2020 American Control Conference
Denver, CO, USA, July 1-3, 2020

Optimal Periodic Multi-Agent Persistent Monitoring of a Finite Set of Targets with Uncertain States

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 investigate the problem of persistently monitoring a finite set of targets with internal states that evolve with linear stochastic dynamics using a finite set of mobile agents. We approach the problem from the infinite-horizon perspective, looking for periodic movement schedules for the agents. Under linear dynamics and some standard assumptions on the noise distribution, the optimal estimator is a Kalman-Bucy filter. It is shown that when the agents are constrained to move only over a line and that they can see at most one target at a time, the optimal movement policy is such that the agent is always either moving with maximum speed or dwelling at a fixed position. Periodic trajectories of this form admit finite parameterization, and we show how to compute a stochastic gradient estimate of the performance with respect to the parameters that define the trajectory using Infinitesimal Perturbation Analysis. A gradient-descent scheme is used to compute locally optimal parameters. This approach allows us to deal with a very long persistent monitoring horizon using a small number of parameters.**

I. INTRODUCTION

As autonomous cyber-physical systems are continuously increasing their importance in our society, the topic of long term autonomy is gaining more interest. In this context, short term goals are not as important as planning behaviors that will be efficient over large horizons. One class of problems of interest in the context of long term autonomy is where one has a collection of points of interest (denoted as "targets") and a set of moving agents that can visit these targets and perform some form of estimation or control to their internal state. This paradigm finds applications in very diverse contexts, such as traffic surveillance in critical points of a city, sea temperature estimation, and tracking of nanometer-scale particles in optical microscopy. While for static systems the estimation or control error does not grow over time, in dynamic and stochastic systems this error may grow very fast as time increases. Therefore, if there are not enough agents to continuously estimate or control these targets, then the mobile agents must travel over the environment with trajectories which can visit the targets infinitely often in order to avoid unbounded errors as time goes to infinity. Persistent monitoring is the term used to refer to this class of problems.

While the persistent monitoring problem has already been studied in the literature [1]–[6], these works focused on analyzing the transient behavior of the system. Motivated by the prospects of long term autonomy, we tackle the problem from the infinite horizon point of view, where continuous estimation of internal states of the targets is performed. While the idea of periodicity of the solution of the persistent monitoring problem has already been explored in [5], [6], these works did not provide tools for analyzing the behavior of the solution in steady state. Therefore, in order to apply these techniques over the long term one would either need to optimize over a very long period or always recompute the solution for the next cycle. Both approaches have excessive computational overhead. Periodicity naturally fits into the persistent monitoring paradigm since targets need to be visited infinitely often and, although a periodic structure of the solution is not necessarily optimal, results in the transient case show that the trajectories tend to converge to oscillatory behavior [7]. On top of that, previous results show that in the discrete time case, periodic schedules can approximate arbitrarily well the cost of an optimal schedule [8].

In this work, we provide tools for analyzing and optimizing a periodic trajectory in order to minimize the steady state estimation error. We assume that agents can observe the targets' internal states with a linear observation model with Gaussian additive noise, and hence, the optimal estimator is a Kalman-Bucy filter. The differential Riccati equation then expresses the dynamics of the covariance matrix and, naturally, the mean quadratic estimation error. We extend the work in [7] in which we considered targets distributed in a 1-D environment and where the agent could see at most one target at a time. In that scenario we are able to show that there is a parameterization of the optimal solution of the finite-time version of the problem considered here. In this paper, we still assume the environment to be 1-D, however we consider the infinite horizon version of the problem and restrict ourselves to periodic trajectories for which we show that, under some

This work was supported in part by NSF under grants ECCS-1509084, ECCS-1931600, 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 by the RevealFlight Concerted Research Action (ARC) of the Federation Wallonie-Bruxelles, by the Incentive Grant for Scientific Research (MIS) Learning from Pairwise Data of the F.R.S.-FNRS, and by a WBI World Excellence Fellowship.

978-1-5386-8266-1/$31.00 ©2020 AACC

assumptions, the covariance matrix converges to a limit cycle. We then use Infinitesimal Perturbation Analysis (IPA) in a centralized gradient descent scheme to obtain locally optimal trajectories. This approach not only allows the shape of the trajectory to be optimized, but also its period. It is worth noticing that in many interesting applications that can be modeled as a persistent monitoring problem, agents are constrained to (possibly multiple) uni-dimensional mobility, such as powerline inspection agents, cars on streets, and autonomous vehicles in rivers.

II. PROBLEM FORMULATION

We consider an environment with M fixed targets located at positions $x_1, ..., x_M \in \mathbb{R}$. Each target has an internal state $\phi_i \in \mathbb{R}^{L_i}$ with dynamics

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

where $w_i,\ i = 1, \ldots, M$, are mutually independent, zero mean, white, Gaussian distributed processes with $E[w_i(t)w_i(t)'] = Q_i$ with Q_i a positive definite matrix for every i.

We have N mobile agents, whose positions at time t are denoted by $s_1(t), ..., s_N(t) \in \mathbb{R}$, equipped with sensing capabilities. These agents move with the kinematic model

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

where their speed is constrained by $|u_j(t)| \leq 1$, after proper scaling. Note that even though we only consider first order dynamics in this paper, extensions to second order dynamics would likely follow similar results, as discussed in [9]. The internal state of target i can be observed by agent j according to the following linear model

$$z_{i,j}(t) = \gamma_j\left(s_j(t) - x_i\right) H_i\phi_i(t) + v_{i,j}(t), \tag{3}$$

where $v_{i,j},\ i = 1, \ldots, M,\ j = 1, \ldots, N$ are mutually independent zero mean, white, Gaussian distributed noise processes, independent of the w_i, with $E[v_{i,j}(t)v'_{i,j}(t)] = R_i$, R_i positive definite, and $\gamma_j(\cdot)$ is a scalar function. In this model, the noise power is constant but the sensed signal level varies as a function of the distance to the target. Even though the analysis conducted in this paper is valid for any unimodal $\gamma_j(\cdot)$ that has finite support, we use the following definition for concreteness:

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

Under this model, the instantaneous signal to noise ratio (SNR) of a single measurement made by agent j is given by

$$\frac{E\left[(z_{i,j}(t) - v_{i,j}(t))'(z_{i,j}(t) - v_{i,j}(t))\right]}{E[v'_{i,j}(t)v_{i,j}(t)]} = \max\left(0, 1 - \frac{|s_j - x_i|}{r_j}\right)\frac{\phi'_i(t)H'_iH_i\phi_i(t)}{\mathrm{tr}(R_i)}, \tag{5}$$

where $\mathrm{tr}(\cdot)$ is the trace of a matrix. The term $\phi'_i(t)H'_iH_i\phi_i(t)(\mathrm{tr}(R_i))^{-1}$ is deterministic and scalar and cannot be influenced by the relative position between the agent and the target. On the other hand, the max function (along with the SNR) is maximum when the agent's position coincides with that of the target, linearly decreases as it moves farther, and is zero if the distance is greater than r_j. The motivation behind this definition is to model sensors which have a finite sensing range and within that range, the sensing quality is higher the closer the agent is to the measurement target.

The instantaneous joint observations performed by all the agents of the same target can be written as a vector of observations,

$$z_i(t) = [z'_{i,1}, ..., z'_{i,N}]' = \tilde{H}_i(s_1, ..., s_n)\phi_i(t) + \tilde{v}_i(t) \tag{6}$$

where

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

$$\tilde{v}_i(t) = [v'_{i,1}(t), ..., v'_{i,N}(t)]', \tag{8}$$

$$E[\tilde{v}'_i(t)\tilde{v}_i(t)] = \tilde{R}_i = diag(R_i, ..., R_i). \tag{9}$$

Note that (1) and (6) define a linear, time-varying, stochastic system if the trajectories are already pre-defined. The optimal estimator for the states $\phi_i(t)$ is then a Kalman-Bucy Filter [10]. A proof of this result is omitted here for space reasons, but the derivation is analogous to a similar result in [6], where it is shown that the Kalman-Bucy filter is indeed optimal, considering targets with internal states with the same dynamics as in (1) and a general agent dependent time-varying observation model, similar to (3).

Let $\hat{\phi}_i(t)$ denote the estimate of the current state of $\phi_i(t)$, and let $e_i(t) = \hat{\phi}_i(t) - \phi_i(t)$ be the estimation error and $\Omega_i = E[e_i(t)e'_i(t)]$ the error covariance matrix. Then, the Kalman-Bucy filter equations are

$$\dot{\hat{\phi}}_i(t) = A_i\hat{\phi}_i(t) + \Omega(t)_i\tilde{H}'_i(t)\tilde{R}_i^{-1}\left(\tilde{z}_i(t) - \tilde{H}_i(t)\hat{\phi}_i(t)\right), \tag{10a}$$

$$\dot{\Omega}_i(t) = A_i\Omega_i(t) + \Omega_i(t)A'_i + Q_i - \Omega_i(t)\tilde{H}'_i\tilde{R}_i^{-1}\tilde{H}_i\Omega_i(t). \tag{10b}$$

Substituting (4), (7), and (8) into (10b) yields

$$\dot{\Omega}_i(t) = A_i\Omega_i(t) + \Omega_i(t)A'_i + Q_i - \Omega_i(t)G_i\Omega_i(t)\eta_i(t), \tag{11}$$

where $G_i = H'_iR_i^{-1}H_i$ and

$$\eta_i(t) = \sum_{j \in C_i(t)} \gamma_j(s_j(t) - x_i). \tag{12}$$

The overall goal is to minimize the mean squared estimation error over an infinite time horizon. Formally, for the set of inputs $u(t)$ where the following limit exists, the objective is to find the optimal cost $J^\star$ (where the input dependence on the time is ommited for the sake of notation conciseness):

$$J^\star = \min_{u_1,\ldots,u_N} \lim_{t\to\infty} \frac{1}{t}\int_0^t \left(\sum_{i=1}^M E\left[e'_i(\xi)e_i(\xi)\right]\right) d\xi. \tag{13}$$

Using the fact that

$$E\left[e'_i(t)e_i(t)\right] = \mathrm{tr}(E\left[e_i(t)e'_i(t)\right]) = \mathrm{tr}(\Omega_i)$$

the optimization in (13) can be rewritten as

$$\min_{u_1,\dots,u_N} J = \lim_{t\to\infty} \frac{1}{t}\int_0^t \left(\sum_{i=1}^{M} \operatorname{tr}\left(\Omega_i(\xi)\right)\right) d\xi, \tag{14}$$

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

In the context of minimizing the mean squared estimation error over an infinite horizon, we focus on periodic movement schedules for the agents. We can show that the optimal policy over a period has the property that $u_j(t) \in \{-1, 0, 1\}$. To present this result, we first define a target i as *isolated* if the following holds:

$$\min_{k\neq i} |x_i - x_k| > 2r_{\max}, \quad r_{\max} = \max\{r_1, ..., r_N\}.$$

We also define the minimum distance between regions for where targets can be visited, ($d_{\min}$), as

$$d_{\min} = \min_{i,k} |x_i - x_k| - 2r_{\max} > 0.$$

Proposition 1: In an environment where all the targets are isolated, given any policy $u_j(\xi)$, $j = 1, ..., N$, $\xi \in [0, 1]$, then there is a policy $\tilde{u}_j(\xi)$ where $\tilde{u}_j(\xi) \in \{-1, 0, 1\}$ $\forall \xi \in [0, 1]$, where $J(u_1, ..., u_N) \geq J(\tilde{u}_1, ..., \tilde{u}_N)$ and the number of control switches is upper bounded by $2\frac{t}{d_{\min}} + 4$.

Due to space limitations, the proof to Proposition 1 is omitted. It can be found in [11].

III. STEADY STATE PERIODIC SCHEDULES

As stated in Sec. I, in the present work we analyze the behavior of the steady state covariance matrices, $\bar{\Omega}_i$. This approach contrasts with [4], [6], [7], where only the transient behavior was studied and, therefore, the number of parameters necessary to represent the trajectory grew as the time-horizon grew. The approach here presented is particularly interesting because it captures the long-term mean squared estimation error while only needing to optimize the parameters that describe a single period of the trajectory.

If the agents' trajectories are constrained to be periodic, we know that $\eta_i(t)$, as defined in (12), will also be periodic and, therefore, the Ricatti equation for this model, as presented in (11), is periodic. Before proceeding to the computation of the steady state covariance, we give a few natural assumptions on the system.

Assumption 1: The pair (A_i, H_i) is detectable, for every $i \in \{1, ..., M\}$.

Assumption 2: Q_i and the initial covariance matrix $\Sigma_i(0)$ are positive definite, for every $i \in \{1, ..., M\}$.

The first assumption is needed in order to ensure that observations are able to maintain a bounded estimation error as time goes to infinity and the second one ensures that the covariance matrix is always positive definite, a property that will be used on the proof of Prop. 3.

Following a procedure similar to the one used in the proof of *Lemma 9* in [12], we show that, when target i is visited for at least a finite amount of time, the Riccati equation for that target (11) converges to a unique periodic solution. A solution $\bar{\Omega}_i$ to (11) is said to be stabilizing if, for any solution Ω_i of (11) with symmetric non-negative initial conditions, $\lim_{t\to\infty} \lambda_{\max}\left(\bar{\Omega}_i - \Omega_i\right) = 0$, where $\lambda_{\max}(.)$ is the eigenvalue of maximum absolute value of a matrix.

Proposition 2: If $\eta_i(t) > 0$ for some interval $[a, b] \in [0, T]$ with $b > a$, then, under Assumption 1, there exists a non-negative stabilizing T-periodic solution to (11).

The proof is omitted here and is available in [11].

Notice that we can always design a periodic trajectory such that every target is visited for at least a finite time interval and therefore, $\eta_i(t) > 0$ for some interval. Defining $\bar{\Omega}_i(t)$ as the unique periodic solution to (11) and $\Omega_i(t)$ as the solution for some non negative initial conditions $\Omega_i(0)$ we know that, since $\bar{\Omega}_i(t)$ is the unique stabilizing solution of (11),

$$\forall \delta > 0, \exists\ t_0\ s.t.\ \left\|\bar{\Omega}_i(t) - \Omega_i(t)\right\| \leq \delta,\ \forall t \geq t_0,$$

and then we have that

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \left|\operatorname{tr}(\bar{\Omega}_i(\xi) - \Omega_i(\xi))\right| d\xi \leq \delta. \tag{15}$$

Equation (15) implies that, for any initial condition on the covariance matrix, if we apply a periodic schedule for the agents such that every target is visited at least once, after sufficient time, the cost given by (14) will become arbitrarily close to the mean cost over time of the steady state periodic solution associated to that same periodic trajectory. Therefore, if we optimize the steady state solution $\bar{\Omega}_i$, the cost of the solution starting at any arbitrary initial condition will asymptotically approach that of the steady state one as time evolves.

Consider now the motion of the agents. The result in Proposition 1 implies that when the targets are isolated there is always a control policy such that $u_j(t) \in \{-1, 0, 1\}$ that leads to lower or equal cost compared to one where $u_j(t) \neq \{-1, 0, 1\}$ for some time interval. This property leads us to restrict ourselves to periodic trajectories where the movement of each agent j consists of a sequence of dwelling at the same position for some duration of time followed by moving at maximum speed to another location. Then, one period of the trajectory of an agent j can be described by the following parameters: T, the period of the trajectory; $s_j(0)$, the initial position; $\omega_{j,p}$, $p = 1, ..., P_j$, the normalized dwelling times for agent j, i.e., the agent dwells for $\omega_{j,p}T$ units of time before it moves with maximum speed for the p-th time in the cycle; $\tau_{j,p}$, $p = 1, ..., P_j$, the normalized movement times for agent j, i.e., the agent j moves for $\tau_{j,p}T$ units of time to the right (if p is odd) or to the left (if p is even) after dwelling for $\omega_{j,p}T$ units of time in the same position.

The constraints in (16) below ensure periodicity and consistency of the trajectory. Note that the last two constraints ensure that the total time that the agents spend moving will be less than or equal to a period and that over the course of a period agents will return to their initial position.

$$\begin{gathered}\tau_{j,m} \geq 0,\ \omega_{j,m} \geq 0,\ T \geq 0,\ \sum_{m=1}^{P_j} (\tau_{j,m} + \omega_{j,m}) \leq 1 \\ \sum_{m=1}^{P_j} (-1)^m \tau_{j,m} = 0.\end{gathered} \tag{16}$$

Notice that this description does not exclude transitions of u_j of the kind $\pm1 \to \mp1$ and $\pm1 \to 0 \to \pm1$, since it allows $\omega_{j,m} = 0$ and $\tau_{j,m} = 0$. This parameterization defines a hybrid system in which the dynamics of the agents remain unchanged between events and abruptly switch when an event occurs. Events are given by a change in control value at the completion of movement and dwell times. Note that these may occur simultaneously, for instance, if the dwell time is zero (representing a switch of control from ±1 to ∓1). Given this parameterization, we use an approach analogous to [4], [7] where IPA is used to calculate the stochastic gradient estimate of the cost function with respect to the parameters defining the trajectories and then the gradient is used in a gradient descent scheme to optimize the cost function.

IV. OPTIMIZATION OF THE PERIODIC TRAJECTORY

In this section we take advantage of the convergence of the Ricatti equation to a steady state solution in order to compute the derivative of this limit cycle solution with respect to all the parameters that define the trajectory. These can be used in a gradient descent scheme to obtain a locally optimal steady state solution. In this work, we use Infinitesimal Perturbation Analysis (IPA) to compute these gradients. IPA is a tool for estimating stochastic gradients of hybrid system states and event times with respect to given system parameters. These estimates, under mild assumptions on the distribution of the random processes involved, have the interesting property of being unbiased and distribution invariant [13]. IPA is particularly attractive due to its event driven nature, i.e., the equations used in the computation of the parameters only need to be updated when some event (e.g. a transition of the discrete mode of the system) happens, which means that effort for updating the equations scales linearly with the number of events (rather than exponentially with the number of targets and agents).

A. IPA Formulation

By defining $q = t/T$, (11) can be rescaled as

$$\dot{\Omega}_i(q) = \frac{d\Omega_i(q)}{dq} = T(A\Omega_i(q) + \Omega_i(q)A' + Q - \eta_i(q)\Omega_i(q)G\Omega_i(q)). \quad (17)$$

In order to optimize the parameters of the agent trajectories using gradient descent, we need the gradient of the cost with respect to these parameters. Taking the partial derivative of (14), we have that for any parameter θ

$$\frac{\partial J}{\partial \theta} = \sum_{i=1}^{M} \int_0^1 \frac{\partial \mathrm{tr}(\Omega_i(q))}{\partial \theta} dq. \quad (18)$$

Using IPA, we derive the ordinary differential equations for which the desired gradient $\frac{\partial \Omega_i(t)}{\partial \theta}$ is a solution. Note that in this paper we sidestep the issue of whether or not these gradients exist. We know that there are sets of parameters for which the gradient does not exist (imagine, for instance, a set of parameters for which one of the targets is never visited and the dynamics of this target are unstable, therefore, its covariance diverges as time goes to infinity). However, experience and simulations support the assumption that these gradients do indeed exist in the interior of the set of parameters for which each target is visited at least once.

Computing the derivative of Ω_i with respect to any parameter θ yields

$$\frac{\partial \dot{\bar{\Omega}}_i(q)}{\partial \theta} - T\bigg(A\frac{\partial \bar{\Omega}_i(q)}{\partial \theta} + \frac{\partial \bar{\Omega}_i(q)}{\partial \theta}A' - \eta_i(q)\bar{\Omega}_i(q)G\frac{\partial \bar{\Omega}_i(q)}{\partial \theta} - \eta_i(q)\frac{\partial \bar{\Omega}_i(q)}{\partial \theta}G\bar{\Omega}_i(q)\bigg) = T\frac{\partial \eta_i(q)}{\partial \theta}\bar{\Omega}_i(q)G\bar{\Omega}_i(q) + \frac{\partial T}{\partial \theta}\frac{\dot{\bar{\Omega}}_i(q)}{T}, \quad (19)$$

where one should look at $\frac{\partial \bar{\Omega}_i}{\partial \theta}$ as the unknown function which we are trying to solve for. In this expression, the term $\eta_i(q)$ is fully determined by the agent's trajectory parameters. The computation of the steady state covariance matrix $\bar{\Omega}_i(q)$ is described in the previous section and explicit expressions for $\frac{\partial \bar{\Omega}_i}{\partial \theta}$ will be given in the next subsection.

Since (19) does not fully determine a unique solution (different initial conditions $\frac{\partial \bar{\Omega}_i}{\partial \theta}(0)$ will yield different solutions), we need extra conditions to determine the partial derivatives of the covariance matrix. Because $\bar{\Omega}_i(q)$ is periodic, $\frac{\partial \bar{\Omega}_i}{\partial \theta}$ must also be periodic. This property will allow us to uniquely determine the initial conditions for computing the derivative $\frac{\partial \bar{\Omega}_i}{\partial \theta}$, as discussed in the following.

Define the problem:

$$\dot{\Sigma}_H(q) - T\left(A - \eta_i(q)\bar{\Omega}_i(q)G\right)\Sigma_H(q) = 0, \ \Sigma_H(0) = I \quad (20)$$

and let Σ_{ZI} be the solution of (19) with the zero matrix as the initial conditions. Also, let Σ_H denote the solution of the homogeneous version of (19) with the identity matrix as the initial condition. Then, the initial conditions matrix Λ that yields a periodic solution of (19) is such that [14]:

$$\Lambda = \Sigma_H(1)\Lambda\Sigma_H'(1) + \Sigma_{ZI}(1), \quad (21)$$

which has at least one solution Λ if $\frac{\partial \bar{\Omega}_i}{\partial \theta}$ exists. The following proposition states sufficient conditions for uniqueness.

Proposition 3: Assume that Σ_H is a solution of (20), Assumptions 1 and 2 hold, target i is observed at least once in the period T, and there exists a solution to (21). Then, the solution to (21) is unique.

The proof of Prop. 3 is omitted here for the sake of space and is available at [11]. The Lyapunov equation in (21) can be efficiently solved for low-dimensional systems using the algorithm in [15] and implemented in MATLAB function *dlyap*. The partial derivatives can then be computed as:

$$\frac{\partial \bar{\Omega}_i(q)}{\partial \theta} = \Sigma_H'(q)\Lambda\Sigma_H(q) + \Sigma_{ZI}(q). \quad (22)$$

For computing the entire gradient, (22) should be used to compute the partial with respect to the parameters that define the trajectory of each agent.

B. Computation of $\frac{\partial \eta_i(q)}{\partial \theta}$

Looking back to (19), in order to give a complete procedure for computing the derivative $\frac{\partial J}{\partial \theta}$ when it exists, the only component left is to compute the derivative $\frac{\partial \eta_i(q)}{\partial \theta}$. Using (12), we know that

$$\frac{\partial \eta_i(q)}{\partial \tau_{j,m}} = -\frac{I_j(s_j - x_i)}{r_j}\frac{\partial s_j(q)}{\partial \tau_{j,m}}, \tag{23}$$

$$I_j(\alpha) = \begin{cases} +1, & 0 < \alpha < r_j, \\ -1, & -r_j < \alpha < 0, \\ 0, & |\alpha| > r_j. \end{cases} \tag{24}$$

As a side note, since $\gamma_{i,j}$ is not differentiable at $\alpha = 0$, we can use the concept of subgradient and use any value between -1 and 1 for $I_j(0)$. Similarly,

$$\frac{\partial \eta_i(q)}{\partial \omega_{j,m}} = -\frac{I_j(s_j - x_i)}{r_j}\frac{\partial s_j(q)}{\partial \omega_{j,m}}, \tag{25}$$

$$\frac{\partial \eta_i(q)}{\partial s_j(0)} = -\frac{I_j(s_j - x_i)}{r_j}\frac{\partial s_j(q)}{\partial s_j(0)}, \tag{26}$$

$$\frac{\partial \eta_i(q)}{\partial T} = -\sum_{j=1}^{N}\frac{I_j(s_j - x_i)}{r_j}\frac{\partial s_j(q)}{\partial T}. \tag{27}$$

In order to compute $\frac{\partial s_j(q)}{\partial \theta}$ for some parameter θ we will explicitly write the position $s_j(q)$ as a function of this parameter. As already discussed, IPA is event-driven in nature. For our parameterization, these events are the instants when the trajectory presents a change in the velocity, at the end of dwell times or movement times. The dynamics of the derivatives may experience discontinuities at these specific event times. The order of the events is defined in such a way that initially the agent dwells, then it moves right, then dwells again, followed by moving left and repeat this sequence until the number of events reaches $2P_j$, where P_j is a designer-defined parameter that indicates the maximum number of direction switches the agent j can experience in its trajectory. Note that the value of P_j is upper bounded when the targets are isolated and Proposition 1 gives an upper bound for P_j as a function of the period T. Also, notice that under this definition events are agent-specific and can happen at different times for different agents.

The position of agent j at normalized time q, after the k-th event and before the $k+1$-th is

$$s_j(q) - s_j(0) = \begin{cases} T\Bigg((-1)^{k/2+1}\Big(q - \sum_{p=1}^{k/2-1}(\tau_{j,p} + \omega_{j,p}) \\ \quad + \omega_{j,\frac{k}{2}}\Big) + \sum_{p=1}^{k/2}(-1)^{p+1}\tau_p\Bigg), \ k \text{ even}, \\ T\sum_{p=1}^{\frac{k-1}{2}}(-1)^{p+1}\tau_{j,p}, \ k \text{ odd}. \end{cases} \tag{28}$$

Therefore,

$$\frac{\partial s_j}{\partial \tau_{j,m}} = \begin{cases} \left((-1)^{\frac{k}{2}+1} + (-1)^p\right)T, & m < \frac{k}{2}, \ k \text{ even}, \\ (-1)^{m+1}T, \ m \le \frac{k-1}{2}, & k \text{ odd}, \end{cases} \tag{29}$$

$$\frac{\partial s_j}{\partial \omega_{j,m}} = \begin{cases} 1, \ m < \frac{k}{2}, & k \text{ even}, \\ 0 & , \text{ otherwise}, \end{cases} \tag{30}$$

$$\frac{\partial s_j(q)}{\partial T} = \frac{s_j(q) - s_j(0)}{T}, \tag{31}$$

$$\frac{\partial s_j}{\partial s_j(0)} = 1. \tag{32}$$

Finally, we note that when optimizing the agent trajectories, we use the procedure to compute the gradient described in this section in conjunction with a projected gradient approach, i.e.,

$$\theta_j^{l+1} = \text{proj}\left(\theta_j^l - \kappa_l \frac{\partial J}{\partial \theta_j}\right), \tag{33}$$

where *proj* represents the projection of the parameters into the convex set defined by the constraints in (16), the upper index l refers to the step number and κ_l is the gradient descent step size. An algorithm that summarizes the entire optimization procedure is available on the extended version of this manuscript [11].

In this paper, a procedure for obtaining θ_j^0 is not discussed. One essential condition for this initial configuration is that every target is visited at least once for a finite amount of time, as discussed in Sec. III, otherwise the covariance matrices will not converge to a steady-state solution. Although providing efficient initial parameters for the optimization is a topic that we are still investigating, one possible way to address it would be to use the transient analysis given in [7].

V. Simulation Results

In this section, we demonstrate the results of our approach in a scenario with five targets and two agents. All targets i have the same state dynamics evolving according to (1) with parameters

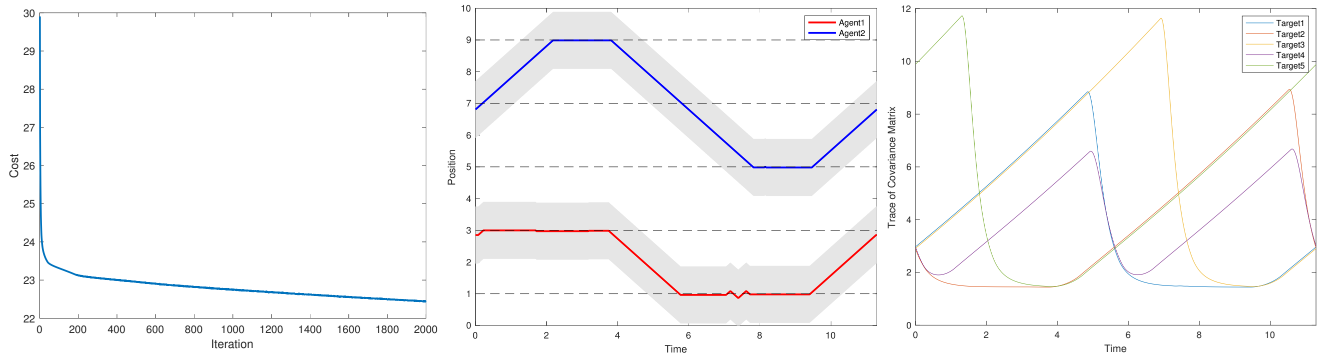
$$A_i = \begin{bmatrix} -1 & -0.1 \\ -0.1 & 0.01 \end{bmatrix}, \quad Q_i = \text{diag}(1,1),$$

and observation model as in (3) with parameters

$$H_i = \text{diag}(1,1), \quad R_i = \text{diag}(1,1), \quad r_j = 0.9.$$

A constant descent stepsize was used ($\kappa_l = \kappa_0 = 0.02$) and the targets were placed in positions $x_i = 1 + 2i,\ i = 1, .., 5$. The initial parameters were the following: $s_1^0(0) = 2.7,\ s_2(0) = 6.8,\ T^0 = 6,\ P_1 = P_2 = 11,\ \tau_1^0 = \tau_2^0 = 0.1[1, 0.1, 1, 1, 0.1, 1, 0.1, 1, 1, 0.1, 1],\ \omega_1^0 = \omega_2^0 = 0.0125[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]$.

Figure 1 shows the results of the optimization in this scenario. Notice that even though both agents and all the targets have the same dynamic models, the solution at the last iteration of the optimization was such that one of the agents visits three of the targets and the other two of them. One interesting aspect of the trajectories of the targets in Fig. 1b is that, while in the period between times 6 and 8 agent 1 makes a movement with small amplitude around target 1, the effects of this oscillatory movement are hard



(a) Cost vs. gradient descent iteration

(b) Agent trajectories at final iteration

(c) Trace of the covariance for each target

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

to notice in the trace of the covariance of target 1 in Fig. 1c. Therefore, even though it is intuitively clear that staying still rather than moving with this oscillatory behavior will lead to a lower cost solution, the difference in terms of cost is minor. Also, notice that the solution has not yet fully converged, as can be seen in Fig. 1a. The number of iterations was selected to highlight interesting aspects of the process and thus not run to convergence. The effect of the gradient descent step size (or, more generally, the descent algorithm applied) and its effect on the convergence rate, are topics of future research. Finally, note that while the maximum number of switches in a direction allowed to each agent was set to 11, the final solution appears to have fewer because some of the movement and dwelling times in the final solution are essentially zero.

VI. Conclusions and Future Work

In this paper, we developed a technique both to analyze and to optimize the steady state mean squared estimation error of a finite set of targets being monitored by a finite set of moving agents. The structure of the optimal solution allowed us to represent it in a parametric way and we provided numerical tools to optimize it in a scalable manner. Some simulation examples were provided in order to demonstrate the proposed technique.

In future work, we plan to approach the question of whether gradients of $\bar{\Omega}_i$ with respect to the parameters that define the trajectory always exist in the interior of the set where they lead to a convergent $\bar{\Omega}_i$. Moreover, we intend to study how to efficiently generate initial trajectories in order to converge to global optimal points or, at least, good local optima. We also plan to extend the results presented here to scenarios where the agents are not constrained to a single dimension, possibly using suboptimal parameterizations for the trajectory, as in [16].

References

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

[2] X. Yu, S. B. Andersson, N. Zhou, and C. G. Cassandras, "Optimal Visiting Schedule Search for Persistent Monitoring of a Finite Set of Targets," in *American Control Conference (ACC)*, 2018, pp. 4032–4037.

[3] 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.

[4] 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.

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

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

[7] S. C. Pinto, S. B. Andersson, J. M. Hendrickx, and C. G. Cassandras, "Optimal Multi-Agent Persistent Monitoring of the Uncertain State of a Finite Set of Targets," in *Proc. of 58th Control and Decision Conference*, 2019.

[8] L. Zhao, W. Zhang, J. Hu, A. Abate, and C. J. Tomlin, "On the Optimal Solutions of the Infinite-horizon Linear Sensor Scheduling Problem," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, pp. 2825–2830, 2014.

[9] Y. Wang, Y. Wei, X. Liu, N. Zhou, and C. G. Cassandras, "Optimal Persistent Monitoring Using Second-Order Agents With Physical Constraints," *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3239–3252, Aug 2019.

[10] 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.

[11] S. C. Pinto, S. B. Andersson, J. M. Hendrickx, and C. G. Cassandras, "Optimal Periodic Multi-Agent Persistent Monitoring of the Uncertain State of a Finite Set of Targets," in *Arxiv*. Available online, 2019.

[12] J. Le Ny, E. Feron, and M. A. Dahleh, "Scheduling Continuous-Time Kalman Filters," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1381–1394, 2010.

[13] 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.

[14] W. T. Reid, "A Matrix Differential Equation of Riccati Type," *American Journal of Mathematics*, vol. 68, no. 2, pp. 237–246, 1946.

[15] A. Barraud, "A Numerical Algorithm to Solve AT XA-X=Q," in *IEEE Conf. on Decision and Control and 16th Symposium on Adaptive Processes and Special Symposium on Fuzzy Set Theory and Applications*, no. 16, 1977, pp. 420–423.

[16] X. Lin and C. G. Cassandras, "An Optimal Control Approach to the Multi-Agent Persistent Monitoring Problem in Two-dimensional Spaces," *IEEE Transactions on Automatic Control*, vol. 60, no. 6, pp. 1659–1664, 2014.