Location Secrecy Enhancement in Adversarial Networks via Trajectory Control

Mohammad Javad Khojasteh, *Member, IEEE*, Augustin A. Saucan, *Member, IEEE*,
Zhenyu Liu, *Graduate Student Member, IEEE*, Andrea Conti, *Fellow, IEEE*, Moe Z. Win, *Fellow, IEEE*

***Abstract*— In networked environments, adversaries may exploit location information to perform carefully crafted attacks on cyber-physical systems. To prevent such security breaches, this letter develops a network localization and navigation (NLN) paradigm that accounts for network secrecy in the control of mobile agents. We consider a scenario in which a mobile agent is tasked with maneuvering through an adversarial network, based on a nominal control policy, and we aim to reduce the ability of the adversarial network to infer the mobile agent's position. Specifically, the Fisher information of the agent's position obtained by the adversarial network is adopted as a secrecy metric. We propose a new control policy that results from an optimization problem and achieves a compromise between maximizing location secrecy and minimizing the deviation from the nominal control policy. Results show that the proposed optimization-based control policy significantly improves the secrecy of the mobile agent.**

***Index Terms*— Localization, Fisher information, sensor network, secrecy, control.**

I. Introduction

LOCATION-AWARENESS plays an essential role in a myriad of network applications, including Internet-of-Things (IoT) [1], Internet-of-Battlefield-Things (IoBT) [2], and autonomy [3]. The distributed nature of these systems can be a source of vulnerability, and adversaries may exploit location information of targets to perform crafted attacks on these systems [4]–[7]. Therefore, studying and preventing such security breaches by enhancing agents' location secrecy is vital. Location secrecy is also critical in applications such as adversarial search-and-rescue missions and pursuit-evasion in mobile robotic settings [8], where an agent tries to actively avoid being detected. In this context, there is an increasing interest in active defense where the target performs evasive maneuvers to reduce the ability of the adversarial network to infer its position.

The fundamental research described in this letter was supported in part by the Office of Naval Research under Grants N00014-16-1-2141 and N62909-22-1-2009, in part by the European Union's Horizon 2020 Research and Innovation Programme under Grant 871249, and in part by the Army Research Office through the MIT Institute for Soldier Nanotechnologies, under Contract W911NF-13-D-0001. *Corresponding author: Moe Z. Win* (e-mail: moewin@mit.edu).

M. J. Khojasteh, A. A. Saucan, and Z. Liu are with the Wireless Information and Network Sciences Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

A. Conti is with the Department of Engineering and CNIT, University of Ferrara, Via Saragat 1, 44122 Ferrara, Italy.

M. Z. Win is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

The scope of this letter lies at the intersection of network localization and navigation (NLN) [9], control of mobile robotic systems [10], and wireless network secrecy [11]. We consider a scenario in which a mobile agent is tasked with maneuvering through an adversarial network. The adversarial sensing nodes attempt to infer the location of the mobile agent (robot) through measurements, e.g., range-only measurements. In particular, we consider a static network of adversarial sensing nodes and a mobile robot described with known system dynamics.

The fundamental limits of NLN provide performance benchmarks and assist the design of the signal processing chain and the architecture of the network [12]–[14]. In particular, the work in [12] determines the Fisher information matrix (FIM) of agents' positions with different types of measurements without considering the control of agents. Related works on the design of control policies or optimal way-points for mobile sensing nodes can be found in [15], [16]. We aim to design a control policy for the mobile agent to enhance its location secrecy.

The literature has considered entropy regularized reinforcement learning (RL), which finds policies with the highest entropy, to improve exploration strategies utilized in RL [17]. In the context of networked control systems, metrics based on Fisher information and mutual information were used in [18], [19] to improve the secrecy of the system dynamics parameters by hampering the adversary's learning process. A two-player game modeling the adversarial information acquisition in robotics was investigated in [20], where each robot aims to design a control policy that maximizes information gain while minimizing information loss with respect to the adversary robot. In particular, the setup of [20] suffers from curse of dimensionality and can be computationally expensive. Secure linear-quadratic optimal control for systems in the absence of disturbances was considered in [21], where the controller aims to drive the system to a set of final states while deteriorating the adversary's ability to estimate the final state.

This work considers a hierarchical control design [22], [23] for the agent in which a nominal control policy is given for a robot in accordance with a specific task. This nominal control policy can be calculated by a high-level planner, such as model predictive control (MPC) [24], RL [25], or motion planning via a neural control contraction metric [26]. This high-level planner is usually computationally expensive and operates at a low frequency. We aim to modify the nominal control policy, during online operation, in a minimum invasive manner to improve the location secrecy of the agent. This requires identifying and solving a tractable optimization problem to

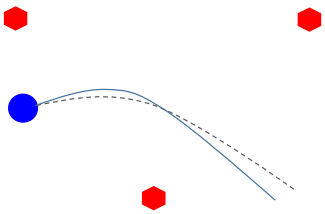


Fig. 1. The agent (blue dot) moves in a network consisting of static adversarial sensing nodes (red hexagons). The trajectory of the agent with nominal control policy is represented by the dashed black curve. We modify the control policy of the mobile agent to reduce the Fisher information obtained by the adversarial network, where the agent follows the path showed by the solid blue curve.

reduce the ability of the adversarial network in inferring the agent's position as it maneuvers through the network.

This letter considers a localization network consisting of a mobile agent and static adversarial sensing nodes. The agent knows its own dynamics as well as the number and the positions of the sensing nodes. The goal is to design policies for controlling the movement of the agent in order to improve its secrecy in a minimally invasive manner during online operation. To achieve this goal, we design an optimization-based controller that reduces the Fisher information obtained by the adversarial network. The key contributions of this letter are as follows:

- we characterize the impact of control policies on the Fisher information of the position for a mobile agent obtained by a network of adversarial nodes;
- we design optimization-based control policies, which can be solved efficiently, for improving agent secrecy with minimal modifications to its nominal control policy; and
- we define a secrecy parameter and quantify the trade-off between the adversary's Fisher information and the change in control policy.

Notations: Random variables are displayed in sans serif, upright fonts; their realizations in serif, italic fonts. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. For a square matrix $\boldsymbol{A}$, notations $\boldsymbol{A}^{-1}$ and $\det(\boldsymbol{A})$ represent its inverse and determinant, respectively. $\mathrm{diag}(x_1, x_2, \ldots, x_n)$ represents an n-by-n diagonal matrix with x_i being the entry on its ith row and column. Expression $\boldsymbol{A} \succcurlyeq \boldsymbol{B}$ represents that matrix $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinte. The zero and identity matrices are denoted by $\mathbf{0}$ and $\boldsymbol{I}$, respectively. The notation $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2.

II. Problem Formulation

Consider a 2D localization network comprised of N_s adversarial sensing nodes that are static. Let set $\mathcal{N}_\mathrm{s} = \{1, 2, \ldots, N_\mathrm{s}\}$ represent the indices of the adversarial sensing nodes with the position of node j represented by $\boldsymbol{q}^{(j)} = [q_\mathrm{x}^{(j)}, q_\mathrm{y}^{(j)}]^\top$. As depicted in Figure 1, the objective of the adversarial sensing nodes is to estimate the position $\mathbf{p}_t = [\mathsf{x}_t, \mathsf{y}_t]^\top$ of an agent of interest at time t, which is a robot, based on the following range measurements

$$\mathbf{r}_t = \left[d_t^{(1)}, d_t^{(2)}, \ldots, d_t^{(N_\mathrm{s})}\right]^\top + \mathbf{n}_t \tag{1}$$

where $d_t^{(j)} = \|\mathbf{p}_t - \boldsymbol{q}^{(j)}\|$ represents the true distance between the agent and the j-th sensing node, and $\mathbf{n}_t$ represents the measurement noise. In particular, the discrete-time process $\{\mathbf{n}_t\}_{t\geqslant 1}$ consists of independent, identically distributed (IID) zero-mean Gaussian random vectors with covariance matrix $\mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_{N_\mathrm{s}}^2)$. The number and the positions of the sensing nodes in the adversarial network are known to the mobile agent.[1] The agent tries to avoid being detected by the adversarial sensing nodes, and estimates its location via a separate *authenticated and secure* network of sensing nodes or using prior knowledge and intra-node measurements obtained by inertial measurement units (IMUs).

The agent dynamics are modeled as

$$\begin{bmatrix} \mathbf{p}_{t+1} \\ \mathbf{v}_{t+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \Delta_\mathrm{s}\boldsymbol{I} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \mathbf{p}_t \\ \mathbf{v}_t \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \boldsymbol{I} \end{bmatrix} \mathbf{u}_t + \mathbf{w}_t \tag{2}$$

where Δ_s is the sampling period, which is set to be 1 s for ease of exposition, and meter is used as the measurement unit for the position. Here, $\mathbf{u}_t$ is the control input applied to the agent at time t, $\mathbf{v}_t = [\mathsf{v}_{\mathrm{x},t}, \mathsf{v}_{\mathrm{y},t}]^\top$ represents the velocity of the agent at time t, and $\{\mathbf{w}_t\}_{t\geqslant 1}$ represents the dynamic disturbances modeled as IID zero-mean Gaussian random vectors with covariance matrix $\boldsymbol{Q}$. Also, for $k \in \{1, 2, 3, 4\}$, $\mathsf{w}_{k,t}$ represents the kth entry of $\mathbf{w}_t$. As discussed in [28], such dynamics are capable of modeling *general robot kinematics*.[2]

A. Fisher Information Inequality

Let $\hat{\mathbf{p}}$ be any unbiased estimator of the unknown deterministic position $\boldsymbol{p}$. Under regularity assumptions [29, Ch.3], the non-Bayesian Fisher information inequality (FII) [12] states that

$$\mathbb{E}\{(\boldsymbol{p} - \hat{\mathbf{p}})(\boldsymbol{p} - \hat{\mathbf{p}})^\top\} \succcurlyeq \boldsymbol{J}^{-1}(\boldsymbol{p})$$

where the FIM $\boldsymbol{J}(\boldsymbol{p})$ is given by

$$\boldsymbol{J}(\boldsymbol{p}) \triangleq \sum_{j\in\mathcal{N}_\mathrm{s}} \lambda^{(j)} \begin{bmatrix} \cos^2(\phi^{(j)}) & \cos(\phi^{(j)})\sin(\phi^{(j)}) \\ \cos(\phi^{(j)})\sin(\phi^{(j)}) & \sin^2(\phi^{(j)}) \end{bmatrix}.$$

Here, $\lambda^{(j)}$ represents the range information intensity (RII) between the agent and the j-th adversarial sensing node, and $\phi^{(j)}$ is the relative angle between the agent and the same node.

Several criteria for assessing the performance of the localization network can be written as a function of the eigenvalues of $\boldsymbol{J}^{-1}(\boldsymbol{p})$ [30, p. 387]. In particular, the D-optimality criterion corresponds to the minimization of $\det(\boldsymbol{J}^{-1}(\boldsymbol{p}))$, the A-optimality criterion corresponds to the minimization of $\mathrm{tr}(\boldsymbol{J}^{-1}(\boldsymbol{p}))$, while the E-optimality criterion corresponds to the minimization of the largest eigenvalue of $\boldsymbol{J}^{-1}(\boldsymbol{p})$. Specifically, the mean-square error (MSE) of the position estimator is lower bounded by $\mathrm{tr}(\boldsymbol{J}^{-1}(\boldsymbol{p}))$ [13]; hence, the

[1] By relying on robust optimization techniques [13], [27], extensions of the proposed method are possible in the case when the agent has only soft-information on the localization of the adversarial nodes (e.g., their belonging to a finite union of compact sets) as opposed to exact knowledge of their positions.

[2] For ease of exposition, we consider the linear dynamics of (2). Linearization can easily be employed to extend our proposed method to more general nonlinear dynamics.

$$\mathscr{P}_1: \underset{\boldsymbol{\delta}_{t-1}}{\text{minimize}} \ \big(\delta_{\mathrm{x},t-1}(\hat{\boldsymbol{\rho}}_{t-1})\big)^2 + \big(\delta_{\mathrm{y},t-1}(\hat{\boldsymbol{\rho}}_{t-1})\big)^2 + \alpha\, \mathbb{E}\big\{ \det\big(\boldsymbol{J}(\mathbf{p}_{t+1})\big) \,\big|\, \hat{\mathbf{p}}_{t-1} = \tilde{\boldsymbol{p}}_{t-1}, \hat{\mathbf{v}}_{t-1} = \tilde{\boldsymbol{v}}_{t-1} \big\} \tag{5}$$

A-optimality criterion is related to the MSE of the position estimator.

Let λ be the maximum of $\{\lambda^{(j)}, j \in \mathcal{N}_\mathrm{s}\}$. To consider a more favorable scenario for the adversarial sensing nodes (and a less favorable scenario for the mobile agent), according to the A-optimiality criterion, we assume that RIIs between the agent and all sensing nodes are equal to λ, that is $\lambda^{(j)} = \lambda$ for all $j \in \mathcal{N}_\mathrm{s}$. In this case, the D-optimality, A-optimality, and E-optimality are equivalent [16]. Moreover, from Theorem 1 of [31] (see also [15]), it follows

$$\det\big(\boldsymbol{J}(\boldsymbol{p})\big) = \lambda^2 \sum_{(i,j)\in\mathcal{S}} \sin^2\big(\phi^{(j)} - \phi^{(i)}\big) \tag{3}$$

where $\mathcal{S} \triangleq \big\{(i,j) : i,j \in \mathcal{N}_\mathrm{s},\, j > i\big\}$. The above results are obtained in a non-Bayesian setting. In Bayesian settings, an average information metric can be obtained by averaging the quantity in (3) across the distribution of the possible agent's location. Other metrics can be constructed directly from the agent's posterior distribution, such as posterior covariance, or posterior entropy (cf. [32]). In the following, $\mathbb{E}\big\{\det\big(\boldsymbol{J}(\mathbf{p})\big)\big\}$ is adopted as the performance metric for localizing the agent.

Given the network of adversarial sensing nodes, the agent modifies its trajectory in a minimally invasive manner to enhance its location secrecy. Let $\hat{\mathbf{p}}_t$ and $\hat{\mathbf{v}}_t$ denote the estimators of the agent position $\mathbf{p}_t$ and its velocity $\mathbf{v}_t$ at time t, respectively, constructed by the adversarial network. That is, the estimators $\hat{\mathbf{p}}_t$ and $\hat{\mathbf{v}}_t$ are constructed using the range measurements (1), up to time t, by the adversarial sensing nodes. We also let $\hat{\boldsymbol{\rho}}_t$ be the agent's estimator of its own position. The estimator $\hat{\boldsymbol{\rho}}_t$ may be constructed using prior knowledge and intra-node measurements obtained by IMUs, or it could be provided to the agent by a separate authenticated and secure network of sensing nodes.

We consider a hierarchical control design as follows. The agent is assumed to follow a *nominal control policy* $\boldsymbol{\pi}(\hat{\boldsymbol{\rho}}_t)$. The policy $\boldsymbol{\pi}(\hat{\boldsymbol{\rho}}_t)$ can be provided by a high-level planner which is usually computationally expensive and operates at a low frequency. The goal is to minimally modify the nominal policy for improving the secrecy of the agent during online operation. At time t, the nominal control policy $\boldsymbol{\pi}(\hat{\boldsymbol{\rho}}_t) = \big(\pi_\mathrm{x}(\hat{\boldsymbol{\rho}}_t), \pi_\mathrm{y}(\hat{\boldsymbol{\rho}}_t)\big)$ is modified as follows:

$$\begin{aligned} \mathsf{u}_{\mathrm{x},t} &= \pi_\mathrm{x}(\hat{\boldsymbol{\rho}}_t) + \delta_{\mathrm{x},t}(\hat{\boldsymbol{\rho}}_t) \\ \mathsf{u}_{\mathrm{y},t} &= \pi_\mathrm{y}(\hat{\boldsymbol{\rho}}_t) + \delta_{\mathrm{y},t}(\hat{\boldsymbol{\rho}}_t) \end{aligned} \tag{4}$$

where $\delta_{\mathrm{x},t}(\hat{\boldsymbol{\rho}}_t)$ and $\delta_{\mathrm{y},t}(\hat{\boldsymbol{\rho}}_t)$ represent the change of the control at time t along the x-axis and the y-axis, respectively. The compact notation is introduced $\boldsymbol{\delta}_t(\hat{\boldsymbol{\rho}}_t) \triangleq \big(\delta_{\mathrm{x},t}(\hat{\boldsymbol{\rho}}_t), \delta_{\mathrm{y},t}(\hat{\boldsymbol{\rho}}_t)\big)$. As the agent aims to reduce the localization accuracy achieved by the adversarial network, it will maximize the uncertainty of the location perceived by the adversarial network by minimizing $\mathbb{E}\big\{\det\big(\boldsymbol{J}(\mathbf{p})\big)\big\}$. In general, $\hat{\mathbf{p}}_{t-1}$ and $\hat{\mathbf{v}}_{t-1}$ are not available to the mobile agent. In this case, we assume the agent to be capable of simulating or approximating the inference process of the adversarial network and to quantify the localization information obtained by it. Let $\tilde{\boldsymbol{p}}_{t-1}$ and $\tilde{\boldsymbol{v}}_{t-1}$ denote the simulated values by the mobile agent when mimicking the inference process of the adversarial network.[3] To increase the location secrecy, the optimization-based refinement for the control policy is formulated in the optimization Problem $\mathscr{P}_1$ at the top of this page. In other words, the Fisher information obtained by the adversarial network is minimized in a regularized fashion ensuring that the squared L_2-norm of the change $\boldsymbol{\delta}_t(\hat{\boldsymbol{\rho}}_t)$ is minimal, where $\boldsymbol{\delta}_t(\hat{\boldsymbol{\rho}}_t)$ is defined after (4). The *secrecy parameter* $\alpha \geqslant 0$ adjusts the trade-off between the Fisher information obtained by the adversarial network of sensing nodes and the deviation from the nominal control policy. For more discussions on the effect of the secrecy parameter α on our design see Section IV. Based on the control policy obtained by $\mathscr{P}_1$, the agent proceeds to estimate its own position and to alter its trajectory in order to reduce the localization information obtained by the adversarial network. In the next section, we present the proposed online controller that addresses Problem $\mathscr{P}_1$.

III. Optimization-Based Controller for Fisher Information Regularized Navigation

Using (2) and (4), at time $t+1$, the j-th relative angle is random and it can be calculated as

$$\varphi_{t+1}^{(j)} = \tan^{-1}\left(\frac{\mathsf{c}_{\mathrm{y},t}^{(j)} - \delta_{\mathrm{y},t-1}(\hat{\boldsymbol{\rho}}_{t-1})}{\mathsf{c}_{\mathrm{x},t}^{(j)} - \delta_{\mathrm{x},t-1}(\hat{\boldsymbol{\rho}}_{t-1})}\right) \tag{6}$$

where

$$\mathsf{c}_{\mathrm{x},t}^{(j)} \triangleq q_\mathrm{x}^{(j)} - \mathsf{x}_t - \mathsf{v}_{\mathrm{x},t-1} - \pi_\mathrm{x}(\hat{\boldsymbol{\rho}}_{t-1}) - \mathsf{w}_{3,t-1} - \mathsf{w}_{1,t} \tag{7a}$$

$$\mathsf{c}_{\mathrm{y},t}^{(j)} \triangleq q_\mathrm{y}^{(j)} - \mathsf{y}_t - \mathsf{v}_{\mathrm{y},t-1} - \pi_\mathrm{y}(\hat{\boldsymbol{\rho}}_{t-1}) - \mathsf{w}_{4,t-1} - \mathsf{w}_{2,t}\,. \tag{7b}$$

Here, $\mathsf{c}_{\mathrm{x},t}^{(j)}$ and $\mathsf{c}_{\mathrm{y},t}^{(j)}$ are the distances between the agent and the j-th sensing node at time t along the x-axis and the y-axis, respectively, if the system follows the nominal control policy $\boldsymbol{\pi}$. Using (3) and (6), the determinant of the FIM as a function of the control input is found to be

$$\det\big(\boldsymbol{J}(\mathbf{p}_{t+1})\big) = \lambda^2 \sum_{(i,j)\in\mathcal{S}} h_{\left(\mathsf{c}_{\mathrm{y},t}^{(j)}, \mathsf{c}_{\mathrm{x},t}^{(j)}, \mathsf{c}_{\mathrm{y},t}^{(i)}, \mathsf{c}_{\mathrm{x},t}^{(i)}\right)}\big(\boldsymbol{\delta}_{t-1}(\hat{\boldsymbol{\rho}}_{t-1})\big) \tag{8}$$

where

$$\begin{aligned} &h_{(a,b,c,d)}(r_1, r_2) \\ &\quad \triangleq \sin^2\Big(\tan^{-1}\Big(\frac{a - r_2}{b - r_1}\Big) - \tan^{-1}\Big(\frac{c - r_2}{d - r_1}\Big)\Big). \end{aligned} \tag{9}$$

[3] In Section IV, a particle filter [33] is employed by the network of adversarial sensing nodes to estimate the agent's position and velocity; and the agent is capable of emulating these calculations.

$$\breve{h}_{\left(\mathsf{c}_{\mathrm{y},t}^{(j)},\mathsf{c}_{\mathrm{x},t}^{(j)},\mathsf{c}_{\mathrm{y},t}^{(i)},\mathsf{c}_{\mathrm{x},t}^{(i)}\right)}(0,0) \triangleq h_{\left(\mathsf{c}_{\mathrm{y},t}^{(j)},\mathsf{c}_{\mathrm{x},t}^{(j)},\mathsf{c}_{\mathrm{y},t}^{(i)},\mathsf{c}_{\mathrm{x},t}^{(i)}\right)}(0,0) + \sin\left(2\left[\tan^{-1}\left(\frac{\mathsf{c}_{\mathrm{y},t}^{(j)}}{\mathsf{c}_{\mathrm{x},t}^{(j)}}\right) - \tan^{-1}\left(\frac{\mathsf{c}_{\mathrm{y},t}^{(i)}}{\mathsf{c}_{\mathrm{x},t}^{(i)}}\right)\right]\right)$$
$$\left[\delta_{\mathrm{x},t-1}(\hat{\boldsymbol{\rho}}_{t-1})\left(\frac{\mathsf{c}_{\mathrm{y},t}^{(j)}}{(\mathsf{c}_{\mathrm{y},t}^{(j)})^2+(\mathsf{c}_{\mathrm{x},t}^{(j)})^2} - \frac{\mathsf{c}_{\mathrm{y},t}^{(i)}}{(\mathsf{c}_{\mathrm{y},t}^{(i)})^2+(\mathsf{c}_{\mathrm{x},t}^{(i)})^2}\right) + \delta_{\mathrm{y},t-1}(\hat{\boldsymbol{\rho}}_{t-1})\left(\frac{\mathsf{c}_{\mathrm{x},t}^{(i)}}{(\mathsf{c}_{\mathrm{y},t}^{(i)})^2+(\mathsf{c}_{\mathrm{x},t}^{(i)})^2} - \frac{\mathsf{c}_{\mathrm{x},t}^{(j)}}{(\mathsf{c}_{\mathrm{y},t}^{(j)})^2+(\mathsf{c}_{\mathrm{x},t}^{(j)})^2}\right)\right] \quad (10)$$

$$z_{\mathrm{x},t}(\hat{\boldsymbol{\rho}}_{t-1}) \triangleq \lambda^2 \sum_{(i,j)\in\mathcal{S}} \left(\frac{\hat{c}_{\mathrm{y},t}^{(j)}}{(\hat{c}_{\mathrm{y},t}^{(j)})^2+(\hat{c}_{\mathrm{x},t}^{(j)})^2} - \frac{\hat{c}_{\mathrm{y},t}^{(i)}}{(\hat{c}_{\mathrm{y},t}^{(i)})^2+(\hat{c}_{\mathrm{x},t}^{(i)})^2}\right) \sin\left(2\left[\tan^{-1}\left(\frac{\hat{c}_{\mathrm{y},t}^{(j)}}{\hat{c}_{\mathrm{x},t}^{(j)}}\right) - \tan^{-1}\left(\frac{\hat{c}_{\mathrm{y},t}^{(i)}}{\hat{c}_{\mathrm{x},t}^{(i)}}\right)\right]\right) \quad (11)$$

$$z_{\mathrm{y},t}(\hat{\boldsymbol{\rho}}_{t-1}) \triangleq \lambda^2 \sum_{(i,j)\in\mathcal{S}} \left(\frac{\hat{c}_{\mathrm{x},t}^{(i)}}{(\hat{c}_{\mathrm{y},t}^{(i)})^2+(\hat{c}_{\mathrm{x},t}^{(i)})^2} - \frac{\hat{c}_{\mathrm{x},t}^{(j)}}{(\hat{c}_{\mathrm{y},t}^{(j)})^2+(\hat{c}_{\mathrm{x},t}^{(j)})^2}\right) \sin\left(2\left[\tan^{-1}\left(\frac{\hat{c}_{\mathrm{y},t}^{(j)}}{\hat{c}_{\mathrm{x},t}^{(j)}}\right) - \tan^{-1}\left(\frac{\hat{c}_{\mathrm{y},t}^{(i)}}{\hat{c}_{\mathrm{x},t}^{(i)}}\right)\right]\right) \quad (12)$$

Note that $\det\big(\boldsymbol{J}(\mathbf{p}_{t+1})\big)$ is a highly non-linear function of the variable $\boldsymbol{\delta}_{t-1}(\hat{\boldsymbol{\rho}}_{t-1})$. Hence, a first-order Taylor approximation of $\det\big(\boldsymbol{J}(\mathbf{p}_{t+1})\big)$ is undertaken with respect to the change $\boldsymbol{\delta}_{t-1}(\hat{\boldsymbol{\rho}}_{t-1})$ in the control policy. The Taylor series approximation is denoted by $T\big(\delta_{\mathrm{x},t-1}(\hat{\boldsymbol{\rho}}_{t-1}), \delta_{\mathrm{y},t-1}(\hat{\boldsymbol{\rho}}_{t-1})\big)$ and is given by

$$\begin{aligned}\det\big(\boldsymbol{J}(\mathbf{p}_{t+1})\big) &\approx T\big(\delta_{\mathrm{x},t-1}(\hat{\boldsymbol{\rho}}_{t-1}), \delta_{\mathrm{y},t-1}(\hat{\boldsymbol{\rho}}_{t-1})\big)\\ &\triangleq \lambda^2 \sum_{(i,j)\in\mathcal{S}} \breve{h}_{\left(\mathsf{c}_{\mathrm{y},t}^{(j)},\mathsf{c}_{\mathrm{x},t}^{(j)},\mathsf{c}_{\mathrm{y},t}^{(i)},\mathsf{c}_{\mathrm{x},t}^{(i)}\right)}(0,0)\,.\end{aligned}$$

Here, $\breve{h}_{\left(\mathsf{c}_{\mathrm{y},t}^{(j)},\mathsf{c}_{\mathrm{x},t}^{(j)},\mathsf{c}_{\mathrm{y},t}^{(i)},\mathsf{c}_{\mathrm{x},t}^{(i)}\right)}(0,0)$ is the first order Taylor series of $h_{\left(\mathsf{c}_{\mathrm{y},t}^{(j)},\mathsf{c}_{\mathrm{x},t}^{(j)},\mathsf{c}_{\mathrm{y},t}^{(i)},\mathsf{c}_{\mathrm{x},t}^{(i)}\right)}$ around point $(0,0)$ and is given by (10) (at the top of this page). The function $\det\big(\boldsymbol{J}(\mathbf{p}_{t+1})\big)$ is highly non-linear with respect to the measurement noise and system disturbance. Hence, $\mathbb{E}\big\{\det\big(\boldsymbol{J}(\mathbf{p}_{t+1})\big) \,\big|\, \hat{\mathbf{p}}_{t-1} = \tilde{\boldsymbol{p}}_{t-1}, \hat{\mathbf{v}}_{t-1} = \tilde{\boldsymbol{v}}_{t-1}\big\}$ is approximated by using methods such as certainty equivalence or Monte Carlo simulations [25]. Here, certainty equivalence is employed, where the expected determinant is replaced with the determinant evaluated at the predicted position of the agent, obtained by simulating the inference process of the adversarial network.

Conditioned on a value of $\hat{\boldsymbol{\rho}}_{t-1}$ and after combining the Taylor series approximation of (10) with the principle of certainty equivalence, $\mathscr{P}_1$ has the following surrogate problem

$$\begin{aligned}\mathscr{P}_2:\ \underset{\boldsymbol{\delta}_{t-1}}{\text{minimize}}\ \ &(\delta_{\mathrm{x},t-1})^2 + (\delta_{\mathrm{y},t-1})^2\\ &+ \alpha\,\lambda^2 \sum_{(i,j)\in\mathcal{S}} h_{\left(\hat{\mathsf{c}}_{\mathrm{y},t}^{(j)},\hat{\mathsf{c}}_{\mathrm{x},t}^{(j)},\hat{\mathsf{c}}_{\mathrm{y},t}^{(i)},\hat{\mathsf{c}}_{\mathrm{x},t}^{(i)}\right)}(0,0)\\ &+ \alpha\,\delta_{\mathrm{x},t-1}\,z_{\mathrm{x},t}(\hat{\boldsymbol{\rho}}_{t-1}) + \alpha\,\delta_{\mathrm{y},t-1}\,z_{\mathrm{y},t}(\hat{\boldsymbol{\rho}}_{t-1})\\ \text{subject to}\ \ &|\delta_{\mathrm{x},t-1}| \leqslant D,\ |\delta_{\mathrm{y},t-1}| \leqslant D\,.\end{aligned} \quad (13)$$

The constraints in $\mathscr{P}_2$ are required to ensure that the total modification in each coordinate of the control policy does not exceed D which is a parameter that limits the error in the Taylor approximation. Also, $z_{\mathrm{x},t}$ and $z_{\mathrm{y},t}$ are defined in (11) and in (12) (at the top of this page). Given the simulated values $\tilde{\boldsymbol{p}}_{t-1}$ and $\tilde{\boldsymbol{v}}_{t-1}$, the instantiations of $\mathsf{c}_{\mathrm{y},t}^{(j)}$ and $\mathsf{c}_{\mathrm{x},t}^{(j)}$ $\forall j \in \mathcal{N}_{\mathrm{s}}$ (defined by (7)) are estimated as follows

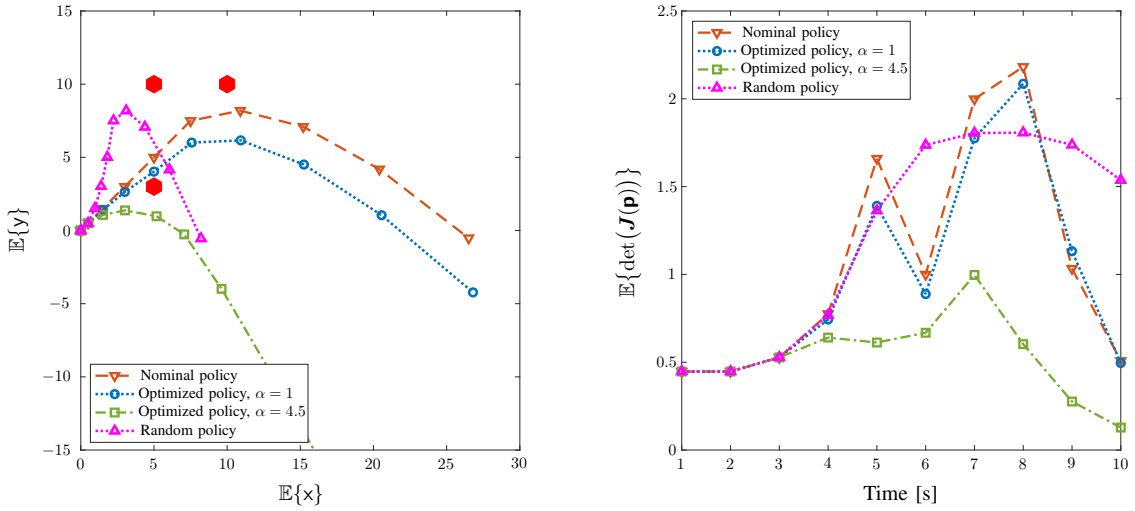
$$\begin{aligned}\hat{c}_{\mathrm{x},t}^{(j)} &= q_{\mathrm{x}}^{(j)} - \mathbb{E}\big\{\mathsf{x}_t \,\big|\, \hat{\mathbf{p}}_{t-1} = \tilde{\boldsymbol{p}}_{t-1}, \hat{\mathbf{v}}_{t-1} = \tilde{\boldsymbol{v}}_{t-1}\big\}\\ &\quad - \tilde{v}_{\mathrm{x},t-1} - \pi_{\mathrm{x}}(\hat{\boldsymbol{\rho}}_{t-1})\\ &= q_{\mathrm{x}}^{(j)} - \tilde{x}_{t-1} - 2\tilde{v}_{\mathrm{x},t-1} - \pi_{\mathrm{x}}(\hat{\boldsymbol{\rho}}_{t-1})\end{aligned}$$

$$\begin{aligned}\hat{c}_{\mathrm{y},t}^{(j)} &= q_{\mathrm{y}}^{(j)} - \mathbb{E}\big\{\mathsf{y}_t \,\big|\, \hat{\mathbf{p}}_{t-1} = \tilde{\boldsymbol{p}}_{t-1}, \hat{\mathbf{v}}_{t-1} = \tilde{\boldsymbol{v}}_{t-1}\big\}\\ &\quad - \tilde{v}_{\mathrm{y},t-1} - \pi_{\mathrm{y}}(\hat{\boldsymbol{\rho}}_{t-1})\\ &= q_{\mathrm{y}}^{(j)} - \tilde{y}_{t-1} - 2\tilde{v}_{\mathrm{y},t-1} - \pi_{\mathrm{y}}(\hat{\boldsymbol{\rho}}_{t-1})\,.\end{aligned}$$

Also, from (9), it follows that $h_{\left(\hat{c}_{\mathrm{y},t}^{(j)},\hat{c}_{\mathrm{x},t}^{(j)},\hat{c}_{\mathrm{y},t}^{(i)},\hat{c}_{\mathrm{x},t}^{(i)}\right)}(0,0) = \sin^2\Big(\tan^{-1}\big(\hat{c}_{\mathrm{y},t}^{(j)}/\hat{c}_{\mathrm{x},t}^{(j)}\big) - \tan^{-1}\big(\hat{c}_{\mathrm{y},t}^{(i)}/\hat{c}_{\mathrm{x},t}^{(i)}\big)\Big)$. In $\mathscr{P}_2$, the Fisher information obtained by the adversarial network of sensing nodes is minimized in a regularized fashion ensuring that the squared L_2-norm of the modification of the control policy is minimal. Problem $\mathscr{P}_2$ is an instance of quadratic programming (QP), which can be efficiently solved via, e.g., interior-point methods [30, Ch. 11]. Solving $\mathscr{P}_2$ is a tractable alternative to solving $\mathscr{P}_1$.

IV. Case Studies

This section shows the improvements provided by the proposed control policy refinement method via simulation. The agent evolves according to the dynamics of (2) in a 2D network consisting of three adversarial sensing nodes. The three adversarial sensing nodes are located at positions $\boldsymbol{q}^{(1)} = [5,3]^\top$, $\boldsymbol{q}^{(2)} = [5,10]^\top$, and $\boldsymbol{q}^{(3)} = [10,10]^\top$, and the covariance matrix of the noise in range measurements (1) is $\boldsymbol{I}$. Also, the dynamic disturbances $\{\mathbf{w}_t\}_{t\geqslant 1}$ are modeled as IID zero-mean Gaussian random vectors with covariance matrix $10^{-4}\,\boldsymbol{I}$. Each simulation runs for 10 time steps. From time 0 s to time 5 s, the nominal control policy is $\boldsymbol{\pi} = (\pi_{\mathrm{x}}, \pi_{\mathrm{y}}) = (0.5, 0.5)$, while from time 6 s to time 10 s, it is set to $\boldsymbol{\pi} = (\pi_{\mathrm{x}}, \pi_{\mathrm{y}}) = (0.9, -1.8)$. Moreover, the agent path is initiated with zero velocity at the point $(0,0)$. The numerical results reported in this work were obtained by averaging over 100 independent Monte Carlo simulations. In the following case studies, the QP is solved via the *quadprog* function in the optimization toolbox of MATLAB.



(a) (b)

Fig. 2. The average trajectories of the agent and the average FIM determinant of the adversarial sensing nodes for the nominal and modified polices. The nominal policy, defined before equation (4), is the default policy implemented by the agent when no secrecy enhancing controller is applied. (a) The average trajectories (measured in meters) in the xy-plane for the nominal and modified policies. The positions of the adversarial sensing nodes are shown by red hexagons. (b) The average of $\det(\boldsymbol{J}(\mathbf{p}))$ for the mobile agent as a function of time.

In this case study, the control policy $\boldsymbol{\pi}$ is a function of time and not a function of $\hat{\boldsymbol{\rho}}_t$. Hence, for solving $\mathscr{P}_1$, we also assume the change in the control policy $\boldsymbol{\delta}_t$ is only a function of time. In this way, there is no need for calculating $\hat{\boldsymbol{\rho}}_t$ to solve $\mathscr{P}_1$. A particle filter [33] is employed by the adversarial sensing nodes to construct the estimations $(\hat{\mathbf{p}}_t, \hat{\mathbf{v}}_t)$ given a history of range measurements $\mathbf{r}_{1:t} \triangleq (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_t)$. Moreover, consider that the agent can also replicate these estimators to construct $\tilde{\mathbf{p}}_t$ and $\tilde{\mathbf{v}}_t$. Given instantiations of estimations for the agent's position and velocity $(\tilde{\mathbf{p}}_t, \tilde{\mathbf{v}}_t)$, the particle filter constructs a set $\{(\omega_t^{(n)}, (\tilde{\boldsymbol{p}}_t^{(n)}, \tilde{\boldsymbol{v}}_t^{(n)})\}_{n=1}^N$ of particles with associated weights that approximate the posterior distribution of $(\tilde{\mathbf{p}}_t, \tilde{\mathbf{v}}_t)|\mathbf{r}_{1:t}$. The weights of the particles are normalized, i.e., $\sum_{n=1}^N \omega_t^{(n)} = 1$ and based on this weighted particle set, we consider the minimum-mean-square-error (MMSE) estimator for the state of the mobile agent as $\tilde{\boldsymbol{p}}_t \triangleq \sum_{n=1}^N \omega_t^{(n)} \tilde{\boldsymbol{p}}_t^{(n)}$ and $\tilde{\boldsymbol{v}}_t \triangleq \sum_{n=1}^N \omega_t^{(n)} \tilde{\boldsymbol{v}}_t^{(n)}$.

Figure 2(a) shows the average of the mobile agent trajectories in the xy-plane based on the nominal and modified policies. Figure 2(b) shows the average of $\det(\boldsymbol{J}(\mathbf{p}))$ for the mobile agent as a function of time. In addition to the nominal policy, a random policy and two optimized policies, with different values for the secrecy parameter $\alpha = 1$ and $\alpha = 4.5$, are considered. In case of the random policy, IID samples of Gaussian random variables $\mathcal{N}(-0.5, 0.25)$ and $\mathcal{N}(0, 0.01)$ are added to $\pi_{\mathrm{x}}(t)$ and $\pi_{\mathrm{y}}(t)$, respectively. To solve $\mathscr{P}_1$ efficiently, we approximate it as $\mathscr{P}_2$, which can be solved as a QP. Here, $D = \infty$ is used. The proposed optimization-based design modifies the agent trajectory to reduce the average of $\det(\boldsymbol{J}(\mathbf{p}))$. This improves the location secrecy of the agent as the designed control policy reduces the capabilities of the adversarial sensing nodes to localize the agent. By contrast, when the random policy is applied, the average of $\det(\boldsymbol{J}(\mathbf{p}))$ has mostly increased even though the agent trajectory has changed significantly compared to the nominal one.

The proposed optimization-based refinement for the control policy aims to reduce the determinant of the FIM, and consequently improve the location secrecy of the agent. As it can be seen from Figure 2(b), this is not true for the modified policy at all time steps. For instance, toward the end of the simulation, where $\alpha = 1$, the nominal trajectory leads to lower values of the FIM determinant. There are different ways to improve the performance of our optimization-based approach. Firstly, $\mathscr{P}_1$ is myopic in the sense that it tries to refine the control policy solely based on the immediate future. In fact, larger look-outs may improve the performance of our optimization-based solution. Secondly, improvements or alternatives for the employed approximations, i.e., the Taylor series and certainty equivalence, may also enhance the efficacy of the proposed optimization-based solution.

The secrecy parameter α in $\mathscr{P}_1$ regulates the trade-off between the change in the control policy and the localization inaccuracy of the adversarial network. To understand this trade-off, the case study of Figure 2 is repeated for different values of $\alpha = 1, 2, \ldots, 5$, where the effect of the secrecy parameter α is illustrated in Figure 3. The right y-axis of Figure 3 shows the maximum change in the expected value of the agent trajectory $\chi(\boldsymbol{p}^{\mathrm{o}}, \boldsymbol{p}^{\mathrm{n}}) \triangleq \max_t \sqrt{(\mathbb{E}\{\mathrm{x}_t^{\mathrm{o}} - \mathrm{x}_t^{\mathrm{n}}\})^2 + (\mathbb{E}\{\mathrm{y}_t^{\mathrm{o}} - \mathrm{y}_t^{\mathrm{n}}\})^2}$ versus the secrecy parameter α. Here, the superscript o stands for the optimized and superscript n refers to nominal. Also, the left y-axis of Figure 3 illustrates the change in the localization accuracy of the adversarial network versus the secrecy parameter α. The left y-axis of the Figure 3 represents the largest reduction in the expected value of the determinant of FIM, $\psi(\boldsymbol{p}^{\mathrm{o}}, \boldsymbol{p}^{\mathrm{n}}) \triangleq \max_t \mathbb{E}\{\det(\boldsymbol{J}(\mathbf{p}_t^{\mathrm{n}})) - \det(\boldsymbol{J}(\mathbf{p}_t^{\mathrm{o}}))\}$. In this example, as the secrecy parameter α gets larger, the proposed controller leads to a larger reduction in the FIM determinant and, consequently, to a higher obfuscation of

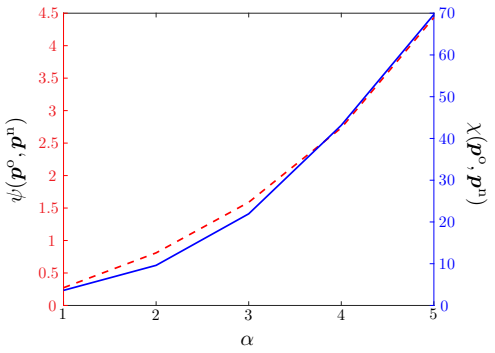


Fig. 3. Effects of secrecy parameter α: the largest reduction in the expected determinant of the FIM is represented by the dashed red curve with values indicated by the left y-axis; and the maximum change in the average trajectories of the agent (measured in meters) is shown by the solid blue curve with values indicated by the right y-axis.

the agent's location with respect to the adversarial network. Alongside the improvement in secrecy, increasing α leads to larger deviations in the agent's position from the trajectory generated by the nominal policy. In safety-critical applications, additional constraints could be added to our optimization-based solution to guarantee proprieties such as stability and collision-free navigation [34], [35].

V. Conclusion

This letter developed an NLN paradigm to enhance the secrecy of a mobile agent maneuvering through an adversarial network of sensing nodes. We designed optimization-based control policies for improving the location secrecy of an agent by reducing the Fisher information obtained by the adversarial network with minimal modifications of the nominal control. Case studies are presented to validate the proposed control policy for the agent as it maneuvers in the adversarial network. The results confirmed the efficacy of performing evasive maneuvers in reducing the ability of the adversarial network to infer the agent's position.

References

[1] M. Z. Win, F. Meyer, Z. Liu, W. Dai, S. Bartoletti, and A. Conti, "Efficient multi-sensor localization for the Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 153–167, Sep. 2018.

[2] T. Abdelzaher *et al.*, "Toward an Internet of Battlefield Things: A resilience perspective," *Computer*, vol. 51, no. 11, pp. 24–36, 2018.

[3] H. Liu, F. Sun, B. Fang, and X. Zhang, "Robotic room-level localization using multiple sets of sonar measurements," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 1, pp. 2–13, Jan. 2017.

[4] M. Weber, B. Jin, G. Lederman, Y. Shoukry, E. A. Lee, S. Seshia, and A. Sangiovanni-Vincentelli, "Gordian: Formal reasoning-based outlier detection for secure localization," *ACM Trans. Cyber-Physical Systems*, vol. 4, no. 4, pp. 1–27, 2020.

[5] J. Jiang, G. Han, C. Zhu, Y. Dong, and N. Zhang, "Secure localization in wireless sensor networks: A survey," *J. Commun.*, vol. 6, no. 6, pp. 460–470, 2011.

[6] S. Yan, R. Malaney, I. Nevat, and G. W. Peters, "Location spoofing detection for VANETs by a single base station in Rician fading channels," in *Proc. Vehicular Tech. Conf.* IEEE, 2015, pp. 1–6.

[7] S. Bartoletti, G. Bianchi, D. Orlando, I. Palamà, and N. Blefari-Melazzi, "Location security under reference signals' spoofing attacks: Threat model and bounds," in *Proc. Inter. Conf. Availability, Reliability and Security*, 2021.

[8] T. H. Chung, G. A. Hollinger, and V. Isler, "Search and pursuit-evasion in mobile robotics," *Auton. Robots*, vol. 31, no. 4, pp. 299–316, 2011.

[9] M. Z. Win, A. Conti, S. Mazuelas, Y. Shen, W. M. Gifford, D. Dardari, and M. Chiani, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.

[10] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to Autonomous Mobile Robots*. MIT press, 2011.

[11] A. Rabbachin, A. Conti, and M. Z. Win, "Wireless network intrinsic secrecy," *IEEE/ACM Trans. Netw.*, vol. 23, no. 1, pp. 56–69, Feb. 2015.

[12] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136–1165, Jul. 2018, special issue on *Foundations and Trends in Localization Technologies*.

[13] M. Z. Win, W. Dai, Y. Shen, G. Chrisikos, and H. V. Poor, "Network operation strategies for efficient localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1224–1254, Jul. 2018, special issue on *Foundations and Trends in Localization Technologies*.

[14] A. Conti, S. Mazuelas, S. Bartoletti, W. C. Lindsey, and M. Z. Win, "Soft information for localization-of-things," *Proc. IEEE*, vol. 107, no. 11, pp. 2240–2264, Nov. 2019.

[15] S. Martínez and F. Bullo, "Optimal sensor placement and motion coordination for target tracking," *Automatica*, vol. 42, no. 4, pp. 661–668, Apr. 2006.

[16] E. Tzoreff and A. J. Weiss, "Path design for best emitter location using two mobile sensors," *IEEE Trans. Signal Process.*, vol. 65, no. 19, pp. 5249–5261, Oct. 2017.

[17] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80. PMLR, 2018, pp. 1861–1870.

[18] I. Ziemann and H. Sandberg, "Parameter privacy versus control performance: Fisher information regularized control," in *American Control Conference*. IEEE, 2020, pp. 1259–1265.

[19] M. J. Khojasteh, A. Khina, M. Franceschetti, and T. Javidi, "Learning-based attacks in cyber-physical systems," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 1, pp. 437–449, 2020.

[20] B. Schlotfeldt, N. Atanasov, and G. J. Pappas, "Adversarial information acquisition," in *Workshop on Adversarial Robotics at RSS*, 2018.

[21] W. A. Malik, N. C. Martins, and A. Swami, "LQ control under security constraints," in *Cont. Cyber-Physical Sys.* Springer, 2013, pp. 101–120.

[22] R. S. Sutton, D. Precup, and S. Singh, "Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning," *Artificial Intelligence*, vol. 112, no. 1-2, pp. 181–211, 1999.

[23] D. Barcelliy, A. Bemporadz, and G. Ripaccioliy, "Hierarchical multi-rate control design for constrained linear systems," in *Proc. IEEE Conf. Decision and Control*. IEEE, 2010, pp. 5216–5221.

[24] F. Borrelli, A. Bemporad, and M. Morari, *Predictive Control for Linear and Hybrid Systems*. Cambridge University Press, 2017.

[25] D. P. Bertsekas, *Reinforcement Learning and Optimal Control*, 1st ed. Belmont, MA: Athena Scientific, 2019.

[26] D. Sun, M. J. Khojasteh, S. Shekhar, and C. Fan, "Uncertain-aware safe exploratory planning using Gaussian process and neural control contraction metric," in *Learning for Dynamics and Control*. PMLR, 2021, pp. 728–741.

[27] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Rev.*, vol. 53, no. 3, pp. 464–501, 2011.

[28] J. Cortés and M. Egerstedt, "Coordinated control of multi-robot systems: A survey," *SICE Journal of Control, Measurement, and System Integration*, vol. 10, no. 6, pp. 495–503, 2017.

[29] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[31] 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, Mar. 2010.

[32] S. Liu, M. Fardad, E. Masazade, and P. K. Varshney, "Optimal periodic sensor scheduling in networks of dynamical systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3055–3068, Jun. 2014.

[33] B. Ristic, M. S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Norwood, MA: Artech House, 2004.

[34] M. J. Khojasteh, V. Dhiman, M. Franceschetti, and N. Atanasov, "Probabilistic safety constraints for learned high relative degree system dynamics," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 781–792.

[35] R. Cheng, M. J. Khojasteh, A. D. Ames, and J. W. Burdick, "Safe multi-agent interaction through robust control barrier functions with learned uncertainties," in *Proc. IEEE Conf. Decision and Control*. IEEE, 2020.