $\mathbf{z} = \frac{\mathbf{x} + \mathbf{y}}{2}$ . Using the same notation as in Lemma 5.4, we know that  $e(CS, \mathbf{x}, J_1)$  is equal to  $e(CS, \mathbf{y}, K_1)$  and  $e(CS, \mathbf{z}, L_1)$ , so

$$e(CS, \mathbf{x}, J_1) + e(CS, \mathbf{y}, K_1) = 2e(CS, \mathbf{z}, L_1).$$

As shown in Lemma 5.4,

$$2e(CS, \mathbf{z}, L_1) \leq e(CS, \mathbf{x}, L_1) + e(CS, \mathbf{y}, L_1).$$

By definition of  $J_1$ ,  $e(CS, \mathbf{x}, L_1) \leq e(CS, \mathbf{x}, J_1)$  and similarly,  $e(CS, \mathbf{y}, L_1) \leq e(CS, \mathbf{y}, K_1)$ . This implies that

$$e(CS, \mathbf{y}, K_1) = e(CS, \mathbf{y}, L_1) = e(CS, \mathbf{x}, J_1) = e(CS, \mathbf{x}, L_1).$$

We can swap  $L_1$  with  $J_1$  in the excess ordering of  $(CS, \mathbf{x})$  without changing the excess vector. This can be done inductively for any  $L_k$ . We conclude that if  $(CS, \mathbf{x}), (CS, \mathbf{y}) \in \mathcal{N}(\mathcal{A}, G)$  then all sets have the same excess in both outcomes.  $\square$

The conservative arbitrator is constant at  $v^*(J)$ , and thus convex. This allows us to use Theorem 5.5 to show that for the conservative arbitration function any two outcomes in the nucleolus correspond to identical payoff vectors.

**COROLLARY 5.6.** *If  $G$  is arbitrated by the conservative arbitrator, then for any  $(CS, \mathbf{x}), (CS, \mathbf{y}) \in \mathcal{N}(\mathcal{A}^c, G)$  and any  $i \in N$ ,  $p_i(CS, \mathbf{x}) = p_i(CS, \mathbf{y})$ .*

**PROOF.** Consider two outcomes  $(CS, \mathbf{x}), (CS, \mathbf{y}) \in \mathcal{N}(\mathcal{A}^c, G)$  and a player  $i$ . By Theorem 5.5,  $\mathcal{A}^{c*}(CS, \mathbf{x}, J) - p_i(CS, \mathbf{x}) = \mathcal{A}^{c*}(CS, \mathbf{x}, J) - p_i(CS, \mathbf{y})$ . On the other hand,  $\mathcal{A}^{c*}(CS, \mathbf{x}, J) = \mathcal{A}^{c*}(CS, \mathbf{y}, J) = v^*(\{i\})$ , so  $p_i(CS, \mathbf{x})$  must equal  $p_i(CS, \mathbf{y})$ .  $\square$

We remark that one can also show that the refined arbitrator is convex. However, as illustrated by Example 5.3, the conclusion of Corollary 5.6 does not hold for the refined arbitrator, since the value of  $\mathcal{A}_z^*(CS, \mathbf{z}, J)$  may depend on the vector  $\mathbf{z}$ .

## 6. THE SHAPLEY VALUE OF OCF GAMES

Introduced by L.S. Shapley in [12], the Shapley value is a central solution concept in classic cooperative game theory. We offer two possible extensions of the Shapley value to OCF games; one assumes a fixed coalition structure and is somewhat similar to the Shapley value for coalition structures defined in [2], while the other takes into account the ability of sets to maximize their profits using coalition structures and is similar to the classic notion defined in [12]. We show that both values are unique with regard to specific sets of axioms<sup>1</sup>. In this section, we denote the Shapley value for crisp games by  $sv$ .

Our first definition assumes that the coalition structure  $CS$  is given; it is possible that the agents have agreed on some division of labor, or one was assigned to them by a central authority. The following definition of a value provides an axiomatic method of assessing the contribution of each player to  $CS$ . Given a game  $G = (N, v)$  and a coalition  $\mathbf{c} \in [0, 1]^n$  we set forth the following axioms for a value  $(\Phi_1(N, v, \mathbf{c}), \dots, \Phi_n(N, v, \mathbf{c}))$ .

- (1) **Coalitional Efficiency:**  $\sum_{i=1}^n \Phi_i(N, v, \mathbf{c}) = v(\mathbf{c})$ .
- (2) **Symmetry:** Two players  $i, j \in N$  are *OCF-symmetric* if for all  $\mathbf{x} \in [0, 1]^n$ ,  $v(\mathbf{x}) = v(\mathbf{x}_{i \sim j})$ , where  $\mathbf{x}_{i \sim j}$  is  $\mathbf{x}$  with the  $i$ -th and  $j$ -th coordinates exchanged. If  $i, j$  are OCF-symmetric, then  $\Phi_i(N, v, \mathbf{c}) = \Phi_j(N, v, \mathbf{c})$ .
- (3) **Dummy Player:** Set  $\mathbf{c}_{-i}$  to be  $\mathbf{c}$  with the  $i^{\text{th}}$  coordinate set to 0. If  $v(\mathbf{c}_{-i}) = v(\mathbf{c})$  then  $\Phi_i(N, v, \mathbf{c}) = 0$ .
- (4) **Additivity:**  $\Phi_i(N, v, \mathbf{c}) + \Phi_i(N, u, \mathbf{c}) = \Phi_i(N, u + v, \mathbf{c})$ .

We define  $\alpha_v : [0, 1]^n \times 2^N \rightarrow \mathbb{R}$  as  $\alpha_v(\mathbf{c}, S) = v(\mathbf{c}|_S)$ . Given a coalition  $\mathbf{c}$ , the *coalitional OCF Shapley value* of  $\mathbf{c}$ , denoted  $SV_i(N, v, \mathbf{c})$ , is  $sv_i(\alpha_v(\mathbf{c}, \cdot))$ . One can verify that

- $\alpha_v(\mathbf{c}, N) = v(\mathbf{c})$ ;
- if two players are OCF-symmetric then they are symmetric in  $\alpha_v(\mathbf{c}, \cdot)$ ;
- if  $i$  is a dummy then  $\alpha_v(\mathbf{c}, S \cup \{i\}) = v(\mathbf{c}_{S \cup \{i\}}) = v(\mathbf{c}_S) = \alpha_v(\mathbf{c}, S)$ , hence  $i$  is dummy in  $\alpha_v(\mathbf{c}, \cdot)$ ;
- $\alpha_u(\mathbf{c}, S) + \alpha_v(\mathbf{c}, S) = u(\mathbf{c}_S) + v(\mathbf{c}_S) = (u + v)(\mathbf{c}_S) = \alpha_{u+v}(\mathbf{c}, S)$ .

<sup>1</sup>For the axiomatization of the classic Shapley value, see [9] chapter 8, pp. 151-179, as well as the detailed review in [16]

This shows that the OCF properties described above naturally translate to their equivalents in non-OCF games. Hence, the coalitional OCF Shapley value satisfies properties (1)–(4). To show uniqueness, we use the following construction: given a function  $u : 2^N \rightarrow \mathbb{R}$ , define  $v : [0, 1]^n \rightarrow \mathbb{R}$  by setting  $v(\mathbf{x}) = u(S)$  if  $\mathbf{x} = \mathbf{c}_S$  for some  $S \subseteq N$  and  $v(\mathbf{x}) = 0$  otherwise. Clearly, we have  $\alpha_v(\mathbf{c}, S) = u(S)$  for any  $S \subseteq N$ . Therefore, uniqueness of the coalitional OCF Shapley value follows from the uniqueness of the classic Shapley value. The coalitional OCF Shapley value can be extended to coalition structures by setting  $SV_i(N, v, CS) = \sum_{j=1}^k SV_i(N, v, \mathbf{c}_j)$ , where  $CS = (\mathbf{c}_1 \dots \mathbf{c}_k)$ . It is immediate that  $SV(N, v, CS)$  is efficient, i.e., the sum of the players' values is  $v(CS)$ , and the value of each coalition is distributed only among those who support it.

An alternative approach for measuring power does not assume a preexisting coalition structure, but rather measures the a-priori marginal contribution of a player, as all players try to maximize social welfare by forming coalition structures. Young [17] gives a characterization of the Shapley value using the notion of *strong monotonicity*; we use a similar notion for a value. We begin by setting forth the desirable axioms.

- (1) **Strong Monotonicity:** if for some  $u, v : [0, 1]^n \rightarrow \mathbb{R}$  and some  $i \in N$  we have  $v^*(\mathbf{c}) - v^*(\mathbf{c}_{-i}) \geq u^*(\mathbf{c}) - u^*(\mathbf{c}_{-i})$  for all  $\mathbf{c} \in [0, 1]^n$ , then  $\Phi_i(N, v) \geq \Phi_i(N, u)$ .
- (2) **Symmetry:** A value  $\Phi$  is *symmetric* if for any two symmetric players  $i, j$ :  $\Phi_i(N, v) = \Phi_j(N, v)$ .
- (3) **Efficiency:**  $\sum_{i=1}^n \Phi_i(N, v) = v^*(N)$ .

Axiom (1) states that if a player  $i$  has higher marginal contribution to  $v^*(\mathbf{c})$  than to  $u^*(\mathbf{c})$  for any  $\mathbf{c}$ , then her value in  $v$  should be higher; this is a generalization of strong monotonicity as defined in [17]. Also note that if two players are OCF-symmetric, then they are symmetric as players in the crisp analogue of the game. Finally, these notions are only well-defined assuming that the game  $G$  has the ECS property. We define the *OCF Shapley value*, denoted  $SV^*(N, v)$ , as

$$SV_i^*(N, v) = sv_i(N, v^*).$$

Strong monotonicity, efficiency and symmetry are inherited from their classic counterparts for  $sv(N, v^*)$ . Note also that the class of crisp analogues of OCF games corresponds to the class of superadditive games. Recall that a function  $u : 2^N \rightarrow \mathbb{R}$  is called *superadditive* if for all disjoint  $S, T \subseteq N$  it holds that  $v(S) + v(T) \leq v(S \cup T)$ . One can verify that the crisp analogue of any OCF game is superadditive. Moreover, given a superadditive  $u : 2^N \rightarrow \mathbb{R}$ , one can define the function  $v : [0, 1]^n \rightarrow \mathbb{R}$  to be  $v(e^S) = u(S)$  and 0 otherwise; for all  $S \subseteq N$ ,  $v^*(S) = u(S)$ . Therefore, the uniqueness of the Shapley value for superadditive games implies its uniqueness for the class of crisp analogues, which in turn implies its uniqueness for OCF games.

The two notions of Shapley value for OCF games considered above do not, in general, coincide, even if the coalition structure for which we compute the coalitional OCF Shapley value is socially optimal.

**EXAMPLE 6.1.** Consider a 3-player TTG with  $w_1 = 5, w_2 = 2, w_3 = 1$ , and two tasks,  $t_1, t_2$ , with  $p(t_1) = 6, p(t_2) = 12$  and  $w(t_1) = 4, w(t_2) = 8$ . Let us compute  $SV^*(N, v)$ . When player 1 is first or second he has marginal contribution of 6. When he is last, his marginal contribution is 12. Therefore,  $SV_1^*(N, v) = 8$ , and, by efficiency and symmetry,  $SV_2^*(N, v) = SV_3^*(N, v) = 2$ . However, consider a coalition structure where players work on two

copies of  $t_1$ , and each of them contributes half of his resources to each copy. Then any player has non-zero marginal contribution only if he is last, in which case he contributes 12. Therefore  $SV_1(N, v, CS) = SV_2(N, v, CS) = SV_3(N, v, CS) = 4$ .

In Example 6.1, player 1 can contribute significantly more than the other players, but his coalitional OCF Shapley value is equal to theirs in  $CS$ . This is because in the *specific* coalition structure  $CS$ , his marginal contribution is the same as his peers'; if any one of them leaves a coalition, the value of the remaining coalition structure becomes zero.

Note also that if two players in a TTG have  $w_i = w_j$ , then  $SV_i^*(N, v) = SV_j^*(N, v)$ . However, this is not necessarily true for the coalitional Shapley value.

EXAMPLE 6.2. Consider a 2-player TTG where both players have weight  $w \geq 2$ , and there are two tasks;  $w(t_1) = 2w - 1$ ,  $w(t_2) = 1$  and  $p(t_1) = M, p(t_2) = x$ , where  $(2w - 1)x \leq M$ . We form  $CS$  so that player 1 contributes all of her weight to  $t_1$ , while player 2 contributes  $w - 1$  to  $t_1$ , and completes  $t_2$  by herself. When player 1 is first, then  $v(\binom{1}{0}) = x$ , while under the current coalition structure, player 2 can gain  $2x$  on her own.  $SV_1(N, v, CS) = \frac{1}{2}(x + M + x - 2x) = \frac{M}{2}$ , and by efficiency  $SV_2(N, v, CS) = \frac{M+2x}{2}$ .

Example 6.2 implies that the difference between  $SV_i^*(N, v)$  and  $SV_i(N, v, CS)$  can be arbitrarily large.

## 7. CONCLUSIONS AND FUTURE WORK

Our work shows significant similarity between concepts from classic cooperative game theory and their OCF counterparts. We can also generalize the notion of the bargaining set [6] to OCF games; the resulting notion shares many properties with the bargaining set in crisp games. We omit these results due to space constraints. Other solution concepts, such as the  $\varepsilon$ -core, can be described using the arbitration function itself, using the freely distributable component mentioned in Remark 3.3.

The arbitrated OCF model is far from being fully explored. We would like to point out a few promising directions for further research. First, it is shown in [13] that the Shapley value is in the core of a convex game. We would like to see if this result extends to the OCF setting. While [4] define convex OCF games and show that their c-core is not empty, it is not clear if there is an outcome in the c-core of a convex game such that each player is paid exactly her OCF Shapley value. In order to do so, we must find some coalition structure that corresponds to such a payoff scheme.

Another promising direction is exploring processes of overlapping coalition formation. While some work has been done on overlapping coalition formation algorithms [8], coalitional stability is yet to be fully explored. Our work assumes that an outcome is exogenously determined, and does not describe a process under which a stable outcome may arise. While [4] proposes a coalition formation procedure for convex OCF games, it is not clear how to extend it to the general case. A decentralized coalition formation algorithm, where agents repeatedly form and dissolve coalitions until a stable coalition structure and payoff division are agreed upon, would be useful in many multi-agent scenarios. The notion of arbitrators may play a significant role in such a process, as the arbitration function can effectively control the degree to which agents will be inclined to deviate.

Finally, we would like to investigate the algorithmic properties of the solution concepts defined in this paper. For example, it is clear that it is generally hard to compute the OCF Shapley value,

even when one can compute  $v^*$  in polynomial time. However, it would be interesting to identify natural classes of games where the Shapley value is tractable.

## 8. ACKNOWLEDGMENTS

This research was supported by National Research Foundation (Singapore) under grant 2009-08, by NTU SUG (Edith Elkind), and by SINGA graduate fellowship (Yair Zick).

## 9. REFERENCES

- [1] J. Aubin. Cooperative fuzzy games. *Mathematics of Operations Research* 6(1):1–13, 1981.
- [2] R. Aumann and J. Drèze. Cooperative games with coalition structures. *International Journal of Game Theory*, 3:217–237, 1974.
- [3] G. Chalkiadakis, E. Elkind, E. Markakis, and N. Jennings. Overlapping coalition formation. In *WINE'08*, pp. 307–321, 2008.
- [4] G. Chalkiadakis, E. Elkind, E. Markakis, M. Polukarov, and N. Jennings. Cooperative games with overlapping coalitions. *Journal of Artificial Intelligence Research* 39:179–216, 2010.
- [5] V. D. Dang, R. K. Dash, A. Rogers, and N. R. Jennings. Overlapping coalition formation for efficient data fusion in multi-sensor networks. In *AAAI'06*, pp. 635–640, 2006.
- [6] M. Davis and M. Maschler. Existence of stable payoff configurations for cooperative games. *Bulletin of the American Mathematical Society* 69:106–108, 1963.
- [7] P. Dubey. On the uniqueness of the Shapley value. *International Journal of Game Theory* 4(3):131–139, 1975.
- [8] C. Lin and S. Hu. Multi-task overlapping coalition parallel formation algorithm. In *AAMAS'07*, pp. 1260–1262, 2007.
- [9] B. Peleg and P. Sudhölter. *Introduction to the Theory of Cooperative Games*, Springer, second edition, 2007.
- [10] T. Michalak, J. Sroka, T. Rahwan, M. Wooldridge, P. McBurney and N. R. Jennings. A distributed algorithm for anytime coalition structure generation. In *AAMAS'10*, pp. 1007-1014, 2010.
- [11] D. Schmeidler. The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics* 17:1163–1170, 1969.
- [12] L. Shapley. A value for  $n$ -person games. In *Contributions to the Theory of Games, vol. 2*, Annals of Mathematics Studies 28 pp. 307–317, 1953.
- [13] L. Shapley. Cores of convex games. *International Journal of Game Theory*, 1:11–26; errata, *ibid.* 1 (1971/72), 199, 1971/72.
- [14] O. Shehory and S. Kraus. Formation of overlapping coalitions for precedence-ordered task-execution among autonomous agents. In *ICMAS-96*, pp. 330–337, 1996.
- [15] T. Shrot, Y. Aumann, S. Kraus. On agent types in coalition formation problems. In *AAMAS'10*, pp. 757–764, 2010.
- [16] E. Winter. The Shapley value. *Handbook of Game Theory with Economic Applications* 3:2025–2054, 2002.
- [17] H. Young. Monotonic solutions of cooperative games. *International Journal of Game Theory* 14(2):65–72, 1985.
- [18] S. Zamir, M. Maschler, and E. Solan. *Game Theory*. Magnes Press, Hebrew University of Jerusalem, Israel, 2008.