Journal of Artificial Intelligence Research 50 (2014) 847-884 Submitted 11/13; published 8/14

Arbitration and Stability in Cooperative Games with Overlapping Coalitions

Yair Zick YAIRZICK@CMU.EDU
School of Computer Science
Carnegie Mellon University, United States

Evangelos Markakis MARKAKIS@GMAIL.COM
Department of Informatics
Athens University of Economics and Business, Greece

Edith Elkind ELKIND@CS.OX.AC.UK
Department of Computer Science
University of Oxford, United Kingdom

Abstract

Overlapping Coalition Formation (OCF) games, introduced by Chalkiadakis, Elkind, Markakis, Polukarov and Jennings in 2010, are cooperative games where players can simultaneously participate in several coalitions. Capturing the notion of stability in OCF games is a difficult task: deviating players may continue to contribute resources to joint projects with non-deviators, and the crucial question is what payoffs the deviators expect to receive from such projects. Chalkiadakis et al. introduce three stability concepts for OCF games—the *conservative core*, the *refined core*, and the *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 stability concepts considered by Chalkiadakis et al. as well as a wide variety of alternative stability concepts. Our approach is based on the notion of *arbitration functions*, which determine the payoff obtained by the deviators, given their deviation and the current allocation of resources. We provide a characterization of stable outcomes under arbitration. We then conduct an in-depth study of four types of arbitration functions, which correspond to four notions of the core; these include the three notions of the core considered by Chalkiadakis et al. Our results complement those of Chalkiadakis et al. and answer questions left open by their work. In particular, we show that OCF games with the conservative arbitration function are essentially equivalent to non-OCF games, by relating the conservative core of an OCF game to the core of a non-overlapping cooperative game, and use this result to obtain a strictly weaker sufficient condition for conservative core non-emptiness than the one given by Chalkiadakis et al.

1. Introduction

Consider a market exchange involving several agents, each owning a certain amount of some goods (say, oil, sugar and flour). The agents trade with each other, and each vendor may offer different prices to different buyers. Suppose that one of the vendors has 3 tons of sugar and has agreed to sell them to buyers 1 and 2 so that buyer 1 receives 1 ton of sugar and pays 500 dollars, whereas buyer 2 receives 2 tons of sugar and pays 900 dollars. The vendor then discovers that he can sell one ton of sugar to a third buyer for 700 dollars. He may decide that he is unhappy with the amount of money he receives from the transaction with buyer 1, and cancel the deal in favor of buyer 3. However, buyer 2, upon hearing that the transaction with buyer 1 has been canceled, may no longer wish to work with this vendor, and cancel her agreement with him as well. Therefore, when deciding whether a deviation from an agreement is profitable, the vendor needs to predict how other parties he is trading with would react to his actions. If it is safe to assume that buyer 2 is willing to upkeep the interaction, then the deviation is profitable; if this is not true, then the seller has to forgo the better deal with buyer 3, or lose more than what the deviation is worth.

This setting possesses several interesting characteristics. First, a vendor may allocate his goods to several buyers, and a buyer may also purchase goods from several vendors. In other words, agents may allocate fractions of their resources to several profit-generating tasks. Second, agents may withdraw some of their resources from some agreements, while keeping others unchanged. For example, an oil vendor may wish to sell less oil to some customer, but not change her interactions with other parties; similarly, a buyer may want to pay less to some vendors, but maintain the same payments to others. Finally, when trying to strategically change an agreement, agents must take into account the impact of their actions on the contracts they still maintain with other (possibly unaffected) parties.

These features are typical of many multi-agent settings, where agents collaborate by allocating parts of their resources to working together in order to generate revenue, and then share the resulting profits. Profit sharing can be done directly, if the exchange of goods itself results in a profit, e.g., if the agents can make a new good using their resources and sell it; profit sharing can also be indirect, e.g., via setting a price for the good sold. In such settings, even if there are no external constraints on how the profits should be distributed, agents should account for individuals or groups of agents who are underpaid. A group of agents that can get more money by deviating from the proposed deal may destabilize the entire agreement, causing a cascade of deviations that results in a less desirable state. However, what constitutes a profitable deviation greatly depends on how the non-deviators respond to the deviating set.

Modeling this system of incentives and reactions is a challenge in itself. Recently this challenge was addressed by Chalkiadakis, Elkind, Markakis, Polukarov, and Jennings (2010), who propose a novel approach to modeling scenarios where agents can divide resources among several coalitions. They introduce *overlapping coalition formation (OCF) games*, which generalize the classic model of *cooperative games* (Peleg & Sudhölter, 2007).

In the classic cooperative game theory model, there is a set of agents N; each subset S of N can form a team and generate profit for its members. In *transferable utility* (TU) games, each subset of agents $S \subseteq N$ is identified with a value $v(S)$. This value can be thought of as the monetary payoff to the set S should its members agree to work together, and can be freely divided among the members of S. As argued above, a desirable property of a payoff distribution scheme is its stability against individual and group deviations, which is captured by the notion of the *core* (Gillies, 1953): a payoff division is said to be in the core if every subset of agents $S \subseteq N$ receives a total payoff that is at least its value $v(S)$.

Classic cooperative game theory assumes that each agent may belong to at most one coalition at any given time. That is, agents form coalition structures by splitting into disjoint groups, with each group working on an individual task. Consequently, the classic approach is not well-suited to handle overlapping coalitions, and the intricacies of deviation in settings where each agent participates in several agreements. Indeed, in the classic cooperative game setting, a set of agents assesses the desirability of a payoff division against what it can make on its own: is the total payoff allotted to S at least $v(S)$? This is because the deviating agents break all ties with the non-deviators, so they cannot expect the non-deviators to collaborate with them in any way. In contrast, when agents are allowed to split their resources among multiple coalitions, the very notion of a profitable deviation must be reexamined, since partial coalitions allow for partial deviations, which means that even though a set of agents decides to deviate from the group agreement, it remains involved in collaborative projects with non-deviators. These non-deviators then decide how much the deviators should receive from such projects. Hence, the deviators' decision whether to withdraw from some tasks (and to what extent) depends on what they expect to receive from their collaboration with non-deviators.

Chalkiadakis et al. (2010) propose a model where agents can split their resources among several coalitions, and study the effect of joint projects with non-deviators on the desirability of deviation in the overlapping coalition setting. They propose and discuss three notions of reaction to deviation. The first reaction, which they term *conservative*, gives the deviators no share of the profits from any coalitions they maintain with non-deviators. The second reaction, termed *refined*, allows the deviators to keep their payoffs from coalitions that were unaffected by the deviation. Finally, under the *optimistic* reaction the deviators may be able to receive some revenue from each coalition they are involved in, as long as the non-deviators in that coalition are guaranteed their pre-deviation payoffs. These reactions to deviation correspond to three notions of stability for OCF games, namely, the *conservative core*, the *refined core*, and the *optimistic core*.

The non-deviators' reaction to deviation can extend beyond these three notions. Many legal contracts (e.g., ones involving a service provider and its clients, or vendors and suppliers, or several companies working on a joint venture) are designed with the explicit aim of detailing the consequences of deviation to all parties. Typically, such contracts impose fines on parties that break the agreement. However, there are also other forms of punishment: for instance, a company that fails to uphold its contractual obligations may be blacklisted by some of its clients, or have its reputation tarnished by "bad press". Alternatively, one may be actively rewarded for deviating, e.g., when an external party wishes to break a certain status-quo, or to encourage project diversification. By adopting the notion of partial deviation and the possible reactions to deviation, we can capture far more nuanced scenarios than those described by classic TU games.

1.1 Our Contribution

Our main contribution is a broad generalization of the overlapping coalition formation (OCF) model proposed by Chalkiadakis et al. (2010). While Chalkiadakis et al. describe three ways in which non-deviators may react to a deviation, we propose a general framework for modeling such reactions. Our framework is based on the notion of *arbitration function* (Section 3). This is a function that, given a deviating group of players and its deviation, outputs, for every coalition containing the deviators, the amount that the deviators can expect to receive from this coalition post-deviation.

Using arbitration functions, we present a class of solution concepts for OCF games that we call *arbitrated cores* (Section 4). We show that the three concepts of the core described by Chalkiadakis et al. (2010) are special cases of our model, and propose a new concept of the core, which we call the *sensitive core*. We propose a criterion for checking whether an outcome is in the arbitrated core, and characterize core outcomes under some important arbitration functions. We then focus on identifying sufficient and necessary conditions for the arbitrated core to be non-empty (Section 5). Building on the work of Chalkiadakis et al., who provide an LP-based criterion for the non-emptiness of the conservative core, we derive criteria for the sensitive and refined core to be non-empty. We use this result to identify an interesting class of OCF games whose refined core is guaranteed to be non-empty. Finally, we show that OCF games with the conservative arbitration function are essentially equivalent to non-OCF games, by relating the conservative core of an OCF game $\mathcal{G}$ to the core of its *discrete superadditive cover*, which is a non-OCF game that is constructed from $\mathcal{G}$ in a natural way. In particular, this means that the conservative core of an OCF game $\mathcal{G}$ is non-empty when the discrete superadditive cover of $\mathcal{G}$ is supermodular. We demonstrate that this condition is strictly weaker than the convexity-based condition for conservative core non-emptiness given by Chalkiadakis et al.

1.2 Related Work

The most direct precursor to our work is the work of Chalkiadakis et al. (2010): we generalize their model to capture a broader class of possible reactions to deviation, and solve some of the problem left open by their work.

Incentives and optimization in collaborative multi-agent environments have received plenty of attention from the multiagent research community (Sims, Corkill, & Lesser, 2008; Airiau & Sen, 2009; Rahwan & Jennings, 2008; Dang, Dash, Rogers, & Jennings, 2006; Shehory & Kraus, 1996; Lin & Hu, 2007; Zhang, Jiang, Su, Qi, & Fang, 2010); see also the PhD thesis of Rahwan (2007) and the recent book of Chalkiadakis, Elkind, and Wooldridge (2011) for an overview. These studies often use cooperative game theory as the modeling framework.

Several authors have considered the problem of optimal coalition structure generation in cooperative settings with overlapping coalitions. Shehory and Kraus (1996) have initiated this line of research; in their work, they describe a distributed coalition formation algorithm where agents may split their resources among several tasks in order to achieve better outcomes. Dang et al. (2006) consider overlapping coalitions in sensor networks, where the agents are sensors that are tasked with tracking objects. Lin and Hu (2007) provide a parallel algorithm for overlapping coalition formation. In all these papers, the underlying assumption is that agents are fully cooperative, i.e., they seek to maximize the social welfare rather than their own gains. In contrast, in our work we focus on game-theoretic aspects of strategic behavior in cooperative settings, taking agent incentives into account.

Our model applies to settings where rational agents distribute their resources among several projects, receiving profits from each. This scenario occurs in a variety of applications. For example, multi-commodity network flows can be easily modeled within the overlapping coalition formation framework (for the analysis of the underlying non-OCF game, see Markakis & Saberi, 2005). This is also the case for several other fractional optimization problems (Deng, Ibaraki, & Nagamochi, 1999). For example, a fractional matching problem can be thought of as an OCF game, where each partial coalition between a pair of agents describes the amount of resources that each of these agents assigns to the respective task.

Another class of settings where overlapping coalitions arise are collaboration networks. Consider a social network where each agent is associated with a weighted node and derives some value from assigning a certain portion of its weight to collaborating with its neighbors. Such settings have received much attention in recent years (for a survey, see Jackson, 2003). For instance, Anshelevich and Hoefer (2010) discuss strategic aspects of network formation in settings where agents may participate in more than one coalition; however, in contrast with our work, the main focus of their analysis is individual rationality and pairwise equilibria in such networks, rather than group stability. Fractional stable matchings were discussed by Aharoni and Fleiner (2003), in a non-transferable utility model. More recently, Ackerman and Brânzei (2014) have analyzed research collaboration using a graphical model, where authors devote some of their efforts to collaborative projects, and receive some credit for their work based on the authorship order model. Ackerman and Brânzei are mainly interested in individually rational and pairwise stable outcomes, much like Anshelevich and Hoefer. The OCF framework has recently been used to model collaborative wireless networks (Wang, Song, Han, & Saad, 2013; Zhang, Song, Han, Saad, & Lu, 2013).

Overlapping coalition formation games, as proposed by Chalkiadakis et al. (2010), share some features with *fuzzy games* (Aubin, 1981). However, there are fundamental differences between the two models. Arguably, the most important difference is that Chalkiadakis et al. assume that several partial coalitions may form, whereas Aubin assumes that all agents pool their resources to work on a single task, and only considers partial coalitions in the context of potential deviations. Assuming the formation of a single coalition eliminates many of the subtleties that games with overlapping coalitions can capture, as was emphasized by Chalkiadakis et al.

While the work presented in this paper is not computational in nature, algorithmic issues that arise in cooperative games with overlapping coalitions have been discussed by Chalkiadakis et al. (2010) and subsequently by Zick, Chalkiadakis, and Elkind (2012); we refer the reader to the companion paper (Zick, Chalkiadakis, Elkind, & Markakis, 2014) for in-depth analysis.

2. Preliminaries

Throughout the paper, we consider games with a finite set of players $N = \{1, \dots, n\}$. We use boldface lowercase letters to denote vectors and uppercase letters to denote sets. Given a set $S \subseteq N$, we write $\mathbf{e}^S$ to denote the indicator vector of S. Furthermore, given a vector $\mathbf{x} = (x^1, \dots, x^n) \in \mathbb{R}^n$ and a subset $S \subseteq N$, we set $x(S) = \sum_{i \in S} x^i$.

2.1 Classic TU Cooperative Games

A *transferable utility (TU) cooperative game* is given by a set of *players*, or *agents*, $N = \{1, \dots, n\}$ and a *characteristic function* $u : 2^N \to \mathbb{R}$ with $u(\emptyset) = 0$; we write $\mathcal{G} = \langle N, u \rangle$. A *payoff vector* for a game $\mathcal{G} = \langle N, u \rangle$ is a vector $\mathbf{x} \in \mathbb{R}^n$ such that $x(N) = u(N)$. A payoff vector $\mathbf{x}$ for a game $\mathcal{G} = \langle N, u \rangle$ is said to be *stable* if $x(S) \geq u(S)$ for each $S \subseteq N$. The set of all stable payoff vectors for $\mathcal{G}$ is called the *core of* $\mathcal{G}$.

It is sometimes assumed that agents are allowed to form *coalition structures*; a coalition structure CS is simply a partition of N, and its value $u(CS)$ is the sum of the values of its constituent coalitions. If a coalition S is a part of some coalition structure CS, then the value of S, $u(S)$, can be freely divided among all members of S, but transfers to agents that are not members of S are not allowed, i.e., $\mathbf{x}$ is a payoff vector for a coalition structure CS if and only if $x(S) = u(S)$ for each $S \in CS$. The *core of a game* $\mathcal{G} = \langle N, u \rangle$ *with coalition structures* is the set of all pairs $(CS, \mathbf{x})$, where CS is a coalition structure, $\mathbf{x}$ is a payoff vector for CS, and $x(S) \geq u(S)$ for each $S \subseteq N$. Cooperative games with coalition structures have been first studied by Aumann and Drèze (1974), and have been the focus of a number of recent papers by the multiagent research community (see Chalkiadakis et al., 2011).

2.2 OCF Games

Given a set of agents $N = \{1 \dots n\}$, a *partial coalition* of players in N is a vector $\mathbf{c} = (c^1, \dots, c^n)$ in $[0, 1]^n$. The i-th coordinate of $\mathbf{c}$ indicates the fraction of player i's resources contributed to $\mathbf{c}$.

In what follows, we will omit the word "partial", and refer to vectors in $[0,1]^n$ as *coalitions*. Of particular interest are the players who actively participate in $\mathbf{c}$. These players may rightfully claim a share of $\mathbf{c}$'s profits, and they are the ones who may potentially get hurt if some agent changes its contribution to $\mathbf{c}$. Formally, the *support of the coalition* $\mathbf{c}$ is the set $supp(\mathbf{c}) = \{i \in N \mid c^i > 0\}$. We now recall the formal definition of OCF games, given by Chalkiadakis et al. (2010).

Definition 2.1. *An* OCF game $\mathcal{G} = \langle N, v\rangle$ *is given by a set of players* $N = \{1,\dots,n\}$ *and a* characteristic function $v : [0,1]^n \to \mathbb{R}_+$*, which assigns a non-negative real value to each partial coalition; we require that* $v(0^n) = 0$*.*

The characteristic function of an OCF game can be quite general; we do not even require v to be monotone. The reasoning behind this apparent leniency is that, given v, the agents in N may form several coalitions in order to optimize their revenue, a process which results in a *coalition structure*. There is, however, one additional assumption on v that we would like to impose; to introduce it, we first need to formally define the concept of a coalition structure in an OCF game.

A *coalition structure over* N is a finite list of coalitions $CS = (\mathbf{c}_1,\dots,\mathbf{c}_m)$ such that $\sum_{j=1}^m \mathbf{c}_j \le 1^n$. The coalition structure CS can be thought of as an $n \times m$ matrix whose rows sum to at most 1. The requirement that the rows sum to no more than 1 means that CS is a valid division of players' resources, i.e., each player gives at most 100% of its resources to CS. For readability, we use set notation when referring to coalition structures: e.g., when referring to a coalition $\mathbf{c}$ that appears in CS, we write $\mathbf{c} \in CS$, and when referring to a coalition structure CS' that is a sublist of CS, we write $CS' \subseteq CS$.

Recall that a coalition structure in a classic cooperative game is a partition of the agent set, i.e., a collection of disjoint subsets of N whose union is N. In particular, no set of agents can appear twice in the same coalition structure. In contrast, in OCF games, it is possible for a coalition structure CS to have the same coalition appearing more than once. This simply means that the respective agents are completing two separate tasks, which require the same resources and generate the same revenue.

Example 2.2. Consider a setting where two researchers collaborate. They can write a single long research paper, which will require all of their time, and will generate a revenue of \$100000 (in, say, grant money or performance payment). Alternatively, if they split their time equally between two research projects, they will produce two shorter papers, which generate \$70000 each. In this setting, it is in the agents' best interest to split into two identical coalitions of the form $\left(\begin{smallmatrix} .5 \\ .5 \end{smallmatrix}\right)$ rather than form a single coalition.

The total revenue generated by the coalition structure CS is simply $\sum_{\mathbf{c}\in CS} v(\mathbf{c})$, and is denoted by $v(CS)$.

Given a set of agents $S \subseteq N$, we say that a coalition structure CS is *over* S if $supp(\mathbf{c}) \subseteq S$ for all $\mathbf{c} \in CS$. We denote the set of all coalition structures over S by $\mathcal{CS}(S)$. The *weight vector of a coalition structure* CS is $\mathbf{w}(CS) = \sum_{\mathbf{c}\in CS} \mathbf{c}$. The vector $\mathbf{w}(CS)$ indicates the total amount of resources that each player in N invests in the coalition structure CS. Note that $\mathbf{w}(CS) \in [0,1]^n$, and if $CS \in \mathcal{CS}(S)$, then $\mathbf{w}(CS) \le \mathbf{e}^S$; we say that a coalition structure CS over S is *efficient* if $\mathbf{w}(CS) = \mathbf{e}^S$. We also write $\mathbf{w}^S(CS)$ to denote the total weight of S in CS; namely, the i-th coordinate of $\mathbf{w}^S(CS)$ equals the i-th coordinate of $\mathbf{w}(CS)$ if i is in S, and is 0 otherwise.

As shown in Example 2.2, agents may form coalition structures in order to increase their revenue. Such considerations naturally give rise to the following definition. The *superadditive cover* *of* v is defined as

$$v^*(\mathbf{c}) = \sup\{v(CS) \mid CS \in \mathcal{CS}(N), \mathbf{w}(CS) \leq \mathbf{c}\}.$$

The value $v^*(\mathbf{c})$ is the most that the players can make using the resources of $\mathbf{c}$. Note that a similar notion of a superadditive cover exists for classic coalitional games (Aumann & Drèze, 1974): given a function $u : 2^N \to \mathbb{R}$, its superadditive cover is the function $u^* : 2^N \to \mathbb{R}$ defined as

$$u^*(S) = \max\{\sum_{T \in \mathcal{P}} v(T) \mid \mathcal{P} \text{ is a partition of } S\}.$$

Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is called *superadditive* if for all $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ we have

$$f(\mathbf{c}) + f(\mathbf{d}) \leq f(\mathbf{c} + \mathbf{d}); \tag{1}$$

if f is defined on a subdomain U of $\mathbb{R}^n$, condition (1) is only imposed on vectors $\mathbf{c}, \mathbf{d}$ such that $\mathbf{c} + \mathbf{d}$ is still in U. This definition has the following discrete analogue: a function $u : 2^N \to \mathbb{R}$ is called *superadditive* if for all sets $S, T \subseteq N$ such that $S \cap T = \emptyset$ we have

$$u(S) + u(T) \leq u(S \cup T). \tag{2}$$

These definitions explain the term "superadditive cover": it is easy to see that v^* is the minimal superadditive function from $[0,1]^n$ to $\mathbb{R}$ such that $v^*(\mathbf{c}) \geq v(\mathbf{c})$ for every $\mathbf{c} \in [0,1]^n$, and, similarly, u^* is the minimal superadditive function from 2^N to $\mathbb{R}$ such that $u^*(S) \geq u(S)$ for every $S \subseteq N$.

We are now ready to formulate the additional requirement that we wish to impose on v.

Definition 2.3. *A characteristic function v has the* efficient coalition structure property *if for every $\mathbf{c} \in [0,1]^n$ there exists a $CS \in \mathcal{CS}(N)$ such that $\mathbf{w}(CS) = \mathbf{c}$ and $v^*(\mathbf{c}) = v(CS)$.*

For characteristic functions that have the efficient coalition structure property, the sup in the definition of v^* can be replaced with max. In what follows, we limit our attention to OCF games whose characteristic functions enjoy this property. To see why the efficient coalition structure property is desirable in the OCF setting, consider the function f given by $f(\mathbf{x}) = \alpha$ for all $\mathbf{x} \in [0,1]^n \setminus \{0^n\}$, where α is some positive constant. This function does not have the efficient coalition structure property and its superadditive cover is $f^*(\mathbf{x}) = \infty$ for all $\mathbf{x} \in [0,1]^n \setminus \{0^n\}$. In particular, this means that no coalition structure is optimal for f: it is always possible to increase the social welfare by splitting some coalition into two non-zero coalitions. We remark that Chalkiadakis et al. (2010) impose a number of conditions on the characteristic function of an OCF game that, taken together, imply the efficient coalition structure property; we find it more convenient to state this property directly.

To provide some intuition for the concepts introduced so far, we now describe a simple class of OCF games. It is often useful to think of the agents as using their resources to complete a given set of tasks. Chalkiadakis et al. (2010) describe a class of OCF games, which they call *Threshold Task Games (TTGs)*, that are based on this idea. A TTG is specified by a finite list of tasks, $T = \{t_1, ..., t_k\}$, where each task t_ℓ requires some weight $\omega(t_\ell) \geq 0$ for its completion and gives a certain payoff $p(t_\ell) \geq 0$, and a list of players' weights, denoted $(w^1, \ldots, w^n)$. Each player can allocate any fraction of its weight to the completion of any task. The worth of a coalition is

$$v(\mathbf{c}) = \max\{p(t_\ell) \mid \omega(t_\ell) \leq \sum_{i=1}^{n} c^i w^i\};$$

that is, the value of a coalition is simply the value of the highest-paying task that the agents can complete using their combined weight. Threshold Task Games are a natural extension of *Weighted Voting Games* (WVGs) (Chalkiadakis et al., 2011). Indeed, in a weighted voting game, each agent $i \in N$ has a non-negative weight w^i, and the value of a set of players $S \subseteq N$ is 1 if the combined weight of the members of S is at least some given non-negative quota q, and 0 otherwise. Thus, weighted voting games are simply threshold task games where there is a single task with weight q and payoff 1, and overlapping coalitions are not allowed.

2.3 Payoff Division

Having split into coalitions and generated revenue, the agents have to decide how to divide this revenue amongst themselves in an agreeable manner. We will now formally describe payoff division in OCF games.

An *imputation* for a coalition structure $CS = (\mathbf{c}_1, ..., \mathbf{c}_m) \in \mathcal{CS}(N)$ is a list of vectors $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_m)$; for each $j = 1, \dots, m$, $\mathbf{x}_j$ is a vector in $\mathbb{R}^n$ that describes how much each agent in N receives from $\mathbf{c}_j$. Given a coalition $\mathbf{c} \in CS$ and an imputation $\mathbf{x}$ for CS, we refer to the way the value of $\mathbf{c}$ is divided as $\mathbf{x}(\mathbf{c})$; thus, the payoff to agent i from coalition $\mathbf{c} \in CS$ under $\mathbf{x}$ is $x(\mathbf{c})^i$. In order for $\mathbf{x}$ to be a valid division of payoffs among the agents, it should satisfy the following conditions.

Individual Rationality: $\sum_{\mathbf{c} \in CS} x(\mathbf{c})^i \geq v^*(\mathbf{e}^{\{i\}})$ for all $i \in N$.

Payoff Distribution: for all $\mathbf{c} \in CS$ it holds that $\sum_{i \in N} x(\mathbf{c})^i \leq v(\mathbf{c})$, and if $c^i = 0$ then $x(\mathbf{c})^i = 0$.

We denote the set of all imputations for CS by $I(CS)$. Observe that our definition does not allow inter-coalitional transfers. That is, if an agent $i \in N$ is not in the support of $\mathbf{c}$, then i cannot expect to receive any payoff from $\mathbf{c}$. We call a tuple $(CS, \mathbf{x})$, where CS is a coalition structure and $\mathbf{x} \in I(CS)$, a *feasible outcome*. Given a set $S \subseteq N$, we denote by $\mathcal{F}(S)$ the set of all feasible outcomes $(CS, \mathbf{x})$ with $CS \in \mathcal{CS}(S)$. Thus, $\mathcal{F}(N)$ refers to all possible ways that agents can form coalition structures and divide payoffs.

Given a feasible outcome $(CS, \mathbf{x}) \in \mathcal{F}(N)$, we define the *payoff* to an agent $i \in N$ as $p^i(CS, \mathbf{x}) = \sum_{\mathbf{c} \in CS} x(\mathbf{c})^i$. This is the total payoff of i from all coalitions in CS. Observe that $p^i(CS, \mathbf{x})$ is uniquely defined: for any given outcome $(CS, \mathbf{x})$, the total payoff to player i depends solely on the amount of payoff received from the coalitions in CS, which is determined by $\mathbf{x}$, and $\mathbf{x}$ alone. Similarly, the total payoff to a set S is given by $p^S(CS, \mathbf{x}) = \sum_{i \in S} p^i(CS, \mathbf{x})$.

Definition 2.4. *Given a set $S \subseteq N$ and a coalition structure $CS \in CS(N)$, the* coalition structure CS reduced to S *is defined as*

$$CS|_S = (\mathbf{c} \in CS \mid supp(\mathbf{c}) \subseteq S) .$$

These are all coalitions that are comprised solely of members of S.

The coalitions in $CS|_S$ are fully controlled by S. Hence, a deviation by the members of S would only affect non-deviators if S changes its contribution to coalitions outside of $CS|_S$.

We are now ready to formally define deviations in OCF games. Loosely speaking, a deviation by a set $S \subseteq N$ from a coalition structure $CS \in \mathcal{CS}(N)$ specifies the amount of resources that S will withdraw from coalitions not in $CS|_S$.

Definition 2.5. *Suppose that* $CS \setminus CS|_S = (\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_k})$. *A coalition structure* $CS' = (\mathbf{d}_1, \ldots, \mathbf{d}_k)$ *is a* deviation of S from CS *if for all* $\ell = 1, \ldots, k$ *it holds that* $\mathbf{d}_\ell \le \mathbf{c}_{j_\ell}$ *and* $\mathbf{d}_\ell \le \mathbf{e}^S$.

The deviation CS' describes how a set S withdraws resources from coalitions in $CS \setminus CS|_S$. The requirement that $\mathbf{d}_\ell \le \mathbf{c}_{j_\ell}$ captures the fact that S cannot withdraw more resources from the coalition $\mathbf{c}_{j_\ell}$ than it has invested in it. We also require that $\mathbf{d}_\ell \le \mathbf{e}^S$ since the deviation CS' can involve members of S only.

It will sometimes be convenient to think of a deviation as a function that, for each coalition $\mathbf{c}$ in $CS \setminus CS|_S$, outputs how much agents in S withdraw from $\mathbf{c}$. That is, given a coalition structure $CS \in \mathcal{CS}(N)$, a set $S \subseteq N$, a deviation CS' of S from CS, and a coalition $\mathbf{c}_{j_\ell} \in CS \setminus CS|_S$, we write $\mathbf{d}_{CS'}(\mathbf{c}_{j_\ell})$ to refer to the coalition d_ℓ in CS' that corresponds to the deviation of S from c_{j_ℓ}. Under this notation, $(d_{CS'}(\mathbf{c}_{j_\ell}))^i$ specifies the amount of resources that agent i withdraws from the coalition $\mathbf{c}_{j_\ell}$ under CS'. When the deviation CS' is understood from the context, we omit the subscript CS' from $\mathbf{d}_{CS'}(\mathbf{c})$ and refer to the deviation of S from $\mathbf{c}$ as $\mathbf{d}(\mathbf{c})$.

Note that if S withdraws all of its resources from $CS \setminus CS|_S$, then the total weight available to S is $\mathbf{w}(CS|_S) + \mathbf{w}^S(CS \setminus CS|_S) = \mathbf{w}^S(CS)$.

Example 2.6. Consider a three-player game described as follows. There are several types of tasks:

- t_{12} requires 50% of player 1's resources and all of player 2's resources, and its payoff is 16;
- t'_{12} requires 40% of player 1's resources and all of player 2's resources, and its payoff is 10;
- t_{13} requires 50% of player 1's resources and all of player 3's resources, and its payoff is 20;
- t_1 requires 10% of player 1's resources and its payoff is 3.

The optimal coalition structure in this case is

$$CS = \left(\mathbf{c}_1 = \begin{pmatrix} 0.5 \\ 1 \\ 0 \end{pmatrix}, \mathbf{c}_2 = \begin{pmatrix} 0.5 \\ 0 \\ 1 \end{pmatrix}\right).$$

The coalitions $\mathbf{c}_1$ and $\mathbf{c}_2$ complete tasks t_{12} and t_{13}, respectively; thus, $v(\mathbf{c}_1) = 16$ and $v(\mathbf{c}_2) = 20$.

A deviation of player 1 from CS would be a coalition structure

$$CS_1 = \left(\begin{pmatrix} \beta \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \gamma \\ 0 \\ 0 \end{pmatrix}\right)$$

such that $\beta \le 0.5$ and $\gamma \le 0.5$. If $\beta = 0$, then player 2 is unaffected by the deviation; if $\gamma = 0$ then player 3 is unaffected by the deviation.

It is possible that player 1 deviates from $\mathbf{c}_1$ while not changing its contribution to $\mathbf{c}_2$, or vice versa. Moreover, it is possible that a coalition from which resources were withdrawn can still generate some revenue. For instance, if player 1 withdraws 10% of its resources from $\mathbf{c}_1$, what remains of $\mathbf{c}_1$ still generates a revenue of 10. This opens opportunities for more nuanced post-deviation behavior: player 3 may decide not to break its agreement with player 1 if $\gamma = 0$, and if $\beta = .1$ and player 2's payment in the original coalition structure was less than 10, player 1 can hope to get some payment from its reduced collaboration with player 2.

As Example 2.6 illustrates, deviations in OCF games may leave some of the non-deviators unaffected by the deviation. Moreover, it may be possible for deviators to ensure that the non-deviators still receive the payoff they were allocated prior to the deviation. If deviators agree to assume the marginal loss of their deviation, it is possible that the non-deviators will agree to maintain coalitions with them. These notions are formalized in Section 3.

3. The Arbitration Function

In classic cooperative game theory, a set S that is unhappy with its payoffs and wants to deviate compares its current payoff with what it can earn on its own. However, as Example 2.6 shows, the complex structure of deviations in OCF games can lead to a variety of ways in which agents may react to a deviation. If a deviating set still keeps some resources invested in coalitions with non-deviators, it may be the case that it will continue to receive payments from these coalitions.

To discuss stability in OCF games, we need to describe how the agents react if some set $S \subseteq N$ deviates from $(CS, \mathbf{x})$. One can think of this process in the following manner: given an outcome $(CS, \mathbf{x})$ and a deviation CS' of $S \subseteq N$ from CS, the set S may use all available resources in $CS|_S$ and CS' in order to generate revenue for itself. Moreover, each coalition $\mathbf{c}$ in $CS \setminus CS|_S$ needs to decide how much payoff it gives to the deviators; this behavior is specified by the *arbitration function*.

Definition 3.1. *Given an outcome $(CS, \mathbf{x})$, a set $S \subseteq N$ and a deviation CS' of S from CS, the* arbitration function $\mathcal{A}$ *is a mapping that assigns a real value $\alpha_{\mathbf{c}}(CS, \mathbf{x}, S, CS')$ to each $\mathbf{c} \in CS \setminus CS|_S$.*

The value $\alpha_{\mathbf{c}}(CS, \mathbf{x}, S, CS')$ represents the amount that the coalition $\mathbf{c}$ will pay to S, given the current outcome $(CS, \mathbf{x})$, the identity of the deviators, and the nature of their deviation.

3.1 Properties of Arbitration Functions

Clearly, we need to impose some restrictions on the amount that the deviators may expect to get from the non-deviators. For example, it would be unreasonable for a deviating set to expect a payoff that exceeds the post-deviation value of the respective coalition. Another natural restriction is that, for deviators to get any payoff from a given coalition, every non-deviator in that coalition should receive the same payoff as before the deviation, i.e., the non-deviators will never agree to pay the deviating set S at the expense of the non-deviating members.

Formally, given an outcome $(CS, \mathbf{x})$, a set $S \subseteq N$, a deviation CS' of S from CS, and an arbitration function $\mathcal{A}$ whose output on $(CS, \mathbf{x}, S, CS')$ is described by the list of values $(\alpha_{\mathbf{c}})_{\mathbf{c} \in CS \setminus CS|_S}$, we require $\mathcal{A}$ to have the following properties:

Accountability: $\alpha_{\mathbf{c}}(CS, \mathbf{x}, S, CS') \leq \max\{v(\mathbf{c} - \mathbf{d}(\mathbf{c})) - \sum_{i \in N \setminus S} x(\mathbf{c})^i, 0\}$.

Deviation-Monotonicity: For every pair of subsets $S, T \subseteq N$ such that $S \subseteq T$ and their respective deviations CS' and CS'' from CS such that $\mathbf{d}_{CS'}(\mathbf{c}) \leq \mathbf{d}_{CS''}(\mathbf{c})$ for each $\mathbf{c} \in CS \setminus CS|_S$, it holds that $\alpha_{\mathbf{c}}(CS, \mathbf{x}, S, CS') \geq \alpha_{\mathbf{c}}(CS, \mathbf{x}, T, CS'')$.

The first condition states a general upper bound on the amount that a deviating set can expect to receive from a coalition: after the deviation, the coalition $\mathbf{c}$ can generate a profit of $v(\mathbf{c} - \mathbf{d}(\mathbf{c}))$, so the most that a deviating set can expect to receive is $v(\mathbf{c} - \mathbf{d}(\mathbf{c}))$, minus the original payments given to non-deviators under $(CS, \mathbf{x})$. The second condition simply states that the punishment imposed by an arbitration function should be monotone in the size of the violation: if the deviators withdraw a smaller amount of resources from each coalition, they should receive at least as much payoff as they received under their original deviation. The rationale behind the first condition is strategic in nature; it stems from the assumption that a set of agents engaged in a task will not agree to pay deviating members if the deviators cannot ensure that each non-deviator is paid the same amount as prior to the deviation. We do stress however, that accountability is not a necessary component in our proofs: our results still hold even if one does not assume accountability. The second condition, while sensible as well, is instrumental in proving Theorem 3.4, and is a necessary assumption. Finally, note that it is possible that a coalition imposes a fine on deviating members, that is, the value of $\alpha_{\mathbf{c}}$ need not be positive.

The total revenue of a deviating set S given its deviation CS' from CS can be written as

$$\mathcal{A}(CS, \mathbf{x}, S, CS') = v^*(\mathbf{w}(CS|_S) + \mathbf{w}(CS')) + \sum_{\mathbf{c}\in CS\setminus CS|_S} \alpha_{\mathbf{c}}(CS, \mathbf{x}, S, CS).$$

Given an outcome $(CS, \mathbf{x})$ and a set $S \subseteq N$, the most that S can get by deviating from $(CS, \mathbf{x})$ is

$$\mathcal{A}^*(CS, \mathbf{x}, S) = \sup\{\mathcal{A}(CS, \mathbf{x}, S, CS') \mid CS' \text{ is a deviation of } S \text{ from } CS\}.$$

Now, in order for S to deem a deviation profitable, each of its members should stand to gain from it. The formal definition that captures this idea is somewhat complicated, as it has to describe how the deviators share the profit/loss allocated to them by the arbitration function.

Definition 3.2. *Given an outcome $(CS, \mathbf{x})$ and a deviation CS' of $S \subseteq N$ from CS, we say that CS' is an $\mathcal{A}$*-profitable deviation *if we can construct*

1. *a coalition structure CS_d such that $\mathbf{w}(CS_d) = \mathbf{w}(CS|_S) + \mathbf{w}(CS')$ together with an imputation $\mathbf{x}_d \in I(CS_d)$, and*
2. *a collection of vectors $\mathbf{y} = (\mathbf{y_c})_{\mathbf{c}\in CS\setminus CS|_S}$, such that for every $\mathbf{c} \in CS \setminus CS|_S$ it holds that*
   - *(a) $\mathbf{y_c} \in (\mathbb{R}_+)^n$ if $\alpha_{\mathbf{c}}(CS, \mathbf{x}, S, CS') \geq 0$ and $\mathbf{y_c} \in (\mathbb{R}_-)^n$ otherwise,*
   - *(b) $(y_{\mathbf{c}})^i = 0$ for $i \notin S \cap supp(\mathbf{c} - \mathbf{d}(\mathbf{c}))$, and*
   - *(c) $\sum_{i=1}^n (\mathbf{y_c})^i = \alpha_{\mathbf{c}}(CS, \mathbf{x}, S, CS')$,*

so that for every $i \in S$ it holds that $p^i(CS_d, \mathbf{x}_d) + \sum_{\mathbf{c}\in CS\setminus CS|_S}(\mathbf{y_c})^i > p^i(CS, \mathbf{x})$.

Simply put, given S's deviation from $(CS, \mathbf{x})$ and the arbitration function $\mathcal{A}$, the agents in S have to agree on three things. First, they have to decide on a coalition structure to form with their post-deviation resources. Second, they have to come up with a way of dividing profits from that coalition structure. Third, they have to agree on a way of dividing the payoffs (or fines) assigned to them by the arbitration function. For the deviation to be profitable, these agreements should be such that all agents in S receive strictly more than what they had received under $(CS, \mathbf{x})$. If no such agreement exists, S cannot $\mathcal{A}$-profitably deviate. Given $(CS_d, \mathbf{x}_d)$ and $\mathbf{y}$, the total payoff of an agent $i \in S$ from deviating is denoted $q^i(\mathbf{x}_d, \mathbf{y})$.

A priori, it seems possible that, even when the total payoff to a deviating set is more than $p^S(CS, \mathbf{x})$, the members of S may be unable to divide the revenue from the deviation in a way that benefits them all due to coalitional restrictions. Chalkiadakis et al. (2010) show that for the conservative arbitration function (see Section 3.2) this is not the case: if there is a way for agents to jointly gain more by deviating, then they can divide the profits in a way that benefits them all. In Theorem 3.4, we will show how to extend this result to general arbitration functions.

Remark 3.3. It is possible to incorporate a *freely divisible value* into the arbitration function. This value is a payoff (or fine) assigned to the deviators by an external entity. Just as the payments from the coalitions in $CS \setminus CS|_S$, this payment may depend on the original outcome, the identity of the deviating set, and the nature of its deviation. All results in this paper carry through with the addition of a freely divisible value. Intuitively, this value allows us to extend notions such as the ε-core (Maschler, Peleg, & Shapley, 1979) to OCF games.

3.2 Some Arbitration Functions

We now present some arbitration functions and briefly discuss their properties. Three of these arbitration functions correspond to the three notions of the core introduced by Chalkiadakis et al. (2010) (the conservative core, the refined core and the optimistic core), and the names we use for these arbitration functions reflect this; in contrast, the sensitive arbitration function is based on a new idea.

3.2.1 The Conservative Arbitration Function

The simplest assumption that one can make with respect to reaction to deviation is that the non-deviators will react by voiding all agreements with the deviators; that is, the deviators do not expect any payment from any of their coalitions with the non-deviators, regardless of their contribution. This reasoning gives rise to the *conservative arbitration function*; formally, it is defined by setting $\alpha_{\mathbf{c}} \equiv 0$ for any given input. Under the conservative arbitration function, a deviating set is not rewarded in any way for continued interaction with the non-deviators; thus, the profit it can hope to gain by deviating is exactly what it can make on its own. Therefore, under the conservative arbitration function, denoted by $\mathcal{A}_c$, we have $\mathcal{A}_c^*(CS, \mathbf{x}, S) = v^*(\mathbf{e}^S)$ for every outcome $(CS, \mathbf{x})$.

3.2.2 The Sensitive Arbitration Function

When reasoning about agents' reaction to deviation, a natural assumption to make is that the non-deviators do not care what the deviating set S does, as long as the deviation does not affect them. This approach is captured by the *sensitive arbitration function*. Under this function the payment that the deviators obtain from a coalition $\mathbf{c} \in CS \setminus CS|_S$ is 0 if there is some $\mathbf{c}' \in CS \setminus CS|_S$ such that $\mathbf{d}(\mathbf{c}') \neq 0$ and $supp(\mathbf{c}) \cap supp(\mathbf{c}') \cap N \setminus S \neq \emptyset$; otherwise, the payment to S is the total payoff that the members of S were receiving from $\mathbf{c}$ prior to the deviation.

Under the sensitive arbitration function, denoted by $\mathcal{A}_s$, the most that a deviating set S can get can be computed as follows. First, it is clear that S does not benefit from investing resources into joint projects with non-deviating agents that will be hurt in any way by the deviation. Thus, S only needs to decide which agents it will keep collaborating with; as for all other agents, it can break all agreements with them. In order to describe payments under the sensitive arbitration function, we will employ the following notation. Given a set $T \subseteq N$, we let CS_T be the set of all coalitions in CS that involve members of T; that is

$$CS_T = (\mathbf{c} \in CS \mid supp(\mathbf{c}) \cap T \neq \emptyset).$$

Now, if a set $S \subseteq N$ chooses to break agreements that involve members of T, it forgoes all payments from coalitions in CS_T, but can use the resources it has invested in CS_T in order to maximize its own profits. That is, we obtain

$$\mathcal{A}_s^*(CS, \mathbf{x}, S) = \max\left\{v^*\left(\mathbf{w}(CS|_S) + \mathbf{w}^S(CS_T)\right) + p^S\left(CS \setminus (CS|_S \cup CS_T), \mathbf{x}\right) \mid T \subseteq N \setminus S\right\}.$$

This definition can be extended as follows: instead of being completely self-interested, each agent $i \in N$ has a set of players $S_i \subseteq N$ that it cares about. If any of the agents in S_i are hurt by a deviation, agent i will not allow the deviating set to have any payoff from the coalitions that it is involved in.

3.2.3 THE REFINED ARBITRATION FUNCTION

The sensitive arbitration function focuses on the impact of the deviation on individual agents; in contrast, the *refined arbitration function* makes its decisions based on the deviation's impact on each coalition. Specifically, under this arbitration function a coalition $\mathbf{c}$ allows the deviators to keep their payoffs from $\mathbf{c}$ if and only if S has not withdrawn any resources from $\mathbf{c}$. Formally, given a deviation CS' of $S \subseteq N$ from CS in the outcome $(CS, \mathbf{x})$, the refined arbitration function, denoted $\mathcal{A}_r$, will let S keep its payoff from a coalition $\mathbf{c} \in CS \setminus CS|_S$ if $\mathbf{d}(\mathbf{c}) = 0$, and allocate S no payoff from $\mathbf{c}$ otherwise. Note that the refined arbitration function is more generous to the deviators than the sensitive arbitration function: if a deviator i and a non-deviator j are both involved in coalitions $\mathbf{c}$ and $\mathbf{c}'$, and $\mathbf{c}'$ is affected by the deviation while $\mathbf{c}$ is not, under the refined arbitration function i will receive its share of $\mathbf{c}$'s payoffs, whereas under the sensitive arbitration function it will not.

The most that a set S can expect to gain by deviating under the refined arbitration function can be computed as follows. First, note that given a coalition $\mathbf{c}$ in $CS \setminus CS|_S$, S should either withdraw all resources from $\mathbf{c}$ or none at all. Thus, S needs to find the best set of coalitions in CS from which to fully withdraw resources. Formally, we have

$$\mathcal{A}_r^*(CS, \mathbf{x}, S) = \max\{v^*(\mathbf{w}^S(\widehat{CS})) + p^S(CS \setminus \widehat{CS}, \mathbf{x}) \mid CS|_S \subseteq \widehat{CS} \subseteq CS\}.$$

3.2.4 THE OPTIMISTIC ARBITRATION FUNCTION

A yet more lenient reaction to deviation is captured by the *optimistic arbitration function*, denoted $\mathcal{A}_o$. Under this arbitration function a coalition is willing to pay the deviating set as long as the non-deviators receive their pre-deviation payoffs. Since we require arbitration functions to have the accountability property, this means that the optimistic arbitration function gives a deviating set the highest possible payoff it can receive under any arbitration function. Formally, given a set S, an outcome $(CS, \mathbf{x})$ and a deviation CS' of S from CS, the payoff that S gets from a coalition $\mathbf{c}$ is $\max\{0, v(\mathbf{c} - \mathbf{d}(\mathbf{c})) - \sum_{i \notin S} x(\mathbf{c})^i\}$.

In order to compute the most that a set S can get by deviating under the optimistic arbitration function, we need to determine the amount of resources that S is going to withdraw from every coalition it is involved in. We obtain

$$\mathcal{A}_o^*(CS, \mathbf{x}, S) = \sup\{v^*(\mathbf{w}^S(\widehat{CS}) + \mathbf{w}(CS')) + \sum_{\mathbf{c} \in CS \setminus \widehat{CS}} v(\mathbf{c} - \mathbf{d}(\mathbf{c})) - p^{N \setminus S}(CS \setminus \widehat{CS}, \mathbf{x})\},$$

where the supremum is taken over all coalition structures $\widehat{CS}$ such that $CS|_S \subseteq \widehat{CS} \subseteq CS$ (intuitively, these are the coalitions of CS from which the members of S withdraw all of their resources) and over all deviations CS' of S from $CS \setminus \widehat{CS}$.

3.3 Redefining $\mathcal{A}$-Profitable Deviations

The definition of an $\mathcal{A}$-profitable deviation is somewhat difficult to work with. To remedy this, we will now provide a simple sufficient condition for the existence of an $\mathcal{A}$-profitable deviation.

Chalkiadakis et al. (2010) prove a similar result for the conservative arbitration function, and we use some of their techniques (most notably, a coloring argument) in our proof.

Theorem 3.4. *Consider an OCF game $\mathcal{G} = \langle N, v\rangle$, an outcome $(CS, \mathbf{x})$ of $\mathcal{G}$ and an arbitration function $\mathcal{A}$. If $\mathcal{A}^*(CS, \mathbf{x}, S) > p^S(CS, \mathbf{x})$ for some $S \subseteq N$, then there is a subset of S that can $\mathcal{A}$-profitably deviate from $(CS, \mathbf{x})$.*

Proof. Consider a set $S \subseteq N$ such that $\mathcal{A}^*(CS, \mathbf{x}, S) > p^S(CS, \mathbf{x})$. First, note that the space of possible deviations of S from the coalition structure CS can be viewed as a subset of $\mathbb{R}^D$ (for an appropriate value of D) that is bounded and can be described by finitely many non-strict linear inequalities. Therefore this space is compact. Thus, there exists a deviation of S from $(CS, \mathbf{x})$ maximizing the total payoff that S can derive from deviating; let CS' be some such deviation. Since v has the efficient coalition structure property, there exists a coalition structure CS_d such that $v(CS_d) = v^*(\mathbf{w}(CS|_S) + \mathbf{w}(CS'))$. For every $\mathbf{c} \in CS \setminus CS|_S$, let $\alpha_{\mathbf{c}}$ be the amount that $S \cap supp(\mathbf{c} - \mathbf{d}(\mathbf{c}))$ receives from $\mathbf{c}$ after the deviation. Our objective is to find an imputation $\mathbf{x}_d \in I(CS_d)$ along with payoff divisions of $(\alpha_{\mathbf{c}})_{\mathbf{c}\in CS\setminus CS|_S}$ that would witness that a subset of S has an $\mathcal{A}$-profitable deviation from $(CS, \mathbf{x})$.

Given a coalition $\mathbf{c} \in CS \setminus CS|_S$, let $I_{\mathbf{c}}$ be the set of all possible ways that $\alpha_{\mathbf{c}}$ can be divided among the members of $S \cap supp(\mathbf{c} - \mathbf{d}(\mathbf{c}))$. Formally, if $\alpha_{\mathbf{c}} \geq 0$, then

$$I_{\mathbf{c}} = \{\mathbf{y}_{\mathbf{c}} \in (\mathbb{R}_+)^n \mid supp(\mathbf{y}_{\mathbf{c}}) \subseteq S \cap supp(\mathbf{c} - \mathbf{d}(\mathbf{c})) \text{ and } \sum_{i=1}^{n} (\mathbf{y}_{\mathbf{c}})^i = \alpha_{\mathbf{c}}\};$$

if $\alpha_{\mathbf{c}} < 0$, then $I_{\mathbf{c}}$ is defined similarly, with all $\mathbf{y}_{\mathbf{c}} \in I_{\mathbf{c}}$ being vectors in $(\mathbb{R}_-)^n$. Given $\mathbf{x}_d \in I(CS_d)$ and $\mathbf{y} \in \prod_{\mathbf{c}\in CS\setminus CS|_S} I_{\mathbf{c}}$, recall that $q^i(\mathbf{x}_d, \mathbf{y})$ is the payoff of agent i from the deviation when the payoffs from CS_d and the arbitration function payoffs are divided according to $\mathbf{x}_d$ and $\mathbf{y}$, respectively. We now define a function $TL : I(CS_d) \times \prod_{\mathbf{c}\in CS\setminus CS|_S} I_{\mathbf{c}} \to \mathbb{R}$ as follows:

$$TL(\mathbf{x}_d, \mathbf{y}) = \sum_{i\in S} \min\left\{p^i(CS, \mathbf{x}) - q^i(\mathbf{x}_d, \mathbf{y}), 0\right\}.$$

The function TL measures the total loss incurred by the members of S who are hurt by the deviation; these are the members of S who, despite joining the deviators, do not enjoy a profit from deviating. Observe that TL is continuous, and that both $I(CS_d)$ and $\prod_{\mathbf{c}\in CS\setminus CS|_S} I_{\mathbf{c}}$ are compact sets; thus, TL attains its minimum value on its domain. Among all minimizers of this function, pick one for which the number of agents i with $p^i(CS, \mathbf{x}) < q^i(\mathbf{x}_d, \mathbf{y})$ is the largest; denote this point by $(\mathbf{x}_0, \mathbf{y}_0)$.

Now, let us color the agents in S as follows: an agent $i \in S$ is green if $p^i(CS, \mathbf{x}) < q^i(\mathbf{x}_0, \mathbf{y}_0)$; red if $p^i(CS, \mathbf{x}) > q^i(\mathbf{x}_0, \mathbf{y}_0)$ and white if $p^i(CS, \mathbf{x}) = q^i(\mathbf{x}_0, \mathbf{y}_0)$. Green agents derive a strictly positive benefit from the proposed deviation, red agents suffer a loss, and white agents break even.

Now, consider a green agent g, and suppose that g can legally transfer payoff to some agent i under $\mathbf{x}_0$ and $\mathbf{y}_0$, i.e., (a) both g and i are in the support of a coalition $\mathbf{c}$ in CS_d and g receives a positive payoff from $\mathbf{c}$, or (b) both g and i are in the support of a coalition $\mathbf{c} - \mathbf{d}(\mathbf{c})$ such that $\mathbf{c} \in CS \setminus CS|_S$, $\alpha_{\mathbf{c}} > 0$, and g receives a positive payoff from $\mathbf{c}$, or (c) both g and i are in the support of a coalition $\mathbf{c} - \mathbf{d}(\mathbf{c})$ such that $\mathbf{c} \in CS \setminus CS|_S$, $\alpha_{\mathbf{c}} < 0$, and i receives a negative payoff from $\mathbf{c}$. If i is red, g can transfer a small positive payoff to i while remaining green, and the resulting payoff division would have a strictly lower value of TL than $(\mathbf{x}_0, \mathbf{y}_0)$, a contradiction. Similarly, if i is white, g can transfer a small amount to i, making it green as well, without changing the value of TL; this is a contradiction, since we chose $(\mathbf{x}_0, \mathbf{y}_0)$ so as to maximize the number of green agents. Thus, if both g and i are in the support of some coalition after the deviation (i.e., either a coalition in CS_d or a surviving coalition in $CS \setminus CS|_S$), then g cannot receive a positive payoff from that coalition. By the same argument, if a coalition in $CS \setminus CS|_S$ assigns a negative payoff to the deviating agents under $\mathcal{A}$, then only the green agents incur a loss from that coalition.

Let S_g be the set of green agents in S. We have argued that the agents in S_g receive no payoff from coalitions they form with non-green agents in S. It follows that S_g can $\mathcal{A}$-profitably deviate from $(CS, \mathbf{x})$. Indeed, suppose that the agents in $S \setminus S_g$ do not deviate, and the agents in S_g deviate by only withdrawing resources they need for forming the coalitions among themselves in CS_d; the resources that S_g withdrew from CS in the deviation CS' in order to invest in coalitions with members of $S \setminus S_g$ in CS_d are kept as is. Since $\mathcal{A}$ is deviation-monotone, the payment that S_g receives from the arbitration function under this new deviation is weakly higher than under CS'. Furthermore, suppose that S_g forms the same coalitions as in $CS_d|_{S_g}$, divides payoffs from these coalitions according to $\mathbf{x}_0$, and distributes the arbitration function payoffs according to $\mathbf{y}_0$, with any surplus payments from the arbitration function shared arbitrarily. Then the payoffs of all members of S_g are at least as high as under CS', $\mathbf{x}_0$ and $\mathbf{y}_0$; since these agents were green, each of them benefits from this new deviation. Finally, note that $S_g \neq \emptyset$, since $p^S(CS, \mathbf{x}) < \mathcal{A}^*(CS, \mathbf{x}, S)$. This concludes the proof. □

Simply put, Theorem 3.4 states that if the most that a set S can get by deviating from $(CS, \mathbf{x})$ is strictly greater than its payoff from $(CS, \mathbf{x})$, then there is a non-empty subset of S that can $\mathcal{A}$-profitably deviate. Note that the distinction between coalitional value and the profitability of deviating exists in cooperative games with coalition structures as well.

Example 3.5. Consider a modified version of the three player 2-majority game: the agent set is $N = \{1, 2, 3\}$, the value of any coalition $S \subseteq N$ of size at least 2 is 6, and the value of each singleton is 1. Suppose that agents form the coalition N, and divide payoffs so that each agent receives 2. The set N can earn more by forming the coalition structure $(\{1, 2\}, \{3\})$, whose value is 7. However, there is no way to divide payoffs from this coalition structure so that every agent receives more than 2: agent 3 can only receive 1, so it is not better off. Thus, N cannot deviate as a whole, but the coalition $\{1, 2\}$ or the singleton $\{3\}$ (or both, depending on how the value of the grand coalition is distributed) have a profitable deviation.

Theorem 3.4 provides us with a justification for using $\mathcal{A}^*(CS, \mathbf{x}, S)$ as a measure of the deviators' satisfaction with the outcome. Indeed, if S can $\mathcal{A}$-profitably deviate from $(CS, \mathbf{x})$ via some deviation CS', then clearly $\mathcal{A}^*(CS, \mathbf{x}, S) > p^S(CS, \mathbf{x})$. On the other hand, if $\mathcal{A}^*(CS, \mathbf{x}, S) > p^S(CS, \mathbf{x})$, then there is some $T \subseteq S$ that can $\mathcal{A}$-profitably deviate from $(CS, \mathbf{x})$.

4. The Arbitrated Core

Given an OCF game $\mathcal{G} = \langle N, v \rangle$ and an arbitration function $\mathcal{A}$, we say that an outcome $(CS, \mathbf{x})$ is $\mathcal{A}$-*stable* if there is no set $S \subseteq N$ that can $\mathcal{A}$-profitably deviate from $(CS, \mathbf{x})$. The *arbitrated core of* $\mathcal{G}$ *with respect to an arbitration function* $\mathcal{A}$, or the $\mathcal{A}$-*core* of $\mathcal{G}$ (denoted $Core(\mathcal{G}, \mathcal{A})$), is the set of all $\mathcal{A}$-stable outcomes of $\mathcal{G}$. Using Theorem 3.4, we obtain the following characterization of $\mathcal{A}$-stable outcomes.

Theorem 4.1. *Let $\mathcal{G} = \langle N, v \rangle$ be an OCF game. An outcome $(CS, \mathbf{x})$ is in the $\mathcal{A}$-core of $\mathcal{G}$ if and only if for every $S \subseteq N$ we have $p^S(CS, \mathbf{x}) \geq \mathcal{A}^*(CS, \mathbf{x}, S)$.*

Proof. First, suppose that for every $S \subseteq N$ we have $p^S(CS, \mathbf{x}) \geq \mathcal{A}^*(CS, \mathbf{x}, S)$. Let CS' be an arbitrary deviation of S from CS. We have $p^S(CS, \mathbf{x}) \geq \mathcal{A}(CS, \mathbf{x}, S, CS')$. Thus, no matter how S divides the deviation payoffs, there would be at least one agent $i \in S$ who gets no more than $p^i(CS, \mathbf{x})$. As this is true for an arbitrary deviation CS', it follows that S cannot $\mathcal{A}$-profitably deviate.

Conversely, suppose that $p^S(CS, \mathbf{x}) < \mathcal{A}^*(CS, \mathbf{x}, S)$. By Theorem 3.4, there is a subset of S that can $\mathcal{A}$-profitably deviate from $(CS, \mathbf{x})$, thus $(CS, \mathbf{x})$ is not $\mathcal{A}$-stable. □

Let $\mathcal{A}$ and $\bar{\mathcal{A}}$ be two arbitration functions. Intuitively, it is clear that if $\mathcal{A}$ is more lenient than $\bar{\mathcal{A}}$, then the $\mathcal{A}$-core is contained in the $\bar{\mathcal{A}}$-core. Indeed, if an outcome $(CS, \mathbf{x})$ is stable with respect to $\mathcal{A}$ and $\mathcal{A}$ always pays more to the deviators than $\bar{\mathcal{A}}$ (thus making the deviation more tempting), then $(CS, \mathbf{x})$ should be $\bar{\mathcal{A}}$-stable as well. To formalize this intuition, we use the following notation. Given two arbitration functions $\mathcal{A}$ and $\bar{\mathcal{A}}$, we write $\mathcal{A} \geq \bar{\mathcal{A}}$ if for every set $S \subseteq N$ and every outcome $(CS, \mathbf{x})$ it holds that $\mathcal{A}^*(CS, \mathbf{x}, S) \geq \bar{\mathcal{A}}^*(CS, \mathbf{x}, S)$. We can now state our observation as follows.

Corollary 4.2. *If $\mathcal{A} \geq \bar{\mathcal{A}}$ then $Core(\mathcal{G}, \mathcal{A}) \subseteq Core(\mathcal{G}, \bar{\mathcal{A}})$.*

Proof. Suppose that $(CS, \mathbf{x})$ is in $Core(\mathcal{G}, \mathcal{A})$. This means that for all $S \subseteq N$ we have $p^S(CS, \mathbf{x}) \geq \mathcal{A}^*(CS, \mathbf{x}, S)$. Since $\mathcal{A} \geq \bar{\mathcal{A}}$, it follows that $p^S(CS, \mathbf{x}) \geq \bar{\mathcal{A}}^*(CS, \mathbf{x})$ for all $S \subseteq N$ as well, which implies $(CS, \mathbf{x}) \in Core(\mathcal{G}, \bar{\mathcal{A}})$. □

The conservative, sensitive, refined and optimistic arbitration functions described in Section 3.2, denoted $\mathcal{A}_c$, $\mathcal{A}_s$, $\mathcal{A}_r$ and $\mathcal{A}_o$, respectively, satisfy $\mathcal{A}_o \geq \mathcal{A}_r \geq \mathcal{A}_s \geq \mathcal{A}_c$. Thus, Corollary 4.2 implies that the conservative core contains the sensitive core, which contains the refined core, which contains the optimistic core. These observations (with respect to the conservative, refined, and optimistic core only) have been made by Chalkiadakis et al. (2010) as well; in fact, Chalkiadakis et al. show that these containments can be strict, i.e., there are games for which the conservative core strictly contains the refined core, and there are games for which the refined core strictly contains the optimistic core. A similar separation can be shown for the sensitive core.

4.1 Some Arbitrated Cores

Using the characterization result of Theorem 4.1, we now proceed to describe some arbitrated cores, namely, those corresponding to the arbitration functions introduced in Section 3.2.

4.1.1 The Conservative Core

As argued in Section 3.2, for every outcome $(CS, \mathbf{x})$ and every set $S \subseteq N$, the most that S can get under the conservative arbitration function is $v^*(\mathbf{e}^S)$. Thus, the conservative core is the set of all outcomes $(CS, \mathbf{x})$ such that for every set $S \subseteq N$ it holds that

$$p^S(CS, \mathbf{x}) \geq v^*(\mathbf{e}^S). \tag{3}$$

4.1.2 THE SENSITIVE CORE

In order for an outcome $(CS, \mathbf{x})$ to be stable under the sensitive arbitration function, for every set $S \subseteq N$ and every $T \subseteq N \setminus S$ it should hold that

$$p^S(CS, \mathbf{x}) \geq v^*(\mathbf{w}(CS|_S) + \mathbf{w}^S(CS_T)) + p^S(CS \setminus (CS|_S \cup CS_T), \mathbf{x});$$

recall that CS_T is the set of all coalitions in CS that involve members of T. Subtracting payments associated with coalitions in $CS \setminus (CS|_S \cup CS_T)$ from both sides of this inequality, we get

$$p^S(CS|_S, \mathbf{x}) + p^S(CS_T, \mathbf{x}) \geq v^*(\mathbf{w}(CS|_S) + \mathbf{w}^S(CS_T)).$$

Chalkiadakis et al. (2010) show that if CS is not an optimal coalition structure then $(CS, \mathbf{x})$ is not in the conservative core and consequently it is not in the sensitive core either. Thus, we can assume that CS is an optimal coalition structure, and hence $CS|_S$ is optimal for any $S \subseteq N$, i.e., $v(CS|_S) = v^*(\mathbf{w}(CS|_S))$. We also have $p^S(CS|_S, \mathbf{x}) = v(CS|_S)$. This implies that, in order for $(CS, \mathbf{x})$ to be in the sensitive core, it must be the case that CS is an optimal coalition structure and for every $S \subseteq N$ and every $T \subseteq N \setminus S$ it holds that

$$p^S(CS_T, \mathbf{x}) \geq v^*(\mathbf{w}(CS|_S) + \mathbf{w}^S(CS_T)) - v^*(\mathbf{w}(CS|_S)), \tag{4}$$

i.e., the total payoff to S from CS_T must be at least the marginal benefit of using the resources S has invested in CS_T. Note that if we take $T = N \setminus S$, we obtain the conservative core condition.

4.1.3 THE REFINED CORE

For the refined core, we have

$$\mathcal{A}_r^*(CS, \mathbf{x}, S) = \max\{v^*(\mathbf{w}^S(\widehat{CS})) + p^S(CS \setminus \widehat{CS}, \mathbf{x}) \mid CS|_S \subseteq \widehat{CS} \subseteq CS\}.$$

Thus, an outcome $(CS, \mathbf{x})$ is in the refined core if and only if for every $S \subseteq N$ and for every coalition structure $\widehat{CS}$ containing $CS|_S$ it holds that

$$p^S(CS, \mathbf{x}) \geq v^*(\mathbf{w}^S(\widehat{CS})) + p^S(CS \setminus \widehat{CS}, \mathbf{x}).$$

Consequently, $(CS, \mathbf{x})$ is in the refined core if and only if for every such S and $\widehat{CS}$ it holds that

$$p_S(\widehat{CS}, \mathbf{x}) \geq v^*(\mathbf{w}^S(\widehat{CS})). \tag{5}$$

In other words, the payoff to S from every coalition structure $\widehat{CS}$ containing $CS|_S$ should be at least as large as the most that S can get from the resources it has invested in $\widehat{CS}$. Note that if we limit ourselves to coalition structures $\widehat{CS}$ such that $\widehat{CS} = CS|_S \cup CS_T$ for some $T \subseteq N \setminus S$, we obtain the sensitive core condition.

4.1.4 THE OPTIMISTIC CORE

For the optimistic arbitration function, $\mathcal{A}_o$, we have shown that

$$\mathcal{A}_o^*(CS, \mathbf{x}, S) = \sup\{v^*(\mathbf{w}^S(\widehat{CS}) + \mathbf{w}(CS')) + \sum_{\mathbf{c} \in CS \setminus \widehat{CS}} v(\mathbf{c} - \mathbf{d}(\mathbf{c})) - p^{N \setminus S}(CS \setminus \widehat{CS}, \mathbf{x})\},$$

where the supremum is taken over all coalition structures $\widehat{CS}$ such that $CS|_S \subseteq \widehat{CS} \subseteq CS$ and over all deviations CS' of S from $CS \setminus \widehat{CS}$. Thus, in order for an outcome $(CS, \mathbf{x})$ to be in the optimistic core, it must be the case that for every set $S \subseteq N$, every coalition structure $\widehat{CS} \subseteq CS$ containing $CS|_S$ and every deviation CS' of S from $CS \setminus \widehat{CS}$ it holds that

$$p^S(CS, \mathbf{x}) \geq v^*(\mathbf{w}^S(\widehat{CS}) + \mathbf{w}(CS')) + \sum_{\mathbf{c} \in CS \setminus \widehat{CS}} v(\mathbf{c} - \mathbf{d}(\mathbf{c})) - p^{N \setminus S}(CS \setminus \widehat{CS}, \mathbf{x}). \quad (6)$$

Let $\Delta(CS \setminus \widehat{CS}, CS')$ denote the coalition structure formed by what remains of $CS \setminus \widehat{CS}$ after S has withdrawn resources from CS according to CS'. We obtain $\sum_{\mathbf{c} \in CS \setminus \widehat{CS}} v(\mathbf{c} - \mathbf{d}(\mathbf{c})) = v(\Delta(CS \setminus \widehat{CS}, CS'))$. Thus, subtracting $p^S(CS \setminus \widehat{CS}, \mathbf{x})$ from both sides of inequality (6) and using the fact that $p^N(CS \setminus \widehat{CS}, \mathbf{x}) = v(CS \setminus \widehat{CS})$, we can rewrite inequality (6) as

$$p^S(\widehat{CS}, \mathbf{x}) \geq v^*(\mathbf{w}^S(\widehat{CS}) + \mathbf{w}(CS')) - \Big(v(CS \setminus \widehat{CS}) - v(\Delta(CS \setminus \widehat{CS}, CS'))\Big). \quad (7)$$

The optimistic core stability condition is thus similar to the refined core stability condition; however, S is also allowed to withdraw resources from coalitions outside of $\widehat{CS}$ if it is willing to assume the marginal costs of this withdrawal.

Example 4.3. Consider the following example. We have three players, $N = \{1, 2, 3\}$, and the following characteristic function:

$$v\begin{pmatrix} 0 \\ \frac{1}{4} \\ 0 \end{pmatrix} = 1, \qquad v\begin{pmatrix} 1 \\ \frac{1}{2} \\ 0 \end{pmatrix} = 40, \qquad v\begin{pmatrix} 0 \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix} = 20,$$

and $v(\mathbf{c}) = 0$ for every other coalition $\mathbf{c} \in [0, 1]^3$. First, we observe that the only optimal coalition structure (up to the order of coalitions) is $CS = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$, where

$$\mathbf{c}_1 = \begin{pmatrix} 1 \\ \frac{1}{2} \\ 0 \end{pmatrix}, \qquad \mathbf{c}_2 = \begin{pmatrix} 0 \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix}, \qquad \mathbf{c}_3 = \begin{pmatrix} 0 \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix}.$$

In words, players 1 and 2 collaborate on one project (which requires 100% of player 1's resources, and 50% of player 2's resources), and generate a revenue of 40; players 2 and 3 work on two identical projects (each requiring 25% of player 2's resources and 50% of player 3's resources), generating a revenue of 40 in total as well. Now, suppose that payoffs are divided according to $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, where

$$\mathbf{x}_1 = \begin{pmatrix} 40 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 0 \\ 20 \end{pmatrix}, \qquad \mathbf{x}_3 = \begin{pmatrix} 0 \\ 4 \\ 16 \end{pmatrix}.$$

Under the conservative arbitration function, this payoff division is stable: no subset of players receives less than the most it can make on its own. However, under the sensitive arbitration function, this is no longer the case: player 2 must receive a payoff of at least 2 from $\mathbf{c}_1$, and a combined payoff of at least 2 from $\mathbf{c}_2$ and $\mathbf{c}_3$.

Now, let us suppose that the following payoff division is proposed instead: $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$, where

$$\mathbf{y}_1 = \begin{pmatrix} 38 \\ 2 \\ 0 \end{pmatrix}, \qquad \mathbf{y}_2 = \begin{pmatrix} 0 \\ 0 \\ 20 \end{pmatrix}, \qquad \mathbf{y}_3 = \begin{pmatrix} 0 \\ 2 \\ 18 \end{pmatrix}.$$

The outcome $(CS, \mathbf{y})$ is stable with respect to the sensitive arbitration function; however, it is not stable with respect to the refined arbitration function: player 2 can withdraw its contribution to $\mathbf{c}_2$ and earn strictly more. Thus, in order for an imputation to be r-stable, player 2 must receive at least 2 from $\mathbf{c}_1$, and at least 1 from each of $\mathbf{c}_2$ and $\mathbf{c}_3$.

Example 4.3 illustrates an interesting phenomenon: more lenient arbitration functions force payments to be more spread out: player 2 receives the same payoff, but the payoff must be received in a more distributed manner when more lenient reactions to deviations are assumed.

5. Non-emptiness of Arbitrated Cores

Core outcomes are highly desirable both in classic cooperative games and in OCF games. Thus, it is important to find conditions on the characteristic function of an OCF game that would guarantee non-emptiness of its arbitrated core, as well as to identify classes of OCF games whose arbitrated core is always non-empty, for natural arbitration functions. We note that both of these questions can be—and have been—looked at through the lens of computational complexity. That is, one can ask whether there exist *polynomial-time* algorithms that decide whether the arbitrated core is non-empty or find an outcome in the arbitrated core, either for the general model or for specific classes of OCF games. While this question has been explored by Chalkiadakis et al. (2010) and Zick et al. (2012), in this paper we do not consider computational complexity issues and focus instead on characterization results (see, however, the companion paper by Zick et al., 2014).

Our characterization results are similar in spirit to the celebrated Bondareva–Shapley condition for core non-emptiness in classic cooperative games (Bondareva, 1963; Shapley, 1967). We will now briefly state this condition. Given a set $N = \{1, \dots, n\}$, a collection of weights $(\delta_S)_{S \subseteq N}$ is called *balanced* if $\delta_S \geq 0$ for all $S \subseteq N$ and $\sum_{S:i \in S} \delta_S = 1$ for all $i \in N$. We can view a balanced collection of weights as a way for an agent i to partially participate in all sets that contain i, with its contribution specified by the respective weight.

Given a cooperative game $\mathcal{G} = \langle N, u \rangle$ where $u : 2^N \to \mathbb{R}_+$, we say that $\mathcal{G}$ is *balanced* if for every balanced collection of weights $(\delta_S)_{S \subseteq N}$ it holds that $\sum_{S \subseteq N} \delta_S u(S) \leq u(N)$.

Theorem 5.1 (Bondareva–Shapley Theorem). *A cooperative game $\mathcal{G} = \langle N, u \rangle$ with $u : 2^N \to \mathbb{R}_+$ has a non-empty core if and only if it is balanced.*

An important feature of this characterization is that it is stated in terms of the characteristic function itself, and does not explicitly refer to outcomes of the game. In what follows, we describe similar characterizations for OCF games, for the conservative, sensitive and refined arbitration functions.[1] Specifically, for each of these arbitration functions, given an OCF game $\mathcal{G} = \langle N, v \rangle$ and a coalition structure $CS \in \mathcal{CS}(N)$, we characterize the set of imputations $\mathbf{x} \in I(CS)$ such that $(CS, \mathbf{x})$ is in the respective arbitrated core of $\mathcal{G}$. We limit our attention to optimal coalition structures, i.e., we assume that $v(CS) = v^*(e^N)$: if CS is not optimal, then no outcome of the form $(CS, \mathbf{x})$ can be in the arbitrated core of $\mathcal{G}$ for any of our arbitration functions, since the grand coalition itself can profitably deviate from $(CS, \mathbf{x})$ (Chalkiadakis et al., 2010).

1. The case of the conservative arbitration function has been considered by Chalkiadakis et al. (2010); we reproduce their results for completeness and to facilitate the comparison with the other two cases.

5.1 The Conservative Core

Given an OCF game $\mathcal{G} = \langle N, v\rangle$ and a coalition structure $CS = (\mathbf{c}_1, \dots, \mathbf{c}_m)$, we say that a collection of non-negative weights $\{(r_j)_{j=1}^m; (\delta_S)_{S\subseteq N}\}$ is *c-balanced with respect to CS* if for every $i \in N$ and every coalition $\mathbf{c}_j$ such that $i \in \mathit{supp}(\mathbf{c}_j)$ it holds that $r_j + \sum_{S:i\in S}\delta_S = 1$. Chalkiadakis et al. (2010) show the following result.

Theorem 5.2 (Chalkiadakis et al., 2010, Thm. 2)**.** *Given an OCF game $\mathcal{G} = \langle N, v\rangle$ and an optimal coalition structure $CS \in \mathcal{CS}(N)$, there exists an outcome $\mathbf{x} \in I(CS)$ such that $(CS, \mathbf{x})$ is in the conservative core of $\mathcal{G}$ if and only if for every collection of non-negative weights $\{(r_j)_{j=1}^m; (\delta_S)_{S\subseteq N}\}$ that is c-balanced with respect to CS it holds that*

$$\sum_{j=1}^m r_j v(\mathbf{c}_j) + \sum_{S\subseteq N} \delta_S v^*(\mathbf{e}^S) \le v^*(\mathbf{e}^N).$$

Our first observation is that the proof of Theorem 5.2, as given in the work of Chalkiadakis et al. (2010), has a gap. Appendix A contains a correction to the proof.

We will now provide an alternative characterization of OCF games with non-empty conservative core, by establishing a connection between the conservative core of an OCF game and the core of a related classic cooperative game.

Given an OCF game $\mathcal{G} = \langle N, v\rangle$, its *discrete superadditive cover* is a classic cooperative game $\bar{\mathcal{G}} = \langle N, U_v\rangle$, where $U_v(S) = v^*(\mathbf{e}^S)$. Simply put, the value of a set $S \subseteq N$ in $\bar{\mathcal{G}}$ is the most that it can make under the function v by forming a coalition structure. We will now show that the conservative core of $\mathcal{G}$ is non-empty if and only if the core of $\bar{\mathcal{G}}$ is non-empty.

Theorem 5.3. *The conservative core of an OCF-game $\mathcal{G} = \langle N, v\rangle$ is non-empty if and only if the core of its discrete superadditive cover $\bar{\mathcal{G}} = \langle N, U_v\rangle$ is non-empty.*

Proof. We have argued that if the conservative core of $\mathcal{G}$ is non-empty, then there exists an outcome $(CS, \mathbf{x})$ such that for all $S \subseteq N$ we have $p^S(CS, \mathbf{x}) \ge v^*(\mathbf{e}^S)$. Note that CS is an optimal coalition structure for $\mathcal{G}$, since $v(CS) = p^N(CS, \mathbf{x}) \ge v^*(\mathbf{e}^N)$. Consider the payoff vector $\mathbf{p} = (p^1, \dots, p^n)$ where $p^i = p^i(CS, \mathbf{x})$ for all $i \in N$. It follows that $\mathbf{p}$ is in the core of $\bar{\mathcal{G}}$. Indeed, since CS is optimal, we have $\sum_{i=1}^n p^i = v(CS) = v^*(\mathbf{e}^N) = U_v(N)$, and for every $S \subseteq N$ we have $p(S) = p^S(CS, \mathbf{x}) \ge v^*(\mathbf{e}^S) = U_v(S)$.

Conversely, let $\mathbf{p} = (p^1, \dots, p^n)$ be a payoff vector in the core of $\bar{\mathcal{G}}$. Fix a coalition structure $CS = (\mathbf{c}_1, \dots, \mathbf{c}_m)$ such that $v(CS) = v^*(\mathbf{e}^N)$. We will now use $\mathbf{p}$ to construct an imputation $\mathbf{x} \in I(CS)$ such that $(CS, \mathbf{x})$ is in the conservative core of $\mathcal{G}$.

Given an imputation $\mathbf{z} \in I(CS)$, we color the agents in N as follows: i is green if $p^i(CS, \mathbf{z}) > p^i$; red if $p^i(CS, \mathbf{z}) < p^i$ and white if $p^i(CS, \mathbf{z}) = p^i$. Green agents are better off under $(CS, \mathbf{z})$ than under $\mathbf{p}$, red agents are worse off under $(CS, \mathbf{z})$ than under $\mathbf{p}$, and white agents are indifferent.

Now, suppose that there is some $\mathbf{z} \in I(CS)$ such that no agent in the respective coloring is green. Since CS is optimal, this means that no agent is red either. Thus, all agents are white. This in turn implies that for all $S \subseteq N$ we have $p^S(CS, \mathbf{z}) = p(S) \ge U_v(S) = v^*(\mathbf{e}^S)$, so $(CS, \mathbf{z})$ is in the conservative core. Thus, all we need to show is that there exists some $\mathbf{z} \in I(CS)$ such that no agent in the respective coloring is green.

Let $F : I(CS) \to \mathbb{R}$ be defined as follows:

$$F(\mathbf{z}) = \sum_{i=1}^n \max\{0, p^i - p^i(CS, \mathbf{z})\}.$$

The function F measures the agents' total unhappiness with the outcome $(CS, \mathbf{z})$ as compared to $\mathbf{p}$. F is a continuous function (it is the sum of the maxima of functions that are continuous in $\mathbf{z}$), and $I(CS)$ is a compact set. Thus, there exists an imputation $\mathbf{x} \in I(CS)$ that minimizes the value of F. We pick $\mathbf{x} \in \arg\min F(\mathbf{z})$ so as to minimize the number of white agents in the respective coloring of N.

Now suppose that there exist a green agent g, a non-green agent i and a coalition $\mathbf{c} \in CS$ such that $g, i \in supp(\mathbf{c})$ and $x(\mathbf{c})^g > 0$. Then there is some $\varepsilon > 0$ such that setting $\bar{x}(\mathbf{c})^g = x(\mathbf{c})^g - \varepsilon$ and $\bar{x}(\mathbf{c})^i = x(\mathbf{c})^i + \varepsilon$ (i.e., transferring an amount of ε from g to i within the coalition $\mathbf{c}$) results in a valid payoff division for CS and keeps g green. If i was white prior to this transfer, it now becomes green (and the value of F does not change). This is a contradiction to the fact that $\mathbf{x}$ minimizes the number of white agents. On the other hand, if i was red, the transfer decreases the contribution of i to F (while g's contribution to F remains 0), which is a contradiction to the fact that $\mathbf{x} \in \arg\min F(\mathbf{z})$. This means that green agents get zero payoff from coalitions with non-green agents. Let us denote the set of green agents by S_g. We have

$$p^{S_g}(CS, \mathbf{x}) = p^{S_g}(CS|_{S_g}, \mathbf{x}) \leq v^*(\mathbf{e}^{S_g}).$$

On the other hand, if $S_g \neq \emptyset$, since for each $i \in S_g$ we have $p^i(CS, \mathbf{x}) > p^i$, we obtain

$$p^{S_g}(CS, \mathbf{x}) > p(S_g) \geq U_v(S_g) = v^*(\mathbf{e}^{S_g}).$$

This contradiction shows that $S_g = \emptyset$, and we have argued that in this case all agents are white. This completes the proof. □

The following corollary is immediately implied by the proof of Theorem 5.3.

Corollary 5.4. *Consider an OCF game $\mathcal{G} = \langle N, v \rangle$. For every payoff vector $\mathbf{p} = (p^1, \ldots, p^n)$ in the core of $\bar{\mathcal{G}}$ and every optimal coalition structure $CS \in \mathcal{CS}(N)$, there exists an $\mathbf{x} \in I(CS)$ such that $p^i = p^i(CS, \mathbf{x})$ for all $i \in N$.*

Corollary 5.4 enables us to generalize the following well-known result for classic cooperative games. Aumann and Drèze (1974) show that if CS and CS' are optimal (non-overlapping) coalition structures for $\mathcal{G} = \langle N, u \rangle$ and $(CS, \mathbf{p})$ is in the core of $\mathcal{G}$ then $(CS', \mathbf{p})$ is also in the core of $\mathcal{G}$. In other words, if $\mathcal{CS}_{\text{opt}}$ is the set of all optimal coalition structures for $\mathcal{G}$, and I_{stab} is the set of all stable payoff divisions for $\mathcal{G}$, then the set of core outcomes for $\mathcal{G}$ is $\mathcal{CS}_{\text{opt}} \times I_{\text{stab}}$. Using Corollary 5.4, we can extend this result to OCF games under the conservative arbitration function.

Corollary 5.5. *Consider an OCF game $\mathcal{G} = \langle N, v \rangle$. For every pair of optimal coalition structures $CS, CS' \in \mathcal{CS}(N)$ and every imputation $\mathbf{x} \in I(CS)$ such that $(CS, \mathbf{x})$ is in the conservative core, there exists an imputation $\mathbf{x}' \in I(CS')$ such that $(CS', \mathbf{x}')$ is in the conservative core of $\mathcal{G}$ and $p^i(CS, \mathbf{x}) = p^i(CS', \mathbf{x}')$ for all $i \in N$.*

5.1.1 CONVEXITY IN OCF GAMES REVISITED

For classic cooperative games, *convexity*, or *supermodularity*, of the characteristic function is well-known to be a sufficient condition for non-emptiness of the core (Shapley, 1971). In more detail, recall that a cooperative game $\mathcal{G} = \langle N, v \rangle$ is said to be *supermodular* if for every pair of sets S, T such that $S \subseteq T \subseteq N$ and every $R \subseteq N \setminus T$ it holds that

$$v(T \cup R) - v(T) \geq v(S \cup R) - v(S).$$

Supermodular games are often referred to as convex games in the literature; however, to avoid confusion with other notions of convexity considered in this paper, we use the term "supermodularity". Shapley (1971) proves the following result.

Theorem 5.6. *If a cooperative game $\mathcal{G}$ is supermodular, then its core is non-empty.*

In order to extend Theorem 5.6 to OCF games, Chalkiadakis et al. (2010) propose the following notion of convexity for OCF games.

Definition 5.7 (OCF convexity, Chalkiadakis et al., 2010)**.** *An OCF game is* OCF-convex *if for every pair of sets S, T such that $S \subseteq T \subseteq N$, every $R \subseteq N \setminus T$, every outcome $(CS_S, \mathbf{x}_S) \in \mathcal{F}(S)$, every outcome $(CS_T, \mathbf{x}_T) \in \mathcal{F}(T)$, and every outcome $(CS_{S\cup R}, \mathbf{x}_{S\cup R}) \in \mathcal{F}(S \cup R)$ the following condition holds: if $p^i(CS_S, \mathbf{x}_S) \leq p^i(CS_{S\cup R}, \mathbf{x}_{S\cup R})$ for all $i \in S$ then there is an outcome $(CS_{T\cup R}, \mathbf{x}_{T\cup R}) \in \mathcal{F}(T \cup R)$ such that*

(1) $p^i(CS_T, \mathbf{x}_T) \leq p^i(CS_{T\cup R}, \mathbf{x}_{T\cup R})$ for all $i \in T$;

(2) $p^i(CS_{S\cup R}, \mathbf{x}_{S\cup R}) \leq p^i(CS_{T\cup R}, \mathbf{x}_{T\cup R})$ for all $i \in S \cup R$.

Intuitively, Definition 5.7 simply means that larger coalitions can offer more lucrative options to players who join them, compared to smaller coalitions.

Chalkiadakis et al. (2010) then show that OCF convexity is a sufficient condition for non-emptiness of the conservative core.

Theorem 5.8 (Chalkiadakis et al., 2010, Thm. 3)**.** *If an OCF game $\mathcal{G}$ is OCF-convex, then its conservative core is non-empty.*

On the other hand, combining Theorem 5.3 with Theorem 5.6, we obtain the following sufficient condition for conservative core non-emptiness.

Proposition 5.9. *Consider an OCF game $\mathcal{G}$. If $\bar{\mathcal{G}} = \langle N, U_v \rangle$ is supermodular, then the conservative core of $\mathcal{G}$ is non-empty.*

We will now argue that Proposition 5.9 is strictly stronger than Theorem 5.8, by showing that supermodularity of $\bar{\mathcal{G}}$ is a strictly weaker condition than OCF-convexity of $\mathcal{G}$. We will first show that OCF-convexity of $\mathcal{G}$ implies supermodularity of $\bar{\mathcal{G}}$. We then present an example of a game $\mathcal{G} = \langle N, v \rangle$ such that $\bar{\mathcal{G}}$ is supermodular, but $\mathcal{G}$ is not OCF-convex (Example 5.13).

To implement the first step of this plan, we first establish the following crucial proposition, which may seem counterintuitive at first sight. Its proof is based on coloring arguments similar to the ones used in the proofs of Theorems 3.4 and 5.3.

Proposition 5.10. *For every pair of sets S, T such that $S \subseteq T \subseteq N$ and every pair of optimal coalition structures $CS_S \in \mathcal{CS}(S), CS_T \in \mathcal{CS}(T)$, there exist imputations $\mathbf{x} \in I(CS_S)$ and $\mathbf{y} \in I(CS_T)$ such that $p^i(CS_S, \mathbf{x}) = p^i(CS_T, \mathbf{y})$ for all $i \in S$.*

Proof. Fix the sets S and T and optimal coalition structures $CS_S \in \mathcal{CS}(S)$, $CS_T \in \mathcal{CS}(T)$. We will first establish that we can find imputations $\mathbf{x}_0 \in I(CS_S)$ and $\mathbf{y}_0 \in I(CS_T)$ with $p^i(CS_S, \mathbf{x}_0) \leq p^i(CS_T, \mathbf{y}_0)$ for all $i \in S$. We will then show how to use this fact to prove our original claim.

Lemma 5.11. *There exist imputations $\mathbf{x}_0 \in I(CS_S)$ and $\mathbf{y}_0 \in I(CS_T)$ such that $p^i(CS_S, \mathbf{x}_0) \leq p^i(CS_T, \mathbf{y}_0)$ for all $i \in S$.*

Proof. Given a pair of imputations $\mathbf{x} \in I(CS_S)$, $\mathbf{y} \in I(CS_T)$, we color the agents in T as follows: an agent i is green if $i \in S$ and $p^i(CS_T, \mathbf{y}) > p^i(CS_S, \mathbf{x})$ or if $i \in T \setminus S$ and $p^i(CS_T, \mathbf{y}) > 0$; red if $i \in S$ and $p^i(CS_T, \mathbf{y}) < p^i(CS_S, \mathbf{x})$; and white otherwise.

We define a mapping $D : I(CS_S) \times I(CS_T) \to \mathbb{R}$ by setting

$$D(\mathbf{x}, \mathbf{y}) = \sum_{i \in S} \max\{0, p^i(CS_S, \mathbf{x}) - p^i(CS_T, \mathbf{y})\}.$$

The function D is continuous over a compact set, and thus attains its minimum value at some point $(\mathbf{x}, \mathbf{y}) \in I(CS_S) \times I(CS_T)$. Among all such points, pick one with the smallest number of white agents and denote it by $(\mathbf{x}_0, \mathbf{y}_0)$. Let $q^i = p^i(CS_T, \mathbf{y}_0)$ for $i \in T$, $p^i = p^i(CS_S, \mathbf{x}_0)$ for $i \in S$. We will now show that $q^i \geq p^i$ for all $i \in S$.

Assume for the sake of contradiction that this is not the case. Then $D(\mathbf{x}_0, \mathbf{y}_0) > 0$ and there is at least one red agent in T. Since $v^*(T) \geq v^*(S)$, there is also a green agent in T.

Now, consider a green agent g and a non-green agent i. Suppose that

(a) there exists a coalition $\mathbf{c} \in CS_T$ such that $g, i \in supp(\mathbf{c})$ and $y_0(\mathbf{c})^g > 0$, or

(b) there exists a coalition $\mathbf{c} \in CS_S$ such that $g, i \in supp(\mathbf{c})$ and $x_0(\mathbf{c})^i > 0$.

In case (a) we can modify $\mathbf{y}_0$ by making g transfer a small amount of payoff to i, and in case (b) we can modify $\mathbf{x}_0$ by making i transfer a small amount of payoff to g; we choose the transfer to be small enough that g remains green. If i was white, it now becomes green, and if i was red, this lowers its contribution to D. In both cases, we get a contradiction with our choice of $(\mathbf{x}_0, \mathbf{y}_0)$. Thus, no such coalition $\mathbf{c}$ exists.

Let S_g denote the set of all green agents in S, and let $S_{ng} = S \setminus S_g$. Under $\mathbf{x}_0$, every $i \in S_{ng}$ gets no payoff from coalitions in CS_S whose support contains members of S_g. Similarly, under $\mathbf{y}_0$ green agents in T get no payoff from coalitions in CS_T that contain members of S_{ng}.

Now, suppose that we modify CS_T by having the agents in S_{ng} abandon their existing coalitions and form the coalitions they form among themselves in CS_S instead. Denote the resulting coalition structure by CS'. We define an imputation $\mathbf{z}_0$ for CS' as follows. The imputation $\mathbf{z}_0$ coincides with $\mathbf{y}_0$ for each $\mathbf{c}$ with $supp(\mathbf{c}) \cap S_{ng} = \emptyset$, and it coincides with $\mathbf{x}_0$ for each coalition $\mathbf{c}'$ that is formed by the members of S_{ng}. The remaining coalitions in CS' involve both agents in S_{ng} and agents in $T \setminus S_{ng}$, and may have suffered a reduction in their value relative to CS_T after S_{ng} withdrew resources from them; for such coalitions $\mathbf{z}_0$ distributes their value arbitrarily among the members of S_{ng}.

As argued above, in $(CS_S, \mathbf{x}_0)$ the members of S_{ng} only receive payoffs from coalitions that they fully control, so $p^i(CS', \mathbf{z}_0) \geq p^i(CS_S, \mathbf{x}_0)$ for all $i \in S_{ng}$. Further, since at least one agent in S is red, we have $p^{S_{ng}}(CS_S, \mathbf{x}_0) > p^{S_{ng}}(CS_T, \mathbf{y}_0)$. On the other hand, in $(CS_T, \mathbf{y}_0)$ the members of $T \setminus S_{ng}$ receive no payoff from coalitions with members of S_{ng}, so letting S_{ng} receive all of the payoffs from those coalitions does not affect the payoff that the members of $T \setminus S_{ng}$ receive, relative to what they received in $(CS_T, \mathbf{y}_0)$. Thus, for all i in $T \setminus S_{ng}$ it holds that $p^i(CS', \mathbf{z}_0) = p^i(CS_T, \mathbf{y}_0)$. Hence, we obtain

$$\begin{aligned} v(CS') &= p^T(CS', \mathbf{z}_0) \\ &= p^{T \setminus S_{ng}}(CS', \mathbf{z}_0) + p^{S_{ng}}(CS', \mathbf{z}_0) \\ &> p^{T \setminus S_{ng}}(CS_T, \mathbf{y}_0) + p^{S_{ng}}(CS_T, \mathbf{y}_0) \\ &= v(CS_T) = v^*(\mathbf{e}^T). \end{aligned} \tag{8}$$

This contradiction shows that the set of red agents is empty and hence $p^i(CS_T, \mathbf{y}_0) \geq p^i(CS_S, \mathbf{x}_0)$ for all $i \in S$. □

We will now use Lemma 5.11 to complete the proof of Proposition 5.10. Define

$$P(CS_S, CS_T) = \{(\mathbf{x}, \mathbf{y}) \in I(CS_S) \times I(CS_T) \mid p^i(CS_T, \mathbf{y}) \geq p^i(CS_S, \mathbf{x}) \text{ for all } i \in S\}.$$

Lemma 5.11 implies that $P(CS_S, CS_T)$ is not empty. Moreover, $P(CS_S, CS_T)$ is compact; thus, it contains a point $(\mathbf{x}, \mathbf{y})$ that minimizes the value $p^S(CS_T, \mathbf{y})$ over $P(CS_S, CS_T)$.

Let $M(CS_S, CS_T)$ be the set of points $(\mathbf{x}, \mathbf{y})$ in $P(CS_S, CS_T)$ that minimize $p^S(CS_T, \mathbf{y})$; to complete the proof, we will identify a point $(\mathbf{x}_1, \mathbf{y}_1)$ in $M(CS_S, CS_T)$ such that $p^S(CS_T, \mathbf{y}_1) = p^S(CS_S, \mathbf{x}_1) = v^*(\mathbf{e}^S)$.

Given a point $(\mathbf{x}, \mathbf{y}) \in M(CS_S, CS_T)$, we color the agents in T as follows: i is green if $i \in S$ and $p^i(CS_T, \mathbf{y}) > p^i(CS_S, \mathbf{x})$, and white otherwise. Note that, since $(\mathbf{x}, \mathbf{y}) \in P(CS_S, CS_T)$, if S contains a white agent i, we have $p^i(CS_T, \mathbf{y}) = p^i(CS_S, \mathbf{x})$. Let $(\mathbf{x}_1, \mathbf{y}_1)$ be a point in $M(CS_S, CS_T)$ that maximizes the number of green agents in T.

Now, consider a green agent g and a white agent i. Suppose that

(a) $g \in S$, $i \in T$, and there is a coalition $\mathbf{c} \in CS_T$ such that $g, i \in supp(\mathbf{c})$ and $y(\mathbf{c})^g > 0$, or

(b) $g, i \in S$, and there is a coalition $\mathbf{c} \in CS_S$ such that $g, i \in supp(\mathbf{c})$ and $x(\mathbf{c})^i > 0$.

In case (a) we could transfer a small amount of payoff from g to i in CS_T so as to either make i green and thereby increase the number of green vertices (if $i \in S$) or lower the total payoff of S in CS_T (if $i \in T \setminus S$), and in case (b) we could transfer a small amount of payoff from i to g in CS_S so as to make i green (again, we choose the transfer to be small enough that g remains green). In both cases, we get a contradiction with our choice of $(\mathbf{x}_1, \mathbf{y}_1)$.

Let S_g be the set of all green agents in S, and let $S_w = S \setminus S_g$ be the set of all white agents in S. We have shown that in $(CS_S, \mathbf{x}_1)$ the members of S_w do not receive payoffs from their joint coalitions with members of S_g, and in $(CS_T, \mathbf{y}_1)$ the members of S_g do not receive payoffs from their joint coalitions with members of S_w, so we have $p^{S_w}(CS_S, \mathbf{x}_1) \leq v^*(\mathbf{e}^{S_w})$ and $p^{S_g}(CS_T, \mathbf{y}_1) \leq v^*(\mathbf{e}^{S_g})$. Thus, if $S_g \neq \emptyset$, we have

$$\begin{aligned} v^*(\mathbf{e}^S) &= p^S(CS_S, \mathbf{x}_1) \\ &= p^{S_g}(CS_S, \mathbf{x}_1) + p^{S_w}(CS_S, \mathbf{x}_1) \\ &< p^{S_g}(CS_T, \mathbf{y}_1) + v^*(\mathbf{e}^{S_w}) \\ &\leq v^*(\mathbf{e}^{S_g}) + v^*(\mathbf{e}^{S_w}). \end{aligned} \tag{9}$$

Thus, if $S_g \neq \emptyset$, then $v^*(\mathbf{e}^S) < v^*(\mathbf{e}^{S_g}) + v^*(\mathbf{e}^{S_w})$, a contradiction with the superadditivity of v^*. We conclude that $S_g = \emptyset$, which implies that $p^i(CS_S, \mathbf{x}_1) = p^i(CS_T, \mathbf{y}_1)$ for all $i \in S$. □

Armed with Proposition 5.10, we are now ready to prove the following theorem.

Theorem 5.12. *If $\mathcal{G} = \langle N, v \rangle$ is OCF-convex then $\bar{\mathcal{G}} = \langle N, U_v \rangle$ is supermodular.*

Proof. Consider sets $S, T, R \subseteq N$ such that $S \subseteq T$ and $R \subseteq N \setminus T$; we will demonstrate that $U_v(S \cup R) - U_v(S) \leq U_v(T \cup R) - U_v(T)$. Set $S' = S \cup R$, $T' = T \cup R$, and consider coalition structures $CS_S \in \mathcal{CS}(S)$, $CS_{S'} \in \mathcal{CS}(S')$ such that $v(CS_S) = v^*(\mathbf{e}^S)$, $v(CS_{S'}) = v^*(\mathbf{e}^{S'})$. By

Proposition 5.10, there exist imputations $\mathbf{x}_S$ and $\mathbf{x}_{S'}$ such that $p^i(CS_S, \mathbf{x}_S) = p^i(CS_{S'}, \mathbf{x}_{S'})$ for all $i \in S$; thus, the total payoff to R from $(CS_{S'}, \mathbf{x}_{S'})$ is $v^*(\mathbf{e}^{S\cup R}) - v^*(\mathbf{e}^S)$.

Consider an outcome $(CS_T, \mathbf{x}_T) \in \mathcal{F}_T$ such that $v(CS_T) = v^*(\mathbf{e}^T)$; since $\mathcal{G}$ is OCF-convex, there is an outcome $(CS_{T'}, \mathbf{x}_{T'})$ that is better for all members of S' than $(CS_{S'}, \mathbf{x}_{S'})$, and also pays T a total of at least $v^*(\mathbf{e}^T)$. The payoff to R under $(CS_{T'}, \mathbf{x}_{T'})$ is $v(CS_{T'}) - p^T(CS_{T'}, \mathbf{x}_{T'})$. We have $v(CS_{T'}) \leq v^*(\mathbf{e}^{T\cup R})$ and $p^T(CS_{T'}, \mathbf{x}_{T'}) \geq v^*(\mathbf{e}^T)$, so the payoff to R in $(CS_{T'}, \mathbf{x}_{T'})$ is at most $v^*(\mathbf{e}^{T\cup R}) - v^*(\mathbf{e}^T)$. Finally, OCF convexity of $\mathcal{G}$ implies that the payoff to R in $(CS_{T'}, \mathbf{x}_{T'})$ is at least as large as its payoff in $(CS_{S'}, \mathbf{x}_{S'})$. Combining these inequalities, we get

$$\begin{aligned} U_v(S\cup R) - U_v(S) &= v^*(\mathbf{e}^{S\cup R}) - v^*(\mathbf{e}^S) \\ &= p^R(CS_{S'}, \mathbf{x}_{S'}) \\ &\leq p^R(CS_{T'}, \mathbf{x}_{T'}) \\ &\leq v^*(\mathbf{e}^{T\cup R}) - v^*(\mathbf{e}^T) \\ &= U_v(T\cup R) - U_v(T), \end{aligned}$$

which concludes the proof. □

Theorem 5.12 shows that if an OCF game is OCF-convex, then its discrete superadditive cover is supermodular. We will now show that the converse does not always hold, i.e., supermodularity is a strictly weaker property than OCF convexity.

Example 5.13. Consider a game $\mathcal{G} = \langle N, v\rangle$ where $N = \{1, 2, 3\}$ and v is defined as follows:

- $v\begin{pmatrix}1\\0\\0\end{pmatrix} = 1$, $v\begin{pmatrix}0\\1\\0\end{pmatrix} = 1$;
- $v\begin{pmatrix}1\\.5\\0\end{pmatrix} = 6$, $v\begin{pmatrix}0\\.5\\1\end{pmatrix} = 3$;
- $v\begin{pmatrix}1\\0\\1\end{pmatrix} = 4$, $v\begin{pmatrix}0\\1\\1\end{pmatrix} = 4$;
- $v(\mathbf{c}) = 0$ for any other partial coalition $\mathbf{c} \in [0, 1]^3$.

We have

- $U_v(\{1\}) = U_v(\{2\}) = 1$, $U_v(\{3\}) = 0$;
- $U_v(\{1, 2\}) = 6$, $U_v(\{1, 3\}) = U_v(\{2, 3\}) = 4$;
- $U_v(N) = 9$.

One can check that U_v is indeed supermodular. However, $\mathcal{G}$ is not OCF-convex. Indeed, set $S = \{3\}$, $T = \{1, 3\}$, $R = \{2\}$, with $CS_S = \left(\begin{pmatrix}0\\0\\1\end{pmatrix}\right)$, $CS_T = \left(\begin{pmatrix}1\\0\\1\end{pmatrix}\right)$, $CS_{S\cup R} = \left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right)$. Assume that the players in T and $S\cup R$ share the payoffs according to $\mathbf{x}_T = ((0, 0, 4))$ and $\mathbf{x}_{S\cup R} = ((0, 4, 0))$, respectively. For $\mathcal{G}$ to be OCF-convex, there has to exist a coalition structure CS where player 3 earns at least 4 and player 2 earns at least 4. However, this is impossible: if player 3 earns at least 4, then CS contains either $\mathbf{c} = \begin{pmatrix}1\\0\\1\end{pmatrix}$ or $\mathbf{c}' = \begin{pmatrix}0\\1\\1\end{pmatrix}$. If $\mathbf{c}$ is formed, then player 2 can get at most 1; if $\mathbf{c}'$ is formed, then players 2 and 3 together can get at most 4. Thus, there is no way to satisfy all of the agents' demands.

5.2 The Sensitive Core

Recall that under the sensitive arbitration function, an agent will withhold payments from the deviators if any of the coalitions it participates in are hurt by the deviation. In Section 4.1, we obtained the following characterization of the sensitive core: an outcome $(CS, \mathbf{x})$ is in the sensitive core if and only if CS is an optimal coalition structure, and for each set $S \subseteq N$ it holds that for each $T \subseteq N \setminus S$ the total payoff that S receives from investing its resources in coalitions involving T is at least as large as the marginal returns of investing the same resources in working on its own (see formula (4)).

Let $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ be an optimal coalition structure. Using the characterization above, we conclude that deciding whether the sensitive core contains an outcome of the form $(CS, \mathbf{x})$ is equivalent to determining whether the value of the following linear program equals $v^*(e^N)$.

$$\begin{aligned} \min \quad & \sum_{j=1}^{m} \sum_{i \in supp(\mathbf{c}_j)} x_j^i && (10)\\ \text{s.t.} \quad & \sum_{i \in supp(\mathbf{c}_j)} x_j^i \geq v(\mathbf{c}_j) && \forall \mathbf{c}_j \in CS \\ & p^S(CS_T, \mathbf{x}) \geq v^*(\mathbf{w}^S(CS|_S) + \mathbf{w}^S(CS_T)) - v^*(\mathbf{w}^S(CS|_S)) && \forall S \subseteq N, \forall T \subseteq N \setminus S \end{aligned}$$

We note that in the linear program (10) we do not require that $x_j^i \geq 0$ for each $\mathbf{c}_j \in CS$ and each $i \in supp(\mathbf{c}_j)$. This is because the argument presented in Appendix A shows that these constraints can be safely omitted. We will revisit this point in the proof of Theorem 5.15 in the context of the refined core.

Consider the dual of this linear program.

$$\begin{aligned} \max \quad & \sum_{j=1}^{m} r_j v(\mathbf{c}_j) + \sum_{S \subseteq N} \sum_{T \subseteq N \setminus S} \delta_{S,T}(v^*(\mathbf{w}^S(CS|_S) + \mathbf{w}^S(CS_T)) - v^*(\mathbf{w}^S(CS|_S))) && (11)\\ \text{s.t.} \quad & r_j + \sum_{S : i \in S} \sum_{\substack{T \subseteq N \setminus S: \\ supp(\mathbf{c}_j) \cap T \neq \emptyset}} \delta_{S,T} = 1 && \forall \mathbf{c}_j \in CS, \forall i \in supp(\mathbf{c}_j) \\ & r_j \geq 0 && \forall \mathbf{c}_j \in CS \\ & \delta_{S,T} \geq 0 && \forall S \subseteq N, \forall T \subseteq N \setminus S \end{aligned}$$

We say that a collection of non-negative weights $\{(r_j)_{j=1}^m; (\delta_{S,T})_{S \subseteq N; T \subseteq N \setminus S}\}$ is *s-balanced with respect to* $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ if $r_j + \sum_{S : i \in S} \sum_{T \subseteq N \setminus S : supp(\mathbf{c}_j) \cap T \neq \emptyset} \delta_{S,T} = 1$ for all $\mathbf{c}_j \in CS$ and all $i \in supp(\mathbf{c}_j)$. Applying linear programming duality to linear programs (10) and (11), we obtain the following theorem.

Theorem 5.14. *The sensitive core of a game $\mathcal{G} = \langle N, v \rangle$ is not empty if and only if there exists a coalition structure $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ with $v(CS) = v^*(\mathbf{e}^N)$ such that for every collection of non-negative weights $\{(r_j)_{j=1}^m; (\delta_{S,T})_{S \subseteq N; T \subseteq N \setminus S}\}$ that is s-balanced with respect to CS it holds that*

$$\sum_{j=1}^{m} r_j v(\mathbf{c}_j) + \sum_{S \subseteq N} \sum_{T \subseteq N \setminus S} \delta_{S,T}(v^*(\mathbf{w}^S(CS|_S) + \mathbf{w}^S(CS_T)) - v^*(\mathbf{w}^S(CS|_S))) \leq v^*(\mathbf{e}^N).$$

Theorem 5.14 is similar to Theorem 5.2, which presents a criterion for non-emptiness of the conservative core. The main difference is that Theorem 5.2 considers collections of weights that contain a single weight δ_S for each set $S \subseteq N$, whereas for the sensitive core we consider collections of weights that contain a weight $\delta_{S,T}$ for each pair of disjoint sets $S, T \subseteq N$.

5.3 The Refined Core

Under the sensitive arbitration function, the deviating players only need to decide which of the non-deviating players they do not wish to work with any more. In contrast, under the refined arbitration function, the deviators have to look at the coalitions with the non-deviators one by one, and decide which of them are worth keeping. We will now provide a characterization of games with non-empty refined core that is similar to the characterizations of games with non-empty conservative and sensitive cores (Theorem 5.2 and Theorem 5.14, respectively). Employing this characterization, we then describe a class of OCF games with a non-empty refined core.

In Section 4.1, we obtained the following characterization of outcomes in the refined core: given a coalition structure CS, there is an imputation $\mathbf{x} \in I(CS)$ such that $(CS, \mathbf{x})$ is in the refined core if and only if $p^S(CS', \mathbf{x}) \geq v^*(\mathbf{w}^S(CS'))$ for every $S \subseteq N$ and every coalition structure $CS' \subseteq CS$ containing $CS|_S$.

In what follows, given a subset of agents $S \subseteq N$ and a coalition structure CS, we write $[CS]_S = \{CS' \subseteq CS \mid CS|_S \subseteq CS'\}$. Also, given a coalition structure $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$, we say that a collection of non-negative weights $\{(r_j)_{j=1}^m; (\delta_{S,CS'})_{S\subseteq N; CS' \in [CS]_S}\}$ is *r-balanced with respect to* CS if for each $\mathbf{c}_j \in CS$ and each $i \in supp(\mathbf{c}_j)$ it holds that

$$r_j + \sum_{S:i\in S} \sum_{\substack{CS' \in [CS]_S: \\ \mathbf{c}_j \in CS'}} \delta_{S,CS'} = 1.$$

We are now ready to present our characterization.

Theorem 5.15. *The refined core of an OCF game $\mathcal{G} = \langle N, v \rangle$ is non-empty if and only if there exists a coalition structure $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ with $v(CS) = v^*(\mathbf{e}^N)$ such that for every collection of weights $\{(r_j)_{j=1}^m; (\delta_{S,CS'})_{S\subseteq N; CS' \in [CS]_S}\}$ that is r-balanced with respect to CS it holds that*

$$\sum_{j=1}^m r_j v(\mathbf{c}_j) + \sum_{S\subseteq N} \sum_{CS' \in [CS]_S} \delta_{S,CS'} v^*(\mathbf{w}^S(CS')) \leq v^*(\mathbf{e}^N).$$

Proof. Fix a coalition structure $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ with $v(CS) = v^*(\mathbf{e}^N)$, and consider the following linear program.

$$\begin{aligned} \min \quad & \sum_{j=1}^m \sum_{i\in supp(\mathbf{c}_j)} x_j^i & & (12) \\ \text{s.t.} \quad & \sum_{i \in supp(\mathbf{c}_j)} x_j^i \geq v(\mathbf{c}_j) & \forall \mathbf{c}_j \in CS & \\ & p^S(CS', \mathbf{x}) \geq v^*(\mathbf{w}^S(CS')) & \forall S \subseteq N, \forall CS' \in [CS]_S & \end{aligned}$$

We claim that the refined core of $\mathcal{G}$ contains an outcome of the form $(CS, \mathbf{x})$ if and only if the value of this linear program is $v^*(e^N)$.

To see why this is the case, it suffices to observe that, by the argument in Appendix A, we can omit constraints of the form $x^i_j \geq 0$ for each $\mathbf{c}_j \in CS$ and each $i \in supp(\mathbf{c}_j)$. In fact, it turns out that for the refined core, a much simpler explanation holds: Non-negativity of x^i_j is implied by the stability constraints. To see this, consider a coalition $\mathbf{c}_j$ such that $i \in supp(\mathbf{c}_j)$. If $supp(\mathbf{c}_j) = \{i\}$, $x^i_j = v(\mathbf{c}_j) \geq 0$ and we are done. Otherwise, consider the coalition structure $CS' = (CS|_{\{i\}}, \mathbf{c}_j)$. The total payoff to i from CS' is $p^i(CS', \mathbf{x}) = p^i(CS|_{\{i\}}, \mathbf{x}) + x^i_j = v(CS|_{\{i\}}) + x^i_j$. The constraint that corresponds to $S = \{i\}$ and CS' states that this payoff must be at least $v^*(\mathbf{w}^{\{i\}}(CS'))$, which is at least $v(CS|_{\{i\}})$, and hence $x^i_j \geq 0$.

Consider now the dual of linear program (12).

$$\begin{aligned}
\max \quad & \sum_{j=1}^{m} r_j v(\mathbf{c}_j) + \sum_{\substack{S \subseteq N \\ CS' \in [CS]_S}} \delta_{S,CS'} v^*(\mathbf{w}^S(CS')) && (13)\\
\text{s.t.} \quad & r_j + \sum_{\substack{S: i \in S \\ CS' \in [CS]_S : \mathbf{c}_j \in CS'}} \delta_{S,CS'} = 1 && \forall \mathbf{c}_j \in CS, \forall i \in supp(\mathbf{c}_j)\\
& \delta_{S,CS'} \geq 0 && \forall S \subseteq N, \forall CS' \in [CS]_S\\
& r_j \geq 0 && \forall \mathbf{c}_j \in CS
\end{aligned}$$

Observe that the dual constraints in (13) are equalities since x^i_j are unconstrained in (12). Note also that every feasible solution of (13) corresponds to a collection of non-negative weights that is r-balanced with respect to CS. Our claim now follows by the standard linear programming duality argument. □

Theorem 5.15 enables us to identify an interesting class of OCF games with non-empty refined core. Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is *homogeneous of degree* k, or k*-homogeneous*, if $f(\alpha \mathbf{x}) = \alpha^k f(\mathbf{x})$ for all $\alpha \in \mathbb{R}_+$. Intuitively, this means that f scales in a consistent manner: if all players invest twice as much resources in a coalition, they will get 2^k times the profit. The returns to scale are positive if $k \geq 1$ and negative if $k < 1$.

Corollary 5.16. *Consider an OCF game $\mathcal{G} = \langle N, v \rangle$. If v^* is homogeneous of degree $k \geq 1$, then the refined core of $\mathcal{G}$ is non-empty.*

Proof. Consider a coalition structure $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ with $v(CS) = v^*(e^N)$ and a collection of non-negative weights $\{(r_j)_{j=1}^m; (\delta_{S,CS'})_{S \subseteq N; CS' \in [CS]_S}\}$ that is r-balanced with respect to CS. According to Theorem 5.15, it suffices to show that

$$\sum_{j=1}^{m} r_j v(\mathbf{c}_j) + \sum_{\substack{S \subseteq N \\ CS' \in [CS]_S}} \delta_{S,CS'} v^*(\mathbf{w}^S(CS')) \leq v^*(\mathbf{e}^N). \tag{14}$$

First, since v^* is k-homogeneous, we have

$$\sum_{\substack{S \subseteq N \\ CS' \in [CS]_S}} \delta_{S,CS'} v^*(\mathbf{w}^S(CS')) \leq v^* \left(\sum_{\substack{S \subseteq N \\ CS' \in [CS]_S}} \delta_{S,CS'} \mathbf{w}^S(CS') \right). \tag{15}$$

Denote the i-th coordinate of $\sum_{\substack{S\subseteq N\\ CS'\in[CS]_S}} \delta_{S,CS'}\mathbf{w}^S(CS')$ by ω^i. We have

$$\begin{aligned}
\omega^i &= \sum_{S:i\in S}\sum_{CS'\in[CS]_S}\delta_{S,CS'}\sum_{\mathbf{c}_j\in CS'}c_j^i\\
&= \sum_{j=1}^m c_j^i\sum_{S:i\in S}\sum_{\substack{CS'\in[CS]_S:\\ \mathbf{c}_j\in CS'}}\delta_{S,CS'}\\
&= \sum_{j=1}^m c_j^i(1-r_j),
\end{aligned}$$

where the first transition is derived from the fact that $\mathbf{w}^i(CS') = \sum_{\mathbf{c}_j\in CS'} c_j^i$ if $i \in S$, and the second transition uses the fact that $\{(r_j)_{j=1}^m; (\delta_{S,CS'})_{S\subseteq N; CS'\in[CS]_S}\}$ is r-balanced with respect to CS. We conclude that the right-hand side of equation (15) is upper-bounded by

$$v^*(\sum_{j=1}^m(1-r_j)\mathbf{c}_j).$$

We now consider the first summand in the left-hand side of (14). By k-homogeneity of v^*, we have

$$\sum_{j=1}^m r_j v(\mathbf{c}_j) \le \sum_{j=1}^m r_j v^*(\mathbf{c}_j) \le v^*(\sum_{j=1}^m r_j\mathbf{c}_j).$$

Using the superadditivity of v^*, we obtain that the left-hand side of (14) can be upper-bounded by

$$v^*(\sum_{j=1}^m r_j\mathbf{c}_j + \sum_{j=1}^m(1-r_j)\mathbf{c}_j) = v^*(\sum_{j=1}^m \mathbf{c}_j) = v^*(\mathbf{e}^N).$$

This concludes the proof. □

In fact, the proof of Corollary 5.16 shows a stronger claim: if v^* is k-homogeneous, then every optimal coalition structure CS admits an imputation $\mathbf{x} \in I(CS)$ such that $(CS, \mathbf{x})$ is in the refined core. This is not the case in general; there exist games where some optimal coalition structures cannot be stabilized under the refined arbitration function, while others can. This is illustrated by the following example.

Example 5.17. Consider the following three-player game. There are four types of tasks: A task of type t_1 can be completed by player 1 alone, requires all of its resources, and is worth 5. A task of type t_{12} requires 50% of both player 2 and player 1's resources and is worth 10. A task of type T_{12} requires all of the resources of players 1 and 2 and is worth 20. Finally, a task of type t_{23} requires all of player 3's resources and 50% of player 2's resources, and is worth 9. Consider coalition structures $CS = (\mathbf{c}_1, \mathbf{c}_2)$ and $CS' = (\mathbf{c}'_1)$, where

$$\mathbf{c}_1 = \begin{pmatrix}.5\\.5\\0\end{pmatrix}, \mathbf{c}_2 = \begin{pmatrix}.5\\.5\\0\end{pmatrix}, \mathbf{c}'_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}.$$

It is easy to see that both CS and CS' are optimal for this game. Simply put, it is best for players 1 and 2 to work together and earn a total of 20 (by completing t_{12} twice or T_{12} once), while player 3 makes no profit.

First, we claim that CS cannot be stabilized with respect to the refined arbitration function. The reason is that for an outcome $(CS, \mathbf{x})$ to be in the refined core, it must be the case that player 2 gets at least 9 from each of the coalitions $\mathbf{c}_1$ and $\mathbf{c}_2$. However, this means that player 1 gets at most 2 from working with player 2, while he can get 5 by working alone. On the other hand, let $\mathbf{y}$ be the imputation for CS' that splits the payoff from T_{12} evenly between players 1 and 2. Then $(CS', \mathbf{y})$ is in the refined core.

Finally, we note that one cannot immediately employ an LP duality argument for characterizing OCF games with non-empty optimistic core. The difficulty is that under the optimistic arbitration function the number of possible deviations is infinite: a deviating set needs to specify the *amount* of resources it withdraws from each coalition with non-deviators. In other words, the linear program that describes the optimistic core has infinitely many constraints, so its dual has infinitely many variables.

6. Conclusions and Future Work

The main contribution of our work is the concept of arbitration functions, and the analysis of the impact of arbitration on the stability of cooperative games with overlapping coalitions. This concept allows us to put the three notions of stability proposed by Chalkiadakis et al. (2010) in a broader context and to derive new notions of stability such as the sensitive core.

Perhaps the most interesting of our results is the connection between OCF games with the conservative arbitration function and their discrete superadditive covers. In a sense, it shows that if the players are so pessimistic as to assume that their post-deviation payoffs will be given by the conservative arbitration function, there is little benefit to employing the OCF framework. In other words, the main value of the OCF approach lies in its ability to model non-trivial post-deviation interactions. Thus, the OCF framework is eminently suitable for settings where agents have a reason to expect that deviating from some of their collaborative projects will not cause other agents to boycott them. We believe that such settings are becoming increasingly more common in the modern interconnected society, where multitasking is the norm and collaboration is necessary to succeed.

The reaction of non-deviators to deviation plays a decisive role in the analysis of many strategic interactions. For example, Ackerman and Brânzei (2014) use the concept of arbitration function in the analysis of Nash equilibrium and pairwise equilibrium. Also, Brânzei, Michalak, Rahwan, Larson, and Jennings (2013) study matchings with externalities; their model can be viewed as an application of the idea of arbitration functions to matching problems. In fact, in any strategic setting where a deviating set must still interact with the non-deviators one has to reason about the non-deviators' behavior after a deviation occurs. While it is common to make the worst-case assumption, this is not always appropriate or realistic. Thus, a notion similar to arbitration functions may be useful beyond the setting of cooperative games. For instance, when players interact in a market, their behavior is governed by contracts (think of a wireless service provider and a consumer buying a new data plan), which clearly specify penalties for failing to fulfill one's obligations (in the case of wireless service providers, it is typically the case that a withdrawal from an agreed-upon contract entails a monetary penalty, i.e. a worse-than-conservative reaction to deviation). Such contracts can be viewed as arbitration functions. Thus, the concept of arbitration function may be useful for modeling market interactions under contracts.

6.1 Future Work

In Section 5, we have provided criteria for core non-emptiness under several arbitration functions. For the conservative and refined core, we have also identified simple sufficient conditions for the core to be non-empty (supermodularity of the discrete superadditive cover and first-order homogeneity of v^*, respectively). It would be useful to obtain similar results for other natural notions of arbitration, such as the sensitive arbitration function and the optimistic arbitration function.

Another interesting research direction is to consider the "inverse" of this problem: given an OCF game, what is the "most generous" arbitration function for which the arbitrated core of that game is not empty? More formally, given a game $\mathcal{G}$, let us call an arbitration function $\mathcal{A}$ *max-stable* for $\mathcal{G}$ if $Core(\mathcal{G}, \mathcal{A}) \neq \emptyset$, and for all $\mathcal{A}'$ such that $\mathcal{A}' \geq \mathcal{A}$ we have $Core(\mathcal{G}, \mathcal{A}') = \emptyset$. Simply put, the arbitration function $\mathcal{A}$ is max-stable if it provides the largest possible payoffs to the deviators without destabilizing $\mathcal{G}$. It would be interesting to prove that some arbitration function is max-stable for a class of OCF games, or, more ambitiously, to completely characterize the set of max-stable arbitration functions. This line of enquiry is similar in spirit to the concepts of the least core (Maschler et al., 1979), the cost of stability (Bejan & Gómez, 2009; Bachrach, Elkind, Meir, Pasechnik, Zuckerman, Rothe, & Rosenschein, 2009), and taxation (Zick, Polukarov, & Jennings, 2013).

We note that in our model deviators are allowed to withdraw resources from coalitions with non-deviators, but not to add resources to such coalitions. The implications of a more general definition, where deviators can both add and withdraw resources to their coalitions with non-deviators, are worth exploring as well.

Our analysis focuses on the core of cooperative games with overlapping coalitions. This is, to an extent, a reflection of the state of the art in the study of classic cooperative games: the core and related notions (i.e., the ε-core and the least core) have been studied in greater detail than other cooperative solution concepts (barring, perhaps, the Shapley value). Further analysis of the various solution concepts for OCF games, such as the nucleolus, the kernel, and the bargaining set, and their relations to one another, would greatly improve our understanding of overlapping coalition formation. While some of these solution concepts for OCF games have been briefly discussed by Zick and Elkind (2011), their analysis is far from complete.

Finally, it would be interesting to study models of dynamic coalition formation in cooperative games with overlapping coalitions. While such models have been explored for classic cooperative games (see, e.g., Arnold & Schwalbe, 2002; Lehrer & Scarsini, 2013), or for games with overlapping coalitions and fully cooperative agents (Shehory & Kraus, 1996), it is not immediately clear how to extend prior work to arbitrated OCF games. In such settings, arbitration functions can be used as a method of controlling the coalition formation process: If an outcome at some stage of the coalition formation process is stable with respect to the currently used arbitration function, more permissive reaction to deviation is allowed in future rounds; if the outcome is unstable, players can be assumed to be less tolerant towards deviators. Thus, by controlling the extent to which players are allowed to deviate, one can trade off the speed of the coalition formation process and the quality of the resulting outcome.

Acknowledgments

Part of this work was done while Y. Zick and E. Elkind were affiliated with Nanyang Technological University, Singapore. Y. Zick was supported by the Singapore International Graduate Award (SINGA), provided by the Agency for Science, Technology and Research (A*STAR). E. Markakis was supported by the European Union (European Social Fund--ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF)—Research Funding Program: THALES. Investing in knowledge society through the European Social Fund. E. Elkind was supported by the National Research Foundation (Singapore) under grant NRF RF2009-08. The authors would like to thank the anonymous JAIR reviewers as well as the anonymous reviewers of earlier versions of this work (Zick & Elkind, 2011; Zick, Markakis, & Elkind, 2012). Y. Zick would also like to thank the anonymous reviewers of his Ph.D thesis (Zick, 2014), parts of which are incorporated into this paper.

Appendix A. A Note on the Characterization of the Conservative Core

The purpose of this appendix is to fill a gap in the proof of the characterization of the conservative core that was given by Chalkiadakis et al. (2010) (see Theorem 5.2). Our argument also has implications for the characterization of the sensitive core (Section 5.2).

The proof of Theorem 5.2 proceeds as follows. Given an optimal coalition structure $CS = (\mathbf{c}_1, \dots, \mathbf{c}_m)$, consider the following linear program:

$$\begin{array}{lll} \min & \sum_{j=1}^{m} \sum_{i \in supp(\mathbf{c}_j)} x_j^i & \\ \text{s.t.} & \sum_{i \in supp(\mathbf{c}_j)} x_j^i \geq v(\mathbf{c}_j) & \forall j \in \{1, \dots, m\} \\ & \sum_{i \in S} \sum_{j=1}^{m} x_j^i \geq v^*(\mathbf{e}^S) & \forall S \subseteq N \end{array} \tag{16}$$

The dual of LP (16) is

$$\begin{array}{lll} \max & \sum_{j=1}^{m} r_j v(\mathbf{c}_j) + \sum_{S \subseteq N} \delta_S v^*(\mathbf{e}^S) & \\ \text{s.t.} & r_j + \sum_{S: i \in S} \delta_S = 1 & \forall j, \forall i \in supp(\mathbf{c}_j) \\ & r_j \geq 0 & \forall j \in \{1, \dots, m\} \\ & \delta_S \geq 0 & \forall S \subseteq N \end{array} \tag{17}$$

Chalkiadakis et al. (2010) argue that LP (16) describes the constraints of the conservative core, with the exception that $\sum_{i \in supp(\mathbf{c}_j)} x_j^i$ is required to be at least $v(\mathbf{c}_j)$ rather than equal to $v(\mathbf{c}_j)$. Thus, if $\bar{\mathbf{x}}$ is an optimal solution to LP (16), then $(CS, \bar{\mathbf{x}})$ is in the conservative core if and only if $\sum_{j=1}^{m} \sum_{i \in supp(\mathbf{c}_j)} x_j^i = v(CS) = v^*(\mathbf{e}^N)$. LP duality then implies that the value of LP (17) is at most $v^*(\mathbf{e}^N)$ if and only if CS can be stabilized with respect to the conservative arbitration function.

The problem with this argument is that LP (16) does not require the variables x_j^i to be non-negative, which allows us to have $r_j + \sum_{S: i \in S} \delta_S = 1$ rather than $r_j + \sum_{S: i \in S} \delta_S \leq 1$ in the dual LP. However, if we allow the variables x_j^i to take negative values, we are effectively breaking the payoff distribution requirement for imputations. We mention that, in contrast, in the Bondareva–Shapley theorem it is acceptable not to impose non-negativity constraints explicitly, since stability constraints imply $p_i \geq u(\{i\}) \geq 0$. However, for the conservative core we only get

$\sum_{j:i\in supp(\mathbf{c}_j)} x_j^i \geq v^*(\mathbf{e}^{\{i\}})$ for all $i \in N$, which does not imply that individual x_j^i are non-negative. In other words, Chalkiadakis et al. (2010) show that every c-balanced OCF game and every optimal coalition structure $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ for this game admit a *pre-imputation* $\mathbf{x}$ such that $(CS, \mathbf{x})$ is in the conservative core (recall that $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$ is a pre-imputation for $CS = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ if for each $j = 1, \ldots, m$ it holds that $\mathbf{x}_j$ is a vector in $\mathbb{R}^n$ such that $\sum_{i=1}^n x_j^i = v(\mathbf{c}_j)$ and if $c_j^i = 0$ then $x_j^i = 0$, i.e., a preimputation satisfies the same conditions as an imputation, except for non-negativity). We will now argue that one can transform this pre-imputation into an imputation.

Theorem A.1. *If there exists a pre-imputation* $\mathbf{x}$ *that satisfies the constraints of LP* (16)*, then there exists an imputation* $\bar{\mathbf{x}} \in I(CS)$ *such that* $p_i(CS, \mathbf{x}) = p_i(CS, \bar{\mathbf{x}})$ *for all* $i \in N$.

Proof. Let $I_{pre}(CS)$ denote the set of all pre-imputations for a coalition structure CS. Given a point $\mathbf{x} \in I_{pre}(CS)$, we define a graph $\Gamma(\mathbf{x}) = \langle V, E \rangle$ as follows: the vertices of $\Gamma(\mathbf{x})$ are

$$V = \{(i, j) \mid \mathbf{c}_j \in CS, i \in supp(\mathbf{c}_j)\},$$

and the set of edges is $E = E_1 \cup E_2$, where

$$E_1 = \left\{\big((i, j), (i, j')\big) \mid i \in supp(\mathbf{c}_j) \cap supp(\mathbf{c}_{j'}), j \neq j'\right\},$$

and

$$E_2 = \left\{\big((i, j), (i', j)\big) \mid i, i' \in supp(\mathbf{c}_j), x_j^i > 0\right\}.$$

That is, E_1 consists of all pairs $(i, j), (i, j')$ such that i is in both $supp(\mathbf{c}_j)$ and $supp(\mathbf{c}_{j'})$, and E_2 consists of all pairs $(i, j), (i', j)$ such that $i, i' \in supp(\mathbf{c}_j)$ for some $\mathbf{c}_j \in CS$ and $x_j^i > 0$. An edge in E_2 means that i can transfer a small amount of payoff to i' in the coalition $\mathbf{c}_j$, and still maintain a positive payoff from $\mathbf{c}_j$.

Further, given a point $\mathbf{x} \in I_{pre}(CS)$, set

$$TN(\mathbf{x}) = \sum_{j=1}^{m} \sum_{i=1}^{n} \min\{0, x_j^i\}.$$

Fix an arbitrary point $\mathbf{x} \in I_{pre}(CS)$, and consider the set

$$I_{pre}(CS, \mathbf{x}) = \{\mathbf{y} \in I_{pre}(CS) \mid p_i(CS, \mathbf{y}) = p_i(CS, \mathbf{x}) \text{ for all } i \in N \text{ and } TN(\mathbf{y}) \geq TN(\mathbf{x})\}.$$

The set $I_{pre}(CS, \mathbf{x})$ is compact and TN is continuous, so there exists a point $\bar{\mathbf{x}} \in I_{pre}(CS, \mathbf{x})$ that maximizes TN over $I_{pre}(CS, \mathbf{x})$.

Suppose that there exists a coalition $\mathbf{c}_{j^*}$ and an agent $i^* \in supp(\mathbf{c}_{j^*})$ such that $\bar{x}_{j^*}^{i^*} < 0$. Note that in $\Gamma(\bar{\mathbf{x}})$ there must be an edge in E_1 that leaves (i^*, j^*): otherwise, i^* is only being paid by $\mathbf{c}_{j^*}$, in which case we would have

$$p_{i^*}(CS, \bar{\mathbf{x}}) = \sum_{j=1}^{m} \bar{x}_j^{i^*} = \bar{x}_{j^*}^{i^*} < 0 \leq v^*(e^{\{i^*\}}),$$

violating the respective constraint of LP (16).

Let C be a minimum-length directed cycle in $\Gamma(\bar{\mathbf{x}})$ that contains an edge $((i^*, j^*), (i^*, j_0))$ for some $c_{j_0} \in CS$ as well as some edges from E_2. Note that C contains no path of the form

$$(i, j_1) \to (i, j_2) \to (i, j_3);$$

this is because the edge $((i, j_1), (i, j_3))$ is in E_1, a contradiction to C being of minimum length. Moreover, C cannot contain a path of the form

$$(i_1, j) \to (i_2, j) \to (i_3, j);$$

indeed, if $((i_1, j), (i_2, j))$ is in E_2, then agent i_1 receives a positive payoff from $\mathbf{c}_j$, and hence the edge $((i_1, j), (i_3, j))$ is also in E_2, a contradiction with our choice of C.

We conclude that a minimum-length cycle that passes through (i^*, j^*) and intersects E_2 must be of the form

$$(i^*, j^*) \to (i^*, j_0) \to (i_1, j_0) \to (i_1, j_1) \to (i_2, j_1) \to \cdots \to (i_t, j^*) \to (i^*, j^*).$$

This means that i^* receives a positive payoff from $\mathbf{c}_{j_0}$, i_1 receives a positive payoff from $\mathbf{c}_{j_1}$, i_2 receives a positive payoff from $\mathbf{c}_{j_2}$ and, in general, i_ℓ receives a positive payoff from the coalition $\mathbf{c}_{j_\ell}$ for all $1 \le \ell \le t-1$. Finally, the player i_t receives a positive payoff from $\mathbf{c}_{j^*}$.

Now, pick an ε satsifying

$$0 < \varepsilon < \min\{\bar{x}^{i^*}_{j_0}, \bar{x}^{i_1}_{j_1}, \ldots, \bar{x}^{i_{t-1}}_{j_{t-1}}, \bar{x}^{i_t}_{j^*}\},$$

and let $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_m)$ be the pre-imputation obtained from $\bar{\mathbf{x}}$ by transferring a payoff of ε from i^* to i_1 in $\mathbf{c}_{j_0}$, from i_ℓ to $i_{\ell+1}$ in $\mathbf{c}_{j_\ell}$ for $\ell = 1, \ldots, t-1$, and from i_t to i^* in $\mathbf{c}_{j^*}$. We have $p_i(CS, \mathbf{y}) = p_i(CS, \bar{\mathbf{x}})$ for all $i \in N$, and $\sum_{i \in supp(\mathbf{c}_j)} y^i_j = \sum_{i \in supp(\mathbf{c}_j)} \bar{x}^i_j$ for all $\mathbf{c}_j \in CS$. Therefore, $\mathbf{y}$ is an optimal solution to LP (16). However, $TN(\mathbf{y}) > TN(\bar{\mathbf{x}})$, a contradiction to our choice of $\bar{\mathbf{x}}$. We conclude that $\Gamma(\bar{\mathbf{x}})$ contains no cycles that pass through (i^*, j^*) and intersect E_2.

Let us define

$$N_p = \{i \in N \mid \text{there is a path from } (i^*, j^*) \text{ to } (i, j) \text{ for some } \mathbf{c}_j\}.$$

Then for every $i \in N_p$ and every $i' \in N \setminus N_p$, our graph contains no edge of the form $((i, j), (i', j))$. Hence, if $i, i' \in supp(\mathbf{c}_j)$ then $\bar{x}^i_j \le 0$. That is, in every coalition that involves both agents in N_p and agents in $N \setminus N_p$ the payoffs are split so that the share of each agent in N_p is non-positive. Hence, $p_{N_p}(CS, \bar{\mathbf{x}}) \le p_{N_p}(CS|_{N_p}, \bar{\mathbf{x}})$.

Moreover, we will now argue that for c_{j^*} the total share of the agents in N_p is, in fact, negative, and therefore $p_{N_p}(CS, \bar{\mathbf{x}}) < p_{N_p}(CS|_{N_p}, \bar{\mathbf{x}})$. Indeed, recall that $\bar{x}^{i^*}_{j^*} < 0$. Suppose that $\bar{x}^i_{j^*} > 0$ for some $i \in N_p$; note that this implies that the edge $e = ((i, j^*), (i^*, j^*))$ is in E_2. Consider a coalition c_j such that the graph contains a path P from (i^*, j^*) to (i, j). If $j = j^*$, by adding e to P we obtain a cycle that passes through (i^*, j^*) and intersects E_2. Otherwise, the edge $e' = ((i, j), (i, j^*))$ is in E_1, and adding e' and e to P results in a cycle that passes through (i^*, j^*) and intersects E_2. In both cases, we get a contradiction, as we have argued that no such cycle exists. Thus, all players in N_p get a non-positive payoff from $\mathbf{c}_{j^*}$, and i^* gets a strictly negative payoff from this coalition. Consequently, we have $p_{N_p}(CS, \bar{\mathbf{x}}) < p_{N_p}(CS|_{N_p}, \bar{\mathbf{x}})$. On the other hand, we have

$$p_{N_p}(CS|_{N_p}, \bar{\mathbf{x}}) = v(CS|_{N_p}) \le v^*(\mathbf{e}^{N_p});$$

combining the inequalities we obtain that $p_{N_p}(CS, \bar{\mathbf{x}}) < v^*(\mathbf{e}^{N_p})$, which is a contradiction to $(CS, \bar{\mathbf{x}})$ satisfying the inequalities in LP (16).

We conclude that $\bar{\mathbf{x}}$ is an imputation of CS such that $p_i(CS, \mathbf{x}) = p_i(CS, \bar{\mathbf{x}})$ for all $i \in N$ and $x^i_j \ge 0$ for all $\mathbf{c}_j \in CS$ and all $i \in supp(\mathbf{c}_j)$, which concludes the proof. $\square$

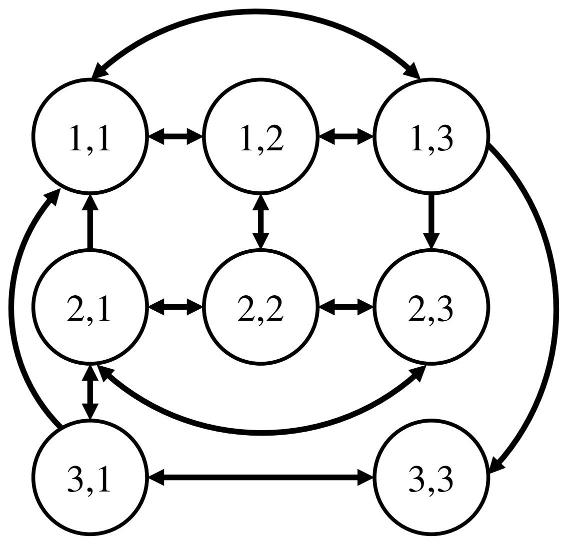

Figure 1: The graph $\Gamma(\mathbf{x})$ formed for the coalition structure and the imputation described in Example A.2.

Example A.2. Suppose the players form the coalition structure $CS = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$ given by

$$\mathbf{c}_1 = \begin{pmatrix} 0.6 \\ 0.3 \\ 0.3 \end{pmatrix}, \qquad \mathbf{c}_2 = \begin{pmatrix} 0.3 \\ 0.6 \\ 0 \end{pmatrix}, \qquad \mathbf{c}_3 = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.7 \end{pmatrix},$$

and use the imputation $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, where

$$\mathbf{x}_1 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}, \qquad \mathbf{x}_2 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \qquad \mathbf{x}_3 = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}.$$

The resulting graph is shown in Figure 1, and the shortest cycle starting from $(3, 3)$ including edges in E_2 is

$$(3, 3) \to (3, 1) \to (1, 1) \to (1, 3) \to (3, 3).$$

That is, the negative payoff to player 3 from $\mathbf{c}_3$ can be reduced if player 3 transfers a small amount of payoff to player 1 under $\mathbf{c}_1$, and player 1 transfers the same amount of payoff to player 3 under $\mathbf{c}_3$; for instance, moving 1 unit of payoff in this manner would result in a valid imputation.

Theorem A.1 implies that even if we assume that x_j^i are unconstrained, if there is a solution to LP (16) with a value of $v^*(\mathbf{e}^N)$, then there is a solution with no negative payoffs from coalitions and the same value, i.e., an optimal solution that is in $I(CS)$. In particular, this means that there is no loss of generality in assuming that the variables x_j^i in LP (16) are unconstrained, and therefore Theorem 5.2 holds.

References

Ackerman, M., & Brânzei, S. (2014). Research quality, fairness, and authorship order. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-14)*, pp. 1487–1488.

Aharoni, R., & Fleiner, T. (2003). On a lemma of Scarf. *Journal of Combinatorial Theory, Series B*, *87*, 72–80.

Airiau, S., & Sen, S. (2009). A fair payoff distribution for myopic rational agents. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-09)*, pp. 1305–1306.

Anshelevich, E., & Hoefer, M. (2010). Contribution games in social networks. In *Proceedings of the 18th European Symposium on Algorithms (ESA-10)*, pp. 158–169.

Arnold, T., & Schwalbe, U. (2002). Dynamic coalition formation and the core. *Journal of Economic Behavior & Organization*, *49*(3), 363–380.

Aubin, J. (1981). Cooperative fuzzy games. *Mathematics of Operations Research*, *6*(1), 1–13.

Aumann, R., & Drèze, J. (1974). Cooperative games with coalition structures. *International Journal of Game Theory*, *3*, 217–237.

Bachrach, Y., Elkind, E., Meir, R., Pasechnik, D., Zuckerman, M., Rothe, J., & Rosenschein, J. (2009). The cost of stability in coalitional games. In *Proceedings of the 2nd International Symposium on Algorithmic Game Theory (SAGT-09)*, pp. 122–134.

Bejan, C., & Gómez, J. C. (2009). Core extensions for non-balanced TU-games. *International Journal of Game Theory*, *38*(1), 3–16.

Bondareva, O. (1963). Some applications of linear programming methods to the theory of cooperative games. *Problemy kibernetiki*, *10*, 119–139.

Brânzei, S., Michalak, T., Rahwan, T., Larson, K., & Jennings, N. R. (2013). Matchings with externalities and attitudes. In *Proceedings of the 12th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-13)*, pp. 295–302.

Chalkiadakis, G., Elkind, E., Markakis, E., Polukarov, M., & Jennings, N. (2010). Cooperative games with overlapping coalitions. *Journal of Artificial Intelligence Research*, *39*, 179–216.

Chalkiadakis, G., Elkind, E., & Wooldridge, M. (2011). *Computational Aspects of Cooperative Game Theory*. Morgan and Claypool.

Dang, V. D., Dash, R. K., Rogers, A., & Jennings, N. R. (2006). Overlapping coalition formation for efficient data fusion in multi-sensor networks. In *Proceedings of the 21st AAAI Conference on AI (AAAI-06)*, pp. 635–640.

Deng, X., Ibaraki, T., & Nagamochi, H. (1999). Algorithmic aspects of the core of combinatorial optimization games. *Mathematics of Operations Research*, *24*(3), 751–766.

Gillies, D. (1953). *Some Theorems on n-Person Games*. Ph.D. thesis, Princeton University.

Jackson, M. O. (2003). A survey of models of network formation—stability and efficiency. In Demange, G., & Wooders, M. (Eds.), *Group Formation in Economics: Networks, Clubs and Coalitions*, chap. 1. Cambridge University Press.

Lehrer, E., & Scarsini, M. (2013). On the core of dynamic cooperative games. *Dynamic Games and Applications*, *3*(3), 359–373.

Lin, C., & Hu, S. (2007). Multi-task overlapping coalition parallel formation algorithm. In *Proceedings of the 6th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-07)*, p. 211.

Markakis, E., & Saberi, A. (2005). On the core of the multicommodity flow game. *Decision Support Systems*, *39*(1), 3–10.

Maschler, M., Peleg, B., & Shapley, L. S. (1979). Geometric properties of the kernel, nucleolus, and related solution concepts. *Mathematics of Operations Research*, *4*(4), 303–338.

Peleg, B., & Sudhölter, P. (2007). *Introduction to the Theory of Cooperative Games* (Second edition)., Vol. 34 of *Theory and Decision Library. Series C: Game Theory, Mathematical Programming and Operations Research*. Springer, Berlin.

Rahwan, T. (2007). *Algorithms for Coalition Formation in Multi-Agent Systems*. Ph.D. thesis, University of Southampton.

Rahwan, T., & Jennings, N. (2008). An improved dynamic programming algorithm for coalition structure generation. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-08)*, pp. 1417–1420.

Shapley, L. S. (1967). On balanced sets and cores. *Naval Research Logistics Quarterly*, *14*(4), 453–460.

Shapley, L. S. (1971). Cores of convex games. *International Journal of Game Theory*, *1*, 11–26.

Shehory, O., & Kraus, S. (1996). Formation of overlapping coalitions for precedence-ordered task-execution among autonomous agents. In *Proceedings of the 2nd International Conference on Multi-Agent Systems (ICMAS-96)*, pp. 330–337.

Sims, M., Corkill, D., & Lesser, V. (2008). Automated organization design for multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, *16*, 151–185.

Wang, T., Song, L., Han, Z., & Saad, W. (2013). Overlapping coalitional games for collaborative sensing in cognitive radio networks. In *Proceedings of 2013 Wireless Communications and Networking Conference (WCNC-13)*, pp. 4118–4123.

Zhang, G., Jiang, J., Su, Z., Qi, M., & Fang, H. (2010). Searching for overlapping coalitions in multiple virtual organizations. *Information Sciences*, *180*, 3140–3156.

Zhang, Z., Song, L., Han, Z., Saad, W., & Lu, Z. (2013). Overlapping coalition formation games for cooperative interference management in small cell networks. In *Proceedings of 2013 Wireless Communications and Networking Conference (WCNC-13)*, pp. 643–648.

Zick, Y. (2014). *Arbitration, Fairness and Stability: Revenue Division in Collaborative Settings*. Ph.D. thesis, Nanyang Technological University.

Zick, Y., Chalkiadakis, G., & Elkind, E. (2012). Overlapping coalition formation games: Charting the tractability frontier. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-12)*, pp. 787–794.

Zick, Y., Chalkiadakis, G., Elkind, E., & Markakis, E. (2014). Cooperative games with overlapping coalitions: Charting the tractability frontier. arXiv 1407:0420.

Zick, Y., & Elkind, E. (2011). Arbitrators in overlapping coalition formation games. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-11)*, pp. 55–62.

Zick, Y., Markakis, E., & Elkind, E. (2012). Stability via convexity and LP-duality in OCF games. In *Proceedings of the 26th AAAI Conference on AI (AAAI-12)*.

Zick, Y., Polukarov, M., & Jennings, N. R. (2013). Taxation and stability in cooperative games. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-13)*, pp. 523–530.