

Reducing Optimism Bias in Incomplete Cooperative Games

Filip Úradník
Charles University
Prague, Czechia
uradnik@kam.mff.cuni.cz

Jakub Černý
Columbia University
New York, USA
cerny@disroot.org

David Sychrovský
Charles University
Prague, Czechia
sychrovsky@kam.mff.cuni.cz

Martin Černý
Charles University
Prague, Czechia
cerny@kam.mff.cuni.cz

ABSTRACT

Cooperative game theory has diverse applications in contemporary artificial intelligence, including domains like interpretable machine learning, resource allocation, and collaborative decision-making. However, specifying a cooperative game entails assigning values to exponentially many coalitions, and obtaining even a single value can be resource-intensive in practice. Yet simply leaving certain coalition values undisclosed introduces ambiguity regarding individual contributions to the collective grand coalition. This ambiguity often leads to players holding overly optimistic expectations, stemming from either inherent biases or strategic considerations, frequently resulting in *collective claims exceeding the actual grand coalition value*. In this paper, we present a framework aimed at optimizing the sequence for revealing coalition values, with the overarching goal of efficiently closing the gap between players' expectations and achievable outcomes in cooperative games. Our contributions are threefold: (i) we study the individual players' optimistic completions of games with missing coalition values along with the arising gap, and investigate its analytical characteristics that facilitate more efficient optimization; (ii) we develop methods to minimize this gap over classes of games with a known prior by disclosing values of additional coalitions in both offline and online fashion; and (iii) we empirically demonstrate the algorithms' performance in practical scenarios, together with an investigation into the typical order of revealing coalition values.

KEYWORDS

incomplete cooperative games; superadditive set functions; active learning; value querying; Shapley value

ACM Reference Format:

Filip Úradník, David Sychrovský, Jakub Černý, and Martin Černý. 2024. Reducing Optimism Bias in Incomplete Cooperative Games. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024)*, Auckland, New Zealand, May 6 – 10, 2024, IFAAMAS, 20 pages.



This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

1 INTRODUCTION

Cooperative game theory provides a versatile framework for modeling coalition formation and collective payoff distributions in multiagent interactions. Its far-reaching applications span diverse fields, including supply chain management [16], communication networks [20], logistics and resource allocation [12], or environmental agreements [7]. However, beneath its promising facade lies a fundamental challenge – specifying a cooperative game requires assigning a value to each possible coalition, i.e., a subset of players, which can be a daunting process as the number of coalitions is exponential in the number of interacting agents.

In practice, acquiring even a single value for a single coalition can be a resource-intensive endeavor. Take, for instance, the realm of machine learning, where determining the value of a feature subset that represents a coalition in the celebrated explainable approach SHAP [13] corresponds to retraining an entire model, consuming time and computational resources. What is more, these contributions can have ripple effects on subsequent financial outlays, such as acquiring new training samples.

In the corporate world, estimating an employee's contribution to collective performance may facilitate their fair evaluation [15], but obtaining the value of a coalition may involve the intricate process of rearranging teams of employees, incurring operational costs and potentially causing disruptions.

Yet, simply leaving the values of many coalitions undetermined further compounds the problem by opening the doors to ambiguity. When dealing with incomplete information, humans tend to exhibit a natural bias towards optimism, defined as the disposition to be overly optimistic about the probability of positive future events and to downplay the likelihood of negative future events [11, 26, 27, 29]. For instance, optimism bias can lead individuals to underestimate the effort needed for retirement savings, potentially resulting in insufficient funds [19], or to downplay the likelihood of catastrophic events like natural disasters, leading to inadequate preparedness and potentially life-threatening situations [30]. In the context of cooperative games, where the players lack precise information about coalition values, such optimism can lead to inflated expectations. The individuals believe their contributions are more significant than they objectively are. This optimism translates into unrealistic demands for a larger share of the grand, i.e. all-agents, coalition value, even to the point where the sum of individual claims surpasses the actual value of the grand coalition. For example, the companies might sometimes demand exorbitant prices for their

data, just as employees may occasionally request wages that are unrealistically high. Moreover, even when players do not possess inherent optimism biases, strategically adopting an optimistic facade can prove advantageous, influencing negotiations and outcomes in cooperative games. This discrepancy creates a critical gap, which we refer to as the *cumulative utopian gap*, between what the players expect and what is feasible within the game.

To narrow this gap, we further assume the existence of an external principal. In the examples we mentioned, this role could be reserved for the company manager or the machine learning engineer. This principal possesses the unique ability to determine the sequence in which coalitions are revealed, and they are provided with a limited number of opportunities to exercise this control. In this manner, the principal mitigates the ambiguity within the system, thereby diminishing the utopian gap as a consequence. Importantly, we assume that each revelation step carries a roughly equivalent cost, ensuring that there is no inherent preference in which coalition is unveiled next besides its effect on bridging the gap. The primary objective of the principal hence reduces solely to minimizing this gap under the budget constraint. Although we do not explicitly outline the methodology for resolving situations when a non-zero gap exists, our underlying assumption is that a lower gap is favorable. In practical terms, a diminished gap may have implications for reducing the bargaining power of individual players, or in more extreme scenarios, it may limit the additional financial incentives required by the principal to encourage players to participate in the grand coalition.

1.1 Organization and Contributions

We begin by formally defining the framework of cooperative games, with a particular emphasis on the Shapley value¹ [25] as a mechanism for fairly distributing the grand coalition value among the individual players. We then explain how to extend the framework to encompass incomplete cooperative games, which feature missing coalition values, a prevalent occurrence in real-world scenarios.

Afterward, we delve into our main contributions. Central to our study is the introduction of superadditive utopian games, designed to incorporate players' optimism biases, resulting in the emergence of the cumulative utopian gap. We establish fundamental theoretical properties of the utopian gap, including its monotonicity, additivity, and circumstances under which it becomes zero. Additionally, we offer an alternative geometric interpretation of the utopian gap, viewing it as a quantification of the remaining uncertainty within the space of potential true cooperative games, considering the coalition values that have been determined thus far.

Building upon these foundations, we formulate both offline and online problems for the principal aiming to minimize the utopian gap, offering a suite of heuristic and approximative algorithms for each scenario. Our empirical analysis provides valuable insights into algorithms' performance and the coalitions typically revealed early in the online and offline setups. Importantly, our findings also illustrate the non-linear nature of the utopian gap's decrease

¹While our work draws from the properties of the Shapley value, it is important to note that our findings extend beyond the scope of cooperative game theory alone. For example, they have potential applicability within the theory of capacities [6], where the Shapley value carries alternative interpretations. For a broader generalization of our approach, please refer to the related work section.

with the number of revealed coalition values, offering nuanced perspectives on what a principal can expect at various stages of the revelation process. Our results further indicate that for specific classes of monotone supermodular games, the gap can be nearly completely reduced by revealing just $\mathcal{O}(n)$ coalitions. This stands in contrast to the general requirement of exploring all possible values to minimize it to zero, as demonstrated in Appendix A.

1.2 Related Work

To the best of our knowledge, there is no existing research directly addressing the reduction of ambiguity in the context of cooperative game theory. Nevertheless, the process of querying coalition values can be seen as an online construction of a compact function representation, an approach supported by promising results in the field of cooperative game theory [2–4, 14]. These results indicate the feasibility of such an approach, with some efforts demonstrating a substantial exponential reduction in the number of values required to represent superadditive functions. A comprehensive survey and detailed presentation of many of these findings can be found in Chalkiadakis' book [5]. While tailored representations have achieved significant reductions in specific cases [10], no general approach for constructing such representation when given a subclass of games has been identified.

Our work adopts an approach reminiscent of active learning [23], where an algorithm actively queries an oracle for labels to new data points to construct the most informative dataset, particularly in situations where labeling is resource-intensive. In our context, we seek coalition values that minimize the utopian gap. To approximate the optimal querying strategy based on this concept, we employ reinforcement learning [28].

2 PRELIMINARIES

Here, we present fundamentals of cooperative games (the reader is encouraged to see [18] for a more extensive introduction). First, we define cooperative games, introduce their classes and the Shapley value. Then we talk about the generalized model of incomplete games where only some values of the game are known. We define sets of \mathbb{S}^n -extensions and the *lower/upper games*. These notions delimit possible values of \mathbb{S}^n -extensions of incomplete games.

Definition 2.1. A *cooperative game* is an ordered pair (N, v) where $N = \{1, \dots, n\}$ and $v: 2^N \rightarrow \mathbb{R}$ is the characteristic function of the cooperative game. Further, $v(\emptyset) = 0$.

It is convenient to view the characteristic function v as a point in 2^n -dimensional space. We say a cooperative game (N, v) is

- *additive* if for every $S \subseteq N$,

$$v(S) = \sum_{i \in S} v(\{i\}), \quad (1)$$
- *superadditive* if for every $S, T \subseteq N, S \cap T = \emptyset$,

$$v(S) + v(T) \leq v(T \cup S) \quad (2)$$
- *supermodular* if for every $S \subseteq N \setminus \{i, j\}$,

$$v(S \cup \{j\}) - v(S) \leq v(S \cup \{i, j\}) - v(S \cup \{i\}). \quad (3)$$

By \mathbb{S}^n and \mathbb{C}^n , we denote the sets of superadditive, respectively supermodular games on $|N| = n$ players. Further we refer to $S \subseteq N$ as a *coalition* and N as the *grand coalition*. The following definition

contains one of the most studied solution concepts in cooperative game theory, which models fair distribution of the value of the grand coalition [25].

Definition 2.2. The Shapley value $\phi : \mathbb{R}^{2^n} \rightarrow \mathbb{R}^n$ of a cooperative game (N, v) is defined for $i \in N$ as

$$\phi_i(v) := \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} (v(S \cup \{i\}) - v(S)). \quad (4)$$

The Shapley value is *linear* and *efficient*, which captures the fact that exactly $v(N)$ is distributed among the players. Specifically,

$$\sum_{i \in N} \phi_i(v) = v(N). \quad (5)$$

The characteristic function v is represented by 2^n real values. In many applications, however, obtaining all the values might be too expensive. To deal with this problem, $\mathcal{K} \subseteq 2^N$ is introduced in the definition of an *incomplete game* which represents only coalitions with known value. In our application, we model obtaining new information about the unknown (but well defined) values by extending the set \mathcal{K} . Therefore, we view \mathcal{K} as a ‘masking set’ applied to some complete game and use the characteristic function of that complete game in our definition.² Another distinction from the standard definition is that we require knowledge of at least *minimal information*³

$$\mathcal{K}_0 := \{\emptyset, N\} \cup \{\{i\} \mid i \in N\}. \quad (6)$$

Definition 2.3. An *incomplete cooperative game* is (N, \mathcal{K}, v) where $N = \{1, \dots, n\}$, $\mathcal{K}_0 \subseteq \mathcal{K} \subseteq 2^N$, $v : 2^N \rightarrow \mathbb{R}$ is the characteristic function, and $v(\emptyset) = 0$. We say that (N, v) is the *underlying game* of (N, \mathcal{K}, v) . Further, an incomplete cooperative game (N, \mathcal{K}, v) is *minimal* if $\mathcal{K} = \mathcal{K}_0$.

If further properties of the underlying game are assumed, one may impose restrictions on values of $S \notin \mathcal{K}$ even though they are not known exactly. The property we consider is superadditivity, which is natural in many situations and common throughout the literature on exchange economies, cost-distribution problems, or supermodular optimization [14, 18].⁴ Under this assumption and based on partial knowledge represented by \mathcal{K} , one can describe the set of *candidates* for the underlying game. The set consists of extensions of the partial set function, which satisfy superadditivity.

Definition 2.4. Let (N, \mathcal{K}, v) be an incomplete cooperative game. Then (N, w) is a \mathbb{S}^n -extension of (N, \mathcal{K}, v) if $(N, w) \in \mathbb{S}^n$ and

$$v(S) = w(S), \quad S \in \mathcal{K}. \quad (7)$$

We say (N, \mathcal{K}, v) is \mathbb{S}^n -extendable if it has a \mathbb{S}^n -extension and we denote the set of \mathbb{S}^n -extensions by $\mathbb{S}^n(\mathcal{K}, v)$.

Since the set of \mathbb{S}^n -extensions is given by a system of linear inequalities, it forms a convex polyhedron in \mathbb{R}^{2^n} . Furthermore, assuming non-negativity of the values, set $\mathbb{S}^n(\mathcal{K}, v)$ can be tightly

²This is in contrast with the standard definition where only $v : \mathcal{K} \rightarrow \mathbb{R}$ is used.

³In the standard definition, only $\emptyset \in \mathcal{K}$ is required [2, 14].

⁴The immediacy of this property is not always guaranteed. For instance, SHAP may not consistently yield a superadditive game. However, within this particular context, the property is observed, for example, in uncorrelated models within an ensemble.

enclosed by an hyper-rectangle given by the so called *lower/upper* games [14]. Specifically, the *lower game* (N, \underline{v}) of (N, \mathcal{K}, v) is

$$\underline{v}_{\mathcal{K}}(S) := \max_{\substack{S_1, \dots, S_k \in \mathcal{K} \\ \bigcup_i S_i = S \\ S_i \cap S_j = \emptyset}} \sum_{i=1}^k v(S_i), \quad (8)$$

and the *upper game* (N, \bar{v}) of (N, \mathcal{K}, v) is

$$\bar{v}_{\mathcal{K}}(S) := \min_{T \in \mathcal{K}: S \subseteq T} v(T) - \underline{v}(T \setminus S). \quad (9)$$

Within this specific context, we restate the following established results that will hold importance for our later findings.

THEOREM 2.5. *Let (N, \mathcal{K}, v) be an \mathbb{S}^n -extendable incomplete game with non-negative values. Then for every \mathbb{S}^n -extension (N, w) of (N, \mathcal{K}, v) it holds*

$$\underline{v}_{\mathcal{K}}(S) \leq w(S) \leq \bar{v}_{\mathcal{K}}(S), \quad \forall S \subseteq N. \quad (10)$$

Further, $\forall S \notin \mathcal{K}$, there are \mathbb{S}^n -extensions (N, w_1) , (N, w_2) such that

$$w_1(S) = \underline{v}_{\mathcal{K}}(S) \quad \text{and} \quad w_2(S) = \bar{v}_{\mathcal{K}}(S).$$

PROOF. The first part of the theorem is equivalent to Theorem 1 in [14]. The second part follows from Theorem 3 in [14], which states that any cooperative game with a characteristic function defined for a non-empty coalition S as

$$v^S(T) := \begin{cases} \bar{v}_{\mathcal{K}}(T) & S \subseteq T, \\ \underline{v}_{\mathcal{K}}(T) & S \not\subseteq T \end{cases} \quad (11)$$

is an \mathbb{S}^n -extension. For $S \notin \mathcal{K}$, choose $w_1 = v^N$ and $w_2 = v^S$. \square

Note that the hyper-rectangle given by lower/upper games might contain non-superadditive games.

3 OPTIMISM BIAS AND ITS MINIMIZATION

The uncertainty arising from \mathcal{K} in an incomplete cooperative game allows each player to speculate about their payment. As players experience the bias towards optimism, they can consider every \mathbb{S}^n -extension of the incomplete game and choose the one which benefits them the most. Such extension can be easily derived based on the properties of the Shapley value.

Definition 3.1. Let (N, \mathcal{K}, v) be a \mathbb{S}^n -extendable incomplete cooperative game and $i \in N$. Then player i 's *utopian game* is (N, v_i) , where

$$v_i(S) := \begin{cases} \bar{v}_{\mathcal{K}}(S) & i \in S, \\ \underline{v}_{\mathcal{K}}(S) & i \notin S. \end{cases} \quad (12)$$

PROPOSITION 3.2. *Let (N, \mathcal{K}, v) be a \mathbb{S}^n -extendable incomplete game. Then the utopian game $(N, v_i) \in \mathbb{S}^n(\mathcal{K}, v)$ for every $i \in N$.*

PROOF. From (8), (9), it is immediate (N, v_i) extends (N, \mathcal{K}, v) . Further, as mentioned earlier, Theorem 3 in [14] states that any cooperative game with a characteristic function (11) is superadditive. We obtain the result by setting $S = \{i\}$. \square

When each player demands the payoff given to them in their utopian game, it leads to an efficiency violation due to the players' cumulative demands surpassing the value of $v(N)$. We refer to the disparity between the players' expectations and the attainable outcome as the *cumulative utopian gap*.

Definition 3.3. Let $(N, \hat{\mathcal{K}}, v)$ be a \mathbb{S}^n -extendable incomplete cooperative game and $\mathcal{K} = \hat{\mathcal{K}} \setminus \mathcal{K}_0$. Then the *cumulative utopian gap* of $(N, \hat{\mathcal{K}}, v)$ is

$$\mathcal{G}_{(N,v)}(\mathcal{K}) := \sum_{i \in N} \left(\max_{w \in \mathbb{S}^n(\hat{\mathcal{K}}, v)} \phi_i(w) \right) - v(N). \quad (13)$$

Due to the properties of utopian games, the cumulative utopian gap can be expressed in terms of individual v_i 's. As a consequence, one can express the gap as an affine combination of difference between the values of the upper and the lower functions.

PROPOSITION 3.4. *Let $(N, \hat{\mathcal{K}}, v)$ be an \mathbb{S}^n -extendable incomplete game. Then the cumulative utopian gap is*

$$\mathcal{G}_{(N,v)}(\mathcal{K}) = \sum_{i \in N} \phi_i(v_i) - v(N) = \sum_{S \subseteq N} \alpha_S \Delta_{\hat{\mathcal{K}}}(S) - v(N), \quad (14)$$

where $\alpha_S = \frac{|S|!(|N|-|S|)!}{|N|!}$ and $\Delta_{\hat{\mathcal{K}}}(S) = \bar{v}_{\hat{\mathcal{K}}}(S) - \underline{v}_{\hat{\mathcal{K}}}(S)$.

PROOF SKETCH. Notice from (4) and the definition of $\mathbb{S}^n(\hat{\mathcal{K}}, v)$ that $\phi_i(w)$ is bounded from above by

$$\sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} (\bar{v}_{\hat{\mathcal{K}}}(S \cup \{i\}) - \underline{v}_{\hat{\mathcal{K}}}(S)),$$

which is exactly $\phi_i(v_i)$. The proof of the second equality is merely a technicality left for the Appendix G. \square

Now, we shall demonstrate several key properties of the utopian gap, starting with its non-negativity.

PROPOSITION 3.5. *Let $(N, \hat{\mathcal{K}}, v)$ be an \mathbb{S}^n -extendable incomplete game. Then $\mathcal{G}_{(N,v)}(\mathcal{K}) \geq 0$. Moreover, $\mathcal{G}_{(N,v)}(\mathcal{K}) = 0$ if and only if $\mathbb{S}^n(\hat{\mathcal{K}}, v)$ is a singleton.*

PROOF. Non-negativity follows from the fact that for $(N, v) \in \mathbb{S}^n(\hat{\mathcal{K}}, v)$ it holds that

$$\sum_{i \in N} \phi_i(v_i) \geq \sum_{i \in N} \phi_i(\underline{v}_{\hat{\mathcal{K}}}) = v(N). \quad (15)$$

The inequality in (15) is strict if and only if there is more than one \mathbb{S}^n -extension as $\phi_i(v_i) > \phi_i(\underline{v}_{\hat{\mathcal{K}}})$ for at least one $i \in N$. \square

Note that the set of extensions can be non-trivial even if we are missing any single value of the underlying game, see Appendix A.

PROPOSITION 3.6. *The cumulative utopian gap \mathcal{G} is in the set \mathcal{K}*

(1) *monotonically non-increasing, i.e., for any $S \in 2^N$*

$$\mathcal{G}_{(N,v)}(\mathcal{K}) \geq \mathcal{G}_{(N,v)}(\mathcal{K} \cup \{S\}), \quad (16)$$

(2) *subadditive, i.e., for $\mathcal{K}, \mathcal{L} \subseteq 2^N$, $\mathcal{K} \cap \mathcal{L} = \emptyset$,*

$$\mathcal{G}_{(N,v)}(\mathcal{K}) + \mathcal{G}_{(N,v)}(\mathcal{L}) \geq \mathcal{G}_{(N,v)}(\mathcal{K} \cup \mathcal{L}). \quad (17)$$

PROOF. To show (16), for an incomplete game $(N, \hat{\mathcal{K}}, v)$ and a coalition $S \subseteq N$, consider the projection of $\mathbb{S}^n(\hat{\mathcal{K}}, v) \subseteq \mathbb{R}^{2^N}$ onto the axis, which is given by S . By Theorem 2.5 and supermodularity of the set of \mathbb{S}^n -extensions, such projection is $[\underline{v}_{\hat{\mathcal{K}}}(S), \bar{v}_{\hat{\mathcal{K}}}(S)]$. By Proposition 3.4 and the definition of $\phi_i(v_i)$, an ε change in $\bar{v}_{\hat{\mathcal{K}}}(S)$ leads to

$$\varepsilon \cdot |S| \cdot \frac{(|S|-1)! (|N|-|S|)!}{|N|!}$$

change in \mathcal{G} , and an ε change in $\underline{v}(S)$ leads to

$$-\varepsilon \cdot |S| \cdot \frac{|S|!(|N|-|S|-1)!}{|N|!}$$

change in \mathcal{G} . This means that \mathcal{G} is increasing in $\bar{v}_{\hat{\mathcal{K}}}(S)$ and decreasing in $\underline{v}_{\hat{\mathcal{K}}}(S)$. As $\mathbb{S}^n(\hat{\mathcal{K}} \cup \{S\}, v) \subseteq \mathbb{S}^n(\hat{\mathcal{K}}, v)$, the projection of $\mathbb{S}^n(\hat{\mathcal{K}} \cup \{S\}, v)$ on axis S is contained in the projection $\mathbb{S}^n(\hat{\mathcal{K}} \cup \{S\}, v)$ on axis S , which implies (16). Finally, from non-negativity and monotonicity follows (17). \square

Finally, let us touch on the geometric interpretation of the utopian gap. As stated in Theorem 2.5, the upper and lower games define a hyper-rectangle which tightly bounds the set of extensions $\mathbb{S}^n(\hat{\mathcal{K}}, v)$. Since the utopian gap is a linear function of the values of the lower/upper games, it can be seen as a ‘‘measure’’ of the size of $\mathbb{S}^n(\hat{\mathcal{K}}, v)$. Uncertainty in each coalition is additionally weighted according to its contribution to the Shapley value.

3.1 Principal's Optimization Problems

A large utopian gap makes it difficult to distribute payoff, as each player tends to their utopian game. The more we know about the underlying game, the smaller the cumulative utopian gap gets, being zero for a game with a single \mathbb{S}^n -extension. However, to obtain all the necessary unknown values might be too expensive, since there are exponentially many of them.

To minimize the utopian gap, we hence assume the existence of a non-affiliated party who we refer to as the *principal*. Her task is to choose which coalitions should be investigated to reduce the utopian gap the most. We further assume the principal holds a level of expertise that guides the selection process. This expertise is formalized as a *prior distribution* over a set of potential characteristic functions. For example, in a medical context, a doctor acting as the principal might seek to assess a patient's response to a combination of drugs (represented as a coalition) and base their expectations on past clinical experience. Similarly, a machine learning engineer could rely on their prior knowledge of feature importance gained from previous problem-solving experiences. Consequently, we assume that each problem instance can be viewed as a sample drawn from this known prior distribution, denoted as \mathcal{F} . To put bluntly, the principal is aware of the prior distribution, but not of the specific instance of the underlying game.

There are two basic approaches to choosing which coalitions to investigate, *online* and *offline*. In the online approach, the principal operates sequentially, utilizing information from previously revealed coalition values.

Definition 3.7(Online Principal's Problem). Let $t \in \mathbb{N}$, \mathcal{F} , $\text{supp } \mathcal{F} \subseteq \mathbb{S}^n$ be a distribution of superadditive games. Then $\mathcal{K}_t^* \subseteq 2^N \setminus \mathcal{K}_0$ is a solution of the *online principal's problem* of size t if

$$\mathcal{K}_t^* \in \underset{\mathcal{K}_t \subseteq 2^N \setminus \mathcal{K}_0, |\mathcal{K}_t|=t}{\text{argmin}} \left\{ \mathbb{E}_{v \sim \mathcal{F}} [\mathcal{G}_{(N,v)}(\mathcal{K}_t)] \right\} \quad (18)$$

where \mathcal{K}_τ , $\tau \leq t$ is such that $\mathcal{K}_\tau = \mathcal{K}_{\tau-1} \cup \{S_\tau\}$, $S_\tau = \pi(N, \mathcal{K}_{\tau-1}, v)$, and π is a policy function that chooses S_τ based on the known values of v , i.e. values of coalitions in $\mathcal{K}_{\tau-1}$.

In contrast, the offline approach entails a lack of such information.

Definition 3.8 (Offline Principal's Problem). Let $t \in \mathbb{N}$, \mathcal{F} , $\text{supp } \mathcal{F} \subseteq \mathbb{S}^n$ be a distribution of superadditive games. Then $\mathcal{K}_t^* \subseteq 2^N \setminus \mathcal{K}_0$ is a solution of the *offline principal's problem* of size t if

$$\mathcal{K}_t^* \in \underset{\mathcal{K} \subseteq 2^N \setminus \mathcal{K}_0, |\mathcal{K}|=t}{\text{argmin}} \left\{ \mathbb{E}_{v \sim \mathcal{F}} [\mathcal{G}_{(N,v)}(\mathcal{K})] \right\}. \quad (19)$$

3.2 Algorithms Solving the Principal's Problems

In this section, we discuss various methods for finding (approximate) solutions to the principal's problems defined above. We defer further details about all algorithms to Appendix B.

3.2.1 Offline Algorithms. At each step t , the OFFLINE OPTIMAL algorithm chooses coalitions $\{S_i\}_{i=1}^t$ which minimize the expected utopian gap under \mathcal{F} . We estimate the expectation w.r.t. \mathcal{F} in Eq. (19) by κ samples, see Algorithm 1.

Algorithm 1: OFFLINE OPTIMAL

Input: distribution of superadditive functions \mathcal{F} , number of steps t , number of samples κ

- 1 $\overline{\mathcal{K}} \leftarrow 2^N \setminus \mathcal{K}_0$
- 2 $G \leftarrow \{\}$ // trajectories and their average gap
- 3 **for** $S \subseteq \overline{\mathcal{K}} : |S| = t$ **do** // for each trajectory
- 4 $\mu \leftarrow 0$
- 5 **for** $j \in \{1, \dots, \kappa\}$ **do** // approximate expectation
- 6 $v \sim \mathcal{F}$
- 7 $\mu \leftarrow \mu + \mathcal{G}_{(N,v)}(S)$
- 8 $\mu \leftarrow \mu / \kappa$
- 9 $G[S] \leftarrow \mu$
- 10 $\{S_i\}_{i=1}^t \leftarrow \underset{S \subseteq \overline{\mathcal{K}} : |S|=t}{\text{argmin}} G[S]$
- 11 **return** $\{S_i\}_{i=1}^t$

A computationally less demanding variant of the OFFLINE OPTIMAL is the OFFLINE GREEDY algorithm. It chooses the next coalition S_t such that, given the previous trajectory $\{S_i\}_{i=1}^{t-1}$, it minimizes the expected utopian gap. Consequently, it can perform no better than the OFFLINE OPTIMAL. We again estimate the expectation in Eq. (19) by κ samples, see Algorithm 2.

3.2.2 Online Algorithm. In comparison to the offline problem, solving the online problem poses a significantly greater challenge. Intuitively, one key reason is that an algorithm for the online problem must compute (or approximate) a restriction of \mathcal{F} that remains consistent with the values it has uncovered in prior steps. However, this can be particularly challenging, especially in case the only access to \mathcal{F} is through sampling. To tackle the online problem and derive an approximate solution, we employ reinforcement learning [28], specifically, the proximal policy optimization (PPO) [21].

At each step τ , PPO receives values of coalitions $\mathcal{K}_{\tau-1} = \{S_i\}_{i=1}^{\tau-1}$ it uncovered in the past and chooses the next coalition S_τ . To get a strategy which efficiently minimizes the utopian gap, we define the reward (which is maximized by the PPO algorithm) as the negative expected utopian gap averaged over $\tau \leq t$.

As previously mentioned, the greedy algorithm is significantly more computationally efficient, with a linear complexity in the

Algorithm 2: OFFLINE GREEDY

Input: distribution of superadditive functions \mathcal{F} , number of steps t , number of samples κ

- 1 $\{S_i\}_{i=1}^{t-1} \leftarrow \text{Offline Greedy}(\mathcal{F}, t-1)$
- 2 $\overline{\mathcal{K}} \leftarrow 2^N \setminus (\mathcal{K}_0 \cup \{S_i\}_{i=1}^{t-1})$
- 3 $G \leftarrow \{\}$ // trajectories & their average gap
- 4 **for** $S \in \overline{\mathcal{K}}$ **do** // for each trajectory
- 5 $\mu \leftarrow 0$
- 6 **for** $j \in \{1, \dots, \kappa\}$ **do** // approximate expectation
- 7 $v \sim \mathcal{F}$
- 8 $\mu \leftarrow \mu + \mathcal{G}_{(N,v)}(\{S_i\}_{i=1}^{t-1} \cup \{S\})$
- 9 $\mu \leftarrow \mu / \kappa$
- 10 $G[S] \leftarrow \mu$
- 11 $S_t \leftarrow \underset{S \in \overline{\mathcal{K}}}{\text{argmin}} G[S]$
- 12 **return** $\{S_i\}_{i=1}^t$

number of coalitions, while the optimal variant exhibits an exponential time complexity. A question which naturally arises is under which conditions the local greedy search can reliably yield a globally optimal solution. While our empirical observations indicate similar performance between both approaches, the greedy method is not optimal in general. To illustrate, we provide a simple example in Appendix C, where the local search fails to identify the optimum. However, it is known that, if the optimized function is supermodular, the locally optimal steps are guaranteed to yield a $(1 - 1/e)$ -approximation of the global optimum [8, Proposition 3.4]. In our case, the utopian gap is supermodular for all small games.

PROPOSITION 3.9. *For $|N| \leq 4$, the gap of an incomplete \mathbb{S}^n -extendable game is supermodular.*

PROOF SKETCH. For $|N| = 3$, the gap is modular due to invariant bounds post-revelation. For $|N| = 4$, the supermodularity is demonstrated through a more technical case analysis. The full proof can be found in Appendix H.1. \square

However, already for $|N| = 5$, it can be shown the gap is not supermodular for several significant subsets of superadditive games, such as totally monotonic games (a subset of convex games, which under the name *belief measures* forms the foundation on the *evidence theory* [9, 24]), symmetric games (often studied for their robust properties with respect to solution concepts [18]), or graph games (games with a compact structure represented by $O(n^2)$ values with application in scenarios where only bilateral collaboration between players occurs [5]).

OBSERVATION. *For $|N| = 5$, there exist a totally monotonic game, convex game, symmetric superadditive game, and graph game for which the utopian gap is not supermodular.*

For even larger games, we can formulate a precise criterion to ascertain the non-supermodularity of the utopian gap.

PROPOSITION 3.10. *For $|N| \geq 6$, the gap of an incomplete \mathbb{S}^n -extendable game is not supermodular if there exist $i, j, k, l \in N$ such*

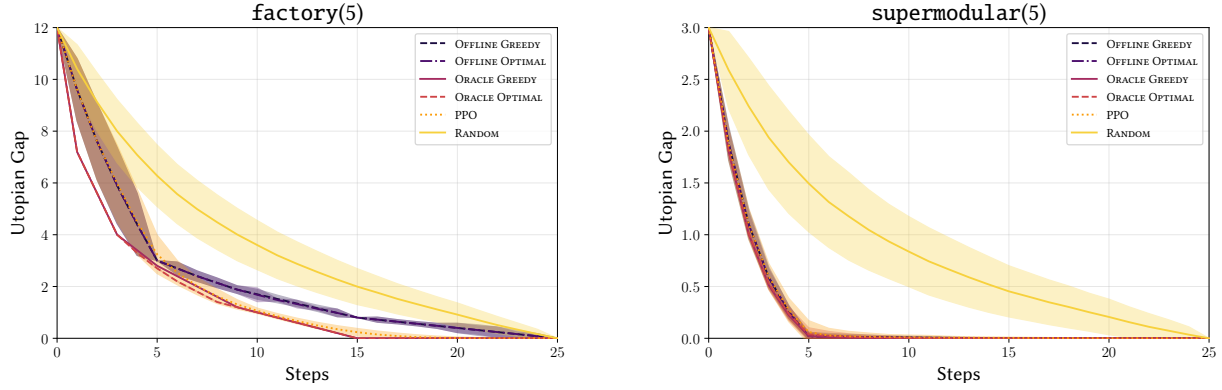


Figure 1: The utopian gap as a function of number of revealed coalitions (i.e. steps of the Principals problem) for different algorithms. We show `factory(5)` (left), and `supermodular(5)` (right) games. All algorithms outperform the `RANDOM` benchmark considerably. The greedy versions of each algorithm exhibit similar performance to the optimal variants. The PPO algorithm is initially close to the offline algorithms, and uses the online information to approach the oracle algorithms.

that $\tilde{v}(ij) \leq \tilde{v}(jk) \leq \tilde{v}(kl)$, where $\tilde{v}(S) = v(S) - \sum_{i \in S} v(\{i\})$, and

$$\tilde{v}(kl) < \left[\binom{n}{2} \binom{n}{\lfloor n/2 \rfloor}^{-1} (2^{n-3} - n + 2) - 1 \right] \tilde{v}(ij). \quad (20)$$

PROOF SKETCH. We rephrase the supermodularity condition (3) in relation to the utopian gap, and demonstrate that this condition is never met when Eq. (20) holds for the pairs in some $\{i, j, k, l\}$. The full proof can be found in Appendix H.3. \square

This criterion allows to construct examples also for all earlier mentioned classes of games. The proofs and examples are available in Appendix H. Despite these findings, empirically, the greedy algorithm demonstrates similar performance to its optimal counterpart, as illustrated in the following section.

4 EMPIRICAL EVALUATION

Finally, we demonstrate the performance of our algorithms on practical examples. We outline the domains used for the evaluation, introduce baseline methods for comparison, detail the algorithmic setups, and conduct a comprehensive analysis of the gathered results. All the details that could not be accommodated within the main text have been addressed in the relevant appendices.

4.1 Experimental Domains

We conduct our experiments on two representative families of cooperative games. Their main difference is the degree to which the values of different coalitions are correlated.

For the tightly correlated scenario, we use a simple model of a factory. The set of players N , $|N| \geq 2$ consists of a single factory owner o and $n - 1$ identical workers. The value of a coalition S is

$$v(S) = \begin{cases} |S| - 1 & \text{if } o \in S, \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

In words, each worker contributes equally to the value of a coalition as long as he has somewhere to work. It is easy to verify that the resulting game is superadditive. We denote the distribution of

factory games of n players where the owner is selected uniformly at random as `factory(n)`. We show that the gap is not supermodular already for `factory(5)` in Appendix C.

The second family is the broad class of supermodular games. These games are among the most studied in the cooperative game theory [9] – some authors consider this property so important they even impose it in the definition of a cooperative game [17]. In our context, the problem of identifying a subset of coalitions which represents a supermodular function well has been studied in a broader context of set functions [22]. In [1], the authors derived an efficient algorithm for uniform sampling of monotone supermodular functions. We refer to the distribution of corresponding games of n players as `supermodular(n)`.

We conduct further experiments on several other classes of cooperative games as well, and present these results in Appendix D. We observe similar trends as those presented here in the main text.

4.2 Benchmarks

We compare our algorithms introduced in the previous section to three baselines: a random algorithm and two oracle algorithms. The random algorithm selects the next coalition uniformly at random. We refer to it in the results as `RANDOM`. Next, we introduce two oracle methods that are not deployable in practice because they assume the knowledge of the underlying true characteristic function. However, they provide an upper bound on what an optimal online algorithm could possibly achieve.

The `ORACLE OPTIMAL` algorithm operates similarly to `OFFLINE OPTIMAL`, with the key difference being that it leverages an oracle to acquire values of the underlying game (N, v) before making the next coalition selection. Consequently, `ORACLE OPTIMAL` can harness complete knowledge of the underlying game to minimize the gap effectively. Its pseudocode as Algorithm 3 is in Appendix E.

Similar to the `OFFLINE GREEDY`, the `ORACLE GREEDY` algorithm selects next coalition S_t such that, in combination with the previous trajectory $\{S_i\}_{i=1}^{t-1}$, it minimizes the utopian gap. It also uses the

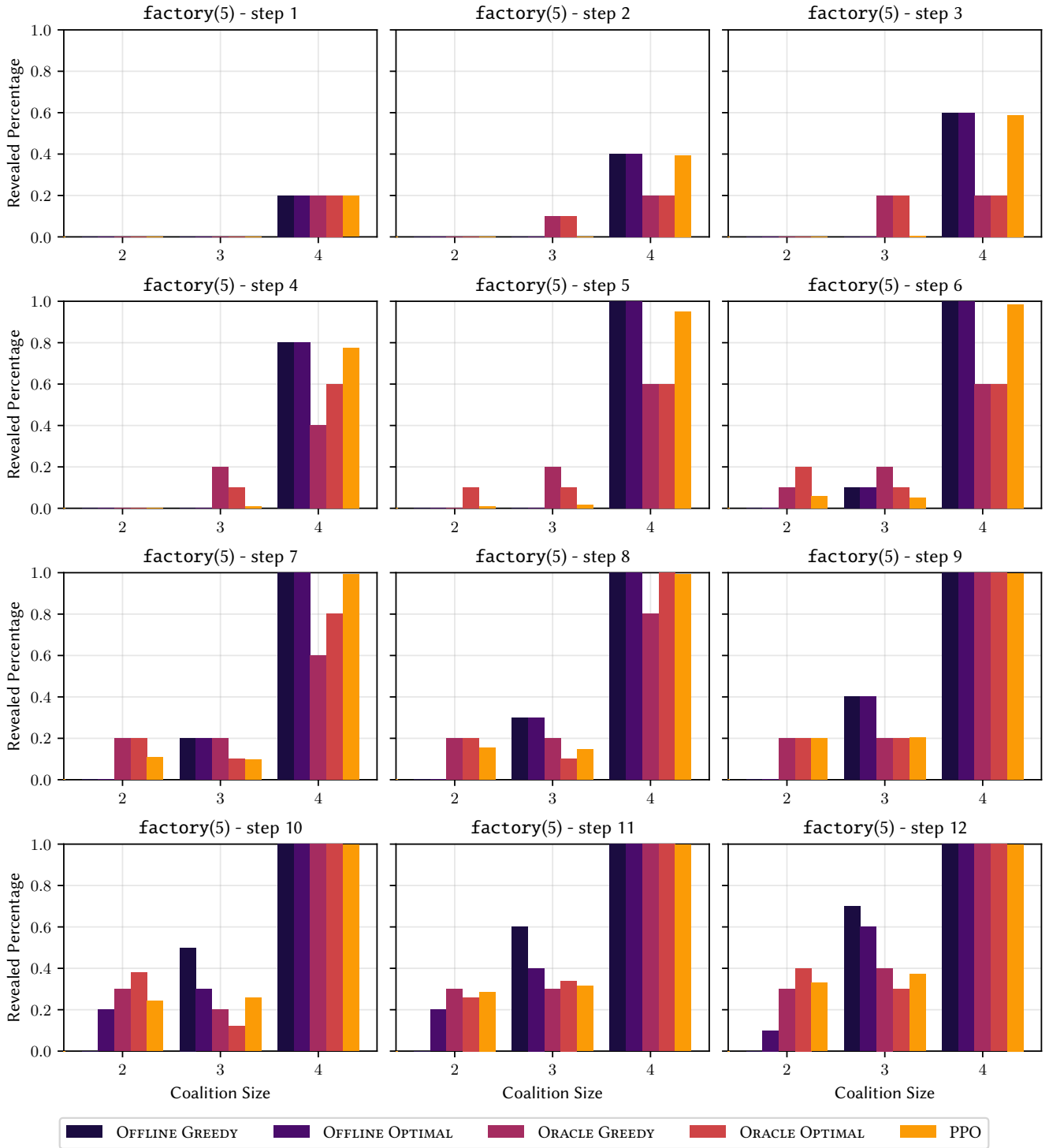


Figure 2: Percentage of coalitions of the same size selected up to step twelve for factory(5) and each algorithm. Results show clear preference for larger coalitions, i.e. they contribute more information about the cooperative game on average. The oracle algorithms favor smaller coalitions earlier, suggesting the representation of a specific game can efficiently use even smaller coalitions. PPO initially behaves similarly to the offline algorithms. At later steps, it uses the previously obtained values and its selections resemble the oracle methods. See Figure 7 in Appendix D for a plot showing individual coalitions.

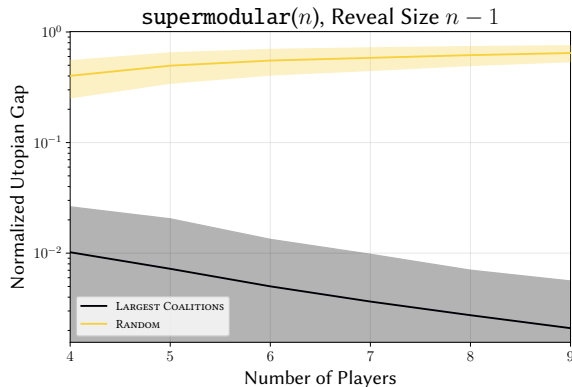


Figure 3: The utopian gap as a function of the number of players. The figure compares expected gap of $\text{supermodular}(n)$ when choosing coalitions randomly, and when all coalitions of size $n - 1$ are selected. Not only is the utopian gap in the latter case small, it decreases with the number of players.

oracle to gather all information about the underlying game (N, v) . Its pseudocode is depicted as Algorithm 4 in Appendix E.

4.3 Experimental Setup

During the training phase, we execute the PPO algorithm for a total of 2,000,000 time-steps, which, in the case of 12-step instances, translates to 40 iterations. In each iteration, we repeatedly collect 4,096 trajectories by sampling batches from the distribution \mathcal{F} . Following each iteration, we optimize the PPO surrogate objective for 10 epochs. To ensure uniformity, we normalize the input values $v(S)$ to a unit interval, as elaborated in Appendix F. For the two optimal algorithms, we employ $\kappa = 10$ samples to estimate the mean value in Eq. (19). See Appendix B for more details.

Among the algorithms we have proposed, both OFFLINE OPTIMAL and ORACLE OPTIMAL come with significantly higher computational costs (their complexity scales as $O(2^{2^n})$) when compared to their more computationally efficient greedy counterparts (which scale $O(2^n)$). To provide some perspective, our evaluation for $n = 5$ with twelve steps and ten samples necessitated approximately 60 hours to finish. Extrapolating, we estimate that completing the evaluations for the remaining steps in this setup, across all ten samples, would require roughly 200 hours. Furthermore, extending these experiments to $n = 6$ would demand over 100 years.

Hardware and Software. All experiments are conducted on a computational cluster with AMD EPYC 7532 CPUs running at 2.4 GHz. When running algorithms on five players games, we utilize 15 cores and 12 GB of RAM. The code was implemented in Python 3.10 using pytorch 2.0, stable_baselines3 2.0, and gymnasium 0.28. Additional details are presented in Appendix B, and is available at GitHub [31].

4.4 Results

First, we study the $\text{factory}(5)$ distribution. Figure 1 shows the dependence of the utopian gap on the number of revealed coalitions. As expected, the oracle algorithms outperform their offline variants,

especially at the beginning. PPO is initially similar to the offline algorithms, and later uses the online information to reach similar values of the utopian gap as the oracle methods. The greedy algorithms exhibit similar performance to their optimal counterparts.

To further illustrate which coalitions are important, Figure 2 shows the percentage of coalitions of a given size selected by each algorithm. The offline algorithms favor the largest coalitions, suggesting they carry the most information overall. The oracle algorithms can gain small improvements by selecting smaller coalitions. This signals that, for a specific game, a tailored representation including also smaller coalitions is better.

Next, we focus on the $\text{supermodular}(5)$ distribution. We show the evolution of the utopian gap as a function of the number of revealed coalitions in Figure 1. Surprisingly, the utopian gap can be minimized very fast, by about 99% by step five. We observed similar trends for $\text{supermodular}(4)$ presented in Appendix D.1. In each case, the algorithms preferred coalitions of size $n - 1$, suggesting they are the most important among the unknown coalitions. We further evaluate the impact of the largest coalitions, by computing the utopian gap as a function of n when all coalitions of size $n - 1$ are revealed. We compare this simple heuristic with RANDOM in Figure 3. The utopian gap is consistently very low and, surprisingly, decreases with the number of players. This suggests that most of the information about a supermodular game can be captured by values of $O(n)$ coalitions.

5 CONCLUSION

In this paper, we study strategies for efficiently mitigating uncertainty within incomplete cooperative games. We introduce the concept of the “utopian gap”, which quantifies the disparity between players’ expectations and a realistic game outcome. We show fundamental properties of the utopian gap that enables us to compute it more efficiently. Geometrically, the utopian gap reflects the number of potential extensions of an incomplete game. We focus on reducing the utopian gap through well-informed queries about the unknown values within the incomplete game, effectively constructing a representation tailored to capture a maximum information about the game in both online and offline fashion. Our findings indicate that our approach significantly outperforms random queries and approaches optimality. Particularly noteworthy is our heuristic for supermodular games, which reduces the utopian gap by orders of magnitude while requiring only $O(n)$ queries – an amount logarithmic in the number of coalitions.

Future Work. We posit that our approach is not confined solely to the boundaries of cooperative game theory and can be extended to more general set functions. This extension assumes the existence of criteria involving functions (in this work, it is the Shapley value) that map the powerset of a set N to subspaces of dimensions linear in the size of N . Still in the context of games, we see applications of our approach within the SHAP [13] framework, as the original formulation can only handle comparatively small models. Finally, we would also like to strengthen our result by providing guarantees by working within the regret minimization framework.

ACKNOWLEDGMENTS

The authors would like to thank Martin Loebl and Milan Hladík for their insightful comments. David Sychrovský and Martin Černý received support from the Charles University Grant Agency (GAUK 206523). This work was supported by the Czech Science Foundation grant no. GA22-26655S, the CoSP Project grant no. 823748, and the Charles University project UNCE 24/SCI/008. Computational resources were supplied by the project e-Infrastruktura CZ (e-INFRA LM2018140) provided within the program Projects of Large Research, Development and Innovations Infrastructures.

REFERENCES

- [1] Gleb Beliakov. 2022. On Random Generation of Supermodular Capacities. *IEEE TRANSACTIONS ON FUZZY SYSTEMS* 30, 1 (2022), 293–295.
- [2] Jan Bok and Martin Černý. 2023. 1-convex extensions of incomplete cooperative games and the average value. *Theory and Decisions* (2023).
- [3] Martin Černý. 2023. Bounds on solution concepts of incomplete cooperative games. arXiv:arXiv:2212.04748 [cs.GT]
- [4] Martin Černý and Michel Grabisch. 2024. Incomplete cooperative games with player-centered information. *Discrete Applied Mathematics* 346 (2024), 62–79.
- [5] Georgios Chalkiadakis, Edith Elkind, and Michael Wooldridge. 2012. *Computational Aspects of Cooperative Game Theory* (1 ed.). Springer Cham. XVI, 150 pages.
- [6] Gustave Choquet. 1954. Theory of capacities. In *Annales de l'institut Fourier*, Vol. 5. 131–295.
- [7] Michael Finus. 2008. Game theoretic research on the design of international environmental agreements: insights, critical remarks, and future challenges. *International Review of environmental and resource economics* 2, 1 (2008), 29–67.
- [8] Nemhauser George, Wolsey Laurence, and Fisher Marshall. 1978. An analysis of approximations for maximizing submodular set functions I. *Mathematical Programming* 14 (1978), 265–294.
- [9] Michel Grabisch. 2016. *Set Functions, Games and Capacities in Decision Making*. Number 978-3-319-30690-2 in Theory and Decision Library C. Springer.
- [10] Samuel Ieong and Yoav Shoham. 2005. Marginal Contribution Nets: A Compact Representation Scheme for Coalitional Games.
- [11] Francis W. Irwin. 1953. Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality* (1953).
- [12] Sebastián Lozano, Plácido Moreno, Belarmino Adenso-Díaz, and Encarnación Algaba. 2013. Cooperative game theory approach to allocating benefits of horizontal cooperation. *European Journal of Operational Research* 229, 2 (2013), 444–452.
- [13] Scott M. Lundberg and Su-In Lee. 2017. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.). Curran Associates, Inc., 4765–4774.
- [14] Satoshi Masuya and Masahiro Inuiguchi. 2016. A fundamental study for partially defined cooperative games. *Fuzzy Optimization Decision Making* 15, 1 (2016), 281–306.
- [15] Kevin R. Murphy and Jeanette N. Cleveland. 1995. *Understanding performance appraisal: Social, organizational, and goal-based perspectives*. Sage.
- [16] Mahesh Nagarajan and Greys Sošić. 2008. Game-theoretic analysis of cooperation among supply chain agents: Review and extensions. *European journal of operational research* 187, 3 (2008), 719–745.
- [17] Guillermo Owen. 2013. *Game Theory* (4th ed.). Emerald Group Publishing, Bingley, U.K.
- [18] Bezalel Peleg and Peter Sudhölter. 2007. *Introduction to the Theory of Cooperative Games*. Springer.
- [19] Manju Puri and David T Robinson. 2007. Optimism and economic choice. *Journal of financial economics* 86, 1 (2007), 71–99.
- [20] Walid Saad, Zhu Han, Mèrouane Debbah, Are Hjorungnes, and Tamer Basar. 2009. Coalitional game theory for communication networks. *Ieee signal processing magazine* 26, 5 (2009), 77–97.
- [21] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal Policy Optimization Algorithms. *CoRR* abs/1707.06347 (2017).
- [22] Comandur Seshadhri and Jan Vondrák. 2014. Is Submodularity Testable? *Algorithmica* 69 (2014), 1–25.
- [23] Burr Settles. 2009. Active Learning Literature Survey. *University of Wisconsin-Madison Department of Computer Sciences* (2009).
- [24] Glenn Shafer. 1976. *A Mathematical Theory of Evidence*. Princeton University Press.
- [25] Lloyd Shapley. 1953. A value for n-person game. *Annals of Mathematical Studies* 28 (1953), 307–317.
- [26] Paul Slovic. 1987. Perception of Risk. *Science* 236, 4799 (1987), 280–285.
- [27] Paul Slovic, Baruch Fischhoff, and Sarah Lichtenstein. 2016. Facts and fears: Understanding perceived risk. In *The perception of risk*. Routledge, 137–153.
- [28] Richard S. Sutton and Andrew G. Barto. 2014. *Reinforcement Learning: An Introduction*.
- [29] Neil D. Weinstein. 1980. Unrealistic optimism about future life events. *Journal of personality and social psychology* 39, 5 (1980), 806.
- [30] Neil D. Weinstein. 1989. Optimistic biases about personal risks. *Science* 246, 4935 (1989), 1232–1233.
- [31] Filip Úradník. 2024. Incomplete Cooperative. <https://github.com/furadnik/IncompleteCooperative/releases/tag/AAMAS24>. Accessed: February 10, 2024.

A ALL VALUES ARE NEEDED FOR ZERO UTOPIAN GAP IN GENERAL

We remark that throughout the paper, we do not aim to achieve a zero gap, as this is impossible for some games until all of the coalition values are known. To see this, consider strictly superadditive incomplete game (N, \mathcal{K}, v) , for which a single coalition value is unknown. Formally, let (N, \mathcal{K}, v) with $\mathcal{K} = 2^N \setminus \{S^*\}$ where $S^* \in 2^N \setminus \mathcal{K}_0$, be such that

$$v(S) + v(T) < v(S \cup T)$$

for all $S, T \subseteq N, S \cap T = \emptyset$. Then as $v(S^*) < v(S^* \cup T) - v(T)$ for every $T \subseteq N \setminus S^*$ and $v(X) + v(S^* \setminus X) < v(S^*)$ for every $X \subseteq S^*$, it follows from the definition of the lower/upper games that

$$\underline{v}_{\mathcal{K}}(S^*) < v(S^*) < \bar{v}_{\mathcal{K}}(S^*).$$

This means there is more than one \mathbb{S}^n -extension, which by Proposition 3.5 yields $\mathcal{G}_{(N,v)}(\mathcal{K}) > 0$.

B ALGORITHM SPECIFICATIONS

Oracle and Offline Algorithms. Greedy strategies are computationally straightforward compared to their optimal counterparts. A single step using a greedy approach demands $\mathcal{O}(g \cdot 2^n)$ time for a single sample. Here, g represents the time required to compute the utopian gap.

On the other hand, optimal algorithms prove to be computationally intensive due to the necessity of examining every sequence of actions, that is, every subset of 2^N . For each sample, the time complexity amounts to $\mathcal{O}(g \cdot 2^{2^n})$. Despite our efforts to parallelize the computation where possible, given our available resources, we were only able to compute the complete optimal strategies for up to 5 players. The computation for 5 players took roughly 60 hours. We estimate, that to compute the optimal strategies for 6 players, for all steps, would takeover 100 years.

The online variants are easier to parallelize, since each sample can be computed and evaluated independently, in contrast to the offline variants, where all the samples need to be computed, put together, averaged, and then finally evaluated.

When estimating the expectation with respect to \mathcal{F} , we used 3,000 samples for 4 players. For 5 players, we ended up using only 10 samples.

Reinforcement Learning. We apply reinforcement learning [28] to approximate the optimal strategy of the online principal’s problem. Namely, we use the Proximal policy optimization (PPO) [21]. We want to find a strategy of the principal which efficiently minimizes the average cumulative utopian gap. As such, we train PPO to minimize the utopian gap

$$r(\mathcal{K}_{\tau-1}, S_\tau) = -\mathcal{G}_{(N,v)}(\mathcal{K}_{\tau-1} \cup \{S_\tau\}).$$

at every step, which provides a stronger learning signal compared to the final reward. This is equivalent to training over a distribution of the online principal’s problems with uniformly distributed size t .

In our implementation, we parametrize both actor and critic of the PPO algorithm with a two-layer fully-connected neural network with 64 hidden units and ReLU activation each. To optimize the surrogate PPO objective, we used the Adam optimizer. The rest of the hyperparameteres can be found in Table 1.

Parameter	Value	Description
α_a	$3 \cdot 10^{-4}$	Actor learning rate
α_c	$1.5 \cdot 10^{-4}$	Critic learning rate
β	0.1	Entropy regularization
γ	1	Reward discounting rate
λ	0.95	Generalized advantage estimate
ϵ	0.2	Surrogate clip range
B	$5 \cdot 10^4$	Rollout buffer size
M	0.5	Max gradient norm
n_e	10	Number of training epochs

Table 1: Hyperparameters used during training.

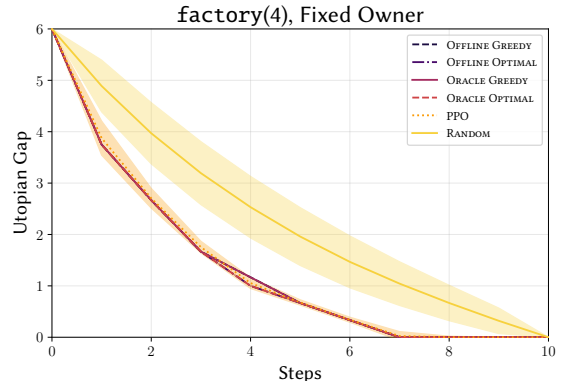


Figure 4: The cumulative utopian gap as a function of the number of revealed coalitions on a factory(4) where the owner is fixed. The greedy algorithms fail to find the global optimum at step four, see Appendix C. The performance of oracle and offline algorithms is the same in this case, because \mathcal{F} includes just a single game. Finally, PPO finds a strategy which is close to optimal, not greedy.

Random Algorithm. The RANDOM algorithm is computationally simple, requiring only $\mathcal{O}(g)$ time to compute a single sample. As such, it took only a few minutes to compute for 5 players. We again used 3,000 samples to approximate the expectation and the standard deviation, for both 4 and 5 players.

C EXAMPLE: LOCAL OPTIMUM IS NOT GLOBAL OPTIMUM

In this section, we present an example of a sequence of locally optimal steps on the principal’s problem which does *not* leading to a global optimum. We will use a single factory(4) game, i.e. the position of the owner is not randomized over. We will refer to the owner as o , and the workers as w_1, w_2, w_3 . Note that since the game is fixed, OFFLINE OPTIMAL has the same information as ORACLE OPTIMAL and thus behaves the same. Similar argument can be used to equate OFFLINE GREEDY and ORACLE GREEDY. Thus, we only analyze the ORACLE OPTIMAL and ORACLE GREEDY algorithms. The cumulative utopian gap as a function of the number of revealed coalitions is given in Figure 4.

The ORACLE GREEDY algorithm selects the coalitions $\{w_1, w_2, w_3\}$, $\{o, w_1, w_2\}$, and $\{o, w_3\}$ at the first three steps, which matches the ORACLE OPTIMAL. The resulting final utopian gap is $\frac{5}{9}$ for both algorithms. At step four, the OFFLINE GREEDY chooses $\{o, w_1, w_2\}$. However, because it chose $\{o, w_3\}$ as step three, it now doesn't match the trajectory of the ORACLE OPTIMAL, which is to choose all the coalitions of size three. As a result, the ORACLE GREEDY strategy achieves utopian gap $\frac{7}{18}$, while the ORACLE OPTIMAL strategy achieves $\frac{1}{3}$.

Importantly, the ORACLE GREEDY strategy strictly prefers its choice at $t = 3$, resulting in $\frac{5}{9}$. Choosing another one of the bigger coalitions, would result in utopian gap $\frac{7}{12}$.

Finally, the PPO algorithm outperforms ORACLE GREEDY. As the latter bounds performance of all greedy online algorithms, it shows PPO is not a greedy algorithm⁵. This is consistent with it optimizing for the expected average utopian gap, see Appendix B.

C.1 Non-Supermodularity for $|N| > 4$

In order for the utopian gap to be supermodular, it needs to satisfy $\forall S, Z \in 2^N$, and $\forall \mathcal{K} \subseteq 2^N, \mathcal{K}_0 \subseteq \mathcal{K}$

$$\mathcal{G}_\Delta = \mathcal{G}_{(N,v)}(\mathcal{K} \cup S \cup Z) - \mathcal{G}_{(N,v)}(\mathcal{K} \cup S) - \mathcal{G}_{(N,v)}(\mathcal{K} \cup Z) + \mathcal{G}_{(N,v)}(\mathcal{K}),$$

However, when the owner is player 1 and one selects $\mathcal{K} = \mathcal{K}_0 \cup \{\{1, 2, 3\}, \{1, 4\}\}$, $S = \{1, 2\}$, and $Z = \{1, 2, 3, 5\}$, the condition reads $7.3 - 7.9 - 8.1 + 8.6 = -0.1 < 0$, which shows factory (5) is not supermodular. Similar argument can be made for $n > 5$.

D ADDITIONAL EXPERIMENTAL RESULTS

D.1 Factory and Supermodular

In this section, we present studies of classes of superadditive cooperative games which were omitted from the main text due to space constraints. In Section D.1, we present results extending Section 4. In the remainder of this section, we present focus on other class of superadditive games. All results were obtained in the same way as those presented in Section 4.

In this section, we extend the results presented in Section 4 by showing how the utopian gap depends on the number of revealed coalitions for four players. Our results are presented in Figure 5. The same general trends seen in the main text are present here. The oracle algorithms outperform their offline counterparts, especially at for lower values of t . At $t = 4$, the optimal strategy is to uncover values of all coalitions of size three, which is impossible for ORACLE GREEDY, see Appendix C. In general, PPO initially performs similarly to the offline algorithms since it has no information about the underlying game. At later steps, it can leverage the known values to outperform the offline algorithms, nearly matching the performance of the ONLINE OPTIMAL.

For supermodular(4), revealing values of the largest coalition first rapidly decreases the utopian gap. In this case, revealing all coalitions of size three decreases the utopian gap by $\approx 99\%$.

Next, we show in Figure 7 a variation of Figure 2, where we do not sum over coalitions of the same size. The results show that all algorithms prefer to select the largest coalitions first. In later steps, a weak preference towards coalitions of size three emerges. Since

⁵This is because ORACLE GREEDY outperforms all greedy algorithms.

there are more coalitions of size three than two, Figure 8 suggested there is a preference for those coalitions.

Finally, we present the study of percentage of selected coalitions for the supermodular(5) class of games. Our results are presented in Figure 8 and 9 All algorithms show preference towards larger coalitions initially. At later steps, PPO deviates from the other strategies and selects coalitions pseudo-uniformly, while the rest favors coalitions of size three. We speculate this is because the utopian gap is very small at this stage, making the actions seem near equal in terms of the reward, i.e. the utopian gap.

D.2 Graph Games

Consider a complete weighted graph $G = (N, E)$ on the set of players N . Then we define the value of characteristic function for a coalition S as the sum of weights of the subgraph induced by S

$$v(S) = \sum_{i,j \in S} w_{i,j},$$

where $w_{i,j} \geq 0$ is the weight of $e_{i,j}$. We construct several distributions \mathcal{F} based on how w 's are generated; in each case, i.i.d.

We investigate three distributions in particular. The graph(n) generates weights from the uniform distribution $w \sim \mathcal{U}(0, 1)$. The graph_decreasing(n) generates weights from a distribution, whose probability density function is $2 - 2x$ on the interval $[0, 1]$. Finally, the graph_poisson(n) generates weights from the Poisson distribution with $\lambda = 1$.

We focus on $n = 4$ for these experiments. The evolution of the utopian gap as a function of the number of revealed coalitions for each algorithm is presented in Figure 10. While all algorithms outperform the RANDOM strategy, they all give similar results. The oracle methods are slightly more efficient at minimizing the utopian gap compared to the offline algorithms. PPO is initially close to the offline methods, and later outperforms them, showing it can utilize the limited knowledge of the underlying game.

D.3 Modifications of Factory

In this section, we discuss three modification of the factory class defined in the main text. In the factory_square(n) all values of coalitions are squared

$$v(S) = \begin{cases} (|S| - 1)^2 & \text{if } o \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, the factory_exp(n) makes all values exponentially larger

$$v(S) = \begin{cases} \exp(|S| - 1) & \text{if } o \in S, \\ 1 & \text{otherwise.} \end{cases}$$

Finally, in the the noisy_factory(n), each worked w_i has a random productivity $p_i \sim \mathcal{U}(0, 1)$. The value of a coalition $S \in 2^N$ is

$$v(S) = \begin{cases} \sum_{i \in S} p_i & \text{if } o \in S, \\ 0 & \text{otherwise,} \end{cases}$$

where the owner's productivity is $p_o = 0$.

We can draw similar conclusions as in the main text.

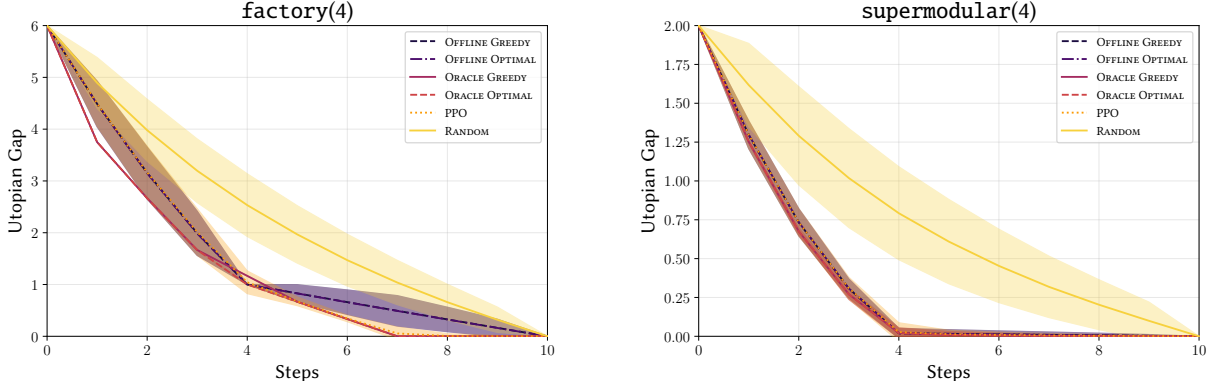


Figure 5: The utopian gap as a function of number of revealed coalitions (i.e. steps of the Principals problem) for different algorithms. We show factory(4) (left), and supermodular(4) (right) games. All algorithms outperform the RANDOM benchmark considerably. The greedy versions of each algorithm exhibit similar performance to the optimal variants. The PPO algorithm is initially close to the offline algorithms, and uses the online information to approach the oracle algorithms.

E ORACLE ALGORITHMS

Following are the pseudocodes of Oracle algorithms discussed in Section 4.

Algorithm 3: ORACLE OPTIMAL

Input: characteristic function $v \in \mathbb{S}^n$, number of steps t

- 1 $\overline{\mathcal{K}} \leftarrow 2^N \setminus \mathcal{K}_0$
 - 2 $\{S_i\}_{i=1}^t \leftarrow \operatorname{argmin}_{S \subseteq \overline{\mathcal{K}}: |S|=t} \mathcal{G}_{(N,v)}(S)$
 - 3 **return** $\{S_i\}_{i=1}^t$
-

Algorithm 4: ORACLE GREEDY

Input: characteristic function $v \in \mathbb{S}^n$, number of steps t

- 1 $\{S_i\}_{i=1}^{t-1} \leftarrow \text{Oracle Greedy}(v, t-1)$
 - 2 $\overline{\mathcal{K}} \leftarrow 2^N \setminus (\mathcal{K}_0 \cup \{S_i\}_{i=1}^{t-1})$
 - 3 $S_t \leftarrow \operatorname{argmin}_{S \in \overline{\mathcal{K}}} \mathcal{G}_{(N,v)}(\{S_i\}_{i=1}^{t-1} \cup \{S\}, v)$
 - 4 **return** $\{S_i\}_{i=1}^t$
-

F VALUE FUNCTION NORMALIZATION

In our application, it is convenient to transform a cooperative game (N, v) by an affine mapping such that after the transformation, the values of the singletons are equal to 0 and the value of the grand coalition is equal to 1. Affine transformations correspond to *strategic equivalence* in cooperative game theory. We say two games (N, v) and (N, w) are *strategically equivalent* if there is $\alpha > 0$ and $\beta \in \mathbb{R}^n$ such that

$$w(S) = \alpha \cdot v(S) + \sum_{i \in S} \beta_i. \quad (22)$$

By considering

$$\beta_i = \frac{v(\{i\})}{v(N) - \sum_{i \in N} v(\{i\})} \quad \text{and} \quad \alpha = \frac{1}{v(N)},$$

we achieve the desired transformation. Lemma F.1 describes the effect on the utopian gap, when such a transformation is applied.

LEMMA F.1. *The cumulative utopian gap \mathcal{G} is in the characteristic function invariant under the strategic equivalence, meaning*

$$\mathcal{G}_{(\alpha v + \beta)}(\mathcal{K}) = \alpha \cdot \mathcal{G}_{(N,v)}(\mathcal{K}) \quad (23)$$

where $\alpha > 0$, $(\alpha v + \beta)(S) := \alpha v(S) + \sum_{i \in N} \beta_i$ and $\beta_i \in \mathbb{R}$ for $i \in N$.

PROOF. It is a standard result [18] that

- (1) $(N, v) \in \mathbb{S}^n \implies (N, \alpha v + \beta) \in \mathbb{S}^n$,
- (2) $\phi_i(\alpha v + \beta) = \alpha \phi_i(v) + \beta_i$ for every $i \in N$.

From these two results, we have

$$\begin{aligned} \mathcal{G}_{(N, \alpha v + \beta)}(\mathcal{K}) &= \sum_{i \in N} \left(\max_{w \in \mathbb{S}^n(\mathcal{K}, \alpha v + \beta)} \phi_i(w) \right) - (\alpha v + \beta)(N) \\ &= \sum_{i \in N} [\phi_i((\alpha v + \beta)_i) - \phi_i(\alpha v + \beta)] \\ &= \sum_{i \in N} [\alpha \phi_i(v_i) + \beta_i - (\alpha \phi_i(v) + \beta_i)] \\ &= \sum_{i \in N} \alpha [\phi_i(v_i) - \phi_i(v)] \\ &= \alpha \cdot \mathcal{G}_{(N,v)}(\mathcal{K}). \end{aligned}$$

□

G EXPLICIT FORM OF THE GAP

Denote $\hat{\mathcal{K}}$, such that $\mathcal{K}_0 \subseteq \hat{\mathcal{K}} \subseteq 2^N$ and $\mathcal{K} = \hat{\mathcal{K}} \setminus \mathcal{K}_0$ and recall the definition of the gap, i.e.

$$\mathcal{G}_{(N,v)}(\mathcal{K}) = \sum_{i \in N} \left(\max_{w \in \mathbb{S}^n(\mathcal{K} \cup \mathcal{K}_0, v)} \phi_i(w) \right) - v(N).$$

In Proposition 3.4, we showed it holds $\max_{w \in \mathbb{S}^n(\mathcal{K} \cup \mathcal{K}_0, v)} \phi_i(w) = \phi_i(v_i)$, thus

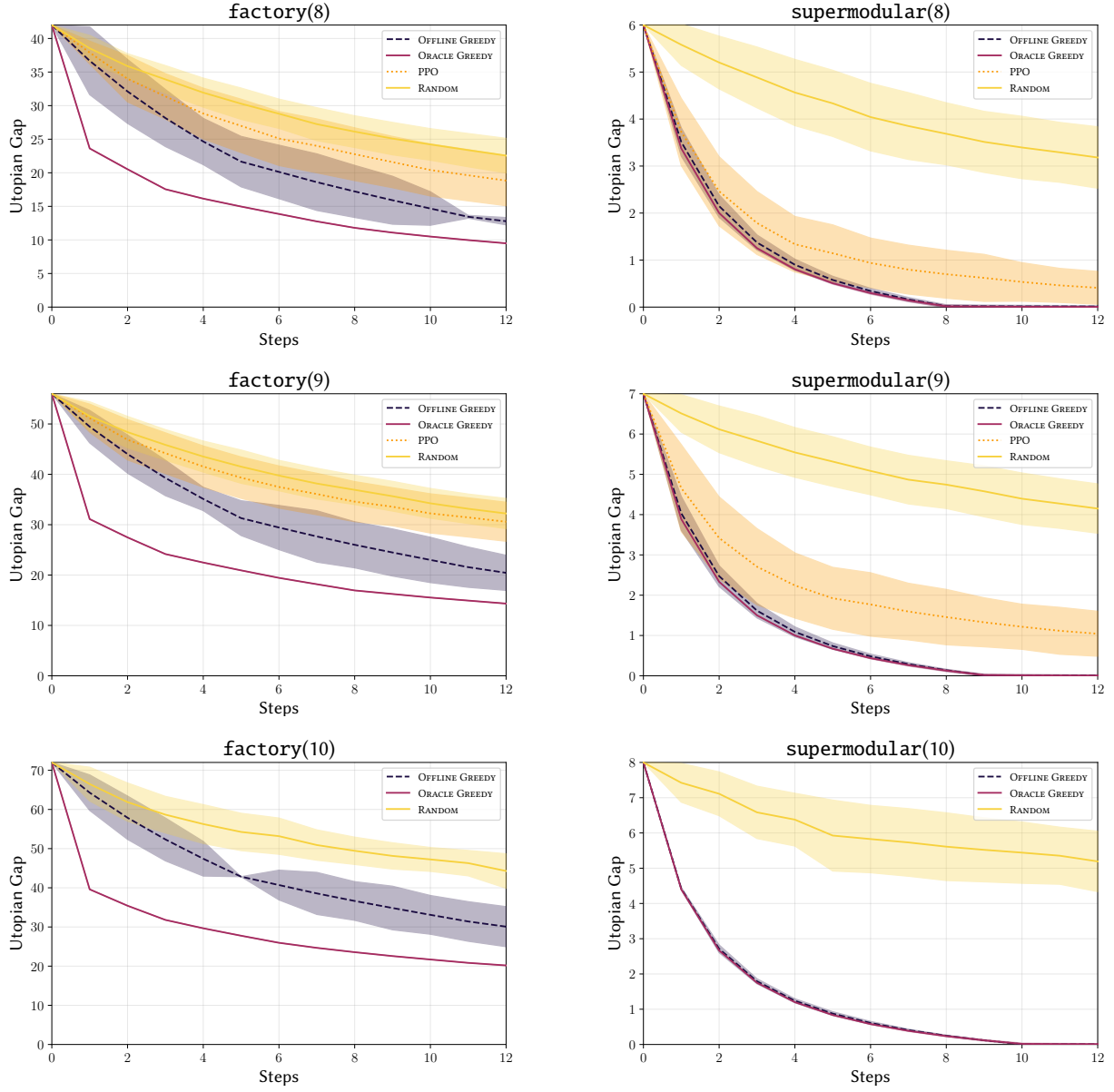


Figure 6: The utopian gap as a function of number of revealed coalitions (i.e. steps of the Principals problem) for different sizes. Note that the PPO algorithm was trained on the same number of time steps as for the 4 player problem. Even though the action space size has increased dramatically (it grows exponentially with n), it still beats the trivial random approach. Better performance of the PPO approach can still be reached by increasing the learning period. Note also, that due to time constraints, when gathering the data for the plots, we used 100 samples for sizes 8 and 9, and 10 samples for size 10, while also dropping the PPO, as the training process was too time demanding.

$$\mathcal{G}_{(N,v)}(\mathcal{K}) = \sum_{i \in N} \sum_{S \subseteq N \setminus i} \beta_{|S|} \left(\bar{v}_{\mathcal{K}}(S \cup i) - v_{\mathcal{K}}(S) \right) - v(N), \quad (24)$$

where $\beta_{|S|} = \frac{|S|!(|N|-|S|-1)!}{|N|!}$. We rewrite this double sum as a sum of $S \subseteq N$. First, $\bar{v}_{\mathcal{K}}(S)$ appears in (24) for every $i \in S$ with

coefficient $\beta_{|S \setminus i|}$. Second, $v_{\mathcal{K}}(S)$ appears in (24) for every $i \in N \setminus S$ with coefficient $\beta_{|S|}$. We get for any $i \in S$,

$$\mathcal{G}_{(N,v)}(\mathcal{K}) = \sum_{S \subseteq N} \left(\gamma_{|S|} \bar{v}_{\mathcal{K}}(S) - \delta_{|S|} v_{\mathcal{K}}(S) \right) - v(N),$$

where

- $\gamma_{|S|} = |S| \frac{(|S|-1)! (|N|-|S|)!}{|N|!} = \frac{|S|! (|N|-|S|)!}{|N|!}$,
- $\delta_{|S|} = (|N|-|S|) \frac{|S|! (|N|-|S|-1)!}{|N|!} = \frac{|S|! (|N|-|S|)!}{|N|!}$.

This concludes the proof of Proposition 3.4, as

$$\mathcal{G}_{(N,v)}(\mathcal{K}) = \sum_{S \subseteq N} \frac{|S|! (|N|-|S|)!}{|N|!} \left(\bar{v}_{\hat{\mathcal{K}}}(S) - \underline{v}_{\hat{\mathcal{K}}}(S) \right) - v(N).$$

H SUPERMODULARITY OF GAP

Supermodularity of the utopian gap is a desirable property, as it is known that greedy algorithms perform well on supermodular function [8, Proposition 3.4]. We show that as long as $|N| \leq 4$, the gap is supermodular, however, when $|N| \geq 5$, the supermodularity does not have to hold for different subsets of superadditive games. For $|N| = 5$, we give examples from different important subsets of superadditive games that do not satisfy supermodularity of the gap. Among these are totally monotonic (thus also supermodular), symmetric, and graph games. For $|N| \geq 6$, we derive a criterion, which, when satisfied, implies non-supermodularity of the gap. As a corollary, we show all three of the above mentioned classes contain games with non-supermodular gap.

Throughout this section, we implicitly assume that the value function $v : 2^N \rightarrow \mathbb{R}$ is normalized as described in Appendix F. This is without the loss of generality, as this preserves the value of the gap, see Lemma F.1. As a consequence, the lower function $\underline{v}(S)$ is zero for $\mathcal{K} = \mathcal{K}_0$ for any $S \in 2^N \setminus \mathcal{K}_0$.

H.1 Supermodularity for $|N| \leq 4$

Before we prove Proposition 3.4, we state several useful lemmata.

LEMMA H.1. *Let $\mathcal{K}_0 \subseteq \hat{\mathcal{K}} \subseteq 2^N \setminus S$ and $(N, v) \in \mathbb{S}^n$. Then $\forall T \in 2^N$*

$$\underline{v}_{\hat{\mathcal{K}}}(T) \leq \underline{v}_{\hat{\mathcal{K}} \cup S}(T) \quad \text{and} \quad \bar{v}_{\hat{\mathcal{K}}}(T) \geq \bar{v}_{\hat{\mathcal{K}} \cup S}(T).$$

PROOF. By definition of the lower game, a larger set of known coalitions can only increase the value of the lower game. Consequently, the upper game cannot increase if the set of known values becomes larger. \square

LEMMA H.2. *Let $\mathcal{K}_0 \subseteq \hat{\mathcal{K}} \subseteq 2^N \setminus S$ and $T \in 2^N$. Then*

$$T \subseteq S \implies \bar{v}_{\hat{\mathcal{K}}}(T) = \bar{v}_{\hat{\mathcal{K}} \cup S}(T),$$

and

$$S \subseteq T \implies \underline{v}_{\hat{\mathcal{K}}}(T) = \underline{v}_{\hat{\mathcal{K}} \cup S}(T).$$

PROOF. The value of T for $\bar{v}_{\hat{\mathcal{K}}}$, resp. $\bar{v}_{\hat{\mathcal{K}} \cup S}$ is given by minimal partition of T in \mathcal{K} , resp. T in $\hat{\mathcal{K}} \cup S$. As $T \subseteq S$, revealing $f(S)$ does not appear in any partition of T , thus $\bar{v}_{\hat{\mathcal{K}}}(T) = \bar{v}_{\hat{\mathcal{K}} \cup S}(T)$.

The value of T for $\underline{v}_{\hat{\mathcal{K}}}$, resp. $\underline{v}_{\hat{\mathcal{K}} \cup S}$ is given by considering maximum over $X \in \hat{\mathcal{K}}, T \subseteq X$. As $S \subseteq T$, it does not appear as a possible X . The only possibility for the change of the upper value would be that by revealing S , we get $\bar{v}_{\hat{\mathcal{K}} \cup S}(X \setminus T) \neq \bar{v}_{\hat{\mathcal{K}}}(X \setminus T)$ for some X . This means $S \subseteq X \setminus T$, but this is not possible as $S \subseteq T$. \square

LEMMA H.3. *Let $\mathcal{K}_0 \subseteq \hat{\mathcal{K}} \subseteq 2^N \setminus S$ and $T \in 2^N$, such that $\emptyset \neq S \cap T, S \subseteq T$ and $T \not\subseteq S$. Then*

$$\bar{v}_{\hat{\mathcal{K}}}(T) = \bar{v}_{\hat{\mathcal{K}} \cup S}(T), \quad \text{and} \quad \underline{v}_{\hat{\mathcal{K}}}(T) = \underline{v}_{\hat{\mathcal{K}} \cup S}(T).$$

PROOF. Follows immediately from the definition of the upper/lower games. \square

PROPOSITION 3.9. *For $|N| \leq 4$, the gap of an incomplete \mathbb{S}^n -extendable game is supermodular.*

PROOF. Denote $\hat{\mathcal{K}}$, such that $\mathcal{K}_0 \subseteq \hat{\mathcal{K}} \subseteq 2^N$ and $\mathcal{K} = \hat{\mathcal{K}} \setminus \mathcal{K}_0$. As $\mathcal{G}_{(N,v)}(\mathcal{K}) = \sum_{T \subseteq N} \alpha_T \Delta_{\hat{\mathcal{K}}}(T) - v(N)$ where $\alpha_T = \frac{|T|! (|N|-|T|)!}{|N|!}$, the condition of supermodularity is

$$\sum_{T \subseteq N} \alpha_T \left(\Delta_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \Delta_{\hat{\mathcal{K}} \cup S}(T) \right) \geq \sum_{T \subseteq N} \alpha_T \left(\Delta_{\hat{\mathcal{K}} \cup Z}(T) - \Delta_{\hat{\mathcal{K}}}(T) \right).$$

To prove this proposition, we show an even a stronger condition holds, i.e., $\forall T \in 2^N$

$$\alpha_T \left(\Delta_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \Delta_{\hat{\mathcal{K}} \cup S}(T) \right) \geq \alpha_T \left(\Delta_{\hat{\mathcal{K}} \cup Z}(T) - \Delta_{\hat{\mathcal{K}}}(T) \right),$$

or since $\alpha_T \geq 0$ for every T , this is equivalent to

$$\bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \bar{v}_{\hat{\mathcal{K}} \cup S}(T) + \underline{v}_{\hat{\mathcal{K}} \cup S}(T) \geq \bar{v}_{\hat{\mathcal{K}} \cup Z}(T) - \underline{v}_{\hat{\mathcal{K}} \cup Z}(T) - \bar{v}_{\hat{\mathcal{K}}}(T) + \underline{v}_{\hat{\mathcal{K}}}(T). \quad (25)$$

For $T \in \hat{\mathcal{K}} \cup S$, condition (25) holds for any superadditive $v : 2^N \rightarrow \mathbb{R}$ with N of arbitrary size. To see this, notice $\bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) = \underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T)$ and $\bar{v}_{\hat{\mathcal{K}} \cup S}(T) = \underline{v}_{\hat{\mathcal{K}} \cup S}(T)$, thus (25) reduces to

$$\bar{v}_{\hat{\mathcal{K}} \cup S}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S}(T) \geq \bar{v}_{\hat{\mathcal{K}}}(T) - \underline{v}_{\hat{\mathcal{K}}}(T).$$

This inequality follows from Lemma H.1.

First, consider case $|N| = 3$. Since, only coalitions of size 2 are possibly unknown, their lower and upper bounds do not change when another subset of size 2 is revealed. Thus, for $\mathcal{K}_0 \subseteq \hat{\mathcal{K}} \subseteq 2^N \setminus \{S, Z\}$ and $T \notin \hat{\mathcal{K}} \cup S$, it holds

$$\bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) = \bar{v}_{\hat{\mathcal{K}} \cup S}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S}(T)$$

and similarly

$$\bar{v}_{\hat{\mathcal{K}} \cup Z}(T) - \underline{v}_{\hat{\mathcal{K}} \cup Z}(T) = \bar{v}_{\hat{\mathcal{K}}}(T) - \underline{v}_{\hat{\mathcal{K}}}(T),$$

which means (25) holds.

Now, consider case $|N| = 4$. To prove (25) for $\mathcal{K}_0 \subseteq \hat{\mathcal{K}} \subseteq 2^N \setminus \{S, Z\}$ and $T \notin \hat{\mathcal{K}} \cup S$, we distinguish several cases based on the relation between Z and T . All of the cases follow a similar pattern.

(1) $Z \subseteq T$:

Since $|N| = 4$ and Z, T are unknown in $\hat{\mathcal{K}} \cup S$, they must be of form $Z = \{i, j\}$ and $T = \{i, j, k\}$. By Lemma H.2, $\underline{v}_{\hat{\mathcal{K}} \cup Z}(T) = \underline{v}_{\mathcal{K}}(T)$ and $\underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) = \underline{v}_{\hat{\mathcal{K}} \cup S}(T)$, thus (25) reduces to

$$\bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \bar{v}_{\hat{\mathcal{K}} \cup S}(T) \geq \bar{v}_{\hat{\mathcal{K}} \cup Z}(T) - \bar{v}_{\mathcal{K}}(T).$$

For a contradiction, suppose the converse holds. As, by Lemma H.1, $\bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \bar{v}_{\hat{\mathcal{K}} \cup Z}(T) \geq 0$, this means $\bar{v}_{\hat{\mathcal{K}} \cup Z}(T) > \bar{v}_{\mathcal{K}}(T)$, which leads, by Lemma H.1, to a contradiction.

(2) $T \subseteq Z$: Similarly to the first case, we have $T = \{i, j\}$, $Z = \{i, j, k\}$ and by Lemma H.2, (25) reduces to

$$\underline{v}_{\hat{\mathcal{K}} \cup S}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) \geq \underline{v}_{\mathcal{K}}(T) - \underline{v}_{\hat{\mathcal{K}} \cup Z}(T).$$

For a contradiction, suppose the converse holds. It means $\underline{v}_{\mathcal{K}}(T) > \underline{v}_{\hat{\mathcal{K}} \cup Z}(T)$, which by Lemma H.1 leads to a contradiction.

- (3) $Z = T$: It holds $\bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) = \underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T)$ and $\bar{v}_{\hat{\mathcal{K}} \cup Z}(T) = \underline{v}_{\hat{\mathcal{K}} \cup Z}(T)$, thus (25) reduces to

$$\bar{v}_{\hat{\mathcal{K}}}(T) - \underline{v}_{\mathcal{K}}(T) \geq \bar{v}_{\hat{\mathcal{K}} \cup S}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S}(T)$$

which holds by Lemma H.1. In this case, $Z = \{i, j\}$, $T = \{k, \ell\}$ and (25) reduces to

$$\underline{v}_{\hat{\mathcal{K}} \cup S}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) \geq \underline{v}_{\mathcal{K}}(T) - \underline{v}_{\hat{\mathcal{K}} \cup Z}(T)$$

as $\bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) = \bar{v}_{\hat{\mathcal{K}} \cup S}(T)$ and $\bar{v}_{\hat{\mathcal{K}} \cup Z}(T) = \bar{v}_{\hat{\mathcal{K}}}(T)$. For a contradiction, if $<$ holds, similarly to previous cases, $\underline{v}_{\hat{\mathcal{K}}}(T) > \underline{v}_{\hat{\mathcal{K}} \cup S}(T)$, a contradiction with Lemma H.1.

- (4) $Z \cap T \neq \emptyset$ and $Z \not\subseteq T$ and $T \not\subseteq Z$: By Lemma H.3, condition (25) reduces to $0 \geq 0$. \square

H.2 Non-supermodularity for $|N| = 5$

All of the examples of games with non-supermodular gap are based on the same principle. We fix $\mathcal{K} = \mathcal{K}_0 \cup \{1, 2, 3\}$, $S = \{1, 2\}$ and $Z = \{3, 4\}$. Then if we denote

$$\mathcal{G}_\Delta = \mathcal{G}_{(N,v)}(\mathcal{K} \cup S \cup Z) - \mathcal{G}_{(N,v)}(\mathcal{K} \cup S) - \mathcal{G}_{(N,v)}(\mathcal{K} \cup Z) + \mathcal{G}_{(N,v)}(\mathcal{K}),$$

a necessary condition of supermodularity of $\mathcal{G}_{(N,v)}$ is thus $\mathcal{G}_\Delta \geq 0$. We note that a necessary, however not sufficient condition for violation of this condition is $v(\{1, 2\}) + v(\{3, 4\}) > v(\{1, 2, 3\})$.

Example H.4. (Totally monotonic game)

Totally monotonic games are defined as non-negative combinations of so called *unanimity games*. Unanimity game (N, u_S) for $S \subseteq N$ is defined as

$$u_S(T) = \begin{cases} 1 & S \subseteq T, \\ 0 & \text{otherwise.} \end{cases}$$

Any totally monotonic game can be then expressed as $\sum_{S \subseteq N} \alpha_S u_S$ where $\alpha_S \geq 0$ for every S . They form a subset of convex games, which under the name *belief measures* forms the foundation on the *evidence theory* [9, 24]

An example of a totally monotonic game with non-supermodular gap is $u_{\{3,4\}} + u_{\{1,2,3\}}$, for which $\mathcal{G}_\Delta = 0.1 \not\geq 0$.

Example H.5. (Symmetric game) Symmetric cooperative game (N, v) satisfies for every $S, T \subseteq N$, $|S| = |T|$ that $v(S) = v(T)$. Symmetric game of n players can be thus defined by $s_1, \dots, s_n \in \mathbb{R}$ corresponding to values of coalitions of different sizes. These games are often studied for their robust properties with respect to solution concepts [18]).

It holds that if symmetric game on 5 players satisfy $s_2 = s_3 = 1$ and $s_5 = 2$, it follows $\mathcal{G}_\Delta = -0.1 \not\geq 0$.

Example H.6. (Graph game) Graph game (N, v) is defined using a complete weighted graph $G = (N, E)$ as

$$v(S) = \sum_{i,j \in S} w_{i,j},$$

where $w_{i,j} \geq 0$ is the weight of $e_{i,j}$. These games have interesting properties from the point of view of algorithmic complexity and have been thoroughly studied in the past decades [5].

For a complete graph (N, E) , where $N = \{1, 2, 3, 4, 5\}$, $w_{1,2} = 1 = w_{3,4} = 1$, it holds $\mathcal{G}_\Delta = -0.1 \not\geq 0$.

H.3 Criterion for $|N| \geq 6$

PROPOSITION 3.10. *For $|N| \geq 6$, the gap of an incomplete \mathbb{S}^n -extendable game (N, w) is not supermodular if there exist $i, j, k, l \in N$ such that $v(ij) \leq v(jk) \leq v(kl)$, where $v(S) = w(S) - \sum_{i \in S} w(\{i\})$, and*

$$v(kl) < \left[\binom{n}{2} \binom{n}{\lfloor n/2 \rfloor}^{-1} (2^{n-3} - n + 2) - 1 \right] v(ij). \quad (26)$$

PROOF. By Lemma F.1 in Appendix F, the gap of (N, w) is supermodular if and only if the gap of (N, v) is supermodular. We focus on the gap (N, v) as this makes our analysis simple and at the same time, the condition easy to check.

The supermodularity condition (3) for the gap is

$$\mathcal{G}_{(N,v)}(\hat{\mathcal{K}} \cup S \cup Z) - \mathcal{G}_{(N,v)}(\hat{\mathcal{K}} \cup S) \geq \mathcal{G}_{(N,v)}(\hat{\mathcal{K}} \cup Z) - \mathcal{G}_{(N,v)}(\hat{\mathcal{K}}), \quad (27)$$

where $\hat{\mathcal{K}} \subseteq 2^N \setminus \mathcal{K}_0$.

We have shown in Appendix G that the utopian gap is

$$\mathcal{G}_{(N,v)}(\mathcal{K}) = \sum_{S \subseteq N} \gamma_S (\bar{v}_{\hat{\mathcal{K}}}(S) - \underline{v}_{\hat{\mathcal{K}}}(S)) - v(N), \quad (28)$$

where $\gamma_S = \frac{|S|!(|N|-|S|)!}{|N|!}$. This equation can be viewed as a weighted l_1 -norm of the difference between the upper and lower games.⁶

Using this equality, we can rewrite Eq. (27) as

$$\sum_{T \subseteq N} \gamma_T (\bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \bar{v}_{\hat{\mathcal{K}} \cup S}(T) + \underline{v}_{\hat{\mathcal{K}} \cup S}(T)) \geq \sum_{T \subseteq N} \gamma_T (\bar{v}_{\hat{\mathcal{K}} \cup Z}(T) - \underline{v}_{\hat{\mathcal{K}} \cup Z}(T) - \bar{v}_{\hat{\mathcal{K}}}(T) + \underline{v}_{\hat{\mathcal{K}}}(T)). \quad (29)$$

We obtain the result by setting $\mathcal{K} = \mathcal{K}_0 \cup \{\{j, k\}\}$, $S = \{i, j\}$, and $Z = \{k, l\}$. Denote $v(ij) = v(jk) - \varepsilon$ and $v(kl) = v(jk) + \delta$ for some $\varepsilon, \delta \geq 0$. For these values, we investigate each of the weighted terms separately. Let

$$L_T = \bar{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \underline{v}_{\hat{\mathcal{K}} \cup S \cup Z}(T) - \bar{v}_{\hat{\mathcal{K}} \cup S}(T) + \underline{v}_{\hat{\mathcal{K}} \cup S}(T), \quad (30)$$

$$R_T = \bar{v}_{\hat{\mathcal{K}} \cup Z}(T) - \underline{v}_{\hat{\mathcal{K}} \cup Z}(T) - \bar{v}_{\hat{\mathcal{K}}}(T) + \underline{v}_{\hat{\mathcal{K}}}(T). \quad (31)$$

In the reminder of the proof, we distinguish different cases based on the relation of T and $\{i, j, k, l\}$ for all $T \in 2^N \setminus \mathcal{K}_0$.

- (1) $\{i, j, k, l\} \subseteq T \subseteq N$:

Using Lemma H.2, we get

$$L_T = \underline{v}_{\mathcal{K} \cup S}(T) - \underline{v}_{\mathcal{K} \cup S \cup Z}(T),$$

$$R_T = \underline{v}_{\mathcal{K}}(T) - \underline{v}_{\mathcal{K} \cup Z}(T),$$

which can be expressed as

$$L_T = v(jk) - v(jk) + \varepsilon - v(jk) - \delta,$$

$$R_T = v(jk) - v(jk) - \delta.$$

It holds $L_T = R_T + \varepsilon - v(jk) = R_T - v(ij)$.

- (2) $\{i, j, k, l\} \subseteq N \setminus T$:

In this case, the lower bound on $v(T)$ is unaffected by the knowledge of S, Z , so

$$L_T = \bar{v}_{\mathcal{K} \cup S \cup Z}(T) - \bar{v}_{\mathcal{K} \cup S}(T),$$

$$R_T = \bar{v}_{\mathcal{K} \cup Z}(T) - \bar{v}_{\mathcal{K}}(T).$$

⁶Assuming they are viewed as vectors in \mathbb{R}^{2^n} .

Since the only superset of T in $\mathcal{K} \cup S \cup Z$ is N , the upper bound (9) reduces to

$$\bar{v}_{\mathcal{L}}(T) = v(N) - \underline{v}_{\mathcal{L}}(N \setminus T) = 1 - \underline{v}_{\mathcal{L}}(N \setminus T),$$

for every $\mathcal{L} \in \{\mathcal{K}, \mathcal{K} \cup S, \mathcal{K} \cup Z, \mathcal{K} \cup S \cup Z\}$. Thus

$$L_T = \underline{v}_{\mathcal{K} \cup S}(N \setminus T) - \underline{v}_{\mathcal{K} \cup S \cup Z}(N \setminus T),$$

$$R_T = \underline{v}_{\mathcal{K}}(N \setminus T) - \underline{v}_{\mathcal{K} \cup Z}(N \setminus T).$$

Further, $i, j, k, l \in N \setminus T$, which means

$$L_T = v(jk) - v(jk) + \varepsilon - v(jk) - \delta,$$

$$R_T = v(jk) - v(jk) - \delta.$$

As in the previous case, $L_T = R_T + \varepsilon - v(jk) = R_T - v(ij)$.

(3) $S = T$:

In this case, since $\underline{v}_{\mathcal{L}}(T) = \bar{v}_{\mathcal{L}}(T) = v(T)$ if $T \in \mathcal{L}$ for every $\mathcal{L} \in \{\mathcal{K}, \mathcal{K} \cup S, \mathcal{K} \cup Z, \mathcal{K} \cup S \cup Z\}$ we get

$$L_T = 0,$$

$$R_T = \bar{v}_{\mathcal{K} \cup Z}(T) - \bar{v}_{\mathcal{K}}(T) = \underline{v}_{\mathcal{K}}(N \setminus T) - \underline{v}_{\mathcal{K} \cup Z}(N \setminus T).$$

This means $R_T = 0 - v(jk) - \delta$, therefore it holds $L_T = R_T + v(jk) + \delta = R_T + v(kl)$.

(4) $Z = T$:

In this case, by the fact that $\underline{v}_{\mathcal{L}}(T) = \bar{v}_{\mathcal{L}}(T) = v(T)$ if $T \in \mathcal{L}$, it holds

$$L_T = \underline{v}_{\mathcal{K} \cup S}(T) - \bar{v}_{\mathcal{K} \cup S}(T),$$

$$R_T = \underline{v}_{\mathcal{K}}(T) - \bar{v}_{\mathcal{K}}(T).$$

Since $S \not\subseteq T$, it holds $\underline{v}_{\mathcal{K} \cup S}(T) = \underline{v}_{\mathcal{K}}(T)$, so we can rewrite

$$L_T = \underline{v}_{\mathcal{K} \cup S}(N \setminus T) - v(N) = v(jk) - \varepsilon - v(N),$$

$$R_T = \underline{v}_{\mathcal{K}}(N \setminus T) - v(N) = -v(N).$$

Thus it holds $L_T = R_T + v(jk) - \varepsilon = R_T + v(ij)$.

(5) $\emptyset \neq Z \cap T, Z \not\subseteq T$, and $T \not\subseteq Z$ or $\emptyset \neq S \cap T, S \not\subseteq T$, and $T \not\subseteq S$:

By Lemma H.1, and Lemma H.3, one immediately arrives at $L_T = R_T = 0$.

A necessary condition for supermodularity of the gap is

$$\sum_{T \subseteq N} \gamma_{|T|} (L_T - R_T) \geq 0.$$

This means, that if any upper bound L ,

$$\sum_{T \subseteq N} \gamma_{|T|} (L_T - R_T) \leq L$$

satisfies $L < 0$, the gap is not supermodular. To construct the upper bound, we focus on terms of the sum corresponding to those from the first and the second case above, i.e. $T \subseteq N$ s.t. $\{i, j, k, l\} \subseteq T \subseteq N$ or $T \subseteq N \setminus \{i, j, k, l\}$. In these cases, $L_T - R_T \leq 0$. Since $\gamma_{|T|} \leq \gamma_{\lfloor n/2 \rfloor}$ for every $T \subseteq N$, we bound the corresponding terms by

$$\gamma_{|T|} (L_T - R_T) \leq \gamma_{\lfloor n/2 \rfloor} (L_T - R_T).$$

The upper bound thus sums to

$$L = -\gamma_{\lfloor n/2 \rfloor} (c_1(n) + c_2(n))v(ij) + \gamma_2(v(kl) + v(ij)),$$

where $c_1(n) = 2^{n-4} - 1$ and $c_2(n) = 2^{n-4} - n + 3$ is the number of terms satisfying conditions 1 and 2, respectively. Thus, $L < 0$ is equal to

$$v(kl) < \left(\frac{\gamma_{\lfloor n/2 \rfloor}}{\gamma_2} (c_1(n) + c_2(n)) - 1 \right) v(ij),$$

or since $c_1(n) + c_2(n) = 2^{n-3} - n + 2$, it is equal to

$$v(kl) < \left[\binom{n}{2} \binom{n}{\lfloor n/2 \rfloor}^{-1} (2^{n-3} - n + 2) - 1 \right] v(ij). \quad (32)$$

The coefficient for $v(ij)$ is equal $2, \frac{28}{5}, \frac{47}{5}, \dots$ for $n = 6, 7, 8, \dots$ and the sequence can be shown to be increasing. For $n = 5$, the coefficient is equal to 0, which makes the criterion useless as it translates to $0 \leq v(kl) < 0$. \square

A corollary of the criterion is that, in any game where $|N| \geq 6$ and there exist $i, j, k, l \in N$ such that $v(ij) \leq v(jk) \leq v(kl)$ and $v(kl) < 2v(ij)$ does not satisfy supermodularity of the gap. The criterion is satisfied by any game satisfying $v(ij) = v(jk) = v(kl) \neq 0$ for some i, j, k, l . Examples of totally monotonic, symmetric superadditive, and graph games are immediate. Note that almost every symmetric superadditive ($v(ij) \neq 0$) as well as almost every additive game (different from zero game) has a non-supermodular gap.

H.4 Simpler criterion for $|N| \geq 7$

Using the same technique as in H.3, one might derive weaker, however, simpler criterion. E.g. since $-\gamma_{\lfloor n/2 \rfloor} (c_1(n) + c_2(n)) \leq -\frac{1}{10}$ for every $n \geq 5$, it holds $L \leq -\frac{1}{10}v(ij) + \gamma_2(v(kl) + v(ij))$, which is strictly smaller than 0 if $v(kl) < (1 - 10\gamma_2)v(ij)/10\gamma_2$, or

$$v(kl) < \frac{n^2 - n - 20}{20} v(ij). \quad (33)$$

However, this criterion is useful only when $n \geq 7$, as for $n = 5, 6$, it holds $\alpha_n = \frac{n^2 - n - 20}{20} < 1$, which implies

$$0 \leq v(ij) \leq v(kl) < \alpha_n v(ij),$$

thus $v(ij) = v(kl) = 0$ and the criterion simplifies to $0 < \alpha_n 0$, which cannot be satisfied. However, for $n \geq 7$, it holds

$$\frac{5}{4} v(ij) \leq \alpha_n v(ij),$$

thus any game satisfying

$$0 < v(ij) \leq v(kl) < \frac{5}{4} v(ij)$$

cannot have a supermodular gap.

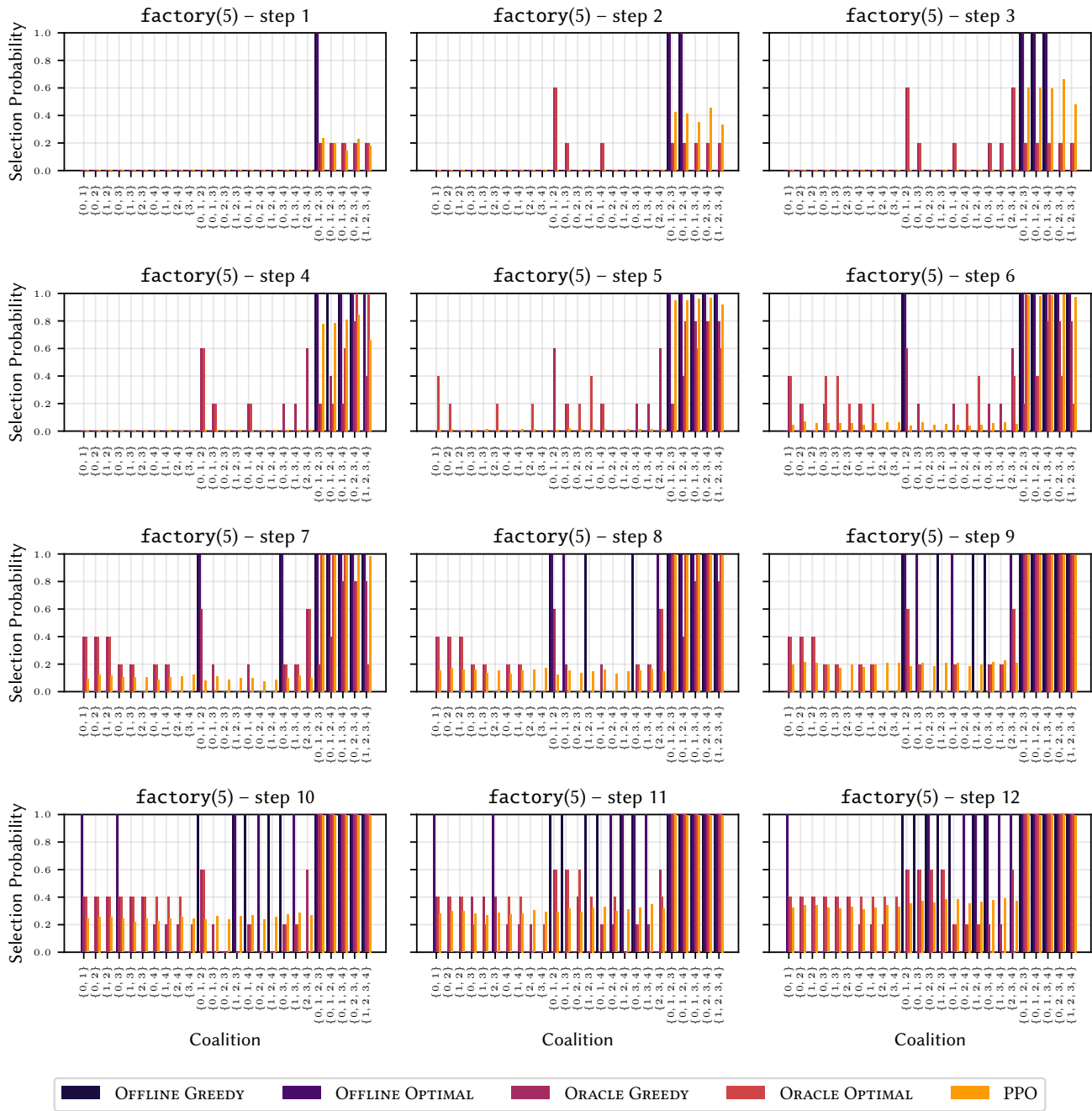


Figure 7: Percentage of coalitions selected up to step twelve for factory(5) and each algorithm. Results show clear preference for larger coalitions, i.e. they contribute more information about the cooperative game on average. The oracle algorithms favor smaller coalitions earlier, suggesting the representation of a specific game can efficiently use even smaller coalitions. PPO initially behaves similarly to the offline algorithms. At later steps, it uses the previously obtained values and its selections resemble the oracle methods. See Figure 2 for plot showing individual coalitions.

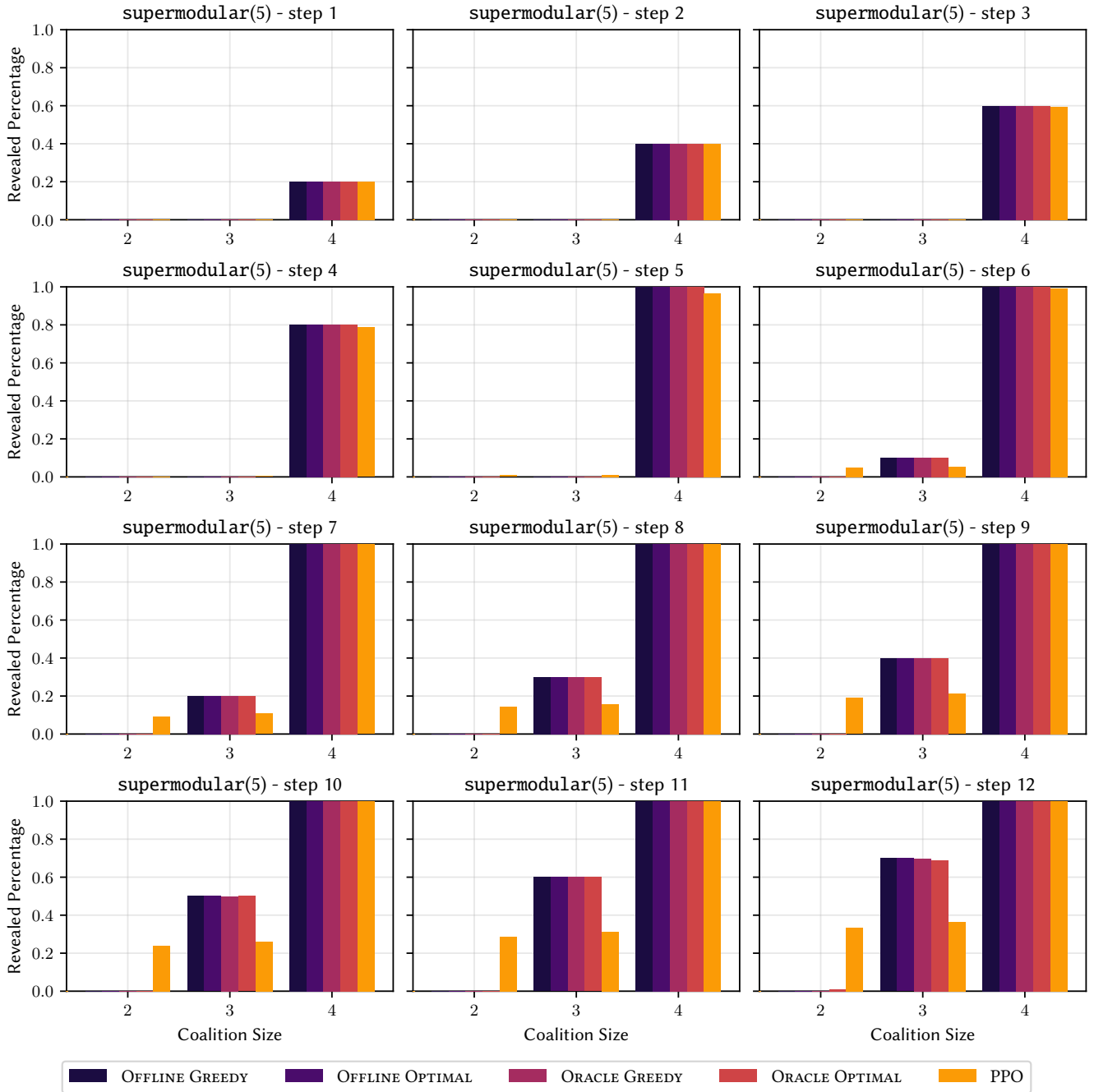


Figure 8: Percentage of coalitions selected up to step twelve for supermodular(5) and each algorithm. Results show clear preference for larger coalitions, i.e. they contribute more information about the cooperative game on average. The oracle algorithms favor smaller coalitions earlier, suggesting the representation of a specific game can efficiently use even smaller coalitions. PPO initially behaves similarly to the offline algorithms. At later steps, it uses the previously obtained values and its selections resemble the oracle methods. See Figure 9 for plot showing individual coalitions.

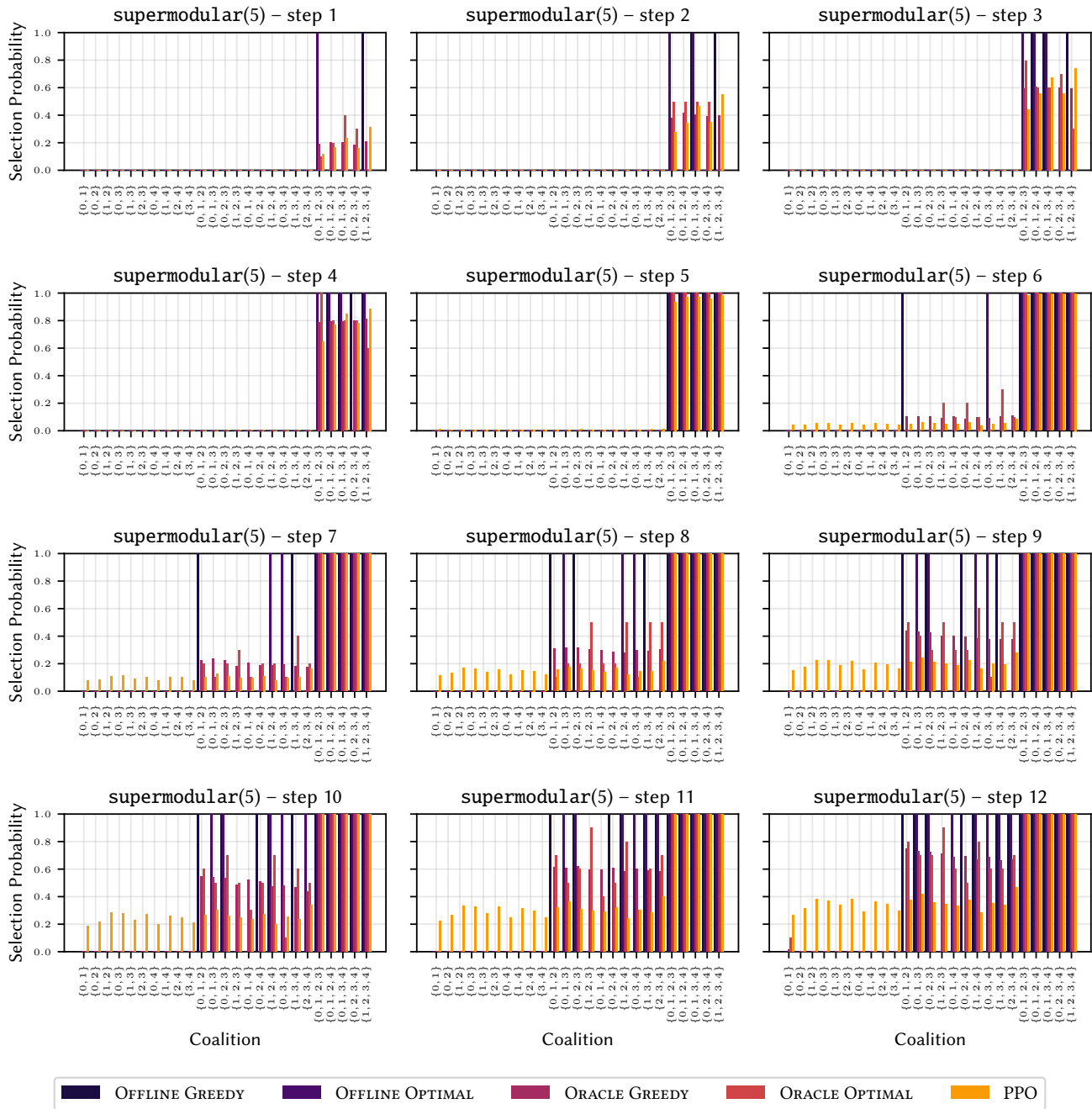


Figure 9: Percentage of coalitions selected up to step twelve for supermodular(5) and each algorithm. Results show clear preference for larger coalitions, i.e. they contribute more information about the cooperative game on average. The oracle algorithms favor smaller coalitions earlier, suggesting the representation of a specific game can efficiently use even smaller coalitions. PPO initially behaves similarly to the offline algorithms. At later steps, it uses the previously obtained values and its selections resemble the oracle methods. See Figure 8 for plot showing individual coalitions.

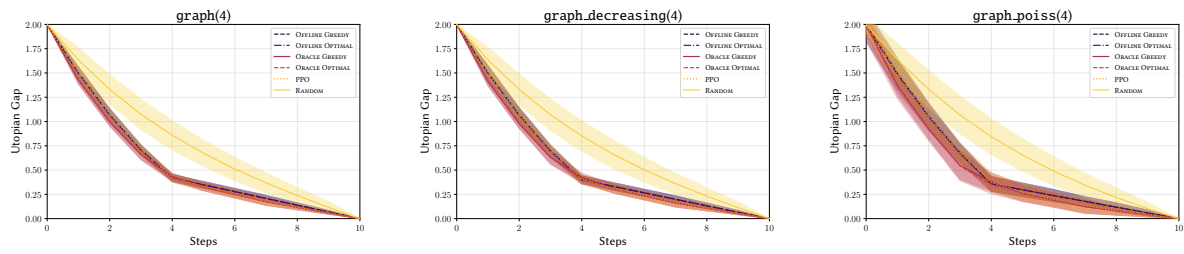


Figure 10: The cumulative utopian gap as a function of the number of revealed coalitions on graph cooperative games. We use graph(4) (left), graph_decreasing(4) (middle) and graph_pois(4) (right) distributions to generate the weights of the graph. See Appendix D.2 for more details.

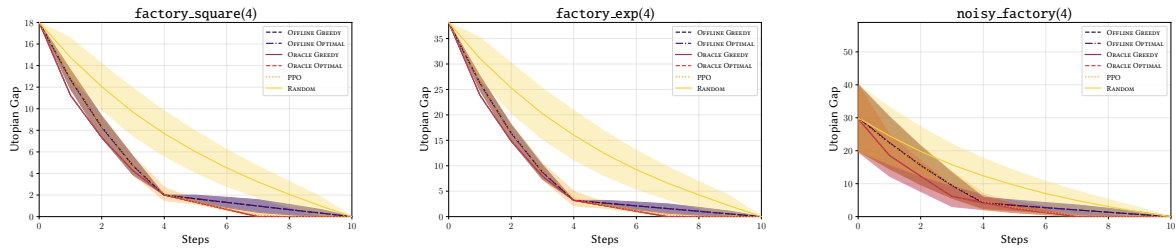


Figure 11: The cumulative utopian gap as a function of the number of revealed coalitions on graph cooperative games. We use factory_square(4) (left), factory_exp(4) (middle) and noisy_factory(4) (right) distributions to generate the weights of the graph. See Appendix D.3 for more details.