Autonomous Localization of an Unknown Number of Targets Without Data Association Using Teams of Mobile Sensors

Philip Dames, *Student Member, IEEE,* and Vijay Kumar, *Fellow, IEEE*

***Abstract*—This paper considers situations in which a team of mobile sensor platforms autonomously explores an environment to detect and localize an unknown number of targets. Individual sensors may be unreliable, failing to detect objects within the field of view, returning false positive measurements to clutter objects, and being unable to disambiguate true targets. In this setting, data association is difficult. We utilize the PHD filter for multi-target localization, simultaneously estimating the number of objects and their locations within the environment without the need to explicitly consider data association. Using sets of potential actions generated at multiple length scales for each robot, the team selects the joint action that maximizes the expected information gain over a finite time horizon. This is computed as the mutual information between the set of targets and the binary events of receiving no detections, effectively hedging against uninformative actions in a computationally tractable manner. We frame the controller as a receding-horizon problem. We demonstrate the real-world applicability of the proposed autonomous exploration strategy through hardware experiments, exploring an office environment with a team of ground robots. We also conduct a series of simulated experiments, varying the planning method, target cardinality, environment, and sensor modality.**

***Note to Practitioners*—Teams of small robots have the potential to automate many information gathering tasks, relaying data back to a base station or human operator from multiple vantage points within an environment. The information gathering tasks we consider in this work are those in which the number of objects being sought is not known at the onset of exploration. Such tasks are common in security and surveillance, where the number of such objects is often zero; search and rescue, where, for example, the number of people trapped due to a natural disaster can be large; or smart building/smart city applications, where the data collection needs may be on an even larger scale. This paper seeks to address the problem of automating this data collection process, so that a team of mobile sensor platforms are able to autonomously explore a given environment in order to determine the number of objects of interest and their locations, while avoiding any explicit data association, *i.e.,* matching individual measurements to targets.**

***Index Terms*—Robots, Robot sensing systems, Distributed tracking, Information theory, Cooperative systems**

WE are interested in applications such as search and rescue, security and surveillance, and smart buildings and smart cities, in which teams of mobile robots can be used to explore an environment to search for a large number of objects of interest. Concrete examples include using thermal imaging to locate individuals trapped in a building after a natural disaster, using cameras to locate suspicious packages in a shopping center, or using wireless pings to locate sensors within a smart building or smart city. Real-world examples of such smart building scenarios include [1], which features thermostats, microphones, access points, and bluetooth-enabled actuators within a building, and [2], which describes low-power sensors embedded within construction materials. In each of these examples, the number of objects is not known a priori, and can potentially be very large. The sensors can be noisy, there can be many false positive or false negative detections, and it may not be possible to uniquely identify and label individual objects.

We model the distribution of objects using the probability hypothesis density (PHD), which describes the spatial density of objects in the environment. The PHD filter, which propagates the first moment of the posterior, multi-target probability distribution in time, was first derived by Mahler [3]. The robots jointly plan actions that maximize the mutual information between the resulting target estimate and the future binary events indicating whether a sensor detects any targets or not, effectively hedging against uninformative actions in a computationally tractable manner. Our approach offers several key advantages: scalability in the number of targets, avoidance of any explicit data association, and the ability to handle a variable number of measurements at each time step.

The main contribution of this paper is the development of a new information-based, receding horizon control law that allows small teams of robots to perform autonomous information gathering tasks. We also propose a criterion, based on the entropy of the multi-target distribution, to terminate the active information gathering process despite the uncertainty in the target cardinality. We demonstrate the real-world applicability of the proposed control law and termination criterion through a series of hardware experiments using a team of ground robots equipped with bearing-only sensors seeking tens of targets in an indoor office environment. We then validate the performance of our simulation environment and use this to demonstrate that the proposed control law performs well across different environments, target cardinalities that span orders of magnitude (1 to 100), and different sensor modalities. In all of these cases, the robot team is able to accurately estimate the number of targets and their locations.

I. Related Work

Situations with multiple targets and the problem of simultaneously estimating the number of targets and their locations

P. Dames and V. Kumar are with the Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, PA, 19104 USA e-mail: {pdames,kumar}@seas.upenn.edu.

are not a trivial extension of the single target case. In general, the sensor can experience false negatives (*i.e.,* missing targets that are actually there), false positives (*i.e.,* seeing targets that are not actually there), and even when a true detection is made it is potentially a very noisy estimate. Another challenge faced in multi-target tracking, such as Ong et al. [4], is the task of data association. The number of such associations grows combinatorially with the number of targets and measurements, and it is often not possible to uniquely identify individual targets. In the robotics community, this problem is typically solved using heuristics, *e.g.,* Dissanayake et al. [5]. The data fusion and tracking community provides us with several other tools, including the Multiple Hypothesis Tracker (MHT), Joint Probabilistic Data Association (JPDA), and the Probabilistic Multi-Hypothesis Tracker (PMHT) [6]. The MHT and PMHT track the maximum likelihood data association in a recursive and batch fashion, respectively. JPDA considers "soft" associations, looking at the likelihood of each measurement originating from each target. However, these approaches still require heuristic methods to initialize and remove target tracks and solve the association and tracking problems independently.

Our approach is based on the mathematical construct of a random finite set [7]. In this scenario, the general Bayes filter avoids explicit data association by averaging over all possible associations [7], but is intractable to use in practice. To solve this problem, Mahler created the probability hypothesis density (PHD) [3] filter, which tracks the first statistical moment of the joint distribution of target cardinality and positions. The PHD models the density of targets in the environment [7], and Erdinc et al. [8] show that this is equivalent to the bin-occupancy filter, in the limit as the bin size goes to zero. We utilize the PHD filter to perform the multi-target localization task, and the proposed information-based control law builds upon this mathematical framework.

Information-based control has seen a lot of attention in recent years as a way of driving robots to localize and track targets. Hoffmann and Tomlin [9] and Julian et al. [10] use mutual information to localize a stationary target and explore unknown environments using a team of robots, assuming limited dependence between robots to achieve scalability. Hollinger et al. [11] use an information-based objective function to perform autonomous ship inspection with an AUV platform. Julian et al. [12] and Souza et al. [13] utilize mutual information as an objective to drive a single robot to explore an unknown environment in order to build a map. Charrow et al. use mutual information to drive a team of robots equipped with range-only sensors to track a single moving target in real-time [14] and to detect and localize an unknown number of targets, but with known data association [15].

Our control policy for active perception for multi-target tracking builds on the literature on receding horizon control and model predictive control. Mayne and Michalska [16] provide a survey of receding horizon control and Mayne et al. [17] provide a survey of model predictive control, including applications in a variety of domains. The work of Ryan [18] is particularly relevant as they use model predictive control in an information gathering setting, using a small team of UAVs to localize and track a moving target. We adapt this work to the multi-target, active estimation problem to consider actions over an extended time horizon, rather than a simple myopic exploration strategy.

There is a relatively limited body of work on active control for target localization based on the RFS framework, with the exception of work by Ristic and Vo [19] and Ristic et al. [20] to maximize information using Rényi's definition. In this work, the measurement model involves a summation over all possible data associations and the authors present simulation results of a single robot seeking three targets in an open environment. Dames et al. [21] use Shannon's information to track a small number of targets, but do not assume that the target positions are independent. In all of these works, the resulting control calculations quickly become computationally intractable for large numbers of targets and measurements. This paper significantly improves upon the control strategy proposed in [22], which focuses on communication rather than control, by planning over a finite horizon and by considering actions across multiple length scales. This avoids the issue of getting stuck in local information minima, without resorting to stochastic exploration strategies such as in [22]. To the best of our knowledge, we also present the first experimental results of an active exploration strategy based on the PHD filter.

II. Problem Formulation

The problem we consider in this work is of a team of R robots exploring an environment E in search of targets. The robots are assumed to be able to localize themselves within the environment, or at least with sufficiently high accuracy so that any errors will have a negligible effect on the target localization. At time t the robot r has pose q_t^r and receives a set of measurements $Z_t^r = \{z_{1,t}^r, \ldots, z_{m_t^r,t}^r\}$, which has m_t^r measurements. A set of n target locations is given by $X_t = \{x_{1,t}, \ldots, x_{n,t}\}$, where each $x_{i,t} \in E$. Here Z and X are random finite sets (RFSs), where an RFS is a set containing a random number of random elements, *e.g.,* each of the n elements x_i in the set $X = \{x_1, \ldots, x_n\}$ is a vector indicating the position of a single target.

A. Random Finite Sets

A random finite set (RFS) is a set containing a random number of random objects, *e.g.,* random vectors representing the locations of targets, as shown in Fig. 1. This differs from random vectors in several key ways: realizations of an RFS may have different cardinalities, so they cannot be added as a random vector would; sets are equivalent under permutations of the elements while random vectors are not; and the expected value of an RFS is not a set, but a density function. These points will be elaborated on slightly here, but Mahler [7] provides a more thorough handling of the subject. We require three concepts: a probability distribution over RFS, a set integrals, and the probability hypothesis density (PHD).

The derivation of a probability distribution over RFSs has its roots in point process theory, see Daley and Vere-Jones [23] for an overview of the subject. In this paper, we assume that all RFSs have independently and identically distributed (i.i.d.)

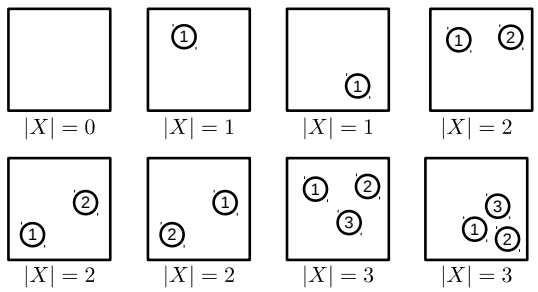


Fig. 1. Examples of random finite sets with 0 to 3 elements drawn from the square environment. The two sets in the lower left are identical, as sets are equivalent under permutations of their elements, *i.e.,* $X = \{1, 2\} = \{2, 1\}$.

elements. The likelihood of such an RFS X is

$$p(X) = |X|!\, p(|X|) \prod_{x \in X} \frac{v(x)}{\langle 1,\, v \rangle}, \tag{1}$$

where the leading $|X|!$ is the number of permutations of elements in the set, $p(|X|)$ is the cardinality distribution, and the term in the product is the normalized PHD, or the probability of a target being at a location x. As a matter of notation, we define the inner product between two real-valued functions $\langle a,\, b \rangle$ to be

$$\langle a,\, b \rangle = \int_E a(x) b(x)\, dx,$$

or $\langle a,\, b \rangle = \sum_{k=0}^{\infty} a(k) b(k)$ for real-valued sequences.

Since two sets cannot be added in the same manner as vectors, the notion of integration in the space of RFSs requires care. A set integral is defined as

$$\int f(X)\, \delta X = \sum_{n=0}^{\infty} \frac{1}{n!} \int f(\{x_1, \ldots, x_n\})\, dx_1 \ldots dx_n. \tag{2}$$

Note the use of δ as the differential element for an RFS, and the sum over the set cardinality. The $n!$ term takes into account the equivalence of a set under permutations of its elements.

The final key concept is that of a probability hypothesis density (PHD). This is defined as a density function over the environment such that the integral over any region S gives the expected cardinality of an RFS X in that region, *i.e.,*

$$\int |X \cap S|\, p(X)\, \delta X = \int_S v(x)\, dx, \tag{3}$$

where $|\cdot|$ denotes set cardinality and $v(x)$ is the PHD. Note that the PHD is *not* a probability density function, but rather describes the spatial density of targets within the environment. Using the idea of a set integral, one may show that the PHD is the first statistical moment (*i.e.,* the mean) of a distribution over random finite sets.

B. Sensor Models

Each robot is equipped with a sensor that is able to detect targets within its field of view. The probability of a sensor with pose q detecting a target at x is given by $p_d(x; q)$ and is identically zero outside of the field of view. Note the dependence on the sensor's position, denoted by the argument q. If a target at x is detected by a robot at q then it returns a measurement $z \sim g(z \mid x;\, q)$. The false positive, or clutter, model consists of a PHD $c(z; q)$ describing the likelihood of clutter measurements in measurement space and the expected clutter cardinality. Note that in general the clutter may depend on environmental factors.

C. Bayes Filter

The general Bayes filter is a set of recursive equations for updating the probability distribution over targets given a collected measurement. For RFSs, the equation is

$$p(X \mid Z) = \frac{p(Z \mid X) p(X)}{\int p(Z \mid X) p(X)\, \delta X}. \tag{4}$$

The distribution over the targets is given in (1), and the measurement model is

$$p(Z \mid X) = e^{-\mu} \left(\prod_{z \in Z} c(z) \right) \left(\prod_{x \in X} (1 - p_d(x)) \right) \times \left(\sum_{\theta} \prod_{j \mid \theta(j) \neq 0} \frac{p_d(x_j) g(z_{\theta(j)} \mid x_j)}{(1 - p_d(x_j)) c(z_{\theta(j)})} \right), \tag{5}$$

where $\theta : \{1, \ldots, n\} \to \{0, 1, \ldots, m\}$ is a data association [7]. Note that $\theta(j) = 0$ means that target j is not detected, and any element of $\{1, \ldots, m\}$ not in the range of $\theta(\{1, \ldots, n\})$ is a false positive. This features a sum over all possible data associations, which grows combinatorially in the number of measurements and targets, quickly becoming intractable.

D. PHD Filter

The PHD filter is a computationally tractable set of equations to track the first moment of the distribution over RFSs. In this paper, we assume that all targets are stationary and that the number of targets does not vary as a function of time. In deriving the filter, Mahler [3] makes the following assumptions about the target and measurement sets:

- the clutter and true measurement RFSs are independent;
- the clutter RFS is Poisson;
- prior and predicted multi-target RFSs are Poisson.

The independence assumption is standard for target localization tasks. However, the assumption that the clutter and target cardinalities follow Poisson distributions is less common within robotics. This Poisson assumption on the cardinalities stems from the use of Poisson point processes, which has the desirable property that the number of points in each finite region is independent if the regions do not overlap [23]. In this case, the mean number of targets is $\lambda = \langle 1,\, v \rangle$ and the target set likelihood (1) simplifies to

$$p(X) = e^{-\lambda} \prod_{x \in X} v(x). \tag{6}$$

Such an RFS is said to be Poisson.

TABLE I
TABLE OF SYMBOLS

R	Number of robots	z	Measurement
q	Robot pose	Z	Measurement set
Q	Action set	y	Binary Measurement
x	Target pose	$p_d(x;q)$	Probability of detection
$v(x)$	Target PHD	$g(z \mid x;q)$	Measurement likelihood
λ	Expected # targets	$c(z;q)$	Clutter PHD
T	Time horizon	μ	Expected clutter rate
ϵ	Termination criterion	L	Number of length scales

The PHD filter update equation is

$$v_t(x) = (1 - p_d(x;q))v_{t-1}(x) + \sum_{z \in Z_t} \frac{\psi_{z,q}(x)v_{t-1}(x)}{c(z;q) + \langle \psi_{z,q},\, v_{t-1} \rangle}, \tag{7}$$

$$\psi_{z,q}(x) = g(z \mid x;q)p_d(x;q), \tag{8}$$

where $\psi_{z,q}(x)$ is the probability of measurement z coming from a target at x.

E. Mutual Information

The mutual information between X and Z is defined as

$$I[\mathcal{X};\mathcal{Z}] = \iint p(X,Z) \log \frac{p(X,Z)}{p(X)p(Z)}\, \delta X \delta Z \tag{9}$$

$$= H[\mathcal{Z}] - H[\mathcal{Z} \mid \mathcal{X}], \tag{10}$$

where $H[\mathcal{Z}]$ is the entropy and $H[\mathcal{Z} \mid \mathcal{X}]$ is the conditional entropy of the measurements [24]. Intuitively, mutual information quantifies the amount of dependence between two random variables, in this case the set of target locations X and set of measurements Z. Throughout the paper, we use lowercase letters for realizations of random numbers and vectors (*e.g.,* x is the position of an individual target), capital letters for realizations of RFSs (*e.g.,* X is a set of target locations), and script letters for random variables (*e.g.,* $\mathcal{X}$ is the random variable representing target locations).

III. Information-Based Receding Horizon Control

In Sec. II-C we saw that the general Bayes filter was intractable due to the number of possible data associations, necessitating the development of the PHD filter. The same sum over data associations also appears in the expression for the mutual information between the target and measurement sets, making it prohibitively expensive to compute. To get around this, we instead consider the binary event of receiving an empty measurement set,

$$y = \begin{cases} 0 & Z = \varnothing \\ 1 & \text{else.} \end{cases} \tag{11}$$

Here $y = 0$ is the event that the robot receives no measurements to any (true or clutter) objects while $y = 1$ is the complement of this, *i.e.,* the robot receives at least one measurement. Mahler proposes a similar idea [25], where the objective is to maximize the mutual information between the target set and the empty measurement set, *i.e.,* when $p(Z = \varnothing) = 1$, so

$$q^* = \operatorname*{argmax}_q I[\mathcal{X};\mathcal{Z}(q) = \varnothing]. \tag{12}$$

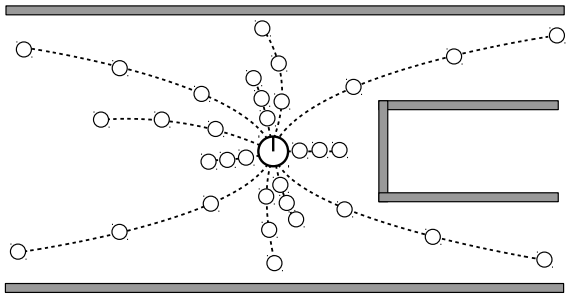


Fig. 2. Example action set with a horizon of $T = 3$ steps and three length scales. Each action is a sequence of T poses at which the robot will take a measurement, denoted by the hollow circles.

This objective is chosen because it hedges against the highly non-informative empty measurement set. Kreucher et al. [26] take a similar approach, using a binary sensor model and an information-based objective function to schedule sensors over a receding horizon to track an unknown number of targets.

This paper considers the information gathering problem in a receding horizon framework, planning T actions into the future. Let the time horizon $\tau = \{t+1, \ldots, t+T\}$. The mutual information is

$$q_\tau^* = \operatorname*{argmax}_{q_\tau \in Q_\tau^{1:R}} I[\mathcal{X}_{t+T}; \mathcal{Y}_\tau^{1:R}(q_\tau)], \tag{13}$$

where $\mathcal{X}_{t+T}$ is the predicted location of the targets at time $t+T$ and $\mathcal{Y}_\tau^{1:R}(q_\tau)$ is the collection of binary measurements for robots 1 to R from time steps $t+1$ to $t+T$, which depend on the future locations of the robots $q_\tau = [q_{t+1}^1, \ldots, q_{t+T}^1, \ldots, q_{t+T}^R]$. Note that the robot poses q are *not* random variables themselves, but the random variable $\mathcal{Y}_t^r$ depends on the value q_t^r through the detection model $p_d(x;q)$.

A. Action Set Generation

The possible future measurements of the robots depend upon their future locations within the environment, so the action set for the team, $Q_\tau^{1:R}$, must be sufficiently rich for the robots to explore the environment. Simultaneously, it must be kept as small as possible to reduce the computational complexity of (13).

Over a short time horizon, an individual robot may move in a small neighborhood around its current location. If one were to naïvely chain such actions, there would be an exponentially growing number of possible actions. However, many of these would be redundant, *i.e.,* the robots would traverse the same region. To curb the number of actions while maintaining diversity, the robot selects a number of candidate points at a given length scale from its current location, plans paths to those goals, and interpolates the paths to get T intermediate points. See Algorithm 1 for details on this process and Fig. 2 for an example action set. This forms the basis of actions Q^r for an individual robot r at a particular length scale.

It is advantageous to plan over multiple length scales, as there are some instances where a lot of information may be gained by staying in a small neighborhood, while other times it is more beneficial to travel to a distant, unexplored area.

To allow for this diversity, each robot generates actions over a range of L length scales. The number of planning steps, T, at each length scale is kept constant so that meaningful comparisons between the information values can be made.

1) Concurrent: Ideally, the team would plan over all possible actions for all robots. Individual robot action sets consist of all actions over all length scales, and the joint action set is the Cartesian product of the individual action sets, $Q^{1:R} = Q^1 \times \ldots \times Q^R$. This leads to individual action sets that grow linearly in the number of length scales, $|Q^r| = \mathcal{O}(L|Q_\ell|)$, and a joint action set that grows exponentially in the number of robots, $|Q^{1:R}| = \mathcal{O}((L|Q_\ell|)^R)$, where $|Q_\ell|$ is the number of actions at an individual length scale. This makes the concurrent computations prohibitively expensive for all but small teams of robots with small action sets.

2) Sequential: To alleviate some of the computational load, we may apply a sequential, but approximate, approach to select the best action for the team. Robots sequentially optimize over length scales, individual robots, or both. This reduces the size of a joint action set to $\mathcal{O}(|Q_\ell|^R)$, $\mathcal{O}(L|Q_\ell|)$, or $\mathcal{O}(|Q_\ell|)$, respectively, and there are L, $\mathcal{O}(R)$, or $\mathcal{O}(LR)$ action sets. However, there is no guarantee that the resulting joint action computed using any of the sequential methods is identical to the fully concurrent plan.

Length Scales: To sequentially plan over length scales, the objective changes to

$$q_\tau^* = \underset{\ell}{\operatorname{argmax}} \underset{q_{\tau,\ell} \in Q_{\tau,\ell}^{1:R}}{\operatorname{argmax}} I[\mathcal{X}_{t+T}; \mathcal{Y}_\tau^{1:R}(q_{\tau,\ell})], \tag{14}$$

where $Q_{\tau,\ell}^{1:R}$ is the set of joint actions at length scale ℓ.

Robots: To sequentially plan over robots, the first robot plans its action independently of all other agents, and each subsequent agent plans its action conditioned on all of the other robots' paths. This cycle is repeated until the robots reach a consensus, *i.e.,* robots have a chance to update their original plans given the new plans of other agents. The sequence of joint action sets, $Q^{1:R}$, is

$$\begin{gathered}
Q^1 \times \varnothing \times \varnothing \times \varnothing \times \ldots \times \varnothing \\
q^{*1} \times Q^2 \times \varnothing \times \varnothing \times \ldots \times \varnothing \\
q^{*1} \times q^{*2} \times Q^3 \times \varnothing \times \ldots \times \varnothing \\
\vdots \\
q^{*1} \times q^{*2} \times \ldots \times q^{*R-1} \times Q^R \\
Q^1 \times q^{*2} \times \ldots \times q^{*R-1} \times q^{*R} \\
q^{*1} \times Q^2 \times q^{*3} \times \ldots \times q^{*R} \\
\vdots
\end{gathered}$$

where q^{*r} is the element of Q^r with the highest expected information gain given the paths of all other robots at the time of planning. This is similar to the idea of Adaptive Sequential Information Planning from Charrow et al. [15].

This sequential optimization over robots is like the idea of coordinate descent in optimization, where a utility function is optimized over each coordinate in sequence reaching until a local optimum. In practice, the number of cycles through the team until reaching consensus was typically one, and never more than three.

Both: To sequentially plan over both robots and length scales, we use the objective (14), where the inner argmax is over the sequence of action sets from above.

3) Planning Modes: We use the following shorthand to describe the different planning modalities:

- Mode 0: MI, concurrent robots, concurrent length scales
- Mode 1: MI, concurrent robots, sequential length scales
- Mode 2: MI, sequential robots, concurrent length scales
- Mode 3: MI, sequential robots, sequential length scales
- Mode 4: Random, concurrent length scales
- Mode 5: Random, sequential length scales

In modes 0–3, the action is selected by maximizing the mutual information, as described in this section, while in modes 4 and 5 an action is selected randomly from the joint action set. For modes 1, 3, and 5, the length scale with the highest expected information gain is selected.

Algorithm 1 Action Set Generation

procedure ACTIONSET(L, T, q, M) ▷ Action set at length scale L over horizon T for a robot at q in map M
  Let $d(x, y)$ be the distance between x and y
  $P \leftarrow \{x \in M \mid d(x, q) = L\}$ ▷ Find all points in the map a distance L from the robot
  $G \leftarrow \{x\}$ ▷ Pick an arbitrary initial goal $x \in P$
  $Q \leftarrow \varnothing$ ▷ Initialize action set
  while not empty(P) **do**
    $x^* = \operatorname{argmin}_{x \in P} \min_{y \in G} d(x, y)$
    $G \leftarrow G \cup \{x^*\}$ ▷ Add to list of goals
    Path $\leftarrow$ path from q to x^* ▷ Found using A*
    $Q \leftarrow Q \cup \{[T$ evenly spaced points along Path$]\}$
    $P \leftarrow P \setminus \{x \in P \mid d(x, x^*) < R\}$ ▷ Remove all points sufficiently close to the goal
  return Q

B. Receding Horizon

As this is a receding horizon control law, the robots replan their action after executing a fraction of the current action. In this work, the team replans an action after each of the robots has completed at least one of the T actions, *i.e.,* after all robots have traversed $1/T$ of the planned path length. This allows robots acting at larger distance scales than others to visit at least one of their planned locations, even if the robots acting at shorter distance scales have completed their actions.

However, it is worth noting that if robots are acting at very different length scales, it is possible for one robot to have completed its full action (*i.e.,* reached all T waypoints) before one of the other robots has reached its first waypoint. In this case the first robot would sit idly, waiting for the second to trigger the replanning.

C. Computing the Objective Function

We utilize the factorization of mutual information from (10) to compute the objective function in (13).

1) Entropy: We begin by computing the binary measurement likelihoods, $p(y)$, first for an individual robot and then for a team of robots. The only way for the sensor to get no detections is for it to have zero clutter detections and to not detect any target, so

$$p(y=0 \mid X) = e^{-\mu} \prod_{x \in X} (1 - p_d(x_i)), \tag{15}$$

where $\mu = \langle 1,\, c \rangle$ is the expected number of clutter detections, $c(z)$ is the clutter PHD from Sec. II-B, and $e^{-\mu}$ is the probability of receiving no clutter detections given the assumption of Poisson clutter cardinality. Using this, we get that

$$\begin{aligned} p(y=0) &= \int p(y=0 \mid X) p(X)\, \delta X \\ &= e^{-\mu} \sum_{n=0}^{\infty} \frac{1}{n!} e^{-\lambda} \underbrace{\langle 1 - p_d,\, v \rangle^n}_{=\lambda^n (1-\alpha)^n} \\ &= e^{-\mu - \alpha\lambda} \end{aligned} \tag{16}$$

where $\lambda = \langle 1,\, v \rangle$ is the expected number of targets and α is the expected fraction of the targets detected

$$\alpha = 1 - \lambda^{-1} \langle 1 - p_d,\, v \rangle = \lambda^{-1} \langle p_d,\, v \rangle . \tag{17}$$

This is easily extended to the multi-robot case. Let C_0 be the set of robots with $y^r = 0$ and C_1 the set of robots with $y^r = 1$. Then

$$\begin{aligned} p(y^1, \ldots, y^R) &= \int \prod_{r \in C_0} p(y^r = 0 \mid X) \\ &\quad \times \prod_{r \in C_1} (1 - p(y^r = 0 \mid X)) p(X)\, \delta X \\ &= \int \sum_{C \subseteq C_1} (-1)^{|C|} \prod_{r \in C_0 \cup C} p(y^r = 0 \mid X) \\ &\quad \times p(X)\, \delta X \\ &= \sum_{C \subseteq C_1} (-1)^{|C|} e^{-|C_0 \cup C| \mu - \alpha(C_0 \cup C)\lambda}, \end{aligned} \tag{18}$$

where

$$\alpha(C) = 1 - \lambda^{-1} \left\langle \prod_{r \in C} (1 - p_d^r),\, v \right\rangle \tag{19}$$

is the expected fraction of targets detected by at least one robot in group C. We substitute this in to the standard definition of entropy,

$$H[\mathcal{Y}] = - \langle p(y),\, \ln p(y) \rangle , \tag{20}$$

where there are 2^{RT} possible binary measurement combinations for R robots and T time steps.

2) Conditional Entropy: The conditional entropy is simpler as measurement sets are conditionally independent of one another given the target set, *i.e.,*

$$p(y_\tau^{1:R} \mid X) = \sum_{k \in \tau} \sum_{r=1}^{R} p(y_k^r \mid X). \tag{21}$$

Thus the conditional entropy of the joint measurements is simply the sum of the conditional entropies of the individual measurements, so we only need the single measurement equation

$$H[\mathcal{Y} \mid \mathcal{X}] = - \int \left(\sum_{y \in \{0,1\}} p(y \mid X) \ln p(y \mid X) \right) p(X)\, \delta X. \tag{22}$$

We separate the two cases for y, beginning with $y = 0$.

$$\begin{aligned} &\int p(y=0 \mid X) p(X) \ln p(y=0 \mid X)\, \delta X \\ &\quad = e^{-\mu} \ln e^{-\mu} \sum_{n=0}^{\infty} \frac{1}{n!} e^{-\lambda} \lambda^n (1-\alpha)^n \\ &\qquad + e^{-\mu} \sum_{n=0}^{\infty} n \frac{1}{n!} e^{-\lambda} \lambda^{n-1} (1-\alpha)^{n-1} \\ &\qquad\quad \times \underbrace{\langle (1 - p_d) \ln (1 - p_d),\, v \rangle}_{-\lambda\beta} \\ &\quad = -(\mu + \lambda\beta) e^{-\mu - \alpha\lambda}. \end{aligned} \tag{23}$$

The negative sign in β is due to the entropy-like definition.

Next we examine the $y = 1$ case, using the Taylor series $\ln(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}$, where $\{r\}^k$ is a set with k copies of the robot r.

$$\begin{aligned} &\int p(y=1 \mid X) p(X) \ln p(y=1 \mid X)\, \delta X \\ &\quad = \int (1 - p(y=0 \mid X)) p(X) \ln (1 - p(y=0 \mid X))\, \delta X \\ &\quad = \int (1 - p(y=0 \mid X)) p(X) \sum_{k=1}^{\infty} \frac{-p(y=0 \mid X)^k}{k}\, \delta X \\ &\quad = \int \left(-p(y=0 \mid X) + \sum_{k=2}^{\infty} \frac{p(y=0 \mid X)^k}{k(k-1)} \right) p(X)\, \delta X \\ &\quad = -e^{-\mu - \alpha\lambda} + \sum_{k=2}^{\infty} \frac{1}{k(k-1)} e^{-k\mu - \alpha(\{r\}^k)\lambda} \end{aligned} \tag{24}$$

Note that $\alpha(\{r\}^k)$ may be computed using (19) and we use the first 10 terms in the Taylor series.

3) Computational Complexity: The computational complexity of the entropy computations is $\mathcal{O}(|Q^{1:R}| 2^{2RT})$, where R is the number of robots, T is the planning horizon, and $|Q^{1:R}|$ is the action set (described in further detail in Sec III-A).

D. Exploration Termination Criterion

It is unclear when to terminate the exploration in the multi-target localization problem, since the number of targets being sought is unknown. Ideally the robots should identify the exact number of targets and their locations within the environment, *i.e.,* if there are N targets in the environment then the PHD should be N Dirac delta functions, each of unit weight, centered at the true target locations. In reality, the estimates will never be as precise, but the difference between the estimated PHD and its idealized counterpart may be used to detect when the robots have sufficient confidence in their estimate. Note that there is no way to determine the team's confidence in the cardinality estimate, as the covariance is equal to the mean for a Poisson distribution. Thus we

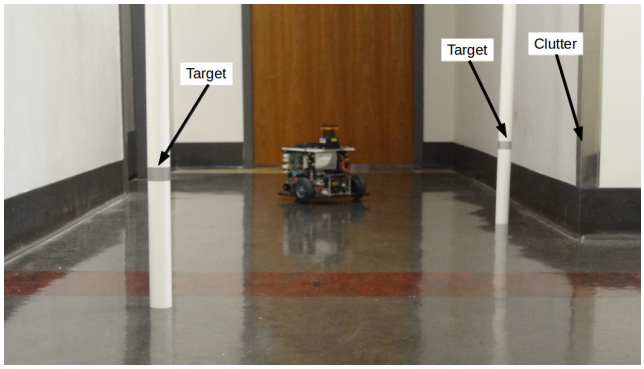


Fig. 3. A Scarab robot with two targets in the experimental environment.

assume that the team is able to accurately estimate the target cardinality, *i.e.,* $\lambda \to N$ as $t \to \infty$.

We turn our attention to the entropy of a Poisson RFS,

$$H[\mathcal{X}] = \lambda - \int v(x) \log v(x)\, dx \tag{25}$$

$$= \lambda + \lambda\big(H[\bar{v}(x)] - \log \lambda\big), \tag{26}$$

where $\bar{v}(x) = \lambda^{-1} v(x)$ is a probability distribution created by normalizing the PHD. This comes from plugging the likelihood of a Poisson RFS, (6), into the formula for Shannon's entropy, (20). The ideal PHD $v^*(x)$ consists of λ particles of unit weight, so $\bar{v}^*(x)$ has λ particles of weights λ^{-1} and has an entropy of $\log \lambda$. The term in parentheses in (26) can thus be interpreted as the difference between the current normalized PHD and its idealized counterpart. The proposed termination criterion can thus be written in the two equivalent forms,

$$H[\bar{v}(x)] - \log \lambda \leq \epsilon \tag{27}$$

$$\lambda^{-1} H[\mathcal{X}] - 1 \leq \epsilon. \tag{28}$$

IV. Experimental System and Results

We conduct a series of experiments using a team of ground robots (Scarabs), pictured in Fig. 3, to validate the performance of the proposed control algorithm. The Scarabs are differential drive robots with an onboard computer with an Intel i5 processor and 4 GB of RAM, running Ubuntu 12.04. They are equipped with a Hokuyo UTM-30LX laser scanner, used for self-localization and for target detection. The robots communicate with a central computer, a laptop with an Intel i7 processor and 16 GB of RAM running ROS on Ubuntu 12.04, via an 802.11n network. The communication requirements are very modest: individual agents upload their measurement sets, which consist of bearing values, and pose estimates to the central server, and the central server sends out actions to each robot, which consist of a sequence of T poses. The team explores in an indoor hallway in the Levine building, shown in Fig. 4, seeking the reflective targets pictured with the robot in Fig. 3.

We use a simple sensor to test the fundamental idea of the paper, the information-based control law from Sec. III: converting the Hokuyo into a bearing-only sensor. This may be thought of as a proxy to a camera, but it avoids common problems with visual sensors such as variable lighting conditions

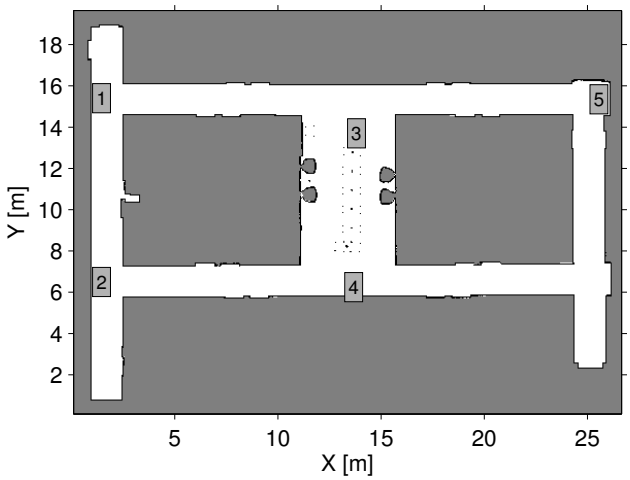


Fig. 4. A floorplan of the Levine environment used in the hardware experiments. This map was generated using a manually driven Scarab robot and the `gmapping` package from ROS [27]. Different starting locations for the robots are labeled in the map.

and distortions. Atanasov et al. [28] use the RFS framework to perform semantic self-localization of a robot equipped with a camera, using bearing-only measurements to landmarks.

The targets are 1.625 in outer diameter PVC pipes with attached 3M 7610 reflective tape. The tape provides high intensity returns to the laser scanner, allowing us to pick out targets from the background environment. However, there is no way to uniquely identify individual targets, making this the ideal setting to use the PHD filter. The hallway features a variety of building materials such as drywall, wooden doors, painted metal (door frames), glass (office windows), and bare metal (chair legs, access panels, and drywall corner protectors, like that in the right side of Fig. 3). The reflective properties of the environment vary according to the material and the angle of incidence of the laser. The intensity of bare metal and glass surfaces at low angles of incidence is similar to that of the reflective tape, leading to false positive detections.

To turn a laser scan into a set of bearing measurements, we first prune the points based on the laser intensity threshold, retaining only those with sufficiently high intensity returns. The points are clustered spatially using the range and bearing information, with each cluster having a maximum diameter d_t. The range data is otherwise discarded. The bearings to each of the subsequent clusters form a measurement set Z.

A. Sensor Models

We now develop the detection, measurement, and clutter models necessary to utilize the PHD filter. See Dames and Kumar [29] for the experimental procedure used to determine the sensor parameters used in this work.

1) Detection Model: The detection model can be determined using simple geometric reasoning due to the nature of the laser scanner, as Fig. 5 shows. Each beam in a laser scan intersects a target that is within $d_t/2$ of the beam. The arc length between two beams at a range r is $r\theta_{\text{sep}}$, and the covered space is d_t. Using the small angle approximation for

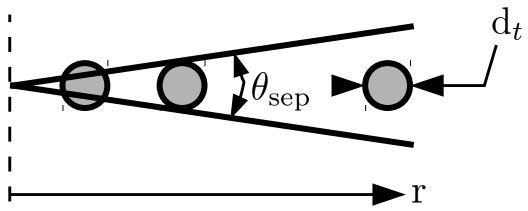


Fig. 5. A pictogram of the laser detection model, where d_t is the diameter of the target, θ_{sep} is the angular separation between beams, and r is the range.

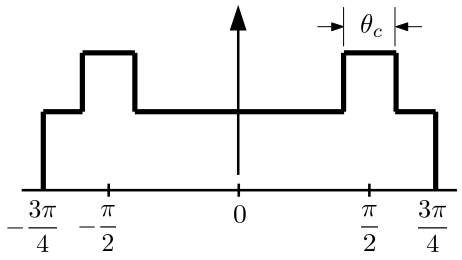


Fig. 6. A pictogram of the clutter model, where θ_c is the width of the clutter peaks centered at $\pm\frac{\pi}{2}$, and the bearing falls within the range $[-\frac{3\pi}{4}, \frac{3\pi}{4}]$.

tangent, the probability of detection is

$$p_d(x;q) = (1-p_{\text{fn}}) \min\left(1, \frac{d_t}{r(x,q)\theta_{\text{sep}}}\right) \times \mathbf{1}\left(b(x,q) \in [b_{\min}, b_{\max}]\right) \mathbf{1}\left(r(x,q) \in [0, r_{\max}]\right) \quad (29)$$

where $r(x,q)$ and $b(x,q)$ are the target range and bearing in the local sensor frame (computed using the robot pose q and the target position x), p_{fn} is the probability of a false negative, and $\mathbf{1}(\cdot)$ is an indicator function. The bearing is limited to fall within $[b_{\min}, b_{\max}]$ and the range to be less than some maximum value $r_{\max}$ (here due to the intensity threshold on the laser). For our sensor, $b_{\max} = -b_{\min} = \frac{3\pi}{4}$, $r_{\max} = 5\,\text{m}$, $p_{\text{fn}} = 0.210$, and $d_t = 1.28\,\text{in}$. Note that the effective target diameter is less than the true target diameter, since the reflective tape does not provide high intensity returns at extreme angles of incidence.

2) Measurement Model: The sensor returns a bearing measurement to each detected target. We assume that bearing measurements are corrupted by Gaussian noise with covariance σ^2, which is independent of the robot pose and the range and bearing to the target. In other words,

$$g(z \mid x;q) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\big(z - b(x,q)\big)^2}{2\sigma^2}\right), \quad (30)$$

where $b(x,q)$ is the bearing of the target in the sensor frame. For our system, $\sigma = 2.25°$.

3) Clutter Model: As previously noted, clutter (*i.e.,* false positive) measurements arise due to reflective surfaces within the environment, such as glass and bare metal, only at low angles of incidence. For these materials, this most often happens while driving down a hallway, so there will be a higher rate of clutter detections near $\pm\frac{\pi}{2}$ rad in the laser scan. For objects such as table and chair legs, there is no clear relationship between the relative pose of the object and robot, so we assume that such detections occur uniformly across the field of view of the sensor. This leads to a clutter model of the form shown in Fig. 6.

Let θ_c be the width of the clutter peaks centered at $\pm\frac{\pi}{2}$ and let p_u be the probability that a clutter measurements was generated from a target in the uniform component of the clutter model. The clutter PHD is

$$c(z) = \frac{p_u \mu}{b_{\max} - b_{\min}} \mathbf{1}\left(z \in [b_{\min}, b_{\max}]\right) + \frac{(1-p_u)\mu}{2\theta_c} \mathbf{1}\left(\big||z| - \pi/2\big| \le \theta_c/2\right), \quad (31)$$

where $\mathbf{1}(\cdot)$ is an indicator function and μ is the expected number of clutter measurements per scan. For our system, $b_{\max} = -b_{\min} = 3\pi/4$, $p_u = 0.725$, $\theta_c = 0.200\pi$, and $\mu = 0.532$. The clutter cardinality is assumed to follow a Poisson distribution with mean μ, as noted in Sec. II-D.

B. PHD Filter Implementation Details

The PHD filter is typically implemented as either a weighted particle set, see Vo et al. [30], or a mixture of Gaussians, see Vo and Ma [31]. We use the particle representation as it allows for nonlinear measurement models, such as the bearing-only case described above. Since we assume no knowledge of the initial positions of targets, the particles are initialized with equal weight on a uniformly-spaced grid at all locations at which a target is visible, *e.g.,* in free space for a bearing-only sensor. Ideally, the grid size should be set to a similar length scale as the sensor noise, as below this scale the sensor cannot disambiguate targets. Particles are stationary during the course of the experiment as we assume that targets are stationary. The complexity of the control objective (13) is $\mathcal{O}(|Q^{1:R}|2^{RT}(PRT + 2^{RT}))$, where P is the number of particles in the PHD representation.

To reduce the computational complexity of the controller, we subsample the PHD estimate from the filter. We create a uniform grid over the environment, merging all particles within a grid cell into a super-particle with weight equal to the total weight of the merged particles, and position equal to the weighted mean of the merged particles. This is similar to the idea from Charrow et al. [14].

C. Validation

To evaluate the real-world performance of the proposed algorithm, we run a set of 10 trials in which 3 robots seek 15 targets placed within the office environment from Fig. 4. The robots all start near location 1, separated by 0.5 m. The team initially believes there are 30 targets in the environment. The team uses a time horizon of $T = 3$ actions and uses planning mode 1 (concurrently planning over the robots and sequentially over length scales). The robots search over length scales starting at $\ell = 3\,\text{m}$, and increasing by a factor of 1.2 until some robot no longer has any possible destinations due to the limited size of the environment. The termination criterion is set to $\epsilon = 0.05$. Fig. 7a shows the target cardinality

estimates for the team, with the exploration taking 300–500 s to complete.

Fig. 7b shows the average performance of the team across the trials. The average expected number of targets approaches the true cardinality after approximately 150 s, and stays close for the remainder of the time, showing that the estimator is accurate. The shaded region shows one standard deviation from the mean (where the standard deviation is computed across trials at a given time step) and generally decreases over time, showing that the estimates are also consistent. We scale each run to be of the median time to completion, to be able to compute the standard deviation at a given time instant across runs of different lengths.

Fig. 7c shows the true target locations and the localization estimate from a single representative trial. There are 15 unique dots on the map, corresponding to the 15 target locations. The size of the dots is proportional to the expected number of targets at that location. Some targets are better identified than others, *e.g.,* the target in the top middle of the map is larger than some of the other targets. There is also a false positive target near (12, 12) m of low weight compared to the true targets. The accompanying video contains a visualization of the data collected by the robots during the course of a representative trial.

The computational complexity is relatively low, with control actions taking an average of 1.01 s, and a maximum of 3.37 s, to compute. The team spent 4.7% of the total exploration time stationary, planning their next actions: a small, but not negligible, fraction of the time.

D. Team Size Comparison

We conduct a series of 10 trials each with 1, 3, and 5 robots in the same environment as above to explore the effect of team size on the exploration performance. For the 5 robot trials, one robot starts at each of the labeled locations in Fig. 4; for the 3 robot trials they begin at locations 1–3; and for the single robot trials it begins at location 3. The robots use planning mode 3 (sequentially over both robots and length scales), as concurrently planning over robots is prohibitively expensive for 5 robots. All of the other parameters are identical to the previous trials.

The team is given a time budget of 400 s to complete the exploration task, with Fig. 8a showing the statistics of the completion times. Within the time budget, the single robot never completes the task, the three robot team completes it in 3 of the 10 runs, and the 5 robot team completes it every time, in a median of 270 s and a maximum of 373 s. As is expected, adding more robots improves completion time, as they are able to simultaneously gather measurements from more locations than a smaller team. Figs. 8b and 8c show the average cardinality estimates and target set entropies for each of the team sizes. The 5 robot trials have the highest rate of entropy reduction and the 3 robot teams nearly finish exploring the environment, with the entropy approaching the terminal value at the end of the trials. The single robot case is furthest from convergence, with the final entropy at the same level that a 3 robot team achieves in 34% of the time.

With planning mode 3, the computational load is minor: taking an average of 0.029 s for 1 robot, 0.092 s for 3 robots, and 0.351 s for 5 robots. This is 0.18% of the total time for 1 robot, 0.42% for 3 robots, and 1.45% for 5 robots, all less than the Mode 1 planning in the previous experiments. However, mode 3 is not guaranteed to return plans with as high of an expected information gain as mode 1.

V. Simulation

We also conduct a series of simulation experiments to further explore the performance of the proposed control strategy (13), varying the planning method, target density, environment, and sensing modality.

A. Simulator Validation

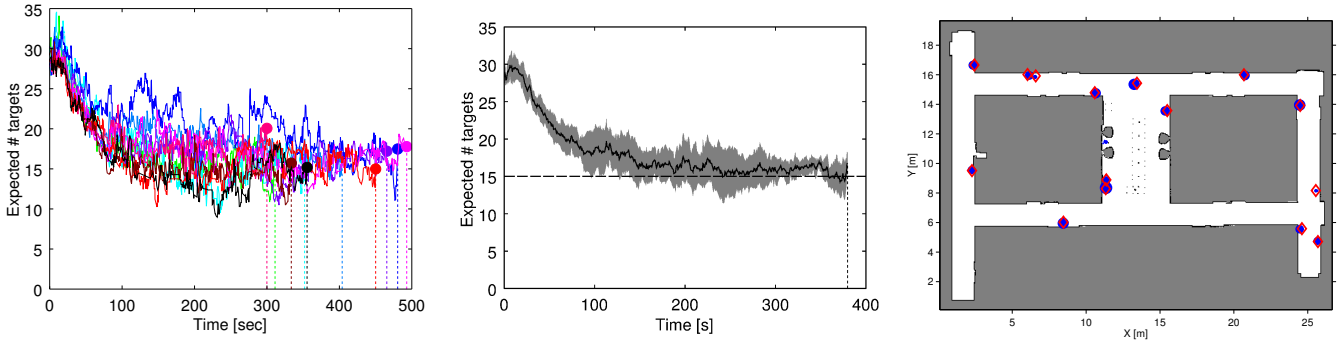
We wish to verify that the simulation environment behaves similarly to the experimental system before conducting a long series of trials in simulation. To do this, we mimic the setup from Sec. IV-C as closely as possible, using identical sensor parameters, controller parameters, team size, planning method, etc. The target locations for the simulation are set to the true locations shown in Fig. 7c.

Overall, the results in Fig 9 show that the two systems are comparable, with both systems able to accurately and consistently estimate the target set cardinality and reach the desired level of confidence in their estimate. The experimental data is more consistent across runs, both in terms of completion time and for inter-run estimates of the target set cardinality and entropy. However, the simulated system has a lower median time of completion, at 338 s compared to 392 s. While there are some differences, overall the systems are similar enough to trust that further simulated results will not differ significantly from experimental results.

B. Planning Method Comparison

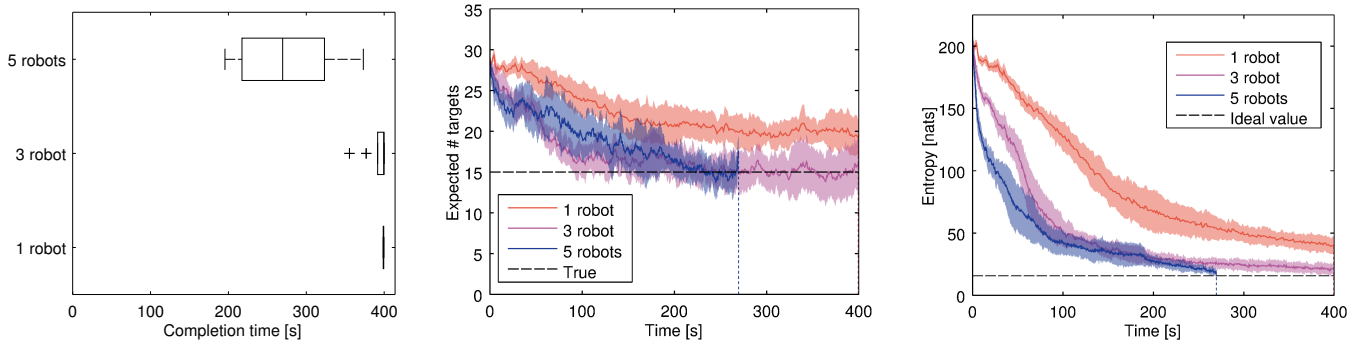
In Sec. IV, we use planning modes 1 and 3, but could not make direct comparisons between the two due to the different team setups. We now wish to see how the different planning methods affect the team's performance, and verify that taking intelligent actions (*i.e.,* maximizing mutual information) outperforms a naïve random walk. In theory, mode 0 leads to plans with the highest expected information gain, but the plans would take longer to compute, potentially causing the actual information gain over time to decrease. Modes 1 to 3 are all approximations, sequentially planning over the length scales, team members, or both, and robots using modes 4 and 5 randomly select actions.

We use the same setup as Sec.IV-C, but vary the planning modality and set a time budget of 900 s. Fig. 10 shows the results of the trials. Information-based control of any kind significantly outperforms the random walk in terms of completion time and estimation accuracy. The information-based methods (modes 1–3) all converge to the desired target set entropy in all of the trials, mode 4 (random walk with concurrent length scales) completes the exploration before the time budget expires in 2 of the 10 trials, and mode 5 (random



(a) Expected cardinality over time (b) Average expected cardinality as a function of normalized time (c) Estimated target locations

Fig. 7. Plots of the performance of a team of three real-world robots exploring the Levine environment using planning mode 1. (a) Shows the expected cardinality of the team over time, with the final cardinality in each run marked by a circle and the final time as a dotted vertical line. (b) Shows the mean (solid black line) and standard deviation (shaded region) of the expected cardinality across runs over time, with the true cardinality shown (dashed black line). (c) Shows the true (red diamonds) and estimated target locations (blue dots), with the dot size proportional to the estimated number of targets at that location.



(a) Completion times (b) Average expected cardinality as a function of time (c) Average target set entropy as a function of time

Fig. 8. Plots of the performance for teams of 1, 3, and 5 real-world robots exploring the Levine environment using planning mode 3. (a) Shows the spread of time to completion. (b) Shows the mean (solid lines) and standard deviation (shaded regions) of the expected cardinality across runs over time budget, with the true cardinality shown (dashed black line). (c) Shows the mean (solid lines) and standard deviation (shaded regions) of the entropy across runs over time with the ideal value shown (dashed black line).

walk with sequential length scales) never completes the task. Mode 5 is also much less consistent than all of the other modes in terms of the rate of entropy reduction.

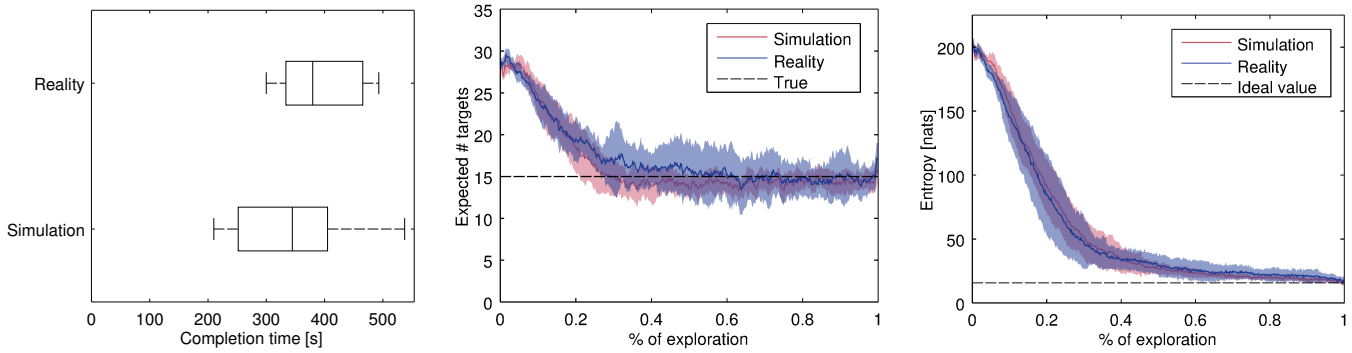
It is not surprising that mode 2 (planning sequentially over robots and concurrently over length scales) has the lowest median completion time and that the spread is narrower compared to modes 1 and 3. When robots are allowed to plan over different length scales, some robots explore local regions of high uncertainty while other robots move across the environment to search for new targets, allowing the team to more efficiently explore. Using modes 1 and 3, all robots act at the same length scale so some robots must occasionally act on a undesirable length scale for the benefit of other team members. It is surprising that mode 1 leads to the most inconsistent completion times, as the expected information gain is an upper bound for the gain in mode 3. The differences could have been due to chance, as there are only 10 trials with each planning modality.

Both of the random exploration methods (modes 4 and 5) perform significantly worse than the information-based planning. Not only does mode 5 never complete the exploration in the given time budget, its rate of entropy reduction is significantly slower than all of the other methods, including mode 4. While the planning times for the random methods are negligibly small, as Fig. 10c shows, this does not counteract the fact that the actions are not being selected in an intelligent manner.

C. Target Cardinality Comparison

We next test the performance of the system in situations with variable target cardinalities, and, correspondingly, variable target densities in the environment. The PHD filter functions for any target cardinality and the exploration controller is agnostic to the target cardinality. We conduct a series of experiments with the same setup, except we use 1, 15, and 100 targets in the Levine environment, shown Fig. 4, which is approximately $144\,\mathrm{m}^2$. Three robots explore, starting at location 1 in the map and using planning mode 2.

Fig. 11 shows the completion times, cardinality estimates, and target set entropies. As expected, the low target density is the fastest to complete, since the team simply needs to sweep out the mostly empty space. With the high target density the robots need to observe the many targets from a variety of vantage points to localize them with sufficient confidence,

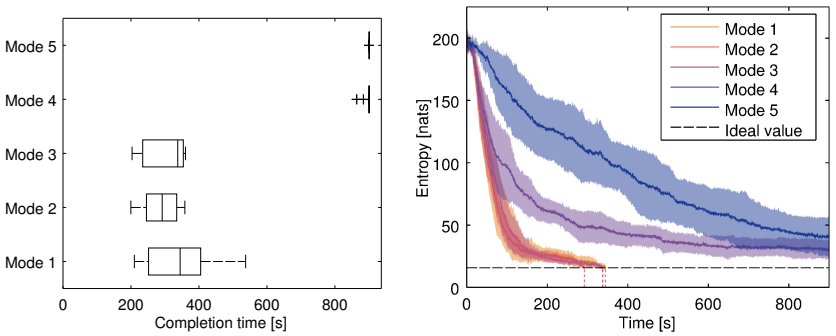


(a) Completion times

(b) Average expected cardinality as a function of normalized time

(c) Average target set entropy as a function of normalized time

Fig. 9. Plots of the performance for teams of three real and simulated robots exploring the Levine environment using planing mode 1. (a) Shows the spread of time to completion. (b) Shows the mean (solid lines) and standard deviation (shaded regions) of the expected cardinality across runs as a fraction of the total time with the true cardinality shown (dashed black line). (c) Shows the mean (solid lines) and standard deviation (shaded regions) of the entropy across runs as a fraction of the total time with the ideal value shown (dashed black line).



Mode	Avg. time	Max time	% total time
1	1240	1790	9.50
2	162	188	0.86
3	211	246	1.63
4	0.3	0.4	0.002
5	0.8	0.9	0.01

(a) Completion times

(b) Average target set entropy as a function of time

(c) Table of computation times in ms

Fig. 10. Plots of the performance for a team of three simulated robots exploring the Levine environment using planning modes 1–5. (a) Shows the spread of time to completion. (b) Shows the mean (solid lines) and standard deviation (shaded regions) of the entropy across over time budget with the ideal value shown (dashed black line). (c) Shows the computation times in ms, and the percentage of the total time spent computing.

a process which involves sweeping across the environment multiple times. The variability in completion time, the time to correctly estimate the true cardinality, and the inconsistency of the cardinality estimates also increases with the target cardinality. In fact, on average, the high target cardinality runs do not reach the correct cardinality until about 95% of the way through exploration.

Fig. 11c shows that the team reaches the desired level of uncertainty at the end of each run. Note that for the high cardinality runs, the initial rate of target discovery is higher than the initial rate of target localization, resulting in an increase in entropy for the first 40 s of the run.

D. Second Environment

We next conduct a series of simulations in the larger, more complex indoor environment shown in Fig. 12a. This environment features many rooms for the robots to explore, and is nearly four times the area of the Levine environment. There are 40 targets in the environment, and the termination criterion is increased to $\epsilon = 1$. Three robots begin in room 1 in the map and use planning mode 2. All of the other system parameters are identical.

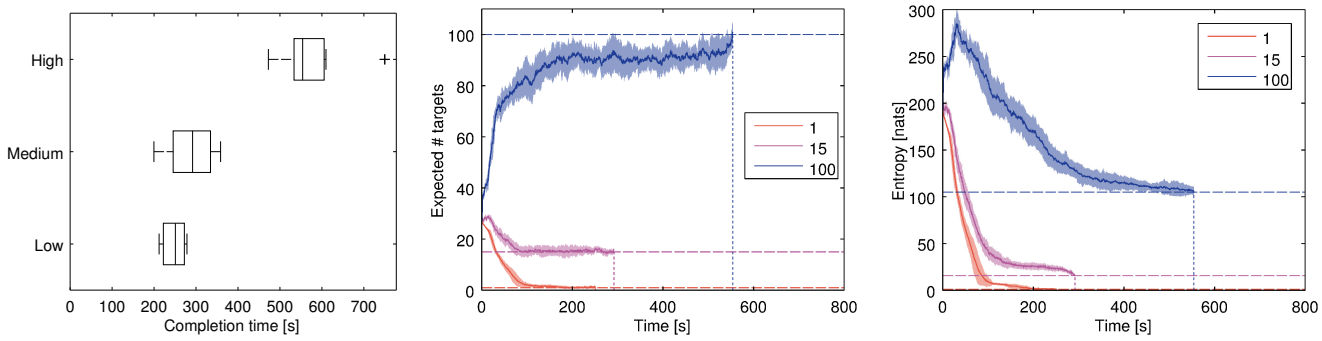
Fig. 12 shows that the team is able to accurately and consistently estimate the true target cardinality and reach the desired level of confidence in the target estimate. The robots take significantly longer to complete the exploration compared to the Levine environment, with a median time of 2220 s, 6.56 times as high. We expect the time to be at least 4 times as high due to the increase in area, with the extra time likely caused by the increased complexity, as the robots must enter many individual rooms.

Since the environment is larger, there are more length scales for the robots to consider, and thus more actions. This increases the planning time. Planning mode 2 takes an average of 1.63 s per plan, and in total is 9.2% of the exploration time.

E. Range-Only Sensing

So long as we are able to create detection, measurement, and clutter models for the sensor, nothing about the estimation or control framework relies upon the specific sensor modality. To verify this, we conduct a final series of simulation experiments in which robots are equipped with noisy, range-only sensors.

1) Sensor Models: The range-only sensor parameters used in this case are not based on a particular physical sensor,



(a) Completion times

(b) Average expected cardinality as a function of normalized time

(c) Average target set entropy as a function of normalized time

Fig. 11. Plots of the performance for a team of three simulated robots exploring the Levine environment for 1, 15, or 100 targets using planning mode 2. (a) Shows the spread of time to completion. (b) Shows the mean (solid lines) and standard deviation (shaded regions) of the expected cardinality across runs as a fraction of the total time with the true cardinality shown (dashed black line). (c) Shows the mean (solid lines) and standard deviation (shaded regions) of the entropy across runs as a fraction of the total time with the ideal value shown (dashed black line).

but rather seek to capture the general behavior of an RF-based range sensor. Fig. 13a shows the detection model for the sensors, $p_d(x;q)$, which decays steadily with distance. The measurements have zero-mean Gaussian noise, so $z \sim \mathcal{N}(|x-q|, \sigma^2)$, where $\sigma = 1\,\mathrm{m}$. The measurement noise is relatively high compared to the bearing-only sensor, so we expect the rate of information gain to be lower. Clutter detections occur uniformly over the sensor footprint, with a clutter PHD $c(z) = \mu/r_{\max}$, where $\mu = 0.1$ is the clutter rate and $r_{\max} = 5\,\mathrm{m}$ is the maximum range of the sensor.

2) Results: Most of the simulation parameters are kept constant: a team of 3 robots begins at location 1 in Levine, and uses planning mode 2 with the same length scales as the bearing-only sensor. The termination criterion is $\epsilon = 3$ to account for the much coarser localization that the range-only sensor is able to achieve, due to the high measurement noise.

Fig. 13 shows the resulting completion times, cardinality estimates, and target set entropies, as well as an example localization result. As is expected, the system takes longer to complete the localization task, and the resulting target estimates are not as precise as with the bearing-only sensor. Despite the errors in target localization, the team is still able to accurately estimate the true target cardinality and to discover the approximate locations of all of the targets.

VI. Conclusion

In this paper, we propose a novel receding-horizon, information-based controller for actively detecting and localizing an unknown number of targets using a small team of autonomous mobile robots. The robots are equipped with unreliable sensors, which may fail to detect targets within the field of view, may return false positive detections, and may be unable to uniquely identify true targets. Despite this, the PHD filter simultaneously estimates the number of targets and their locations, avoiding the need to explicitly consider data association and providing a scalable approach for various team sizes, sensor modalities, and environments.

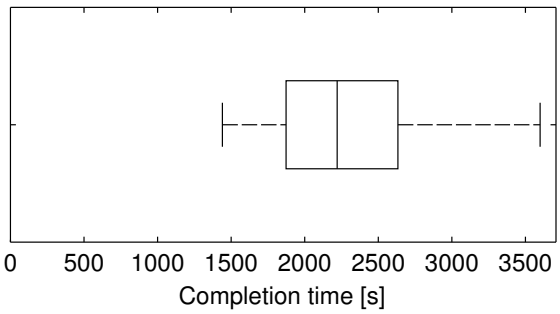
The controller, which maximizes the mutual information between the target set and the future binary measurements of the team, hedges against highly uninformative actions in a computationally tractable manner. We provide several variations on the controller, concurrently or sequentially planning across robots in the team and length scales of actions. We demonstrate the effectiveness of our control strategy through a series of hardware experiments with small teams of ground robots exploring an indoor office environment. A series of simulated experiments show that the proposed approach performs well in a variety of settings: with low and high target cardinality, in multiple environments, and with multiple sensor modalities. The proposed control law also significantly outperforms a random walk through the environment without significantly increasing the computational load. The team is able to autonomously cease exploration once its confidence in the target estimates is sufficiently high.

Future work on this subject will aim to extend the active control framework presented in this paper to other scenarios. There are many information gathering problems where targets are not stationary, *e.g.,* surveillance and security scenarios and environmental monitoring. Such scenarios will be the subject of future work. We previously introduced a communication architecture to perform this multi-target estimation and active control in a decentralized manner [22]. Future work will address other communication issues, such as missing and erroneous information, that may arise in more complex environments.

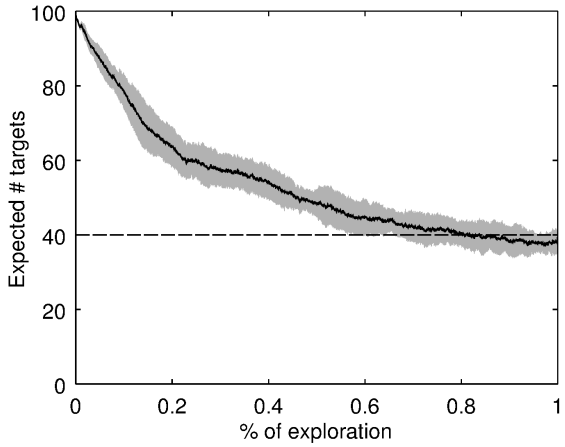
Acknowledgment

This work was funded in part by ONR MURI Grants N00014-07-1-0829, N00014-09-1-1051, and N00014-09-1-1031, the SMART Future Mobility project, and TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA. Philip Dames was supported by the Department of Defense through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

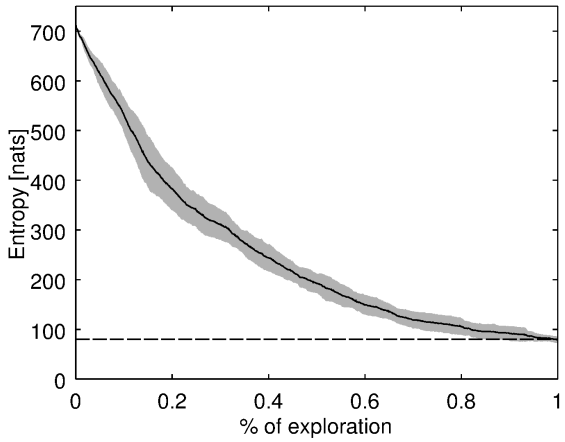
(a) Floorplan



(b) Completion times



(c) Average expected cardinality as a function of normalized time



(d) Average target set entropy as a function of normalized time

Fig. 12. Plots of the performance for a team of three simulated robots exploring a second environment using planning mode 2. (a) Shows the floorplan of the complex, indoor environment used in simulations. The robots begin in room 1 in the upper left corner. (b) Shows the spread of time to completion. (c) Shows the mean (solid lines) and standard deviation (shaded regions) of the expected cardinality across runs as a fraction of the total time with the true cardinality shown (dashed black line). (d) Shows the mean (solid lines) and standard deviation (shaded regions) of the entropy across runs as a fraction of the total time with the ideal value shown (dashed black line).

REFERENCES

[1] A. Rowe, M. E. Berges, G. Bhatia, E. Goldman, R. Rajkumar, J. H. Garrett, J. M. F. Moura, and L. Soibelman, "Sensor Andrew: Large-scale campus-wide sensing and actuation," *IBM Journal of Research and Development*, vol. 55, no. 1.2, pp. 6:1–6:14, Jan. 2011.

[2] K. Fu. (2014, Jan.) RFID-scale devices in concrete. [Online]. Available: https://spqr.eecs.umich.edu/moo/apps/concrete/

[3] R. Mahler, "Multitarget bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[4] L.-L. Ong, T. Bailey, H. Durrant-Whyte, and B. Upcroft, "Decentralised particle filtering for multiple target tracking in wireless sensor networks," in *IEEE International Conference on Data Fusion*, 2008.

[5] M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, "A Solution to the Simultaneous Localization and Map Building (SLAM) Problem," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 3, pp. 229–241, 2001.

[6] L. D. Stone, R. L. Streit, T. L. Corwin, and K. L. Bell, *Bayesian Multiple Target Tracking*. Artech House, 2013.

[7] R. Mahler, *Statistical multisource-multitarget information fusion*. Artech House Boston, 2007, vol. 685.

[8] O. Erdinc, P. Willett, and Y. Bar-Shalom, "The bin-occupancy filter and its connection to the PHD filters," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4232–4246, 2009.

[9] G. Hoffmann and C. Tomlin, "Mobile sensor network control using mutual information methods and particle filters," *IEEE Transactions on Automatic Control*, pp. 1–16, 2010.

[10] B. Julian, M. Angermann, M. Schwager, and D. Rus, "Distributed robotic sensor networks: An information-theoretic approach," *The International Journal of Robotics Research*, vol. 31, no. 10, pp. 1134–1154, Aug. 2012.

[11] G. A. Hollinger, B. Englot, F. S. Hover, U. Mitra, and G. S. Sukhatme, "Active planning for underwater inspection and the benefit of adaptivity," *The International Journal of Robotics Research*, vol. 32, no. 1, pp. 3–18, Nov. 2012.

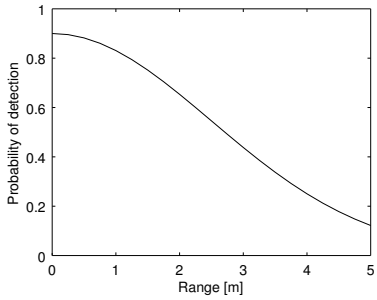
[12] B. J. Julian, S. Karaman, and D. Rus, "On mutual information-based control of range sensing robots for mapping applications," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Nov. 2013, pp. 5156–5163.

[13] J. R. Souza, R. Marchant, L. Ott, D. F. Wolf, and F. Ramos, "Bayesian Optimisation for Active Perception and Smooth Navigation," in *IEEE International Conference on Robotics and Automation*, 2014.

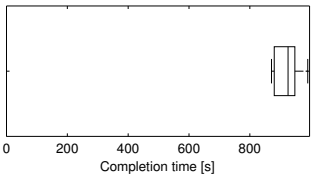
[14] B. Charrow, V. Kumar, and N. Michael, "Approximate representations for multi-robot control policies that maximize mutual information," *Autonomous Robots*, vol. 37, no. 4, pp. 383–400, 2014.

[15] B. Charrow, N. Michael, and V. Kumar, "Active control strategies for discovering and localizing devices with range-only sensors," in *Workshop on the Algorithmic Foundations of Robotics*, 2014.

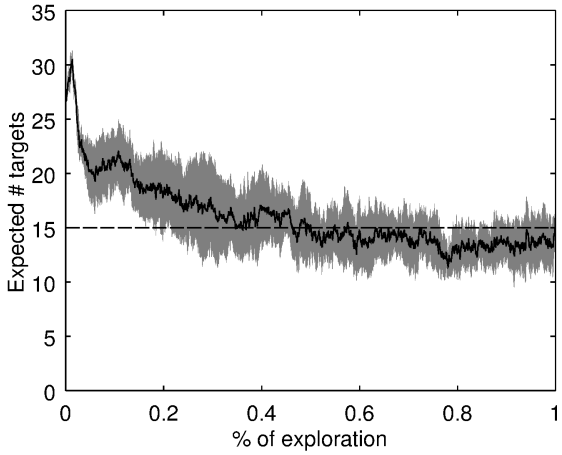
[16] D. Q. Mayne and H. Michalska, "Receding horizon control of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 7, pp.



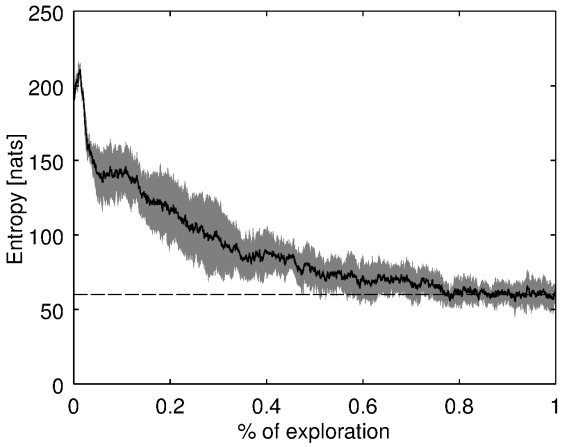
(a) Range-only detection model



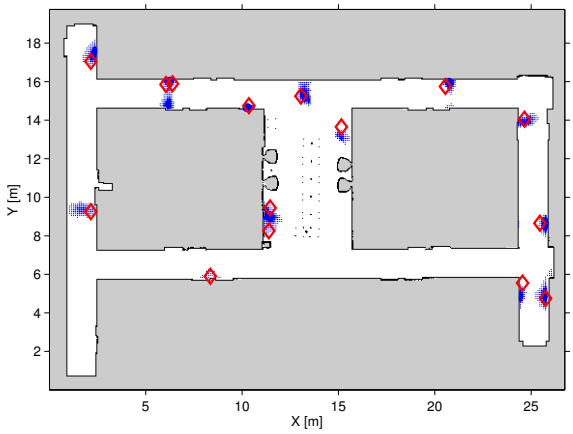
(b) Completion times



(c) Average expected cardinality as a function of normalized time



(d) Average target set entropy as a function of normalized time



(e) Estimated target locations

Fig. 13. Plots of the performance for a team of three simulated robots equipped with range-only sensors exploring the Levine environment using planning mode 2. (a) Shows the detection model used in the simulation trials. (b) Shows the spread of time to completion. (c) Shows the mean (solid lines) and standard deviation (shaded regions) of the expected cardinality across runs as a fraction of the total time with the true cardinality shown (dashed black line). (d) Shows the mean (solid lines) and standard deviation (shaded regions) of the entropy across runs as a fraction of the total time with the ideal value shown (dashed black line). (e) Shows the true (red diamonds) and estimated (blue dots) target locations, with dot size proportional to the number of targets at that location.

814–824, 1990.

[17] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[18] A. Ryan, "Information-theoretic control for mobile sensor teams," Ph.D. dissertation, University of California, Berkeley, 2008.

[19] B. Ristic and B.-N. Vo, "Sensor control for multi-object state-space estimation using random finite sets," *Automatica*, vol. 46, no. 11, pp. 1812–1818, Nov. 2010.

[20] B. Ristic, B.-N. Vo, and D. Clark, "A note on the reward function for PHD filters with sensor control," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1521–1529, 2011.

[21] P. Dames, M. Schwager, V. Kumar, and D. Rus, "A decentralized control policy for adaptive information gathering in hazardous environments," in *IEEE Conference on Decision and Control (CDC)*, Dec. 2012, pp. 2807–2813.

[22] P. Dames and V. Kumar, "Cooperative multi-target localization with noisy sensors," in *IEEE International Conference on Robotics and Automation*, May 2013, pp. 1877–1883.

[23] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes*. Springer, 2003, vol. 2.

[24] T. Cover and J. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[25] R. Mahler, "Objective functions for bayesian control-theoretic sensor management, 1: multitarget first-moment approximation," in *IEEE Aerospace Conference Proceedings*, vol. 4. Ieee, 2003.

[26] C. Kreucher, K. Kastella, and A. O. Hero III, "Sensor management using an active sensing approach," *Signal Processing*, vol. 85, no. 3, pp. 607–624, 2005.

[27] B. Gerkey. (2014, Jul.) gmapping. [Online]. Available: http://wiki.ros.org/gmapping

[28] N. Atanasov, M. Zhu, K. Daniilidis, and G. J. Pappas, "Semantic Localization Via the Matrix Permanent," in *Robotics: Science and Systems*, 2014.

[29] P. Dames and V. Kumar, "Experimental Characterization of a Bearing-only Sensor for Use With the PHD Filter," 2015, available at: arXiv:1502.04661 [cs:RO].

[30] B.-N. Vo, S. Singh, and A. Doucet, "Sequential monte carlo methods for multi-target filtering with random finite sets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, Oct. 2005.

[31] B.-N. Vo and W.-K. Ma, "The Gaussian Mixture Probability Hypothesis Density Filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

Philip Dames is a Ph.D. candidate in the Department of Mechanical Engineering and Applied Mechanics at the University of Pennsylvania. He received his B.S. and M.S. in Mechanical Engineering from Northwestern University, both in 2006. His research interests lie at the intersection of estimation, control, and communication in multi-agent systems.

Vijay Kumar is the UPS Foundation Professor in the School of Engineering and Applied Science at the University of Pennsylvania. He received his Bachelors of Technology from the Indian Institute of Technology, Kanpur and his Ph.D. from the Ohio State University in 1987. He has been on the Faculty in the Department of Mechanical Engineering and Applied Mechanics since 1987 with secondary appointments in the Departments of Computer and Information Science and of Electrical and Systems Engineering. He was the assistant director for robotics and cyber physical systems at the White House Office of Science and Technology Policy from 2012–2014.

Dr. Kumar's research interests are in robotics, specifically multi-robot systems, and micro aerial vehicles. He is a Fellow of the American Society of Mechanical Engineers (2003), a Fellow of the Institution of Electrical and Electronic Engineers (2005) and a member of the National Academy of Engineering (2013).