Existence of Stability in Hedonic Coalition Formation Games

Haris Aziz
Department of Informatics
Technische Universität München
85748 Garching bei München, Germany
aziz@in.tum.de

Florian Brandl
Department of Mathematics
Technische Universität München
85748 Garching bei München, Germany
f.brandl@mytum.de

ABSTRACT

In this paper, we examine *hedonic coalition formation games* in which each player's preferences over partitions of players depend only on the members of his coalition. We present three main results in which restrictions on the preferences of the players guarantee the existence of stable partitions for various notions of stability. The preference restrictions pertain to *top responsiveness* and *bottom responsiveness* which model optimistic and pessimistic behavior of players respectively. The existence results apply to natural subclasses of *additively separable hedonic games* and *hedonic games with $\mathcal{B}$-preferences*. It is also shown that our existence results cannot be strengthened to the case of stronger known stability concepts.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity; I.2.11 [**Distributed Artificial Intelligence**]: Multiagent Systems; J.4 [**Computer Applications**]: Social and Behavioral Sciences - Economics

General Terms

Economics, Theory and Algorithms

Keywords

Game theory (cooperative and non-cooperative), teamwork, coalition formation, and coordination

1. INTRODUCTION

In many models of multiagent interaction such as roommate matching and exchange of discrete goods, deviations from one outcome to another can cycle and it may well be possible that no stable outcome is guaranteed. This leads to one of the most fundamental questions in game theory: *what are the necessary and sufficient conditions for the existence of stable outcomes?* This question has been examined extensively by researchers working in market design, multiagent systems, and operations research. We address this question in the context of coalition formation games in which outcomes are partitions of the players. We focus

on *hedonic coalition formation games* in which each player's preferences over partitions depend only on the members of his coalition. Hedonic games are a rich and versatile class of coalition formation games which also encapsulate various stable matching scenarios [see e.g., 5, 6, 7, 14].

In game theory and multiagent systems, understanding the conditions under which systems and social outcomes are guaranteed to be in equilibrium is a fundamental research problem. In this paper, we advance the state of the art on existence results for hedonic games. We strengthen the recently introduced stability concept *strong Nash stability* [15] to *strict strong Nash stability* and show that *top responsiveness* and *mutuality*—conditions different from ones in [15]—are sufficient for the existence of a strictly strong Nash stable partition in any hedonic game. The result applies to natural subclasses of *additively separable hedonic games* [6]. It is also shown that top responsiveness and mutuality together do not guarantee the non-emptiness of the set of *perfect* partitions—a natural concept stronger than strict strong Nash stability.

We then consider a recently introduced property of hedonic games called *bottom refuseness* [17] which we will refer to as *bottom responsiveness*. A new stability notion called *strong individual stability* is formulated which is stronger than both core stability and individual stability. It is shown that bottom responsiveness guarantees the existence of a strong individually stable partition. Also, the combination of *strong bottom responsiveness* and mutuality guarantees the existence of a strong Nash stable partition. Our results concerning bottom responsive games cannot be strengthened to any stronger known stability concept. They also apply to *'aversion to enemies' games* introduced in [11].

Outline.

In Section 2, we present the backdrop of our results and discuss related work. We then introduce hedonic games and the stability concepts considered for these games in Section 3. The relationships between the stability concepts are expounded and clarified in Section 4. We then proceed to Sections 5 and 6 in which the main results are presented. Section 5 concerns hedonic games satisfying top responsiveness whereas in Section 6, existence results concerning bottom responsive games are presented. In Section 7, well-studied subclasses of hedonic games such as additively separable hedonic games and hedonic games with $\mathcal{B}$-preferences are considered and it is shown how existence results apply to these games. Finally, we conclude the discussion in Section 8.

2. RELATED WORK

Identifying sufficient and necessary conditions for the existence of stability in coalition formation has been active area of research. Perhaps the most celebrated result in this field is the existence of a (core) stable matching for the stable marriage problem via the Gale-Shapley algorithm [13]. Later, Banerjee et al. [5] proved that if a hedonic game satisfies a condition called *weak top coalition property*, then the core is non-empty. Banerjee et al. [5] also showed that for various restrictions over preferences, stability is still not guaranteed.

In another important paper, Bogomolnaia and Jackson [6] formalized Nash stability and individual stability in the context of hedonic games and presented a number of sufficient conditions for the existence of various stability concepts. For instance, they showed that *symmetric additively separable preferences* guarantee the existence of a Nash stable partition. A hedonic game is additively separable if each player has a cardinal value for every other player and the player's utility in a partition is the sum of his values for the players in his coalition. The strict core and core is also non-empty for 'appreciation of friends' and 'aversion to enemies' games respectively—two simple classes of additively separable games [11].

Alcalde and Revilla [1] proposed a natural preference restriction called top responsiveness which is based on the idea that players value other players on how they could complement them in research teams. They showed that there exists an algorithm called the *Top Covering Algorithm* which finds a core stable partition for top responsive hedonic games. The Top Covering Algorithm can be seen as a generalization of *Gale's Top Trading Cycle algorithm* [16]. Dimitrov and Sung [9, 10] simplified the Top Covering Algorithm and proved that top responsiveness implies non-emptiness of the strict core and if mutuality is additionally satisfied, then a Nash stable partition exists.

In a follow-up paper, Suzuki and Sung [17] introduced bottom refuseness in an analogous way to top responsiveness. They showed that for hedonic games satisfying bottom refuseness, the *Bottom Avoiding Algorithm* returns a core stable partition. Suzuki and Sung [17] noted that 'appreciation of friends' and 'aversion to enemies' games satisfy top responsiveness and bottom responsiveness respectively, thereby explaining the results in [11].

Very recently, Karakaya [15] proposed a new stability concept called *strong Nash stability* which is stronger than Nash stability and core stability combined. He showed that strong-Nash is non-empty if the *weak top choice property* (stronger than the weak top coalition property) is satisfied or if preferences are 'descending separable'. We will prove three different results in which natural restrictions on the player preferences guarantee the existence of stable partitions where stability is strong Nash stability or its generalization or variant.

3. HEDONIC GAMES & STABILITY CONCEPTS

In this section, we review the terminology, notation, and concepts related to hedonic games.

Hedonic games.

A *hedonic coalition formation game* is a pair $(N, \succsim)$ where N is a set of players and $\succsim$ is a *preference profile* which specifies for each player $i \in N$ the preference relation $\succsim_i$, a reflexive, complete and transitive binary relation on set $\mathcal{N}_i = \{S \subseteq N : i \in S\}$. $S \succ_i T$ denotes that i strictly prefers S over T and $S \sim_i T$ that i is indifferent between coalitions S and T. A *partition* π is a partition of players N into disjoint coalitions. By $\pi(i)$, we denote the coalition in π which includes player i.
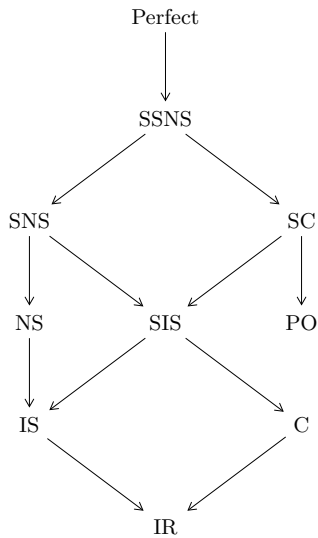
Stability Concepts.

We present the various stability concepts for hedonic games. Nash stability, strict core stability, Pareto optimality, core stability, and individual rationality are classic stability concepts. Individual stability was formulated in [6]. Strong Nash stability was introduced by Karakaya [15] and perfect partitions were considered in [2]. In this paper, we also introduce strict strong Nash stability and strong individual stability which imply strong Nash stability and core stability respectively.

- A partition π is *individually rational (IR)* if no player has an incentive to become alone, i.e., for all $i \in N$, $\pi(i) \succsim_i \{i\}$.

- A partition is *perfect* if each player is in one of his most preferred coalition [2].

- A partition is *Nash stable (NS)* if no player can benefit by moving from his coalition to another (possibly empty) coalition T.

- A partition is *individually stable (IS)* if no player can benefit by moving from his coalition to another existing (possibly empty) coalition T while not making the members of T worse off.

- A coalition $S \subseteq N$ *blocks* a partition π, if each player $i \in S$ strictly prefers S to his current coalition $\pi(i)$ in the partition π. A partition which admits no blocking coalition is said to be in the *core (C)*.

- A coalition $S \subseteq N$ *weakly blocks* a partition π, if each player $i \in S$ weakly prefers S to $\pi(i)$ and there exists at least one player $j \in S$ who strictly prefers S to his current coalition $\pi(j)$. A partition which admits no weakly blocking coalition is in the *strict core (SC)*.

- A partition π is *Pareto optimal (PO)* if there is no partition π' with $\pi'(j) \succsim_j \pi(j)$ for all players j and $\pi'(i) \succ_i \pi(i)$ for at least one player i.

- For partition π, $\pi' \neq \pi$ is called *reachable* from π by movements of players $H \subseteq N$, denoted by $\pi \xrightarrow{H} \pi'$, if $\forall i, j \in N \setminus H, i \neq j : \pi(i) = \pi(j) \Leftrightarrow \pi'(i) = \pi'(j)$.

  A subset of players $H \subseteq N, H \neq \emptyset$ *strong Nash blocks* π if a partition $\pi' \neq \pi$ exists with $\pi \xrightarrow{H} \pi'$ and $\forall i \in H : \pi'(i) \succ_i \pi(i)$.

  If a partition π is not strong Nash blocked by any set $H \subseteq N$, π is called *strong Nash stable (SNS)* [15].

- A subset of players $H \subseteq N, H \neq \emptyset$ *weakly Nash blocks* π if a partition $\pi' \neq \pi$ exists with $\pi \xrightarrow{H} \pi'$, $\forall i \in H : \pi'(i) \succsim_i \pi(i)$ and $\exists i \in H : \pi'(i) \succ_i \pi(i)$.

  A partition which admits no weakly Nash blocking coalition is said to satisfy *strict strong Nash stability (SSNS)*.

- A non-empty set of players $H \subseteq N$ is *strongly individually blocking* a partition π, if a partition π' exists such that:

  1. $\pi \xrightarrow{H} \pi'$ (as for SNS),
  2. $\forall i \in H : \pi'(i) \succ_i \pi(i)$, and
  3. $\forall j \in \pi'(i)$ for some $i \in H : \pi'(j) \succsim_j \pi(j)$.

A partition for which no strongly individually blocking set exists is *strongly individually stable (SIS)*.[1]

Figure 1: Inclusion relationships between stability concepts for hedonic games. For e.g, every NS partition is also IS. NS, SC, PO, C and IR are classic stability concepts. IS was formulated in [6]; SNS in [15]; and perfect partitions in [2]. We also introduce SSNS and SIS in this paper.

Depending on the context, we will utilize abbreviations like SIS, SNS, SSNS, IS etc. either for adjectives (for e.g. IS for individually stable) or for nouns (for e.g. IS for individual stability).

4. RELATIONS BETWEEN STABILITY CONCEPTS

In this section, we will explore and clarify the inclusion relationships between the stability concepts. The inclusion relationships between stability concepts are depicted in Figure 1.

PROPOSITION 1. *Strict core stability implies strong individual stability which implies individual stability and also core stability.*

PROOF. Strict core stability implies strong individual stability. Assume that a partition π is strict core stable but

---

[1]SIS is a natural intermediate stability concept which is implied by strong Nash stability and strict core stability respectively and it also implies individual stability and core stability.

not strong individually stable. Then, there exists a coalition $S \subseteq N$ such that $S \notin \pi$ and each player in S is at least as happy as in π and one player in S is strictly happier than in π. But this means that π is not strict core stable.

Strong individual stability trivially implies individual stability.

Finally, we show that strong individual stability implies core stability. Assume that a partition π is strong individually stable but not core stable. Then there is a core deviating coalition S. But this would mean that each player $i \in S$ is strictly better off than in $\pi(i)$. But this means that π is not strong individually stable. This completes the proof. □

Strong Nash stability as introduced by Karakaya [15] is quite a strong stability notion as seen by the following simple proposition.

PROPOSITION 2. *Strong Nash stability implies Nash stability and also core stability.*

Furthermore, even if a partition is both strict core stable and Nash stable, it is not necessarily strong Nash stable.

PROOF. The first statement follows from the definitions of the stability concepts and was already pointed out by Karakaya [15]. In fact, it can also easily be shown that SNS implies SIS. If a partition is SNS, then there is no strong Nash blocking set. This implies that there does not exist any strongly individually blocking set.

We now show that even if a partition is both strict core stable and Nash stable, it is not necessarily strong Nash stable. The following example shows a game, that admits a strict core and Nash stable partition but no strong Nash stable partition.

Let $(N, \succsim)$ be a game with $N = \{1, 2, 3, 4\}$ and the preference profile specified as follows:

$$\{1,2\} \succ_1 \{1,4\} \succ_1 \{1\} \succ_1 \ldots$$
$$\{2,3\} \succ_2 \{1,2\} \succ_2 \{2\} \succ_2 \ldots$$
$$\{3,4\} \succ_3 \{2,3\} \succ_3 \{3\} \succ_3 \ldots$$
$$\{1,4\} \succ_4 \{3,4\} \succ_4 \{4\} \succ_4 \ldots$$

It is easy to check, that the partitions $\pi = \{\{1,2\}, \{3,4\}\}$ and $\pi' = \{\{1,4\}, \{2,3\}\}$ are both (even strictly) core stable and Nash stable. But neither of them is strong Nash stable since $\{2,4\}$ is blocking π and π' is blocked by $\{1,3\}$. Obviously any partition containing a coalition with 3 or more players is not even Nash stable, since each player prefers being alone to any coalition with more than 2 players. Also $\{\{1,3\}, \{2,4\}\}$ is not even Nash stable, since it is not individually rational.

Nash and core stability prevent single players from moving to another (possibly empty) coalition or several players forming a new coalition respectively. In the given partitions π and π', it is possible for a pair of players to improve by switching coalitions and therefore prevent π and π' from being strong Nash stable. □

In the next proposition, we show that although strong Nash stability is a strong stability concept, it implies neither strict core stability nor Pareto optimality.

PROPOSITION 3. *Strict strong Nash stability implies strong Nash stability, strict core stability and Pareto optimality.*

On the other hand, strong Nash stability implies neither strict core stability nor Pareto optimality.

PROOF. Strict strong Nash stability trivially implies strong Nash stability. Strict strong Nash stability also implies strict core stability. If a partition is strong Nash stable, there exists no new coalition H, in which each player at least as happy and one player is strictly better off. Therefore, the partition is also strict core stable.

Now, we will show that strong Nash stability implies neither strict core stability nor Pareto optimality. Since, it is well-known that strict core stability implies Pareto optimality, it is sufficient to show that strong Nash stability does not imply Pareto optimality.

Strong Nash stability does not imply Pareto optimality. Consider the following four-player hedonic game:

$$\{1,2\} \sim_1 \{1,3\} \sim_1 \{1,4\} \succ_1 \cdots$$

$$\{1,2\} \sim_2 \{2,3\} \sim_2 \{2,4\} \succ_2 \cdots$$

$$\{2,3\} \sim_3 \{3,4\} \succ_3 \cdots$$

$$\{1,4\} \sim_4 \{2,4\} \succ_4 \{3,4\} \succ_4 \cdots$$

Then, the partition $\{\{1,2\},\{3,4\}\}$ is strong Nash stable. However it is Pareto dominated by $\{\{2,3\},\{1,4\}\}$. □

In the next sections, we will present the central results of the paper.

5. TOP RESPONSIVENESS

Top responsiveness [1, 9, 10] and bottom responsiveness [17] are natural restrictions that are imposed on the individual preferences and not on the whole preference profile. The idea is that a player's preference for a coalition depends on the best and worst subcoalitions respectively. In this section, we present a result that a partition fulfilling SSNS exists for hedonic games satisfying top responsiveness and an additional property called mutuality (with respect to top responsiveness).

Top responsiveness.

Top responsiveness is based on *choice sets*—sets of players which each player wants to be with. Let $Ch(i,S)$—the *choice sets* of player i in coalition S—be defined as follows:

$$Ch(i,S) = \{S' \subseteq S : (i \in S') \land (S' \succsim_i S'' \; \forall S'' \subseteq S)\}.$$

A game satisfies *top responsiveness* if for each $i \in N$, the following three conditions hold:

1. for each $X \in \mathcal{N}_i$, $|Ch(i,X)| = 1$, (we denote by $ch(i,X)$ the unique maximal set of player i on X under $\succsim_i$),

2. for each pair $X,Y \in \mathcal{N}_i$, $X \succ_i Y$ if $ch(i,X) \succ_i ch(i,Y)$;

3. for each pair $X,Y \in \mathcal{N}_i$, $X \succ_i Y$ if $ch(i,X) = ch(i,Y)$ and $X \subset Y$.

A hedonic game satisfying top responsiveness additionally satisfies *mutuality* if

$$\forall i,j \in N, X \in \mathcal{N}_i \cap \mathcal{N}_j : i \in ch(j,X) \Leftrightarrow j \in ch(i,X).$$

We will also specify similar notion of mutuality with respect to hedonic games satisfying strong bottom responsiveness. When the context is clear, we will refer to the condition simply as mutuality.

EXAMPLE 1. *Let* $(N,\succsim)$ *be a game with* $N = \{1,2,3\}$ *and the preference profile specified as follows:*

$$\{1,2\} \succ_1 \{1,2,3\} \succ_1 \{1\} \succ_1 \{1,3\}$$

$$\{1,2,3\} \succ_2 \{1,2\} \sim_2 \{2,3\} \succ_2 \{2\}$$

$$\{2,3\} \succ_3 \{1,2,3\} \succ_3 \{3\} \succ_3 \{1,3\}$$

Then, $(N,\succsim)$ *satisfies top responsiveness and mutuality.*

We are now in a position to present our first result.

THEOREM 1. *Top responsiveness and mutuality together guarantee the existence of an SSNS partition.*

We prove Theorem 1 by showing that if a hedonic game satisfies top responsiveness and mutuality, then the Top Covering Algorithm of [1, 9, 10] returns an SSNS partition. Therefore, we identify conditions different than the ones identified by Karakaya [15] for which strong Nash stability is guaranteed. Since SSNS is stronger than SNS (Proposition 3) which in turn is stronger than even the combination of Nash stability and strict core stability (Proposition 2), Theorem 1 simultaneously strengthens the result in [9] and [10] in which it was shown that top responsiveness and mutuality together guarantee the existence of a Nash stable and strict core partition.

It can also be proved that Theorem 1 is optimal in the sense that it does not extend to perfect partitions. To be precise, we show that top responsiveness and mutuality together do not guarantee the existence of a perfect partition.

PROPOSITION 4. *Top responsiveness and mutuality together do not guarantee the existence of a perfect partition.*

PROOF. By counter example. In the game in Example 1, top responsiveness and mutuality are satisfied but no perfect partition exists. □

Now that we have stated Theorem 1 and its complementing Proposition 4, we will present the proof of Theorem 1.

Proof of Theorem 1.

Firstly, we need additional definitions and a description of the Top Covering Algorithm. For each $X \subseteq N$, we denote by $\frown_X$ the relation on $X \times X$ where $i \frown_X j$ if and only if $j \in ch(i,X)$. In this case j is called a *neighbor* of i in X. Note that in the preference profiles satisfies top responsiveness mutuality, then $\frown_X$ is a symmetric relation.

The connected component $CC(i,X)$ of i with respect to X is defined as follows:

$$CC(i,X) = \{k \in X : \exists j_1,\ldots,j_l \in X : i = j_1 \frown_X \cdots \frown_X j_l = k\}.$$

If $j \in CC(i,X)$, j is called *reachable* from i in X. Also note that $CC(j,X) \subseteq CC(i,X)$ if j is reachable from i and if mutuality is satisfied, then the following holds: $\forall X \subseteq N$, $i,j \in X : i \in CC(j,X) \Leftrightarrow j \in CC(i,X)$.

Now we are ready to present the simplified Top Covering Algorithm provided by Dimitrov and Sung [9, 10], adapted to the notation defined above. The algorithm is specified as Algorithm 1.

The following lemma will be used in the proof to Theorem 1.

LEMMA 1. *Let $(N, \succsim)$ be a game satisfying top responsiveness and mutuality and π be the partition resulting by applying the simplified Top Covering Algorithm to it. Then*

$$\forall i \in N : ch(i, N) \subseteq \pi(i)$$

PROOF. First we show by induction over the iterations of the algorithm that $ch(i, R^k) = ch(i, N)$ $\forall i \in R^k$, $k = 1, 2, ...$. For $k = 1$, this is obviously true, because $R^1 = N$. Assume by induction, that $ch(i, R^k) = ch(i, N)$ $\forall i \in R^k$. Let i' be the player selected in the k-th iteration of Step 3 and $j \in R^{k+1}$. Therefore $j \notin ch(i, R^k)$ $\forall i \in CC(i', R^k)$. Because of mutuality $i \notin ch(j, R^k)$ $\forall i \in CC(i', R^k)$. So $ch(j, R^k) \subseteq R^{k+1}$ and therefore $ch(j, R^{k+1}) = ch(j, R^k) = ch(j, N)$.

Now take an arbitrary player $i \in N$ and denote by k the iteration of the algorithm in which i was added to his coalition, i.e. $i \in S^k$. Let i' be the player selected in the k-th iteration of Step 3, so $i \in CC(i', R^k)$. Because of mutuality, $CC(i', R^k) = CC(i, R^k)$ and clearly $ch(i, R^k) \subseteq CC(i, R^k)$. From above, we know that $ch(i, N) = ch(i, R^k) \subseteq CC(i, R^k) = S^k = \pi(i)$. $\square$

Algorithm 1 Top Covering Algorithm

Input: A hedonic game $(N, \succsim)$ satisfying top responsiveness.

1: $R^1 \leftarrow N$; $\pi \leftarrow \emptyset$.
2: **for** $k = 1$ to $|N|$ **do**
3:      Select $i \in R^k$ such that $|CC(i, R^k)| \leq |CC(j, R^k)|$ for each $j \in R^k$.
4:      $S^k \leftarrow CC(i, R^k)$; $\pi \leftarrow \pi \cup \{S^k\}$; and $R^{k+1} \leftarrow R^k \setminus S^k$
5:      **if** $R^{k+1} = \emptyset$ **then**
6:          **return** π
7:      **end if**
8: **end for**
9: **return** π

We note here that Lemma 1 may not hold, if mutuality is violated.

As shown by Dimitrov and Sung [9, 10] the resulting partition of the simplified Top Covering Algorithm is strict core stable as well as Nash stable if preferences as mutual. We are now ready to present the proof of Theorem 1.

PROOF. Let π be the resulting partition and suppose it is not strictly strong Nash stable. Then a pair (H, π') exists where $H \subseteq N$ is the set of deviators and π' is the partition resulting after the deviation, i.e. $\pi \xrightarrow{H} \pi'$. Firstly, by Lemma 1, $ch(i, N) \subseteq CC(i, N)$ $\forall i \in N$. Since H is a coalition blocking strict strong Nash stability, the following holds:

$$\forall i \in H : \pi'(i) \succsim_i \pi(i) \quad and$$
$$\exists j \in H : \pi'(j) \succ_j \pi(j).$$

Now consider the player j, who is better off in his new coalition $\pi'(j)$. Assume that $\pi(j) \cap \pi'(j) \subseteq H$, which means only deviators in $\pi(j) \cap \pi'(j)$. For $i \in \pi(j) \cap \pi'(j)$: $ch(i, \pi'(i)) \succsim_i ch(i, \pi(i))$, since $i \in H$ by assumption. We also know that $ch(i, \pi(i)) = ch(i, N)$ by Lemma 1. Therefore, for $i \in \pi(j) \cap \pi'(j)$: $ch(i, \pi'(i)) \succsim_i ch(i, \pi(i)) = ch(i, N)$. Because of uniqueness of choice sets in the definition of top responsiveness, $ch(i, \pi'(i)) = ch(i, N)$. So

$ch(i, N) \subseteq \pi(i) \cap \pi'(i) = \pi(j) \cap \pi'(j)$.

$$\implies \forall i \in \pi(j) \cap \pi'(j) : (\pi'(j) \cap \pi(j)) \succsim_i \pi'(j).$$

Due to assumption $\pi(j) \cap \pi'(j) \subseteq H$, the following holds:

$\forall i \in \pi(j) \cap \pi'(j) : (\pi(j) \cap \pi'(j)) \succsim_i \pi'(j) \succsim_i \pi(j) = \pi(i)$ &

$(\pi(j) \cap \pi'(j)) \succsim_j \pi'(j) \succ_j \pi(j)$

So $\pi(j) \cap \pi'(j)$ would be a coalition blocking strict core stability, but Dimitrov and Sung [10] proved that π as produced by the simplified Top Covering Algorithm has to be strict core stable. Therefore $\pi'(j) \cap \pi(j) \nsubseteq H$ and there is at least one non-deviator in $\pi(j) \cap \pi'(j)$. Let us call this player i'.

Now take a look at the players in $\pi(j) \setminus \pi'(j)$. Note that this is not an empty set, because otherwise $\pi'(j) \supset \pi(j) \implies \pi'(j) \precsim_j \pi(j)$. If one of them is not in H, then he was in the same coalition as i' in π, namely $\pi(j)$, and is now in a different, which is not consistent with $\pi \xrightarrow{H} \pi'$. So $(\pi(j) \setminus \pi'(j)) \subseteq H$. Because $\pi(j)$ is a connected component, at least one player k in $\pi(j) \setminus \pi'(j)$ has a friend l in $\pi'(j)$, meaning they are in each other's choice sets and as mentioned $l \in ch(k, N) \subseteq \pi(k)$. But now $l \notin \pi'(k)$ and therefore $\pi'(k) \prec_k \pi(k)$ which contradicts k being a deviator. $\square$

6. BOTTOM RESPONSIVENESS

In this section, we present the central results concerning hedonic games which satisfy bottom responsiveness.

Bottom responsiveness.

Bottom responsiveness is a restriction on the preferences of each player in a hedonic game which models conservative or pessimistic agents. In contrast to top responsiveness, bottom responsiveness is based on *avoid sets*—sets of players which each player wants to avoid having in his coalition.

For any player $i \in N$ and $S \in \mathcal{N}_i$, $Av(i, S)$—the set of *avoid sets* of player i in coalition S—is defined as follows:

$$Av(i, S) = \{S' \subseteq S : (i \in S') \wedge (S' \precsim_i S'' \ \forall S'' \subseteq S)\}.$$

A game satisfies *bottom responsiveness* if for each $i \in N$, the following conditions hold:

1. for each pair $X, Y \in \mathcal{N}_i$, $X \succ_i Y$ if $X' \succ_i Y'$ for each $X' \in Av(i, X)$ and each $Y' \in Av(i, Y)$; and

2. for each $i \in N$ and $X, Y \in \mathcal{N}_i$, $Av(i, X) \cap Av(i, Y) \neq \emptyset$ and $|X| \geq |Y|$ implies $X \succsim_i Y$.

A hedonic game $(N, \succsim)$ satisfies *strong bottom responsiveness* if it is bottom responsive and if for each $i \in N$ and $X \in \mathcal{N}_i$, $|Av(i, X)| = 1$. By $av(i, X)$, we denote the unique minimal set of player i on X under $\succsim_i$. The strong part of bottom responsiveness is analogous to Property 1 in the definition of top responsiveness. A hedonic game $(N, \succsim)$ satisfying strong bottom responsiveness additionally satisfies *mutuality* if for all $i, j \in N$, and X such that $i, j \in X$, $i \in av(j, X)$ if and only if $j \in av(i, X)$.

EXAMPLE 2. *Let $(N, \succsim)$ be a game with $N = \{1, 2, 3\}$ and the preference profile specified as follows:*

$$\{1, 3\} \succ_1 \{1\} \succ_1 \{1, 2, 3\} \succ_1 \{1, 2\}$$
$$\{2, 3\} \succ_2 \{2\} \succ_2 \{1, 2, 3\} \succ_2 \{1, 2\}$$
$$\{1, 2, 3\} \succ_3 \{1, 3\} \sim_3 \{2, 3\} \succ_3 \{3\}$$

Then, $(N, \succsim)$ satisfies strong bottom responsiveness and also mutuality (with respect to strong bottom responsiveness).

For bottom responsive games, we prove that an SIS partition is guaranteed to exist even in the absence of mutuality.

THEOREM 2. *Bottom responsiveness guarantees the existence of an SIS partition.*

As a corollary, a core stable partition and an individually stable partition is guaranteed to exist. Previously, it was only known that the core is non-empty for bottom responsive games [17]. In contrast to the result by Suzuki and Sung [17], the proof of Theorem 2 does not require the Bottom Avoiding Algorithm. We associate with each IR partition a vector of coalition sizes in decreasing order. It is then shown via lexicographic comparisons between the corresponding vectors that arbitrary deviations between partitions are acyclic. With an additional natural constraint, even SNS is guaranteed (Theorem 3).

THEOREM 3. *Strong bottom responsiveness and mutuality together guarantee the existence of an SNS partition.*

We point out that Theorem 2 cannot be extended any further to take care of strict core stability and Theorem 3 cannot be extended to SSNS. The reason is that *symmetric 'aversion to enemies' games*—a subclass of strong bottom responsive games which satisfy mutuality—may not admit a strict core stable partition (Example 4, [11]).

Now that we have stated our results concerning hedonic games satisfying bottom responsiveness, we sketch the proofs.

Proof of Theorem 2.

For the use of further proofs, we introduce an ordering relation on the partitions. The definition will also apply to the proof of Theorem 3.

DEFINITION 1. *Let $N = \{1, ..., n\}$ be a set of players and π, π' two partitions of N, where $\pi = (S_1, ..., S_k)$ and $\pi' = (T_1, ..., T_l)$ with $|S_i| \geq |S_{i+1}| \ \forall i \in \{1, ..., k-1\}$ and $|T_j| \geq |T_{j+1}| \ \forall j \in \{1, ..., l-1\}$ respectively. We say, that*

$$\pi \mathrel{\dot{>}} \pi' \Leftrightarrow \exists i \leq \min\{k, l\} : |S_i| > |T_i| \text{ and } |S_j| = |T_j| \ \forall j < i$$

& $\pi \mathrel{\dot{=}} \pi' \Leftrightarrow k = l$ and $\forall i \leq k : |S_i| = |T_i|$.

The relation $\dot{>}$ is complete, transitive and asymmetric, and places an ordering on the set of partitions. We now present the proof of Theorem 2 in which we utilize the relation $\dot{>}$.

PROOF. To simplify the presentation, we prove that every bottom responsive game admits an IS partition. The same argument can also be used to show that every bottom responsive game admits an SIS partition.

We show individual stability for each maximum element according to $\dot{>}$ of the set of individual rational coalitions. Consider the set $P = \{\pi' : \pi' \ partitions \ N \ and \ \forall S \in \pi', i \in S : \{i\} \in Av(i, S)\}$. Note that $P \neq \emptyset$, because the partition consisting of only singletons is in P and that P is a finite set because the number of partitions is finite. Denote by π a maximal element of P according to $\dot{>}$, i.e. $\pi \mathrel{\dot{\geq}} \pi' \ \forall \pi' \in P$. By definition, π is individually rational.

Now assume π is not individually stable. Then, there exists a player $i \in N$ and a coalition $S \in \pi \cup \{\emptyset\}$, such that $S \cup \{i\} \succ_i \pi(i)$ and $\forall j \in S : S \cup \{i\} \succsim_j S$. Now we show that the partition π resulting after the deviation of i is still individually rational and therefore an element of P. Clearly $S \neq \emptyset$ because of individual rationality of π. Furthermore $\{j\} \in Av(j, S \cup \{i\}) \ \forall j \in S$, because if not $S \cup \{i\} \prec_j S$ for some $j \in S$.
Consider a player $j \in \pi(i) \setminus \{i\}$. Due to individual rationality of π, $\{j\} \in Av(j, \pi(i))$, which implies $T \succsim_j \{j\} \ \forall T \subseteq \pi(i)$ with $j \in T$. So $\pi(i) \setminus \{i\} \succsim_j \{j\}$. All other players $j \in N \setminus (\pi(i) \cup S)$ are not affected by the deviation of i because of the hedonic game setting. Therefore π' is individually rational and $\pi' \in P$.

The last step is to show $\pi' \mathrel{\dot{>}} \pi$, which contradicts the maximality of π in P. Because player i improves by changing, $|S \cup \{i\}| > |\pi(i)|$ follows from condition 2) of bottom responsiveness. So $(S \cup \{i\}, \pi(i) \setminus \{i\}) \mathrel{\dot{>}} (S, \pi(i))$ and all other coalitions are identical in π and π'. This contradicts $\pi \mathrel{\dot{\geq}} \pi'$ and finishes the proof. $\square$

The proof also highlights a decentralized way to compute an IS or SIS partition. Start from the partition of singletons and enable arbitrary deviations. For each partition π_k, the new partition π_{k+1} is such that $\pi_{k+1} \mathrel{\dot{>}} \pi_k$. Therefore, in a finite number of deviations, an IS or SIS partition is achieved.

Proof of Theorem 3.

We now present the proof of Theorem 3.

PROOF. We show strong Nash stability for each maximal element according to $\dot{\geq}$ of the set of individual rational coalitions (please see Definition 1). Consider the set $P = \{\pi' : \pi' \ partitions \ N \ and \ \forall S \in \pi', i \in S : \{i\} = av(i, S)\}$. Note that $P \neq \emptyset$, because the partition consisting of only singletons is in P and P is a finite set, because the number of partitions is finite. Denote by π a maximal element of P according to $\dot{\geq}$, i.e. $\pi \mathrel{\dot{\geq}} \pi' \ \forall \pi' \in P$.

Now assume π is not strong Nash stable. Then a set of players $H \subseteq N$ and a partition π' exist, such that

(1) $\pi \xrightarrow{H} \pi'$

(2) $\forall i \in H : \pi'(i) \succ_i \pi(i)$.

We show that the partition π' resulting after the deviation is still individually rational and therefore an element of P. Clearly $av(i, \pi'(i)) = \{i\} \ \forall i \in H$, because otherwise $\pi'(i) \succ_i \pi(i)$ would not hold. Now consider a player j such that $\pi'(j) \cap H \neq \emptyset$. $\forall i \in H \cap \pi'(j) : j \notin av(i, \pi'(j))$. Mutuality implies $i \notin av(j, \pi'(j))$ and therefore $av(j, \pi'(j)) = av(j, \pi(j)) = \{j\}$. All other players $j \in N$ are either not affected by any changes $(\pi(j) = \pi'(j))$ or they are left by some players in H $(\pi'(j) \subset \pi(j))$. In both cases $av(j, \pi'(j)) = av(j, \pi(j)) = \{j\}$, so π' is an element of P.

The last step is to show $\pi' \mathrel{\dot{>}} \pi$, which contradicts the maximality of π in P. Because each player $i \in H$ improves, $|\pi'(i)| > |\pi(i)| \ \forall i \in H$, which follows from condition (iii) of bottom responsiveness. Take the largest coalition $S \in \pi$ such that $S \cap H \neq \emptyset$. Obviously any coalition bigger than S in π at least does not get smaller after the deviation, because it contains no players from H. Then one of following two cases holds:

Case 1: at least one coalition $T \in \pi$ with $|T| > |S|$ gets joined by some player $i \in H$. But then $\pi' \overset{.}{>} \pi$, since T increases in size and any larger coalition in π does not get smaller.

Case 2: If Case 1 does not hold, we know that no coalition in π larger than S is joined by a player in H and therefore stays the same. But one player $i \in S \cap H$ is part of a coalition $S' \in \pi'$ with $|S'| > |S|$. Since all coalitions in π, which are larger than S also exist in π', we can again conclude $\pi' \overset{.}{>} \pi$.

In both cases $\pi' \overset{.}{>} \pi$ which contradicts the maximality of π in P and finishes the proof. $\square$

7. EXISTENCE OF STABILITY FOR SPECIFIC CLASSES OF GAMES

In this section, we highlight some natural subclasses of additively separable hedonic games [see e.g., 3, 6, 12, 14] and hedonic games with $\mathcal{B}$-preferences [see e.g., 8, 14] which guarantee top responsiveness or bottom responsiveness. Consequently, our existence results in Sections 5 and 6 and an existence result in the literature [10] applies to these settings.

Additively separable hedonic games.

Additively separable hedonic games are one of the most well-studied and natural class of hedonic games [see e.g., 3, 6, 12, 14]. In an *additively separable hedonic game (ASHG)* $(N, \succsim)$, each player $i \in N$ has value $v_i(j)$ for player j being in the same coalition as i and if i is in coalition $S \in \mathcal{N}_i$, then i gets utility $\sum_{j \in S \setminus \{i\}} v_i(j)$. For coalitions $S, T \in \mathcal{N}_i$, $S \succsim_i T$ if and only if $\sum_{j \in S \setminus \{i\}} v_i(j) \geq \sum_{j \in T \setminus \{i\}} v_i(j)$. Therefore an ASHG can be represented as (N, v). An ASHG is *symmetric* if $v_i(j) = v_j(i)$ for any two players $i, j \in N$ and is *strict* if $v_i(j) \neq 0$ for all $i, j \in N$.

We now formally introduce two classes of additively separable hedonic games which also satisfy top responsiveness and bottom responsiveness respectively. Both classes were introduced by Dimitrov et al. [11].

- An ASGH (N, v) is *appreciation of friends* if for all $i, j \in N$ such that $i \neq j$, the following holds: $v_i(j) \in \{-1, +n\}$.

- An ASGH (N, v) is *aversion to enemies* if for all $i, j \in N$ such that $i \neq j$, the following holds: $v_i(j) \in \{-n, +1\}$.

It is clear that 'appreciation of friends' and 'aversion to enemies' games are ASHGs with strict preferences. Suzuki and Sung [17] noted that 'appreciation of friends' and 'aversion to enemies' games satisfy top responsiveness and bottom responsiveness respectively. As a consequence, our main results apply to these games.

COROLLARY 1. *There exists an SSNS partition for each symmetric 'appreciation of friends' game.*

PROOF. An 'appreciation of friends' game satisfies top responsiveness. Furthermore, if the game is (additively separable) symmetric, then it also satisfies mutuality with respect to top responsiveness. Then, as a result of Theorem 1, we get the corollary. $\square$

COROLLARY 2. *There exists an SIS partition for each 'aversion to enemies' game.*

PROOF. The statement follows from Theorem 2 and the fact that 'aversion to enemies' games satisfy bottom responsiveness. $\square$

COROLLARY 3. *There exists an SNS partition for each symmetric 'aversion to enemies' game.*

PROOF. It is already known that 'aversion to enemies' games satisfy bottom responsiveness. Since 'aversion to enemies' are additively separable hedonic games with strict preferences, they not only satisfy bottom responsiveness but also strong bottom responsiveness. If 'aversion to enemies' have symmetric preferences, then they not only satisfy strong bottom responsiveness but also (bottom responsive) mutuality. Therefore, we can apply Theorem 3 to derive the corollary. $\square$

$\mathcal{B}$-hedonic games.

Finally, we show another important subclass of hedonic games called $\mathcal{B}$-hedonic games [8, 7] satisfies top responsiveness. In $\mathcal{B}$-hedonic games, players express preferences over players and these preferences over players are naturally extended to preferences over coalitions. We will assume that $\max_i(\emptyset) = \{i\}$. In *hedonic games with $\mathcal{B}$-preferences* (in short $\mathcal{B}$-hedonic games), for $S, T \in \mathcal{N}_i$, $S \succ_i T$ if and only if one of the following conditions hold:

1. for each $s \in \max_i(S \setminus \{i\})$ and $t \in \max_i(T \setminus \{i\})$, $s \succ_i t$, or

2. for each $s \in \max_i(S \setminus \{i\})$ and $t \in \max_i(T \setminus \{i\})$, $s \sim_i t$ and $|S| < |T|$.

A $\mathcal{B}$-hedonic has strict preferences for each $i \in N$ and $j, k \in N$, the following holds: $j \neq k \Rightarrow j \not\sim_i k$. Then, we have the following proposition.

PROPOSITION 5. *$\mathcal{B}$-hedonic games with strict preferences satisfy top responsiveness.*

PROOF. We show that $\mathcal{B}$-hedonic games with strict preferences satisfy all the three conditions of top responsiveness.

1. Firstly, for each $X \in \mathcal{N}_i$, $Ch(i, X) = \{\max_i X \cup \{i\}\}$ and thus $|Ch(i, X)| = 1$.

2. For a pair $X, Y \in \mathcal{N}_i$, assume that $ch(i, X) \succ_i ch(i, Y)$. This means that $\{\max_i(X)\} \cup \{i\} \succ_i \{\max_i Y\} \cup \{i\}$. Since the best player in X is more preferred by i than the best player in Y, then by the definition of $\mathcal{B}$-hedonic games, $X \succ_i Y$.

3. Finally, for each pair $X, Y \in \mathcal{N}_i$, assume that $ch(i, X) = ch(i, Y)$ and $X \subset Y$. Then, the player most preferred by i in X is the same as the player player most preferred by i in Y. Therefore, by the definition of $\mathcal{B}$-hedonic games, $X \succ_i Y$.

This completes the proof. $\square$

Therefore, as a corollary we get the following statement which was proved by Cechlárová and Romero-Medina [8].

COROLLARY 4. *For each $\mathcal{B}$-hedonic game with strict preferences, a strict core stable partition is guaranteed to exist.*

PROOF. Dimitrov and Sung [10] showed that for hedonic games satisfying top responsiveness admit a strict core stable partition. Since $\mathcal{B}$-hedonic games satisfy top responsiveness, they admit a strict core stable partition. □

It will be interesting to see whether there are any natural restrictions on $\mathcal{B}$-hedonic games with strict preferences such that not only top responsiveness is satisfied but also (top responsive) mutuality is satisfied. In that case, we can apply Theorem 1 concerning SSNS to $\mathcal{B}$-hedonic games.

Our demonstrated connection between $\mathcal{B}$-hedonic games and top responsiveness goes deeper. The essential fact behind previous results concerning $\mathcal{B}$-hedonic games with strict preferences is that they satisfy top responsiveness. It turns out that the Top Covering Algorithm in [1] generalizes the B-STABLE algorithm in [8] and in fact Theorems 4.4 and 5.2 in [1] imply Theorem 1 and Theorem 2 in [8] respectively. This connection seems to have been unnoticed in the literature.

8. CONCLUSIONS

To conclude, we tried to paint a clearer picture of the landscape of stability concepts used in coalition formation games. The concepts ranged from standard ones such as the core to recently introduced concepts such as strong Nash stability. The core and strong Nash stability were generalized to strong individual stability and strict strong Nash stability respectively. The basic inclusion relationships between the stability concepts are depicted in Figure 1. Since hedonic games generalize various matching settings, the relations between the stability concepts also hold in matching settings such as two-sided matching, roommate matching etc.

We then examined restrictions on the preferences of agents which guarantee stable outcomes for the new stability concepts. Three main existence results (Theorems 1, 2 and 3) pertaining to top responsiveness and bottom responsiveness were presented. Our results strengthen or complement a number of results in the literature. We also showed that none of our existence results can be extended to a stronger known stability concept. It was seen that the theorems apply to some natural subclasses of hedonic games which have already been of interest among game-theorists. It will be interesting to find further applications of our existence results.

Identifying the impact of preference restrictions on stability also has algorithmic consequences. Recently, hedonic games have attracted research from an algorithmic and computational complexity point of view. There are various algorithmic questions such as checking the existence of and computing stable partitions for different representations of hedonic games (see e.g., [7, 14]). A general framework of preference restrictions and their impact on stability of partitions promises to be useful in devising generic algorithmic techniques to compute stable partitions. For example, we noted that the Top Covering Algorithm in [1] generalizes the B-STABLE algorithm in [8] by utilizing the insight that $\mathcal{B}$-hedonic games with strict preferences satisfy top responsiveness. We also mention the following interesting algorithmic questions. For hedonic games represented by individually rational lists of coalitions [4], what is the computational complexity of testing whether the game satisfies top reponsiveness or bottom responsiveness?

Our focus in the paper has been on sufficient conditions which guarantee the existence of stable outcomes. It will be interesting to see what additional conditions are required to ensure uniqueness of stable partitions for different notions of stability. Finally, characterizing the conditions for the existence of stability remains an open problem.

REFERENCES

[1] J. Alcalde and P. Revilla. Researching with whom? Stability and manipulation. *Journal of Mathematical Economics*, 40(8):869–887, 2004.

[2] H. Aziz, F. Brandt, and P. Harrenstein. Pareto optimality in coalition formation. In G. Persiano, editor, *Proceedings of the 4th International Symposium on Algorithmic Game Theory (SAGT)*, Lecture Notes in Computer Science (LNCS), pages 93–104. Springer-Verlag, 2011.

[3] H. Aziz, F. Brandt, and H. G. Seedig. Optimal partitions in additively separable hedonic games. In T. Walsh, editor, *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 43–48. AAAI Press, 2011.

[4] C. Ballester. NP-completeness in hedonic games. *Games and Economic Behavior*, 49(1):1–30, 2004.

[5] S. Banerjee, H. Konishi, and T. Sönmez. Core in a simple coalition formation game. *Social Choice and Welfare*, 18:135–153, 2001.

[6] A. Bogomolnaia and M. O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

[7] K. Cechlárová. Stable partition problem. In *Encyclopedia of Algorithms*, pages 885–888. Springer, 2008.

[8] K. Cechlárová and A. Romero-Medina. Stability in coalition formation games. *International Journal of Game Theory*, 29:487–494, 2001.

[9] D. Dimitrov and S. C. Sung. Top responsiveness and Nash stability in coalition formation games. *Kybernetika*, 42(4):453–460, 2006.

[10] D. Dimitrov and S. C. Sung. On top responsiveness and strict core stability. *Journal of Mathematical Economics*, 43(2):130–134, 2007.

[11] D. Dimitrov, P. Borm, R. Hendrickx, and S. C. Sung. Simple priorities and core stability in hedonic games. *Social Choice and Welfare*, 26(2):421–433, 2006.

[12] M. Gairing and R. Savani. Computing stable outcomes in hedonic games with voting-based deviations. In *Proceedings of the 10th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 559–566, 2011.

[13] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[14] J. Hajduková. Coalition formation games: A survey. *International Game Theory Review*, 8(4):613–641, 2006.

[15] M. Karakaya. Hedonic coalition formation games: A new stability notion. *Mathematical Social Sciences*, 2011.

[16] L. S. Shapley and H. Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37, 1974.

[17] K. Suzuki and S. C. Sung. Hedonic coalition formation in conservative societies. Technical Report 1700921, Social Science Research Network, 2010.