# Information Acquisition with Sensing Robots: Algorithms and Error Bounds

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

*Abstract*—Utilizing the capabilities of configurable sensing systems requires addressing difficult information gathering problems. Near-optimal approaches exist for sensing systems without internal states. However, when it comes to optimizing the trajectories of mobile sensors the solutions are often greedy and rarely provide performance guarantees. Notably, under linear Gaussian assumptions, the problem becomes deterministic and can be solved *off-line*. Approaches based on submodularity have been applied by ignoring the sensor dynamics and *greedily* selecting informative locations in the environment. This paper presents a *non-greedy* algorithm with *suboptimality guarantees*, which relies on concavity instead of submodularity and takes the sensor dynamics into account. Coupled with linearization and model predictive control, the algorithm can be used to generate adaptive policies for mobile sensors with non-linear sensing models. Applications in gas concentration mapping and target tracking are presented.

## I. INTRODUCTION

Miniaturization, wireless communication, and sensor technology have advanced remarkably in recent years. Autonomous vehicles instrumented with various sensors and interconnected by a network are deployed in large numbers, leading to configurable networked sensing systems. In order to utilize their capabilities, some important information gathering problems such as environmental monitoring [1], [2], surveillance and reconnaissance [3], [4], search and rescue [5], active perception and active SLAM [6], [7], [8] need to be addressed.

*Sensor management* [9] offers a formal methodology to control the degrees of freedom of sensing systems in order to improve the information acquisition process. Much research in the field has been focused on sensors without internal states by studying the problem of *sensor scheduling*. Efficient non-myopic approaches have been proposed [10], [11], [12]. However, when it comes to mobile sensors, which posses internal states, the results are limited. The main complication is that the evolution of the states depends on the management decisions and affects the measurements in the long run. Concretely, whereas in radar management a sensor can switch instantaneously between targets [13], a feasible and informative path needs to be designed for a sensing robot [14]. Due to this complication, most approaches for mobile sensor management

are myopic [15], [16] or use short planning horizons [17], [18]. However, it is precisely the presence of an internal state that makes multi-step optimization important. The behavior of the observed phenomenon needs to be predicted at an early stage to facilitate effective control of the mobile sensor.

A key insight is that under linear Gaussian assumptions the problem can be formulated as deterministic optimal control [19]. As a result, informative sensing paths can be planned *off-line*. Some of the successful approaches rely on a *submodular* function to quantify the informativeness of the sensor paths [14], [20]. The sensing locations are partitioned into independent clusters. Submodularity is used to greedily select informative locations within clusters. An orienteering problem is solved to choose the sequence of clusters to visit. The drawback is that within clusters the movement of the sensor is restricted to a graph, essentially ignoring the sensor dynamics. As a result, the cluster sizes cannot be increased much without affecting the quality of the solution. We address this limitation by considering the sensor dynamics and planning non-greedily.

The **contribution** of this paper is an approach for discrete-time dynamic sensors with linear Gaussian sensing models to track a target with linear dynamics. An approximate non-myopic algorithm is developed to decrease the complexity of obtaining an optimal policy while providing strong performance guarantees. The key idea is to discard any sensing paths which are close in space and dominated in informativeness by other paths in the planning space. Coupled with linearization and model predictive control, the generated policy can be used with non-linear sensing models and performs provably better than a greedy approach. Our work can be considered a search-based method for planning in information space. Related work in this area includes [21], [19]. Sampling-based methods have been proposed as well [22], [23]. Just as in traditional planning, they are able to find feasible solutions quickly but provide no finite-time guarantees on optimality. Approaches which do not make linear Gaussian assumptions and use non-linear filters exist as well [24], [25], [26]. They can handle general sensing and target models but are computationally demanding and have no performance guarantees.

The rest of the paper is organized as follows. The information acquisition problem is formulated precisely in Sec. II. A separation principle, which allows computing the optimal sensor path off-line, is proved in Sec. III. The approximate algorithm with suboptimality guarantees is developed in Sec. IV to reduce the exponential complexity. Applications in gas concentration mapping and target tracking are presented in Sec. V. All proofs are provided in the Appendix.

## II. Problem Formulation

Consider a sensor mounted on a robot or vehicle, whose dynamics are governed by the following *sensor motion model*:

$$x_{t+1} = f(x_t, u_t), \qquad (1)$$

where $x \in \mathcal{X} \cong \mathbb{R}^{n_x}$ is the sensor state, $\mathcal{X}$ is an $n_x$-dimensional state space with metric $d_{\mathcal{X}}$, $u \in \mathcal{U}$ is the control input, and $\mathcal{U}$ is a *finite* space of admissible controls. The task of the sensor is to track the evolution of the state of a target, whose dynamics are governed by a linear *target motion model*:

$$y_{t+1} = A y_t + w_t, \qquad w_t \sim \mathcal{N}(0, W), \qquad (2)$$

where $y \in \mathbb{R}^{n_y}$ is the target state, $A \in \mathbb{R}^{n_y \times n_y}$, and $w_t$ is a white Gaussian noise with covariance $\mathbb{E} w_t w_\tau^T = W \delta(t - \tau)$. The minimum eigenvalue of $W$ is denoted by $\underline{\lambda}_W$.

The operation of the sensor is governed by the following *sensor observation model*:

$$z_t = H(x_t) y_t + v_t(x_t), \qquad v_t(x_t) \sim \mathcal{N}(0, V(x_t)), \quad (3)$$

where $z_t \in \mathbb{R}^{n_z}$ is the measurement signal, $H(x_t) \in \mathbb{R}^{n_z \times n_y}$, and $v_t(x_t)$ is a sensor-state-dependent Gaussian noise, whose values at any pair of times are independent. The measurement noise is independent of the target noise $w_t$ as well. Note that the observation model is *linear in the target state* but might depend nonlinearly on the sensor state. The information available to the sensor at time $t$ to compute the control $u_t$ is:

$$\mathcal{I}_0 := z_0 \quad \mathcal{I}_t := (z_{0:t}, u_{0:(t-1)}) \in \left(\mathbb{R}^{n_z}\right)^{t+1} \times \mathcal{U}^t, \ t > 0.$$

**Problem** (Active Information Acquisition). *Given an initial sensor pose $x_0 \in \mathcal{X}$, a prior distribution of the target state $y_1$, and a planning horizon $T$, the task of the sensor is to choose a sequence of functions $\mu_0 : \mathbb{R}^{n_z} \to \mathcal{U}$, $\mu_t : \mathbb{R}^{(t+1)n_z} \times \mathcal{U}^t \to \mathcal{U}$ for $t = 1, \ldots, T - 1$, which maximize the mutual information between the final target state $y_{T+1}$ and the measurement set $z_{1:T}$:*

$$\max_{\mu_0, \ldots, \mu_{T-1}} \mathbb{I}(y_{T+1}; z_{1:T} \mid x_{1:T}) \qquad (4)$$

$$\begin{aligned}
\text{s.t.} \quad & x_{t+1} = f(x_t, \mu_t(\mathcal{I}_t)), && t = 0, \ldots, T-1, \\
& y_{t+1} = A y_t + w_t, && t = 1, \ldots, T, \\
& z_t = H(x_t) y_t + v_t(x_t), && t = 1, \ldots, T.
\end{aligned}$$

*Remark:* To simplify notation, we work with autonomous models but all results hold for time-varying $f_t$, $A_t$, $W_t$, $H_t$, and $V_t$ with $\underline{\lambda}_W > 0$ such that $\underline{\lambda}_W I_{n_y} \preceq W_t$ for all $t \in [1, T]$.

Problem (4) subsumes the static case where sensors do not have internal states. In particular, if $f(x, u) := u$ and we think of the control $u$ as choosing a subset of sensors to be activated, then (4) becomes a sensor scheduling problem. Similarly, if the space of sensor configurations, $\mathcal{X}$, is restricted to a graph and revisiting is not allowed, we get the problem in [14]. Numerous information gathering tasks can be cast in the form of (4). For example, environmental monitoring [19], active target tracking (Sec. V, [24]), and active mapping with gas (Sec. V), stereo [27], laser, or any other sensor, whose operation is captured reasonably by a linearized model.

To motivate the discussion consider the methane emission monitoring problem, which was addressed by the Best Service Robotics paper [2] at ICRA 2013. The task is to estimate the gas concentration in a landfill using a remote methane leak detector (RMLD) based on tunable diode laser absorption spectroscopy. The RMLD sensor is mounted on a robotic platform, Gasbot, which follows an exploration path *pre-specified by hand*. In this paper, we would like to automatically generate the most informative path for the Gasbot.

In detail, suppose that state of the Gasbot consists of its 2D position $(x^1, x^2) \in \mathbb{R}^2$ and the orientation of the RMLD sensor $\theta \in SO(2)$ so that $x := (x^1, x^2, \theta)^T$. At each time period, the Gasbot can move on a grid and choose the orientation of the sensor in $30°$ increments: $\Theta := \{-\pi, -5\pi/6, \ldots, 5\pi/6\}$,

$$\mathcal{U} := \left\{ (u^1, u^2, u^3)^T \middle| \binom{u^1}{u^2} \in \{0, \pm e_1, \pm e_2\}, u^3 \in \Theta \right\},$$

where $e_1$ and $e_2$ are the standard basis vectors in $\mathbb{R}^2$. With $u_t \in \mathcal{U}$, the sensor motion model is:

$$x_{t+1} = f(x_t, u_t) := (x_t^1 + u_t^1, \ x_t^2 + u_t^2, \ u_t^3)^T.$$

Given a time horizon $T$, we would like to choose a control policy for the Gasbot in order to obtain a good map $y_{T+1} \in \mathbb{R}^{n_y}$ of the gas concentration in the landfill. The $i$th component, $y_{T+1}^i$, represents the estimate of the concentration in parts per million (ppm) in the $i$th cell of the map. We assume a static methane field so that $A = I_{n_y}, W = 0$. It was experimentally verified in [2] that a good sensor observation model is:

$$z_t = H(x_t) y_t + v_t = \sum_{i=1}^{n_y} l_i(x_t) y_t^i + v_t, \quad v_t \sim \mathcal{N}(0, V),$$

where the $i$th component of $H(x_t) \in \mathbb{R}^{1 \times n_y}$ is the distance $l_i(x_t)$ traveled by sensor laser beam in cell $y_t^i$ for the given sensor pose $x_t$. Solving problem (4) will provide an automatic way to control the Gasbot in order to obtain an accurate map of the methane concentration.

*Notation*: The sets of $n \times n$ symmetric positive definite and semidefinite matrices are denoted $S_{++}^n$ and $S_+^n$ respectively. Given a control sequence $\sigma = u_t, \ldots, u_{T-1} \in \mathcal{U}^{T-t}$ at time $t$, the corresponding sensor trajectory is $\pi(\sigma) := x_{t+1}, \ldots, x_T \in \mathcal{X}^{T-t}$. Also, let $\sigma_i := u_{t+i}$, $\pi_{i+1} := x_{t+i+1}$, and $\pi(i+1) := x_{t+i+1}, \ldots, x_T \in \mathcal{X}^{T-t-i}$ for $i = 0, \ldots, (T - t - 1)$.

## III. A Separation Principle

Active Information Acquisition (4) is a stochastic optimal control problem and in general for such problems adaptive (closed-loop) control policies have a significant advantage over non-adaptive (open-loop) ones. However, due to the linearity of the observation model (3) in the target state and the Gaussian-noise assumptions, it can be shown that (4) reduces to a deterministic optimal control problem, for which open-loop policies are optimal.

**Theorem 1.** *If the prior distribution of $y_1$ is Gaussian with covariance $\Sigma_0 \in S_+^{n_y}$, there exists an open-loop control sequence $\sigma = u_0, \ldots, u_{T-1} \in \mathcal{U}^T$, which is optimal in (4).*

*Moreover, (4) is equivalent to the following deterministic optimal control problem:*

$$\min_{\sigma \in \mathcal{U}^T} \ \log \det(\Sigma_T) \tag{5}$$

$$\text{s.t.} \quad x_{t+1} = f(x_t, \sigma_t), \quad t = 0, \ldots, T-1,$$
$$\Sigma_{t+1} = \rho_{x_{t+1}}(\Sigma_t), \quad t = 0, \ldots, T-1,$$

*where $\rho_x(\cdot) := \rho^p(\rho_x^e(\cdot))$ is the Kalman filter Riccati map:*

**Update:** $\rho_x^e(\Sigma) := (\Sigma^{-1} + M(x))^{-1} = C_x(\Sigma)\Sigma$

$$M(x) := H(x)^T V(x)^{-1} H(x) \tag{6}$$
$$C_x(\Sigma) := I - K_x(\Sigma)H(x) = (I + \Sigma M_x)^{-1}$$
$$K_x(\Sigma) := \Sigma H(x)^T R_x^{-1}(\Sigma)$$
$$R_x(\Sigma) := H(x)\Sigma H(x)^T + V(x)$$

**Predict:** $\rho^p(\Sigma) := A\Sigma A^T + W$

*and $M(\cdot) \in \mathbb{R}^{n_y \times n_y}$ is called the sensor information matrix.*

*Remark*: Our solution to (5) is applicable to any *monotone concave* cost function but for clarity we use $\log \det(\cdot)$.

Separation principles like Thm. 1 have been exploited since the early work on sensor management [28], [29]. The result is significant because the state space of a deterministic control problem is smaller than the stochastic version and an open-loop policy can be computed *off-line*. As discussed by Le Ny and Pappas [19], the main bottleneck for computations is the large dimension of the state $(x_t, \Sigma_t)$ and it is beneficial to use forward value iteration (FVI, Alg. 1). The advantage is that the set of reachable target covariances is built progressively starting from the initial state rather than discretizing the high-dimensional space of all possible covariances as required by the more commonly used backward value iteration.

---

**Algorithm 1** Forward Value Iteration

1: $S_0 \leftarrow \{(x_0, \Sigma_0)\}, \quad S_t \leftarrow \emptyset$ for $t = 1, \ldots, T$
2: **for** $t = 1 : T$ **do**
3:    **for all** $(x, \Sigma) \in S_{t-1}$ **do**
4:       **for all** $u \in \mathcal{U}$ **do**
5:          $x_t \leftarrow f(x, u)$
6:          $S_t \leftarrow S_t \cup \{(x_t, \rho_{x_t}(\Sigma))\}$
7: **return** $\min \{\log \det(\Sigma) \mid (x, \Sigma) \in S_T\}$

---

Alg. 1 is constructing a *search tree*, $\mathcal{T}_t$, with nodes at stage $t \in [0, T]$ corresponding to the reachable states $(x_t, \Sigma_t)$. Starting from node $(x_t, \Sigma_t)$, there is an edge for each control in $\mathcal{U}$ leading to a node $(x_{t+1}, \Sigma_{t+1})$ obtained from the dynamics in (5). FVI is guaranteed to find the optimal control sequence $\sigma^*$ and provides a significant computational advantage over the backward version but still the number of nodes in $\mathcal{T}_t$ corresponds to the number of sensor trajectories of length $t$ and grows exponentially, $\mathcal{O}(|U|^T)$. The other extreme is the greedy approach, which discards all but the best (lowest cost) node at level $t$ of the tree $\mathcal{T}_t$. The greedy policy $\sigma^g$ is

$$\sigma_t^g \in \arg\min_{u \in \mathcal{U}} \left( \log \det\big(\rho_{f(x_t, u)}(\Sigma_t)\big) \right), \ t \in [0, T-1] \tag{7}$$

and is computationally very efficient, $\mathcal{O}(|\mathcal{U}|T)$, but no guarantees exist for its performance.

## IV. REDUCED VALUE ITERATION

### A. Optimality-preserving reductions

The goal of this paper is to provide an algorithm, with complexity lower than that of FVI and performance better than that of the greedy policy (7). This can be achieved by discarding some but not all of the nodes at level $t$ of $\mathcal{T}_t$. Intuitively, if two sensor trajectories cross, i.e. there are two nodes at level $t$ of $\mathcal{T}_t$ with the same vehicle configuration $x$ but different target covariances, and one is clearly less informative, i.e. its uncertainty about the target is "larger" than that of the other one, then it is not necessary to keep it. In the same vein, the covariance of a sensor path at time $t$ may be dominated by a combination of the covariances of several other paths. The following definition makes this notion precise.

**Definition 1** (Algebraic redundancy [10]). *Let $\{\Sigma_i\}_{i=1}^K \subset S_+^n$ be a finite set. A matrix $\Sigma \in S_+^n$ is algebraically redundant with respect to $\{\Sigma_i\}$, if there exist nonnegative constants $\{\alpha_i\}_{i=1}^K$ such that:*

$$\textstyle\sum_{i=1}^K \alpha_i = 1 \quad and \quad \Sigma \succeq \sum_{i=1}^K \alpha_i \Sigma_i.$$

The next theorem shows that when several sensor paths cross at time $t$, i.e. there are several nodes at level $t$ of $\mathcal{T}_t$ with the same vehicle state, algebraically redundant ones can be discarded without removing the optimal one.

**Theorem 2** (Optimal reduction). *For $t \in [1, T]$, let $(x, \Sigma) \in S_t$ be a node in level $t$ of the search tree $\mathcal{T}_t$. If there exist a set $\{x^i, \Sigma^i\} \subseteq S_t \setminus \{(x, \Sigma)\}$ such that $x = x^i$, $\forall i$ and $\Sigma$ is algebraically redundant with respect to $\{\Sigma^i\}$, then $(x, \Sigma)$ can be removed from $S_t$ without eliminating the optimal trajectory.*

### B. $\epsilon$-Suboptimal reductions

At the expense of losing optimality, we can discard even more of the crossing trajectories by using a relaxed notion of algebraic redundancy.

**Definition 2** ($\epsilon$-Algebraic redundancy [10]). *Let $\epsilon \geq 0$ and let $\{\Sigma_i\}_{i=1}^K \subset S_+^n$ be a finite set. A matrix $\Sigma \in S_+^n$ is $\epsilon$-algebraically redundant with respect to $\{\Sigma_i\}$, if there exist nonnegative constants $\{\alpha_i\}_{i=1}^K$ such that:*

$$\textstyle\sum_{i=1}^K \alpha_i = 1 \quad and \quad \Sigma + \epsilon I_n \succeq \sum_{i=1}^K \alpha_i \Sigma_i.$$

Let $\pi^* = x_1^*, \ldots, x_T^* \in \mathcal{X}^T$ be the optimal sensor trajectory in $\mathcal{T}_T$ with covariance sequence $\Sigma_1^*, \ldots, \Sigma_T^*$ and cost $J_T^* := \log \det(\Sigma_T^*)$. Denote the search tree at time $t$ with all $\epsilon$-algebraically redundant nodes removed by $\mathcal{T}_t^\epsilon$. Let $\pi^\epsilon = x_1^\epsilon, \ldots, x_T^\epsilon \in \mathcal{X}^T$ be the trajectory obtained by forward value iteration on the reduced tree $\mathcal{T}_T^\epsilon$ with a corresponding covariance sequence $\Sigma_1^\epsilon, \ldots, \Sigma_T^\epsilon$ and cost $J_T^\epsilon := \log \det(\Sigma_T^\epsilon)$. The following theorem provides an upper bound on the gap, $(J_T^\epsilon - J_T^*)$, between the performances of $\pi^\epsilon$ and $\pi^*$.

**Theorem 3** ($\epsilon$-Suboptimal reduction). *Let $\beta_* < \infty$ be the peak estimation error of the optimal policy, $\Sigma_t^* \preceq \beta_* I_{n_y}$, for $t \in [1, T]$. Then,*

$$0 \leq J_T^\epsilon - J_T^* \leq \epsilon \Delta_T, \tag{8}$$

*where*

$$\Delta_T := \frac{n_y}{\underline{\lambda}_W}\left(1 + \frac{\beta_*^2}{\underline{\lambda}_W^2}\left(1 - \eta_*^{T-1}\right)\right), \quad \eta_* := \frac{\beta_*}{\beta_* + \underline{\lambda}_W} < 1.$$

*Remark*: If the peak estimation error $\beta_*$ remains bounded as $T \to \infty$, i.e. the sensor performs well when using the optimal policy, then $\Delta_T \to \frac{n_y}{\underline{\lambda}_W}\left(1 + \frac{\beta_*^2}{\underline{\lambda}_W^2}\right)$. In other words, the suboptimality gap of $\pi^\epsilon$ is bounded regardless of the length $T$ of the planning horizon and grows linearly in $\epsilon$!

### C. $(\epsilon, \delta)$-*Suboptimal reductions*

If the motion of the sensor is restricted to a graph many of the planned trajectories will be crossing and the results from Thm. 3 are very satisfactory. However, if the space of sensor configurations, $\mathcal{X}$, is continuous, depending on the sensor motion model (e.g. differential drive), it is possible that no two sensor states reachable at time $t$ are exactly equal. Then, the reductions from Thm. 3 cannot be applied. To avoid this case, we can relax the notion of trajectory crossing at time $t$.

**Definition 3** (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 $\delta \geq 0$.*

Instead of discarding $\epsilon$-algebraically redundant trajectories which cross, we can discard those which $\delta$-cross. Let $\mathcal{T}_t^{\epsilon,\delta}$ be the reduced tree at time $t$, with all $\epsilon$-algebraically redundant $\delta$-crossing nodes removed. Some continuity assumptions are necessary in order to provide suboptimality guarantees for searching within $\mathcal{T}_t^{\epsilon,\delta}$.

**Assumption 1** (**Motion Model Continuity**). *The sensor motion model is Lipschitz continuous in $x$ with Lipschitz constant $L_f \geq 0$ for every fixed $u \in \mathcal{U}$:*

$$d_{\mathcal{X}}(f(x_1, u), f(x_2, u)) \leq L_f d_{\mathcal{X}}(x_1, x_2), \quad \forall x_1, x_2 \in \mathcal{X}.$$

**Assumption 2** (**Observation Model Continuity**). *There exists a real constant $L_m \geq 0$ such that:*

$$\begin{aligned} M(x_1) &\preceq \left(1 + L_m d_{\mathcal{X}}(x_1, x_2)\right) M(x_2) \\ M(x_2) &\preceq \left(1 + L_m d_{\mathcal{X}}(x_1, x_2)\right) M(x_1) \end{aligned}, \quad \forall x_1, x_2 \in \mathcal{X},$$

*where $M(\cdot)$ is the sensor information matrix (6).*

Assumption 1 says that when two sensor configurations are close and the same sequence of controls is applied, then the resulting trajectories will remain close. Assumption 2 says that sensing from similar configurations results in similar information gain. This gives the intuition that when two trajectories $\delta$-cross their future informativeness will be similar. We make this intuition precise in the next theorem. Let $\pi^{\epsilon,\delta} \in \mathcal{X}^T$ be the sensor trajectory obtained by searching the reduced tree $\mathcal{T}_T^{\epsilon,\delta}$ with corresponding cost $J_T^{\epsilon,\delta}$. The gap, $(J_T^{\epsilon,\delta} - J_T^*)$, between the performances of $\pi^{\epsilon,\delta}$ and $\pi^*$ is bounded as follows.

**Theorem 4** ($(\epsilon, \delta)$-Suboptimal reduction). *Let $\beta_* < \infty$ be the peak estimation error of the optimal policy, $\Sigma_t^* \preceq \beta_* I_{n_y}$, for $t \in [1, T]$. Then,*

$$0 \leq J_T^{\epsilon,\delta} - J_T^* \leq (\zeta_T - 1)\left[J_T^* - \log \det(W)\right] + \epsilon \Delta_T, \quad (9)$$

*where* $\quad \zeta_t := \prod_{\tau=1}^{t-1}\left(1 + \sum_{s=1}^{\tau} L_f^s L_m \delta\right) \geq 1,$

$$\Delta_T := \frac{n_y}{\underline{\lambda}_W}\left(1 + \frac{\beta_*}{\underline{\lambda}_W}\sum_{\tau=1}^{T-1}\frac{\zeta_T}{\zeta_\tau}\eta_*^{T-\tau}\right), \quad \eta_* := \frac{\beta_*}{\beta_* + \underline{\lambda}_W} < 1.$$

Notice that Thm. 3 is a special case of Thm. 4 because if $\delta = 0$, then $\zeta_t = 1, \forall t \in [1, T]$ and (9) reduces to (8). Then, the suboptimality gap remains bounded regardless of the time horizon and grows linearly with $\epsilon$. If $\delta > 0$, then $\lim_{T \to \infty} \zeta_T = \infty$ and the suboptimality gap grows with $T$ and $\delta$. The bound is loose, however, because it uses a worst-case analysis. The worst case happens when a trajectory, which was discarded from $\mathcal{T}_T^{\epsilon,\delta}$, persistently obtains more information than a trajectory, which remains very close in space and is still in the search tree. Even then, if the optimal policy performs well, the term $J_T^* - \log \det(W)$ gets smaller as $\zeta_T$ increases and the suboptimality gap remains small.

### D. $(\epsilon, \delta)$-*Reduced value iteration*

The approaches for reducing the search tree, developed in the previous subsections, can be used to significantly decrease the complexity of the FVI algorithm (Alg. 1), while providing suboptimality guarantees (Thm. 4). The $(\epsilon, \delta)$-Reduced Value Iteration (RVI) is summarized in Alg. 2.

---

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

---

1: $S_0 \leftarrow \{(x_0, \Sigma_0)\}, \quad S_t \leftarrow \emptyset$ for $t = 1, \ldots, T$
2: **for** $t = 1 : T$ **do**
3:     **for all** $(x, \Sigma) \in S_{t-1}$ **do**
4:         **for all** $u \in \mathcal{U}$ **do**
5:             $x_t \leftarrow f(x, u)$
6:             $S_t \leftarrow S_t \cup \{(x_t, \rho_{x_t}(\Sigma))\}$
7:     Sort $S_t$ in ascending order according to $\log \det(\cdot)$
8:     $S_t' \leftarrow S_t[1]$           % State with lowest cost
9:     **for all** $(x, \Sigma) \in S_t \setminus S_t[1]$ **do**
10:         % Find all nodes in $S_t'$, which $\delta$-cross $x$:
11:         $Q \leftarrow \{\Sigma' \mid (x', \Sigma') \in S_t', \ d_{\mathcal{X}}(x, x') \leq \delta\}$
12:         **if** isempty($Q$) **or not**( $\Sigma$ is $\epsilon$-alg. redundant wrt $Q$ ) **then**
13:             $S_t' \leftarrow S_t' \cup (x, \Sigma)$
14:     $S_t \leftarrow S_t'$
15: **return** $\min\{\log \det(\Sigma) \mid (x, \Sigma) \in S_T\}$

---

The most computationally demanding operation is checking $\epsilon$-algebraic redundancy (Line 12). It is a feasibility problem for a linear matrix inequality (LMI) and off-the-shelf solvers exist [30]. An appealing property of Alg. 2 is that at stage $t$ at least the node with currently lowest cost is retained (Line 8). This guarantees that the control sequence obtained from RVI *performs at least as well as* the greedy policy (7).

## V. APPLICATIONS

### A. *Gas distribution mapping and leak localization*

In this subsection, we apply the $(\epsilon, \delta)$-RVI to the methane monitoring problem introduced in Sec. II. Since the movement of the Gasbot is restricted to a grid, the planned sensor paths will be crossing frequently and we can use $\delta = 0$. The dimension $n_y$ of the target is the number of cells in the gas concentration map and would typically be very large. Checking

algebraic redundancy requires solving an $n_y$-dimensional LMI feasibility problem, which is computationally very demanding. To avoid this, we let $\epsilon = \infty$. This means that when several paths cross at time $t$, only the most informative one is kept in $\mathcal{T}_t^{\epsilon,\delta}$. Thus, the number of nodes in $\mathcal{T}_t^{\epsilon,\delta}$ remains bounded by the number of reachable sensor states. Trajectories of length $T = 40$ were planned using RVI and the greedy algorithm (GREEDY). The results (Fig. 1) reveal an important drawback of GREEDY. It remains trapped in a local region of relatively high variance and fails to see that there are more interesting regions which should be explored during the limited available time. Fig. 1 also shows that the growth of the search tree is much slower for RVI compared to FVI, while the quality of the RVI solution is better than the greedy one.

### B. Target localization and tracking

In a lot of applications, the linear assumptions (2), (3) are reasonable for the target motion model but very limiting for the sensor observation model. A lot more problems can fit in a framework with the following non-linear observation model:

$$z_t = h(x_t, y_t) + v(x_t, y_t), \ v(x_t, y_t) \sim \mathcal{N}(0, V(x_t, y_t)). \quad (10)$$

In this subsection, we show that our method can be coupled with linearization and model predictive control to generate an adaptive policy for a mobile sensor with the model in (10).

To motivate the discussion we consider a target tracking application, in which both the sensor motion and observation models are highly non-linear. Suppose that the sensor is mounted on a vehicle with differential-drive dynamics, which are discretized using a sampling period $\tau$ as follows:

$$\begin{pmatrix} x_{t+1}^1 \\ x_{t+1}^2 \\ \theta_{t+1} \end{pmatrix} = \begin{pmatrix} x_t^1 \\ x_t^2 \\ \theta_t \end{pmatrix} + \begin{cases} \begin{pmatrix} \tau\nu\cos(\theta_t + \tau\omega/2) \\ \tau\nu\sin(\theta_t + \tau\omega/2) \\ \tau\omega \end{pmatrix}, \ |\tau\omega| < 0.001 \\ \begin{pmatrix} \frac{\nu}{\omega}(\sin(\theta_t + \tau\omega) - \sin\theta_t) \\ \frac{\nu}{\omega}(\cos\theta_t - \cos(\theta_t + \tau\omega)) \\ \tau\omega \end{pmatrix}, \ \text{else.} \end{cases}$$

The vehicle is controlled using the motion primitives $\mathcal{U} = \{(\nu, \omega) \mid \nu \in \{0, 1, 2, 3\} \text{ m/s}, \ \omega \in \{0, \pm\pi/2, \pm\pi\} \text{ rad/s}\}$. The task of the sensor is to track the position $(y^1, y^2) \in \mathbb{R}^2$ and velocity $(\dot{y}^1, \dot{y}^2) \in \mathbb{R}^2$ of another vehicle with discretized double integrator dynamics driven by Gaussian noise:

$$y_{t+1} = \begin{bmatrix} I_2 & \tau I_2 \\ 0 & I_2 \end{bmatrix} y_t + w_t, \ \ w_t \sim \mathcal{N}\left(0, q \begin{bmatrix} \tau^3/3I_2 & \tau^2/2I_2 \\ \tau^2/2I_2 & \tau I_2 \end{bmatrix}\right),$$

where $y = [y^1, y^2, \dot{y}^1, \dot{y}^2]^T$ is the target state and $q$ is a scalar diffusion strength measured in $(\frac{m}{sec^2})^2 \frac{1}{Hz}$. The sensor takes noisy range and bearing measurments of the target's position:

$$h(x, y) = \begin{bmatrix} r(x, y) \\ \alpha(x, y) \end{bmatrix} := \begin{bmatrix} \sqrt{(y^1 - x^1)^2 + (y^2 - x^2)^2} \\ \arctan\left((y^2 - x^2)/(y^1 - x^1)\right) \end{bmatrix} \quad (11)$$

The target needs to be tracked during a period $T_{max}$ in a wooded area (Fig. 2), which affects the covariance of the measurement noise. The noise in the range measurement grows linearly with the distance between the sensor and the target

but trees along the way make the growth faster. The bearing measurement noise increases linearly with the speed of the sensor. The sensor has a *maximum range* of 15 meters, after which the noise covariance is infinite.

To employ RVI, the observation model (11) needs to be linearized about a predicted target trajectory during planning. Linearized about an arbitrary target state $y \neq x$, the model is:

$$H(x, y) := \nabla_y h(x, y) = \frac{1}{r(x, y)} \begin{bmatrix} (y^1 - x^1) & (y^2 - x^2) \\ -\sin\alpha(x, y) & \cos\alpha(x, y) \end{bmatrix}.$$

The linearization depends on the predicted target trajectory, which in turn depends on the measurements obtained on-line. As a result, it is necessary to re-plan the sensor path on-line after every new measurement. In general, re-planning would be feasible only for a short horizon $T < T_{max}$ before the plan is needed. Alg. 3 shows how to use the $(\epsilon, \delta)$-RVI and model predictive control to generate an adaptive policy.

---

**Algorithm 3** Model predictive control via the $(\epsilon, \delta)$-RVI
___

1: **Input**: $T_{max}, x_0, \hat{y}_0, \hat{\Sigma}_0, f, \mathcal{U}, H, V, A, W, T, \epsilon, \delta$
2: **for** $t = 0 : T_{max}$ **do**
3:     Predict a target trajectory of length $T$: $\hat{y}_t, \ldots, \hat{y}_{t+T}$
4:     Linearize the observation model: $H_\tau(\cdot) \leftarrow H(\cdot, \hat{y}_{t+\tau}), \tau = 0, .., T$
5:     *% Plan a sensor trajectory of length $T$*
6:     $\sigma \leftarrow (\epsilon, \delta)$-RVI$(x_t, \hat{y}_t, \hat{\Sigma}_t, \{H_\tau\}_{\tau=0}^T, V, f, \mathcal{U}, A, W, T)$
7:     Move the sensor: $x_{t+1} \leftarrow f(x_t, \sigma_1)$
8:     Get a new observation: $z_{t+1} \leftarrow h(x_{t+1}, y_{t+1}) + v_{t+1}(x_{t+1})$
9:     Update the target estimate: $(\hat{y}_{t+1}, \hat{\Sigma}_{t+1}) \leftarrow Filter(\hat{y}_t, \hat{\Sigma}_T, z_{t+1})$

---

We emphasize that the linearized sensing models are utilized merely for determining the next configuration (Line 6), while the target state inference (Line 9) can still be performed with *any non-linear filter*. Alg. 3 was implemented with $T_{max} = 100$, $T = 7$, $\epsilon = 0.1$, $\delta = 1$, $\tau = 0.5$, $q = 0.2$ and 100 Monte-Carlo simulations were carried out. The tracking performance is compared to the greedy solution in Fig. 2. We can see that the greedy policy goes in straight line to keep the speed low, i.e. the bearing noise small, but cannot predict in advance that the range noise will increase as the target goes further away.

## VI. CONCLUSION

Under linear Gaussian assumptions, information acquisition with mobile sensors can be planned off-line. Most previous approaches are greedy or neglect the sensor dynamics and rarely provide performance guarantees. In this paper, we developed a non-greedy algorithm (RVI), which takes the sensor dynamics into account and has suboptimality guarantees. It provides parameters $(\epsilon, \delta)$ to control the trade-off between complexity and optimality. Coupled with linearization and model predictive control, RVI can generate an adaptive policy for a non-linear mobile sensor. Our method can also be applied to a multi-sensor problem by stacking all motion and observation models into single centralized versions. Unfortunately, the complexity of this solution is exponential in the number of sensors. Future work will focus on a decentralized method for the multi-sensor, multi-target active information acquisition. In detail, we would like to address cooperative control, distributed target estimation, and noisy self-localization.
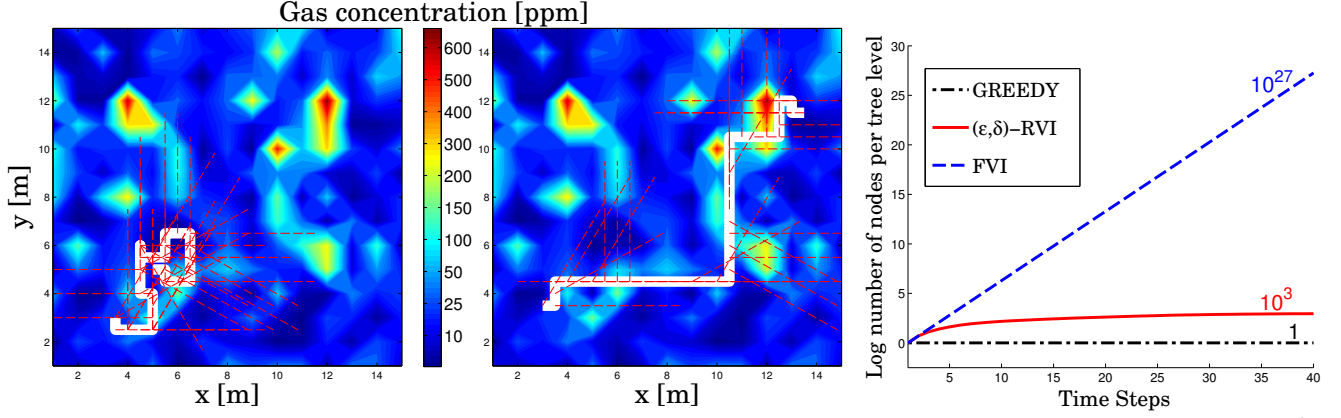
Fig. 1. Comparison of the sensor trajectories (white) obtained by the greedy algorithm (left) and the reduced value iteration (middle) with $\epsilon = \infty$ and $\delta = 0$ after 40 time steps. A typical realization of the methane field is shown. The red lines indicate the orientation of the gas sensor during the execution. On the right, the log number of nodes maintained in the search tree by the two approaches is compared to the complete tree maintained by forward value iteration.
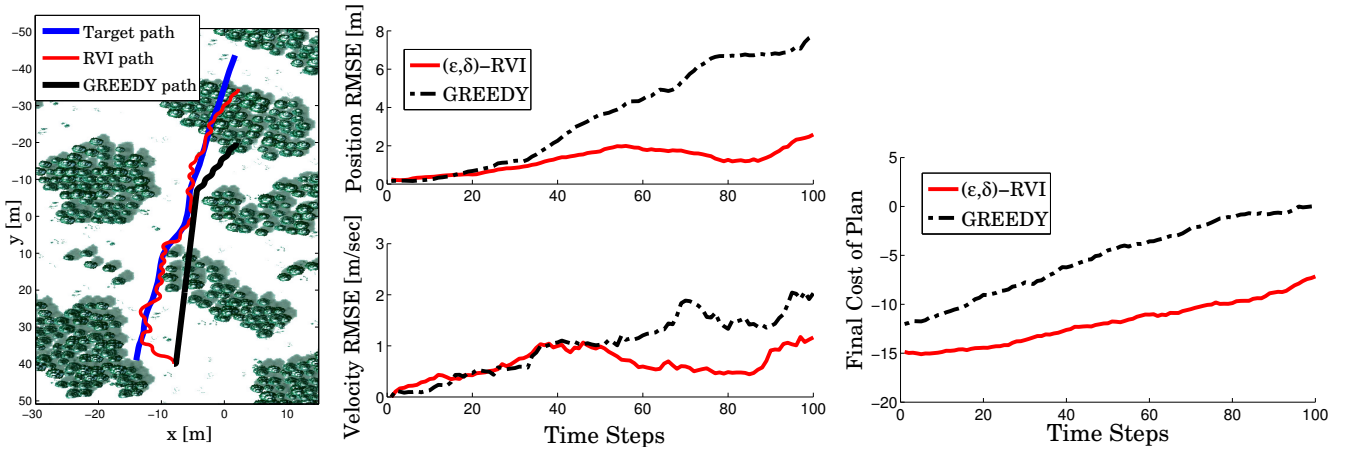


Fig. 2. Simulation results from 100 Monte-Carlo runs of the target tracking scenario. A typical realization is shown on the left. The average root-mean-square error (RMSE) of the estimated target position and velocity is shown in the middle. The $\log \det$ of the predicted target covariance is shown on the right.

## APPENDIX A: PROOF OF THEOREM 1

**Lemma 1.** *Let $X \sim \mathcal{N}(\mu, \Sigma)$ be $n$-dimensional. Then, its differential entropy is $\mathbf{h}(X) = \left( n \log(2\pi e) + \log \det(\Sigma) \right)/2$.*

*Proof of Thm 1:* By definition of mutual information:

$$\mathbb{I}(y_{T+1}; z_{1:T} \mid x_{1:T}) = \mathbf{h}(y_{T+1} \mid x_{1:T}) - \mathbf{h}(y_{T+1} \mid z_{1:T}, x_{1:T}).$$

Since $y_{T+1}$ is independent of the sensor path the first term is constant with respect to the optimization in (4). Due to linearity of the observation model (3) in the target state, the distribution of $y_{t+1}$ given $z_{1:t}$ and $x_{1:t}$ remains Gaussian for $t > 1$. Its covariance $\Sigma_t$ can be obtained from the Bayes filter, which due to the linear Gaussian assumptions is equivalent to the Kalman filter. Thus, $\Sigma_{t+1} = \rho_{x_t}(\Sigma_t)$ for $t = 0, \dots, T-1$, which is *independent* of the measurements $z_{1:t}$. By Lemma 1:

$$\mathbf{h}(y_{T+1} \mid z_{1:T}, x_{1:T}) = \mathbb{E}_{\hat{z}_{1:T}} \mathbf{h}(y_{T+1} \mid z_{1:T} = \hat{z}_{1:T}, x_{1:T}) \quad (12)$$
$$= \frac{1}{2} \mathbb{E}_{\hat{z}_{1:T}} \left( \log(2\pi e)^{n_y} + \log |\Sigma_T| \right) = \frac{1}{2} \left( \log(2\pi e)^{n_y} + \log |\Sigma_T| \right).$$

Let $\mu^* = \{\mu_0^*, \dots, \mu_{T-1}^*\}$ be optimal in (4) with cost $J^*$. Fix a realization $\hat{z}_{1:T}$ of the measurements and let $\sigma$ be the open-loop policy induced by $\mu^*$ given $\hat{z}_{1:T}$ with cost $J^\sigma$. From (12), $J^*$ is independent of $\hat{z}_{1:T}$, hence $J^* = J^\sigma$ for any $\hat{z}_{1:T}$. ∎

## APPENDIX B: PROOF OF THEOREM 2

**Definition 4** (*$t$-step Riccati map*). *For $\pi \in \mathcal{X}^T$, $\Sigma \in S_+^n$ let*

$$\phi_\pi^0(\Sigma) := \Sigma, \qquad \phi_\pi^t(\Sigma) := \rho_{\pi_t} \circ \dots \circ \rho_{\pi_1}(\Sigma), \quad t \in [1, T].$$

**Lemma 2** ([10])**.** *For any $t \in [0, T]$, the $t$-step Riccati map is operator monotone and operator concave.*

*Proof of Thm 2:* Let $\sigma \in U^{T-t}$ be any admissible control sequence. Starting at $(x, \Sigma)$, by Lemma 2 and Definition 1, there exist nonnegative constants $\{\alpha_i\}_{i=1}^K$ such that

$$\phi_{\pi(\sigma)}^{T-t}(\Sigma) \succeq \phi_{\pi(\sigma)}^{T-t} \left( \sum_{i=1}^K \alpha_i \Sigma^i \right) \succeq \sum_{i=1}^K \alpha_i \phi_{\pi(\sigma)}^{T-t}(\Sigma^i).$$

Then, from monotonicity and concavity of $\log \det(\cdot)$:

$$\log \det \left( \phi_{\pi(\sigma)}^{T-t}(\Sigma) \right) \geq \log \det \left( \sum_{i=1}^K \alpha_i \phi_{\pi(\sigma)}^{T-t}(\Sigma^i) \right)$$
$$\geq \sum_{i=1}^K \alpha_i \log \det \left( \phi_{\pi(\sigma)}^{T-t}(\Sigma^i) \right) \geq \log \det \left( \phi_{\pi(\sigma)}^{T-t}(\Sigma^{i^*}) \right).$$

The last inequality holds because a convex combination of scalars is lower bounded by the smallest one $i^*$. ∎

**Lemma 3** ([10])**.** *For any $\pi \in \mathcal{X}^T$, $Q \in S_+^n$, and $\epsilon \geq 0$ the directional derivative of the $t$-step Riccati map is:*

$$g_\pi^t(\Sigma; Q) := \frac{d}{d\epsilon} \phi_\pi^t(\Sigma + \epsilon Q)\Big|_{\epsilon=0}$$

$$= \prod_{\tau=t}^{1} A C_{\pi_\tau}(\phi_\pi^\tau(\Sigma)) Q \prod_{\tau=1}^{t} C_{\pi_\tau}(\phi_\pi^\tau(\Sigma))^T A^T.$$

**Lemma 4.** *For any $t \in [0, T]$, $\pi \in \mathcal{X}^T$, $\Sigma, Q_1, Q_2 \in S_+^n$, $a, b \in \mathbb{R}$*

$$g_\pi^t(\Sigma; aQ_1 + bQ_2) = a g_\pi^t(\Sigma; Q_1) + b g_\pi^t(\Sigma; Q_2)$$

*because a directional derivative is linear in the perturbation. In addition, by operator concavity of the $t$-step Riccati map:*

$$\phi_\pi^t(\Sigma + \epsilon Q) \preceq \phi_\pi^t(\Sigma) + \epsilon g_\pi^t(\Sigma; Q)$$

**Lemma 5.** *For all $t \in [1, T]$, $\pi \in \mathcal{X}^T$, and $Q \in S_+^n$, if there exists a constant $\beta < \infty$ such that $\phi_\pi^t(\Sigma) \preceq \beta I_{n_y}$, then*

$$\operatorname{tr}\big(g_\pi^t(\Sigma; Q)\big) \leq \beta \eta^t \operatorname{tr}(\Sigma^{-1} Q), \qquad \eta := \beta/(\beta + \underline{\lambda}_W) < 1.$$

*Proof:* We follow the steps of [10, Thm. 5] but use the tighter bound $\hat{A}_t \hat{\Sigma}_t \hat{A}_t^T \preceq (\beta - \lambda_{\bar{W}}) I_n$ in (A.4), which leads to $L_t^{(l)} - L_{t+1}^{(l)} \geq \frac{\lambda_{\bar{W}}}{\beta} L_t^{(l)}$. Also, since $Q \in S_+^n$ we can decompose it as $Q = \sum_{l=1}^{n} \lambda_Q^l q_l q_l^T$ and let $\xi^{(l)}(0) = \sqrt{\lambda_Q^l} q_l$. ∎

**Lemma 6.** *Consider two nodes $(x_1, \Sigma)$ and $(x_2, \Sigma)$ with $d_\mathcal{X}(x_1, x_2) \leq \delta$. Let $\Sigma_1$ and $\Sigma_2$ be the updated covariance matrices after applying $u \in \mathcal{U}$ at each node. Then,*

$$\Sigma_1 \succeq \gamma \Sigma_2 + (1 - \gamma) W \quad and \quad \Sigma_2 \succeq \gamma \Sigma_1 + (1 - \gamma) W,$$

*where $0 < \gamma := (1 + L_m L_f \delta)^{-1} \leq 1$.*

*Proof:* Let $M_i := M(f(x_i, u))$ and $\rho_i(\cdot) := \rho_{f(x_i,u)}(\cdot)$ for $i = 1, 2$. From Assumption 1, $d_\mathcal{X}(f(x_1, u), f(x_2, u)) \leq L_f \delta$ and from Assumption 2, $M_1 \preceq (1 + L_m L_f \delta) M_2$. Then,

$$\rho_1(\Sigma) = A(\Sigma^{-1} + M_1)^{-1} A^T + W$$
$$\succeq A(\Sigma^{-1} + \gamma^{-1} M_2)^{-1} A^T + W$$
$$= \gamma \rho_2(\gamma^{-1} \Sigma) + (1 - \gamma) W \succeq \gamma \rho_2(\Sigma) + (1 - \gamma) W$$

The second result follows analogously. ∎

**Lemma 7.** *For $t \in [1, T]$, $\epsilon \geq 0$, $\delta \geq 0$, the reduced tree $\mathcal{T}_t^{\epsilon,\delta}$ contains a set of nodes $\{(x_t^i, \Sigma_t^i) \mid i = 1, \ldots, K\}$ such that:*

$$d_\mathcal{X}(x_t^*, x_t^i) \leq \sum_{\tau=0}^{t-1} L_f^\tau \delta, \quad \forall i, \tag{13}$$

$$\Sigma_t^* + \epsilon \left( \Gamma_t I_{n_y} + \sum_{\tau=1}^{t-1} \Gamma_\tau g_{\pi^*(\tau+1)}^{t-\tau}(\Sigma_\tau^*; I_{n_y}) \right) \tag{14}$$

$$\succeq \Gamma_t \sum_{i=1}^{K} \alpha_i \Sigma_t^i + \sum_{\tau=1}^{t-1} \Gamma_\tau (1 - \gamma_\tau) \phi_{\pi^*(\tau+1)}^{t-1-\tau}(W),$$

*where $\pi^* = x_1^*, \ldots, x_T^*$ is the optimal trajectory in 5,*

$$0 < \gamma_t := (1 + \textstyle\sum_{s=1}^{t} L_f^s L_m \delta)^{-1} \leq 1, \quad and \quad \Gamma_t := \textstyle\prod_{s=1}^{t-1} \gamma_s.$$

*Proof of Lemma 7:* We proceed by induction.
**Base Case**: At time 1, (13), (14) follow from Def. 2 and 3.
**Hypothesis**: Suppose that (13) and (14) hold for some set $\{(x_t^j, \Sigma_t^j) \mid j = 1, \ldots, J\}$ of nodes in $\mathcal{T}_t^{\epsilon,\delta}$.
**Induction**: At time $t + 1$, there are sets $\{(x_{t+1}^{ji}, \Sigma_{t+1}^{ji})\}_{i=1}^{K_j}$ in $\mathcal{T}_{t+1}^{\epsilon,\delta}$ corresponding to each node $j$ from time $t$ and satisfying

$$d_\mathcal{X}(x_{t+1}^j, x_{t+1}^{ji}) \leq \delta \text{ and } \Sigma_{t+1}^j + \epsilon I_{n_y} \succeq \sum_{i=1}^{K_j} \alpha_{ji} \Sigma_{t+1}^{ji}. \tag{15}$$

From Lemma 3 for every $\tau = 1, \ldots, t$:

$$g_{\pi^*(t+1)}^1 \left( \Sigma_t^*; g_{\pi^*(\tau+1)}^{t-\tau}(\Sigma_\tau^*; I_{n_y}) \right) = g_{\pi^*(\tau+1)}^{t+1-\tau}(\Sigma_\tau^*; I_{n_y}).$$

From this and Lemma 4:

$$\Sigma_{t+1}^* + \epsilon \sum_{\tau=1}^{t} \Gamma_\tau g_{\pi^*}^{t+1-\tau}(\Sigma_\tau^*; I_{n_y})$$

$$= \rho_{\pi_{t+1}^*}(\Sigma_t^*) + \epsilon g_{\pi^*(t+1)}^1 \left( \Sigma_t^*; \sum_{\tau=1}^{t} \Gamma_\tau g_{\pi^*(\tau+1)}^{t-\tau}(\Sigma_\tau^*; I_{n_y}) \right)$$

$$\succeq \rho_{\pi_{t+1}^*} \left( \Sigma_t^* + \epsilon \sum_{\tau=1}^{t-1} \Gamma_\tau g_{\pi^*(\tau+1)}^{t-\tau}(\Sigma_\tau^*; I_{n_y}) + \epsilon \Gamma_t I_{n_y} \right)$$

Note that $\sum_{\tau=1}^{t-1} \Gamma_\tau (1 - \gamma_\tau) + \Gamma_t = 1$. Thus, the terms $(1 - \gamma_1), \gamma_1(1 - \gamma_2), \ldots, \Gamma_{t-1}(1 - \gamma_{t-1}), \Gamma_t$ are the coefficients of a convex combination. Using the hypothesis and Lemma 2:

$$\rho_{\pi_{t+1}^*} \left( \Sigma_t^* + \epsilon \sum_{\tau=1}^{t-1} \Gamma_\tau g_{\pi^*(\tau+1)}^{t-\tau}(\Sigma_\tau^*; I_{n_y}) + \epsilon \Gamma_t I_{n_y} \right)$$

$$\succeq \rho_{\pi_{t+1}^*} \left( \Gamma_t \sum_{j=1}^{J} \alpha_j \Sigma_t^j + \sum_{\tau=1}^{t-1} \Gamma_\tau (1 - \gamma_\tau) \phi_{\pi^*(\tau+1)}^{t-1-\tau}(W) \right)$$

$$\succeq \Gamma_t \sum_{j=1}^{J} \alpha_j \rho_{\pi_{t+1}^*}(\Sigma_t^j) + \sum_{\tau=1}^{t-1} \Gamma_\tau (1 - \gamma_\tau) \phi_{\pi^*(\tau+1)}^{t-\tau}(W).$$

By hypothesis, $d_\mathcal{X}(x_t^*, x_t^j) \leq \sum_{\tau=0}^{t-1} L_f^\tau \delta$, and from Lemma 6:

$$\rho_{\pi_{t+1}^*}(\Sigma_t^j) \succeq \gamma_t \Sigma_{t+1}^j + (1 - \gamma_t) W.$$

The nodes $\{(x_{t+1}^j, \Sigma_{t+1}^j)\}$ might not be in $\mathcal{T}_{t+1}^{\epsilon,\delta}$ but from (15):

$$\rho_{\pi_{t+1}^*}(\Sigma_t^j) + \gamma_t \epsilon I_{n_y} \succeq \gamma_t \sum_{i=1}^{K_j} \alpha_{ji} \Sigma_{t+1}^{ji} + (1 - \gamma_t) W.$$

Combining the previous results, we have:

$$\Sigma_{t+1}^* + \epsilon \sum_{\tau=1}^{t} \Gamma_\tau g_{\pi^*}^{t+1-\tau}(\Sigma_\tau^*; I_{n_y}) + \epsilon \Gamma_{t+1} I_{n_y}$$

$$\succeq \Gamma_t \sum_{j=1}^{J} \alpha_j \left( \rho_{\pi_{t+1}^*}(\Sigma_t^j) + \gamma_t \epsilon I_{n_y} \right) + \underbrace{\sum_{\tau=1}^{t-1} \Gamma_\tau (1 - \gamma_\tau) \phi_{\pi^*(\tau+1)}^{t-\tau}(W)}_{(*)}$$

$$\succeq \Gamma_t \sum_{j=1}^{J} \alpha_j \left( \gamma_t \sum_{i=1}^{K_j} \alpha_{ji} \Sigma_{t+1}^{ji} + (1 - \gamma_t) W \right) + (*)$$

$$= \Gamma_{t+1} \sum_{j=1}^{J} \sum_{i=1}^{K_j} \alpha_j \alpha_{ji} \Sigma_{t+1}^{ji} + \sum_{\tau=1}^{t} \Gamma_\tau (1 - \gamma_\tau) \phi_{\pi^*(\tau+1)}^{t-\tau}(W).$$

Thus, the set $\bigcup_{j=1}^{J} \bigcup_{i=1}^{K_j}\{(x_{t+1}^{ji}, \Sigma_{t+1}^{ji})\}$ satisfies (14) at time $t+1$. It satisfies (13) at $t+1$ from (15) and Assumption 1. ∎

*Proof of Thm 4:* As defined in Lemma 7 $\Gamma_T = \zeta_T^{-1}$. Define

$$J(\cdot) := \log\det(\cdot) \text{ and } G := \Gamma_T I_{n_y} + \sum_{\tau=1}^{T-1} \Gamma_\tau g_{\pi^*(\tau+1)}^{T-\tau}(\Sigma_\tau^*; I_{n_y}).$$

By monotonicity of $J(\cdot)$ and the result in Lemma 7:

$$J(\Sigma_T^* + \epsilon G) \geq J\left(\Gamma_T \sum_{i=1}^{K} \alpha_i \Sigma_T^i + \sum_{\tau=1}^{T-1} \Gamma_\tau (1-\gamma_\tau) \phi_{\pi^*(\tau+1)}^{T-1-\tau}(W)\right)$$

for some set of nodes $\{(x_T^i, \Sigma_T^i) \mid i = 1, \ldots, K\}$ in the reduced tree $\mathcal{T}_T^{\epsilon,\delta}$. By definition, $\phi_\pi^t(W) \succeq W$ for any $t, \pi$ and $\sum_{\tau=1}^{T-1} \Gamma_\tau(1-\gamma_\tau) = 1 - \Gamma_T$. By concavity of $J(\cdot)$:

$$J(\Sigma_T^* + \epsilon G) \geq \Gamma_T \sum_{i=1}^{K} \alpha_i J(\Sigma_T^i) + (1-\Gamma_T)J(W) \qquad (16)$$

$$\geq \Gamma_T J(\Sigma_T^{i^*}) + (1-\Gamma_T)J(W) \geq \Gamma_T J_T^{\epsilon,\delta} + (1-\Gamma_T)J(W).$$

The second inequality holds because a convex combination of scalars is lower bounded by the smallest one $i^*$. The last inequality holds because $\pi^{\epsilon,\delta}$ is the optimal path in the reduced tree. Next, by concavity of $\log\det(\cdot)$:

$$J(\Sigma_T^* + \epsilon G) \leq J(\Sigma_T^*) + \epsilon \frac{d}{d\epsilon} J\left(\Sigma_T^* + \epsilon G\right)\Big|_{\epsilon=0}$$

$$= J_T^* + \epsilon \operatorname{tr}\left((\Sigma_T^*)^{-1} G\right) \leq J_T^* + \epsilon \frac{1}{\underline{\lambda}_W} \operatorname{tr}(G). \qquad (17)$$

From Lemma 5 and since $\operatorname{tr}((\Sigma_T^*)^{-1}) \leq n_y/\underline{\lambda}_W$:

$$\operatorname{tr}(G) = \Gamma_T \operatorname{tr}(I_{n_y}) + \sum_{\tau=1}^{T=1} \Gamma_\tau \operatorname{tr}\left(g_{\pi^*(\tau+1)}^{T-\tau}(\Sigma_\tau^*; I_{n_y})\right)$$

$$\leq n_y \Gamma_T + \sum_{\tau=1}^{T=1} \Gamma_\tau \beta_* \eta_*^{T-\tau} \operatorname{tr}((\Sigma_T^*)^{-1}) \leq \Gamma_T \Delta_T \quad (18)$$

Finally, by combining (16), (17), and (18) we get:

$$\Gamma_T J_T^{\epsilon,\delta} + (1-\Gamma_T)J(W) \leq J_T^* + \epsilon \Gamma_T \Delta_T$$

$$0 \leq \Gamma_T(J_T^{\epsilon,\delta} - J_T^*) \leq (1-\Gamma_T)(J_T^* - J(W)) + \epsilon \Gamma_T \Delta_T.$$

Multiplying by $\zeta_T = \Gamma_T^{-1}$ gives the result in (9). ∎

## REFERENCES

[1] H. Choi, "Adaptive Sampling and Forecasting With Mobile Sensor Networks," Ph.D. dissertation, MIT, 2009.

[2] V. Hernandez Bennetts, A. Lilienthal, A. Khaliq, V. Pomareda Sese, and M. Trincavelli, "Towards Real-World Gas Distribution Mapping and Leak Localization Using a Mobile Robot with 3D and Remote Gas Sensing Capabilities," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2013.

[3] P. Rybski, S. Stoeter, M. Erickson, M. Gini, D. Hougen, and N. Papanikolopoulos, "A Team of Robotic Agents for Surveillance," in *Proc. Int. Conf. on Autonomous Agents*, 2000.

[4] A. Hilal, "An Intelligent Sensor Management Framework for Pervasive Surveillance," Ph.D. dissertation, University of Waterloo, 2013.

[5] V. Kumar, D. Rus, and S. Singh, "Robot and Sensor Networks for First Responders," *IEEE Pervasive Computing*, vol. 3, no. 4, 2004.

[6] R. Sim and N. Roy, "Global A-Optimal Robot Exploration in SLAM," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2005.

[7] V. Karasev, A. Chiuso, and S. Soatto, "Controlled Recognition Bounds for Visual Learning and Exploration," in *Advances in Neural Information Processing Systems (NIPS)*, 2012.

[8] N. Atanasov, B. Sankaran, J. Le Ny, T. Koletschka, G. Pappas, and K. Daniilidis, "Hypothesis Testing Framework for Active Object Detection," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2013.

[9] A. Hero III and D. Cochran, "Sensor Management: Past, Present, and Future," *IEEE Sensors Journal*, vol. 11, no. 12, 2011.

[10] 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.

[11] S. Joshi and S. Boyd, "Sensor Selection via Convex Optimization," *IEEE Trans. on Signal Processing*, vol. 57, no. 2, 2009.

[12] J. Williams, "Information Theoretic Sensor Management," Ph.D. dissertation, MIT, 2007.

[13] V. Krishnamurthy and R. Evans, "Hidden Markov Model Multiarm Bandits: A Methodology for Beam Scheduling in Multitarget Tracking," *IEEE Trans. on Signal Processing*, vol. 49, no. 12, 2001.

[14] A. Singh, A. Krause, C. Guestrin, and W. Kaiser, "Efficient Informative Sensing Using Multiple Robots," *Journal of Artificial Intelligence Research (JAIR)*, vol. 34, no. 1, 2009.

[15] B. Grocholsky, "Information Theoretic Control of Multiple Sensor Platforms," Ph.D. dissertation, University of Sidney, 2002.

[16] T. Chung, J. Burdick, and R. Murray, "A Decentralized Motion Coordination Strategy for Dynamic Target Tracking," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2006.

[17] C. Kreucher, "An Information-based Approach to Sensor Resource Allocation," Ph.D. dissertation, University of Michigan, 2005.

[18] M. Huber, "Probabilistic framework for sensor management," Ph.D. dissertation, Universität Karlsruhe (TH), 2009.

[19] 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.

[20] A. Singh, A. Krause, and W. Kaiser, "Nonmyopic Adaptive Informative Path Planning for Multiple Robots," in *21st Int. Joint Conf. on Artificial Intelligence (IJCAI)*, 2009.

[21] J. Vander Hook, P. Tokekar, and V. Isler, "Bearing-Only Active Target Localization Strategies for a System of Two Communicating Mobile Robots: Full Technical Report," University of Minnesota, Tech. Rep. TR-13012, 2013.

[22] X. Lan and M. Schwager, "Planning Periodic Persistent Monitoring Trajectories for Sensing Robots in Gaussian Random Fields," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2013.

[23] G. Hollinger and G. Sukhatme, "Sampling-based Motion Planning for Robotic Information Gathering," in *Proc. Robotics: Science and Systems (RSS)*, 2013.

[24] B. Charrow, V. Kumar, and N. Michael, "Approximate Representations for Multi-Robot Control Policies that Maximize Mutual Information," in *Proc. Robotics: Science and Systems (RSS)*, 2013.

[25] G. Hoffmann and C. Tomlin, "Mobile Sensor Network Control Using Mutual Information Methods and Particle Filters," *IEEE Trans. on Automatic Control*, vol. 55, no. 1, 2010.

[26] S. LaValle, "Sensing and Filtering: A Fresh Perspective Based on Preimages and Information Spaces," *Foundations and Trends in Robotics*, vol. 1, no. 4, 2012.

[27] C. Freundlich, P. Mordohai, and M. Zavlanos, "A Hybrid Control Approach to the Next-Best-View Problem using Stereo Vision," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2013.

[28] M. Athans, "On the Determination of Optimal Costly Measurement Strategies for Linear Stochastic Systems," *Automatica*, vol. 8, no. 4, 1972.

[29] L. Meier, J. Peschon, and R. Dressler, "Optimal Control of Measurement Subsystems," *IEEE Trans. on Automatic Control*, vol. 12, no. 5, 1967.

[30] MATLAB, *Robust Control Toolbox*. Natick, Massachusetts, United States: The MathWorks Inc., 2012.