Joint Estimation and Control for Multi-Target Passive Monitoring with an Autonomous UAV Agent

Savvas Papaioannou, Christos Laoudias, Panayiotis Kolios,
Theocharis Theocharides, and Christos G. Panayiotou

***Abstract*— This work considers the problem of passively monitoring multiple moving targets with a single Unmanned Aerial Vehicle (UAV) agent equipped with a direction-finding radar (i.e., a bearings-only sensor). This is in general a challenging problem due to the unobservability of the target states, and the highly non-linear measurement process. In addition to these challenges, in this work we also consider: a) environments with multiple obstacles where the targets need to be tracked as they manoeuvre through the obstacles, and b) multiple false-alarm measurements caused by the cluttered environment. To address these challenges we first design a model predictive guidance controller which is used to plan hypothetical target trajectories over a rolling finite planning horizon. We then formulate a joint estimation and control problem where the trajectory of the UAV agent is optimized to achieve optimal multi-target monitoring through stochastic-filtering which accounts for the target model uncertainty, the noisy measurement process, and the false-alarms. Extensive simulation results demonstrate the effectiveness of the proposed approach.**

I. Introduction

Increased mobility, flexibility, and rapid deployment are highly desirable properties in many application domains. Nowadays, Unmanned Aerial Vehicles (UAV) have demonstrated their potential in a wide variety of applications spanning from consumer to disaster management scenarios. These include supporting wireless communication networks [1], delivering goods [2], surveillance and security tasks [3]–[6], search and rescue missions [7]–[9], and situational awareness for first responders in emergency operations [10]. To meet the requirements in the above-mentioned applications, UAV agents may carry on-board various sensors including cameras (e.g., optical, thermal, multi-spectral, LiDAR, etc.), ranging devices that measure the agent-target distance using their radio frequency communication link (e.g., through timing, angle, or signal strength measurements), direction-finders that passively scan the spectrum to detect and resolve the direction of a target transmission through signal processing, radars that provide the direction and/or distance by processing the reflections on objects by either purposefully transmitted signals (i.e., *active* radar) or by ambient signals of opportunity such as FM radio, DVB-T, etc. (i.e., *passive* radar).

This work is co-funded by the Ministry of Defence of the Republic of Cyprus, the European Union's Horizon 2020 research and innovation programme under grant agreement No. 739551 (KIOS CoE), and the Government of the Republic of Cyprus through the Cyprus Deputy Ministry of Research, Innovation and Digital Policy. The authors are with the KIOS Research and Innovation Centre of Excellence (KIOS CoE) and the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, 1678, Cyprus. {papaioannou.savvas, laoudias, pkolios, ttheocharides, christosp}@ucy.ac.cy

Passive systems, e.g., based on radar technology [11], are less power demanding as no signal transmissions are required, thus extending the UAV agent's flight time. While this makes them preferable in many practical scenarios, such systems typically provide bearing-only measurements, i.e., the *angle* between the target-agent line and a reference direction (e.g., magnetic north). In this case several challenges appear due to the highly nonlinear measurement process, and the unobservability of the target states, especially when a single UAV agent in considered; see [12] for an overview of methods for single and multi-sensor bearing-only tracking.

Several existing solutions assume multiple agents to perform the intended task which leads to a standard multi-target tracking problem [13], where unobservability is no longer an issue. We rather focus on applications that necessitate the use of a single agent due to various scenario-specific limitations. For instance, in emergency management scenarios where swift response is critical, the deployment of multiple UAVs may introduce delays in the mission. In addition, using a group of UAVs has inherent computational and communication overhead for multi-agent control and coordination. In other scenarios where the budget is important, using more than one agents may be cost-prohibitive if the payload includes specialized sensing devices or the UAV itself is expensive. Finally, in military and defence applications it may be desirable to deploy a single agent to reduce the risk of being detected by the enemy forces.

Furthermore, in such scenarios the agent is envisioned to operate autonomously in complex settings with multiple physical obstacles. Therefore, the targets need to be monitored accurately as they manoeuvre through the obstacles, while the cluttered environment introduces noisy and *false-alarm* measurements due to multi-path signal propagation and reflections, which makes the joint estimation and control even more challenging. On top of that, *spatial constraints* may exist that do not allow the agent to move freely inside the area of interest, but rather operate within a bounded sub-region. For instance, an agent operating for situational awareness may need to stay behind *no-fly zones*, avoid moving through critical infrastructure facilities, or flying over crowds. Similarly, in a defence scenario the agent needs to fly at a safe distance behind the enemy lines to avoid detection. Under such constraints, it may not be possible to optimally monitor multiple targets; yet, it is important to meet the mission objectives (e.g., deliver adequate target tracking accuracy).

To our knowledge, existing works do not sufficiently address togetner these challenges, i.e., using a single UAV agent to monitor multiple targets that manoeuvre through obstacles

as well as estimating the target states and controlling the agent in the presence of noisy and false-alarm measurements. To this end, our contribution is this work is threefold:

- We introduce a flexible obstacle model and formulate the problem of target manoeuvring through various obstacles as a model predictive control problem. In particular, the proposed mixed integer guidance controller is used to design hypothetical target trajectories, which are then refined through stochastic filtering.
- We formulate a joint estimation and control problem, where the trajectory of the UAV agent is optimized for minimizing the estimation error of the target states in environments with obstacles resulting in multiple noisy and false-alarm measurements.
- We consider spatial constraints in the agent's movements and investigate various pre-computed manoeuvres that can be performed to maximize the monitoring performance, while the agent is moving within specific boundaries. We compare these manoeuvres against the optimal trajectory produced by our controller and report our findings in a realistic simulation environment.

The rest of the paper is structured as follows. Section II overviews related works on passive target monitoring using bearings-only measurements with a focus on the control aspects of autonomous agents. Section III introduces the system model including the target and agent mobility dynamics, the agent sensing model for bearings-only measurements, and the obstacle model. The proposed model predictive control approach for planning realistic trajectories for multiple targets is described in Section IV. Section V presents the proposed joint estimation and control approach for passive multi-target monitoring using a single UAV agent, and experimental results in a simulated environment are reported and discussed in Section VI. Finally, Section VII provides concluding remarks.

II. Related Work

Over the past years a plethora of approaches have been proposed in the literature for the problem of passive target monitoring/tracking. A good starting point on the concept of passive radar, and passive direction-finding systems can be found in [11]. In this section we will focus mainly on the task of passive target monitoring with a single UAV agent which utilizes angle measurements (i.e., bearings), for estimating the target's state. This is considered a challenging problem due to the unobservability of the target states, the stochastic target dynamics, and the highly nonlinear measurement process. A recent survey paper on this topic can be found in [12], while the general problem of monitoring and tracking single and multiple targets is covered in [13].

Regarding the bearings-only passive target monitoring, the authors in [14] provide a thorough analysis of 3 different state-estimation approaches for tracking a single target with a single sensor, and examine the effects of nonlinearity and non-observability on this problem. The authors in [15] design a particle-filter based estimator that uses multiple radar measurements with glint noise in order to passively monitor a single moving target. The work in [16] investigates the concept of fuzzy uncertainty in passive location monitoring systems, and proposes a robust fuzzy extended Kalman filter for monitoring a moving target.

With respect to the agent/observer control aspect which appears in the passive target monitoring applications, the work in [17] proposes a greedy algorithm for optimally choosing the measurement locations in order to localize a stationary target in the least amount of time. Similarly, in [18] the authors first use the geometric dilution of precision to characterize the uncertainty of passive target localization using angle measurements, and then they propose a measurement gathering strategy that jointly minimizes the target localization error of a stationary target, and the time spend in gathering the measurements. Authors in [19] examine the geometry of various navigation schemes in guiding a pursuer agent from a fixed initial position to a final target position using successive bearing measurements, and the work in [20] presents an optimality analysis of various sensor-target localization geometries for passively localizing a static target. Authors in [21] formulate the problem of observer control for bearings-only target localization as an optimal control problem and they propose a trajectory optimization scheme based on the maximization of the determinant of the Fisher information matrix.

The problem of optimally controlling an autonomous agent/observer for accurate passive monitoring of a moving target is further investigated in [22], [23]. Specifically, in [22] various particle-filter estimators are proposed based on the multiple model jump Markov system framework to tackle the various maneuvers of the target, whereas in [23] the observer control is posed as stochastic optimal control problem which aims at maximizing the tracking accuracy. In [24] the authors use the theory of random finite sets [25], and specifically the Bernoulli filter [26] in order to passively track a single moving target, with a control strategy which maximizes the information gain based on the Renyi divergence.

III. System Modelling

A. Target Dynamics

In this work we assume that a known number of M ground targets $\boldsymbol{x}^j,\ j \in [1,..,M]$ operate inside a bounded surveillance environment $\mathcal{E} \subset \mathbb{R}^3$ according to the following stochastic discrete-time dynamical model:

$$\boldsymbol{x}_t^j = A\boldsymbol{x}_{t-1}^j + B\boldsymbol{u}_{t-1}^j + \boldsymbol{\nu}_{t-1},\ j \in [1,..,M] \qquad (1)$$

where $\boldsymbol{x}_t^j = [x_t^j(x), x_t^j(y), x_t^j(z), \dot{x}_t^j(x), \dot{x}_t^j(y), \dot{x}_t^j(z)]^\top \in \mathbb{R}^6$ denotes the state of the j_{th} target at time t, which is composed of the target's position $(x_t^j(x), x_t^j(y), x_t^j(z))$, and velocity $(\dot{x}_t^j(x), \dot{x}_t^j(y), \dot{x}_t^j(z))$ components in 3D Cartesian coordinates. The control input $\boldsymbol{u}_t^j \in \mathbb{R}^3$ denotes the applied control force which allows the target to change its direction and speed, and the term $\boldsymbol{\nu}_t$ is the process noise which models the uncertainty on the target's state, and which is distributed according to a zero mean multi-variate Gaussian distribution with covariance matrix Q, i.e., $\boldsymbol{\nu}_t \sim \mathcal{N}(0, Q)$. Without loss of generality we assume that the process noise profile is the same for all targets. The matrices A and B are defined as:

$$A = \begin{bmatrix} I_{3\times3} & \Delta t \cdot I_{3\times3} \\ 0_{3\times3} & (1-\varepsilon)\cdot I_{3\times3} \end{bmatrix},\ B = \begin{bmatrix} 0_{3\times3} \\ \frac{\Delta t}{m} \cdot I_{3\times3} \end{bmatrix}, \qquad (2)$$

where Δt is the sampling interval, $\varepsilon \in [0,1]$ models the effect of friction on the target's velocity, and m is the target mass which for brevity we assume to be the same for all targets. Moreover, $I_{3\times3}$, and $0_{3\times3}$ denote the identity and zero matrices of size 3-by-3 respectively. Finally, it is assumed that during a reconnaissance phase, the approximate target initial location, and final destination have been acquired and made available to the UAV agent. Therefore, we assume that: a) target's j initial state $\boldsymbol{x}_0^j$ is distributed according to $\boldsymbol{x}_0^j \sim \mathcal{N}(\boldsymbol{\mu}_0^j, \Sigma_0^j)$, and b) the target j is moving towards a goal region on the ground denoted hereafter as $\mathcal{G}^j \subset \mathbb{R}^3$.

B. Agent Dynamics

An autonomous UAV agent/observer, equipped with a passive direction-finding radar which is calibrated for a certain altitude h, is deployed inside the surveillance environment $\mathcal{E}$ with the purpose of monitoring the trajectories of the M targets on the ground. The state of the UAV agent at time-step t i.e., $\boldsymbol{s}_t = [s_t(x), s_t(y), s_t(z)]^\top \in \mathcal{E}$ which is composed of the agent's position in cartesian coordinates, evolves in time according to:

$$\begin{bmatrix} s_t(x) \\ s_t(y) \\ s_t(z) \end{bmatrix} = \begin{bmatrix} s_{t-1}(x) \\ s_{t-1}(y) \\ h \end{bmatrix} + \begin{bmatrix} \lambda\Delta_r \cos(\kappa\Delta_\theta) \\ \lambda\Delta_r \sin(\kappa\Delta_\theta) \\ 0 \end{bmatrix}, \begin{array}{l} \lambda \in [0,..,N_r] \\ \kappa \in [1,..,N_\theta], \end{array} \quad (3)$$

where Δ_r is the radial step size, $\Delta_\theta = 2\pi/N_\theta$, and the parameters (N_r, N_θ) specify the set $\mathcal{S}_t$ containing all possible states $\boldsymbol{s}_t \in \mathcal{S}_t$ which the agent can take at time-step t. Therefore, the set $\mathcal{S}_t$ is given by: $\mathcal{S}_t = \{(s_{t-1}(x) + \lambda\Delta_r\cos(\kappa\Delta_\theta), s_{t-1}(y) + \lambda\Delta_r\sin(\kappa\Delta_\theta), h)\}$, $\forall\lambda \in [0,..,N_r],\ \forall\kappa \in [0,..,N_\theta]$.

C. Agent Sensing Model

As already mentioned the UAV agent is equipped with a passive radar (i.e., a direction-finder) which is used for monitoring nearby ground targets operating inside its sensing range. Specifically, at each time-step t, the UAV agent receives a set of noisy angular measurements (i.e., bearings) from each target j, denoted as $\Phi_t^j = \{\phi_{t,1}^j, .., \phi_{t,|\Phi_t^j|}^j\},\ \phi_{t,i}^j \in (-\pi, \pi]$ rad, where the number of total received measurements, i.e., $|\Phi_t^j|$ ($|.|$ denotes the set cardinality), is random. In particular, it is assumed that due to various obstacles and clutter in the environment the UAV agent receives at each time-step t: a) with a Poisson rate Λ multiple false-alarm measurements (denoted as $\tilde{\phi}_{t,i}^j \in \Phi_t^j$) which are distributed over the measurement space according to the probability distribution $p_{\tilde{\phi}}(\tilde{\phi}_{t,i}^j)$, and b) a single bearing measurement $\hat{\phi}_t^j \in \Phi_t^j$ from the target j with probability p_D. The target generated measurement $\hat{\phi}_t^j$ is related to the target and agent states according to the measurement model $\hat{\phi}_t^j = \ell(\boldsymbol{x}_t^j, \boldsymbol{s}_t) + w_t$, where:

$$\ell(\boldsymbol{x}_t^j, \boldsymbol{s}_t) = \tan^{-1}\left(\frac{x_t^j(x) - s_t(x)}{x_t^j(y) - s_t(y)}\right), \quad (4)$$

and w_t is a Gaussian random variable which models the measurement noise, and which is distributed according to $w_t \sim \mathcal{N}(0, \sigma_\phi^2)$. Without loss of generality we assume that the same target detection probability, false-alarm rate, and the measurement noise applies for all targets, since all targets are sensed by the same radar equipment. In addition, we assume in this work that the targets are sensed by the UAV agent through different communication channels.

D. Obstacle Model

We consider the existence of multiple convex obstacles $\xi_n \in \Xi,\ n \in [1,..,|\Xi|]$ inside the surveillance area $\mathcal{E}$, which are represented in this work as cuboids of arbitrary sizes. In particular, a regular cuboid ξ is a box-shaped object with six rectangular faces, and 8 right angles; therefore, a point $\boldsymbol{p} = [p(x), p(y), p(z)]^\top \in \mathbb{R}^3$ that resides inside the convex-hull of cuboid ξ_n must satisfy the following 6 linear inequalities:

$$\begin{aligned} a_1^n(x)p(x) + a_1^n(y)p(y) + a_1^n(z)p(z) &\leq b_1^n, \\ a_2^n(x)p(x) + a_2^n(y)p(y) + a_2^n(z)p(z) &\leq b_2^n, \\ &\vdots \\ a_6^n(x)p(x) + a_6^n(y)p(y) + a_6^n(z)p(z) &\leq b_6^n, \end{aligned}$$

where $\boldsymbol{a}_i^n = [a_i^n(x), a_i^n(y), a_i^n(z)],\ i \in [1,..,6]$ is the outward unit normal vector on the i_{th} face of the n_{th} cuboid obstacle, and b_i^n is a constant obtained from the dot product between $\boldsymbol{\alpha}_i^n$ and a known point on the plane which contains the i_{th} face. This obstacle model has the flexibility to create 3D objects of varying dimensions, thus adequately representing real-world settings.

Suppose now that $\boldsymbol{p}$ describes the position of a target $\boldsymbol{x}_t^j$ at time-step t. This target, can avoid a potential collision with obstacle ξ_n when the following condition holds:

$$\exists\ i \in [1,..,6] : \text{dot}(\boldsymbol{a}_i^n, \boldsymbol{p}) > b_i^n, \quad (5)$$

where $\text{dot}(a, b)$ is the dot product between vectors a and b. In essence, we require that the target's position resides outside the convex-hull of obstacle $\xi_n,\ n \in [1,..,|\Xi|]$.

IV. Target Trajectory Planning

As we have already mentioned in Sec. III-A, for each target j we consider the availability of the following information: a) its approximate initial location, i.e., we know that the state of target j is initially distributed according to $\boldsymbol{x}_0^j \sim \mathcal{N}(\boldsymbol{\mu}_0^j, \Sigma_0^j)$, and b) its final destination, i.e., we know that target's j objective is to move towards, and reach a specific goal region $\mathcal{G}^j$. Based on these information, and in combination with a known map of the environment (i.e., in this work we use information regarding the position, and dimensions of various obstacles), the objective is to generate a hypothetical trajectory for each target, which can then be passively monitored through sensing, i.e., via the received bearing measurements, as discussed in Sec. III-C.

To do that, the target trajectory hypothesis generation is formulated in this work as a model predictive control problem, where we seek to find target's j hypothetical control inputs $U_t^j = \{\boldsymbol{u}_{t+\tau|t}^j\}, \forall\tau \in [0,..,T-1]$ inside a rolling finite planning horizon of length T time-steps, which enable the guidance of the target to its goal region, subject to kinematic and collision avoidance constraints.

Let us denote the future hypothetical trajectory of target j over a planning horizon of length T time-steps, as $X_t^j = \{\boldsymbol{x}_{t+\tau+1|t}^j\}, \forall\tau \in [0,..,T-1]$, where the notation $\boldsymbol{x}_{t'|t}$ is

used here to denote the predicted target state at time-step t' which was generated at time-step t. Now, based on Eq. (1), observe that the target trajectory X_t^j is in fact a stochastic process, with each future target state $\boldsymbol{x}_{t+\tau+1|t}^j, \forall \tau$, to be distributed according to $\boldsymbol{x}_{t+\tau+1|t}^j \sim \mathcal{N}(\boldsymbol{\mu}_{t+\tau+1|t}^j, \Sigma_{t+\tau+1|t}^j)$, where $\boldsymbol{\mu}_{t+\tau+1|t}^j$, and $\Sigma_{t+\tau+1|t}^j$ are given by:

$$\begin{aligned} \boldsymbol{\mu}_{t+\tau+1|t}^j &= A^{\tau+1}\boldsymbol{\mu}_t^j + \sum_{k=0}^{\tau} A^{\tau-k} B \boldsymbol{u}_{t+k|t}^j, \\ \Sigma_{t+\tau+1|t}^j &= A^{\tau+1}\Sigma_t^j (A^\top)^{\tau+1} + \sum_{k=0}^{\tau} A^{\tau-k} Q (A^\top)^{\tau-k}. \end{aligned} \tag{6}$$

Observe that Eq. (6), has been obtained from the recursive application of Eq. (1). The parameters $\boldsymbol{\mu}_t^j$, and Σ_t^j are respectively the mean, and covariance matrix of the target state at time-step t, which for time-step $t = 0$ are given by $\boldsymbol{\mu}_0^j$ and Σ_0^j respectively. In order to generate the trajectory which guides target j to its goal region $\mathcal{G}^j$, the following cost function is minimized for the control inputs $U_t^j = \{\boldsymbol{u}_{t+\tau|t}^j\}, \forall \tau \in [0, .., T-1]$:

$$\begin{aligned} \arg\min_{U_t} \mathbb{E}[\mathcal{J}^j(X_t^j, U_t^j)] = \|\boldsymbol{\mu}_{t+T|t}^{j,\text{pos}} - \mathcal{G}_o^j\|_2^2 \\ + \varrho \sum_{\tau=1}^{T-1} \|\boldsymbol{u}_{t+\tau|t}^j - \boldsymbol{u}_{t+\tau-1|t}^j\|_2^2, \end{aligned} \tag{7}$$

where $\mathbb{E}$ is the expectation operator, $\|.\|_2$ is the 2-norm, $\boldsymbol{\mu}_{t+T|t}^{j,\text{pos}}$ is the predicted mean of the target's position at the end of the planning horizon computed with Eq. (6), and $\mathcal{G}_o^j$ is the centroid point of the goal region $\mathcal{G}^j$ on the ground. The second term in Eq. (7) is used in order to minimize abrupt changes in the target's direction and speed, and thus produce more realistic smooth trajectories. Therefore, the tuning weight ϱ controls the emphasis given to this secondary objective.

The predicted target trajectory for agent j is then generated with the guidance controller shown in Problem (P1). As shown in Problem (P1), at each time-step t the optimal control inputs $U_t^j = \{\boldsymbol{u}_{t+\tau|t}^j\}, \forall \tau \in [0, .., T-1]$ are computed over a rolling planning horizon of length T time-steps, by solving an open-loop optimal control problem shown, which essentially drives the target to its goal region, while at the same time considering obstacle avoidance constraints. Once the sequence of control inputs is determined, the first control input $\boldsymbol{u}_{t|t}^j$ in the sequence is applied to the target, and the procedure described above is repeated for the next time-step.

Specifically, in Problem (P1) the constraint in Eq. (8b) computes the expected state of target j (i.e., $\boldsymbol{\mu}_{t+\tau+1|t}^j$) inside the planning horizon, which has an associated covariance matrix $\Sigma_{t+\tau+1|t}^j$. Observe that the covariance matrix does not depend on the generated control inputs, and thus can be pre-computed as shown in Eq. (6). The constraints shown in Eq. (8c) state that the target distribution for time-step t is initialized with the posterior distribution of the previous time-step i.e., $\mathcal{N}(\hat{\boldsymbol{\mu}}_{t|t-1}^j, \hat{\Sigma}_{t|t-1}^j)$ which is computed via stochastic filtering as discussed in Sec. V.

The constraints in Eq. (8d)-(8e) enable the generation of collision-free trajectories, by making sure that all targets

Problem (P1) : `Guidance Controller`

$$\min_{U_t^j} \mathbb{E}[\mathcal{J}^j(X_t^j, U_t^j)] \qquad \forall j \quad \text{(8a)}$$

subject to $\tau \in [0, .., T-1]$**:**

$$\boldsymbol{\mu}_{t+\tau+1|t}^j = A^{\tau+1}\boldsymbol{\mu}_t^j + \sum_{k=0}^{\tau} A^{\tau-k} B \boldsymbol{u}_{t+k|t}^j, \qquad \forall \tau, j \quad \text{(8b)}$$

$$\boldsymbol{\mu}_t^j = \hat{\boldsymbol{\mu}}_{t|t-1}^j, \Sigma_t^j = \hat{\Sigma}_{t|t-1}^j, \qquad \forall j \quad \text{(8c)}$$

$$\text{dot}(\boldsymbol{a}_i^n, \boldsymbol{\mu}_{t+\tau+1|t}^{j,\text{pos}}) > b_i^n - H y_{\tau,i}^{j,n}, \qquad \forall \tau, j, n, i \quad \text{(8d)}$$

$$\sum_{i=1}^{6} y_{\tau,i}^{j,n} \leq 5, \qquad \forall \tau, j, n \quad \text{(8e)}$$

$$X_t^j \in \mathcal{X}, U_t^j \in \mathcal{U} \quad \text{(8f)}$$

$$y_{\tau,i}^{j,n} \in \{0, 1\},\ n = [1, .., |\Xi|],\ i = [1, .., 6]$$

avoid collisions with the obstacles in the environment. As a reminder, a collision with some obstacle $\xi_n, n = [1, .., |\Xi|]$, which is represented as a cuboid, is avoided at time-step t when the target state (i.e., its position coordinates) resides outside the convex hull of ξ_n as explained in Sec. III-D. In order to enable this functionality we use the binary variable $y_{\tau,i}^{j,n} \in \{0, 1\}$ which is activated i.e., $y_{\tau,i}^{j,n} = 1$ when the inequality $\text{dot}(\boldsymbol{a}_i^n, \boldsymbol{\mu}_{t+\tau+1|t}^{j,\text{pos}}) > b_i^n$ is not satisfied for target j with position $\boldsymbol{\mu}_{t+\tau+1|t}^{j,\text{pos}}$ at time-step $t+\tau+1|t$, and the i_{th} face of the n_{th} obstacle. In such cases the activation of $y_{\tau,i}^{j,n}$ makes the constraint shown in Eq. (8d) valid with the utilization of a large positive constant $H \in \mathbb{Z}^+$. Now, as discussed in Sec. III-D a collision is avoided at time-step $t+\tau+1|t$ between the target j with state $\boldsymbol{\mu}_{t+\tau+1|t}^{j,\text{pos}}$, and the obstacle ξ_n when $\exists\ i \in [1, .., 6] : \text{dot}(\boldsymbol{a}_i^n, \boldsymbol{\mu}_{t+\tau+1|t}^{j,\text{pos}}) > b_i^n$, which is achieved via the constraint in Eq. (8e) by enforcing the binary variable $y_{\tau,i}^{j,n}$ to take the value of zero for at least one face i.e., $\exists i \in [1, .., 6] : y_{\tau,i}^{j,n} = 0$. Finally, the constraints in Eq. (8f) restrict the target's speed and control inputs within the desired limits. We should point out that Problem (P1) is a mixed integer quadratic program (MIQP), which can be solved efficiently using off-the-shelf optimization tools [27].

V. Autonomous UAV Control for Passive Multi-Target Monitoring

A. Target State Estimation

In order to improve the monitoring performance of the targets while at the same time remaining undetected, the UAV agent utilizes passive sensing. Subsequently, the agent uses its direction-finding radar to obtain at each time-step t, and for each target j a set of bearing measurements Φ_t^j as discussed in Sec. III-C. These measurements are then utilized to improve the estimation of the target states through stochastic filtering [28].

More specifically, for each target j the UAV agent maintains a Bayes filter [29], which uses in order to compute, and recursively update over time its belief (i.e., a probability distribution) on the state of each target. This is shown in Eq. (9) where we denote as $bel(\boldsymbol{x}_{t+1}^j)$ the agent's initial belief on the state of target j for the next time-step $t+1$, and with

$\hat{bel}(\boldsymbol{x}^j_{t+1})$ we denote the posterior belief on the target's state after incorporating the received target measurements.

$$bel(\boldsymbol{x}^j_{t+1}) = \int f(\boldsymbol{x}^j_{t+1}|\boldsymbol{x}^j_t, \boldsymbol{u}^j_t)\hat{bel}(\boldsymbol{x}^j_t)d\boldsymbol{x}^j_t \tag{9a}$$

$$\hat{bel}(\boldsymbol{x}^j_{t+1}) = \eta^{-1} g(\Phi^j_{t+1}|\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1})bel(\boldsymbol{x}^j_{t+1}) \tag{9b}$$

The agent's initial belief $bel(\boldsymbol{x}^j_{t+1})$ is computed through the prediction step shown in Eq. (9a), where $f(\boldsymbol{x}^j_{t+1}|\boldsymbol{x}^j_t, \boldsymbol{u}^j_t)$ is the target state transition density which is governed by the target dynamics in Eq. (1), and therefore is given by $f(\boldsymbol{x}^j_{t+1}|\boldsymbol{x}^j_t, \boldsymbol{u}^j_t) = \mathcal{N}(A\boldsymbol{x}^j_t + B\boldsymbol{u}^j_t, Q)$. On the other hand, $\hat{bel}(\boldsymbol{x}^j_t)$ is the posterior belief of the current time-step i.e., $\hat{bel}(\boldsymbol{x}^j_t) = \mathcal{N}(\hat{\boldsymbol{\mu}}^j_t, \hat{\Sigma}^j_t)$, and thus $bel(\boldsymbol{x}^j_{t+1}) = \mathcal{N}(A\hat{\boldsymbol{\mu}}^j_t + B\boldsymbol{u}^j_t, A\hat{\Sigma}^j_t A^\top + Q)$. Observe that this result is also obtained from Eq. (6) by setting $\tau = 0$, to obtain the one step look-ahead predictive density for the state of target j computed at time-step t i.e., $\boldsymbol{x}^j_{t+1|t} \sim \mathcal{N}(\boldsymbol{\mu}^j_{t+1|t}, \Sigma^j_{t+1|t}) = bel(\boldsymbol{x}^j_{t+1})$.

Subsequently, at time-step $t+1$ the agent with state $\boldsymbol{s}_{t+1}$ receives from each target j the measurement set Φ^j_{t+1}, and updates its belief by computing $\hat{bel}(\boldsymbol{x}_{t+1})$ with the update step shown in Eq.(9b). Specifically, $\eta = \int g(\Phi^j_{t+1}|\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1})bel(\boldsymbol{x}^j_{t+1})d\boldsymbol{x}^j_{t+1}$ is a normalizing constant, and the measurement likelihood function $g(\Phi^j_{t+1}|\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1})$ gives the likelihood that the agent with state $\boldsymbol{s}_{t+1}$ will receive at time-step $t+1$ the measurement set Φ^j_{t+1} from target j with state $\boldsymbol{x}^j_{t+1}$.

To compute this likelihood function, first observe that the measurement set Φ^j_{t+1} contains a random number of random measurements i.e., multiple false-alarm measurements $\tilde{\phi}^j_{t+1,i} \in \Phi^j_{t+1}$ coming with a Poisson rate Λ, which are distributed according to $p_{\tilde{\phi}}(\tilde{\phi}^j_{t+1,i})$, and up to one target measurement $\hat{\phi}^j_{t+1} \in \Phi^j_{t+1}$ which is received with probability p_D, and which is distributed according to $c(\hat{\phi}^j_{t+1}) = \mathcal{N}(\hat{\phi}^j_{t+1}; \ell(\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1}), \sigma^2_\phi)$ as discussed in Sec. III-C. That said, the measurement likelihood function is derived as:

$$g(\Phi^j_{t+1}|\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1}) = (1-p_D)n^j!\Psi(n^j;\Lambda)\prod_{\phi\in\Phi_{t+1}} p_{\tilde{\phi}}(\phi) + (n^j-1)!\Psi(n^j-1;\Lambda)p_D \sum_{\phi\in\Phi_{t+1}} c(\phi)\prod_{\substack{\varphi\in\Phi_{t+1}\\ \varphi\neq\phi}} p_{\tilde{\phi}}(\varphi) \tag{10}$$

where $n^j = |\Phi^j_{t+1}|$ is the total number of received measurements, and $\Psi(n^j;\Lambda)$ is probability mass function of the Poisson distribution with rate parameter Λ, and input argument n^j. Therefore, the first term in Eq. (10) computes the event of receiving at time-step $t+1$ exactly n^j false-alarm measurements (i.e., $\Psi(n^j;\Lambda)\prod_{\phi\in\Phi_{t+1}} p_{\tilde{\phi}}(\phi)$), and no measurement from target j, i.e., the target is not detected with probability $(1-p_D)$; and the factor $n^j!$ accounts for all possible permutations of the measurements in the set. On the other hand, the second term in Eq. (10) accounts for the event where the measurement set Φ^j_{t+1} contains a single target measurement $\hat{\phi}$ with likelihood $p_D c(\hat{\phi})$, and $(n-1)$ false-alarm measurements. Finally, the posterior mean and covariance of the state of target j for time-step $t+1$ is extracted from $\hat{bel}(\boldsymbol{x}^j_{t+1})$ as:

$$\begin{aligned}\hat{\boldsymbol{\mu}}^j_{t+1} &= \int \boldsymbol{x}^j_{t+1}\hat{bel}(\boldsymbol{x}^j_{t+1})d\boldsymbol{x}^j_{t+1},\\ \hat{\Sigma}^j_{t+1} &= \mathbb{E}[(\boldsymbol{x}^j_{t+1} - \hat{\boldsymbol{\mu}}^j_{t+1})(\boldsymbol{x}^j_{t+1} - \hat{\boldsymbol{\mu}}^j_{t+1})^\top].\end{aligned} \tag{11}$$

The posterior distribution computed above i.e., $\mathcal{N}(\hat{\boldsymbol{\mu}}^j_{t+1}, \hat{\Sigma}^j_{t+1})$ is used to initialize the guidance controller for the next time-step, and subsequently, the recursion shown in Eq. (9) is repeated.

B. Monitoring Control

As shown in Eq. (9b) the quality of target monitoring at the next time-step $t+1$ is directly related to the posterior belief $\hat{bel}(\boldsymbol{x}^j_{t+1}), \forall j$, through the measurement likelihood function $g(\Phi^j_{t+1}|\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1})$. Essentially, the measurement set Φ^j_{t+1} obtained at time-step $t+1$, from the agent with state $\boldsymbol{s}_{t+1}$, has the potential to increase the target observability which in turn can be reflected in the reduced uncertainty on the target state captured by the posterior belief $\hat{bel}(\boldsymbol{x}^j_{t+1})$. Moreover, note that the received measurement set Φ^j_{t+1} is linked to the agent state $\boldsymbol{s}_{t+1}$ through Eq. (4) i.e., the agent receives relative angular measurements (bearings) from each target j. Therefore, in order to optimize the monitoring performance at time-step $t+1$ for a particular target j it suffices to select the agent's next state $\hat{\boldsymbol{s}}_{t+1} \in \mathcal{S}_{t+1}$, which will result in the gathering of such measurement set Φ^j_{t+1}, which maximizes the target state observability. This strategy, however, cannot be applied directly since the measurement set Φ^j_{t+1} becomes available only after the agent moves to its new state $\hat{\boldsymbol{s}}_{t+1}$. To overcome this limitation, we follow the procedure described next: For each admissible agent state $\boldsymbol{s}_{t+1} \in \mathcal{S}_{t+1}$, we generate for each target j the hypothetical ideal (i.e., noise-free, no false-alarms) measurement set $Z^j_{t+1} = \{z^j_{t+1}\}$, which would have been received if the agent moves at time-step $t+1$ to state $\boldsymbol{s}_{t+1}$, and target j is distributed according to $bel(\boldsymbol{x}^j_{t+1})$ (computed with Eq. (9a)), with expected position denoted as $\mu^{j,\text{pos}}_{t+1}$. That said, the hypothetical measurement set $Z^j_{t+1} = \{z^j_{t+1}\}$ is generated as:

$$z^j_{t+1} = \tan^{-1}\left(\frac{\mu^{j,\text{pos}}_{t+1}(x) - s_{t+1}(x)}{\mu^{j,\text{pos}}_{t+1}(y) - s_{t+1}(y)}\right). \tag{12}$$

Then, for each pair $(\boldsymbol{s}_{t+1}, z^j_{t+1})_i,\ i \in [1, .., |\mathcal{S}_{t+1}|]$ we compute the pseudo-posterior distribution $\tilde{bel}(\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1}, z^j_{t+1})_i$ according to Eq. (9b), where the measurement likelihood function $g(z^j_{t+1})$ is now given by $g(z^j_{t+1}|\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1}) = \mathcal{N}(z^j_{t+1}; \ell(\boldsymbol{x}^j_{t+1}, \boldsymbol{s}_{t+1}), \sigma^2_\phi)$. Finally, we extract the pseudo-posterior target state mean and covariance $(\tilde{\boldsymbol{\mu}}^j_{t+1}, \tilde{\Sigma}^j_{t+1})_i$ as shown in Eq. (11). The optimal state $\hat{\boldsymbol{s}}_{t+1}$ of the UAV agent for time-step $t+1$ which achieves optimized monitoring performance is then obtained as:

$$\hat{\boldsymbol{s}}_{t+1} = \underset{\boldsymbol{s}\in\mathcal{S}_{t+1}}{\arg\min} \sum_{j=1}^{M} \text{tr}\left(\tilde{\Sigma}^j_{t+1}(\boldsymbol{s})\right), \tag{13}$$

where $\text{tr}(\Sigma)$ is the trace of matrix Σ, and the notation $\tilde{\Sigma}^j_{t+1}(\boldsymbol{s})$ denotes the covariance matrix associated with the pseudo-posterior distribution of the state of the j_{th} target,

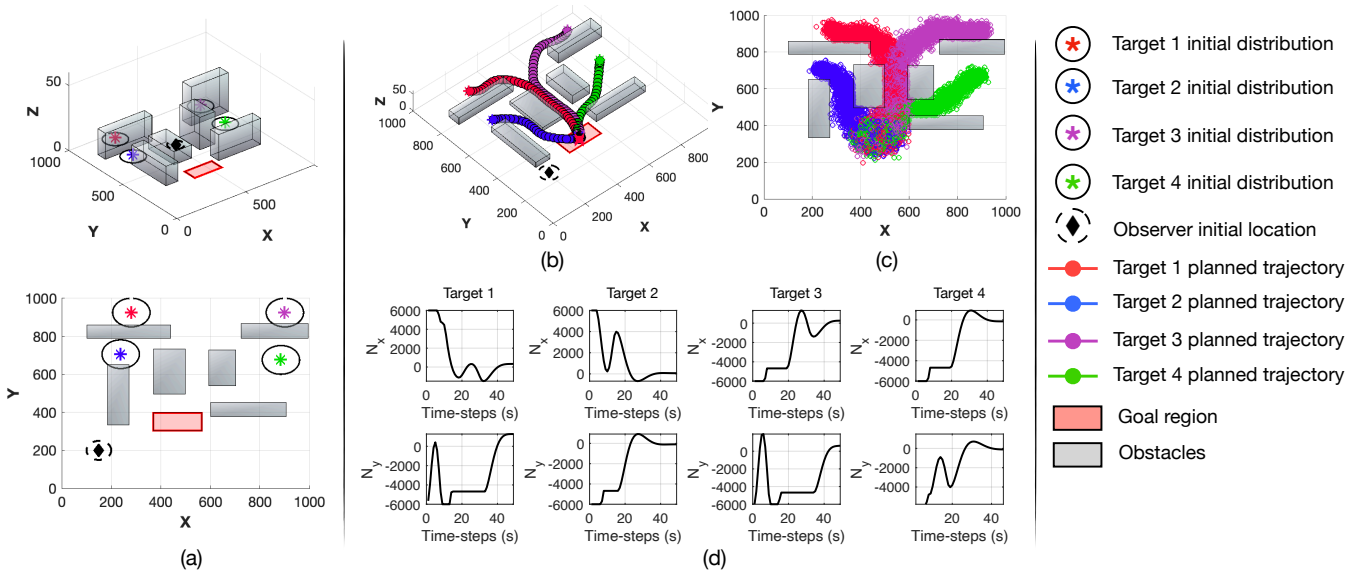

Fig. 1. The figure illustrates the proposed target trajectory hypothesis generation approach, which is realized with the guidance controller shown in Problem (P1), and allows the 4 targets to be guided to the goal region while avoiding collisions with the obstacles in the environment.

which was obtained under the assumption that the agent moved to state $\boldsymbol{s}$ at time-step $t+1$. Once the the optimization problem of Eq. (13) is solved, the agent moves to its new state $\hat{\boldsymbol{s}}_{t+1}$, where the actual target measurements $\Phi^j_{t+1}, \forall j$ are received, and subsequently the posterior distribution on the target states is computed with Eq. (9b) as explained earlier. We should mention here that the UAV obstacle avoidance is handled by making sure that the agent's next state does not reside within the convex-hull of any obstacle i.e., $\boldsymbol{s}_{t+1} \notin \xi_n, \forall n \in [1, .., |\Xi|]$. To achieve this we follow the procedure discussed in Sec. III-D, where all candidate agent states $\boldsymbol{s} \in \mathcal{S}_{t+1}$ which result in a collision are removed from the set $\mathcal{S}_{t+1}$ before solving Eq. (13).

VI. Evaluation

A. Simulation Setup

To evaluate the proposed approach we have used the following simulation setup. The surveillance area $\mathcal{E} \subset \mathbb{R}^3$ is given by a cube with a total volume of 1km^3. The target dynamics are given by Eq. (1) with $\Delta t = 1\text{s}$, $\varepsilon = 0.2$, and $m = 1300\text{kg}$, and are the same for all $M = 4$ targets. The process noise $\boldsymbol{\nu}_t$ is distributed according to $\boldsymbol{\nu}_t \sim \mathcal{N}(0, Q)$, with $Q = \text{diag}([30\ 30\ eps\ 3\ 3\ eps])$, where eps is a very small number i.e., $eps = 1\text{E}-10$, which indicates our knowledge that the targets evolve on the ground plane. Initially it is assumed that the four targets are distributed according to $\boldsymbol{x}^1_0 \sim \mathcal{N}(\boldsymbol{\mu}^1_0, \Sigma_0)$, $\boldsymbol{x}^2_0 \sim \mathcal{N}(\boldsymbol{\mu}^2_0, \Sigma_0)$, $\boldsymbol{x}^3_0 \sim \mathcal{N}(\boldsymbol{\mu}^3_0, \Sigma_0)$, and $\boldsymbol{x}^4_0 \sim \mathcal{N}(\boldsymbol{\mu}^4_0, \Sigma_0)$, where $\boldsymbol{\mu}^1_0 = [281, 925, 0]\text{m}$, $\boldsymbol{\mu}^2_0 = [238, 706, 0]\text{m}$, $\boldsymbol{\mu}^3_0 = [901, 925, 0]\text{m}$, and $\boldsymbol{\mu}^4_0 = [885, 676, 0]\text{m}$. The covariance matrix Σ_0 is given by $\Sigma_0 = \text{diag}([200\ 200\ eps\ 20\ 20\ eps])$ for all targets.

The control input $\boldsymbol{u}_t$ is bounded in the x, and y dimensions inside the interval $[-6000, 6000]\text{N}$ for all targets, and in the z dimension is zero. The targets can reach a ground speed of up to 16m/s. The agent dynamics are given by Eq. (3) with $\Delta_r = 5\text{m}$, $N_\theta = 15$, $N_r = 4$, and $h = 40\text{m}$. In addition, the measurement noise w_t is distributed according to $w_t \sim \mathcal{N}(0, \sigma^2_\phi)$, with $\sigma_\phi = 1\,\text{deg}$, and the target detection probability is set to $p_D = 0.95$. The false-alarms are uniformly distributed inside the measurement space $(-\pi, \pi]$, and arrive with a Poisson rate $\Lambda = 1$. Finally, we note that the stochastic filtering recursion in Eq. (9) has been implemented as a particle filter [30] mainly for handling the non-linear measurement model i.e., Eq. (4), and the guidance problem i.e., Problem (P1), was solved with the Gurobi's MIQP solver [31].

B. Performance Evaluation

In the first part of the evaluation we illustrate in Fig. 1 the target trajectory hypothesis generation as discussed in Sec. IV. Specifically, Fig. 1(a) shows in 3D and top-down views, the initial position of targets $\boldsymbol{x}^1$, $\boldsymbol{x}^2$, $\boldsymbol{x}^3$, and $\boldsymbol{x}^4$ which are marked with a red, blue, purple and green $\star$ respectively. The initial covariance of the target states is drawn as an error-ellipse around the mean of the target location as shown. The obstacles in the environment are shown as gray colored cuboids, and the goal region, which in this scenario is the same for all targets, is shown with the red rectangular region in Fig. 1(a). The UAV agent is initialized in this example at $\boldsymbol{s}_0 = [150, 200, 40]\text{m}$, as shown with the black $\diamond$. Figure 1(b) shows the output of the proposed guidance controller as depicted in Problem (P1), which has generated the hypothetical planned trajectories for the 4 targets over a planning horizon $T = 50$ time-steps. Specifically, the guidance controller which runs on the UAV agent has been initialized at $t = 1$ with the initial probability distributions on the target states $\mathcal{N}(\boldsymbol{\mu}^j_0, \Sigma_0), \forall j$, and using the formulation shown in Eq. (8b)-(8f) computed the control inputs $U^j_1 = \{\boldsymbol{u}^j_{1+\tau|1}\}, \forall \tau = [0, .., T-1]$, for all targets j which minimize the cost function in Eq. (7) with the parameter ρ set to $\rho = 0.01$. As shown in the figure, this optimization allows the targets to avoid the obstacles in the environment, and based on their mobility capabilities to reach the goal region as soon as possible. As

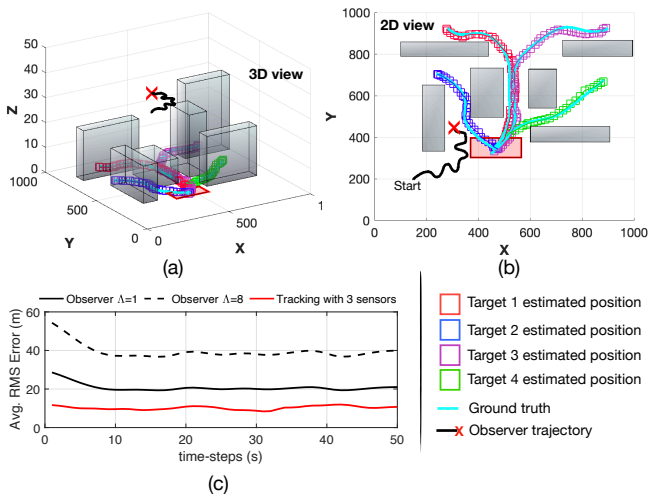

Fig. 2. (a)(b) The figure shows the optimized trajectory of the UAV agent which maximizes the target monitoring performance, (c) the average positional RMSE obtained during 100 Monte-Carlo trials.

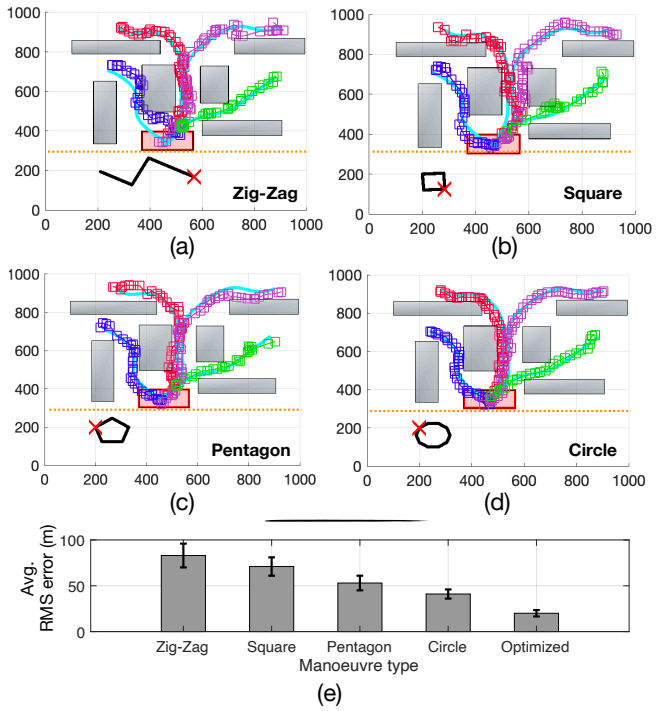

Fig. 3. The figure illustrates the effectiveness of various UAV manoeuvre types on the target monitoring performance.

discussed in Sec. V this optimization is repeated at each time-step t in a rolling horizon fashion after incorporating the target measurements which are used to compute the posterior distribution of the target states. Fig. 1(c) shows the uncertainty on the targets' states over the planning horizon as computed at time-step $t = 1$, with the Eq. (6). In particular, the figure shows the target position as particles sampled from $\mathcal{N}(\boldsymbol{\mu}^j_{1+\tau+1|1}, \Sigma^j_{1+\tau+1|1}), \forall\tau, \forall j$. Finally, Fig. 1(d), shows the optimal control inputs (x and y dimensions) over the planning horizon that guide the targets to the goal region, while producing smooth trajectories without abrupt changes in the speed and direction.

Next we demonstrate the performance of the proposed approach for the task of passively monitoring the four ground targets. In order to evaluate the proposed approach we first assume that the targets move along the noise-free path generated by the guidance controller over the planning horizon of $T = 50$ time-steps, as shown in Fig. 1(b). On the other hand, the UAV agent only observes a noisy version of the target states as indicated by Eq. (1). Therefore, the objective now becomes the selection of the optimal UAV control inputs at each time-step such that the collective uncertainty on the target states is minimized. To achieve this, the UAV agent makes a prediction on the target next states as discussed in Sec. IV, and then uses the received target measurements to update those predictions using the filtering procedure discussed in Sec. V. Figure 2(a)(b) show the result of the optimization problem in Eq. (13) i.e., the UAV's optimal trajectory which maximizes the monitoring performance i.e., minimizes the uncertainty on the target states. Essentially, the UAV agent seeks at each time-step to select its next state from which will obtain the most informative bearing measurement, and which in turn will allow the estimation of the target state. We define the root mean square error (RMSE) on the target position at time-step t as $\epsilon_t = \sqrt{N^{-1}\sum_{n=1}^{N} ||\hat{\boldsymbol{x}}^{\text{pos}}_t(n) - \boldsymbol{x}^{\text{pos}}_t||_2^2}$, where $||.||_2^2$ is the squared 2-norm, N is the number of Monte-Carlo trials, $\hat{\boldsymbol{x}}^{\text{pos}}_t(n)$ denotes the estimated (x, y) target coordinates at time-step t on the n_{th} trial, and $\boldsymbol{x}^{\text{pos}}_t$ is the true target position at the same time-step. Figure 2(c) shows the average positional RMSE obtained for tracking the four targets during the scenario depicted in Fig. 2(a)(b). This scenario was simulated for $N = 100$ trials, where in each trial the UAV initial position was randomly initialized inside the surveillance area. This result is then compared with the positional error obtained from a 3-sensor tracking system. Specifically, we assume that three fixed direction-finding sensors located at $[150, 200]$m, $[800, 200]$m, and $[500, 900]$m, receive three bearing measurements from each target at each time-step, and localize the targets according to the procedure discussed in Sec. V by combining their individual measurement likelihood functions. The measurement noise profile in this case is as discussed in Sec. VI-A, however without false-alarms. As shown in the graph, although the 3-sensor system achieves better results (note that in this case the target state is fully observable), the proposed single sensor system by optimizing the measurement collection process, achieves comparable performance (i.e., solid black line) despite the presence of false-alarms. Finally, the black dotted-line shows what is the achievable performance of the proposed approach in scenarios with higher false-alarm rates i.e., $\Lambda = 8$. Although, the rate of false-alarms degrades the overall monitoring performance as shown in the figure, the targets can still be tracked with a reasonable accuracy, which can be adequate for certain application domains.

Finally, we consider the scenario where the UAV agent is not allowed to move freely inside the surveillance area, rather it can only operate inside a specific sub-region which allows certain mission objectives to be achieved e.g., to remain undetected from enemy radar systems, while passively monitoring the opposing forces activities. In such scenarios the

control optimization procedure discussed in Sec. V i.e., Eq. (13), cannot be applied directly due to the aforementioned mission constraints. For this reason, we investigate various types of pre-computed manoeuvres which the UAV agent can utilize in order to maximize the monitoring performance, while operating within the allowed area boundaries. This is shown in Fig. 3(a)-(d), where we investigate the target monitoring performance during 4 different types of manoeuvres i.e., Zig-Zag, Square, Pentagon, and Circular as shown in Fig. 3(a), Fig. 3(b), Fig. 3(c), and Fig. 3(d), respectively. In this scenario we assume that the UAV agent is only allowed to operate within the sub-region under the orange dotted line. The average positional RMSE at each time-step is shown in Fig. 3(e) for the four manoeuvres types. This result, was obtained with a Monte-Carlo simulation, where the UAV initial position is randomly sampled inside the allowed operating area 100 times, and then the designated manoeuvre type is executed. The results indicate that the circular manoeuvre outperforms the rest, as the higher degree of manoeuvrability that it offers allows the UAV agent to collect more informative bearing measurements from the targets. As a comparison, we also show in Fig. 3(e), labelled "Optimized", the achievable performance obtained when the agent is allowed to move freely inside the surveillance environment guided by the control optimization approach shown in Eq. (13).

VII. Conclusion

In this work we propose a joint estimation and control approach for passively monitoring multiple targets of interest in challenging conditions (i.e., environments with obstacles, and false-alarm measurements) with a single UAV agent equipped with a direction-finding sensor. Model predictive control is used for generating hypothetical target trajectories inside a rolling finite planning horizon, which are then refined through stochastic filtering. In particular, we show how the agent's path can be optimized in order to minimize the collective uncertainty over the target states. Finally, we analyze the performance of various types of pre-computed manoeuvres, which can be utilized in missions with constraints on the agent's movements.

References

[1] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[2] K. Peng, J. Du, F. Lu, Q. Sun, Y. Dong, P. Zhou, and M. Hu, "A hybrid genetic algorithm on routing and scheduling for vehicle-assisted multi-drone parcel delivery," *IEEE Access*, vol. 7, pp. 49 191–49 200, 2019.

[3] H. Huang and A. V. Savkin, "An algorithm of reactive collision free 3-d deployment of networked unmanned aerial vehicles for surveillance and monitoring," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 132–140, 2020.

[4] S. Papaioannou, P. Kolios, C. G. Panayiotou, and M. M. Polycarpou, "Cooperative simultaneous tracking and jamming for disabling a rogue drone," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 7919–7926.

[5] S. Papaioannou, P. Kolios, and G. Ellinas, "Downing a rogue drone with a team of aerial radio signal jammers," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 2555–2562.

[6] ——, "Distributed estimation and control for jamming an aerial target with multiple agents," *IEEE Transactions on Mobile Computing*, 2022, doi:10.1109/TMC.2022.3207589.

[7] M. Erdelj, E. Natalizio, K. R. Chowdhury, and I. F. Akyildiz, "Help from the sky: Leveraging uavs for disaster management," *IEEE Pervasive Computing*, vol. 16, no. 1, pp. 24–32, 2017.

[8] S. Papaioannou, P. Kolios, T. Theocharides, C. G. Panayiotou, and M. M. Polycarpou, "Towards automated 3D search planning for emergency response missions," *Journal of Intelligent & Robotic Systems*, vol. 103, no. 1, p. 2, 2021.

[9] ——, "3D trajectory planning for UAV-based search missions: An integrated assessment and search planning approach," in *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2021, pp. 517–526.

[10] J. Moon, S. Papaioannou, C. Laoudias, P. Kolios, and S. Kim, "Deep reinforcement learning multi-uav trajectory control for target tracking," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15 441–15 455, 2021.

[11] H. D. Griffiths and C. J. Baker, *An introduction to passive radar*. Artech House, 2022.

[12] B. Sindhu, J. Valarmathi, and S. Christopher, "Bearing only target tracking using single and multisensor: A review." *Journal of Engineering Science & Technology Review*, vol. 12, no. 1, 2019.

[13] B.-n. Vo, M. Mallick, Y. Bar-Shalom, S. Coraluppi, R. Osborne, R. Mahler, and B.-t. Vo, "Multitarget tracking," *Wiley encyclopedia of electrical and electronics engineering*, no. 2015, 2015.

[14] B. La Scala and M. Morelande, "An analysis of the single sensor bearings-only tracking problem," in *2008 11th International Conference on Information Fusion*. IEEE, 2008, pp. 1–6.

[15] H.-W. Li and J. Wang, "Particle filter for manoeuvring target tracking via passive radar measurements with glint noise," *IET Radar, Sonar & Navigation*, vol. 6, no. 3, pp. 180–189, 2012.

[16] X. Yang, G. Liu, J. Guo, H. Wang, and B. He, "The robust passive location algorithm for maneuvering target tracking," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[17] J. Vander Hook, P. Tokekar, and V. Isler, "Cautious greedy strategy for bearing-only active localization: Analysis and field experiments," *Journal of Field Robotics*, vol. 31, no. 2, pp. 296–318, 2014.

[18] H. Bayram, J. Vander Hook, and V. Isler, "Gathering bearing data for target localization," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 369–374, 2016.

[19] A. N. Bishop and P. N. Pathirana, "Optimal trajectories for homing navigation with bearing measurements," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 12 117–12 123, 2008.

[20] A. N. Bishop, B. Fidan, B. D. Anderson, K. Doğançay, and P. N. Pathirana, "Optimality analysis of sensor-target localization geometries," *Automatica*, vol. 46, no. 3, pp. 479–492, 2010.

[21] Y. Oshman and P. Davidson, "Optimization of observer trajectories for bearings-only target localization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 3, pp. 892–902, 1999.

[22] M. S. Arulampalam, B. Ristic, N. Gordon, and T. Mansell, "Bearings-only tracking of manoeuvring targets using particle filters," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, pp. 1–15, 2004.

[23] P. Skoglar, U. Orguner, and F. Gustafsson, "On information measures based on particle mixture for optimal bearings-only tracking," in *2009 IEEE Aerospace conference*. IEEE, 2009, pp. 1–14.

[24] B. Ristic and S. Arulampalam, "Bernoulli particle filter with observer control for bearings-only tracking in clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2405–2415, 2012.

[25] R. Mahler, "Statistics 102 for multisource-multitarget detection and tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 376–389, 2013.

[26] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina, "A tutorial on bernoulli filters: Theory, implementation and applications," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3406–3430, 2013.

[27] L. M. Hvattum, A. Løkketangen, and F. Glover, "Comparisons of commercial mip solvers and an adaptive memory (tabu search) procedure for a class of 0–1 integer programming problems," *Algorithmic Operations Research*, vol. 7, no. 1, pp. 13–20, 2012.

[28] A. Bain and D. Crisan, *Fundamentals of stochastic filtering*. Springer, 2009, vol. 3.

[29] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge university press, 2013, no. 3.

[30] F. Gustafsson, "Particle filter theory and practice with positioning applications," *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, no. 7, pp. 53–82, 2010.

[31] T. Achterberg, "What's new in Gurobi 9.0," *Webinar Talk url: https://www.gurobi.com/wp-content/uploads/2019/12/Gurobi-90-Overview-Webinar-Slides-1.pdf*, 2019.