A Best-Response Algorithm With Voluntary Communication and Mobility Protocols for Mobile Autonomous Teams Solving the Target Assignment Problem

Sarper Aydın, *Member, IEEE*, and Ceyhun Eksin, *Member, IEEE*

***Abstract*— We consider a team of mobile autonomous robots with the aim to cover a given set of targets. Each robot aims to select a target to cover and physically reach it by the final time in coordination with other robots given the locations of targets. Robots are unaware of which targets other robots intend to cover. Each robot can control its mobility and who to send information to. We assume communication happens over a wireless channel that is subject to fading and failures. Given the setup, we propose a decentralized algorithm based on decentralized fictitious play (DFP) in which robots reason about the selections and locations of other robots to decide which target to select, whether to communicate or not, who to communicate with, and where to move. Specifically, the communication actions of the robots are learning-aware, and their mobility actions are sensitive to the success probability of communication. We show that the decentralized algorithm guarantees that robots will cover the targets in finite time. Numerical simulations and experiments using a team of mobile robots confirm the target coverage in finite time and show that mobility control for communication and learning-aware communication protocols reduce the number of communication attempts in comparison to a benchmark distributed algorithm that relies on communication at each decision epoch.**

***Index Terms*— Autonomous robots, decentralized control, multiagent systems, networked control systems.**

I. Introduction

With advances in robotics and wireless communication, autonomous systems are deployed in many different areas ranging from unmanned aerial vehicles (UAVs) [1] to watercraft systems [2], [3] and self-driving cars [4]. In practice, typical goals of multiagent robot teams can be search, rescue, and patrolling missions. Accomplishment of these goals includes, but not limited to, scanning and covering physical locations in hazardous environments, where communication abilities are limited—see [5], [6], [7] for more current applications of autonomous systems. In such team missions, autonomous robots are put together to collaboratively achieve a common goal using wireless communication and their physical abilities. Collaboration entails each team member gathering data and resolving differences with others efficiently via rapid communication to produce a joint action profile. Here, we posit that communication and mobility capabilities need to be managed by the team members based on the occurrence of a need for additional information to maximize team performance.

We specifically consider a team of robots tasked with covering a given set of targets. Target assignment problems are combinatorial optimization problems seeking assignments between resources (robots) and tasks (targets) to maximize utilities or minimize costs—see [8], [9] for detailed surveys of centralized approaches on target assignment problems.

Among the multiagent approaches to solving the variants of target assignment problem are the utility-based [10], [11] and action-based game-theoretical learning algorithms [12], [13], [14], [15], [16], auction-based algorithms [17], [18], [19], [20], dynamic partitioning and coalition formation methods [21], [22], distributed Hungarian algorithm [23], and maximal-matching algorithms [24]. While the mentioned studies so far provide (almost and near) optimal solutions to variants of the target assignment problems, they do not address communication failure under realistic conditions. Some assume perfect information provided to agents or shared between agents [10], [11], [12], [15], whereas others [13], [14], [17], [18], [19], [21], [23] rely on (strongly) connected network formed in a bounded time interval without a specific communication model. The remaining ones specifically consider limited-range communication [24], faulty communication with packet loss [22], and communication failure [16], [20]. However, those operating without connectivity assumption assume unrealistic coordination between agents, e.g., existence of rendezvous points, determination of leader agents [24], or assume unrealistic communication models [16], [20], [22].

In this article, robots have limited communication resources per decision epoch, and communication is subject to failures due to path loss and fading. Fig. 1 shows an example of a team of three robots that wants to cover three targets. Here, we model the target assignment problem as a game played among team members [12]. The game appears in this setup because the individual payoffs depend on other agents' actions (target selections) that are unknown and determined by their individual payoffs. If agents are strategic and rational, then

Manuscript received 17 January 2023; accepted 13 June 2023. This work was supported by NSF under Grant CCF-2008855. Recommended by Associate Editor G. Hu. *(Corresponding author: Ceyhun Eksin.)*

The authors are with the Industrial and Systems Engineering Department, Texas A&M University, College Station, TX 77843 USA (e-mail: sarper.aydin@tamu.edu; eksinc@tamu.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCST.2023.3296003.

Digital Object Identifier 10.1109/TCST.2023.3296003

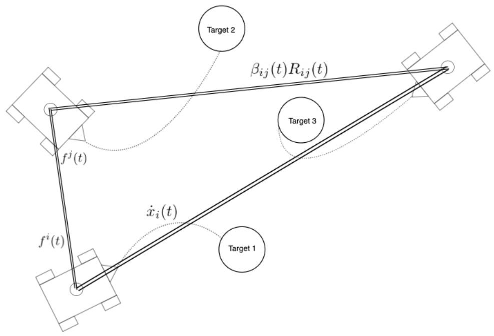

Fig. 1. Team of robots are tasked to solve a target assignment problem. Each robot relies on the local estimates of the possible target selection of other robots $f^i(t)$ to select targets. The local estimates are updated based on information received from other robots over a noisy wireless channel subject to failures. The probability of a successful communication from robot i to j $[\beta_{ij}(t)R_{ij}(t)]$ depends on their locations $[x_i(t), x_j(t)]$ and flow rate $[\beta_{ij}(t)]$ determined by robot i.

an optimal action profile is a Nash equilibrium (NE). In the target assignment game, the (pure) Nash equilibria correspond to robots covering all the targets. Given the setup, along the lines of the aforementioned vision for team collaboration, we propose a decentralized game-theoretic learning algorithm in which agents learn to cover the targets as a team by making *learning-aware communication* and *communication-aware mobility* decisions.

In particular, we generalize a decentralized form of the fictitious play (FP) algorithm, so that it is suitable for realistic communication and mobility settings and is able to manage limited communication resources (see Section III). The proposed algorithm has three main parts that operate in tandem.

1) *FP:* Agents keep estimates of the intended target selections of other agents to select best available targets.
2) *Intermittent and Voluntary Communication:* Agents use their current estimates to make voluntary communication attempts with other agents.
3) *Communication-Aware Mobility:* Agents take movement actions considering the tradeoff between covering their selected targets in a given time and increasing the chance of successful communication.

FP is a best-response-type distributed game-theoretic learning algorithm [25], [26], [27], [28]. In the standard FP algorithm [29], each agent takes an action that maximizes its expected utility assuming other agents select their actions randomly from a stationary distribution. In FP, agents assume that this stationary distribution is given by the past empirical frequency of past actions. FP is not a decentralized algorithm, since agents need to observe the past actions of *all the agents* to be able to form these distributions and compute their expected utilities. In the decentralized FP (DFP) algorithm, agents form estimates on empirical frequencies of other agents' actions by averaging the estimates of their neighbors received over a communication network. The fast convergence rate of averaging updates guarantee convergence of the DFP algorithm for potential games, i.e., games with payoffs that admit potential functions [13], [14], [26]. Here, the generalization of the DFP algorithm allows for communication failures, and intermittent and voluntary communication attempts.

Learning-aware communication refers to agents assessing novelty of their information and the information need of other agents who are potential receivers of information. Indeed, if an agent has no new information to share, it can save energy by abstaining from communication without hampering team performance. Based on this premise, recent studies in distributed optimization [30], [31], [32] propose local threshold-based communication protocols that rely on the novelty of information measured by, e.g., change in local gradient. These communication protocols are also referred to as censoring implying that a sender agent self-censors if it deems its information as stale. Along the same lines, here we consider a communication protocol that gives full autonomy to agents in deciding whether to communicate or not based on the changes in their estimates of empirical frequencies. Our protocol departs from the past approaches by the feature that agents also determine who to communicate with by assessing information need of other agents, in addition to the novelty of local information. Specifically, agents keep an estimate of the similarity between others' estimates of their empirical frequencies. If agent i assesses that agent j has a quality estimate of i's empirical frequency, it may choose to not transmit its updated empirical frequency to agent j. Moreover, each agent allocates its communication resources based on a statistic that measures the similarity of target selections. That is, if an agent is more likely to select the same target with another agent, then it is more urgent for these two agents to coordinate their selections. These features in which agents determine who to communicate with and allocate communication resources based on the urgency of information exchange make the communication protocol learning-aware.

Communication-aware mobility refers to agents determining their heading directions not only based on the selected targets and target locations but also based on their need to communicate in the presence of fading. In the target assignment game, if robots move toward their selected targets, they may quickly lose connectivity due to fading, i.e., increasing interagent distances. This may lead to certain robots committing early to their target selections without spending the time needed to coordinate their actions with all the other robots. Since the team would need to resolve robot–target allocations eventually, this mobility would be highly inefficient as some robots may take many steps toward their selected targets only to change their selections. In addressing some of these issues, recent works in mobility and communication control in autonomous teams propose mobility decisions that mind communication failures [33], [34], [35], [36]—also see [37] for a survey on the relationship of different communication setups and mobility. However, in these studies network connectivity is treated as a constraint to be satisfied by the team. Ensuring connectivity as mobile robots move to reach their selected targets can significantly hamper team performance and cause the target assignment problem to be infeasible since some targets may never be covered to remain connected. Recent studies on mobile robotic teams account for intermittent communication for distributed state estimation problems [38], [39]. Along similar lines, we incorporate information exchange needs of agents and fading effects into their mobility decisions. Our goal is to reduce the total effort spent by the team by increasing the chances of successful communication attempts.

It is worth noting that we considered similar communication schemes for the DFP algorithm in [16], [40], and [41]. In [40], we propose a self-censoring protocol based only on the novelty of information for the target assignment game. The communication protocol in this article extends it to consider the value of local information for the receiving agents. In addition, we propose a novel communication-aware mobility protocol where agents anticipate the role of their future positions in the chance of communication. In [16] and [41], we provide a theoretical framework for the convergence of the communication-censored DFP algorithms in weakly acyclic games by providing a sufficient learning condition that needs to be satisfied by a communication protocol. In this article, we propose communication protocols particular to the target assignment game and rely on the theoretical framework developed in [16] to prove convergence. We detail the contributions of this article next.

A. Contributions

The contributions of the article are threefold.

1) We formulate a novel target assignment game for which we show that the pure Nash equilibria correspond to target coverage solutions.
2) We design a decentralized inertial best-response algorithm enabling agents to select targets autonomously, decide who they send information to, and where they move without a centralized controller or access to any global information. The algorithm design entails novel learning-aware communication and communication-aware mobility protocols considering a faulty communication network.
3) We show that the decentralized algorithm will guarantee a physical coverage of targets in finite time, i.e., converge to a feasible local optimal solution of the centralized problem. The simulations and experiments demonstrate the benefit of learning aware-communication and mobility-aware communication protocols in reducing communication attempts and improving the communication success rate.

B. Organization

In Section II, we first provide the centralized target assignment model with mobility constraints. Second, we present the target-assignment game stating individual payoffs and actions. We also show that any pure Nash equilibria of the target assignment game are a one-to-one assignment of the robots to targets (see Lemma 1). In Section III, we propose the DFP with learning-aware voluntary communication and communication-aware mobility protocols, denoted as MC-DFP. In Section IV, we show that MC-DFP converges to a pure NE of the target assignment game given appropriate choices of the threshold parameters in the learning-aware communication protocol, which implies that all the targets are covered by some finite time (see Theorem 2). In Section V, numerical simulations demonstrate the reduction in the number of communication attempts due to learning-aware communication and the increased likelihood of covering the targets by a given final time. We demonstrate the practical applicability, the effectiveness, and the scalability of the decentralized decision-making scheme in experiments with a team of mobile-wheeled robots. We conclude in Section VI.

II. Target Assignment Problem and a Game Formulation

A. Target Assignment Model

We consider a team of N robots denoted with $\mathcal{N} = \{1, 2, \ldots, N\}$ that move on a 2-D surface. There are N targets denoted using $\mathcal{K} := \{1, 2, \ldots, N\}$. The goal of the team is to cover all the targets. For robot $i \in \mathcal{N}$ to cover a target $k \in \mathcal{K}$, it has to select that target. We define the selection variable $a_{ik} \in \{0, 1\}$ which is equal to 1 if robot i selects target k, and it is equal to 0, otherwise. Then the team goal to cover all the targets is achieved, when the following equations are satisfied:

$$\sum_{k\in\mathcal{K}} a_{ik} = 1, \quad i \in \mathcal{N}, \quad \text{and} \quad \sum_{i\in\mathcal{N}} a_{ik} = 1, \quad k \in \mathcal{K}. \tag{1}$$

If the conditions above are satisfied, there is a one-to-one matching between the robot–target pairs.

1) Mobility Dynamics and Coverage Constraints: Each robot starts at position $x_i(0) \in \mathbb{R}^2$ and moves to $x_i(t) \in \mathbb{R}^2$ with a chosen velocity $\dot{x}_i(t) \in \mathbb{R}^2$ for $t \in \mathcal{T} := \{1, \ldots, T_f\}$ where T_f is some final time. Assuming uniform time intervals Δt, we have the following mobility dynamics:

$$x_i(0) + \sum_{s=1}^{t} \dot{x}_i(s)\Delta t = x_i(t), \quad (i, t) \in \mathcal{N} \times \mathcal{T}. \tag{2}$$

Agent i's velocity is bounded with $V \in \mathbb{R}^+$

$$\|\dot{x}_i(t)\| \le V, \quad (i, t) \in \mathcal{N} \times \mathcal{T}. \tag{3}$$

Robots determine their velocities to reach their selected targets by the final time, i.e.,

$$x_i(T_f) = a_i^\top q := \sum_{k\in\mathcal{K}} a_{ik} q_k, \quad i \in \mathcal{N} \tag{4}$$

where $q_k \in \mathbb{R}^2$ denotes the target k's static location, the selection vector of robot i is defined as $a_i := [a_{i1}, \ldots, a_{iN}]^\top$, and q is the target location matrix that is a concatenation of the locations of all the targets. The equality in (4), when satisfied, ensures that robots are at their selected targets by time T_f.

Each robot i has to exert an effort to cover target k physically. This effort may stem from distance traversed, energy consumption, or existing preferences among targets. We denote the effort/cost required for robot i to select target k as $d_{ik} > 0$. Then the team objective can be written as to minimize total effort to cover all the targets while satisfying the conditions above

$$\min_{a_{ik}, \dot{x}_i(t)} \sum_{i\in\mathcal{N}} \sum_{k\in\mathcal{K}} d_{ik} a_{ik} \quad \text{s.t. } (1) - (4). \tag{5}$$

In the above formulation, the objective only depends on the action selections a_{ik}, not on the mobility dynamics. That is, once the robots determine their action selections that optimize the objective in (5), they can determine their velocities to satisfy (2)–(4)–see Remark 1 for further discussion on the case that the mobility dynamics are coupled with the objective. Our approach here is to provide decentralized solutions to the above problem when there is no centralized coordinator and robots have partial access to other agents' beliefs and selections due to random communication.

Indeed, the problem in (5) is relatively easy to solve when robots have complete information, i.e., all the robots know the costs of other agents d_{ik} for $i \in \mathcal{N}$ and $k \in \mathcal{K}$. In such a scenario, each robot can compute the optimal solution to (5) and implement its portion of the optimal selection and mobility dynamics. Robots cannot be sure of other robots' costs to cover the targets, e.g., when it depends on local energy consumption. This means robots need to solve (5) using their local information. Because robots have different and partial information, robots need to reason about each other's selections to make optimal selections. Here, we model the reasoning and decision-making of robots via a decentralized game-theoretic learning framework. We first define the target assignment game and then present the decentralized algorithm.

Remark 1: The problem of finding optimal paths, e.g., shortest path, and the target assignment in (5) can be coupled in the objective. For instance, the action selections could be such that the total distance traversed by the agents is minimized. This problem is a dynamic program where agents traverse the shortest path (Euclidean distance from initial locations) to their selected targets on the optimal path. That is, the dynamic program could be simplified to (5) by defining $d_{ik} := x_i(0) - a_i^T q$. Accordingly, we define the cost parameters (d_{ik}) as the Euclidean distance of initial agent positions to target locations in the numerical simulations.

B. Target Assignment Game

In a game, robot i, who knows its cost associated with covering each target $\{d_{ik}\}_{k\in\mathcal{K}}$, has to compare among its target options $\mathcal{K}$ without the knowledge of the selections of other robots. Here, we use the selection vector $a_i \in \mathbb{R}^N$ to denote the action of robot i. The action of robot i belongs to the space of canonical vectors $\mathbf{e}_k \in \mathbb{R}^{N\times 1}$ for $k = 1, \ldots, N$, which has only a single element equal to 1 and the rest of the elements are zero. That is, we define the action space as $\mathcal{A} := \{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$, meaning that only a single target can be selected. We denote the kth element of the action by $a_{ik} \in \{0, 1\}$ as per the definition of the selection variable in (1). Given the action space, we represent the utility function of robot i as follows:

$$\min_{a_i\in\mathcal{A}} \; u_i(a_i, a_{-i}) = \sum_{k\in\mathcal{K}} d_{ik}\bar{a}_{-ik}a_{ik} \tag{6}$$

where $-i := \mathcal{N}\backslash\{i\}$ denotes the set of agents other than agent i, $a_{-i} := \{a_j\}_{j\in -i}$, and $\bar{a}_{-ik} := \max\{a_{jk}\}_{j\in -i}$. The definition of $\bar{a}_{-ik}$ implies that if there exists a robot $j \in -i$ that selects target k, then $\bar{a}_{-ik} = 1$, and otherwise if none of the robots selects k, then $\bar{a}_{-ik} = 0$. As per this definition and (6), robot i's cost from selecting target k is d_{ik} if there exists another robot selecting the target. Accordingly, the cost of selecting a target k is zero, if no other agent is selecting that target.

The utility of robot i depends on the actions of all the other robots via the term $\bar{a}_{-ik}$ in (6). This dependence sets up a target assignment game among the robots defined by the tuple $\Gamma := (\mathcal{N}, \{\mathcal{A}, u_i\}_{i\in\mathcal{N}})$. In the target assignment game, the structure of objective function $u_i(a_i, a_{-i})$ together with the action space $\mathcal{A}$ assume the role of the assignment constraints in (1). Once robot i selects its target, it can determine its path toward the chosen action as per (2)–(4) to satisfy the mobility constraints.

1) Game Theory Preliminaries: A mixed strategy of robot i, denoted with σ_i, is a probability distribution over the action space, i.e., $\sigma_i \in \Delta(\mathcal{A})$. The set of joint mixed strategies is given by $\Delta^N(\mathcal{A}) = \prod_{i=1}^N \Delta(\mathcal{A})$ where we assume the individual strategies are independent. An NE of the game Γ is a strategy profile such that no individual has a unilateral profitable deviation.

Definition 1 (NE): The joint strategy profile $\sigma^* = (\sigma_i^*, \sigma_{-i}^*) \in \Delta^N(\mathcal{A})$ is an NE of the game Γ if and only if for all $i \in \mathcal{N}$

$$u_i(\sigma_i^*, \sigma_{-i}^*) \le u_i(\sigma_i, \sigma_{-i}^*), \quad \text{for all } \sigma_i \in \Delta(\mathcal{A}). \tag{7}$$

An NE strategy profile σ^* is a pure NE if $\sigma^* = (\sigma_i^*, \sigma_{-i}^*) \in \Delta^N(\mathcal{A})$ is a degenerate probability distribution, i.e., gives weight 1 on a single action profile $a = (a_i, a_{-i}) \in \mathcal{A}^N$.

A game Γ is a best-response potential game [42] if there exists a best-response potential function $u : \prod_{i\in\mathcal{N}} \mathcal{A} \to \mathbb{R}$, such that the following holds for any actions $a_i \in \mathcal{A}$ and $a_{-i} \in \mathcal{A}^{N-1} := \prod_{i\in\mathcal{N}\backslash\{i\}} \mathcal{A}$:

$$\operatorname*{argmin}_{a_i\in\mathcal{A}} \; u_i(a_i, a_{-i}) = \operatorname*{argmin}_{a_i\in\mathcal{A}} \; u(a_i, a_{-i}). \tag{8}$$

A best-response potential game with a finite set of actions $\mathcal{A}$ is weakly acyclic, i.e., a pure-strategy NE exists, and starting from any $a \in \mathcal{A}^N$, there exists a finite best-response path to a pure-strategy NE.

2) Properties of the Target Assignment Game: We show that the target assignment game is a best-response potential game with pure Nash equilibria corresponding to one-to-one assignments between robots and targets.

Lemma 1: The target assignment game defined by the tuple $\Gamma = (\mathcal{N}, \{\mathcal{A}, u_i\}_{i\in\mathcal{N}})$ with the utility function defined in (6) is a best-response potential game, and an action profile a^* is a pure NE if and only if it is a one-to-one assignment between robots and targets.

Proof: Consider the best-response potential function $u(a_i, a_{-i}) = \sum_{k\in\mathcal{K}} \bar{a}_{-ik}a_{ik}$, where $\bar{a}_{-ik}$ and a_{ik} are defined as per (6). Since the cardinality of the set of targets and robots is the same, $|\mathcal{N}| = |\mathcal{K}| = N$, there is always at least one target k uncovered given $a_{-i} \in \mathcal{A}^{N-1}$. Suppose robot i selects to cover one of the uncovered targets $\bar{k} \in \mathcal{K}$, so that $a_i = \mathbf{e}_{\bar{k}}$. Then, it holds that $u(a_i, a_{-i}) = u_i(a_i, a_{-i}) = 0$ and $\min_{a_i\in\mathcal{A}} u(a_i, a_{-i}) = \min_{a_i\in\mathcal{A}} u_i(a_i, a_{-i}) = 0$. Thus, it satisfies the condition in (8).

Suppose $a = (a_i, a_{-i}) \in \mathcal{A}^N$ is not an one-to-one assignment. For the same reason, $|\mathcal{N}| = |\mathcal{K}| = N$, any agent can unilaterally find a better solution and minimize its utility value by selecting an uncovered target. By contradiction, therefore such an assignment is not a pure NE. When it is a one-to-one assignment, no agent can unilaterally find a better solution. It concludes that only one-to-one assignments are the pure NE of this game. ■

Lemma 1 implies that any pure NE solution is a feasible solution to the target assignment problem satisfying constraints (1). Next, we show that agents cannot be indifferent between other targets at a pure NE.

Lemma 2: For any pure NE action profile $a^* \in \mathcal{A}^N$ of the target assignment game Γ, it holds that

$$\{a_i^*\} = \underset{a_i \in \mathcal{A}}{\operatorname{argmin}}\ u_i(a_i, a_{-i}^*) \tag{9}$$

that is, the minimizer is a singleton.

Proof: The statement follows by the fact that the set of pure NE action profiles of the target assignment game constitutes of action profiles that covers all the targets. At a pure NE in which all the agents cover a single target, agent i does not have an action profile that it can deviate to which will lead to an equivalent cost of zero given that $d_{ik} > 0$ for all $k \in \mathcal{K}$. ■

Remark 2: The utility function in (6) ensures that each robot incurs a cost of zero for selecting a target when each target is selected by only a single robot. Utility functions with a similar rationale are considered in [12], [43], and [44]. In the utility functions considered in [12] and [44] for the target assignment game, the payoff from selecting a target k reduces proportional to the total number of players selecting that target. [43] defines binary-valued agent utilities that are equal to 1 if agents are away by a certain distance threshold from each other considering the area coverage problem where the goal is to cover the nodes of a graph.

III. Decentralized Game-Theoretic Learning in the Target Assignment Game

We assume that robots do not have time to coordinate their actions a priori. Thus, they learn to select the optimal (equilibrium) actions in the target assignment game via repeated interaction as they move to reach their current target selections. Robots' interactions are determined by the communication model that is subject to fading and path loss. In such a setting, if robots only move toward the actions they select, the chance of successful information exchanges may significantly diminish depending on the target locations. In the following, we propose a decentralized game-theoretic learning algorithm where robots determine who to talk to and their mobility actions according to the tradeoff between the need to communicate and the goal to reach their selected targets as per (4).

DFP With Inertia: We denote the target selection of robot i at time $t \in \mathbb{N}^+$ by $a_i(t) \in \mathcal{A}$. In making its target selection, robot i needs to form estimates on the current selection of other robots to evaluate its utility in (6). Similar to standard FP, robot i assumes that other robots act according to a stationary distribution that is determined by their empirical frequency of past actions. We define the empirical frequency of robot i as follows:

$$f_i(t) = (1-\rho_1) f_i(t-1) + \rho_1 a_i(t) \tag{10}$$

where $\rho_1 \in (0, 1)$ is a fading memory constant that measures the importance of current actions. In the standard FP algorithm, robots best respond, i.e., take the action that minimizes their expected utilities computed with respect to others' empirical frequencies $f_{-i}(t) = \{f_j(t)\}_{j\in -i}$. However, in a random communication setting, it is not possible for the robots to observe the past action of all the robots at every time instant. Thus, robot i cannot correctly compute $f_j(t)$ at every time instant.

Instead, robot i needs to form local estimates of the empirical frequencies based on the information received from others. We define the estimate of robot i on robot j's empirical frequency in (10) as $f_j^i(t)$. The estimate $f_j^i(t)$ belongs to the space of probability distributions on $\mathcal{A}$ denoted as $\Delta(\mathcal{A})$. Then the expected utility of robot i with respect to its estimates $f_{-i}^i(t) := \{f_j^i(t)\}_{j\in\mathcal{N}\setminus\{i\}}$ is given by

$$u_i(a_i, f_{-i}^i(t)) = \sum_{a_{-i}} u_i(a_i, a_{-i}) f_{-i}^i(t)(a_{-i}). \tag{11}$$

In DFP with inertia, robots best respond to the estimated empirical frequencies with some probability $\epsilon_{\text{inertia}} \in (0, 1)$ for all $t \geq 2$

$$a_i(t) = \begin{cases} \underset{a_i \in \mathcal{A}}{\operatorname{argmin}}\ u_i(a_i, f_{-i}^i(t)), & \text{w.pr. } 1-\epsilon_{\text{inertia}} \\ a_i(t-1), & \text{w.pr. } \epsilon_{\text{inertia}}. \end{cases} \tag{12}$$

In the following, we describe how robots update their estimates about others' empirical frequencies $f_{-i}^i(t)$ based on message exchanges with other robots.

Information Exchange and Estimate Updates: At each time step t, robots update their individual empirical frequency $f_i(t)$ according to (10) and let $f_i^i(t) = f_i(t)$. After updating its individual empirical frequency, robots attempt to exchange their empirical frequencies with each other. Robot i updates its estimate about robot j's empirical frequency as follows:

$$f_j^i(t) = \begin{cases} (1-\rho_2) f_j^i(t-1) + \rho_2 f_j^j(t), & \text{if } c_{ji}(t) = 1 \\ f_j^i(t-1), & \text{otherwise} \end{cases} \tag{13}$$

where $\rho_2 \in (0, 1)$ is a learning rate, and $c_{ji}(t)$ is a Bernoulli random variable that indicates whether the communication attempt by robot j at time t is successful or not. If the communication is successful, then agent i takes a weighted combination of its last belief $[f_j^i(t-1)]$, and the information received $[f_j^j(t)]$. Otherwise, agent i assigns its prior belief as the current belief, i.e., $f_j^i(t) = f_j^i(t-1)$. The success probability of a communication attempt at time t, i.e., $\mathbb{P}(c_{ji}(t) = 1)$, depends on the channel model and the communication protocol that we describe next.

Communication Model: We consider point-to-point communication among robots i and j with a rate function $R_{ij}(x_i(t), x_j(t))$ that determines the amount of information robot i can send to robot j at time t. Robot i can choose the routing rate $\beta_{ij}(t) \in [0, 1]$ that controls the fraction of time robot i spends to send information to robot j at communication epoch t. The probability of existence of a communication link is given by a Bernoulli random variable, which depends on the rate function and the routing rate

$$\mathbb{P}(c_{ij}(t) = 1) = \beta_{ij}(t) R_{ij}(t) = \beta_{ij}(t)\, e^{-r\|x_i(t)-x_j(t)\|_2^2} \tag{14}$$

where $r > 0$ is the channel fading constant.

Remark 3: The communication model considered in (14) is a probabilistic model similar to the models presented in [33], [35], and [45]. For hazardous environments, in which wireless communication between robots is challenging, probabilistic models are better suited–see [46], [47] for general surveys of probabilistic wireless communication models and their usage areas.

Learning-Aware Voluntary Communication: Robot i decides on whether it wants to communicate with robot j based on two metrics: novelty of its information $H_{ii}(t) := \|f_i^i(t) - a_i(t)\|$ and the error that robot j makes in estimating i's empirical frequency $H_{ij}(t) := \|f_i^i(t) - f_i^j(t)\|$. In particular, robot i assigns a communication weight $w_{ij}(t)$ to robot j that is equal to zero if novelty and distance conditions are, respectively, below certain threshold constants $\eta_1 \in (0, 1)$ and $\eta_2 \in (0, 1)$; otherwise, the weight is equal to the inverse of the empirical frequency overlap between the two robots defined as $\Delta_{ij}(t) := \max(\delta_1, \|f_i^i(t) - f_j^i(t)\|)$, where $\delta_1 \in (0, 1)$ is a positive lower bound on $\Delta_{ij}(t)$

$$w_{ij}(t) = \begin{cases} 0, & \text{if } H_{ii}(t) \le \eta_1 \text{ and } H_{ij}(t) \le \eta_2 \\ \dfrac{1}{\Delta_{ij}(t)}, & \text{otherwise.} \end{cases} \tag{15}$$

We provide an intuition for the above threshold rule. The novelty $H_{ii}(t)$ measures the change in the empirical frequency of robot i. $H_{ii}(t)$ gets smaller when robot i repeatedly selects the same target as per (10). Together with the condition that $H_{ij}(t)$ needs to be smaller than η_2, the above threshold rule checks whether robot j needs further information from i in predicting i's target selection accurately. In the case that these thresholds are not met, i.e., $H_{ii}(t) > \eta_1$ or $H_{ij}(t) > \eta_2$, then the communication weight depends on the overlap metric $\Delta_{ij}(t)$. The overlap metric is the estimated similarity between the empirical frequencies of robots i and j. If $\Delta_{ij}(t)$ is small, then two robots are likely to select the same targets according to robot i. The smaller the $\Delta_{ij}(t)$, the more important it is for robot i to coordinate the selection of the targets with robot j so that robots i and j do not select to cover the same target. The constant δ_1 puts a cap on the communication weight that a single robot j can have.

To compute the communication weight (15), robot i needs access to its own empirical frequency $f_i(t)$, its estimate of j's empirical frequency $f_j^i(t)$, and robot j's estimate of i's empirical frequency $f_i^j(t)$. Robot i can locally compute $f_i(t)$ and $f_j^i(t)$ using (10) and (13), respectively. For computing $f_i^j(t)$, here we devise an acknowledgment protocol where we assume the receiving robot (j) sends an "ACK" signal to the sender (i) upon successful communication. Given this protocol and the initial estimate of j on i's empirical frequency $f_i^j(0)$, robot i (sender) can keep track of the value of $f_i^j(t)$ using the update rule in (13) with indices i and j exchanged.

At each step, robot i computes a communication weight for all the robots as per (15). Together these weights $\{w_{ij}(t)\}_{j\in\mathcal{N}\setminus\{i\}}$ determine the relative importance of communicating with other robots. Next we explain how these weights are used in determining flow rates.

Robot i allocates its routing rate by solving the following optimization problem:

$$\max_{\substack{\beta_{ij}(t)\in[0,1] \\ \sum_j \beta_{ij}(t)\le 1}} \prod_{j\in -i} \mathbb{P}(c_{ij}(t) = 1) \tag{16}$$

where the total flow rate is capped by 1, i.e., the sum of routing rates cannot be infinite in a given communication epoch. Given the structure of the communication channel in (14), we incorporate the communication weights $w_{ij}(t)$ into (16). Then, the optimization problem can be reduced to the following by taking the logarithm of the product term and weighting each term by $w_{ij}(t)$:

$$\max_{\substack{\beta_{ij}(t)\in[0,1] \\ \sum_j \beta_{ij}(t)\le 1}} \sum_{j\in\mathcal{N}\setminus\{i\}} w_{ij}(t)\log(\beta_{ij}(t)). \tag{17}$$

Robot i solves the concave optimization problem in (17) at each time step t to determine the flow rates $\beta_{ij}(t)$. Due to the threshold condition in (15), the weights in (16) for some of the robots can be zero, which means robot i eliminates a subset of the robots from the communication all together.

In reducing problem (16) to (17), we observe that the fading terms in communication $r\|x_i(t) - x_j(t)\|_2^2$ disappeared. This means the positions $x_i(t)$ and $x_j(t)$ do not play a role in determining the flow rates; even though the chance of successful communication between robot i and robot j depends on the positions as per (14). Indeed, robots can adjust their positions based on who they need to send information to. In the following, we propose such a mobility decision-making protocol that accounts for both the dependence of communication success on distances, and reaching the selected targets.

Communication-Aware Mobility: Each robot needs to cover the target they selected by the final time T_f. At the same time, as per the communication model in (14), the distance between robots is crucial to the success of a communication attempt. Since the exact location of robot j $[x_j(t)]$ is unknown to each robot i, robot i can use its estimate of agent j's empirical frequency $[f_j^i(t)]$ and target locations q together with (4) to estimate the current location of robot j as shown below

$$\hat{x}_j^i(T_f, t) = f_j^i(t)^\top q. \tag{18}$$

Given the estimates of the locations of other robots $\{\hat{x}_j^i\}_{j\in\mathcal{N}\setminus i}$, robot i selects its next heading direction $x_i^{\text{dir}}(t)$ to jointly optimize communication success and covering the selected target by solving the following problem:

$$\min_{x_i^{\text{dir}}(t)\in\mathbb{R}^2} \sum_{j\in\mathcal{N}\setminus\{i\}} v_{ij}(t)\left\|x_i^{\text{dir}}(t) - \hat{x}_j^i(T_f, t)\right\|^2 + \left\|x_i^{\text{dir}}(t) - \bar{x}_i(T_f, t)\right\|^2 \tag{19}$$

where $\bar{x}_i(T_f, t) := a_i(t)^\top q$ denotes the location of the target selected by robot i, and $v_{ij}(t) > 0$ represents robot i's preference to be close to robot j. The weight $v_{ij}(t)$ is computed using the same threshold condition as the communication weight $w_{ij}(t)$ in (15) but is computed using updated empirical frequency estimates post communication

Algorithm 1 MC-DFP for Robot i

1: **Input:** Physical locations $x_{i0} \in \mathbb{R}^2$ for all $i \in \mathcal{N}$; $q_k \in \mathbb{R}^2$ for all $k \in \mathcal{K}$; the parameters $\rho_1, \rho_2, \alpha(t), \eta_1, \eta_2, \delta_1, T_f$.
2: **for** $t = 1, 2, \cdots, T_f$ **do**
3: Select an action $a_i(t)$ using (12).
4: Update $f_i^i(t)$ with the selected action via (10).
5: Compute weights $w_{ij}(t)$ (15) for all $j \neq i$.
6: If $w_{ij}(t) \neq 0$, each robot i decides the routing variables $\beta_{ij}(t)$ (17) and transmit its empirical frequency $f_i(t)$ with probability of success $\beta_{ij}(t)R_{ij}(t)$.
7: Update $\{f_j^i(t)\}_{j\in\mathcal{N}}$ using (13).
8: Compute the weights $v_{ij}(t)$ using (15).
9: Determine direction (19) and move according to (20).
10: **end for**

phase. In other words, $v_{ij}(t)$ is an updated version of $w_{ij}(t)$ computed after the information exchange and belief updates. After determining the direction $x_i^{\text{dir}}(t)$, robot i's velocity is given by

$$\dot{x}_i(t) = \frac{\alpha(t)\big(x_i^{\text{dir}}(t) - x_i(t-1)\big)}{\Delta t} \tag{20}$$

where $\alpha(t)$ is the speed of robots at time t, respecting physical constraints.

A. Algorithm

Algorithm 1 summarizes DFP proposed in Section III that determines mobility (M) and communication (C) decisions of robot i and thus is referred to as MC-DFP.

Robots start the updates at each time step with the selection of a target in step 3. In steps 4 and 5, robots determine their current empirical frequencies and their communication weights, which they use to determine their flow rates. In step 6, all the robots engage in a round of communication with the determined flow rates. After robots receive new information, they update their estimates about the empirical frequencies in step 7. The updated frequencies are used to determine where robots move next in steps 8 and 9.

MC-DFP has two mechanisms, namely, learning-aware voluntary communication (steps 5 and 6) and communication-aware mobility (steps 8 and 9), which makes it distinct from prior decentralized approaches in team of mobile robots [33], [34], [35]. In contrast to prior approaches that focus on ensuring probabilistic connectivity for all the time steps, the proposed communication and mobility protocols make learning of others' selections the goal. Moreover, the MC-DFP algorithm considers realistic communication and mobility models compared with prior decentralized game-theoretic learning schemes, e.g., [13], [14], [20].

Remark 4: MC-DFP aims to reduce communication attempts by having agents reason about the value of their local information. In this aspect, the protocols are similar to the voluntary communication protocols considered in [48] for DFP. The protocols in this article are tailored to the target assignment problem solved by a team of mobile agents. In particular, we use mobility protocols to aid communication, and thus learning.

Remark 5: The iteration complexity of MC-DFP is polynomial in N. Target selection has polynomial complexity since the computation of probabilities of target selections is polynomial and selecting an element among possible choices can be completed in $O(N)$. For learning-aware communication and communication-aware mobility, again, the computation of weights consists of simple array operations and is polynomial. We solve the voluntary communication problem via a concave maximization problem with linear constraints. The objective is a sum of concave and separable functions and is shown to be solvable in polynomial complexity in terms of problem size and solution accuracy [49]. Communication-aware mobility is an unconstrained and convex quadratic minimization problem. It has a closed-form solution using the first-order condition to find optimal solutions. This means that complexity per iteration is polynomial.

IV. Convergence Analysis

In the following, we show that MC-DFP (Algorithm 1) converges to a rational action profile implying that the constraints of the target assignment problem in (1) and (4) are satisfied.

We begin by stating prior conditions and results that will be used in the convergence analysis.

A. Prior Results

We will use a result from [48] to show the convergence of MC-DFP. The result in [48] hinges on the following condition that an information exchange and belief update protocol needs to satisfy.

Condition 1 (see [48]): There exists a positive probability $\hat{\epsilon} > 0$ and a finite time $\hat{T}$ such that if an agent $j \in \mathcal{N}$ repeats the same action for at least $T > \hat{T}$ times starting from time $t > 0$, i.e., $a_j(s) = \mathbf{e}_k$ for $s = t, t+1, \ldots, t+T-1$ and $\mathbf{e}_k \in \mathcal{A}$, agent $i \in \mathcal{N}$ learns agent j's action with positive probability $\hat{\epsilon} > 0$, i.e., $\mathbb{P}(\|a_j(t+T) - f_j^i(t+T)\| \leq \xi | \mathcal{H}(t)) \geq \hat{\epsilon}$ for any $\xi > 0$, where $\{\mathcal{H}(t)\}_{t\geq 0}$ is a sub-sigma algebra of the Borel sigma algebra $\mathcal{B}$ created by the set $(\mathcal{A}^N \times \mathcal{G})^t$ of actions and the space of all possible networks $\mathcal{G}$.

Condition 1, if satisfied by a communication and belief update protocol, e.g., in MC-DFP, implies that robot i is able to learn the action of another robot j with some positive probability when robot j repeats the same action. Condition 1 requires a communication and belief update protocol to be able to learn given an environment that is static for a finite time horizon. The following result, given in [48], states that any FP-type algorithm with inertia will converge to a pure NE of any weakly acyclic game given that it satisfies this condition.

Theorem 1 (see [48]): Let $\{a(t) = (a_1(t), (a_2(t), \ldots, a_N(t))\}_{t\geq 1}$ be a sequence of actions generated by an FP-type algorithm with inertia where agents best respond as per (12) and update local empirical frequencies as per (10). Suppose the local empirical frequency estimates of agent $i \in \mathcal{N}$ $[f_{-i}^i(t)]$ satisfy Condition 1. If the game $\Gamma = (\mathcal{N}, \{\mathcal{A}, u_i\}_{i\in\mathcal{N}})$ is weakly acyclic, and agents are not indifferent between any two actions at a pure NE, i.e., the set of minimizers in (12) is a singleton $\{a_i^*\} = \operatorname{argmin}_{a_i \in \mathcal{A}} u_i(a_i, a_{-i}^*)$, then the action sequence $\{a(t)\}_{t\geq 1}$ converges to a pure NE a^* of the game Γ, almost surely.

B. Convergence

We establish the convergence of MC-DFP via Theorem 1 by showing that its conditions are satisfied.

We make the following two assumptions on bounded distance between robots and targets and on guaranteed success of acknowledgment signals.

Assumption 1: There exists a positive real number $D > 0$ such that $\|x_i(t) - x_j(t)\| \leq D$, for all $(i, j, t) \in \mathcal{N} \times \mathcal{N}\backslash\{j\} \times \mathcal{T}$, and $\|x_i(t) - q_k\| \leq D$, for all $(i, k, t) \in \mathcal{N} \times \mathcal{K} \times \mathcal{T}$.

Assumption 2: A receiving robot $j \in \mathcal{N}\backslash\{i\}$ can successfully acknowledge if they received the estimates $f_i^i(t)$ from the sender robot i given $c_{ij}(t) = 1$.

Assumption 2 assures that acknowledgment signals are always received without failures. Thus, the metric $H_{ij}(t)$ used in the voluntary communication protocol is accurate. From a theoretical perspective, this assumption allows us to lower bound the probability of successful communication as per (14). The acknowledgment signal has complexity of $O(1)$, so ensuring its transmission is not costly for the receiving agent.

First, we show that the communication and belief update protocol of MC-DFP in Algorithm 1 satisfies Condition 1.

Lemma 3: Suppose Assumptions 1 and 2 hold. Let $\{a(t) = (a_1(t), (a_2(t), \ldots, a_N(t))\}_{t\geq 1}$ be a sequence of actions generated by MC-DFP (Algorithm 1). Then, Condition 1 is satisfied for any $\xi > 0$ given small enough $0 \leq \eta_1 < \xi/2$, and $0 \leq \eta_2 \leq \xi/2$ such that if an agent $j \in \mathcal{N}$ repeats the same action for at least $T > \hat{T}$ times starting from time $t > 0$, i.e., $a_j(s) = \mathbf{e}_k$ for $s = t, t+1, \ldots, t+T-1$ and $\mathbf{e}_k \in \mathcal{A}$, agent $i \in \mathcal{N}$ learns agent j's action with positive probability $\hat{\epsilon} > 0$, i.e., $\mathbb{P}(\|a_j(t+T) - f_j^i(t+T)\| \leq \xi | \mathcal{H}(t)) \geq \hat{\epsilon}$.

Proof: See Appendix. ■

Next, we state the main convergence result for MC-DFP.

Theorem 2: Suppose Assumptions 1 and 2 hold. Then, at some finite time t, robots implementing MC-DFP [Algorithm 1 achieve the team goal (1) and cover targets physically (4)].

Proof: By Lemma 1, there exists a finite best-response path to reach pure NE $a = (a_i^*, a_{-i}^*) \in \mathcal{A}^N$ from any joint action profile $(a_i, a_{-i}) \in \mathcal{A}^N$ in target assignment game defined by the tuple $\Gamma = (\mathcal{N}, \{\mathcal{A}, u_i\}_{i\in\mathcal{N}})$. Lemma 2 assures that agents are not indifferent between any two actions at a pure NE. Finally, Lemma 3 satisfies Condition 1. Thus, by Theorem 1, the sequence of joint pure action profiles $\{a_t = (a_{it}, a_{-it})\}_{t\geq 1}$ converges to a pure NE (a_i^*, a_{-i}^*) almost surely in finite time.

At a pure NE, there is a one-to-one assignment of robots to targets. Then, if robots converge to an NE a^* in finite time t, the mobility weights $v_{ij}(t)$ are zero, i.e., $v_{ij}(t) = 0$ for all $j \in \mathcal{N}$. Thus, each robot goes in the direction of their selected target (q_k), without changing $a_i^* = \mathbf{e}_k$. By Assumption 1, robots arrive at their selected target locations by following the mobility dynamics in (2) and (3) satisfying (4) in some finite time. ■

The result above shows that MC-DFP is guaranteed to reach a feasible solution of (5). Theorem 2 provides convergence guarantees for MC-DFP despite the fact that robots can choose to cutoff communication based on local statistics as per (15) or move toward other robots to increase communication as per (19) and (20).

In Section V, we assess the effects of voluntary communication and learning-aware mobility dynamics in MC-DFP in terms of convergence time and number of communication attempts using simulations and experiments.

V. Simulation and Experiments

In both the simulations and experiments, we assume that the cost of robot i to cover target k is proportional to the distance that it has to traverse, defined as $d_{ik} := \|x_i(0) - q_k\|_2^2$.

A. Simulation Setup

We consider $N = 5$ robots and targets, in which robots and targets are positioned according to two different scenarios. In Scenario 1, robots start at the origin $(0, 0)$ and targets are located at $(0, 1), (1, 1,), (1, -1), (-1, 1), (-1, -1)$. For Scenario 2, robots are positioned at different starting points $(-0.5, 0,), (-0.5, -0.5), (-0.5, 0.5), (0.5, 0.5), (0.5, -0.5)$, and also targets are given as $(0, 0), (-0.5, 1.5), (-0.5, -1.5), (0.5, 1.5), (0.5, -1.5)$.

The algorithmic parameters ρ_1, ρ_2, and $\epsilon_{\text{inertia}}$ are chosen as 0.4, 1, and 0.05, respectively. The initial empirical frequencies and their estimates $(f_i^i(t)$ and $f_j^i(0)$ for all $(i, j) \in \mathcal{N} \times \mathcal{N}\backslash\{j\})$ are assigned as uniformly distributed over five targets so that $f_i^i(t) = [0.2, \ldots, 0.2]$ and $f_j^i(0) = [0.2, \ldots, 0.2]$. The channel fading constant r is determined as 0.65. Moreover, each scenario is experimented with different constant speed values over time $\alpha(t)$, which are, respectively, 0.1 and 0.05 for Scenario 1 and 0.05 and 0.025 for Scenario 2. Communication threshold constants (η_1, η_2) are given as $(0.1, 0.4)$. We explore the MC-DFP performance with respect to parameters ρ_1, ρ_2, η_1, and η_2 in Section V-E. Finally, the upper bound for Δ_{ij} in (15) is selected as $\delta = 10$. Targets are assumed to be covered if the Euclidean distance to final positions of robots is within 0.1.

Given the setup, we compare the performance of MC-DFP with respect to two decentralized benchmark learning schemes. The first benchmark learning scheme only uses learning-aware voluntary communication and does not use communication-aware mobility, i.e., it only moves toward the selected target. We denote this learning algorithm as the C-DFP algorithm. The second benchmark algorithm only implements DFP without learning-aware voluntary communication and communication-aware mobility. We denote this learning algorithm as DFP. In DFP, we further replace the voluntary communication protocol in C-DFP by a fixed communication protocol where robot i attempts to communicate at all the time steps with equal flow rates for all the robots, i.e., $\beta_{ij}(t) = (1/(N-1)) = 0.25$.

B. Rate of Convergence to an NE and Estimation Errors

Fig. 2 (left) illustrates the convergence to equilibrium in Scenario 1 with top and bottom figures corresponding to

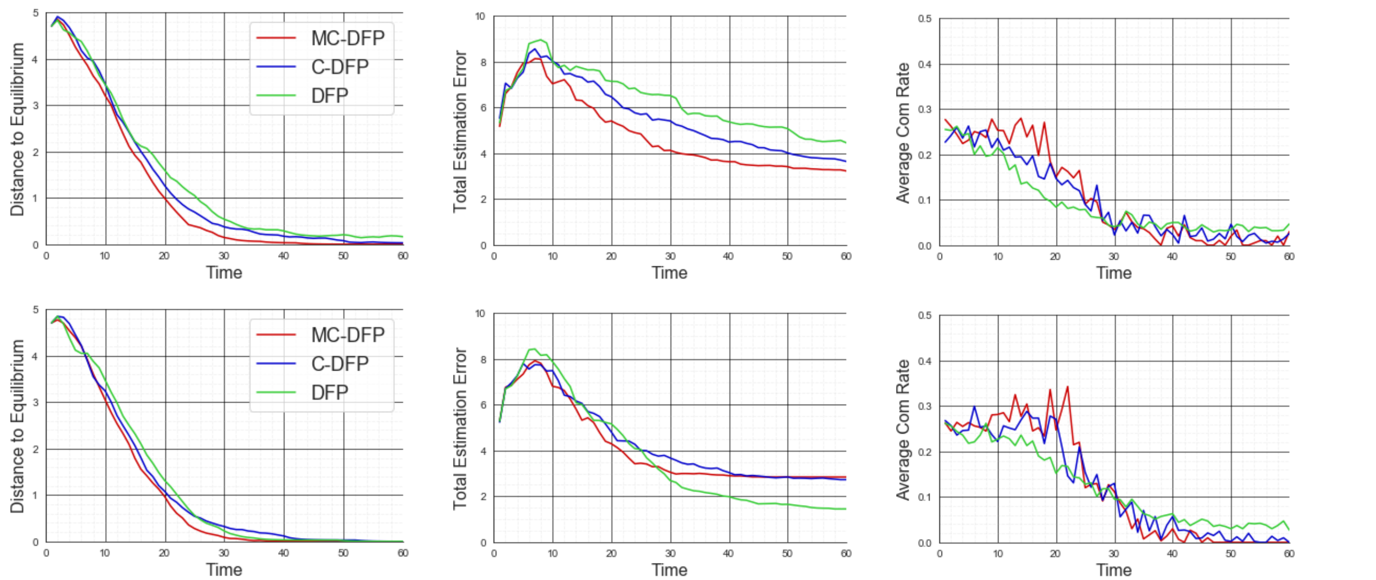

Fig. 2. Convergence results over 50 replications for Scenario 1 for speeds (top row) $\alpha = 0.1$ and (bottom row) $\alpha = 0.05$. (Left) Convergence of empirical frequencies to pure NE $\sum_{i\in\mathcal{N}} \|f_i^i(t) - a_i^*\|$. (Middle) Convergence of estimation errors $\sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}\setminus\{j\}} \|f_i^i(t) - f_i^j(t)\|$. (Right) Success ratio of communication attempts over time $(\sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}\setminus\{j\}}((c_{ij}(t))/(\mathbf{1}_{\beta_{ij}(t)>0})))$. Robots play a pure NE action profile after $t = 40$ on average.

speeds 0.1 and 0.05, respectively. All the three algorithms converge to a pure NE in all the 50 cases within the time frame T_f. MC-DFP has a slightly faster average convergence rate. We do not observe a significant effect of robot speed in convergence to NE while it has some effect on communication success as we discuss in Section V-D.

Note that only the benchmark DFP has positive communication weights at all times. This means the total estimation error of robots estimating each other's empirical will go to zero. Bechmark DFP is the only algorithm among the three that guarantees convergence to zero in estimation errors. However, given the communication failures due to fading, diminishing of estimation errors may take a long time to be practically relevant as is evident from the similarity of the estimation errors among the three algorithms in Fig. 2 (middle). Fig. 2 (middle) shows the total error robots make in estimating each other's empirical frequencies. Combined with the fact that all the learning algorithms converge, i.e., the action profile is an NE, before the final time T_f, we can conclude that robots can converge to a pure NE even when there remain gaps between actual and estimated empirical frequencies. That is, the sustained communication attempts in DFP do not provide an advantage over C-DFP or MC-DFP. In summary, DFP comes with unnecessary communication attempts incurring significant energy costs to robots as we explore next.

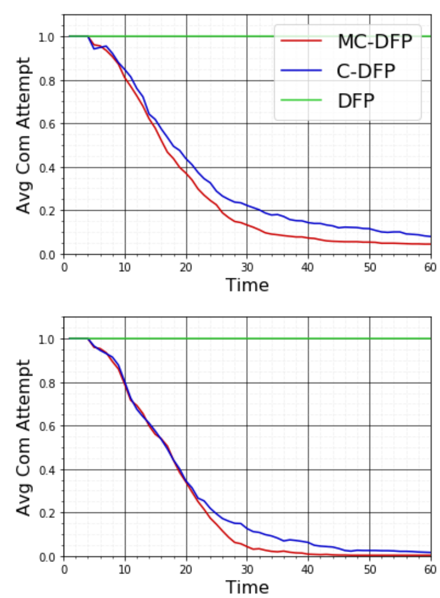

Fig. 3. Average communication attempts per link in Scenario 1 for speeds (top) 0.1 and (bottom) 0.05. Average communication attempt per link is obtained by dividing the total number of communication attempts at each step by the total number of possible communication attempts, which is equal to 20. The results show average over all 50 runs. MC-DFP reduces the total communication attempts over the entire horizon by a factor of 3 on average compared with sustained communication in DFP.

C. Effects of Learning-Aware Voluntary Communication

The total estimation errors with respect to time follow a similar shape for all the algorithms—see Fig. 2 (middle). There is an initial increase in the estimation error starting from an uninformative common prior $f_j^i(0)$ as robots begin to make target selections using best response with inertia. After reaching a peak around $t = 8$, the total estimation error decreases implying that robots learn the empirical frequencies of others. In C-DFP and MC-DFP, as robots successfully transmit their empirical frequencies, they begin to reduce communication attempts as per (15). Indeed, after time $t = 5$, robots begin to reduce communication attempts in both C-DFP and MC-DFP. By time $t = 18$, communication attempt per link drops below 0.5 for both C-DFP and MC-DFP. That is, robots attempt to use a link less than 50% of the time. The average communication attempt per link shown in Fig. 3 highlights the relative reduction in total cost of communication energy.

The cease of communication attempts leads to a slow down in descent of total estimation errors in C-DFP and MC-DFP compared with DFP [see Fig. 2 (middle)]. Nevertheless, the slow down does not prohibit convergence to an NE as discussed in Section V-B. Moreover, when robots are moving faster, we observe that robots have higher total estimation errors in DFP due to fading becoming an important factor early on [compare top and bottom rows of Fig. 2 (middle)]. The intuition for this is as follows. In contrast to DFP, robots allocate communication rates by prioritizing robots based on their need for information in C-DFP and MC-DFP. This helps in obtaining smaller estimation errors faster when fading is important as in the case when robot speeds are fast.

Fig. 2 (right) shows the average success ratio of communication attempts with respect to time in the three learning schemes. All the learning models start with similar success rates as neither prioritization nor mobility has any effect on communication success. Over time, there is a gradual decrease in chance of communication success for all the models due to robots moving away from each other toward their selected targets. However, this gradual decrease is faster at the beginning ($t \in (0, 20]$) for DFP as robots do not allocate their communication rates by prioritization as they do in C-DFP. After time $t = 30$, communication success ratio drops to zero for C-DFP and MC-DFP while DFP retains a small chance of success around 0.05. This is because we let communication success be equal to zero by convention if a communication attempt between two robots is ceased. Overall, the voluntary communication protocol in (15) saves energy without hampering team performance with appropriately chosen communication threshold constants.

D. Effects of Communication-Aware Mobility

Fig. 2 (right) also demonstrates the effect of mobility on communication success ratio. Specifically, at the beginning $t \in (0, 20]$, robots' attempts to overcome fading by moving toward their intended communication targets (receiving robots) yield higher success rate for communication in MC-DFP compared with other algorithms. This high success rate results in lower average communication attempt per link in Fig. 3.

Fig. 4 demonstrates the effects of communication-aware mobility on the team movement for Scenarios 1 and 2. In Scenario 1 [see Fig. 4 (top)], robots start from the same location which means communication failure due to fading is not likely. In Fig. 4 (top left), robots stay close due to the communication-aware direction selections as per (19). In contrast, when robots move toward their selected targets in DFP, we observe robots heading away from each other early on following their best target selections followed by sharp direction changes in Fig. 4 (top right). In Scenario 2—Fig. 4 (bottom), three robots on the left are close to each other but are far from the two robots on the right who are also close to each other. This implies that the robots on the left are highly unlikely to communicate with the robots on the right at the beginning. In MC-DFP (bottom left), all the robots move toward the center target for some time increasing the chance of successful communication between the initially disconnected robots. This behavior that minds communication highly increases the team's chance to cover each target by the final time. In contrast, robot movements are driven by target selections in DFP (bottom right). This reduces the chance of communication between robots on the left with robots on the right, leading to some targets not being covered by the final time.

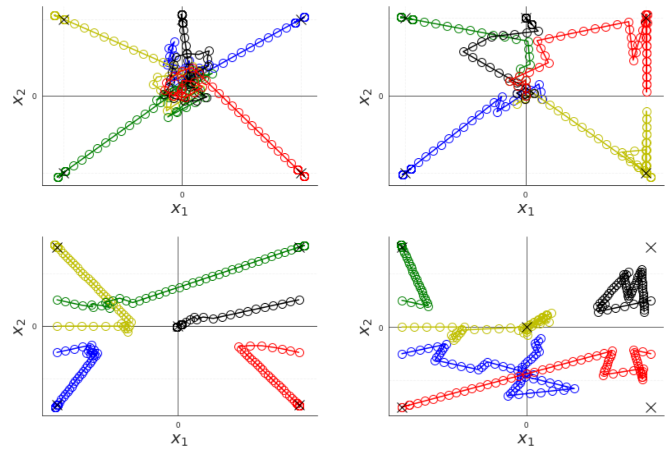

Fig. 4. Positions of robots over time in (left) MC-DFP and (right) DFP in (top) Scenarios 1 and (bottom) 2. In Scenarios 1 and 2, robots move at speeds $\alpha = 0.1$ and $\alpha = 0.05$, respectively. (Left) All the robots arrive at targets by time T_f in MC-DFP for both the scenarios. (Right) Targets remain uncovered in DFP for both the scenarios. Mobility-aware communication allows quick dissemination of information by evading failures due to path loss.

TABLE I
CHANCE OF SUCCESSFUL PHYSICAL COVERAGE BY FINAL TIME

		Coverage		
	Speed	MC-DFP	C-DFP	DFP
Scenario 1	0.1	1.00	0.96	0.86
	0.05	0.98	0.92	0.90
Scenario 2	0.05	0.96	0.92	0.94
	0.025	0.74	0.58	0.42

We further analyze the effect of speed on team's likelihood of covering every target in different scenarios. Table I shows that with decreasing speed, convergence is less likely. In particular for Scenario 2 where subsets of robots start distant from each other (high initial fading), the likelihood of covering all the targets by final time drops for all the algorithms. This drop is higher in C-DFP and DFP compared with MC-DFP. Overall, MC-DFP obtains superior performance in more challenging scenarios, e.g., when robot speeds are slow, or robots start distant from each other. We also note that MC-DFP does not incur additional computational burden because the optimization problem in (19) that determines direction admits a closed-form solution.

E. Parameter Sensitivity

We analyze the effects of fading memory constants ρ_1 and ρ_2, and threshold constants η_1 and η_2 in MC-DFP for Scenario 1. We consider large $(\rho_1, \rho_2) = (0.5, 1)$ and small $(\rho_1, \rho_2) = (0.1, 0.2)$ fading constant values along with large $(\eta_1, \eta_2) = (0.2, 1.5)$ and small $(\eta_1, \eta_2) = (0.1, 0.4)$ communication threshold constants. As fading memory constants take larger

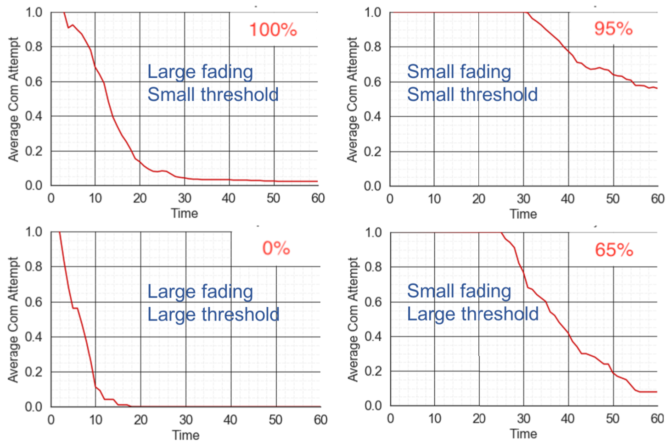

Fig. 5. Average communication attempts per link over time with different parameters in MC-DFP. Selected values of parameters (ρ_1, ρ_2, η_1, η_2) are for: 1) top left: (0.5, 1, 0.1, 0.4); 2) top right: (0.1, 0.2, 0.1,0.4); 3) bottom left: (0.5, 1, 0.2,1.5); and 4) bottom right: (0.1, 0.2, 0.2,1.5). For each set of parameters, we show average communication attempt per link over 20 runs of Scenario 1 with speed $\alpha = 0.1$. Percentage values in red in each figure show the success rate of NE convergence. The case with large fading constants combined with small threshold constants (top left) is both effective and efficient.

values, robots dismiss past information faster. As threshold constants take small values, robots are less likely to cut communication as per (15). We observe that as threshold constants increase, the likelihood of successful convergence to NE drops significantly [compare percentage values in Fig. 5 (top and bottom)]. Moreover, if threshold constants are low enough, then it is better to have high fading constants in terms of saving communication energy [compare Fig. 5 (top left and top right)]. However, if threshold constants are high, then it is better to have small fading constants so that communication is not cut very early to prohibit convergence to NE [compare Fig. 5 (bottom left and bottom right)]. Overall, small communication threshold values combined with high fading constants guarantee convergence while reducing communication attempts by threefold compared with DFP. An intuition for this is that agents can weigh most recent information more heavily, if the communication attempts are more likely to continue.

F. Experiments

We use GRITSBot X model mobile-wheeled robots, each 11-cm wide, 10-cm long, and 7-cm tall, made available by the Robotarium project [50]. The robots operate in a 3.2 × 2-m area with a maximum speed of 20 cm/s linearly and a maximum rotational speed of about 3.6 rad/s. For more technical details, see [50].

We coded the implementation in MATLAB and use the same set of parameter values used in Fig. 5 (top left), i.e., ($\rho_1 = 0.5, \rho_2 = 1, \eta_1 = 0.1, \eta_2 = 0.4$). Similarly, we select the inertia probability $\epsilon_{\text{inertia}}$ and the channel fading constant r as 0.1 and 0.60, respectively. We let Δ_{ij} in (15) be equal to 10 as before. The random communication model is simulated within MATLAB.

We consider randomly assigned initial robot and target positions for a team of size $N = 5$, $N = 10$, and $N = 18$ (maximum number allowed in Robotarium). The links for the experiments for each case can be found as below:

1) https://youtu.be/wuJLI6CNGgU (five robots);
2) https://youtu.be/KyoPGJNql2c (ten robots);
3) https://youtu.be/PSn_osWDGXA (18 robots).

We observe that robots are able to cover a sequence of randomly assigned targets for all the scenarios. The average time to cover targets is approximately 30, 39, and 49 s, respectively, for $N = 5$, 10, and 18.

There are several differences in the movements of robots between the simulations and the experiments. Due to linear and rotational speed limitations, robots are not always able to turn and go through the directions assigned by the algorithm. Moreover, local collision avoidance protocols, which prohibit robots getting too close, limit the mobility decisions. We observe that these differences do not affect the overall performance of MC-DFP. Robots successfully reach one-to-one assignment with targets and cover targets physically.

VI. Conclusion

We proposed decentralized mobility and communication protocols for a team of robots solving a target assignment problem by best responding to the intended target selection of other robots. Each robot learns about others' intended selections by keeping track of others' frequency of past actions. For keeping such estimates, robots need to be able to transmit their empirical frequencies to each other over a wireless network subject to path loss and fading. The proposed communication protocol relies on metrics that measure novelty of information and information need of the robots to decide whether to transmit or not and how to allocate available communication resources. Moreover, robots may alter their mobility to overcome fading in communication depending on their assessment of the need to communicate certain robots. We stated sufficient conditions for convergence to an NE and presented the numerical and experimental results that demonstrated the benefits of the proposed learning-aware voluntary communication and the communication-aware mobility protocols on reducing communication need while retaining convergence guarantees.

Appendix

Proof of Lemma 3

Let us define the following events E_1 and E_2 to show Condition 1 holds:

$$E_1(t) = \{\|a_j(t+T) - f_j(t+T)\| \leq \xi/2\} \quad (21)$$
$$E_2(t) = \{\|f_j(t+T) - f_j^i(t+T)\| \leq \xi/2\}. \quad (22)$$

By the triangle inequality, we have

$$\mathbb{P}(\|a_j(t+T) - f_j^i(t+T)\| \leq \xi|\mathcal{H}(t)) \geq \mathbb{P}(E_1(t), E_2(t))|\mathcal{H}(t)). \quad (23)$$

Hence, it is enough to show that the intersection of the events E_1 and E_2 has positive probability to assure Condition 1. Given the repetition of actions by an agent j, for any fading rate $\rho_1 \in (0, 1]$, there exists a finitely long enough $\hat{T}$ as the lower bound on agent j repeating the same action $a_j \in \mathcal{A}$ such that the following holds:

$$\mathbb{P}(E_1(t))|\mathcal{H}(t)) = 1. \tag{24}$$

Observe that the model (17) always admits optimal solutions $\beta_{ij}^*(t) > b$ where $b > 0$, as long as the weights $w_{ij}(t) > 0$. Combined with Assumption 1, it holds

$$\beta_{ij}^*(t)\, e^{-r\|x_i(t)-x_j(t)\|_2^2} \geq \epsilon_{\text{com}} = be^{(-rD^2)} > 0 \tag{25}$$

when $w_{ij}(t) > 0$ with small enough $0 \leq \eta_1 < \xi/2$, and $0 \leq \eta_2 \leq \xi/2$ by the definition (15). Further note that Assumption 2 lets agent j to acknowledge its successful communication so that agent j can also locally compute agent i's information on itself $f_j^i(t)$ for any time t. For any $\rho_2 \in (0, 1]$, during consecutive repetition of the same action $a_j = \mathbf{e}_k$ from time t to $t+T$, starting at time $t+T_1$ where $T = T_1 + T_2$ and $T_1 < T$, agent j needs to send f_j for T_2 times ending at $t+T$. This provides that the event $E_2(t)$ has a positive probability

$$\mathbb{P}(E_2(t))|\mathcal{H}(t)) \geq \epsilon_{\text{com}}^{T_2}. \tag{26}$$

Thus, there exists a positive bound $\hat{\epsilon}$ on the probability of the given event in Condition 1

$$\mathbb{P}\left(\left\|a_j(t+T) - f_j^i(t+T)\right\| \leq \xi|\mathcal{H}(t)\right) \geq \mathbb{P}(E_1(t), E_2(t)|\mathcal{H}(t)) \geq \epsilon_{\text{com}}^{T_2} = \hat{\epsilon} > 0. \tag{27}$$

References

[1] R. Dutta, L. Sun, and D. Pack, "A decentralized formation and network connectivity tracking controller for multiple unmanned systems," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 6, pp. 2206–2213, Nov. 2018.

[2] N. T. Hung, F. F. C. Rego, and A. M. Pascoal, "Cooperative distributed estimation and control of multiple autonomous vehicles for range-based underwater target localization and pursuit," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 4, pp. 1433–1447, Jul. 2022.

[3] B.-B. Hu, H.-T. Zhang, B. Liu, H. Meng, and G. Chen, "Distributed surrounding control of multiple unmanned surface vessels with varying interconnection topologies," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 1, pp. 400–407, Jan. 2022.

[4] A. Carron, F. Seccamonte, C. Ruch, E. Frazzoli, and M. N. Zeilinger, "Scalable model predictive control for autonomous mobility-on-demand systems," *IEEE Trans. Control Syst. Technol.*, vol. 29, no. 2, pp. 635–644, Mar. 2021.

[5] A. Dorri, S. S. Kanhere, and R. Jurdak, "Multi-agent systems: A survey," *IEEE Access*, vol. 6, pp. 28573–28593, 2018.

[6] P. Shi and B. Yan, "A survey on intelligent control for multiagent systems," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 1, pp. 161–175, Jan. 2021.

[7] Y. Rizk, M. Awad, and E. W. Tunstel, "Decision making in multiagent systems: A survey," *IEEE Trans. Cognit. Develop. Syst.*, vol. 10, no. 3, pp. 514–529, Sep. 2018.

[8] R. A. Murphey, *Target-Based Weapon Target Assignment Problems*. Boston, MA, USA: Springer, 2000, pp. 39–53, doi: 10.1007/978-1-4757-3155-2_3.

[9] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer Programming*, vol. 271. Cham, Switzerland: Springer, 2014.

[10] J. R. Marden, H. P. Young, G. Arslan, and J. S. Shamma, "Payoff-based dynamics for multiplayer weakly acyclic games," *SIAM J. Control Optim.*, vol. 48, no. 1, pp. 373–396, Jan. 2009.

[11] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation," *Games Econ. Behav.*, vol. 75, no. 2, pp. 788–808, Jul. 2012.

[12] G. Arslan, J. R. Marden, and J. S. Shamma, "Autonomous vehicle-target assignment: A game-theoretical formulation," *J. Dyn. Syst., Meas., Control*, vol. 129, no. 5, pp. 584–596, Sep. 2007.

[13] C. Eksin and A. Ribeiro, "Distributed fictitious play for multiagent systems in uncertain environments," *IEEE Trans. Autom. Control*, vol. 63, no. 4, pp. 1177–1184, Apr. 2018.

[14] B. Swenson, C. Eksin, S. Kar, and A. Ribeiro, "Distributed inertial best-response dynamics," *IEEE Trans. Autom. Control*, vol. 63, no. 12, pp. 4294–4300, Dec. 2018.

[15] J. Song and S. Gupta, "CARE: Cooperative autonomy for resilience and efficiency of robot teams for complete coverage of unknown environments under robot failures," *Auto. Robots*, vol. 44, nos. 3–4, pp. 647–671, Mar. 2020.

[16] S. Aydın and C. Eksin, "Decentralized inertial best-response with voluntary and limited communication in random communication networks," *Automatica*, vol. 145, Nov. 2022, Art. no. 110566.

[17] H.-L. Choi, L. Brunet, and J. P. How, "Consensus-based decentralized auctions for robust task allocation," *IEEE Trans. Robot.*, vol. 25, no. 4, pp. 912–926, Aug. 2009.

[18] L. Luo, N. Chakraborty, and K. Sycara, "Provably-good distributed algorithm for constrained multi-robot task assignment for grouped tasks," *IEEE Trans. Robot.*, vol. 31, no. 1, pp. 19–30, Feb. 2015.

[19] L. Luo, N. Chakraborty, and K. Sycara, "Distributed algorithms for multirobot task assignment with task deadline constraints," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 3, pp. 876–888, Jul. 2015.

[20] M. Otte, M. J. Kuhlman, and D. Sofge, "Auctions for multi-robot task allocation in communication limited environments," *Auto. Robots*, vol. 44, nos. 3–4, pp. 547–584, Mar. 2020.

[21] L. Liu and D. A. Shell, "Large-scale multi-robot task allocation via dynamic partitioning and distribution," *Auto. Robots*, vol. 33, no. 3, pp. 291–307, Oct. 2012.

[22] P. Mazdin and B. Rinner, "Distributed and communication-aware coalition formation and task assignment in multi-robot systems," *IEEE Access*, vol. 9, pp. 35088–35100, 2021.

[23] S. Chopra, G. Notarstefano, M. Rice, and M. Egerstedt, "A distributed version of the Hungarian method for multirobot assignment," *IEEE Trans. Robot.*, vol. 33, no. 4, pp. 932–947, Aug. 2017.

[24] S. L. Smith and F. Bullo, "Monotonic target assignment for robotic networks," *IEEE Trans. Autom. Control*, vol. 54, no. 9, pp. 2042–2057, Sep. 2009.

[25] J. S. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria," *IEEE Trans. Autom. Control*, vol. 50, no. 3, pp. 312–327, Mar. 2005.

[26] B. Swenson, S. Kar, and J. Xavier, "Empirical centroid fictitious play: An approach for distributed learning in multi-agent games," *IEEE Trans. Signal Process.*, vol. 63, no. 15, pp. 3888–3901, Aug. 2015.

[27] F. Salehisadaghiani, W. Shi, and L. Pavel, "Distributed Nash equilibrium seeking under partial-decision information via the alternating direction method of multipliers," *Automatica*, vol. 103, pp. 27–35, May 2019.

[28] F. Parise, B. Gentile, and J. Lygeros, "A distributed algorithm for almost-Nash equilibria of average aggregative games with coupling constraints," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 2, pp. 770–782, Jun. 2020.

[29] D. Monderer and L. S. Shapley, "Fictitious play property for games with identical interests," *J. Econ. Theory*, vol. 68, no. 1, pp. 258–265, Jan. 1996.

[30] Y. Chen, R. S. Blum, and B. M. Sadler, "Ordering for communication-efficient quickest change detection in a decomposable graphical model," *IEEE Trans. Signal Process.*, vol. 69, pp. 4710–4723, 2021.

[31] T. Chen, G. Giannakis, T. Sun, and W. Yin, "LAG: Lazily aggregated gradient for communication-efficient distributed learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 5050–5060.

[32] Y. Liu, W. Xu, G. Wu, Z. Tian, and Q. Ling, "Communication-censored ADMM for decentralized consensus optimization," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2565–2579, May 2019.

[33] J. Fink, A. Ribeiro, and V. Kumar, "Robust control for mobility and wireless communication in cyber–physical systems with application to robot teams," *Proc. IEEE*, vol. 100, no. 1, pp. 164–178, Jan. 2012.

[34] J. Stephan, J. Fink, V. Kumar, and A. Ribeiro, "Concurrent control of mobility and communication in multirobot systems," *IEEE Trans. Robot.*, vol. 33, no. 5, pp. 1248–1254, Oct. 2017.

[35] Y. Kantaros and M. M. Zavlanos, "Distributed communication-aware coverage control by mobile sensor networks," *Automatica*, vol. 63, pp. 209–220, Jan. 2016.

[36] Y. Yan and Y. Mostofi, "To go or not to go: On energy-aware and communication-aware robotic operation," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 3, pp. 218–231, Sep. 2014.

[37] A. Muralidharan and Y. Mostofi, "Communication-aware robotics: Exploiting motion for communication," *Annu. Rev. Control, Robot., Auto. Syst.*, vol. 4, no. 1, pp. 115–139, May 2021.
[38] Y. Kantaros, M. Guo, and M. M. Zavlanos, "Temporal logic task planning and intermittent connectivity control of mobile robot networks," *IEEE Trans. Autom. Control*, vol. 64, no. 10, pp. 4105–4120, Oct. 2019.
[39] R. Khodayi-Mehr, Y. Kantaros, and M. M. Zavlanos, "Distributed state estimation using intermittently connected robot networks," *IEEE Trans. Robot.*, vol. 35, no. 3, pp. 709–724, Jun. 2019.
[40] S. Aydin and C. Eksin, "Communication censoring in decentralized fictitious play for the target assignment problem," in *Proc. IEEE Conf. Control Technol. Appl. (CCTA)*, Aug. 2020, pp. 334–339.
[41] S. Aydin and C. Eksin, "Decentralized fictitious play with voluntary communication in random communication networks," in *Proc. 59th IEEE Conf. Decis. Control (CDC)*, Dec. 2020, pp. 337–342.
[42] M. Voorneveld, "Best-response potential games," *Econ. Lett.*, vol. 66, no. 3, pp. 289–295, Mar. 2000.
[43] A. Y. Yazicioglu, M. Egerstedt, and J. S. Shamma, "Communication-free distributed coverage for networked systems," *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 3, pp. 499–510, Sep. 2017.
[44] S. Kalam, M. Gani, and L. Seneviratne, "A game-theoretic approach to non-cooperative target assignment," *Robot. Auto. Syst.*, vol. 58, no. 8, pp. 955–962, Aug. 2010.
[45] J. Fink, A. Ribeiro, and V. Kumar, "Robust control of mobility and communications in autonomous robot teams," *IEEE Access*, vol. 1, pp. 290–309, 2013.
[46] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[47] W. Wang, S. L. Capitaneanu, D. Marinca, and E.-S. Lohan, "Comparative analysis of channel models for industrial IoT wireless communication," *IEEE Access*, vol. 7, pp. 91627–91640, 2019.
[48] S. Aydın and C. Eksin, "Decentralized inertial best-response with voluntary and limited communication in random communication networks," 2021, *arXiv:2106.07079*.
[49] S. Chubanov, "A polynomial-time descent method for separable convex optimization problems with linear constraints," *SIAM J. Optim.*, vol. 26, no. 1, pp. 856–889, Jan. 2016.
[50] S. Wilson et al., "The Robotarium: Globally impactful opportunities, challenges, and lessons learned in remote-access, distributed control of multirobot systems," *IEEE Control Syst. Mag.*, vol. 40, no. 1, pp. 26–44, Feb. 2020.

Sarper Aydın (Member, IEEE) received the B.Sc. degree in industrial engineering from Bilkent University, Ankara, Turkey, in 2017, and the Ph.D. degree from Lehigh University, Bethlehem, PA, USA, in 2019. He is currently pursuing the Ph.D. degree with the Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA.

During the B.Sc. degree, he spent one semester as an Exchange Student at the Queensland University of Technology, Brisbane, QLD, Australia. He joined Texas A&M University in 2019. His current research interests include decentralized control of multiagent systems and game-theoretic learning with applications to assignment problems in autonomous robot teams.

Ceyhun Eksin (Member, IEEE) received the B.S. degree in control engineering from Istanbul Technical University, Istanbul, Turkey, in 2005, the M.S. degree in industrial engineering from Boğaziçi University, Istanbul, in 2008, the M.A. degree in statistics from the Wharton Statistics Department, University of Pennsylvania, Philadelphia, PA, USA, in 2015, and the Ph.D. degree in electrical and systems engineering from the Department of Electrical and Systems Engineering, University of Pennsylvania, in 2015.

He was a Post-Doctoral Researcher jointly affiliated with the School of Biological Sciences and the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. He is currently an Assistant Professor with the Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA. His research interests include networks, game theory, control theory, and distributed optimization. His current research interests include game-theoretic learning and decentralized optimization with applications to autonomous teams, epidemics, and energy systems.

Dr. Eksin was a recipient of the NSF CAREER Award in 2023.