Decentralized Planning for Active Information Gathering on Targets with Probabilistic Model

Dabin Kim[1], YoungSeok Jang[1], and H. Jin Kim[1]

Abstract— **Motion strategies for multiple robots actively acquiring information in a dynamic environment have been widely studied. However, existing active information gathering algorithms are restricted by assumption of linear target process, or deals with only limited number of agents and targets because of high cost of reward computation. In this paper, we formulated the active information gathering problem with the belief distribution of desired process, instead of a specific model. The reward function is derived based on mutual information of measurement and belief distribution, which can be efficiently computed under the Gaussian assumption on belief distribution. A decentralized path planner is designed to maximize reward function, which scales well in both numbers of agents and targets. We apply the proposed planner to a cooperative localization scenario and validate the performance and scalability in numerical simulation.**

I. INTRODUCTION

Active information gathering problems which require a motion strategy to efficiently sense desired physical quantities with limited sensors, such as environment monitoring [1], exploring subterranean environment [2], and search and rescue [3], are important applications of multi-robot systems. However, they are challenging because mobile robots have finite sensing range and it is impossible to visit many areas simultaneously. Therefore, selection of search trajectory and multi-agent coordination become especially important in aspects of avoiding redundant measurement and inter-collision.

The information gathering problems can be characterized into two groups according to the type of information. The first type observes spatio-temporal data such as indoor temperature distribution and models the spatial correlation of environment as a stochastic process such as Gaussian Process(GP) [4]. The other type acquires information about discrete events or targets, such as active target tracking [5] and persistent surveillance [6] scenarios. Unlike spatially correlated quantities, there is no measurement when the position of the target is out of the sensing range. Our work is focused on the latter type.

Many previous researches on active information gathering assumed the existence of *a priori* knowledge about the full model of underlying dynamics [5] or process [7]. But there are circumstances where only probabilistic information can be provided. For instance, in search-and-rescue, deployed search robots may collect environmental data which help to infer possible locations of survivors [8]. In target tracking, probabilistic search methods such as Probability Hypothesis Density(PHD) filter [9] can be employed to build a stochastic model about cardinality and belief state of target based on measurements. Also, in cooperative localization, each agent has its own belief of state and transmits to other agents [10]. Similarly, our work does not rely on the assumption of full knowledge of the target model and replaces it with a probability distribution over states of the target. Note that in the rest of the paper, information sources are collectively called *targets*, regardless of applications.

Here, we formulate the problem of maximizing gathered information about the desired event or target with measurements of mobile robots in a given time horizon. We derive an information-theoretic reward function and suggest an efficient computation. By exploiting properties of the reward function, a decentralized greedy path planner is proposed which is scalable in robot and target numbers. As an application, an active cooperative localization problem is investigated.

The contributions of the paper are as follows.

- Information-theoretic reward for stochastic events or targets, with rigorous derivation and efficient computation.
- Decentralized greedy path planning algorithm which is scalable in the number of agents and targets.
- Practical implementation in an active cooperative localization problem, validated with numerical simulation.

II. RELATED WORKS

Information gathering problems have been investigated both in theory and application domains. Information-theoretic methods have been used in various active perception problems. [11] used a gradient of mutual information to guide robots toward possible target positions in a grid environment. In sensor scheduling to estimate a stochastic process, [12] used the approximation algorithm with conditional entropy reward function, similar to our work. However, we focus on both the detection and measurement of discrete targets with mobile robots rather than spatially correlated processes. Non-myopic planners are developed under the linear target process assumption sampling-based [13] and search-based planners [5]. [14] used reinforcement learning to avoid assumptions on the model. [15] formulated an optimal control problem and developed event-driven controllers. [16], [17] applied Voronoi coverage control to ensure global coverage in a region of interest. [18] proposed a centralized planner to track a single moving target with approximated mutual information between the range sensor and target estimation with the particle filter. The Gaussian

[1]Dabin Kim, YoungSeok Jang, and H. Jin Kim is with Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul, South Korea {dabin404, duscjs59, hjinkim}@snu.ac.kr

belief assumption sacrifices the capability to express multiple hypothesis compared with non-parametric estimation, but it enables to deal with a larger number of mobile targets.

Active target tracking is one of the most widely studied problems among applications of information gathering. [19] used a task allocation approach for target tracking with local communication. Active target tracking strategies combined with PHD filters are studied in [20] and [21]. In [20], a sequential greedy planner is proposed with submodular reward functions and sequential Monte-Carlo PHD(SMC-PHD) filter. [21] combined particle PHD filter with coverage control by dividing the global environment into Voronoi partitions and exchanging information with Voronoi neighbors.

Active cooperative localization (CL), aimed to to improve the navigation quality with inter-robot measurements, is another application for an information gathering planner. Recently, motion strategies for actively improving CL performance have gained attention from researchers. [22] studied sensor placement to improvement the navigation of robots with pre-defined trajectories. In multi-robot setting, [23] selected the subset of robots which works as landmarks at each planning step. In [24], follower robots' trajectories are optimized to help localization of leaders. For UGV-UAV cooperation, [25] proposed a path where a ground robot provides better localization quality while an aerial robot exploits its high mobility and sensing range.

III. Problem Formulation

In an environment $\mathcal{E} \subset \mathbb{R}^d$, $A = \{A^1, \dots, A^N\}$ is a set of agents and $\mathbf{x}_k^i \in \mathbb{R}^n$ is the state of agent i at time step k. Robots follow dynamics $\mathbf{x}_{k+1}^i = f_d(\mathbf{x}_k^i, \mathbf{u}_k^i)$. $\mathbf{u}_k^i \in \mathbb{R}^d$ is the control input and $U_k^A = \{\mathbf{u}_k^1, \dots \mathbf{u}_k^N\}$ is the set of control input for set A. The state of target τ, $\mathbf{y}_k^\tau \in \mathbb{R}^m$ follows $\mathbf{y}_{k+1}^\tau = g(\mathbf{y}_k^\tau)$ where function g and true $\mathbf{y}_k^\tau$ are both unknown. The estimation of unknown target state, $\bar{\mathbf{y}}_k^\tau$, is given as a random variable T_k^τ which is estimation of true $\mathbf{y}_k^\tau$ and assumed to follow Gaussian distribution, i.e., $T_k^\tau = \mathcal{N}(\bar{\mu}_k^\tau, \bar{\Sigma}_k^\tau)$. Gaussian assumption is reasonable in many settings such as Gaussian Mixture PHD (GM-PHD) filter [9] of target tracking, state estimation with Extended Kalman Filter (EKF) in cooperative localization [10], and GP regression of target motion model [26]. $T_k = \{T_k^1, \dots, T_k^M\}$ is total set of target estimations.

Measurement of the t-th target from the i-th agent is $\mathbf{z}_k^{i\tau} \in \mathbb{R}^l$. Sensor model is $\mathbf{z}_k^{i\tau} = \mathbf{1}^{i\tau} \cdot (h(\mathbf{x}_k^i, \mathbf{y}_k^\tau) + v_k^i)$ where $\mathbf{1}^{i\tau}$ is the indicator function of the τ-th target detected by the i-th agent. For the sake of simplified derivation, measurement noise is approximated to Gaussian $v_k^i \sim \mathcal{N}(\mathbf{0}, \Sigma_n(\mathbf{x}_k^i))$. $Z_k^{G,\tau}$ is measurement set of target τ by agent set G. With assumption of no false-positive and false-negative measurement, belief distribution for the target state is updated as

$$\bar{\mu}_k^\tau = \mathbf{y}_k^\tau, \; \bar{\Sigma}_k^\tau = \Sigma_n(\mathbf{x}_k^i) \tag{1}$$

when target τ is detected and measurement is acquired by agent. It is assumed that all robots are equipped with enough sensors to observe target states for every target within sensing range. Control inputs to the agent set A for finite horizon K, $U_{1:K}^A$, are selected to maximize the total information through measurements obtained by sensing trajectories, represented by the objective function $f(U_{1:K}^A, T)$. Dependent on the task and specific purpose of motion planner, the objective function can be varied as optimality conditions of experiment design(i.e. A-opt, D-opt, E-opt) or information-theoretic quantity such as mutual information and conditional entropy [27]. As a result, problem formulation becomes

$$\max_{U_{1:K}^A} \sum_{k=1}^{K} f(U_k^A, T_k) \tag{2}$$

$$\mathbf{x}_{k+1}^i = f_d(\mathbf{x}_k^i, \mathbf{u}_k^i) \tag{3}$$

$$\bar{\mathbf{y}}_k^\tau \sim T_k^\tau = \mathcal{N}(\bar{\mu}_k^\tau, \bar{\Sigma}_k^\tau) \tag{4}$$

$$\mathbf{z}_k^{i\tau} = \mathbf{1}^{i\tau} \cdot (h(\mathbf{x}_k^i, \mathbf{y}_k^\tau) + v_k^i) \tag{5}$$

$$\text{for } i \in \{1, ..., N\}, \tau \in \{1, ..., M\}$$

In order to find trajectory with more information, designing the reward function f is important to lead robots to information-rich area and also achieve computational efficiency for real-time implementation. In section IV, the mutual information based reward is derived and an efficient computation algorithm is proposed.

IV. Information-Theoretic Reward

A. Derivation of the reward function

The reward function of agent set $G \subset A$ and target set T is the sum of all targets where measurement about each target is independent.

$$f(U_k^G, T_k) = \sum_{\tau=1}^{M} f(U_k^G, T_k^\tau) \tag{6}$$

This simplifies derivation of the reward function for a single target. Note that, on the contrary, the reward function for agent set is not a sum of each agent's reward. The reward function for a single agent, single target case is first examined, and then the multi-agent case is explained. The reward function is evaluated for one time step, thus time subscripts are omitted for conciseness.

Let the binary random variable $D_k^{G,\tau}$ represent the detection of target τ by the agent set G. Information reward function $f(U_k^G, T_k^\tau)$ is defined as expectation of mutual information of measurement set $Z_k^{G,\tau}$ and target random variable T_k^τ with respect to $D_k^{G,\tau}$.

$$\begin{aligned} f(U_k^G, T_k^\tau) &= \mathbb{E}_D[I(Z_k^G; T_k^\tau | U_k^G)] \\ &= p(D_k^{G,\tau}) I(Z_k^G; T_k^\tau | D_\tau^G, U_k^G) \end{aligned} \tag{7}$$

$\mathbb{E}$ is expectation and $I(X;Y)$ is mutual information of random variables X and Y.

1) Single Agent, Single Target: Let agent i gather information from the belief distribution for the target τ, $T_k^\tau = \mathcal{N}(\bar{\mu}_k^\tau, \bar{\Sigma}_k^\tau)$. For brevity of notation, index superscripts and time subscripts are omitted for this subsection. In (7), the reward function $f(U,T)$ is composed with probability of detection $p(D)$ and conditional mutual information $I(Z;T|D,U)$.

With the field of view(FoV) $S \subset \mathbb{R}^{d_S}$ of the agent, $p(D)$ is integral of belief distribution T_k^τ on $\mathcal{Y} \subset S \times \mathbb{R}^{m-d_S}$.

$$p(D) = \int_{\mathcal{Y}} p(t)d\tau = \int_{\mathcal{Y}} \mathcal{N}(t; \bar{\mu}, \bar{\Sigma})d\tau \tag{8}$$

The conditional mutual information is

$$I(Z;T|D,U) = H(Z|D,U) - H(Z|T,D,U) \tag{9}$$

where $H(A|B)$ is the conditional differential entropy of the random variable A given B [28].

With the state update equation (1), two terms of (9) become

$$H(Z|D,U) = -\frac{1}{p(D)} \int_{\mathcal{Z}} q(z)w(z) \log \frac{q(z)w(z)}{p(D)} dz \tag{10}$$

$$H(Z|T,D,U) = -\frac{1}{p(D)} \int_{\mathcal{Z}} q(z)v(z)dz \tag{11}$$

where $\mathcal{Z} \subset S \times \mathbb{R}^{l-d_S}$ is measurement space and

$$q(z) = \mathcal{N}(z; \bar{\mu}, \bar{\Sigma} + \Sigma_n) \tag{12}$$

$$w(z) = \int_{\mathcal{Y}} \mathcal{N}(t; \hat{\mu}, \hat{\Sigma})d\tau \tag{13}$$

$$v(z) = \int_{\mathcal{Y}} \mathcal{N}(t; \hat{\mu}, \hat{\Sigma}) \log \mathcal{N}(z; t, \Sigma_n)dt \tag{14}$$

$$\hat{\mu} = \bar{\mu} + H(z-t) \tag{15}$$

$$\hat{\Sigma} = (I - H)\bar{\Sigma} \tag{16}$$

$$H = \bar{\Sigma}(\bar{\Sigma} + \Sigma_n)^{-1} \tag{17}$$

Detailed derivations of (10), (11) are in Appendix A. As the result, the reward function for single agent and single target is

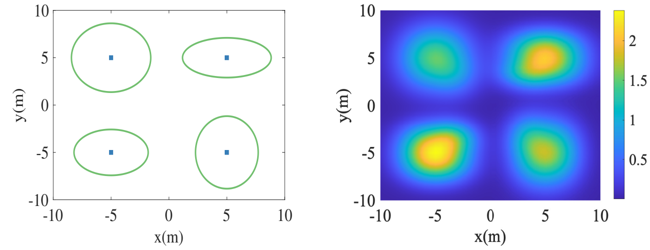
$$f(U,T) = -\int_{\mathcal{Z}} q(z)\Big(p(D)w(z) \log \frac{q(z)w(z)}{p(D)} - v(z)\Big)dz \tag{18}$$

2) Multiple Agents, Single Target: With two considerations about single agent function (18), it is possible to avoid integration on complex joint measurement domain, $\mathcal{Z}^A = \mathcal{Z}^1 \times \cdots \times \mathcal{Z}^N$. First, it is reasonable to assume that regardless of the number of times detected at a particular time step, the amount of information obtained is the same. $I(Z^{\{i,j\}}; T^\tau | D^{\{i,j\}}, U^{\{i,j\}}) \approx I(Z^{\{i\}}; T^\tau | D^{\{i\}}, U^{\{i\}})$ for any $A^i, A^j \in A$. The other is that detection probability of A is $p(D^{A,\tau}) = 1 - \prod_{i \in A}(1 - p(D^{i,\tau}))$. Therefore, the reward function for multi-agent is same as the single target case (18) beside that the detection probability is $p(D^{G,\tau})$.

For $A^j \in A$,

$$f(U^A, T^\tau) \approx \Big\{1 - \prod_{A^i \in A}(1 - p(D^{i,\tau}))\Big\} I(Z^j; T^\tau | D^{j,\tau}, U^j) \tag{19}$$

The reward function for multi-agents and multi -targets problem is followed from (6) and (19).



(a) Belief distributions for targets

(b) Heatmap of reward function

Fig. 1. Information-theoretic reward function for single agent and four targets on 2D plane. (a) the blue rectangle is the mean position and the green ellipse is the covariance matrix represented with confidence interval of $\chi^2 = 1.211$. (b) a heatmap of the reward function over 2D position validates the locality property (Lemma 1).

B. Properties of the reward function

In this subsection, two properties of the proposed reward function are provided. Proofs of the following two lemmas are provided in the full version of the paper, which can be found in author's personal webpage [29].

Lemma 1 (Locality): *If the covariance matrix Σ^τ is finite and measurement domain $\mathcal{Z}$ is bounded, following statements are satisfied where $d(x,\mu)_\Sigma \triangleq \sqrt{(x-\mu)^T\Sigma^{-1}(x-\mu)}$, p^i is position of agent i, ($\bar{\mu}_p^\tau$, $\bar{\Sigma}_p^\tau$) is the belief distribution on position state of target τ.*

1) If $\bar{\mu}_p^\tau \notin S^i$, information reward function f is upper bounded as $f(U^i, T^\tau) \leq \xi \exp\left(-\frac{1}{2}\left(d(p^i, \bar{\mu}_p^\tau)_{\bar{\Sigma}_p^\tau} - \psi\right)^2\right)$

2) $f(U^i, T^\tau) \to 0$ as $d(p^i, \bar{\mu}_p^\tau)_{\bar{\Sigma}_p^\tau} \to \infty$

Lemma 1 means that the reward f is governed by targets in local area. The locality property significantly eases the computation by considering targets near to the agent only. Fig. 1 shows a heatmap of reward function with four targets on 2D space.

The second property of reward function is submodularity which is widely used in combinatorial optimization [30], machine learning [31], and informative path planning [32].

Lemma 2 (Submodularity): *Information reward $f(U^A, T)$ is monotone submodular function with respect to action set U_A.*

The connection of the submodular maximization and the current problem is further discussed in section V.

C. Numerical computation

The procedure of numerical evaluation of (19) is listed in Algorithm 1. For simple explanation, we assume that the sensor is omnidirectional although it is possible to apply Algorithm 1 to sensors with limited FoV. We consider a target which has two-dimensional position and orientation as its state $\mathbf{y}^\tau$. In this case, measurement domain $\mathcal{Z}$ and target belief domain $\mathcal{Y}$ are identical. $\mathcal{Z} = \mathcal{Y} = S \times [-\pi, \pi]$

It is able to find approximations of (8), (13), and (14) in closed-form. From singular value decomposition of Σ in a integrand ($\bar{\Sigma}$ in (8), $\hat{\Sigma}$ in (13) and (14)), eigenvector matrix

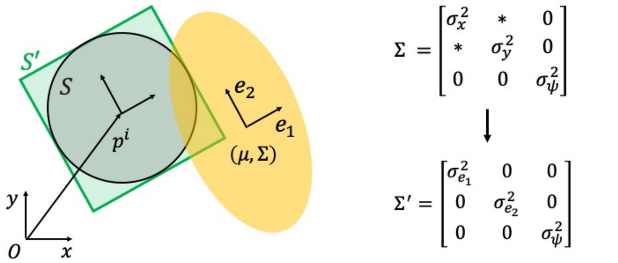


Fig. 2. Sensing range S (grey circle) is approximated to S' (green square) as (20). Axes of S' are parallel to eigenvectors e_1, e_2 of matrix Σ (yellow ellipse). Σ', the representation of Σ in eigenvector frame, is diagonal matrix.

E can be obtained as $\Sigma = EDF^T$ where D is a diagonal matrix of singular values and F is an orthonormal matrix. Let the sensing range S be approximated to S' as

$$S' = \{s \in \mathbb{R}^2 | s = E'(s_0 - p^i) + p^i, s_0 \in S\} \tag{20}$$

where $E = \begin{bmatrix} E' & 0_{2\times1} \\ 0_{1\times2} & 1 \end{bmatrix}$ and $p^i = [x^i, y^i]^T$ is current position of agent i. In the eigenvector frame (column vectors of E), covariance matrix Σ becomes diagonal matrix Σ' as in Fig. 2. Hence, by changing integration domain to $\mathcal{Y} = S' \times [-\pi, \pi]$, triple integrations in (8) and (13) are divided into multiplication of single integrals for each axis. Same holds for (14) if Σ_n is approximated to ${\Sigma_n}' = ED_nE^T$ where D_n is from $\Sigma_n = E_nD_nE_n^T$.

Each single integral is composed with multiplications and linear combinations of definite integrals $\int \exp x^2$ and $\int x^2 \exp x^2$. Since these integrals have analytic solutions in terms of polynomials and standard normal cumulative density functions (CDF), (8), (13) and (14) have closed-form solutions.

For each target from T, detection probability (8) for agents in G is computed and combined for the multi-agent reward function [line 5, Alg. 1]. Then, Monte-Carlo integration is used to evaluate the reward function in the measurement domain $\mathcal{Z}$ [lines 6-13, Alg. 1]. The integration resolution n_Z, n_ψ is the parameter for the trade-off between accuracy and efficiency. The total reward function is the sum of reward function for each target [line 14, Alg. 1].

Remark 1. *Computational complexity of Algorithm 1 is* $\mathcal{O}(M(N + n_Z^2 n_\psi))$ where n_Z is the integration resolution of S' and n_ψ is of $[-\pi, \pi]$.

Proof: For each target, Line 5 computes the detection probability (8) for every agent in $\mathcal{O}(N)$. Since the standard normal CDF can be obtained from a look-up table, evaluation of functions in line 8 is $\mathcal{O}(1)$. Therefore, for loops from Line 6 to Line 13 compute the single target reward function in $\mathcal{O}(n_Z^2 n_\psi)$. As a result, the total computation complexity is $\mathcal{O}(M(N + n_Z^2 n_\psi))$. ■

V. Decentralized Greedy Planner

In this section, a decentralized greedy planning algorithm is proposed based on the information reward function from the previous section. We adapted the greedy approximation algorithm since optimizing the mutual information reward is known as NP-Complete problem [33].

Instead of planning in the continuous space, agents generate a set of motion primitives, $L = \{l^1, \ldots l^{N_L}\}$, for every step. Motion primitive is a collision-free local trajectory that satisfies dynamics (3). Then, agents evaluate reward at the endpoint of each motion primitive as an approximation of trajectory reward.

Due to the submodularity of reward function (Lemma 2), the sequential greedy planner (Algorithm 2) has the suboptimality bound.

Theorem 1: *The sequential greedy assignment of multi-robot trajectory with the reward function* (19) *guarantees 1/2 approximation suboptimality.*

Proof: Proof directly follows from Theorem 1 of [20] and Lemma 2. (19) is a monotone, submodular function and multi-robot trajectories form a partition-matroid constraint. ■

Even though it has a guarantee on worst-case performance, the sequential greedy planner has exponential complexity with respect to the number of agents. [34] showed that a decentralized planner with coordinate descent also has the same performance guarantee as the sequential greedy algorithm and scales linearly with respect to a number of agents. Still, it requires full connectivity for multihop communication through the whole robot team and the linear complexity is intractable for large robot teams.

However, due to the locality of reward function stated in Lemma 1, it is possible to lower the number of targets and agents considered in the planning process. In the proposed decentralized planner, each agent considers only nearby targets and agents. This significantly lowers computational burden and improves scalability, which we show empirically in section VI. Also, direct communications only with local agents are required. It is especially useful when the communication range is limited and constraint of full connectivity

Algorithm 1 Evaluation of reward function (19)

1: **Input**: $T = \{T^1, \ldots T^M\}$, $U^A = \{U^1, \ldots U^M\}$
2: **Output**: $f(U^A, T)$
3: $\Delta \leftarrow (R/n_Z)^2(2\pi/n_\psi)$, $f(U^A, T) \leftarrow 0$
4: **for** $\tau = 1$ **to** M **do**
5: $\quad f(U^A, T^\tau) \leftarrow 0$
6: $\quad p^\tau \leftarrow \{1 - \prod_{i\in A}(1 - p(D^{i,\tau})\}$ $\Rightarrow p(D^{G,\tau})$
7: $\quad$ **for** $i = 1$ **to** n_Z **do**
8: $\quad\quad$ **for** $j = 1$ **to** n_Z **do**
9: $\quad\quad\quad$ **for** $k = 1$ **to** n_ψ **do**
10: $\quad\quad\quad\quad z \leftarrow z_{ijk}$ $\Rightarrow$ *sampled point in* $\mathcal{Z}$
11: $\quad\quad\quad\quad f(U^A, T^\tau) \leftarrow f(U^A, T^\tau) +$ $p^\tau \times q(z)\big(p^\tau w(z) \log \frac{q(z)w(z)}{p} - v(z)\big)\Delta$
12: $\quad\quad\quad$ **end for**
13: $\quad\quad$ **end for**
14: $\quad$ **end for**
15: $\quad f(U^A, T) \leftarrow f(U^A, T) + f(U^A, T^\tau)$
16: **end for**

Algorithm 2 Sequential Greedy Planner [20]

Input: Agents A, Targets T
Output: Actions U_A
Update motion primitive sets $L^1, ..., L^N$
$L = \cup_{i=1:N} L^i$
$U_A = \emptyset$ $\Rightarrow$ *Assigned action set*
$A = \{A_1, \ldots, A_N\}$ $\Rightarrow$ *Agents left for assignment*
while $A \neq \emptyset$ **do**
  $(\mathbf{u}, i) = \arg\max_{\mathbf{u}_i \in L} f(U_A \cup \{\mathbf{u}_i\}, T)$
  $U_A \leftarrow U_A \cup \{\mathbf{u}\}$
  $A \leftarrow A \backslash A_i$
end while

reduces team performance.

The proposed decentralized planning algorithm is explained in detail in Algorithm 3. Prior to deployment, each robot is assigned ID, to distinguish and decide priority among neighbors. At every time step, each agent generates a fixed number of motion primitives [Line 4, Alg. 3]. Also, the neighbor set and local target set are updated [Line 5-6, Alg. 3]. The neighbor set of agent i, A^i_o, is defined as agents whose euclidean distance from agent i is smaller than threshold ϵ_e. The local target set T^i_o is set of targets whose euclidean distance from mean μ^τ to agent i is smaller than threshold ϵ_m. The neighbor agent set A^i_o is decomposed into disjoint sets A^i_{pre} and $A^i_{post} = A^i_o \backslash A^i_{pre}$. A^i_{pre} is a set of agents with prior ID compared to the agent i. A^i_{post} is the rest of the agents in A^i_o.

Before making its own decision, the agent i requests actions of agents in A^i_{pre} [Line 7, Alg. 3]. If there is no element in A^i_{pre}, agent makes its decision without considering other agents. For each motion primitive, the reward is evaluated and action with the maximum reward is selected as the successive input [Line 8, Alg. 3]. Then, the agent transmits the selected action to robots in A^i_{post}.

Although the greedy planner has an advantage in fast computation, it can get stuck in local maxima and lead agents to cover only the local area. Therefore, heuristics that lead to global coverage can be added to the reward function for better performance. In the next section, a decentralized information gathering planner is applied to active cooperative localization.

VI. Active Cooperate Localization

The proposed planner can be applied in order to find an effective motion strategy to improve performance in CL. In this problem, we define two sets of robots that compose CL. One set is leader robots (targets), which are performing predefined tasks for higher level goals. The other robots are follower robots (agents) which try to search and track leader robots to provide better localization information for targets. Targets transmit their belief states to agents, and agents transmit updated target state to the target when it is detected. First, to simplify the problem, we assume that agents have perfect localization. This seemingly strong assumption is reasonable in cases when agents have map information or

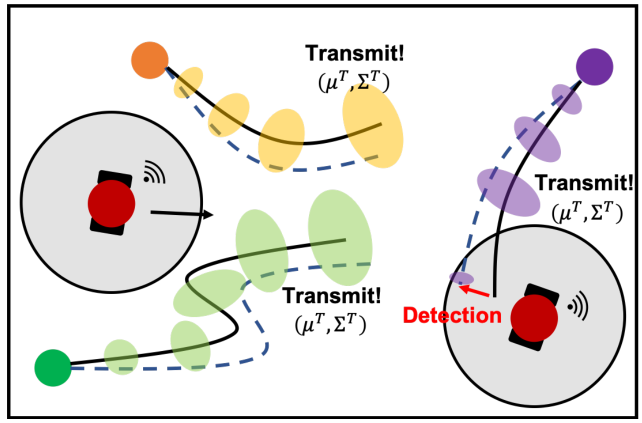


Fig. 3. Active cooperative localization problem. Three target robots (circle) have estimation of ego motion. Uncertainty (ellipse) and error between the estimated state (solid line) and the ground-truth state (dashed line) gradually increase. Each target transmits own belief distribution to agents. Two mobile robots with limited sensing range (grey circle) are deployed to detect and measure target states. When detection occurs (red arrow), agent transmits its target state measurement to corresponding target.

agents are equipped with high-performance sensors. This is one of the examples of [10] which considers distributed leader-follower CL. Generalizing problems with localization uncertainty of agents is left as future work. The problem setup is explained graphically in Fig. 3.

The goal of cooperative localization is to minimize the uncertainty and error of target robots. Log-det of covariance matrix is used as a quantity to express the localization uncertainty. Minimizing log-det of covariance is equivalent to maximizing the reward function (19) since the differential entropy of the Gaussian random variable is an affine form of log-det of covariance matrix [33].

As mentioned in the previous section, adding heuristic value in reward function may improve global coverage, although it makes reward function no longer submodular. In CL application, we designed heuristic function h in case agents have empty local target set. If the target set of agent

Algorithm 3 Decentralized Greedy Planner for Agent i

1: **Input**: A, R, T, ϵ_e, ϵ_m
2: **Output**: U_A
3: **for** k=1 to K **do**
4:   Update $L^i = \{l^1, ..., l^{N_L}\}$
5:   Update Neighbor Set $A^i_o = A^i_{pre} \cup A^i_{post}$
  $A^i_o = \{A_a | A_a \in A, d(\mathbf{x}^i, \mathbf{x}^a) \leq \epsilon_e\}$
6:   Update Local Target Set T^i_o
  $T^i_o = \{T^\tau | T^\tau \in T, d(\mathbf{x}^i, \bar{\mu}^\tau) \leq \epsilon_m\}$
7:   $U_{pre} \leftarrow$ request actions of prior robots A_{pre}
8:   $\mathbf{u}^i \leftarrow \arg\max_{\mathbf{u} \in L^i} f(U_{pre} \cup \{\mathbf{u}\}, T_o)$
9:   Transmit $\mathbf{u}^i_k$ to A_{post}.
10: **end for**

i, T_o^i, is empty, heuristic reward is defined as

$$h = \lambda \cdot \max_{\tau \in T} \frac{\det \bar{\Sigma}^{\tau}}{d(\mathbf{x}^i, \bar{\mu}^{\tau})}, \qquad \lambda > 0 \tag{21}$$

A. Simulation Setup

In numerical simulation, algorithms are evaluated in the two dimension environment with 50 m $\times$ 50 m size. Target robots have position x, y and orientation ψ states. They follow double-integrator dynamics for position states and single-integrator dynamics for orientation state, which is simplified dynamics of multirotor. Each target runs Kalman Filter(KF) with noisy linear/angular velocity measurements. For every time step, belief gaussian distribution $\mathcal{N}(\bar{\mu}_{\tau}, \bar{\Sigma}_{\tau})$ is transmitted to agents. Agent robots also follow a differential drive model and motion primitives are generated from finite set of forward velocity $v = \{0.5,\ 1,\ 1.5,\ 2,\ 2.5\}$m/s and angular velocity $\omega = \{0,\ \pm 0.5,\ \pm 1.0\}$rad/s. Agents are equipped with an omnidirectional sensor to get relative position measurements of the target such as range and bearing sensors. Also, relative orientation of target is extracted from camera image as [35]. Agents do not have prior information about dynamics of targets, and only the received belief distributions are used to generate informative trajectories. With sensing radius $R = 2.5$ m and threshold $\epsilon_e = 10$ m, $\epsilon_m = 5$ m, each agent actively plans their trajectory from Algorithm 3.

Two algorithms, a decentralized greedy planner with coordinate descent (DGP-CD) and a coverage planner (CP) with lawn-mowing trajectory, are compared with the proposed algorithm (P). Additionally, the proposed algorithm with heuristic reward (21) is evaluated (P+H). Every algorithm except for CP, the information-theoretic reward (19) is used to appraise local trajectories. Simulations were performed on MATLAB with Intel Core i7-7700 @ 3.60Hz CPU and 16GB RAM.

B. Result

Fig. 4. is the results of applying two proposed planners (P, P+H) and baseline algorithms (DGP-CD, CP) on cooperative localization problem. With 8 agents and 16 targets, the proposed planner persistently minimize target uncertainty and error which is the goal of cooperative localization. Also, the proposed algorithm outperforms CP and performs similar to DGP-CD even it only considers local agents and targets. Applying heuristics to the proposed planner consistently shows better result, especially when the proposed planner is stuck in local maxima. Heuristic reward (21) leads agents to informative region and helps to find better solution.

Scalability of the proposed algorithm is analyzed with two experiments. First, computation time of the proposed planner using local target set, T_o, and global target set, T, are compared. Table I shows the results with two types of agent set and four types of target set. When planner uses global target set, computation time linearly increases with respect to the target number. Since evaluation of reward function takes majority of the total computation time, considering whole targets makes planner intractable with large number of targets. Besides, using local target set significantly reduces computation time and planner remains tractable even in case of 64 targets. Second scalability analysis compares between the proposed algorithm and DGP-CD. DGP-CD performs single agent planner sequentially for all agents and single planning step is terminated when the last agent is finished. Accordingly, complexity of DGP-CD scales linearly with respect to the number of agents. However, the proposed algorithm enables simultaneous planning of non-neighboring agents in distributed manner. In Table II computation time of single planning step of proposed planner and DGP-CD are compared. Two planners both evaluate reward function (7) with local target set of threshold $\epsilon_m = 5$ m. Computation time of the proposed planner and DGP-CD are similar when agent set and target set are relatively small. But proposed algorithm scales well as size of agents increases. Gap between two algorithms increase as the number of agents increases, since the size of the neighbor set A_o does not increase at the same rate. The scalability of the proposed planner makes it possible to implement to larger robot teams compare to DGP-CD.

TABLE I: Computation time of the proposed planner per agent with local and global target sets. (AN: N agents, TM: M targets)

Targets	Computaion Time(ms)							
	A4				A8			
	T8	T16	T32	T64	T8	T16	T32	T64
Local	44.0	58.8	73.0	96.0	47.4	57.5	87.1	117
Global	308	580	1104	2141	287	575	1149	2008

TABLE II: Computation time of planning with proposed algorithm and DGP-CD (AN: N agents, TM: M targets)

Planner	Computaion Time(ms)							
	A2		A4		A8		A16	
	T4	T64	T8	T64	T16	T64	T32	T64
Proposed	30.4	96.8	104	361	139	405	368	583
DGP-CD	33.2	165	137.1	505	359	745	1007	1643

VII. CONCLUSIONS

In this paper, we proposed a decentralized path planner for active information gathering tasks. Especially, the target model is replaced by belief distribution of the target state. The information-theoretic reward function which leads agents toward the desired target is derived and an efficient computation procedure is introduced. The proposed decentralized planner is designed by exploiting the locality and submodularity properties of the reward function. The planner scales well with the number of agents and targets. In an application of active cooperative localization, numerical simulations validated performance of the proposed planner.

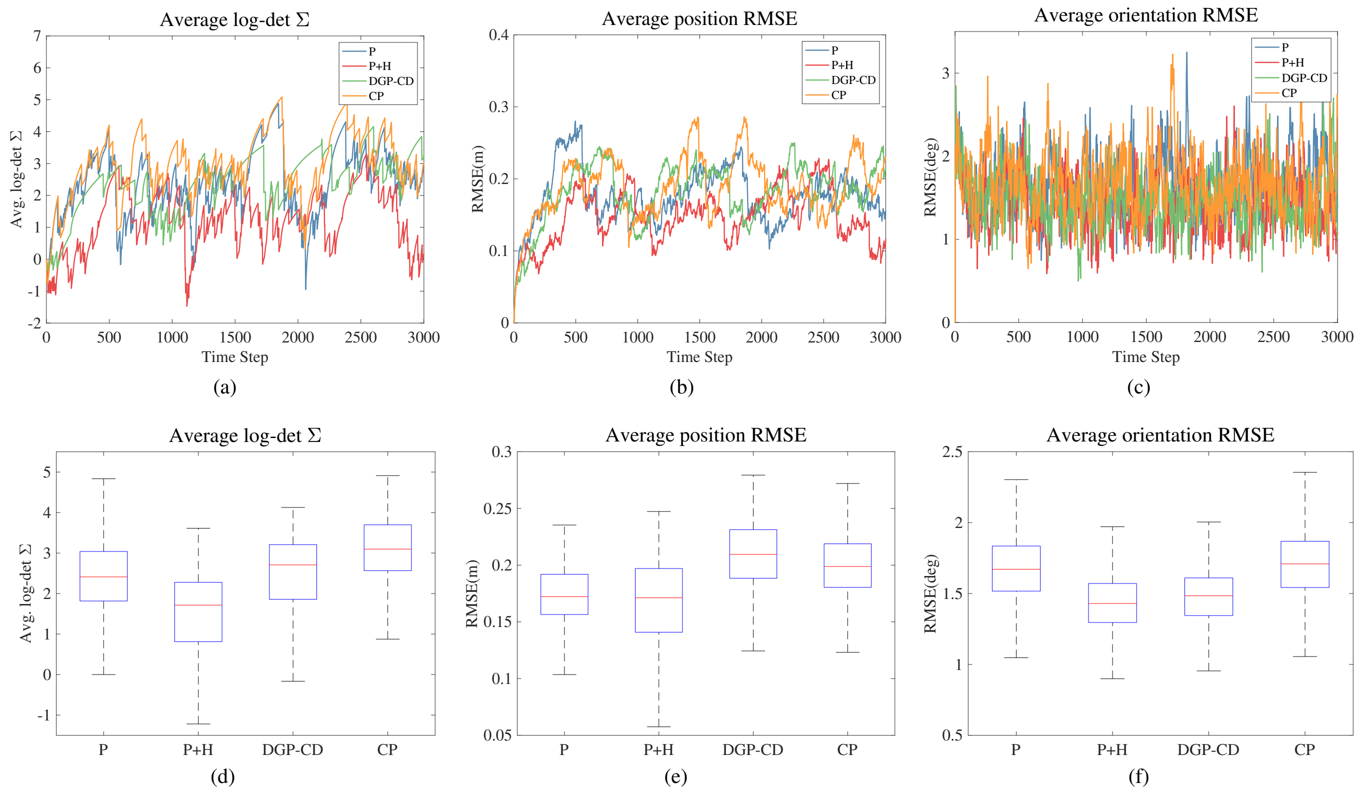


Fig. 4. With 8 agents and 16 targets, the proposed planner(P), proposed planner with heuristic(P+H), decentralized greedy planner with coordinate descent(DGP-CD), and coverage planner(CP) are applied to cooperatie localization problem. (a) Log-det of covariance matrix of belief distribution averaged over targets. (b) root mean square error(RMSE) of position averaged over targets. (c) RMSE of orientation averaged over targets. (d) Box plot of average log-det of covariance matrix. (e) Box plot of average position RMSE. (f) Box plot of average orientation RMSE.

APPENDIX A.

A. *Derivation of* (10)

Conditional differential entropy $H(Z|D,U)$ is

$$H(Z|D,U) = -\int_{\mathcal{Z}} p(z|D,U)p(D)\log p(z|D,U)dz \quad (22)$$

and integral domain is $\mathcal{Z}$ since $p(z|D,U) = 0$ for $z \notin \mathcal{Z}$.

$p(Z|D,U)$ is given as marginalization on target state variable t:

$$\begin{aligned} p(Z|D,U) &= \int_{\mathbb{R}^m} p(Z|t,D,U)p(t|D,U)dt \\ &= \frac{1}{p(D)}\int_{\mathbb{R}^m} p(Z|t,D,U)p(t,D|U)dt \\ &= \frac{1}{p(D)}\int_{\mathbb{R}^m} p(Z|t,D,U)p(D|t,U)p(t)dt. \end{aligned} \quad (23)$$

Since $p(D|t,U) = 1$ if $t \in S$ and 0 otherwise, we have

$$p(Z|D,U) = \frac{1}{p(D)}\int_{\mathcal{Y}} p(Z|t,D,U)p(t)dt \quad (24)$$

From (1) and Lemma 2 of [9],

$$\begin{aligned} &p(Z|t,D,U)p(t) = \mathcal{N}(z;t,\Sigma_n)\mathcal{N}(t;\bar{\mu},\bar{\Sigma}) \\ &= q(z)\mathcal{N}(t;\hat{\mu},\hat{\Sigma}) \end{aligned} \quad (25)$$

where

$$q(z) = \mathcal{N}(z;\bar{\mu},\bar{\Sigma}+\Sigma_n) \quad (26)$$

$$\hat{\mu} = \bar{\mu} + H(z-\tau) \quad (27)$$

$$\hat{\Sigma} = (I-H)\bar{\Sigma} \quad (28)$$

$$H = \bar{\Sigma}(\bar{\Sigma}+\Sigma_n)^{-1} \quad (29)$$

Substituting (25) to (24), we obtain

$$\begin{aligned} p(Z|D,U) &= \frac{1}{p(D)}q(z)\int_{\mathcal{Y}} \mathcal{N}(t;\hat{\mu},\hat{\Sigma})dt \\ &= \frac{1}{p(D)}q(z)w(z) \end{aligned} \quad (30)$$

where $w(z) = \int_{\mathcal{Y}} \mathcal{N}(t;\hat{\mu},\hat{\Sigma})dt$. Substitution of (30) into (22) results in

$$H(Z|D,U) = -\int_{\mathcal{Z}} q(z)w(z)\log\left(\frac{q(z)w(z)}{p(D)}\right)dz \quad (31)$$

B. *Derivation of* (11)

The second term of (9) is

$$H(Z|T,D,U) = -\int_{\mathcal{Z}}\int_{\mathbb{R}^m} p(z,t|D,U)\log p(z|t,D,U)dtdz \quad (32)$$

Probability $p(z,t|D,U)$ is

$$p(z,t|D,U) = p(z|t,D,U)p(t|D,U) \tag{33}$$
$$= \frac{p(t)}{p(D)}p(z|t,D,U)p(D|t,U) \tag{34}$$

Substitute (33) to (22) to yield

$$H(Z|T,D,U)$$
$$= -\int_{\mathcal{Z}}\int_{\mathbb{R}^m}\frac{p(t)}{p(D)}p(z|t,D,U)p(D|t,U)\log p(z|t,D,U)dtdz$$
$$= -\int_{\mathcal{Z}}\int_{\mathcal{Y}}\frac{p(t)}{p(D)}p(z|t,D,U)\log p(z|t,D,U)dtdz \tag{35}$$

Using (25) to (35), we arrive at

$$H(Z|T,D,U) = -\frac{1}{p(D)}\int_{\mathcal{Z}}q(z)v(z)dz \tag{36}$$

where

$$v(z) = \int_{\mathcal{Y}}\mathcal{N}(t;\hat{\mu},\hat{\Sigma})\log\mathcal{N}(z;t,\Sigma_n)dt \tag{37}$$

APPENDIX B.

Lemma 1 (Locality): *If the covariance matrix Σ^τ is finite and measurement domain $\mathcal{Z}$ is bounded, following statements are satisfied where $d(x,\mu)_\Sigma \triangleq \sqrt{(x-\mu)^T\Sigma^{-1}(x-\mu)}$, p^i is position of agent i, ($\bar{\mu}_p^\tau$, $\bar{\Sigma}_p^\tau$) is the belief distribution on position state of target τ.*

1) *If $\bar{\mu}_p^\tau \notin S^i$, information reward function f is upper bounded as $f(U^i,T^\tau) \leq \xi\exp\left(-\frac{1}{2}\left(d(p^i,\bar{\mu}_p^\tau)_{\bar{\Sigma}_p^\tau}-\psi\right)^2\right)$*
2) *$f(U^i,T^\tau) \to 0$ as $d(p^i,\bar{\mu}_p^\tau)_{\bar{\Sigma}_p^\tau} \to \infty$*

Proof:

1) Reward function for single agent, single target is $f(U^i,T^\tau) = p(D^{i,\tau})I(Z^i;T^\tau|D^{i,\tau},U^i)$. For brevity of notation, superscript i and τ are omitted. First, conditional mutual information between measurement and target varialbe given detection is upper bounded. $I(Z;T|D,U) \leq \eta$

Mutual information is non-negative[28].

$$I(Z;T|D,U) = H(Z|D,U) - H(Z|T,D,U) \geq 0 \tag{38}$$

$H(Z|T,D,U) \geq 0$ since $p(z|\tau,D,U) = \mathcal{N}(z;\tau,\Sigma_n)$. Hence, $I(Z;T|D,U)$ is upper bounded if $H(Z|D,U)$ has a finite upper bound.

Let $g(z) = 1/p(z|D,U)$ and $G = \int p(z)dz$, a cumulative probability function. Note that $p(z|D,U) > 0$ in domain $\mathcal{Z}$. Then, the entropy of $p(z|D)$ is

$$H(Z|D,U) = \int_{\mathcal{Z}}\log g(z)dG$$
$$\leq \log(\int_{\mathcal{Z}}gdG) = \log(\int_{\mathcal{Z}}1dz) = \log(\text{Vol}(\mathcal{Z})) \tag{39}$$

where Jensen's inequality with the concavity of $\log(x)$ on $x \in (0,\infty)$ is used.

The volume of measurement space depends on the target states. For example, if sensor measures a two-dimensional position vector, $\text{Vol}(\mathcal{Z}) = \frac{\pi}{2}R^2$. The volumes corresponding to complex states such as $SE(3)$ can be found in [36]. From the assumption that measurement domain $\mathcal{Z}$ is bounded,

$$\therefore I(Z;T|D,U) \leq \log(\text{Vol}\mathcal{Z}) = \eta \tag{40}$$

The probability of detection is upper bounded as

$$p(D) = \int_S p(t)dt$$
$$= \int_S \frac{1}{2\pi(\det\bar{\Sigma}_p)^{1/2}}\exp -\frac{1}{2}d(t,\bar{\mu}_p)^2_{\bar{\Sigma}_p}dt$$
$$\leq \frac{\text{Vol}(S)}{2\pi(\det\bar{\Sigma}_p)^{1/2}}\exp(-\frac{1}{2}d(t^*,\bar{\mu}_p)^2_{\bar{\Sigma}_p}) \tag{41}$$

where $t^* = \underset{t\in S}{\arg\min}\ d(t,\bar{\mu}_p)^2_{\bar{\Sigma}_p}$.

If $\bar{\mu}_p \notin S$, from triangle inequality,

$$d(t,\bar{\mu}_p)_{\bar{\Sigma}_p} \geq d(p,\bar{\mu}_p)_{\bar{\Sigma}_p} - d(p,t)_{\bar{\Sigma}_p} \tag{42}$$

Since $\bar{\Sigma}_p^t$ is positive definite and $||p-t||^2 \leq R^2$ by limited sensing range R,

$$d(p,t)_{\bar{\Sigma}_p} = (p-t)^T(\bar{\Sigma}_p)^{-1}(p-t)$$
$$= (p-t)^T VDV^T(p-t)$$
$$= \sum_{j=1}^{d_S}\sigma_j^2||p-t||^2 \leq \sum_{j=1}^{d_S}\sigma_j^2R^2 \tag{43}$$

where V is eigenvector matrix, D is diagonal matrix whose entries are eigenvalues, and σ_j are singular values of $(\bar{\Sigma}_p)^{-1}$.

Therefore, $d(t,\bar{\mu}_p)_{\bar{\Sigma}_p} \geq d(p,\bar{\mu}_p)_{\bar{\Sigma}_p} - \sum_{i=1}^{d_S}\sigma_i^2R^2$ and from (41), we get

$$p(D) \leq \frac{\text{Vol}(S)}{2\pi(\det\bar{\Sigma}_p)^{1/2}}\exp\Big(-\frac{1}{2}(d(p,\bar{\mu}_p)_{\bar{\Sigma}_p} - \sum_{j=1}^{d_S}\sigma_j^2R^2)^2\Big) \tag{44}$$

As a result, from (40) and (44), the reward function is upper bounded as

$$f(U^i,T^\tau) \leq \xi\exp(-\frac{1}{2}(d(p^i,\bar{\mu}_p^\tau)_{\bar{\Sigma}_p^\tau}-\psi)^2) \tag{45}$$

2) Since $\bar{\Sigma}^\tau$ is finite and positive definite, singular values σ_j are finite. Thus, ψ of (45) is also finite. Then, from (45), as $d(p,\bar{\mu}_p^\tau)_{\bar{\Sigma}_p^\tau} \to \infty$, upper bound of $f(U^i,T^\tau)$ goes to zero. $f(U^i,T^\tau) \geq 0$ since $p(D)$ and mutual information are non-negative. Therefore, $f(U^i,T^\tau) \to 0$ as $d(p,\bar{\mu}_p^\tau)_{\bar{\Sigma}_p^\tau} \to \infty$.

Lemma 2 (Submodularity): *Information reward $f(U^A,T)$ is monotone submodular function with respect to action set U^A.*

Proof.

1) For agent set $G \subset A$, information reward function $f(U^G,T)$ is

$$f(U^G,T) = \sum_{\tau=1}^{M}p(D^{G,\tau})I(Z^G;T^\tau|D^{G,\tau},U^G). \tag{46}$$

By adding agent r to set G, the detection probability changes as

$$p(D^{G\cup\{A_r\},\tau}) - p(D^{G,\tau}) = \prod_{i\in G}(1-p(D^{i,\tau}))\cdot p(D^{r,\tau}) \geq 0 \tag{47}$$

With the fact that mutual information is non-negative, reward function increases as

$$\begin{aligned}&f(U^{G\cup\{A_r\}},T) - f(U^G,T)\\ =&\sum_{\tau=1}^{M}\prod_{i\in G}(1-p(D^{i,\tau}))\cdot p(D^{r,\tau})I(Z^G;T^\tau|D^{G,\tau},U^G) \geq 0\end{aligned} \tag{48}$$

Therefore, reward function f is monotonically increasing with respect to agent set.

2) Let $G \subset F \subset A$ and $|G| = n$, $|F| = n + l$, $n, l \in \mathbb{N}$. From equation (19), information reward for single target T^τ with agent set G is

$$f(U^G,T^\tau) = \Big\{1-\Pi_{i=1}^{n}(1-p(D^{i,\tau}))\Big\}I(Z^G;T^\tau|D^{G,\tau},U^G) \tag{49}$$

With additional agent r which $A^r \notin G$, reward function becomes

$$\begin{aligned}f(U^{G\cup\{r\}},T^\tau) =&\Big\{1-(1-p(D^{r,\tau}))\Pi_{i=1}^{n}(1-p(D^{i,\tau}))\Big\}\\ &I(Z^{G\cup\{r\}};T^\tau|D^{G\cup\{r\},\tau},U^{G\cup\{r\}})\end{aligned} \tag{50}$$

As stated in section IV, if detection occurs, gathered information can be assumed as constant $k \geq 0$ with respect to the agent set. Therefore, additional reward for set G is

$$\begin{aligned}\Delta^G &= f(U^{G\cup\{r\}},T^\tau) - f(U^G,T^\tau)\\ &= p(D^{r,\tau})\Pi_{i=1}^{n}(1-p(D^{i,\tau}))k\end{aligned} \tag{51}$$

Similarly, additional reward for set F is

$$\Delta^F = f(U^{F\cup\{r\}},T^\tau) - f(U^F,T^\tau) \tag{52}$$

$$= p(D^{r,\tau})\Pi_{i=1}^{n+l}(1-p(D^{i,\tau}))k \tag{53}$$

By subtracting (52) to (51), we obtain

$$\begin{aligned}\Delta^G - \Delta^F =& p(D^{r,\tau})\big\{1-\Pi_{i=n+1}^{n+l}(1-p(D^{i,\tau}))\big\}\\ &\Pi_{i=1}^{n}(1-p(D^{i,\tau}))k \geq 0\end{aligned} \tag{54}$$

As a result, information reward $f(U^A,T)$ is submodular function with respect to action set U^A.

References

[1] S. Li, Y. Guo, and B. Bingham, "Multi-robot cooperative control for monitoring and tracking dynamic plumes," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 67–73.

[2] W. Tabib, M. Corah, N. Michael, and R. Whittaker, "Computationally efficient information-theoretic exploration of pits and caves," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 3722–3727.

[3] J. L. Baxter, E. Burke, J. M. Garibaldi, and M. Norman, "Multi-robot search and rescue: A potential field based approach," in *Autonomous robots and agents*. Springer, 2007, pp. 9–16.

[4] A. Viseras, T. Wiedemann, C. Manss, L. Magel, J. Mueller, D. Shutin, and L. Merino, "Decentralized multi-agent exploration with online-learning of gaussian processes," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 4222–4229.

[5] B. Schlotfeldt, D. Thakur, N. Atanasov, V. Kumar, and G. J. Pappas, "Anytime planning for decentralized multirobot active information gathering," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1025–1032, 2018.

[6] C. Baykal, G. Rosman, S. Claici, and D. Rus, "Persistent surveillance of events with unknown, time-varying statistics," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 2682–2689.

[7] S. L. Smith, M. Schwager, and D. Rus, "Persistent robotic tasks: Monitoring and sweeping in changing environments," *IEEE Transactions on Robotics*, vol. 28, no. 2, pp. 410–426, 2011.

[8] K.-M. Yang, J.-B. Han, and K.-H. Seo, "A target position decision system based on eyewitness account for exploring disaster environment," in *2019 IEEE International Conference on Advanced Robotics and its Social Impacts (ARSO)*. IEEE, pp. 13–19.

[9] B.-N. Vo and W.-K. Ma, "The gaussian mixture probability hypothesis density filter," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[10] T. R. Wanasinghe, G. K. Mann, and R. G. Gosine, "Distributed leader-assistive localization method for a heterogeneous multirobotic system," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 3, pp. 795–809, 2015.

[11] J.-P. Ramirez-Paredes, E. A. Doucette, J. W. Curtis, and N. R. Gans, "Distributed information-based guidance of multiple mobile sensors for urban target search," *Autonomous Robots*, vol. 42, no. 2, pp. 375–389, 2018.

[12] V. Tzoumas, N. A. Atanasov, A. Jadbabaie, and G. J. Pappas, "Scheduling nonlinear sensors for stochastic process estimation," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 580–585.

[13] Y. Kantaros, B. Schlotfeldt, N. Atanasov, and G. J. Pappas, "Asymptotically optimal planning for non-myopic multi-robot information gathering," *Proceedings of Robotics: Science and Systems (RSS), Freiburg, Germany*, pp. 22–26, 2019.

[14] H. Jeong, B. Schlotfeldt, H. Hassani, M. Morari, D. D. Lee, and G. J. Pappas, "Learning q-network for active information acquisition," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 6822–6827.

[15] N. Zhou, X. Yu, S. B. Andersson, and C. G. Cassandras, "Optimal event-driven multiagent persistent monitoring of a finite set of data sources," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4204–4217, 2018.

[16] A. Mellone, G. Franzini, L. Pollini, and M. Innocenti, "Persistent coverage control for teams of heterogeneous agents," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 2114–2119.

[17] J. M. Palacios-Gasós, E. Montijano, C. Sagüés, and S. Llorente, "Distributed coverage estimation and control for multirobot persistent tasks," *IEEE transactions on Robotics*, vol. 32, no. 6, pp. 1444–1460, 2016.

[18] B. Charrow, V. Kumar, and N. Michael, "Approximate representations for multi-robot control policies that maximize mutual information," *Autonomous Robots*, vol. 37, no. 4, pp. 383–400, 2014.

[19] Y. Sung, A. K. Budhiraja, R. K. Williams, and P. Tokekar, "Distributed assignment with limited communication for multi-robot multi-target tracking," *Autonomous Robots*, vol. 44, no. 1, pp. 57–73, 2020.

[20] P. Dames, P. Tokekar, and V. Kumar, "Detecting, localizing, and tracking an unknown number of moving targets using a team of mobile robots," *The International Journal of Robotics Research*, vol. 36, no. 13-14, pp. 1540–1553, 2017.

[21] P. Dames, "Distributed multi-target search and tracking using the phd filter," in *2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*. IEEE, 2017, pp. 1–8.

[22] M. P. Vitus and C. J. Tomlin, "Sensor placement for improved robotic navigation," *Robotics: science and Systems VI*, p. 217, 2011.

[23] Q. Yan, L. Jiang, and S. Kia, "Measurement scheduling for cooperative localization in resource-constrained conditions," *IEEE Robotics and Automation Letters*, 2020.

[24] L. Zhang, S. Chowdhury, R. Siegwart, and J. J. Chung, "Optimized motion strategy for active target localization of mobile robots with time-varying connectivity," in *2019 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*. IEEE, 2019, pp. 185–187.

[25] J. Butzke, K. Gochev, B. Holden, E.-J. Jung, and M. Likhachev, "Planning for a ground-air robotic system with collaborative localization,"

in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 284–291.
[26] N. Wahlström and E. Özkan, "Extended target tracking using gaussian processes," *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4165–4178, 2015.
[27] D. A. Paley and A. Wolek, "Mobile sensor networks and control: Adaptive sampling of spatiotemporal processes," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, 2019.
[28] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
[29] [Online]. Available: https://dabinkim-lgom.github.io/
[30] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—i," *Mathematical programming*, vol. 14, no. 1, pp. 265–294, 1978.
[31] F. Bach *et al.*, "Learning with submodular functions: A convex optimization perspective," *Foundations and Trends® in Machine Learning*, vol. 6, no. 2-3, pp. 145–373, 2013.
[32] A. Singh, A. Krause, and W. J. Kaiser, "Nonmyopic adaptive informative path planning for multiple robots," in *Twenty-First International Joint Conference on Artificial Intelligence*, 2009.
[33] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research*, vol. 9, no. Feb, pp. 235–284, 2008.
[34] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, "Decentralized active information acquisition: Theory and application to multi-robot slam," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 4775–4782.
[35] H.-Y. Tseng, P.-C. Wu, M.-H. Yang, and S.-Y. Chien, "Direct 3d pose estimation of a planar target," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2016, pp. 1–9.
[36] L. J. Boya, E. Sudarshan, and T. Tilma, "Volumes of compact manifolds," *Reports on Mathematical Physics*, vol. 52, no. 3, pp. 401–422, 2003.