

Distributed Learning Dynamics for Coalitional Games

Aya Hamed¹ and Jeff S. Shamma¹

Abstract—In the framework of transferable utility coalitional games, a scoring (characteristic) function determines the value of any subset/coalition of agents. Agents decide on both which coalitions to form and the allocations of the values of the formed coalitions among their members. An important concept in coalitional games is that of a core solution, which is a partitioning of agents into coalitions and an associated allocation to each agent under which no group of agents can get a higher allocation by forming an alternative coalition. We present distributed learning dynamics for coalitional games that converge to a core solution whenever one exists. In these dynamics, an agent maintains a state consisting of (i) an aspiration level for its allocation and (ii) the coalition, if any, to which it belongs. In each stage, a randomly activated agent proposes to form a new coalition and changes its aspiration based on the success or failure of its proposal. The coalition membership structure is changed, accordingly, whenever the proposal succeeds. Required communications are that: (i) agents in the proposed new coalition need to reveal their current aspirations to the proposing agent, and (ii) agents are informed if they are joining the proposed coalition or if their existing coalition is broken. The proposing agent computes the feasibility of forming the coalition. We show that the dynamics hit an absorbing state whenever a core solution is reached. We further illustrate the distributed learning dynamics on a multi-agent task allocation setting.

I. INTRODUCTION

In strategic interactions with self-interested agents, cooperating with other agents can be the optimal strategic decision an agent can make to maximize its own benefits. In some interactive situations, the whole can be bigger than the sum of its parts, and agents forming coalitions can collectively gain higher payoffs. Coalitional games, within the field of cooperative game theory, enable us to study such situations.

The authors in [1]–[3] use coalitional game theory concepts to study the problem of forming coalitions of microgrids for local power exchange. Such collaboration between the microgrids increases their revenues and improves the autonomy of the system by decreasing their reliance on the main grid. In communications and networks fields, coalitional game theory offers a suitable framework to address a spectrum of problems. The tutorial paper [4] provides a classification of coalitional games and a variety of applications for each. One such application is the rate allocation problem presented in [5], where the users of a multiaccess channel are modeled as agents in a coalitional game. The value of

any coalition of agents depends on the maximum sum-rate achieved by the coalition given that the agents outside the coalition will act as jammers. The authors motivate various allocation methods for the agents including two of the most common solution concepts in coalitional game theory, the core and the Shapley value.

Coalitional game theory is applied to various problems in multi-agent systems. Coalitional games framework was utilized in [6] in the control of a multi-agent system. Coalitions between agents dictate the communication between them and, hence, the collective optimization that they solve. The grand coalition, i.e., the union of all agents, represents a centralized optimization problem, and the partition of agents into singletons represents the fully decentralized optimization. The cost of coalitions depends on both the cost resulting from the coalition optimization as well as the cost of forming the coalition. Even though centralized optimization yields the least cost for the optimization problem, it has the highest overhead of communication between agents and complexity of the optimization problem. The authors in [7] and [8] use coalitional game theory solution concepts to solve the multi-robot task allocation problem. The Shapley value solution is used in [7] to group the robots into coalitions to fulfill the available tasks and specify the reward of each robot. The authors in [8] use coalition formation as an intermediary step to group robots and tasks into smaller coalitions, after which they solve for the optimal task allocation within these coalitions.

In multi-agent systems, a centralized agent is not always accessible to mediate computing a solution with satisfactory payoffs for all agents. In addition, reliance on a centralized agent means having a single point of failure, which threatens the robustness of the system. Furthermore, privacy concerns and communication constraints may limit the information an agent is willing or capable of sharing. Our proposed distributed learning algorithm addresses these concerns. The presented algorithm converges to a core solution, a central solution concept in coalitional game theory, whenever one exists. The algorithm does not require the agents to learn about the full state of the environment. Alternatively, at any time instance, an agent is randomly activated which then needs to know only the current payoff aspirations of a subset of its neighboring agents. Furthermore, the algorithm has small memory and computation requirements from the agents. They only need to retain their current payoff aspirations, coalition membership, and the values of the coalitions that they can form or a means of calculating such values online.

The core of Transferable Utility (TU) coalitional games

*This work was supported by the University of Illinois Urbana-Champaign

¹Aya Hamed and Jeff S. Shamma are with the Department of Industrial and Enterprise Systems Engineering, the Grainger College of Engineering, University of Illinois at Urbana-Champaign, 117 Transportation Building, MC-238, 104 S. Mathews Ave. Urbana, IL 61801-3080, USA ayah2@illinois.edu, jshamma@illinois.edu

is defined by a distribution of the payoff resulting from forming the grand coalition. The core allocation respects the individual rationality of the agents such that no agent gains less in the grand coalition than what it can gain on its own. In addition, the core guarantees stability against group deviations where no group of agents can gain more by jointly deviating from the grand coalition. Finally, the core represents efficiency in that the total payoffs gained by the agents equals the value of the grand coalition. However, in some settings, the grand coalition is not the optimal coalition to form; specifically, when there is an associated cost with coalition formation that grows with the size of the coalition. The review paper [9] mentions the lack of literature in coalitional game theory for solution concepts that consider these settings, which readily occur in energy applications. In this paper, we consider a generalized core solution concept, similarly defined in [10] and [11], which applies to such settings. This generalized core solution still preserves individual rationality and stability against deviation but with an extension to the efficiency concept such that the sum of the payoffs gained by the agents equals the maximum welfare that can be gained by any partitioning of the agents. Our proposed dynamics converge to this generalized core solution for general TU games, whenever one exists.

Classical core and Shapley value calculations do not consider the dynamics of coalition formation and dissolution. In addition to the assumption of the grand coalition optimality in most of the literature, the presented solutions and algorithms come short in connecting the reached solutions with real-life bargaining and negotiation setups. Our proposed dynamics exhibit feasible coalition formation dynamics throughout the horizon of iterations. Similar distributed aspiration-based algorithms were introduced in [12]–[14]. Our algorithm builds upon the blind matching algorithm in [12] and [13]. It uses similar aspiration-based states and negotiation mechanism to reach the final allocations of the agents. The algorithm in [14] considers superadditive TU games, games where the value of a coalition is at least as good as the sum of values of any disjoint set of its subcoalitions. In superadditive TU games, the grand coalition is always optimal to form. The algorithm thus converges to the standard core solution of these games. For general TU games, [10] and [11] provide similar convergence results to ours. However, they use best reply response, which imposes higher requirements on the agents' knowledge of their opponents' demands in addition to a global knowledge of the evolving coalition structure.

The rest of this paper is organized as follows. Section II presents the TU games setup and relevant propositions. Section III introduces the Coalition Proposal algorithm. Section IV discusses the convergence proof. Section V exhibits simulation results using the Coalition Proposal algorithm in a multi-agent task allocation setting. Finally, Section VI concludes the paper. For the sake of brevity, proofs and further discussion are included in the arXiv version of the paper [15].

II. TRANSFERABLE UTILITY GAMES

In this section, we present the TU games setup along with some definitions and propositions pertaining to it. We are following Hans Peters' book [16] in the fundamental definitions of the TU games, Definitions 1 and 2 below. However, we are generalizing the solution concept in Definition 4 to better suit general TU games where grand coalition formation does not necessarily result in the optimal welfare.

Definition 1. A *transferable utility coalitional game* is defined by the pair (N, v) , such that $N = \{1, 2, \dots, n\}$ is the set of players and $v : 2^N \rightarrow \mathbb{R}$, where $v(\emptyset) = 0$, is the characteristic function defining the value of each coalition of players. An *allocation* $\mathbf{x} \in \mathbb{R}^n$ is a vector of real numbers representing the payoff distribution among the players.

Definition 2. For a TU game (N, v) , an allocation \mathbf{x} is *individually rational* if $x_i \geq v(\{i\})$ for all $i \in N$ and *coalitionally rational* if $\sum_{i \in S} x_i \geq v(S)$ for all $S \subseteq N$.

Definition 3. A collection ρ of subsets of N is defined as a *partition* of N if $S \cap S' = \emptyset$ for all $S, S' \in \rho$ and $\cup_{S \in \rho} S = N$. We denote the set of all partitions of N by $\mathcal{P}(N)$.

Definition 4. A *core solution* of a TU game (N, v) is a pair (\mathbf{x}, ρ) , where \mathbf{x} is an allocation vector and $\rho \in \mathcal{P}(N)$ is a partition of players, satisfying that \mathbf{x} is coalitionally rational, $\sum_{i \in S} x_i \geq v(S)$ for all $S \subseteq N$, and $\sum_{i \in S} x_i = v(S)$ for all $S \in \rho$. We denote the set of core solutions of a game (N, v) by $\Pi(v)$.

Definition 5. For a TU game (N, v) , we denote the *maximum welfare value*, $\max_{\rho \in \mathcal{P}(N)} \sum_{S \in \rho} v(S)$, by K_v .

Proposition 1. For any TU game (N, v) , if $\Pi(v) \neq \emptyset$, then $\sum_{i \in N} x_i = K_v$ for all $(\mathbf{x}, \rho) \in \Pi(v)$.

Proposition 2. If $(\mathbf{x}, \rho) \in \Pi(v)$, $\mathbf{y} \in \mathbb{R}^n$ such that \mathbf{y} is coalitionally rational, $\sum_{i \in S} y_i \geq v(S)$ for all $S \subseteq N$, and $\sum_{i \in N} y_i = K_v$, then $(\mathbf{y}, \rho) \in \Pi(v)$.

III. COALITION PROPOSAL ALGORITHM

Algorithm 1 presents a pseudo code for our proposed algorithm. Informally, the algorithm proceeds as follows:

- 1) Players come with arbitrary initial aspirations. The aspirations can be set to the players' singleton coalitions valuations or any value on a grid of width δ , a chosen discretization value.
- 2) A player is activated uniformly at random, which in turn chooses, using a uniform distribution, a set of other players to propose forming a coalition with.
- 3) The proposing player asks for and receives the current aspirations of the other players in the proposed coalition. Then, if the total of their aspirations in addition to its own aspiration raised by δ , is less than or equal to the coalition valuation, the coalition is formed, otherwise, the proposal fails.
- 4) If a coalition is successfully formed, the proposing player increases its aspiration by δ , and all the previous

coalitions that had any player from the new coalition are dissolved.

- 5) If a coalition cannot be formed, the proposing player decreases its aspiration by δ if it is not in any other non-singleton coalition and if it is not already at its singleton coalition valuation. Otherwise, the player's aspiration stays the same.
- 6) Steps 2–5 are then repeated for a set horizon H .

Algorithm 1: Coalition Proposal Algorithm

Input: the set of players, N , and the characteristic function, v .

Set: δ to be a discretization value where the $v(S)$ values for all $S \subseteq N$ lie on the δ -discretized grid.

Initialize: for all $i \in N$,
the player aspiration $a_i \leftarrow a_i^0$, an arbitrary value on the δ -discretized grid, and
the player coalition $C_i \leftarrow \emptyset$.

```

1 for  $h=1:H$  do
2   Activate a player  $i \in N$  uniformly at random.
3   Player  $i$  randomly chooses  $S \subseteq N \setminus \{i\}$ .
4   Player  $i$  proposes to form coalition  $J = S \cup \{i\}$ .
5   if  $\sum_{j \in J} a_j + \delta \leq v(J)$  then
6      $a_i \leftarrow a_i + \delta$ 
7     Break old coalitions of players in  $J$ :
8     For all  $j \in J$ , for all  $k \in C_j, k \neq j$ ,  $C_k \leftarrow \emptyset$ 
9     Form coalition  $J$ :
10     $C_j \leftarrow J$  for all  $j \in J$ 
11  else
12    if  $C_i = \emptyset$  then
13       $a_i \leftarrow \max(v(\{i\}), a_i - \delta)$ 
14    end
15  end
16  if  $a_i = v(\{i\})$  and  $C_i = \emptyset$  then
17     $C_i \leftarrow \{i\}$ 
18  end
19 end

```

IV. ANALYSIS OF THE COALITION PROPOSAL ALGORITHM

In this section, we define the environment state at each iteration of the algorithm and prove how the environment state converges to a core solution of the input game (N, v) , whenever one exists.

Definition 6. The *environment state* $(\mathbf{a}, \mathcal{C})$ at any iteration is the *vector of aspirations* \mathbf{a} and the *coalition structure* \mathcal{C} , where \mathcal{C} is the set of disjoint formed coalitions at this iteration, i.e. $\mathcal{C} = \{C_i : i \in N\}$.

Definition 7. A *feasible environment state* is a pair of an aspiration vector and a coalition structure $(\mathbf{a}, \mathcal{C})$, where $\mathbf{a} \in \mathbb{R}^n$ and \mathcal{C} is a set of disjoint subsets of N such that $a_i \geq v(\{i\})$ for all $i \in N$ and $\sum_{i \in S} a_i \leq v(S)$ for all $S \in \mathcal{C}$. We denote the set of all feasible environment states by $\Omega(v)$.

A. Convergence Cases

We divide the TU games into two cases. The first case is when the game (N, v) has an empty core, $\Pi(v) = \emptyset$, and the second is when the game (N, v) has a non-empty core, $\Pi(v) \neq \emptyset$.

Assumption 1. The discretization value, δ , is chosen such that all the values of $v(S)$ for all $S \subseteq N$ lie on the δ -discretized grid. When $\Pi(v) \neq \emptyset$, the choice of δ must guarantee, as well, that there exists a state $(\mathbf{x}, \rho) \in \Pi(v)$ where all the values of \mathbf{x} lie on the δ -discretized grid.

1) Games with empty core $\Pi(v) = \emptyset$:

Proposition 3. Following the Coalition Proposal algorithm for a game (N, v) with $\Pi(v) = \emptyset$, under Assumption 1, and starting from any environment state $(\mathbf{a}, \mathcal{C}) \in \Omega(v)$, the environment state never converges.

Proof: Let $\mathcal{C}' = \mathcal{C} \cup \{\{i\} : C_i = \emptyset \text{ and } a_i = v(\{i\})\}$.

First, consider the instances where $\mathcal{C}' \in \mathcal{P}(N)$. Accordingly, $\sum_{i \in S} a_i \leq v(S)$ for all $S \in \mathcal{C}'$. Yet, $(\mathbf{a}, \mathcal{C}') \notin \Pi(v)$, then there exists a set $S \subseteq N$ such that $\sum_{i \in S} a_i + \delta \leq v(S)$. Hence, if any player $i \in S$ is randomly activated, there is a positive probability that i will propose S to form, and hence, player i will increase its aspiration by δ .

Second, consider the remaining instances where $\mathcal{C}' \notin \mathcal{P}(N)$. Then, there exists a player $i \in N$ such that $C_i = \emptyset$ and $a_i \geq v(\{i\}) + \delta$. If any such player i is activated, then for any choice of a proposed coalition, a_i will either increase or decrease its aspiration by δ , depending on the success or failure of the proposal. Consequently, for all instances of the environment state $(\mathbf{a}, \mathcal{C})$, there is a positive probability of generating a proposal that would lead to a change in a player's aspiration.

2) Games with non-empty core $\Pi(v) \neq \emptyset$:

Theorem 1. Following the Coalition Proposal algorithm for a game (N, v) with $\Pi(v) \neq \emptyset$, under Assumption 1, and starting from any environment state $(\mathbf{a}, \mathcal{C}) \in \Omega(v)$, the environment state converges to some state $(\mathbf{x}, \rho) \in \Pi(v)$ with probability one.

The next two subsections establish the necessary propositions and arguments to prove this theorem. Henceforth in this section, we will only discuss the games where $\Pi(v) \neq \emptyset$.

B. Feasible States Partition

Definition 8. Given a game (N, v) , we define the following subsets of the feasible states, $\Omega(v)$.

- $\Gamma(v) = \{(\mathbf{a}, \mathcal{C}) \in \Omega(v) : \sum_{i \in N} a_i = K_v, \sum_{i \in S} a_i \geq v(S) \text{ for all } S \subseteq N, \text{ and } \mathcal{C} \notin \mathcal{P}(N)\}$
- $\Upsilon(v) = \{(\mathbf{a}, \mathcal{C}) \in \Omega(v) : \sum_{i \in N} a_i > K_v, \text{ and } \sum_{i \in S} a_i \geq v(S) \text{ for all } S \subseteq N\}$
- $\Psi(v) = \{(\mathbf{a}, \mathcal{C}) \in \Omega(v) : \exists S \subseteq N \text{ such that } \sum_{i \in S} a_i < v(S)\}$

Proposition 4. $\{\Pi(v), \Gamma(v), \Upsilon(v), \Psi(v)\}$ is a partition of $\Omega(v)$.

C. The Steering Sequences

In this subsection, we exhibit possible positive probability sequences of proposals that steer any state $(\mathbf{a}, \mathcal{C}) \in \Omega(v)$ to some state $(\mathbf{x}, \rho) \in \Pi(v)$. The following definitions and propositions are needed for constructing the aforementioned sequences.

Definition 9. For any $\mathbf{a}, \mathbf{x} \in \mathbb{R}^n$, define the **lower-valued indices set** as $L_{\mathbf{a}, \mathbf{x}} = \{i : a_i < x_i\}$, the **upper-valued indices set** as $U_{\mathbf{a}, \mathbf{x}} = \{i : a_i > x_i\}$, and the **equal-valued indices set** as $E_{\mathbf{a}, \mathbf{x}} = \{i : a_i = x_i\}$.

Definition 10. For a given $(\mathbf{a}, \mathcal{C}) \in \Omega(v)$, the **banded set of players** is $B(\mathcal{C}) = \{i : \exists S \in \mathcal{C} \text{ such that } i \in S\}$ and the **free set of players** is $F(\mathcal{C}) = N \setminus B(\mathcal{C})$.

Proposition 5. Assuming $\Pi(v) \neq \emptyset$, then for all $(\mathbf{a}, \mathcal{C}) \in \Psi(v)$ and $(\mathbf{x}, \rho) \in \Pi(v)$, if $S \subseteq N$ is such that $\sum_{i \in S} a_i < v(S)$, then $L_{\mathbf{a}, \mathbf{x}} \cap S \neq \emptyset$.

Proposition 6. If $\Pi(v) \neq \emptyset$, then for all $(\mathbf{a}, \mathcal{C}) \in \Upsilon(v)$ and $(\mathbf{x}, \rho) \in \Pi(v)$, $U_{\mathbf{a}, \mathbf{x}} \cap F(\mathcal{C}) \neq \emptyset$.

The steering sequences are used to show that there is a positive probability of picking a sequence of proposals that lead any feasible state $(\mathbf{a}, \mathcal{C})$ to a core solution of the game when the set of core solutions is non-empty. The steering sequence is constructed upon choosing and fixing any $(\mathbf{x}^*, \rho^*) \in \Pi(v)$ that lies on the δ -discretized grid, which is assumed to exist by Assumption 1. The sequence has two main stages. If $(\mathbf{a}, \mathcal{C}) \in \Upsilon(v) \cup \Psi(v)$, the sequence will follow the first stage proposals till it reaches $\Gamma(v) \cup \Pi(v)$. When the state reaches or starts in $\Gamma(v)$, the sequence will follow the second stage proposals.

First stage: If the state $(\mathbf{a}, \mathcal{C}) \in \Upsilon(v) \cup \Psi(v)$, we choose a feasible proposal (i, S) that transitions this state to a state $(\mathbf{a}^+, \mathcal{C}^+)$ with $\|\mathbf{a}^+ - \mathbf{x}^*\|_1 = \|\mathbf{a} - \mathbf{x}^*\|_1 - \delta$. Choosing (i, S) means that player i proposes to form S . Specifically, the chosen proposals will be as follows:

- If $(\mathbf{a}, \mathcal{C}) \in \Psi(v)$, pick $S \subseteq N$ such that $\sum_{i \in S} a_i < v(S)$. From Proposition 5, $L_{\mathbf{a}, \mathbf{x}^*} \cap S \neq \emptyset$. Pick any $i \in L_{\mathbf{a}, \mathbf{x}^*} \cap S$. Then, **the proposal (i, S) is chosen**. S will be successfully formed and only player i 's aspiration will increase by δ , $a_i^+ = a_i + \delta$. Since $i \in L_{\mathbf{a}, \mathbf{x}^*}$, $\|\mathbf{a}^+ - \mathbf{x}^*\|_1 = \|\mathbf{a} - \mathbf{x}^*\|_1 - \delta$.
- If $(\mathbf{a}, \mathcal{C}) \in \Upsilon(v)$, from Proposition 6, $U_{\mathbf{a}, \mathbf{x}^*} \cap F(\mathcal{C}) \neq \emptyset$. Pick any $i \in U_{\mathbf{a}, \mathbf{x}^*} \cap F(\mathcal{C})$ and any $S \subseteq N$ such that $i \in S$. Then, **the proposal (i, S) is chosen**. S will fail to form and player i is free and has an aspiration $a_i \geq v(\{i\}) + \delta$. Hence, the proposal failure will cause the player's aspiration to decrease by δ , $a_i^+ = a_i - \delta$. Since $i \in U_{\mathbf{a}, \mathbf{x}^*}$, $\|\mathbf{a}^+ - \mathbf{x}^*\|_1 = \|\mathbf{a} - \mathbf{x}^*\|_1 - \delta$.

Note that if $(\mathbf{a}, \mathcal{C}) \in \Upsilon(v) \cup \Psi(v)$, the above proposals will consistently reduce the 1-norm between the new state and \mathbf{x}^* by δ . We can continue performing these proposals until the state reaches $\Gamma(v) \cup \Pi(v)$. This will happen within a finite number of steps, which is at most the 1-norm, between the starting state and the chosen core solution, divided by δ .

Second stage: If the state $(\mathbf{a}, \mathcal{C}) \in \Gamma(v)$, a sequence of proposals is selected to transition said state to a new state $(\mathbf{a}, \rho^*) \in \Pi(v)$, which is proven to exist by Proposition 2.

- If $(\mathbf{a}, \mathcal{C}) \in \Gamma(v)$, then $F(\mathcal{C}) \neq \emptyset$. Pick any $i \in F(\mathcal{C})$ and pick the set $S \in \rho^*$ where $i \in S$. **Choose the proposal (i, S)** . From Proposition 2, $(\mathbf{a}, \rho^*) \in \Pi(v)$. Hence, $\sum_{i \in S} a_i = v(S)$ for all $S \in \rho^*$. Thus, the proposal will fail resulting in $a_i^+ = a_i - \delta$ and $v(S) - \sum_{i \in S} a_i^+ = \delta$. Then **pick the proposal (i, S) again**. Now, the proposal will succeed and $a_i^{++} = a_i^+ + \delta = a_i$ and $v(S) - \sum_{i \in S} a_i^{++} = 0$. Now, $\mathbf{a}^{++} = \mathbf{a}$, hence, the new state is either in $\Gamma(v)$ or $\Pi(v)$. If the state is still in $\Gamma(v)$, then make another iteration of the second stage.

Note that this stage ends in a finite number of steps. That is because if $S \in \rho^*$ is chosen as a part of a proposal in one iteration, then following the second stage proposals, none of the players in S become free again. Hence, the number of iterations of the above two proposals is bounded by the number of elements of ρ^* .

The sequence achieves its target whenever the state reaches $\Pi(v)$. As demonstrated next, once a state reaches $\Pi(v)$, the state can never change using the Coalition Proposal algorithm.

Proposition 7. Any state $(\mathbf{x}, \rho) \in \Pi(v)$ is absorbing in the Coalition Proposal algorithm.

Proof: Since $(\mathbf{x}, \rho) \in \Pi(v)$, then $\sum_{i \in S} x_i \geq v(S)$ for all $S \subseteq N$. Hence, no proposal can succeed, and consequently, no player will be able to increase its aspiration nor can any change in the coalitions happen. Furthermore, ρ is a partition of N ; thus, all players are in non-singleton coalitions, or at their individual valuations, and hence, will not decrease their aspirations from the failed proposals. Accordingly, no proposal at this state can change the aspirations or the coalition structure.

In conclusion, Theorem 1 is proven following the proof of the existence of a finite steering sequence from any state $(\mathbf{a}, \mathcal{C}) \in \Omega(v)$ to a state $(\mathbf{x}, \rho) \in \Pi(v)$. By the Borel-Cantelli lemma, the probability that one such sequence is followed at some iteration goes to 1 as the number of iterations goes to infinity. Finally, Proposition 7 shows that any state $(\mathbf{x}, \rho) \in \Pi(v)$ is an absorbing state.

V. AN ILLUSTRATIVE EXAMPLE

In this section, we apply our algorithm to a multi-agent task allocation setting. We consider self-interested and heterogeneous agents. In task fulfillment problems, it is justifiable to assume that a number of heterogeneous agents are needed to complete a task because of the different resources and capabilities that each agent has [17]. Self-interested agents aim to maximize their own benefits be it through collaborations or individual actions. In such settings, the core solution provides a satisfactory allocation for said agents. Given a core solution, no group of agents can gain more by deviating from the proposed allocation and coalition structure. In addition, the core solution guarantees

the optimal social welfare, the total allocations to the agents equals the maximum welfare value.

A. Setup

Formulate the multi-agent task allocation problem as a game $G = (N, v)$. The set of players $N = A \cup T$, where A is a set of m agents and T is a set of n tasks. F is a set of k features. Each agent is assumed to be equipped with a non-empty subset of these features. Matrix Q is an $m \times k$ binary matrix such that $Q_{a,f} = 1$ if feature f is present in agent a and $Q_{a,f} = 0$ otherwise. For the tasks, R is an $n \times k$ binary matrix of task requirements such that $R_{t,f} = 1$ if task t requires feature f to be present in the group of agents fulfilling the task and $R_{t,f} = 0$ otherwise. The worth of any task t , is characterized by the function $W : T \rightarrow \mathbb{R}$, which is the base value of fulfilling the task. Here, we assume this value to be proportional to the complexity of the task, i.e. the number of features required to fulfill the task. The agents and tasks are set to have a location on a bounded grid. The location is specified by an $(m + n) \times 2$ matrix L .

The characteristic function $v(S)$ for $S \subseteq N$ is formulated as follows:

$v(S) = 0$ if

- $|S| = 1$, the coalition is a singleton, or
- $|T \cap S| \neq 1$, the coalition does not have exactly one task in it, or
- $\min_{f \in F} (\sum_{a \in A \cap S} Q_{a,f} - \sum_{t \in T \cap S} R_{t,f}) < 0$, the union of the features that the agents in the coalition have does not cover all of the required features to fulfill the coalition's task.

Otherwise,

$$v(S) = \max\{0, W(t) - \sum_{a \in A \cap S} \|L_a - L_t\|_1\},$$

where $t = T \cap S$ and L_i is the location vector of player i .

The following are additional assumptions on the setup and the agents' implementation of the algorithm.

- We are assuming a static task allocation setup where the agents are required to fulfill only a subset of the available tasks.
- The current setup assumes full communication, however, the dynamics are readily applied to limited communication setups so long as agents that are part of any positive-valued coalition can communicate with each other.
- To avoid unnecessary proposals, we assume that agents propose only to form coalitions of positive values.
- Tasks are assumed to be passive players, they do not propose or have payoff aspirations.
- Proposing agents broadcast formed coalitions so that other agents know when tasks in their coalitions leave and cause the dissolution of said coalitions.

B. Simulation

Parameters: The setup used to produce the succeeding runs is as follows. On a 9×9 grid, we randomly generated the L matrix for 10 players and 20 tasks. The Q and R matrices were randomly generated accounting for 5 features.

1) *Sample run:* On the left of Fig. 1, a visualization of a randomly generated scenario with the above parameters and the coalition structure produced from the attained core solution after running the Coalition Proposal algorithm is shown. Squares represent tasks and circles represent agents. Agents and tasks in the same coalition have the same color and are connected by a dashed line. Grey circles, whenever they exist, are agents that have no benefit of becoming a part of any coalition and grey squares are tasks that were not chosen to be fulfilled. The right plot of Fig. 1 displays the total aspirations of the players over the communication rounds resulting from running the Coalition Proposal algorithm. The dotted line is the optimal welfare value as solved by an LP representation of the problem.

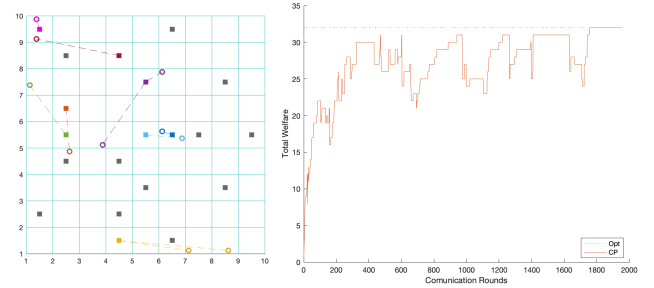


Fig. 1. A visualization of a sample configuration and a plot of the total aspirations of the agents obtained from running the Coalition Proposal algorithm.

2) *Empirical performance for configurations with non-empty core:* To illustrate the performance of our algorithm in the specified setting. We considered 50 random configurations that produced games with non-empty sets of restricted core solutions. The restricted core solutions are core solutions that allocate zero payoffs to all the tasks. This restriction follows our assumption about the tasks being passive players. We ran our algorithm in addition to three best reply algorithms from [10] and [11]. We set the experimentation parameter of the best reply algorithms to 0.05 and Bernoulli agent activation probability to 0.1. The best reply algorithms with experimentation were proven to converge to the core in [10]. The best reply algorithm without experimentation converges very quickly, but possibly, to suboptimal solutions outside the core. However, for the best reply algorithms with experimentation, an activated agent will not experiment before searching for a coalition among all of its possible coalitions outside the current coalition structure that can strictly increase the agent's payoff. This search is computationally expensive. To be able to calculate the best reply, the states of the agents must include the global coalition structure state and the demands of all other agents.

We ran our proposed algorithm, the best reply algorithm from [10], the best reply algorithm with experimentation from [10], and the best reply algorithm with experimentation using only feasible demands from [11] on the aforementioned 50 configurations. Fig. 2(a) shows plots of the average of the relative welfare for the four algorithms. The relative welfare for each configuration is the total aspirations of the players

divided by the optimal welfare value.

3) Empirical performance for general configurations:

Using the specified parameters, the configuration resulted in games with non-empty core around half of the time. However, since our algorithm produces a feasible outcome whenever it is terminated, we explored its performance on 50 general configurations, configurations that may produce games with empty core. Even though the algorithm is proven to cycle whenever the core is empty, it may cycle within allocations close to the optimal values. Hence, the algorithm can still produce a good feasible solution if negotiations are set to be terminated after a specified time. Fig. 2(b) shows the plots of the average relative welfare using our algorithm in addition to the previously discussed best reply algorithms.

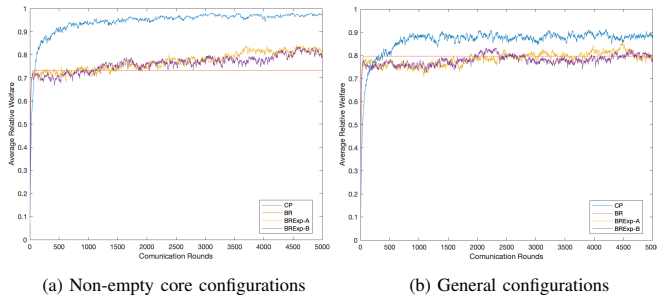


Fig. 2. Plots of the average relative welfare of the agents obtained by running four different dynamics for 50 randomly generated configurations. The “CP” lines are for the Coalition Proposal algorithm, the “BR” lines are for the best reply dynamics [10], the “BRExp-A” lines are for the best reply dynamics with experimentation [10], and the “BRExp-B” lines are for the best reply dynamics with experimentation using only feasible demands [11].

C. Discussion

An additional consideration in multi-agent systems is communication failures. In many systems, a small percentage of message drops are inevitable. To explore our proposed algorithm’s tolerance against communication drops, we have allowed messages that inform agents of the dissolution of a coalition to be dropped with some percentage. For up to 5% of message drops, the algorithm performance on average did not show any significant changes from the perfect communication case. Even though these message drops affect the convergence guarantee, they allow for faster growth of welfare. Specifically in the initial rounds, a false assumption of being in a coalition dissuades the agents from decreasing their aspirations from failed proposals. Hence, the total welfare gets close to the optimal welfare fast.

The empirical simulations that we have performed illustrated the advantages of using our algorithm in this multi-agent task allocation setting. Our proposed algorithm allows for negotiation-based distributed decision-making. The algorithm only involves simple computations from one agent in each round of communication and requires limited local knowledge of the environment. Furthermore, we have observed empirically that the algorithm can still reach the optimal welfare value even in the presence of a small percentage of communication failures.

VI. CONCLUSION

We introduced distributed learning dynamics for coalitional games. We discussed a core solution concept for general TU games. The core solution provides the agents with payoff allocations that preserve individual and coalitional rationality and achieve the optimal social welfare. We proved the convergence of our proposed dynamics to a core solution, whenever one exists. Finally, we illustrated the learning dynamics on a multi-agent task allocation setting and compared it to best reply algorithms and the optimal social welfare value. Our dynamics exhibited desirable performance in simulation for convergence in perfect communication setups as well as in the presence of small percentages of communication failures.

REFERENCES

- [1] J. Mei, C. Chen, J. Wang, and J. L. Kirtley, “Coalitional game theory based local power exchange algorithm for networked microgrids,” *Applied Energy*, vol. 239, pp. 133–141, 2019.
- [2] Y. Du *et al.*, “A cooperative game approach for coordinating multi-microgrid operation within distribution systems,” *Applied Energy*, vol. 222, pp. 383–395, 2018.
- [3] W. Saad, Z. Han, and H. V. Poor, “Coalitional game theory for cooperative micro-grid distribution networks,” in *2011 IEEE International Conference on Communications Workshops (ICC)*, Kyoto, Japan, 2011, pp. 1–5.
- [4] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Başar, “Coalitional game theory for communication networks,” *IEEE Signal Processing Magazine*, vol. 26, no. 5, pp. 77–97, 2009.
- [5] R. J. La and V. Anatharam, “A game-theoretic look at the Gaussian multiaccess channel,” *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 66, pp. 87–105, 2004.
- [6] F. Fele, E. Debada, J. M. Maestre, and E. F. Camacho, “Coalitional control for self-organizing agents,” *IEEE Transactions on Automatic Control*, vol. 63, no. 9, pp. 2883–2897, 2018.
- [7] M. Li, H. Wu, and L. Chen, “Research on multi-robot cooperative task assignment based on cooperative game,” in *2021 4th International Conference on Intelligent Autonomous Systems (ICoIAS)*, Wuhan, China, 2021, pp. 424–429.
- [8] J. G. Martin, F. J. Muros, J. M. Maestre, and E. F. Camacho, “Multi-robot task allocation clustering based on game theory,” *Robotics and Autonomous Systems*, vol. 161, 2023.
- [9] A. Churkin, J. Bialek, D. Pozo, E. Sauma, and N. Korgin, “Review of cooperative game theory applications in power system expansion planning,” *Renewable and Sustainable Energy Reviews*, vol. 145, 2021.
- [10] T. Arnold and U. Schwalbe, “Dynamic coalition formation and the core,” *Journal of Economic Behavior and Organization*, vol. 49, no. 3, pp. 363–380, 2002.
- [11] F. Bistaffa and A. Farinelli, “A fast approach to form core-stable coalitions based on a dynamic model,” in *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Player Technologies (IAT)*, Atlanta, GA, USA, 2013, pp. 122–129.
- [12] D. Hamza and J. S. Shamma, “BLMA: A blind matching algorithm with application to cognitive radio networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 302–316, 2017.
- [13] D. Hamza and J. S. Shamma, “Many-to-one blind matching for device-to-device communications,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, Melbourne, VIC, Australia, 2017, pp. 4988–4993.
- [14] H. H. Nax, “Uncoupled aspiration adaptation dynamics into the core,” *German Economic Review*, vol. 20, no. 2, pp. 243–256, 2019.
- [15] A. Hamed and J. S. Shamma, “Distributed learning dynamics for coalitional games,” 2023, *arXiv:2309.06705 [cs.GT]*.
- [16] H. Peters, “Cooperative games with transferable utility,” in *Game theory: A Multi-Levelled Approach*. Germany: Springer Berlin, Heidelberg, 2008.
- [17] S. Airiau, “Cooperative games and multiagent systems,” *The Knowledge Engineering Review*, vol. 28, no. 4, pp. 381–424, 2013.