Proficiency-Driven Decision-Making for Networks of Autonomous Agents

Anna Guerra*, Francesco Guidi†, Siwei Zhang‡, Pau Closas§, Davide Dardari* and Petar M. Djurić¶

*DEI "Guglielmo Marconi", University of Bologna, WiLAB-CNIT, Italy. Email: {anna.guerra3, davide.dardari}@unibo.it
†IEIIT, National Research Council of Italy, Bologna, Italy. Email: francesco.guidi@cnr.it
‡Institute of Communications and Navigation, German Aerospace Center (DLR), Germany. Email: siwei.zhang@dlr.de
§ Electrical and Computer Engineering Department, Northeastern University, Boston, USA. Email: closas@northeastern.edu
¶Electrical and Computer Engineering Department, Stony Brook University, USA. Email: petar.djuric@stonybrook.edu

***Abstract*—Autonomous agents play a crucial role in various modern fields, including emergency response and urban security. Their ability to operate effectively without direct human supervision is essential, especially in high-stakes situations. A key challenge is enabling these agents to evaluate their proficiency in completing tasks and use this evaluation for informed decision-making. This paper explores the use of a metric based on the assessment of autonomous agents' proficiency and applies it to improve their decision-making at run time. In this context, proficiency self-assessment will improve agent navigation, enabling agents to more effectively complete their mission tasks, such as reaching a destination area and enhancing estimation accuracy.**

I. Introduction

The deployment of autonomous agents in critical applications such as emergency response, urban security, and environmental monitoring has increased significantly in recent years [1]–[4]. These agents, i.e., unmanned aerial vehicles (UAVs) or mobile robots, must operate in complex and dynamic environments with minimal human supervision where the ability to make timely and effective decisions becomes crucial for mission success. However, a significant challenge in multi-agent networks lies in ensuring that each agent can self-assess its own proficiency in completing tasks and use it to make more informed and adaptive decisions [5]–[8].

Traditional decision-making frameworks in autonomous systems often assume that all agents have homogeneous capabilities, following predefined strategies or learned policies without assessing their proficiency in completing a given task [9]. Indeed, the existing approaches typically rely on static decision models, which do not account for the varying levels of expertise and uncertainty that arise from differences in sensing accuracy, computational resources, and environmental conditions [10].

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART"), under the NRPP of NextGenerationEU (Mision 4 – Component 2 -Investment 1.1) Prin 2022 (No. 104, 2/2/2022, CUP J53C24002790006), under ERC Grant no. 101116257 (project CUE-GO: Contextual Radio Cues for Enhancing Decision Making in Networks of Autonomous Agents), and the National Science Foundation under Awards 2212506, 1845833 and 2326559.

We can categorize self-assessment methods into three categories [6]. The *knowledge-based* category relies on human expertise in human-agent interactions [8]. The *learning-based* category uses machine learning like classifiers [11]. Lastly, the *test-based* category applies probabilistic and statistical methods [12]. Here, we focus on the test-based approach, where proficiency assessment is rooted in probabilistic modeling and Bayesian estimation techniques. Using these statistical methods, agents can quantify the uncertainty associated with their models and refine their decision-making accordingly.

This paper presents a proficiency-driven decision-making framework that enables autonomous agents to evaluate their proficiency and dynamically select the most suitable observation model to improve navigation and estimate the position of a moving source. In this regard, we consider a Bayesian proficiency metric inspired by the Cramér-Rao Lower Bound (CRLB), enabling agents to assess the accuracy of their available observation models and select the most suitable one for real-time decision-making. Additionally, to mitigate the impact of measurement noise, we propose a team-based strategy in which agents coordinate their decisions and collectively adopt a common observation model, enhancing overall system performance.

II. System Model

We consider a reference scenario in which a source moves within an area $\mathcal{A}$, as shown in Fig. 1, and its position and velocity are estimated in real time by a team of autonomous agents.

The source state at time step k is denoted by $\mathbf{s}_s^{(k)} = [\mathbf{p}_s^{(k)}, \dot{\mathbf{p}}_s^{(k)}] \in \mathbb{R}^4$ and includes position and velocity information. We assume that the positions of the agents are known due to the availability of global navigation satellite system (GNSS) information [9], [13], and, at each time instant, they can be written as $\mathbf{p}_i^{(k)} = \mathbf{p}_i^{(k-1)} + \mathbf{u}_i^{(k)}$ where $i \in \mathcal{T} = \{1, \dots, N\}$, with $\mathcal{T}$ being the set of agents and N its cardinality, $\mathbf{p}_i^{(k)} \in \mathbb{R}^2$ the position of the i-th agent, and $\mathbf{u}_i^{(k)} \in \mathbb{R}^2$ its navigation command input [10], [14]. Such a command is found as the solution to a navigation problem that will be described in Sec. IV.

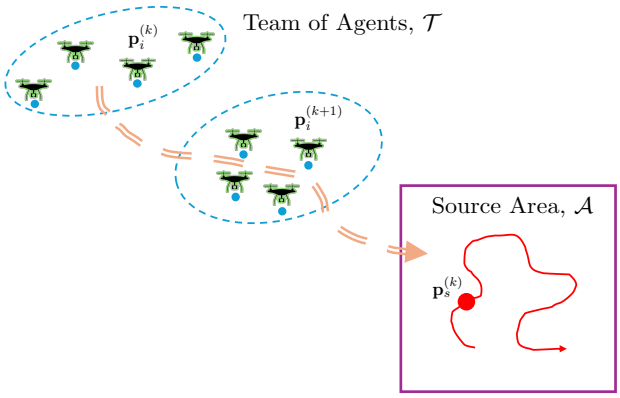

Fig. 1. Example of the considered scenario with four agents (UAVs in green) navigating towards a source area, denoted with $\mathcal{A}$.

Each agent is equipped with a library of statistical models $\mathcal{L}_i = \{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_m \ldots, \mathcal{M}_{N_\mathrm{M}}\}$, with N_M being the number of available models, which includes the generative model used to simulate the true source trajectory and its corresponding observations. In the next, we assume $\mathcal{L}_i = \mathcal{L}, \forall i \in \mathcal{T}$.

For each agent, the models are defined by two probability density functions (pdfs)

$$\mathcal{M}_m : \left\{ p\left(\mathbf{s}_s^{(k)} | \mathbf{s}_s^{(k-1)}; \mathcal{M}_{m,\mathrm{s}}\right),\ p\left(\mathbf{z}_i^{(k)} | \mathbf{s}_s^{(k)}; \mathcal{M}_{m,\mathrm{z}}\right) \right\}, \quad (1)$$

where $p\left(\mathbf{s}_s^{(k)} | \mathbf{s}_s^{(k-1)}; \mathcal{M}_{m,\mathrm{s}}\right)$ is the state transition pdf, and $p\left(\mathbf{z}_i^{(k)} | \mathbf{s}_s^{(k)}; \mathcal{M}_{m,\mathrm{z}}\right)$ is the observation pdf with $\mathbf{z}_i^{(k)}$ representing the observations available at the i-th agent at time instant k. From (1), we observe that $\mathcal{M}_m = \{\mathcal{M}_{m,\mathrm{s}}, \mathcal{M}_{m,\mathrm{z}}\}$. Within a Bayesian framework, typically Bayesian filters are adopted to efficiently combine prior information given the transition model with the measurement model to track the state of the source $\mathbf{s}_s^{(k)}$ [13]. The choice of the models is essential for accurate source localization. We now describe the transition and observation models.

Regarding the transition model, the motion of the source at time k is modeled as

$$\mathbf{s}_s^{(k)} = f_k\left(\mathbf{s}_s^{(k-1)}\right) + \mathbf{w}_s^{(k)}, \quad (2)$$

where $f_k(\cdot)$ is the transition function and $\mathbf{w}_s^{(k)}$ is the transition noise process.

Each agent collects a measurement vector that contains ranging (i.e., $z_{i,r}^{(k)}$) and bearing (i.e., $z_{i,\alpha}^{(k)}$) measurements ($\mathbf{z}_i^{(k)} = \left[z_{i,r}^{(k)}, z_{i,\alpha}^{(k)}\right] \in \mathbb{R}^2$), with

$$z_{i,r}^{(k)} = d_i^{(k)} + b_i^{(k)} + v_{i,r}^{(k)}, \quad z_{i,\alpha}^{(k)} = a_i^{(k)} + v_{i,\alpha}^{(k)}, \quad (3)$$

where the true distance and angle between the i-th agent and the source are, respectively, $d_i^{(k)} = d\left(\mathbf{p}_i^{(k)}, \mathbf{p}_s^{(k)}\right) = \left\|\mathbf{p}_s^{(k)} - \mathbf{p}_i^{(k)}\right\|$ and $a_i^{(k)} = a\left(\mathbf{p}_i^{(k)}, \mathbf{p}_s^{(k)}\right) = \arctan\left(y_s^{(k)} - y_i^{(k)}, x_s^{(k)} - x_i^{(k)}\right)$. The ranging bias $b_i^{(k)} = b\left(\mathbf{p}_i^{(k)}, \mathbf{p}_s^{(k)}\right)$ accounts for line-of-sight (LOS)/non-line-of-sight (NLOS) propagation between the source and the i-th agent [15]. The ranging noise, denoted with $v_{i,r}^{(k)}$, is drawn from a zero-mean Gaussian distribution with variance

$$\left(\sigma_{i,r}^{(k)}\right)^2 = \sigma_r^2 \left(d_i^{(k)} / d_\mathrm{ref}\right)^\beta, \quad (4)$$

where β is the path-loss exponent and σ_r^2 is the ranging variance measured at the reference distance of one meter (denoted by d_ref) [16]. The bearing noise is $v_{i,\alpha}^{(k)} \sim \mathcal{N}\left(0, \left(\sigma_{i,\alpha}^{(k)}\right)^2\right)$.

For simplicity, we assume that the team moves in a compact formation without significant spatial extension with respect to obstacles, such that all the agents consider the same ranging bias model. Moreover, we assume mild NLOS propagation effects, allowing for the assumption of a Gaussian distribution [15], [17], [18]. Specifically, the ranging bias follows a Gaussian distribution with mean $\mu_b^{(k)}$ and variance $\left(\sigma_b^{(k)}\right)^2$. The LOS locations are characterized by $\mu_b = 0$ and, typically, by small values of the variance σ_b^2, whereas the NLOS locations by $\mu_b > 0$ and larger values of σ_b^2. It is worth mentioning that the presence of unknown bias in range measurements is the main source of significant errors in localization systems [16].

In the next section, we introduce the concept of model proficiency and describe how it can be quantified and used to support decision making.

III. Proficiency in Multi-Agent Networks

We now introduce the concept of self-proficiency inspired by the definition of the CRLB applied to the distribution used to predict the measurements [5]. Here, proficiency is intended as a metric that quantifies an agent's ability to complete a given task (e.g., estimating the state of the source) using its model to predict measurements. Once each agent evaluates its proficiency, it can collaborate with others to derive the optimal navigation control.

We start by defining the measurement prediction bound (MPB) of a model $\mathcal{M}_m$ for the i-th agent at time k as [5]

$$\boldsymbol{\mathcal{P}}_i^{(k)}\left(\mathcal{M}_m | \mathbf{z}_i^{(1:k)}\right) = \mathbb{E}\left[-\frac{\partial^2 \ln p\left(\mathbf{z}_i^{(k)} | \mathbf{z}_i^{(1:k-1)}; \mathcal{M}_m\right)}{\partial \mathbf{z}_i^{(k)} \partial \left(\mathbf{z}_i^{(k)}\right)^T} \right], \quad (5)$$

where the statistical expectation $\mathbb{E}[\cdot]$ is computed according to the predictive distribution of the observations, that is, $p\left(\mathbf{z}_i^{(k)} | \mathbf{z}_i^{(1:k-1)}; \mathcal{M}_m\right)$. We can claim that [5], [19]

$$\boldsymbol{\mathcal{P}}_i^{(k)}\left(\mathcal{M}_m | \mathbf{z}_i^{(1:k)}\right) \geq \left(\mathbb{E}[(\hat{\mathbf{z}}_i^{(k)} - \mathbf{z}_i^{(k)})^T (\hat{\mathbf{z}}_i^{(k)} - \mathbf{z}_i^{(k)})]\right)^{-1}, \quad (6)$$

where the MPB represents the lower bound on the covariance matrix of the measurement prediction error.

To derive the MPB for the state-space model in (2)-(3) and considering the use of an extended Kalman filter (EKF)[1], we recall that the state covariance matrix is given by

$$\mathbf{P}_i^{(k|k)} = \mathbf{P}_i^{(k|k-1)} - \mathbf{K}_i^{(k)}\,\mathbf{S}_i^{(k|k-1)}\left(\mathbf{K}_i^{(k)}\right)^T, \tag{7}$$

with $\mathbf{P}_i^{(k|k-1)}$ being the covariance matrix of the predicted state, $\mathbf{K}_i^{(k)}$ being the Kalman gain and $\mathbf{S}_i^{(k|k-1)}$ being the innovation covariance matrix, which depends on the agent's position $\mathbf{p}_i^{(k)}$, and it is given by [15]

$$\mathbf{S}_i^{(k|k-1)}\left(\mathbf{p}_i^{(k)}\right) = \mathbf{J}_i^{(k|k-1)}\mathbf{P}_i^{(k|k-1)}\left(\mathbf{J}_i^{(k|k-1)}\right)^T + \mathbf{R}_i^{(k)}, \tag{8}$$

where $\mathbf{R}_i^{(k)}$ is the noise covariance matrix, using the model $\mathcal{M}_m$. The Jacobian matrix $\mathbf{J}_i^{(k|k-1)} \triangleq \mathbf{J}_i\left(\hat{\mathbf{p}}_{s;i}^{(k|k-1)}; \mathbf{p}_i^{(k)}\right)$ relates the measurements to the coordinates of the state and is evaluated at the predicted state estimate, i.e., $\hat{\mathbf{p}}_{s;i}^{(k|k-1)}$. The elements of the Jacobian are reported in Appendix A.

Consequently, the MPB becomes [5]

$$\boldsymbol{\mathcal{P}}_i^{(k)}\left(\mathcal{M}_m|\mathbf{z}_i^{(1:k)}\right) = \left(\mathbf{S}_i^{(k|k-1)}\right)^{-1}. \tag{9}$$

Subsequently, we introduce a metric that measures the agreement between the MPB and the mean squared error (MSE) (or squared error (SE) in the case of a single measurement) in measurement prediction.

For each measurement component $\xi \in \{r, \alpha\}$, we define the proficiency discrepancy as

$$\eta_{i,\xi}^{(k)}(\mathcal{M}_m) = \frac{\left|\left[\boldsymbol{\mathcal{P}}_i^{(k)}(\mathcal{M}_m \mid \mathbf{z}_i^{(1:k)})\right]_{\xi,\xi}^{-1} - \left(z_{i,\xi}^{(k)} - \hat{z}_{i,\xi}^{(k)}\right)^2\right|}{\left[\boldsymbol{\mathcal{P}}_i^{(k)}(\mathcal{M}_m \mid \mathbf{z}_i^{(1:k)})\right]_{\xi,\xi}^{-1}} \tag{10}$$

where $[\cdot]_{\xi,\xi}$ selects the diagonal entry associated with modality $\xi \in \{r, \alpha\}$, i.e., position $(1,1)$ for r, and $(2,2)$ for α.

Equation (10) expresses the relative deviation between the model-implied uncertainty (via the inverse of the MPB) and the squared prediction error for each modality ξ. This normalized metric allows each agent to assess how well model $\mathcal{M}_m$ explains its current observations. The total proficiency score aggregates range and bearing discrepancies as

$$\eta_{i,\text{MPB}}^{(k)}(\mathcal{M}_m) = \eta_{i,r}^{(k)}(\mathcal{M}_m) + \eta_{i,\alpha}^{(k)}(\mathcal{M}_m). \tag{11}$$

In the following, we show how proficiency can be effectively leveraged to enhance navigation and, consequently, the state estimation.

IV. Proficiency-Driven Navigation

Based on the concept of proficiency, each agent performs the following steps to derive its navigation command.

- P_1: Each agent, independently from the others, identifies the model that minimizes the discrepancy between its MPB and the squared error.
- P_2: Based on a consensus strategy achieved by sharing their identified models, the agents agree on a model to be used by everyone.
- P_3: According to the chosen model, the agent infers the control to approach the source area $\mathcal{A}$, and that minimizes the uncertainty of the source location.

For P_1, the i-th agent identifies the most proficient model as

$$\mathsf{P}_1 : \hat{\mathcal{M}}_{m;i}^{(k)} = \arg\min_{\mathcal{M}_m \in \mathcal{L}} \left\{\eta_{i,\text{MPB}}^{(k)}(\mathcal{M}_m)\,;\, \mathcal{M}_m \in \mathcal{L}\right\}. \tag{12}$$

Then, in P_2, each agent shares with the network its identified model $\hat{\mathcal{M}}_{m;i}^{(k)}$, and a simple solution is to select the model $\hat{\mathcal{M}}_m^{(k)}$ identified by the majority of agents. Finally, for P_3, the i-th agent estimates its control signal as $\mathbf{u}_i^{(k+1)} = \hat{\mathbf{p}}_i^{(k+1)} - \mathbf{p}_i^{(k)}$, where $\hat{\mathbf{p}}_i^{(k+1)}$ is found as the solution of P_3:

$$\mathsf{P}_3 : \hat{\mathbf{p}}_i^{(k+1)} = \arg\min_{\mathbf{p}_i^{(k+1)}} \mathcal{C}\left(\mathbf{p}_i^{(k+1)}; \hat{\mathbf{p}}_{s;i}^{(k+1|k)}\right), \tag{13}$$

with $\hat{\mathbf{p}}_{s;i}^{(k+1|k)}$ being the predicted state according to the chosen model $\hat{\mathcal{M}}_m^{(k)}$. The cost function in (13) is defined as [9], [10]

$$\mathcal{C}_i^{(k+1)} = \mathcal{C}\left(\mathbf{p}_i^{(k+1)}; \hat{\mathbf{p}}_{s,i}^{(k+1|k)}\right) = \tag{14}$$

$$= \underbrace{\omega_\text{g}\, f_\text{g}\left(\mathbf{p}_i^{(k+1)}; \hat{\mathbf{p}}_{s;i}^{(k+1|k)}\right)}_{\text{goal approaching}} + \underbrace{\omega_\text{p}\, f_\text{p}\left(\mathbf{p}_i^{(k+1)}; \hat{\mathbf{p}}_{s;i}^{(k+1|k)}\right)}_{\text{information seeking}}$$

where ω_g and ω_p are empirically determined weights. The two functions composing the cost function relate to the source approach and prediction accuracy goals, and they are defined as

$$f_\text{g}\left(\mathbf{p}_i^{(k+1)}; \hat{\mathbf{p}}_{s;i}^{(k+1|k)}\right) = \left\|\bar{\mathbf{p}}^{(k+1)} - \hat{\mathbf{p}}_{s;i}^{(k+1|k)}\right\|^2 \tag{15}$$

$$f_\text{p}\left(\mathbf{p}_i^{(k+1)}; \hat{\mathbf{p}}_{s;i}^{(k+1|k)}\right) = \text{tr}\left(\left[\mathbf{P}_i^{(k+1|k+1)}\right]_{1:2,1:2}\right). \tag{16}$$

The vector $\bar{\mathbf{p}}^{(k)}$ denotes the position centroid of the network and it is given by $\bar{\mathbf{p}}^{(k)} = (1/N)\sum_i \mathbf{p}_i^{(k)}$. Moreover, we will consider the following constraints:

$$\|\mathbf{p}_i^{(k+1)} - \mathbf{p}_i^{(k)}\| \le v_\text{max} \cdot \Delta t, \quad \forall i \in \mathcal{T} \tag{17}$$

$$d_\text{min} \le \|\mathbf{p}_i^{(k+1)} - \mathbf{p}_j^{(k+1)}\| \le d_\text{max}, \quad \forall i, j \in \mathcal{T}, i \ne j, \tag{18}$$

where v_max is the maximum allowed velocity, d_min and d_max define the acceptable inter-agent spacing. Finally, the solution of (13) for the i-th agent can be found as

$$\hat{\mathbf{p}}_i^{(k+1)} = -\,\nu\,\mathbf{M}_i\,\nabla_{\mathbf{p}_i^{(k+1)}}\mathcal{C}_i^{(k+1)} - \mathbf{N}_i(\mathbf{N}_i^T\mathbf{N}_i)^{-1}\mathbf{c}, \tag{19}$$

where ν represents the spatial step. The projection matrix is denoted with $\mathbf{M}_i = \mathbf{I} - \mathbf{N}_i\left(\mathbf{N}_i^T\mathbf{N}_i\right)^{-1}\mathbf{N}_i^T$ with $\mathbf{I}$ being the identity matrix and $\mathbf{N}_i = \left(\nabla_{\mathbf{p}_i^{(k+1)}}\mathbf{c}_i^T\right)$ being the gradient of the active constraints in $\mathbf{c}_i = \left[\mathbf{c}_{i,1}^T\ \mathbf{c}_{i,2}^T\ \mathbf{c}_{i,3}^T\right]^T$, where

$$\mathbf{c}_{i,1} = d_\text{v} - \mathsf{d}_\text{V}, \qquad \mathsf{d}_\text{V} = \left\{\Delta\, d_i^{(k)} : \Delta\, d_i^{(k)} \ge d_\text{v}\right\}, \tag{20}$$

$$\mathbf{c}_{i,2} = \mathsf{d}_\text{m} - d_\text{min}, \quad \mathsf{d}_\text{m} = \left\{d_{ij}^{(k)} : d_{ij}^{(k)} \le d_\text{min}\right\}, \tag{21}$$

$$\mathbf{c}_{i,3} = d_\text{max} - \mathsf{d}_\text{M}, \quad \mathsf{d}_\text{M} = \left\{d_{ij}^{(k)} : d_{ij}^{(k)} \ge d_\text{max}\right\}, \tag{22}$$

[1]The EKF is chosen because of its balance of computational efficiency and estimation accuracy in nonlinear systems.

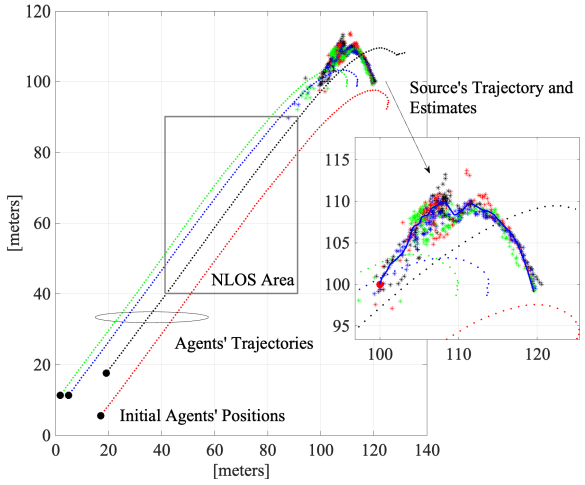

Fig. 2. Example of a simulated scenario (see its description in the text).

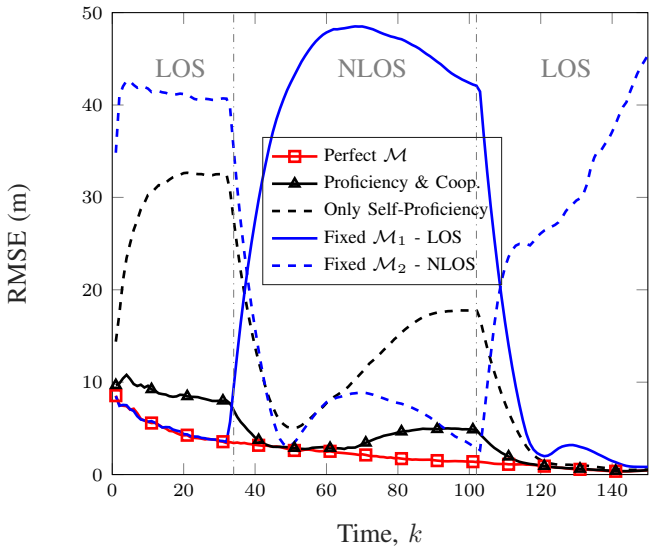

Fig. 3. RMSE on source position as a function of time and considering different navigation approaches.

where $\Delta\, d_i^{(k)} = \|\mathbf{p}_i^{(k+1)} - \mathbf{p}_i^{(k)}\|$, $d_{ij}^{(k)} = ||\mathbf{p}_j^{(k+1)} - \mathbf{p}_i^{(k+1)}||$, and $d_{\text{v}} = v_{\max} \cdot \Delta t$. The computations are reported in Appendix B.

In the following, we evaluate the performance of the proposed proficiency-driven navigation framework.

V. Results

a) Simulation Parameters: We considered a scenario with four agents, initially in the positions indicated by the black circled markers in Fig. 2. Each agent has access to a library of two models, i.e., $\mathcal{L}_i = \{\mathcal{M}_1, \mathcal{M}_2\}$, where these models correspond to the observation model for LOS and NLOS conditions, respectively, and are defined as $\mathcal{M}_m = \mathcal{M}_{m,\text{z}} : \boldsymbol{\theta}_{m,\text{z}} = \left\{\mu_b^{(m)}, \sigma_b^{(m)}\right\}$, $m = 1, 2$. The true generative model depends on the actual LOS/NLOS link between each agent and the source, as per Fig. 2. In our simulations, we set $\mathcal{M}_{1,\text{z}} : \boldsymbol{\theta}_{1,\text{z}} = \left\{\mu_b^{(1)}, \sigma_b^{(1)}\right\} = \left\{0, 10^{-3}\right\}$ m to generate LOS ranging biases, while for NLOS biases, we used $\mathcal{M}_{2,\text{z}} : \boldsymbol{\theta}_{2,\text{z}} = \left\{\mu_b^{(2)}, \sigma_b^{(2)}\right\} = \{40, 1\}$ m. As depicted in Fig. 2, the grey square represents an NLOS area of $50 \times 50\,\text{m}^2$. The source is initially positioned at $[100, 100]$ m, marked by the red dot in Fig. 2, and its motion is generated according to the transition model in [9, Eq. 40], with a time step $\Delta t = 1$ s and process noise covariance matrix $\mathbf{W} = \text{diag}(w_x, w_y) = \text{diag}(0.1, 2.5) \cdot 10^{-3}$. In Fig. 2, the estimated source positions are represented by circled markers, with each marker's color indicating the agent responsible for the estimate. For the model in (4), we considered $\sigma_r = 1\,\text{m}$ and $\sigma_\alpha = 1\,\text{deg}$ and, for source tracking, we employed an EKF, initialized as $\hat{\mathbf{p}}_{s;i}^{(0|0)} = \mathbf{p}_s^{(0)}$ and with an initial covariance matrix $\mathbf{P}^{(0|0)} = \text{diag}\left(20^2 \cdot \mathbf{I}_{2\times 2}, 10^2 \cdot \mathbf{I}_{2\times 2}\right)$. We considered $K = 150$ discrete time instants and $N_{\text{MC}} = 100$ Monte Carlo runs. Regarding navigation, we set the weights as $\omega_{\text{g}} = 1$ and $\omega_{\text{p}} = 1$, the spatial step $\nu = 5$, and the minimum and maximum inter-agent distances to $d_{\min} = 2\,\text{m}$ and $d_{\max} = 100\,\text{m}$, respectively. The maximum velocity was set to $v_{\max} = 1\,\text{m/step}$.

b) Results: In Fig. 3, we compare five different configurations. The benchmark, i.e., *Perfect* $\mathcal{M}$ (red solid line with square markers), where each agent has perfect knowledge of the model generating the measurements, i.e., the correct LOS or NLOS model is selected at each time instant based on the actual channel condition. Then, in the proficiency-driven navigation with cooperation, i.e., *Proficiency & Coop.* (black solid line with triangle markers), we assumed that each agent follows the three-step approach presented in Sec. IV, where the agreement is made by picking the model selected by the majority of agents. Based on this selection, navigation commands are derived using the analytical solution of (13). Then, for the *Only Self-Proficiency* approach (black dashed line), we consider the previous case but without a network consensus on the chosen model. Finally, in the *Fixed* $\mathcal{M}_m$ approach (blue lines), the agents use a fixed model (either LOS or NLOS), regardless of proficiency or actual channel conditions. Fig 3 illustrates the team root mean squared error (RMSE) on the source position estimation error for the different approaches. In particular, proficiency-driven navigation with a cooperative approach exhibits a trend similar to that of the benchmarks. This is because the agents can correctly identify the model by exploiting the proficiency, whereas cooperation allows them to discard wrong self-assessment caused by noisy measurements. In fact, when cooperation is absent, some agents may misidentify the model, leading to larger errors. As expected, the worst performance is observed in the *Fixed* $\mathcal{M}_m$ approach. For instance, considering the blue solid line (corresponding to the $\mathcal{M}_1$/LOS model), the RMSE is low when the agents operate in LOS conditions (i.e., before time step 30 and after time step 100), but significantly deteriorates in the NLOS area.

VI. Conclusions

This paper introduced a proficiency-driven decision-making framework for networks of autonomous agents, enabling them to self-assess their proficiency for improving navigation and the position estimation of a moving source. Using a Bayesian-based proficiency metric, the agents were able to evaluate the reliability of their observation models and select the most suitable one for real-time decision making. Simulation results demonstrated that proficiency-driven navigation, especially when combined with cooperative model selection, significantly improves localization accuracy compared to traditional approaches employing fixed models.

Appendix A

Considering the measurement model in (3) and a 2D scenario, the elements of the Jacobian can be written as

$$\begin{aligned}
&\left[\mathbf{J}_i^{(k|k-1)}\right]_{1,1} = \frac{\Delta\hat{x}_i^{(k)}}{\hat{d}_i^{(k)}}, \quad \left[\mathbf{J}_i^{(k|k-1)}\right]_{1,2} = \frac{\Delta\hat{y}_i^{(k)}}{\hat{d}_i^{(k)}},\\
&\left[\mathbf{J}^{(k|k-1)}\right]_{1,3} = \left[\mathbf{J}^{(k|k-1)}\right]_{1,4} = 0,\\
&\left[\mathbf{J}_i^{(k|k-1)}\right]_{2,1} = -\frac{\Delta\hat{y}_i^{(k)}}{\left(\hat{d}_i^{(k)}\right)^2}, \quad \left[\mathbf{J}_i^{(k|k-1)}\right]_{2,2} = \frac{\Delta\hat{x}_i^{(k)}}{\left(\hat{d}_i^{(k)}\right)^2},\\
&\left[\mathbf{J}^{(k|k-1)}\right]_{2,3} = \left[\mathbf{J}^{(k|k-1)}\right]_{2,4} = 0,
\end{aligned} \tag{23}$$

with $\Delta\hat{x}_i^{(k)} = \hat{x}_{s;i}^{(k|k-1)} - x_i^{(k)}$, $\Delta\hat{y}_i^{(k)} = \hat{y}_{s;i}^{(k|k-1)} - y_i^{(k)}$, and $\hat{d}_i^{(k)} = \sqrt{(\Delta\hat{x}_i^{(k)})^2 + (\Delta\hat{y}_i^{(k)})^2}$.

Appendix B

To sketch a closed-form solution of (13), we simplify the notation by omitting the temporal index and focusing solely on the x-coordinate, as the computations for the y-coordinate are similar. The first derivative of the cost function taken with respect to the x-coordinate of the i-th agent is

$$D_x = \omega_{\mathrm{g}} \frac{\partial f_{\mathrm{g}}(\mathbf{p}_i)}{\partial x_i} + \omega_{\mathrm{p}} \frac{\partial f_{\mathrm{p}}(\mathbf{p}_i)}{\partial x_i}. \tag{24}$$

The first addend of D_x can be written as

$$\frac{\partial f_{\mathrm{g}}(\mathbf{p}_i)}{\partial x_i} = \frac{(1/N)\sum_i (x_i - \hat{x}_{s;i})}{N}. \tag{25}$$

For the second, we consider the approach in [10] and take the information form as

$$\mathbf{Y}_i^{(k+1|k+1)} = \left(\mathbf{P}_i^{(k+1|k+1)}\right)^{-1} = \mathbf{Y}_i^{(k+1|k)} + \mathbf{I}_i^{(k+1)}, \tag{26}$$

where $\mathbf{Y}_i^{(k+1|k)}$ is independent of the current position, and

$$\mathbf{I}_i^{(k+1)} = \left(\mathbf{J}_i^{(k+1|k)}\right)^T \left(\mathbf{R}_i^{(k+1))}\right)^{-1} \mathbf{J}_i^{(k+1|k)} \tag{27}$$

is the expression of the snapshot CRB. Considering the derivative chain rule $\partial\mathbf{A}^{-1} = -\mathbf{A}^{-1}\,\partial\mathbf{A}\,\mathbf{A}^{-1}$ and the trace property $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, we have [10]

$$\frac{\partial f_{\mathrm{p}}(\mathbf{p}_i)}{\partial x_i} = \frac{\partial\ \mathrm{tr}\left(\mathbf{Y}_i^{-1}\right)}{\partial x_i} = -\,\mathrm{tr}\left(\mathbf{Y}_i^{-1}\ \frac{\partial \mathbf{Y}_i}{\partial x_i}\ \mathbf{Y}_i^{-1}\right). \tag{28}$$

In the next, we assume that the noise covariance matrix $\mathbf{R}_i$ is independent from the agent's position. By denoting with $\mathbf{C}_i = \left[(\mathbf{J}_i(\mathbf{p}_i))^T\,\mathbf{R}_i^{-1}\mathbf{J}_i(\mathbf{p}_i)\right]$, we can write

$$\frac{\partial}{\partial x_i}\left[\mathbf{C}_i\right]_{11} = \frac{2\,\Delta\hat{x}_i\,\Delta\hat{y}_i^2}{\hat{d}_i^4}\left[\frac{2}{\sigma_{i,\alpha}^2\,\hat{d}_i^2} - \frac{1}{\sigma_{i,r}^2}\right], \tag{29}$$

$$\frac{\partial}{\partial x_i}\left[\mathbf{C}\right]_{22} = \frac{2\,\Delta\hat{x}_i}{\hat{d}_i^4}\left(\frac{\Delta\hat{y}_i^2}{\sigma_{i,r}^2} + \frac{\Delta\hat{x}_i^2 - \Delta\hat{y}_i^2}{\sigma_{i,\alpha}^2\,\hat{d}_i^2}\right), \tag{30}$$

$$\frac{\partial}{\partial x_i}\left[\mathbf{C}_i\right]_{12} = \frac{\Delta\hat{y}_i}{\hat{d}_i^4}\left[\frac{\Delta\hat{x}_i^2 - \Delta\hat{y}_i^2}{\sigma_{i,r}^2} - \frac{3\,\Delta\hat{x}_i^2 - \Delta\hat{y}_i^2}{\sigma_{i,\alpha}^2\,\hat{d}_i^2}\right]. \tag{31}$$

References

[1] A. Guerra *et al.*, "Networks of UAVs of low complexity for time-critical localization," *IEEE Aerosp. Elect. Sys. Mag.*, vol. 37, no. 10, pp. 22–38, 2022.

[2] J. Wang, Y. Liu, and H. Song, "Counter-unmanned aircraft system(s) (C-UAS): State of the art, challenges, and future trends," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 3, pp. 4–29, 2021.

[3] I. Guvenc *et al.*, "Detection, tracking, and interdiction for amateur drones," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 75–81, 2018.

[4] A. Guerra *et al.*, "Reinforcement learning for joint detection and mapping using dynamic UAV networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 3, pp. 2586–2601, 2024.

[5] P. M. Djurić and P. Closas, "On self-assessment of proficiency of autonomous systems," in *Proc. Int. Conf. Acoustics, Speech, Signal Process.*, 2019, pp. 5072–5076.

[6] N. Conlon, N. R. Ahmed, and D. Szafir, "A survey of algorithmic methods for competency self-assessments in human-autonomy teaming," *ACM Comput. Surveys*, vol. 56, no. 7, pp. 1–31, 2024.

[7] X. Cao *et al.*, "Robot proficiency self-assessment using assumption-alignment tracking," *IEEE Trans. Robot.*, 2023.

[8] A. Norton *et al.*, "Metrics for robot proficiency self-assessment and communication of proficiency in human-robot teams," *ACM Trans. Human-Robot Interact.*, vol. 11, no. 3, pp. 1–38, 2022.

[9] A. Guerra, D. Dardari, and P. M. Djurić, "Dynamic radar network of UAVs: A joint navigation and tracking approach," *IEEE Access*, vol. 8, pp. 116 454–116 469, 2020.

[10] S. Zhang *et al.*, "Self-aware swarm navigation in autonomous exploration missions," *Proc. IEEE*, vol. 108, no. 7, pp. 1168–1195, 2020.

[11] X. Cao, J. W. Crandall, and M. A. Goodrich, "Improving robot proficiency self-assessment via meta-assessment," *IEEE Robot. Autom. Lett*, vol. 8, no. 11, 2023.

[12] B. Israelsen *et al.*, "Machine self-confidence in autonomous systems via meta-analysis of decision processes," in *Proc. Int. Conf. Applied Human Factors Erg. Energy*. Springer, Jul. 2020, pp. 213–223.

[13] D. Dardari, P. Closas, and P. M. Djurić, "Indoor tracking: Theory, methods, and technologies," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1263–1278, 2015.

[14] S. Zhang *et al.*, "Multi-agent navigation with Reinforcement Learning enhanced information seeking," in *Proc. 30th European Signal Process. Conf.*, 2022, pp. 982–986.

[15] E. Arias-de Reyna *et al.*, "Enhanced indoor localization through crowd sensing," in *Proc. Int. Conf. Acoustics, Speech and Signal Process.*, 2017, pp. 2487–2491.

[16] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position error bound for UWB localization in dense cluttered environments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 2, pp. 613–628, 2008.

[17] I. Guvenc, C.-C. Chong, and F. Watanabe, "NLOS identification and mitigation for UWB localization systems," in *Proc. IEEE Wireless Commun. and Netw. Conf.*, 2007, pp. 1571–1576.

[18] D.-H. Kim, A. Farhad, and J.-Y. Pyun, "Uwb positioning system based on lstm classification with mitigated nlos effects," *IEEE Internet of Things J.*, vol. 10, no. 2, pp. 1822–1835, 2023.

[19] H. L. Van Trees and K. L. Bell, "Bayesian bounds for parameter estimation and nonlinear filtering/tracking," *AMC*, vol. 10, no. 12, pp. 10–1109, 2007.