

# Epistemic Prediction and Planning with Implicit Coordination for Multi-Robot Teams in Communication Restricted Environments

Lauren Bramblett, Shijie Gao, and Nicola Bezzo

**Abstract**—In communication restricted environments, a multi-robot system can be deployed to either: i) maintain constant communication but potentially sacrifice operational efficiency due to proximity constraints or ii) allow disconnections to increase environmental coverage efficiency, challenges on how, when, and where to reconnect (rendezvous problem). In this work we tackle the latter problem and notice that most state-of-the-art methods assume that robots will be able to execute a predetermined plan; however system failures and changes in environmental conditions can cause the robots to deviate from the plan with cascading effects across the multi-robot system. This paper proposes a coordinated epistemic prediction and planning framework to achieve consensus without communicating for exploration and coverage, task discovery and completion, and rendezvous applications. Dynamic epistemic logic is the principal component implemented to allow robots to propagate belief states and empathize with other agents. Propagation of belief states and subsequent coverage of the environment is achieved via a frontier-based method within an artificial physics-based framework. The proposed framework is validated with both simulations and experiments with unmanned ground vehicles in various cluttered environments.

## I. INTRODUCTION

Multi-robot systems (MRS) have the potential to assist in many safety-critical applications such as search and rescue, military intelligence and surveillance, and inspection operations where it may be hazardous and costly to deploy humans. Looking to the state-of-the-art, we note that most MRS research assumes constant communication between robots [1]–[3]. However, within the aforementioned application space, long-range communication is often unreliable or unavailable. Humans adequately cope with such problems, performing these tasks collaboratively by extrapolating and empathizing with what other actors might believe if the local plan must change at run-time. This subconscious process can be modally represented as epistemic planning, computing and reasoning about multiple predictions and actions while accounting for a priori beliefs, current observations, and other actors’ sensing and mobility capabilities.

In this work, we insist that if the robots in a team could perform similar reasoning without communication then we could relax the typical connectivity constraints, while increasing autonomy (i.e., decrease human intervention) and mission performance (i.e., more coverage, faster task discovery and completion). For this reason, we propose a novel epistemic planning framework for multi-robot systems, allowing each robot to cooperate without constant communication by reasoning about teammates’ performance and states, through the propagation of beliefs states about each others. As a reader can note, calculating a distributed

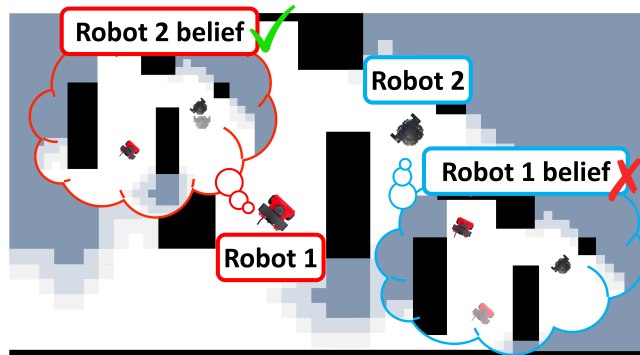


Fig. 1. Pictorial depiction of the problem. The proposed framework enables a robot to reason from other agents’ perspectives as it experiences a behavior change or observes that another robot is not where expected.

plan for coverage while accounting for any combination of robot system failures, changes in the environments, or deviations is intractable. Instead, constructing a finite set of possibilities and implementing a reasoning framework for each robot can reduce computational complexity and allow for more efficient operations. Thus, we introduce a coordinated epistemic prediction and planning method in which a robot propagates a finite set of *belief* states representing possible states of other agents in the system and *empathy* states representing a finite set of possible states from other agents’ perspectives. Subsequently, using epistemic planning, we can formulate a consensus strategy such that every distributed belief in the system achieves consensus. For example, consider Fig. 1 where two robots are canvassing an environment. During disconnection, Robot 1 maintains a set of possible (belief) states for Robot 2 and also a set of (empathy) states that Robot 2 might believe about Robot 1. Once Robot 2 experiences a failure, it tracks another state in its empathy set. We reason that though Robot 1 holds a false belief about Robot 2’s state, there exists an epistemic strategy that can allow robot 1 to find robot 2 (i.e., updating its belief after observing robot 1’s believed state).

The contribution of our approach is two-fold: i) an epistemic planning formulation using dynamic epistemic logic, formalizing beliefs and knowledge for robot control and ii) a generalized task assignment and artificial potential field-based model for belief propagation and coverage of an environment with considerations for connectivity constraints and team member dynamics.

## II. RELATED WORK

Heterogeneous multi-robot exploration, foraging, and coverage have been widely studied in robotics literature [4]–[6]. Recent works consider communication restricted or intermittent connectivity by modeling ways to maintain connectivity while exploring [1], to account for momentary disconnection

Lauren Bramblett, Shijie Gao, and Nicola Bezzo are with the Departments of Engineering Systems and Environment and Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904, USA. Email: {qbr5kx sg9dn nb6be}@virginia.edu

[7], or dropping intermediary notifications to other robots [8]. Authors in [9] use a decentralized Markov decision process to estimate future locations of robots when inter-robot messages are delayed at a stochastic rate. Several works have also included system failures or disturbances in multi-robot exploration policies, such as in the work [10]; however, the policy assumes robots are able to communicate these disruptions. In most approaches, the execution policy is static and communication is a constraint that either needs to be satisfied for all time or at a defined location [11], [12].

A separate, but related, field to multi-robot coverage is multi-robot task allocation (MRTA). MRTA is solved by uniquely assigning a subset of robots to optimize the completion of an objective [13]. Authors in [14] include connection limitations and allocate tasks to connected robots using a bundling algorithm. Another example in [15] allocated targets to individual teams and plans rendezvous with team members to reduce uncertainty of targets over time.

Though recent works in multi-robot task allocation and coverage have included realistic constraints, there is little consideration for the combination of prolonged disconnection and system failures. In our previous work [16], we define rendezvous points at known locations to coordinate roles for any events during exploration; however, back-tracking to this location reduced efficiency of exploration. In contrast, this work applies dynamic epistemic logic (DEL) [17] to allow a robot to reason about beliefs among actors in a multi-robot system while disconnected and converge to a dynamic rendezvous location. While classically used to describe how knowledge and information changes for players in a game, DEL is not a new concept in robotics. Using robot and human actors, the framework in [18] recreates the Sally-Anne psychological test where a robot must reason about the human's beliefs. Typical multi-robot applications use DEL to solve for a cooperative set of actions in multi-player games [19]. We extend DEL for a realistic multi-robot application allowing each robot to reason about the system's state considering system failures, task discovery, and partially known environments.

### III. PRELIMINARIES

#### A. Notation & Communication

Let us consider a multi-robot system of  $N_a$  robots in the set  $\mathcal{A}$ , noting that our approach is suitable for heterogeneous robots of differing capabilities (e.g., dynamics, sensing, etc.). The system's connectivity graph is denoted as  $G = (\mathcal{A}, \mathcal{E})$  where the set  $\mathcal{E} \subset \mathcal{A} \times \mathcal{A}$  represents edge connections between robots and an edge  $(i, j) \in \mathcal{E}$  indicates that robots  $i$  and  $j$  are connected. Additionally,  $N_t$  tasks in the set  $\mathcal{T}$  are located in unknown positions within the operating environment. We assume the tasks are stationary and completed once any robot  $i \in \mathcal{A}$  navigates within a radius  $r_t > 0$ , considering that the number of tasks  $N_t$  is initially unknown to the robots.

The robots are assigned to search for the tasks in an environment that is partitioned into  $N_m$  cells, which we define as an occupancy map  $\mathcal{M} \subseteq \mathbb{R}^2$ . When robots navigate to observe unexplored cells  $\mathcal{M}_u \subseteq \mathcal{M}$ ,  $\mathcal{M}$  is updated using Recursive Bayesian estimation [20], though any method can be used. Subsequently, we define the frontier set  $\mathcal{F}$  as the set of explored cells adjacent to unknown cells.

#### B. Epistemic Logic

In this work, epistemic and doxastic logic [21] is used to model distributed knowledge and reasoning for non-catastrophic system changes during disconnectivity. To propagate uncertain states of robots, we define  $N_b$  beliefs in the set  $\mathcal{B}$  as the behaviors likely to occur to agents during operation. The set  $\mathcal{P} = \{\mathcal{P}_1, \dots, \mathcal{P}_{N_a}\}$  holds the distributed beliefs of all agents, where an element in  $\mathcal{P}_i$  represents possible states from an agent  $i$ 's perspective of robots  $j \in \mathcal{A}$ .  $\Psi$  is a set of functions that describe the current state of the system. For this application, the epistemic language,  $\mathcal{L}(\Psi, \mathcal{P}, \mathcal{A})$ , is obtained as follows in Backus-Naur form [22]:

$$\phi ::= H(\omega) \mid i < j \mid \phi \wedge \phi \mid \neg \phi \mid K_i \phi \mid B_i \phi$$

where  $i, j \in \mathcal{A}$ ,  $H \in \Psi$  is a function to describe a system state, and  $\omega$  broadly indicates function arguments.  $\neg \phi$  and  $\phi \wedge \phi$  denote that propositions can be negated and form logical conjunctions.  $B_i \phi$  and  $K_i \phi$  are interpreted as "agent  $i$  believes  $\phi$ " and "agent  $i$  knows  $\phi$ ."  $i < j$  is an observability atom that reads "agent  $i$  is within communication range of agent  $j$ ".

A pointed Kripke model ( $M$ ) represents an epistemic model [23] consisting of the set of possible states of the system referred to as worlds, accessibility relations ( $R$ ) between worlds,  $w$ , that are possible for an agent, and a valuation ( $V$ ) that labels true propositions in each world. Dynamic epistemic logic formalizes Kripke model transforms given an event,  $a$ . In this work, we use dynamic epistemic logic and Kripke models to modally represent a distributed strategy using logical belief-based reasoning.

### IV. PROBLEM FORMULATION

In this paper, we consider a scenario in which a multi-robot system must coordinate in a decentralized fashion to efficiently search for tasks at unknown locations in a communication restricted, partially-known environment. There are several challenges that arise to allow efficient and cooperative behavior given limited communication including: 1) how to efficiently explore, search, and cover a partially known environment while remaining disconnected for extended periods of time and 2) how to properly plan to take into account uncertainties that could happen during disconnection that lead to different behaviors of the robots in the systems. Formally, we define this problem as:

**Problem 1 (Communication restricted coverage):** Find a distributed policy to enable a multi-robot system to quickly perform cooperative search of an environment for tasks with intermittent communication accounting for uncertainties. The policy should consider minimizing mission time while enabling periodic information sharing to cooperatively allocate portions of the environment if a robot's behavior changes.

### V. APPROACH

In this section, we present the approach for the coordinated epistemic prediction and planning framework which propagates belief and empathy states to inform frontier assignment and robot control, all while considering failures, task discovery, and unknown obstacles. For ease of discussion let us consider two robots  $i$  and  $j$ . From robot  $i$ 's perspective, a *belief state*,  $p_{i,j,b} \in \mathcal{P}_i$ , represents a possible state of a robot  $j$  and an *empathy state*,  $p_{ii,b} \in \mathcal{P}_i$ , describes robot  $i$ 's belief

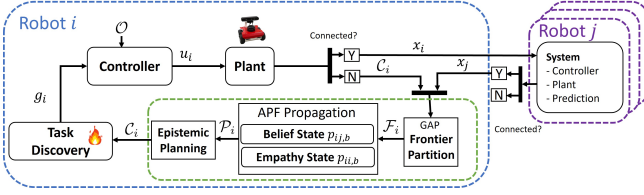


Fig. 2. Diagram of the proposed approach. The contributions of this paper are within the green box.

of robot  $j$ 's belief about robot  $i$ 's state. With this knowledge, robot  $i$  predicts and tracks empathy states to ensure that a robot  $j$  holds one true belief of the state of robot  $i$ . The diagram in Fig. 2 summarizes this architecture.

As shown in Fig. 2, the robot  $i$  initially assesses whether communication is successful with a robot  $j$ . If communication is successful, robot  $i$  uses its current state  $x_i$  and the state of robot  $j$ ,  $x_j$ , to partition its frontiers using a generalized assignment problem (GAP) [24] and to predict future states of robot  $j$  using an APF method. When connected, epistemic planning is reduced to direct communication of states. If the robots disconnect, a common belief set,  $C_i$ , acts as the state for any robot  $j \in \mathcal{A}$  from  $i$ 's perspective. Predictions for these belief and empathy states are accomplished using the same GAP and APF methods. A robot  $i$  then uses these predicted states to plan considering its belief about robot  $j$ .

In both connected and disconnected conditions, the robot's objective is to search for tasks. If connected and a task is discovered, the robots bid on and accomplish tasks. If disconnected, the robots will deviate to accomplish the task and subsequently continue to track its empathy state.

In the following sections, we lay out the key components of the planner including: i) belief propagation, ii) MRS coverage assignments, iii) epistemic planning for belief consensus, and iv) obstacle avoidance and task discovery.

### A. Belief & Empathy Propagation

In our coordinated epistemic prediction and planning framework, the robots propagate belief and empathy states for all robots in the multi-robot system. This allows a robot  $i$  to plan according to its belief of other robots and reason about what other robots' expect robot  $i$  to accomplish while disconnected. As previously noted, to account for uncertainties over long periods of disconnection, it is important to have a finite number of these states. With this goal in mind, we define a finite set of particles,  $\mathcal{P}_i$ , to represent these belief and empathy states for the  $i^{\text{th}}$  robot:

$$\mathcal{P}_i = \{p_{i,j,b} \mid \forall j \in \mathcal{A}, \forall b \in \mathcal{B}\}. \quad (1)$$

The  $i^{\text{th}}$  robot defines its empathy particles as  $\mathcal{P}_i^e = \{p_{i,b} \mid \forall b \in \mathcal{B}\}$  and its belief particles about other robots as  $\mathcal{P}_i^r = \{p_{i,j,b} \mid \forall j \in \mathcal{A} \setminus \{i\}, \forall b \in \mathcal{B}\}$  where  $\mathcal{P}_i = \mathcal{P}_i^e \cup \mathcal{P}_i^r$ . For each robot  $j \in \mathcal{A}$ , the robot  $i$  orders its belief and empathy particles 1 through  $N_b$  by likelihood of occurrence (i.e., from largest to smallest). The order is initialized prior to deployment and each robot  $i$  initially tracks its first empathy particles.

While not in communication range of other robots, each robot  $i$  has a common belief about each robot  $j$  and itself. We define robot  $i$ 's common belief as  $C_i \subseteq \mathcal{P}_i$  and refer to it as the *common belief set*. All robots track their first empathy particle upon disconnection,  $C_i = \{p_{i,j,1} \mid \forall j \in \mathcal{A}\}$ .

If a robot experiences a failure, choosing which of its next empathy particles to track is nontrivial, but we assume each robot  $i$  is capable of computing the set of empathy states that are suitable to track, denoted by  $\mathcal{P}_i^f \subseteq \mathcal{P}_i^e$ . The robot chooses to track the particle in  $\mathcal{P}_i^f$  with the highest likelihood. If all robots are within communication range, the first particle becomes the robot's current state and subsequent particles are propagated based on the updated common belief.

Since the robot will be tracking an empathy particle, these states must propagate in a manner that allows the robot to safely and efficiently accomplish the coverage objective with considerations for intentional information sharing. Thus, we propagate the particles using an artificial potential field (APF) that leverages four main objectives: 1) attraction to frontier, 2) cooperative rendezvous, 3) obstacle avoidance, and 4) task completion. In this propagation method, the total force acting on particle  $p_{i,j,b}$  is formulated generally as:

$$F_{i,j,b}^{\text{total}} = \beta_1 F_{i,j,b}^1 + \beta_2 F_{i,j,b}^2 + \beta_3 F_{i,j,b}^3 + \beta_4 F_{i,j,b}^4 \quad (2)$$

considering  $\beta_n$  is a weighting coefficient for force  $F_{i,j,b}^n$  where each  $F_{i,j,b}^n$  corresponds to the  $n^{\text{th}}$  objective listed previously and will be discussed in detail. Local minima is avoided using an A\* path planner [25].

### B. Frontier Attraction

A frontier-based exploration method is proposed here due to its completeness and simplicity. To begin, the force  $F_{i,j,b}^1$  in (2) is an attraction to a frontier set  $\mathcal{F} \subset \mathcal{M}$ . However, a robot should only traverse unique portions of the environment to reduce redundancy and minimize completion time. So, a decentralized GAP assigns particles to a unique subset of  $\mathcal{F}$  using its belief of each robots' capabilities.

For particle  $p_{i,j,b}$ , we allocate frontiers based on the cost of assigning the particle or any other robot. Cost and binary assignment are denoted as  $\Lambda$  and  $\Gamma$ , respectively. The corresponding GAP is formulated as:

$$\begin{aligned} \mathcal{F}_{i,j,b} = \min & \sum_{k \in \mathcal{A}} \sum_{z \in \mathcal{F}} \lambda_{zk} \gamma_{zk} \\ \text{s.t.} & \sum_{k \in \mathcal{A}} \gamma_{zk} = 1, \quad \forall z \in \mathcal{F} \\ & \sum_{z \in \mathcal{F}} \gamma_{zk} \leq u, \quad \forall k \in \mathcal{A} \\ & \sum_{z \in \mathcal{F}} \gamma_{zk} \geq \ell, \quad \forall k \in \mathcal{A} \\ & \gamma_{zk} \in \{0, 1\}, \quad \forall k \in \mathcal{A}, \forall z \in \mathcal{F} \end{aligned} \quad (3)$$

where the elements  $\lambda_{zk} \in \mathbb{R}_{\geq 0}$  and  $\gamma_{zk}$  represent the  $z^{\text{th}}$  frontier and  $k^{\text{th}}$  particle in the matrices  $\Lambda$  and  $\Gamma$ , respectively. Cost generally refers to any traversal metric (i.e., energy, time, etc.) to a frontier point  $z \in \mathcal{F}$ . The variables  $u \in \mathbb{N}$  and  $\ell \in \mathbb{N}$  are the upper and lower bounds on the number of frontier points that can be assigned to any particle. Each robot  $i$  calculates the frontier assignment for each particle in the set  $\mathcal{P}_i$  and the assigned frontier set is denoted as  $\mathcal{F}_{i,j,b}$ .

Subsequently, to utilize the GAP solution, the force  $F_{i,j,b}^1$  controls the  $b^{\text{th}}$  particle  $p_{i,j,b}$  towards its assigned frontiers

$$F_{i,j,b}^1 = \frac{1}{|\mathcal{F}_{i,j,b}|} \sum_{z \in \mathcal{F}_{i,j,b}} \frac{s_z - p_{i,j,b}}{\|s_z - p_{i,j,b}\|^3} \quad (4)$$

where  $|\cdot|$  indicates the set's cardinality and the coordinate of a  $z^{\text{th}}$  frontier is designated as  $s_z$ . The force computed in (4) encourages particle motion to the unexplored regions of the environment  $\mathcal{M}_u$  based on their frontier assignment in (3).

### C. Epistemic Planning

Intentional information sharing allows an agent to communicate any environmental or capability changes with other robots. For this purpose, we introduce  $F_{ij,b}^2$  in (2) to control each particle  $p_{ij,b}$  according to robot  $i$ 's common belief set:

$$F_{ij,b}^2 = \varphi_i \sum_{k \in \mathcal{A}} \frac{c_k - p_{ij,b}}{\|c_k - p_{ij,b}\|^3} \quad (5)$$

$$\varphi_i = \begin{cases} -h(t_r, \tau), & t_r < \tau \\ h(t_r, \tau), & t_r \geq \tau \end{cases} \quad (6)$$

where  $c_k$  denotes the  $k^{\text{th}}$  element in  $C_i$ . Variable  $\tau$  is a time-based threshold for rendezvous and  $h : (t_r, \tau) \mapsto \mathbb{R}_{\geq 0}$  where  $t_r$  is the time lapsed since the last successful communication.

Fig. 3 shows the effect of  $\varphi_i$ . Given the partitioned frontier from (4), the robots' particles are incentivized to travel i) away from  $c_k$  when  $\varphi_i < 0$ , ii) towards its assigned frontier when  $\varphi_i = 0$ , or iii) towards  $c_k$  when  $\varphi_i > 0$ . We denote the line between common belief particles as the *anchor line*. In this way, (5) controls all of robot  $i$ 's particles to all beliefs in  $C_i$  when  $t_r > \tau$ . This is formalized in the following lemma:

**Lemma 1:** If  $F_{ij,b}^2 = 0$ , as  $t \rightarrow \infty$  all particles in the set  $\{p_{ij,b} \mid p_{ij,b} \in \mathcal{P}\}$  will converge to  $c_k, \forall k \in \mathcal{A}$ .

*Proof:* Given the  $\lim_{t \rightarrow \infty} F_{ij,b}^1 + F_{ij,b}^4 = 0$  once all area has been covered, the only force acting on each particle will be  $F_{ij,b}^2$ . Also,  $\varphi_i > 0$  since  $t_r > \tau$  when  $t \rightarrow \infty$ , all particles in the set  $\{p_{ij,b} \mid p_{ij,b} \in \mathcal{P}_i\}$  will converge to  $c_k \in C_i$ . ■

Considering that the  $c_k \in \mathcal{P}_i$  is controlled via (5), all common belief states in  $C_i$  converge and so, all particles in  $\mathcal{P}_i$  converge to the same rendezvous location. Thus, it is imperative to ensure that while the robots are not communicating they can reach consensus such that

$$C_i \equiv C_j, \quad \forall i, j \in \mathcal{A}. \quad (7)$$

We coordinate this consensus using dynamic epistemic logic.

Referring to the previously established semantics for DEL in Sec. III, we introduce the set  $\Psi$  consisting of a binary and tertiary function, `track` and `anchor`, noting the argument for the  $i^{\text{th}}$  robot is denoted as  $A_i$  for readability in the epistemic model. The function `track`( $A_i, p_{ii,b}$ ) is read as “robot

$i$  is tracking empathy particle  $b$ ” and `anchor`( $A_i, p_{ii,b}, p_{ij,b}$ ) is read as “robot  $i$  is using belief particle  $p_{ij,b}$  and empathy particle  $p_{ii,b}$  as its common belief.”

Since a robot's failures can affect a robot's capabilities in disparate ways and at any time, we formulate a strategy for rendezvous that accounts for all possible combinations of failures. Given that robots are connected given a communication graph  $G$ , we also assume particles can observe each other similarly (e.g., based on range, line-of-sight), denoted logically as  $p_{ij,b} \prec p_{ik,b}$ . If two common belief particles have observed each other, we know that one of four events have occurred for two robots: i)  $A_i$  has observed that  $A_j$  is not tracking the common belief particle, ii)  $A_j$  has observed that  $A_i$  is not tracking the common belief particle, iii) neither agents observe either common belief particle, or iv) both agents communicate. In events (i)-(iii), neither robot knows the true system state since the robots did not communicate.

Given that our current common belief is the  $b^{\text{th}}$  particle and all particles are within observation range, we define the consensus-based policy sequence formally as:

$$\begin{aligned} a_0 &= p_{ij,b} \prec p_{ji,b} \quad \forall i, j \in \mathcal{A} \\ a_1 &= \text{anchor}(A_j, p_{ij,b+1} \quad \forall j \in \mathcal{A}) \\ a_s &= a_0 \otimes a_1 \end{aligned}$$

such that the frame-policy update is:

$$f_{b+1} = f_b \otimes a_s \quad (8)$$

where  $b \in \mathcal{B} \setminus \{N_b\}$  and  $\otimes$  is a modal product. This strategy is used until all robots communicate at which time the particles are reset to the robots' poses and dynamics are updated.

**Example:** Consider  $f_0, f'_0, f_1$  in the Kripke model shown in Fig. 4 where the true world is denoted by the black vertex, edges represent accessibility relation ( $R$ ), and new propositions ( $V$ ) are denoted in each worlds' sets. From the perspective of robot  $A_1$ , each robot initially tracks particles  $p_{11,1}, p_{12,1}$  and  $p_{13,1}$ . All robots propagates four possible belief states for each teammate and four empathy states of itself. We denote robot  $A_1$ 's knowledge and belief of his own state as  $K_1 \text{ track}(A_1, p_{11,1})$ . Similarly, robot  $A_1$ 's knowledge and belief about his teammate's state is represented as  $B_1 \text{ track}(A_j, p_{1j,1}), j \in \{2, 3\}$  and empathy from robot  $A_1$ 's perspective is shown as  $B_1 B_j \text{ track}(A_1, p_{j1,1}), j \in \{2, 3\}$  and can be read as “robot  $A_1$  believes that robot  $j$  believes that robot  $A_1$  is tracking particle 1”. We denote the initial common belief propositions for all three agents as `anchor`( $A_i, p_{i1,1}, p_{i2,1}, p_{i3,1}$ ),  $i \in 1, 2, 3$ .

After disconnection (as shown in  $f'_0$  in Fig. 4), the robots can no longer assume to have knowledge about the true state of the system. In this scenario,  $A_3$  experiences a failure such that `track`( $A_3, p_{33,2}$ ) and updates its common belief ( $w_1$ ). However,  $A_3$  subsequently reasons that  $A_1$  and  $A_2$  hold a false belief that  $A_3$  is still tracking the first particle. Thus,  $A_3$  reverts its common belief in  $w_2$  to mirror the initial proposition such that `anchor`( $A_i, p_{i1,1}, p_{i2,1}, p_{i3,1}$ ).

When  $A_1 \prec p_{13,1}$  in  $f'_0$ ,  $A_1$  and  $A_2$  are are connected and so  $A_1$  relays that  $A_1 \prec p_{13,1} \wedge A_1 \prec \neg A_3$ , reasoning that  $A_3$  does not know that both agents are still tracking their respective first particle.  $A_3$  also knows that all three first particles are within communication range, but cannot observe any

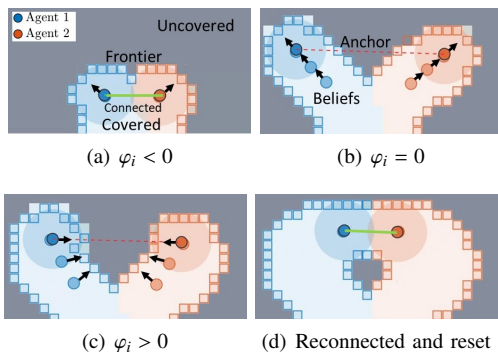


Fig. 3. Depiction of APF forces for each particle given different  $\varphi_i$  over time. The color of the anchor line indicates communication (green) or no communication (red).

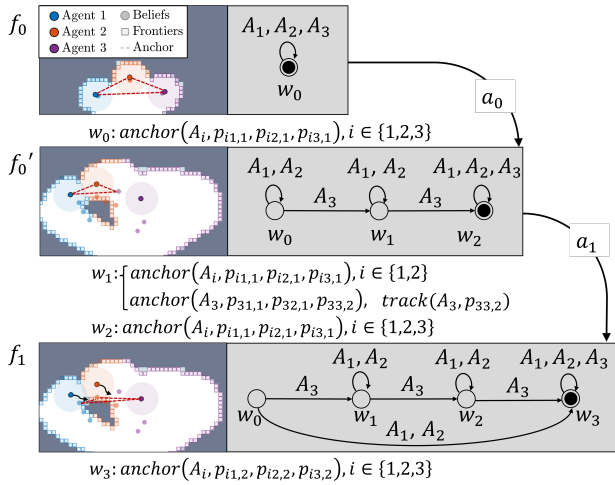


Fig. 4. Example scenario where  $A_3$  experiences a failure. The Kripke models are shown in the gray boxes with frame transitions denoted as  $a_n$ .

additional information. Thus, all three agents update their common belief such that  $\text{anchor}(A_i, p_{i1,2}, p_{i2,2}, p_{i3,2})$  and all particles begin converging to the updated common belief set.

#### D. Obstacle Avoidance & Task Completion

Finally, we consider the last two forces in (2). To avoid obstacles, force  $F_{ij,b}^3$  is formulated as:

$$F_{ij,b}^3 = \frac{1}{|\mathcal{O}|} \sum_{o \in \mathcal{O}} \frac{s_o - p_{ij,b}}{\|s_o - p_{ij,b}\|^3} \quad (9)$$

where  $s_o \in \mathcal{C}$  is the coordinate for an obstacle  $o \in \mathcal{O}$  in the environment. We note that  $\mathcal{O}$  is a set of the commonly known obstacles for all agents. For example, if a robot individually encounters an obstacle, but has not communicated its location to all teammates, the particles' motions are not affected by this new obstacle which is unknown to other robots.

For attraction to tasks, the force  $F_{ij,b}^4$  is only active when all agents are connected to propagate particles towards any commonly known tasks. These tasks are centrally auctioned and each robot submits a bid based on estimated traversal time to the task. Assigned tasks placed in a queue set  $\mathcal{Q}_i$ , and distributed to robot  $i$ 's particle task queues,  $\mathcal{Q}_{ij,b}$ . Attraction to the first task,  $\mathcal{Q}_{ij,b}[1]$ , in a particle's queue is formulated as:

$$F_{ij,b}^4 = \mathcal{Q}_{ij,b}[1] - p_{ij,b} \quad (10)$$

where the coefficients for  $F_{ij,b}^1$  and  $F_{ij,b}^4$  depend on the particle's queue  $\mathcal{Q}_i$ , such that:

$$\begin{cases} \beta_1 = 0 \text{ and } \beta_4 > 0, & \text{if } \mathcal{Q}_i \neq \emptyset, \\ \beta_1 > 0 \text{ and } \beta_4 = 0, & \text{otherwise.} \end{cases} \quad (11)$$

Once the task queue is empty, the particle force for tasks is set to zero and coverage resumes.

#### E. Particle Tracking

After particle propagation has occurred, a robot must predict and track its empathy particle considering the possibility of a new obstacle or discovering a task while disconnected. Though any constrained tracking algorithm can be employed, we use a nonlinear receding horizon controller (RHC) to minimize the distance to the particle while maintaining a radius from any obstacle [26]. Additionally, if the team is

disconnected and a robot discovers a task within its observing radius  $r_o$ , the tasks are re-indexed by monotonically increasing cost and placed in the set  $\mathcal{Q}_i$ . The controller's goal is updated to the first queued task location. Once the robot has traveled within the completion radius,  $r_r$ , the robot continues tracking its respective particle and the task is removed from the robot's queue. An example for both obstacle avoidance and task discovery are shown in Fig. 5.

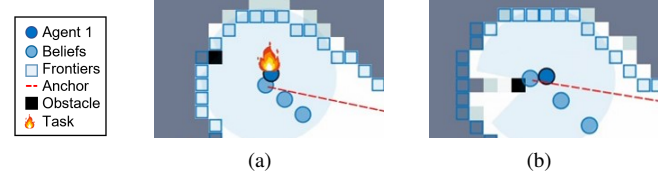


Fig. 5. Examples of necessary deviation. In (a) the robot spots a task and updates its RHC goal. In (b) the robot must avoid a discovered obstacle while minimizing distance to the particle.

## VI. SIMULATIONS

In this section, we provide results and comparisons from MATLAB simulations with our approach implemented on a two robot team. Simulations were performed on 15 random  $50\text{m} \times 50\text{m}$  environments with 5-15 initially unknown obstacles and a maximum of seven tasks. The robots start by assuming that the environment has no obstacles and do not know the location of the tasks. We compare the results between: 1) no failures, 2) one failure by one robot, and 3) one failure by each robot at random times.

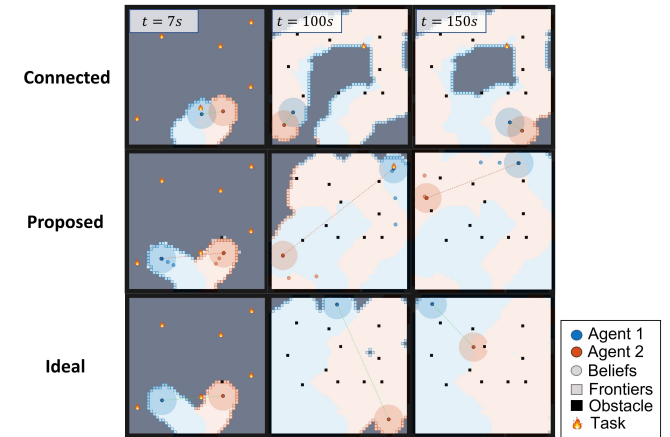


Fig. 6. Snapshots of example simulations. The robots experience a failure each at random times with 11 unknown obstacles and 7 tasks.

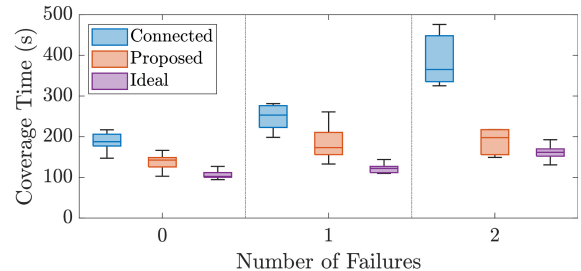


Fig. 7. Comparison between methods given 0, 1, and 2 failures.

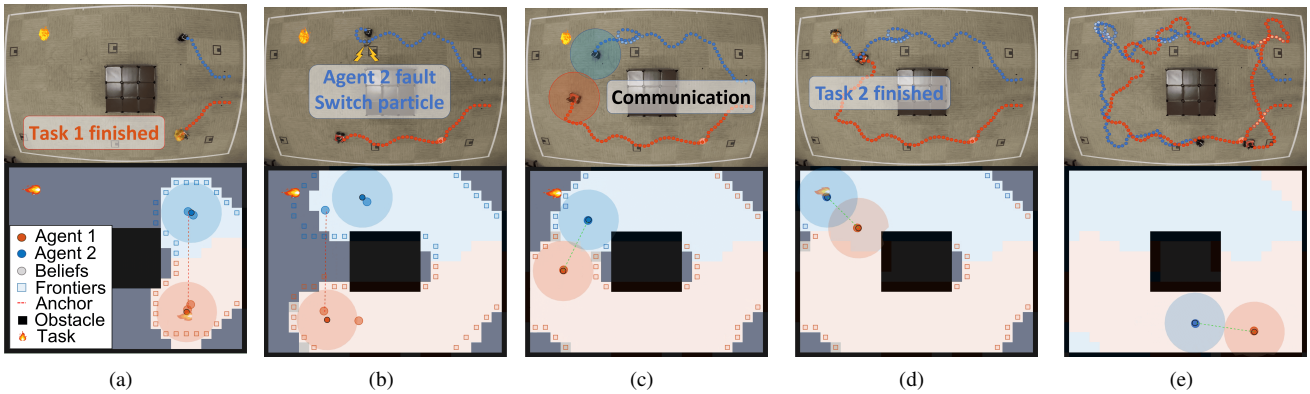


Fig. 8. Snapshots and results of an experimental case study.

Each robot propagates three particles traveling at 2m/s, 1m/s and 0.5m/s. A failure can cause the robot to track either the second or third particle. The particles propagate according to these failure velocities and the maximum communication range is 10m from the center of the robot.

The proposed approach is compared against two other methods. The first method applies a constant connectivity constraint, not allowing agents to travel outside of a 10m communication range. The second method assumes ideal conditions, where robots can communicate across the entire environment. Both methods use an artificial potential field technique for controlling the robots towards uncovered regions and away from obstacles. In all methods, the initial maximum velocity is 2m/s, the simulated LiDAR range is 5m, and the robots' motion is modeled using unicycle dynamics. Fig. 6 shows an example at various time steps for the three methods, displaying the coverage disparity of the connected method versus our proposed method and showing the similarities between our method and the ideal method.

As shown in Fig. 7, the proposed method outperforms the fully connected method in all scenarios. Additionally, the median coverage time for the proposed method is similar to the ideal method, even with the communication limitation.

## VII. EXPERIMENTS

The proposed approach was also validated through laboratory experiments with a two-robot team. The team consists of a Husarion ROSbot 2.0 UGV and a Turtlebot3 Burger UGV using a Vicon motion capture system. The two-robot experiments effectively demonstrate all parts of the proposed approach, including intentional disconnections, searching, and rendezvousing behaviors. In all experiments, the UGVs start within communication range and are tasked to cover the environment and complete any discovered tasks.

Experiments were performed in a 4m×5.5m space containing convex obstacles considering, as a proof of concept, a sensing and communication range for each robot of 1m. Displayed in Fig. 8 are the results from the two-robot experiment in which the vehicles are required to search for and complete two tasks unknown a priori. The columns of Fig. 8 correspond to different instances within the experiment, and each row from top to bottom shows snapshots of the robots at different times throughout the experiment and the current map of the environment covered by the team. In Fig. 8(a), the UGVs start to cover the map in search of tasks and

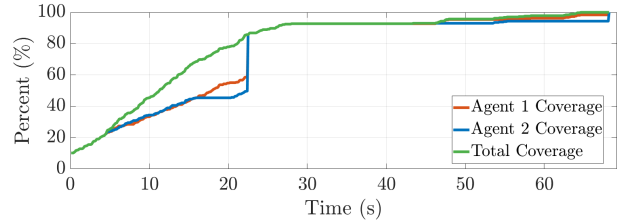


Fig. 9. Graphical coverage over time for the experimental case study.

disconnect. Robot 1 (ROSbot) finds a task and completes it. In Fig. 8(b), robot 2 (Turtlebot) experiences a fault and begins following the second empathy particle. In Fig. 8(c-d) the robots connect, share fault information, and bid on the discovered task. Robot 1 receives a larger share of the frontier as a result of Robot 2's failure and Robot 1 is assigned the task based on proximity. Finally, once all tasks are completed and no frontiers remain, the agents rendezvous and the final results are shown in Fig. 8(e). Final coverage over time for each robot is shown in Fig. 9. More lab experiments with two-robots are included in the supplementary material.

## VIII. CONCLUSION & FUTURE WORK

In this work, we have presented a novel framework for multi-robot systems to use epistemic planning to propagate and behave according to a set of beliefs. The proposed method promotes disconnection using frontier-based exploration and includes a dynamic rendezvous approach to reconnect and share data among the multi-robot system. The extensive simulations and experiment results show the validity, applicability, generality of the proposed method. Using this framework, we also demonstrate improved task completion and coverage time of partially known environments with respect to standard coverage methods. From here, future theoretical work includes addressing the challenges of dynamic task lengths and optimal strategies for intentional information sharing in complex or unknown environments. Further modeling of epistemic planning using epistemic MDPs to reach a more informed consensus for faster information dissemination is also on our agenda. Additionally, we plan to apply this framework in outdoor experiments using robots equipped with short-range communication devices.

## IX. ACKNOWLEDGEMENT

This work is based on research sponsored by Northrop Grumman through the University Basic Research Program and DARPA under Contract No. FA8750-18-C-0090.

## REFERENCES

- [1] B. Capelli and L. Sabattini, "Connectivity maintenance: Global and optimized approach through control barrier functions," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 5590–5596.
- [2] A. Hussein, M. Adel, M. Bakr, O. M. Shehata, and A. Khamis, "Multi-robot task allocation for search and rescue missions," in *Journal of Physics: Conference Series*, vol. 570, no. 5. IOP Publishing, 2014, p. 052006.
- [3] D. Morilla-Cabello, L. Bartolomei, L. Teixeira, E. Montijano, and M. Chli, "Sweep-your-map: Efficient coverage planning for aerial teams in large-scale environments," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10810–10817, 2022.
- [4] T. X. Lin, E. Yel, and N. Bezzo, "Energy-aware persistent control of heterogeneous robotic systems," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 2782–2787.
- [5] X. Zhou, W. Wang, T. Wang, Y. Lei, and F. Zhong, "Bayesian reinforcement learning for multi-robot decentralized patrolling in uncertain environments," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 11 691–11 703, 2019.
- [6] B. Reily, T. Mott, and H. Zhang, "Adaptation to team composition changes for heterogeneous multi-robot sensor coverage," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 9051–9057.
- [7] G. Best, M. Forrai, R. R. Mettu, and R. Fitch, "Planning-aware communication for decentralised multi-robot coordination," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1050–1057.
- [8] G. A. Cardona, D. Yanguas-Rojas, M. F. Arevalo-Castiblanco, and E. Mojica-Nava, "Ant-based multi-robot exploration in non-convex space without global-connectivity constraints," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 2065–2070.
- [9] L. Matignon, L. Jeanpierre, and A.-I. Mouaddib, "Coordinated multi-robot exploration under communication constraints using decentralized markov decision processes," in *Twenty-sixth AAAI conference on artificial intelligence*, 2012.
- [10] S. Al-Hussaini, J. M. Gregory, and S. K. Gupta, "Generation of context-dependent policies for robot rescue decision-making in multi-robot teams," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4317–4324.
- [11] K. Cesare, R. Skeele, S.-H. Yoo, Y. Zhang, and G. Hollinger, "Multi-uav exploration with limited communication and battery," in *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 2230–2235.
- [12] A. El Shenawy, K. Mohamed, and H. Harb, "Hdec-posmdps mrs exploration and fire searching based on iot cloud robotics," *International Journal of Automation and Computing*, vol. 17, no. 3, pp. 364–377, 2020.
- [13] J. Kim, C. Ju, and H. I. Son, "A multiplicatively weighted voronoi-based workspace partition for heterogeneous seeding robots," *Electronics*, vol. 9, no. 11, p. 1813, 2020.
- [14] J. Chen, X. Qing, F. Ye, K. Xiao, K. You, and Q. Sun, "Consensus-based bundle algorithm with local replanning for heterogeneous multi-uav system in the time-sensitive and dynamic environment," *The Journal of Supercomputing*, vol. 78, no. 2, pp. 1712–1740, 2022.
- [15] R. Khodayi-mehr, Y. Kantaros, and M. M. Zavlanos, "Distributed state estimation using intermittently connected robot networks," *IEEE Transactions on Robotics*, vol. 35, no. 3, pp. 709–724, 2019.
- [16] L. Bramblett, R. Peddi, and N. Bezzo, "Coordinated multi-agent exploration, rendezvous, & task allocation in unknown environments with limited connectivity," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (To Appear)*. IEEE, 2022.
- [17] H. Van Ditmarsch, W. van Der Hoek, and B. Kooi, *Dynamic epistemic logic*. Springer Science & Business Media, 2007, vol. 337.
- [18] T. Bolander, L. Dissing, and N. Herrmann, "Del-based epistemic planning for human-robot collaboration: Theory and implementation," in *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning*, vol. 18, no. 1, 2021, pp. 120–129.
- [19] B. Maubert, S. Pinchinat, F. Schwarzentruber, and S. Stranieri, "Concurrent games in dynamic epistemic logic," in *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, 2021, pp. 1877–1883.
- [20] A. Asgharivaskasi and N. Atanasov, "Active bayesian multi-class mapping from range and semantic segmentation observations," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 1–7.
- [21] R. Fagin, J. Y. Halpern, Y. Moses, and M. Vardi, *Reasoning about knowledge*. MIT press, 2004.
- [22] D. E. Knuth, "Backus normal form vs. backus naur form," *Communications of the ACM*, vol. 7, no. 12, pp. 735–736, 1964.
- [23] M. C. Browne, E. M. Clarke, and O. Grumberg, "Characterizing finite kripke structures in propositional temporal logic," *Theoretical computer science*, vol. 59, no. 1-2, pp. 115–131, 1988.
- [24] T. Öncan, "A survey of the generalized assignment problem and its applications," *INFOR: Information Systems and Operational Research*, vol. 45, no. 3, pp. 123–141, 2007.
- [25] M. Juliá, A. Gil, L. Payá, and O. Reinoso, "Local minima detection in potential field based cooperative multi-robot exploration," *International Journal of Factory Automation, Robotics and Soft Computing*, vol. 3, 2008.
- [26] D. Q. Mayne and H. Michalska, "Receding horizon control of nonlinear systems," in *Proceedings of the 27th IEEE Conference on Decision and Control*. IEEE, 1988, pp. 464–465.