

Distributed Submodular Maximization with Limited Information

Bahman Ghahesifard Stephen L. Smith

Abstract—We consider a class of distributed submodular maximization problems in which each agent must choose a single strategy from its strategy set. The global objective is to maximize a submodular function of the strategies chosen by each agent. When choosing a strategy, each agent has access to only a limited number of other agents’ choices. For each of its strategies, an agent can evaluate its marginal contribution to the global objective given its information. The main objective is to investigate how this limitation of information about the strategies chosen by other agents affects the performance when agents make choices according to a local greedy algorithm. In particular, we provide lower bounds on the performance of greedy algorithms for submodular maximization, which depend on the clique number of a graph that captures the information structure. We also characterize graph-theoretic upper bounds in terms of the chromatic number of the graph. Finally, we demonstrate how certain graph properties limit the performance of the greedy algorithm. Simulations on several common models for random networks demonstrate our results.

I. INTRODUCTION

There has been a recent surge of activity in developing efficient algorithms for large-scale optimization problems, by partitioning the objective into smaller subtasks, each of which is optimized separately. A well-studied class of such problems is the distributed optimization problem, where the main objective is to optimize a sum of functions, each one available to an individual agent, by means of limited information sharing between individuals [2], [3], [4], [5], [6], [7], [8], [9], [10]. Unlike this setting, where strategy sets are often assumed to be continuous sets, here we consider a collaborative scenario, where agents choose a strategy from a private *discrete* strategy set and the goal is to maximize a common objective function defined over the set of strategies chosen by each agent. In this sense, this work can be considered as one possible way to extend the literature on studying distributed optimization with limited information and continuous strategy sets to combinatorial scenarios. We consider situations where the total strategy set is partitioned into subsets, each available only to an individual agent, and the common objective function is additionally assumed to be *submodular*. When choosing a strategy, each agent has access to only a limited number

of other agents’ choices, and can evaluate its marginal contribution to the global objective given its information. This problem arises in a variety of applications of interest, including problems in distributed sensor coverage [11], [12], information gathering [13], [14], and facility location [15].

Building on the early classic work [16], [17], [18], [19], there has been an emerging body of work on submodular optimization, mainly due to its wide set of applications to areas of computer science, see for example [15] and references therein. It is well known that the submodular maximization problem is NP-hard [18], unlike submodular minimization, which can be solved in polynomial time [18]. However, good approximation algorithms exist; in particular, when the submodular function is normalized and monotone, a simple greedy algorithm, yields a solution that is within a multiplicative factor of $(1 - 1/e)$ of the optimal. The distributed problem we consider in this paper can be captured as a partition matroid constraint; in particular, there is a bound on the cardinality of the intersection of the solution set and the strategy set of each individual agent. It has been shown [16] that maximizing a normalized and monotone submodular function subject to a matroid constraint can be approximated to within $1/2$ using the simple greedy algorithm. More recently [20] a randomized algorithm has been proposed for this problem, which yields an approximation of $(1 - 1/e)$.

The key feature that differentiates the problem under study in this paper from the classic setting of submodular maximization is the limitations on the information available to a decision maker about the strategies chosen by others. Similar to the literature on distributed optimization, in order to achieve any reasonable performance, agents need to share information about their actions. The main objective of this paper is to investigate how much information is required to guarantee a certain performance, and what limitations are put on the performance by the information structure between agents.

It is worth mentioning that parallel computations in submodular optimization has recently generated a lot of interest [21], [22], [23], [24], [25]. These works focus primarily on large-scale maximization or minimization problems, solving them by partitioning the problem into subsets, each of which is optimized separately. The overall solution is aggregated and refined through additional computations by a central node, often in a MapReduce framework. In our work, however, we envision scenarios where the limitations imposed on individual agents is physical (for example, being able to only estimate data in a neighbourhood). This paper is also somewhat related to the recent work [11], where the role of limitation of information in submodular optimization is studied in the context

Bahman Ghahesifard is with the Department of Mathematics and Statistics at Queen’s University, Kingston, ON, Canada ghareb@queensu.ca. Stephen L. Smith is with the Department of Electrical and Computer Engineering at the University of Waterloo, Waterloo, ON, Canada stephen.smith@uwaterloo.ca.

The work of the authors was partially supported by the Natural Sciences and Engineering Research Council of Canada. A preliminary version of this paper appeared in the Proceedings of the 2016 American Control Conference [1].

of coverage problem, for the cases where agents have full information or no information. Here, we address the limitations imposed by the information network topology. Finally, part of our work is related to the so-called “local greedy algorithms”, which is studied in the classical paper [17], as we describe in details later.

Statement of Contributions: We consider a class of distributed submodular maximization problems in which each agent has a strategy set from which it must choose a single strategy. The global objective is to maximize a submodular function of the set of strategies chosen by each agent. The group of agents take decisions sequentially, having available to them only limited information about the strategies chosen by previous agents. Each agent can evaluate its marginal contribution to the global objective given its information. The information structure is cast as a directed acyclic graph (DAG) and our main objective is to characterize the fundamental limitations that information limitations impose on the performance of local greedy algorithms. We first show that the well-known $\frac{1}{2}$ lower bound on the performance of greedy algorithms for submodular maximization can be obtained when the information structure is a complete DAG. We provide a general lower bound on the performance, which depends on the clique number of the graph, and also provide lower bounds for graph topologies with multiple interconnected cliques. We then characterize graph-theoretic upper bounds on the performance of the local greedy algorithm. We tackle this objective by considering two problem statements. In the first one, we characterize fundamental limitations on the performance of the greedy algorithms on a given graph topology in terms of its chromatic number. Our second result demonstrates how achieving a certain guaranteed performance imposes limitations on the topology of the underlying graph. We characterize the gap between these two bounds, and show how they can be used to efficiently compute limitations of distributed greedy algorithms for scenarios with limited information. Case studies on several types of randomly generated graphs demonstrate the relationship between the performance bounds provided and the true performance of the distributed greedy algorithm.

A preliminary version of this paper appeared as [1]. Relative to this early work, key contributions include proofs of all results; a treatment of intersecting strategy sets and synchronous agent updates; applications to sensor coverage; and, extensive simulation results for several random graph models.

II. PRELIMINARIES

Many combinatorial optimization problems can be phrased as a *submodular maximization problem*, which can be stated as follows. Consider a base set of elements E , and let 2^E be the power set of E , containing all of its subsets. Then a function $f : 2^E \rightarrow \mathbb{R}_{\geq 0}$ is *submodular* if it possesses the property of diminishing returns: For all $A \subseteq B \subseteq S$, and for all $x \in S \setminus B$ we have

$$f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B).$$

We refer to $f(A \cup \{x\}) - f(A)$ as the *marginal reward* of x given A , and denote it by $\Delta(x|A)$. This definition is equivalent

to stating that for all $X, Y \subseteq S$, we have that

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y). \quad (1)$$

In addition to submodularity, we will consider functions that possess two further properties:

- (i) *Monotonicity:* For all $A \subseteq B \subseteq E$, we have $f(B) \geq f(A)$; and
- (ii) *Normalization:* $f(\emptyset) = 0$.

Given a normalized submodular function f defined over a base set E containing $|E| = n$ elements, along with a positive integer $k < n$, the submodular maximization problem can be stated as

$$\max_{S \subseteq E} f(S) \quad (2)$$

$$\text{subject to } |S| \leq k. \quad (3)$$

It is well known that this optimization problem is NP-hard. It contains as a special case several well known combinatorial optimization problems including MAX CUT and FACILITY LOCATION. Even when the function is monotone, the problem is still NP-hard, and contains as a special case MAX k -COVER.

When the submodular function is normalized and monotone, a simple greedy algorithm, in which a solution is built incrementally by adding the element x to S that maximizes $\Delta(x|S)$ yields a solution that is within a multiplicative factor of $(1 - 1/e) \approx 0.63$ of the optimal. In addition, it is shown in [26] that MAX k -COVER cannot be approximated to a factor better than $(1 - 1/e)$, unless $P = NP$. This implies that the greedy algorithm provides the best possible approximation.

In place of the constraint (3), one can use a more general form of constraint known as a matroid. The constraint in (3) defines the *uniform matroid* in which S must belong to $\{A \subseteq E : |A| \leq k\}$. Another commonly used matroid constraint in optimization is the *partition matroid* defined as follows: Let E be a base set of elements, and partition E into n disjoint sets E_1, \dots, E_n and let k_1, \dots, k_n be positive integers. Then allowable sets are members of $\{A \subseteq E : |A \cap E_i| \leq k_i\}$. It has been shown [16] that maximizing a normalized and monotone submodular function subject to a matroid constraint can be approximated to within $1/2$ using the simple greedy algorithm.

III. PROBLEM STATEMENT

We begin by introducing a centralized version of the submodular maximization problem studied in this paper.

Consider the following submodular maximization problem. We are given n disjoint sets X_i , $i \in \{1, \dots, n\}$, $X = \cup_i X_i$, and a submodular monotone and normalized function $f : 2^X \rightarrow \mathbb{R}_{\geq 0}$. We wish to solve

$$\max_{S \subseteq X} f(S)$$

subject to

$$|S \cap X_i| \leq 1 \quad \text{for each } i \in \{1, \dots, n\}. \quad (4)$$

This is a submodular maximization problem over a partition matroid, and thus from Section II, the optimal solution can

be approximated to within a multiplicative factor of $1/2$ using the simple greedy algorithm [16].

A. Distributed Optimization Problem

We consider a *collaborative* submodular analogue of (4).

Problem 3.1 (Distributed submodular maximization): We are given n agents, or players $V = \{1, \dots, n\}$. Each agent has an *strategy set* X_i , and must choose one strategy (or action) $x_i \in X_i$. We define $X = \cup_i X_i$ and are given a normalized and monotone increasing submodular function $f : 2^X \rightarrow \mathbb{R}_{\geq 0}$. The goal is for the agents to each choose an $x_i \in X_i$ such that $f(x_1, \dots, x_n)$ is maximized.

We consider the case where agents choose their strategies in a sequential manner, starting with agent 1 and ending with agent n . Each agent $i \in V$ has access to a subset of the strategies that have been chosen by agents $\{1, \dots, i-1\}$. This information is encoded in a directed acyclic graph (DAG) $\mathcal{G} = (V, E)$ where all edges $(i, j) \in E$ satisfy $i < j$. Note that every DAG can be topologically sorted, and thus there always exists a labeling of the vertices of \mathcal{G} in which $i < j$ for all $(i, j) \in E$. A *complete DAG* is a DAG for which no edge can be added without creating a cycle. A *clique* in \mathcal{G} is a subgraph of \mathcal{G} that is complete. The *clique number* of a DAG \mathcal{G} is the number of vertices in its largest clique, and is denoted $\omega(\mathcal{G})$.

We define the in-neighbors of agent i as

$$\mathcal{N}(i) = \{j \in V \mid (j, i) \in E\},$$

and thus the information available to agent i when choosing its strategy is

$$X_{\text{in}}(i) = \{x_j \mid j \in \mathcal{N}(i)\}.$$

In this paper we study the performance of the greedy algorithm in which agent i chooses its strategy x_i to maximize its marginal reward relative to its limited information $X_{\text{in}}(i)$:

$$x_i = \operatorname{argmax}_{x \in X_i} \Delta(x \mid X_{\text{in}}(i)). \quad (5)$$

In the case where multiple strategies have equal marginal reward, the greedy algorithm selects one maximizer arbitrarily.

Given a set of strategies $\{x_1, \dots, x_n\}$ for the n agents, the overall objective can be written as a sum of marginal rewards as

$$f(x_1, \dots, x_n) = \sum_{i=1}^n \Delta(x_i \mid \{x_1, \dots, x_{i-1}\}), \quad (6)$$

where $\{x_1, \dots, x_0\}$ is defined to be the empty set. It is worth emphasizing that individual agents do not require access to the function f given in (6). They require only the ability to calculate their own marginal rewards using the limited information they have access to in order to compute (5). Note that while x_i that maximizes agent i 's marginal reward relative to $X_{\text{in}} \subseteq \{x_1, \dots, x_{i-1}\}$, its contribution to the overall objective value is given by the marginal reward relative to $\{x_1, \dots, x_{i-1}\}$. The main goal of this paper is to investigate how this lack of information affects the performance of the greedy algorithm.

Remark 3.2 (Topological ordering): In general, there are multiple topological orderings of a DAG $\mathcal{G} = (V, E)$, each

defining a different sequence in which the agents select their strategies. However, the strategy chosen by each agent will be identical in all valid orderings, since each agent i must select its strategy after all agents in $X_{\text{in}}(i)$. •

Example 3.3 (Sensor coverage): As an illustrative example, consider the setting shown in the left panel of Figure 1. There are five agents, and each agent $i \in \{1, \dots, 5\}$ has a strategy set X_i consisting of two elements, which in this example are disks. Each agent must choose one of its two disks, and thus as an example, agent 2 must choose one of the two grey disks labeled with the number 2. The submodular function f gives the total area covered by the union of the disks chosen by each agent. The information structure \mathcal{G} is shown in the DAG below the panel. The center panel shows the disk chosen by each agent under the greedy algorithm (5). The right panel shows the disks chosen when each agent has complete information of the choices made before it (i.e., when \mathcal{G} is a complete DAG). We can see that significantly more area is covered when the agents have complete information. •

Remark 3.4 (Synchronous agent updates): An alternative model is that each agent $i \in V$ synchronously updates its strategy based on the strategies of its in-neighbors. That is, let $x_i(t)$ be the strategy of agent i at the synchronous iteration t . Then, agent i greedily chooses $x_i(t)$ to maximize its marginal reward relative to its in-neighbors choices at $t-1$:

$$x_i(t) = \operatorname{argmax}_{x \in X_i} \Delta(x \mid X_{\text{in}}(i, t-1)),$$

where $X_{\text{in}}(i, t-1) = \{x_j(t-1) \mid j \in \mathcal{N}(i)\}$. In this case, it is straightforward to verify that agent k 's strategy is fixed after iteration k , and its choice of strategy is equal to that in the sequential setup. Thus, the performance of the synchronous update is identical to that of the sequential update. •

B. Disjoint vs. Intersecting Strategy Sets

The main results in this paper are bounds on the performance of the greedy algorithm as functions of the information topology \mathcal{G} and the number of agents n . In this section we show that without loss of generality we can restrict our attention to disjoint strategy sets. A reader can skip directly to Section IV and revisit the formal argument later.

Consider Problem 3.1 with information structure \mathcal{G} and strategy sets $X_1, \dots, X_n \subseteq X$ that are possibly intersecting. If the agents run the greedy algorithm in (5), then the solution $\bar{S} = (x_1, \dots, x_n)$ produced (written as a tuple to keep track of the choice of each agent) has two notable properties: It may contain multiple and/or repeated elements from a single strategy set. The value of \bar{S} is evaluated by defining the set S containing the unique elements of \bar{S} , i.e., $S = \{x \in X \mid \bar{S}(i) = x \text{ for some } i \in \{1, \dots, n\}\}$, and evaluating $f(S)$. As a simple example, if two agents pick the same strategy x , resulting in the solution $\bar{S} = (x, x)$, then $S = \{x\}$ and the value of the solution is $f(\{x\})$.

Now, we define a corresponding problem with the same number of agents n and information topology \mathcal{G} , but with a submodular function f' and disjoint strategy sets X'_1, \dots, X'_n in $X' = \cup_i X'_i$.

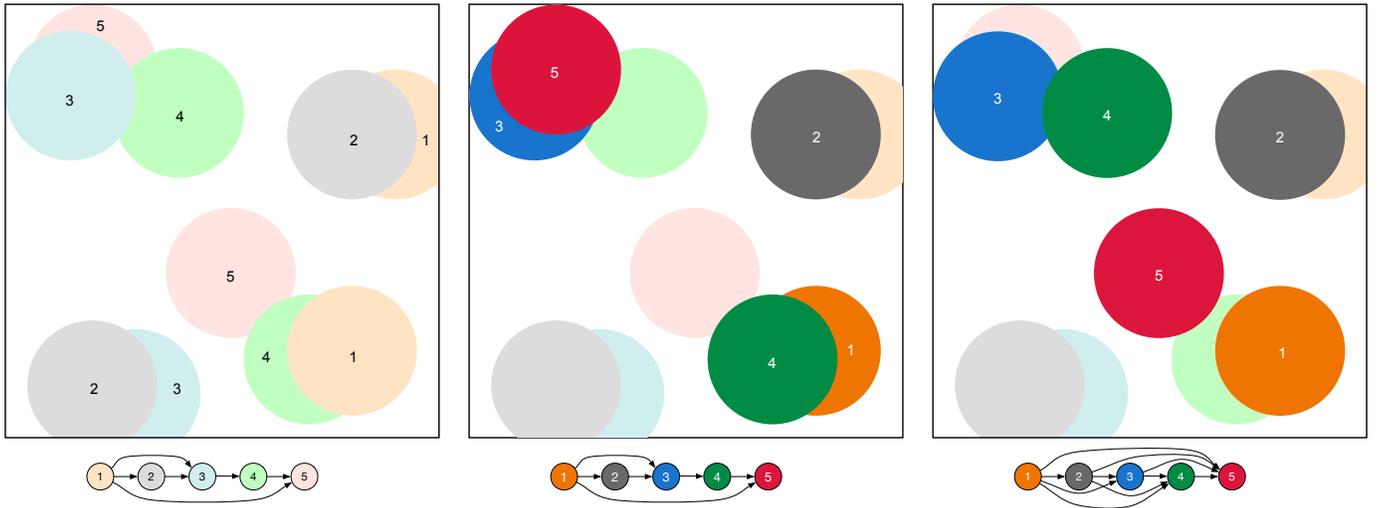


Fig. 1. Left: Five agents, each with the choice of two disks. The objective function is the total area covered by the union of disks. Center: The dark numbered disks show the choice of each agent under limited information. Right: The dark numbered disks show the choice of each agent given complete information of the choices of prior agents.

Define X'_i as a set containing a unique copy of each element in X_i . We then define a function $g : 2^{X'} \rightarrow 2^X$ that for a set of elements $A \subseteq X'$, returns the corresponding set of elements in X . Formally, for any set $A \subseteq X'$, we have

$$g(A) = \{x \in X \mid g(a) = x \text{ for some } a \in A\}.$$

Note that $g(X'_i) = X_i$. By the definition of g , we immediately see that for any $A, B \subseteq X'$,

$$g(A) \cup g(B) = g(A \cup B), \text{ and } g(A) \cap g(B) = g(A \cap B).$$

Next, we define the function $f' : 2^{X'} \rightarrow \mathbb{R}_{\geq 0}$ as $f' = f \circ g$ so that for any $A \in X'$,

$$f'(A) = f(g(A)).$$

To see that the function is submodular, note that for any $A, B \subset X'$, we have

$$\begin{aligned} f'(A) + f'(B) &= f(g(A)) + f(g(B)) \\ &\leq f(g(A) \cup g(B)) + f(g(A) \cap g(B)) \\ &= f(g(A \cup B)) + f(g(A \cap B)) \\ &= f'(A \cup B) + f'(A \cap B), \end{aligned}$$

where the inequality holds by submodularity of f . If f is monotone and normalized, then so is f' .

Suppose the greedy algorithm (5) constructs a solution $S = \{x_1, \dots, x_n\}$ on the non-disjoint instance f . Consider any $S' = \{x'_1, \dots, x'_n\}$ such that $x'_i \in X'_i$ and $g(x'_i) = x_i$. What remains is to show that S' is a solution of the greedy algorithm (5) when run on the disjoint instance f' . Also, since $g(S') = S$ the solutions have equal value, $f'(S') = f(S)$.

By recalling the definition of $X_{\text{in}}(i)$ and defining $X'_{\text{in}}(i) = \{x'_j \mid j \in \mathcal{N}(i)\}$, we have

$$X_{\text{in}}(i) = g(X'_{\text{in}}(i)).$$

Now, we need to verify that x'_i satisfies

$$x'_i = \operatorname{argmax}_{x' \in X'_i} \Delta'(x' \mid X'_{\text{in}}(i)), \quad (7)$$

where Δ' is the marginal reward for f' . From the definition that $f'(A) = f(g(A))$, we have that

$$\Delta'(x'_i \mid X'_{\text{in}}(i)) = \Delta(x_i \mid X_{\text{in}}(i)).$$

Suppose now, by way of contradiction, that x'_i is not a maximizer of (7) and there exists a $y' \in X'_i$ such that

$$\Delta'(y' \mid X'_{\text{in}}(i)) > \Delta'(x'_i \mid X'_{\text{in}}(i)).$$

This would imply that $\Delta(g(y') \mid X_{\text{in}}(i)) > \Delta(x_i \mid X_{\text{in}}(i))$, and since $g(y') \in X_i$, we conclude that x_i is not a valid choice of the greedy algorithm on f , which is a contradiction. Thus, for every intersecting instance, there is a corresponding disjoint instance on the same information structure, with the same optimal solution value, and such that the greedy algorithm produces solutions with equal value.

IV. THE SEQUENTIAL DISTRIBUTED GREEDY ALGORITHM: LOWER BOUND

Throughout this section, we assume that the strategy sets are disjoint. The agents take their decisions sequentially in increasing order according to their index. We start with a scenario where the agents *do not observe* the strategies of the agents that have taken action prior to them (see Figure 2(a)). The decision of agent $i \in V$ is then

$$x_i = \operatorname{argmax}_{x \in X_i} \Delta(x \mid \emptyset).$$

Suppose that $\{x_1^*, x_2^*, \dots, x_n^*\}$ be the solution of (4), where $x_i^* \in X_i$ for each $i \in \{1, \dots, n\}$. We have that

$$f(x_1, \dots, x_n) \geq f(x_i) \geq f(x_i^*),$$

for all $i \in \{1, \dots, n\}$, and hence

$$f(x_1, \dots, x_n) \geq \frac{1}{n} \sum_{i=1}^n f(x_i^*) \geq \frac{1}{n} f(x_1^*, \dots, x_n^*),$$

where the last inequality follows from submodularity of f given by (1). It is easy to observe that this lower bound is

tight and cannot be improved. A natural problem is hence to investigate if this lower bound on the performance can be improved when the agents can observe (perhaps partially) the decisions of the preceding agents before making decisions. Consider the scenario where agent i observes the decision of all agents in the set $\mathcal{I}_i = \{1, \dots, i-1\}$. The information that each agent has access to before taking its decision is best represented by Figure 2(b). In this case, the decision for agent i from (5) becomes

$$x_i = \operatorname{argmax}_{x \in X_i} \Delta(x | \{x_1, \dots, x_{i-1}\}). \quad (8)$$

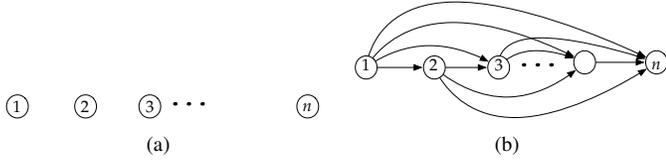


Fig. 2. (a) shows a scenario with no observation and (b) shows a scenario where each agents observes the decisions of all the preceding agents.

The performance of the strategy proposed by (8) can be deduced from the so-called “local greedy algorithms”, which is studied in the classical paper [17, Theorem 4.1]. The proof presented in [17] relies on a clever use of linear programming. Here we present an independent proof that allows us to extend this result to more general information topologies. We start with the following result.

Lemma 4.1: Consider Problem 3.1 and let $(x_1^*, x_2^*, \dots, x_n^*)$ be an optimal solution. Suppose that players sequentially update their strategies (x_1, \dots, x_n) according to (8). Then, for all $1 \leq k \leq n$, we have that

$$f(x_1, \dots, x_k) \geq f(x_1^*, \dots, x_k^*) - f(x_1, \dots, x_{k-1}). \quad (9)$$

Proof: We prove the following slightly stronger result,

$$f(x_1, \dots, x_{k+1}) \geq f(x_1, \dots, x_k, x_1^*, \dots, x_{k+1}^*) - f(x_1, \dots, x_k).$$

Note that the result in (9) immediately follows after proving this, since

$$f(x_1, \dots, x_k, x_1^*, \dots, x_{k+1}^*) \geq f(x_1^*, \dots, x_{k+1}^*),$$

by monotonicity. The proof will proceed by induction. Throughout the proof, we use the “greedy choice property”, i.e., the fact that

$$\Delta(x_{k+1} | x_1, \dots, x_k) \geq \Delta(x_{k+1}^* | x_1, \dots, x_k).$$

First, consider the base case of $k = 0$. By the greedy choice property, we have that

$$f(x_1) \geq f(x_1^*),$$

and the base case holds. Next, suppose the result holds for

k and we show it holds for $k+1$. We have

$$\begin{aligned} f(x_1, \dots, x_{k+1}) &= f(x_1, \dots, x_k) + \Delta(x_{k+1} | x_1, \dots, x_k) \\ &\geq f(x_1, \dots, x_{k-1}, x_1^*, \dots, x_k^*) - f(x_1, \dots, x_{k-1}) \\ &\quad + \Delta(x_{k+1}^* | x_1, \dots, x_k) \\ &= f(x_1, \dots, x_{k-1}, x_1^*, \dots, x_k^*) - f(x_1, \dots, x_{k-1}) \\ &\quad + f(x_1, \dots, x_k, x_{k+1}^*) - f(x_1, \dots, x_k), \end{aligned} \quad (10)$$

where in the first inequality we applied the inductive hypothesis and the greedy choice property. Next, using submodularity, we have

$$\begin{aligned} f(x_1, \dots, x_{k-1}, x_1^*, \dots, x_k^*) + f(x_1, \dots, x_k, x_{k+1}^*) \\ \geq f(x_1, \dots, x_k, x_1^*, \dots, x_{k+1}^*) + f(x_1, \dots, x_{k-1}). \end{aligned} \quad (11)$$

Substituting (11) into the right hand side of (10) and cancelling the common $f(x_1, \dots, x_{k-1})$ term we get

$$\begin{aligned} f(x_1, \dots, x_{k+1}) &\geq f(x_1, \dots, x_k, x_1^*, \dots, x_{k+1}^*) \\ &\quad - f(x_1, \dots, x_k), \end{aligned}$$

which proves the result. \blacksquare

Theorem 4.2: Consider Problem 3.1 and let $(x_1^*, x_2^*, \dots, x_n^*)$ be an optimal solution. Suppose that players sequentially update their strategies (x_1, \dots, x_n) according to (8). Then

$$f(x_1, \dots, x_n) \geq \frac{1}{2} f(x_1^*, \dots, x_n^*).$$

Proof: Using Lemma 4.1, we have

$$f(x_1, \dots, x_n) \geq f(x_1^*, \dots, x_n^*) - f(x_1, \dots, x_{n-1}).$$

Also, by monotonicity, we have $f(x_1, \dots, x_n) \geq f(x_1, \dots, x_{n-1})$, and hence the result follows. \blacksquare

The following consequence of this result holds for any general directed acyclic graph.

Corollary 4.3: Consider Problem 3.1 and let $(x_1^*, x_2^*, \dots, x_n^*)$ be an optimal solution. Suppose that players sequentially update their strategies (x_1, \dots, x_n) according to (5) on a directed acyclic graph \mathcal{G} . Then,

$$f(x_1, \dots, x_n) \geq \frac{1}{(n - \omega(\mathcal{G})) + 2} f(x_1^*, \dots, x_n^*),$$

where $\omega(\mathcal{G})$ is the clique number of \mathcal{G} .

Proof: Let us partition the nodes in \mathcal{G} into the set of nodes in a maximum clique, whose strategy sets are denoted by S_{clique} with its complement \bar{S}_{clique} . Consider now the smaller optimization problem over S_{clique} . By Theorem 4.2,

$$f(S_{\text{clique}}) \geq \frac{1}{2} f(S_{\text{clique}}^*),$$

where we have again used the star notation to denote the optimal solution. On the other hand, the remaining $n - \omega(\mathcal{G})$ agents cannot have a performance worst than $\frac{1}{n - \omega(\mathcal{G})}$ of their optimal, i.e.,

$$f(\bar{S}_{\text{clique}}) \geq \frac{1}{n - \omega(\mathcal{G})} f(\bar{S}_{\text{clique}}^*).$$

Using these two inequalities, we conclude that

$$\begin{aligned} f(S_{\text{clique}}^*) + f(\bar{S}_{\text{clique}}^*) &\leq 2f(S_{\text{clique}}) + (n - \omega(\mathcal{G}))f(\bar{S}_{\text{clique}}) \\ &\leq (2 + (n - \omega(\mathcal{G})))f(S), \end{aligned}$$

where $S = S_{\text{clique}} \cup \bar{S}_{\text{clique}}$, and the last inequality follows by monotonicity of f . Using the submodularity of f , we have that $f(S^*) \leq f(S_{\text{clique}}^*) + f(\bar{S}_{\text{clique}}^*)$ and hence

$$f(S^*) \leq (2 + (n - \omega(\mathcal{G})))f(S),$$

concluding the proof. \blacksquare

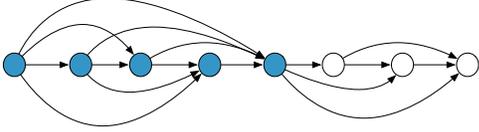


Fig. 3. A DAG with 8 nodes is depicted, where $\omega(\mathcal{G}) = 5$. By Corollary 4.3, we conclude that under (5), we have $f(x_1, \dots, x_n) > \frac{1}{6}f(x_1^*, \dots, x_n^*)$.

Figure 3 shows an example where Corollary 4.3 is used to conclude the performance of (5). Note that if we take a complete DAG and delete a single edge, its clique number reduces from n to $n - 1$, and from Corollary 4.3, our lower bound decreases from $1/2$ to $1/3$. The reason for this is that we have not made any assumptions on the relative contribution of each player to the total reward. The one agent that is removed from the clique may contribute more to the total reward than all $n - 1$ other agents combined.

A. Interconnected Cliques of Full Information

It is not clear how to characterize the impact of the availability or lack of information on the performance of local greedy algorithms for arbitrary graph topologies. In this subsection, however, we extend Theorem 4.2 to a scenario with multiple cliques, where the agent's in each clique have access to the decision of the last agent in the last clique that takes decision prior to them.

Theorem 4.4: Consider Problem 3.1 and let $(x_1^*, x_2^*, \dots, x_n^*)$ be an optimal solution. Suppose that agents are partitioned into $\kappa \in \mathbb{Z}_{\geq 1}$ cliques as

$$\{1, \dots, m_1\}, \{m_1 + 1, \dots, m_2\}, \dots, \{m_{\kappa-1} + 1, \dots, n\},$$

with the only information available from other cliques to the agent's in $(i + 1)$'s clique being the choice of agent m_i , $i \in \{1, \dots, \kappa - 1\}$. Suppose that players sequentially update their strategies (x_1, \dots, x_n) according to (5). Then

$$f(x_1, \dots, x_n) \geq \frac{1}{2\kappa} \left(f(x_1^*, \dots, x_n^*) + \sum_{i=1}^{\kappa-1} f(x_{m_i}^*) \right).$$

Proof: We proceed by induction on κ . The base case, where $\kappa = 1$, follows from Theorem 4.2. Suppose now that the statement holds for $\kappa - 1$. We prove that it also holds for κ . For notational convenience, let us define $m_\kappa := n$. First, by Theorem 4.2,

$$f(x_{m_{\kappa-1}}, \dots, x_{m_\kappa}) \geq \frac{1}{2} f(\tilde{x}_{m_{\kappa-1}}^*, \dots, \tilde{x}_{m_\kappa}^*),$$

where $(\tilde{x}_{m_{\kappa-1}}^*, \dots, \tilde{x}_{m_\kappa}^*)$ is the optimal strategy for clique κ . Hence, we also have that

$$f(x_{m_{\kappa-1}}, \dots, x_{m_\kappa}) \geq \frac{1}{2} f(x_{m_{\kappa-1}}^*, \dots, x_{m_\kappa}^*),$$

where $(x_{m_{\kappa-1}}^*, \dots, x_{m_\kappa}^*)$ is the strategy for clique κ in the true optimal. By using monotonicity, we have that

$$2f(x_1, \dots, x_n) \geq f(x_{m_{\kappa-1}}^*, \dots, x_{m_\kappa}^*).$$

On the other hand, by the induction assumption, along with monotonicity, we have that

$$2(\kappa - 1)f(x_1, \dots, x_{m_{\kappa-1}}) \geq f(x_1^*, \dots, x_{m_{\kappa-1}}^*) + \sum_{i=1}^{\kappa-2} f(x_{m_i}^*).$$

Summing the last two inequalities, and noting that by monotonicity $f(x_1, \dots, x_n) \geq f(x_1, \dots, x_{m_{\kappa-1}})$ since $n = m_\kappa$, we have that

$$\begin{aligned} 2\kappa f(x_1, \dots, x_n) &\geq f(x_1^*, \dots, x_{m_{\kappa-1}}^*) \\ &+ f(x_{m_{\kappa-1}}^*, \dots, x_{m_\kappa}^*) + \sum_{i=1}^{\kappa-2} f(x_{m_i}^*). \end{aligned}$$

Using submodularity, we have that

$$\begin{aligned} 2\kappa f(x_1, \dots, x_n) &\geq f(x_1^*, \dots, x_{m_\kappa}^*) \\ &+ f(x_{m_{\kappa-1}}^*) + \sum_{i=1}^{\kappa-2} f(x_{m_i}^*) \\ &= f(x_1^*, \dots, x_n^*) + \sum_{i=1}^{\kappa-1} f(x_{m_i}^*), \end{aligned}$$

concluding the proof. \blacksquare

As noted in Remark 3.4, the results of this section are extend to the case of synchronous updates. As a final observation, consider the setting of Theorem 4.4 and suppose that there is no communications between agents in the cliques. Then, since

$$\begin{aligned} f(x_1, \dots, x_{m_1}) &\geq \frac{1}{2} f(x_1^*, \dots, x_{m_1}^*), \quad \text{and} \\ f(x_{m_i+1}, \dots, x_{m_{i+1}}) &\geq \frac{1}{2} f(x_{m_i+1}^*, \dots, x_{m_{i+1}}^*), \end{aligned}$$

for all $i \in \{1, \dots, \kappa - 1\}$, using submodularity and monotonicity, we conclude that $f(x_1, \dots, x_n) \geq \frac{1}{2\kappa} f(x_1^*, \dots, x_n^*)$. In this sense, Theorem 4.4 captures a scenario where the overall performance can only be enhanced by having access to more information. In particular, the performance is guaranteed to improve by at least

$$\frac{1}{2\kappa} \sum_{i=1}^{\kappa-1} f(x_{m_i}^*) \geq \frac{1}{2\kappa} f(x_{m_1}^*, \dots, x_{m_{\kappa-1}}^*) \geq 0,$$

where the first inequality holds by submodularity.

V. UPPER BOUNDS USING COLOURING

Our main objective in this section is to determine upper bounds on the performance of the sequential greedy algorithm. We proceed with studying two cases: 1) we construct a graph dependent submodular function that exploits information on graph topology to yield an upper bound; and 2) we design a global submodular function that provides bounds on all graph topologies. We will see that these two cases are essentially equivalent to computing an optimal and a greedy coloring [27] of the graph, respectively. The second case also allows us to extract graph properties that limit performance.

A. Background on Graph Coloring

Given a graph $G = (V, E)$ with $|V| = n$, a *coloring* is a function $c : V \rightarrow \{1, \dots, n\}$ such that for every $(u, v) \in V$, we have $c(u) \neq c(v)$. The *vertex-coloring problem* is to find a coloring $c : V \rightarrow \{1, \dots, k\}$ such that k is minimized. That is, the goal is to color the vertices of the graph with the minimum number of colors such that no pair of adjacent vertices have the same color. The optimum value of k is called the *chromatic number* of the graph G and is often denoted by $\chi(G)$. The problems of computing a minimum coloring or equivalently of determining the chromatic number of a graph are NP-hard [27].

B. Upper Bound via Adversarial Choice of Function

Consider the following problem: Given a graph $\mathcal{G} = (V, E)$, can we construct a submodular function $f_{\mathcal{G}}$, exploiting information of the graph for which we can upper bound the performance of greedy algorithm? We call such a function *adversarial* since the function can be designed to exploit the weaknesses of the information graph. Given a graph $\mathcal{G} = (V, E)$, we construct strategy sets and a submodular function as follows. Let $c : V \rightarrow \{1, \dots, k\}$ be an optimal coloring of \mathcal{G} and where $\chi(\mathcal{G}) = k$ is the chromatic number of \mathcal{G} . Given a coloring, we define the set of vertices of color $\ell \leq k$ as

$$V_{\ell} = \{v \in V \mid c(v) = \ell\}.$$

We use this coloring to construct strategy sets X_i for each agent $i \in V$ and the submodular function $f : 2^X \rightarrow \mathbb{R}_{\geq 0}$. For each agent i , we define $X_i = \{a_i, b_i\}$, and let $X = \cup_i X_i$. We define the function f through its marginal rewards. To that end, let $S \subset X$ be a set such that $|S \cap X_i| \leq 1$ containing decision for a subset of the n agents in V . Then, the marginal reward of a_i is defined as

$$\Delta(a_i|S) = \begin{cases} 0 & \text{if } a_j \in S \text{ for some } j \in V_{c(i)}, \\ 1 & \text{otherwise.} \end{cases}$$

The marginal reward of b_i is defined is $\Delta(b_i|S) = 1$, for all $i \in V$. Then, given a set $S = \{s_1, \dots, s_m\} \subset X$ with $|S \cap X_i| \leq 1$, for all $i \in V$, we define

$$f(S) = \sum_{i=1}^m \Delta(s_i|\{s_1, \dots, s_{i-1}\}). \quad (12)$$

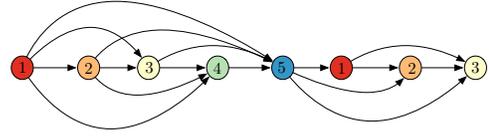


Fig. 4. A minimum coloring of the graph of Figure 3 using five colors. Using the lower bounds on the performance and Proposition 5.1, we can conclude that $\frac{1}{6}f(x_1^*, \dots, x_n^*) < f(x_1, \dots, x_n) < \frac{5}{8}f(x_1^*, \dots, x_n^*)$.

Note that we can equivalently define this function over non-disjoint strategy sets. Let $X = \{a, b_1, \dots, b_n\}$ and define $X_i = \{a, b_i\}$. Then given a set $S \subset X$, we define $f(S) = |S|$. Notice that this function is modular. Given a tuple containing agent strategies (x_1, \dots, x_n) , we define the set of unique strategies as $S = \{x_1, \dots, x_n\}$ and evaluate $f(S)$ as the reward. This definition is equivalent to the disjoint definition above as discussed in Section III-B.

Using the function (12), we obtain the following result.

Proposition 5.1: Consider Problem 3.1, where f is given by (12), and let $(x_1^*, x_2^*, \dots, x_n^*)$ be an optimal solution. Suppose that players sequentially update their strategies (x_1, \dots, x_n) according to (5). Then,

$$f(x_1, \dots, x_n) \leq \frac{\chi(G)}{n} f(x_1^*, \dots, x_n^*).$$

Proof: An optimal solution for maximizing (12) is $x_i^* = b_i$ for each $i \in V$, which yields a reward of $f(x_1^*, \dots, x_n^*) = n$. Now, considering the greedy algorithm in (5), we see that for each agent $i \in V$, we have $\mathcal{N}(i) \cap V_{c(i)} = \emptyset$, and thus no agent in $V_{c(i)}$ is visible to agent i . Thus, for each agent i we have

$$\operatorname{argmax}_{x \in X_i} \Delta(x \mid X_{\text{in}}(i)) = \{a_i, b_i\},$$

both of which achieve a marginal reward of 1. A possible evolution of the greedy algorithm is $x_i = a_i$ for each $i \in V$. The value of this solution is

$$f(a_1, \dots, a_n) = \sum_{i=1}^n \Delta(a_i \mid a_1, \dots, a_{i-1}) = \chi(\mathcal{G})$$

since only the first agent of each color contributes 1 to the total reward, and all others contribute zero. ■

Figure 4 depicts the DAG used in Figure 3 along with an optimal coloring yielding a chromatic number of 5. Using Proposition 5.1, we conclude that under (5), we have $f(x_1, \dots, x_n) < \frac{5}{8}f(x_1^*, \dots, x_n^*)$.

C. Upper Bound via Universal Function

The upper bound in the previous section required a submodular function f that depended on the graph. In the following, we propose a single function, independent of the graph topology, and analyze its performance for all graphs. This allows us to state a few simple properties of graphs that limit its performance.

The base set of the function is $X = \{e_1, \dots, e_m\}$ where $m \geq n$ and we let $X_i = X$ for each agent i . Given a choice

$x_i \in X_i$ for each agent i , the value of the submodular function is

$$f(x_1, \dots, x_n) = |\cup_i \{x_i\}|. \quad (13)$$

That is, the value is given by the number of unique elements of X chosen by the agents. Clearly, an optimal solution is any one in which each agent chooses a different element from X , yielding a value of n . Under the greedy algorithm, agent i will compute its marginal reward as

$$\Delta(e_j | X_{\text{in}}(i)) = \begin{cases} 1 & \text{if } e_j \notin X_{\text{in}}(i) \\ 0 & \text{if } e_j \in X_{\text{in}}(i), \end{cases}$$

and thus agent i will choose any strategy e_j such that $e_j \notin X_{\text{in}}(i)$. Suppose that each agent i breaks ties by choosing the strategy with lowest index. Then, we can write the greedy choice for agent i as

$$x_i = \min\{e_j \in X \mid e_j \notin X_{\text{in}}(i)\}. \quad (14)$$

Next we relate the performance of the greedy algorithm on this function to properties of the underlying graph.

Proposition 5.2: Consider the submodular function in (13) and any graph $\mathcal{G} = (V, E)$. If the greedy algorithm finds a solution within k/n of the optimal for some $k > 0$ then each of the following properties hold:

- (i) there is a vertex in \mathcal{G} with in-degree of at least $k - 1$;
- (ii) for each $\ell \in \{1, \dots, k\}$ there are at least ℓ agents with in-degree of at least $k - \ell$;
- (iii) the number of edges in \mathcal{G} is at least $k(k - 1)/2$.

Proof: By Equation (14), if $x_i = e_j$, then $\{e_1, \dots, e_{j-1}\} \subseteq X_{\text{in}}(i)$. From this we immediately see that for an agent to choose e_k , it must have an in-degree of at least $k - 1$, proving property (i). Property (ii) also follows by noticing that each agent that chooses a strategy with index $\geq j$ must have in-degree of at least $j - 1$. To achieve a value of k , each strategy from e_1 to e_k must be chosen by an agent.

To see property (iii), we notice that the sum of in-degrees of all agents must be at least $\sum_{j=1}^k j = k(k - 1)/2$. ■

Notice that the sequence of strategy choices in (14) provides a simple and efficient algorithm for computing a performance upper bound for a given graph. For completeness, we give the details in Algorithm 1.

Algorithm 1 Greedy Algorithm Upper Bound

Input: A directed acyclic graph $\mathcal{G} = (V, E)$.

Output: An upper bound on the approximation ratio of greedy algorithm on \mathcal{G} .

- 1: Topologically sort the vertices V
 - 2: **for** each $v \in V$:
 - 3: Set $\text{value}[v] = 1$
 - 4: **for** each $v \in V$ in topological order :
 - 5: Set $\text{value}[v]$ to smallest integer k such that for each $u \in \mathcal{N}(i)$, $\text{value}[u] \neq k$.
 - 6: **return** $\frac{1}{|V|} \max_{v \in V} \text{value}[v]$
-

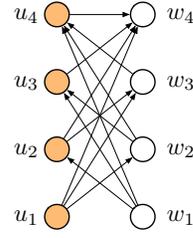


Fig. 5. A bipartite graph on eight vertices. The graph is colored using just two colors, and its chromatic number is two. Under (14), vertex w_4 will select strategy e_5 , and thus Algorithm 1 will output $5/8$.

Complexity of Algorithm 1: The complexity of Algorithm 1 is $O(|V| + |E|)$. The topological sort can be performed in $O(|V| + |E|)$ time. The only detail is to implement line 5: in linear time, which essentially computes the smallest element not in an array. This can be done using two passes through the array of in-neighbor values. In the first pass, we populate a Boolean array of length $|\mathcal{N}(i)|$. All entries of the array are initialized to `false`, and the j th entry is then reset to `true` if and only if there is a vertex $u \in \mathcal{N}(i)$ with $\text{value}[u] = j$. In the second pass, we scan the Boolean array for the first false entry. This is the smallest value that is not chosen by an in-neighbor. Thus, the total computation time for the for-loop in line 4: is $O(|E|)$.

D. Gap Between Adversarial and Universal Upper Bounds

The strategies chosen by each agent for the submodular function in (13) provides a coloring of the graph \mathcal{G} . That is, the vertices $V_\ell := \{i \in V \mid x_i = e_\ell\}$ are those colored with color ℓ . By construction, there are no edges between vertices of the same color. This implies that the adversarial upper bound is tighter than the universal upper bound, which utilizes the minimum number of colors.

A key question is how large the gap can be between the two upper bounds. In general it can be arbitrarily large. To see this, consider the following bipartite directed graph $\mathcal{G} = (V, E)$ consisting of $n = 2m$ vertices, where

$$V = \{u_1, \dots, u_m\} \cup \{w_1, \dots, w_m\}.$$

The graph contains all edges (u_i, w_j) such that $i < j$ and all edges (w_i, u_j) such that $i < j$ along with the edge (u_m, w_m) . An example is shown in Figure 5. We fix the topological ordering of the vertices to be $u_1, w_1, u_2, w_2, \dots, u_m, w_m$. For this graph, under Algorithm 1, the vertex w_m chooses strategy e_{m+1} . Thus

- (i) the chromatic number of \mathcal{G} is 2, and thus the adversarial upper bound is $2/n$; and
- (ii) Algorithm 1 returns $1/2 + 1/(2m)$.

However, the advantage of Algorithm 1 is that it is a simple linear-time algorithm for computing an upper bound on achievable performance of a given graph topology. In contrast, the adversarial function of Section V-B provides a much tighter upper bound, but requires solving an NP-hard problem. An interesting connection to note is that Algorithm 1 is essentially

computing a greedy coloring of \mathcal{G} , where vertices are considered in topological order. It is known that in general a greedy coloring does not provide a constant factor approximation to the minimum coloring [27].

Remark 5.3 (Gap between bounds for random graphs):

Given a probability $p \in [0, 1]$ and a set V of n vertices, the Erdős-Rényi graph $\mathcal{G}(n, p)$ on V is a random graph in which each edge (i, j) , where $i, j \in V$ and $i \neq j$ is included with probability p . From [28], the chromatic number of almost every $\mathcal{G}(n, p)$ is

$$\left(\frac{1}{2} + o(1)\right) \log(1/(1-p)) \frac{n}{\log n}.$$

In addition, for almost every graph $\mathcal{G}(n, p)$, the greedy coloring uses at most twice as many colors [28]. Consider now random directed acyclic graphs generated as follows: we first generate an Erdős-Rényi graph, and then we use any correct method to assign directions to each edge such that the resulting directed graph becomes acyclic. Then, using the result above on the chromatic number of the underlying Erdős-Rényi graph, Algorithm 1 provides

$$\left(\frac{1}{2} + o(1)\right) \log(1/(1-p)) \frac{1}{\log n}$$

as an upper bound on the performance. In addition, for almost every graph $\mathcal{G}(n, p)$, the greedy coloring uses at most twice as many colors [28], implying that for such graphs the gap between the two upper bounds is at most two. •

E. Comparison of Lower and Upper Bounds

In Section IV, we provided lower bounds on the performance of greedy algorithm (8) for any monotone, normalized, submodular function. This consisted of results for several graph topologies, and a general result based on the clique number of the graph. Table I compares these lower bounds with the upper bounds obtained from Proposition 5.1 and Algorithm 1. Given a graph topology, the lower bound provides a minimum performance guarantee for all submodular functions. In contrast, the upper bounds provide limitations on performance for a specific (worst-case) submodular function.

Graph	Lower	$\chi(\mathcal{G})$ Upper	Alg. 1 Upper
Empty	$1/n$	$1/n$	$1/n$
Complete acyclic	$1/2$	1	1
Interconnected cliques	$1/(2\kappa)$	$1/\kappa + 1/n$	$1/\kappa + 1/n$
General DAG	$\frac{1}{n - \omega(\mathcal{G}) + 2}$	$\chi(\mathcal{G})/n$	Alg. 1

TABLE I

COMPARISON BETWEEN THE LOWER BOUND IN SECTION IV AND THE TWO UPPER BOUNDS FOR THREE GRAPH TOPOLOGIES.

The entries in Table I for the interconnected cliques graph assume κ cliques, each of size n/κ .

VI. SIMULATION RESULTS

In this section we explore the relation between the lower bound, the two upper bounds, and the true performance of the greedy algorithm on several types of randomly generated graphs. The classes of random graphs chosen here are Erdős-Rényi, Barabási-Albert preferential attachment, and Watts-Strogatz model graphs, which are commonly used to model social networks [29], as well as physical networks [30]. Given a graph \mathcal{G} , we compute the chromatic number $\chi(\mathcal{G})$ and the clique number $\omega(\mathcal{G})$ using their standard integer programming formulations [31], and solve each integer program using Gurobi [32].

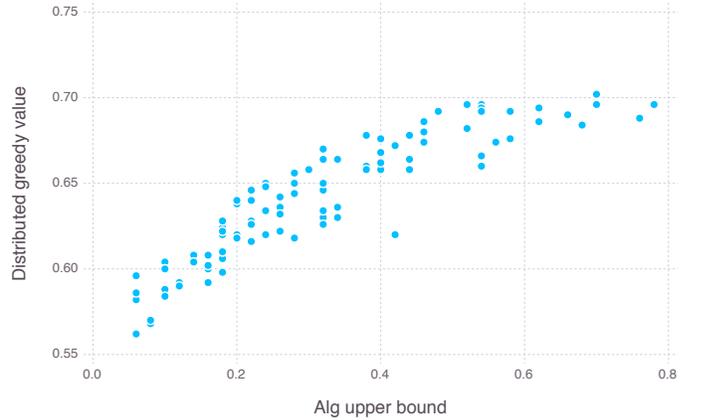


Fig. 6. Comparison between the upper bound from Algorithm 1 and the performance on the distributed greedy algorithm over 100 random graphs. The value of the upper bound correlates strongly with the true performance of the distributed greedy algorithm (Spearman's rank correlation coefficient of 0.92).

First, we explored the relation between the true performance of the distributed greedy algorithm and the upper bound given by Algorithm 1. We generated a submodular function using the disk coverage problem illustrated in Figure 1 with 50 agents. Each agent has the choice of three disks each of radius $r = 0.07$, uniformly randomly distributed in the unit square. We then generated 100 random graphs describing the information structure between the agents. Each graph is generated as follows. We select a probability p uniformly randomly from $[0, 1]$, generate a directed Erdos-Renyi graph with probability p , fix a random ordering of the vertices, and delete all edges that are directed to an earlier vertex in the ordering, creating a random DAG. Figure 6 displays the results on these 100 graphs. The upper bound, which is an approximation ratio relative to the global optimal, and the value obtained by the distributed greedy algorithm, which is the fraction of the unit square covered by the selected disks, are highly correlated.

Next, Figure 7 compares the lower bound to the two upper bounds for graphs generated according to the Barabási-Albert preferential attachment model [33] a function of the number of vertices in the graph (i.e., the number of agents in the network). Each graph is created by starting with a complete graph on five vertices, and then adding vertices using preferential attachment. An acyclic graph is then created by

generating a random ordering of the vertices and deleting cyclic edges as described above. Each data point shows the mean of 20 preferential attachment graphs, where error bars give the standard deviation. The figure gives insight into the

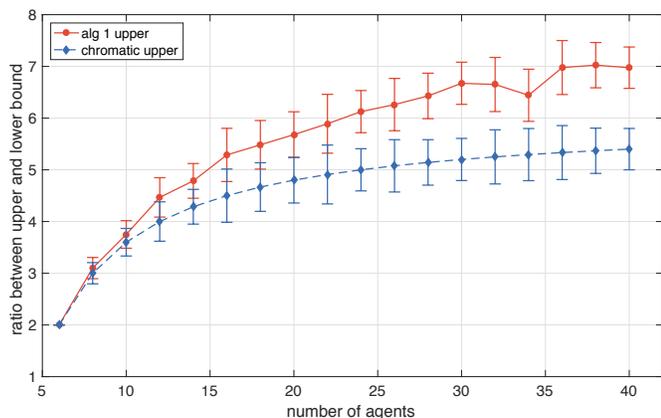


Fig. 7. Ratio between upper and lower bounds as a function of the number of agents for networks formed via preferential attachment. Each graph is grown from a complete directed graph on 5 nodes, and each new node is connected to five neighbors using the preferential attachment rule. Each data point shows the mean of 20 trials, and error bars give the standard deviation.

gap between lower and upper bounds, and the growth of this gap with the number of agents. Note, this plot shows the dependence of the bounds on the information structure, and holds for all functions f .

Finally, Figure 8 compares the lower bound to the two upper bounds for small-world graphs created according to the Watts-Strogatz model [34]. In this experiment, each graph begins as a ring lattice over 25 vertices, where each vertex is connected to K neighbors on each side. The value of K is plotted on the horizontal axis in Figure 8. Then, each edge is rewired with a probability $\beta = 0.25$, by shifting one of its endpoints to a new vertex at random, while avoiding self-loops and duplication of edges. A DAG is then created by fixing a random ordering of the vertices, and removing acyclic edges, as described above.

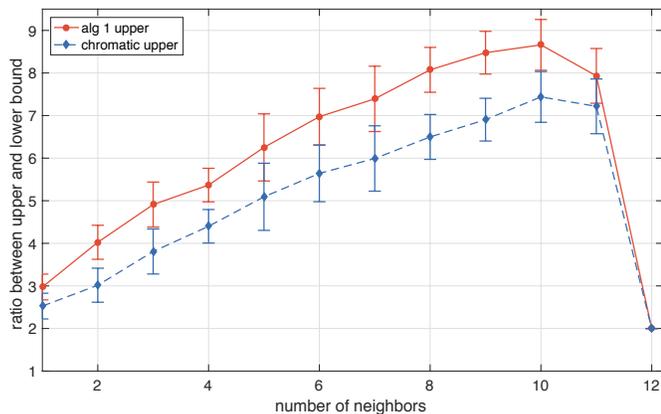


Fig. 8. Ratio between upper and lower bounds as a function of the number of neighbors for each agent prior to the rewiring operation. Each graph contains 25 agents, and the rewiring probability $\beta = 0.25$. Each data point shows the mean of 20 trials, and error bars give the standard deviation.

Each data point shows the mean of 20 small-world graphs, where error bars give the standard deviation. As the number of neighbors K increases, the graphs become more connected. When $K = 12$, each vertex is connected to all others, and thus the resulting DAG is complete. Of note in this figure is the relatively small gap between the two upper bounds, which gives support to the use of Algorithm 1 as an efficient approximation to the chromatic number bound. Also of note is the large change in the ratio of the bounds as K is increased from 11 to 12. Recall that for a complete DAG (i.e., $K = 12$), the lower bound is $1/2$, while the upper bounds are both 1. On the other hand, as discussed immediately after Corollary 4.3, removing even just a single edge from a complete DAG has a large impact on the lower bound, which is observed in the large change in the ratio.

VII. CONCLUSIONS AND FUTURE WORK

We have studied a class of submodular maximization problems under matroid constraint, where the strategy set is partitioned into private strategy sets, assigned to a group of agents; in this sense, the maximization task is distributed among them. The agents take actions sequentially and have limited information about the actions taken by agents prior to them. We have investigated the limitations that the lack of information about the actions of other agents can impose by characterizing upper and lower bounds on the performance of local greedy algorithms. Our lower bound depend on the clique number of a graph that captures the information structure, whereas our upper bounds are given in terms of the chromatic number of the underlying graph. Studying scenarios where agents are allowed to modify their choices in multiple rounds, obtaining tighter upper bounds for more homogeneous classes of submodular functions, investigating under what additional assumptions on the submodular functions one can generalize the lower bounds provided, and exploring other applications of distributed submodular optimization are interesting avenues of future research.

ACKNOWLEDGMENTS

The authors wish to thank Mohd. Shabbir Ali and Professor Jason Marden for their useful comments on an earlier version of this paper. We also thank the associate editor and the reviewers for their helpful comments.

REFERENCES

- [1] B. Ghahsifard and S. L. Smith, "On distributed submodular maximization with limited information," in *American Control Conference*, pp. 1048–1053, July 2016.
- [2] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 803–812, 1986.
- [3] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *Symposium on Information Processing of Sensor Networks*, (Berkeley, CA), pp. 20–27, Apr. 2004.
- [4] L. Xiao and S. Boyd, "Optimal scaling of a gradient method for distributed resource allocation," *Journal of Optimization Theory & Applications*, vol. 129, no. 3, pp. 469–488, 2006.

- [5] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.
- [6] P. Wan and M. D. Lemmon, "Event-triggered distributed optimization in sensor networks," in *Symposium on Information Processing of Sensor Networks*, (San Francisco, CA), pp. 49–60, 2009.
- [7] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.
- [8] B. Johansson, M. Rabi, and M. Johansson, "A randomized incremental subgradient method for distributed optimization in networked systems," *SIAM Journal on Control and Optimization*, vol. 20, no. 3, pp. 1157–1170, 2009.
- [9] B. Ghahsifard and J. Cortés, "Distributed continuous-time convex optimization on weight-balanced digraphs," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 781–786, 2014.
- [10] A. Nedic and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 601–615, 2015.
- [11] J. R. Marden, "The role of information in distributed resource allocation," *IEEE Transactions on Control of Network Systems*, 2016. To appear.
- [12] S. T. Jawaid and S. L. Smith, "Submodularity and greedy algorithms in sensor scheduling for linear dynamical systems," *Automatica*, vol. 61, pp. 282–288, 2015.
- [13] A. A. Ageev and M. I. Sviridenko, "Pipage rounding: A new method of constructing algorithms with proven performance guarantee," *Journal of Combinatorial Optimization*, vol. 8, no. 3, pp. 307–328, 2004.
- [14] S. T. Jawaid and S. L. Smith, "Informative path planning as a maximum traveling salesman problem with submodular rewards," *Discrete Applied Mathematics*, vol. 186, pp. 112–127, 2015.
- [15] A. Krause and D. Golovin, "Submodular function maximization," *Tractability: Practical Approaches to Hard Problems*, vol. 3, p. 19, 2012.
- [16] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions-I," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.
- [17] M. L. Fisher, G. L. Nemhauser, and L. A. Wolsey, "An analysis of approximations for maximizing submodular set functions-II," in *Mathematical Programming Study 8*, pp. 73–87, North-Holland Publishing Company, 1978.
- [18] L. Lovász, "Submodular functions and convexity," in *Mathematical Programming The State of the Art*, pp. 235–257, Springer, 1983.
- [19] S. Fujishige, *Submodular functions and optimization*, vol. 58. Elsevier, 2005.
- [20] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák, "Maximizing a monotone submodular function subject to a matroid constraint," *SIAM Journal on Computing*, vol. 40, no. 6, pp. 1740–1766, 2011.
- [21] B. Mirzasoleiman, A. Karbasi, R. Sarkar, and A. Krause, "Distributed submodular maximization: Identifying representative elements in massive data," in *Advances in Neural Information Processing Systems*, pp. 2049–2057, 2013.
- [22] B. Mirzasoleiman, A. Karbasi, R. Sarkar, and A. Krause, "Distributed submodular maximization," *arXiv preprint arXiv:1411.0541*, 2014.
- [23] P. Stobbe, *Convex analysis for minimizing and learning submodular set functions*. PhD thesis, California Institute of Technology, 2013. Ph.D. thesis.
- [24] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, "Scalable and distributed submodular maximization with matroid constraints," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2015 13th International Symposium on*, pp. 435–442, 2015.
- [25] R. d. P. Barbosa, A. Ene, H. L. Nguyen, and J. Ward, "A new framework for distributed submodular maximization," in *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*, pp. 645–654, IEEE, 2016.
- [26] U. Feige, "A threshold of $\ln(n)$ for approximating set cover," *Journal of the ACM*, vol. 45, no. 4, pp. 634–652, 1998.
- [27] B. Korte and J. Vygen, *Combinatorial Optimization: Theory and Algorithms*, vol. 21 of *Algorithmics and Combinatorics*. Springer, 4 ed., 2007.
- [28] B. Bollobás and P. Erdős, "Cliques in random graphs," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 80, no. 03, pp. 419–427, 1976.
- [29] N. Bof, G. Baggio, and S. Zampieri, "On the role of network centrality in the controllability of complex networks," *IEEE Transactions on Control of Network Systems*, 2016. to appear.
- [30] S. Soltan, D. Mazaauric, and G. Zussman, "Analysis of failures in power grids," *IEEE Transactions on Control of Network Systems*, 2016. to appear.
- [31] J. Pattillo and S. Butenko, "Clique, independent set, and graph coloring," *Encyclopedia of Operations Research and Management Science*, pp. 3150–3163, 2011.
- [32] Gurobi Optimization, Inc., "Gurobi optimizer reference manual," 2016.
- [33] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [34] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.



Bahman Ghahsifard (M'13) received the B.Sc. degree in Mechanical Engineering, in 2002, and the M.Sc. degree in Control and Dynamics, in 2005, from Shiraz University, Iran. He received the Ph.D. degree in Mathematics, in 2009, from Queen's University, Canada.

He held postdoctoral positions with the Department of Mechanical and Aerospace Engineering at University of California, San Diego 2009-2012 and with the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign from 2012-2013. He held a visiting faculty position at the Institute for Systems Theory and Automatic Control at the University of Stuttgart in summers of 2016 and 2017. He is currently an Assistant Professor with the Department of Mathematics and Statistics at Queen's University. He received the Engineering and Applied Science First Year Instructor Teaching Award in 2015 and 2017, and was shortlisted for a Frank Knox Award for Excellence in Teaching in 2014. His research interests include systems and controls, distributed control and optimization, social and economic networks, game theory, geometric control and mechanics, and Riemannian geometry.



Stephen L. Smith (S'05–M'09–SM'15) received the B.Sc. degree in engineering physics from Queen's University, Canada, in 2003, the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Canada, in 2005, and the Ph.D. degree in mechanical engineering from the University of California, Santa Barbara, USA, in 2009.

He is currently an Associate Professor in the Department of Electrical and Computer Engineering at the University of Waterloo, Canada, where he holds a Canada Research Chair in Autonomous Systems. From 2009 to 2011 he was a Postdoctoral Associate in the Computer Science and Artificial Intelligence Laboratory at the Massachusetts Institute of Technology. His main research interests lie in control and optimization for autonomous systems, with a particular emphasis on robotic motion planning and coordination.

Prof. Smith received the Early Researcher Award from the Ontario Ministry of Research and Innovation in 2016, the NSERC Discovery Accelerator Supplement Award in 2015, and an Outstanding Performance Award from the University of Waterloo in 2015.