2019 IEEE 58th Conference on Decision and Control (CDC)
Palais des Congrès et des Expositions Nice Acropolis
Nice, France, December 11-13, 2019

Optimal Multi-Agent Persistent Monitoring of the Uncertain State of a Finite Set of Targets

Samuel C. Pinto[1], Sean B. Andersson[1,2], Julien M. Hendrickx[3], and Christos G. Cassandras[2,4]
[1]Dept. of Mechanical Engineering, [2]Division of Systems Engineering,
[4]Dept. of Electrical and Computer Engineering
Boston University, Boston, MA 02215, USA
[3]ICTEAM Institute, UCLouvain, Louvain-la-Neuve 1348, Belgium
{samcerq,sanderss,cgc}@bu.edu, julien.hendrickx@uclouvain.be

Abstract— **We approach the problem of persistent monitoring of a finite set of fixed targets located in a one-dimensional environment with internal, linear, stochastic dynamics. Monitoring is performed by a set of agents with limited sensing range and range-dependent sensing quality. The optimal estimator of the target dynamics from the agent measurements is the Kalman-Bucy Filter. We formulate an optimal control problem to minimize the estimation error across all the targets as a function of the trajectories of the agents. Using Hamiltonian analysis, the structure of the optimal controller is defined and, given this structure, we reformulate the problem as a hybrid systems optimization problem. Using Infinitesimal Perturbation Analysis (IPA), stochastic gradient estimates of the hybrid system are computed and gradient descent is used in order to achieve a locally optimal solution.**

I. INTRODUCTION

The general problem of multi-agent persistent monitoring involves a collection of mobile agents moving through a spatial domain to interact with targets at specific locations to, in some sense, control or monitor some state of those targets. This paradigm finds applications across a wide range of domains, ranging from smart cities, where one example goal may be to measure the evolving length of a traffic jam at specific intersections, to optical microscopy, where one may wish to track the location of individual biological macromolecules in a sample [1]–[4]. The dynamic and stochastic nature of these variables implies that they cannot be measured just once but rather must be monitored over time. We focus on situations where the number of available agents to do this monitoring is lower than the number of targets to be monitored, aiming to design an optimal motion policy for the agents that minimizes a measure of the uncertainty in the estimates of the states of the targets.

This cooperative persistent monitoring objective may be accomplished by assigning the agents to targets dynamically or by a periodic scheduling approach in which agent schedules define the sequence of targets to be visited and how long to spend at their current assignment [5]–[7]. So long as the number of targets and agents remains low, the scheduling approach can yield a globally optimal solution. However, off-the-shelf schemes for the periodic schedule approach, such as those developed for traveling salesman or vehicle routing problems, do not scale well as those numbers increase.

In this paper, motivated by the approach developed in our earlier works [8], [9], we use an optimal control framework whose objective is to control the motion of the agents so as to minimize a measure of the overall uncertainty. We no longer assume that the internal dynamics are represented by simple linear growth and decay and consider instead the more general and practically relevant situation where those dynamics are represented by a linear system driven by a Gaussian white noise process, as well as a sensing model which involves linear measurements corrupted by Gaussian noise whose signal to noise ratio depends on the range to the target being measured. This makes the model more reflective of real-world concerns, and is a step towards persistent monitoring of moving, dynamic targets.

Modeling the target state dynamics as linear, stochastic systems naturally leads to the cost function involving the uncertainty in the state estimates and places the problem in the realm of distributed filtering. A similar problem was explored in [10], [11] where a single agent with a position-dependent measurement model was used to minimize the estimation error of a spatiotemporal scalar field represented using a finite number of basis functions defined at specific locations in the environment, analogous to the targets considered in our work. In [10], the trajetory was planned with a cyclic version of rapidly exploring random trees while [11] used an optimal control approach, leading to a two-point boundary value problem. However, solving such boundary value problems is computationally expensive and does not scale well.

While we also consider an optimal control framework, rather than solving the corresponding boundary value problem, we restrain ourselves to targets distributed in a 1-D environment where the agent can see at most one target at a time in order to identify the *structure* of the optimal control and to reduce it to a parametric optimization problem.

This work was supported in part by NSF under grants ECCS-1509084, DMS-1664644, CNS-1645681, and CMMI-1562031, by ARPA-E's NEXTCAR program under grant DE-AR0000796, by AFOSR under grant FA9550-19-1-0158, and by the MathWorks. The work of J. Hendrickx was supported in part by Communaté française de Belgique - Actions de Recherche Concertées and a WBI World Excellence Fellowship.

We then apply Infinitesimal Perturbation Analysis (IPA) [12] to determine the gradient of the cost function with respect to the parameters defining the optimal control and to then obtain a (possibly local) optimal solution through a gradient descent scheme. Note that in many real world applications the agent's mobility is restricted to (possible multiple) single dimension spaces, such as cars on roads, underwater vehicles in rivers/waterways and powerline inspection agents.

A. Notation and Mathematical Preliminaries

To avoid confusion between scalar variables and vector variables, vectors are marked with an underscore and matrices are denoted in capital letters. $X^{(i,j)}$ denotes the $(i,j)^{th}$ entry of a matrix, the trace of a matrix X is denoted by $\mathrm{tr}(X)$ and $\mathrm{sgn}(x)$ is the signum function that returns the sign of its argument (or zero when the argument is zero). Since matrix derivatives will be used throughout the paper, we briefly recall their properties. The derivative $\frac{\partial \mathrm{tr}(G)}{\partial X}$ is a square matrix whose $(i,j)^{th}$ element is $\frac{\partial \mathrm{tr}(G)}{\partial X^{(i,j)}}$. Given a constant, square matrix A and a time-varying square matrix X, then we have [13]:

$$\frac{\partial \mathrm{tr}(AX)}{\partial X} = A^T, \ \frac{\partial \mathrm{tr}(X^T AX)}{\partial X} = AX + X^T A^T,$$

II. PROBLEM FORMULATION

Consider a collection of M fixed targets located at positions $x_1, ..., x_M \in \mathbb{R}$. Each target has an internal state $\underline{\phi} \in \mathbb{R}^{L_i}$ that evolves in time according to

$$\dot{\underline{\phi}}_i(t) = A_i \underline{\phi}_i(t) + \underline{w}_i(t), \tag{1}$$

where the w_i, $i = 1, \ldots, M$, are mutually independent, zero mean Gaussian white noise processes with $E[\underline{w}_i(t)\underline{w}_i(t)^T] = Q_i$ with Q_i a positive definite matrix.

In addition to the targets, we have N mobile agents whose positions at time t are denoted by $s_1(t), ..., s_N(t) \in \mathbb{R}$. These agents can move in the mission space following the dynamics

$$\dot{s}_j(t) = u_j(t), \ j = 1, ..., N, \tag{2}$$

with their speeds, after proper scaling, constrained by $|u_j(t)| \leq 1$. Agent j can observe the internal states of target i according to the model

$$\underline{z}_{i,j}(t) = \gamma_j \left(s_j(t) - x_i\right) H_i \underline{\phi}_i(t) + \underline{v}_{i,j}(t), \tag{3}$$

where the $\underline{v}_{i,j}$, $i = 1, \ldots, M$, $j = 1, \ldots, N$ are mutually independent zero mean Gaussian white noise processes, independent of $\underline{w}_i$, with $E[\underline{v}_{i,j}(t)\underline{v}_{i,j}^T(t)] = R_i$, R_i positive definite, and $\gamma_j(\cdot)$ a scalar function. The intuition behind the model described is that the noise power is constant but that the signal level varies as a function of the distance to the target. Although the analysis conducted is not heavily dependent on the specific form of $\gamma_j(\cdot)$, as long as it is unimodal (i.e. has only one peak) and has a finite support, we use the following definition for concreteness:

$$\gamma_j(\alpha) = \begin{cases} 0, & |\alpha| > r_j, \\ \sqrt{1 - \frac{|\alpha|}{r_j}}, & |\alpha| \leq r_j, \end{cases} \tag{4}$$

where r_j is the sensing radius of agent j. Then, the instantaneous signal to noise ratio (SNR) of a single measurement made by agent j is given by

$$\frac{E\left[(\underline{z}_{i,j}(t) - \underline{v}_{i,j}(t))^T(\underline{z}_{i,j}(t) - \underline{v}_{i,j}(t))\right]}{E[\underline{v}_{i,j}^T(t)\underline{v}_{i,j}(t)]} = \max\left(0, 1 - \frac{|s_j - x_i|}{r_j}\right) \frac{\underline{\phi}_i^T(t) H_i^T H_i \underline{\phi}_i(t)}{\mathrm{tr}(R_i)} \tag{5}$$

Notice that the term $\underline{\phi}_i^T(t) H_i^T H_i \underline{\phi}_i(t) (\mathrm{tr}(R_i))^{-1}$ is a deterministic scalar that depends on the current state of target i. The dependence on the position of the agent is completely captured by the max function and as a result the SNR is maximum when the agent is on top of the target and linearly decreases as it moves away, reaching zero when the distance is greater than r_j. Thus each agent has a finite range beyond which no useful information about a target can be acquired and a measurement quality that improves the closer the agent gets to the target. At a given instant, the observations of the same target performed by different agents can be written as a vector of observations as:

$$\underline{z}_i(t) = [\underline{z}_{i,1}^T, ..., \underline{z}_{i,N}^T]^T = \tilde{H}_i(s_1, ..., s_n)\underline{\phi}_i(t) + \tilde{\underline{v}}_i(t), \tag{6}$$

where the following variables are defined as

$$\tilde{H}_i = [\gamma_1(s_1 - x_i)H_i^T, \cdots, \gamma_N(s_N - x_i)H_i^T]^T, \tag{7}$$

$$\tilde{\underline{v}}_i(t) = [\underline{v}_{i,1}^T(t), ..., \underline{v}_{i,N}^T(t)]^T, \tag{8}$$

$$\tilde{R}_i = E[\tilde{\underline{v}}_i^T(t)\tilde{\underline{v}}_i(t)]. \tag{9}$$

Given pre-defined trajectories for the agents, the combination of (1) and (6) define a linear, time-varying, stochastic system. It can easily be shown that the optimal (minimum mean square error) estimator for $\underline{\phi}_i(t)$ is the Kalman-Bucy Filter [14]. The proof of this is omitted for space reasons but the derivation is analogous to Theorem 2 in [11].

Let $\hat{\underline{\phi}}_i(t)$ denote the estimate of the current state of $\underline{\phi}_i(t)$, $\underline{e}_i(t) = \hat{\underline{\phi}}_i(t) - E[\hat{\underline{\phi}}_i(t)]$ as the estimation error and $\Omega_i = E[\underline{e}_i(t)\underline{e}_i^T(t)]$ is the estimate covariance matrix. Then the Kalman-Bucy filter is given by

$$\dot{\hat{\underline{\phi}}}_i(t) = A_i\hat{\underline{\phi}}_i(t) + \Omega(t)_i \tilde{H}_i^T(t)\tilde{R}_i^{-1}\left(\tilde{\underline{z}}_i(t) - \tilde{H}_i(t)\hat{\underline{\phi}}_i(t)\right), \tag{10a}$$

$$\dot{\Omega}_i(t) = A_i\Omega_i(t) + \Omega_i(t)A_i^T + Q_i - \Omega_i(t)\tilde{H}_i^T\tilde{R}_i^{-1}\tilde{H}_i\Omega_i(t). \tag{10b}$$

Substituting (4), (7), and (9) into (10b) yields

$$\dot{\Omega}_i(t) = A_i\Omega_i(t) + \Omega_i(t)A_i^T + Q_i - \Omega_i(t)G_i\Omega_i(t) \sum_{j \in C_i(t)} \left(1 - \frac{|s_j - x_i|}{r_j}\right), \tag{11}$$

where $G_i = H_i^T R_i^{-1} H_i$ and $C_i(t)$ is the *agent neighborhood* of target i, defined to be the indices of all agents within their sensing range to target i at time t.

The overall goal is to minimize the mean estimation error over a given time horizon T. Formally,

$$\min_{u_1,...,u_N} J = \frac{1}{T}\int_0^T \left(\sum_{i=1}^M E\left[\underline{e}_i^T(t)\underline{e}_i(t)\right]\right) dt. \quad (12)$$

Using the fact that $E\left[\underline{e}_i^T(t)\underline{e}_i(t)\right] = \mathrm{tr}(E\left[\underline{e}_i(t)\underline{e}_i^T(t)\right]) = \mathrm{tr}(\Omega_i)$, the optimization in (12) can be rewritten as

$$\min_{u_1,...,u_N} J = \frac{1}{T}\int_0^T \left(\sum_{i=1}^M \mathrm{tr}\left(\Omega_i(t)\right)\right) dt, \quad (13)$$

subject to the dynamics in (2) and (11).

III. OPTIMAL CONTROL SOLUTION

In this section we use a Hamiltonian approach and leverage the Pontryagin Minimum Principle (PMP) [15] to derive properties of the optimal control of the agents to minimize (13). The state $\underline{\chi}$ and control variables $\underline{u}$ can be defined as

$$\underline{\chi}(t) = [s_1, ..., s_N, \mathrm{vec}^T(\Omega_1), ..., \mathrm{vec}^T(\Omega_M)]^T, \quad (14)$$

$$\underline{u}(t) = [u_1(t), ..., u_N(t)]^T. \quad (15)$$

where $\mathrm{vec}(\cdot)$ is the (columnwise) vectorization of a matrix into a column vector. The corresponding costate is given by

$$\underline{\lambda}(t) = [\alpha_1, ..., \alpha_N, \mathrm{vec}^T(\Gamma_1), ..., \mathrm{vec}^T(\Gamma_M)]^T, \quad (16)$$

where Γ_i has the same dimensions of Ω_i. Given the cost function in (13) and state dynamics in (2) and (11), the Hamiltonian is

$$H(\underline{\chi}, \underline{\lambda}, \underline{u}, t) = \frac{1}{T}\sum_{i=1}^M \mathrm{tr}(\Omega_i(t)) + \sum_{j=1}^N \alpha_j(t)u_j(t) + \sum_{i=1}^M \underbrace{\sum_{m=1}^{L_i}\sum_{n=1}^{L_i} \Gamma_i^{(m,n)}(t)\dot{\Omega}_i^{(m,n)}(t)}_{\mathrm{tr}\left(\Gamma_i\dot{\Omega}_i\right)}. \quad (17)$$

Applying the PMP, we have the following necessary conditions for the optimal control.

$$\underline{u}^\star(t) = \arg\min_{\underline{u}} H(\underline{\chi}^\star, \underline{\lambda}^\star, \underline{u}, t), \ \forall\ t \quad (18)$$

together with the costate dynamics

$$\dot{\Gamma}_i^\star(t) = -\frac{\partial H}{\partial \Omega_i}(\underline{\chi}^\star, \underline{\lambda}^\star, \underline{u}^*, t) = -\frac{1}{T}\mathbb{I} - (\Gamma_i^\star)^T\left(A + A^T\right) + (\Gamma_i^\star)^T\left(G_i\Omega_i^\star + \Omega_i^\star G_i\right)\sum_{j\in C_i(t)}\left(1 - \frac{|s_j^\star - x_i|}{r_j}\right), \quad (19)$$

$$\dot{\alpha}_j^\star(t) = -\frac{\partial H}{\partial s_j}(\underline{\chi}^\star, \underline{\lambda}^\star, \underline{u}^*, t) = -\frac{1}{r_j}\sum_{i\in D_j(t)} \mathrm{tr}\left(\Gamma_i^\star\Omega_i^\star G_i\Omega_i^\star\right)\mathrm{sgn}(s_j^\star - x_i), \quad (20)$$

with boundary conditions

$$\Gamma_i^\star(T) = 0_{L_i\times L_i}, \ \alpha_j^\star(T) = 0. \quad (21)$$

Here $\mathbb{I}$ is the identity matrix, $C_i(t)$ is the agent neighborhood of target i defined earlier and $D_j(t)$ is the *target neighborhood* of agent j, that is the collection of all targets within the sensing range of the agent at time t.

From (17) and (18), we can conclude that the optimal control policy for agent j when $\alpha_j^\star \neq 0$ is

$$u_j^\star(t) = \begin{cases} -1, & \alpha_j^\star(t) > 0, \\ 1, & \alpha_j^\star(t) < 0. \end{cases} \quad (22)$$

In order to fully characterize the structure of the optimal solution, it is necessary to know the optimal policy of agent j along a singular arc, that is when $\alpha_j^\star = 0$ over a finite interval. To do so, we first establish the positive definiteness of the matrix $\Gamma_i^\star$ along optimal trajectories.

Proposition 1: $\Gamma_i^\star(t)$ is positive definite for $t \in [0, T)$.

Proof: First, notice that $\Gamma_i^\star$ is a symmetric matrix, since it has a symmetric terminal condition and symmetric dynamics. Therefore, it has real eigenvalues. From Thm. 1.e in [16], since $\Gamma_i^\star(t)$ is a C^1 matrix, its eigenvalues can be C^1 time parameterized. Let μ_n denote the n^{th} eigenvalue of $\Gamma_i^\star(t)$ and $x_n(t)$ the corresponding unit norm eigenvector. Then, from Thm. 5 in [17] we have

$$\dot{\mu}_n = x_n^T\dot{\Gamma}_i^\star x_n.$$

Substituting the costate dynamics of $\Gamma_i^\star$ yields

$$\dot{\mu}_n = x_n^T\left(-\frac{\mathbb{I}}{T} + \Gamma_i^\star(A + A^T) + \Gamma_i^\star(G_i\Omega_i^\star + \Omega_i^\star G_i)\sum_{j\in C_i(t)}\left(1 - \frac{|s_j^\star - x_i|}{r_j}\right)\right)x_n.$$

From the terminal condition on the costate equation, we have that at time T, $\Gamma_i^\star = 0_{L_i\times L_i}$. Therefore, $\mu_n(T) = 0$ and

$$\dot{\mu}_n(T) = -\frac{1}{T} < 0.$$

This implies that $\exists\,\delta > 0$ such that $\forall\, t \in\ (T-\delta, T),\ \mu_n > 0$, and, hence, all the eigenvalues of $\Gamma_i^\star(t)$ have to be positive in this interval preceding the terminal time T. Now suppose that $\mu_n(t') = 0$ for some time $t' \in (0, T)$. Then, since x_n is the corresponding eigenvector, $\Gamma_i^\star(t')x_n(t') = 0$ and, by symmetry of $\Gamma_i^\star$, $x_n^T(t')\Gamma_i^\star(t') = 0$. Thus $\dot{\mu}_n(t') = -\frac{1}{T} < 0$. Since the derivative of the eigenvalue is negative when its value reaches zero, it must stay negative until time T. However, we have already established that the eigenvalue must be positive in an interval ending at the terminal time. Thus, $\mu_n > 0\ \forall\ t \in [0, T)$, which proves the proposition. ■

Define an isolated target i as one for which

$$\min_{k\neq i}|x_i - x_k| > r_{\max}, \quad r_{\max} = \max\{r_1, ..., r_N\}.$$

Then, we show that a singular arcs with isolated targets can only occur if the agent's position coincides with the position of the target.

Proposition 2: Consider an isolated target x_i and an agent j at position s_j such that $0 < |s_j(t') - x_i| < r_j$, $t' \in (0, T)$. Suppose further that $G_i = H^TR_iH \neq 0$ and that Ω_i is positive definite in $(0, T)$. Then $\dot{\alpha}_j^\star(t') \neq 0$.

Proof: Since $0 < |s_j - x_i| < r_j$, i.e., the agent is within sensing range of the isolated target i but not on top of it, and recalling (20), then

$$\dot{\alpha}_j^\star = -\frac{\text{sgn}(s_j^\star - x_i)}{r_j}\text{tr}\left(\Gamma_i^\star \Omega_i^\star G_i \Omega_i^\star\right). \quad (23)$$

Using the inequality $\text{tr}(BC) \geq \mu_{\min}(B)\text{tr}(C)$ with B and C positive semi-definite matrices and $\mu_{\min}(\cdot)$ denoting the smallest eigenvalue of its argument [18] we have that

$$\text{tr}\left(\Gamma_i^\star \Omega_i^\star G_i \Omega_i^\star\right) \geq \mu_{\min}(\Gamma_i^*)\text{tr}(\Omega_i^\star G_i \Omega_i^\star). \quad (24)$$

Since $G_i \neq 0$, $\Omega_i^\star G_i \Omega_i^\star$ is positive semidefinite with at least one positive eigenvalue. Finally, from Prop. 1 we have that all the eigenvalues of $\Gamma_i^\star$ are positive. Therefore,

$$\text{tr}\left(\Gamma_i^\star \Omega_i^\star G_i \Omega_i^\star\right) > 0. \quad (25)$$

By hypothesis, $s_k^\star \neq x_i$, then, using (20), we get that $\dot{\alpha}_j^\star \neq 0$ and the proposition is established. ■

Proposition 2 immediately implies that agent j cannot experience a singular arc when visiting an isolated target unless directly on top of a target. When on top of a target, the only way to stay there for a finite interval of time (and thus have a singular arc) is for $u_j^\star = 0$. Also, suppose that for some optimal policy $u_j^\star(t) \notin \{-1, 0, 1\}$ when the agent is not visiting any target in a finite interval $t \in (a, b)$, then the alternative policy $\tilde{u}_j(t)$

$$\tilde{u}_j^\star(t) = \begin{cases} sgn(s_j(b) - s_j(a)) & t \in (a, a + |s(b) - s(a)|) \\ 0 & t \in (a + |s(b) - s(a)|, b) \\ u_j^\star(t) & otherwise \end{cases}$$

is feasible and also optimal, since in both policies the agent does not have any effect on any targets' covariance for $t \in (a, b)$. Hence, in an environment where all targets are isolated, there is an optimal control where $u_j^*(t) \in \{-1, 0, 1\}$. This control structure is the same as in [8], even though the performance metric and underlying dynamics are different. We note that establishing a similar result for the case of non-isolated targets remains a topic of ongoing research.

IV. INFINITESIMAL PERTURBATION ANALYSIS

The results of Sec. III showed that there is an optimal control policy (at least in the case of isolated targets) whose control only takes values in $\{-1, 0, 1\}$. Any trajectory of agent j under such a control law can be fully described by the initial position $s_j(0)$ and the parameters

$$\underline{\theta}_j = [\theta_{j,1}, ..., \theta_{j,K_j}],\ \underline{\omega}_j = [\omega_{j,1}, ..., \omega_{j,K_j}], \quad (26)$$

where $\underline{\theta}_j$ are the *switching points*, that is $\theta_{j,m}$ is the position where agent j had its m-th direction change, and $\underline{\omega}_j$ are the *dwell times*, that is $\omega_{j,m}$ is the time agent j spent stopped at position $\theta_{j,m}$ before changing direction. K_j is the total number of events for agent j. While the number of events is not typically known a priori, it is often possible to determine an upper bound based on the system dynamics and constraints; see [8] for details.

This parameterization defines a hybrid system in which the dynamics of the agents only change when an event occurs. Events are given by a change in control value at a switching point and then the completion of a dwell time. These may occur simultaneously if the dwell time is zero (representing a switch of control from ± 1 to ∓ 1). Given this parameterization, we use an approach analogous to [8] where IPA is used to calculate the gradient of the cost function with respect to the parameters defining the trajectories and a gradient descent scheme to optimize the cost function.

A. IPA Review

IPA is a tool for computing stochastic gradient estimates that are unbiased and distribution invariant when there is uncertainty on the model parameters under mild stochastic assumptions on the distribution. Also, IPA is naturally event-driven, i.e. it specifies how the occurrence of an event influences the state and the event times and, consequently, the cost function. This subsection provides a very brief review of IPA; more details can be found in [12].

Let $\psi \in \Psi$ denote a parameter vector in a compact, convex set Ψ. Define $\{\tau_k(\theta)\}$, $k = 1, \dots, K$ to be the times of all event occurrences of a hybrid system with dynamics $\dot{\xi}(t) = f_k(\xi, t, \psi)$ over the time interval $[\tau_k(\psi), \tau_{k+1}(\psi))$. It can be shown that in the interval $[\tau_k(\psi), \tau_{k+1}(\psi))$,

$$\frac{d}{dt}\left(\frac{\partial \xi(t)}{\partial \psi}\right) = \frac{\partial f_k(t)}{\partial \xi}\frac{\partial \xi(t)}{\partial \psi} + \frac{\partial f_k(t)}{\partial \psi} \quad (27)$$

with the boundary condition

$$\frac{\partial \xi}{\partial \psi}(\tau_k^+) = \frac{\partial \xi}{\partial \psi}(\tau_k^-) + [(f_{k-1}(\tau_k^-)) - f_k(\tau_k^+)]\frac{\partial \tau_k}{\partial \psi}. \quad (28)$$

To evaluate the boundary condition (28), it is necessary to give a procedure to compute $\frac{\partial \tau_k}{\partial \psi}$. If the event transition at time τ_k does not depend on the parameter ψ, then $\frac{\partial \tau_k}{\partial \psi} = 0$. Otherwise, if there exists a function $g_k(\xi(\psi, t), t)$ such that $\tau_k = \inf\{t > \tau_{k-1},\ g_k(\xi(\psi, t), t) \leq 0\}$, then

$$\frac{\partial \tau_k}{\partial \psi} = -\left[\frac{\partial g_k}{\partial \psi} f_k(\tau_k^-)\right]^{-1}\left(\frac{\partial g_k}{\partial \psi} + \frac{\partial g_k}{\partial \xi}\frac{\partial \xi}{\partial \psi}(\tau_k^-)\right). \quad (29)$$

B. IPA formulation for finding an optimal trajectory

Recall that the underlying idea for online optimization of the agent trajectories is to use gradient descent on the cost function. We therefore need the gradient of that cost with respect to the parameters. From (13) we have

$$\frac{\partial J}{\partial \theta} = \frac{1}{T}\sum_{i=1}^{M}\int_0^T \frac{\partial \text{tr}(\Omega_i(t))}{\partial \theta}dt. \quad (30)$$

To compute $\frac{\partial \text{tr}(\Omega_i(t))}{\partial \theta}$ we use IPA to derive the ordinary differential equations whose solution will yield the desired gradient. Applying (27), first with respect to the switching points parameter yields

$$\frac{d}{dt}\left(\frac{\partial \Omega_i(t)}{\partial \theta_{j,m}}\right) = A\frac{\partial \Omega_i(t)}{\partial \theta_{j,m}} + \frac{\partial \Omega_i(t)}{\partial \theta_{j,m}}A^T - \Bigg(\Omega_i(t)G_i\frac{\partial \Omega_i(t)}{\partial \theta_{j,m}} + \frac{\partial \Omega_i(t)}{\partial \theta_{j,m}}G_i\Omega_i(t)\Bigg)\sum_{j \in C_i(t)}\left(1 - \frac{|s_j - x_i|}{r_j}\right) + \Omega_i(t)G_i\Omega_i(t)\frac{I_j(s_j - x_i)}{r_j}\frac{\partial s_j(t)}{\partial \theta_{j,m}}, \quad (31)$$

$$I_j(\alpha) = \begin{cases} +1, & 0 < \alpha < r_j, \\ -1, & -r_j < \alpha < 0, \\ 0, & \text{otherwise}. \end{cases} \tag{32}$$

The ODE for the dwelling time can be obtained by replacing $\theta_{j,m}$ by $\omega_{j,m}$. The initial conditions for (31) and its analogous $\omega_{j,m}$ version are

$$\frac{\partial \Omega_i(0)}{\partial \theta_{j,m}} = \frac{\partial \Omega_i(0)}{\partial w_{j,m}} = 0_{L_i \times L_i}. \tag{33}$$

Notice that the differential equations (31) and its analogous $\omega_{j,m}$ version depend on the gradients $\frac{\partial s(t)}{\partial \theta_{j,m}}$ and $\frac{\partial s(t)}{\partial \omega_{j,m}}$. Details of this calculation can be found in [8]; here we present only the resulting forms. Note that m indexes a change in control value while n indexes the total number of events (changes in control and completion of a (possibly zero) dwell period).

Case 1: $u_j(\tau_{j,n}^-) = \pm 1,\ u_j(\tau_{j,n}^+) = 0$.

$$\frac{\partial s_j}{\partial \theta_{j,m}}(\tau_{j,n}^+) = \begin{cases} 1, & \text{if } n = 2m-1, \\ 0, & \text{otherwise}. \end{cases} \tag{34}$$

$$\frac{\partial s_j}{\partial w_{j,m}}(\tau_{j,n}^+) = 0, \forall n. \tag{35}$$

Case 2: $u_j(\tau_{j,n}^-) = 0,\ u_j(\tau_{j,n}^+) = \pm 1$.

$$\frac{\partial s_j}{\partial \theta_{j,m}}(\tau_{j,n}^+) = \begin{cases} \frac{s_j}{\partial \theta_{j,m}}(\tau_{j,n}^-) + 1, & \text{if } n = 2m, \\ \frac{s_j}{\partial \theta_{j,m}}(\tau_{j,n}^-) + 2(-1)^m, & \text{if } n > 2m, \\ 0, & \text{otherwise}. \end{cases} \tag{36}$$

$$\frac{\partial s_j}{\partial w_{j,m}}(\tau_{j,n}^+) = \begin{cases} -u_j(\tau_{k,n}^+), & \text{if } n \geq 2m \\ 0, & \text{otherwise}. \end{cases} \tag{37}$$

Case 3: $u_j(\tau_{j,n-2}^-) = \pm 1,\ u_j(\tau_{j,n}^+) = \mp 1$ and $\tau_{j,n-1} = \tau_{j,n}$.

$$\frac{\partial s_j}{\partial \theta_{j,m}}(\tau_{j,n}^+) = \begin{cases} 2, & \text{if } n = 2m, \\ -\frac{\partial s_j}{\partial \theta_{j,m}}(\tau_{j,n}^-), & \text{otherwise}. \end{cases} \tag{38}$$

$$\frac{\partial s_j}{\partial w_{j,m}}(\tau_{j,n}^+) = 0, \forall\ n \tag{39}$$

Notice that (31) is a first order linear ODE in which $\Omega_i(t)$, $s_j(t)$, $\partial s_j(t)/\partial \theta_{j,m}$ should be interpreted as an inputs. One interesting feature of these partial derivatives is that they are constant between events and they are zero at $t = 0$. We thus need only evaluate the gradients at the event times $t = \tau_{j,m}$.

The complete optimization is summarized in Algorithm 1, where $proj$ is the projection onto the (convex) feasible set, η_k is the gradient descent step size, $\underline{s}_0 = [s_{0,1}, ..., s_{0,N}]$ is the set of initial positions of the agents, $\Omega(0) = [\Omega_1(0), ...\Omega_M(0)]$ is the set of initial covariance matrices of the estimators, $\underline{\theta} = [\underline{\theta}_1^T, ..., \underline{\theta}_N^T]^T$ is the complete vector of switching points and $\underline{\omega} = [\underline{\omega}_1^T, ..., \underline{\omega}_N^T]^T$ is the complete vector of dwell times.

Algorithm 1 Agents' Trajectory Optimization

procedure GRADIENT DESCENT
  Input: $\underline{s}_0, \Omega(0), \underline{\theta}_1^0, ..., \underline{\theta}_N^0, \underline{\omega}_1^0, ..., \underline{\omega}_N^0$,
  $||\nabla J|| \leftarrow \infty$
  $k \leftarrow 0$
  while $||\nabla J|| > \epsilon$ **do**
    $k \leftarrow k + 1$
    for j ranging from 1 to N **do**
      for m ranging from 1 to K_j **do**
        $\left[\frac{\partial J}{\partial \theta_{j,m}}, \frac{\partial J}{\partial \omega_{j,m}}\right]$ =IPA$(j, m, \underline{s_0}, \Omega(0), \underline{\theta}, \underline{\omega})$
      $\underline{\theta}_j^k \leftarrow$ proj$(\underline{\theta}_j^{k-1} - \eta_k \frac{\partial J}{\partial \underline{\theta}_j})$
      $\underline{\omega}_j^k \leftarrow$ proj$(\underline{\omega}_j^{k-1} - \eta_k \frac{\partial J}{\partial \underline{\omega}_j})$
    $||\nabla J|| = \frac{1}{\eta_k} \left|\left| \left(\underline{\theta}_k^T - \underline{\theta}_{k-1}^T, \underline{\omega}_k^T - \underline{\omega}_{k-1}^T\right)^T \right|\right|$
  Output: $\underline{\theta}_1^0, ..., \underline{\theta}_N^0, \underline{\omega}_1^0, ..., \underline{\omega}_N^0$
procedure IPA
  Input: $j, m, \underline{s}_0, \Omega(0), \underline{\theta}_1, ..., \underline{\theta}_N, \underline{\omega}_1, ..., \underline{\omega}_N$
  Compute $s_1(t), ..., s_N(t)$ from the parameterization
  Compute $\Omega_i(t)$ according to Eqs. (11) initial condition $\Omega_i(0)$ for every $i = 1, ..., M$
  Compute $\frac{\partial s_j}{\partial \theta_{j,m}}(\tau_n^+)$ and $\frac{\partial s_j}{\partial \theta_{j,m}}(\omega_n^+)$ for $n = 1, .., K$
  Solve Diff. Eqs. (31) with initial conditions in Eq (33) and using $s_j(t)$, $\Omega_i(t)$ and $\frac{\partial s}{\partial \theta_{j,m}}(\tau_n^+)$ and $\frac{\partial s}{\partial \omega_{j,m}}(\tau_n^+)$ as inputs for every target $i = 1, ..., M$
  $\frac{\partial J}{\partial \theta_{j,m}} = \int_0^T \sum_{i=1}^M \text{tr}\left(\frac{\partial \Omega_i}{\partial \theta_{j,m}}\right) dt$
  $\frac{\partial J}{\partial \omega_{j,m}} = \int_0^T \sum_{i=1}^M \text{tr}\left(\frac{\partial \Omega_i}{\partial \omega_{j,m}}\right) dt$
  Output: $\frac{\partial J}{\partial \theta_{j,m}}, \frac{\partial J}{\partial \omega_{j,m}}$

V. SIMULATION AND RESULTS

To demonstrate this scheme, we simulated a scenario with two agents and four targets over a time horizon of $T = 25$ units. The initial positions of the agents were set to $s_1(0) = s_2(0) = 0$ and each had a sensing radius $r_1 = 1$. The targets were located at $(x_1, x_2, x_3, x_4) = (1, 3, 5, 7)$. For each target, the dynamics of the state $\underline{\Phi}_i(t) \in \mathbb{R}^2$ evolved according to (1) with

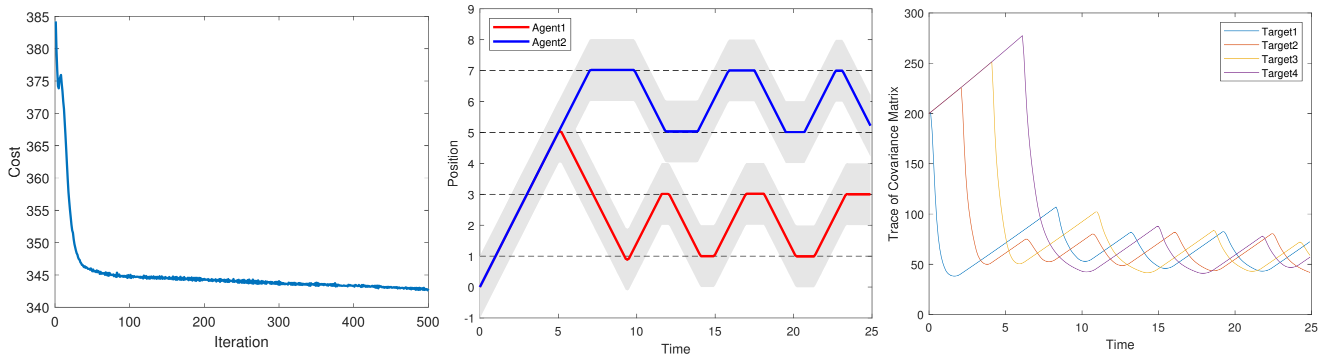
$$A_i = 10^{-3} \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix}, \quad Q_i = \text{diag}(5, 5).$$

The observations from each agent were given by (3) with

$$H_i = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad R_i = \text{diag}(10, 10).$$

A constant gradient descent step size $\eta_k = 2 \times 10^{-4}$ was used and the parameters of the trajectory at the initial step of the optimization were

$$\underline{\theta}_1^0 = \begin{bmatrix} 4, & 0.5, & 4, & 0.5, & 4, & 0.5, & 4 \end{bmatrix},$$
$$\underline{\theta}_2^0 = \begin{bmatrix} 4, & 7, & 4, & 7, & 4, & 7 & 4 \end{bmatrix},$$
$$\omega_1^0 = 0.5 \begin{bmatrix} 1, & 1, & 1, & 1, & 1, & 1 & 1 \end{bmatrix},$$
$$\omega_2^0 = 0.5 \begin{bmatrix} 1, & 1, & 1, & 1, & 1, & 1 \end{bmatrix}.$$



(a) Cost vs. iteration number (b) Agent trajectories at final iteration (c) Trace of the covariance for each target

Fig. 1: Results of a simulation with two agents and four targets. (a) Evolution of the overall cost as a function of iteration number on the gradient descent. (b) Trajectories of the two agents at the final iteration. The dashed lines indicate the positions of the targets and the grey shaded area the visibility region of each agent. (c) Evolution of the trace of the estimation covariance matrices of the four targets.

The results of the simulation are shown in Fig. 1. The cost in (13), shown in Fig. 1a, improved rapidly in the first few iterations. The optimal trajectories for the agents found after the algorithm converged are shown in Fig. 1b. After 10 units of time, the mission space is divided into two with each agent cycling between two adjacent targets, dwelling at each one before moving to the other. The traces of the target covariances along these trajectories are shown in Fig. 1c. In the first cycle, the two agents go over the first three targets, and after that the first agent stays cycling between the first two targets and the second agent between the last two. The covariances of all the targets rapidly decrease and are then held below a value of 80. While the agents converge to a repeating sequence that splits the space, the dwell times over this time horizon vary from visit to visit.

VI. CONCLUSION AND FUTURE WORK

In this work, we modeled the problem of monitoring a finite set of targets, each one of them with internal states that evolve according to a linear, stochastic system, using a set of mobile agents equipped with noisy sensors with range-dependent performance. We established that for the case of isolated targets, the optimal control can be written in a parametric form. Using IPA, this cost function can be optimized online to determine optimal trajectories for the agents. Finally, this result was demonstrated through a simple simulation scenario.

In future work, we plan to investigate the case where the targets are not isolated where the primary challenge is understanding the singular arcs in the optimal trajectories as well as conditions for stability of the estimation error. Also, we are going to investigate the extension of this technique to situations where the agent and the targets are not constrained to be in a 1D environment.

REFERENCES

[1] N. Michael, E. Stump, and K. Mohta, "Persistent Surveillance with a Team of MAVs," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2011, pp. 2708–2714.

[2] J. Yu, S. Karaman, and D. Rus, "Persistent Monitoring of Events With Stochastic Arrivals at Multiple Stations," *IEEE Transactions on Robotics*, vol. 31, no. 3, pp. 521–535, 2015.

[3] J. Yu, M. Schwager, and D. Rus, "Correlated Orienteering Problem and its Application to Persistent Monitoring Tasks," *IEEE Transactions on Robotics*, vol. 32, no. 5, pp. 1106–1118, 2016.

[4] T. Alam, G. M. Reis, L. Bobadilla, and R. N. Smith, "A Data-Driven Deployment Approach for Persistent Monitoring in Aquatic Environments," in *IEEE International Conference on Robotic Computing*. IEEE, 2018, pp. 147–154.

[5] E. Stump and N. Michael, "Multi-robot Persistent Surveillance Planning as a Vehicle Routing Problem," in *IEEE International Conference on Automation Science and Engineering*. IEEE, 2011, pp. 569–575.

[6] X. Yu, S. B. Andersson, N. Zhou, and C. G. Cassandras, "Optimal Dwell Times for Persistent Monitoring of a Finite Set of Targets," in *American Control Conference*. IEEE, 2017, pp. 5544–5549.

[7] ——, "Optimal Visiting Schedule Search for Persistent Monitoring of a Finite Set of Targets," in *American Control Conference (ACC)*. IEEE, 2018, pp. 4032–4037.

[8] C. G. Cassandras, X. Lin, and X. Ding, "An Optimal Control Approach to the Multi-agent Persistent Monitoring Problem," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 947–961, 2013.

[9] N. Zhou, X. Yu, S. B. Andersson, and C. G. Cassandras, "Optimal Event-Driven Multiagent Persistent Monitoring of a Finite Set of Data Sources," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4204–4217, 2018.

[10] X. Lan and M. Schwager, "Planning Periodic Persistent Monitoring Trajectories for Sensing Robots in Gaussian Random Fields," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 2415–2420.

[11] ——, "A Variational Approach to Trajectory Planning for Persistent Monitoring of Spatiotemporal Fields," in *2014 American Control Conference*. IEEE, 2014, pp. 5627–5632.

[12] 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.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*. JHU press, 2012, vol. 3.

[14] R. S. Bucy and P. D. Joseph, "Filtering for Stochastic Processes with Applications to Guidance," University of Southern California Los Angeles, Dept. of Aerospace Engineering, Tech. Rep., 1968.

[15] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation and Control*. CRC Press, 1975.

[16] A. Kriegl, P. W. Michor, and A. Rainer, "Denjoy–Carleman Differentiable Perturbation of Polynomials and Unbounded Operators," *Integral Equations and Operator Theory*, vol. 71, no. 3, p. 407, 2011.

[17] P. Lancaster, "On Eigenvalues of Matrices Dependent on a Parameter," *Numerische Mathematik*, vol. 6, no. 1, pp. 377–387, 1964.

[18] Y. Fang, K. A. Loparo, and X. Feng, "Inequalities for the Trace of Matrix Product," *IEEE Transactions on Automatic Control*, vol. 39, no. 12, pp. 2489–2490, 1994.