An Optimal Control Approach to the Multi-agent Persistent Monitoring Problem

Christos.G. Cassandras, Xuchao Lin and **Xu Chu Ding**
Division of Systems Engineering
and Center for Information and Systems Engineering
Boston University, cgc@bu.edu, mmxclin@bu.edu, xcding@bu.edu

Abstract—We present an optimal control framework for persistent monitoring problems where the objective is to control the movement of multiple cooperating agents to minimize an uncertainty metric in a given mission space. In a one-dimensional mission space, we show that the optimal solution is for each agent to move at maximal speed from one switching point to the next, possibly waiting some time at each point before reversing its direction. Thus, the solution is reduced to a simpler parametric optimization problem: determining a sequence of switching locations and associated waiting times at these switching points for each agent. This amounts to a hybrid system which we analyze using Infinitesimal Perturbation Analysis (IPA) to obtain a complete on-line solution through a gradient-based algorithm. We also show that the solution is robust with respect to the uncertainty model used. This establishes the basis for extending this approach to a two-dimensional mission space.

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 robotics and sensor networks regarding coverage control [1], [2], [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 covered over time optimally (in some well-defined sense) while still satisfying sensing and motion constraints. Examples of persistent monitoring missions include surveillance, patrol missions with unmanned vehicles, and environmental applications where routine sampling of an area is involved.

In this paper, we address the persistent monitoring problem by proposing 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 this model to a function of both space and time to capture uncertainty in the environment. We assume that uncertainty at a point grows in time if it is not covered by any agent sensors. 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 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],[9], and by stochastic flow models [10], we assign sampling points of the environment to be monitored persistently (this is equivalent to partitioning 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" content 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 [6], [11] a similar metric of uncertainty is used to model knowledge of a dynamic environment. In [11], the sampling points in a one-dimensional 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 [6], the authors proposed a stabilizing speed controller for a single agent so that the accumulated uncertainty over a given path in the environment is bounded, along with an optimal controller that minimizes the maximum steady-state uncertainty, 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], [14].

Although one-dimensional persistent monitoring problems are of interest in their own right (e.g., see [15]), our ultimate goal is to optimally control a team of cooperating agents in a two or three-dimensional environment. The contribution of this paper is to take a first step toward this goal by formulating and solving an optimal control problem for a team of agents moving in a one-dimensional mission space described by an interval $[0, L] \subset \mathbb{R}$ 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 computationally hard, as seen in [16] where the single-agent case was first considered. However, we show that the problem can be reduced to a *parametric* optimization problem. In particular, the optimal trajectory of each agent is to move at full speed until it reaches some switching point, dwell on the switching point for some time (possibly zero), and then switch directions. In addition, we prove that all agents should never reach the end points of the mission space $[0, L]$. Thus, each agent's optimal trajectory is fully described by a set of switching points $\{\theta_1, \ldots, \theta_K\}$ and associated waiting times at these points, $\{w_1, \ldots, w_K\}$. As a result, we show that the behavior of the agents operating under optimal control is described by a hybrid system. This allows us to make use of generalized Infinitesimal Perturbation Analysis (IPA), as presented in [17],[18], to determine gradients of the objective function with respect to these parameters and subsequently obtain optimal switching locations and waiting times that fully characterize an optimal solution. It also allows us to exploit robustness properties of IPA to extend this solution approach to a stochastic uncertainty model. Our analysis establishes the basis for extending this approach to a two-dimensional mission space (in ongoing research). In a broader context, our approach brings together optimal control, hybrid systems, and perturbation analysis techniques in solving a class of problems which, under optimal control, can be shown to behave like hybrid systems characterized by a set of parameters whose optimal values deliver a complete optimal control solution.

The rest of the paper is organized as follows. Section II formulates the optimal control problem. Section III characterizes the solution of the problem in terms of two parameter vectors specifying switching points in the mission space and associated dwelling times at them. Using IPA in conjunction with a gradient-based algorithm, a complete solution is also provided. Section IV provides some numerical results and Section V concludes the paper.

II. PERSISTENT MONITORING PROBLEM FORMULATION

We consider N mobile agents moving in a 1-dimensional mission space of length L, for simplicity taken to be an interval $[0, L] \subset \mathbb{R}$. Let the position of the agents at time t be $s_n(t) \in [0, L]$, $n = 1, \ldots, N$, following the dynamics:

$$\dot{s}_n(t) = u_n(t) \qquad (1)$$

i.e., we assume that the agent can control its direction and speed. Without loss of generality, after some rescaling with the size of the mission space L, we further assume that the speed is constrained by $|u_n(t)| \leq 1$, $n = 1, \ldots, N$. For the sake of generality, we include the additional constraint:

$$a \leq s(t) \leq b, \ \ a \geq 0, \ \ b \leq L \qquad (2)$$

over all t to allow for mission spaces where the agents may not reach the end points of $[0, L]$, possibly due to the presence of obstacles. We also point out that the agent dynamics in (1) can be replaced by a more general model of the form $\dot{s}_n(t) = g_n(s_n) + b_n u_n(t)$ without affecting the main results of our analysis (see also Remark 1 in the next section.) Finally, an additional constraint may be imposed if we assume that the agents are initially located so that $s_n(0) < s_{n+1}(0)$, $n = 1, \ldots, N-1$, and we wish to prevent them from subsequently crossing each other over all t:

$$s_n(t) - s_{n+1}(t) \leq 0 \qquad (3)$$

We associate with every point $x \in [0, L]$ a function $p_n(x, s_n)$ that measures the probability that an event at location x is detected by agent n. We also assume that $p_n(x, s_n) = 1$ if $x = s_n$, and that $p_n(x, s_n)$ is monotonically nonincreasing in the distance $|x - s_n|$ between x and s_n, thus capturing the reduced effectiveness of a sensor over its range which we consider to be finite and denoted by r_n (this is the same as the concept of "sensor footprint" found in the robotics literature.) Therefore, we set $p_n(x, s_n) = 0$ when $|x - s_n| > r_n$. Although our analysis is not affected by the precise sensing model $p_n(x, s_n)$, we will limit ourselves to a linear decay model as follows:

$$p_n(x, s_n) = \begin{cases} 1 - \frac{|x - s_n|}{r_n}, & \text{if } |x - s_n| \leq r_n \\ 0, & \text{if } |x - s_n| > r_n \end{cases} \qquad (4)$$

Next, 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 $p_{n,i}(s_n(t)) \equiv p_n(\alpha_i, s_n(t))$. 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 $\alpha_i = \frac{(2i-1)L}{2M}$, $i = 1, \ldots, M$. We can then set $p_n(x, s_n(t)) = p_{n,i}(s_n(t))$ for all $x \in [\alpha_i - \frac{L}{2M}, \alpha_i + \frac{L}{2M}]$. Therefore, the joint probability of detecting an event at location $x \in [\alpha_i - \frac{L}{2M}, \alpha_i + \frac{L}{2M}]$ by all the N agents simultaneously (assuming detection independence) is:

$$P_i(\mathbf{s}(t)) = 1 - \prod_{n=1}^{Q} [1 - p_{n,i}(s_n(t))] \qquad (5)$$

where we set $\mathbf{s}(t) = [s_1(t), \ldots, s_N(t)]^{\mathrm{T}}$. We define uncertainty functions $R_i(t)$ associated with the intervals $[\alpha_i - \frac{L}{2M}, \alpha_i + \frac{L}{2M}]$, $i = 1, \ldots, M$, so that they have the following properties: (i) $R_i(t)$ increases with a prespecified rate A_i if $P_i(\mathbf{s}(t)) = 0$, (ii) $R_i(t)$ decreases with a fixed rate B if $P_i(\mathbf{s}(t)) = 1$ 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
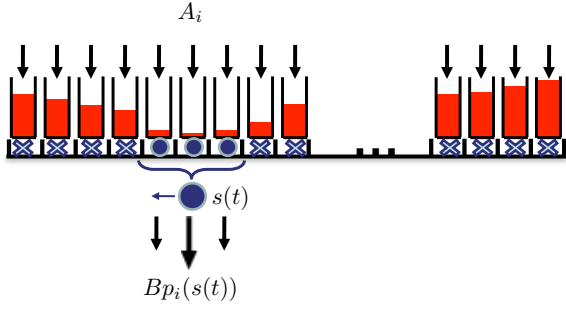
A_i

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

probability of detection. In particular, we model the dynamics of $R_i(t)$, $i = 1, \ldots, M$, as follows:

$$\dot{R}_i(t) = \begin{cases} 0 & \text{if } R_i(t) = 0, \ A_i \leq BP_i(\mathbf{s}(t)) \\ A_i - BP_i(\mathbf{s}(t)) & \text{otherwise} \end{cases}$$

(6)

where we assume that initial conditions $R_i(0)$, $i = 1, \ldots, M$, are given and that $B > A_i > 0$ (thus, the uncertainty strictly decreases when there is perfect sensing $P_i(\mathbf{s}(t)) = 1$.)

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 (similar models have been used in some data havesting problems, e.g., [19]). The service rate of this queue is time-varying and given by $BP_i(\mathbf{s}(t))$, controllable through the agent position at time t. Figure 1 illustrates this polling system when $N = 1$. This interpretation is convenient for characterizing the *stability* of such a system over a mission time T: For each queue, we may require that $\int_0^T A_i < \int_0^T Bp_i(s(t))dt$. 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 two or three-dimensional settings.

The goal of the optimal persistent monitoring problem we consider is to control the movement of the N agents through $u_n(t)$ in (1) so that the cumulative uncertainty over all sensing points $\{\alpha_i\}$, $i = 1, \ldots, M$ is minimized over a fixed time horizon T. Thus, setting $\mathbf{u}(t) = [u_1(t), \ldots, u_N(t)]$ we aim to solve the following optimal control problem **P1**:

$$\min_{\mathbf{u}(t)} \quad J = \frac{1}{T} \int_0^T \sum_{i=1}^M R_i(t) dt$$

(7)

subject to the agent dynamics (1), uncertainty dynamics (6), control constraint $|u_n(t)| \leq 1$, $t \in [0, T]$, and state constraints (2), $t \in [0, T]$. Note that we require $a \leq r_n$ and $b \geq L - r_m$, for at least some $n, m = 1, \ldots, N$; this is to ensure that there are no points in $[0, L]$ which can never be sensed, i.e., any i such that $\alpha_i < a - r_n$ or $\alpha_i > b + r_n$ would always lie outside any agent's sensing range. We will omit the additional constraint (3) from our initial analysis, but we will show that, when it is included, the optimal solution never allows it to be active.

III. Optimal Control Solution

We first characterize the optimal control solution of problem **P1** and show that it can be reduced to a parametric optimization problem. This allows us to utilize an Infinitesimal Perturbation Analysis (IPA) gradient estimation approach [17] to find a complete optimal solution through a gradient-based algorithm. We define the state vector $\mathbf{x}(t) = [s_1(t), \ldots, s_N(t), R_1(t), \ldots, R_M(t)]^\mathsf{T}$ and the associated costate vector $\lambda(t) = [\lambda_{s_1}(t), \ldots, \lambda_{s_N}(t), \lambda_1(t), \ldots, \lambda_M(t)]^\mathsf{T}$. In view of the discontinuity in the dynamics of $R_i(t)$ in (6), the optimal state trajectory may contain a boundary arc when $R_i(t) = 0$ for some i; otherwise, the state evolves in an interior arc. We first analyze the system operating in such an interior arc and omit the constraint (2) as well. Using (1) and (6), the Hamiltonian is

$$H(\mathbf{x}, \lambda, \mathbf{u}) = \sum_{i=1}^M R_i(t) + \sum_{n=1}^N \lambda_{s_n}(t) u_n(t) + \sum_{i=1}^M \lambda_i(t) \dot{R}_i(t)$$

(8)

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

$$\dot{\lambda}_i(t) = -\frac{\partial H}{\partial R_i(t)} = -1, \quad i = 1, \ldots, M$$

(9)

$$\dot{\lambda}_{s_n}(t) = -\frac{\partial H}{\partial s_n(t)}$$

$$= -\frac{B}{r_n} \sum_{i \in F_n^-(t)} \lambda_i(t) \prod_{d \neq n} [1 - p_{d,i}(s_d(t))]$$

$$+ \frac{B}{r_n} \sum_{i \in F_n^+(t)} \lambda_i(t) \prod_{d \neq n} [1 - p_{d,i}(s_d(t))]$$

(10)

where we have used (4), and the sets $F_n^-(t)$ and $F_n^+(t)$ are defined as

$$F_n^-(t) = \{i : s_n(t) - r_n \leq \alpha_i \leq s_n(t)\}$$
$$F_n^+(t) = \{i : s_n(t) < \alpha_i \leq s_n(t) + r_n\}$$

(11)

for $n = 1, \ldots, N$. Note that $F_n^-(t)$, $F_n^+(t)$ identify all points α_i to the left and right of $s_n(t)$ respectively that are within agent n's sensing range. Since we impose no terminal state constraints, the boundary conditions are $\lambda_i(T) = 0$, $i = 1, \ldots, M$ and $\lambda_{s_n}(T) = 0$, $n = 1, ..., N$. Applying the Pontryagin minimum principle to (8) with $\mathbf{u}^\star(t)$, $t \in [0, T)$, denoting an optimal control, we have

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

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

$$u_n^\star(t) = \begin{cases} 1 & \text{if } \lambda_{s_n}(t) < 0 \\ -1 & \text{if } \lambda_{s_n}(t) > 0 \end{cases}$$

(12)

This condition excludes the possibility that $\lambda_{s_n}(t) = 0$ over some finite singular intervals [20]. We will show that if $s_n(t) = a > 0$ or $s_n(t) = b < L$, then $\lambda_{s_n}(t) = 0$ for some $n \in \{1, \ldots, N\}$ may in fact be possible for some finite arc; otherwise $\lambda_{s_n}(t) = 0$ can arise only when $u_n(t) = 0$.

The implication of (9) with $\lambda_i(T) = 0$ is that $\lambda_i(t) = T - t$ for all $t \in [0, T]$ and all $i = 1, \ldots, M$ and that $\lambda_i(t)$ 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., all $R_i(t) \geq 0$ constraints for all $i = 1, \ldots, M$ remain inactive. On the other hand, looking at (10), observe that when the two end points, 0 and L, are not within the range of an agent, we have $|F_n^-(t)| = |F_n^+(t)|$, since the number of indices i satisfying $s_n(t) - r_n \leq \alpha_i \leq s_n(t)$ is the same as that satisfying $s_n(t) < \alpha_i \leq s_n(t) + r_n$. Consequently, for the one-agent case $N = 1$, (10) becomes

$$\dot{\lambda}_{s_1}(t) = -\frac{B}{r_1} \sum_{i \in F_1^-(t)} \lambda_i(t) + \frac{B}{r_1} \sum_{i \in F_1^+(t)} \lambda_i(t) \quad (13)$$

and $\dot{\lambda}_{s_1}(t) = 0$ since the two terms in (13) will cancel out, i.e., $\lambda_{s_1}(t)$ remains constant as long as this condition is satisfied and, in addition, none of the state constraints $R_i(t) \geq 0$, $i = 1, \ldots, M$, is active. Thus, for the one agent case, as long as the optimal trajectory is an interior arc and $\lambda_{s_1}(t) < 0$, the agent moves at maximal speed $u_1^\star(t) = 1$ in the positive direction towards the point $s_1 = b$. If $\lambda_{s_1}(t)$ switches sign before any of the state constraints $R_i(t) \geq 0$, $i = 1, \ldots, M$, becomes active or the agent reaches the end point $s_1 = b$, then $u_1^\star(t) = -1$ and the agent reverses its direction or, possibly, comes to rest.

In what follows, we examine the effect of the state constraints which significantly complicates the analysis, leading to a challenging two-point-boundary-value problem. However, we will establish the fact that the complete solution boils down to determining a set of switching locations over $[a, b]$ and waiting times at these switching points, with the end points, 0 and L, being always infeasible on an optimal trajectory. This is a much simpler problem that we are subsequently able to solve.

We begin by recalling that the dynamics in (6) indicate a discontinuity arising when the condition $R_i(t) = 0$ is satisfied while $\dot{R}_i(t) = A_i - BP_i(\mathbf{s}(t)) < 0$ for some $i = 1, \ldots, M$. Thus, $R_i = 0$ defines an interior boundary condition which is not an explicit function of time. Following standard optimal control analysis [20], if this condition is satisfied at time t for some $j \in \{1, \ldots, M\}$,

$$H\left(\mathbf{x}(t^-), \lambda(t^-), \mathbf{u}(t^-)\right) = H\left(\mathbf{x}(t^+), \lambda(t^+), \mathbf{u}(t^+)\right) \quad (14)$$

where we note that one can choose to set the Hamiltonian to be continuous at the entry point of a boundary arc or at the exit point. Using (8) and (6), (14) implies:

$$\sum_{n=1}^{N} \lambda_{s_n}^\star(t^-) u_n^\star(t^-) + \lambda_j^\star(t^-)[A_j(t) - BP_j(\mathbf{s}(t))]$$

$$= \sum_{n=1}^{N} \lambda_{s_n}^\star(t^+) u_n^\star(t^+) \quad (15)$$

In addition, $\lambda_{s_n}^\star(t^-) = \lambda_{s_n}^\star(t^+)$ for all $n = 1, \ldots, N$ and $\lambda_i^\star(t^-) = \lambda_i^\star(t^+)$ for all $i \neq j$, but $\lambda_j^\star(t)$ may experience a discontinuity so that:

$$\lambda_j^\star(t^-) = \lambda_j^\star(t^+) - \pi_j \quad (16)$$

where $\pi_j \geq 0$ is a multiplier associated with the constraint $-R_j(t) \leq 0$. Recalling (12), since $\lambda_{s_n}^\star(t)$ remains unaffected, so does the optimal control, i.e., $u_n^\star(t^-) = u_n^\star(t^+)$. Moreover,

since this is an entry point of a boundary arc, it follows from (6) that $A_j - BP_j(\mathbf{s}(t)) < 0$. Therefore, (15) and (16) imply that $\lambda_j^\star(t^-) = 0$ and $\lambda_j^\star(t^+) = \pi_j \geq 0$. Thus, $\lambda_i(t)$ always decreases with constant rate -1 until $R_i(t) = 0$ is active, at which point $\lambda_i(t)$ jumps to a non-negative value π_i and decreases with rate -1 again. The value of π_i is determined by how long it takes for the agents to reduce $R_i(t)$ to 0 once again. Obviously,

$$\lambda_i(t) \geq 0, \; i = 1, \ldots, M, \, t \in [0, T] \quad (17)$$

with equality holding only if $t = T$, or $t = t_0^-$ with $R_i(t_0) = 0$, $R_i(t') > 0$, where $t' \in [t_0 - \delta, t_0)$, $\delta > 0$. The actual evaluation of the costate vector over the interval $[0, T]$ requires solving (10), which in turn involves the determination of all points where the state variables $R_i(t)$ reach their minimum feasible values $R_i(t) = 0$, $i = 1, \ldots, M$. 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 (12) which we have seen to remain unaffected by the presence of boundary arcs when $R_i(t) = 0$ for one or more $i = 1, \ldots, M$. We will next prove some additional structural properties of an optimal trajectory, based on which we show that it is fully characterized by a set of non-negative scalar parameters. Determining the values of these parameters is a much simpler problem that does not require the solution of a two-point-boundary-value problem.

Let us turn our attention to the constraints $s_n(t) \geq a$ and $s_n(t) \leq b$ and consider first the case where $a = 0$, $b = L$, i.e., the agents can move over the entire $[0, L]$. We shall make use of the following technical condition:

Assumption 1: For any $n = 1, \ldots, N$, $i = 1, \ldots, M$, $t \in (0, T)$, and any $\epsilon > 0$, if $s_n(t) = 0$, $s_n(t - \epsilon) > 0$, then either $R_i(\tau) > 0$ for all $\tau \in [t - \epsilon, t]$ or $R_i(\tau) = 0$ for all $\tau \in [t - \epsilon, t]$; if $s_n(t) = L$, $s_n(t - \epsilon) < L$, then either $R_i(\tau) > 0$ for all $\tau \in [t - \epsilon, t]$ or $R_i(\tau) = 0$ for all $\tau \in [t - \epsilon, t]$.

This condition excludes the case where an agent reaches an endpoint of the mission space at the exact same time that any one of the uncertainty functions reaches its minimal value of zero. Then, the following proposition asserts that neither of the constraints $s_n(t) \geq 0$ and $s_n(t) \leq L$ can become active on an optimal trajectory. The assumption is used only in Proposition III.1 for technical reasons and does not alter the structure of the optimal controller.

Proposition III.1: Under Assumption 1, if $a = 0$, $b = L$, then on an optimal trajectory: $s_n^\star(t) \neq 0$ and $s_n^\star(t) \neq L$ for all $t \in (0, T)$, $n \in \{1, \ldots, N\}$.

Proof. Suppose at $t = t_0 < T$ an agent reaches the left endpoint, i.e., $s_n^\star(t_0) = 0$, $s_n^\star(t_0^-) > 0$. We will then establish a contradiction. Thus, assuming $s_n^\star(t_0) = 0$, we first show that $\lambda_{s_n}^\star(t_0^-) = 0$ by a contradiction argument. Assume that $\lambda_{s_n}^\star(t_0^-) \neq 0$, in which case, since the agent is moving toward $s_n = 0$, we have $u_n^\star(t_0^-) = -1$ and $\lambda_{s_n}^\star(t_0^-) > 0$ from (12). Then, $\lambda_{s_n}^\star(t)$ may experience a discontinuity so that

$$\lambda_{s_n}^\star(t_0^-) = \lambda_{s_n}^\star(t_0^+) - \pi_n \quad (18)$$

where $\pi_n \geqslant 0$ is a scalar constant. It follows that $\lambda_{s_n}^\star(t_0^+) = \lambda_{s_n}^\star(t_0^-) + \pi_n > 0$. Since the constraint $s_n(t) = 0$ is not an

explicit function of time, we have

$$\lambda_{s_n}^\star \left(t_0^-\right) u_n^\star \left(t_0^-\right) = \lambda_{s_n}^\star \left(t_0^+\right) u_n^\star \left(t_0^+\right) \qquad (19)$$

On the other hand, $u_n^\star \left(t_0^+\right) \geqslant 0$, since agent n must either come to rest or reverse its motion at $s_n = 0$, hence $\lambda_{s_n}^\star \left(t_0^+\right) u_n^\star \left(t_0^+\right) \geqslant 0$. This violates (19), since $\lambda_{s_n}^\star \left(t_0^-\right) u_n^\star \left(t_0^-\right) < 0$. This contradiction implies that $\lambda_{s_n}^\star \left(t_0^-\right) = 0$. Next, consider (10) and observe that in (11) we have $F_n^-(t_0) = \varnothing$, since $\alpha_i > s_n^\star (t_0) = 0$ for all $i = 1, \ldots, M$. Therefore, recalling (17), it follows from (10) that

$$\dot{\lambda}_{s_n} \left(t_0^-\right) = \frac{B}{r_n} \sum_{i \in F_n^+ \left(t_0^-\right)} \lambda_i \left(t_0^-\right) \prod_{d \neq n} \left[1 - p_{d,i}(s_d \left(t_0^-\right))\right] \geq 0$$

Under Assumption 1, there exists $\delta_1 > 0$ such that during the interval $(t_0 - \delta_1, t_0)$ no $R_i(t) \geq 0$ becomes active, hence no $\lambda_i(t)$ encounters a jump for $i = 1, \ldots, M$. It follows that $\lambda_i^\star(t) > 0$ for $i \in F_n^+(t)$ and $\dot{\lambda}_{s_n}^\star(t)$ is continuous with $\dot{\lambda}_{s_n}^\star(t) > 0$ for $t \in (t_0 - \delta_1, t_0)$. Again, since $s_n^\star(t_0) = 0$, there exists some $\delta_2 \leq \delta_1$ such that for $t \in (t_0 - \delta_2, t_0)$, we have $u_n^\star(t) < 0$ and $\lambda_{s_n}^\star(t) \geq 0$. Thus, for $t \in (t_0 - \delta_2, t_0)$, we have $\lambda_{s_n}^\star(t) \geq 0$ and $\dot{\lambda}_{s_n}^\star(t) > 0$. This contradicts the established fact that $\lambda_{s_n}^\star \left(t_0^-\right) = 0$ and we conclude that $s_n^\star(t) \neq 0$ for all $t \in [0, T]$, $n = 1, \ldots, N$. Using a similar line of argument, we can also show that $s_n^\star(t) \neq L$. ∎

Proposition III.2: If $a > 0$ and (or) $b < L$, then on an optimal trajectory there exist finite length intervals $[t_0, t_1]$ such that $s_n(t) = a$ and (or) $s_n(t) = b$, for some $n \in \{1, \ldots, N\}$, $t \in [t_0, t_1]$, $0 \leq t_0 < t_1 \leq T$.

Proof. Proceeding as in the proof of Proposition III.1, when $s_n^\star(t_0) = a$ we can establish (19) and the fact that $\lambda_{s_n}^\star \left(t_0^-\right) = 0$. On the other hand, $u_n^\star \left(t_0^+\right) \geqslant 0$, since the agent must either come to rest or reverse its motion at $s_n(t_0) = a$. In other words, when $s_n(t_0) = a$ on an optimal trajectory, (19) is satisfied either with the agent reversing its direction immediately (in which case $t_1 = t_0$ and $\lambda_{s_n}^\star \left(t_0^+\right) = 0$) or staying on the boundary arc for a finite time interval (in which case $t_1 > t_0$ and $u_n^\star(t) = 0$ for $t \in [t_0, t_1]$). The exact same argument can be applied to $s_n(t) = b$. ∎

The next result establishes the fact that on an optimal trajectory, every agent either moves at full speed or is at rest.

Proposition III.3: On an optimal trajectory, either $u_n^\star(t) = \pm 1$ if $\lambda_{s_n}^\star(t) \neq 0$, or $u_n^\star(t) = 0$ if $\lambda_{s_n}^\star(t) = 0$ for $t \in [0, T]$, $n = 1, \ldots, N$.

Proof. When $\lambda_{s_n}^\star(t) \neq 0$, we have shown in (12) that $u_n^\star(t) = \pm 1$, depending on the sign of $\lambda_{s_n}^\star(t)$. Thus, it remains to consider the case $\lambda_{s_n}^\star(t) = 0$ for some $t \in [t_1, t_2]$, where $0 \leq t_1 < t_2 \leq T$. Since the state is in a singular arc, $\lambda_{s_n}^\star(t)$ does not provide information about $u_n^\star(t)$. On the other hand, the Hamiltonian in (8) is not a explicit function of time, therefore, setting $H(\mathbf{x}^\star, \lambda^\star, \mathbf{u}^\star) \equiv H^\star$, we have $\frac{dH^\star}{dt} = 0$,

which gives

$$\frac{dH^\star}{dt} = \sum_{i=1}^M \dot{R}_i^\star(t) + \sum_{n=1}^N \dot{\lambda}_{s_n}^\star(t) u_n^\star(t)$$
$$+ \sum_{n=1}^N \lambda_{s_n}^\star(t) \dot{u}_n^\star(t) + \sum_{i=1}^M \dot{\lambda}_i^\star(t) \dot{R}_i^\star(t)$$
$$+ \sum_{i=1}^M \lambda_i^\star(t) \ddot{R}_i^\star(t) = 0 \qquad (20)$$

Define $S(t) = \{n | \lambda_{s_n}(t) = 0, n = 1, \ldots, N\}$ as the set of indices of agents that are in a singular arc and $\bar{S}(t) = \{n | \lambda_{s_n}(t) \neq 0, n = 1, \ldots, N\}$ as the set of indices of all other agents. Thus, $\lambda_{s_n}^\star(t) = 0$, $\dot{\lambda}_{s_n}^\star(t) = 0$ for $t \in [t_1, t_2]$, $n \in S(t)$. In addition, agents move with constant full speed, either 1 or -1, so that $\dot{u}_n^\star(t) = 0$, $n \in \bar{S}(t)$. Then, (20) becomes

$$\frac{dH^\star}{dt} = \sum_{i=1}^M [1 + \dot{\lambda}_i^\star(t)]\dot{R}_i^\star(t) + \sum_{n \in \bar{S}(t)} \dot{\lambda}_{s_n}^\star(t) u_n^\star(t)$$
$$+ \sum_{i=1}^M \lambda_i^\star(t) \ddot{R}_i^\star(t) = 0 \qquad (21)$$

From (9), $\dot{\lambda}_i^\star(t) = -1$, $i = 1, \ldots, M$, so $1 + \dot{\lambda}_i^\star(t) = 0$, leaving only the last two terms above. Note that $\dot{\lambda}_{s_n}^\star(t) = -\frac{\partial H^\star}{\partial s_n^\star(t)}$ and writing $\ddot{R}_i^\star(t) = \frac{d\dot{R}_i^\star(t)}{dt}$ we get:

$$-\sum_{n \in \bar{S}(t)} u_n^\star(t) \frac{\partial H^\star}{\partial s_n^\star(t)} + \sum_{i=1, R_i \neq 0}^M \lambda_i^\star(t) \frac{d\dot{R}_i^\star(t)}{dt} = 0$$

Recall from (6) that when $R_i(t) \neq 0$ we have $\dot{R}_i(t) = A_i - B[1 - \prod_{n=1}^N [1 - p_i(s_n(t))]]$, so that

$$\frac{\partial H^\star}{\partial s_n^\star(t)}$$
$$= -B \sum_{i=1, R_i \neq 0}^M \lambda_i^\star(t) \frac{\partial p_i(s_n^\star(t))}{\partial s_n^\star(t)} \prod_{d \neq n}^N (1 - p_i(s_d^\star(t))) \quad (22)$$

$$\frac{d\dot{R}_i^\star(t)}{dt}$$
$$= -B \sum_{n=1}^N u_n^\star(t) \frac{\partial p_i(s_n^\star(t))}{\partial s_n^\star(t)} \prod_{d \neq n}^N (1 - p_i(s_d^\star(t))) \quad (23)$$

which results in

$$B \sum_{\substack{i=1 \\ R_i \neq 0}}^{M} \lambda_i^\star(t) \sum_{n \in \bar{S}(t)} u_n^\star(t) \frac{\partial p_i(s_n^\star(t))}{\partial s_n^\star(t)} \prod_{d \neq n}^{N} (1 - p_i(s_d^\star(t)))$$

$$- B \sum_{\substack{i=1 \\ R_i \neq 0}}^{M} \lambda_i^\star(t) \sum_{n=1}^{N} u_n^\star(t) \frac{\partial p_i(s_n^\star(t))}{\partial s_n^\star(t)} \prod_{d \neq n}^{N} (1 - p_i(s_d^\star(t)))$$

$$= -B \sum_{\substack{i=1 \\ R_i \neq 0}}^{M} \lambda_i^\star(t) \sum_{n \in S(t)} u_n^\star(t) \frac{\partial p_i(s_n^\star(t))}{\partial s_n^\star(t)}$$

$$\times \prod_{d \neq n}^{N} (1 - p_i(s_d^\star(t))) = 0 \qquad (24)$$

Note that $\frac{\partial p_i(s_1^\star(t))}{\partial s_n^\star(t)} = \pm\frac{1}{r_1}$ or 0, depending on the relative position of $s_1^\star(t)$ with respect to α_i. Moreover, (24) is invariant to M or the precise way in which the mission space $[0, L]$ is partitioned, which implies that

$$\lambda_i^\star(t) \sum_{n \in S(t)} u_n^\star(t) \frac{\partial p_i(s_n^\star(t))}{\partial s_n^\star(t)} \prod_{d \neq n}^{N} (1 - p_i(s_d^\star(t))) = 0$$

for all $i = 1, \ldots, M$, $t \in [t_1, t_2]$. Since $\dot{\lambda}_i^\star(t) = -1$, $i = 1, \ldots, M$, it is clear that to satisfy this equality we must have $u_n^\star(t) = 0$ for all $t \in [t_1, t_2]$, $n \in S(t)$. In conclusion, in a singular arc with $\lambda_{s_n}^\star(t) = 0$ for some $n \in \{1, \ldots, N\}$, the optimal control is $u_n^\star(t) = 0$. ∎

Next, we consider the case where the additional state constraint (3) is included. We can then prove that this constraint is never active on an optimal trajectory, i.e., agents reverse their direction before making contact with any other agent.

Proposition III.4: If the constraint (3) is included in problem **P1**, then on an optimal trajectory, $s_n^\star(t) \neq s_{n+1}^\star(t)$ for $t \in (0, T]$, $n = 1, \ldots, N-1$.

Proof. Suppose at $t = t_0 < T$ we have $s_n^\star(t_0) = s_{n+1}^\star(t_0)$, for some $n = 1, \ldots, N-1$. We will then establish a contradiction. First assuming that both agents are moving (as opposed to one being at rest) toward each other, we have $u_n^\star(t_0^-) = 1$ and $u_{n+1}^\star(t_0^-) = -1$. From (12) and Prop III.3, we know $\lambda_{s_n}^\star(t_0^-) < 0$ and $\lambda_{s_{n+1}}^\star(t_0^-) > 0$. When the constraint $s_n(t) - s_{n+1}(t) \leq 0$ is active, $\lambda_{s_n}^\star(t)$ and $\lambda_{s_{n+1}}^\star(t_0^-)$ may experience a discontinuity so that

$$\lambda_{s_n}^\star(t_0^-) = \lambda_{s_n}^\star(t_0^+) + \pi$$
$$\lambda_{s_{n+1}}^\star(t_0^-) = \lambda_{s_{n+1}}^\star(t_0^+) - \pi \qquad (25)$$

where $\pi \geq 0$ is a scalar constant. It follows that $\lambda_{s_n}^\star(t_0^+) = \lambda_{s_n}^\star(t_0^-) - \pi < 0$ and $\lambda_{s_{n+1}}^\star(t_0^+) = \lambda_{s_{n+1}}^\star(t_0^-) + \pi > 0$. Since the constraint $s_n(t) - s_{n+1}(t) \leq 0$ is not an explicit function of time, we have

$$\lambda_{s_n}^\star(t_0^-) u_n^\star(t_0^-) + \lambda_{s_{n+1}}^\star(t_0^-) u_{n+1}^\star(t_0^-)$$
$$= \lambda_{s_n}^\star(t_0^+) u_n^\star(t_0^+) + \lambda_{s_{n+1}}^\star(t_0^+) u_{n+1}^\star(t_0^+) \qquad (26)$$

On the other hand, $u_n^\star(t_0^+) \leq 0$ and $u_{n+1}^\star(t_0^+) \geq 0$, since agents n and $n+1$ must either come to rest or reverse their motion after making contact, hence $\lambda_{s_n}^\star(t_0^+) u_n^\star(t_0^+) + \lambda_{s_{n+1}}^\star(t_0^+) u_{n+1}^\star(t_0^+) \geq 0$. This violates

(26), since $\lambda_{s_n}^\star(t_0^-) u_n^\star(t_0^-) + \lambda_{s_{n+1}}^\star(t_0^-) u_{n+1}^\star(t_0^-) < 0$. This contradiction implies that $s_n(t) - s_{n+1}(t) = 0$ cannot be active and we conclude that $s_n^\star(t) \neq s_{n+1}^\star(t)$ for $t \in [0, T]$, $n = 1, \ldots, N-1$. Moreover, if one of the two agents is at rest when $s_n^\star(t_0) = s_{n+1}^\star(t_0)$, the same argument still holds since it is still true that $\lambda_{s_n}^\star(t_0^-) u_n^\star(t_0^-) + \lambda_{s_{n+1}}^\star(t_0^-) u_{n+1}^\star(t_0^-) < 0$. ∎

Based on this analysis, the optimal control $u_n^\star(t)$ depends entirely on the sign of $\lambda_{s_n}^\star(t)$ and, in light of Propositions III.1-III.3, the solution of the problem reduces to determining: (i) *switching points* in $[0, L]$ where an agent switches from $u_n^\star(t) = \pm 1$ to either ∓ 1 or 0; or from $u_n^\star(t) = 0$ to either ± 1, and (ii) if an agent switches from $u_n^\star(t) = \pm 1$ to 0, *waiting times* until the agent switches back to a speed $u_n^\star(t) = \pm 1$. In other words, the full solution is characterized by two parameter vectors for each agent n: $\theta_n = [\theta_{n,1}, \ldots, \theta_{n,\Gamma_n}]^T$ and $w_n = [w_{n,1} \ldots, w_{n,\Gamma_n}]^T$, where $\theta_{n,\xi} \in (0, L)$ denotes the ξth location where agent n changes its speed from ± 1 to 0 and $w_{n,\xi} \geq 0$ denotes the time (which is possibly null) that agent n dwells on $\theta_{n,\xi}$. Note that Γ_n is generally not known a priori and depends on the time horizon T. In addition, we always assume that agent n reverses its direction after leaving the switching point $\theta_{n,\xi}$ with respect to the one it had when reaching $\theta_{n,\xi}$. This seemingly excludes the possibility of an agent's control following a sequence $1, 0, 1$ or $-1, 0, -1$. However, these two motion behaviors can be captured as two adjacent switching points approaching each other: when $|\theta_{n,\xi} - \theta_{n,\xi+1}| \to 0$, the agent control follows the sequence $1, 0, 1$ or $-1, 0, -1$, and the waiting time associated with $u_n^\star(t) = 0$ is $w_{n,\xi} + w_{n,\xi+1}$.

For simplicity, we will assume that $s_n(0) = 0$, so that it follows from Proposition III.1 that $u_n^\star(0) = 1$, $n = 1, \ldots, N$. Therefore, $\theta_{n,1}$ corresponds to the optimal control switching from 1 to 0. Furthermore, $\theta_{n,\xi}$ with ξ odd (even) always corresponds to $u_n^\star(t)$ switching from 1 to 0 (-1 to 0.) Thus, we have the following constraints on the switching locations for all $\xi = 2, \ldots, \Gamma_n$:

$$\begin{cases} \theta_{n,\xi} \leq \theta_{n,\xi-1}, & \text{if } \xi \text{ is even} \\ \theta_{n,\xi} \geq \theta_{n,\xi-1}, & \text{if } \xi \text{ is odd.} \end{cases} \qquad (27)$$

It is now clear that the behavior of each agent under the optimal control policy is that of a *hybrid system* whose dynamics undergo switches when $u_n^\star(t)$ changes from ± 1 to 0 and from 0 to ∓ 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 [17],[18] to conveniently obtain the gradient of the objective function J in (7) with respect to the vectors θ and w, and therefore, determine (generally, locally) optimal vectors $\theta^\star$ and $w^\star$ through a gradient-based optimization approach. Note that this is done on line, i.e., the gradient is evaluated by observing a trajectory with given θ and w over $[0, T]$ based on which θ and w are adjusted until convergence is attained using standard gradient-based algorithms.

Remark 1. If the agent dynamics in (1) are replaced by a model such as $\dot{s}_n(t) = g_n(s_n) + b_n u_n(t)$, observe that

(12) still holds. The difference lies in (10) which would involve a dependence on $\frac{dg_n(s_n)}{ds_n}$ and further complicate the associated two-point-boundary-value problem. However, since the optimal solution is also defined by parameter vectors $\theta_n = [\theta_{n,1}, \ldots, \theta_{n,\Gamma_n}]^{\mathrm{T}}$ and $w_n = [w_{n,1} \ldots, w_{n,\Gamma_n}]^{\mathrm{T}}$ for each agent n, we can still apply the IPA approach presented in the next section.

A. Infinitesimal Perturbation Analysis (IPA)

Our analysis thus far has shown that, on an optimal trajectory, the agent moves at full speed, dwells on a switching point (possibly for zero time) and never reaches either boundary point, i.e., $0 < s_n^\star(t) < L$. Thus, the nth agent's movement can be parameterized through $\theta_n = [\theta_{n,1}, \ldots, \theta_{n,\Gamma_n}]^{\mathrm{T}}$ and $w_n = [w_{n,1} \ldots, w_{n,\Gamma_n}]^{\mathrm{T}}$ where $\theta_{n,\xi}$ is the ξth control switching point and $w_{n,\xi}$ is the waiting time for this agent at the ξth switching point. Therefore, the solution of problem **P1** reduces to the determination of optimal parameter vectors $\theta_n^\star$ and $w_n^\star$, $n = 1, \ldots, N$. As we pointed out, the agent's optimal behavior defines a hybrid system, and the switching locations translate to switching times between particular modes of this system. This is similar to switching-time optimization problems, e.g., [21],[22],[23], except that we can only control a subset of mode switching times. We make use of IPA in part to exploit robustness properties that the resulting gradients possess [24]; specifically, we will show that they *do not depend on the uncertainty model parameters* A_i, $i = 1, \ldots, M$, and may therefore be used without any detailed knowledge of how uncertainty affects the mission space.

1) Single-agent solution with $a = 0$ and $b = L$: To maintain some notational simplicity, we begin with a single agent who can move on the entire mission space $[0, L]$ and will then provide the natural extension to multiple agents and a mission space limited to $[a, b] \subset [0, L]$. We present the associated hybrid automaton model for this single-agent system operating on an optimal trajectory. Our goal is to determine $\nabla J(\theta, w)$, the gradient of the objective function J in (7) with respect to θ and w, which can then be used in a gradient-based algorithm to obtain optimal parameter vectors $\theta_n^\star$ and $w_n^\star$, $n = 1, \ldots, N$. We will apply IPA, which provides a formal way to obtain state and event time derivatives with respect to parameters of hybrid systems, from which we can subsequently obtaining $\nabla J(\theta, w)$.

Hybrid automaton model. We use a standard definition of a hybrid automaton (e.g., see [25]) as the formalism to model the system described above. Thus, let $q \in Q$ (a countable set) denote the discrete state (or mode) and $x \in X \subseteq \mathbb{R}^n$ denote the continuous state. Let $\upsilon \in \Upsilon$ (a countable set) denote a discrete control input and $u \in U \subseteq \mathbb{R}^m$ a continuous control input. Similarly, let $\delta \in \Delta$ (a countable set) denote a discrete disturbance input and $d \in D \subseteq \mathbb{R}^p$ a continuous disturbance input. The state evolution is determined by means of (i) a vector field $f : Q \times X \times U \times D \to X$, (ii) an invariant (or domain) set $Inv : Q \times \Upsilon \times \Delta \to 2^X$, (iii) a guard set $Guard : Q \times Q \times \Upsilon \times \Delta \to 2^X$, and (iv) a reset function $r : Q \times Q \times X \times \Upsilon \times \Delta \to X$. The system remains at a discrete state q as long as the continuous (time-driven) state x does not leave the set $Inv(q, \upsilon, \delta)$. If x reaches a set $Guard(q, q', \upsilon, \delta)$ for some $q' \in Q$, a discrete transition can take place. If this transition does take place, the state instantaneously resets to (q', x') where x' is determined by the reset map $r(q, q', x, \upsilon, \delta)$. Changes in υ and δ are discrete events that either *enable* a transition from q to q' by making sure $x \in Guard(q, q', \upsilon, \delta)$ or *force* a transition out of q by making sure $x \notin Inv(q, \upsilon, \delta)$. We will classify all events that cause discrete state transitions in a manner that suits the purposes of IPA. Since our problem is set in a deterministic framework, δ and d will not be used.

We show in Fig. 2 a partial hybrid automaton model of the single-agent system where $a = 0$ and $b = L$. Since there is only one agent, we set $s(t) = s_1(t)$, $u(t) = u_1(t)$ and $\theta = \theta_1$ for simplicity. Due to the size of the overall model, Fig. 2 is limited to the behavior of the agent with respect to a single $\alpha_i, i \in \{1, \ldots, M\}$ and ignores modes where the agent dwells on the switching points (these, however, are included in our extended analysis in Section III-A2.) 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 $u(t) = -1$. States where $u(t) = 0$ are omitted since we do not include the waiting time parameter $w = w_1$ here. 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 (6). 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 $s(t) = \theta_\xi$ for any $\xi = 1, \ldots, \Gamma$. In these cases, a transition results from a state z to $z+7$ if $z = 1, \ldots, 6$ and to $z-7$ otherwise. (iii) The agent position reaches one of several critical values that affect the dynamics of $R_i(t)$ while $R_i(t) > 0$. Specifically, 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 $s(t) = \theta_\xi$, ξ odd, when a switch must occur and a transition to state 13 takes place (similarly for state 8).

IPA review. Before proceeding, we provide a brief review of the IPA framework for general stochastic hybrid systems as presented in [17]. 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 (i) *Exogenous* if it causes a discrete state transition independent of θ and satisfies $\frac{d\tau_k}{d\theta} = 0$; (ii) *Endogenous*, if there exists a continuously differentiable function $g_k : \mathbb{R}^n \times \Theta \to \mathbb{R}$ such that $\tau_k = \min\{t > \tau_{k-1} : g_k(x(\theta, t), \theta) = 0\}$; and (iii)

where $f_1 = A_i - B\left(1 - \frac{(\alpha_i - s)}{r}\right)$, $f_2 = A_i - B\left(1 - \frac{(s - \alpha_i)}{r}\right)$, $\beta_1 = \alpha_i - r\left(1 - \frac{K_1}{K_2}\right)$, $\beta_2 = \alpha_i + r\left(1 - \frac{K_1}{K_2}\right)$

Figure 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.

Induced if it is triggered by the occurrence of another event at time $\tau_m \leq \tau_k$. IPA specifies how changes in θ influence the state $x(\theta, t)$ and the event times $\tau_k(\theta)$ and, ultimately, how they influence interesting performance metrics which are generally expressed in terms of these variables.

Given $\theta = [\theta_1, \ldots, \theta_\Gamma]^{\mathrm{T}}$, we use the Jacobian matrix notation: $x'(t) \equiv \frac{\partial x(\theta, t)}{\partial \theta}$, $\tau_k' \equiv \frac{\partial \tau_k(\theta)}{\partial \theta}$, $k = 1, \ldots, K$, for all state and event time derivatives. It is shown in [17] that $x'(t)$ satisfies:

$$\frac{d}{dt} x'(t) = \frac{\partial f_k(t)}{\partial x} x'(t) + \frac{\partial f_k(t)}{\partial \theta} \qquad (28)$$

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

$$x'(\tau_k^+) = x'(\tau_k^-) + [f_{k-1}(\tau_k^-) - f_k(\tau_k^+)] \tau_k' \qquad (29)$$

for $k = 0, \ldots, K$. In addition, in (29), the gradient vector for each τ_k is $\tau_k' = 0$ if the event at τ_k is exogenous and

$$\tau_k' = -\left[\frac{\partial g_k}{\partial x} f_k(\tau_k^-)\right]^{-1} \left(\frac{\partial g_k}{\partial \theta} + \frac{\partial g_k}{\partial x} x'(\tau_k^-)\right) \qquad (30)$$

if the event at τ_k is endogenous (i.e., $g_k(x(\theta, \tau_k), \theta) = 0$), 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; $\xi = 1, \ldots, \Gamma$ indexes the components of the parameter vector; and $k = 1, \ldots, K$ indexes event times. In order to apply the three fundamental IPA equations (28)-(30) to our system, we use the state vector $x(t) = [s(t), R_1(t), \ldots, R_M(t)]^{\mathrm{T}}$ and parameter vector $\theta = [\theta_1, \ldots, \theta_\Gamma]^{\mathrm{T}}$. 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 (28) to $s(t)$ with $f_k(t) = 1$ or -1 due to (1) and (12), 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

$$\frac{\partial s}{\partial \theta_\xi}(t) = \frac{\partial s}{\partial \theta_\xi}(\tau_k^+), \text{ for } t \in [\tau_k, \tau_{k+1}) \qquad (31)$$

for all $k = 1, \ldots, K$, i.e., for all states $z(t) \in \{1, \ldots, 14\}$. 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 (6) that $f_k(t) = 0$ for states $z(t) \in Z_1 \equiv \{7, 14\}$; $f_k(t) = A_i$ for states $z(t) \in Z_2 \equiv \{1, 6, 8, 13\}$; and $f_k(t) = A_i - Bp_i(s(t))$ for all other states which we further classify into $Z_3 \equiv \{2, 3, 11, 12\}$ and $Z_4 \equiv \{4, 5, 9, 10\}$. Thus, solving (28) and using (31) gives:

$$\nabla R_i(t) = \nabla R_i(\tau_k^+)$$
$$- \begin{cases} 0 & \text{if } z(t) \in Z_1 \cup Z_2 \\ B\left(\frac{\partial p_i(s)}{\partial s}\right) \nabla s(\tau_k^+) \cdot (t - \tau_k) & \text{otherwise} \end{cases}$$
$$(32)$$

where $\frac{\partial p_i(s)}{\partial s} = \pm \frac{1}{r}$ as evaluated from (4) depending on the sign of $\alpha_i - s(t)$ at each associated automaton state. Details on the derivation of a simple recursive expression for $\nabla R_i(t)$ above can be found in Appendix A.

Objective Function Gradient Evaluation. Based on our analysis, the objective function (7) in problem **P1** can now be written as $J(\theta)$, a function of θ instead of $u(t)$ and we can

rewrite it as

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

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

$$\nabla J(\theta) = \frac{1}{T} \sum_{i=1}^{M} \sum_{k=0}^{K} \Big(\int_{\tau_k}^{\tau_{k+1}} \nabla R_i(t)\, dt$$
$$+ R_i(\tau_{k+1}) \nabla \tau_{k+1} - R_i(\tau_k) \nabla \tau_k \Big) \quad (33)$$

Observing the cancelation of all terms of the form $R_i(\tau_k)\nabla\tau_k$ for all k (with $\tau_0 = 0$, $\tau_{K+1} = T$ fixed), we finally get

$$\nabla J(\theta) = \frac{1}{T} \sum_{i=1}^{M} \sum_{k=0}^{K} \int_{\tau_k(\theta)}^{\tau_{k+1}(\theta)} \nabla R_i(t)\, dt. \quad (34)$$

The evaluation of $\nabla J(\theta)$ therefore depends entirely on $\nabla R_i(t)$, which is obtained from (47)-(48) in Appendix A and the observable event times τ_k, $k = 1, \ldots, K$, given initial conditions $s(0) = 0$, $R_i(0)$ for $i = 1, \ldots, M$ and $\nabla R_i(0) = 0$. Since $\nabla R_i(t)$ itself depends only on the event times τ_k, $k = 1, \ldots, K$, the gradient $\nabla J(\theta)$ is obtained by observing the switching times in a trajectory over $[0, T]$ characterized by the vector θ.

2) Multi-agent solution where $a \geq 0$ and $b \leq L$: Next, we extend the results obtained in the previous section to the general multi-agent problem where we also allow $a \geq 0$ and $b \leq L$. Recall that we require $0 \leq a \leq r_n$ and $L - r_m \leq b \leq L$, for at least some $n, m = 1, \ldots, N$ since, otherwise, controlling agent movement cannot affect $R_i(t)$ for all α_i located outside the sensing range of agents. We now include both parameter vectors $\theta_n = [\theta_{n,1}, \ldots, \theta_{n,\Gamma_n}]^{\mathrm{T}}$ and $w_n = [w_{n,1}, \ldots w_{n,\Gamma_n}]^{\mathrm{T}}$ for each agent n and, for notational simplicity, concatenate them to construct $\theta = [\theta_1, \ldots, \theta_N]^{\mathrm{T}}$ and $w = [w_1, \ldots, w_N]^{\mathrm{T}}$. The solution of problem **P1** reduces to the determination of optimal parameter vectors $\theta^\star$ and $w^\star$ and we will use IPA to evaluate $\nabla J(\theta, w) = [\frac{dJ(\theta, w)}{d\theta} \frac{dJ(\theta, w)}{dw}]^{\mathrm{T}}$. Similar to (34), it is clear that this depends on $\nabla R_i(t) = \left[\frac{\partial R_i(t)}{\partial \theta} \frac{\partial R_i(t)}{\partial w}\right]^{\mathrm{T}}$ and the event times τ_k, $k = 1, \ldots, K$, observed on a trajectory over $[0, T]$ with given θ and w.

IPA equations. We begin by recalling the dynamics of $R_i(t)$ in (6) which depend on the relative positions of all agents with respect to α_i and change at time instants τ_k such that either $R_i(\tau_k) = 0$ with $R_i(\tau_k^-) > 0$ or $A_i > BP_i(\mathbf{s}(\tau_k))$ with $R_i(\tau_k^-) = 0$. Moreover, using (1) and our earlier Hamiltonian analysis, the dynamics of $s_n(t)$, $n = 1, \ldots, N$, in an optimal trajectory can be expressed as follows. Define $\Theta_{n,\xi} = (\theta_{n,\xi-1}, \theta_{n,\xi})$ if ξ is odd and $\Theta_{n,\xi} = (\theta_{n,\xi}, \theta_{n,\xi-1})$ if ξ is even to be the ξth interval between successive switching points for any $n = 1, \ldots, N$, where $\theta_{n,0} = s_n(0)$. Then, for $\xi = 1, 2, \ldots,$

$$\dot{s}_n(t) = \begin{cases} 1 & s_n(t) \in \Theta_{n,\xi}, \ \xi \text{ odd} \\ -1 & s_n(t) \in \Theta_{n,\xi}, \ \xi \text{ even} \\ 0 & \text{otherwise} \end{cases} \quad (35)$$

where transitions for $s_n(t)$ from ± 1 to ∓ 1 are incorporated by treating them as cases where $w_{n,\xi} = 0$, i.e., no dwelling

at a switching point $\theta_{n,\xi}$ (in which case $\dot{s}_n(t) = 0$.) We can now concentrate on all events causing switches either in the dynamics of any $R_i(t)$, $i = 1, \ldots, M$, or the dynamics of any $s_n(t)$, $n = 1, \ldots, N$. From (29), any other event at some time τ_k in this hybrid system cannot modify the values of $\nabla R_i(t) = \left[\frac{\partial R_i(t)}{\partial \theta} \frac{\partial R_i(t)}{\partial w}\right]^{\mathrm{T}}$ or $\nabla s_n(t) = \left[\frac{\partial s_n(t)}{\partial \theta_n} \frac{\partial s_n(t)}{\partial w_n}\right]^{\mathrm{T}}$ at $t = \tau_k$.

First, applying (28) to $s_n(t)$ with $f_k(t) = 1$, -1 or 0 due to (35), the solution yields

$$\nabla s_n(t) = \nabla s_n(\tau_k^+), \text{ for } t \in [\tau_k, \tau_{k+1}) \quad (36)$$

for all $k = 1, \ldots, K$, $n = 1, \ldots, N$. Similarly, applying (28) to $R_i(t)$ and using (6) gives:

$$\frac{\partial R_i}{\partial \theta_{n,\xi}}(t) = \frac{\partial R_i}{\partial \theta_{n,\xi}}(\tau_k^+)$$
$$- \begin{cases} 0 & \text{if } R_i(t) = 0, \ A_i < BP_i(\mathbf{s}(t)) \\ G \frac{\partial s_n(\tau_k^+)}{\partial \theta_{n,\xi}} & \text{otherwise} \end{cases} \quad (37)$$

and

$$\frac{\partial R_i}{\partial w_{n,\xi}}(t) = \frac{\partial R_i}{\partial w_{n,\xi}}(\tau_k^+)$$
$$- \begin{cases} 0 & \text{if } R_i(t) = 0, \ A_i < BP_i(\mathbf{s}(t)) \\ G \frac{\partial s_n(\tau_k^+)}{\partial w_{n,\xi}} & \text{otherwise} \end{cases} \quad (38)$$

where $G = B \prod_{d \neq n} (1 - p_i(s_d(t))) \left(\frac{\partial p_i(s_n)}{\partial s_n}\right)(t - \tau_k)$

Details on the derivation of simple recursive expressions for the components of $\nabla s_n(\tau_k^+)$ and $\nabla R_i(\tau_k^+)$ in (36)-(38) can be found in Appendix B.

Objective Function Gradient Evaluation. Proceeding as in the evaluation of $\nabla J(\theta)$ in Section III-A1, we are now interested in minimizing the objective function $J(\theta, w)$ in (7) with respect to θ and w and we can obtain $\nabla J(\theta, w) = [\frac{dJ(\theta, w)}{d\theta} \frac{dJ(\theta, w)}{dw}]^{\mathrm{T}}$ as

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

This depends entirely on $\nabla R_i(t)$, which is obtained from (37) and (38) and the event times τ_k, $k = 1, \ldots, K$, given initial conditions $s_n(0) = a$ for $n = 1, \ldots, N$, and $R_i(0)$ for $i = 1, \ldots, M$. In (37), $\frac{\partial R_i}{\partial \theta_{n,\xi}}(\tau_k^+)$ is obtained through (50) and (52) found in Appendix B, whereas $\frac{\partial s_n(\tau_k^+)}{\partial \theta_{n,\xi}}$ is obtained through (36) and (49), (57), (63) found in Appendix B. In (38), $\frac{\partial R_i}{\partial w_{n,\xi}}(\tau_k^+)$ is again obtained through (50) and (52), whereas $\frac{\partial s_n(\tau_k^+)}{\partial w_{n,\xi}}$ is obtained through (49) and (67) which are also found in Appendix B.

Remark 2. Observe that the evaluation of $\nabla R_i(t)$, hence $\nabla J(\theta, w)$, is *independent* of A_i, $i = 1, \ldots, M$, i.e., the values in our uncertainty model. In fact, the dependence of $\nabla R_i(t)$ on A_i, $i = 1, \ldots, M$, manifests itself through the event times τ_k, $k = 1, \ldots, K$, that do affect this evaluation, but they, unlike A_i which may be unknown, are directly observable during the gradient evaluation process. Thus, the IPA approach possesses an inherent *robustness* property: there is no need

to explicitly model how uncertainty affects $R_i(t)$ in (6). Consequently, we may treat A_i as unknown without affecting the solution approach (the values of $\nabla R_i(t)$ are obviously affected). We may also allow this uncertainty to be modeled through random processes $\{A_i(t)\}$, $i = 1, \ldots, M$; in this case, however, the result of Proposition III.3 no longer applies without some conditions on the statistical characteristics of $\{A_i(t)\}$ and the resulting $\nabla J(\theta, w)$ is an *estimate* of a stochastic gradient.)

B. Objective Function Optimization

We now seek to obtain $\theta^\star$ and $w^\star$ minimizing $J(\theta, w)$ through a standard gradient-based optimization scheme of the form

$$[\theta^{l+1}, w^{l+1}]^{\mathrm{T}} = [\theta^l, w^l]^{\mathrm{T}} - [\eta_\theta, \eta_w] \tilde{\nabla} J(\theta^l, w^l) \qquad (39)$$

where $\{\eta_\theta^l\}, \{\eta_w^l\}$ are appropriate step size sequences and $\tilde{\nabla} J(\theta^l, w^l)$ is the projection of the gradient $\nabla J(\theta^l, w^l)$ onto the feasible set (the set of θ^{l+1} satisfying the constraint (27), $a \leq \theta^{l+1} \leq b$, and $w^l \geq 0$). The optimization scheme terminates when $|\tilde{\nabla} J(\theta, w)| < \varepsilon$ (for a fixed threshold ε) for some θ and w. Our IPA-based algorithm to obtain $\theta^\star$ and $w^\star$ minimizing $J(\theta, w)$ is summarized in Algorithm 1 where we have adopted the Armijo method in step-size selection (see [26]) for $\{[\eta_\theta^l, \eta_w^l]\}$.

One of the unusual features in (39) is the fact that the dimension $\Gamma_n^\star$ of $\theta_n^\star$ and $w_n^\star$ is a priori unknown (it depends on T). Thus, the algorithm must implicitly determine this value along with $\theta_n^\star$ and $w_n^\star$. One can search over feasible values of $\Gamma_n \in \{1, 2, \ldots\}$ by starting either with a lower bound $\Gamma_n = 1$ or an upper bound to be found. The latter approach results in much faster execution and is followed in Algorithm 1. An upper bound is determined by observing that $\theta_{n,\xi}$ is the switching point where agent n changes speed from 1 to 0 for ξ odd and from -1 to 0 for ξ even. By setting these two groups of switching points so that their distance is sufficiently small and waiting times $w_n = \mathbf{0}$ for each agent, we determine an approximate upper bound for Γ_n as follows. First, we divide the feasible space $[a, b]$ evenly into N intervals: $[a + \frac{n-1}{N}(b-a), a + \frac{n}{N}(b-a)]$, $n = 1, \ldots, N$. Define $D_n = a + \frac{2n-1}{2N}(b-a)$ to be the geometric center of each interval and set $\theta_{n,\xi} = D_n - \sigma$ if ξ is even and $\theta_{n,\xi} = D_n + \sigma$ if ξ is odd, so that the distance between switching points $\theta_{n,\xi}$ for ξ odd and even is 2σ, where $\sigma > 0$ is an arbitrarily small number, $n = 1, \ldots, N$. In addition, set $w_n = \mathbf{0}$. Then, T must satisfy

$$\theta_{n,1} - s_n(0) + 2\sigma(\Gamma_n - 1) \leq T \leq \theta_{n,1} - s_n(0) + 2\sigma\Gamma_n \quad (40)$$

$n = 1, \ldots, N$, where Γ_n is the number of switching points agent n can reach during $(0, T]$, given $\theta_{n,\xi}$ as defined above. From (40) and noting that Γ_n is an integer, we have

$$\Gamma_n = \left\lceil \frac{1}{2\sigma} [T - \theta_{n,1} + s_n(0)] \right\rceil \qquad (41)$$

where $\lceil \cdot \rceil$ is the ceiling function. Clearly, reducing σ increases the initial number of switching points Γ_n assigned to agent n and $\Gamma_n \to \infty$ as $\sigma \to 0$. Therefore, σ is selected sufficiently

small while ensuring that the algorithm can be executed sufficiently fast.

As Algorithm 1 repeats steps 3-6, $w_{n,\xi} \geq 0$ and distances between $\theta_{n,\xi}$ for ξ odd and even generally increase, so that the number of switching points agent n can actually reach within T decreases. In other words, as long as σ is sufficiently small (hence, Γ_n is sufficiently large), when the algorithm converges to a local minimum and stops, there exists $\zeta_n < \Gamma_n$, such that θ_{n,ζ_n} is the last switching point agent n can reach within $(0, T]$, $n = 1, \ldots, N$. Observe that there generally exist ξ such that $\zeta_n < \xi \leq \Gamma_n$ which correspond to points $\theta_{n,\xi}$ that agent n cannot reach within $(0, T]$; the associated derivatives of the cost with respect to such $\theta_{n,\xi}$ are obviously 0, since perturbations to these $\theta_{n,\xi}$ will not affect $s_n(t)$, $t \in (0, T]$ and thus the cost $J(\theta, w)$. When $|\tilde{\nabla} J(\theta, w)| < \epsilon$, we achieve a local minimum and stop, at which point the dimension of $\theta_n^\star$ and $w_n^\star$ is ζ_n.

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

1: Pick $\sigma > 0$ and $\epsilon > 0$.
2: Define $D_n = a + \frac{2n-1}{2N}(b-a)$, $n = 1, \ldots, N$, and set
$$\begin{cases} \theta_{n,\xi} = D_n - \sigma & \text{if } \xi \text{ even} \\ \theta_{n,\xi} = D_n + \sigma & \text{if } \xi \text{ odd} \end{cases}.$$
Set $w = [w_1, \ldots, w_N] = 0.$, where $w_n = [w_{n,1}, \ldots, w_{n,\xi_n}]$ and $\Gamma_n = \left\lceil \frac{1}{2\sigma}[T - \theta_{n,1} + s_n(0)] \right\rceil$
3: **repeat**
4:    Compute $s_n(t)$, $t \in [0, T]$ using $s_n(0)$, (12), θ and w for $n = 1, \ldots, N$
5:    Compute $\tilde{\nabla} J(\theta, w)$ and update θ, w through (39)
6: **until** $|\tilde{\nabla} J(\theta, w)| < \epsilon$
7: Set $\theta_n^\star = \left[\theta_{n,1}^\star, \ldots, \theta_{n,\zeta_n}^\star\right]$ and $w_n^\star = \left[w_{n,1}^\star, \ldots, w_{n,\zeta_n}^\star\right]$, where ζ_n is the index of θ_{n,ζ_n}, which is the last switching point agent n can reach within $(0, T]$, $n = 1, \ldots, N$

IV. NUMERICAL EXAMPLES

In this section we present some examples of persistent monitoring problems in which agent trajectories are determined using Algorithm 1. The first four are single-agent examples with $L = 20$, $M = 21$, $\alpha_1 = 0$, $\alpha_M = 20$, and the remaining sampling points are evenly spaced over $[0, 20]$. The sensing range in (4) is set to $r = 4$, the initial values of the uncertainty functions in (6) are $R_i(0) = 4$ for all i, and the time horizon is $T = 400$. In Fig. 3(a) we show results where the agent is allowed to move over the entire space $[0, 20]$ and the uncertainty model is selected so that $B = 3$ and $A_i = 0.1$ for all $i = 0, \ldots, 20$, whereas in Fig. 3(b) the feasible space is limited to $[a, b]$ with $a = r = 4$ and $b = L - r = 16$. The top plot in each example shows the optimal trajectory $s^\star(t)$ obtained, while the bottom shows the cost $J(\theta^l, w^l)$ as a function of iteration number. In Fig. 4, the trajectories in Fig. 3(a),(b) are magnified for the interval $t \in [0, 75]$ to emphasize the presence of strictly positive waiting times at the switching points.

In Fig. 3(c) we show results for a case similar to Fig. 3(a) except that the values of A_i are selected so that $A_0 = A_{20} = $
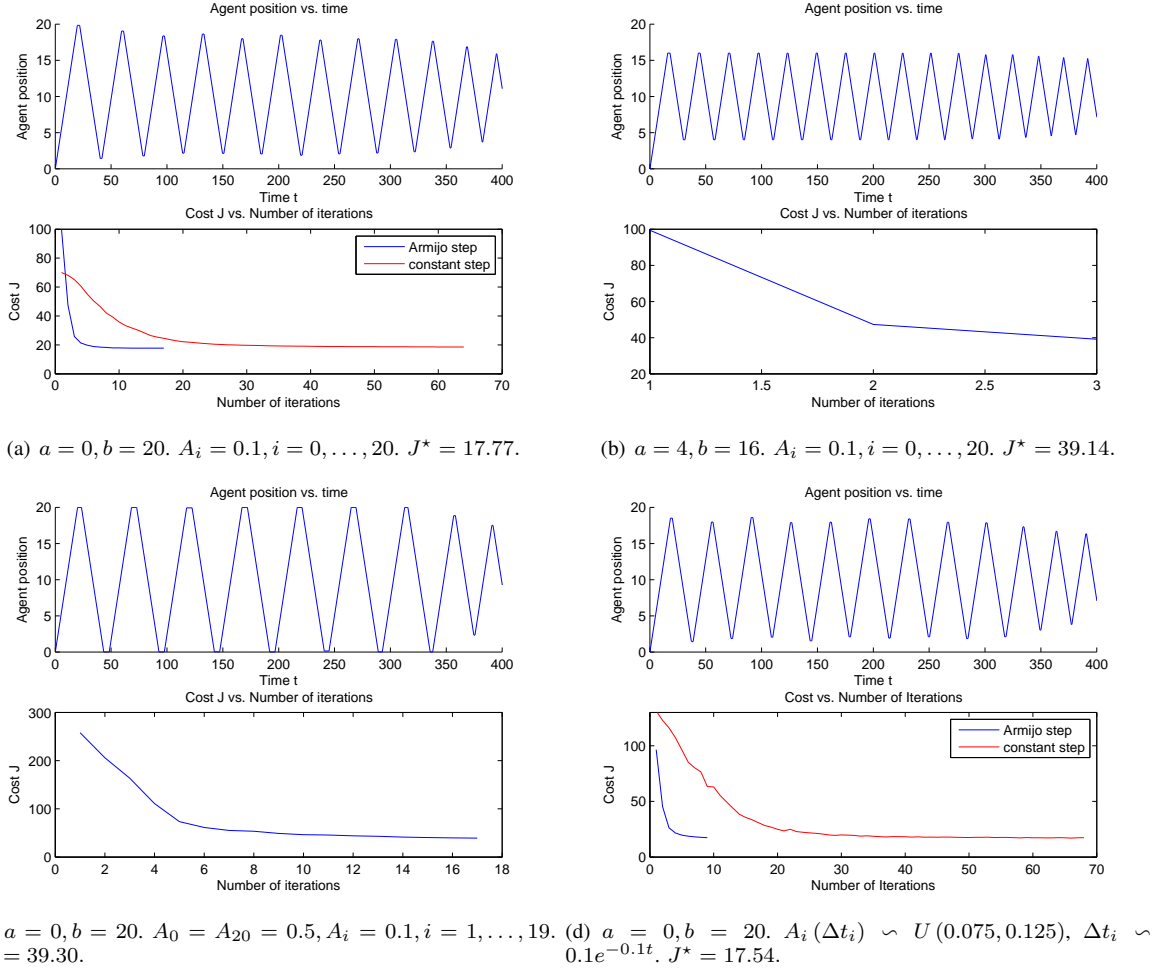
(a) $a = 0, b = 20$. $A_i = 0.1, i = 0, \ldots, 20$. $J^\star = 17.77$.

(b) $a = 4, b = 16$. $A_i = 0.1, i = 0, \ldots, 20$. $J^\star = 39.14$.

(c) $a = 0, b = 20$. $A_0 = A_{20} = 0.5, A_i = 0.1, i = 1, \ldots, 19$. $J^\star = 39.30$.

(d) $a = 0, b = 20$. $A_i(\Delta t_i) \curvearrowright U(0.075, 0.125)$, $\Delta t_i \curvearrowright 0.1 e^{-0.1t}$. $J^\star = 17.54$.

Figure 3. One agent example. $L = 20, T = 400$. For each example, top plot: optimal trajectory; bottom plot: J versus iterations.

0.5, while $A_i = 0.1, i = 1, \ldots, 19$. Note that the waiting times at the switching points are now longer and even though it seems that the switching points are at the two end points, they are actually very close but not equal to these end points, consistent with Proposition III.1. In Fig. 3(d), on the other hand, the values of A_i are allowed to be *random*, thus dealing with a persistent monitoring problem in a stochastic mission space, where we can test the robustness of the proposed approach. In particular, each A_i is treated as a piecewise constant random process $\{A_i(t)\}$ such that $A_i(t)$ takes on a fixed value sampled from an uniform distribution over $(0.075, 0.125)$ for an exponentially distributed time interval with mean 10 before switching to a new value. Note that the behavior of the system in this case is very similar to Fig. 3(a) where $A_i = 0.1$ for all $i = 0, \ldots, 20$ without any change in the way in which $\nabla J(\theta^l, w^l)$ is evaluated in executing (39). As already pointed out, this exploits a robustness property of IPA which makes the evaluation of $\nabla J(\theta^l, w^l)$ independent of the values of A_i. In general, however, when $A_i(t)$ is time-varying, Proposition III.3 may no longer apply, since an extra term $\sum_i \dot{A}_i(t)$ would be present in (24). In such a case, $u_n^\star(t)$ may be nonzero when $\lambda_n^\star(t) = 0$ and the determination of an optimal trajectory through switching points and waiting times alone may no longer be possible. In the case of 3(d),

$A_i(t)$ changes sufficiently slowly to maintain the validity of Proposition III.3 over relatively long time intervals, under the assumption that w.p. 1 no event time coincides with the jump times in any $\{A_i(t)\}$.

In all cases, we initialize the algorithm with $\sigma = 5$ and $\varepsilon = 2 \times 10^{-10}$. The running times of Algorithm 1 are approximately 10 sec using Armijo step-sizes. Note that although the number of iterations for the examples shown may vary substantially, the actual algorithm running times do not. This is simply because the Armijo step-size method may require several trials per iteration to adjust the step-size in order to achieve an adequate decrease in cost. In Fig. 3(a),(d), the red line shows the cost as a function of iteration number using a constant step size and the two lines converge to the same approximate optimal value. Non-smoothness in Fig. 3(d) comes from the fact that it is a stochastic process. Note that in all cases the initial cost is significantly reduced indicating the importance of optimally selecting the values of the switching points and associated waiting times (if any).

Figure 5 shows two two-agent examples with $L = 40, M = 41$ and evenly spaced sampling points over $[0, L]$, $A_i = 0.01$, $B = 3$, $r = 4$, $R_i(0) = 4$ for all i and $T = 400$. In Fig. 5(a) the agents are allowed to move over the whole mission space $[0, L]$, while in Fig. 5(b) they are only allowed to move over

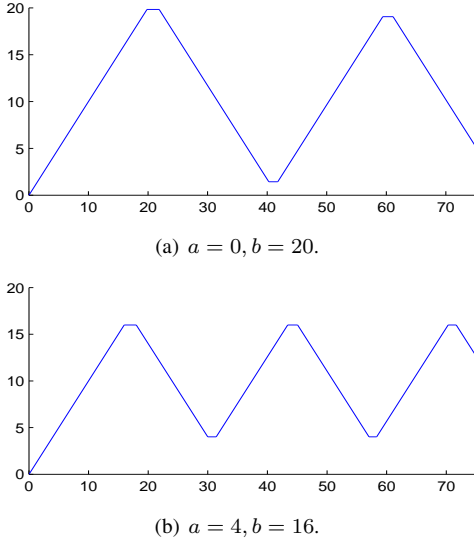(a) $a = 0, b = 20$.

(b) $a = 4, b = 16$.

Figure 4. Magnified trajectory for sub-figure (a) and (b) in Fig. 3, $t \in [0, 75]$.

$[a, b]$ where $a = r$ and $b = L - r$. We initialize the algorithm with the same σ and ε as before. The algorithm running time is approximately 15 sec using Armijo step-sizes, and we observe once again significant reductions in cost.

V. CONCLUSION

We have formulated an optimal persistent monitoring problem with the objective of controlling the movement of multiple cooperating agents to minimize an uncertainty metric in a given mission space. In a one-dimensional mission space, we have shown that the optimal solution is reduced to the determination of two parameter vectors for each agent: a sequence of switching locations and associated waiting times at these switching points. We have used Infinitesimal Perturbation Analysis (IPA) to obtain sensitivities of the objective function with respect to all the parameters and, therefore, obtain a complete on-line (locally optimal) solution through a gradient-based algorithm. We have also shown that the solution is robust with respect to the uncertainty model used. Our ongoing work aims at incorporating constraints such as $R_i(T) \leq R_{\max}$ to the problem formulation, thus ensuring that an optimal persistent monitoring solution provides certain performance guarantees. We are also investigating the use of receding horizon controllers that provide computationally fast approximate solutions. Finally, our work to date has established the basis for extending this approach to a two-dimensional mission space. Specifically, one idea is to decompose such a two-dimensional mission space into regions each one of which is monitored by agents moving on a one-dimensional trajectory, thus taking direct advantage of the results in this paper.

APPENDIX A
IPA DERIVATION FOR SINGLE-AGENT SOLUTION

In order to determine $\nabla s\left(\tau_k^+\right)$ and $\nabla R_i(\tau_k^+)$ which are needed to evaluate $\nabla R_i(t)$ in (32), we use (29), which involves the event time gradient vectors $\nabla \tau_k = [\frac{\partial \tau_k}{\partial \theta_1}, \ldots, \frac{\partial \tau_k}{\partial \theta_\Gamma}]^{\mathrm{T}}$ for $k = 1, \ldots, K$ (the value of K depends on T.) Looking at

(a) $a = 0, b = 20$. $J^\star = 17.77$.

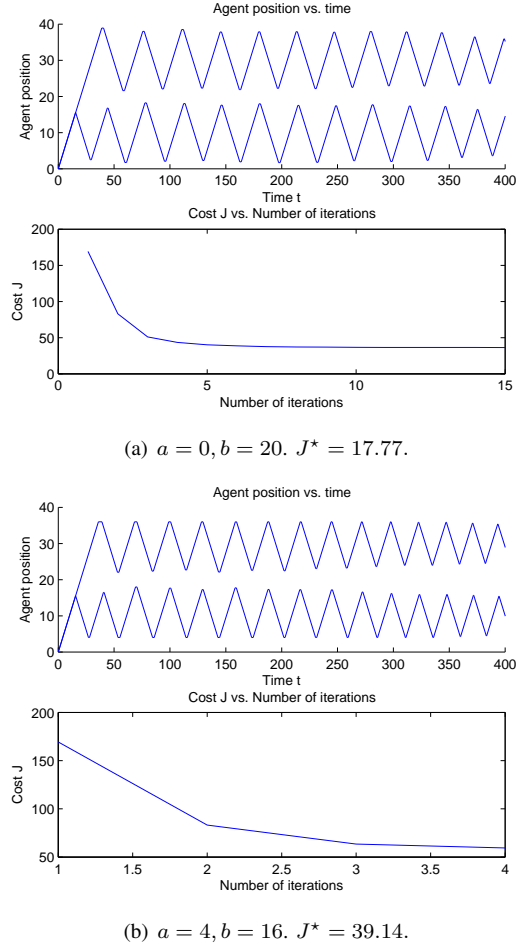(b) $a = 4, b = 16$. $J^\star = 39.14$.

Figure 5. Two agent example. $L = 40, T = 400$. Top plot: optimal trajectory. Bottom plot: J versus iterations.

Fig. 2, there are three readily distinguishable cases regarding the events that cause discrete state transitions:

Case 1: An event at time τ_k which is neither $R_i = 0$ nor $s = \theta_\xi$, for any $\xi = 1, \ldots, \Gamma$. In this case, it is easy to see that the dynamics of both $s(t)$ and $R_i(t)$ are continuous, so that $f_{k-1}(\tau_k^-) = f_k(\tau_k^+)$ in (29) applied to $s(t)$ and $R_i(t)$, $i = 1, \ldots, M$ gives:

$$\begin{cases} \nabla s\left(\tau_k^+\right) = \nabla s\left(\tau_k^-\right) \\ \nabla R_i(\tau_k^+) = \nabla R_i(\tau_k^-), & i = 1, \ldots, M \end{cases} \quad (42)$$

Case 2: An event $R_i = 0$ at time τ_k. This corresponds to transitions $3 \to 7$, $4 \to 7$, $10 \to 14$ and $11 \to 14$ in Fig. 2 where the dynamics of $s(t)$ are still continuous, but the dynamics of $R_i(t)$ switch from $f_{k-1}(\tau_k^-) = A_i - Bp_i(s(\tau_k^-))$ to $f_k(\tau_k^+) = 0$. Thus, $\nabla s\left(\tau_k^-\right) = \nabla s\left(\tau_k^+\right)$, but we need to evaluate τ_k' to determine $\nabla R_i(\tau_k^+)$. Observing that this event is endogenous, (30) applies with $g_k = R_i = 0$ and we get

$$\frac{\partial \tau_k}{\partial \theta_\xi} = -\frac{\frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^-\right)}{A_i - Bp_i(s(\tau_k^-))}, \quad \xi = 1, \ldots, \Gamma, \ k = 1, \ldots, K$$

It follows from (29) that

$$\frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^+\right) = \frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^-\right) - \frac{[A_i - Bp_i(s(\tau_k^-))]\frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^-\right)}{A_i - Bp_i(s(\tau_k^-))} = 0$$

Thus, whenever an event occurs at τ_k such that $R_i(\tau_k)$ becomes zero, $\frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^+\right)$ is always reset to 0 regardless of $\frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^-\right)$.

Case 3: An event at time τ_k due to a control sign change at $s = \theta_\xi$, $\xi = 1, \ldots, \Gamma$. This corresponds to any transition between the upper and lower part of the hybrid automaton in Fig. 2. In this case, the dynamics of $R_i(t)$ are continuous and we have $\frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^+\right) = \frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^-\right)$ for all i, ξ, k. On the other hand, we have $\dot{s}(\tau_k^+) = u(\tau_k^+) = -u(\tau_k^-) = \pm 1$. Observing that any such event is endogenous, (30) applies with $g_k = s - \theta_\xi = 0$ for some $\xi = 1, \ldots, \Gamma$ and we get

$$\frac{\partial \tau_k}{\partial \theta_\xi} = \frac{1 - \frac{\partial s}{\partial \theta_\xi}\left(\tau_k^-\right)}{u(\tau_k^-)} \tag{43}$$

Combining (43) with (29) and recalling that $u(\tau_k^+) = -u(\tau_k^-)$, we have

$$\frac{\partial s}{\partial \theta_\xi}(\tau_k^+) = \frac{\partial s}{\partial \theta_\xi}(\tau_k^-) + [u\left(\tau_k^-\right) - u(\tau_k^+)]\frac{1 - \frac{\partial s}{\partial \theta_\xi}\left(\tau_k^-\right)}{u(\tau_k^-)} = 2$$

where $\frac{\partial s}{\partial \theta_\xi}\left(\tau_k^-\right) = 0$ because $\frac{\partial s}{\partial \theta_\xi}(0) = 0 = \frac{\partial s}{\partial \theta_\xi}(t)$ for all $t \in [0, \tau_k)$, since the position of the agent cannot be affected by θ_ξ prior to this event.

In this case, we also need to consider the effect of perturbations to θ_j for $j < \xi$, i.e., prior to the current event time τ_k (clearly, for $j > \xi$, $\frac{\partial s}{\partial \theta_j}(\tau_k^+) = 0$ since the current position of the agent cannot be affected by future events.) Observe that since $g_k = s - \theta_\xi = 0$, we have $\frac{\partial g_k}{\partial \theta_j} = 0$ for $j \neq \xi$ and (30) gives $\frac{\partial \tau_k}{\partial \theta_j} = \frac{-1}{u(\tau_k^-)}\frac{\partial s}{\partial \theta_j}\left(\tau_k^-\right)$, so that using this in (29) we get:

$$\frac{\partial s}{\partial \theta_j}(\tau_k^+) = \frac{\partial s}{\partial \theta_j}(\tau_k^-) - \frac{[u\left(\tau_k^-\right) - u(\tau_k^+)]\frac{\partial s}{\partial \theta_j}\left(\tau_k^-\right)}{u\left(\tau_k^-\right)}$$
$$= -\frac{\partial s}{\partial \theta_j}\left(\tau_k^-\right) \tag{44}$$

Combining the above results, the components of $\nabla s(\tau_k^+)$ where τ_k is the event time when $s(\tau_k) = \theta_\xi$ for some ξ, are given by

$$\frac{\partial s}{\partial \theta_j}(\tau_k^+) = \begin{cases} -\frac{\partial s}{\partial \theta_j}\left(\tau_k^-\right) & \text{if } j = 1, \ldots, \xi - 1 \\ 2 & \text{if } j = \xi \\ 0 & \text{if } j > \xi \end{cases} \tag{45}$$

It follows from (31) and the analysis of all three cases above that $\frac{\partial s}{\partial \theta_\xi}(t)$ for all ξ is constant throughout an optimal trajectory except at transitions caused by control switching locations (*Case 3*). In particular, for the kth event corresponding to $s(\tau_k) = \theta_\xi$, $t \in [\tau_k, T]$, if $u(t) = 1$, then $\frac{\partial s}{\partial \theta_\xi}(t) = -2$ if ξ is odd, and $\frac{\partial s}{\partial \theta_\xi}(t) = 2$ if ξ is even; similarly, if $u(t) = -1$, then $\frac{\partial s}{\partial \theta_\xi}(t) = 2$ if ξ is odd and $\frac{\partial s}{\partial \theta_\xi}(t) = -2$ if ξ is even. In summary, we can write:

$$\frac{\partial s}{\partial \theta_\xi}(t) = \begin{cases} (-1)^\xi \cdot 2u(t) & t \geq \tau_k \\ 0 & t < \tau_k \end{cases}, \quad \xi = 1, \ldots, \Gamma \tag{46}$$

Finally, we can combine (46) with our results for $\frac{\partial R_i}{\partial \theta_\xi}(t)$ in all three cases above. Letting $s(\tau_l) = \theta_\xi$, we obtain the following

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

$$\frac{\partial R_i}{\partial \theta_\xi}(t) = \frac{\partial R_i}{\partial \theta_\xi}\left(\tau_k^+\right)$$
$$+ \begin{cases} 0 & \text{if } z(t) \in Z_1 \cup Z_2 \\ (-1)^{\xi+1}\frac{2B}{r}u\left(\tau_k^+\right) \cdot (t - \tau_k) & \text{if } z(t) \in Z_3 \\ -(-1)^{\xi+1}\frac{2B}{r}u\left(\tau_k^+\right) \cdot (t - \tau_k) & \text{if } z(t) \in Z_4 \end{cases} \tag{47}$$

with boundary condition

$$\frac{\partial R_i}{\partial \theta_\xi}(\tau_k^+) = \begin{cases} 0 & \text{if } z\left(\tau_k^+\right) \in Z_1 \\ \frac{\partial R_i}{\partial \theta_\xi}(\tau_k^-) & \text{otherwise} \end{cases} \tag{48}$$

Appendix B
IPA Derivation for Multi-Agent Solution

The evaluation of the components of $\nabla s_n\left(\tau_k^+\right)$ and $\nabla R_i(\tau_k^+)$ in (36)-(38) using (29) involves the event time gradient vectors $\nabla \tau_k = \left[\frac{\partial \tau_k}{\partial \theta} \frac{\partial \tau_k}{\partial w}\right]^{\mathrm{T}}$ for $k = 1, \ldots, K$, which will be determined through (30). There are three possible cases regarding the events that cause switches in the dynamics of $R_i(t)$ or $s_n(t)$ as mentioned above:

Case 1: An event at time τ_k such that $\dot{R}_i(t)$ switches from $\dot{R}_i(t) = 0$ to $\dot{R}_i(t) = A_i - BP_i(\mathbf{s}(t))$. In this case, it is easy to see that the dynamics of both $s_n(t)$ and $R_i(t)$ are continuous, so that $f_{k-1}(\tau_k^-) = f_k(\tau_k^+)$ in (29) applied to $s_n(t)$ and $R_i(t)$, $i = 1, \ldots, M$, $n = 1, \ldots, N$, and we get

$$\nabla s_n\left(\tau_k^+\right) = \nabla s_n\left(\tau_k^-\right), \quad n = 1, \ldots, N \tag{49}$$
$$\nabla R_i(\tau_k^+) = \nabla R_i(\tau_k^-), \quad i = 1, \ldots, M \tag{50}$$

Case 2: An event at time τ_k such that $\dot{R}_i(t)$ switches from $\dot{R}_i(t) = A_i - BP_i(\mathbf{s}(t))$ to $\dot{R}_i(t) = 0$, i.e., $R_i(\tau_k)$ becomes zero. In this case, we need to first evaluate $\nabla \tau_k$ from (30) in order to determine $\nabla R_i(\tau_k^+)$ through (29). Observing that this event is endogenous, (30) applies with $g_k = R_i = 0$ and we get

$$\nabla \tau_k = -\frac{\nabla R_i(\tau_k^-)}{A_i\left(\tau_k^-\right) - BP_i\left(\mathbf{s}(\tau_k^-)\right)} \tag{51}$$

It follows from (29) that

$$\nabla R_i(\tau_k^+) = \nabla R_i(\tau_k^-) - \frac{[A_i\left(\tau_k^-\right) - BP_i(\mathbf{s}(t))]\nabla R_i\left(\tau_k^-\right)}{A_i\left(\tau_k^-\right) - BP_i\left(\tau_k^-\right)}$$
$$= 0 \tag{52}$$

Thus, $\nabla R_i(\tau_k^+)$ is always reset to 0 regardless of $\nabla R_i(\tau_k^-)$. In addition, (49) holds, since the the dynamics of $s_n(t)$ are continuous at time τ_k.

Case 3: An event at time τ_k such that the dynamics of $s_n(t)$ switch from ± 1 to 0, or from 0 to ± 1. Clearly, (50) holds since the the dynamics of $R_i(t)$ are continuous at this time. However, determining $\nabla s_n\left(\tau_k^+\right)$ is more elaborate and requires us to consider its components separately, first $\frac{\partial s_n\left(\tau_k^+\right)}{\partial \theta_n}$ and then $\frac{\partial s_n\left(\tau_k^+\right)}{\partial w_n}$.

Case 3.1: Evaluation of $\frac{\partial s_n\left(\tau_k^+\right)}{\partial \theta_n}$.

Case 3.1.1: An event at time τ_k such that the dynamics of $s_n(t)$ in (35) switch from ± 1 to 0. This is an endogenous

event and (30) applies with $g_k = s_n - \theta_{n,\xi} = 0$ for some $\xi = 1, \ldots, \Gamma_n$ and we have:

$$\frac{\partial \tau_k}{\partial \theta_{n,\xi}} = \frac{1 - \frac{\partial s_n}{\partial \theta_{n,\xi}}\left(\tau_k^-\right)}{u_n(\tau_k^-)} \tag{53}$$

and (29) yields

$$\frac{\partial s_n}{\partial \theta_{n,\xi}}(\tau_k^+) = \frac{\partial s_n}{\partial \theta_{n,\xi}}(\tau_k^-) + [u_n\left(\tau_k^-\right) - 0]\frac{1 - \frac{\partial s_n}{\partial \theta_{n,\xi}}\left(\tau_k^-\right)}{u_n(\tau_k^-)}$$
$$= 1 \tag{54}$$

As in Case 3 of Section III-A1, we also need to consider the effect of perturbations to θ_j for $j < \xi$, i.e., prior to the current event time τ_k (clearly, for $j > \xi$, $\frac{\partial s_n}{\partial \theta_j}(\tau_k^+) = 0$ since the current position of the agent cannot be affected by future events.) Observe that $\frac{\partial g_k}{\partial \theta_j} = 0$, therefore, (30) becomes

$$\frac{\partial \tau_k}{\partial \theta_{n,j}} = -\frac{\frac{\partial s_n}{\partial \theta_{n,j}}\left(\tau_k^-\right)}{u_n(\tau_k^-)} \tag{55}$$

and using this in (29) gives:

$$\frac{\partial s_n}{\partial \theta_{n,j}}(\tau_k^+) = \frac{\partial s_n}{\partial \theta_{n,j}}(\tau_k^-) - \frac{[u_n\left(\tau_k^-\right) - 0]\frac{\partial s_n}{\partial \theta_{n,j}}\left(\tau_k^-\right)}{u_n\left(\tau_k^-\right)} = 0 \tag{56}$$

Thus, combining the above results, when $s_q(\tau_k) = \theta_{q,\xi}$ for some ξ and the agent switches from ± 1 to 0, we have

$$\frac{\partial s_n}{\partial \theta_{n,j}}(\tau_k^+) = \begin{cases} 0, & \text{if } j \neq \xi \\ 1, & \text{if } j = \xi \end{cases} \tag{57}$$

Case 3.1.2: An event at time τ_k such that the dynamics of $s_n(t)$ in (35) switch from 0 to ± 1. This is an induced event since it is triggered by the occurrence of some other endogenous event when the agent switches from ± 1 to 0 (see *Case 3.1.1* above.) Suppose the agent starts from an initial position $s_n(0) = a$ with $u_n(0) = 1$ and τ_k is the time the agent switches from the 0 to ± 1 at the switching point $\theta_{n,\xi}$. If $\theta_{n,\xi}$ is such that $u_n\left(\tau_k^+\right) = 1$, then ξ is even and τ_k can be calculated as follows:

$$\tau_k = (\theta_{n,1} - a) + w_{n,1} + (\theta_{n,1} - \theta_{n,2}) + w_{n,2} + \ldots$$
$$+ (\theta_{n,\xi-1} - \theta_{n,\xi}) + w_{n,\xi}$$
$$= 2\left(\sum_{v=1,\ v\ \text{odd}}^{\xi-1} \theta_{n,v} - \sum_{v=2,\ v\ \text{even}}^{\xi-2} \theta_{n,v}\right) + \sum_{v=1}^{\xi} w_{n,v}$$
$$- \theta_{n,\xi} \tag{58}$$

Similarly, if $\theta_{n,\xi}$ is the switching point such that $u_n\left(\tau_k^+\right) = -1$, then ξ is odd and we get:

$$\tau_k = 2\left(\sum_{v=1,\ v\ \text{odd}}^{\xi-2} \theta_{n,v} - \sum_{v=2,\ v\ \text{even}}^{\xi-1} \theta_{n,v}\right) + \sum_{v=1}^{\xi} w_{n,v} + \theta_{n,\xi} \tag{59}$$

We can then directly obtain $\frac{\partial \tau_k}{\partial \theta_{n,\xi}}$ as

$$\frac{\partial \tau_k}{\partial \theta_{n,\xi}} = -sgn(u\left(\tau_k^+\right)) \tag{60}$$

Using (60) in (29) gives:

$$\frac{\partial s_n}{\partial \theta_{n,\xi}}(\tau_k^+) = \frac{\partial s_n}{\partial \theta_{n,\xi}}(\tau_k^-)$$
$$+ \left[0 - u\left(\tau_k^+\right)\right] \cdot \left[-sgn(u\left(\tau_k^+\right))\right]$$
$$= \frac{\partial s_n}{\partial \theta_{n,\xi}}(\tau_k^-) + 1 \tag{61}$$

Once again, we need to consider the effect of perturbations to θ_j for $j < \xi$, i.e., prior to the current event time τ_k (clearly, for $j > \xi$, $\frac{\partial s_n}{\partial \theta_j}(\tau_k^+) = 0$.) In this case, from (58)-(59), we have

$$\begin{cases} \frac{\partial \tau_k}{\partial \theta_{n,j}} = 2, & \text{if } j \text{ odd} \\ \frac{\partial \tau_k}{\partial \theta_{n,j}} = -2, & \text{if } j \text{ even} \end{cases} \tag{62}$$

and it follows from (29) that for $j < \xi$:

$$\frac{\partial s_n}{\partial \theta_{n,j}}(\tau_k^+)$$
$$= \begin{cases} \frac{\partial s_n}{\partial \theta_{n,j}}(\tau_k^-) + 2, & \text{if } u_n\left(\tau_k^+\right) = 1,\ j \text{ even,} \\ & \quad \text{or } u_n\left(\tau_k^+\right) = -1,\ j \text{ odd} \\ \frac{\partial s_n}{\partial \theta_{n,j}}(\tau_k^-) - 2, & \text{if } u_n\left(\tau_k^+\right) = 1,\ j \text{ odd,} \\ & \quad \text{or } u_n\left(\tau_k^+\right) = -1,\ j \text{ even} \end{cases} \tag{63}$$

Case 3.2: Evaluation of $\frac{\partial s_n(\tau_k^+)}{\partial w_n}$.
Case 3.2.1: An event at time τ_k such that the dynamics of $s_n(t)$ in (35) switch from ± 1 to 0. This is an endogenous event and (30) applies with $g_k = s_n - \theta_{n,\xi} = 0$ for some $\xi = 1, \ldots, \Gamma_n$. Then, for any $j \leq \xi$, we have:

$$\frac{\partial \tau_k}{\partial w_{n,j}} = \frac{-\frac{\partial s_n}{\partial w_{n,j}}\left(\tau_k^-\right)}{u_n(\tau_k^-)} \tag{64}$$

Combining (64) with (29) and since $u_n\left(\tau_k^-\right) = \pm 1$, we have

$$\frac{\partial s_n}{\partial w_{n,j}}(\tau_k^+) = \frac{\partial s_n}{\partial w_{n,j}}(\tau_k^-) + [u_n\left(\tau_k^-\right) - 0]\frac{-\frac{\partial s_n}{\partial w_{n,j}}\left(\tau_k^-\right)}{u_n(\tau_k^-)}$$
$$= 0 \tag{65}$$

Case 3.2.2: An event at time τ_k such that the dynamics of $s_n(t)$ in (35) switch from 0 to ± 1. As in *Case 3.1.2*, τ_k is given by (58) or (59), depending on the sign of $u_q\left(\tau_k^+\right)$. Thus, we have $\frac{\partial \tau_k}{\partial w_{n,j}} = 1$, for $j \leq \xi$. Using this result in (29) and observing that $\frac{\partial s_n}{\partial w_{n,j}}(\tau_k^-) = 0$ from (65), we have

$$\frac{\partial s_n}{\partial w_{n,j}}(\tau_k^+) = \frac{\partial s_n}{\partial w_{n,j}}(\tau_k^-) + [0 - u_n\left(\tau_k^+\right)] \cdot 1$$
$$= -u_n\left(\tau_k^+\right), \text{ for } j \leq \xi \tag{66}$$

Combining the above results, we have for *Case 3.2*:

$$\frac{\partial s_n}{\partial w_{n,j}}(\tau_k^+) = \begin{cases} 0, & \text{if } u_n\left(\tau_k^-\right) = \pm 1,\ u_n\left(\tau_k^+\right) = 0 \\ \mp 1, & \text{if } u_n\left(\tau_k^-\right) = 0,\ u_n\left(\tau_k^+\right) = \pm 1 \end{cases} \tag{67}$$

Finally, note that $\frac{\partial s_n}{\partial w_{n,\xi}}(t) = 0$ for $t \in [0, \tau_k)$, since the position of the agent n cannot be affected by $w_{n,\xi}$ prior to such an event.

REFERENCES

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

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

[3] W. Li and C. G. 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 Conf. Decision and Control*, vol. 1, 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] S. Smith, M. Schwager, and D. Rus, "Persistent robotic tasks: Monitoring and sweeping in changing enviroments," *IEEE Transactions on Robotics*, 2012, to appear.

[7] D. Paley, F. Zhang, and N. Leonard, "Cooperative control for ocean sampling: The glider coordinated control system," *IEEE Trans. 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] R. Cooper, *Introduction to queuing theory*. Edward Arnold, 1981.

[10] G. Sun, C. G. Cassandras, Y. Wardi, C. G. 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.

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

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

[13] Y. Elmaliach, A. Shiloni, and G. Kaminka, "A realistic model of frequency-based multi-robot polyline patrolling," in *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 2008, pp. 63–70.

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

[15] D. Kingston, R. Beard, and R. Holt, "Decentralized perimeter surveillance using a team of uavs," *IEEE Transactions on Robotics*, vol. 24, pp. 1394–1404, 2008.

[16] C. G. Cassandras, X. C. Ding, and X. Lin, "An optimal control approach for the persistent monitoring problem," in *Proc. of 50th IEEE Conf. Decision and Control*, 2011, pp. 2907–2912.

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

[18] Y. Wardi, R. Adams, and B. Melamed, "A unified approach to infinitesimal perturbation analysis in stochastic flow models: the single-stage case," *IEEE Trans. on Automatic Control*, vol. 55, no. 1, pp. 89–103, 2009.

[19] J. Le Ny, M. Dahleh, E. Feron, and E. Frazzoli, "Continuous path planning for a data harvesting mobile server," in *47th IEEE Conference on Decision and Control*. IEEE, 2008, pp. 1489–1494.

[20] A. Bryson and Y. Ho, *Applied optimal control*. Wiley N.Y., 1975.

[21] M. Egerstedt, Y. Wardi, and H. Axelsson, "Transition-time optimization for switched-mode dynamical systems," *Automatic Control, IEEE Transactions on*, vol. 51, no. 1, pp. 110–115, 2006.

[22] M. Shaikh and P. Caines, "On the hybrid optimal control problem: Theory and algorithms," *IEEE Trans. on Automatic Control*, vol. 52, no. 9, pp. 1587–1603, 2007.

[23] X. Xu and P. Antsaklis, "Optimal control of switched systems based on parameterization of the switching instants," *IEEE Trans. on Automatic Control*, vol. 49, no. 1, pp. 2–16, 2004.

[24] C. Yao and C. G. Cassandras, "Perturbation analysis of stochastic hybrid systems and applications to resource contention games," *Frontiers of Electr. and Electr. Engin. in China*, vol. 6,3, pp. 453–467, 2011.

[25] C. G. Cassandras and J. E. Lygeros, *Stochastic hybrid systems*. Taylor and Francis, 2007.

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

Christos G. Cassandras received the B.S. degree from Yale University, the M.S.E.E degree from Stanford University, and the S.M. and Ph.D. degrees from Harvard University in 1977, 1978, 1979, and 1982, respectively. From 1982 to 1984 he was with ITP Boston, Inc. where he worked on the design of automated manufacturing systems. From 1984 to 1996 he was a faculty member at the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst. Currently, he is Head of the Division of Systems Engineering and Professor of Electrical and Computer Engineering at Boston University, Boston, MA and a founding member of the Center for Information and System Engineering (CISE). He specializes in the areas of discrete event and hybrid systems, cooperative control, stochastic optimization, and computer simulation, with applications to computer and sensor networks, manufacturing systems, and transportation systems. He has published over 300 papers in these areas, and five books. Dr. Cassandras was Editor-in-Chief of the IEEE Transactions on Automatic Control from 1998 through 2009 and has served on several editorial boards and as Guest Editor for various journals. He is the 2012 President of the IEEE Control Systems Society and the recipient of several awards, including the 2011 IEEE Control Systems Technology Award, the Distinguished Member Award of the IEEE Control Systems Society (2006), the 1999 Harold Chestnut Prize (IFAC Best Control Engineering Textbook) for Discrete Event Systems: Modeling and Performance Analysis, a 2012 Kern Fellowship, and a 1991 Lilly Fellowship. He is a member of Phi Beta Kappa and Tau Beta Pi, a Fellow of the IEEE and a Fellow of the IFAC.

Xuchao Lin received the B.E. degree in electrical engineering from Zhejiang University, Hangzhou, China and the M.S. degree from Boston University, Boston, MA in 2009 and 2012, respectively. He is currently a PhD candidate in System Engineering at Boston University, Boston, MA. His research interests include discrete event systems and optimal control of hybrid systems, with applications to persistent monitoring, communication networks, sensor networks, intelligent vehicle systems and robotics. He is a student member of the IEEE.

Xu Chu (Dennis) Ding received his B.S., M.S. and Ph.D. degree in Electrical and Computer Engineering from the Georgia Institute of Technology, Atlanta, in 2004, 2007 and 2009, respectively. During 2010 to 2011 he was a postdoctoral research fellow at Boston University. He has recently joined United Technologies Research Center as a Senior Research Engineer. His research interests include formal methods in control synthesis, optimal control of hybrid systems, coordination and control of multi-agent networked systems, and intelligent and persistent surveillance.