

# Decentralized Active Information Acquisition: Theory and Application to Multi-Robot SLAM

Nikolay Atanasov, Jerome Le Ny, Kostas Daniilidis, and George J. Pappas

**Abstract**—This paper addresses the problem of controlling mobile sensing systems to improve the accuracy and efficiency of gathering information autonomously. It applies to scenarios such as environmental monitoring, search and rescue, surveillance and reconnaissance, and simultaneous localization and mapping (SLAM). A *multi-sensor active information acquisition problem*, capturing the common characteristics of these scenarios, is formulated. The goal is to design sensor control policies which minimize the entropy of the estimation task, conditioned on the future measurements. First, we provide a *non-greedy centralized solution*, which is computationally fast, since it exploits linearized sensing models, and memory efficient, since it exploits sparsity in the environment model. Next, we decentralize the control task to obtain linear complexity in the number of sensors and provide suboptimality guarantees. Finally, our algorithms are applied to the multi-robot active SLAM problem to enable a decentralized nonmyopic solution that exploits sparsity in the planning process.

## I. INTRODUCTION

The remarkable advances in sensing and mobility of robotic systems allow us to address important information acquisition problems such as environmental monitoring [1], [2], search and rescue [3], source seeking [4], active perception [5], [6], and active SLAM [7]–[9]. In all these scenarios, robots are deployed to gather information about a physical phenomenon (target) of interest. For example, in environmental monitoring, they are used to build spatiotemporal models of pollution propagation [10]. While the estimation aspect of these problems has received significant attention, there are far less results for controlling the sensing platforms to acquire information *actively*. Near-optimal methods exist for sensor placement and scheduling [11]–[13] but the approaches for optimizing the trajectories of mobile sensors are often *greedy* (optimize only with respect to the next time step) and rarely provide performance guarantees. This paper formulates a *multi-sensor active information acquisition* problem in which both the sensors and the targets are modeled as dynamical systems and allowed to evolve over time. The goal is to design *nonmyopic* control policies for the sensors such that the target uncertainty, conditioned on the future measurements, is minimized.

N. Atanasov, G. Pappas, and K. Daniilidis are with GRASP Lab, University of Pennsylvania, Philadelphia, PA 19104, USA, {atanasov, pappasg}@seas.upenn.edu, kostas@cis.upenn.edu.

J. Le Ny is with the Department of Electrical Engineering, Polytechnique de Montreal, and GERAD, Montreal, QC H3T1J4, Canada, jerome.le-ny@polymtl.ca.

We gratefully acknowledge support by TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA and the following grants: NSF-IIA-1028009, ARL MAST-CTA W911NF-08-2-0004, ARL RCTA W911NF-10-2-0016, ONR-N000141310778, NSF-DGE-0966142, NSF-IIS-1317788, NSF-IIP-1439681, and NSF-IIS-1426840.

**Contributions.** In previous work [14], we developed an efficient solution, with performance guarantees, for the *single-sensor* active information acquisition problem. In this work:

- the algorithm is improved by using a square-root information filter to exploit sparsity in the target information matrix,
- a decentralized solution, based on coordinate descent, is proposed for the *multi-sensor* active information acquisition problem and guarantees are provided for its performance with respect to the optimal centralized solution,
- the multi-robot active SLAM problem is reduced to an instance of the active information acquisition problem and our algorithms are applied to it.

**Related work.** Approaches for mobile sensor trajectory optimization include [15]–[22]. Charrow et al. [20] choose trajectories based on a fast and accurate approximation of the mutual information (MI) between nonlinear Gaussian measurements and a particle representation of the target distribution. Dames et al. [19] and Meyer et al. [21] have the sensors follow the gradient of MI and conditional entropy, respectively. Lauri and Ritala [22] use Monte Carlo tree search to solve a finite-horizon stochastic control problem with MI as the reward. A common theme in these works is that the length of the planning horizon is sacrificed in favor of using the (typically) nonlinear sensor and target models during planning.

We take the opposite stance and give up some model accuracy (via linearization) in order to plan efficiently with long horizons. An additional benefit is that any sparsity in the target information matrix can be exploited during the planning process. Recent advances in large-scale finite and infinite dimensional estimation (e.g. Gaussian Processes [23], [24], graph-based SLAM [25], [26]) take advantage of such structure inherent in a graphical-model representation of the environment. Since the information planning process is based on the estimation layer, it is natural that it utilizes a sparse representation too. Interestingly, we also show that when the target dynamics are controlled (e.g. needed for active self-localization), MI might not be an adequate value function.

To alleviate the computational burden of multi-robot deployments, it is mandatory to decentralize the information-seeking control task. Hoffmann and Tomlin [17] compute MI only for pairs of sensors, thus decreasing the dimension of the required integration. Similarly, Dames et al. [19] achieve scalability by assuming that MI approximately decouples among robot cliques. Fully-distributed approaches based on belief consensus are proposed in [18], [21]. Decentralization based on coordinate descent has been used by Singh et al. [15] and a similar approach based on Gauss-Seidel Relaxation was proposed by Zhou and Roumeliotis [27]. Of these works, only Singh et al. [15] provide guarantees on the performance

of their approach, valid for static targets and discrete sensing locations, without revisiting. Here, we allow revisiting and extend the guarantees to dynamic targets.

To demonstrate the performance of our algorithms, we consider the multi-robot active SLAM problem. The advantage of our approach over existing work in active SLAM is that it is decentralized, nonmyopic, and memory-efficient. Trajectory optimization in SLAM is particularly challenging due to the coupling between the localization accuracy of the robots, the quality of the map, and the exploration of the environment. Leung et al. [28] propose a nonmyopic approach for active SLAM with a single robot. They use attractors for exploration and an exhaustive search for trajectory optimization. Sim et al. [8] improve the efficiency by discretizing the environment and planning only for trajectories without self-intersections. There are also information-theoretic approaches which rely on non-Gaussian models (e.g. occupancy grids) but typically resort to greedy planning [7], [29], [30]. Recently, attention has been devoted to active visual SLAM as well [31], [32]. When it comes to multi-robot active SLAM, there are very few approaches [9], [21], [33] and most use greedy control.

**Paper organization.** The multi-sensor active information acquisition problem is formulated in Sec. II. The optimal centralized solution, with exponential complexity in the planning horizon and the number of sensors, is reviewed in Sec. III. Also, our previous algorithm [14, Alg. 2], which reduces the complexity in the planning horizon, is reviewed and improved by exploiting sparsity in the target representation. The decentralization scheme, which reduces the complexity in the number of sensors, is introduced in Sec. IV. Finally, in Sec. V, the algorithms are applied to multi-robot active SLAM.

## II. PROBLEM FORMULATION

Consider a team of  $n$  sensors with states  $x_{i,t} \in \mathcal{X}_i \cong \mathbb{R}^{d_{x_i}}$ ,  $i = 1, \dots, n$  at time  $t$  and dynamics governed by the following *sensor motion model*:

$$x_{i,t+1} = f_i(x_{i,t}, u_{i,t}), \quad u_{i,t} \in \mathcal{U}_i, \quad (1)$$

where  $u_{i,t}$  is the control input and  $\mathcal{U}_i$  is a *finite* space of admissible controls. We use the notation  $x_t := [x_{1,t}^T \cdots x_{n,t}^T]^T$ ,  $u_t := [u_{1,t}^T \cdots u_{n,t}^T]^T$ ,  $f(x, u) := [f_1(x_1, u_1)^T \cdots f_n(x_n, u_n)^T]^T$ ,  $\mathcal{U} := \mathcal{U}_1 \times \dots \times \mathcal{U}_n$ , and  $\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_n$  with metric  $d_{\mathcal{X}}$ . The task of the sensors is to track the evolution of a phenomenon of interest, whose dynamics are governed by a linear Gaussian *target motion model*:

$$\begin{aligned} y_{t+1}^s &= y_t^s, \\ y_{t+1}^d &= A_t(x_t, u_t)y_t^d + w_t(x_t, u_t), \\ w_t(x_t, u_t) &\sim \mathcal{N}(0, W_t(x_t, u_t)), \end{aligned} \quad (2)$$

where  $y_t := [(y_t^s)^T (y_t^d)^T]^T \in \mathbb{R}^{d_y}$  is the target state at time  $t$ , composed of a dynamic part  $y_t^d$  and a static part  $y_t^s$ , and  $w_t$  is uncorrelated Gaussian noise.

*Remark:* The deterministic sensor transitions (1) simplify the planning process (Sec. III) but do not limit the applicability of the models (1), (2). As will be seen in the active-SLAM context (Sec. V), nondeterministic robot dynamics can be handled by our approach. The (deterministic) mean of the robot pose estimates will be captured by  $x$ , while the (nondeterministic)

error in those estimates will be captured by  $y^d$ . The landmark locations in the environment will be represented by  $y^s$ .

Further, suppose that the dynamic part of the target state decomposes as  $y_t^d := [(y_{0,t}^d)^T (y_{1,t}^d)^T \cdots (y_{n,t}^d)^T]^T$ , where  $y_{0,t}^d$  captures the evolution of an exogenous process, which the sensors need to track collaboratively, and  $y_{i,t}^d$  captures the evolution of an endogenous process, only of interest to sensor  $i$ . The decomposition of the dynamics is:

$$\begin{aligned} y_{0,t+1}^d &= A_{0,t}y_{0,t}^d + w_{0,t}, \quad w_{0,t} \sim \mathcal{N}(0, W_{0,t}), \\ y_{i,t+1}^d &= A_{i,t}(x_{i,t}, u_{i,t})y_{i,t}^d + w_{i,t}, \\ w_{i,t} &\sim \mathcal{N}(0, W_{i,t}(x_{i,t}, u_{i,t})). \end{aligned} \quad (3)$$

The target states that are of interest to sensor  $i$  are  $y_{i,t} := [(y_t^s)^T (y_{0,t}^d)^T (y_{i,t}^d)^T]^T$ . The operation of each sensor is governed by the following *sensor observation model*:

$$z_{i,t} = H_{i,t}(x_{i,t})y_{i,t} + v_{i,t}(x_{i,t}), \quad v_{i,t}(x_{i,t}) \sim \mathcal{N}(0, V_{i,t}(x_{i,t})) \quad (4)$$

where  $z_{i,t} \in \mathbb{R}^{d_{z_i}}$  is the measurement signal and  $v_{i,t}(x_{i,t})$  is a sensor-state-dependent Gaussian noise, whose values are independent at any pair of times and across the sensors. The measurement noise is independent of the target noise  $w_t$  as well. To simplify notation, let  $z_t := [z_{1,t}^T \cdots z_{n,t}^T]^T$ ,  $v_t(x_t) := [v_{1,t}(x_{1,t})^T \cdots v_{n,t}(x_{n,t})^T]^T$ , and define  $H_t(\cdot)$  appropriately so that  $z_t = H_t(x_t)y_t + v_t(x_t)$ . Note that the observation model is *linear in the target state* but might depend nonlinearly on the sensor state. The information available to the sensors at time  $t$  to compute the control  $u_t$  is:

$$\mathcal{I}_{i,0} := z_{i,0}, \quad \mathcal{I}_{i,t} := (z_{i,0:t}, u_{i,0:(t-1)}), \quad t > 0, \quad i = 1, \dots, n. \quad (5)$$

**Problem (Active Information Acquisition).** *Given initial sensor states  $x_0 \in \mathcal{X}$ , a prior distribution of the target states  $y_0$ , and a finite planning horizon  $T$ , choose a sequence of functions  $\mu_{i,t}(\mathcal{I}_t) \in \mathcal{U}_i$  for  $i = 1, \dots, n$  and  $t = 1, \dots, T-1$ , which optimizes:*

$$\begin{aligned} \min_{\mu_{1:n,0:(T-1)}} & \frac{1}{T} \sum_{t=1}^T \mathbf{h}(y_t | z_{1:t}) \\ \text{s.t.} & \quad u_t := \mu_{1:n,t}(\mathcal{I}_t), \quad \text{eq.(2)}, \quad t = 0, \dots, T-1, \\ & \quad x_{t+1} = f(x_t, u_t), \quad t = 0, \dots, T-1, \\ & \quad z_t = H_t(x_t)y_t + v_t(x_t), \quad t = 1, \dots, T. \end{aligned} \quad (6)$$

where  $\mathbf{h}(y_t | z_{1:t})$  is the differential entropy<sup>1</sup> of the target state  $y_t$  at time  $t$  conditioned on the measurement set  $z_{1:t}$ .

To emphasize the dependence of problem (6) on the number of sensors  $n$  and the planning horizon  $T$ , we use the following shorthand notation:

$$\min_{\mu_{1:n,0:(T-1)}} J_T^{(n)}(\mu_{1,0:(T-1)}, \dots, \mu_{n,0:(T-1)}). \quad (7)$$

## III. CENTRALIZED SOLUTION

Active Information Acquisition (6) is a stochastic optimal control problem and in general for such problems adaptive (closed-loop) control policies have a significant advantage over nonadaptive (open-loop) ones. However, we showed in [14]

<sup>1</sup>The differential entropy of a continuous random variable  $Y$  with probability density function  $p$  is defined as  $h(Y) := -\int p(y) \log p(y) dy$ .

that due to the linearity of the models (2), (4) in the target state, the Gaussian-noise assumptions, and the deterministic sensor transitions (1), the cost function in (6) becomes independent of any particular measurement realization and (6) reduces to a deterministic optimal control problem.

**Theorem 1** ([14]). *If the prior distribution of  $y_0$  is Gaussian with covariance  $\Sigma_0 \succeq 0$ , there exists an open-loop control sequence  $\sigma \in \mathcal{U}^T$ , which is optimal in (6). Moreover, (6) is equivalent to the deterministic optimal control problem:*

$$\begin{aligned} \min_{\sigma \in \mathcal{U}^T} J_T^{(n)}(\sigma) &:= \frac{1}{T} \sum_{t=1}^T \gamma_t(\Sigma_t, x_t) \\ \text{s.t. } x_{t+1} &= f(x_t, \sigma_t), \quad t=0, \dots, T-1, \\ \Sigma_{t+1} &= \rho_{t+1}^e(\rho_t^p(\Sigma_t, x_t, \sigma_t), x_{t+1}), \quad t=0, \dots, T-1, \end{aligned} \quad (8)$$

where  $\gamma_t(\Sigma_t, x_t) := \log \det(\Sigma_t)$  for all  $t$ ,  $\rho_t^e(\Sigma, x)$  is the Kalman filter measurement update with matrices  $H_t(x), V_t(x)$ , and  $\rho_t^p(\Sigma, x, u)$  is the prediction step with matrices  $\text{diag}(A_t(x, u), I)$  and  $\text{diag}(W_t(x, u), 0)$ .

Due to this separation principle (Thm. 1), an optimal control policy for problem (6) can be obtained via forward value iteration (FVI, Alg. 1). FVI constructs a search tree,  $\mathcal{T}_t$ , with nodes at stage  $t \in [0, T]$  corresponding to the reachable states  $(x_t, \Sigma_t, J_t)$ . At each such node, there are edges (one for each control in  $\mathcal{U}$ ) leading to nodes  $(x_{t+1}, \Sigma_{t+1}, J_{t+1})$ , obtained from the dynamics in (8). Unfortunately, FVI has exponential complexity,  $O(|\mathcal{U}_1 \times \dots \times \mathcal{U}_n|^T)$ , in both the time horizon  $T$  and the number of sensors  $n$ , since the number of nodes in  $\mathcal{T}_t$  equals the number of sensor trajectories of length  $t$ .

---

#### Algorithm 1 Forward Value Iteration

---

```

1:  $J_0 \leftarrow 0, \quad S_0 \leftarrow \{(x_0, \Sigma_0, J_0)\}, \quad S_t \leftarrow \emptyset$  for  $t = 1, \dots, T$ 
2: for  $t = 1 : T$  do
3:   for all  $(x, \Sigma, J) \in S_{t-1}$  do
4:     for all  $u \in \mathcal{U}$  do
5:        $x_t \leftarrow f(x, u), \quad \Sigma_t \leftarrow \rho_{t+1}^e(\rho_t^p(\Sigma, x, u), x_t)$ 
6:        $J_t \leftarrow \frac{t-1}{t} J + \frac{1}{t} \gamma_t(\Sigma_t, x_t)$ 
7:        $S_t \leftarrow S_t \cup \{(x_t, \Sigma_t, J_t)\}$ 
8: return  $\min \{J \mid (x, \Sigma, J) \in S_T\}$ 

```

---

#### A. Reduced Value Iteration

In [14] we showed how to significantly reduce the complexity in  $T$  of Alg. 1 while providing guarantees on the performance under the following assumption<sup>2</sup>.

**Assumption.** *The cost  $\gamma_t(\Sigma, x)$  incurred at time  $t$  is concave and monotone in  $\Sigma$ .*

In short, we defined notions of trajectory crossing (Def. 1) and domination in informativeness (Def. 2) and proved that if several sensor trajectories cross at time  $t$  and one is dominated in informativeness by the rest, then its corresponding dominated node in the search tree  $\mathcal{T}_t$  can be discarded with a minimal loss in performance (see [14, Thm. 2, 3, 4]).

**Definition 1** (Trajectory  $\delta$ -crossing). *Trajectories  $\pi^1, \pi^2 \in \mathcal{X}^T$   $\delta$ -cross at time  $t \in [1, T]$  if  $d_{\mathcal{X}}(\pi_t^1, \pi_t^2) \leq \delta$  for  $\delta \geq 0$ .*

<sup>2</sup>The assumption is satisfied for conditional differential entropy and the performance guarantees in [14] apply here.

**Definition 2** ( $\epsilon$ -Algebraic redundancy [11]). *Let  $\epsilon \geq 0$  and let  $\{\Sigma_i\}_{i=1}^K$  be a finite set of symmetric positive semidefinite matrices. A matrix  $\Sigma \succeq 0$  is  $\epsilon$ -algebraically redundant with respect to  $\{\Sigma_i\}$ , if there exist nonnegative constants  $\{\alpha_i\}_{i=1}^K$  such that:  $\sum_{i=1}^K \alpha_i = 1$  and  $\Sigma + \epsilon I_n \succeq \sum_{i=1}^K \alpha_i \Sigma_i$ .*

This enabled a reduced value iteration (RVI, Alg. 2) algorithm for solving (6). RVI is a powerful technique because it provides a trade-off between complexity and optimality which can be controlled via the parameters  $(\epsilon, \delta)$ . In particular, if both parameters are zero, we obtain the optimal solution with exponential complexity in  $T$ , while if both are infinite, we obtain a greedy solution with linear complexity in  $T$ , which keeps only the best node at level  $t$  of the tree  $\mathcal{T}_t$ .

---

#### Algorithm 2 $(\epsilon, \delta)$ -Reduced Value Iteration

---

```

1:  $J_0 \leftarrow 0, \quad S_0 \leftarrow \{(x_0, \Sigma_0, J_0)\}, \quad S_t \leftarrow \emptyset$  for  $t = 1, \dots, T$ 
2: for  $t = 1 : T$  do
3:   for all  $(x, \Sigma, J) \in S_{t-1}$  do
4:     for all  $u \in \mathcal{U}$  do
5:        $x_t \leftarrow f(x, u), \quad \Sigma_t \leftarrow \rho_{t+1}^e(\rho_t^p(\Sigma, x, u), x_t)$ 
6:        $J_t \leftarrow \frac{t-1}{t} J + \frac{1}{t} \gamma_t(\Sigma_t, x_t)$ 
7:        $S_t \leftarrow S_t \cup \{(x_t, \Sigma_t, J_t)\}$ 
8:   Sort  $S_t$  in ascending order according to  $\max\{J_t, \gamma_T(\Sigma_t, x_t)\}$ 
9:    $S_t' \leftarrow S_t[1]$  % State with lowest cost
10:  for all  $(x, \Sigma, J) \in S_t \setminus S_t[1]$  do
11:    % Find all nodes in  $S_t'$ , which  $\delta$ -cross  $x$ :
12:     $Q \leftarrow \{\text{diag}(\Sigma', J') \mid (x', \Sigma', J') \in S_t', d_{\mathcal{X}}(x, x') \leq \delta\}$ 
13:    if  $\text{isempty}(Q)$  or not ( $\text{diag}(\Sigma, J)$  is  $\epsilon$ -alg. red. wrt  $Q$ ) then
14:       $S_t' \leftarrow S_t' \cup (x, \Sigma, J)$ 
15:   $S_t \leftarrow S_t'$ 
16: return  $\min \{J \mid (x, \Sigma, J) \in S_T\}$ 

```

---

The bottleneck of Alg. 2 is checking  $\epsilon$ -algebraic redundancy (Line 13), which requires solving a linear matrix inequality (LMI) feasibility problem in  $d_y$  dimensions. If a solution is needed quickly (e.g., in applications with nonlinear models, it is necessary to re-linearize the models and re-plan the trajectories after every few measurements), the algebraic redundancy can be verified only approximately, i.e., for a finite number of values for  $\{\alpha_i\}_{i=1}^K$ . In particular, choosing  $\alpha_j = 1$  and  $\alpha_i = 0, i \neq j$  for  $j = 1, \dots, K$ , requires verifying only  $K$  positive-semidefinite inequalities:

$$\Sigma + \epsilon I_{d_y} \succeq \Sigma_i, \quad i = 1, \dots, K, \quad (9)$$

instead of solving the LMI feasibility problem.

#### B. Exploiting Sparsity

A fact, which does not play a role in the problem formulation but is important for the solution, is that if the interactions among the target states are local (e.g. static landmarks), the information matrix,  $\Omega_0 := \Sigma_0^{-1}$ , is sparse, whereas the covariance matrix is not [16]. For example, the information structure in SLAM is naturally sparse [25], since it encodes constraints from measurements that connect only two variables (two robot poses or a robot pose and a landmark location). Since the planning process (8) uses prior information from the estimation layer, it is natural that it utilizes a sparse representation too. Yet, Alg. 2 does not exploit this.

The sparsity of  $\Omega$  can be leveraged by replacing the Kalman filter in Alg. 2 with a square-root information filter (SRIF) [34]. Since most existing formulations require invertible noise covariances, we derive SRIF equations that specifically handle

the case that part of the target state is static and the associated covariance is zero. Let the Cholesky factor (informally, square root) of the information matrix be the upper triangular matrix  $C$ , such that  $\Omega = C^T C$ , and decompose it into blocks:

$$C_{t|t} = \begin{bmatrix} D_{t|t} & F_{t|t} \\ 0 & G_{t|t} \end{bmatrix}, \quad (10)$$

where  $D_{t|t}$  corresponds to the dynamic target states ( $y_t^d$ ) and  $G_{t|t}$  - to the static target states ( $y_t^s$ ).

**Proposition 1.** *The square-root information filter equations, with prior square-root matrix  $C_{t|t}$  in (10), process model in (2), measurement model in (4), and known  $x_t, x_{t+1}, u_t$ , are:*

$$\begin{aligned} \text{Predict:} \quad & \begin{bmatrix} W_t^{-1/2} & 0 & 0 \\ D_{t|t} A_t^{-1} & D_{t|t} A_t^{-1} & F_{t|t} \\ 0 & 0 & G_{t|t} \end{bmatrix} = Q \begin{bmatrix} * & * \\ 0 & C_{t+1|t} \end{bmatrix} \\ \text{Update:} \quad & \begin{bmatrix} C_{t+1|t} \\ V_{t+1}^{-1/2} H_{t+1} \end{bmatrix} = \bar{Q} \begin{bmatrix} C_{t+1|t+1} \\ 0 \end{bmatrix} \\ \text{Combined:} \quad & \begin{bmatrix} W_t^{-1/2} & 0 & 0 \\ D_{t|t} A_t^{-1} & D_{t|t} A_t^{-1} & F_{t|t} \\ 0 & 0 & G_{t|t} \\ 0 & V_{t+1}^{-1/2} H_{t+1} \end{bmatrix} = \hat{Q} \begin{bmatrix} * & * \\ 0 & C_{t+1|t+1} \\ 0 & 0 \end{bmatrix} \end{aligned}$$

where  $Q, \bar{Q}, \hat{Q}$  are orthogonal matrices such that the other matrices on the right-hand sides are upper triangular. The dependence of  $A_t, W_t$  on  $x_t, u_t$  and of  $H_{t+1}, V_{t+1}$  on  $x_{t+1}$  is not explicit to simplify notation.

Proposition 1 states that the update and prediction steps can be computed via QR decomposition. To use the SRIF and avoid computing the covariance matrices in Alg. 2, we express the cost  $\gamma_t$  and condition (9) in terms of  $C$ . The cost becomes:

$$\gamma_t(\Sigma, x) = \log \det(\Sigma) = -\log(\det(C)^2) = -2 \sum_{i=1}^{d_y} \log |C_{ii}|,$$

while (9) can be modified via a Schur complement:

$$\begin{aligned} \Omega_i - (\Sigma + \epsilon I_{d_y})^{-1} &\succeq 0, & i = 1, \dots, K, \\ C_i^T C_i - C^T (I_{d_y} + \epsilon C C^T)^{-1} C &\succeq 0, & i = 1, \dots, K, \\ \begin{bmatrix} I_{d_y} + \epsilon C C^T & C \\ C^T & C_i^T C_i \end{bmatrix} &\succeq 0, & i = 1, \dots, K. \end{aligned}$$

#### IV. DECENTRALIZED SOLUTION

##### A. Coordinate Descent

Reduced value iteration decreases the exponential complexity of forward value iteration in the time horizon  $T$  but the size  $|\mathcal{U}|$  of the control set still scales exponentially with the number of sensors  $n$ . Instead of solving the active information acquisition problem (7) jointly over all sensors, consider the following coordinate descent approach. Suppose that sensor 1 plans its own trajectory, individually, without taking the other sensors into account. In other words, it solves a single-sensor active information acquisition problem:

$$\mu_{1,0:(T-1)}^c \in \arg \min_{\hat{\mu} \in \mathcal{U}_1^T} J_T^{(1)}(\hat{\mu}).$$

Then, it passes its chosen control sequence (and if necessary the models (1), (3), (4) evaluated along it) on to sensor 2. Sensor 2 takes this information into account and plans its

own trajectory by solving a two-sensor active information acquisition problem but with a fixed policy for sensor 1:

$$\mu_{2,0:(T-1)}^c \in \arg \min_{\hat{\mu} \in \mathcal{U}_2^T} J_T^{(2)}(\mu_{1,0:(T-1)}^c, \hat{\mu}).$$

Then, sensor 2 passes its information along to sensor 3, which solves a 3-sensor problem with fixed policies for the first 2 sensors. In general, sensor  $i$  needs the control sequences (and the models (1), (3), (4) evaluated along them), chosen by sensors 1 to  $i-1$ , and solves an  $i$ -sensor active information acquisition problem with fixed policies for the first  $i-1$  sensors

$$\mu_{i,0:(T-1)}^c \in \arg \min_{\hat{\mu} \in \mathcal{U}_i^T} J_T^{(i)}(\mu_{1:(i-1),0:(T-1)}^c, \hat{\mu}). \quad (11)$$

Since the sensor labels are arbitrary, the coordinate descent can be carried out in any order. Because subproblem  $i$  in (11) has a search space of size  $|\mathcal{U}_i|^T$ , the coordinate descent scheme reduces the complexity in  $n$  from exponential to linear, i.e. from  $O(|\mathcal{U}_1 \times \dots \times \mathcal{U}_n|^T)$  to  $O(\sum_{i=1}^n |\mathcal{U}_i|^T)$ . Note that the approach is not fully-distributed because multihop communication is needed to pass the information but all-to-all communication is not required either because the optimization naturally occurs in communication hops.

##### B. Performance Guarantees

To gain intuition about the performance of the coordinate descent scheme with respect to the optimal centralized policy, we temporarily change the cost function in the active information acquisition problem (6) from  $\frac{1}{T} \sum_{t=1}^T \mathbf{h}(y_t | z_{1:t})$  to the more-commonly-used mutual information,  $\mathbb{I}(y_{1:T}; z_{1:T})$ .

**Theorem 2.** *Consider problem (6) with negative mutual information,  $-\mathbb{I}(y_{1:T}; z_{1:T})$ , as cost function. Let  $\mu^*$  be the optimal policy with associated cost  $J_T^*$ . Let  $\mu^c$  be the coordinate descent policy with arbitrary sensor order and cost  $J_T^c$ . Suppose that the subproblems (11) are solved optimally. Then, the following guarantee holds for the performance of  $\mu^c$ :*

$$J_T^c \geq J_T^* \geq 2J_T^c. \quad (12)$$

Since the cost is negative, this means that  $\mu^c$  obtains at least half of the optimal value  $-J^*$ .

*Remark:* The proof relies on the fact that mutual information, when viewed as a function of the sensor set, is submodular. This result is different from [15, Thm.1] because it applies to controlled dynamic Gaussian-Markov random fields and allows revisiting sensor states.

Next, we provide performance guarantees with respect to the original cost function in (6) - conditional differential entropy.

**Corollary 1.** *Let  $\mu^*$  be the optimal policy in (6) with cost  $J_T^*$ . Let  $\mu^c$  be a coordinate descent policy with arbitrary sensor order and cost  $J_T^c$ . Suppose that the subproblems (11) are solved optimally. Then:*

$$J_T^c \geq J_T^* \geq 2J_T^c - \frac{1}{T} \sum_{t=1}^T \mathbf{h}(y_t) - \Delta_T,$$

where  $\Delta_T := \sum_{i=1}^n \frac{1}{T} \left( \sum_{t=1}^T \mathbf{h}(y_{1:i,t}) - \mathbf{h}(y_{1:i,1:T}) \right) \geq 0$ . If the dynamic target states are controlled (i.e.,  $A_{i,t}, W_{i,t}$  depend on  $x_{i,t}, u_{i,t}$ ), then  $\Delta_T$  depends on  $\mu^*$ .

Although the coordinate descent has a theoretically appealing guarantee with respect to mutual information, Table I below shows that, when the target transitions can be controlled (as in the active SLAM application in Sec. V), mutual information might be inadequate as a value function. The conditional differential entropy criterion in (6) is more appropriate. Note that the two are equivalent if  $y^d$  is not controlled.

TABLE I: Suppose that the sensor state is  $x_0$ , the target covariance is  $\Sigma_0 = 1$ , and there are two available controls  $u^{(1)}, u^{(2)}$ . The table shows an example in which the target state is a lot more uncertain after  $u^{(1)}$  than after  $u^{(2)}$  but nevertheless  $u^{(1)}$  is considered more informative by the mutual information value function. In the SLAM context, mutual information might prefer very uncertain controls (e.g. high velocity) even if they provide the same measurement information (captured by  $M(x)$  below) as more certain ones.

	$u^{(1)}$	$u^{(2)}$
$A(x_0, u)$	1	1
$W(x_0, u)$	99999	1
$M(x_1) := H(x_1)^T V(x_1)^{-1} H(x_1)$	0.01	0.01
$\Sigma_1^p := \rho_1^p(\Sigma_0, x_0, u) = A \Sigma_0 A^T + W$	100000	2
$\Sigma_2^s := \rho_2^s(\Sigma_1^p, x_1) = ((\Sigma_1^p)^{-1} + M(x_1))^{-1}$	99.9	1.96
$\mathbf{h}(y_1   z_1) := \frac{1}{2} (\log(2\pi e) + \log \det(\Sigma_2^s))$	3.72	<b>1.76</b>
$\mathbb{I}(y_1; z_1) := \mathbf{h}(y_1) - \mathbf{h}(y_1   z_1)$	<b>3.45</b>	0.01

## V. APPLICATION TO MULTI-ROBOT ACTIVE SLAM

In this section, we show that the active SLAM problem with  $n$  robots can be reduced to an instance of the information acquisition problem (6). Consequently, the algorithms we developed so far enable a decentralized nonmyopic solution that exploits sparsity. Since the focus of the paper is on the selection of informative controls, rather than the estimation aspect of the problem, we use an existing graph SLAM approach along the lines of [25]. In our simulations, the estimation is centralized but the control is decentralized. Decentralized estimation can be achieved as well, via [35], [36] for example.

Let the robot states at time  $t$  be  $r_{i,t}$  for  $i = 1, \dots, n$  and let the dynamics of the  $i$ -th robot be:

$$r_{i,t+1} = f_i(r_{i,t}, u_{i,t}, \eta_{i,t}), \quad u_{i,t} \in \mathcal{U}_i, \quad (13)$$

where  $u_{i,t}$  is the control input,  $\mathcal{U}_i$  is a finite space of admissible controls (e.g. motion primitives), and  $\eta_{i,t}$  is zero-mean Gaussian noise with covariance  $E_i$ . The robots evolve in an environment with  $M$  landmarks with positions  $m := [m_1^T, \dots, m_M^T]^T$ . The task is to explore the environment, autonomously and efficiently, and create an accurate map of the landmark locations while staying well-localized. Each robot, depending on its pose  $r_{i,t}$  can obtain measurements  $z_{i,t}$  (typically nonlinear, e.g., range and bearing) of the visible landmarks according to the sensing model:

$$z_{i,t} = h_i(r_{i,t}, m) + v_{i,t}, \quad v_{i,t} \sim \mathcal{N}(0, V_i) \quad (14)$$

where  $v_{i,t}$  is a Gaussian measurement noise, whose values are independent at any pair of times and across robots. We assume that the SLAM estimation layer provides a Gaussian prior on the robot poses  $r_0$  and on the locations  $m_0$  of the landmarks that have been discovered so far:

$$\begin{bmatrix} r_0 \\ m_0 \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \bar{r}_0 \\ \bar{m}_0 \end{bmatrix}, \begin{bmatrix} \Sigma_0^{rr} & \Sigma_0^{rm} \\ \Sigma_0^{mr} & \Sigma_0^{mm} \end{bmatrix} \right).$$

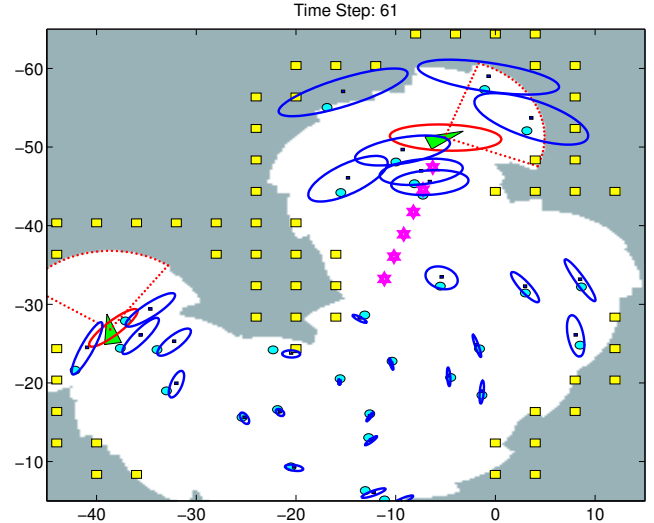


Fig. 2: A snapshot of the multi-robot active SLAM simulation. The cyan circles show the true landmark positions, the blue squares and ellipses show the estimated landmark positions and covariances, the green triangles show the true robot poses, the red ellipses show the estimated robot position covariances, the dotted red sectors indicate the robots' fields of view, the yellow squares show the exploration landmarks, and the magenta stars show localization attractors associated with the top robot. The gray area represents unexplored space.

**Problem (Multi-Robot Active SLAM).** Given a planning horizon  $T$ , choose a sequence of functions  $\mu_t(\mathcal{I}_t) \in \mathcal{U}$  for  $t = 1, \dots, T-1$ , which optimizes:

$$\begin{aligned} \min_{\mu_{0:(T-1)}} \quad & \frac{1}{T} \sum_{t=1}^T \mathbf{h}(r_t, m_t | z_{1:t}) \\ \text{s.t.} \quad & r_{t+1} = f(r_t, \mu_t(\mathcal{I}_t), \eta_t), \quad t = 0, \dots, T-1, \\ & m_{t+1} = m_t, \quad t = 0, \dots, T-1, \\ & z_t = h(r_t, m_t) + v_t, \quad t = 1, \dots, T, \end{aligned} \quad (15)$$

where  $\mathcal{I}_t$  is defined as in (5).

Because the robots neither know the total number  $M$  of landmarks, nor have prior information about the locations of the undiscovered landmarks, it is necessary to encourage them to explore the environment. We introduce dummy “exploration” landmarks with locations  $l := [l_1^T, \dots, l_{N_i}^T]^T$  at the current map frontiers [37] (see Fig. 2) and specify a Gaussian prior on their locations with mean  $\bar{l} := l$  and block diagonal covariance  $\Sigma^l$  with  $N_l$  blocks. This fake uncertainty in the exploration-landmark locations promises information gain to the robots. If it happens that there are no exploration landmarks within the reachable fields of view of the robots, the exploration process will stop because the algorithm we developed so far is unable to perceive information about the environment (even if known to the robots) beyond the planning horizon. Following Leung et al. [28], we include attractor landmarks (see Fig. 2), which incorporate global information about the environment in the local planning process. For each robot we use a state machine as discussed in [28, Sec.V] to decide the attractor state from  $\{None, Explore, Improve Map, Improve Localization\}$ , with *None* having the lowest priority and *Improve Localization* - the highest. For instance, if the entropy of robot  $i$ 's pose is larger than a threshold, we place localization attractors

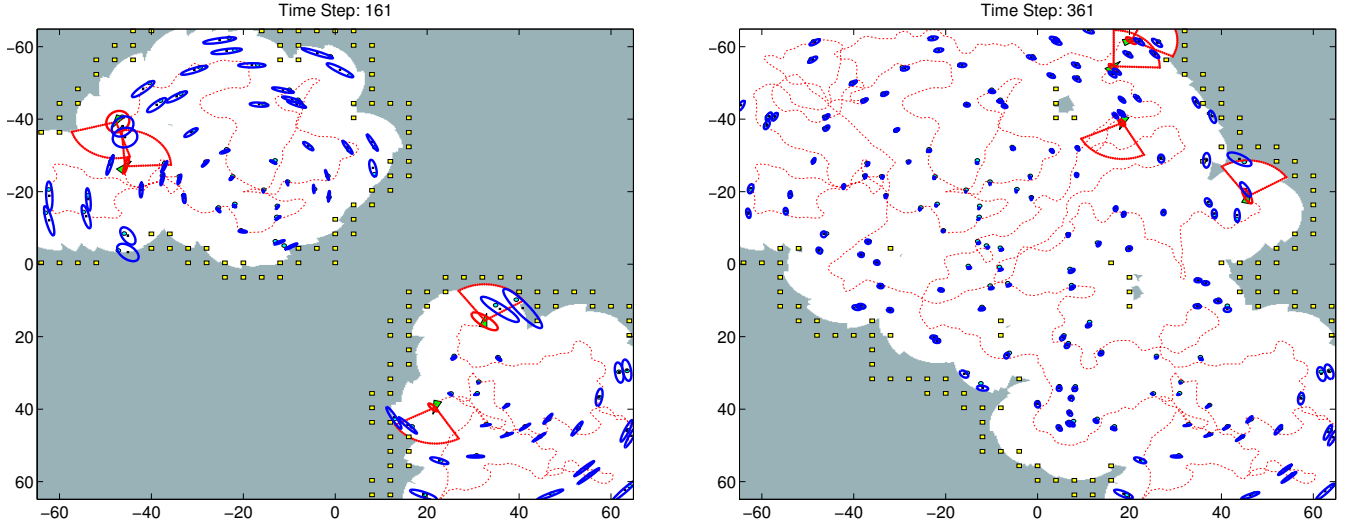


Fig. 1: Two instances of a four-robot active SLAM simulation which demonstrate the estimation quality and the exploration progress in an environment with 200 landmarks. The red dotted curves show the estimated robot trajectories, while the other symbols are described in the caption of Fig. 2. The robots had differential drive dynamics with maximum velocity  $3 \text{ m/s}$ , standard deviations  $0.1 \text{ m/s}$  and  $5^\circ/\text{s}$  in linear and angular velocities, respectively,  $10 \text{ m}$  sensing ranges,  $94^\circ$  fields of view, and used range and bearing measurements with standard deviations  $0.15 \text{ m}$  and  $5^\circ$ , respectively. See [http://www.seas.upenn.edu/~atanasov/vid/Atanasov\\_ActiveSLAM\\_ICRA15.mp4](http://www.seas.upenn.edu/~atanasov/vid/Atanasov_ActiveSLAM_ICRA15.mp4) for more details.

along the shortest path from the robot's estimated pose to the best localized landmark (see Fig. 2). We specify a Gaussian prior on the attractor locations  $a := [a_1^T \ \dots \ a_{N_a}^T]^T$  with mean  $\bar{a} := a$  and block diagonal covariance  $\Sigma^a$  with  $N_a$  blocks. To simplify notation, let  $q_0 := [m_0^T \ l^T \ a^T]^T$  be the combined locations of the discovered landmarks, the exploration landmarks, and the attractors with prior:

$$\begin{bmatrix} r_0 \\ q_0 \end{bmatrix} = \begin{bmatrix} r_0 \\ m_0 \\ l \\ a \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \bar{r}_0 \\ \bar{m}_0 \\ \bar{l} \\ \bar{a} \end{bmatrix}, \begin{bmatrix} \Sigma_0^{rr} & \Sigma_0^{rm} & 0 & 0 \\ \Sigma_0^{mr} & \Sigma_0^{mm} & 0 & 0 \\ 0 & 0 & \Sigma^l & 0 \\ 0 & 0 & 0 & \Sigma^a \end{bmatrix} \right)$$

Finally, to reduce the active SLAM problem (15) to an instance of the active information acquisition problem (6), we need to deal with the noisy robot dynamics (the sensor dynamics in (1) were deterministic) and the nonlinear measurement models. We take a model-predictive-control approach (Alg. 3), in which we linearize the motion and observation models, plan sensor trajectories, and apply the first set of control inputs. Due to the nonlinearities, after every few measurements, the models need to be re-linearized and the trajectories - re-planned. The models in (15) are linearized about an open-loop predicted trajectory of the (deterministic) mean  $(\bar{r}_t, \bar{q}_t)$  of the joint robot-landmark state  $(r_t, q_t)$ . Let  $\delta r_t := r_t - \bar{r}_t$ ,  $\delta q_t := q_t - \bar{q}_t$ , and  $\delta z_t := z_t - h(\bar{r}_t, \bar{q}_t)$  be the deviations from the mean. We obtain the following linearized version of (15):

$$\begin{aligned} \min_{\mu_{0:(T-1)}} & \frac{1}{T} \sum_{t=1}^T \mathbf{h}(r_t, q_t \mid z_{1:t}) \\ \text{s.t.} & \text{ for } t = 0, \dots, T-1 \\ & \bar{r}_{t+1} = f(\bar{r}_t, \mu_t(\mathcal{I}_t), 0), \quad \bar{q}_{t+1} = \bar{q}_t, \quad \delta q_{t+1} = \delta q_t \\ & \delta r_{t+1} \approx \left[ \frac{\partial f}{\partial r}(\bar{r}_t, \mu_t(\mathcal{I}_t), 0) \right] \delta r_t + \left[ \frac{\partial f}{\partial \eta}(\bar{r}_t, \mu_t(\mathcal{I}_t), 0) \right] \eta_t \\ & \delta z_{t+1} \approx \left[ \frac{\partial h}{\partial r}(\bar{r}_{t+1}, \bar{q}_{t+1}) \right] \delta r_{t+1} + \left[ \frac{\partial h}{\partial m}(\bar{r}_{t+1}, \bar{q}_{t+1}) \right] \delta q_{t+1} + v_t, \end{aligned} \quad (16)$$

### Algorithm 3 Multi-Robot Active SLAM

- 
- |         |   |  |
|---------|---|--|
| repeat: | { | <ol style="list-style-type: none"> <li>1) Receive measurements and update the SLAM estimate</li> <li>2) Remove any exploration landmarks within the robots' fields of view and add new ones at the map frontiers</li> <li>3) Remove the old attractors and add new ones if necessary</li> <li>4) Plan <math>T</math>-step trajectories by solving (16) via coordinate descent (11) and reduced value iteration (Alg. 2)</li> <li>5) Apply the first control inputs to move each robot</li> </ol> |
|---------|---|--|
- 

which is an instance of problem (6) with target state  $y_t := (\delta r_t^T \ \delta q_t^T)^T$ , sensor state  $x_t := (\bar{r}_t^T \ \bar{q}_t^T)^T$ , target system matrix  $A(x, u) := \frac{\partial f}{\partial r}(x, u, 0)$ , target noise covariance  $W(x, u) := \frac{\partial f}{\partial \eta}(x, u, 0) E \frac{\partial f}{\partial \eta}(x, u, 0)^T$ , sensor observation matrix  $H(x) := \begin{bmatrix} \frac{\partial h}{\partial r}(x) & \frac{\partial h}{\partial m}(x) \end{bmatrix}$  and sensor noise covariance  $V$ .

Our solution to the active SLAM problem is summarized in Alg. 3 and is illustrated in Fig. 1. Cooperation among the robots is achieved via the coordinate descent scheme, which allows the robots to take information from their teammates into account during the planning process. Details about the robot motion and measurement models are presented in the caption of Fig. 1. The performance is quantified in Fig. 3. Importantly, the entropy in the landmark positions is decreasing over time, although new landmarks are continuously being discovered. The robot pose entropies fluctuate because the robots need to repeatedly sacrifice localization accuracy in order to explore the environment. The plots indicate that the robots successfully explore the environment, create an accurate map of the landmark positions, and remain localized well in the process. The performance with a 12-step planning horizon is, as expected, better than greedy planning. However, the message here is not that our algorithm is better than a greedy approach, but that, if time for computation is available, it should be used to improve the planned trajectories. The  $(\epsilon, \delta)$  parameters of the reduced value iteration algorithm allow us to utilize the computation time effectively and the extra work is guaranteed to improve the performance, compared to the greedy approach baseline.

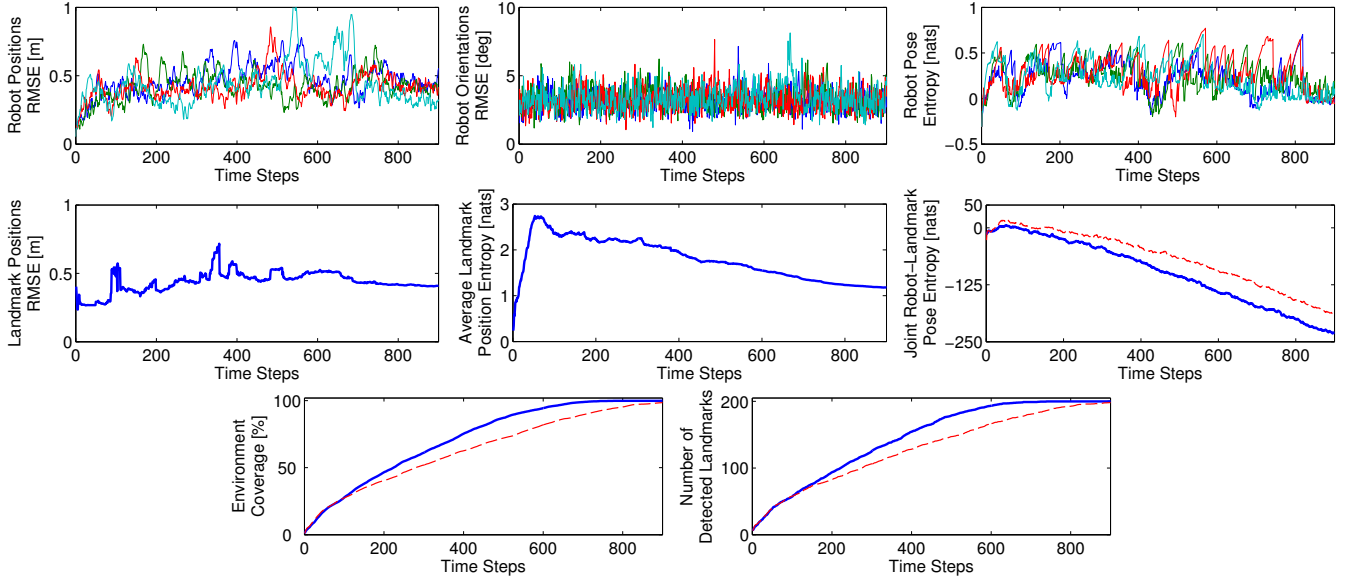


Fig. 3: The top three plots show the root mean square error (RMSE) in the robot position estimates, the RMSE in the robot orientation estimates, and the entropy of the robot pose estimates, average over 15 repetitions of the four-robot active SLAM scenario in Fig. 1. The middle row shows the RMSE in the landmark position estimates, the average entropy in the landmark position estimates, and the entropy of the joint robot-landmark-pose estimate. The last two plots show the percentage of the environment covered by the robots and the number of detected landmarks over time. The robot trajectories were planned using RVI (Alg. 2) with planning horizon  $T = 12$ ,  $\epsilon = \infty$ , and  $\delta = 1.5$ . The red dotted curves on the last three plots show the performance when the trajectories are obtained via a greedy policy (RVI with  $T = 1$ ).

## VI. CONCLUSION

The optimal centralized algorithm for active information acquisition has exponential complexity in both the planning horizon and the number of sensors. This paper discussed an approximation algorithm with suboptimality guarantees, which provides parameters to control the complexity in the time horizon. Further, it introduced a decentralization scheme to achieve linear complexity in the number of sensors. The combination of these techniques is an effective and scalable approach for controlled information acquisition with multiple sensors. For example, when applied to multi-robot active SLAM, it led to a decentralized nonmyopic solution, which demonstrated good performance and exploited sparsity in the planning process. Our formulation, however, cannot handle discrete target state spaces, which are needed particularly for hypothesis testing and classification problems. Future work will focus a discrete formulation of the active information acquisition problem. We also plan to compare our approach to methods that do not linearize the sensing models and carry out experiments with mobile targets and real robots.

### APPENDIX A: PROOF OF THEOREM 2

**Lemma 1.** Let  $\mathcal{F} \subseteq \{1, \dots, n\}$ ,  $y_{\mathcal{F}} := \{y_{i,1:T} \mid i \in \mathcal{F}\}$ , and  $z_{\mathcal{F}} := \{z_{i,1:T} \mid i \in \mathcal{F}\}$ . The mutual information  $g(\mathcal{F}) := \mathbb{I}(y_{\mathcal{F}}; z_{\mathcal{F}})$  between the target states  $y_{\mathcal{F}}$  and the measurement set  $z_{\mathcal{F}}$  viewed as a function of the set of sensors  $\mathcal{F}$  is submodular. In other words, given sets  $\mathcal{A} \subseteq \mathcal{B} \subseteq \{1, \dots, n\}$  and  $\mathcal{C} \subseteq \{1, \dots, n\} \setminus \mathcal{B}$ , the following is satisfied:  $g(\mathcal{A} \cup \mathcal{C}) - g(\mathcal{A}) \geq g(\mathcal{B} \cup \mathcal{C}) - g(\mathcal{B})$ .

*Proof:* The proof follows the steps of [12, Lemma 2.1] but to handle the varying dimension of  $y_{\mathcal{F}}$  we use that for  $\mathcal{A} \subseteq \mathcal{B}$ :

$$\begin{aligned} \mathbb{I}(y_{\mathcal{B}}; z_{\mathcal{A}}) &= \mathbf{h}(z_{\mathcal{A}}) - \mathbf{h}(z_{\mathcal{A}} \mid y_{\mathcal{B}}) \\ &= \mathbf{h}(z_{\mathcal{A}}) - \mathbf{h}(z_{\mathcal{A}} \mid y_{\mathcal{A}}) = \mathbb{I}(y_{\mathcal{A}}; z_{\mathcal{A}}) \end{aligned}$$

since when conditioned on  $y_{\mathcal{A}}$ , the measurements  $z_{\mathcal{A}}$  from known sensor states  $x_{\mathcal{A}} := \{x_{i,1:T} \mid i \in \mathcal{A}\}$  are independent of all other target states. ■

Let the measurement sets obtained by  $\mu^*$  and  $\mu^c$  be  $z_{1:T}^*$  and  $z_{1:T}^c$ , respectively. Then,

$$\begin{aligned} \mathbb{I}(y_{1:T}; z_{1:T}^*) &\stackrel{(a)}{\leq} \mathbb{I}(y_{1:T}; z_{1:T}^*) + \sum_{i=1}^n \left( \mathbb{I}(y_{1:T}; z_{1:i,1:T}^c, z_{i+1:n,1:T}^*) \right. \\ &\quad \left. - \mathbb{I}(y_{1:i-1,1:T}, y_{i+1:n,1:T}; z_{1:i-1,1:T}^c, z_{i+1:n,1:T}^*) \right) \\ &= \mathbb{I}(y_{1:T}; z_{1:T}^c) + \sum_{i=1}^n \left( \mathbb{I}(y_{1:T}; z_{1:i-1,1:T}^c, z_{i:n,1:T}^*) \right. \\ &\quad \left. - \mathbb{I}(y_{1:i-1,1:T}, y_{i+1:n,1:T}; z_{1:i-1,1:T}^c, z_{i+1:n,1:T}^*) \right) \\ &\stackrel{\text{Lemma 1}}{\leq} \mathbb{I}(y_{1:T}; z_{1:T}^c) + \sum_{i=1}^n \left( \mathbb{I}(y_{1:i,1:T}; z_{1:i-1,1:T}^c, z_{i,1:T}^*) \right. \\ &\quad \left. - \mathbb{I}(y_{1:i-1,1:T}; z_{1:i-1,1:T}^c) \right) \\ &= \mathbb{I}(y_{1:T}; z_{1:T}^c) + \mathbb{I}(y_{1:T}; z_{1:n-1,1:T}^c, z_{n,1:T}^*) \\ &\quad + \sum_{i=1}^{n-1} \left( \mathbb{I}(y_{1:i,1:T}; z_{1:i-1,1:T}^c, z_{i,1:T}^*) - \mathbb{I}(y_{1:i,1:T}; z_{1:i,1:T}^c) \right) \end{aligned} \quad (17)$$

where (a) follows from the non-decreasing property of mutual information. By definition of the coordinate descent policy

$$-\mathbb{I}(y_{1:i,1:T}; z_{1:i-1,1:T}^c, z_{i,1:T}^*) \geq -\mathbb{I}(y_{1:i,1:T}; z_{1:i,1:T}^c), \quad \forall i$$

and from (17),  $J_T^* = -\mathbb{I}(y_{1:T}; z_{1:T}^*) \geq J_T^c + J_T^c + 0$ . ■

APPENDIX B: PROOF OF COROLLARY 1

**Lemma 2.** Given (1), (2), (4), and a deterministic sequence  $x_{0:T}$ , the following relationship between mutual information and conditional entropy holds:

$$\begin{aligned} \mathbf{h}(y_{1:T}) - \sum_{t=1}^T \mathbf{h}(y_t \mid z_{1:t}) &\leq \mathbb{I}(y_{1:T}; z_{1:T}) \\ &\leq \sum_{t=1}^T h(y_t) - \sum_{t=1}^T \mathbf{h}(y_t \mid z_{1:t}) \end{aligned}$$

*Proof:* 
$$\begin{aligned} \mathbf{h}(y_{1:T}) - \sum_{t=1}^T \mathbf{h}(y_t \mid z_{1:t}) &\stackrel{(a)}{\leq} \mathbf{h}(y_{1:T}) - \sum_{t=1}^T \mathbf{h}(y_t \mid z_{1:t}, y_{t-1}) \\ &\stackrel{(b)}{=} \mathbf{h}(y_{1:T}) - \mathbf{h}(y_{1:T} \mid z_{1:T}) = \mathbb{I}(y_{1:T}; z_{1:T}) \\ &= \sum_{t=1}^t [\mathbf{h}(y_t \mid z_{1:t-1}) - h(y_t \mid z_{1:t})] \\ &\stackrel{(c)}{\leq} \sum_{t=1}^T \mathbf{h}(y_t) - \sum_{t=1}^T \mathbf{h}(y_t \mid z_{1:t}) \end{aligned}$$

where (a), (c) hold because conditioning decreases entropy, (b) is due to the entropy chain rule. ■

From Lemma 2 and (17), it follows that

$$\begin{aligned} \mathbf{h}(y_{1:T}) - TJ_T^* &\leq \sum_{t=1}^T \mathbf{h}(y_t) - TJ_T^c + \sum_{t=1}^T \mathbf{h}(y_t) - \sum_{t=1}^T \mathbf{h}(y_t \mid z_{1:n-1,1:t}^c, z_{n,1:t}^*) \\ &\quad + \sum_{i=1}^{n-1} \left( \left( \sum_{t=1}^T \mathbf{h}(y_{1:i,t}) \right) - \mathbf{h}(y_{1:i,1:T}) \right) \\ &\quad + \sum_{i=1}^{n-1} \left( \sum_{t=1}^T \mathbf{h}(y_{1:i,t} \mid z_{1:i,1:t}^c) - \sum_{t=1}^T \mathbf{h}(y_{1:i,t} \mid z_{1:i-1,1:t}^c, z_{i,1:t}^*) \right) \\ &\leq 2 \sum_{t=1}^T \mathbf{h}(y_t) - 2TJ_T^c + \left( \sum_{i=1}^{n-1} \left( \sum_{t=1}^T \mathbf{h}(y_{1:i,t}) \right) - \mathbf{h}(y_{1:i,1:T}) \right) \end{aligned}$$

REFERENCES

- [1] H. Choi, "Adaptive Sampling and Forecasting With Mobile Sensor Networks," Ph.D. dissertation, MIT, 2009.
- [2] P. Tokekar, E. Branson, J. Vander Hook, and V. Isler, "Tracking Aquatic Invaders: Autonomous Robots for Monitoring Invasive Fish," *IEEE Robotics and Automation Magazine*, vol. 20, no. 3, 2013.
- [3] V. Kumar, D. Rus, and S. Singh, "Robot and Sensor Networks for First Responders," *IEEE Pervasive Computing*, vol. 3, no. 4, 2004.
- [4] N. Atanasov, J. Le Ny, and G. Pappas, "Distributed Algorithms for Stochastic Source Seeking with Mobile Robot Networks," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 137, no. 3, 2015.
- [5] L. Valente, Y.-H. Tsai, and S. Soatto, "Information-Seeking Control Under Visibility-Based Uncertainty," *Journal of Mathematical Imaging and Vision*, vol. 48, no. 2, pp. 339–358, 2014.
- [6] N. Atanasov, B. Sankaran, J. Le Ny, G. Pappas, and K. Daniilidis, "Nonmyopic View Planning for Active Object Classification and Pose Estimation," *IEEE Trans. on Robotics (TRO)*, vol. 30, no. 5, 2014.
- [7] L. Carlone, J. Du, M. Kaouk Ng, B. Bona, and M. Indri, "Active SLAM and Exploration with Particle Filters Using Kullback-Leibler Divergence," *Intelligent & Robotic Systems*, vol. 75, no. 2, 2014.
- [8] R. Sim and N. Roy, "Global A-Optimal Robot Exploration in SLAM," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2005.
- [9] M. Kontitsis, E. Theodorou, and E. Todorov, "Multi-Robot Active SLAM with Relative Entropy Optimization," in *American Control Conference (ACC)*, 2013.
- [10] R. Marchant and F. Ramos, "Bayesian Optimisation for Informative Continuous Path Planning," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2014.
- [11] M. Vitus, W. Zhang, A. Abate, J. Hu, and C. Tomlin, "On Efficient Sensor Scheduling for Linear Dynamical Systems," *Automatica*, vol. 48, no. 10, 2012.

- [12] J. Williams, "Information Theoretic Sensor Management," Ph.D. dissertation, MIT, 2007.
- [13] A. Krause, "Optimizing sensing: Theory and applications," Ph.D. dissertation, Carnegie Mellon University, 2008.
- [14] N. Atanasov, J. Le Ny, K. Daniilidis, and G. Pappas, "Information Acquisition with Sensing Robots: Algorithms and Error Bounds," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2014.
- [15] A. Singh, A. Krause, C. Guestrin, and W. Kaiser, "Efficient Informative Sensing Using Multiple Robots," *Journal of Artificial Intelligence Research (JAIR)*, vol. 34, no. 2, 2009.
- [16] J. Le Ny and G. Pappas, "On Trajectory Optimization for Active Sensing in Gaussian Process Models," in *Proc. IEEE Conf. on Decision and Control (CDC)*, 2009.
- [17] G. Hoffmann and C. Tomlin, "Mobile Sensor Network Control Using Mutual Information Methods and Particle Filters," *IEEE Trans. on Automatic Control (TAC)*, vol. 55, no. 1, 2010.
- [18] B. Julian, M. Angermann, M. Schwager, and D. Rus, "Distributed robotic sensor networks: An information-theoretic approach," *International Journal of Robotics Research (IJRR)*, vol. 31, no. 10, 2012.
- [19] P. Dames, M. Schwager, V. Kumar, and D. Rus, "A Decentralized Control Policy for Adaptive Information Gathering in Hazardous Environments," in *IEEE Conf. on Decision and Control (CDC)*, 2012.
- [20] B. Chawar, V. Kumar, and N. Michael, "Approximate Representations for Multi-Robot Control Policies that Maximize Mutual Information," in *Robotics: Science and Systems (RSS)*, 2013.
- [21] F. Meyer, H. Wymeersch, M. Fröhle, and F. Hlawatsch, "Distributed Estimation with Information-Seeking Control in Agent Networks," *arXiv preprint:1408.3732*, 2014.
- [22] M. Lauri and R. Ritala, "Stochastic Control for Maximizing Mutual Information in Active Sensing," in *IEEE Int. Conf. on Robotics and Automation (ICRA) Workshop on Robots in Homes and Industry*, 2014.
- [23] J. Quiñero-Candela and C. Rasmussen, "A Unifying View of Sparse Approximate Gaussian Process Regression," *Journal of Machine Learning Research*, vol. 6, 2005.
- [24] T. Barfoot, C. Tong, and S. Sarkka, "Batch Continuous-Time Trajectory Estimation as Exactly Sparse Gaussian Process Regression," in *Robotics: Science and Systems (RSS)*, 2014.
- [25] F. Dellaert and M. Kaess, "Square Root SAM: Simultaneous Localization and Mapping via Square Root Information Smoothing," *International Journal of Robotics Research (IJRR)*, vol. 25, no. 12, 2006.
- [26] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, "iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree," *International Journal of Robotics Research (IJRR)*, vol. 31, no. 2, 2012.
- [27] K. Zhou and S. Roumeliotis, "Multirobot Active Target Tracking With Combinations of Relative Observations," *IEEE Trans. on Robotics (TRO)*, vol. 27, no. 4, 2011.
- [28] C. Leung, S. Huang, and G. Dissanayake, "Active SLAM using Model Predictive Control and Attractor based Exploration," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2006.
- [29] F. Bourgault, A. Makarenko, S. Williams, B. Grocholsky, and H. Durrant-Whyte, "Information Based Adaptive Robotic Exploration," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2002.
- [30] B. Tovar, L. Muñoz-Gómez, R. Murrieta-Cid, M. Alencastre-Miranda, R. Monroy, and S. Hutchinson, "Planning Exploration Strategies for Simultaneous Localization and Mapping," *Robotics and Autonomous Systems*, vol. 54, no. 4, 2006.
- [31] T. Vidal-Calleja, A. Davison, J. Andrade-Cetto, and D. Murray, "Active Control for Single Camera SLAM," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2006.
- [32] C. Forster, M. Pizzoli, and D. Scaramuzza, "Appearance-based Active, Monocular, Dense Reconstruction for Micro Aerial Vehicles," in *Robotics: Science and Systems (RSS)*, 2014.
- [33] V.-C. Pham and J.-C. Juang, "A Multi-Robot, Cooperative, and Active SLAM Algorithm for Exploration," *International Journal of Innovative Computing, Information and Control*, vol. 9, no. 6, 2013.
- [34] G. Bierman, "Sequential Square Root Filtering and Smoothing of Discrete Linear Systems," *Automatica*, vol. 10, no. 2, 1974.
- [35] A. Cunningham, V. Indelman, and F. Dellaert, "DDF-SAM 2.0: Consistent Distributed Smoothing and Mapping," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2013.
- [36] S. Shahrampour, A. Rakhlin, and A. Jadbabaie, "Distributed Detection: Finite-time Analysis and Impact of Network Topology," *arXiv preprint:1409.8606*, 2014.
- [37] B. Yamauchi, "A Frontier-Based Approach for Autonomous Exploration," in *IEEE Int. Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, 1997.