

Efficient Feature Mapping Using a Collaborative Team of AUVs

Benjamin Biggs
Virginia Tech
babiggs@vt.edu

Daniel J. Stilwell
Virginia Tech
stilwell@vt.edu

Harun Yetkin
Virginia Tech
yetkinh@vt.edu

James McMahon
US Naval Research Laboratory
Acoustics Division, Code 7130
james.mcmahon@nrl.navy.mil

Abstract—We present the results of experiments performed using a team of small autonomous underwater vehicles (AUVs) to determine the location of an isobath. The primary contributions of this work are (1) the development of a novel objective function for level set estimation that utilizes a rigorous assessment of uncertainty, and (2) a description of the practical challenges and corresponding solutions needed to implement our approach in the field using a team of AUVs. We combine path planning techniques and an approach to decentralization from prior work that yields theoretical performance guarantees. Experimentation with a team of AUVs provides empirical evidence that the desirable performance guarantees can be preserved in practice even in the presence of limitations that commonly arise in underwater robotics, including slow and intermittent acoustic communications and limited computational resources.

I. INTRODUCTION

In the field of underwater robotics, autonomous underwater vehicles (AUV) can be employed for a variety of tasks such as mine hunting, collection of oceanographic data, and mapping of the ocean floor. The conventional approach to underwater feature mapping (e.g., mine-hunting) relies on an exhaustive search over an area with a single AUV. In this work, we seek to address the problem of deploying a team of AUVs that is able to collaboratively search an area of interest in a decentralized manner, allowing the area to be searched much more quickly and efficiently.

We specifically consider the problem of isobath localization or, in other words, the localization of a depth contour within a body of water. While our approach could be applied to other problems, isobath localization provides a simple way for us to demonstrate our work with a team of AUVs deployed in the field. A motivating application of isobath localization is determining the regions within a waterway that are safely navigable. The primary contributions of this work are (1) the development of a novel objective function for level set estimation that utilizes a rigorous assessment of uncertainty provided by Gaussian processes, and (2) a description of experiments with a team of AUVs that addresses the practical challenges of performing collaborative tasks with underwater robots. In addition, we empirically

evaluate if a team of AUVs deployed in the field obtains the performance guarantees that are due to the use of the multi-agent receding horizon path planning approach from [1] and from recent results on the near-optimality of sequential (greedy) optimization of reward functions that are not necessarily submodular, such as ours, in [2].

Level set estimation is the process of using noisy observations of the function f to determine the region(s) where f exceeds some threshold [3]. Level set estimation has applications in geospatial data [4], bionformatics [5], and environmental studies [6]–[8]. This work is distinct from the existing literature on level set estimation in the following ways: (1) the level set is estimated using the decision-theoretic *benefit of search* objective function presented in this work instead of other methods reported in the literature to measure classification ambiguity [7]–[9]. We note that the work in [9] is closely related to the work we report herein, but it only addresses the case of a single agent and uses a different reward function.

For path planning, we formulate the isobath localization problem as an informative path planning (IPP) problem. Despite a host of techniques developed to provide approximate or suboptimal solutions to the informative path planning problem [10]–[13], receding horizon methods are often used regardless of the planning techniques used within the shorter planning horizon [14]–[16]. However, the sequence of path segments in a receding horizon implementation do not inherit the near-optimality of each individual segment. The theoretical performance guarantees that we seek to leverage in this work rely on the use of a so-called terminal reward in the path planning optimization problem at each step in the construction of a receding horizon path. It is shown in [1] that the appropriate use of terminal rewards ensures that the reward attained by traversing sections of nearly-optimal paths (the complete receding horizon path) is no worse than a desired lower bound.

II. SEARCH OBJECTIVE FOR LEVEL SET ESTIMATION

The uncertain waterway depth f is modeled by a Gaussian process, $f \sim \mathcal{GP}(m(p), k(p, p'))$. An estimate of f is computed from samples $d = (p, z)$ where $p \in \mathcal{P}$ is a location and $z \in \mathbb{R}$ is a measurement. Let \mathcal{D}_t be the set of all samples acquired up to time t . For any input p_i , an associated measurement z_i with

$$z_i = f(p_i) + \epsilon_i, \text{ for } i = 1, \dots, |\mathcal{D}_t| \quad (1)$$

*This work was supported by the Office of Naval Research via grants N00014-18-1-2627, and N00014-19-1-2194. The work of J. McMahon is supported by the Office of Naval Research through the NRL Base Program.

James McMahon is with the US Naval Research Laboratory, Code 7130, Washington D.C., USA

Benjamin Biggs, Harun Yetkin, and Daniel Stilwell are with the Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA, Harun Yetkin is also with Bartin University, Turkey

where $|\cdot|$ denotes cardinality and $\epsilon_i \sim \mathcal{N}(0, \sigma_n^2)$ is drawn from a zero mean Gaussian distribution with standard deviation σ_n .

Given the data set \mathcal{D}_t , one may wish to predict $f(p^*)$ for an arbitrary input p^* . From [17, Chapter 2.7], the posterior predictive distribution of $f(p^*)$ conditioned on \mathcal{D}_t is Gaussian and specified by the mean $\mu_{p^*|\mathcal{D}_t}$ and variance $\sigma_{p^*|\mathcal{D}_t}^2$.

We assume that incorrectly classifying a location as being above or below a threshold incurs a loss. For example, falsely labeling a location as deep enough for ship navigation could lead to a collision while falsely labeling a location as too shallow for ship navigation could lead a vessel to travel further than necessary. Mathematically, we represent the costs incurred by poor estimation using a loss function \mathcal{L} .

In the context of isobath localization, $f(p)$ gives the true depth of the waterway at the input location p . For convenience, we use the notation $f_p \triangleq f(p)$. Let $l \in \mathbb{R}$ be the level of interest. The indicator function

$$g(f_p, l) = \begin{cases} 0 & \text{if } f_p < l \\ 1 & \text{if } f_p \geq l \end{cases} \quad (2)$$

classifies the input location p as being safe ($f_p \geq l$) or unsafe ($f_p < l$) for navigation. Let $\delta \in \{0, 1\}$ be an estimate of $g(f_p, l)$. We define our loss function for level set estimation as $\mathcal{L}(f_p, \delta) = c_i |g(f_p, l) - \delta|$ where $c_i > 0$ and

$$c_i = \begin{cases} c_1 & \text{if } g(f_p, l) < \delta \\ c_2 & \text{if } g(f_p, l) \geq \delta. \end{cases} \quad (3)$$

Intuitively, c_1 gives the cost associated with estimating a location to be safe when it is actually unsafe while c_2 gives the cost of estimating a location to be unsafe when it is actually safe.

A. Bayes' Risk and the Benefit Of Search

Recall that the true value of f_p is not known. Therefore, the estimate δ which minimizes the loss function $\mathcal{L}(f_p, \delta)$ cannot be computed directly. However, because we assume f is modeled by a Gaussian process, a probability distribution for f_p is known. Let \mathcal{D}_t be a data set and δ be an estimate of (2), the expected loss or *risk* at an input point $p \in \mathcal{P}$ is

$$\mathbb{E}[\mathcal{L}(f_p, \delta) | \mathcal{D}_t] = \int_{\mathbb{R}} \pi(f_p | \mathcal{D}_t) \mathcal{L}(f_p, \delta) df_p \quad (4)$$

$$= \begin{cases} c_2 \int_l^\infty \pi(f_p | \mathcal{D}_t) df_p & \text{if } \delta = 0 \\ c_1 \int_{-\infty}^l \pi(f_p | \mathcal{D}_t) df_p & \text{if } \delta = 1. \end{cases} \quad (5)$$

where $\pi(f_p | \mathcal{D}_t)$ is the Gaussian probability distribution of f_p conditioned on the data set \mathcal{D}_t . Equation (5) simply expands the notation and partitions the integral into the regions corresponding to the different regions where $f_p < l$ and $f_p \geq l$.

The Bayes' estimate is the estimate that minimizes the *risk* in (5) and is denoted

$$\delta^*(\mathcal{D}_t) = \arg \min_{\delta \in \{0, 1\}} \mathbb{E}[\mathcal{L}(f_p, \delta) | \mathcal{D}_t].$$

The *Bayes' risk* is the expected loss given the Bayes' estimate

$$\mathbb{E}[\mathcal{L}(f_p, \delta^*(\mathcal{D}_t)) | \mathcal{D}_t] = \min \left(c_1 \int_{-\infty}^l \pi(f_p | \mathcal{D}_t) df_p, c_2 \int_l^\infty \pi(f_p | \mathcal{D}_t) df_p \right). \quad (6)$$

To represent the Bayes' risk, we define

$$r(\mathcal{D}_t) = \mathbb{E}[\mathcal{L}(f_p, \delta^*(\mathcal{D}_t)) | \mathcal{D}_t]. \quad (7)$$

Definition 1 (Benefit of Search): Let \mathcal{D}_1 and \mathcal{D}_2 be data sets. The *benefit of search* is the reduction in Bayes' risk due to obtaining additional data and is written

$$\mathcal{B}(\mathcal{D}_2 | \mathcal{D}_1) = r(\mathcal{D}_1) - r(\mathcal{D}_1 \cup \mathcal{D}_2). \quad (8)$$

Consider that a data set \mathcal{D}_t is available at time t . Let $\mathcal{D}_{t+1:t+n}$ be the data set acquired in the n time increments following time t . And let $\mathcal{D}_{t+n} = \mathcal{D}_t \cup \mathcal{D}_{t+1:t+n}$ be the combined data set. In planning at time t , the input locations $\mathcal{P}_{t+1:t+n}$ corresponding to the future data set $\mathcal{D}_{t+1:t+n}$ are known, but the measurements ($Z_{t+1:t+n}$) are not yet available. In planning, the benefit of search is computed using the expectation of the Bayes' risk

$$\mathbb{E}_{Z_{t+1:t+n}} [r(\mathcal{D}_{t+n})] = \int_{\mathbb{R}^{n_{\mathcal{V}}}} \int_{\mathbb{R}} \pi(f_p | \mathcal{D}_{t+n}) \mathcal{L}(f_p, \delta^*(\mathcal{D}_{t+n})) df_p dZ_{t+1:t+n} \quad (9)$$

where $n_{\mathcal{V}}$ denotes the cardinality of the measurement set $\mathcal{D}_{t+1:t+n}$ and, therefore, the dimensionality of the vector $Z_{t+1:t+n}$.

Theorem 1: The benefit of search in (8) is normalized and monotone in expectation.

Significantly, the normalized and monotone properties indicate that the benefit of search is an appropriate objective function for use with the greedy algorithm described in our prior work [9] which provides a lower bound on greedily selected short-horizon paths when the reward function is normalized, monotone, but not submodular. Due to space limitations, a proof of Theorem 1 is omitted here, but it can be found in [18] Sec. 8.6.1.

B. Computing Expected Bayes' Risk

While the expected Bayes' risk in (9) may be approximated using Monte Carlo methods, a closed form approximation is desirable for real-time implementation. However, no closed-form solution to (9) is known for this problem. Therefore, we propose an alternative approach. We address the computation of the expected Bayes' risk generally assuming a Gaussian process model. We note that measured values influence the posterior predictive mean but have no effect on the posterior predictive variance. Therefore, we consider a data set \mathcal{Q} with $\mathcal{Q} = \mathcal{S} \cup \mathcal{V}$ for which the input locations $\mathcal{P}_{\mathcal{Q}}$ are known, but the measurement vector $Z_{\mathcal{V}}$ associated with the partial data set \mathcal{V} is not yet available.

We first determine the marginal probability of μ ($\pi(\mu_p | \mathcal{Q}) | Z_{\mathcal{S}}, \mathcal{P}_{\mathcal{Q}}$). Given the data set \mathcal{Q} , the posterior predictive distribution of f_p is Gaussian with mean $\mu_p | \mathcal{Q}$, and

variance $\sigma_{p|Q}^2$. Marginalizing a product of Gaussian (see, e.g., [18], Sec. 8.71) yields the marginal probability of $\mu_{p|Q}$ ($\pi(\mu_{p|Q}|Z_S, \mathcal{P}_Q)$) which is a Gaussian distribution with parameters

$$\mu_\mu = \mu_{p|S} \quad (10)$$

$$\sigma_\mu^2 = \sigma_{p|S}^2 - \sigma_{p|Q}^2. \quad (11)$$

Thus, the distribution for $\mu_{p|Q}$ is centered around the posterior predictive mean given the set of available measurements with variance being the reduction in variance due to obtaining the data set \mathcal{V} for which the measurement vector $Z_{\mathcal{V}}$ is not available. The expectation of the Bayes' risk with respect to the marginal distribution of the posterior predictive mean can be expressed

$$\begin{aligned} \mathbb{E}_\mu[r(Q)] &= \int_{\mathbb{R}} \pi(\mu|Z_S, \mathcal{P}_Q) \int_{\mathbb{R}} \pi(f_p|\mu, \sigma_{p|Q}^2) \mathcal{L}(f_p, \delta^*(\mu)) df_p d\mu \\ &= \int_{\mathbb{R}} \pi(\mu|Z_S, \mathcal{P}_Q) \\ &\quad \min \left(c_1 \pi(f_p < l|\mu, \sigma_{p|Q}^2), c_2 \pi(f_p \geq l|\mu, \sigma_{p|Q}^2) \right) d\mu \end{aligned} \quad (12)$$

where $\delta^*(\mu)$ denotes the Bayes' estimate satisfying

$$\delta^*(\mu) = \arg \min_{\delta} \int_{\mathbb{R}} \pi(f_p|\mu, \sigma_{p|Q}^2) \mathcal{L}(f_p, \delta) df_p. \quad (13)$$

To handle the min function in (12), we define the space

$$\Theta_\mu = \{ \mu \in \mathbb{R} | c_1 \pi(f_p < l|\mu, \sigma_{p|Q}^2) \leq c_2 \pi(f_p \geq l|\mu, \sigma_{p|Q}^2) \}$$

or equivalently

$$\Theta_\mu = \left\{ \mu \in \mathbb{R} | \pi(f_p < l|\mu, \sigma_{p|Q}^2) \leq \frac{c_2}{c_1 + c_2} \right\}.$$

Because $\pi(f_p < l|\mu, \sigma_{p|Q}^2)$ is simply the CDF of a Gaussian distribution, we have $\Theta_\mu = (-\infty, \mu^*]$ where μ^* satisfies

$$\pi(f_p < l|\mu^*, \sigma_{p|Q}^2) = \frac{c_2}{c_1 + c_2}. \quad (14)$$

Again due to space limitations, we refer to [18], Sec. 8.6.3 for the straight-forward steps that yields

$$\mu^* = l - \text{erf}^{-1} \left(\frac{c_2 - c_1}{c_1 + c_2} \right) \sigma_{p|Q} \sqrt{2}. \quad (15)$$

The expected value of the Bayes' risk conditioned on the posterior predictive may be expressed

$$\begin{aligned} \mathbb{E}_\mu[r(Q)] &= c_2 \int_{-\infty}^{\mu^*} \pi(\mu|Z_S, \mathcal{P}_Q) \pi(f_p \geq l|\mu, \sigma_{p|Q}^2) d\mu \\ &\quad + c_1 \int_{\mu^*}^{\infty} \pi(\mu|Z_S, \mathcal{P}_Q) \pi(f_p < l|\mu, \sigma_{p|Q}^2) d\mu \end{aligned} \quad (16)$$

Again saving space, we appeal to [18], Sec. 8.6.4, for the long but straightforward derivation of the expression,

$$\begin{aligned} \mathbb{E}_\mu[r(Q)] &= \frac{(c_2 - c_1)}{4} \left[1 + \text{erf} \left(\frac{\mu^* - \mu_\mu}{\sigma_\mu \sqrt{2}} \right) \right] \\ &\quad + \frac{c_1}{2} + \frac{c_1}{2} \text{erf} \left(\frac{l - \mu_\mu}{\sqrt{2\sigma_{p|Q}^2 + 2\sigma_\mu^2}} \right) \\ &\quad - \frac{(c_1 + c_2)}{2} \int_{-\infty}^{\mu^*} \pi(\mu|Z_S, \mathcal{P}_Q) \text{erf} \left(\frac{l - \mu}{\sigma_{p|Q} \sqrt{2}} \right) d\mu. \end{aligned} \quad (17)$$

To the best of the authors' knowledge, the solution to the integral in the final term of (17) is unknown and represents a topic of current research. We use the recently proposed *exponential-quadratic approximation* [19] to provide an approximate solution. Again, due to limited space and due to the derivation of the approximation being straight-forward but long, we refer to [18] Sec. 8.6.5 for details and the final form of the approximation.

III. PATH PLANNING

Given a team of K agents, an n -length joint path $\gamma_n(s_i) = \{s_i, \dots, s_{i+n}\}$ beginning at joint state s_i is a feasible sequence of joint states. A joint state is a collection of individual agent states and is denoted $s_i = \{s_i^1, \dots, s_i^K\}$. A sequence is feasible if there exists a joint action a such that the team of agents will transition from s_i to s_{i+1} by executing the action a .

We assume that a state s_i^k for agent k corresponds to an input location $p_i^k \in \mathcal{P}$ and a joint state s_i corresponds to a set of input locations $p_i = \{p_i^1, \dots, p_i^K\}$. Let $\mathcal{P}_n(s_i)$ be the set of input locations corresponding to the n -length path $\gamma_n(s_i)$. We assume that, in traversing $\gamma_n(s_i)$, the team of agents obtains a vector of measurements $Z_n(s_i)$ corresponding to the input locations $\mathcal{P}_n(s_i)$ resulting in a data set $\mathcal{D}_n(s_i) = (\mathcal{P}_n(s_i), Z_n(s_i))$.

Recall that the benefit of search is defined as a point-wise operation. Therefore, the reward associated with a path $\gamma_n(s_i)$ is a function of the locations at which the benefit of search is evaluated as well as the data set $\mathcal{D}_n(s_i)$. Additionally, the reward associated with a path depends on previously obtained data $\mathcal{D}_{i-1}(s_0)$.

We define the reward associated with a path $\gamma_n(s_i)$ to be the benefit of search at the locations along $\gamma_n(s_i)$ or

$$J(\gamma_n(s_i)) = \sum_{p_j^k \in \mathcal{P}_n(s_i)} \mathcal{B}_j^k(\mathcal{D}_n(s_i) | \mathcal{D}_{i-1}(s_0)) \quad (18)$$

where \mathcal{B}_j^k denotes the benefit of search defined in equation (8) evaluated at the input location p_j^k . Our goal is for a group of agents to jointly maximize (18)

A. Augmented Receding Horizon Path Planning

We consider missions that correspond to path-length l that is much too long to compute in real-time. Thus we adopt a receding horizon approach where we compute m -length

paths ($m \ll l$) continuously. A robot travels one step along a m -length path, and then computes another m -length path. In [1], it is shown that the appropriate use of a terminal reward when computing m -length optimal paths yields lower bound on the value of the corresponding l -length receding horizon path. In this work, we assign the terminal reward to be the reward obtained by completing the mission (total length l) by naively mowing the lawn. The lawn mower heuristic is chosen because it is computationally efficient and is a common approach to subsea search and mapping applications.

The remainder of this subsection describes the assumptions that must be made to directly obtain the lower bound guarantees in [1]. Let $\hat{\gamma}_{l-i-n-1}(s_{i+n+1})$ be a deterministic naive solution to the l -length path planning problem (e.g., mowing the lawn) beginning at the joint state s_{i+n+1} which is feasible from the final state s_{i+n} in $\gamma_n(s_i)$. Then we say that $\bar{\gamma}_n(s_i) = \gamma_n(s_i) \cup \hat{\gamma}_{l-i-n-1}(s_{i+n+1})$ is the naively augmented path from the joint state s_i . We define the reward of a naively augmented path as

$$\bar{J}_n(s_i) \triangleq J(\gamma_n(s_i) \cup \hat{\gamma}_{l-i-n-1}(s_{i+n+1})) \quad (19)$$

Consider the augmented path $\bar{\gamma}_n(s_i) = \{s_i, \dots, s_{i+n}, \hat{s}_{i+n+1}, \dots, \hat{s}_l\}$. Using the property that for data sets \mathcal{S} and \mathcal{V} , $\mathcal{B}(\mathcal{S} \cup \mathcal{V} | \emptyset) = \mathcal{B}(\mathcal{V} | \mathcal{S}) + \mathcal{B}(\mathcal{S} | \emptyset)$, the value of the m -length path $\gamma_m(s_i) \subset \bar{\gamma}_n(s_i)$ is

$$\begin{aligned} J_m(s_i) &= \bar{J}_n(s_i) - \\ &\sum_{p_j^k}^{\mathcal{P}_{l-m-1}(s_{i+m+1})} \mathcal{B}_j^k(\bar{\mathcal{D}}_{n-m-1}(s_{i+m+1}) \cup \mathcal{D}_m(s_i) | \mathcal{D}_{i-1}(s_0)) \\ &- \sum_{p_j^k}^{\mathcal{P}_m(s_i)} \mathcal{B}_j^k(\bar{\mathcal{D}}_{n-m-1}(s_{i+m+1}) | \mathcal{D}_{i+m}(s_0)). \end{aligned} \quad (20)$$

The first term on the right-hand side of (20) gives the reward for the augmented path $\bar{\gamma}_n(s_i)$. The second term gives the total benefit of search along the locations in $\bar{\gamma}_{n-m-1}(s_{i+m+1}) = \bar{\gamma}_n(s_i) \setminus \gamma_m(s_i)$. The third term on the right-hand side of (20) gives the additional reduction in risk at the locations in $\gamma_m(s_i)$ due to the data set acquired while traversing $\bar{\gamma}_{n-m-1}(s_{i+m+1})$.

Considering the third term on the right-hand side of (20), any additional reduction in risk due to the relatively small data set $\bar{\mathcal{D}}_{n-m-1}(s_{i+m+1})$ is likely negligible in practice. Therefore, we assume that

$$\begin{aligned} J_m(s_i) &\approx \bar{J}_n(s_i) - \\ &\sum_{p_j^k}^{\mathcal{P}_{l-m-1}(s_{i+m+1})} \mathcal{B}_j^k(\bar{\mathcal{D}}_{n-m-1}(s_{i+m+1}) \cup \mathcal{D}_m(s_i) | \mathcal{D}_{i-1}(s_0)). \end{aligned} \quad (21)$$

Similarly, we assume that the second term on the right-hand

side of (21) may be appropriately approximated as

$$\begin{aligned} &\sum_{p_j^k}^{\mathcal{P}_{l-m-1}(s_{i+m+1})} \mathcal{B}_j^k(\bar{\mathcal{D}}_{n-m-1}(s_{i+m+1}) \cup \mathcal{D}_m(s_i) | \mathcal{D}_{i-1}(s_0)) \\ &\approx \\ &\sum_{p_j^k}^{\mathcal{P}_{l-m-1}(s_{i+m+1})} \mathcal{B}_j^k(\bar{\mathcal{D}}_{n-m-1}(s_{i+m+1}) | \mathcal{D}_m(s_i) \cup \mathcal{D}_{i-1}(s_0)) \end{aligned} \quad (22)$$

meaning that the measurements obtained while traversing $\gamma_m(s_i)$ have little effect on the benefit obtained while traversing the locations along $\bar{\gamma}_{n-m-1}(s_{i+m+1})$. Given these assumptions,

$$J_m(s_i) \approx \bar{J}_n(s_i) - \bar{J}_{n-m-1}(s_{i+m+1}) \quad (23)$$

where the second term on the right-hand side of (23) gives the reward for the augmented path $\bar{\gamma}_{n-m-1}(s_{i+m+1})$.

Given (23), Proposition 5.2 in [1] guarantees that the value of a receding horizon path produced using augmented paths will be bounded below by $J(\hat{\gamma}_l(s_0))$ under the condition that the n -length augmented path $\bar{\gamma}_n(s_i)$ satisfies

$$\bar{J}_n(s_i) \geq \bar{J}_{n-m-1}(s_i) \quad (24)$$

at each planning step.

The condition in (24) is difficult to guarantee in practice, we discuss this briefly in Section VI. However, it is shown in [1] that a joint receding horizon path produced using the Dec-MCTS algorithm [20] performs comparably to a method specifically designed to satisfy (24) despite no explicit guarantees that (24) is satisfied at each planning step.

IV. ONLINE GAUSSIAN PROCESS REGRESSION

Recall that the matrix inversion required for GP regression has time complexity $\mathcal{O}(|\mathcal{D}_t|^3)$ where $|\mathcal{D}_t|$ is the cardinality of the data set or the number of samples. Therefore, we seek a sparse representation of the environment with a limited number of kernel centers. There are a number of well-known approaches to obtaining a sparse representation, including [21]–[24]. Because the focus of this work is on planning, and field implementation, we choose an effective and simple approach. The maximum density of kernel centers, which can be related to the worst-case accuracy of prediction, is selected to limit the total number of kernel centers in a bounded area. A datum (p, z) is added to the set \mathcal{D}_{t-1} only if p is sufficiently far from all of the locations in \mathcal{D}_{t-1} . Moreover, when computing a prediction at the location p^* , we use the data subset $\hat{\mathcal{D}}_t = \{(p, z) \in \mathcal{D}_t : |p - p^*| \leq d_e\}$ where d_e is related to the kernel length-scale.

V. FIELD TRIALS

In order to demonstrate our augmented receding horizon path planning method in the field, we construct an example problem with the objective of localizing the 15 meter isobath within a bounded area of Claytor Lake in Dublin, VA, USA.

In order to localize the isobath, teams of Virginia Tech 690 AUVs (shown in Figure 1) were deployed and programmed

to autonomously produce receding horizon paths that seek to maximize the path value function given in (19) indirectly using sequential (greedy) optimization of the benefit of search reward function described in equation (8). Each 690 AUV is equipped with a suite of navigation sensors including a KVH 1750 inertial measurement unit (IMU), Teledyne Pathfinder Doppler velocity log (DVL), pressure sensor, WHOI micro-modem 2 acoustic modem, and GPS receiver. The inertial navigation system (INS) on the vehicle uses an unscented Kalman filter on manifolds (UKF-M) [25] to fuse the data from each sensor into an estimate of the vehicle’s attitude, velocity, and position as described in [26]. Additionally, each AUV is able to use its acoustic modem to communicate with other AUVs, although over severely limited bandwidth.

Field trials were conducted using teams of 2 and 3 Virginia Tech 690 AUVs. For each team size, two receding horizon path planning methods are tested. One method utilizes a naive heuristic to augment short horizon paths while the other does not. The naive heuristic paths used in these field trials are simple lawnmower paths. All receding horizon paths produced are 100 steps long. The planning horizon m for the method that includes naive heuristic paths is 3 steps. The planning horizon for the method that does not utilize augmented paths is 10 steps.



Fig. 1: 690 AUVs

The benefit of search gives the reduction in Bayes’ risk due to obtaining a set of depth measurements at locations throughout the operational area. The depth of interest (l in (2)) is 15 meters. The costs of (3) are equal in these experiments ($c_1 = c_2 = 10$). Therefore, the locations with the highest Bayes’ risk throughout the operational area are associated with the expected location of the 15 meter isobath.

To prevent collision between AUVs, each AUV operates at its own fixed depth of 3, 4, and 5 meters for AUVs 1, 2, and 3 respectively.

In planning, the vehicle state is $s = \{x_h, p_n, p_e\}$ where x_h is the heading in radians and p_n and p_e are the North and East components of position. The set of next states available from any current state are based on a Dubins’ type kinematic motion model. The onboard autopilot and guidance

laws enable the AUVs to track kinematic paths. The discrete dynamics defining the next state for an AUV are

$$x_h[t + 1] = x_h[t] + a \quad (25)$$

$$p_n[t + 1] = \cos(-x_h[t])\Delta_x - \sin(-x_h[t])\Delta_y \quad (26)$$

$$p_e[t + 1] = \sin(-x_h[t])\Delta_x + \cos(-x_h[t])\Delta_y. \quad (27)$$

where

$$d = r(\theta_{max} + |a|)$$

$$\Delta_x = \text{sign}(a)[r - r \cos(|a|) + d \sin(|a|)]$$

$$\Delta_y = r \sin(|a|) + d \cos(|a|).$$

for turning radius $r = 15$, a desired change in heading $a \in \mathcal{A} = \{-90, -30, -20, -10, 5, 0, 5, 10, 20, 30, 90\}$, and a maximum change in heading $\theta_{max} = 90$, where all angles are in degrees for convenience.

Joint path planning is accomplished via the sequential optimization (a greedy algorithm) wherein an ordering is assigned to the team of AUVs and each agent plans a path for itself considering only the paths of other AUVs that precede its position in the ordering (see, e.g., [27]). Additionally, because the benefit of search is both normalized and monotone, the greedy algorithm is expected to yield a high quality solution given the theoretical lower bounds presented in [2] despite the benefit of search having no proven submodularity properties.

Each AUV plans paths for itself using the Monte-Carlo tree search method described in [18] and [1] with each AUV seeking to maximize the additional reward that it contributes to the reward obtained by the paths of preceding AUVs per the sequential optimization (greedy) algorithm.

In order to facilitate communication between vehicles underwater using each vehicle’s acoustic modem, a compact packet structure for communication of path and depth measurement data was designed. A time division multiple access (TDMA) scheme is used to prevent acoustic interference between vehicles while transmitting data. Each AUV is periodically allotted a 10 second window of time to transmit a single packet of data. The acoustic packets consist of 252 bytes. The first 14 bytes in a packet describe the initial state in the current best path known to the AUV. The initial state bytes are followed by a sequence of single bytes describing the actions from \mathcal{A} taken to construct the path of the desired length. A terminal byte describes the end of the action sequence. When using naive paths to augment the short horizon paths, the terminal byte takes a value describing the deterministic heuristic used to produce the naive path. In this way, long paths may be communicated quite compactly. Following the terminal byte for the path, depth measurements are appended to the packet as groups of 12 bytes representing the north and east position of the measurement in the local tangent frame as well as the depth at that position. Given the packet size constraints, approximately 18 depth samples are communicated in a packet depending on the length of the short horizon path. To select which measurements to communicate, each AUV maintains a history of which measurements have been communicated and which have not.

From the time-ordered list of samples that have not been communicated, the AUV selects every n^{th} sample where n is selected according to the number of samples that may be transmitted in a single packet.

A. Results of Field Trials

The results for experiments conducted using 3 Virginia Tech 690 vehicles are presented in Figures 2 and 3. In Figure 2, the prior Bayes' risk in the search environment for two experiments is shown in panels the top two panels. The bottom two panels show the corresponding posterior Bayes' risk given all measurements obtained by all AUVs while traversing the joint receding horizon path shown in the top two panels. For each path, all AUVs begin at the location (0,0).

Figure 3 shows the total reward accumulated by the team of AUVs for each planning method. The dashed blue line gives the total reward accumulated by the joint receding horizon path produced without the use of a naive path as a terminal rewards. The solid green line gives the total reward accumulated by the joint receding horizon path produced using a naive path as a terminal reward. The dashed red line gives the time-varying lower bound (the reward of the complete receding horizon path is bounded below by the reward of the naive path — mowing the lawn) while the solid black line gives the maximum value of the time-varying lower bound. The receding horizon path produced using a naive path as a terminal reward performs no worse than the receding horizon path with no terminal rewards. In fact, it performs better. Given the non-idealities that arise when reducing our approach to practice, there is no guarantee that paths produced using the terminal rewards will perform better or worse than receding horizon paths produced using any other method. However, the results of these experiments appear to demonstrate that the lower bound guarantee that the reward of the receding horizon path is bounded below by the reward of the naive path (mowing the lawn) may be achieved despite non-idealities such as those described in Section VI.

VI. CONCLUDING REMARKS

The analysis in [1] suggests that using a terminal reward when computing short-horizon segments of receding horizon paths, as in (19), should yield a total reward for the complete receding horizon path that is no less than would be obtained with no terminal reward. Moreover, the results in [1] suggest that the complete path reward can be greater than when not using terminal rewards for case of computer simulations that are idealized, but do not fully meet the hypotheses in [1]. In this paper, we ask if the same property can be obtained by AUVs that are deployed in the field, in which case additional practical challenges arise. Indeed, a team of AUVs deployed in the field does not meet the requirements assumed in [1] in several ways.

- 1) **The greedy algorithm used in receding horizon path planning is not guaranteed to satisfy the condition of (24):** While the sequential optimization (greedy) algorithm as implemented in these field trials is not

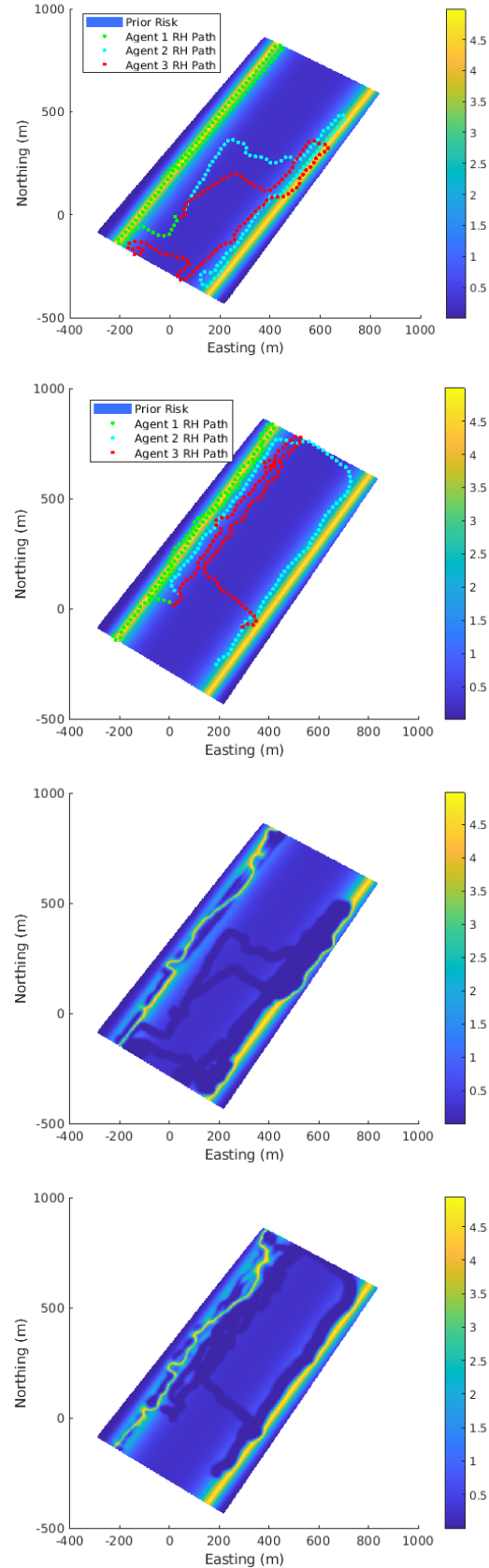


Fig. 2: Paths and Bayes' risk for experiments accomplished using 3 vehicles. Top two show the prior Bayes' risk while the bottom two panels show the posterior Bayes' risk.

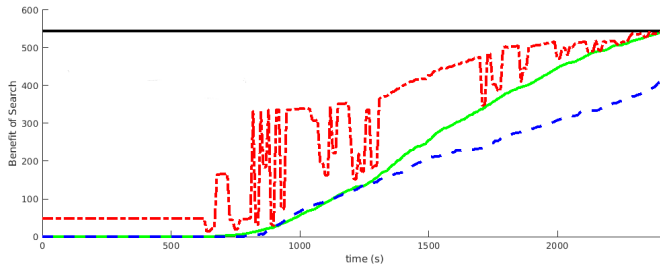


Fig. 3: Results of receding horizon missions conducted with 3 vehicles; the green line is the accumulated reward with receding horizon planning and a terminal reward, the blue dash is the accumulated reward when no terminal reward is used, the black line is the guaranteed lower bound, which is the reward of a naive lawn-mower path, and the red dashed line is the cumulative reward of using receding-horizon planning with terminal reward plus the reward of the naive path.

explicitly guaranteed to satisfy the condition of (24), the algorithm requires very little communication and is guaranteed to produce near optimal solutions when using an appropriate objective function as discussed in [28] and [1].

- 2) **Local approximations of the global objective function:** Each AUV only has access to a local approximation of the global objective function which accounts for all measurements obtained by all AUVs. Each local approximation accounts for all measurements obtained by the corresponding AUV, but the measurements obtained by all other AUVs are only available as they are communicated via acoustic channels. An example of the error between the posterior Bayes' risk for each AUVs vs the true posterior risk given all measurements from all AUVs is shown in Figure 4. Notably, given the limited communication bandwidth and the probability of data not being received at all, the error in each local approximation is small.
- 3) **Slow communication prohibits ideal coordination in path planning:** The use of the TDMA protocol allows each AUV to transmit path planning data only once per TDMA cycle. Thus, a 10 second window ensures that each AUV is able to broadcast only once every 30 seconds given a team of 3 AUVs. In this time, an AUV may have an entirely different best path for planning.
- 4) **Asynchronous path traversal:** One AUV may traverse a short-horizon path faster than another; one AUV may reach the starting waypoint before another AUV and therefore begin the overall mission before the other AUV, etc.

Fortunately, these disparities between theory and practice do not appear to have an appreciable negative effect on the results of these experiments with respect to the expected performance guarantees.

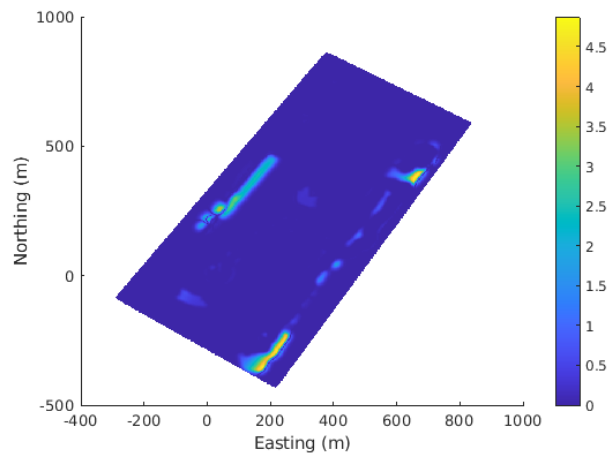
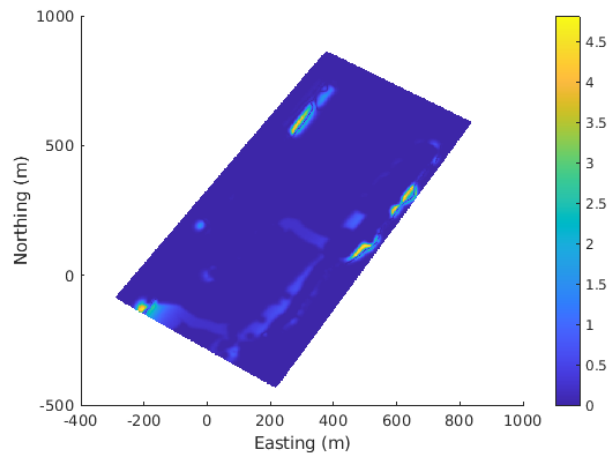
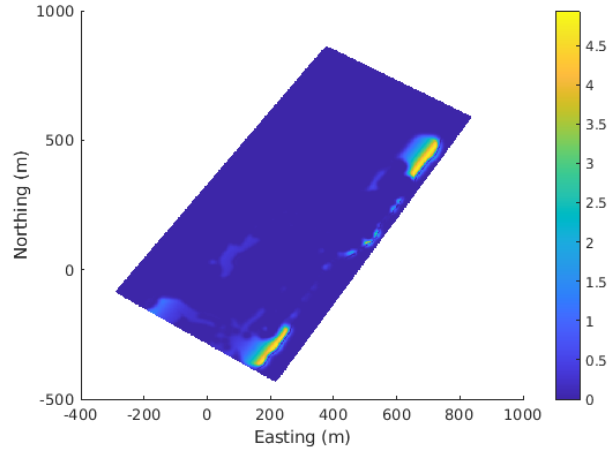


Fig. 4: The top, middle, and bottom panel show the difference between the posterior Bayes' risk given the measurements available to AUVs 1, 2, and 3, respectively, and the posterior Bayes' risk given all measurements shown in Figure 2. Note that all measurements not obtained by the AUV itself must be received via acoustic communications.

REFERENCES

- [1] B. Biggs, J. McMahon, P. Baldoni, and D. J. Stilwell, "Multi-agent receding horizon search with terminal cost," in *Proc. International Conference on Robotics and Automation (ICRA)*, 2021. doi:10.1109/ICRA48506.2021.9561643.
- [2] B. Biggs, J. McMahon, P. Baldoni, and D. J. Stilwell, "Non-submodular maximization via the greedy algorithm and the effects of limited information in multi-agent execution," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, (Kyoto, Japan), 2022. doi.org/10.1109/IROS47612.2022.9982070.
- [3] R. M. Willett and R. D. Nowak, "Minimax optimal level-set estimation," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2965–2979, 2007.
- [4] A. Solé, V. Caselles, G. Sapiro, and F. Aràndiga, "Morse description and geometric encoding of digital elevation maps," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1245–1262, 2004.
- [5] Y. H. Yang, M. J. Buckley, S. Dudoit, and T. P. Speed, "Comparison of methods for image analysis on cdna microarray data," *Journal of computational and graphical statistics*, vol. 11, no. 1, pp. 108–136, 2002.
- [6] R. Szewczyk, E. Osterweil, J. Polastre, M. Hamilton, A. Mainwaring, and D. Estrin, "Habitat monitoring with sensor networks," *Communications of the ACM*, vol. 47, no. 6, pp. 34–40, 2004.
- [7] A. Gotovos, "Active learning for level set estimation," Master's thesis, Eidgenössische Technische Hochschule Zürich, Department of Computer Science., 2013.
- [8] G. Hitz, A. Gotovos, M.-É. Garneau, C. Pradalier, A. Krause, R. Y. Siegwart, *et al.*, "Fully autonomous focused exploration for robotic environmental monitoring," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2658–2664, IEEE, 2014.
- [9] B. Biggs, H. He, J. McMahon, and D. J. Stilwell, "Experiments in underwater feature tracking with performance guarantees using a small AUV," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, (London, England), 2023. in press, pre-print: <https://arxiv.org/abs/2210.02524>.
- [10] A. Singh, A. Krause, and W. J. Kaiser, "Nonmyopic adaptive informative path planning for multiple robots," in *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, 2009.
- [11] 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.
- [12] H. Yetkin, J. McMahon, N. Topin, A. Wolek, Z. Waters, and D. J. Stilwell, "Online planning for autonomous underwater vehicles performing information gathering tasks in large subsea environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6354–6361, IEEE, 2019.
- [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] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, "Information acquisition with sensing robots: Algorithms and error bounds," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6447–6454, IEEE, 2014.
- [15] 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)*, pp. 4775–4782, IEEE, 2015.
- [16] J. Wakulicz, K. M. B. Lee, C. Yoo, T. Vidal-Calleja, and R. Fitch, "Informative planning for worst-case error minimisation in sparse Gaussian process regression," *arXiv preprint arXiv:2203.03828*, 2022.
- [17] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006.
- [18] B. A. Biggs, *Robotic Search Planning In Large Environments with Limited Computational Resources and Unreliable Communications*. PhD thesis, Virginia Polytechnic Institute and State University, 2023.
- [19] G. Pulford, "A 4-term exponential-quadratic approximation for Gaussian Q or error functions accurate to 1.65×10^{-4} ," *REV Journal on Electronics and Communications*, 2022.
- [20] G. Best, O. M. Cliff, T. Patten, R. R. Mettu, and R. Fitch, "Decmcts: Decentralized planning for multi-robot active perception," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 316–337, 2019.
- [21] L. Csató and M. Opper, "Sparse on-line Gaussian processes," *Neural Computation*, vol. 14, p. 641–668, Mar 2002.
- [22] M. Titsias, "Variational learning of inducing variables in sparse Gaussian processes," in *Artificial intelligence and statistics*, pp. 567–574, PMLR, 2009.
- [23] D. Nguyen-Tuong, M. Seeger, and J. Peters, "Model learning with local Gaussian process regression," *Advanced Robotics*, vol. 23, p. 2015–2034, Jan 2009.
- [24] B. Wilcox and M. C. Yip, "Solar-gp: Sparse online locally adaptive regression using Gaussian processes for bayesian robot model learning and control," *IEEE Robotics and Automation Letters*, vol. 5, p. 2832–2839, Apr 2020.
- [25] M. Brossard, A. Barrau, and S. Bonnabel, "A code for unscented Kalman filtering on manifolds (UKF-M)," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5701–5708, 2020.
- [26] S. Krauss and D. J. Stilwell, "Unscented Kalman filtering on manifolds for AUV navigation - experimental results," in *Proc. IEEE/MTS OCEANS*, (Virginia Beach, Virginia, USA), 2022. doi.org/10.1109/OCEANS47191.2022.9977251.
- [27] M. Corah and N. Michael, "Efficient online multi-robot exploration via distributed sequential greedy assignment," in *Robotics: Science and Systems*, vol. 13, 2017.
- [28] B. Biggs, D. J. Stilwell, and J. McMahon, "Extended performance guarantees for receding horizon search with terminal cost," in *Proc. IEEE/RSJ Intelligent Robots and Systems (IROS)*, (Las Vegas, NV, USA), 2020.