2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Decentralized Multi-agent Exploration with Limited Inter-agent Communications

Hans J. He[1], Alec Koppel[2], Amrit Singh Bedi[3], Daniel J. Stilwell[1], Mazen Farhood[4], and Benjamin Biggs[1]

***Abstract*—We consider the problem of decentralized multi-agent environmental learning through maximizing the joint information gain among a team of agents. Inspired by subsea applications where bandwidth is severely limited, we explicitly consider the challenge of restricted communication between agents. The environment is modeled as a Gaussian process (GP), and the global information gain maximization problem in a GP is a set-valued optimization problem involving all agents' locally acquired data. We develop a decentralized method to solve it based on decomposition of information gain and exchange of limited subsets of data between agents. A key technical novelty of our approach is that we formulate the incentives for information exchange among agents as a submodular set optimization problem in terms of the log-determinant of their local covariance matrices. Numerical experiments on real-world data demonstrate the ability of our algorithm to explore trade-off between objectives. In particular, we demonstrate favorable performance on mapping problems where both decentralized information gathering and limited information exchange are essential.**

I. Introduction

We consider the problem of decentralized multi-agent environmental learning. Our work is motivated by the specific challenges of underwater applications where extremely low bandwidth acoustic signals are typically used for communication between agents [1]. Communication constraints pose a challenge for decentralized learning, namely, agents are hampered in their ability to construct a global model due to only knowing a subset of their neighbors' measurements. In addition, limited knowledge about their neighbors' measurements affects an agent's decision-making when attempting to optimize for the objective of information gathering. Because agents make decisions based on information communicated by other agents, we desire agents' local models be statistically close to each other. We approximate this as an additional criterion such that each agent must consider the set of shared data between it and its neighbors when selecting new measurement locations. The problem is posed as a multi-objective submodular set maximization problem, and it is addressed using a greedy algorithm. Our approach is illustrated using simulations on real world data.

We model the environment f as a Gaussian process (GP). Our focus on GP models are due to their ability to model probabilistic hypotheses efficiently in that their posterior parameters admit a closed-form update [2] and they define a natural notion of uncertainty associated with the environment. GP models for environmental characteristics that are approximately spatially continuous, such as water temperature [3], salinity [4], etc., are fairly common in the literature.

The question of where to sample data to train the GP most efficiently has been explored in the related problem of sensor placement [5], [6]. These studies seek to maximize information gain at a fixed set of locations, which can be derived from the GP posterior covariance. In the frequentist setting, one can solve the GP mean representation and learning problem through suitable modifications of stochastic gradient iteration together with consensus protocol [7]–[9]. We focus on Bayesian settings due to their more natural representation of parameter uncertainty conditioned on past information. In applied settings, online distributed learning using GPs have been successfully applied to autonomous air and ground robots [10]–[14], as well as exploration problems for single agent surface vehicles [15].

When resource constraints on communication between agents are present, periodic network disconnection has been put forth as an initial solution [16], [17]. Alternatively, coding schemes that reduce the number of bits needed for fusion of local Gaussian process models have been developed [18], but mandate that periodic synchronization of each local machine to a centralized node is required.

Few works consider bandwidth limitations when exchanging information between decentralized Gaussian process models. Towards allowing full decentralization, a dissimilarity measure is devised in [19] to select which subsets of data agents must communicate in their GP models, but this does not address multi-agent control. Others have developed communication criteria in service of multi-agent objectives outside of exploration. In both [20] and [21], Euclidean distance is used as a criterion for selecting subsets of measurements to communicate in distributed mobile sensor networks. However, their objectives were to minimize prediction error at fixed locations of interest and to aid multi-agent coordination for patrolling missions, respectively.

By contrast, our work considers the setting where agents collaboratively learn an environmental map by selecting measurement locations that seek to maximize information

This material is based upon work supported by the Office of Naval Research (ONR) under Awards N00014-18-1-2627 and N00014-19-1-2194.

[1]Hans J. He, Daniel Stilwell, and Benjamin Biggs are with the Bradley Department of Electrical and Computer Engineering at Virginia Tech, Blacksburg, VA 24060, USA. Email: {hjh2bs, stilwell, babiggs}@vt.edu

[4]Mazen Farhood is with the Kevin T. Crofton Department of Aerospace and Ocean Engineering at Virginia Tech, Blacksburg, VA 24061, USA. Email: farhood@vt.edu

[2]Alec Koppel is with JP Morgan AI Research, New York, NY. Email: alec.koppel@jpmchase.com

[3]Amrit Singh Bedi is with the Institute of Systems Research at the University of Maryland, College Park, MD 20742, USA. Email: amritbd@umd.edu

gain. We introduce a constraint on the difference between neighboring agents' statistical models, then employ results from [22] and [23] to further develop the constraint as a submodular set function on the agents' shared measurement data. We develop an algorithm allowing agents to make decisions by jointly optimizing for information gain and statistical model similarity based on locally acquired information and information that is exchanged with neighbors. The contributions of this work are (1) the formulation of the decentralized exploration problem subject to communication and model similarity constraints, (2) the approximation of the aforementioned problem as a multi-objective submodular optimization problem, and (3) an algorithm to solve the optimization problem.

In Section II, we describe the mathematical background necessary for formulating the problem. In Section III, we describe the multi-agent setting considered, and then we formulate the exploration problem under bandwidth limitations. In Section IV, we derive a local penalty functional for approximating the problem, and provide an algorithm addressing the approximated problem. Experimental results are given in Section V, and concluding remarks are provided in Section VI.

II. Preliminaries

In this section, we summarize the mathematical background needed to formulate our problem. For additional details on Gaussian process regression and information theory, please refer to [24] and [25], respectively.

A. Gaussian Processes

We consider the problem of learning a spacial field $f : \mathbb{R}^d \to \mathbb{R}$ from a sequence of measurement pairs $(\mathbf{x}_t, y_t)$ such that $\mathbf{x}_t \in \mathcal{X} \subset \mathbb{R}^d$ and $y_t \in \mathbb{R}$, where $y_t = f(\mathbf{x}_t) + \varepsilon$ and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ is the measurement noise. We assume that the latent function f is a Gaussian process denoted by $GP(\bar{f}(\mathbf{x}), k(\mathbf{x}, \mathbf{x}))$ where $\bar{f} : \mathcal{X} \to \mathbb{R}$ denotes the mean function and $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ describes the covariance (kernel) function. GPs are collections of random variables, any finite number of which form a joint Gaussian distribution. Therefore, a GP evaluated at a finite set of locations $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathcal{X}$ induces a multivariate normal distribution and with a slight abuse of notation, we can write $f(\mathbf{X}) \sim \mathcal{N}(\bar{f}_\mathbf{X}, \mathbf{K}_\mathbf{X})$. The mean vector $\bar{f}_\mathbf{X} \in \mathbb{R}^N$ and the covariance matrix $\mathbf{K}_\mathbf{X} \in \mathbb{R}^{N \times N}$ are defined element-wise as $[\bar{f}_\mathbf{X}]_i = \bar{f}(\mathbf{x}_i), \mathbf{x}_i \in \mathbf{X}$ and $[\mathbf{K}_\mathbf{X}]_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j), \forall \mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}$.

Gaussian process regression begins by specifying a prior distribution $GP(\bar{f}_0(\cdot), k_0(\cdot, \cdot))$. The mean function $\bar{f}_0$ is typically selected to be zero-mean [24], while we choose the covariance kernel to be the squared exponential

$$k_0(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\{-(1/2)(\mathbf{x} - \mathbf{x}')^T \Lambda^{-1} (\mathbf{x} - \mathbf{x}')\}. \quad (1)$$

Here, σ_f^2 denotes the signal variance, while $\Lambda = \text{diag}(\lambda_1^2, \ldots, \lambda_d^2)$ contains the length-scale parameter associated with each input dimension $1, \ldots, d$. Consider a set of training data $S := \{(\mathbf{x}_n, y_n)\}_{n=1}^N$. In this work, we will use the shorthand $\mathbf{x} \in S$ to denote $(\mathbf{x}, y) \in S$. Under iid assumption of data points in S, the Bayesian update produces the predictive posterior process as $f(\mathbf{x}) \sim GP(\bar{f}_S(\mathbf{x}), k_S(\mathbf{x}, \mathbf{x}'))$ where

$$\bar{f}_S(\mathbf{x}) = \mathbf{k}_N^T(\mathbf{x})(\mathbf{K}_N + \sigma^2 \mathbf{I})^{-1} \mathbf{y}, \quad (2)$$

$$k_S(\mathbf{x}, \mathbf{x}') = k_0(\mathbf{x}, \mathbf{x}') - \mathbf{k}_N^T(\mathbf{x})(\mathbf{K}_N + \sigma^2 \mathbf{I})^{-1} \mathbf{k}_N(\mathbf{x}'). \quad (3)$$

Here, $\mathbf{k}_N(\mathbf{x}) \in \mathbb{R}^N, \mathbf{K}_N \in \mathbb{R}^{N \times N}$, $[\mathbf{k}_N(\mathbf{x})]_n = k_0(\mathbf{x}, \mathbf{x}_n)$ for $\mathbf{x}_n \in S$, $[\mathbf{K}_N]_{n,m} = k_0(\mathbf{x}_n, \mathbf{x}_m)$ for all $\mathbf{x}_n, \mathbf{x}_m \in S$, and $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ [24]. The posterior mean $\bar{f}_S$ in (2) can be viewed as an approximation of f using a finite sum of kernel function evaluations over the training data S

$$\bar{f}_S(\cdot) = \sum_{n=1}^{N} \alpha_n k_0(\mathbf{x}_n, \cdot) \quad (4)$$

where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_N]^T \in \mathbb{R}^N$ is a set of coefficients. From (4), it is clear that $\boldsymbol{\alpha} = (\mathbf{K}_N + \sigma^2 \mathbf{I})^{-1} \mathbf{y}$ in the GP setting. Further, we assume that f belongs to reproducing kernel Hilbert space (RKHS) induced from our selection of covariance kernel, and $\|f\|$ denotes the RKHS norm.

B. Information Gain

One measure of how informative the set of training data S is in learning f is the information gain, or the mutual information between observations $\mathbf{y}_S = \{y \mid (\mathbf{x}, y) \in S\}$ and f defined as

$$F(S) = I(\mathbf{y}_S; f) = H(\mathbf{y}_S) - H(\mathbf{y}_S | f), \quad (5)$$

where $H(\mathbf{y}_S)$ is the entropy of random variables $\mathbf{y}_S$ and $H(\mathbf{y}_S|f)$ is the conditional entropy given f. For a Gaussian distribution, the difference in entropy terms simplifies to

$$F(S) = \frac{1}{2} \log \det(\mathbf{I} + \sigma^{-2} \mathbf{K}_S), \quad (6)$$

where $[\mathbf{K}_S]_{i,j} = k_0(\mathbf{x}_n, \mathbf{x}_m), \ \forall \mathbf{x}_n, \mathbf{x}_m \in S$ [26]. The squared exponential covariance function satisfies the locality property, i.e. if measurement locations $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ are far from each other, then $k_0(\mathbf{x}, \mathbf{x}') \approx 0$. If we assume all measurement locations in S are far enough apart, then the information gain can be written as the sum of the information gain of each location

$$F(S) = I(\mathbf{y}_S; f) = \frac{1}{2} \log \det(\mathbf{I} + \sigma^{-2} \mathbf{K}_S)$$
$$\approx \frac{1}{2} \sum_{n=1}^{N} \log(\sigma^{-2} k_0(\mathbf{x}_n, \mathbf{x}_n) + 1) \approx \sum_{n=1}^{N} F(\mathbf{x}_n). \quad (7)$$

III. Problem Formulation

In this section we define our multi-agent streaming setting, and the optimization problem we seek to solve.

A. Problem Setting

We assume a multi-agent setting which consists of V agents. The communication network of the agents forms an undirected graph $\mathcal{G}$ defined by a finite set of nodes $\mathcal{V} = \{1, 2, \ldots, V\}$ and an edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ indicating communication between agents. The neighbors of agent i are

denoted as $\mathcal{N}_i = \{j \in \mathcal{V} : (i,j) \in \mathcal{E}\}$ and let $|\mathcal{N}_i| = N_i$ where $|\cdot|$ denotes cardinality. We consider fixed edges in this work, but analysis for time-varying graphs are potential topics for future work.

We consider a streaming setting, where agents acquire sequential batches of measurements every iteration. The control variables for the agents are the locations at which to sample the next batch of measurements. For each agent i, at the end of iteration t, let $S^o_{i,t}$ denotes all measurement pairs $\{(\mathbf{x}, y)\}$ agent i has not shared with other agents in its neighborhood, $\tilde{S}_{i,t}$ denotes all measurement pairs agent i has shared, define $S_{i,t} = S^o_{i,t} \cup \tilde{S}_{i,t}$ as set of measurements agent i has acquired, and define $S^c_{i,t} = S_{i,t} \cup \bigcup_{j \in \mathcal{N}_i} \tilde{S}_{j,t}$ as the complete set of measurements agent i knows from itself and its neighbors. At iteration t, each agent does the following: (1) selects the next set of m locations $S^o_{i(new)}$ to collect new measurements, (2) updates $S^o_{i,t} \leftarrow S^o_{i,t-1} \cup S^o_{i(new)}$, (3) communicates a subset of its measurements $\tilde{S}_{i(new)} \subset S^o_{i,t}$, to other agents in its neighborhood such that $|\tilde{S}_{i(new)}| \leq \tilde{m} < m$, and (4) updates its GP models with measurements acquired from sampling and received from neighbors. We remark that after communicating its measurements, agent i removes $\tilde{S}_{i(new)}$ from $S^o_{i,t}$ and adds it to $\tilde{S}_{i,t}$ so that $S^o_{i,t}$ has only unshared measurements at the end of iteration t. Because agents cannot move arbitrarily far during an iteration, we enforce $S^o_{i,(new)} \subset \mathcal{B}_{i,cl}$ where $\mathcal{B}_{i,cl} := \{\mathbf{x} \in \mathcal{X} \mid \|\mathbf{x} - \mathbf{x}_{i,cl}\|_2 \leq R\}$ is a ball centered at agent i's current location $\mathbf{x}_{i,cl}$.

B. Constrained Maximization of Information Gain

At the beginning of iteration t, we are given the set of new measurements $S^o_{i,(new)}$ acquired by each agent i, the measurement pairs $S_{i,t-1}$ acquired by each agent i in all previous iterations, and the complete set of measurement pairs $S^c_{i,t-1}$ each agent knows from sharing data with its neighbors. Additionally, agents can compute the information gain F and function $\bar{f}$ given any of the aforementioned known sets of measurements. The set optimization problem we wish to solve at each iteration t is

$$\arg \max_{\{S^o_{i(new)}\}_{i\in\mathcal{V}} \subset \{\mathcal{B}_{i,cl}\}_{i\in\mathcal{V}}} \quad \sum_{i=1}^{V} F(S_{i,t-1} \cup S^o_{i(new)}) \tag{8}$$

$$\text{subject to} \quad |\tilde{S}_{i,(new)}| \leq \tilde{m} \tag{9}$$

$$\|\bar{f}_{S^c_{i,t}} - \bar{f}_{S^c_{j,t}}\| \leq \varepsilon, \ \forall (i,j) \in \mathcal{E}, \tag{10}$$

where the optimization is over the set of new locations $S^o_{i,(new)}$ to acquire measurements for each agent. The sum in (8) arises from our assumption that measurements from any location $\mathbf{x}$ sampled by agent i and those from any location $\mathbf{x}'$ sampled by agent j are sufficiently distant such that $k_0(\mathbf{x}, \mathbf{x}') \approx 0$. Generally, locality holds when the spatial scale of the data is much larger than the kernel bandwidth, or length scale. Measurements acquired by the same agent, however, need not be uncorrelated. This assumption of locality holds for many multi-agent mapping settings from modeling wind fields [17] to plankton density [27] and our numerical experiments in Section V suggest that agents tend to explore disjoint areas for efficient maximization of information gain.

Information gain has been shown to satisfy the submodularity property for set functions [5]. Although maximization of monotone submodular functions under certain classes of constraints is generally a NP-hard problem [28], results from the seminal work of Nemhauser et al. [29] have shown that sequentially selecting elements to compose S using the greedy algorithm efficiently produces a near optimal solution. Specifically, the set S_G generated by the greedy algorithm maximizing a submodular function F under cardinality constraints $|S| \leq m$ satisfies $F(S_G) \geq (1 - \frac{1}{e}) \max_{S:|S|\leq m} F(S)$.

Due to bandwidth limitations in the underwater environment, the rate at which agents acquire measurements is much higher than the rate at which agents can communicate measurements. The constraint in (9) limits agents to communicating at most $\tilde{m}$ measurements per iteration.

A key novelty in the problem we consider in (8) lies in introducing the RKHS norm constraint in (10). This constraint allows us to select the next location in a manner that the local estimates of the global spatial field by each agent are similar. Enforced consensus ($\varepsilon = 0$) in multi-agent networks can be solved by methods such as distributed gradient descent [7], [9], [30]–[32], but we do not seek consensus due to the bandwidth being severely constrained. Instead, we seek to share only enough data so that environmental models onboard each agent are statistically similar in a useful way. This distinguishes our problem from multi-agent exploration settings considered in works such as [8], [11].

The problem in (8) is challenging because the objective is a submodular set function, but (10) is not a constraint on a set function. Therefore, we next derive an approximation of (8) and provide a solution to the approximate problem.

IV. Proposed Solution

In this section, we show that the norm constraint (10) can be represented by a constraint on a submodular function of $\tilde{S}_{i,t} \cup \tilde{S}_{j,t}$.

A. Submodular Representation

An agent i and its neighbors seek to communicate a subset of their measurements to satisfy the constraint

$$\|\bar{f}_{S^c_{i,t}} - \bar{f}_{S^c_{j,t}}\| \leq \varepsilon. \tag{11}$$

Using the triangle inequality, we can write

$$\left\|\bar{f}_{S^c_{i,t}} - \bar{f}_{S^c_{j,t}}\right\| \leq \left\|\bar{f}_{S^c_{i,t}} - \bar{f}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}}\right\| + \left\|\bar{f}_{S^c_{j,t}} - \bar{f}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}}\right\|. \tag{12}$$

Note that $\tilde{S}_{i,t} \cup \tilde{S}_{j,t}$ is a subset of both $S^c_{i,t}$, and $S^c_{j,t}$. The communication problem for agent i is to minimize $\left\|\bar{f}_{S^c_{i,t}} - \bar{f}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}}\right\|$ by selecting measurements from $S^o_{i,t}$ to share. For example, the constraint in (8) is satisfied if $\left\|\bar{f}_{S^c_{i,t}} - \bar{f}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}}\right\| \leq \frac{\varepsilon}{2}$ and $\left\|\bar{f}_{S^c_{j,t}} - \bar{f}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}}\right\| \leq \frac{\varepsilon}{2}$.

To obtain a submodular representation of the norm constraint (10), we utilize a result from [22], which is restated as follows.

Lemma 1: Given a finite set of training data S, and function $\bar{f}_S = \sum_{i=1}^{|S|} \alpha_i k(\cdot, \mathbf{x}_i)$, then any subset $\tilde{S} \subset S$ satisfies the criterion $\left\| \bar{f}_S - \bar{f}_{\tilde{S}} \right\| \leq \varepsilon$ if $\tilde{S}$ satisfies the inequality

$$\log\det(\mathbf{K}_{\tilde{S}} + \delta\mathbf{I}) \geq \log\left[\int_{\mathcal{X}} d\mathbf{x} - \varepsilon\right]. \tag{13}$$

Using Lemma 1, we conclude that any set $\tilde{S}_{i,t} \cup \tilde{S}_{j,t}$ satisfying

$$\log\det(\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}} + \sigma^2\mathbf{I}) \geq \log\left[\int_{\mathcal{X}} d\mathbf{x} - \frac{\varepsilon}{2}\right], \tag{14}$$

also satisfies $\left\| \bar{f}_{S^c_{i,t}} - \bar{f}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}} \right\| \leq \frac{\varepsilon}{2}$ and $\left\| \bar{f}_{S^c_{j,t}} - \bar{f}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}} \right\| \leq \frac{\varepsilon}{2}$. At every iteration t, agent i selects data measurements $(\mathbf{x}, y) \in S^o_{i,t}$ to share, such that (14) is satisfied when $(\mathbf{x}, y)$ is appended to $\tilde{S}_{i,t-1} \cup \tilde{S}_{j,t-1}$. We subtract the term $\log\det(2\pi e\sigma^2\mathbf{I})$ from (14) to obtain the equivalent inequality

$$\begin{aligned}\log\det(\sigma^{-2}\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}} + \mathbf{I}) \geq &\log\left[\int_{\mathcal{X}} d\mathbf{x} - \frac{\varepsilon}{2}\right] \\ &- \log\det(2\pi e\sigma^2\mathbf{I}),\end{aligned} \tag{15}$$

where the left side of (15) can be interpreted as the information gain given the set $\tilde{S}_{i,t} \cup \tilde{S}_{j,t}$. Therefore, the problem in (8)-(10) can be rewritten as

$$\begin{aligned}\arg\max_{\{S^o_{i(\text{new})}\}_{i\in\mathcal{V}} \subset \{\mathcal{B}_{i,cl}\}_{i\in\mathcal{V}}} \quad & \sum_{i=1}^{V} F(S_{i,t-1} \cup S^o_{i(new)}) \\ \text{subject to} \quad & |\tilde{S}_{i,(new)}| \leq \tilde{m} \\ & \log\det(\sigma^{-2}\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}} + \mathbf{I}) \geq \\ & \log\left[\int_{\mathcal{X}} d\mathbf{x} - \frac{\varepsilon}{2}\right] \\ & - \log\det(2\pi e\sigma^2\mathbf{I}), \forall (i,j) \in \mathcal{E},\end{aligned} \tag{16}$$

where we now have a constraint on a submodular function instead of a constraint on the norm difference.

B. Information Sharing as Submodular Optimization

Because measurement pairs in $\tilde{S}_{i,t}$ are selected from $S^o_{i,t}$, we must consider the subset $\tilde{S}_{i,t} \cup \tilde{S}_{j,t}$ when selecting locations of measurement pairs to add to $S^o_{i,t}$. Additionally, there is no guarantee that a feasible subset $\tilde{S}_{i,t-1} \cup \tilde{S}_{j,t-1} \cup \tilde{S}_{i,(new)}$ satisfying the last inequality in (16) exists for arbitrary $\varepsilon > 0$. Therefore, we aim to solve (16) approximately by moving the constraint (15) into the objective function via penalty method. The penalty functional is

$$\psi_t(S) = \sum_{i=1}^{V}\left(F(S^c_{i,t}) + \sum_{j\in\mathcal{N}_i} \log\det(\sigma^{-2}\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}} + \mathbf{I})\right) \tag{17}$$

Algorithm 1 Decentralized Multi-agent Exploration with Limited Communications (DME-LC)

Require: T number of total iterations

1: $S^c_{i,0} \leftarrow \emptyset$ for each agent $i \in \mathcal{V}$; $t \leftarrow 1$;
2: **while** $t \leq T$ **do**
3: **Loop in parallel** for agent $i \in \mathcal{V}$
4: Compose $S^o_{i,(new)}$ by greedily selecting m locations according to $\arg\max_{\mathbf{x}\in\mathcal{B}_{i,cl}} \mathbf{w}_i^T \boldsymbol{\psi}_{i,t}(S^c_{i,t} \cup \{\mathbf{x}\})$, append locations $S^o_{i,t} \leftarrow S^o_{i,t-1} \cup S^o_{i(new)}$;
5: Select $\tilde{m}$ points to share $\tilde{S}_{i(new)} \subset S^o_{i,t}$ using criterion $\arg\max_{\mathbf{x}\in S_{i,t}} \sum_{j\in\mathcal{N}_i} w_j \log\det(\sigma^{-2}\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}\cup\{\mathbf{x}\}} + \mathbf{I})$;
6: Share $\tilde{S}_{i(new)}$ with neighbors, receive $\tilde{S}_{j(new)}$ from all neighbors j;
7: $\tilde{S}_{i,t} \leftarrow \tilde{S}_{i,t-1} \cup \tilde{S}_{i(new)}$; $\tilde{S}_{j,t} \leftarrow \tilde{S}_{j,t-1} \cup \tilde{S}_{j(new)} \forall j \in \mathcal{N}_i$;
8: $S^o_{i,t} \leftarrow S^o_{i,t} \backslash \tilde{S}_{i,t}$; $S_{i,t} \leftarrow S^o_{i,t} \cup \tilde{S}_{i,t}$;
9: $S^c_{i,t} \leftarrow S^o_{i,t} \cup \tilde{S}_{i,t} \bigcup_{j\in\mathcal{N}_i} \tilde{S}_{j,t}$;
10: Update $\bar{f}_{S^c_{i,t}}$ and posterior covariance with $S^c_{i,t}$ using (2),(3);
11: $t \leftarrow t + 1$;
12: **end while**

where we are summing the term $\log\det(\sigma^{-2}\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}} + \mathbf{I})$ for each neighbor of agent i. The local penalty functional is

$$\psi_{i,t}(S^c_{i,t}) = F(S^c_{i,t}) + \sum_{j\in\mathcal{N}_i} \log\det(\sigma^{-2}\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}} + \mathbf{I}) \tag{18}$$

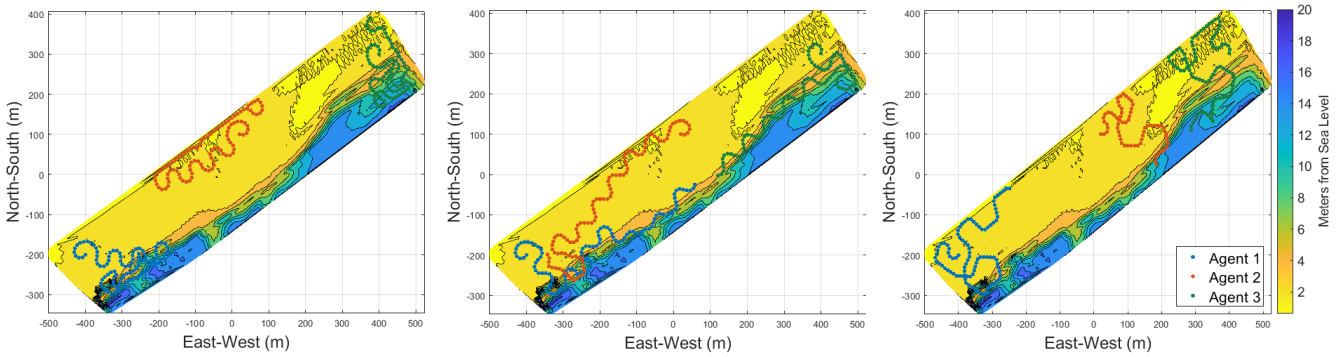
where it follows that $\psi_t(S) = \sum_{i=1}^{V} \psi_{i,t}(S^c_{i,t})$. The local penalty functional creates a multi-objective maximization problem for each agent i. A user defined weight w_k is commonly appended to each k^{th} objective to control the priority of each objective and enable trade-off analysis [33]. Appending weights $\mathbf{w}_i = \begin{bmatrix} w_{i,1} & \dots & w_{i,N_i+1} \end{bmatrix}^T$ to each objective in the local penalty functional, we obtain

$$\begin{aligned}\mathbf{w}_i^T \boldsymbol{\psi}_{i,t}(S^c_{i,t}) = &w_{i,1} \log\det(\mathbf{I} + \sigma^{-2}\mathbf{K}_{S^c_{i,t}}) \\ &+ \sum_{j\in\mathcal{N}_i} w_{i,j} \log\det(\mathbf{I} + \sigma^{-2}\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}}),\end{aligned} \tag{19}$$

where

$$\boldsymbol{\psi}_{i,t}(S^c_{i,t}) = \begin{bmatrix} \log\det(\mathbf{I} + \sigma^{-2}\mathbf{K}_{S^c_{i,t}}) \\ \log\det(\mathbf{I} + \sigma^{-2}\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}}) \\ \vdots \end{bmatrix}.$$

Because each term in (19) is submodular, then the sum of the terms is submodular as well. As noted in [28], there exist branch and bound algorithms for maximizing submodular functions, but their scalability is limited. Therefore, we choose the greedy algorithm to solve (19) because of its



(a) Weights maximizing information gain (b) Equal weights (c) Weights minimizing model differences

Fig. 1: Measurement locations taken by three agents with (a) weights biased towards information gain, (b) balanced weights, and (c) weights biased towards reducing model difference. The choice of weights affects DME-LC performance as expected, e.g. in (a) and (b), the agents have sampled more locations than in (c).

efficiency and near optimal guarantees. The greedy criterion for selecting measurements to compose $S^o_{i,(new)}$ is

$$\begin{aligned}\mathbf{x}_l &= \arg\max_{\mathbf{x}\in\mathcal{B}_{i,cl}} \psi_{i,t}(S^c_{i,t}\cup\{\mathbf{x}_1,\dots,\mathbf{x}_{l-1}\})\\ &= \arg\max_{\mathbf{x}\in\mathcal{B}_{i,cl}} w_1 \log\det(\mathbf{I}+\sigma^{-2}\mathbf{K}_{S^c_{i,t}\cup\{\mathbf{x}_1,\dots,\mathbf{x}_{l-1}\}})\\ &\quad+\sum_{j\in\mathcal{N}_i} w_j \log\det(\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}\cup\{\mathbf{x}_1,\dots,\mathbf{x}_{l-1}\}}+\sigma^2\mathbf{I}),\\ &\quad\forall l=1,\dots m.\end{aligned} \tag{20}$$

Although $\mathcal{B}_{i,cl}$ is infinite, in practice the agent will select from a finite set of candidate locations within $\mathcal{B}_{i,cl}$ based on the agent's real world dynamics.

The key steps in our method are summarized in Algorithm 1, which we refer to as DME-LC. In line 4 of DME-LC, we utilize the greedy selection criterion from (20). In line 5, agent i greedily selects measurements to share from $S^o_{i,t}$. Note that we are jointly optimizing for both information gain and model difference when selecting locations to sample in line 4, but only optimize for model difference when selecting measurements to communicate in line 5. Because we consider $\log\det(\mathbf{K}_{\tilde{S}_{i,t}\cup\tilde{S}_{j,t}}+\sigma^2\mathbf{I})$ in (19), we expect the difference between local models to not increase dramatically per iteration. In the remaining lines 7-10, the agents update their respective sets from acquired and received measurements.

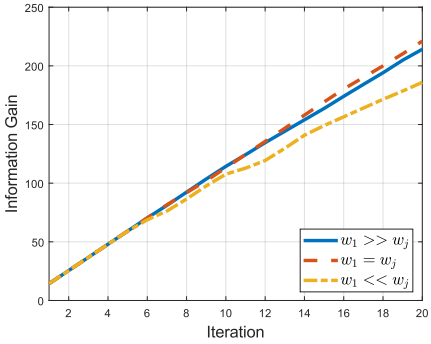
V. Empirical Assessment

In this section, we evaluate the performance of DME-LC across two numerical experiments over a bathymetric dataset collected from Claytor Lake (CL), Virginia, USA (shown in Fig. 1). First, we demonstrate how selection of weights in DME-LC affects the joint objective of maximizing information gain and minimizing model difference. In the second experiment, we benchmark our method by comparing it with a distance-based data exchange selection criterion. Note that in practice, our assumption of locality may be occasionally broken. However, this occurrence is infrequent as agents gravitate to disjoint regions after receiving data from neighbors and thus, the total information gain can still be reasonably approximated as a sum.

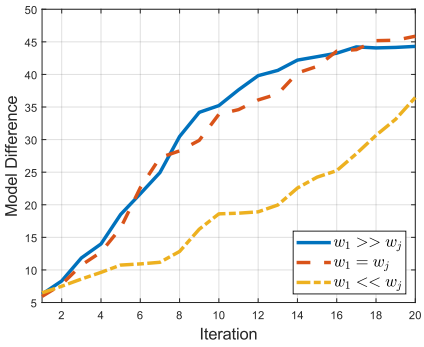
A. Real World Bathymetric Data

We run simulations with three agents acquiring measurements over the CL dataset. For the communication network, we choose $\mathcal{V}=\{1,2,3\}$ and $\mathcal{E}=\{(1,2),(2,3)\}$. The radius of $\mathcal{B}_{i,cl}$ is 10 meters, and we choose a finite number of candidate locations, including the agent's current location, within $\mathcal{B}_{i,cl}$ at which to evaluate $\mathbf{w}_i^T\boldsymbol{\psi}_{i,t}$. At each iteration, agents collect $m=5$ measurements and share $\tilde{m}=1$ measurement with their neighbors. The hyperparameters are $\{\sigma^2=0.2,\sigma_f^2=0.36,\Lambda=190\mathbf{I}_d\}$ where $\mathbf{I}_d$ is the d-dimensional identity matrix. The hyperparameters $\lambda^2=190$ and $\sigma_f^2=0.36$ ensure that pairs of measurements are approximately uncorrelated when distances between measurement locations increase beyond 30 meters. We consider three sets of weights for each agent k. The first set prioritizes maximizing information gain by setting $w_1=0.99, w_j=0.1,\ \forall j=2,\dots,N_k+1$, the second contains equivalent weights, i.e. $w_i=w_j,\ \forall i,j=1,\dots,N_k+1$, and the third prioritizes minimizing model difference, setting $w_1=0.1, w_j=0.99,\ \forall j=2,\dots,N_k+1$.

Comparing Fig. 1a, 1b with 1c, we note that the total number of locations visited by the three agents is fewer when prioritizing model difference. This is due to agents selecting their current location as the best candidate for minimizing model difference. The ratio of shared to unshared data increases when agents remain in place, further minimizing model differences. Figures 2a and 2b confirm the expected trade-off between maximizing information gain and minimizing model differences when weights are extremely biased for one objective. In general, however, selecting measurement locations beneficial to model similarity facilitates the total joint information gain, evidenced by the slight improvement in joint information gain when weights are balanced and lack of a steep drop-off in information gain even when weights are biased for model similarity, as shown in Fig. 2a.

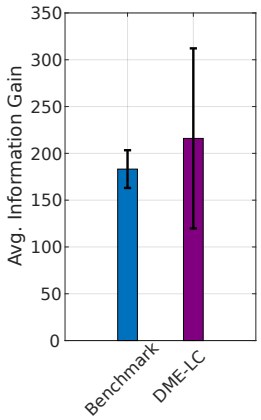


(a) Joint information gain from three agents

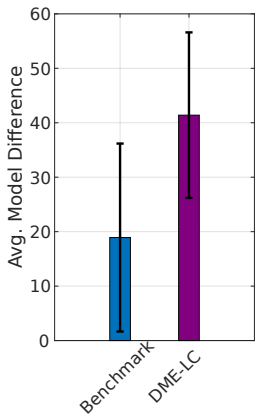


(b) Avg. model difference between neighbor pairs

Fig. 2: Performance metrics for DME-LC given three sets of weights. In (a), the joint information gain from three agents is lower when weights prioritize minimizing model difference (shown by the yellow line), but in (b) we see that this prioritization greatly reduces the rate of increase of the average model difference between neighbors.



(a) Avg. info. gain



(b) Avg. model difference

Fig. 3: In (a), DME-LC leads to improved information gathering, but with lower agent model similarity (b). However, we do not seek increasing agent similarity, but rather seek agent models that are sufficiently similar to enable improved joint information gathering. As indicated in Fig. 2, selection of weights can be used to enforce more or less model similarity, which also affects joint information gathering performance.

B. Benchmark Assessment

Although there exist criteria for selecting subsets of information to communicate, we are unaware of any algorithm directly addressing the problem of collaborative exploration under bandwidth limitations. Thus, to benchmark our solution, we modify our weighted local penalty functional in (19) to consider the Euclidean distance criterion devised in [20] instead of the log determinant of the covariance matrix. The choice of using distance is based on the intuition that agents sampling near each other will naturally have similar function approximations. Modifying (19), we obtain

$$\mathbf{w}_i^T \boldsymbol{\psi}_{i,t}^{(dist)}(S_{i,t}^c) = w_{i,1} \log\det(\mathbf{I} + \sigma^{-2}\mathbf{K}_{S_{i,t}^c}) - \sum_{j\in\mathcal{N}_i} g(w_{i,j}, \mathbf{x}_{i,cl}, \tilde{S}_{j,t}) dist(\tilde{S}_{i,t}, \tilde{S}_{j,t}), \tag{21}$$

where $dist(\tilde{S}_{i,t}, \tilde{S}_{j,t})$ is the minimum distance between any two locations in $\tilde{S}_{i,t}, \tilde{S}_{j,t}$. To incentive locality we define

$$g(w_{i,j}, \mathbf{x}_{i,cl}, \tilde{S}_{j,t}) = \begin{cases} w_{i,j} & dist(\mathbf{x}_{i,cl}, \tilde{S}_{j,t}) > \delta \\ 0 & dist(\mathbf{x}_{i,cl}, \tilde{S}_{j,t}) \leq \delta \end{cases}.$$

We compare DME-LC with the functional based on distance (21) by running Monte-Carlo simulations over the CL dataset. The number of agents, network configuration, hyperparameters, and starting location of the agents are the same as described in Section V-A. For each simulation, the weights $\mathbf{w}_i$ are chosen randomly from a uniform distribution over the interval of $[0, 10]$. Both methods are evaluated for joint information gain and average model difference between all neighbor pairs at the end of simulation. Results from each algorithm are averaged over 100 simulations.

From Fig. 3a and 3b, we observe that on average DME-LC achieves higher information gain, but also higher model difference. However, variance in performance for DME-LC is higher for information gain and similar for model difference than that achieved by (21), indicating that the choice of weights has a greater effect on objective prioritization for DME-LC. Loosely speaking, we can think of the objective of sampling to keep models similar as being orthogonal to the objective of sampling to maximize information gain. For example, a naive way for models to remain consistent among agents is to acquire uninformative samples. However, coordinated exploration among agents requires similar models. Indeed, in our approach, maintaining similar models among agents implicitly facilitates coordination in exploration.

VI. Conclusion

The problem of distributed environment learning under communication constraints is formulated as a multi-objective submodular maximization problem. We have empirically demonstrated that our multi-objective criteria is able to facilitate maximizing information gain while also optimizing for model similarity between agents. To further generalize our strategy, we seek principled methods for addressing this problem while relaxing our assumptions of a fixed communication network and identical hyperparameters between agents.

REFERENCES

[1] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE communications magazine*, vol. 47, no. 1, pp. 84–89, 2009.

[2] C. E. Rasmussen, "Gaussian processes in machine learning," in *Summer school on machine learning*. Springer, 2003, pp. 63–71.

[3] M. Jadaliha, Y. Xu, J. Choi, N. S. Johnson, and W. Li, "Gaussian process regression for sensor networks under localization uncertainty," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 223–237, 2012.

[4] K.-C. Ma, L. Liu, and G. S. Sukhatme, "Informative planning and online learning with sparse gaussian processes," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 4292–4298.

[5] A. Krause and C. E. Guestrin, "Near-optimal nonmyopic value of information in graphical models," *arXiv preprint arXiv:1207.1394*, 2012.

[6] 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. 2, 2008.

[7] A. Koppel, S. Paternain, C. Richard, and A. Ribeiro, "Decentralized online learning with kernels," *IEEE Transactions on Signal Processing*, vol. 66, no. 12, pp. 3240–3255, 2018.

[8] D. Jang, J. Yoo, C. Y. Son, and H. J. Kim, "Fully distributed informative planning for environmental learning with multi-robot systems," *arXiv preprint arXiv:2112.14433*, 2021.

[9] R. Dixit, A. S. Bedi, and K. Rajawat, "Online learning over dynamic graphs via distributed proximal gradient algorithm," *IEEE Transactions on Automatic Control*, vol. 66, no. 11, pp. 5065–5079, 2021.

[10] M. Corah and N. Michael, "Distributed matroid-constrained submodular maximization for multi-robot exploration: Theory and practice," *Autonomous Robots*, vol. 43, pp. 485–501, 2019.

[11] 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.

[12] D. Jang, J. Yoo, C. Y. Son, D. Kim, and H. J. Kim, "Multi-robot active sensing and environmental model learning with distributed gaussian process," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5905–5912, 2020.

[13] J. Banfi, A. Q. Li, N. Basilico, I. Rekleitis, and F. Amigoni, "Multirobot online construction of communication maps," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 2577–2583.

[14] A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser, "Efficient informative sensing using multiple robots," *Journal of Artificial Intelligence Research*, vol. 34, pp. 707–755, 2009.

[15] N. Karapetyan, J. Moulton, J. S. Lewis, A. Q. Li, J. M. O'Kane, and I. Rekleitis, "Multi-robot dubins coverage with autonomous surface vehicles," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2373–2379.

[16] M. N. Rooker and A. Birk, "Multi-robot exploration under the constraints of wireless networking," *Control Engineering Practice*, vol. 15, no. 4, pp. 435–445, 2007.

[17] A. Viseras, Z. Xu, and L. Merino, "Distributed multi-robot cooperation for information gathering under communication constraints," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1267–1272.

[18] M. Tavassolipour, S. A. Motahari, and M. T. M. Shalmani, "Learning of gaussian processes in distributed and communication limited systems," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 8, pp. 1928–1941, 2019.

[19] M. E. Kepler and D. J. Stilwell, "An approach to reduce communication for multi-agent mapping applications," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 4814–4820.

[20] Y. Xu, J. Choi, and S. Oh, "Mobile sensor network navigation using gaussian processes with truncated observations," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1118–1131, 2011.

[21] A. Marino, G. Antonelli, A. P. Aguiar, A. Pascoal, and S. Chiaverini, "A decentralized strategy for multirobot sampling/patrolling: Theory and experiments," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 1, pp. 313–322, 2014.

[22] H. Pradhan, A. Koppel, and K. Rajawat, "On submodular set cover problems for near-optimal online kernel basis selection," in *ICASSP 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 4168–4172.

[23] M. Schlegel, Y. Pan, J. Chen, and M. White, "Adapting kernel representations online using submodular maximization," in *International Conference on Machine Learning*. PMLR, 2017, pp. 3037–3046.

[24] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

[25] T. M. Cover and J. A. Thomas, *Elements of information theory (2. ed.)*. Wiley, 2006. [Online]. Available: http://www.elementsofinformationtheory.com/

[26] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for gaussian process optimization in the bandit setting," *IEEE transactions on information theory*, vol. 58, no. 5, pp. 3250–3265, 2012.

[27] R. Ouyang, K. H. Low, J. Chen, and P. Jaillet, "Multi-robot active sensing of non-stationary gaussian process-based environmental phenomena," 2014.

[28] A. Krause and D. Golovin, "Submodular function maximization." *Tractability*, vol. 3, pp. 71–104, 2014.

[29] 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.

[30] K. Yuan, Q. Ling, and W. Yin, "On the convergence of decentralized gradient descent," *SIAM Journal on Optimization*, vol. 26, no. 3, pp. 1835–1854, 2016.

[31] A. S. Bedi, A. Koppel, and K. Rajawat, "Asynchronous saddle point algorithm for stochastic optimization in heterogeneous networks," *IEEE Transactions on Signal Processing*, vol. 67, no. 7, pp. 1742–1757, 2019.

[32] ——, "Beyond consensus and synchrony in decentralized online optimization using saddle point method," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, 2017, pp. 293–297.

[33] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.