Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Multi-Agent Infinite Horizon Persistent Monitoring of Targets with Uncertain States in Multi-Dimensional Environments ⋆

Samuel C. Pinto * **Sean B. Andersson** *,*** **Julien M. Hendrickx** **** **Christos G. Cassandras** **,***

* *Department of Mechanical Engineering*
** *Department of Electrical and Computer Engineering*
*** *Division of Systems Engineering*
Boston University, Boston, MA 02215 USA.
**** *ICTEAM Institute, UCLouvain, Louvain-la-Neuve 1348, Belgium.*
e-mail:{samcerq,sanderss,cgc}@bu.edu,julien.hendrickx@uclouvain.be

Abstract: This paper investigates the problem of persistent monitoring, where a finite set of mobile agents persistently visits a finite set of targets in a multi-dimensional environment. The agents must estimate the targets' internal states and the goal is to minimize the mean squared estimation error over time. The internal states of the targets evolve with linear stochastic dynamics and thus the optimal estimator is a Kalman-Bucy Filter. We constrain the trajectories of the agents to be periodic and represented by a truncated Fourier series. Taking advantage of the periodic nature of this solution, we define the infinite horizon version of the problem and explore the property that the mean estimation squared error converges to a limit cycle. We present a technique to compute online the gradient of the steady state mean estimation error of the targets' states with respect to the parameters defining the trajectories and use a gradient descent scheme to obtain locally optimal movement schedules. This scheme allows us to address the infinite horizon problem with only a small number of parameters to be optimized.

Keywords: Persistent Monitoring, Multi-Agent, Trajectory Optimization.

1. INTRODUCTION

In the problem of persistent monitoring, we consider a set of targets which have internal states that accumulate uncertainty over time. A finite set of mobile agents with sensing capabilities moves around the environment with the goal of keeping the mean uncertainty of the target states as low as possible. It is often the case that not all the targets can be continuously monitored, hence, the motion policy of the agents has to be carefully planned in order to reduce the mean uncertainty and to make sure that it does not increase without bound as time goes to infinity.

This paradigm finds applications across a wide range of domains, such as trajectory planning of underwater vehicles to measure ocean temperature (Lan and Schwager, 2013; Alam et al., 2018), surveillance in smart cities (Kim et al., 2018) and tracking of multiple macromolecules by an optical microscope (Shen and Andersson, 2010). A very wide range of approaches to address this problem has been proposed by different researchers (Nigam, 2014; Yu et al., 2015; Zhou et al., 2018). While different works consider different ways of modeling the uncertainty of the target states, the approach in the present paper uses the model presented in (Pinto et al., 2019), where the targets are fixed in space, each one of them has an internal state that evolves with linear stochastic dynamics, and the measurement quality of an observing agent decays as that agent moves away from the target. For this model, it has been shown that the optimal estimator is a Kalman-Bucy filter and that the objective of reducing the mean uncertainty can be directly associated to the covariance matrix of this estimator.

In our previous work, (Pinto et al., 2020) the infinite horizon persistent monitoring problem was addressed by restricting the movement schedules to be periodic and computing the steady state estimation error of the target states. This work restricted both the agents and targets to be in a one-dimensional environment, for which, under some assumptions on the spatial distribution of the targets, it was shown that an optimal periodic policy can be represented by a finite set of parameters. However, finding a finite parameterization for an optimal policy in a general multi dimensional environment is a much more complex problem, as discussed, e.g., in (Lin and Cassandras, 2014).

⋆ This work was supported in part by NSF under grants 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.

In this work, we extend the investigation of the infinite horizon persistent monitoring problem to the multi-dimensional setting. Instead of looking for an exact representation of an optimal control policy, we restrict ourselves to a parameterized family of curves by representing the periodic movement in each coordinate (e.g., x, y and z, for 3D spaces) of the multidimensional space as a truncated Fourier series. This representation can provide very general motion curves using only a small number of coefficients. Also, even though this work is limited to simple agent dynamics, this parameterization produces smooth curves that can be used as reference trajectories for a wide range of dynamics of the agents, as long as these trajectories are used in conjunction with trajectory tracking controllers.

We use results originally derived in (Pinto et al., 2020) to compute gradients of the steady state estimation error with respect to the coefficients that represent the trajectory and optimize them using a centralized gradient descent scheme. However, in order to optimize the trajectory, it is necessary to find an initial periodic schedule that visits all the targets in a period, which avoids unbounded growth in the uncertainty. In this paper, we introduce the idea of combining these conditions along with a heuristic solution of the Multiple Traveling Salesman Problem (Bektas, 2006) to generate initial conditions that guarantee a feasibile and efficient solution of the optimization.

2. PROBLEM FORMULATION

Consider an environment with a set of M points of interest (targets) at fixed positions $x_i \in \mathbb{R}^P$, $i = 1, ..., M$. Each of these targets has an internal state $\phi_i \in \mathbb{R}^{L_i}$ that needs to be monitored and that evolves according to linear time-invariant stochastic dynamics

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

where $w_i(t)$ is a white noise process distributed according to $w_i(t) \sim \mathcal{N}(0, Q_i)$, $i = 1, \ldots, M$, and w_i and w_j are statistically independent if $i \neq j$.

Suppose that there is a collection of N mobile agents at positions $s_i(t) \in \mathbb{R}^P$ that can move with the following kinematic model:

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

where u_j is a controllable input. Each of these agents is equipped with sensors that can observe the targets according to the following 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 $v_{i,j}(t)$ is a white noise process distributed according to $v_{i,j}(t) \sim \mathcal{N}(0, R_i)$ with $v_{i,j}(t)$ independent of $v_{k,l}$ if $i \neq k$ or $j \neq l$, and $\gamma_{i,j} : \mathbb{R}^N \mapsto \mathbb{R}$ is a function that captures the coupling between measurement quality and the relative position from a given agent to the target. It is worth noting that in most of the applications of mobile agents to sensing it is assumed that there is a limited sensing range or that the quality of the measurement gets worse as the agent moves farther away from the target. The general model of $\gamma_{i,j}$ is capable of capturing both the finite range and the dependence between measurement quality and relative position of the target from the agent. Even though the analysis in this paper does not depend on the specific $\gamma_{i,j}$, for the sake of concreteness we use the following expression:

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

The intuition behind this specific form is that the best measurement quality is achieved when the agent is on top of the target with the quality decaying as it moves farther, until the agents reaches the distance of its sensing radius r_j, from where only noise can be observed.

In this paper we approach the problem from a centralized perspective. Therefore, at a given instant, the combined observations from all the agents of a single target can be grouped in a vector $\tilde{z}(t)$ as:

$$z_i(t) = \begin{bmatrix} z_{i,1}^T & \ldots & z_{i,N}^T \end{bmatrix}^T = \tilde{H}_i(s_1, ..., s_n)\phi_i(t) + \tilde{v}_i(t) \tag{5}$$

where

$$\tilde{H}_i = \begin{bmatrix} \gamma_1(s_1 - x_i)H_i^T & \cdots & \gamma_N(s_N - x_i)H_i^T \end{bmatrix}^T, \tag{6}$$

$$\tilde{v}_i(t) = \begin{bmatrix} v_{i,1}^T(t) & \ldots & v_{i,N}^T(t) \end{bmatrix}^T, \tag{7}$$

and

$$E[\tilde{v}_i^T(t)\tilde{v}_i(t)] = \tilde{R}_i = diag(R_i, ..., R_i). \tag{8}$$

The overall goal is to obtain estimators $\hat{\phi}_i(t, z(t))$ and open-loop control inputs $u_j(t)$ to minimize the following infinite horizon cost, when the limit exists:

$$J = \lim_{t\to\infty} \frac{1}{t} \int_0^t \left(\sum_{i=1}^M E[e_i^T(\zeta)e_i(\zeta)] + \xi \sum_{j=1}^N u_j^T(\zeta)u_j(\zeta) \right) d\zeta, \tag{9}$$

where $e_i(t) = \hat{\phi}_i(t) - \phi_i(t)$. Note that this formulation has a minor but important difference from our previous work (Pinto et al., 2020), where there was no penalization ξ for the control effort in the cost but the control was bounded such that $u_j(t) \in [-1, 1]$. The penalization of the control effort avoids unbounded growth in the speed commanded to the agents by optimizing the mean power consumption over the trajectory.

The models in (6) and (7) define a linear time-varying stochastic system and it has been shown in (Lan and Schwager, 2014) that, for a finite time version of the cost in (12), the optimal estimator is a Kalman-Bucy filter. For the sake of conciseness, the proof of the fact that the Kalman-Bucy filter is still the optimal estimator in the infinite horizon setting of (12) is omitted, but the derivation is analogous to the one in (Lan and Schwager, 2014). The dynamics of the filter are given by:

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

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

where $\Omega_i(t)$ is the covariance matrix of the estimator. Using (6) and (7), we can rewrite (10b) as:

$$\dot{\Omega}_i(t) = A_i\Omega_i(t) + \Omega_i(t)A_i^T + Q_i - \Omega_i(t)G_i\Omega_i(t)\sum_{j=1}^N \gamma_{i,j}^2(t), \tag{11}$$

where $G_i = H_i^T R_i^{-1} H_i$ and $\gamma_{i,j}(t) = \gamma_{i,j}(s_j(t) - x_i)$. Using the fact that

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

we can rewrite the cost function in (9) as

$$J = \lim_{t\to\infty} \frac{1}{t} \int_0^t \left(\sum_{i=1}^{M} \mathrm{tr}(\Omega_i(\zeta)) + \xi \sum_{j=1}^{N} u_j^T(\zeta) u_j(\zeta) \right) d\zeta. \tag{12}$$

Notice that the goal is to minimize the cost (12) subject to the dynamics in (11) and (2). In other words, the goal is to design a trajectory that minimizes a weighted sum of the mean control effort and the mean estimation error. The estimation error and the trajectory are linked through the dynamics of the covariance matrix (11).

3. PERIODIC PERSISTENT MONITORING

The concept of persistent monitoring is inherently linked to the idea of visiting the targets infinitely often. In this work, we constraint ourselves to look for periodic trajectories since under some natural assumptions the covariance matrices of the estimation error converge to steady-state periodic matrices. Therefore, as time goes to infinity the mean estimation error approaches the steady state estimation error and transient effects can be neglected. Hence it suffices to take into consideration the steady state behavior of the estimation error, simplifying the trajectory planning.

In order to proceed to a more complete discussion of the steady state behavior, we state two very natural assumptions.

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 guarantees that sensing can make the uncertainty bounded even for long horizons. The second one imposes that the covariance matrix is always positive definite, a property used to prove Prop. 2. We also define

$$\eta_i(t) = \sum_{j=1}^{N} \gamma_{i,j}^2(t), \tag{13}$$

which represents the instantaneous power level of the signal, combining all the agents' observations of the same target i. In (Pinto et al., 2020) the following proposition was established:

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

Prop. 1 implies that, if $\eta_i(t)$ is periodic, given any initial covariance matrix $\Omega_i(0)$, the estimation covariance for target i converges to a T-periodic matrix $\bar{\Omega}_i(t)$, as long as target i is visited for some non-zero amount of time in the periodic trajectory. Therefore,

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

which implies that

$$\lim_{t\to\infty} \frac{1}{t} \int_{t_0}^{t} |\mathrm{tr}(\tilde{\Omega}_i(\zeta) - \Omega_i(\zeta))| d\zeta \le \delta,$$

and, since $\lim_{t\to\infty} \frac{1}{t} \int_0^{t_0} |\mathrm{tr}(\tilde{\Omega}_i(\zeta) - \Omega_i(\zeta))| d\zeta = 0$, we can add up the integration limits and conclude that

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t |\mathrm{tr}(\tilde{\Omega}_i(\zeta) - \Omega_i(\zeta))| d\zeta \le \delta. \tag{14}$$

This discussion implies that if we run a periodic trajectory long enough, the mean estimation error will become arbitrarily close to the mean steady state estimation error. Therefore, if we plan only (one period of) the steady state trajectory, the actual estimation error will be arbitrairly close to the one we planned for the case where time goes to infinity. Note that even though Prop. 1 states that the solution of the periodic Riccati equation is globally attractive, it does not provide any convergence rate for its numerical computation. However, the problem of computing numerical solutions to this equation has been studied in other works and we refer the reader to (Varga, 2013) for a good review and discussion of these methods.

Using similar ideas to the one-dimensional case in (Pinto et al., 2020), we plan to use gradient descent to optimize parameters that describe the periodic trajectory of the agents in order to reduce the steady state estimation error of the target states. In the rest of this section, we will provide a procedure to compute these gradients, along with a procedure to provide an initial parameter configuration for the optimization.

3.1 Steady State Gradients

Assuming that the trajectory is periodic and all the targets are visited, we introduce the change of variables $q = t/T$, where T is the period of the trajectory. The cost can be rewritten as:

$$J = \int_0^1 \left(\sum_{i=1}^{M} \mathrm{tr}(\bar{\Omega}_i(q)) + \xi \sum_{j=1}^{N} \bar{u}_j^T(q) \bar{u}_j(q) \right) dq, \tag{15}$$

where $\bar{u}(q) = u(qT)$. The dynamics of $\bar{\Omega}_i(q)$ are

$$\begin{aligned}\dot{\bar{\Omega}}_i(q) = \frac{d\bar{\Omega}_i(q)}{dq} = T(A\bar{\Omega}_i(q) + \bar{\Omega}_i(q)A^T \\ + Q - \eta_i(q)\bar{\Omega}_i(q)G\bar{\Omega}_i(q)).\end{aligned} \tag{16}$$

Now, suppose that the trajectory of the agents is represented by a finite set of parameters $\theta_1, ..., \theta_{\mathcal{O}}$. If we want to compute the gradient with respect to a parameter, we get the following dynamics:

$$\begin{aligned}\frac{\partial\dot{\bar{\Omega}}_i(q)}{\partial\theta} - T\Bigg(A\frac{\partial\bar{\Omega}_i(q)}{\partial\theta} + \frac{\partial\bar{\Omega}_i(q)}{\partial\theta}A^T \\ - \eta_i(q)\bar{\Omega}_i(q)G\frac{\partial\bar{\Omega}_i(q)}{\partial\theta} - \eta_i(q)\frac{\partial\bar{\Omega}_i(q)}{\partial\theta}G\bar{\Omega}_i(q)\Bigg) = \\ T\frac{\partial\eta_i(q)}{\partial\theta}\bar{\Omega}_i(q)G\bar{\Omega}_i(q) + \frac{\partial T}{\partial\theta}\frac{\dot{\bar{\Omega}}_i(q)}{T}.\end{aligned} \tag{17}$$

Notice that if the partial derivative $\partial\bar{\Omega}_i(q)/\partial\theta$ exists, it is a 1-periodic solution of (17), since $\bar{\Omega}_i$ itself is periodic with period one. If we define the following auxiliary problems:

$$\dot{\Sigma}_H - T\left(A - \eta_i\bar{\Omega}_i G\right)\Sigma_H = 0,\ \Sigma_H(0) = I, \tag{18}$$

$$\begin{aligned}\dot{\Sigma}_{ZI} - T\left(A - \eta_i\bar{\Omega}_i G\right)\Sigma_{ZI} - T\Sigma_{ZI}^T\left(A - \eta_i\bar{\Omega}_i G\right)^T \\ = T\frac{\partial\eta_i}{\partial\theta}\bar{\Omega}_i G\bar{\Omega}_i + \frac{\partial T}{\partial\theta}\frac{\dot{\bar{\Omega}}_i(q)}{T},\ \Sigma_{ZI}(0) = 0,\end{aligned} \tag{19}$$

then we can use the following proposition, introduced in (Pinto et al., 2020), to compute the partial derivatives:

Proposition 2. Suppose Σ_H is a solution of (18), Σ_{ZI} is a solution of (19), Assumptions 1 and 2 hold and that target i is observed at least once in the period T. Then, if the solution Λ of the equation

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

exists, it is unique, and

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

4. FOURIER CURVE REPRESENTATION FOR AGENT TRAJECTORIES

While in previous work (Pinto et al., 2019, 2020) we derived a parameterization with a finite number of parameters of the optimal solution, the same result does not extend to the multi-dimensional scenario. Therefore, instead of looking for an exact representation of the optimal trajectory, we focus on a family of parameterized curves that can approximate very general curves. Since periodicity is an essential feature of the paradigm discussed in this work, a natural choice is to use a truncated Fourier series to represent the movement of the agents in each of the coordinates e_p, $p = 1, ..., P$, i.e.

$$s_j^{e_p}(q) = s_{j,0}^{e_p} + \sum_{k=1}^{K} a_{j,k}^{e_p} \sin(2\pi f_k q) + b_{j,k}^{e_p}(\cos(2\pi f_k q) - 1), \tag{20}$$

where f_k are integer frequencies and, therefore, $s_j^{e_p}(q)$ is periodic with period 1. The set of parameters that fully characterize all the agents trajectories is $\Theta = \{\{a_{j,k}^{e_p}\}, \{b_{j,k}^{e_p}\}, \{s_{j,0}^{e_p}\}, T\}$, $j = 1, .., N$, $p = 1, ..., P$, $k = 1, ..., K$. In order to compute the derivative of the covariance matrix using the procedure introduced in Prop. 2, we still need to give a procedure to compute $\frac{\partial\eta_i}{\partial\theta}$, where η_i is defined in (13). For any parameter $\theta \in \Theta$,

$$\frac{\partial\eta_i(q)}{\partial\theta} = \sum_{j=1}^{N}\sum_{p=1}^{P} \frac{\partial\eta_j(q)}{\partial s_j^{e_p}} \frac{\partial s_j^{e_p}(q)}{\partial\theta}, \tag{21}$$

and, using (4) and the fact that $\eta_i(t) = \sum_{j=1}^{N} \gamma_{i,j}^2(t)$,

$$\frac{\partial\eta_j}{\partial s_j^{e_i}} = \begin{cases} \frac{s_j^{e_p} - x_i^{e_p}}{r_j\|s_j - x_i\|}, & \text{if } \|s_j - x_i\| < r_j, \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

Moreover, in this parameterization,

$$\frac{\partial s_j^{e_p}}{\partial a_{m,k}^{e_r}} = \begin{cases} \sin(2\pi f_k q), & \text{if } j = m \text{ and } p = r, \\ 0, & \text{otherwise,} \end{cases} \tag{23a}$$

$$\frac{\partial s_j^{e_p}}{\partial b_{m,k}^{e_r}} = \begin{cases} \cos(2\pi f_k q) - 1, & \text{if } j = m \text{ and } p = r, \\ 0, & \text{otherwise,} \end{cases} \tag{23b}$$

$$\frac{\partial s_j^{e_p}}{\partial s_{m,0}^{e_r}} = \begin{cases} 1, & \text{if } j = m \text{ and } p = r, \\ 0, & \text{otherwise,} \end{cases} \tag{23c}$$

$$\frac{\partial s_j^{e_p}}{\partial T} = 0. \tag{23d}$$

The equations above give enough information to compute the partial derivatives of the steady state covariance matrix as indicated in Prop. 2. In order to compute the gradient of the cost function, the following expression can be used:

$$\frac{\partial J}{\partial\theta} = \int_0^1 \sum_{i=1}^{N} \mathrm{tr}\left(\frac{\partial\Omega_i}{\partial\theta}\right) dq + \xi \frac{\partial}{\partial\theta} \sum_{j=1}^{N} \int_0^1 \left\| \frac{ds_j}{dt} \right\|^2 dq. \tag{24}$$

Note that

$$\frac{ds_j}{dq} = T\frac{ds_j}{dt}. \tag{25}$$

Using (20), we can compute

$$\sum_{j=1}^{N} \int_0^1 \left\| \frac{ds_j}{dt} \right\|^2 dq = \sum_{j=1}^{N}\sum_{p=1}^{P}\sum_{k=1}^{K} \frac{(2\pi f_k)^2}{2T^2} \left(\left(a_{j,k}^{e_p}\right)^2 + \left(b_{j,k}^{e_p}\right)^2 \right), \tag{26}$$

and, therefore,

$$\frac{\partial}{\partial a_{j,k}^{e_p}} \sum_{j=1}^{N} \int_0^1 \left\| \frac{ds_j}{dt} \right\|^2 dq = \frac{(2\pi f_k)^2}{2T^2} a_{j,k}^{e_p}, \tag{27a}$$

$$\frac{\partial}{\partial b_{j,k}^{e_p}} \sum_{j=1}^{N} \int_0^1 \left\| \frac{ds_j}{dt} \right\|^2 dq = \frac{(2\pi f_k)^2}{2T^2} b_{j,k}^{e_p}, \tag{27b}$$

$$\frac{\partial}{\partial s_{j,0}^{e_p}} \sum_{j=1}^{N} \int_0^1 \left\| \frac{ds_j}{dt} \right\|^2 dq = 0, \tag{27c}$$

$$\frac{\partial}{\partial T} \sum_{j=1}^{N} \int_0^1 \left\| \frac{ds_j}{dt} \right\|^2 dq = \sum_{j=1}^{N}\sum_{p=1}^{P}\sum_{k=1}^{K} \frac{-(2\pi f_k)^2}{T^3} \left(\left(a_{j,k}^{e_p}\right)^2 + \left(b_{j,k}^{e_p}\right)^2 \right). \tag{27d}$$

Algorithm 1 summarizes the procedure to compute the gradient of the cost function of the parameterized trajectory.

Algorithm 1 Agents' Trajectory Optimization

```
procedure GRADIENT DESCENT
    Input: Θ^0,
    ||∇J|| ← ∞
    l ← 0
    while ||∇J|| > ϵ do
        ∇J ← ComputeGradient(Θ^l)
        Θ^{l+1} ← Θ^l − κ_l ∇J
        l ← l + 1
    Output: Θ^l

procedure COMPUTEGRADIENT
    Input: Θ
    Compute s_1(q), ..., s_N(q) from the parameterization
    for i ranging from 1 to M do
        Compute the steady state covariance Ω̄_i(q)
    for every θ in Θ do
        for i ranging from 1 to M do
            Compute ∂Ω_i(q)/∂θ as indicated in Prop. 2
        Compute ∂J/∂θ using (24) and (27)
    Output: ∇J
```

4.1 Initial Parameter Computation

Alg. 1 describes a gradient descent approach to locally minimize the cost function. However, in order to minimize this cost, it is necessary to provide an initial parameter configuration for the algorithm. Prop. 1 states that, if every target is visited at least once, then the steady-state covariance matrix exists. However, if in a periodic trajectory one of the targets is never visited and its internal state dynamics are unstable, then the estimation error will grow without bound, and, therefore, the steady-state covariance matrix would not exist. Therefore, such initial condition must be excluded by ensuring that for any parameters we select, the corresponding agent trajectories visit all the targets. In this section, we discuss a method for finding these initial trajectories that will always lead to a feasible initial configuration. Note that while we rely on the gradient descent optimization to provide optimized solutions, the problem described in this paper is non-convex and different initial conditions can lead to different local optima. We, therefore, leverage intuition about the problem to provide reasonable initial solutions with the hope that they will converge to good local optima. Approaching the problem of initialization from a more systematic point of view is a topic of current research.

The idea of finding a schedule where all the targets are visited fits naturally into a graph search paradigm, where the targets are modelled as nodes and the edge weights between nodes are the distances between the targets. The problem of finding a feasible schedule can be translated to the one of finding N sequences (that represent the schedule of each agent) of nodes where each target is at least in one of these sequences. One can add to that a cost function that guides the way in which these sequences are created. A goal that intuitively will lead to reasonable initial solutions is to minimize the distance of the agent that has the longest travel path. This is the well known Multiple Traveling Salesman Problem (MTSP). (Bektas, 2006) provides a good overview of this problem and approaches to solve it. It is worth mentioning that the MTSP is NP-Hard, and therefore intractable to be solved exactly in all but the simplest scenarios. However, meta-heuristic approaches can provide feasible, though not necessarily optimal, solutions. In this work, we use the genetic algorithm described in (Tang et al., 2000) to find heuristic solutions. This approach is interesting because it finds a feasible solution in the first iteration and refines it as the number of iterations increases. Therefore, one can decide how much computation time to spend in this solution, leveraging the tradeoff between optimality and computation effort spent in this initial trajectory.

The solution of the MTSP problem gives, for each agent j, a cyclic schedule of targets $\mathcal{S}_j = \{y_j^1, ..., y_j^{Y_j}, y_j^1\}$. However, it is still necessary to obtain the parameters $\Theta = \{\{a_{j,k}^{e_p}\}, \{b_{j,k}^{e_p}\}, \{s_{j,0}^{e_p}\}, T\}$ from this schedule. We define d_j^m as the cumulative distance that the agent has traveled when it reaches the m-th target in the schedule $\mathcal{S}_j$, and D_j as the total distance traveled by an agent in one cycle. We then look for a feasible truncated Fourier series trajectory such that at the normalized time $q = d_j^m/(D_j T)$, the agent is at a distance $(s_j(d_j^m/(D_j T)) - x_{y_j^m})$ lower or equal than the sensing radius (multiplied by a factor $1-\delta$, $0 < \delta < 1$, in order to give some distance margin) from the target. The position of the agent in the beginning of the cycle is set to be the position of the first target in the schedule $\mathcal{S}_j$, and the period T set to any positive number. For each of the agents, the following optimization problem gives a set of feasible $\{a_{j,k}^{e_p}\}, \{b_{j,k}^{e_p}\}$. The constraint in the optimization problem represents the fact that the agent must be close enough to see the target in a time analogous to the one it visited the same target in the MTSP solution.

$$\begin{aligned} \min_{a_{j,k}^{e_p},\, b_{j,k}^{e_p}} \quad & \sum_{p=1}^{P}\sum_{k=1}^{K} f_k|a_{j,k}^{e_p}| + f_k|b_{j,k}^{e_p}| \\ \text{s.t.} \quad & \left\| s_j\left(\frac{d_j^m}{D_j}\right) - x_{y_j^m} \right\|_2 \leq (1-\delta) r_j,\ m = 1,..,Y_j \end{aligned} \tag{28}$$

Substituting the definition in (20) into the constraint (28), we see that this optimization can be formulated as a Quadratically Constrained Program, which is a convex optimization problem and there exist efficient algorithms to solve it. The one norm in the objective function in (28) was selected because in our experience this yielded smooth trajectories. Other objectives functions could be used instead.

The trajectory generated by the heuristic solution of the MTSP problem consists of segments of straight lines that visit each of the targets in the schedule $\mathcal{S}_j$. Note that this trajectory, as a function of time, composed by a sequence of straight lines can be projected onto each of the axis e_p and the projection in that axis will still be a sequence of segments of straight lines. Since piecewise linear functions can be represented by Fourier series, there always exists a K large enough such that there is a solution to (28) because for that K there is a representation of the trajectory that would be close enough to the original MTSP solution such that it satisfies the constraint in (28). Therefore, we can always find feasible solutions to (28) if we have a MTSP solution.

5. SIMULATION RESULTS

In this section, we demonstrate the results of the algorithm in simulated 2D scenarios, with three agents and 15 targets. All the internal states of the targets have the same state dynamics, evolving according to (1) with

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

and the agents observation models (3) with

$$H_i = R_i = \text{diag}(1,1), \quad r_j = 0.5, \quad \xi = 10^{-3}.$$

For each of the agents, their trajectories had the first five harmonics in each axis, i.e., $f_k = k$, $k = 1, ..., K_j$ and $K_j = 5, \forall\, j$. In the initial step of the optimization, the period T^0 was set to 1. The initial coefficients $a_{j,k}^{e_p}, b_{j,k}^{e_p}$ were obtained by solving the optimization problem in (28), using the solution after 3000 iterations of the genetic algorithm for solving the associated MTSP. The initial position of each agent was set to coincide with the position of the first target in the solution of the MTSP. For simplicity, a constant descent stepsize $\kappa_l = 10^{-4}$ was used in the gradient descent.

The positions of targets were generated randomly from independent uniform distributions ranging from -5 to 5 in both axis. Fig. 1 compares the trajectories of the agents in the first and last step of the gradient descent optimization,

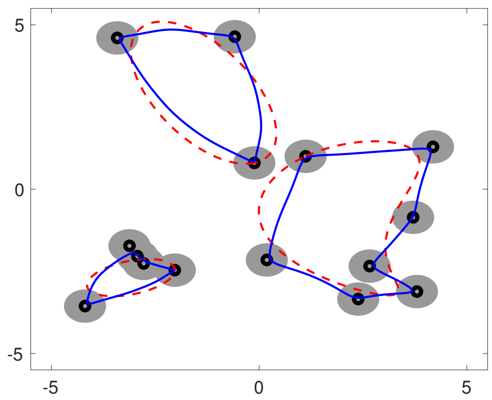

Fig. 1. Trajectories of the targets in the first (red, dashed) and last (blue, solid) iterations of the gradient descent optimization. The target locations are marked in black and the grey shaded represent the regions where the target can be sensed by an agent.

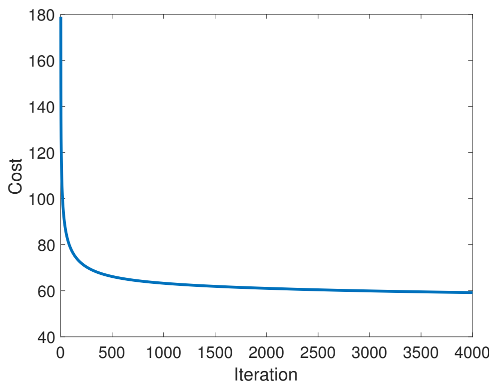

Fig. 2. Evolution of the cost function in the gradient descent optimization.

while Fig. 2 shows the evolution of the cost as a function of the gradient descent step. The results of the optimization show that the solution of (28) led to smooth trajectories that still visited all the targets. The gradient descent changed the trajectory geometry but did not change the visiting order. As can be observed in Fig. 2, the cost has an abrupt reduction in the beginning of the optimization and then the convergence speed decreases significantly. The optimization process leads to very significant reductions of the cost, reducing it to less than one third of its initial value.

6. CONCLUSION AND FUTURE WORK

In this work, we addressed the infinite horizon persistent monitoring problem where the agents and targets lie in a multi-dimensional environment, by constraining their trajectories in each axis to be represented by truncated Fourier series. We discussed conditions for convergence of the covariance matrix as time goes to infinity and described a procedure to compute the gradient of the cost with respect to each agent's trajectory parameters. We linked the problem to the MTSP and used it to provide feasible initial solutions to a gradient descent optimization. Simulation results illustrated the effectiveness of the proposed procedure.

In future work, we plan to study in more depth the effect of the initial trajectories in the optimization. Also, in our results we noted that each target was observed always by a single agent. In this context, we do not need to constrain all the agents to have the same period and plan to allow different agents to have different periods in their trajectories in future works. Finally, we also plan to approach this problem from a decentralized perspective.

REFERENCES

Alam, T., Reis, G.M., Bobadilla, L., and Smith, R.N. (2018). A Data-Driven Deployment Approach for Persistent Monitoring in Aquatic Environments. In *Int. Conference on Robotic Computing*, 147–154. IEEE.

Bektas, T. (2006). The Multiple Traveling Salesman Problem: an Overview of Formulations and Solution Procedures. *Omega*, 34(3), 209 – 219.

Kim, H., Mokdad, L., and Ben-Othman, J. (2018). Designing UAV Surveillance Frameworks for Smart City and Extensive Ocean with Differential Perspectives. *IEEE Communications Magazine*, 56(4), 98–104.

Lan, X. and Schwager, M. (2013). Planning Periodic Persistent Monitoring Trajectories for Sensing Robots in Gaussian Random Fields. In *International Conference on Robotics and Automation*, 2415–2420. IEEE.

Lan, X. and Schwager, M. (2014). A Variational Approach to Trajectory Planning for Persistent Monitoring of Spatiotemporal Fields. In *American Control Conference*, 5627–5632. IEEE.

Lin, X. and Cassandras, C.G. (2014). An Optimal Control Approach to the Multi-Agent Persistent Monitoring Problem in Two-dimensional Spaces. *IEEE Transactions on Automatic Control*, 60(6), 1659–1664.

Nigam, N. (2014). The Multiple Unmanned Air Vehicle Persistent Surveillance Problem: A Review. *Machines*, 2(1), 13–72.

Pinto, S.C., Andersson, S.B., Hendrickx, J.M., and Cassandras, C.G. (2019). Optimal Multi-Agent Persistent Monitoring of the Uncertain State of a Finite Set of Targets. In *Conference on Decision and Control*. IEEE.

Pinto, S.C., Andersson, S.B., Hendrickx, J.M., and Cassandras, C.G. (2020). Optimal Periodic Multi-Agent Persistent Monitoring of a Finite Set of Targets with Uncertain States. In *American Control Conf.* IEEE.

Shen, Z. and Andersson, S.B. (2010). Tracking Nanometer-scale Fluorescent Particles in Two Dimensions with a Confocal Microscope. *IEEE Transactions on Control Systems Technology*, 19(5), 1269–1278.

Tang, L., Liu, J., Rong, A., and Yang, Z. (2000). A Multiple Traveling Salesman Problem Model for Hot Rolling Scheduling in Shanghai Baoshan Iron & Steel Complex. *Europ. Journal of Operational Research*, 124(2), 267 – 282.

Varga, A. (2013). Computational Issues for Linear Periodic Systems: Paradigms, Algorithms, Open Problems. *International Journal of Control*, 86(7), 1227–1239.

Yu, J., Karaman, S., and Rus, D. (2015). Persistent Monitoring of Events With Stochastic Arrivals at Multiple Stations. *IEEE Tran. on Robotics*, 31(3), 521–535.

Zhou, N., Yu, X., Andersson, S.B., and Cassandras, C.G. (2018). Optimal Event-Driven Multiagent Persistent Monitoring of a Finite Set of Data Sources. *IEEE Transactions on Automatic Control*, 63(12), 4204–4217.