2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

An Optimal Control Approach for the Persistent Monitoring Problem

Christos G. Cassandras, Xu Chu Ding, Xuchao Lin

Abstract— **We propose an optimal control framework for persistent monitoring problems where the objective is to control the movement of mobile agents to minimize an uncertainty metric in a given mission space. For a single agent in a one-dimensional space, we show that the optimal solution is obtained in terms of a sequence of switching locations, thus reducing it to a parametric optimization problem. Using Infinitesimal Perturbation Analysis (IPA) we obtain a complete solution through a gradient-based algorithm. We illustrate our approach with numerical examples.**

I. INTRODUCTION

Enabled by recent technological advances, the deployment of autonomous agents that can cooperatively perform complex tasks is rapidly becoming a reality. In particular, there has been considerable progress reported in the literature on sensor networks that can carry out coverage control [1]–[3], surveillance [4], [5] and environmental sampling [6], [7] missions. In this paper, we are interested in generating optimal control strategies for *persistent monitoring* tasks; these arise when agents must monitor a dynamically changing environment which cannot be fully covered by a stationary team of available agents. Persistent monitoring differs from traditional coverage tasks due to the perpetual need to cover a changing environment, i.e., all areas of the mission space must be visited infinitely often. The main challenge in designing control strategies in this case is in balancing the presence of agents in the changing environment so that it is optimally covered over time while still satisfying sensing and motion constraints. Examples of persistent monitoring missions include surveillance in a museum to prevent unexpected events or thefts, unmanned vehicles for border patrol missions, and environmental applications where routine sampling of an area is involved.

In this paper, we address the persistent monitoring problem through an optimal control framework to drive agents so as to minimize a metric of uncertainty over the environment. In coverage control [2], [3], it is common to model knowledge of the environment as a non-negative density function defined over the mission space, and usually assumed to be fixed over time. However, since persistent monitoring tasks involve dynamically changing environments, it is natural to extend it to a function of both space and time to model uncertainty in the environment. We assume that uncertainty at a point grows in time if it is not covered by any agent sensors; for simplicity, we assume this growth is linear. To model sensor coverage, we define a probability of detecting events at each point of the mission space by agent sensors. Thus, the uncertainty of the environment decreases (for simplicity, linearly) with a rate proportional to the event detection probability, i.e., the higher the sensing effectiveness is, the faster the uncertainty is reduced.

While it is desirable to track the value of uncertainty over all points in the environment, this is generally infeasible due to computational complexity and memory constraints. Motivated by polling models in queueing theory, e.g., spatial queueing [8], and by stochastic flow models [9], we assign sampling points of the environment to be monitored persistently (equivalently, we partition the environment into a discrete set of regions). We associate to these points "uncertainty queues" which are visited by one or more servers. The growth in uncertainty at a sampling point can then be viewed as a flow into a queue, and the reduction in uncertainty (when covered by an agent) can be viewed as the queue being visited by mobile servers as in a polling system. Moreover, the service flow rates depend on the distance of the sampling point to nearby agents. From this point of view, we aim to control the movement of the servers (agents) so that the total accumulated "uncertainty queue" is minimized.

Control and motion planning for agents performing persistent monitoring tasks have been studied in the literature. In [1] the focus is on sweep coverage problems, where agents are controlled to sweep an area. In [10], [11] a similar metric of uncertainty is used to model knowledge of a dynamic environment. In [11], the sampling points in a 1-D environment are denoted as cells, and the optimal control policy for a two-cell problem is given. Problems with more than two cells are addressed by a heuristic policy. In [10], the authors proposed a stabilizing speed controller for a single agent so that the accumulated uncertainty over a set of points along a given path in the environment is bounded, and an optimal controller that minimizes the maximum steady-state uncertainty over points of interest, assuming that the agent travels along a closed path and does not change direction. The persistent monitoring problem is also related to robot patrol problems, where a team of robots are required to visit points in the workspace with frequency constraints [12], [13].

Our ultimate goal is to optimally control a team of cooperating agents in a 2 or 3-D environment. The contribution of this paper is to take a first step toward this goal by formulating and solving an optimal control problem for one agent moving in a 1-D mission space in which we minimize the accumulated uncertainty over a given time horizon and over an arbitrary number of sampling points. Even in this simple case, determining a complete explicit solution is

This work was supported in part by NSF under Grant EFRI-0735974, by AFOSR under grants FA9550-07-1-0361 and FA9550-09-1-0095, by DOE under grant DE-FG52-06NA27490, and by ONR under grant N00014-09-1-1051.

The authors are with the Division of Systems Engineering and Center for Information and Systems Engineering, Boston University, Boston, MA; e-mail:{*cgc,xcding,mmxclin*}@*bu.edu*.

computationally hard. However, we show that the optimal trajectory of the agent is to oscillate in the mission space: move at full speed, then switch direction before reaching either end point. Thus, we show that the solution is reduced to a parametric optimization problem over the switching locations. We then use generalized Infinitesimal Perturbation Analysis (IPA) [14] to determine these optimal switching locations, which fully characterize the optimal control for the agent. This establishes the basis for extending this approach, first to multiple agents and then to a 2-D mission space. It also provides insights that motivate a receding horizon approach, bypassing the computational complexity for real-time control. These next steps are the subjects of ongoing research.

II. Persistent Monitoring Problem Formulation

We consider a mobile agent in a 1-D mission space of length L. Let the position of the agent be $s(t) \in [0,L]$ with dynamics:

$$\dot{s}(t) = u(t), \quad s(0) = 0 \tag{1}$$

i.e., we assume that the agent can control its direction and speed. We assume that the speed is constrained by $|u(t)| \leq 1$.

We associate with every point $x \in [0,L]$ a function $p(x,s)$ at state $s(t)$ that captures the probability of detecting an event at this point. We assume that $p(x,s) = 1$ if $x = s$, and that $p(x,s)$ decays when the distance between x and s (i.e., $|x-s|$) increases. Assuming a finite sensing range r, we set $p(x,s) = 0$ when $|x-s| > r$. In this paper, we use a linear decay model shown below as our event detection probability function:

$$p(x,s) = \begin{cases} 1 - \frac{|x-s|}{r} & \text{if } |x-s| \leq r \\ 0 & \text{if } |x-s| > r \end{cases} \tag{2}$$

We consider a set of points $\{\alpha_i\}$, $i = 1,\ldots,M$, $\alpha_i \in [0,L]$, and associate a time-varying measure of uncertainty with each point α_i, which we denote by $R_i(t)$. Without loss of generality, we assume $0 \leq \alpha_1 \leq \cdots \leq \alpha_M \leq L$ and, to simplify notation, we set $p_i(s(t)) \equiv p(\alpha_i, s(t))$. This set may be selected to contain points of interest in the environment, or sampled points from the mission space. Alternatively, we may consider a partition of $[0,L]$ into M intervals whose center points are $\alpha_i = (2i-1)L/2M$, $i = 1,\ldots,M$. We can then set $p(x,s) = p_i(s)$ for all $x \in [\alpha_i - \frac{L}{2M}, \alpha_i + \frac{L}{2M}]$. The uncertainty functions $R_i(t)$ are defined to have the following properties: (*i*) $R_i(t)$ increases with a fixed rate dependent on α_i, if $p_i(s(t)) = 0$, (*ii*) $R_i(t)$ decreases with a fixed rate if $p_i(s(t)) = 1$, and (*iii*) $R_i(t) \geq 0$ for all t. It is then natural to model uncertainty so that its decrease is proportional to the probability of detection. In particular, we model the dynamics of $R_i(t)$, $i = 1,\ldots,M$, as follows:

$$\dot{R}_i(t) = \begin{cases} 0 & \text{if } R_i(t) = 0,\ A_i < Bp_i(s(t)) \\ A_i - Bp_i(s(t)) & \text{otherwise} \end{cases} \tag{3}$$

where we assume that initial conditions $R_i(0)$, $i = 1,\ldots,M$, are given and that $B > A_i > 0$ for all i (thus, the uncertainty strictly decreases when $s(t) = \alpha_i$).

Viewing persistent monitoring as a polling system, each point α_i (equivalently, ith interval in $[0,L]$) is associated with a "virtual queue" where uncertainty accumulates with inflow rate A_i. The service rate of this queue is time-varying and given by $Bp_i(s(t))$, controllable through the agent position at time t, as shown in Fig. 1. This interpretation is convenient for characterizing the *stability* of this system: For each queue, we may require that $A_i < \frac{1}{T}\int_0^T Bp_i(s(t))dt$. Alternatively, we may require that each queue becomes empty at least once over $[0,T]$. We may also impose conditions such as $R_i(T) \leq R_{\max}$ for each queue as additional constraints for our problem so as to provide bounded uncertainty guarantees, although we will not do so in this paper. Note that this analogy readily extends to multi-agent and 2 or 3-D settings. Also, note that B can also be made location dependent without affect the analysis in this paper.

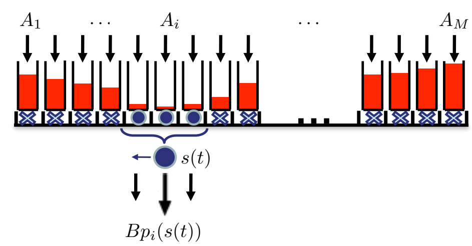

Fig. 1. A queueing system analog of the persistent monitoring problem.

The goal of the optimal persistent monitoring problem we consider is to control the mobile agent direction and speed $u(t)$ so that the cumulative uncertainty over all sensor points $\{\alpha_i\}$, $i = 1,\ldots,M$ is minimized over a fixed time horizon T. Thus, we aim to solve the following optimal control problem:

$$\textbf{Problem P1}: \quad \min_{u(t)} J = \frac{1}{T}\int_0^T \sum_{i=1}^M R_i(t)dt \tag{4}$$

subject to the agent dynamics (1), uncertainty dynamics (3), state constraint $0 \leq s(t) \leq L$, $t \in [0,T]$, and control constraint $|u(t)| \leq 1$, $t \in [0,T]$.

III. Optimal Control Solution

In this section we first characterize the optimal control solution of **Problem P1** and show that it is reduced to a parametric optimization problem. This allows us to utilize the IPA method [14] to find a complete optimal solution.

A. Hamiltonian analysis

We define the state vector $x(t) = [s(t), R_1(t), \ldots, R_M(t)]^{\mathrm{T}}$ and the associated costate vector $\lambda(t) = [\lambda_s(t), \lambda_1(t), \ldots, \lambda_M(t)]^{\mathrm{T}}$. In view of the discontinuity in the dynamics of $R_i(t)$ in (3), the optimal state trajectory may contain a boundary arc when $R_i(t) = 0$ for any i; otherwise, the state evolves in an interior arc. We first analyze the system operating in such an interior arc. Due to (1) and (3), the Hamiltonian is:

$$H(x,\lambda,u) = \sum_{i=1}^M R_i(t) + \lambda_s(t)u(t) + \sum_{i=1}^M \lambda_i(t)(A_i - Bp_i(s)) \tag{5}$$

and the costate equations $\dot{\lambda} = -\frac{\partial H}{\partial x}$ are:

$$\begin{aligned}\dot{\lambda}_s(t) &= -\frac{\partial H}{\partial s} = -B\sum_{i=1}^{M}\lambda_i(t)\frac{\partial p_i(s)}{\partial s}\\ &= -\frac{B}{r}\sum_{i\in F^-(t)}\lambda_i(t) + \frac{B}{r}\sum_{i\in F^+(t)}\lambda_i(t)\\ \dot{\lambda}_i(t) &= -\frac{\partial H}{\partial R_i} = -1, \qquad i = 1,\dots,M, \end{aligned} \tag{6}$$

where we have used (2), and the sets $F^-(t)$ and $F^+(t)$ are defined as:

$$\begin{cases} F^-(t) = \{i : s(t) - r \leq \alpha_i \leq s(t)\} \\ F^+(t) = \{i : s(t) < \alpha_i \leq s(t) + r\}, \end{cases}$$

so that they identify all points α_i within the agent's sensing range. Since we impose no terminal state constraints, the boundary conditions are $\lambda_s(T) = 0$ and $\lambda_i(T) = 0$, $i = 1,\dots,M$. Applying the Pontryagin minimum principle to (5) with $u^\star(t)$, $t \in [0,T)$, denoting an optimal control, we have

$$H(x^\star, \lambda^\star, u^\star) = \min_{u\in[-1,1]} H(x, \lambda, u)$$

and it is immediately obvious that it is necessary for an optimal control to satisfy:

$$u^\star(t) = \begin{cases} 1 & \text{if } \lambda_s(t) < 0 \\ -1 & \text{if } \lambda_s(t) > 0 \end{cases} \tag{7}$$

This condition excludes the case where $\lambda_s(t) = 0$ over some finite "singular intervals" [15]. It turns out this can arise only in some pathological cases which we shall not discuss in this paper.

The implication of (6) with $\lambda_i(T) = 0$ is that $\lambda_i(t) = T - t$ for all $t \in [0,T]$ and all $i = 1,\dots,M$ and $\lambda_i(t)$ is monotonically decreasing starting with $\lambda_i(0) = T$. However, this is only true if the entire optimal trajectory is an interior arc, i.e., the state constraints remain inactive. On the other hand, looking at (6), observe that when the two end points, 0 and L, are not within the range of the agent, we have $|F^-(t)| = |F^+(t)|$, since the number of indices i satisfying $s(t) - r \leq \alpha_i \leq s(t)$ is the same as that satisfying $s(t) < \alpha_i \leq s(t) + r$. Consequently, $\dot{\lambda}_s(t) = 0$, i.e., $\lambda_s(t)$ remains constant as long as this condition is satisfied and, in addition, none of the state constraints $R_i(t) \geq 0$, $i = 1,\dots,M$, is active.

Thus, as long as the optimal trajectory is an interior arc and $\lambda_s(t) < 0$, the agent moves at maximal speed $u^\star(t) = 1$ in the positive direction towards the point $s = L$. If $\lambda_s(t)$ switches sign before any of the state constraints $R_i(t) \geq 0$, $i = 1,\dots,M$, becomes active or the agent reaches the end point $s = L$, then $u^\star(t) = -1$ and the agent reverses its direction or, possibly, comes to rest. In what follows, we examine the effect of the state constraints and will establish the fact that the complete solution of this problem boils down to determining a set of switching locations over $(0,L)$ with the end points being infeasible on an optimal trajectory.

The dynamics in (3) indicate a discontinuity arising when the condition $R_i(t) = 0$ is satisfied while $\dot{R}_i(t) = A_i - Bp_i(s(t)) < 0$ for some $i = 1,\dots,M$. Thus, $R_i = 0$ defines an interior boundary condition which is not an explicit function of time. Following standard optimal control analysis [15], if this condition is satisfied at time t for some $k \in \{1,\dots,M\}$,

$$H\left(x(t^-), \lambda(t^-), u(t^-)\right) = H\left(x(t^+), \lambda(t^+), u(t^+)\right) \tag{8}$$

where we note that one can make a choice of setting the Hamiltonian to be continuous at the entry point of a boundary arc or at the exit point. Using (5) and (3), (8) implies:

$$\lambda_s^\star(t^-)\, u^\star(t^-) + \lambda_k^\star(t^-)\left(A_i - Bp_k(s(t^-))\right) = \lambda_s^\star(t^+)\, u^\star(t^+) \tag{9}$$

In addition, $\lambda_s^\star(t^-) = \lambda_s^\star(t^+)$ and $\lambda_i^\star(t^-) = \lambda_i^\star(t^+)$ for all $i \neq k$, but $\lambda_k^\star(t)$ may experience a discontinuity so that:

$$\lambda_k^\star(t^-) = \lambda_k^\star(t^+) - \pi_k \tag{10}$$

where $\pi_k \geq 0$. Recalling (7), since $\lambda_s^\star(t)$ remains unaffected, so does the optimal control, i.e., $u^\star(t^-) = u^\star(t^+)$. Moreover, since this is an entry point of a boundary arc, it follows from (3) that $\dot{R}_k(t^-) = A_i - Bp_k(s(t^-)) < 0$. Therefore, (9) and (10) imply that

$$\lambda_k^\star(t^-) = 0, \quad \lambda_k^\star(t^+) = \pi_k \geq 0.$$

The actual evaluation of the costate vector over the interval $[0,T]$ requires solving (6), which in turn involves the determination of all points where the state variables $R_i(t)$ reach their minimum feasible values $R_i(t) = 0$, $i = 1,\dots,M$. This generally involves the solution of a two-point-boundary-value problem. However, our analysis thus far has already established the structure of the optimal control (7) which we have seen to remain unaffected by the presence of boundary arcs where $R_i(t) = 0$ for one or more $i = 1,\dots,M$. Let us now turn our attention to the constraints $s(t) \geq 0$ and $s(t) \leq L$. The following proposition asserts that neither of these can become active on an optimal trajectory.

Proposition 1: On an optimal trajectory, $s^\star(t) \neq 0$ and $s^\star(t) \neq L$ for all $t \in (0,T]$.

Proof: See [16]. ■

Based on this analysis, the optimal control in (7) depends entirely on the points where $\lambda_s(t)$ switches sign and, in light of Prop. 1, the solution of the problem reduces to the determination of a parameter vector $\theta = [\theta_1,\dots,\theta_N]^{\mathrm{T}}$, where $\theta_j \in (0,L)$ denotes the jth location where the optimal control changes sign. Note that N is generally not known a priori and depends on the time horizon T.

Since $s(0) = 0$, from Prop. 1 we have $u^\star(0) = 1$, thus θ_1 corresponds to the optimal control switching from 1 to -1. Furthermore, θ_j, j odd, always correspond to $u^\star(t)$ switching from 1 to -1, and vice versa if j is even. Thus, we have the following constraints on the switching locations for all $j = 2,\dots,N$:

$$\begin{cases} \theta_j \leq \theta_{j-1}, & \text{if } j \text{ is even} \\ \theta_j \geq \theta_{j-1}, & \text{if } j \text{ is odd.} \end{cases} \tag{11}$$

It is now clear that the behavior of the agent under the optimal control policy (7) is that of a hybrid system whose dynamics undergo switches when $u^\star(t)$ changes between 1 and -1 or when $R_i(t)$ reaches or leaves the boundary value $R_i = 0$. As a result, we are faced with a parametric optimization problem for a system with hybrid dynamics.

This is a setting where one can apply the generalized theory of Infinitesimal Perturbation Analysis (IPA) in [14] to obtain the gradient of the objective function J in (4) with respect to the vector θ and, therefore, determine an optimal vector $\theta^\star$ through a gradient-based optimization approach.

Remark 1: If the agent dynamics are replaced by a model such as $\dot{s}(t) = g(s) + bu(t)$, observe that (7) still holds, as does Prop. 1. The only difference lies in (6) which would involve a dependence on $\frac{dg(s)}{ds}$ and further complicate the associated two-point-boundary-value problem. However, since the optimal solution is also defined by a parameter vector $\theta = [\theta_1, \ldots, \theta_N]^{\mathrm{T}}$, we can still apply the IPA approach presented in the next section.

B. Infinitesimal Perturbation Analysis (IPA)

Our analysis has shown that, for an optimal trajectory, the agent always moves at full speed and never reaches either boundary point, i.e., $0 < s^\star(t) < L$ (excluding certain pathological cases as mentioned earlier.) Thus, the agent's movement can be parametrized through $\theta = [\theta_1, \ldots, \theta_N]^{\mathrm{T}}$ where θ_i is the ith control switching point and the solution of **Problem P1** reduces to the determination of an optimal parameter vector $\theta^\star$. As we pointed out, the agent's behavior on an optimal trajectory defines a hybrid system, and the switching locations translate to switching times between particular modes of the hybrid system. To describe an IPA treatment of the problem, we first present the hybrid automaton model corresponding to the system operating on an optimal trajectory.

Hybrid automaton model. We use a standard definition of a hybrid automaton (e.g., see [17]) and show in Fig. 2 a partial model of the system limited to the behavior of the agent with respect to a single $\alpha_i, i \in \{1, \ldots, M\}$. The model consists of 14 discrete states (modes) and is symmetric in the sense that states $1-7$ correspond to the agent operating with $u(t) = 1$, and states $8-14$ correspond to the agent operating with $u(t) = -1$. The events that cause state transitions can be placed in three categories: (i) The value of $R_i(t)$ becomes 0 and triggers a switch in the dynamics of (3). This can only happen when $R_i(t) > 0$ and $\dot{R}_i(t) = A_i - Bp_i(s(t)) < 0$ (e.g., in states 3 and 4), causing a transition to state 7 in which the invariant condition is $R_i(t) = 0$. (ii) The agent reaches a switching location, indicated by the guard condition $s(t) = \theta_j$ for any $j = 1, \ldots, N$. In these cases, a transition results from a state q to $q+7$ if $q = 1, \ldots, 6$ and to $q-7$ otherwise. (iii) The agent position reaches one of several critical values that affect the dynamics of $R_i(t)$ while $R_i(t) > 0$. When $s(t) = \alpha_i - r$, the value of $p_i(s(t))$ becomes strictly positive and $\dot{R}_i(t) = A_i - Bp_i(s(t)) > 0$, as in the transition $1 \to 2$. Subsequently, when $s(t) = \alpha_i - r(1 - A_i/B)$, as in the transition $2 \to 3$, the value of $p_i(s(t))$ becomes sufficiently large to cause $\dot{R}_i(t) = A_i - Bp_i(s(t)) < 0$ so that a transition due to $R_i(t) = 0$ becomes feasible at this state. Similar transitions occur when $s(t) = \alpha_i$, $s(t) = \alpha_i + r(1 - A_i/B)$, and $s(t) = \alpha_i + r$. The latter results in state 6 where $\dot{R}_i(t) = A_i > 0$ and the only feasible event is $s(t) = \theta_j$, j odd, when a switch must occur and a transition to state 13 takes place (similarly for state 8).

Before proceeding, we provide a brief review of the IPA framework for general stochastic hybrid systems as presented in [14]. The purpose of IPA is to study the behavior of a hybrid system state as a function of a parameter vector $\theta \in \Theta$ for a given compact, convex set $\Theta \subset \mathbb{R}^l$. Let $\{\tau_k(\theta)\}$, $k = 1, \ldots, K$, denote the occurrence times of all events in the state trajectory. For convenience, we set $\tau_0 = 0$ and $\tau_{K+1} = T$. Over an interval $[\tau_k(\theta), \tau_{k+1}(\theta))$, the system is at some mode during which the time-driven state satisfies $\dot{x} = f_k(x, \theta, t)$. An event at τ_k is classified as *Exogenous* if it causes a discrete state transition independent of θ and satisfies $\frac{d\tau_k}{d\theta} = 0$; otherwise, it is *Endogenous* and there exists a continuously differentiable function $g_k : \mathbb{R}^n \times \Theta \to \mathbb{R}$ such that $\tau_k = \min\{t > \tau_{k-1} : g_k(x(\theta, t), \theta) = 0\}$. Given $\theta = [\theta_1, \ldots, \theta_N]^{\mathrm{T}}$, we use the notation for Jacobian matrices: $x'(t) \equiv \frac{\partial x(\theta, t)}{\partial \theta}$, $\tau'_k \equiv \frac{\partial \tau_k(\theta)}{\partial \theta}$, $k = 1, \ldots, K$, for all state and event time derivatives. It is shown in [14] that $x'(t)$ satisfies:

$$\frac{d}{dt} x'(t) = \frac{\partial f_k(t)}{\partial x} x'(t) + \frac{\partial f_k(t)}{\partial \theta} \tag{12}$$

for $t \in [\tau_k, \tau_{k+1})$ with boundary condition:

$$x'(\tau_k^+) = x'(\tau_k^-) + \left[f_{k-1}(\tau_k^-) - f_k(\tau_k^+) \right] \tau'_k \tag{13}$$

for $k = 0, \ldots, K$. In addition, in (13), the gradient vector for each τ_k is $\tau_k = 0$ if the event at τ_k is exogenous and

$$\tau'_k = -\left[\frac{\partial g_k}{\partial x} f_k(\tau_k^-) \right]^{-1} \left(\frac{\partial g_k}{\partial \theta} + \frac{\partial g_k}{\partial x} x'(\tau_k^-) \right) \tag{14}$$

if the event at τ_k is endogenous and defined as long as $\frac{\partial g_k}{\partial x} f_k(\tau_k^-) \neq 0$.

IPA equations. To clarify the presentation, we first note that $i = 1, \ldots, M$ is used to index the points where uncertainty is measured; $j = 1, \ldots, N$ indexes the components of the parameter vector; and $k = 1, \ldots, K$ indexes event times. In order to apply the three fundamental IPA equations (12)-(14) to our system, we use the state vector $x(t) = [s(t), R_1(t), \ldots, R_M(t)]^{\mathrm{T}}$ and parameter vector $\theta = [\theta_1, \ldots, \theta_N]^{\mathrm{T}}$. We then identify all events that can occur in Fig. 2 and consider intervals $[\tau_k(\theta), \tau_{k+1}(\theta))$ over which the system is in one of the 14 states shown for each $i = 1, \ldots, M$. Applying (12) to $s(t)$ with $f_k(t) = 1$ or -1 due to (1) and (7), the solution yields the gradient vector $\nabla s(t) = [\frac{\partial s}{\partial \theta_1}(t), \ldots, \frac{\partial s}{\partial \theta_M}(t)]^{\mathrm{T}}$, where

$$\frac{\partial s}{\partial \theta_j}(t) = \frac{\partial s}{\partial \theta_j}(\tau_k^+), \text{ for } t \in [\tau_k, \tau_{k+1}) \tag{15}$$

for all $k = 1, \ldots, K$, i.e., for all states $q(t) \in \{1, \ldots, 14\}$. Similarly, let $\nabla R_i(t) = [\frac{\partial R_i}{\partial \theta_1}(t), \ldots, \frac{\partial R_i}{\partial \theta_M}(t)]^{\mathrm{T}}$ for $i = 1, \ldots, M$. We note from (3) that $f_k(t) = 0$ for states $q(t) \in Q_1 \equiv \{7, 14\}$; $f_k(t) = A_i$ for states $q(t) \in Q_2 \equiv \{1, 6, 8, 13\}$; and $f_k(t) = A_i - Bp_i(s(t))$ for all other states which we further classify into $Q_3 \equiv \{2, 3, 11, 12\}$ and $Q_4 \equiv \{4, 5, 9, 10\}$. Using (12)-(15) and omitting details (note that the full detail of this derivation can be found in [16]), we obtain the following

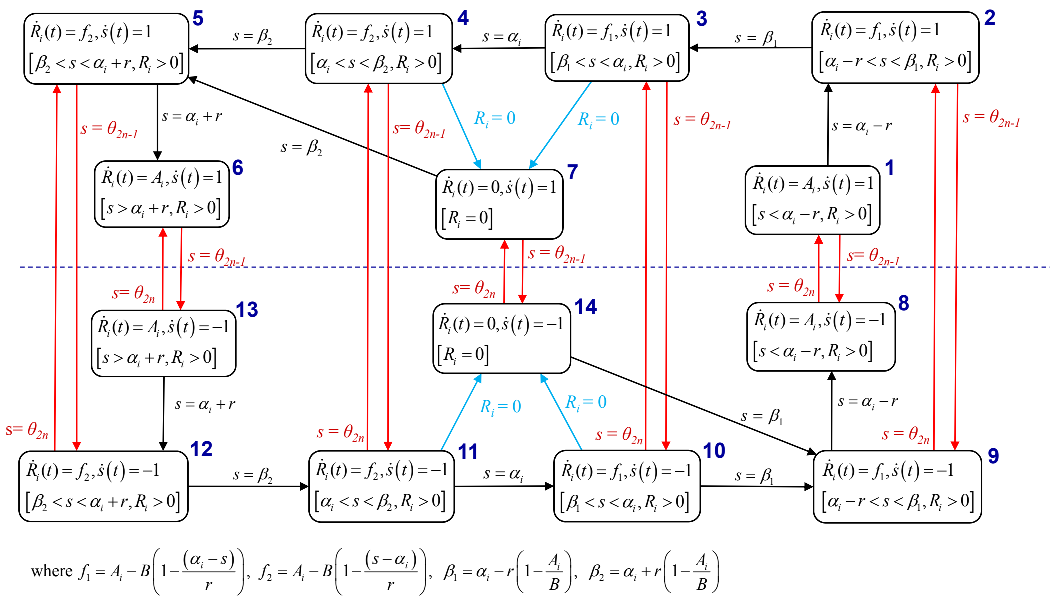

Fig. 2. Hybrid automaton for each α_i. Red arrows represent events when the control switches between 1 and -1. Blue arrows represent events when R_i becomes 0. Black arrows represent all other events.

expression for $\frac{\partial R_i}{\partial \theta_j}(t)$ for all $k \geq l$, $t \in [\tau_k, \tau_{k+1})$:

$$\frac{\partial R_i}{\partial \theta_j}(t) = \frac{\partial R_i}{\partial \theta_j}\left(\tau_k^+\right) \tag{16}$$

$$+\begin{cases} 0 & \text{if } q(t) \in Q_1 \cup Q_2 \\ (-1)^{j+1}\frac{2B}{r}u\left(\tau_k^+\right)\cdot(t-\tau_k) & \text{if } q(t) \in Q_3 \\ -(-1)^{j+1}\frac{2B}{r}u\left(\tau_k^+\right)\cdot(t-\tau_k) & \text{if } q(t) \in Q_4 \end{cases}$$

with boundary condition

$$\frac{\partial R_i}{\partial \theta_j}(\tau_k^+) = \begin{cases} 0 & \text{if } q\left(\tau_k^+\right) \in Q_1 \\ \frac{\partial R_i}{\partial \theta_j}(\tau_k^-) & \text{otherwise} \end{cases} \tag{17}$$

Objective Function Gradient Evaluation. Since we are ultimately interested in minimizing the objective function $J(\theta)$ (now a function of θ instead of u) in (4) with respect to θ, we first rewrite:

$$J(\theta) = \frac{1}{T}\sum_{i=1}^{M}\sum_{k=0}^{K}\int_{\tau_k(\theta)}^{\tau_{k+1}(\theta)} R_i(t,\theta)\,dt$$

where we have explicitly indicated the dependence on θ. We then obtain:

$$\nabla J(\theta)$$
$$= \frac{1}{T}\sum_{i=1}^{M}\sum_{k=0}^{N}\left(\int_{\tau_k}^{\tau_{k+1}} \nabla R_i(t)\,dt + R_i(\tau_{k+1})\nabla\tau_{k+1} - R_i(\tau_k)\nabla\tau_k\right)$$

With cancellations of $R_i(\tau_k)\nabla\tau_k$ for all k, we finally get

$$\nabla J(\theta) = \frac{1}{T}\sum_{i=1}^{M}\sum_{k=0}^{N}\int_{\tau_k}^{\tau_{k+1}} \nabla R_i(t)\,dt. \tag{18}$$

The evaluation of $\nabla J(\theta)$ therefore depends entirely on $\nabla R_i(t)$, which is obtained from (16)-(17) and the event times τ_k, $k=1,\ldots,K$, given initial conditions $s(0)=0$, $R_i(0)$ for $i=1,\ldots,M$ and $\nabla R_i(0)=0$.

Objective Function Optimization. We now seek to obtain $\theta^\star$ minimizing $J(\theta)$ through a standard gradient-based optimization scheme of the form

$$\theta^{l+1} = \theta^l - \eta_l \tilde{\nabla} J(\theta^l) \tag{19}$$

where $\{\eta_l\}$ is an appropriate step size sequence and $\tilde{\nabla}J(\theta)$ is the projection of the gradient $\nabla J(\theta)$ onto the feasible set (the set of θ satisfying the constraint (11)). The optimization scheme terminates when $|\tilde{\nabla}J(\theta)| < \varepsilon$ (for a fixed threshold ε) for some θ. Our IPA-based algorithm to obtain $\theta^\star$ minimizing $J(\theta)$ is summarized in Alg. 1 where we have adopted the Armijo step-size (see [18]) for $\{\eta_l\}$.

Algorithm 1 : IPA-based optimization algorithm to find $\theta^\star$

1: Set $N = \lfloor \frac{T}{L} \rfloor$ ($\lfloor\cdot\rfloor$ is the floor function), and set $\theta = [\theta_1,\ldots,\theta_N]^{\mathrm{T}}$ satisfying constraint (11)
2: **repeat**
3: Compute $s(t)$, $t \in [0,T]$ using θ
4: Compute $\tilde{\nabla}J(\theta)$ and update θ through (19)
5: **until** $|\tilde{\nabla}J(\theta)| < \varepsilon$
6: **if** θ satisfies Prop. 1 **then**
7: **Stop**, return θ as $\theta^\star$
8: **else**
9: Set $N+1 \to N$ and set $\theta_N = s(T)$
10: Go to Step 2
11: **end if**

Recalling that the dimension N of $\theta^\star$ is unknown (it depends on T), a distinctive feature of Alg. 1 is that we vary N by possibly increasing it after a vector θ locally minimizing J is obtained, if it does not satisfy the necessary optimality condition in Prop. 1. We start the search for a feasible N by setting it to $\lfloor \frac{T}{L} \rfloor$, the minimal N for which θ

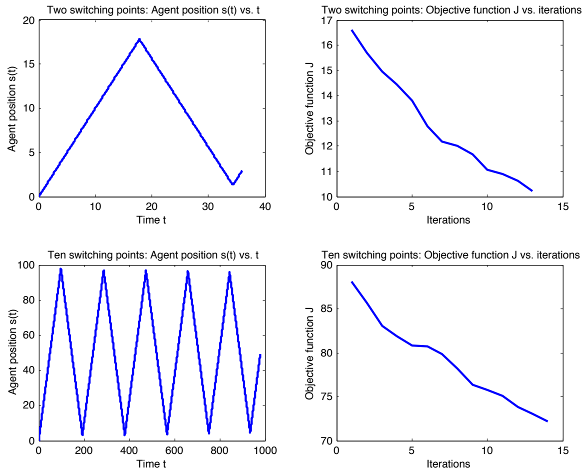

Fig. 3. Numerical results. Top figures correspond to $L = 20$, $T = 36$, 21 sampling points in $[0, L]$. Bottom figures correspond to $L = 100$, $T = 980$, 101 sampling points in $[0, L]$. Left plots: optimal trajectories. Right plots: J versus iterations.

can satisfy Prop. 1, and only need to increase N if the locally optimal θ vector violates Prop. 1. Due to non-convexity of the problem in terms of θ and N, it is possible to increase N further after Alg. 1 stops, and obtain a local optimal θ vector with a lower cost. In practice, this computation can take place in the background while the agent is in operation.

IV. Numerical results

In this section we present two numerical examples where we have used Alg. 1 to obtain an optimal persistent monitoring trajectory. The results are shown in Fig. 3. The top two figures correspond to an example with $L = 20$, $M = 21$, $\alpha_1 = 0$, $\alpha_M = 20$, and the remaining sampling points are evenly spaced between each other. Moreover, $A_i = 0.01$ for all i, $B = 3, r = 4, R_i(0) = 2$ for all i and $T = 36$. We start the algorithm with $\theta = [12]^\mathrm{T}$ and $\varepsilon = 2 \times 10^{-10}$. The algorithm stopped after 13 iterations (about 9 sec) using Armijo step-sizes, and the cost, J, was decreased from 16.63 to $J^\star = 10.24$ with $\theta^\star = [17.81, 1.29]^\mathrm{T}$, i.e., the dimension increased by 1. In the top-left, the optimal trajectory $s^\star(t)$ is plotted; in the top-right, J is plotted against iterations. We also increased N to 3 with initial $\theta = [12, 16, 4]$; Alg. 1 converged to a local minimum $J = 13.27 > J^\star = 10.24$ under $N = 2$.

The bottom two figures correspond to an example with $L = 100$, $M = 101$ and evenly spaced sampling points over $[0, L]$, $A_i = 0.01$ for all i, $B = 3$, $r = 4$, $R_i(0) = 2$ for all i and $T = 980$. We start the algorithm with $N = 9$, $\theta = [95, 95, 95, 95, 95, 5, 5, 5, 5]^\mathrm{T}$. The algorithm stopped after 14 iterations with the same ε (in about 10 min) using Armijo step-sizes, and J was decreased from 88.10 to $J^\star = 70.49$ with $\theta^\star = [98.03, 96.97, 96.65, 96.35, 95.70, 2.94, 3.21, 3.61, 4.08, 4.57]^\mathrm{T}$ where $N = 10$. Note that the cost is much higher in this case due to the larger number of sampling points. Moreover, none of the optimal switching locations is at 0 or L, consistent with Prop. 1.

V. Conclusions

We have formulated a persistent monitoring problem where we consider a dynamic environment with uncertainties at points changing depending on the proximity of the agent. We obtained an optimal control solution that minimizes the accumulated uncertainty over the environment, in the case of a single agent and 1-D mission space. The solution is characterized by a sequence of switching points, and we use an IPA-based gradient algorithm to compute the solution. Ongoing work aims at solving the problem with multiple agents and a richer dynamical model for each agent, as well as addressing the persistent monitoring problem in 2-D and 3-D mission spaces.

References

[1] I. Rekleitis, V. Lee-Shue, A. New, and H. Choset, "Limited communication, multi-robot team based coverage," in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, vol. 4. IEEE, 2004, pp. 3462–3468.

[2] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *Robotics and Automation, IEEE Transactions on*, vol. 20, no. 2, pp. 243–255, 2004.

[3] W. Li and C. Cassandras, "A cooperative receding horizon controller for multivehicle uncertain environments," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 242–257, 2006.

[4] A. Girard, A. Howell, and J. Hedrick, "Border patrol and surveillance missions using multiple unmanned air vehicles," in *43rd IEEE Conference on Decision and Control*, vol. 1. IEEE, 2005, pp. 620–625.

[5] B. Grocholsky, J. Keller, V. Kumar, and G. Pappas, "Cooperative air and ground surveillance," *IEEE Robotics & Automation Magazine*, vol. 13, no. 3, pp. 16–25, 2006.

[6] R. Smith, S. Mac Schwager, D. Rus, and G. Sukhatme, "Persistent ocean monitoring with underwater gliders: Towards accurate reconstruction of dynamic ocean processes," in *IEEE Conf. on Robotics and Automation, Shanghai, China (to appear)*, 2011.

[7] D. Paley, F. Zhang, and N. Leonard, "Cooperative control for ocean sampling: The glider coordinated control system," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 4, pp. 735–744, 2008.

[8] D. Bertsimas and G. Van Ryzin, "Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles," *Operations Research*, pp. 60–76, 1993.

[9] G. Sun, C. Cassandras, Y. Wardi, C. Panayiotou, and G. Riley, "Perturbation analysis and optimization of stochastic flow networks," *Automatic Control, IEEE Transactions on*, vol. 49, no. 12, pp. 2143–2159, 2004.

[10] M. S. S. L. Smith and D. Rus, "Persistent monitoring of dynamic environments using a robot with limited range sensing," in *IEEE Conf. on Robotics and Automation, Shanghai, China, May 2011*, 2011.

[11] N. Nigam and I. Kroo, "Persistent surveillance using multiple unmanned air vehicles," in *IEEE Aerospace Conference*. IEEE, 2008, pp. 1–14.

[12] P. Hokayem, D. Stipanovic, and M. Spong, "On persistent coverage control," in *Decision and Control, 2007 46th IEEE Conference on*. IEEE, 2008, pp. 6130–6135.

[13] Y. Elmaliach, N. Agmon, and G. Kaminka, "Multi-robot area patrol under frequency constraints," in *Robotics and Automation, 2007 IEEE International Conference on*. IEEE, 2007, pp. 385–390.

[14] C. Cassandras, Y. Wardi, C. Panayiotou, and C. Yao, "Perturbation analysis and optimization of stochastic hybrid systems," *European Journal of Control*, vol. 16, no. 6, pp. 642–664, 2010.

[15] A. Bryson and Y. Ho, *Applied optimal control*. Wiley New York, 1975.

[16] C. G. Cassandras, X. C. Ding, and X. Lin, "An optimal control approach for the persistent monitoring problem," Boston University, Tech. Rep., August 2011, available at http://arxiv.org/abs/1108.3221.

[17] C. Cassandras, J. Lygeros, and C. Press, *Stochastic hybrid systems*. CRC/Taylor & Francis, 2007.

[18] E. Polak, *Optimization: algorithms and consistent approximations*. Springer Verlag, 1997.