2016 IEEE 55th Conference on Decision and Control (CDC)
ARIA Resort & Casino
December 12-14, 2016, Las Vegas, USA

Optimal Event-Driven Multi-Agent Persistent Monitoring of a Finite Set of Targets

Nan Zhou[1], Xi Yu[2], Sean B. Andersson[1,2], and Christos G. Cassandras[1,3]
[1]Division of Systems Engineering, [2]Department of Mechanical Engineering, [3]Department of Electrical and Computer Engineering
Boston University, Boston, MA 02215, USA
E-mail:{nanzhou,xyu,sanderss,cgc}@bu.edu

Abstract— **We consider the problem of controlling the movement of multiple cooperating agents so as to minimize an uncertainty metric associated with a finite number of targets. In a one-dimensional mission space, we adopt an optimal control framework and show that the solution is reduced to a simpler parametric optimization problem: determining a sequence of locations where each agent may dwell for a finite amount of time and then switch direction. This amounts to a hybrid system which we analyze using Infinitesimal Perturbation Analysis (IPA) to obtain a complete on-line solution through an event-driven gradient-based algorithm which is also robust with respect to the uncertainty model used. The resulting controller depends on observing the events required to excite the gradient-based algorithm, which cannot be guaranteed. We solve this problem by proposing a new metric for the objective function which creates a potential field guaranteeing that gradient values are non-zero. This approach is compared to an alternative graph-based task scheduling algorithm for determining an optimal sequence of target visits. Simulation examples are included to demonstrate the proposed methods.**

I. Introduction

Systems consisting of cooperating mobile agents are often used to perform tasks such as coverage control [1], surveillance, and environmental sampling. The *persistent monitoring* problem arises when agents must monitor a dynamically changing environment which cannot be fully covered by a stationary team of agents. Thus, persistent monitoring differs from traditional coverage tasks due to the perpetual need to cover a changing environment [2], [3]. A result of this exploration process is the eventual discovery of various "points of interest" which, once detected, become "targets" or "data sources" which need to be monitored. This setting arises in multiple application domains ranging from surveillance, environmental monitoring, and energy management [4], [5] down to nano-scale systems tasked to track fluorescent or magnetic particles for the study of dynamic processes in bio-molecular systems and in nano-medical research [6]. In contrast to [2], [3] where *every* point in a mission space must be monitored, the problem we address here involves a *finite number* of targets (typically larger than the number of agents) which the agents must cooperatively monitor through periodic visits.

Each target may be viewed as a dynamic system in itself whose state is observed by agents equipped with sensing capabilities (e.g., cameras) and which are normally dependent upon their physical distance from the target. The objective of cooperative persistent monitoring in this case is to minimize an overall measure of uncertainty about the target states. This may be accomplished by assigning agents to specific targets or by designing motion trajectories through which agents reduce the uncertainty related to a target by periodically visiting it (and possibly remaining at the target for a finite amount of time). As long as the numbers of agents and targets are small, it is possible to identify sequences that yield a globally optimal solution; in general, however, this is computationally intensive and does not scale well [7].

Rather than viewing this problem as a scheduling task which eventually falls within the class of traveling salesman or vehicle routing problems [8], in this paper we follow earlier work in [2], [3] and introduce an optimal control framework whose objective is to control the movement of agents so as to collect information from targets (within agent sensing ranges) and ultimately minimize an average metric of uncertainty over all targets. An important difference between the persistent monitoring problem in previous work [2] and the current setting is that there is now a finite number of targets that agents need to monitor as opposed to every point in the mission space. In a one-dimensional mission space, we show that the optimal control problem can be reduced to a parametric optimization problem. As a result, the behavior of agents under optimal control is described by a hybrid system. This allows us to make use of Infinitesimal Perturbation Analysis (IPA) [9], [10] to determine on-line the gradient of the objective function with respect to these parameters and to obtain a (possibly local) optimal trajectory. Our approach exploits an inherent property of IPA which allows virtually arbitrary stochastic effects in modeling target uncertainty. Moreover, IPA's event-driven nature renders it *scalable* in the number of events in the system and not its state space.

A potential drawback of event-driven control methods is that they obviously depend on the events which "excite" the controller being observable. However, this is not guaranteed under every feasible control: it is possible that no such events are excited, in which case the controller may be useless. The crucial events in our case are "target visits" and it is possible that such events may never occur for a large number of feasible agent trajectories which IPA uses to estimate a gradient on-line. This lack of event excitation is a serious problem in many trajectory planning and optimization tasks [11], [12]. We solve this problem using a new cost metric introduced in [13] which creates a potential field guaranteeing that gradient values are generally non-zero throughout the mission space and ensures that all events are ultimately excited.

* The work of Cassandras and Zhou is supported in part by NSF under grants CNS-1239021, ECCS-1509084, and IIP-1430145, by AFOSR under grant FA9550-15-1-0471, and by ONR under grant N00014-09-1-1051. The work of Andersson and Yu is supported in part by the NSF through grant ECCS-1509084.

II. Persistent Monitoring Problem Formulation

We consider N mobile agents moving in a one dimensional mission space $[0,L] \subset \mathbb{R}$. Let the position of the agents at time t be $s_j(t) \in [0,L]$, $j = 1,\ldots,N$, following the dynamics:

$$\dot{s}_j(t) = u_j(t), |u_j(t)| \leq 1, \forall j = 1,\ldots,N \tag{1}$$

An additional constraint may be imposed if we assume that agents are initially located so that $s_j(0) < s_{j+1}(0)$, $j = 1,\ldots,N-1$, and we wish to prevent them from crossing each other over all t: $s_j(t) - s_{j+1}(t) \leq 0$. The ability of an agent to sense its environment is modeled by a function $p_j(x,s_j)$ that measures the probability that an event at location $x \in [0,L]$ is detected by agent j. We assume that $p_j(x,s_j) = 1$ if $x = s_j$, and that $p_j(x,s_j)$ is monotonically nonincreasing in the distance $|x - s_j|$, thus capturing the reduced effectiveness of a sensor over its range. We consider this range to be finite and denoted by r_j. Therefore, we set $p_j(x,s_j) = 0$ when $|x - s_j| > r_j$. Although our analysis is not affected by the precise sensing model $p_j(x,s_j)$, we will limit ourselves to a linear decay model as follows:

$$p_j(x,s_j) = \max\left\{1 - \frac{|s_j - x|}{r_j}, 0\right\} \tag{2}$$

Unlike the persistent monitoring problem setting in [2], here we consider a known finite set of targets located at $x_i \in (0,L)$, $i = 1,\ldots,M$ (we assume $M > N$ to avoid uninteresting cases where there are at least as many agents as targets, in which case every target can be assigned to at least one agent). We can then set $p_j(x_i,s_j(t)) \equiv p_{ij}(s_j(t))$ to represent the effectiveness with which agent j can sense target i when located at $s_j(t)$. Accordingly, the joint probability that $x_i \in (0,L)$ is sensed by all N agents simultaneously (assuming detection independence) is

$$P_i(\mathbf{s}(t)) = 1 - \prod_{j=1}^{N}[1 - p_{ij}(s_j(t))] \tag{3}$$

where we set $\mathbf{s}(t) = [s_1(t),\ldots,s_N(t)]^{\mathrm{T}}$. Next, we define uncertainty functions $R_i(t)$ associated with targets $i = 1,\ldots,M$, so that they have the following properties: (i) $R_i(t)$ increases with a prespecified rate A_i if $P_i(\mathbf{s}(t)) = 0$ (we will later allow this to be a random process $\{A_i(t)\}$), (ii) $R_i(t)$ decreases with a fixed rate B_i if $P_i(\mathbf{s}(t)) = 1$ and (iii) $R_i(t) \geq 0$ for all t. It is then natural to model uncertainty dynamics associated with each target as follows:

$$\dot{R}_i(t) = \begin{cases} 0 & \text{if } R_i(t) = 0, A_i \leq B_i P_i(\mathbf{s}(t)) \\ A_i - B_i P_i(\mathbf{s}(t)) & \text{otherwise} \end{cases} \tag{4}$$

where we assume that initial conditions $R_i(0)$, $i = 1,\ldots,M$, are given and that $B_i > A_i > 0$.

Our goal is to control the movement of the N agents through $u_j(t)$ in (1) so that the cumulative average uncertainty over all targets $i = 1,\ldots,M$ is minimized over a fixed time horizon T. Thus, setting $\mathbf{u}(t) = [u_1(t),\ldots,u_N(t)]^{\mathrm{T}}$ we aim to solve the following optimal control problem **P1**:

$$\min_{\mathbf{u}(t)} \quad J = \frac{1}{T}\int_0^T \sum_{i=1}^{M} R_i(t)dt \tag{5}$$

subject to the agent dynamics (1), uncertainty dynamics and control constraint (4), and state constraints (e.g. no crossing).

III. Optimal control solution

In this section, we derive properties of the optimal control solution of problem **P1** and show that it can be reduced to a parametric optimization problem. This will allow us to utilize an IPA gradient estimation approach [9] to find a complete optimal solution through a gradient-based algorithm. We begin by defining the state vector $\mathbf{x}(t) = [R_1(t),...R_M(t), s_1(t)...s_N(t)]$ and associated costate vector $\lambda = [\lambda_1(t),...,\lambda_M(t), \lambda_{s_1}(t),...,\lambda_{s_N}(t)]$. Due to the discontinuity in the dynamics of $R_i(t)$ in (4), 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 such an interior arc. Using (1) and (4), the Hamiltonian and costate equations are

$$H(\mathbf{x},\lambda,\mathbf{u}) = \sum_{i=1}^{M} R_i(t) + \sum_{i=1}^{M}\lambda_i(t)\dot{R}_i(t) + \sum_{j=1}^{N}\lambda_{s_j}(t)u_j(t) \tag{6}$$

Applying the Pontryagin Minimum Principle to (6) with $\mathbf{u}^\star(t)$, $t \in [0,T)$, denoting an optimal control, a necessary condition for optimality is

$$u_j^*(t) = \begin{cases} 1 & \text{if } \lambda_{s_j}(t) < 0 \\ -1 & \text{if } \lambda_{s_j}(t) > 0 \end{cases} \tag{7}$$

Note that there exists a possibility that $\lambda_{s_j}(t) = 0$ over some finite singular intervals [14], in which case $u_j^*(t)$ may take values in $\{-1,0,1\}$.

Similar to the case of the persistent monitoring problem studied in [2], the complete solution requires solving the state and costate equations, which in turn involves the determination of all points where $R_i(t) = 0$, $i = 1,\ldots,M$. This generally involves the solution of a two-point-boundary-value problem. However, we will next prove some structural properties of an optimal trajectory, based on which we show that it is fully characterized by a set of parameters, thus reducing the optimal control problem to a much simpler parametric optimization problem.

We begin by assuming that targets are ordered according to their location so that $x_1 < \cdots < x_M$. Let $r = \max_{j=1,\ldots,N}\{r_j\}$ and $a = \max\{0, x_1 - r\}$, $b = \min\{L, x_M + r\}$. Thus, if $s_j(t) < x_1 - r$ or $s_j(t) > x_M + r$, then it follows from (2) that $p_{ij}(s_j(t)) = 0$ for all targets $i = 1,\ldots,M$. Clearly, this implies that the effective mission space is $[a,b]$. We will show next that on an optimal trajectory every agent is constrained to move within the interval $[x_1, x_M]$. To establish this and subsequent results, we will make a technical assumption that no two events altering the dynamics in this system can occur at the exact same time.

Assumption 1: Suppose that an agent switches direction at $\theta \in [a,b]$. For any $j = 1,\ldots,N$, $i = 1,\ldots,M$, $t \in (0,T)$, and any $\epsilon > 0$, if $s_j(t) = \theta$, $s_j(t-\epsilon) > \theta$ or if $s_j(t) = \theta$, $s_j(t-\epsilon) < \theta$, then either $R_i(\tau) > 0$ for all $\tau \in [t-\epsilon, t]$ or $R_i(\tau) = 0$ for all $\tau \in [t-\epsilon, t]$.

Proposition 1: In an optimal trajectory, $x_1 \leq s_j^*(t) \leq x_M$, $t \in [0,T]$, $j = 1,\ldots,N$.

Due to space limitations all proofs (some of which have many technical details) are omitted but may be found in [15].

The next result excludes singular arcs from an agent's trajectory while this agent has no target in its sensing range.

Lemma 1: If $|s_j(t) - x_i| > r_j$ for any $i = 1,\ldots,M$, then $u_j^*(t) \neq 0$.

The next lemma establishes the fact that if the agent is visiting an *isolated* target and experiences a singular arc, then the corresponding optimal control is $u_j^*(t) = 0$. An isolated target with position x_i is defined to be one that satisfies $|x_i - x_j| > 2r$, for all $j \neq i$ where r was defined earlier as $r = \max_{j=1,\dots,N}\{r_j\}$. Accordingly, the subset $I \subseteq \{1,\dots,M\}$ of isolated targets is defined as $I = \{i : |x_i - x_j| > 2r, j \neq i \in \{1,\dots,M\}, r = \max_{j=1,\dots,N}\{r_j\}\}$

Lemma 2: Let $|s_j^*(t) - x_k| < r_j$ for some $j = 1,\dots,N$ and isolated target $k \in I$. If $\lambda_{s_j}^*(t) = 0$, $t \in [t_1, t_2]$, then $u_j^*(t) = 0$.

We can further establish the fact that if an agent j experiences a singular arc while sensing an isolated target k, then the optimal point to stop is such that $s_j^*(t) = x_k$.

Proposition 2: Let $|s_j^*(t) - x_k| < r_j$ for some $j = 1,\dots,N$ and isolated target $k \in I$. If $\lambda_{s_j}^*(t) = 0, t \in [t_1, t_2]$, and $u_j^*(t_1^-) = u_j^*(t_2^+)$, then $s_j^*(t) = x_k$, $t \in [t_1, t_2]$.

Finally, we consider the case with the state constraint (no agent crossing). We can then prove that this constraint is never active on an optimal trajectory, i.e., agents reverse their directions before making contact with any other agent.

Proposition 3: Under the constraint $s_j(t) \leq s_{j+1}(t)$, on an optimal trajectory, $s_j(t) \neq s_{j+1}(t)$ for all $t \in (0, T)$.

Based on this analysis, we can parameterize **P1** so that the cost in (5) depends on a set of (i) *switching points* where an agent switches its control from $u_j(t) = \pm 1$ to ∓ 1 or possibly 0, and (ii) *dwelling times* if an agent switches from $u_j(t) = \pm 1$ to 0. In other words, the optimal trajectory of each agent j is totally characterized by two parameter vectors: switching points $\boldsymbol{\theta}_j = [\theta_{j1}, \theta_{j2} \dots \theta_{j\Gamma}]$ and dwelling times $\boldsymbol{\omega}_j = [\omega_{j1}, \omega_{j2} \dots \omega_{j\Gamma'}]$ where Γ and Γ' are prior parameters depending on the given time horizon. This defines a hybrid system with state dynamics (1), (4). The dynamics remain unchanged in between events, i.e., the points $\theta_{j1}, \dots, \theta_{j\Gamma}$ above and instants when $R_i(t)$ switches from > 0 to 0 or vice versa. Therefore, the overall cost function (5) can be parametrically expressed as $J(\boldsymbol{\theta}, \boldsymbol{\omega})$ and rewritten as the sum of costs over corresponding inter-event intervals over a given time horizon:

$$J(\boldsymbol{\theta}, \boldsymbol{\omega}) = \frac{1}{T} \sum_{k=0}^{K} \int_{\tau_k(\boldsymbol{\theta},\boldsymbol{\omega})}^{\tau_{k+1}(\boldsymbol{\theta},\boldsymbol{\omega})} \sum_{i=1}^{M} R_i(t) dt \qquad (8)$$

where τ_k is the k-th event time. This will allow us to apply IPA to determine a gradient $\nabla J(\boldsymbol{\theta}, \boldsymbol{\omega})$ with respect to these parameters and apply any standard gradient-based optimization algorithm to obtain a (locally) optimal solution.

IV. Infinitesimal Perturbation Analysis

As concluded in the previous section, optimal agent trajectories may be selected from the family $\{\boldsymbol{s}(\boldsymbol{\theta}, \boldsymbol{\omega}, t, \boldsymbol{s}_0)\}$ with parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ and a given initial condition $\boldsymbol{s}_0$. Along these trajectories, agents are subject to dynamics (1) and targets are subject to (4). An *event* (e.g., an agent stopping at some target x_i) occurring at time $\tau_k(\boldsymbol{\theta}, \boldsymbol{\omega})$ triggers a switch in these state dynamics. IPA specifies how changes in $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ influence the state $\boldsymbol{s}(\boldsymbol{\theta}, \boldsymbol{\omega}, t, \boldsymbol{s}_0)$, as well as event times $\tau_k(\boldsymbol{\theta}, \boldsymbol{\omega})$, $k = 1, 2, \dots$, and, ultimately the cost function (8). We briefly review next the IPA framework for general stochastic hybrid systems as presented in [9].

Let $\{\tau_k(\theta)\}$, $k = 1, \dots, K$, denote the occurrence times of all events in the state trajectory of a hybrid system with dynamics $\dot{x} = f_k(x, \theta, t)$ over an interval $[\tau_k(\theta), \tau_{k+1}(\theta))$, where $\theta \in \Theta$ is some parameter vector and Θ is a given compact, convex set. For convenience, we set $\tau_0 = 0$ and $\tau_{K+1} = T$. We use the Jacobian matrix notation: $x'(t) \equiv \frac{\partial x(\theta,t)}{\partial \theta}$ and $\tau_k' \equiv \frac{\partial \tau_k(\theta)}{\partial \theta}$, for all state and event time derivatives. It is shown in [9] that

$$\frac{d}{dt} x'(t) = \frac{\partial f_k(t)}{\partial x} x'(t) + \frac{\partial f_k(t)}{\partial \theta}, \qquad (9)$$

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

$$x'(\tau_k^+) = x'(\tau_k^-) + [f_{k-1}(\tau_k^-) - f_k(\tau_k^+)]\tau_k' \qquad (10)$$

for $k = 0, \dots K$. In order to complete the evaluation of $x'(\tau_k^+)$ in (10), we need to determine τ_k'. If the event at τ_k is *exogenous*, $\tau_k' = 0$. However, if the event is *endogenous*, 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

$$\tau_k' = -[\frac{\partial g_k}{\partial x} f_k(\tau_k^-)]^{-1} (\frac{\partial g_k}{\partial \theta} + \frac{\partial g_k}{\partial x} x'(\tau_k^-)) \qquad (11)$$

as long as $\frac{\partial g_k}{\partial x} f_k(\tau_k^-) \neq 0$. (details may be found in [9]).

In our setting, the time-varying cost along a given trajectory is $\sum_{i=1}^{M} R_i(t)$ from (8), which is not an explicit function of the state $\mathbf{x}(t) = [R_1(t), \dots R_M(t), s_1(t) \dots s_N(t)]$. The gradient $\nabla J(\boldsymbol{\theta}, \boldsymbol{\omega}) = [\frac{\partial J(\boldsymbol{\theta},\boldsymbol{\omega})}{\partial \boldsymbol{\theta}}, \frac{\partial J(\boldsymbol{\theta},\boldsymbol{\omega})}{\partial \boldsymbol{\omega}}]^{\mathrm{T}}$ reduces to

$$\nabla J(\boldsymbol{\theta}, \boldsymbol{\omega}) = \frac{1}{T} \sum_{k=0}^{K} \sum_{i=1}^{M} \int_{\tau_k(\boldsymbol{\theta},\boldsymbol{\omega})}^{\tau_{k+1}(\boldsymbol{\theta},\boldsymbol{\omega})} \nabla R_i(t) dt \qquad (12)$$

where $\nabla R_i(t) = [\frac{\partial R_i(t)}{\partial \boldsymbol{\theta}}, \frac{\partial R_i(t)}{\partial \boldsymbol{\omega}}]^{\mathrm{T}}$.

Applying (9), (10), (11), we can evaluate $\nabla R_i(t)$. In contrast to [2], in our problem agents are allowed to dwell on every target and IPA will optimize these dwelling times. Therefore, we need to consider all possible forms of control sequences: (i) $\pm 1 \to 0$, (ii) $0 \to \pm 1$, and (iii) $\pm 1 \to \mp 1$. We can then obtain from (4):

$$\frac{\partial R_i(t)}{\partial \theta_{j\xi}} = \frac{\partial R_i(\tau_k^+)}{\partial \theta_{j\xi}} - \begin{cases} 0 & \text{if } R_i(t) = 0, A_i < B_i P_i(\mathbf{s}(t)) \\ B_i \frac{\partial p_{ij}(s_j)}{\partial s_j} \frac{\partial s_j(\tau_k^+)}{\partial \theta_{j\xi}} G(t) & \text{otherwise} \end{cases} \qquad (13)$$

where $G(t) = \int_{\tau_k}^{t} \prod_{d \neq j} [1 - p_{id}(s_d(t))] dt$ and $\frac{\partial p_{ij}(s_j)}{\partial s_j} = \pm \frac{1}{r_j}$ or 0. A similar derivation applies to $\frac{\partial R_i(t)}{\partial \omega_{j\xi}}$.

First, let us consider the events that cause switches in $\dot{R}_i(t)$ in (4) at time τ_k. For these events, the dynamics of $s_j(t)$ are continuous so that $\nabla s_j(\tau_k^-) = \nabla s_j(\tau_k^+)$. For target i,

$$\nabla R_i(\tau_k^+) = \begin{cases} \nabla R_i(\tau_k^-) & \text{if } \dot{R}_i(\tau_k^-) = 0, \\ & \dot{R}_i(\tau_k^+) = A_i - B_i P_i(\boldsymbol{s}(\tau_k^+)). \\ 0 & \text{if } \dot{R}_i(\tau_k^-) = A_i - B_i P_i(\boldsymbol{s}(\tau_k^-)), \\ & \dot{R}_i(\tau_k^+) = 0. \end{cases} \qquad (14)$$

Second, let us consider events that cause switches in $\dot{s}_j(t) = u_j(t)$ at time τ_k. For these events, the dynamics of $R_i(t)$ are continuous so that $\nabla R_i(\tau_k^-) = \nabla R_i(\tau_k^+)$. In order to evaluate (13) and $\frac{\partial R_i(t)}{\partial \omega_{j\xi}}$, we need $\frac{\partial s_j(\tau_k^+)}{\partial \theta_{j\xi}}$ and $\frac{\partial s_j(\tau_k^+)}{\partial \omega_{j\xi}}$. Clearly, these are not affected by future events and we only have to consider the current and prior control switches. Let $\theta_{j\xi}$ and $\omega_{j\xi}$ be the current switching point and dwelling time. Again, applying (9), (10), (11), we have

Case 1: $u_j(\tau_k^-) = \pm 1, u_j(\tau_k^+) = 0$

$$\frac{\partial s_j}{\partial \theta_{jl}}(\tau_k^+) = \begin{cases} 1 & \text{if } l = \xi \\ 0 & \text{if } l < \xi \end{cases} \tag{15}$$

$$\frac{\partial s_j}{\partial \omega_{jl}}(\tau_k^+) = 0 \quad \text{for all } l \leq \xi \tag{16}$$

Case 2: $u_j(\tau_k^-) = 0, u_j(\tau_k^+) = \pm 1$

$$\frac{\partial s_j}{\partial \theta_{jl}}(\tau_k^+) = \begin{cases} \frac{\partial s_j}{\partial \theta_{jl}}(\tau_k^-) - u_j(\tau_k^+) sgn\big(\theta_{j\xi} - \theta_{j(\xi-1)}\big) \text{ if } l = \xi \\ \frac{\partial s_j}{\partial \theta_{jl}}(\tau_k^-) - u_j(\tau_k^+)\Big[sgn(\theta_{jl} - \theta_{j(l-1)}) \\ \qquad -sgn(\theta_{j(l+1)} - \theta_{jl})\Big] \text{ if } l < \xi \end{cases} \tag{17}$$

$$\frac{\partial s_j}{\partial \omega_{jl}}(\tau_k^+) = -u_j(\tau_k^+) \quad \text{for all } l \leq \xi \tag{18}$$

Case 3: $u_j(\tau_k^-) = \pm 1, u_j(\tau_k^+) = \mp 1$

$$\frac{\partial s_j}{\partial \theta_{jl}}(\tau_k^+) = \begin{cases} 2 & \text{if } l = \xi \\ -\frac{\partial s_j}{\partial \theta_{jl}}(\tau_k^-) & \text{if } l < \xi \end{cases} \tag{19}$$

Details of these derivations can be found in [2]. An important difference arises in *Case 2* above, where $\tau_k = |\theta_{j1} - a| + \omega_{j1} + ... + |\theta_{j\xi} - \theta_{j(\xi-1)}| + \omega_{j\xi}$. We eliminate the constraints on the switching location that $\theta_{j\xi} \leq \theta_{j(\xi-1)}$ if ξ is even and $\theta_{j\xi} \geq \theta_{j(\xi-1)}$ if ξ is odd.

The event excitation problem. Note that all derivative updates above happen exclusively at events occurring at times $\tau_k(\boldsymbol{\theta}, \boldsymbol{\omega})$, $k = 1, 2, \ldots$. Thus, this approach scales with the number of events characterizing the hybrid system, not its state space. While this is a distinct advantage, it also involves a potential drawback. In particular, it assumes that the events involved in IPA updates are observable along a state trajectory. However, if the current trajectory never reaches the vicinity of any target so as to be able to sense it and affect the overall uncertainty cost function, then any small perturbation to the trajectory will have no effect on the cost. As a result, IPA will fail as illustrated in Fig. 1: here, the single agent trajectory $s_1(\boldsymbol{\theta}, \boldsymbol{\omega}, t)$ is initially limited (yellow segment) to include no event. Thus, if a gradient-based procedure is initialized with such $s_1(\boldsymbol{\theta}, \boldsymbol{\omega}, t)$, no event involved in the evaluation of $\nabla R_i(t)$ is "excited" and the cost gradient remains zero.

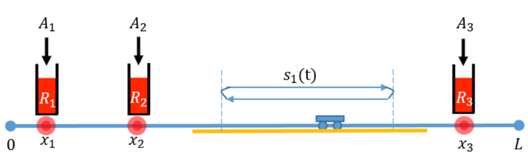

Fig. 1: An example of no event excitation.

In order to overcome this problem, we propose a modification of our cost metric by introducing a function $V(\cdot)$ with the property of "spreading" the value of some $R_i(t)$ over all points $w \in \Omega \equiv [0, L]$. Recalling Proposition 1, we limit ourselves to the subset $\mathcal{B} = [x_1, x_M] \subset \Omega$. Then, for all points $w \in \mathcal{B}$, we define $V(w,t)$ as a continuous density function which results in a total value equivalent to the weighted sum of the target values $\sum_{i=1}^{M} R_i(t)$. We impose the condition that $V(w,t)$ be monotonically decreasing in the Euclidean distance $\|w - x_i\|$. More precisely, we define $d_i^+(w) = \max\big(\|w - x_i\|, r\big)$ where $r = \min_{j=1,\ldots,N}\{r_j\}$ which ensures that $d_i^+(w) \geq r$. Thus, $d_i^+(w) = r > 0$ is fixed for all points within the target's vicinity, $w \in [x_i - r, x_i + r]$. We define

$$V(w,t) = \sum_{i=1}^{M} \frac{\alpha_i R_i(t)}{d_i^+(w)} \tag{20}$$

Note that $V(w,t)$ corresponds to the "total weighted reward density" at $w \in \mathcal{B}$. The weight α_i may be included to capture the relative importance of targets, but we shall henceforth set $\alpha_i = 1$ for all i for simplicity. In order to differentiate points $w \in \mathcal{B}$ in terms of their location relative to the agents states $s_j(t)$, $j = 1, \ldots, N$, we also define the travel cost function

$$Q(w, \mathbf{s}(t)) = \sum_{j=1}^{N} \|s_j(t) - w\| \tag{21}$$

Using these definitions we introduce a new objective function component, which is added to the objective function in (5):

$$J_2(t) = \int_{\mathcal{B}} Q(w, \mathbf{s}(t)) V(w,t) dw \tag{22}$$

The significance of $J_2(t)$ is that it accounts for the movement of agents through $Q(w, \mathbf{s}(t))$ and captures the target state values through $V(w,t)$. Introducing this term in the objective function in the following creates a non-zero gradient even if the agent trajectories are not passing through any targets. We now incorporate the new metric $J_2(t)$ into (8) as follows:

$$\min_{\boldsymbol{\theta} \in \Theta, \boldsymbol{\omega} \geq 0} J(\boldsymbol{\theta}, \boldsymbol{\omega}, T) = \frac{1}{T} \int_0^T \Big[J_1(\boldsymbol{\theta}, \boldsymbol{\omega}, t) + e^{-\beta t} J_2(\boldsymbol{\theta}, \boldsymbol{\omega}, t)\Big] dt \tag{23}$$

where $J_1(\boldsymbol{\theta}, \boldsymbol{\omega}, t) = \sum_{i=1}^{M} R_i(t)$ is the original uncertainty metric. This creates a continuous potential field for the agents which ensures a non-zero cost gradient even when the trajectories do not excite any events. The factor $e^{-\beta t}$ with $\beta > 0$ is included so that as the number of IPA iterations increases, the effect of $J_2(\boldsymbol{\theta}, \boldsymbol{\omega}, t)$ is diminished and the original objective is ultimately recovered. The IPA derivative of $J_2(\boldsymbol{\theta}, \boldsymbol{\omega}, t)$ is

$$\frac{\partial J_2}{\partial \boldsymbol{\theta}} = \int_{\mathcal{B}} \Big[\frac{\partial Q(w, \boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{s}(t), t)}{\partial \boldsymbol{\theta}} V(w, \boldsymbol{\theta}, \boldsymbol{\omega}, t) + Q(w, \boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{s}(t), t) \frac{\partial V(w, \boldsymbol{\theta}, \boldsymbol{\omega}, t)}{\partial \boldsymbol{\theta}}\Big] dw \tag{24}$$

where the derivatives of $Q(w, \boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{s}(t), t)$ and $V(w, \boldsymbol{\theta}, \boldsymbol{\omega}, t)$ are obtained following the same procedure described previously. Before making this modification, the lack of event excitation in a state trajectory results in the total derivative (12) being zero. In (24) we observe that if no events occur, the second part in the integral, which involves $\frac{\partial V(\cdot)}{\partial \boldsymbol{\theta}}$ is zero, since $\sum_{i=1}^{M} \frac{\partial R_i(t)}{\partial \boldsymbol{\theta}} = 0$ at all t. However, the first part in the integral does not depend on events, but only the sensitivity of $Q(w, \boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{s}(t), t)$ in (21) with respect to the parameters $\boldsymbol{\theta}, \boldsymbol{\omega}$. As a result, agent trajectories are adjusted so as to eventually excite desired events and any gradient-based procedure we use in conjunction with IPA is no longer limited by the absence of event excitation.

IPA robustness to uncertainty modeling. Observe that the evaluation of $\nabla R_i(t)$, hence $\nabla J(\boldsymbol{\theta}, \boldsymbol{\omega})$, is *independent* of A_i, $i = 1, \ldots, M$, i.e., the parameters in our uncertainty model. In fact, the dependence of $\nabla R_i(t)$ on A_i 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 (4). 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 $\{A_i(t)\}$. Under mild technical conditions on the statistical characteristics of $\{A_i(t)\}$ [9], the resulting $\nabla J(\boldsymbol{\theta}, \boldsymbol{\omega})$ is an unbiased estimate of a stochastic gradient.

V. Graph-based scheduling method

While the IPA-driven gradient-based approach described in Sec. IV offers several compelling advantages, it is not guaranteed to find a global optimum. In addition, it has been shown that in mission spaces of dimension greater than one, optimal trajectories cannot be described parametrically [3]. This motivates the use of an alternative approach where the targets are viewed as discrete tasks, leading naturally to a graph-based description of the problem [5], [16]–[19]. This higher level of abstraction allows one to guarantee an optimal solution, though at the cost of a significant increase in computational complexity.

Our approach to the discrete setting is to divide the overall planning time horizon T for agent j into a sum of K_j consecutive time steps $\{t_j^1, t_j^2, ..., t_j^{K_j}\}$, $j = 1, \ldots, N$, with $t_j^1 = 0$. The dependence on j indicates that each agent may have a different discretization. We denote the end of the K-th step as $t_j^{K+1} = T$. Each step $k \in \{1, ..., K_j\}$ begins with a travel stage where the agent moves to a particular target i. Under the assumption that during the transition between targets each agent moves at its maximum speed, the travel time is $\Delta t_j^k = |s_j^k(t_j^k) - x_i|$. Upon arriving at a target, the agent dwells for a time Δd_j^k. Due to the range-based nature of the sensing, the uncertainty actually decreases before the arrival of the agent at the target and after the agent has departed until the target is out of the sensing range.

The problem of optimizing u_j to minimize (5) can be translated into a mixed integer programming (MIP) problem to select the sequence of targets and the dwell time at each target. Letting a_{ji}^k be a binary variable denoting whether agent j is assigned to target i at time step k, this MIP is

$$\min_{a_{ji}^k, \Delta d_j^k} J = \frac{1}{T} \sum_{i=1}^{M} \int_0^T R_i(t) dt \tag{25}$$

subject to $a_{ji}^k \in \{1, 0\}$, $\sum_{i=1}^{M} a_{ji}^k = 1, \forall j, k$, and $\sum_{k=1}^{K} \Delta t_j^k + \Delta d_j^k \leq T, \forall j$. Assuming that each agent is assigned to a maximum of only one target at any one time, we break the solution of this problem into three parts: enumeration of all feasible trajectories, calculation of the cost of the feasible trajectories, and selection of the optimal trajectory.

The first part is to determine feasible trajectories. Given the fixed time horizon T, the target locations, the initial locations of the agent, and the maximum speed, a feasible trajectory is one where the sequence of targets can all be visited within the time horizon.

In the second part, the cost of each feasible trajectory must be determined. Suppose we have a given feasible trajectory with K targets in its sequence. Note that K may be larger than m (and may be much larger for large time horizons). Let $\{i_1, i_2, \ldots, i_K\}$ denote the indices of the targets in the sequence. From (25), the cost of this trajectory optimization is then only constrained by $\sum_{k=1}^{K} \Delta t_j^k + \Delta d_j^k \leq T, \forall j$.

Our approach to solving this relaxed problem is to set up a recursive calculation. The travel times Δt_i are completely determined by the sequence alone. Assume for the moment that the switching times through t_{K-1} have been determined (and thus the first $K-2$ dwell times, $\Delta d^1, \ldots, \Delta d^{K-2}$ are known). The two final dwell times are completely determined by selecting the time t_K at which to switch the agent from target i_{K-1} to target i_K. This gives us a simple single variable optimization problem $\min_{\Delta T_K} J = \frac{1}{\Delta T} \int_{t^{K-1}}^{T} (R_{i_{K-1}}(t) + R_{i_K}(t)) dt$ where $\Delta T = T - t_{K-1}$, which allows the final switching time to be expressed as a function of the previous time $t_K = f(t_{K-1})$. Repeating this leads to an expression of the optimal switching times as a nested sequence of optimization functions which can be solved numerically.

This same optimization procedure can be generalized to the case of multiple agents. The primary challenge is that the set of feasible trajectories, and the calculation of the cost of those trajectories, quickly becomes intractable since all possible combinations of assignments of multiple agents must be considered. In prior work on linear systems, it was shown that an appropriately defined periodic schedule is sufficient to ensure the system remains controllable [20], [21]. In the current context, this implies keeping $R_i(t), i = 1, ..., M$, close to zero. Thus, we apply this approach over short horizons and, if the resulting trajectories are periodic, we repeat them over longer horizons.

VI. Simulation Examples

To demonstrate the performance of the gradient-based algorithm using the IPA scheme described in Sec. IV, we present two sets of numerical examples. The first set uses deterministic target locations and dynamics. The results are compared against the optimal found by the discrete scheduling algorithm of Sec. V. The second set demonstrates the robustness of the IPA scheme with respect to a stochastic uncertainty model. Additional examples are included in [15].

The first example consists of a single agent performing a persistent monitoring task on three targets over a time horizon of 100 seconds. The targets are located at $x_1 = 5$, $x_2 = 10$, $x_3 = 15$ and their uncertainty dynamics in (4) are defined by the parameters $A_i = 1$, $B_i = 5$, and $R_i(0) = 1$ for all i. The agent has a sensing range of 2 and is initialized with $s(0) = 0$, $u(0) = 1$. The results from the IPA gradient descent approach are shown in Fig. 2. The corresponding result based on the discrete setting of Sec. V is essentially the same with the agent moving through the three targets in a periodic fashion as shown in Fig. 3.

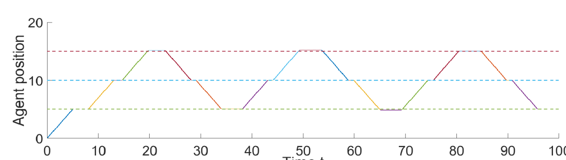

Fig. 2: Optimal trajectory of a single agent monitoring three targets using the IPA gradient descent. The final cost is 26.11.

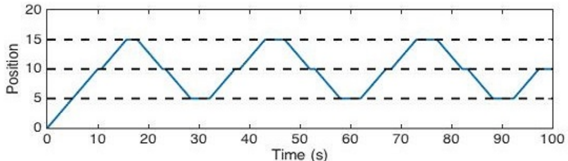

Fig. 3: A single agent monitoring three targets using the optimal discrete assignment and dwelling time. The final cost is 25.07.

The next example involves 2 agents and 5 targets over a time horizon of 500 seconds. The targets are located at $x_1 = 5$, $x_2 = 7$, $x_3 = 9$, $x_4 = 13$, $x_5 = 15$. The uncertainty dynamics are the same as in the single agent, three target case. All agents have a sensing range of 2 and are initialized at $s_1(0) = s_2(0) = 0$, with $u_1(0) = u_2(0) = 1$. The results from the IPA gradient descent approach are shown in Fig. 4.

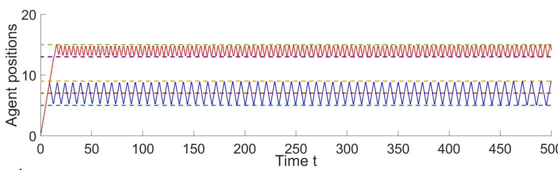

Fig. 4: Optimal trajectory of two agents monitoring five targets using the IPA gradient descent algorithm. The final cost is 4.99.

As mentioned earlier, the IPA robustness property allows us to handle stochastic uncertainty models at targets. We show a one-agent example in Fig. 5 where the uncertainty inflow rate $A_i(t)$ is uniformly distributed over $[0, 2]$ for all targets. Additional examples are included in [15].

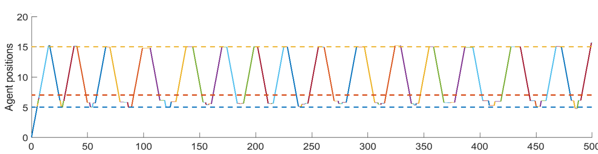

Fig. 5: Example with stochastic uncertainty processes: $A_i \sim U(0, 2)$, $J^*(\boldsymbol{\theta}, \boldsymbol{\omega}) = 42.46$.

The event excitation issue is addressed in Fig. 6, where the agent trajectory is initialized so that it is not close to any of the targets. Using the original problem formulation (without the inclusion of $J_2(\boldsymbol{\theta}, \boldsymbol{\omega}, t)$ in (23)), the initial trajectory and cost remain unchanged. After event excitation, the cost reduces to 30.24 which is close to the optimal cost.

VII. Conclusion

We have formulated a persistent monitoring problem with the objective of controlling the movement of multiple cooperating agents so as to minimize an uncertainty metric associated with a finite number of targets. We have established properties of the optimal control solution which reduce the problem to a parametric optimization one. A complete online solution is given by Infinitesimal Perturbation Analysis (IPA) to evaluate the gradient of the objective function with respect to all parameters. We also address the case when IPA gradient estimation fails because of the lack of event excitation. We solve this problem by proposing a new metric for the objective function which creates a potential field guaranteeing that gradient values are non-zero. This approach is compared to an alternative graph-based task scheduling algorithm for determining an optimal sequence of target visits. Ongoing research is investigating how to extend these methodologies to higher dimensional mission spaces.

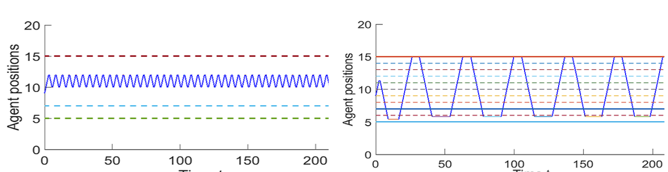

Fig. 6: Left: a trajectory where IPA fails due to lack of event excitation. Right: an optimal trajectory after event excitation.

References

[1] M. Zhong and C. G. Cassandras, "Distributed coverage control and data collection with mobile sensor networks," *IEEE Trans. on Automatic Control*, vol. 56, no. 10, pp. 2445–2455, 2011.

[2] C. G. Cassandras, X. Lin, and X. Ding, "An optimal control approach to the multi-agent persistent monitoring problem," *IEEE Trans. on Automatic Control*, vol. 58, no. 4, pp. 947–961, 2013.

[3] X. Lin and C. G. Cassandras, "An optimal control approach to the multi-agent persistent monitoring problem in two-dimensional spaces," *IEEE Trans. on Automatic Contr.*, vol. 60, no. 6, pp. 1659–1664, 2015.

[4] N. Michael, E. Stump, and K. Mohta, "Persistent surveillance with a team of mavs," in *Proc. IEEE/RSJ Intl. Conf. Intelligent Robots Systems*, 2011, pp. 2708–2714.

[5] S. L. Smith, M. Schwager, and D. Rus, "Persistent monitoring of changing environments using a robot with limited range sensing," in *Proc. IEEE Intl Conf. on Robotics and Automation*, 2011, pp. 5448–5455.

[6] Z. Shen and S. B. Andersson, "Tracking Nanometer-Scale Fluorescent Particles in Two Dimensions With a Confocal Microscope," *IEEE Trans. on Contr. Syst. Tech.*, vol. 19, no. 5, pp. 1269–1278, 2011.

[7] J. Yu, S. Karaman, and D. Rus, "Persistent monitoring of events with stochastic arrivals at multiple stations," *IEEE Trans. on Robotics*, vol. 31, no. 3, pp. 521–535, 2015.

[8] E. Stump and N. Michael, "Multi-robot persistent surveillance planning as a vehicle routing problem," in *Proc. IEEE Conf. on Automation Science and Engineering*, 2011, pp. 569–575.

[9] 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. 642–661, 2010.

[10] 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, 2010.

[11] M. Schwager, D. Rus, and J.-J. Slotine, "Decentralized, adaptive coverage control for networked robots," *The Intl. Journal of Robotics Research*, vol. 28, no. 3, pp. 357–375, 2009.

[12] K.-K. Oh and H.-S. Ahn, "Formation control and network localization via orientation alignment," *IEEE Trans. on Automatic Control*, vol. 59, no. 2, pp. 540–545, 2014.

[13] Y. Khazaeni and C. G. Cassandras, "Event excitation for event-driven control and optimization of multi-agent systems," in *Proc. IEEE Intl. Workshop on Discrete Event Systems*, 2016, pp. 197–202.

[14] A. E. Bryson and Y.-C. Ho, *Applied optimal control: optimization, estimation and control*. CRC Press, 1975.

[15] N. Zhou, X. Yu, S. B. Andersson, and C. G. Cassandras, "Optimal event-driven multi-agent persistent monitoring of a finite set of targets," *preprint arXiv: http://arxiv.org/abs/1603.04128*, 2016.

[16] X. Lan and M. Schwager, "Planning periodic persistent monitoring trajectories for sensing robots in gaussian random fields," in *Proc. IEEE Intl. Conf. on Robotics and Automation*, 2013, pp. 2415–2420.

[17] M. Lahijanian, J. Wasniewski, S. B. Andersson, and C. Belta, "Motion planning and control from temporal logic specifications with probabilistic satisfaction guarantees," in *Proc. IEEE Intl. Conf. on Robotics and Automation*, 2010, pp. 3227–3232.

[18] S. L. Smith, J. Tumova, C. Belta, and D. Rus, "Optimal path planning for surveillance with temporal-logic constraints," *International Journal of Robotics Research*, vol. 30, no. 14, pp. 1695–1708, 2011.

[19] N. Mathew, S. L. Smith, and S. L. Waslander, "A graph-based approach to multi-robot rendezvous for recharging in persistent tasks," in *Proc. IEEE Intl. Conf. on Robotics and Automation*, 2013, pp. 3497–3502.

[20] X. Yu and S. B. Andersson, "Effect of switching delay on a networked control system," in *Proc. IEEE Conf. on Decision and Control*, 2013, pp. 5945–5950.

[21] ——, "Preservation of system properties for networked linear, time-invariant control systems in the presence of switching delays," in *Proc. IEEE Conf. on Decision and Control*, 2014, pp. 5260–5265.