Decentralized Multi-Robot Information Gathering from Unknown Spatial Fields

Abdullah Al Redwan Newaz[1], Murtadha Alsayegh[2], Tauhidul Alam[3], and Leonardo Bobadilla[2]

***Abstract*—We present an incremental scalable motion planning algorithm for finding maximally informative trajectories for decentralized mobile robots. These robots are deployed to observe an unknown spatial field, where the informativeness of observations is specified as a density function. Existing works that are typically restricted to discrete domains and synchronous planning often scale poorly depending on the size of the problem. Our goal is to design a distributed control law in continuous domains and an asynchronous communication strategy to guide a team of cooperative robots to visit the most informative locations within a limited mission duration. Our proposed Asynchronous Information Gathering with Bayesian Optimization (AsyncIGBO) algorithm extends ideas from asynchronous Bayesian Optimization (BO) to efficiently sample from a density function. It then combines them with decentralized reactive motion planning techniques to achieve efficient multi-robot information gathering activities. We provide a theoretical justification for our algorithm by deriving an asymptotic no-regret analysis with respect to a known spatial field. Our proposed algorithm is extensively validated through simulation and real-world experiment results with multiple robots.**

I. INTRODUCTION

THE capabilities of mobile robots to gather critical information from a large area play pivotal roles in their applications to monitor several environmental phenomena. Some of these phenomena are nuclear radiation, greenhouse gas emission, soil parameters (e.g., nutrient levels and pH) in precision agriculture, and chemical or oil spills in coastal areas that have high spatial variability. These phenomena are generally referred to as spatial fields. To assess and study unknown spatial fields, the goal is to maximize an information gain metric while satisfying robot motion and resource constraints [1]. However, it is challenging to gather maximum information by exploring a large geographical space with a thorough search by planning robot trajectories for a specific time period. Such an information gathering process involves the robot physically exploring the environment to collect measurements and is not purely a computational process. Due to physical constraints of robots, we cannot simply "teleport" the robot to a random location, and information must be collected sequentially along a trajectory. An effective information gathering method using multi-robot systems requires deploying and collecting samples from a finite number of locations and predicting the missing values of the field using a statistical model. Significant existing works [1, 2, 3] focus on sampling strategies that decide where measurements should be taken while satisfying constraints on fuel, energy, or time. Most of these works are typically restricted to discrete high-level synchronous planning and are often not suitable for large-scale information gathering problems. Closely related works on multi-robot coordination [4, 5] leverage Bayesian Optimization (BO) for estimating unknown spatial fields characterized by density functions with theoretical guarantees. However, these works present a synchronous planning mechanism that lacks optimal resource utilization. In this synchronous planning mechanism, robots that complete their tasks ahead of others are left idle until the arrival of new tasks, which entails higher mission times. Furthermore, they neither consider inter-robot collisions nor static collisions with obstacles in a continuous state space while gathering samples for a field estimation.

Large-scale information gathering problems require reasoning about realistic planning constraints, e.g., constraints among resources and constraints among robots' motion. To accomplish this task, we propose an Asynchronous Information Gathering with Bayesian Optimization (AsyncIGBO) algorithm that generates a distributed control law in continuous domains and an asynchronous communication strategy to guide a team of cooperative robots to visit the most informative locations within a limited mission duration. Our AsyncIGBO algorithm extends ideas from asynchronous parallel Bayesian optimization [6] to select the most informative locations within a limited mission duration and systematically combines them with decentralized reactive motion planning techniques [7]. Utilizing our proposed AsyncIGBO algorithm, robots can independently collect informative samples across a spatial field and share their information quickly and frequently through asynchronous communication for global decision-making.

In order to leverage the strength of a team of decentralized robots for the information gathering task, we first characterize the *informativeness* of observations from a spatial field as an unknown density function. We then need to estimate the density function from robot sensor measurements. Unfortunately, realistic sensor measurements from a spatial field are noisy and do not provide precise field information. Therefore, we utilize a Gaussian Process (GP) as the surrogate model to estimate an unknown density function because of its potent approximation properties and ability to quantify uncertainty from noisy sensor measurements. This density function is estimated over time by periodically gathering informative samples with multiple robots from the spatial field.

Manuscript received: August 26, 2022; Revised November 27, 2022; Accepted March 14, 2023. This paper was recommended for publication by Editor Hyungpil Moon upon evaluation of the Associate Editor and Reviewers' comments. This work is supported in part by the NSF grants IIS-2034123, IIS-2024733, and the U.S. Dept. of Homeland Security grant 2017-ST-062000002.

[1]A. A. R. Newaz is with the Department of Computer Science, University of New Orleans, New Orleans, LA 70148, USA (email: aredwann@uno.edu).

[2]M. Alsayegh and L. Bobadilla are with the Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, FL 33199, USA (email: {malsa061@, bobadilla@cs.}fiu.edu).

[3]T. Alam is with the Department of Computer Science, Louisiana State University Shreveport, Shreveport, LA 71115, USA (email: talam@lsus.edu).

Digital Object Identifier (DOI): see top of this page.

We make the following contributions in this paper: (i) For fast, frequent updates on the density function, we leverage the AsyncIGBO algorithm to choose the next informative samples for robots subject to their observed knowledge; (ii) Our algorithm learns GP hyperparameters to estimate the unknown density function from gathered samples with independently exploring multiple robots; (iii) Our algorithm generates a distributed control law in continuous domains while considering realistic robotic resource and motion constraints; (iv) We theoretically analyze the asymptotic no-regret properties of our algorithm with regard to a known spatial field; and (v) We evaluate our algorithm using simulated environments and hardware-in-the-loop simulations with physical robots.

II. Prior Work

Information gathering has been studied from a number of different perspectives for various applications. Classically, it is approached as a problem of geometric mapping via sensor networks. Several sensor networks utilize Bayesian Optimization (BO) for environmental monitoring [8]. However, sensor networks are neither accurate enough to accommodate large-scale demands nor cost-effective solutions for the information gathering process. A variant of BO called PLAyBOOK [6] is successful in hyperparameter tuning of Convolutional Neural Networks, where it utilizes asynchronous parallel Bayesian optimization. We herein extend the ideas from PLAyBOOK to efficiently sample from a scalar field and combine them with decentralized reactive motion planning techniques.

The core concept of informative path planning is to select a set of locations to visit at each time step in such a way that the missing information can be predicted using a statistical model. In the literature on informative path planning, random graph search [9], evolutionary optimization [10], heuristic search [11], and sampling-based approaches [1, 12] are applied. A popular statistical model for multi-robot information gathering algorithms is Gaussian processes (GPs) [13, 14, 15, 16, 17]. However, the majority of these works do not consider realistic conditions, where information collection and navigational imperfections for robots are inevitable.

Several prior online learning-enabled methods combine Bayesian Optimization (BO) [18, 19] with robotic motion planning algorithms to intelligently gather more samples over time from an area. These methods are primarily rooted in centralized coordination, where robots collect samples sequentially through continuous communication. In [4], a team of coordinated robots estimates and optimizes the coverage of an initially unknown spatial field characterized by a density function where robots communicate their locations. An adaptive learning goal of the density function in the unexplored areas along with the coverage goal is studied in [5]. However, in realistic multi-robot settings, there are also limits on resources, e.g., fuel or battery life. To deal with a limited mission time, there are some distributed multi-robot information gathering algorithms [20, 21, 22] that offer strategies to intelligently select robotic actions to efficiently gather information, but without theoretical performance guarantees.

Inspired by these works, our AsyncIGBO algorithm plans realistic robot trajectories in a distributed manner to efficiently visit informative locations while minimizing the mission completion time incurred and also provides theoretical performance guarantees.

III. Preliminaries and Problem Formulation

We consider a set of n robots $\mathcal{A}$ with sensing capability operating in a ν dimensional bounded domain $\mathcal{X} \subset \mathbb{R}^{\nu}$. We refer to this bounded domain as a "search space" over a spatial field where we define an a-priori unknown information density function $\varphi : \mathcal{X} \to \mathbb{R}$ based on the measurement model and sensor constraints. The goal of our multi-robot information gathering is to perform in situ sensing of the search space so that the states of the i^{th} robot state $\mathbf{x}_i(t)$ at time t are proportional to the unknown spatial field of $\varphi(.)$.

A. Robot Dynamics and Sensing Model

We want to find a batch of promising trajectories of individual robots in such a way that the statics of robots' trajectories is equivalent to the scalar statics of $\varphi(.)$ through an information metric. Assume the dynamics of the i^{th} robot is governed by a generic motion model as follows:

$$\dot{\mathbf{x}}_i(t) = f(\mathbf{x}_i(t), \mathbf{u}_i(t)), \tag{1}$$

where $\mathbf{u}_i(t)$ is the control to the i^{th} robot at time t. For the particular case of a nonholonomic robot that operates in a 2-D workspace, the robot's motion model is approximated by a unicycle system as:

$$\dot{x} = v\cos(\theta),\ \dot{y} = v\sin(\theta),\ \dot{\theta} = \omega, \tag{2}$$

where x and y correspond to the Cartesian coordinates of the robot, θ is the robot's orientation, v is the robot's forward velocity, and ω is the robot's angular velocity. Thus, the robot's state is defined as $\mathbf{x} = (x, y, \theta)^T$, and the robot's control input is defined as $\mathbf{u} = (v, \omega)^T$. When the i^{th} robot visits a location $x \in \mathcal{X}$ at time t, it can sense the event by obtaining a noisy measurement of the density function according to the sensing model as in [4]:

$$y_i = \varphi(\mathbf{x}_i) + \varepsilon, \tag{3}$$

where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ is the Gaussian white noise with σ standard deviation, and it is independent and identically distributed across time and space.

B. Field Characterization and Bayesian Optimization

In an unknown field, we do not have direct access to the global maximum of the information density function φ. Since noisy sensor readings introduce additional uncertainty to the system, we use a Gaussian Process (GP) as the surrogate model for φ. We define a GP prior ϕ to be $\mathcal{GP}\left(\mu(x;\rho), \kappa(x, x';\theta)\right)$, where μ is the mean function, κ is a kernel function, and ρ, θ are the hyperparameters of the mean and kernel functions, respectively. Now, the estimation of unknown spatial fields can be reduced to the problem of unknown hyperparameters estimation of GP. The posterior distribution of GP at an input location x is Gaussian as:

$$\mathbb{P}(\phi(x)|\rho,\theta) = \mathcal{N}(\phi(x); \mu(x;\rho), \sigma^2(x;\theta)) \tag{4}$$

with mean and variance

$$\mu(x) = \kappa(x, X)K(X, X)^{-1}Y, \quad (5)$$
$$\sigma^2(x) = \kappa(x, x') - \kappa(x, X)K(X, X)^{-1}\kappa(X, x'), \quad (6)$$

where K is the covariance function, X is a matrix with an input location in each row $x_1, x_2, \cdots, x_n$ and Y is a column vector of the corresponding observations $y_1, y_2, \cdots, y_n$. Given observations $D = (X, Y)$, we can condition our GP distribution on D as follows:

$$\mathbb{P}(\phi(x)|D; \rho, \theta) = \mathcal{GP}(\phi; \mu_{\phi|D}(x; \rho), \kappa_{\phi|D}(x, x'; \theta)). \quad (7)$$

To estimate hyperparameters of GP, Bayesian Optimization (BO) proceeds by maintaining a probabilistic belief about ϕ and designing an acquisition function α to determine where to evaluate the function next. Although many acquisition functions can be used in the framework of Bayesian decision, we use the GP Upper Confidence Bound (UCB) in our experiments. The UCB is defined as:

$$\alpha(x|D; \rho, \theta) = \mu(x|D; \rho) + \sqrt{\beta}\sigma(x|D; \theta), \quad (8)$$

where μ and σ are the mean and standard deviation of the GP posterior and $\sqrt{\beta}$ is the exploration constant.

C. Description of Obstacles

We have a set of n robots $\mathcal{A}$ sharing an environment with moving robots and/or static obstacles $\mathcal{O}$. Each static obstacle O_m with $m \in \{1, \cdots, N_o\}$ is assumed to be an a-priori known object in the search space. The free space where robots are allowed to explore in such an occupancy map is defined as follows:

$$\mathcal{X}_{\text{free}} = \mathcal{X} \setminus \bigcup_{m=1}^{N_o} O_m. \quad (9)$$

Next, the articulated robots must avoid inter-robot collisions which are defined in terms of a velocity obstacle [7]. Let $i \in \{1, \cdots, n\}$ be one robot that has a current state $\mathbf{x}_i$, a velocity $\dot{\mathbf{x}}_i$, and a radius ℓ_i, all of which are known to the robot and can be measured by the other robots in the environment. Let j be another robot that has a current state $\mathbf{x}_j$, velocity $\dot{\mathbf{x}}_j$, and radius ℓ_j and is considered as a dynamic obstacle moving in the environment. The velocity obstacle for the robot i presented by a dynamic obstacle j, denoted as $VO_{i|j}$, is the set of all velocities of robot i that will result in a collision between robot i and j at some future moment in time, considering the fact that a dynamic obstacle j maintains a constant velocity $\dot{\mathbf{x}}_j$. Formally, the velocity obstacle for the i^{th} robot presented by a j^{th} robot is written as follows:

$$VO_{i|j} = \{\dot{\mathbf{x}} \mid \exists t > 0 : t(\dot{\mathbf{x}}_i - \dot{\mathbf{x}}_j) \in \mathcal{D}(\mathbf{x}_j - \mathbf{x}_i, \ell_i + \ell_j)\}, \quad (10)$$

where $\mathcal{D}(\mathbf{x}, \ell)$ is an open disc of the radius ℓ centered at $\mathbf{x}$.

D. Behavior Tree

This work uses a Behavior Tree (BT) for multi-robot coordination purposes. The BT is a directed acyclic graph with four primary nodes: root nodes, composite nodes (selector, sequence, parallel), leaf nodes (condition and action), and decorator nodes. In the BT, the execution starts from a root node, then selector nodes execute tasks with priority, whereas sequence nodes execute tasks in a predetermined order. On the other hand, condition nodes are sensing units for action nodes since they activate actions when specific conditions are met, and decorator nodes modify the return value or running frequency. A detailed explanation of the BT can be found in [23].

E. Problem Statement

Gathering information from a spatial scalar field using multi-robot exploration requires making observations $y_1, \cdots, y_n$ over sequential locations $x_1, \cdots, x_n$ in such a way that maximizes the information gain, $\mathcal{I}$. The information gain over the i^{th} robot's visited locations $x_i^{(1:t)}$ is measured by the mutual information between ϕ and its gathered observations $y_i^{(1:t)}$ as:

$$\begin{aligned}\mathcal{I}\left(x^{(1:t)}; \phi\right) &= H\left(x^{(1:t)}\right) - H\left(x^{(1:t)}|\phi\right), \\ &= \frac{1}{2}\log|\mathbf{I} + \sigma^{-2}\mathbf{K}|,\end{aligned} \quad (11)$$

where H is the entropy function such that $H(\mathcal{N}(\mu, \Sigma)) = \frac{1}{2}\log|2\pi e\Sigma|$, $\mathbf{K}$ is a kernel matrix such that $\mathbf{K} = [K(s, s')]_{s,s' \in s^{1:t}}$, and $\mathbf{I}$ is an identity matrix. From a high level perspective, Eqn. (11) quantifies the reduction in uncertainty about ϕ from revealing $y^{(1:t)}$. Solving this task in a distributed and decentralized manner is difficult, which is why most earlier works [5, 4, 19] solve this sequential decision-making problem through either planning with centralized algorithms or planning in the discrete domain of the search space, as mentioned in Section II. In this context, the problem that we are interested in, however, is decentralized control strategies for a multi-robot system to collect maximum information $\mathcal{I}$ for a given period of time T such that $t \in [0, T]$ over the spatial scalar field ϕ. Specifically, at each time step t, we aim to find the optimal choice of control $\mathbf{u}_i(t)$ for the i^{th} robot that not only takes into account our current knowledge over ϕ, but also marginalizes previous observations accumulated over n robot trajectories. Formally, we formulate our problem of interest as follows:

Problem. *Given a team of n robots, find controls $\mathbf{u}_1(t), \cdots, \mathbf{u}_n(t)$ at each time step t to drive robots to a series of states $\mathbf{x}_0(t), \cdots, \mathbf{x}_n(t)$ that maximize the sum of information gained relative to the spatial scalar field ϕ as:*

$$\underset{\mathbf{u}_i(t), \cdots, \mathbf{u}_n(t)}{\arg\max} \sum_{i=1}^{n} \int_{t=1}^{T} \mathcal{I}\left[\mathbf{x}_i(0) + f\left(\mathbf{x}_i(t), \mathbf{u}_i(t)\right) dt; \phi\right],$$

subject to Eqn. (1)-(10).

IV. Decentralized Multi-robot Information Gathering Algorithm

To address the above problem, this paper proposes a distributed control law and an asynchronous communication strategy to guide a team of cooperative robots to visit the most informative locations within a limited mission duration. To deal with unknown parameters, our AsyncIGBO algorithm alternates among (i) the selection of informative batch locations that are evaluated in parallel, (ii) the estimation of parameters from collected observations asynchronously, and (iii) Bayesian inference, assuming that the estimated parameters are true ones. In this section, the time step from equations has been dropped for clarity.

A. Informative Location Sampling

Since it is costly to gather information by visiting all locations of φ, our algorithm selects a set of informative locations to visit at some time steps and uses the GP model to predict the missing values. Inspired by the work [6], we develop an asynchronous Bayesian optimization strategy to select the next sample locations while maximizing the following objective function:

$$x_i^{\text{goal}} = \underset{x_i \in \mathcal{X}^i_{\text{neigh}}}{\arg\max} \left\{ \alpha\left(x_i|D\right) \prod_{j=1}^{n} \psi(x_i|D) \right\}, \tag{12}$$

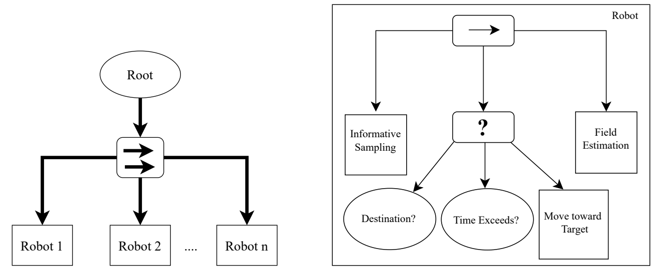
where $\mathcal{X}^i_{\text{neigh}}$ is the i^{th} robot's local search space such that $\mathcal{X}^i_{\text{neigh}} \subset \mathcal{X}^i_{\text{free}}$, α is acquisition function, and ψ is a penalization function which facilitates the sampling diversity among the robots. On a high level, Eqn. (12) assigns a goal location to the i^{th} robot in the vicinity of its neighboring locations in such a way that it maximizes information gain through the acquisition function α and penalizes j^{th} robots such that $j \neq i$ to collect information from these locations. In the cooperative information gathering context, the penalization function needs to eliminate the risk of repeating sampling in the previously gathered informative locations or sampling near a busy location that is currently being evaluated by another robot. To develop an exploration algorithm for a team of decentralized robots, we assume that the unknown density function φ with $\varphi(x^*) = M$ is the Lipschitz function over current gathered observations D with a constant L and a global optimum value M. For the i^{th} robot, let $r_i = \frac{M-\varphi(x_i)}{L}$ be the associated sensing radius to the position x_i. To satisfy cooperative information gathering requirements, we propose a simple yet effective penalization function as:

$$\psi(x_i|x_j, y_j) = \frac{||x_i - x_j||}{\mathbb{E}(r_i) + \sqrt{\beta}\frac{\sigma(x_i)}{L}}, \tag{13}$$

where β is a positive value such that $\beta > 0$, and $\mathbb{E}(r_i)$ is the expected value of r_i over an approximated field ϕ such that $\mathbb{E}(r_i) = \frac{M-\phi(x_i)}{L}$. Section V provides further analysis of Eqn. (12).

B. Generating Collision-free Trajectories

Once we choose informative sampling locations, our next goal is to generate collision-free trajectories for n robots by independently choosing their controls. It is important to note that robots choose their control velocities as local, reactive navigation strategies in a decentralized fashion without any communication or central coordination. One inspiring method we use to implement such navigation strategies is the Hybrid Reciprocal Velocity Obstacle (HRVO) [7] in the presence of static obstacles and other dynamic robots in the environment. In the HRVO method, each robot chooses a new velocity at each time step to compute a trajectory toward its goal avoiding collisions with any other robots or static obstacles.



(a) High-level BT (b) Expanded sub-BT

Fig. 1: BT-based distributed multi-robot control architecture.

Each robot i has a goal state $\mathbf{x}_i^{\text{goal}}$ and a preferred velocity $\dot{\mathbf{x}}_i^{\text{pref}}$ which are unknown to the other robots. The preferred velocity is the velocity that a robot takes regardless of other robots' velocities or obstacles to direct itself toward its goal. Formally, a preferred velocity for each robot i toward its goal centered at $\mathbf{x}_i^{\text{goal}}$ is defined as:

$$\dot{\mathbf{x}}_i^{\text{pref}} = \dot{\mathbf{x}}_i^{\text{pref}} \frac{\mathbf{x}_i - \mathbf{x}_i^{\text{goal}}}{\|\mathbf{x}_i - \mathbf{x}_i^{\text{goal}}\|_2}, \tag{14}$$

where $\dot{\mathbf{x}}_i^{\text{pref}}$ is the constant preferred speed of that robot.

The combined hybrid reciprocal velocity obstacle, denoted as $HRVO_i$, for a robot i presented by all other robots and static obstacles in the environment is the union of all hybrid reciprocal velocity obstacles presented by the other robots individually and all velocity obstacles generated by the obstacles that can be written as

$$HRVO_i = \bigcup_{\substack{j \in \mathcal{A} \\ i \neq j}} HRVO_{i|j} \cup \bigcup_{j \in \mathcal{O}} VO_{i|j}. \tag{15}$$

Therefore, each robot i selects its next control $\mathbf{u}_i^*$ outside of the combined hybrid reciprocal velocity obstacle that is closest to its preferred velocity $\dot{\mathbf{x}}_i^{\text{pref}}$ as:

$$\mathbf{u}_i^* = \underset{\mathbf{u}_i}{\arg\min} \|f(\mathbf{x}_i, \mathbf{u}_i) - \dot{\mathbf{x}}_i^{\text{pref}}\|_2 \text{ s.t. } f(\mathbf{x}_i, \mathbf{u}_i) \notin HRVO_i. \tag{16}$$

C. Asynchronous Information Gathering with Bayesian Optimization (AsyncIGBO) Algorithm

We utilize the Behavior Tree (BT) to capture the teaming and coordination of decentralized robots to collectively observe an unknown spatial field. The BT's inherent hierarchical and modular structure allows us to combine the informative sampling strategy with asynchronous BO and a systematic way of generating collision-free trajectories through HRVO. We now propose the AsyncIGBO algorithm which can simultaneously plan and control using the BT by mapping the high-level sampling of informative locations, which describes which locations to visit, onto a low-level control model, which defines how to generate trajectories for robots.

Fig. 1 shows our BT-based distributed control architecture, where all robots are connected through a parallel node of the BT. Each robot behavior is defined as a sub-BT and shown separately in Fig. 1(b). Each robot performs three sequential actions: 1) sampling of informative locations from a spatial field ϕ, 2) moving toward a target location assigned by the informative sampler, and 3) estimating the field from gathered observations. First, the reactivity of BTs enables a robot to reason about the maximum informative location based on the current knowledge over ϕ. Once a high-level informative location, x_i^{goal}, is planned for the i^{th} robot, the modularity of the BT enables generating low-level controls using HRVO to make progress toward the x_i^{goal} location without needing to plan collision-free trajectories for others. Asynchronous communication among robots is established using a selector node. As it is illustrated in Fig. 1(b), the selector nodes of sub-BTs need to return success to update the field estimation through communication. However, the selector node returns success when a robot reaches its target destination or the budgeted time for executing the assigned task is exceeded. Otherwise, the selector node returns the running status due to the move toward the target node. Finally, when the i^{th} robot accomplishes its task by either visiting the assigned location or exceeding the time budget for that assignment, the reactivity of the BT enables asynchronously updating of the surrogate model ϕ through its current observation y_i and sharing this observation y_i with others through communication. This way robots progressively learn the latent phenomena by updating GP parameters from gathered observations. Formally, after the desired number of robots $c < n$ complete their tasks, we assign new tasks to them without waiting for the remaining $b = (n - c)$ busy robots to finish their tasks by estimating GP parameters as:

$$(\rho, \theta) = \arg\max_{\rho,\theta} \prod_{i=1}^{n} \prod_{l=1}^{b} \mathbb{P}(y_i, y_l | x_i, x_l, D, D_b, D_i; \rho, \theta), \tag{17}$$

where the set D contains the locations associated with observations that robots have gathered so far, the set D_b contains the busy locations and corresponding observations which are not accessible during the time of decision-making, and the set D_i contains the current locations and corresponding observations obtained through communication among all robots. Thus, the interleaved plan-and-control process of the BT enables decentralized mobile robots to find maximally informative trajectories.

V. Algorithm Analysis

In this section, first, we assume that the unknown density function φ can be modeled as the realization of GP [4]. Second, we also assume that φ holds Lipschitz continuity [24] and then theoretically analyze the asymptotic no-regret properties of our proposed AsyncIGBO algorithm with respect to a true density function, ϕ. Using the Lipschitz continuity property, at each step, the i^{th} robot selects a sample from its neighboring locations that eliminates the largest possible portion of the input space while guaranteeing, with high probability, that the eliminated part does not include the maximizer of the function [24].

Lemma 1. *Let the unknown objective function be a Lipschitz continuous function with a constant L which has a global optimum value M. Let r_i be the associated penalization radius of the i^{th} robot to the location x_i such that $r_i = \frac{|\phi(x_i) - M|}{L}$ and let $\mathbb{E}(r_i)$ be the expected value of r_i such that $\mathbb{E}(r_i) = \frac{|\mu(x_i) - M|}{L}$. Then,*

$$\mathbb{P}\left[r_i \leq \mathbb{E}(r_i) + \epsilon\right] \leq \exp\left(-\frac{2\epsilon^2 L^2}{\sigma^2}\right),$$

holds for $\epsilon = \sqrt{\beta}\frac{\sigma(x_i)}{L}$.

Sketch of Proof. We assume the unknown objective function is Lipschitz continuous with a constant L and has a global optimum value M. Hence, the following relationship holds as:

$$|\phi(x_i) - M| \leq L||x_i - x^*||. \tag{18}$$

This implies x^* cannot lie within the circular region centered on x_i with a radius r_i. In the remainder of the proof, we first recall a result from [4, Lemma IV.5] which states that if μ is an approximation of the true density function, ϕ, then

$$|\phi(x_i) - \mu(x_i)| \leq \sqrt{\beta}\sigma(x_i) \tag{19}$$

holds for any x_i any $t > 1$ with probability at least $1 - \delta$. Now, adding and subtracting M in Eqn. (19), we have

$$\begin{aligned} &|M + \phi(x_i) - M - \mu(x_i)| \leq \sqrt{\beta}\sigma(x_i) \\ =&|M - \phi(x_i)| \leq |\mu(x_i) - M| + \sqrt{\beta}\sigma(x_i) \\ =&\frac{|M - \phi(x_i)|}{L} \leq \frac{|\mu(x_i) - M|}{L} + \sqrt{\beta}\frac{\sigma(x_i)}{L} \\ =&r_i \leq \mathbb{E}(r_i) + \sqrt{\beta}\frac{\sigma(x_i)}{L} \end{aligned}$$

Since $\phi(x_i)$ is a normal random variable $\mathcal{N}(\mu, \sigma^2)$, using Hoeffding inequality for all $\epsilon > 0$, we can then say from [24] that

$$\mathbb{P}\left[r_i \leq \mathbb{E}(r_i) + \epsilon\right] \leq \exp\left(-\frac{2\epsilon^2 L^2}{\sigma^2}\right). \qquad \square$$

Thus, the penalization function in (12) guides the i^{th} robot to sample the next asynchronous locations by reducing the value of the acquisition function at those locations centered on x_i with an expected radius $\mathbb{E}(r_i)$.

We now show that, under the assumptions of discrete observation set D and a known field ϕ with respect to hyperparameters, the regret between the optimal locational utility and the utility efficiently obtained through the proposed algorithm is sub-linear in T. Since our locational utility function is a combination of acquisition and penalization functions, we start with the recent regret analysis results in [4, Lemma IV.5] to illustrate a sub-linear property of the acquisition function (8), and then we combine the Lemma 1 to provide a precise bound over the penalization function. At a high level, knowing a penalization region allows us to reduce regrets by avoiding redundant sampling.

Theorem 1. *Let R be the regret at stage $t \in [1, T]$. If the field ϕ is the realization of a GP with mean μ and covariance κ, our proposed algorithm ensures that*

$$\int_{t=1}^{T} R(t)dt \leq \int_{t=1}^{T} c_0 \frac{\sqrt{\beta(t)}\sigma(t-1)(x_i)}{\exp\left(-\frac{2\epsilon^2 L^2}{\sigma(t-1)^2}\right)} dt,$$

where $\beta(t) = 2\log(|D|\pi_t/\delta)$, and $\sum_{t\geq 1} \pi_t^{-1} = 1, \pi_t > 0$.

Sketch of Proof. We can rewrite Eqn. (12) as:

$$x_i^{\text{goal}} = \underset{x_i \in \mathcal{X}^i_{neigh}}{\arg\max}\ \Upsilon(\phi_t, x_i),$$

where Υ is the locational utility that captures the need for the robots to collectively move toward some locations in order to obtain the maximum information based on the current estimation of field ϕ_t as:

$$\Upsilon(\phi_t, x_i) = \alpha_t\left(x_i | y_i = \phi_t(x_i)\right) \prod_{j=1}^{n} \Psi_t(x_i | x_j, y_j = \phi_t(x_j)),$$

Let $\Upsilon^* = \max_{x_i \in \mathcal{X}^i_{neigh}} \Upsilon(\phi, x_i^*)$ be the optimal locational utility and $x_i^* = \arg\max_{x_i \in \mathcal{X}^i_{neigh}} \Upsilon(\phi, x_i)$ be the optimal choice of a sample location with respect to a known density function ϕ. The regret $R(t)$ at time t is then the difference between the locational utility with the current approximation of the field and the optimal locational utility that is written as: $R(t) = \Upsilon(\phi_t, x_i) - \Upsilon^*$, such that

$$\Upsilon(\phi_t, x_i) \leq \Upsilon(\phi, x_i^*) \leq \Upsilon^*.$$

From Eqns. (13) and (12), we get:

$$R(t) = \Upsilon(\phi, x_i) - \Upsilon^* \leq \Upsilon(\phi, x_i) - \Upsilon(\phi_t, x_i) \quad (20)$$

$$= \sum_i \sum_{x\in D} \frac{||x - x_i||}{r_i - \left(\mathbb{E}(r_i) + \sqrt{\beta(t)}\frac{\sigma(t-1)(x_i)}{L}\right)} \left(\alpha(x) - \alpha_t(x)\right) \quad (21)$$

Eqn. (21) shows two primary sources of regrets: estimation error due to field approximation, and penalization error due to approximation of the penalization radius. From [4, Lemma IV.5], we can replace the error in the field approximation as:

$$R(t) \leq \sum_i \sum_{x\in D} ||x - x_i|| \cdot \frac{\sqrt{\beta(t)}\sigma(t-1)(x_i)}{r_i(t) - \left(\mathbb{E}(r_i(t)) + \sqrt{\beta(t)}\frac{\sigma(t-1)(x_i)}{L}\right)}$$

Let $c_0 = |D| \max_{x, x_i} \sum_i \sum_{x\in D} ||x - x_i||$. From Lemma 1, we can replace the penalization error as:

$$R(t) \leq c_0 \frac{\sqrt{\beta(t)}\sigma(t-1)(x_i)}{\exp\left(-\frac{2\epsilon^2 L^2}{\sigma(t-1)^2}\right)}.$$

Hence,

$$\int_{t=1}^{T} R(t)dt \leq \int_{t=1}^{T} c_0 \frac{\sqrt{\beta(t)}\sigma(t-1)(x_i(t))}{\exp\left(-\frac{2\epsilon^2 L^2}{\sigma(t-1)^2}\right)} dt.$$

□

Remark 1. *Since the global Lipschitz constant L of the objective function is unknown, we can approximate L locally for each busy location of robots. In addition, the global optimum M is unknown in practice and this can also be approximated by the best observed function value.*

VI. Simulations and Experiments

In this section, we describe our simulations and real-world experiments deploying our AsyncIGBO algorithm in a variety of spatial fields for information gathering purposes. Our experiments compare the AsyncIGBO algorithm to a synchronous IGBO algorithm, evaluate their performance with different sizes of teams, and study their regrets in various environments.

a) Simulation Results: We simulate a team of cooperative robots in a large-scale environmental monitoring setting. We conduct our simulations in a geographic area which is 50 m long on the x-axis and 20 m long on the y-axis. The environment consists of several static obstacles as shown with black circles in Fig. 2. Inspired by [20], the scalar field is generated according to a sum of Gaussian functions as:

$$\varphi(x) = \sum_{l=1}^{3} \Gamma_l \exp\left(-\frac{||\Lambda_l - x||}{\Omega_l}\right), \quad (22)$$

where Γ_l values are set to $\{0.33, 0.33, 0.33\}$, Λ_l values are set to $\{(-35, 3.0), (-16.5, -2.3), (0, 2.3)\}$, and Ω_l values are set to $\{(20.7, 20.7), (15.0, 15.0), (9.9, 9.9)\}$, for $l \in \{1, 2, 3\}$, respectively. The resultant scalar field is shown with a gray scale background image in Fig. 2. For this simulation, three robots are initiated with random states as shown in Fig. 2(a), and their sensing radii are set to 3 m. Our AsyncIGBO algorithm generates the next best sample location in the proximity of each robot based on the current observations over the field. Each robot asynchronously accomplishes its task by visiting the assigned location and obtaining the sensor measurements from respective locations. Therefore, the AsyncIGBO algorithm utilizes this information without waiting for other robots to finish their tasks. The implementation of the HRVO method helps to generate realistic motions for decentralized robots by avoiding static obstacles and inter-robot collisions, as shown in Fig. 2(b). Finally, each robot tries to minimize the cumulative regrets by visiting the most informative locations first. As we can see from Fig. 2(c), robots prioritize visiting the light gray regions where information gains are higher than in the dark gray regions.

b) Physical Experiment Results: We utilize two autonomous mobile robots to validate our algorithm. Due to the limited space and the size of the robot platform (shown in Fig. 3), these experiments are conducted without any static obstacles in an indoor laboratory environment.

In this experiment, we utilize a real-world ROMS water salinity dataset [25] to generate a spatial scalar field. We then project this scalar field to a 2 m $\times$ 2 m workspace as illustrated in Fig. 3(b). Thus, a query is made with a two-dimensional location as an input to obtain a scalar measurement (output). The three major components that drive our decentralized robots are (a) an Async-BO probabilistic model with the information gain objective, (b) an HRVO for local trajectory planning, and (c) a BT for global coordination. Our Async-BO takes two-dimensional sampling locations as inputs and predicts the next best sampling location based on a robot's current position. Next, our local trajectory planner generates a collision-free trajectory to drive the robot to its current destination. Unlike existing works which only focus on point-based information

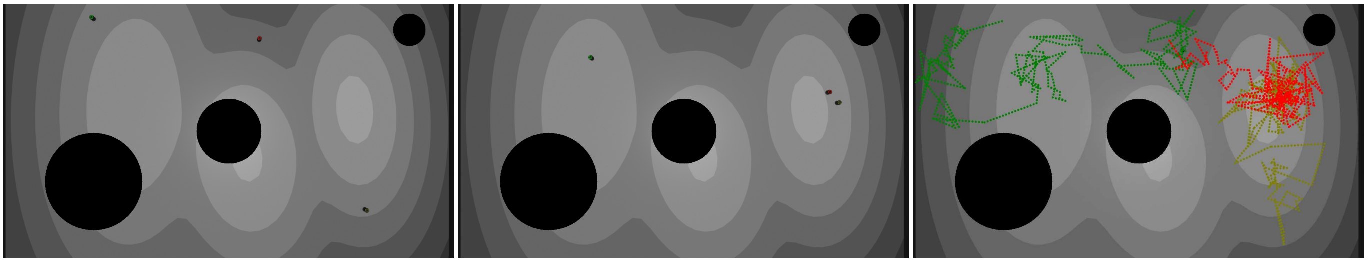


Fig. 2: **Decentralized multi-robot information gathering simulation**: Three robots starting with random initial states are visiting the most informative regions on a spatial field while avoiding not only static obstacles but also inter-robot collisions.

Fig. 3: **Physical experiment with multiple robots**: Two robots are collecting information from a projected spatial field.

gathering [4, 1], our work considers a realistic safe motion model to collect information from a given area. In some cases, an informative sample location from the Async-BO might not be feasible to explore, due to inter-robot collisions. Therefore, we introduce a time-budget in our robotic information gathering procedure using the BT-based coordination mechanism. Once the robot reaches its destination or the robot gets stuck in a loop when it tries to satisfy collision avoidance requirements, the robot communicates its current measurement and location information with others using the BT.

We conduct this experiment multiple times and report the average results. The AsynIGBO quickly identifies the most informative locations with few samples collected by our decentralized robot team, leading to faster convergence in the regret metric. The spatial field is represented with a colored image projected over the geometric space, where yellowish colors represent the most informative regions and bluish colors indicate the least informative regions. We observe that the robots using the AsynIGBO algorithm effectively navigate to the most informative locations to minimize cumulative regrets while generating a good approximation of the field. A video related to this experiment can be found at https://youtu.be/I86su28mdJc.

c) Performance Analysis: In this paragraph, we address two questions: (i) the scalability of our proposed AsyncIGBO algorithm and (ii) a comparison between asynchronous and synchronous IGBO algorithms. We evaluate the performance of our algorithm with global optimization of Eqn. (22). In this setting, we are interested in computing regrets with respect to the global optimum value. To facilitate this comparison, we consider a measure of seconds for performing the same task with different numbers of robots. For each experiment, robots are initialized at random states, and then the experiment is repeated 10 times.

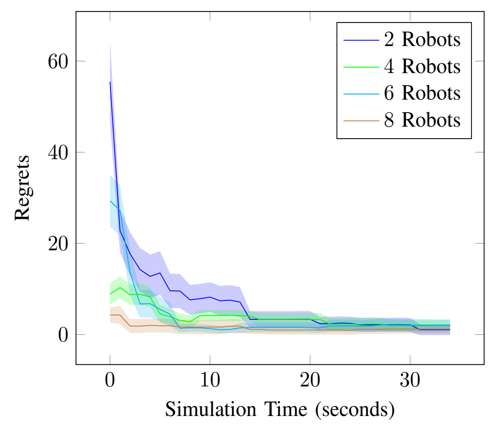


Fig. 4: Cumulative regrets of the AsynIGBO algorithm decrease with an increase in the number of robots.

Fig. 4 shows the regrets of robots with respect to simulation times for teams of $2, 4, 6$, and 8 robots, respectively. As illustrated in Fig. 4, a team of 2 robots takes the longest time to reach a convergence level. Hence, cumulative regrets for a team of 2 robots are higher. On the other hand, a team of 8 robots takes the shortest time to reach a convergence level. Hence, cumulative regrets for a team of 8 robots are lower. Thus, cumulative regrets of the AsyncIGBO algorithm decrease with an increase in the number of robots, which shows that our algorithm is scalable for a large team of robots. To predict a complex spatial field using our AsyncIGBO algorithm, we need to choose a suitable GP kernel to efficiently collect samples over the target area. Table I demonstrates

# robots	Matérn Kernel		RBF Kernel	
	# evaluations	mean regret	# evaluations	mean regret
3	**79**	**0.528**	72	2.52
5	29	**0.093**	**16**	0.54
7	**48**	**2.818**	51	4.26

TABLE I: **Sample efficiency comparison:** Matérn Kernel is more sample efficient – requires fewer numbers of evaluations to converge, than the RBF Kernel.

	2 robots	**4 robots**	**6 robots**	**8 robots**
SyncIGBO	68.5 ± 4.9	17.9 ± 4.1	3.7 ± 2.1	2.2 ± 1.8
AsyncIGBO	3.6 ± 2.3	2.3 ± 1.3	2.8 ± 1.8	1.4 ± 1.0

TABLE II: **Performance comparison:** Our AsyncIGBO outperforms a SyncIGBO in terms of lower mean regrets.

a comparison between two popular GP kernels, Matérn and Radial Basis Function (RBF) kernels. As it is obvious from Table I, the Matérn kernel requires fewer sample evaluations to reach convergence and also achieves lower mean regrets than the RBF kernel. Thus, we utilize the Matérn kernel in all of our experiments.

The recent closest work to our algorithm has utilized a synchronous IGBO (SyncIGBO) algorithm in a discrete domain setting [4]. We compare asynchronous and synchronous IGBO algorithms by performing head-to-head regret comparisons with different numbers of robots. To maximize information gain within a limited mission time, we need to reason about (i) how much cumulative area should be covered by all robots collectively and (ii) how much traffic congestion will be introduced by increasing the number of robots. Although information gathering using multiple robots benefits from our distributed control architecture, inter-robot collision avoidance needs to be taken into account for safe deployments of robots. Our mean regrets do not monotonically decrease with an increase in the number of robots since robots may take longer trajectories to safely collect information. However, our asyncIGBO algorithm outperforms the syncIGBO algorithm in every scenario, as demonstrated in Table II.

VII. Conclusion

This paper presents a scalable planning algorithm to find trajectories for decentralized robots used for gathering the maximum information from unknown spatial fields. The informativeness of observations from the spatial field is characterized as a density function. Our proposed AsyncIGBO algorithm efficiently chooses the next best informative location for a team of decentralized robots. We then present a distributed reactive planning and control technique to generate continuous motions for robots to safely visit the informative locations within a limited mission period. We evaluate our algorithm by theoretically analyzing its asymptotic no-regret properties with respect to a known spatial field. Finally, we validate our algorithm through several simulation and physical experiments with multiple robots. We can conclude from our results that our AsyncIGBO algorithm is scalable to an increasing number of robots and outperforms the SyncIGBO algorithm in terms of mean regrets.

References

[1] G. A. Hollinger and G. S. Sukhatme, "Sampling-based robotic information gathering algorithms," *Int. J. Robot. Res.*, vol. 33, no. 9, pp. 1271–1287, 2014.

[2] R. Almadhoun, T. Taha, L. Seneviratne, and Y. Zweiri, "A survey on multi-robot coverage path planning for model reconstruction and mapping," *SN Appl. Sci.*, vol. 1, no. 8, pp. 1–24, 2019.

[3] A. A. R. Newaz, T. Alam, J. Mondello, J. Johnson, and L. Bobadilla, "Multi-robot information gathering subject to resource constraints," in *Proc. IEEE Int. Conf. Robot Human Interact. Commun.*, pp. 1–6, 2021.

[4] A. Benevento, M. Santos, G. Notarstefano, K. Paynabar, M. Bloch, and M. Egerstedt, "Multi-robot coordination for estimation and coverage of unknown spatial fields," in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 7740–7746, 2020.

[5] M. Santos, U. Madhushani, A. Benevento, and N. E. Leonard, "Multi-robot learning and coverage of unknown spatial fields," in *Proc. IEEE Int. Symp. Multi-Robot Multi-Agent Syst.*, pp. 137–145, 2021.

[6] A. Alvi, B. Ru, J.-P. Calliess, S. Roberts, and M. A. Osborne, "Asynchronous batch Bayesian optimisation with improved local penalisation," in *Proc. Int. Conf. Mach. Learn.*, pp. 253–262, 2019.

[7] J. Snape, J. Van Den Berg, S. J. Guy, and D. Manocha, "The hybrid reciprocal velocity obstacle," *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 696–706, 2011.

[8] R. Marchant and F. Ramos, "Bayesian optimisation for intelligent environmental monitoring," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 2242–2249, 2012.

[9] R. Marchant, F. Ramos, S. Sanner, *et al.*, "Sequential Bayesian optimisation for spatial-temporal monitoring," in *Proc. Conf. Uncert. Artif. Intell.*, pp. 553–562, 2014.

[10] M. Popović, G. Hitz, J. Nieto, I. Sa, R. Siegwart, and E. Galceran, "Online informative path planning for active classification using uavs," in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 5753–5758, 2017.

[11] L. Bottarelli, M. Bicego, J. Blum, and A. Farinelli, "Orienteering-based informative path planning for environmental monitoring," *Engg. Appl. Artif. Intell.*, vol. 77, pp. 46–58, 2019.

[12] M. Ghaffari Jadidi, J. Valls Miro, and G. Dissanayake, "Sampling-based incremental information gathering with applications to robotic exploration and environmental monitoring," *Int. J. Robot. Res.*, vol. 38, no. 6, pp. 658–685, 2019.

[13] A. Viseras, Z. Xu, and L. Merino, "Distributed multi-robot cooperation for information gathering under communication constraints," in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 1267–1272, 2018.

[14] K.-C. Ma, L. Liu, and G. S. Sukhatme, "Informative planning and online learning with sparse gaussian processes," in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 4292–4298, 2017.

[15] K.-C. Ma, L. Liu, H. K. Heidarsson, and G. S. Sukhatme, "Data-driven learning and planning for environmental sampling," *J. Field Robot.*, vol. 35, no. 5, pp. 643–661, 2018.

[16] S. A. Zanlongo, L. Bobadilla, D. McDaniel, and Y. T. Tan, "Development of informative path planning for inspection of the Hanford tank farm," in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 2297–2303, 2019.

[17] A. Singh, A. Krause, C. Guestrin, W. J. Kaiser, and M. A. Batalin, "Efficient planning of informative paths for multiple robots," in *Proc. Int. Joint Conf. Artif. Intell.*, pp. 2204–2211, 2007.

[18] W. Luo and K. Sycara, "Adaptive sampling and online learning in multi-robot sensor coverage with mixture of gaussian processes," in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 6359–6364, 2018.

[19] W. Luo, C. Nam, G. Kantor, and K. Sycara, "Distributed environmental modeling and adaptive sampling for multi-robot sensor coverage," in *Proc. Int. Conf. Auton. Agents Multi-Agent Syst.*, pp. 1488–1496, 2019.

[20] T. X. Lin, S. Al-Abri, S. Coogan, and F. Zhang, "A distributed scalar field mapping strategy for mobile robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 11581–11586, 2020.

[21] G. Best and G. A. Hollinger, "Decentralised self-organising maps for multi-robot information gathering," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2020.

[22] A. Viseras, Z. Xu, and L. Merino, "Distributed multi-robot information gathering under spatio-temporal inter-robot constraints," *Sensors*, vol. 20, no. 2, p. 484, 2020.

[23] M. Colledanchise and P. Ögren, *Behavior Trees in Robotics and AI: An Introduction*. CRC Press, 2018.

[24] A. Jalali, J. Azimi, X. Fern, and R. Zhang, "A lipschitz exploration-exploitation scheme for Bayesian optimization," in *Euro. Conf. Mach. Learn. Know. Disc. Datab.*, pp. 210–224, 2013.

[25] A. F. Shchepetkin and J. C. McWilliams, "The Regional Oceanic Modeling System," *Ocean Modelling*, vol. 9, no. 4, pp. 347–404, 2005.