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

Samuel C. Pinto¹, Sean B. Andersson^{1,2}, Julien M. Hendrickx³, and Christos G. Cassandras^{2,4}

¹Dept. of Mechanical Engineering, ²Division of Systems Engineering,

⁴Dept. of Electrical and Computer Engineering

Boston University, Boston, MA 02215, USA

³ICTEAM 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 and the mean estimation error is a function of its covariance matrix, which evolves as a differential Riccati equation. It is shown that when the agents are constrained to move only over a line and they can see at most one target at a time, the movement policy that minimizes the mean estimation error over time is such that the agent is always either moving with maximum speed or dwelling at a fixed position. This type of trajectory can be fully defined by a finite set of parameters. For periodic trajectories, under some observability conditions, the estimation error converges to a steady state condition and the stochastic gradient estimate of the cost with respect to the trajectory parameters of each agent and the global period can be explicitly computed using Infinitesimal Perturbation Analysis. A gradient-descent approach 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 the context of long term autonomy one is not looking only to accomplish short term goals, but also to plan behaviors that will be efficient in the long term. 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 macro particles in optical microscopy. While for static systems the estimation or control error does not grow over time, in general 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 persistently visit the targets 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]-[8], 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. For a periodic solution, the mean estimation error of these internal states, will, as time goes to infinity, approach the estimation error of a steady state periodic solution independent of the initial conditions. While the idea of periodicity of the solution of the persistent monitoring problem has already been explored in [7], [8], these works did not provide tools for analyzing the behavior of the solution in steady state. Therefore, in order to apply these techniques for in long term one would either need to optimize over a very long period or always recompute the solution for the next cycle and both approaches have an expressive computational overhead. As a way to overcome this issue, instead of minimizing the transient estimation error, we can neglect the transient effects and plan trajectories that minimize the steady state estimation error of a periodic trajectory. In this paradigm, it is only necessary to optimize the parameters that describe one period of the trajectory, which is usually a very small number of parameters. Moreover, as time goes to infinity, the mean estimation error will be arbitrarily close to the state state error that has been planned for.

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 for this model is a Kalman-Bucy filter and the differential Riccati equation expresses the dynamics of the covariance matrix and, naturally, the mean quadratic estimation error.

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 in part by Communaté française de Belgique -Actions de Recherche Concertées and a WBI World Excellence Fellowship.

We extend the work [9], in which we considered targets distributed in a 1-D environment and where the agent can see at most one target at a time, for in this 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, however, we consider the infinite horizon version and restrict ourselves to periodic trajectories for which we show that, under some 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.

The remainder of this paper is organized as follows. Sec. II presents the problem formulation, including target and agent dynamics and the Kalman-Bucy filter for estimating target states from the agent measurements. Sec. III discusses properties of the optimal control, leading to a parameterized representation of an optimal trajectory. In Sec. IV, properties and conditions for convergence of the Riccati equation to a limit cycle solution are discussed. The IPA-driven gradient descent is considered in Sec. V and the entire scheme is demonstrated through simulations in Sec. VI. Finally, Sec. VII gives a conclusion and shares ideas for future works.

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, ..., M, are mutually independent, zero mean, white, Gaussian distributed processes with $E[w_i(t)w_i(t)^T] = Q_i$ 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 can move with the following the kinematic model

$$\dot{s}_j(t) = u_j(t), \ j = 1, ..., N,$$
 (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 [10]. The internal state of target *i* can be observed by agent *j* according to the following linear model.

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

where $v_{i,j}$, i = 1, ..., M, j = 1, ..., N are mutually independent zero mean, white, Gaussian distributed noise processes, independent of the w_i , with $E[v_{i,j}(t)v_{i,j}^T(t)] =$ R_i , R_i positive definite, and $\gamma_j(\cdot)$ is a scalar function. In this model, the noise power is constant but sensed the 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:

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

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

$$\frac{E\left[(z_{i,j}(t) - v_{i,j}(t))^{T}(z_{i,j}(t) - v_{i,j}(t))\right]}{E[v_{i,j}^{T}(t)v_{i,j}(t)]} = \max\left(0, 1 - \frac{|s_{j} - x_{i}|}{r_{j}}\right)\frac{\phi_{i}^{T}(t)H_{i}^{T}H_{i}\phi_{i}(t)}{\operatorname{tr}(R_{i})}, \quad (5)$$

where $\operatorname{tr}(\cdot)$ is the trace of a matrix. The term $\phi_i^T(t)H_i^TH_i\phi_i(t)(\operatorname{tr}(R_i))^{-1}$ is deterministic and scalar and can not 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 . Therefore, useful information can only be acquired within the sensing range of the agent and within this range the measurement quality is higher the closer the agent is to the target. Figure 1 illustrates this dependence.



Fig. 1: Illustration of the dependence of the SNR on the distance between agent j and target i.

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

$$z_i(t) = [z_{i,1}^T, ..., z_{i,N}^T]^T = \tilde{H}_i(s_1, ..., s_n)\phi_i(t) + \tilde{v}_i(t) \quad (6)$$

where

$$\tilde{H}_{i} = [\gamma_{1}(s_{1} - x_{i})H_{i}^{T}, \cdots, \gamma_{N}(s_{N} - x_{i})H_{i}^{T}]^{T}, \quad (7)$$
$$\tilde{v}_{i}(t) = [v_{i,1}^{T}(t), \dots, v_{i,N}^{T}(t)]^{T}, \quad (8)$$

and

$$E[\tilde{v}_{i}^{T}(t)\tilde{v}_{i}(t)] = \tilde{R}_{i} = \begin{bmatrix} R_{i} & 0 & \dots & 0\\ 0 & R_{i} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & R_{i} \end{bmatrix}.$$
 (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 [11]. A proof that this is indeed the optimal estimator is not ommited here for space reasons, but the derivation is analogous to the similar result in [8], where it is shown that the Kalman-Bucy filter is an optimal estimator, considering targets with internal states with the same dynamics as in (1) and a general agent dependent time-varying observation model, similar to (3).

We denote $\hat{\phi}_i(t)$ the estimate of the current state of $\phi_i(t)$, and $e_i(t) = \hat{\phi}_i(t) - E[\hat{\phi}_i(t)]$ the estimation error and $\Omega_i =$ $E[e_i(t)e_i^T(t)]$ the error covariance matrix. Then, the Kalman-Bucy filter equations are

$$\dot{\hat{\phi}}_{i}(t) = A_{i}\hat{\phi}_{i}(t) + \Omega(t)_{i}\tilde{H}_{i}^{T}(t)\tilde{R}_{i}^{-1}\left(\tilde{z}_{i}(t) - \tilde{H}_{i}(t)\hat{\phi}_{i}(t)\right),$$
(10a)
$$\dot{\Omega}_{i}(t) = A_{i}\Omega_{i}(t) + \Omega_{i}(t)A_{i}^{T} + Q_{i} - \Omega_{i}(t)\tilde{H}_{i}^{T}\tilde{R}_{i}^{-1}\tilde{H}_{i}\Omega_{i}(t).$$
(10b)

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

$$\dot{\Omega}_i(t) = A_i \Omega_i(t) + \Omega_i(t) A_i^T + Q_i - \Omega_i(t) G_i \Omega_i(t) \eta_i(t), \quad (11)$$

where $G_i = H_i^T R_i^{-1} H_i$ and

$$\eta_i(t) = \sum_{j \in C_i(t)} \left(1 - \frac{|s_j(t) - x_i|}{r_j} \right).$$
(12)

 $C_i(t)$ is the agent neighborhood of target i, i.e., the indices of all agents with target *i* within their respective sensing range at time t.

The overall goal is to minimize the mean estimation error over an infinite time horizon. Formally, for the set of inputs u 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):

$$J^{\star} = \min_{u_1, \dots, u_N} \lim_{t \to \infty} \frac{1}{t} \int_0^t \left(\sum_{i=1}^M E\left[e_i^T(t') e_i(t') \right] \right) \, \mathrm{dt'}.$$
(13)

Using the fact that

$$E\left[e_i^T(t)e_i(t)\right] = \operatorname{tr}(E\left[e_i(t)e_i^T(t)\right]) = \operatorname{tr}(\Omega_i)$$

the optimization in (13) can be rewritten as

$$\min_{u_1,\dots,u_N} J = \lim_{t \to \infty} \frac{1}{t} \int_0^t \left(\sum_{i=1}^M \operatorname{tr} \left(\Omega_i(t') \right) \right) \, \mathrm{dt}', \qquad (14)$$

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

III. OPTIMAL CONTROL PROPERTIES

The purpose of this section is to establish properties of the optimal control solution in order to be able to describe the optimal trajectory by a finite set of parameters. In previous work [9], we used Hamiltonian Analysis to derive these properties for non-periodic trajectories. However, the same argument cannot be used for periodic trajectory since the latter derivation was deeply dependent on the fact that the terminal position of the agent was not specified, which gave a terminal condition for the optimal costate matrix associated to the covariance. Therefore, we establish a similar results that also contemplates periodic trajectories, where we constraint ourselves to trajectories where the terminal and the initial point coincide over the course of one period. Initially, we introduce the following proposition which is going to be essential in the proof of the properties of an optimal control solution.

Proposition 1: Given $\Omega_1(t)$ and $\Omega_2(t)$, two bounded covariance matrices under the dynamics in (11) with $A = A_1 =$ $A_2, G = G_1 = G_2, Q = Q_1 = Q_2$, then if $\Omega_1(0) - \Omega_2(0)$ is negative semi-definite and $\eta_1(t) \ge \eta_2(t)$, then $\Omega_1(t) - \Omega_2(t)$ is a negative semi definite matrix for $t \ge 0$.

Proof: Defining $\Xi = \Omega_1(t) - \Omega_2(t)$. The dynamics of Ξ is described by the following equation.

$$\dot{\Xi}(t) = A\Xi(t) + \Xi A^T - \eta_1(t)\Omega_1(t)G\Omega_1(t) + \eta_2(t)\Omega_2(t)G\Omega_2(t).$$
(15)

Adding and subtracting the terms $\eta_1(t)\Omega_2(t)G\Omega_2(t)$ and $\eta_1(t)\Omega_1(t)G\Omega_2(t)$ to the equation, we can rewrite the relation in (15) as:

$$\dot{\Xi}(t) = A\Xi(t) + \Xi A^T - \eta_1(t) \left[\Omega_1(t)G\Xi(t) + \Xi(t)G\Omega_2(t)\right] + \left[\eta_2(t) - \eta_1(t)\right]\Omega_2(t)G\Omega_2(t).$$
(16)

From Thm. 1.e in [12], since $\Xi(t)$ is a C^1 matrix, its eigenvalues can be C^1 time parameterized. Let μ_n denote the n^{th} eigenvalue of $\Xi(t)$ and $x_n(t)$ the corresponding unit norm eigenvector. Then, from Thm. 5 in [13] we have that

$$\dot{\mu}_n = x_n^T \dot{\Xi} x_n.$$

Also, notice that by using (16) and the fact that $\lambda_{\min}\left(\frac{D+D^{T}}{2}\right) \leq x^{T}Dx \leq \lambda_{\max}\left(\frac{D+D^{T}}{2}\right) \|x\| = \|D\| \|x\|,$ for any square matrix D,

$$\dot{\mu}_{n} \leq \|A\| \,\mu_{n} - \eta_{1}\xi\mu_{n} + [\eta_{2} - \eta_{1}] \,x_{n}^{T}\Omega_{2}G\Omega_{2}x_{n} \\ \leq \|A\| \,\mu_{n} - \eta_{1}\xi\mu_{n}$$

m

where $\xi = \lambda_{min} \left((\Omega_1 + \Omega_2) G + G(\Omega_1 + \Omega_2) \right)$. Using Gronwall's inequality and the fact that a first order linear ODE does not change sign, we conclude that $\mu_n(t) \leq 0, \forall t \in$ [0,T] and, therefore, $\Xi(t)$ is negative semidefinite. In Prop. 1, even though it does not fully match the notation established in Sec. II, Ω_1 and Ω_2 can also be understood as covariance matrices for the same target but under different agent trajectories and this is the way that they are going to be interpreted in Prop. 2.

Before proceeding to the proposition about an optimal control structure, a few definitions are necessary. An isolated target i is a target such that

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

and the minimum distance between visible areas (d_{\min}) is defined as:

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

and the finite time cost is defined as

$$J(u_1, ..., u_N) = \frac{1}{t} \int_0^t \left(\sum_{i=1}^M \text{tr} \left(\Omega_i(t') \right) \right) \, \mathrm{dt'}.$$
 (17)

We can, then, claim a similar result to Prop. 1 in [9].

Proposition 2: In an environment where all the targets are isolated, given any policy $u_j(t')$, j = 1, ..., N, then there is a policy $\tilde{u}_j(t')$ where $\tilde{u}_j(t') \in \{-1, 0, 1\}$ where $J(u_1, ..., u_N) \geq J(\tilde{u}_1, ..., \tilde{u}_N)$ and the number of control switches is upper bounded by $2\frac{t}{d_{\min}} + 4$. *Proof:* This proposition is going to be proved by

Proof: This proposition is going to be proved by construction: given a policy $u_j(t')$ with $\eta_i(t')$ associated to it (as defined by (12)), we will construct an alternative policy $\tilde{u}_j(t')$ associated with $\tilde{\eta}_i(t')$ such that $\tilde{\eta}_i(t') \ge \eta_i(t') \ \forall t' \in [0, t]$ and i = 1, ..., M, and then use Prop. 1, along with the definition of the cost (17), to show that the alternative policy has lower or equal cost than the original one.

Initially, we focus on the policy $u_j(t')$. Note that an agent j is said to visit a target i if at some time $t', |s_j(t') - x_i(t')| < r_j$. For every agent in the policy $u_j(t')$, there is an ordered collection of targets it visits in [0, t]. Therefore, there must exist a set of indices of all the targets visited by agent j, $y_0^j, ..., y_{K_j}^j \in \{1, ..., M\}$, such that $y_p^j \neq y_{p-1}^j$ and agent j visited no other target in the time between visiting targets y_p^j and y_{p-1}^j . This set is an ordered set of all the targets that agent j visited over [0, t], not counting consecutive visits to the same target. Notice that the same target can be present more than once in the vector $[y_0^j, ..., y_{K_j}^j]$ but if that is the case, it will not be in consecutive positions.

For each of these visits, we can define the initial visitation time t_p^j for $p = 1, ..., K_j$ as

$$t_p^j = \inf\{t' | t' > t_{p-1}^j \text{ and agent } j \text{ visits target } y_p^j \text{ at time } t'\}$$

and $t_0^j = 0$ and $t_{K_j+1}^j = t$ Also, we define the initial visitation positions a_p^j , i.e., the position of the agents when it starts visiting a target, $a_p^j = s_j(t_p^j)$, $p = 0, ..., K_j + 1$. χ_p^j is the position of the target y_p^j when a visit starts, i.e. $\chi_p^j = x_{y_p^j}$. Also note that while $t' \in [t_{p-1}^j, t_p^j)$, agent j only influences the value of $\eta_i(j)$ of the target it is currently visiting.

We propose the following alternative policy, where $\tilde{u}_j(t')$ for $t' \in [t_{p-1}^j, t_p^j)$ is such that:

$$\tilde{u}_{j}(t') = \begin{cases} \frac{a_{p}^{j} - s_{j}(t')}{|a_{p}^{j} - s_{j}(t')|}, & \text{if } t_{p}^{j} - t' \leq |a_{p}^{j} - s_{j}(t)|, \\ 0, & \text{if } t_{p}^{j} - t' > |a_{p}^{j} - s_{j}(t)| & \text{and } s_{j}(t') = \chi_{p}^{j}, \\ \frac{\chi_{p}^{j} - s_{j}(t')}{|a_{p}^{j} - s_{j}(t')|}, & \text{otherwise.} \end{cases}$$

$$(18)$$

The intuition behind the proposed alternative policy is that at the beginning of each visit, the agent moves with maximum speed towards the target y_p^j and if it reaches the target, it dwells on top of it. However, it must move in a way such that it begins the next visit at the same time that the original policy, i.e., the positions of agent j associated to the alternative policy $\tilde{s}_j(t')$ is such that $\tilde{s}_j(t_p^j) = s_j(t_p^j) = a_p^j$.

Notice that the provided construction provides a feasible trajectory, since the original trajectory is assumed feasible. Also, for time $t' \in [t_p^j, t_{p+1}^j]$ both the original and the alternative policies only influence value of η_i for $i = t_p^j$ and, since in the alternative policy all the agents are closer or at least as close to the currently visited target, by looking at (12) we can claim that

$$\tilde{\eta}_i(t') \ge \eta_i(t'), \ \forall t' \in [0, t], \ i \in \{1, ..., M\}.$$

Therefore, using Prop. 1 and the cost definition (17), we get that

$$J(\tilde{u}_1, ..., \tilde{u}_N) - J(u_1, ..., u_N) = \frac{1}{t} \int_a^b \sum_{i=1}^M \operatorname{tr}\left(\tilde{\Omega}_i(t') - \Omega_i(t')\right) \le 0.$$

which shows that the alternative policy has a lower or equal cost compared to the original one. Also, notice that the maximum number of velocity switches in the alternative policy is $2\frac{t}{d_{\min}} + 4$, since there can be a maximum of $\frac{t}{d_{\min}} + 1$ visits to targets with maximum speed equal to one and, for each visited target, the alternative policy has at most 2 velocity switches, plus one switch to match the initial position and another to match the terminal position of the original policy.

One way to interpret this proposition is, if you look ahead the next T units of time (where T might represent the period of a periodic solution), there is an optimal trajectory that has its controls in the set $\{1, 0, 1\}$. This is the kind of trajectory that we will be pursuing in the remainder of this paper. Also, notice that even though we were not able so far to prove that there exists some optimal control solution with the same structure when the targets are not necessarily isolated, the same structure can still be used but without the guarantee of optimality.

IV. STEADY STATE PERIODIC SCHEDULES

As stated in Sec. I, the goal of the present work is to provide tools for analyzing the steady state behavior of the covariance matrices $\overline{\Omega}_i$. This approach contrasts with [5], [8], [9], where only the transient behavior was analyzed 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 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 $i \in \{1, ..., M\}$.

Following a procedure similar to the one used in the proof of *Lemma 9* in [14], 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. Note that a solution $\overline{\Omega}_i$ to (11) is said to be stabilizing if, for any solution Ω_i of (11) with symmetric non-negative initial conditions, $\lim_{t\to\infty} \lambda_{\max} (\overline{\Omega}_i - \Omega_i) = 0$, where $\lambda_{\max}(.)$ is the eigenvalue if maximum absolute value of a matrix.

Proposition 3: 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).

Proof: According to [15, p. 130], a pair $(A_i, \eta_i(t)H_i)$ of a periodic system is detectable if and only if for every eigenpair (x, λ) with $x \neq 0$,

$$A_i x = \lambda x \implies \exists t \in [0, T] \ s.t. \ \eta_i(t) e^{\lambda t} H_i x \neq 0$$

Notice that, due to Assumption 1, for any eigenvector x of A_i , $H_i x \neq 0$, therefore when $\eta_i(t) > 0$ (i.e. any $t \in [a,b]$), $\eta_i(t)e^{\lambda t}H_i x \neq 0$, which implies that $(A_i, \eta_i(t)H_i)$ is detectable. Therefore, the collorary to Theorem 3 in [16, p. 95] states shows that there exists a non negative *T*-periodic solution to (11), $\overline{\Omega}_i(t)$, and

$$\lim_{t \to \infty} (\Omega_i(t) - \bar{\Omega}_i(t)) = 0$$

for any positive definite initial condition $\Omega_i(0)$.

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 $\overline{\Omega}_i(t)$ 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 $\overline{\Omega}_i(t)$ is the unique stabilizing solution of (11),

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

which implies that

$$\lim_{t \to \infty} \frac{1}{t} \int_{t_0}^t |\operatorname{tr}(\tilde{\Omega}_i(t') - \Omega_i(t'))| \, \mathrm{d}t' \le \delta$$

and, since we know that, for a finite period of integration t_0 , $\lim_{t\to\infty} \frac{1}{t} \int_0^{t_0} |\operatorname{tr}(\tilde{\Omega}_i(t') - \Omega_i(t'))| dt' = 0$, we conclude that

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t |\operatorname{tr}(\tilde{\Omega}_i(t') - \Omega_i(t'))| \, \mathrm{dt}' \le \delta.$$
(19)

Equation (19) 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. This implies that if we optimize the steady state solution $\overline{\Omega}_i$, 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 2 implies that when the targets are isolated there is an optimal control policy such that $u_j(t) \in \{-1, 0, 1\}$, even if the trajectory is constrained to be periodic. This property allows the optimal trajectory to be described by a finite set of parameters, as in [5], [9]. In this work in particular we are looking into periodic trajectories and, hence, this property implies that the movement in each period of 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. Therefore, one period of the trajectory of an agent j can fully be described by the following set of parameters:

- 1) T, the period of the trajectory.
- 2) $s_j(0)$, the initial position.
- ω_{j,p}, p = 1,..., P_j, the normalized dwelling times for agent j, i.e., the agent dwells for ω_{j,p}T units of time before it moves with maximum speed for the *p*-th time in the cycle.
- τ_{j,p}, p = 1, ..., P_j, the normalized movement times for agent j, i.e., the agent j moves for τ_{j,p}T units of time to the right (if p is odd) or to the left (if p is even) after dwelling for ω_{j,p}T units of time in the same position.

The following five constraints are enforced to ensure periodicity of the trajectory and consistency of its parameters. The last two constraints ensure, respectively, that the total time that the agents spend moving will be less or equal than a period and that the amount that over the course of a period, each agent will return to its initial position.

$$\tau_{j,m} \ge 0, \ \omega_{j,m} \ge 0, \ T \ge 0, \ \sum_{m=1}^{P_j} (\tau_{j,m} + \omega_{j,m}) \le 1$$

$$\sum_{m=1}^{P_j} (-1)^m \tau_{j,m} = 0.$$
(20)

Notice that this description does not exclude transitions of u_j of the kind $\pm 1 \rightarrow \mp 1$ and $\pm 1 \rightarrow 0 \rightarrow \pm 1$, 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 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 [5], [9] in which 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.

V. OPTIMIZATION OF THE PERIODIC TRAJECTORY

In the previous section, we described how, in an environment with isolated targets, the optimal steady state trajectory computation can be framed as an optimization of a finite set of parameters that represent the trajectory. We also showed that if this trajectory is used from any arbitrary initial condition, the mean cost will become arbitrarily close to the steady state cost. In this section we take advantage of the fact that, if every target is observed at least once, the Riccati equation is globally attractive in order to compute the derivative of the steady state solution of the Riccati equation with respect to all the parameters that are part of it. These can be used in a gradient descent scheme to obtain locally optimal steady state solution. Note that the suboptimality of the gradient descent in the context of persistent monitoring, along with an approach to converge to better local optima, is discussed in [6].

In this work, we take advantage of 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 [17]. 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

$$\dot{\Omega}_i(q) = \frac{d\Omega_i(q)}{dq} = T(A\Omega_i(q) + \Omega_i(q)A^T + Q - \eta_i(q)\Omega_i(q)G\Omega_i(q)), \quad (21)$$

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 derivative of (14), we have that for any parameter θ

$$\frac{\partial J}{\partial \theta} = \sum_{i=1}^{M} \int_{0}^{1} \frac{\partial \operatorname{tr}(\Omega_{i}(q))}{\partial \theta} dq.$$
 (22)

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 prior results 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 (except for T) yields

$$\frac{\partial \bar{\Omega}_i(q)}{\partial \theta} - T \left(A \frac{\partial \bar{\Omega}_i(q)}{\partial \theta} + \frac{\partial \bar{\Omega}_i(q)}{\partial \theta} A^T - \eta_i(q) \bar{\Omega}_i(q) G \frac{\partial \bar{\Omega}_i(q)}{\partial \theta} - \eta_i(q) \frac{\partial \bar{\Omega}_i(q)}{\partial \theta} G \bar{\Omega}_i(q) \right) = T \frac{\partial \eta_i(q)}{\partial \theta} \bar{\Omega}_i(q) G \bar{\Omega}_i(q), \quad (23)$$

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 steady state covariance matrix $\bar{\Omega}_i(q)$ as described in the previous section and explicit expressions for the term $\frac{\partial \bar{\Omega}_i}{\partial \theta}$ will be given in the next subsection.

Since (23) 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. Since $\bar{\Omega}_i(q)$ 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:

$$\dot{\Sigma}_H(q) - T\left(A - \eta_i(q)\bar{\Omega}_i(q)G\right)\Sigma_H(q) = 0, \ \Sigma_H(0) = I$$
(24)

and let Σ_{ZI} be the solution of (23) with the zero matrix as the initial conditions. Also, let Σ_H denote the solution of the homogeneous version of (23) with the identity matrix as the initial condition. Then, the initial conditions matrix Λ that yields a periodic solution of (23) is such that [18]:

$$\Lambda = \Sigma_H(1)\Lambda\Sigma_H^T(1) + \Sigma_{ZI}(1), \qquad (25)$$

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 4: Under the following conditions:

- Σ_H is a solution of (24);
- Assumptions 1 and 2 hold;
- target i is observed at least once in the period T;

• there exists a solution to (25);

Then, the solution to (25) is unique.

Proof: If Λ and Λ are solutions of (25), then

$$\Lambda - \tilde{\Lambda} = \Sigma_H(1) \left(\Lambda - \tilde{\Lambda} \right) \Sigma_H^T(1)$$
(26)

which is equivalent to

$$vec\left(\Lambda - \tilde{\Lambda}\right) = (\Sigma_H(1) \otimes \Sigma_H(1)) vec\left(\Lambda - \tilde{\Lambda}\right)$$
 (27)

Notice that $\Lambda = \tilde{\Lambda}$ is a solution of (27) and it is the unique solution if and only if 1 is not an eigenvalue of $\Sigma_H(1) \otimes \Sigma_H(1)$, and its eigenvalues are in the form $\mu_1\mu_2$, where μ_1 and μ_2 are distinct eigenvalues of $\Sigma_H(1)$ [19].

In the following we show that all the eigenvalues of $\Sigma_H(1)$ have absolute value lower than one. For that, first notice that

since Q is positive definite, then $\overline{\Omega}_i$ must be positive definite and hence, invertible. We can define

$$\mathcal{I} = \bar{\Omega}_i^{-1},$$

and, since $\dot{\mathcal{I}} = -\Omega_i^{-1} \dot{\Omega}_i \Omega_i^{-1} = -\mathcal{I} \dot{\Omega}_i \mathcal{I}$, the dynamics of \mathcal{I} can be expressed as:

$$\dot{\mathcal{I}} = -T(\mathcal{I}A + A^T \mathcal{I} + \mathcal{I}Q\mathcal{I} - \eta_i G).$$
(28)

Therefore, if we define the Lyapunov Function $V = \Sigma_H^T \mathcal{I} \Sigma_H$, we know that:

$$\frac{d}{dq} \left(\Sigma_H^T \mathcal{I} \Sigma_H \right) = \Sigma_H^T \left(\mathcal{I} A + AT \mathcal{I} + 2\eta_i G + \dot{\mathcal{I}} \right) \Sigma_H$$

$$= -\Sigma_H^T \mathcal{I} Q \mathcal{I} \Sigma_H.$$
(29)

By integrating the previous relation and using the fact that \mathcal{I} is periodic with period one, $\Sigma_H(0) = I$ and assumption 2, we know that

$$\Sigma_H(1)\mathcal{I}(0)\Sigma_H(1) - \mathcal{I}(0) = -\int_0^1 \Phi(q,0)^T \mathcal{I}Q\mathcal{I}\Phi(q,0) \,\mathrm{d}q < 0, \quad (30)$$

where $\Phi(q_1, q_2)$ is the transition matrix of the system (24) betwen times q_1 and q_2 . Since \mathcal{I} is positive definite, we know that the system is stable and therefore the absolute value of all the eigenvalues of $\Sigma_H(1)$ are lower than 1. This implies that $\mu_1\mu_2 < 1$ and therefore the solution of (25) has to be unique.

The Lyapunov Equation in (25) can be efficiently solved for low-dimensional systems using the algorithm proposed in [20] and implemented in MATLAB function dlyap. The needed derivative can then be computed as:

$$\frac{\partial \bar{\Omega}_i(q)}{\partial \theta} = \Sigma_H^T(q) \Lambda \Sigma_H(q) + \Sigma_{ZI}(q).$$
(31)

The same discussion also holds for the parameter T, with the differential equation (23) replaced by:

$$\frac{\partial \bar{\Omega}_{i}(q)}{\partial T} - T \left(A \frac{\partial \bar{\Omega}_{i}(q)}{\partial T} + \frac{\partial \bar{\Omega}_{i}(q)}{\partial T} A^{T} - \eta_{i}(q) \bar{\Omega}_{i}(q) G \frac{\partial \bar{\Omega}_{i}(q)}{\partial T} - \eta_{i}(q) \frac{\partial \bar{\Omega}_{i}(q)}{\partial T} G \bar{\Omega}_{i}(q) \right) = 4\bar{\Omega}_{i}(q) + \bar{\Omega}_{i}(q) A^{T} + Q - \left(\eta_{i}(q) + T \frac{\partial \eta_{i}(q)}{\partial T} \right) \bar{\Omega}_{i}(q) G \bar{\Omega}_{i}(q),$$
(32)

which is associated to the same homogeneous equation (24).

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

Looking back to (23), 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

$$\frac{\partial \eta_i(q)}{\partial \tau_{j,m}} = -\frac{I_j(s_j - x_i)}{r_j} \frac{\partial s_j(q)}{\partial \tau_{j,m}},\tag{33}$$

where

$$I_{j}(\alpha) = \begin{cases} +1, & 0 < \alpha < r_{j}, \\ -1, & -r_{j} < \alpha < 0, \\ 0, & |\alpha| > r_{j}. \end{cases}$$
(34)

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,

$$\frac{\partial \eta_i(q)}{\partial \omega_{j,m}} = -\frac{I_j(s_j - x_i)}{r_j} \frac{\partial s_j(q)}{\partial \omega_{j,m}},\tag{35}$$

$$\frac{\partial \eta_i(q)}{\partial s_j(0)} = -\frac{I_j(s_j - x_i)}{r_j} \frac{\partial s_j(q)}{\partial s_j(0)},\tag{36}$$

$$\frac{\partial \eta_i(q)}{\partial T} = -\sum_{j=1}^N \frac{I_j(s_j - x_i)}{r_j} \frac{\partial s_j(q)}{\partial T}.$$
(37)

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. These events, for our parameterization, 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 $2P_i$, where P_i 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 2 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

$$s_{j}(q) - s_{j}(0) = \begin{cases} T\left((-1)^{k/2+1}\left(q - \sum_{p=1}^{k/2-1}(\tau_{j,p} + \omega_{j,p}) + \omega_{j,p}\right) + \omega_{j,p}\right) \\ + \omega_{j,\frac{k}{2}} + \sum_{p=1}^{k/2}(-1)^{p+1}\tau_{p} \\ T \sum_{p=1}^{\frac{k-1}{2}}(-1)^{p+1}\tau_{j,p}, \ k \text{ odd.} \end{cases}$$
(38)

Therefore,

$$\frac{\partial s_j}{\partial \tau_{j,m}} = \begin{cases} \left((-1)^{\frac{k}{2}+1} + (-1)^p \right) T, & m < \frac{k}{2}, \ k \text{ even}, \\ (-1)^{m+1}T, \ m \le \frac{k-1}{2}, & k \text{ odd}, \end{cases}$$
(39)

$$\frac{\partial s_j}{\partial \omega_{j,m}} = \begin{cases} 1, \ m < \frac{k}{2}, & k \text{ even,} \\ 0 & , \text{ otherwise,} \end{cases}$$
(40)

$$\frac{\partial s_j(q)}{\partial T} = \frac{s_j(q) - s_j(0)}{T},\tag{41}$$

$$\frac{\partial s_j}{\partial s_j(0)} = 1. \tag{42}$$

C. Complete Optimization Procedure

In this subsection, we summarize our approach by gathering all the components of the optimization procedure into a single algorithm. First, we define the parameter set Θ_i^l :

$$\Theta_{j}^{l} = [s_{j}(0), \tau_{j,1}^{l}, ..., \tau_{j,P_{j}}^{l}, \omega_{j,1}^{l}, ..., \omega_{j,P_{j}}^{l}],$$
(43)

where the upper index l refers to the step number in the gradient descent, i.e.,

$$\Theta_j^{l+1} = \operatorname{proj}\left(\Theta_j^l - \kappa_l \frac{\partial J}{\partial \Theta_j}\right), \tag{44}$$

where proj represents the projection into the convex set defined by the constraints in (20) and κ_l is the gradient descent step size. Algorithm 1 describes the complete optimization procedure.

In this paper, a procedure for obtaining Θ_i^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. IV, 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 [9], possibly augmenting it with the technique proposed in [6], where the cost function was augmented to provide a larger exploration of the environment by the gradient descent algorithm.

Also, it is worth noting that although we have already shown that the steady state solution of the periodic Riccati equation is globally attractive, no convergence rate was indicated. There are, however, various alternative numerical methods that can provide guaranteed convergence that could replace line 21 in Alg. 1. We refer the reader to [21] for a more complete discussion of these methods.

VI. SIMULATION RESULTS

In this section, we demonstrate the results of Algorithm 1 in a two different scenarios, one with only one agent and two targets and a second one with five targets and two agents. All targets *i* have the same state dynamics evolving according to (1) with

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

and the same observation model as in (3) with

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

A constant descent stepsize was used ($\kappa_l = \kappa_0$) and, order to provide an index convention for the targets, we define that $x_1 < x_2 < \dots < x_M.$

A. One agent, two targets

In the first scenario, with one agent and two targets, the following set of initial conditions was used:

$$s_1^0(0) = 0, \quad T^0 = 6, \quad \tau_1^0 = [0.2, 0.4, 0.2],$$

Algorithm 1 Agents' Trajectory Optimization

1: procedure GRADIENT DESCENT Input: $\Theta_1^0, ..., \Theta_N^0, T^0,$ 2: $||\nabla J|| \leftarrow \infty$ 3: 4: $l \leftarrow 0$ while $||\nabla J|| > \epsilon$ do 5: $\begin{bmatrix} \frac{\partial J}{\partial \Theta_1}, ..., \frac{\partial J}{\partial \Theta_N}, \frac{\partial J}{\partial T} \end{bmatrix} \leftarrow \text{IPA}(\Theta_1^l, ..., \Theta_N^l, T^l)$ **for** j from 1 to N **do** $\Theta_j^{l+1} \leftarrow \text{proj} \left(\Theta_j^l - \kappa_l \frac{\partial J}{\partial \Theta_j}\right)$ 6: 7: 8: $T^{l+1} \leftarrow T^{l} - \kappa_{l} \frac{\partial J}{\partial T}$ $||\nabla J|| \leftarrow \frac{1}{\kappa_{k}} \sqrt{\left(\frac{\partial J}{\partial T}\right)^{2} + \sum_{j=1}^{N} \left\|\Theta_{k}^{l+1} - \Theta_{k}^{l}\right\|^{2}}$ $l \leftarrow l+1$ 9: 10: 11: **Output:** $\underline{\theta}_1^l, \dots, \underline{\theta}_N^l, \underline{\omega}_1^l, \dots, \underline{\omega}_N^l$ 12: 13: 14: procedure IPA **Input**: $\Theta_1, ..., \Theta_N, T$ 15: for j from 1 to N do 16: $\tfrac{\partial J}{\partial \Theta_i} \leftarrow 0$ 17: $\frac{\partial J}{\partial T} \leftarrow 0$ 18: Compute $s_1(q), ..., s_N(q)$ from the parameterization 19: for i from 1 to M do 20: Compute $\overline{\Omega}_i(q)$ by running to (21) until it con-21: verges to a periodic solution Compute Σ_H^i from (24) 22: 23: for j from 1 to N do for every θ in Θ do 24: Solve (23) with zero initial conditions to 25: compute $\Sigma_{ZI}^{\theta,i}$ Compute $\Lambda^{\theta,i}$ using (25) 26:

27:
$$I \leftarrow \int_0^1 \operatorname{tr} \left((\Sigma_H^i)^T \Lambda^{\theta, i} \Sigma_H + \Sigma_{ZI}^{\theta, i} \right) dq$$
28:
$$\frac{\partial J}{\partial \theta_i} \leftarrow \frac{\partial J}{\partial \theta_i} + I$$

28:

Solve (32) with zero initial conditions to compute 29: $\Sigma_{ZI}^{T,i}$

30: Compute
$$\Lambda^{T,i}$$
 using (25)
31: $I \leftarrow \int_0^1 \operatorname{tr} \left((\Sigma_H^i)^T \Lambda^{T,i} \Sigma_H^i + \Sigma_{ZI}^{T,i} \right) dq$
32: $\frac{\partial J}{\partial T} \leftarrow \frac{\partial J}{\partial T} + I$
33: **Output:** $\frac{\partial J}{\partial \Theta_1}, ..., \frac{\partial J}{\partial \Theta_j}, \frac{\partial J}{\partial T}$

$$\omega_1^0 = [0.05, 0.05, 0.05],$$

and the gradient descent step size was set to $\kappa_0 = 0.02$.

Figure 2 illustrates the results of running the simulation in this scenario. By analyzing Figures 2b and 2c, the optimized policy is such that the agent moves between the two targets and dwells on top of them in a symmetric way. The steady state covariances behave with very similar curves, but shifted in time. Also notice that the period is lower than the original one.

B. Two agents and Five targets

In this second scenario, the targets were placed in positions $x_i = 1 + 2i, i = 1, ..., 5$. The initial parameters were the



Fig. 2: Results of a simulation with one agent and two targets. (a) Evolution of the overall cost as a function of iteration number on the gradient descent. (b) Trajectories of the agent 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 covariance matrices of the two targets.

following:

$$s_1^0(0) = 2.7, \quad s_2(0) = 6.8, \quad T^0 = 6,$$

 $\tau_1^0 = \tau_2^0 = 0.1[1, 0.1, 1, 1, 0.1, 1, 0.1, 1, 1, 0.1, 1],$
 $\omega_1^0 = \omega_2^0 = 0.0125[1, 1, 1, 1, 1, 1, 1, 1, 1, 1],$

and the gradient descent step size was set to be constant, $\kappa_0 = \kappa_l = 0.02$.

Figure 3 shows the results of the optimization in this scenario. Notice that even though both agents and all the targets have the same dynamical 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. 3b is that, while between times 6 and 8 agent 1 makes a movement with small amplitude around target 1, the effects of this oscillatory movement are hard to notice in the trace of the covariance of target 1 in Fig. 3c. 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. 3a, the results are shown this way to highlight interesting aspects of the process. 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.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we developed a technique both to analyze and to optimize the steady state mean 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. Among the open questions we plan to address in future work are the following.

- Do the gradients of Ω_i with respect to the parameters that define the trajectory always exist in the interior of the set where they lead to a convergent Ω_i?
- How can we efficiently generate initial trajectories in order to converge to global optimal points or, at least, good local optima?

Also, the simulated results highlight the interesting feature that the locally optimal solution split the set of targets into indepedent sets. That is, no targets were shared by agents. Even though this might not always hold in general, this feature motivates the future investigation of policies where only one agent observes each target and the agents would not necessarily be constrained to the same movement period. We also plan to extend the results here presented to scenarios where the agents are not constrained to a single dimensions, possibly using suboptimal parameterizations for the trajectory, as in [22].

REFERENCES

- 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.
- [2] 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.
- [3] 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.
- [4] —, "Optimal Visiting Schedule Search for Persistent Monitoring of a Finite Set of Targets," in *American Control Conference (ACC)*. IEEE, 2018, pp. 4032–4037.
- [5] 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.



Fig. 3: 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 covariance matrices of the five targets.

- [6] 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.
- [7] 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.
- [8] —, "A Variational Approach to Trajectory Planning for Persistent Monitoring of Spatiotemporal Fields," in 2014 American Control Conference. IEEE, 2014, pp. 5627–5632.
- [9] 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 *Proceedings of the 2019 Control and Decision Conference (to appear)*. IEEE.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] P. Lancaster, "On Eigenvalues of Matrices Dependent on a Parameter," *Numerische Mathematik*, vol. 6, no. 1, pp. 377–387, 1964.
- [14] 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.
- [15] S. Bittanti, P. Colaneri, and G. Guardabassi, "Periodic Solutions of Periodic Riccati Equations," *IEEE Transactions on Automatic Control*, vol. 29, no. 7, pp. 665–667, 1984.
- [16] G. Nicolao, "On the Convergence to the Strong Solution of Periodic Riccati Equations," *International Journal of Control*, vol. 56, no. 1, pp. 87–97, 1992.
- [17] C. G. Cassandras, Y. Wardi, C. G. Panayiotou, and C. Yao, "Perturbation Analysis and Optimization of Stochastic Hybrid Systems," *European Journal of Control*, vol. 16, no. 6, pp. 642–661, 2010.
- [18] W. T. Reid, "A Matrix Differential Equation of Riccati Type," American Journal of Mathematics, vol. 68, no. 2, pp. 237–246, 1946.
- [19] F. Zhang, Matrix Theory: Basic Results and Techniques. Springer Science & Business Media, 2011.
- [20] A. Barraud, "A Numerical Algorithm to Solve AT XA-X=Q," in 1977 IEEE Conference on Decision and Control including the 16th Symposium on Adaptive Processes and A Special Symposium on Fuzzy Set Theory and Applications, no. 16, 1977, pp. 420–423.
- [21] A. Varga, "Computational Issues for Linear Periodic Systems: Paradigms, Algorithms, Open Problems," *International Journal of Control*, vol. 86, no. 7, pp. 1227–1239, 2013.

[22] 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.