The Price of Democracy in Coalition Formation

Georgios Chalkiadakis[*] Edith Elkind[*,†] Maria Polukarov[*]
Nicholas R. Jennings[*]
{gc2,ee,mp3,nrj}@ecs.soton.ac.uk
[*]School of Electronics and Computer Science, University of Southampton, UK
[†]Nanyang Technological University, Singapore

ABSTRACT

Whenever rational agents form coalitions to execute tasks, doing so via a decentralized negotiation process—while more robust and democratic—may lead to a loss of efficiency compared to a centralized solution. To quantify this loss, we introduce the notion of the *Price of Democracy* (PoD), which measures the amount of resources needlessly committed to the task(s) at hand. After defining this concept for general coalitional games, we instantiate it in the setting of weighted voting games, a simple but expressive class of coalitional games that can be used to model resource allocation in multiagent scenarios. We approach the problem of forming winning coalitions in this setting from a non-cooperative perspective, and put forward an intuitive deterministic bargaining process, which exhibits *no delay of agreement* (i.e., the agents are guaranteed to form a winning coalition in round one) and allows for efficient computation of bargaining strategies. We show a tight bound of 3/2 on the PoD of our process if two rounds of bargaining are allowed, and demonstrate that this bound cannot improve with more rounds. We then generalize our bargaining process to settings where multiple coalitions are allowed to be formed, show that this generalization also exhibits no delay of agreement, and discuss the PoD in such settings.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: [multiagent systems]

General Terms

Economics, Measurement, Theory

Keywords

game theory, coalition formation, multilateral bargaining

1. INTRODUCTION

Coalitional games [19] provide a rich framework for the study of cooperation in economics and social sciences, and have also been used widely to model collaboration in multiagent systems (see, e.g., [24, 4, 12]). In such games, rational agents come together into teams (or *coalitions*) to achieve a common goal, and derive individual gains from this activity. Therefore, the study of processes that enable such agents to form coalitions is a matter of great interest. In particular, an important research goal is to design robust coalition formation processes which take into account the opinions of all individuals involved, and are at the same time beneficial from the broader society's perspective. Decentralized negotiation processes to form coalitions are natural candidates to achieve this goal.

Cite as: The Price of Democracy in Coalition Formation, G. Chalkiadakis, E. Elkind, M. Polukarov, N. R. Jennings, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. XXX-XXX.

Consider for example a "Request for Proposals" domain similar to the one presented in [12]. In this domain, as is frequently the case in multiagent systems, there are tasks that await service, and autonomous agents that possess resources and have to come up with proposals to form coalitions to fulfill these tasks. A coalition that is successful in its application is then awarded a payment for the effective completion of its task.

Clearly, several interesting questions arise in decentralized settings such as the one above. First, one can ask which coalitions are most likely to form and how will the agents in these coalitions distribute the payoff obtained by fulfilling the task. Another important issue is how long it will take the agents to form a successful coalition. Finally, a natural parameter of interest is the total amount of resources contributed by the members of winning coalitions. Indeed, from the system designer's perspective, it is desirable that the most resource-efficient solution is chosen, i.e., a coalition that can perform a task using the minimal amount of resources should be awarded the contract. However, in the absence of an all-powerful autocratic center, one cannot just handpick the coalitions expending the minimal amount of resources. Therefore, when agents are selfishly maximizing their own utility, there is likely to be a price to pay in terms of resource efficiency.

In this paper, inspired by the existing work on the price of anarchy pioneered by [11], we propose a concept to substantiate this relative (in)efficiency of coalitions formed by self-interested agents via a decentralized coalition formation approach. In such settings, a coalition cannot form unless it is acceptable to everyone involved, so the decisions to form coalitions are made democratically. Therefore, we call this concept *the price of democracy* (PoD). Intuitively, it compares the value and the cost of outcomes of a coalition formation process to those of an optimal solution. Our definition is applicable both in the case where there is only one coalition to be formed, and in the multiple-coalition scenario.

To illustrate our general framework, we then proceed to study this concept, as well as the rest of the questions above, in the context of *weighted voting games* (WVGs) [26], a simple but expressive class of coalitional games. In these games, each agent has a weight, and a coalition wins if its members' total weight meets a given threshold. Weighted voting games have many applications, but in this context are perhaps most useful for modelling resource allocation in multi-agent systems: an agent's weight can be thought of as the amount of resources available to him, and the threshold indicates the amount of resources required to achieve a task. Thus, a winning coalition naturally corresponds to a team of agents that can successfully cooperate to complete a given task.

The properties of various *cooperative* solutions for WVGs, such as the Shapley value or the core, are well understood. However, while these solutions describe which outcomes, i.e., coalitions and payoff distribution schemes, are stable or fair, they do not explain how these outcomes are likely to arise. Nevertheless, this topic is clearly important from a practical point of view: coalitions do not arise in a vacuum, but are usually the result of some negotiation process among the agents involved. Despite this fact, to date, very few papers have studied coalition formation processes, such as e.g., bargaining, in these games, and, moreover, the existing approaches do not seem to be adequate for our resource allocation scenario (for a more detailed discussion of the related work, see Section 2).

Against this background, we explore the *non-cooperative* approach to coalition formation in weighted voting games. We first focus on the case where there is only one task to be executed. For this scenario, we propose and justify a bargaining process that the agents can use in order to form a winning coalition. This process can be seen as an extensive-form game, and therefore a natural solution concept for it is a subgame-perfect Nash equilibrium (SPNE). We then show that our bargaining process always possesses an efficiently computable SPNE. Moreover, any "reasonable" SPNE of this process (including the one that can be efficiently computed) exhibits no delay of agreement; that is, even though there can be several rounds of negotiation, the outcome proposed in the first round is accepted. Taken together, these two results imply that under this bargaining process, the agents can quickly arrive to an acceptable outcome.

We then study the price of democracy of this bargaining process. We instantiate our generic PoD concept in this domain and demonstrate that for a wide range of parameters of this process its PoD is tighly bounded by 2 if there is only one round of bargaining, and by $3/2$ if two rounds of bargaining are allowed. Interestingly, this bound cannot be improved by adding more rounds. The aforementioned bounds mean that agents may commit twice as much resources as necessary if they can only use one round of bargaining; however, adding just one more round cuts the amount of resources overcommitted to the task by half. Note that even with two rounds of bargaining available, the negotiation will terminate in the first round: intuitively, it is the threat of the first offer being rejected that makes the first proposer suggest a more reasonable—and therefore more efficient—outcome. In other words, the increase in resource efficiency from adding an extra round does not cause any actual loss in time efficiency.

Finally, we generalize our approach to the case where there is more than one task to be executed, and thus multiple coalitions to be formed, extend the description and the analysis of our bargaining process to this setting, and discuss its PoD.

2. BACKGROUND AND RELATED WORK

A *coalitional game with transferable utility* $G = (N, v)$ is given by (i) a set of agents (or *players*) $N = \{1, \dots, n\}$ and (ii) a *characteristic function* $v : 2^N \mapsto \mathbb{R}$ that specifies the *value* $v(C)$ of each *coalition* $C \subseteq N$ [19]. Intuitively, $v(C)$ represents the maximal payoff the members of C can jointly receive by cooperating. A coalitional game is *simple* if $v(C) \in \{0, 1\}$ for any $C \subseteq N$.

A *weighted voting game* (WVG) is a coalitional game G given by a set of agents $N = \{1, \dots, n\}$, their *weights* $\mathbf{w} = (w_1, \dots, w_n)$, $w_i \in \mathbb{R}^+$, and a *threshold* $T \in \mathbb{R}^+$; we write $G = (N; \mathbf{w}; T)$. We use $w(C)$ to denote $\sum_{i \in C} w_i$; we also use similar notation for other n-dimensional vectors considered in the paper. For a coalition $C \subseteq N$, its value $v(C)$ is 1 if $w(C) \geq T$; otherwise, $v(C) = 0$. Without loss of generality, let the value of the grand coalition N be 1 (i.e., $w(N) \geq T$). Throughout, we assume for simplicity that the weights are normalized, i.e., $w_1 + \dots + w_n = 1$.

A thorough discussion of weighted voting games can be found in [26]. From an algorithmic perspective, the problem of computing various cooperative solution concepts for WVGs *without* coalition structures, such as Shapley–Shubik power index [25], Banzhaf power index [2], and the core and other stability-related solution concepts is quite well studied [16, 15, 22, 8]. Also, Elkind *et al.* [7] have recently introduced weighted voting games with coalition structures (i.e., WVGs in which, unlike in the traditional model, several winning coalitions can form simultaneously) and examined the coalitional stability of such games. However, to the best of our knowledge, the only paper that examines non-cooperative aspects of weighted voting is [10]. Moreover, the model in [10] differs from the one we put forward in several important ways, e.g., in [10] agents outside the proposed coalition may also be allocated a share of the total payoff, and have the power to reject a proposal. Clearly, the latter requirement would be very counterintuitive in multi-agent resource allocation domains.

On the other hand, coalitional bargaining in generic coalitional games has been the focus of a growing corpus of research [5, 18, 20, 27]. However, the focus of these papers is usually on the *existence* of an equilibrium rather than its algorithmic properties. Indeed, in some of these papers the proofs are non-constructive, thus providing very little insight into the structure of the equilibria. Moreover, unlike in our work, many of these results rely on discounting and/or the random order of proposers to show no delay of agreement.

Finally, the notion of the price of anarchy was introduced by Koutsoupias and Papadimitriou in their seminal paper [11] in the context of congestion games. In that paper, it is defined as the ratio of the social welfare in a worst-case Nash equilibrium to that in an optimal solution. Subsequently, Anshelevich *et al.* [1] introduced the complementary notion of the price of stability, which compares the performance of the *best* Nash equilibrium to that of the optimal solution. Providing bounds on both of these quantities for various classes of games has been an active stream of research in recent years [6, 23], and has inspired our work on the price of democracy in this paper. However, these concepts were traditionally applied to non-cooperative games and Nash equilibria of those games. In contrast, our concept of the price of democracy, which we will now define, is applicable to *coalitional games* and is not tightly linked to equilibria. Hence, it can be used to measure the efficiency of both cooperative and non-cooperative solutions of such games.

3. THE PRICE OF DEMOCRACY

In this section, we introduce the notion of the *price of democracy* for coalition formation. Our goal is to be able to compare the coalition structure that emerges as the outcome of a "democratic" coalition formation (e.g., bargaining) process to the "optimal" coalition structure, which can be selected in a centralized manner. The important advantage of a negotiation procedure over a centralized solution is that it does not assume the existence of a trusted center, and, furthermore, the resulting outcome must be acceptable to all players involved. However, finding a solution that is broadly acceptable to the players may naturally lead to a loss of efficiency. Our goal here is to quantify this loss.

To build up intuition, we will first consider coalition formation processes that result in a single coalition. Let $\mathcal{G}_N$ be the set of all coalitional games with the set of players N, and let $\mathcal{B} : \mathcal{G}_N \to 2^N$ be a coalition formation process that takes a coalitional game as an input and outputs a coalition $C_{\mathcal{B}} \subseteq N$. For now, to simplify the exposition, we assume that the output of $\mathcal{B}$ is unique: later, we will relax this assumption. A straightforward approach would be to compare the value of the coalition $C_{\mathcal{B}}$ that emerges in $\mathcal{B}$ to $v_{\max} = \max\{v(C) \mid C \subseteq N\}$. While natural in many settings, this approach ignores the issue of overcommitting resources. For example, in simple games, this approach will not be able to distinguish between processes that form a minimal winning coalition and ones that end up forming the grand coalition. Now, in multiagent systems resource efficiency is often an important consideration; so we need a more fine-grained approach, which will allow us to penalize coalition formation processes that form redundant coalitions.

To formally capture this notion of redundancy, we assume that we are given a cost vector $\boldsymbol{c} = (c_1, \ldots, c_n)$. The entries of this vector reflect the cost of employing the players. For example, one can set $c_i = 1$ for all $i = 1, \ldots, n$: this will have the effect of preferring minimal coalitions over non-minimal ones. Alternatively, if the game involves contributing resources to a coalition, one can set $c_i = r_i$, where r_i is the amount of resources available to player i. Now, for each coalition we have two parameters: its value $v(C)$ and its cost $c(C) = \sum_{i \in C} c_i$. Given this, we can measure the *efficiency* $e(C)$ of each coalition by setting $e(C) = \frac{v(C)}{c(C)}$. Then, let $e_{\max} = \max_{C \subseteq N} e(C)$. Using this notation, one can define the price of democracy of a coalitional bargaining process $\mathcal{B}$ on G as

$$PoD(\mathcal{B}, G) = e_{\max}/e(C_{\mathcal{B}}) \tag{1}$$

However, while this approach may be natural in some settings, in many multi-agent resource allocation domains achieving the highest possible value is paramount, and costs are of secondary importance. In this case, it is more natural to use a two-tiered approach, i.e., define $PoD(\mathcal{B}, G)$ as

$$PoD(\mathcal{B}, G) = \begin{cases} +\infty & \text{if } v(C_{\mathcal{B}}) \neq v_{\max} \\ \frac{c(C_{\mathcal{B}})}{c_{\min}} & \text{otherwise,} \end{cases} \tag{2}$$

where $c_{\min} = \min\{c(C) \mid v(C) = v_{\max}\}$. This definition assigns a price of democracy of $+\infty$ to any process that does not produce a coalition with the maximal value. Note that for simple games (and, in particular, WVGs discussed in the next section) both approaches (the efficiency-based and the value-maximizing) are equivalent. However, as we will now see, this is not the case for games with multiple coalitions—even ones based on simple games.

We now consider coalition formation processes whose outputs may consist of more than one coalition. Formally, we say that $CS = (C_1, \ldots, C_k)$ is a *coalition structure* for N if $C_i \subseteq N$, $C_i \cap C_j = \emptyset$ for any $i, j = 1, \ldots, k$, and $\cup_{i=1}^k C_i = N$. In many games, there are restrictions on admissible coalition structures: e.g., coalition structures with more than a fixed number of coalitions or containing coalitions that exceed a certain size are impossible. To formalize this, we assume that, together with a game G with the set of players N, we are given a set of admissible coalition structures over N, denoted by $\mathcal{CS}_N$. We are now interested in coalition formation processes of the form $\mathcal{B} : \mathcal{G}_N \to \mathcal{CS}_N$.

We can straightforwardly extend the notions of cost and value to coalition structures by setting $c(CS) = \sum_{C_i \in CS} c(C_i)$ and $v(CS) = \sum_{C_i \in CS} v(C_i)$ and define $v_{\max}$, $e_{\max}$ and $c_{\min}$ accordingly. Observe, however, that in this setting—even if the underlying game is simple and every formed coalition has a value of 1—the most efficient coalition structure need not be the one with the maximal value: indeed, a coalition structure with 2 coalitions of cost 1 each is more efficient than a coalition structure with 5 coalitions whose costs are 1, 2, 3, 4, and 5. The efficiency-based definition is more appropriate if we are interested in the *average* efficiency of the coalitions formed. However, as argued above, in a typical multi-agent task allocation scenario, maximizing the total number of the tasks completed often has a higher priority than minimizing the resource usage. We therefore define the price of democracy (PoD) as follows:

DEFINITION 1. *Given a game $G = (N, v)$, a set of admissible coalition structures CS_N, and a coalition formation process $\mathcal{B}$, the* price of democracy *of $\mathcal{B}$ with respect to a game G is given by*

$$PoD(\mathcal{B}, G) = \begin{cases} +\infty & \text{if } v(CS_{\mathcal{B}}) \neq v_{\max} \\ \frac{c(CS_{\mathcal{B}})}{c_{\min}} & \text{otherwise,} \end{cases} \tag{3}$$

where $CS_{\mathcal{B}}$ is the coalition structure output by the coalition formation process $\mathcal{B}$, $v_{\max} = \max\{v(CS) \mid CS \in \mathcal{CS}_N\}$ and $c_{\min} = \min\{c(CS) \mid CS \in \mathcal{CS}_N, v(CS) = v_{\max}\}$.

This definition also captures the case of a single coalition: indeed, we can define admissible coalition structures as ones that can contain at most one non-singleton coalition.

Extensions First, if the coalition formation process $\mathcal{B}$ can output more than one coalition (e.g., it is a bargaining process with multiple equilibria), we can define the *pessimistic/optimistic* price of democracy as the maximal/minimal value of PoD, over all possible outputs of $\mathcal{B}$. This distinction is reminiscent of that between the *price of anarchy* and the *price of stability* in non-cooperative games, as discussed in Section 2. In what follows, we focus on the pessimistic variant (which we refer to as "the price of democracy" from now on), as it allows us to quantify the worst-case loss in efficiency and therefore provides reliable performance guarantees in realistic scenarios.

Second, Definition 1 deals with the price of democracy of a fixed coalition formation process with respect to a fixed game. However, one can also define the price of democracy of a process with respect to a certain class of games by taking the maximal value of the PoD of this process over all games in this class. For instance, in what follows we bound the PoD of the bargaining process proposed in Section 4 with respect to the class of all weighted voting games, rather than compute it for individual games. Clearly, such bounds are important as they allow us to estimate the performance of a coalition formation process before we know the details of the game.

Finally, we can easily incorporate time discounting and stochasticity in our model. Due to space constraints, we postpone this discussion to an extended version of this paper.

4. A COALITIONAL BARGAINING PROCESS

We now put forward an intuitive bargaining process for coalition formation in WVG environments. The process is motivated by realistic considerations and requirements prevalent in multiagent settings such as the "Request for Proposals" domain presented in Section 1. In particular, since in such domains there is a need to guarantee the allocation and execution of a task in a timely fashion, it is natural for any such formation process to last for a predefined *finite* number of rounds.

To formally introduce our process, consider a weighted voting game $G = (N; \mathbf{w}; T)$, $|N| = n$. We can assume without loss of generality that $w_1 \geq w_2 \geq \cdots \geq w_n$. For any fixed m, consider the game that consists of m rounds. In the kth round, $k \leq m$, player $k \mod n$ proposes a coalition C^k and a payoff vector $\mathbf{p}^k = (p_1^k, \ldots, p_n^k)$ that satisfies (i) $p_i^k \geq 0$ for all $i = 1, \ldots, n$, (ii) $\sum_{i \in N} p_i^k = v(C^k)$ and (iii) $p_i^k = 0$ for all $i \notin C^k$. Intuitively, $\mathbf{p}^k$ describes a way to distribute the value of the coalition C^k between its members. Then all members of C^k simultaneously and independently decide whether to accept this proposal. If all of them accept, C^k is formed, everyone is paid according to $\mathbf{p}^k$ and the game ends; otherwise, the game moves on to the next round.

Throughout the paper, we focus on the case $m \leq n$, i.e., each player gets at most one chance to propose a coalition. This is a natural restriction, and it substantially simplifies the analysis of the agents' strategic behavior. Observe also that we assume that the agents make proposals in order of decreasing weight. This reflects the fact that in most real-life scenarios the agents who have more power/resources have an advantage during the bargaining process. Bargaining processes with these properties are quite standard in the economics literature[1].

We assume that each agent i has a reservation cost ε, which we essentially consider as the incentive for the agents to participate in the bargaining process.[2] The individual rationality then implies that i will reject any offer in which his payoff p_i is less than ε. We think of the ε value as being relatively small; in particular, we always assume $\varepsilon < 1$, and some (though not all) of our results rely on $\varepsilon \leq 1/n$. Furthermore, we assume that each agent can participate in a winning coalition, i.e., for each agent i there is a coalition C_i^* that satisfies $i \in C_i^*$, $w(C_i^*) \geq T$, $\varepsilon|C_i^*| \leq 1$. The agents are fully strategic, i.e., they know each other's weights as well as the number of rounds m, and choose their strategies so as to optimize their overall utility from the game. Finally, we assume that there is no discounting, but all agents prefer getting a payoff p in the current round to getting the same payoff in a subsequent round. Given a weighted voting game $G = (N; \mathbf{w}; T)$, reservation cost ε, and a number of rounds m, we denote the corresponding bargaining process by $\mathcal{B}(G, \varepsilon, m)$.

Clearly, the process $\mathcal{B}(G, \varepsilon, m)$ can be described as an extensive form game of perfect information. The appropriate equilibrium solution concept for such a game is the Subgame-Perfect Nash Equilibrium (SPNE) [14]. Briefly, a profile of strategies is an SPNE if it induces a Nash equilibrium in every subgame of the extensive form game. Moreover, it is known that every finite game of perfect information has a pure strategy SPNE.

5. NO DELAY OF AGREEMENT

We now show that $\mathcal{B}(G, \varepsilon, m)$ has several desirable properties. Specifically, we show that: (i) $\mathcal{B}(G, \varepsilon, m)$ possesses an efficiently computable SPNE; (ii) in this SPNE, the strategies of all players are *stationary*—i.e., the proposal made in each round does not depend on the details of the proposals rejected in the previous rounds; (iii) *any* stationary SPNE of $\mathcal{B}(G, \varepsilon, m)$ exhibits *no delay of agreement*, i.e., the coalition and the payoff division scheme proposed by the first player are accepted by all members of that coalition.

We start by formalizing the notion of a stationary strategy. Namely, we say a strategy of player j is *stationary* if the outcome he proposes in the jth round is the same no matter what coalitions and payoff division schemes were proposed by players $1, \ldots, j-1$ (observe that if the game has proceeded to the jth round, it means that the proposals made in the previous rounds were rejected). An SPNE is stationary if the strategy of each player in this SPNE is stationary. Focusing on such equilibria—henceforth, *SSPE*s—is a standard approach in the multilateral bargaining literature [17, 21, 5, 20], since stationarity provides a natural and realistic way to narrow the set of proposals that occur in equilibrium. We are now ready to state our main result.

THEOREM 1. *The bargaining process $\mathcal{B}(G, \varepsilon, m)$ described in Section 4 has an SSPE. Moreover, in any SSPE of this process there is no delay of agreement. Finally, suppose that for $i = 1, \ldots, n$ we can represent w_i as $w_i = w'_i/W$, where all w'_i and W are integers. Then we can compute an SSPE in time* $\mathrm{poly}(n, \log W)$.

PROOF. The proof is by backwards induction. We inductively construct a particular SSPE $(s_1, \ldots, s_n)$ of our process; at step j, $j = m, \ldots, 1$, we specify the proposal that player j makes under s_j, and argue that it can be computed within $\mathrm{poly}(n, \log W)$ steps. Furthermore, we show that in an SSPE, if player j makes a proposal $(C, \mathbf{p})$ that is accepted, then $\mathbf{p}$ satisfies $p_i = \varepsilon$ for all $i \in C$, $i < j$. Finally, we show that in any SSPE, the proposal made by player j in round j is accepted by all players in the proposed coalition. Applying the latter result with $k = 1$ implies that there is no delay of agreement in an SSPE.

First consider the case $j = m$. It is clear that m prefers making an acceptable proposal to one that will be rejected. Moreover, any player invited by m will accept the offer as long as he is offered at least ε. Therefore, in equilibrium m will offer exactly ε to every player he invites. Consequently, he can maximize his profit by inviting as few players as possible. This can be achieved by a greedy strategy: m invites $1, \ldots, k_m$, where $k_m = \min\{k \mid \sum_{i=1,\ldots,k, i \neq m} w_i + w_m \geq T\}$. Note that the size of the resulting coalition C^m is at most the size of C_m^* (for definition of C_m^*, see Section 4), so under this proposal m makes positive profit. Note that C^m is not the only proposal m can make as a part of an SSPE strategy, but any proposal $(C, \mathbf{p})$ that m can make in an SSPE satisfies $|C| = |C^m|$, $p_i = \varepsilon$ for $i \in C$. We can now partially specify s_m by stipulating that m selects C^m greedily and offers ε to any $i \in C^m$; in the end, we will complete the description of the strategy s_m by explaining which proposals m would accept at each of the rounds $1, \ldots, m-1$.

Now, suppose that the statement has been proved for all k, $j < k \leq m$, and consider the problem faced by agent j in round j. Fix an SSPE of the bargaining process, and consider the proposal $(C^{j+1}, \mathbf{p}^{j+1})$ that $j+1$ makes in the $(j+1)$st round under this SSPE (as this equilibrium is stationary, this proposal is independent of the one that j is about to make). If this proposal includes j, then he can make an acceptable proposal by proposing $(C^{j+1}, \mathbf{p}^{j+1})$ himself. Indeed, by inductive assumption, $(C^{j+1}, \mathbf{p}^{j+1})$ is accepted in round $j+1$, so it satisfies individual rationality. Moreover, if any agent $i \in C^{j+1}$ rejects this proposal in round j, the game will proceed to round $j+1$, where this proposal would be accepted and i will receive the same payoff that is currently offered, a contradiction. On the other hand, if $j \notin C^{j+1}$, then j can propose coalition $C' = C^{j+1} \setminus \{j+1\} \cup \{j\}$, and a payoff vector $\mathbf{p}'$ such that $p'_i = p_i^{j+1}$ for all $i \neq j, j+1$ and $p'_j = p_{j+1}^{j+1}$, $p'_{j+1} = 0$. As $w_j \geq w_{j+1}$, we have $w(C') \geq T$. On the other hand, as $\varepsilon \leq p_{j+1}^{j+1} = p'_j$, this proposal is individually rational. Moreover, by the same argument as in the previous case, all agents in C' will accept this proposal. We conclude that j can always make a proposal that is accepted and guarantees him at least ε. Moreover, he weakly prefers making such a proposal to one that is rejected. Indeed, by our inductive assumption, if j's proposal is rejected, then the game will proceed to the next round, where player $j+1$ will make a proposal; this proposal will be accepted and the game will stop. Then, by our inductive assumption, player j will receive at most ε under $(j+1)$'s proposal.

Now, suppose that j is considering proposing $(C, \mathbf{p})$ in the current round. Clearly, C will have to satisfy $j \in C$, $w(C) \geq T$,

[1]Of course, such an ordering would be impossible to impose in environments with unknown opponents' weights. Different decentralized bargaining processes, requiring a random order of proposers, would be required in such settings. However, we limit our study here to problems with perfect information regarding opponent weights, as is the norm in the WVGs literature.

[2]The ε reservation cost is not to be confused with the notion of the reservation *value* (the value of an agent's singleton coalition) in generic coalitional games, nor does it represent the "complete costs" sustained by an agent for serving a task—these "complete costs", in fact, correspond to the resources expended by the agents (i.e., the agents' weights).

$\varepsilon|C| \le 1$. Now, for any agent $i \ne j$, $i \in C$, it has to be the case that $p_i = \max\{\varepsilon, p_i^{j+1}\}$. Indeed, any offer of less than ε will be rejected by any $i \in C$, and if $i \in C^{j+1}$ and he is offered less than p_i^{j+1} in round j, he will reject the current proposal, as he expects to get more in the next round. On the other hand, setting $p_i > \max\{\varepsilon, p_i^{j+1}\}$ is not a best response to agent $(j+1)$'s strategy: clearly, i will accept a payoff of $\max\{\varepsilon, p_i^{j+1}\}$ at this stage, so j can improve his own payoff by offering $\max\{\varepsilon, p_i^{j+1}\}$ to i. Note that this implies that j will never offer to pay more than ε to any $i < j$, as, by the inductive assumption, $j+1$ never does. Moreover, it follows that the payment vector $\mathbf{p}$ is uniquely determined by the choice of coalition C. Therefore, given the proposal made by $j+1$, the proposal made by j can be determined by solving an instance of the Knapsack problem, where the goal is to construct a coalition of weight at least T and the smallest possible cost, where the costs are given by $\max\{\varepsilon, p_i^{j+1}\}$, $i = 1, \dots, n$. In particular, assuming that in the $(j+1)$st round $j+1$ behaves according to s_{j+1} (recall that by our inductive assumption, the proposal that $j+1$ should make under s_{j+1} has already been specified), we can define j's proposal under s_j to be the lexicographically first solution to the corresponding instance of Knapsack. Now, by construction it is easy to see that each p_i^{j+1}, $i = 1, \dots, n$, is of the form $x\varepsilon$ or $1 - y\varepsilon$ for some x, y that are polynomial in n, i.e., p_i^{j+1} can take at most polynomially many different values. In such settings, Knapsack can be solved within $\mathrm{poly}(n, \log W)$ steps [13].

Now, to fully describe the strategy profile $(s_1, \dots, s_n)$, it remains to specify how each player should react to the proposals made in rounds $1, \dots, n$. This is done as follows. Let $(C^j, \mathbf{p}^j)$ be the proposal made by player j in round j under s_j. Then in that round player $i \ne j$ should accept an invitation to join a coalition C^j if he is offered at least $\max\{p_i^{j+1}, \varepsilon\}$. It is easy to see that this strategy profile is indeed an SSPE. $\square$

We can extend our result to the case when the agents' reservation costs may differ, i.e., the reservation cost of agent i is given by ε_i. We still assume that each agent can form a winning coalition, i.e., for each i there exists a C, $i \in C$, such that $\varepsilon(C) \le 1$. In this model, we show that there is no delay of agreement if either (i) all ε_i are sufficiently small or (ii) the sequence $\varepsilon_1, \dots, \varepsilon_n$ is non-decreasing; moreover, these assumptions appear to be necessary.

We conclude this section by stating some SSPE strategies' properties which follow directly from the proof of Theorem 1, and will be useful for the analysis presented in the next section.

PROPOSITION 1. *If a proposer j makes a proposal $(C, \mathbf{p})$ that is accepted, then $\mathbf{p}$ satisfies: (i) $p_i = \varepsilon$ for all $i \in C$, $i < j$ and (ii) $|\{i \in C | p_i > \varepsilon\}| \le 1$. Moreover, in any SSPE of $\mathcal{B}(G, \varepsilon, m)$, each player's decision whether to accept a proposal made in the current round only depends on the payment offered to this player (and not on the payments offered to others in this round).*

6. THE PRICE OF DEMOCRACY IN WVGS

We now proceed to instantiate our generic price of democracy concept in the weighted voting games setting. A natural measure of (in)efficiency for a coalition formation process in this domain is the amount of resources needlessly committed to the task at hand. Hence, in this setting it is natural to identify the cost of the ith player c_i with his weight w_i. Furthermore, we identify the outcomes of our bargaining process with the set of its SSPEs: this approach is standard in the economic literature, as the non-stationary SPNE are rather unnatural, and therefore unlikely to arise. By instantiating Definition 1 in this setting, and taking into account the multiplicity of SSPE of $\mathcal{B}$, we obtain the following definition:

DEFINITION 2. *Given a weighted voting game $G = (N, \mathbf{w}, T)$ and the bargaining process $\mathcal{B} = \mathcal{B}(G, \varepsilon, m)$ defined above, the* price of democracy *in G under $\mathcal{B}$ is defined as*

$$PoD(\mathcal{B}, G) = \max_{C \in CEQ(\mathcal{B}, G)} \frac{w(C)}{w_{\min}}, \tag{4}$$

where $w_{\min} = \min\{w(C) \mid w(C) \ge T\}$ and $CEQ(\mathcal{B}, G)$ is the set of all coalitions that can form in SSPE of $\mathcal{B}$.

Observe that the case $PoD = +\infty$ never arises here, since by Theorem 1 a winning coalition always forms. Notice also that, since bargaining *equilibria* are now considered, the PoD in this setting can be viewed as an instantiation of the Price of Anarchy concept.

We can now evaluate the price of democracy of $\mathcal{B}(G, \varepsilon, m)$. First, observe that unless we require $w_i < T$ for all i, no meaningful bound on $PoD(\mathcal{B}, G)$ can be derived. Indeed, if this condition does not hold, it may happen that $w_1 = kT$ for some $k > 1$, $w_2 = T$ (and the weights of all other agents are arbitrary). In this case, agent 1 can propose the coalition that includes himself only, resulting in $PoD = k$ irrespective of the number of rounds. Moreover, in practice, it is often the case that no agent can achieve the goal on his own, so this assumption is quite realistic. Therefore, from now on, we assume that $w_i < T$ for $i = 1, \dots, n$. Furthermore, it is easy to see that in WVGs with $n = 2$ agents, the price of democracy is equal to 1: both agents must be included in a winning coalition, since none of them achieves T on its own. Thus, henceforth we assume the existence of at least 3 agents.

For this setting, we consider the cases $T > 1/2$ and $T \le 1/2$ separately. In the former case, we show that, even though there is no delay of agreement, the number of bargaining rounds is crucial: if $m = 1$, we show a tight upper bound of 2, and if $m > 1$, we show a tight upper bound of 3/2. For $T \le 1/2$, we show an upper bound of 2 which is tight irrespective of the number of rounds.

Given a game G, fix $C_{opt} \in \mathrm{argmin}\{w(C) \mid w(C) \ge T\}$. We start by proving a simple upper bound on $PoD(\mathcal{B}, G)$ for any weighted voting game G.

LEMMA 1. *For any WVG G and any $C \in CEQ(\mathcal{B}, G)$, we have $w(C) - w(C_{opt}) \le w(C) - T < \min_{i \in C} w_i$.*

Proof : Suppose otherwise and fix C such that $w(C) - T \ge \min_{i \in C} w_i$ and $j \in \mathrm{argmin}_{i \in C} w_i$. We have $w(C) - w_j \ge T$, so $C \setminus \{j\}$ is a winning coalition. Hence, instead of proposing C, player 1 (the proposer) can invite players in $C \setminus \{j\}$ and offer to pay them the same amount as they were getting in C. By Proposition 1, this proposal will be accepted. On the other hand, this allows the proposer to save at least ε, a contradiction. $\square$

COROLLARY 1. *If a WVG G satisfies $w_i < \frac{T}{q}$ for all $i \in N$ and some $q \ge 1$, then $PoD(\mathcal{B}, G) \le \frac{1+q}{q}$.*

Indeed, we have

$$PoD(\mathcal{B}, G) = \max_{C \in CEQ(\mathcal{B}, G)} \frac{w(C)}{w(C_{opt})} \le \frac{w(C_{opt}) + \min_{i \in C} w_i}{w(C_{opt})}$$
$$= 1 + \frac{\min_{i \in C} w_i}{w(C_{opt})} \le 1 + \frac{\min_{i \in C} w_i}{T} < 1 + \frac{T/q}{T} = \frac{1+q}{q}.$$

Note that $\frac{1+q}{q}$ decreases in q. Thus, Corollary 1 essentially means that if the weights of the agents are relatively small, then the price of democracy is close to 1. In addition, by substituting $q = 1$ one may conclude the following.

COROLLARY 2. *For any WVG G we have $PoD(\mathcal{B}, G) \le 2$.*

We are now ready to present our main results on $PoD(\mathcal{B}, G)$. First, we consider the case $T > \frac{1}{2}$. We show that the upper bound of 2 is tight if only one bargaining round is allowed (see Example 1). For $m \geq 2$, we prove a tight upper bound of $\frac{3}{2}$ (Theorem 2).

EXAMPLE 1. Consider a WVG G with $T > \frac{1}{2}$, $m = 1$, and weights $w_1 = T - \delta, w_2 = T - 2\delta, w_3 = \ldots = w_n = \frac{3\delta}{n-2}$, for some arbitrarily small $\delta > 0$. To form a winning coalition, player 1 has to propose either to player 2 or to $\frac{n-2}{3}$ players $i > 2$. Player 1 prefers to invite as few other players as possible, so proposes to player 2. Player 2 accepts, since there will be no other opportunity for him to get into a winning coalition; thus, the equilibrium outcome is $\{1, 2\}$. An optimal coalition, however, consists of player 2 and $\frac{2(n-2)}{3}$ players $i > 2$, implying $PoD(\mathcal{B}, G) = \frac{2T-3\delta}{T} \xrightarrow[\delta \to 0]{} 2$.

If more than one round of bargaining is allowed, then player 2 has an opportunity to propose in the second round. Hence, in the equilibrium outcome which is formed with no delay of agreement (by Theorem 1, there exists one), player 1 has to pay player 2 at least the amount the latter can guarantee to himself later on. Thus, player 1 might consider to propose to $\frac{n-2}{3}$ players $i > 2$ rather than to 2. However, if ε, the minimum amount he has to offer to each one of them, is large enough, such a coalition will *not* be formed. For example, if $\varepsilon > \frac{3}{n-2}$, then player 1 would have to pay the players $3, \ldots, n$ more than 1 in total, implying a negative payoff for himself. Hence, for relatively large values of ε it can be the case that $PoD(\mathcal{B}(G, \varepsilon, m), G)$ is close to 2 even for $m > 1$.

However, as discussed before, we interpret ε as a participation-incentivizing reservation cost. Thus, in the context of our setting, it is natural to assume that ε is sufficiently small. We will now derive stronger bounds on the price of democracy under the assumption $\varepsilon \leq \frac{1}{n}$. In particular, we will show that under this assumption the price of democracy can be significantly reduced by allowing only one additional round of bargaining. Moreover, one *cannot* achieve a better result by allowing more than one additional rounds.

THEOREM 2. *Suppose that $\varepsilon \leq \frac{1}{n}$ and G satisfies $T > \frac{1}{2}$ and $m \geq 2$. Then $PoD(\mathcal{B}, G)$ is tightly bounded by $\frac{3}{2}$.*

Proof : Let $C \in \text{argmax}\{w(S) \mid S \in CEQ(\mathcal{B}, G)\}$. By Lemma 1, if C includes some i with $w_i \leq \frac{T}{2}$, then $PoD(\mathcal{B}, G) = \frac{w(C)}{w(C_{opt})} < \frac{w(C_{opt}) + \min_{i \in C} w_i}{w(C_{opt})} \leq 1 + \frac{T/2}{T} = \frac{3}{2}$, as required.

On the other hand, if C only includes agents with $w_i > \frac{T}{2}$, it cannot contain any agent $i \geq 4$, since $w_i \leq w_4 \leq \frac{1}{4} < \frac{T}{2}$. Also notice that C consists of *only two* agents. Indeed, otherwise we have $C = \{1, 2, 3\}$. Since $w_1 \geq w_2 \geq w_3 > \frac{T}{2}$, it follows that any pair of agents from 1,2,3 can form a winning coalition; thus the proposer would be able to increase his payoff by at least ε by excluding one partner, a contradiction. Now, note that $w(\{2, 3\}) \leq w(\{1, 3\}) \leq w(\{1, 2\})$ and, by Theorem 1, there exists an SSPE with no delay of agreement. Thus C is either $\{1, 2\}$ or $\{1, 3\}$, and it forms in the first round of bargaining.

Consider first the case of $C = \{1, 2\}$. Here player 1 has to pay player 2 at least $1 - \varepsilon$ since any other offer will be rejected by 2 (as he can guarantee to himself a payoff of $1 - \varepsilon$ in the second round by offering ε to player 1 who would obviously accept). Note also that any other player $i > 2$ would accept from player 1 any offer of $p_i \geq \varepsilon$ (since player 2, if he gets to propose in the second round, will not invite i at all). Now, since $\varepsilon \leq \frac{1}{n}$, we have that $(n-2)\varepsilon \leq \frac{n-2}{n} \leq 1 - 2\varepsilon$. Hence, by replacing player 2 with a subset of $N \setminus \{1, 2\}$, player 1 could increase his payoff by at least ε. Hence, the only reason for player 1 to invite player 2 is that player 2 is a veto player, i.e., $w(N \setminus \{2\}) < T$. Then, it is also the case that $w(N \setminus \{1\}) < T$. Therefore, $C_{opt} = C = \{1, 2\}$, implying $PoD(\mathcal{B}, G) = 1$.

Consider now the case of $C = \{1, 3\}$. If both agents 1 and 3 belong to C_{opt}, then $C_{opt} = \{1, 3\}$ and $PoD(\mathcal{B}, G) = 1$. If none of them is in C_{opt}, then $w(N \setminus \{1, 3\}) \geq w(C_{opt}) \geq T$, implying $w_1 + w_3 < T$, in contradiction to $\{1, 3\}$ being a winning coalition. So, assume $|C \cap C_{opt}| = 1$.

Recall that $w_1 \geq w_2 \geq w_3 > \frac{T}{2}$ yields $w_2 + w_3 > T$, implying $w(N \setminus \{2, 3\}) < T$. Hence, if a winning coalition contains only one of the agents 1 and 3, it must include agent 2 as well. Thus, since $\{2, 3\}$ is a winning coalition and $w_2 + w_3 \leq w_1 + w_2$, we conclude that $C_{opt} = \{2, 3\}$. We complete the proof by showing that $PoD(\mathcal{B}, G) = \frac{w_1 + w_3}{w_2 + w_3} \leq \frac{3}{2}$.

Assume this is not the case. Then,

$$\frac{w_1 + w_3}{w_2 + w_3} > \frac{3}{2} \Rightarrow w_1 - \frac{3}{2}w_2 > \frac{1}{2}w_3 > \frac{T}{4}. \tag{5}$$

On the other hand,

$$w_3 > \frac{T}{2} \text{ and } \sum_{i \in N} w_i = 1 \Rightarrow w_1 + w_2 < 1 - \frac{T}{2} \Rightarrow w_1 < 1 - \frac{T}{2} - w_2. \tag{6}$$

Now, (5) coupled with (6) implies $1 - \frac{T}{2} - w_2 - \frac{3}{2}w_2 > \frac{T}{4} \Rightarrow w_2 < \frac{4-3T}{10}$, which, by $w_2 > T/2$, implies $T < 1/2$, a contradiction.

To show that this bound is tight, consider the following example. Let $w_1 = T - 4\delta$, $w_2 = \frac{T}{2} + 2\delta$, $w_3 = \frac{T}{2} + \delta$, and $w_4 = \ldots = w_n = \frac{\delta}{n-3}$ (for $n = 3$, set $w_1 = T - 3\delta$). Regardless of the number of rounds allowed, there is an equilibrium outcome consisting of either players 1 and 2 or players 1 and 3. The optimal winning coalition is obviously given by $\{2, 3\}$. Thus, $PoD(\mathcal{B}, G) \geq \frac{\frac{3}{2}T - 3\delta}{T + 3\delta} \xrightarrow[\delta \to 0]{} \frac{3}{2}$. □

We now discuss the complementary case of $T \leq \frac{1}{2}$. In this case, the upper bound of $3/2$ no longer holds. Indeed, we will now show that when T is small, the price of democracy can be arbitrarily close to 2 even with many rounds of bargaining. We present the construction for the case $T \geq \frac{1}{n-1}$. Observe that it is always the case that $T > \frac{1}{n}$: otherwise, $\sum_{i \in N} w_i < nT \leq n \cdot \frac{1}{n} = 1$. The analysis for the case $\frac{1}{n} \leq T < \frac{1}{n-1}$ can be done in a similar fashion, and is omitted for brevity.

EXAMPLE 2. Consider a WVG G with $n \geq 4$, $T \geq \frac{1}{n-1}$ and $m < \frac{1}{T} - 1$, with agents' weights being given by: $w_1 = T - \delta, w_2 = T - 2\delta, \ldots w_{p-1} = T - (p-1)\delta$, and $w_p = 1 - (p-1)T$, $w_{p+1} = \ldots = w_n = \frac{p(p-1)\delta}{2(n-p)}$, where $p \in \arg\min_{3 \leq q \leq n-1}\{q | T \geq \frac{1}{q}\}$, and $\delta > 0$ is arbitrarily small.

Since $m < p - 1$, $\{1, p-1\}$ is an equilibrium winning coalition that can be obtained with no delay of agreement (player $p-1$ will accept from player 1 any offer of at least ε). On the other hand, there is a winning coalition consisting of player $p-1$ and a (sufficiently small) subset of players $i \geq p+1$. Indeed, we have $\frac{p(p-1)}{2} \geq p - 1$, implying $w_{p-1} + \sum_{i=p+1}^{n} w_i \geq T$). Hence, $PoD(\mathcal{B}, G) \geq \frac{2T-3\delta}{C_{opt}} > \frac{2T-3\delta}{T + \frac{p(p-1)\delta}{2(n-p)}} \xrightarrow[\delta \to 0]{} 2$.

However, one can argue that if $T \leq 1/2$, the formation of a single coalition is not the only possible outcome. Indeed, if there is more than one task to fulfill, and the agents have enough resources to handle multiple tasks, it is plausible (and, moreover, desirable from the efficiency perspective) that multiple coalitions will form. This scenario is not captured by the bargaining process of Section 4, or the analysis in this section. Therefore, in the next section we discuss how to extend our approach to the multiple-tasks scenario.

7. WEIGHTED VOTING GAMES WITH MULTIPLE WINNING COALITIONS

In this section, we extend our discussion to deal with multiple winning coalitions in WVGs. We focus on the scenario where there is an *a priori* upper bound on the number of coalitions to be formed (as opposed to agents forming coalitions until the total weight of the remaining agents is lower than T). This is a natural requirement in many multiagent environments with a fixed number of tasks.

We extend the bargaining procedure described in Section 4 to this setting, and show that the resulting procedure also exhibits no delay of agreement. That is, *any round of bargaining results in the formation of a winning coalition*. We then discuss the price of democracy in the context of multiple winning coalitions in WVGs.

7.1 Bargaining to Form Multiple Coalitions

Consider a weighted voting game $G = (N; \mathbf{w}; T)$ with $N = \{1, \ldots, n\}$ and $w_1 \geq \ldots \geq w_n$. Suppose that there are d tasks to be allocated, so, potentially, there can be d disjoint winning coalitions. That is, the set of admissible coalition structures consists of all structures that contain at most d non-singleton coalitions. Given G and d as above, a reservation cost ε, and m rounds, $d \leq m \leq n$, we now define a *bargaining process to form multiple coalitions* $\mathcal{B}_d(G, \varepsilon, d, m)$. $\mathcal{B}_d$ proceeds as follows.

We say that agent i is *active* in round r, $1 \leq r \leq m$, if i has not joined any of the coalitions formed in rounds $1, \ldots, r-1$. Let N^r denote the set of active agents in round r and let $P^r = \{i \in N^r \mid i \text{ has not proposed in rounds } 1, \ldots, r-1\}$ be the set of potential proposers. In the rth round, agent $j^r = \min\{i \mid i \in P^r\}$ has to propose a coalition C^r and a payoff vector $\mathbf{p}^r = (p_i^r)_{i \in N^r}$ that satisfies (i) $p_i^r \geq 0$ for all $i \in N^r$, (ii) $\sum_{i \in N^r} p_i^r = v(C^r)$ and (iii) $p_i^r = 0$ for all $i \notin C^r$. If j^r's offer is accepted by all agents in C^r, then C^r forms, and the set of active agents in the next round is set to be $N^{r+1} = N^r \setminus C^r$. Otherwise the set of active agents does not change, i.e., we have $N^{r+1} = N^r$ and $P^{r+1} = P^r \setminus \{j^r\}$. The process can terminate in less than m rounds if d coalitions have been formed, or no more winning coalitions can form, i.e., $w(N^{r+1}) < T$.

7.2 No Delay of Agreement

Here we show that the bargaining procedure described in this section shares many of the good qualities of $\mathcal{B}(G, \varepsilon, m)$. Specifically, under the same assumptions as before, we prove the following results: (i) the $\mathcal{B}_d$ process has an efficiently computable SSPE, and (ii) any SSPE of $\mathcal{B}_d$ exhibits *no delay of agreement*, i.e., the coalition and the payoff division scheme proposed at each bargaining round are accepted by all members of that coalition.

First we extend the notion of stationary strategy to this setting: A strategy of player j^r (the proposer at round r) is *stationary* if, given the set of currently active players N^r and the number of uncompleted tasks $d^r (\leq d)$, the outcome j^r proposes in the rth round is the same no matter what coalitions and payoff division schemes were proposed by players whose proposals were rejected in rounds $1, \ldots, r-1$. We say that an SPNE is stationary if the equilibrium strategy of each player is stationary.

We now present the main result of this section:

THEOREM 3. *The bargaining process $\mathcal{B}_d(G, \varepsilon, d, m)$ has an SSPE. Moreover, in any SSPE there is no delay of agreement. Finally, suppose that for $i = 1, \ldots, n$ we can represent w_i as $w_i = w_i'/W$, where all w_i's and W are integers. Then, an SSPE can be computed in time $\mathrm{poly}(n, \log W)$.*

Proof Sketch: The proof is by induction (see Theorem 1 for $d = 1$), and uses the following idea. Suppose that the statement holds for any game $\mathcal{B}_{d'}(G', \varepsilon, d', m')$ with $d' \leq d-1$, $m' \leq m$, $N' \subseteq N$. If there is an SSPE in which the player 1's (stationary) strategy in the first round, $(C^1, \mathbf{p}^1)$, is efficiently computable and accepted by the members of C^1, then the winning coalition C^1 forms, and we have to play a bargaining game $\mathcal{B}_{d-1}(G', \varepsilon, d' = d-1, m' = m-1)$, where G' is given by $G(N' = N \setminus C^1; \mathbf{w}\big|_{N'}; T)$, with $\mathbf{w}\big|_{N'}$ denoting the restriction of $\mathbf{w}$ to N'. This game has all the desirable properties, by inductive assumption. Therefore, it suffices to prove that there is an SSPE in which the first winning coalition forms in the first round.

For that, the proof uses backwards induction as follows. Consider the subgame that starts at round j and assume that no winning coalition has formed so far. For any $j = m, \ldots, 1$, we inductively construct a particular SSPE $(s_1, \ldots, s_n)$ of our process; at step j, we specify the proposal that player j makes at this stage under s_j and argue that it can be computed within $\mathrm{poly}(n, \log W)$ steps. Furthermore, we show that in any SSPE, the proposal made by player j in round j is accepted by all agents in the proposed coalition. Applying the latter result with $j = 1$ implies that the first coalition forms in the first round. □

The result is extendable to the case with different reservation costs, where they are either sufficiently small or non-decreasing. We defer the related proofs to an extended version of the paper.

7.3 The PoD in the Multiple-Task Scenario

We can adapt our definition of the price of democracy to the setting of WVGs with multiple tasks as follows:

$$PoD(G, \mathcal{B}) = \begin{cases} +\infty & \text{if } v(CS_{\mathcal{B}}) < v_d \\ c(CS_{\mathcal{B}})/c_{\min} & \text{otherwise,} \end{cases} \tag{7}$$

where $v_d = \min\{d, \max_{CS \in \mathcal{CS}_N} v(CS)\}$ and $c_{\min} = \min\{c(CS) \mid v(CS) = v_d\}$.

It turns out that the bargaining process $\mathcal{B}_d$ described above does not always form the optimal number of winning coalitions, i.e., there are G, d such that $PoD(G, \mathcal{B}_d) = +\infty$. For example, consider a WVG G with $\mathbf{w} = (T - 2\delta, T - 3\delta, 3\delta, \delta, \delta)$ for some small enough δ and $d = 2$. We claim that for any $m \geq 2$ and any small enough ε we have $PoD(G, \mathcal{B}(G, \varepsilon, d, m)) = +\infty$. Indeed, the optimal coalition stucture consists of two coalitions: namely, $\{1, 4, 5\}$ and $\{2, 3\}$. Now, note that if in the first round 1 invites 3 and offers him ε, 3 will accept: otherwise, in the next round 2 will propose to 1, and the game will be over. Moreover, 1 prefers inviting 3 to inviting 4 and 5, as this leaves him with a higher share of payoffs. Hence, in the first round $\{1, 3\}$ will form, thus precluding formation of the optimal coalition structure.

Moreover, this issue is not unique to our bargaining process. Indeed, for *any* negotiation process $\mathcal{B}$ with polynomially computable strategies there is a game G such that $PoD(G, \mathcal{B}(G, \varepsilon, 2, m)) = +\infty$ for any m. To see this, consider a game G where weights correspond to an instance of PARTITION, a classic NP-complete problem given by a list of numbers $\{a_1, \ldots, a_n\}$ with $\sum_{i=1}^n a_i = 2K$; it is a "yes"-instance if there is a $S \subseteq N$ such that $\sum_{i \in S} a_i = K$. It is not hard to see that if $PoD(G, \mathcal{B}(G, \varepsilon, 2, m)) < \infty$ for any such G, one can use $\mathcal{B}$ to solve PARTITION in polynomial time. This result can be generalized to $d > 2$; due to space restrictions, we omit the details.

One can show, however, that if the supply of resources exceeds demand by at least a factor of 2, i.e., $w(N) \geq 2dT$, then the price of democracy is guaranteed to be finite (and, in fact, does not exceed 2). The proof follows immediately from our no delay of agreement result and the fact that the weight of any coalition formed in an equilibrium of $\mathcal{B}_d$ is at most $2T$.

8. CONCLUSIONS AND FUTURE WORK

This paper introduced the notion of the *price of democracy* (PoD), an intuitive and conceptually appealing measure of the relative inefficiency of decentralized processes in forming coalitions to serve tasks. After discussing several approaches to defining the PoD, we focused on one that prioritizes value over cost, by setting $PoD = +\infty$ if the coalition structure output by the coalition formation process does not attain the maximal value, and to the ratio of the cost of the coalition structure output by this process to that of the most cost-effective coalition structure with maximal value otherwise.

To illustrate the applicability of this notion, we studied it in the context of weighted voting games, a simple but expressive class of games that can be used to model resource allocation in multi-agent domains. We approached the coalition formation process in such games from a non-cooperative perspective, and presented an intuitive, multi-round coalitional bargaining process for such games. We showed that this process exhibits several desirable properties: there is no delay of agreement in equilibrium and the equilibrium strategies of all agents are efficiently computable. To quantify the price of democracy of this process, we instantiated our proposed generic PoD metric for this class of games and showed a simple upper bound of 2 for the PoD in a wide range of bargaining scenarios, and a tight bound of $3/2$ for our multi-round process (which, moreover, cannot improve after two rounds of bargaining), when the players' reservation values are relatively small. We then discussed WVGs with multiple coalitions. We presented a generalized version of our bargaining method for a variant of this setting with a fixed target number of winning coalitions and showed that it also results in no delay of agreement. Finally, we argued that in this scenario the PoD can be infinite if there is no surplus of resources, but having abundant resources allows us to bound the PoD by 2.

As weighted voting games are usually viewed as cooperative games, one might think that it is more natural to define the price of democracy in terms of solution concepts characteristic of cooperative games, such as, e.g., the core. In particular, one may want to compare an outcome in the core to the most efficient outcome. However, this approach is problematic for several reasons. First, the traditional definition of the core of the weighted voting games presumes that the grand coalition will form, so comparing the core outcomes to the optimal outcome is not very informative: the price of democracy will be simply the inverse of the weight of the optimal coalition. Second, the core of a weighted voting game can be empty; in fact, it is empty whenever there are no veto agents, a situation that is not unusual for this class of games. These issues can be mitigated somewhat by allowing for coalition structures, as done in [7], where the core for WVGs with coalition structures (CS-core) is introduced. However, the CS-core may still be empty—and, moreover, it is NP-hard to check whether a given outcome is in the CS-core or to generate an outcome in the CS-core [7]. For these reasons, we choose not to take this route.

In terms of future work, we intend to explore the PoD for classes of games other than WVGs. Indeed, the PoD notion is conceptually appealing and provides a useful tool for the analysis of any kind of coalitional games that involve resource allocation, such as, e.g., min-cost spanning tree [3] or network flow games [9]. Also, we intend to study the relationship between the number of rounds and the quality of the outcome in other multi-round bargaining processes. Indeed, our paper illustrates that, even though there is no delay of agreement, having two rounds of bargaining rather than one can improve the outcome considerably, whereas adding more rounds has no effects on the outcome quality. Given this, it would be interesting to see if the same phenomenon occurs in a wider range of settings. In a still broader context, quantifying the relationship between repeated or extensive-form games' outcomes and their respective optimal centralized solutions, is an interesting research topic that has received little attention in the existing literature. We intend to work towards filling this gap.

Acknowledgements This research is supported by the ALADDIN (Autonomous Learning Agents for Decentralised Data and Information Networks) project, which is jointly funded by a BAE Systems and EPSRC strategic partnership (EP/C548051/1), as well as by EPSRC (GR/T10664/01) and ESRC (ES/F035845/1).

9. REFERENCES

[1] E. Anshelevich, A. Dasgupta, J. M. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden. The Price of Stability for Network Design with Fair Cost Allocation. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 295–304, Rome, Italy, 2004.

[2] J. F. Banzhaf. Weighted voting doesn't work: a mathematical analysis. *Rutgers Law Review*, 19:317–343, 1965.

[3] C. Bird. On cost allocation for spanning tree: a game theoretic approach. *Networks*, 6:335–350, 1976.

[4] P. Caillou, S. Aknine, and S. Pinson. Multi-agent models for searching pareto optimal solutions to the problem of forming and dynamic restructuring of coalitions. In *Proceedings of the 15th European Conference on Artificial Intelligence (ECAI'02)*, pages 13 – 17, 2002.

[5] K. Chatterjee, B. Dutta, and K. Sengupta. A Noncooperative Theory of Coalitional Bargaining. *Review of Economic Studies*, 60:463–477, 1993.

[6] G. Christodoulou and E. Koutsoupias. The price of anarchy of finite congestion games. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC-05)*, pages 67–73, Baltimore, MD, USA, 22-24May 2005.

[7] E. Elkind, G. Chalkiadakis, and N. R. Jennings. Coalition structures in weighted voting games. In *Proceedings of the 18th European Conference on Artificial Intelligence (ECAI'08)*, 2008.

[8] E. Elkind, L. Goldberg, P. Goldberg, and M. Wooldridge. Computational complexity of weighted threshold games. In *Proceedings of AAAI'07*, pages 718–723, 2007.

[9] D. Granot and F. Granot. On some network flow games. *Mathematics of Operations Research*, 4:792–841, 1992.

[10] S. A. J.M. Snyder Jr., M.M. Ting. Legislative bargaining under weighted voting. *American Economic Review*, 95(4):981–1004, 2005.

[11] E. Koutsoupias and C. Papadimitriou. Worst-case Equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, pages 404–413, Trier, Germany, 1999.

[12] S. Kraus, O. Shehory, and G. Taase. Coalition Formation with Uncertain Heterogeneous Information. In *Proceedings of the Second International Conference on Autonomous Agents and Multiagent Systems (AAMAS'03)*, 2003.

[13] S. Martello, P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*. John Wiley & Sons, 1990.

[14] A. Mas-Colell, M. Whinston, and J. Green. *Microeconomic Theory*. Oxford University Press, 1995.

[15] T. Matsui and Y. Matsui. A survey of algorithms for calculating power indices of weighted majority games. *J. Oper. Res. Soc. Japan*, 43(1):71–86, 2000.

[16] Y. Matsui and T. Matsui. NP-completeness for calculating power indices of weighted majority games. *Theoretical Comp. Science*, 263(1-2):305–310, 2001.

[17] K. Matsuyama. Perfect Equilibria in a Trade Liberalization Game. *The American Economic Review*, 80(3):480–492, June 1990.

[18] B. Moldovanu and E. Winter. Order Independent Equilibria. *Games and Economic Behavior*, 9:21–34, 1995.

[19] R. Myerson. *Game Theory: Analysis of Conflict*. Harvard U. Press, 1991.

[20] A. Okada. A Noncooperative Coalitional Bargaining Game With Random Proposers. *Games and Economic Behavior*, 16:97–108, 1996.

[21] M. Osborne and A. Rubinstein. *Bargaining and Markets*. Academic Press, San Diego, 1990.

[22] K. Prasad and J. S. Kelly. NP-completeness of some problems concerning voting games. *Internat. J. Game Theory*, 19(1):1–9, 1990.

[23] T. Roughgarden and E. Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.

[24] T. Sandholm and V. Lesser. Coalitions Among Computationally Bounded Agents. *Artificial Intelligence*, 94(1):99 – 137, 1997.

[25] L. S. Shapley and M. Shubik. A Method for Evaluating the Distribution of Power in a Committee System. *American Political Science Review*, 48:787–792, 1954.

[26] A. Taylor and W. Zwicker. *Simple Games: Desirability Relations, Trading, Pseudoweightings*. Princeton University Press, Princeton, 1999.

[27] H. Yan. Noncooperative selection of the core. *International Journal of Game Theory*, 31(4):527–540, Sept. 2003.