# 3D UAV Localization Optimization under Jamming Attacks: A Hybrid Machine Learning and Wireless Jamming Defense Method

Yujiao Zhu*, Mingzhe Chen†, Sihua Wang*, Yuchen Liu‡, Gaolei Li§, and Changchuan Yin*

*Beijing Laboratory of Advanced Information Network, Beijing University of Posts and Telecommunications, Beijing, 100876, China. Emails: {yjzhu, sihuawang, ccyin}@bupt.edu.cn.

†Department of Electrical and Computer Engineering and Institute for Data Science and Computing, University of Miami, Coral Gables, FL, 33146, USA. Email: mingzhe.chen@miami.edu.

‡Department of Computer Science, North Carolina State University, Raleigh, NC, 27695, USA. Email: yuchen.liu@ncsu.edu.

§School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China. Email: gaolei_li@sjtu.edu.cn.

***Abstract*—In this paper, the optimization of unmanned aerial vehicle (UAV) localization under jamming attacks is studied. In the considered network, a base station (BS) coordinates with active UAVs jointly locate a target UAV. Active UAVs transmit signals to the target UAV. Then, the signals are reflected by the target UAV and received by mobile active and passive UAVs which are carrying out missions. Based on received signals, active and passive UAVs calculate the distances from the transmit nodes to receive nodes via the target UAV and transmit distance information to the BS. During this positioning process, the jamming UAV transmits discontinuous signals to interfere with the localization of the target UAV. The BS requires to select the appropriate subset of measured distances for locating target UAV and select a jamming attack defense method to calculate the position of the target UAV. Due to the UAV transmit power energy consumption limit and UAV propulsion energy consumption limit, active UAVs require to determine whether to transmit signals to the target UAV and optimize their trajectories. This three-dimensional UAV localization problem is formulated as an optimization problem whose goal is to minimize the positioning error between the estimated and the ground truth positions of the target UAV while considering jamming attacks. To solve this problem, we propose a reinforcement learning (RL) based jamming attack defense method that jointly uses a) a generate adversarial network (GAN) based anti-attack method and b) a wireless technique based anti-attack method that determines signals transmit scheme of active UAVs and adjusts the trajectories and transmit powers of active UAVs, and the passive UAVs selection scheme to avoid jamming attacks and improve localization accuracy. In particular, the designed RL based method enables active UAVs to determine whether to transmit signals, their transmit powers, and trajectories, and enables the BS to select the most appropriate subsets of distance measurement information and the optimal jamming attack defense method from (i.e., 1) GAN, 2) wireless technique optimization, and 3) the combination of 1) and 2)) according to the movement of UAVs and the unknown jamming attack pattern of the jamming UAV. Simulation results show that the proposed method can reduce the positioning error of the target UAV by up to 33.8% compared to the method that does not consider the GAN based jamming attack defense method.***

## I. Introduction

Unmanned aerial vehicle (UAV) localization has been widely used in military and civilian applications [1]. Radio frequency (RF) based passive localization methods can accurately locate unknown target UAVs in scenarios where the global navigation satellite systems (GNSSs) are not available [2]. However, using passive radio frequency localization methods to locate target UAVs faces many challenges. First, the high-speed mobility of UAVs makes it difficult to estimate the real-time three-dimensional (3D) coordinates of UAVs. Second, the interference of the dynamic wireless environments and attacks of jamming objects affect the transmission signals used for UAV localization.

### A. Related Works

Existing works [3]–[9] have studied several problems of using radio frequency for UAV localization. In particular, in [3] and [4], the authors used the time of arrival of signals to evaluate the localization accuracy for the target UAV with different speeds by using the ground sensors. The authors in [5] measured the angle of signals to estimate the position of the target UAV. In [6], the authors obtained the distance information based on the signal strength for UAV localization. However, the works in [3]–[6] ignored the impacts of the positions of sensors on localization accuracy. The authors in [7] investigated the relationship between the deployment of sensors and the UAV localization accuracy. In [8], the authors optimized the deployment of ground sensors under communication constraints. The authors in [9] optimized the selection of sensors for real-time UAV localization. However, most of these works [3]–[9] used static sensors for UAV localization, which may not be applied for UAVs with high-speed movement. In addition, the above works [3]–[9] did not consider how dynamic jamming attacks affect the UAV localization performance.

A number of existing works such as [10]–[14] have studied the problem of UAV localization while avoiding jamming attacks. In [10], the authors proposed a UAVs grouping and merging scheme to reduce the influence of interference on UAV localization. The authors in [11] studied the relationship between the UAV localization performance and the number of participating BSs under different signal-to-interference-plus-noise radio (SINR) conditions. However, these wireless technique jamming attacks defense methods in [10], [11](i.e. adjusting the position of UAVs or the number of BSs) bring large costs. The authors in [12] proposed a novel deep neural network (DNN) model to generate an image of received signals amplitude and phase to improve the positioning accuracy of the UAV by using the noise and interference in the environment as features. The authors in [13] used a convolutional neural network (CNN) to analyze the received RF signals to prevent interference and estimate the angle of arrival accurately, thus decreasing the positioning error of the UAV. In [14], the authors proposed to use a DNN to recognize the visual information of UAVs under the scenario with high interference and improve the positioning accuracy of the UAV. However, the performance of these machine learning based jamming attacks defense methods in [12]–[14] depends on the training datasets. In particular, these ML-based defense methods learn the features of samples in the training datasets. When the jamming pattern and position of the jamming UAV are unknown, the trained neural network may not work. To this end, we proposed to use a RL method to adaptively select the most appropriate method from the ML based jamming attacks defense method, the traditional wireless technique based jamming attacks defense method, or the joint ML and traditional wireless technique based jamming attacks defense method to improve the localization performance while avoiding jamming attacks.

### *B. Contributions*

The main contribution of this work is to design a framework that enables the base station (BS) and UAVs to jointly locate a target UAV while considering jamming attacks. Our key contributions include:

- We propose a joint BS and UAVs based target localization framework. In particular, to estimate the position of the target UAV, active UAVs can transmit signals to the target UAV. Then, the signals are reflected by the target UAV and received by active and passive UAVs. These UAVs estimate the distances from transmit nodes to receive nodes via the target UAV and transmit the distance information to the BS.
- Since passive UAVs are carrying out their missions at different positions and active UAVs are also moving in real time, the BS requires to select the optimal subsets of distance measurement information for locating the target UAV, where the subset consists of at least four measured UAVs[1]. In addition, the jamming UAV transmits interference signals to interfere with the accuracy of distance measurement estimated by passive UAVs. To improve the localization accuracy while avoiding interference of the jamming UAV under limited UAV energy consumption, active UAVs requires to determine whether to transmit signals to the target UAV, adjust their transmit powers, and optimize their trajectories. And the BS requires to select the subset of distance measurement information to calculate the position of the target UAV. This problem is formulated as an optimization problem whose goal is to minimize the positioning error of the target UAV while considering jamming attacks.
- To solve this problem, we first propose a generative adversarial network (GAN) based method and a wireless technique based method to avoid jamming attacks from the jamming UAV. Since the position and jamming pattern of the jamming UAV are unknown, a distributed reinforcement learning (RL) is proposed to jointly select the appropriate subset of passive UAVs for the target UAV and the optimal jamming attack defense method from the GAN based jamming attack defense method, the traditional wireless technique based jamming attack defense method, and the joint GAN and wireless technique based jamming attack defense method to minimize the positioning error while avoiding jamming attacks.
- The results show that the proposed RL based joint GAN and wireless technique based jamming attack defense method can yield 23.1% and 33.8% gains in terms of the positioning accuracy compared to the method that only uses wireless technique based jamming attack defense method and the method that only uses the GAN based jamming attack defense method to avoid jamming attacks respectively.

[1]Estimating 3D coordinates of the target UAV requires at least four UAVs that receives signals reflected by the target UAV [15].

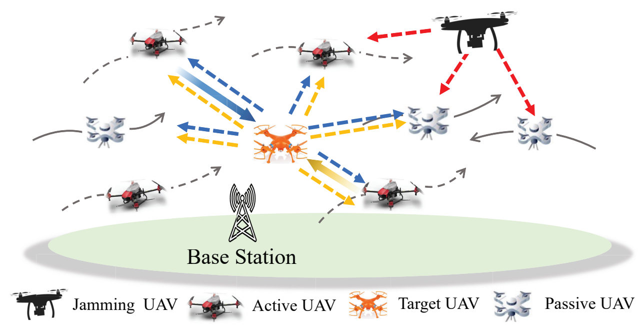


Fig. 1. Illustration of the considered UAV localization network.

The rest of this paper is organized as follows. The system model and problem formulation are described in Section II. The proposed RL based joint jamming attack defense method are introduced in Section III. The optimal localization accuracy under jamming attacks is analyzed in Section IV. Simulation results are analyzed in Section V. Conclusions are drawn in Section VI.

## II. System Model and Problem Formulation

Consider a wireless network in which one BS, a set $\mathcal{K}$ of $k$ active UAVs and a set $\mathcal{U}$ of $U$ passive UAVs cooperatively locate a target UAV under attacks from a jamming UAV. In the considered network, the transmitted signals from active UAVs will be reflected by the target UAV to active and passive UAVs. According to the received signals, active and passive UAVs calculates the signal transmission distance from the

TABLE I. List of Notations

| Notation | Description | Notation | Description |
|---|---|---|---|
| $K$ | Number of active UAVs | $U$ | Number of passive UAVs |
| $i_{k,t}$ | Indicator represents active UAV $k$ transmits signals | $\boldsymbol{l}_{k,t}^{\mathrm{A}}$ | Position of active UAV $k$ |
| $\boldsymbol{l}_{u,t}^{\mathrm{P}}$ | Position of passive UAV $u$ | $v$ | Flight speed of each UAV |
| $\Delta_t$ | Time duration of a time slot | $\alpha_{k,t}$ | Yaw angle of the active UAV |
| $\beta_{k,t}$ | Pitch angle of the active UAV | $\tau_{u,t}$ | Transmit time of signals |
| $P_{k,t}$ | Aerodynamic power consumption of active UAV $k$ | $E_{k,t}$ | Aerodynamic energy consumption of active UAV $k$ |
| $\boldsymbol{l}_t^{\mathrm{J}}$ | Position of the jamming UAV | $P^{\mathrm{J}}$ | Transmit power of the jamming UAV |
| $\boldsymbol{l}_t$ | Position of the target UAV | $\hat{\boldsymbol{l}}_t$ | Estimated position of the target UAV |
| $h_{\mathrm{J},u,t}$ | Path loss from the jamming UAV to passive UAV $u$ | $g$ | Gravitational acceleration |
| $h_{u,t}$ | Path loss from the target UAV to passive UAV $u$ | $h_{k,t}$ | Path loss from active UAV $k$ to the target UAV |
| $p_{k,t}$ | Transmit power of active UAV $k$ | $f_{\mathrm{J}}$ | Probability of the jamming attacks |
| $D_{u,t}^{\mathrm{G}}$ | Data size of passive UAV $u$ | $p_{u,t}$ | Transmit power of passive UAV $u$ |
| $x_{k,u,t}$ | Reflection coefficient from the active $k$ to passive $u$ | $\beta_0$ | LoS path loss at a reference distance |
| $s_{k,u,t}^{\mathrm{P}}$ | SINR of passive UAV $u$ transmitted from active UAV $k$ | $s_{u,t}^{\mathrm{G}}$ | SNR of signals received by the BS from passive UAV $u$ |
| $\bar{g}_{u,t}$ | Path loss from passive UAV $u$ to the BS | $I_{u,t}^{\mathrm{J}}$ | Jamming signals received by passive UAV $u$ |
| $\boldsymbol{l}_{\mathrm{B}}$ | Position of the BS | $\chi_{u,t}$ | Elevation angle of passive UAV $u$ |
| $L_{\mathrm{FS}}$ | Free-space path loss | $g_{u,t}^{\mathrm{LoS}}$ | LoS path loss from UAV $u$ to the BS |
| $\Pr\left(l_{u,t}^{\mathrm{LoS}}\right)$ | Probability of LoS | $g_{u,t}^{\mathrm{NLoS}}$ | NLoS path loss from UAV $u$ to the BS |
| $W_1$ | Bandwidth of UAV links | $\epsilon^2$ | Variance of Gaussian noise |
| $W_2$ | Bandwidth of UAV-BS links | $E_{k,u,t}^{\mathrm{T}}$ | Transmit energy consumption of active UAV $k$ |
| $\hat{\boldsymbol{d}}_t$ | Distance measurement information | $T_{k,t}^{\mathrm{G}}$ | Transmit delay from active UAV $k$ to the BS |
| $T_{u,t}^{\mathrm{G}}$ | Transmission delay from passive UAV $u$ to the BS | $T$ | Number of time slots |
| $e_t$ | Positioning error of the target UAV | $q_{m,t}$ | Indicator for selecting distance $m$ for locating the target UAV |

transmit nodes to receive nodes via the target UAV. Then, these distance information are transmitted to the BS for calculating the position of the target UAV. To reduce the influence of the jamming UAV attacks and the mobility of UAVs, active UAVs require to optimize the trajectories, transmit powers, and determine whether to transmit signals at each time slot. In addition, the BS requires to select the most appropriate subset of distance information to estimate the real-time position of the target UAV accurately. In addition, we assume that when multiple active UAVs are transmitting signals at the same time, UAVs that receives signals can distinguish the signals based on multi-source separation techniques transmitted from different active UAVs [2].

Next, we first introduce the mobility patterns of UAVs and the jamming pattern of the jamming UAV. Then, the signal transmission model and the positioning model used to calculate the position of the target UAV are introduced. Finally, we formulate the optimization problem.

### A. UAV Aerodynamic Model

The mobility patterns and propulsion energy consumption of each UAV depend on its position, pitch angle, yaw angle, and flying speed. Next, we present the UAV movement model and UAV flight energy consumption model, respectively.

*1) UAV movement model:* Given the 3D coordinate $\boldsymbol{l}_{k,t}^{\mathrm{A}}$ of active UAV $k$ at time slot $t$, $\boldsymbol{l}_{k,t+1}^{\mathrm{A}}$ is given by

$$\boldsymbol{l}_{k,t+1}^{\mathrm{A}}\left(\alpha_{k,t},\beta_{k,t}\right)=\boldsymbol{l}_{k,t}^{\mathrm{A}}+v\Delta_t\begin{bmatrix}\cos\alpha_{k,t}\cos\beta_{k,t}\\ \sin\alpha_{k,t}\cos\beta_{k,t}\\ \sin\beta_{k,t}\end{bmatrix}, \tag{1}$$

where $\alpha_{k,t}$ is the yaw angle, $\beta_{k,t}$ is the pitch angle of active UAV $k$, $v$ is the speed, and $\Delta_t$ is the time duration of each time slot.

*2) UAV propulsion energy consumption model:* The aerodynamic power consumption $P_{k,t}\left(\alpha_{k,t},\beta_{k,t}\right)$ of active UAV $k$ at time slot $t$ is given by [16]

$$\begin{aligned}&P_{k,t}\left(\alpha_{k,t},\beta_{k,t}\right)\\&=\frac{C_1}{\sqrt{\left(v_{k,t}^{\mathrm{L}}\right)^2+\sqrt{\left(v_{k,t}^{\mathrm{L}}\right)^4+4\left(v_{k,t}^{\mathrm{H}}\right)^4}}}+Mgv_{k,t}^{\mathrm{Z}}+C_2\left(v_{k,t}^{\mathrm{L}}\right)^3,\end{aligned} \tag{2}$$

where $C_1$ and $C_2$ are coefficients [16], $v_{k,t}^{\mathrm{L}}=v\cos\beta_{k,t}$ is the horizontal flight speed, $v_{k,t}^{\mathrm{Z}}=v\sin\beta_{k,t}$ is the vertical flight speed, $g$ is the acceleration of gravity, $M$ is the weight of each UAV, and $v_{k,t}^{\mathrm{H}}$ is the power needed for hovering [16]. Then, the required instantaneous propulsion energy at time slot $t$ is given by

$$E_{k,t}\left(\alpha_{k,t},\beta_{k,t}\right)=P_{k,t}\left(\alpha_{k,t},\beta_{k,t}\right)\Delta_t. \tag{3}$$

### B. Jamming Model

To interfere with the localization of the target UAV, the jamming UAV transmits discontinuous interference signals to UAVs to interfere with the target UAV localization. We use an indicator $j_t$ to represent whether the jamming UAV transmits signals at time slot $t$. $j_t=1$ implies that the jamming UAV transmits signals and $j_t=0$, otherwise. Let $f_{\mathrm{J}}$ be the probability that the jamming UAV transmits jamming signals. The jamming power received by passive UAV $u$ is given by

$$I_{u,t}^{\mathrm{J}}=j_tP^{\mathrm{J}}|h_{\mathrm{J},u,t}|^2, \tag{4}$$

where $P^{\mathrm{J}}$ is the transmit power of the jamming UAV, $|h_{\mathrm{J},u,t}|^2=\beta_0\left\|\boldsymbol{l}_t^{\mathrm{J}}-\boldsymbol{l}_{u,t}^{\mathrm{P}}\right\|^{-2}$ is the path loss from the jamming UAV to passive UAV $u$ with $\boldsymbol{l}_{u,t}^{\mathrm{P}}$ being the position of passive UAV $u$, $\boldsymbol{l}_t^{\mathrm{J}}$ being the position of the jamming UAV at time slot $t$, and $\beta_0$ being the path loss at a reference distance.

### C. Transmission Model

Signals are transmitted from active UAVs to the target UAV, and then reflected to active and passive UAVs. Based on received signals, active and passive UAVs estimate the transmission distance from the transmit node to themselves via the target UAV. Then, the estimated distances are transmitted to the BS. Due to the interference of the jamming UAV, active UAVs require to determine whether to transmit signals to the target UAV. Next, we introduce the transmission links a) from active UAVs to the target UAV and from the target UAV to active and passive UAVs, and b) from active and passive UAVs to the BS.

*1) UAV Links:* We assume that a link between any two UAVs is line-of-sight (LoS). Due to the interference introduced by channel noise, the jamming UAV, and signals transmitted from other active UAVs, the signal-to-interference-plus-noise ratio (SINR) of the signal transmitted from active UAV $k$, reflected by the target UAV, and received by passive UAV $u$ at time slot $t$ is

$$\begin{aligned} & s_{k,u,t}^{\mathrm{P}}\left(i_{k,t}, \boldsymbol{l}_{k,t}^{\mathrm{A}}, p_{k,t}\right) \\ & = \frac{i_{k,t} p_{k,t} |h_{u,t} x_{k,u,t} h_{k,t}|^2}{\epsilon^2 + I_{u,t}^{\mathrm{J}} + \sum_{k'=1,k'\neq k}^{K} p_{k',t} |h_{u,t} x_{k',u,t} h_{k',t}|^2}, \end{aligned} \tag{5}$$

where $i_{k,t} \in \{0,1\}$ is an indicator that represents whether active UAV $k$ transmits signals with $i_{k,t}=1$ indicating that active UAV $k$ transmits signals to the target UAV, otherwise, $i_{k,t}=0$. $p_{k,t}$ is the transmit power of active UAV $k$, $\epsilon^2$ is the power of the Gaussian noise, $h_{k,t}^2 = \beta_0 \left\|\boldsymbol{l}_{k,t}^{\mathrm{A}} - \boldsymbol{l}_t\right\|^{-2}$ is the path loss from active UAV $k$ to the target UAV with $\boldsymbol{l}_t$ being the position of the target UAV and $\boldsymbol{l}_{k,t}^{\mathrm{A}} = \left[x_{k,t}^{\mathrm{A}}, y_{k,t}^{\mathrm{A}}, z_{k,t}^{\mathrm{A}}\right]^T$ being the position of active UAV $k$, $x_{k,u,t}$ is the reflection coefficient of the target UAV from active UAV $k$ to passive UAV $u$, $h_{u,t}^2 = \beta_0 \left\|\boldsymbol{l}_t - \boldsymbol{l}_{u,t}^{\mathrm{P}}\right\|^{-2}$ represents the path loss from the target UAV to passive UAV $u$ with $\boldsymbol{l}_{u,t}^{\mathrm{P}}$ being the position of passive UAV $u$, and $I_{u,t}^{\mathrm{J}}$ is the interference caused by the jamming UAV.

Similarly, active UAVs can also receive signals to calculate signals transmit distance. The SINR of signals received by active UAV $k'$, transmitted from active UAV $k$, and reflected by the target UAV is given by

$$\begin{aligned} & s_{k,k',t}^{\mathrm{A}}\left(i_{k,t}, \boldsymbol{l}_{k,t}^{\mathrm{A}}, \boldsymbol{l}_{k',t}^{\mathrm{A}}, p_{k,t}\right) \\ & = \frac{i_{k,t} p_{k,t} |h_{k,t}^2 x_{k,k',t}|^2}{\epsilon^2 + I_{u,t}^{\mathrm{J}} + \sum_{k^*=1,k^*\neq k}^{K} i_{k^*,t} p_{k^*,t} |h_{k',t} x_{k^*,k',t} h_{k^*,t}|^2}, \end{aligned} \tag{6}$$

where $x_{k,k',t}$ is the reflection coefficient from active UAV $k$ to active UAV $k'$ via the target UAV. If $k=k'$, $x_{k,k',t}$ represents echo signals transmitted from active UAV $k$, reflected by the target UAV, and received by itself. The transmit energy consumption of signals from active UAV $k$ to passive UAV $u$ can be written as

$$E_{k,m,t}^{\mathrm{T}}\left(i_{k,t}, \boldsymbol{l}_{k,t}^{\mathrm{A}}, p_{k,t}\right) = \frac{i_{k,t} p_{k,t} D_{k,t}^{\mathrm{A}}}{W_1 \log_2\left(1 + s_{k,u,t}^{\mathrm{P}}/i_{k,t}\right)}, \tag{7}$$

where $D_{k,t}^{\mathrm{A}}$ is the data size of active UAV $k$ and $W_1$ is the bandwidth. The transmit energy consumption of signals from active UAV $k$ to active UAV $k'$ can be written as $E_{k,k',t}^{\mathrm{T}}\left(i_{k,t}, \boldsymbol{l}_{k,t}^{\mathrm{A}}, \boldsymbol{l}_{k',t}^{\mathrm{A}}, p_{k,t}\right) = \frac{i_{k,t} p_{k,t} D_{k,t}^{\mathrm{A}}}{W_1 \log_2\left(1 + s_{k,k',t}^{\mathrm{A}}/i_{k,t}\right)}$.

*2) UAV-BS links:* The UAV-BS links are used to transmit distance information from active and passive UAVs to the BS. Given the position $\boldsymbol{l}_{u,t}^{\mathrm{P}}$ of passive UAV $u$, $\boldsymbol{l}_{k,t}^{\mathrm{A}}$ of active UAV $k$ and the position $\boldsymbol{l}_{\mathrm{B}}$ of the BS, the probabilistic LoS and non-line-of-sight (NLoS) channel model is used to model the UAV-BS link. To be specific, the LoS path loss $g_{m,t}^{\mathrm{LoS}}$ and NLoS path loss $g_{m,t}^{\mathrm{NLoS}}$ from active or passive UAV $u$ to the BS at time slot $t$ can be given by

$$g_{u,t}^{\mathrm{LoS}} = L_{\mathrm{FS}}\left(l_0\right) + 10\mu_{\mathrm{LoS}} \log\left(\left\|\boldsymbol{l}_{u,t}^{\mathrm{P}} - \boldsymbol{l}_{\mathrm{B}}\right\|\right) + \varphi_{\sigma_{\mathrm{LoS}}}, \tag{8}$$

$$g_{u,t}^{\mathrm{NLoS}} = L_{\mathrm{FS}}\left(l_0\right) + 10\mu_{\mathrm{NLoS}} \log\left(\left\|\boldsymbol{l}_{u,t}^{\mathrm{P}} - \boldsymbol{l}_{\mathrm{B}}\right\|\right) + \varphi_{\sigma_{\mathrm{NLoS}}}, \tag{9}$$

where $L_{\mathrm{FS}}\left(l_0\right) = 20\log\left(l_0 f_0 4\pi/c\right)$ is the free-space path loss with $l_0$ being the free-space reference distance and $f_0$ being the carrier frequency. $\varphi_{\sigma_{\mathrm{LoS}}}$ and $\varphi_{\sigma_{\mathrm{NLoS}}}$ are the shadowing random variables, which are Gaussian variables in dB with zero mean and $\sigma_{\mathrm{LoS}}^2$, $\sigma_{\mathrm{NLoS}}^2$ dB variances. Given (8) and (9), the path loss from passive UAV $u$ to the BS at time slot $t$ is given by

$$\bar{g}_{u,t} = \Pr\left(g_{u,t}^{\mathrm{LoS}}\right) \times g_{u,t}^{\mathrm{LoS}} + \left(1 - \Pr\left(g_{u,t}^{\mathrm{LoS}}\right)\right) \times g_{u,t}^{\mathrm{NLoS}}, \tag{10}$$

where $\Pr\left(g_{u,t}^{\mathrm{LoS}}\right) = \left(1 + \zeta \exp\left(-\eta\left[\chi_{u,t} - \zeta\right]\right)\right)^{-1}$ is the probability of LoS with $\zeta$ and $\eta$ being constants which depend on the environment factors, and $\chi_{u,t}$ being the elevation angle of passive UAV $u$ at time slot $t$, which satisfies $\sin\left(\chi_{u,t}\right) = \frac{s_{u,t}}{\left\|\boldsymbol{l}_{u,t}^{\mathrm{P}} - \boldsymbol{l}_{\mathrm{B}}\right\|}$.

The signal-to-noise ratio (SNR) of the signal transmitted from passive UAV $u$ to the BS at time slot $t$ is given by

$$s_{u,t}^{\mathrm{G}} = \frac{p_{u,t}}{\epsilon^2} 10^{-\bar{g}_{u,t}/10}, \tag{11}$$

where $p_{u,t}$ is the transmit power of passive UAV $u$ at time slot $t$. The transmission delay from passive UAV $u$ to the BS at time slot $t$ is given by

$$T_{u,t}^{\mathrm{G}} = \frac{D_{u,t}^{\mathrm{G}}}{W_2 \log_2\left(1 + s_{u,t}^{\mathrm{G}}\right)}, \tag{12}$$

where $W_2$ is the bandwidth of each active and passive UAV to transmit distance to the BS and $D_{u,t}^{\mathrm{G}}$ is the data size of the distance information from passive UAV $u$ to the BS. Similarly, the transmit delay $T_{k,t}^{\mathrm{G}}$ from active UAV $k$ to the BS at time slot $t$ can be calculated by substituting $\boldsymbol{l}_{k,t}^{\mathrm{A}}$ for $\boldsymbol{l}_{u,t}^{\mathrm{P}}$ in (8)-(12).

### D. Localization Model

After receiving signals transmitted from active UAVs and reflected by the target UAV, active and passive UAVs estimate the distance measurement information from active UAVs to the target UAV and then from the target UAV to passive UAVs or active UAVs. Then, the measured distances, that is given by $\hat{\boldsymbol{d}}_t = \left\{\hat{d}_{1,1,t}, \cdots, \hat{d}_{k,K,t}, \hat{d}_{k,K+1,t}, \hat{d}_{K,K+U,t}\right\}$, are transmitted to the BS, where $\hat{d}_{k,u,t}$ represents distance from active UAV $k$ to passive UAV $u$. Note that, when active UAV $k$ does not transmit signals, that is $i_{k,t} = 0$, we have $\hat{d}_{k,k',t} = 0$ and $\hat{d}_{k,u,t} = 0$. Due to the mobility of UAVs and the jamming UAV, the BS requires to select a subset of distance information transmitted for locating target UAV. Let $\boldsymbol{q}_t = \left\{q_{1,t}^{\mathrm{A}}, \cdots, q_{K,t}^{\mathrm{A}}, q_{1,t}^{\mathrm{P}}, \cdots, q_{U,t}^{\mathrm{P}}\right\}$ be the select indicator vector of distance measurement information for locating the target UAV, where $q_{k,t}^{\mathrm{A}}, q_{u,t}^{\mathrm{P}} \in \{0,1\}$ with the value equals to 1 indicating distance information is selected to estimate the position of the target UAV, otherwise, we have the value is 0. Based on distance information $\hat{\boldsymbol{d}}_t$, positions $\boldsymbol{l}_t^{\mathrm{P}} = \left\{\boldsymbol{l}_{1,t}^{\mathrm{P}}, \cdots, \boldsymbol{l}_{U,t}^{\mathrm{P}}\right\}$ of passive UAVs, positions $\boldsymbol{l}_t^{\mathrm{A}} = \left\{\boldsymbol{l}_{1,t}^{\mathrm{A}}, \cdots, \boldsymbol{l}_{K,t}^{\mathrm{A}}\right\}$ of active UAVs, the SINR $\boldsymbol{s}_t = \left\{s_{1,1,t}^{\mathrm{A}}, \cdots, s_{K,K,t}^{\mathrm{A}}, s_{1,1,t}^{\mathrm{P}}, \cdots, \boldsymbol{s}_{K,U,t}^{\mathrm{P}}\right\}$ of active and passive UAVs, and the UAVs selection scheme $\boldsymbol{q}_t$, the BS estimates the position $\hat{\boldsymbol{l}}_t$ of the target UAV.

### E. Problem Formulation

After defining the system model, our goal is to accurately estimate the real-time position of the target UAV under the interference introduced by wireless channel noise and the jamming UAV. We formulate an optimization problem whose goal is to minimize the positioning error $e_t\left(\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t\right) = \sqrt{\left\|\hat{\boldsymbol{l}}_t\left(\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t\right) - \boldsymbol{l}_t\right\|^2}$ between the estimated position $\hat{\boldsymbol{l}}_t\left(\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t\right)$ and the ground truth position $\boldsymbol{l}_t$ of the target UAV over $T$ time slots by determining whether active UAVs transmit signals, the trajectories, transmit powers of active UAVs, and distance information selection scheme under the impacts of the jamming UAV. The minimization problem is

$$\min_{\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t} \sum_{t=1}^{T} e_t\left(\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t\right), \tag{13}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} q_{k,t}^{\mathrm{A}} E_{k,t} \leqslant E_{\max}^{\mathrm{F}}, \ \forall k \in \mathcal{K}, \tag{13a}$$

$$\sum_{k=1}^{K}\sum_{u=1}^{U}\sum_{k'=1}^{K} i_{k,t}\left(E_{k,u,t}^{\mathrm{T}} + E_{k,k',t}^{\mathrm{T}}\right) \leqslant E_{\max}^{\mathrm{T}}, \tag{13b}$$

$$q_{k,t}^{\mathrm{A}} T_{k,t}^{\mathrm{G}} \leqslant T_{\max}, \ \forall k \in \mathcal{K}, \tag{13c}$$

$$q_{u,t}^{\mathrm{P}} T_{u,t}^{\mathrm{G}} \leqslant T_{\max}, \ \forall u \in \mathcal{U}, \tag{13d}$$

$$\left\|\boldsymbol{l}_{u,t}^{\mathrm{P}} - \boldsymbol{l}_{k,t}^{\mathrm{A}}\right\| \geqslant L_{\min}, \forall u \in \mathcal{U}, \tag{13e}$$

$$\beta_{\min} \leqslant \beta_{k,t} \leqslant \beta_{\max}, \ \forall u \in \mathcal{U}, \tag{13f}$$

$$\alpha_{\min} \leqslant \alpha_{k,t} \leqslant \alpha_{\max}, \ \forall u \in \mathcal{U}, \tag{13g}$$

$$\sum_{k=1}^{K} i_{k,t} \geqslant 1, \forall k \in \mathcal{K}, \tag{13h}$$

$$\sum_{m=1}^{K+U} q_{m,t} = 4, \tag{13i}$$

where $\boldsymbol{i}_t = \left[i_{1,t}, \cdots, i_{K,t}\right]^T$ is the matrix represents whether active UAVs transmit signals and $\boldsymbol{p}_t = \left[p_{1,t}, \cdots, p_{K,t}\right]$ is the transmit powers of active UAVs. $E_{\max}^{\mathrm{F}}$ is the maximum active UAV propulsion energy consumption, $E_{\max}^{\mathrm{T}}$ is the maximum active UAV transmit energy consumption, $T_{\max}$ is the maximum transmission delay from each passive UAV to the BS, $L_{\min}$ is the safe distance between any two UAVs, and $P_{\max}$ is the maximum transmit power of active UAVs. Constraints (13a) and (13b) are the flight energy constraint and transmit energy constraint for active UAVs. Constraints (13c) and (13d) are the delay requirements for distance measurement information transmission from active and passive UAVs to the BS, and constraint (13e) is the safe distance requirement between any passive UAV and the target UAV. Constraints (13f) and (13g) are the movement constraint of active UAVs. Constraint (13h) represents that at least one active UAV transmits signals at each time slot and constraint (13i) is the required number of distance measurement information to locate the target UAV.

Problem (13) is difficult to solve by traditional convex algorithms due to the following reasons. First, the relationship between the estimated position of the target UAV obtained by the BS and the optimization variables in (13) cannot be accurately characterized due to the discontinuous attacks from the jamming UAV. Second, using traditional optimization algorithms to optimize the trajectories the active UAVs, the transmit powers of active UAVs, and the distance selection scheme needs the BS to calculate the error $e_t\left(\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t\right)$ of the target UAV based on the ground truth position $\boldsymbol{l}_t$ of the target UAV. However, $\boldsymbol{l}_t$ is unknown in practice. Third, the position and the jamming pattern of the jamming UAV are unknown in advance so that the BS may not select the most appropriate passive UAVs subset to avoid attacks. To this end, we investigate a RL based method to jointly optimize the optimal passive UAVs subset and select the jamming attack defense method from the GAN based anti-attack method, the wireless technique based anti-attack method or the joint GAN and wireless technique based anti-attack method to solve problem (13).

## III. RL BASED HYBRID JAMMING DEFENSE METHOD

To solve (13), we introduce a distributional RL based hybrid anti-jamming localization method that combines the GAN based anti-attack method and traditional wireless technique based anti-attack method (i.e., trajectory design, transmit power optimization, and passive UAVs selection). Compared to traditional defense methods, the proposed method uses a probability distribution based RL method to adaptively select the optimal jamming defense method from a) the GAN based jamming attack defense method, b) the wireless technique based jamming attack defense method, and c) joint the GAN and wireless technique based jamming attack defense method to improve the UAV localization performance under jamming attacks. Next, we first introduce the GAN and the wireless technique based jamming attack defense methods. Then, we introduce the use of RL to select the optimal jamming attack defense method.

### A. Introduction of Jamming Attack Defense Methods

*1) GAN based jamming attack defense method:* To accurately locate the target UAV under the attacks of the jamming

UAV, the BS can use a GAN to calculate the position of the target UAV. The input of the GAN is the received distance measurement information $\hat{\boldsymbol{d}}_t$ and positions $\boldsymbol{l}_t$ from passive UAVs. The output of the GAN is the estimated position $\hat{\boldsymbol{l}}_t = [\hat{x}_t, \hat{y}_t, \hat{z}_t]^T$ of the target UAV. However, the localization performance of the GAN method depends on the limited training data samples. Since we can collect only limited number of training data samples, GAN based method may not be applied for the locations that are not in the training dataset. Meanwhile, when the jamming signal is very strong, only using GAN based method may not be able to directly solve the jamming issues.

*2) Wireless technique based jamming attack defense method:* To avoid the impacts of jamming attacks on UAV localization, we can also use a traditional wireless technique based jamming attack defense method. Through changing the trajectory and the transmit power of the target UAV, and optimizing the passive UAVs selection scheme, the accuracy of the distance measurement and the UAV localization performance can be improved. However, this method requires to adjust the trajectory frequently based on the estimated position of the target UAV calculated by a two-stage weight least-square (TSWLS) method at the BS [15], which will increase aerodynamic power consumption.

*3) Joint GAN and wireless technique based jamming attack defense method:* Given the GAN and wireless technique based jamming attack defense methods introduced above, we can also study the joint use of these two methods simultaneously.

### *B. The Distributed RL method for Selecting the Optimal Defense Method*

Since the position and the jamming pattern of the jamming UAV are unknown in practice, we proposed to use a distributional RL method to adaptively select the optimal attack defense method by jointly considering the use of the GAN based method and wireless technique based method under jamming attacks. Next, we introduce the components and the training process of the RL method.

*1) Components of the proposed method:* The proposed method consists of the following five components:

- *Agents*: The agents are the BS and active UAVs. In particular, the BS needs to select the distance measurement information subset and determine the method (the GAN or the TSWLS) to estimate the position of the target UAV. Active UAVs needs to determine whether to transmit signals and optimize their transmit powers and trajectories.
- *States*: A state of the BS is $\boldsymbol{o}_t^{\mathrm{B}} = \left[\hat{\boldsymbol{d}}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{l}_t^{\mathrm{P}}, \boldsymbol{s}_t\right]$ that captures the distance transmitted from active and passive UAVs, the deployment of active and passive UAVs, and the SINR at each UAV. A state of active UAV $k$ is $\boldsymbol{o}_{k,t} = \left[\boldsymbol{l}_{k,t}, s_{k',k,t}^{\mathrm{A}}\right]$ that captures the position and SINR of itself at time slot $t$. Hereinafter, we use $\boldsymbol{o}_t = \left[\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{o}_{1,t}, \cdots, \boldsymbol{o}_{K,t}\right]$ to represent the global states at time slot $t$.
- *Actions*: An action of active UAV $k$ is represented by $\boldsymbol{a}_{k,t} = [i_{k,t}, p_{k,t}, \alpha_{k,t}, \beta_{k,t}]$ that determines whether to transmit signals and optimize its trajectory and the transmit power. An action of the BS is $\boldsymbol{a}_t^{\mathrm{B}} = \left[\boldsymbol{q}_t, g_t^{\mathrm{B}}\right]$, where $\boldsymbol{q}_t$ is the distance subset selection strategy and $g_t^{\mathrm{B}}$ is the position calculation method indicator with $g_t^{\mathrm{B}} = 1$ implying that the BS uses a GAN method to calculate the position of the target UAV and $g_t^{\mathrm{B}} = 0$ implying that the BS uses the TSWLS method. We define the actions of all agents at time slot $t$ as $\boldsymbol{a}_t = \left[\boldsymbol{a}_{1,t}, \cdots, \boldsymbol{a}_{K,t}, \boldsymbol{a}_t^{\mathrm{B}}\right]$.
- *Reward*: The reward of agents can be represented by $r_t(\boldsymbol{o}_t, \boldsymbol{a}_t) = -e_t\left(\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t\right)$, where $e_t\left(\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t\right)$ is the positioning error of the target UAV at time slot $t$. Since the BS calculates the estimated position $\hat{\boldsymbol{l}}_t\left(\boldsymbol{i}_t, \boldsymbol{l}_t^{\mathrm{A}}, \boldsymbol{p}_t, \boldsymbol{q}_t\right)$ of the target UAV after obtaining all distance measurement information from active and passive UAVs, the BS and active UAVs share a reward $r_t(\boldsymbol{o}_t, \boldsymbol{a}_t)$ at time slot $t$.
- *Value function*: Under a given state $\boldsymbol{o}_t$ and a selected action $\boldsymbol{a}_t$, the value function of active UAV $k$ is $v(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t}) = \sum_{t=0}^{\infty} \gamma^t r_t(\boldsymbol{o}_t, \boldsymbol{a}_t)$ with $\gamma$ being the discount factor and the value function of the BS is $v\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}\right) = \sum_{t=0}^{\infty} \gamma^t r_t(\boldsymbol{o}_t, \boldsymbol{a}_t)$. Compared with traditional RL methods that directly estimate the expected future reward to represent the value function of each agent [17], the proposed method aims to estimate the probability distribution of $v(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{u,t})$ and $v\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}\right)$. And then the proposed method calculates the expected value of $v(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t})$ and $v\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}\right)$. Next, we introduce the process of estimating the probability distribution of the value function. Each active UAV estimates the probability distribution of its value functions by using a DNN that is parameterized by $\boldsymbol{w}_k$ and the DNN of the BS is parameterized by $\boldsymbol{w}_{\mathrm{B}}$. The input of the DNN is a set of probability values $\boldsymbol{\zeta} = [\zeta_1, \cdots, \zeta_N]$ that satisfies $\zeta_i = \frac{i}{N}$. The output is the approximated value function of active UAV $k$ and the BS. To be specific, the approximated value function of active UAV $k$ can be written as $\hat{\boldsymbol{v}}(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t}, \boldsymbol{\zeta}) = [\hat{v}(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t}, \zeta_1), \cdots, \hat{v}(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t}, \zeta_N)]$, where $\hat{v}(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t}, \zeta_i)$ represents the value of $v(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t})$ with the cumulative distribution probability of $\zeta_i$. In addition, the relationship between $\hat{v}(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t}, \zeta_i)$ and $\zeta_i$ is
$$\zeta_i = \mathbb{P}\left(v(\boldsymbol{o}_{0,t}, \boldsymbol{a}_{0,t}) \leqslant \hat{v}(\boldsymbol{o}_{0,t}, \boldsymbol{a}_{0,t}, \zeta_i)\right). \tag{14}$$
And the approximated value function of the BS is $\hat{\boldsymbol{v}}\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}, \boldsymbol{\zeta}\right) = \left[\hat{v}\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}, \zeta_1\right), \cdots, \hat{v}\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}, \zeta_N\right)\right]$, where $\hat{v}\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}, \zeta_i\right)$ represents the value of $v\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}\right)$ with the cumulative distribution probability of $\zeta_i$.

*2) Training process of the proposed method:* Here, we introduce how the BS and the active UAV collaboratively perform the proposed method to minimize the positioning error of the target UAV while avoiding jamming attacks. We first introduce the loss function of the proposed method. Then, the whole training process of the proposed method is introduced.

The loss function of the proposed method is given by [18]

$$\rho\left(\boldsymbol{w}_0, \boldsymbol{w}_{\mathrm{B}}\right) = \frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \left|\zeta_i - \mathbb{1}_{\{\iota(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j)\}}\right| \frac{M\left(\iota\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j\right)\right)}{\eta}, \tag{15}$$

where $\eta$ is a coefficient and $\iota\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j\right) = r_t\left(\boldsymbol{o}_t, \boldsymbol{a}_t\right) + \gamma \hat{v}\left(\boldsymbol{o}_{t+1}, \boldsymbol{a}_{t+1}, \zeta_i\right) - \hat{v}\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_j\right)$ with $\boldsymbol{a}_{t+1} = \arg\max_{\boldsymbol{a}'} \frac{1}{N} \sum_{i=1}^{N} \left(\hat{v}\left(\boldsymbol{o}_{0,t}, \boldsymbol{a}_{0,t}, \zeta_i\right) + \hat{v}\left(\boldsymbol{o}_t^{\mathrm{B}}, a_t^{\mathrm{B}}, \zeta_i\right)\right)$. In addition, $\mathbb{1}_{\{x\}} = 0$ when $x \leqslant 0$ and $\mathbb{1}_{\{x\}} = 1$, otherwise. $M\left(\iota\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j\right)\right)$ is given by

$$M\left(\iota\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j\right)\right) = \begin{cases} \frac{1}{2}\left(\iota\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j\right)\right)^2, & \text{if } \left|\iota\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j\right)\right| \leqslant \eta, \\ \eta\left(\left|\iota\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j\right)\right| - \frac{1}{2}\eta\right), & \text{otherwise}, \end{cases} \tag{16}$$

where $M\left(\iota\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \zeta_i, \zeta_j\right)\right)$ is the Huber quantile regression loss [18].

The proposed method is collaboratively trained by the BS and the active UAV. Then, the training process can be divided into the training process at the BS and the training process at active UAVs.

- Training process at the BS: From (15), the total loss requires passive UAVs to transmit a) the distance information measured by passive UAVs and real-time positions of passive UAVs for calculating the reward $r_t\left(\boldsymbol{o}_t, \boldsymbol{a}_t\right)$, and b) value function of the active UAV to the BS for calculating value functions $v_t\left(\boldsymbol{o}_{t+1}, \boldsymbol{a}_{t+1}, \boldsymbol{\zeta}\right)$ and $v_t\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \boldsymbol{\zeta}\right)$. Based on $r_t\left(\boldsymbol{o}_t, \boldsymbol{a}_t\right)$, $v_t\left(\boldsymbol{o}_{t+1}, \boldsymbol{a}_{t+1}, \boldsymbol{\zeta}\right)$, and $v_t\left(\boldsymbol{o}_t, \boldsymbol{a}_t, \boldsymbol{\zeta}\right)$, the BS calculates the loss function $\rho\left(\boldsymbol{w}_0, \boldsymbol{w}_{\mathrm{B}}\right)$ based on (15). Then, the BS updates its DNN parameters and the process is
$$\boldsymbol{w}_{\mathrm{B}} = \boldsymbol{w}_{\mathrm{B}} + b_{\mathrm{B}} \bigtriangledown \rho\left(\boldsymbol{w}_0, \boldsymbol{w}_{\mathrm{B}}\right), \tag{17}$$
where $b_{\mathrm{B}}$ is the step size at the BS. Meanwhile, the BS transmits the value of loss function $\rho\left(\boldsymbol{w}_0, \boldsymbol{w}_{\mathrm{B}}\right)$ to the active UAV.
- Training process at each active UAV: Active UAV $k$ requires to update its DNN based on the loss function $\rho\left(\boldsymbol{w}_k, \boldsymbol{w}_{\mathrm{B}}\right)$. The update of the active UAV is given by
$$\boldsymbol{w}_k = \boldsymbol{w}_k + b_k \bigtriangledown \rho\left(\boldsymbol{w}_k, \boldsymbol{w}_{\mathrm{B}}\right), \tag{18}$$
where $b_k$ is the step size. The specific training process of the proposed method is summarized in Algorithm 1.

### C. Convergence, Implementation, and Complexity Analysis

Here, we analyze the convergence, implementation, and the complexity of training the proposed method.

## IV. Analysis of Target UAV Localization Performance

In this section, our goal is to analyze the influence of active and passive UAV 3D coordinates errors and the jamming attacks on target UAV localization performance. First, we analyze the relationship between the jamming attacks and the positioning error of the target UAV. Then, when the 3D coordinates of active and passive UAVs have errors, we derive the brief form of the relationship between the positioning errors and the coordinates errors of active and passive UAVs under the assumption that the measurement error variance matrix is a identity matrix.

The measured distance from active UAV $k$ to passive UAV $u$ can be written as

$$\hat{d}_{k,u,t} = r_{u,t} + r_{k,t}^{\mathrm{A}} + \Delta d_{k,u,t}, \tag{19}$$

where $r_{u,t} = \sqrt{\left\|\boldsymbol{l}_{u,t}^{\mathrm{P}} - \boldsymbol{l}_t\right\|^2}$ is the actual distance between the target UAV and passive UAV $u$, $r_{k,t}^{\mathrm{A}} = \sqrt{\left\|\boldsymbol{l}_{k,t}^{\mathrm{A}} - \boldsymbol{l}_t\right\|^2}$ is the actual distance between active UAV $k$ to the target UAV, and $\Delta d_{k,u,t}$ is the measurement error of $\hat{d}\left(k,u,t\right)$ caused by the jamming attacks and channel noise. $\Delta d_{k,u,t}$ follows the Gaussian distribution with a mean of zero and a variance dependent on $s_{k,u,t}^{\mathrm{P}}\left(i_{k,t}, \boldsymbol{l}_{k,t}^{\mathrm{A}}, p_{k,t}\right)$. Take the derivative of both

**Algorithm 1** Proposed Method for Solving Problem (13)

1: Initialize the DNN parameters $\boldsymbol{w}_0$ and $\boldsymbol{w}_{\mathrm{B}}$, and a set $\boldsymbol{\zeta}$.
2: **for** each iteration **do**
3. **for** each time slot $t$ **do**
5. Observe the observation $\boldsymbol{o}_{k,t}$ and $\boldsymbol{o}_t^{\mathrm{B}}$.
6: Select an action according to a $\epsilon$-greedy scheme.
7: Calculate value function values $\hat{\boldsymbol{v}}\left(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t}, \boldsymbol{\zeta}\right)$ and $\hat{\boldsymbol{v}}\left(\boldsymbol{o}_t^{\mathrm{B}}, \boldsymbol{a}_t^{\mathrm{B}}, \boldsymbol{\zeta}\right)$ at time slot $t$ and $t+1$.
8: **end for**
9: Active UAV $k$ transmits $\boldsymbol{o}_{k,t}$, $\hat{\boldsymbol{v}}_{k,t}\left(\boldsymbol{o}_{k,t}, \boldsymbol{a}_{k,t}, \boldsymbol{\zeta}\right)$, and $\hat{\boldsymbol{v}}\left(\boldsymbol{o}_{k,t+1}, \boldsymbol{a}_{k,t+1}, \boldsymbol{\zeta}\right)$ to the BS.
10: **end for**
11: The BS calculates the value of loss function and transmits it to the active UAV.
12: **for** each agent $u$ **do**
13: Update $\boldsymbol{w}_k$ and $\boldsymbol{w}_{\mathrm{B}}$ based on (17) and (18).
14: **end for**
15: **end for**

TABLE II. PARAMETERS

| Parameters | Values | Parameters | Values |
|---|---|---|---|
| $\epsilon^2$ | -95 dBm | $\Delta_t$ | 1 s |
| $v_{u,t}^{\mathrm{H}}$ | 9.43 m/s | $p_{u,t}$ | 5 W |
| $C_1$ | 4929 | $W$ | 1 MHz |
| $C_2$ | 0.002 | $M$ | 4 kg |
| $\sigma_{\mathrm{LoS}}^2$ | 8.41 | $\sigma_{\mathrm{NLoS}}^2$ | 33.78 |
| $E_{\max}$ | 500 J | $\xi$ | 1 s |
| $L_{\min}$ | 80 m | $L_{\max}$ | 10 km |
| $\beta_{\min}$ | $-15^o$ | $\beta_{\max}$ | $15^o$ |
| $\alpha_{\min}$ | $-15^o$ | $\alpha_{\max}$ | $15^o$ |
| $D_{\mathrm{A}}$ | 5 bit | $D_{\mathrm{B}}$ | 5 bit |
| $X$ | 11.9 | $Y$ | 0.13 |
| $T$ | 10 | $f_{\mathrm{J}}$ | 0.5 |
| $\mu_{\mathrm{LoS}}^B$ | 2 | $\mu_{\mathrm{NLoS}}^B$ | 2.4 |

sides of (19) with respect to $\boldsymbol{l}_t$, we have

$$\partial \hat{d}_{k,u,t} = \left(\frac{x_t - x^{\mathrm{P}}_{u,t}}{r_{u,t}} + \frac{x_t - x^{\mathrm{A}}_{k,t}}{r^{\mathrm{A}}_{k,t}}\right)\partial x_t + \left(\frac{y_t - y^{\mathrm{P}}_{u,t}}{r_{u,t}} + \frac{y_t - y^{\mathrm{A}}_{k,t}}{r^{\mathrm{A}}_{k,t}}\right)\partial y_t + \left(\frac{z_t - z^{\mathrm{P}}_{u,t}}{r_{u,t}} + \frac{z_t - z^{\mathrm{A}}_{k,t}}{r^{\mathrm{A}}_{k,t}}\right)\partial z_t. \tag{20}$$

Based on four distance measurement information selected by the BS, the position of the target UAV can be estimated. The selected distance subset for calculating the position of the target UAV consists of four measured distance values which are obtained by active and passive UAVs. We denote the selected distance subset as $\hat{\boldsymbol{d}}^{\mathrm{S}}_t = \left[\hat{d}_{k_1,m_1,t}, \hat{d}_{k_2,m_2,t}, \hat{d}_{k_3,m_3,t}, \hat{d}_{k_4,m_4,t}\right]^T \in \hat{\boldsymbol{d}}_t$, where $k_1,k_2,k_3,k_4 \in \mathcal{K}$ are the active UAVs and $m_1,m_2,m_3,m_4 \in \mathcal{K}\cup\mathcal{U}$ are the active or passive UAVs. Based on (20), the relationship between the selected distance subset and the estimated target UAV position can be given by

$$\boldsymbol{\partial}\hat{\boldsymbol{d}}^{\mathrm{S}}_t = \boldsymbol{M}\boldsymbol{\partial}\boldsymbol{l}_t, \tag{21}$$

where $\boldsymbol{\partial}\hat{\boldsymbol{d}}^{\mathrm{S}}_t = \left[\partial\hat{d}_{k_1,m_1,t}, \partial\hat{d}_{k_2,m_2,t}, \partial\hat{d}_{k_3,m_3,t}, \partial\hat{d}_{k_4,m_4,t}\right]^T$, $\boldsymbol{\partial}\hat{\boldsymbol{l}}_t = [\partial x_t, \partial y_t, \partial z_t]^T$, and $\boldsymbol{M}$ is given by (22).

From (21), we have

$$\boldsymbol{\partial}\boldsymbol{l}_t = \left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{\partial}\hat{\boldsymbol{d}}^{\mathrm{S}}_t, \tag{23}$$

where $(\cdot)^{-1}$ is the inverse matrix. The positioning error of the target UAV is $e_t = \sqrt{(\partial x_t)^2 + (\partial y_t)^2 + (\partial z_t)^2}$, then we have $e_t = \mathrm{tr}\left(\mathbb{E}\left[\boldsymbol{\partial}\boldsymbol{l}_t\left(\boldsymbol{\partial}\boldsymbol{l}_t\right)^T\right]\right)$, where $\mathrm{tr}(\cdot)$ is the trace of the matrix $\mathbb{E}\left[\boldsymbol{\partial}\boldsymbol{l}_t\left(\boldsymbol{\partial}\boldsymbol{l}_t\right)^T\right]$. Then the relationship between of the

**Proposition 1**. The relationship between the jamming attacks and the positioning error of the target UAV is

$$e_t = \left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{J}\boldsymbol{M}\left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}, \tag{24}$$

where $\boldsymbol{J}$ is the distance measurement variance matrix and $\boldsymbol{J} = \mathrm{diag}\left(k^1_1 j_t p^{\mathrm{J}}\left|h_{\mathrm{J},m_1,t}\right|^2 + k^2_1, \cdots, k^1_4 j_t p^{\mathrm{J}}\left|h_{\mathrm{J},m_4,t}\right|^2 + k^2_4\right)$ with $k^1_i, k^2_i$ are coefficients in [19].

*Proof:* See Appendix A. ☐

From Proposition 1, we can see that the positioning error of the target UAV relies on $\boldsymbol{J}$ determined by the jamming model of the jamming UAV and $\boldsymbol{M}$ determined by the position of active and passive UAVs.

Next, we analyze the influence of the 3D coordinates errors of active and passive UAVs on the target UAV localization performance. The position of active UAV $k$ with errors can be represented by $\tilde{\boldsymbol{l}}^{\mathrm{A}}_{k,t} = \left[\tilde{x}^{\mathrm{A}}_{k,t}, \tilde{y}^{\mathrm{A}}_{k,t}, \tilde{z}^{\mathrm{A}}_{k,t}\right]^T$, where $\tilde{x}^{\mathrm{A}}_{k,t} = x^{\mathrm{A}}_{k,t} + \Delta x^{\mathrm{A}}_{k,t}$, $\tilde{y}^{\mathrm{A}}_{k,t} = y^{\mathrm{A}}_{k,t} + \Delta y^{\mathrm{A}}_{k,t}$, and $\tilde{z}^{\mathrm{A}}_{k,t} = z^{\mathrm{A}}_{k,t} + \Delta z^{\mathrm{A}}_{k,t}$ with $\Delta x^{\mathrm{A}}_{k,t}$, $\Delta y^{\mathrm{A}}_{k,t}$, and $\Delta z^{\mathrm{A}}_{k,t}$ being the 3D coordinates errors of active UAV $k$. Similarly,the position of passive UAV $m$ with errors can be represented by $\tilde{\boldsymbol{l}}^{\mathrm{P}}_{u,t} = \left[\tilde{x}^{\mathrm{P}}_{u,t}, \tilde{y}^{\mathrm{P}}_{u,t}, \tilde{z}^{\mathrm{A}}_{u,t}\right]^T$. Substituting the position with errors $\tilde{\boldsymbol{l}}^{\mathrm{A}}_{k,t}$ and $\tilde{\boldsymbol{l}}_{k,t}$ into (21), (21) can be rewritten as $\boldsymbol{\partial}\hat{\boldsymbol{d}}^{\mathrm{S}}_t = \widetilde{\boldsymbol{M}}\boldsymbol{\partial}\boldsymbol{l}_t$, where $\widetilde{\boldsymbol{M}} = \boldsymbol{M} + \Delta\boldsymbol{M}$ with the 3D coordinates of active and passive UAVs in $\boldsymbol{M}$ being accurate and the coordinates in $\widetilde{\boldsymbol{M}}$ being with errors, and $\Delta\boldsymbol{M}$ being

$$\Delta\boldsymbol{M} = \begin{bmatrix} \frac{\Delta x_{m_1,t}}{r_{m_1}} + \frac{\Delta x_{k_1,t}}{r_{k_1}} & \frac{\Delta y_{m_1,t}}{r_{m_1}} + \frac{\Delta y_{k_1,t}}{r_{k_1}} & \frac{\Delta z_{m_1,t}}{r_{m_1}} + \frac{\Delta z_{k_1,t}}{r_{k_1}} \\ \frac{\Delta x_{m_2,t}}{r_{m_2}} + \frac{\Delta x_{k_2,t}}{r_{k_2}} & \frac{\Delta y_{m_2,t}}{r_{m_2}} + \frac{\Delta y_{k_2,t}}{r_{k_2}} & \frac{\Delta z_{m_2,t}}{r_{m_3}} + \frac{\Delta z_{k_2,t}}{r_{k_3}} \\ \frac{\Delta x_{m_3,t}}{r_{m_3}} + \frac{\Delta x_{k_3,t}}{r_{k_3}} & \frac{\Delta y_{m_3,t}}{r_{m_3}} + \frac{\Delta y_{k_3,t}}{r_{k_3}} & \frac{\Delta z_{m_3,t}}{r_{m_3}} + \frac{\Delta z_{k_3,t}}{r_{k_3}} \\ \frac{\Delta x_{m_4,t}}{r_{m_4}} + \frac{\Delta x_{k_4,t}}{r_{k_4}} & \frac{\Delta y_{m_4,t}}{r_{m_4}} + \frac{\Delta y_{k_4,t}}{r_{k_4}} & \frac{\Delta z_{m_4,t}}{r_{m_4}} + \frac{\Delta z_{k_4,t}}{r_{k_4}} \end{bmatrix}. \tag{25}$$

Based on Proposition 1, we can also derive the positioning error of the target UAV when the measurement variance matrix $\boldsymbol{J}$ of distance information obtained by each active and passive UAV are the same, which is shown in the following Lemma 1.

**Lemma 1**. When the distance measurement variances of the selected distance subset are the same as a constant $k$, the positioning error can be given by

$$e_t = k\left(\left(\boldsymbol{M}^T - (\Delta\boldsymbol{M})^T\right)(\boldsymbol{M} - \Delta\boldsymbol{M})\right)^{-1}, \tag{26}$$

*Proof:* See Appendix B. ☐

From Lemma 1, we see that when the jamming model are given and the distance measurement variances matrix are a identity matrix, the positioning error depends on the 3D coordinates' errors of active and passive UAVs.

## V. Simulation Results and Analysis

For simulation, we consider that the jamming UAV and five controllable UAVs are randomly distributed in a 3D space. The system parameters of the simulations are listed in Table I. For comparison purpose, we consider three baseline methods: a)

$$\boldsymbol{M} = \begin{bmatrix} \frac{x_t - x_{m_1,t}}{r_{m_1,t}} + \frac{x_t - x^{\mathrm{A}}_{k_1,t}}{r^{\mathrm{A}}_{k_1,t}} & \frac{y_t - y_{m_1,t}}{r_{m_1,t}} + \frac{y_t - y^{\mathrm{A}}_{k_1,t}}{r^{\mathrm{A}}_{k_1,t}} & \frac{z_t - z_{m_1,t}}{r_{m_1,t}} + \frac{z_t - z^{\mathrm{A}}_{k_1,t}}{r^{\mathrm{A}}_{k_1,t}} \\ \frac{x_t - x_{m_2,t}}{r_{m_2,t}} + \frac{x_t - x^{\mathrm{A}}_{k_2,t}}{r^{\mathrm{A}}_{k_2,t}} & \frac{y_t - y_{m_2,t}}{r_{m_2,t}} + \frac{y_t - y^{\mathrm{A}}_{k_2,t}}{r^{\mathrm{A}}_{k_2,t}} & \frac{z_t - z_{m_2,t}}{r_{m_2,t}} + \frac{z_t - z^{\mathrm{A}}_{k_2,t}}{r^{\mathrm{A}}_{k_2,t}} \\ \frac{x_t - x_{m_3,t}}{r_{m_3,t}} + \frac{x_t - x^{\mathrm{A}}_{k_3,t}}{r^{\mathrm{A}}_{k_3,t}} & \frac{y_t - y_{m_3,t}}{r_{m_3,t}} + \frac{y_t - y^{\mathrm{A}}_{k_3,t}}{r^{\mathrm{A}}_{k_3,t}} & \frac{z_t - z_{m_3,t}}{r_{m_3,t}} + \frac{z_t - z^{\mathrm{A}}_{k_3,t}}{r^{\mathrm{A}}_{k_3,t}} \\ \frac{x_t - x_{m_4,t}}{r_{m_4,t}} + \frac{x_t - x^{\mathrm{A}}_{k_4,t}}{r^{\mathrm{A}}_{k_4,t}} & \frac{y_t - y_{m_4,t}}{r_{m_4,t}} + \frac{y_t - y^{\mathrm{A}}_{k_4,t}}{r^{\mathrm{A}}_{k_4,t}} & \frac{z_t - z_{m_4,t}}{r_{m_4,t}} + \frac{z_t - z^{\mathrm{A}}_{k_4,t}}{r^{\mathrm{A}}_{k_4,t}} \end{bmatrix}. \tag{22}$$

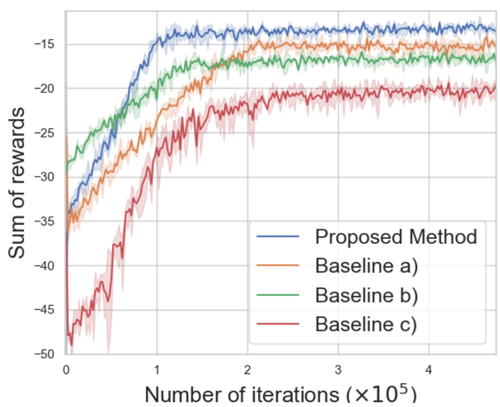

Fig. 2. The convergence of the proposed method.

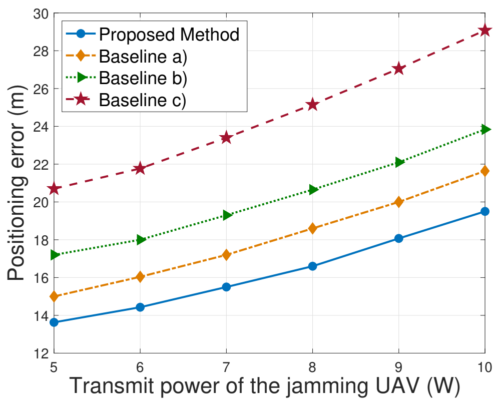

Fig. 3. Value of the positioning error as the jamming power.

an algorithm that selects the optimal jamming attack defense method by using a value decomposition based deep Q network [17], b) an algorithm that uses the proposed RL method while considering the joint GAN and wireless technique based jamming attack defense method to avoid attacks, and c) an algorithm that uses the proposed RL for trajectory and transmit power optimization, and the passive UAVs selection, which does not consider the GAN based jamming attack defense method.

In Fig. 2, we show the convergence of the proposed method and three baseline methods. From this figure, we see that the proposed method can achieve up to 9.1%, 23.1%, and 33.8% gains in terms of the sum of rewards compared to the baselines a), b), and c). The 9.1% gain stems from the fact that the proposed RL uses value probability distribution to estimate the expected value and thus selecting a better defense method. The 23.1% and 33.8% gains stem from the fact that baselines use only one jamming attack method (GAN or trajectory design) but the proposed method dynamically changes the jamming attack defense method according to the positions of UAVs and jamming attack patterns.

Fig. 3 shows how the positioning error of the target UAV changes as the transmit power of the jamming UAV changes. From Fig. 3, we see that as the transmit power of the jamming UAV increases, the positioning error of the target UAV obtained by all methods increase. This is due to the fact that, as the jamming power increases, the SINR of received signals at passive UAVs increases. Fig. 3 also shows that compared to baselines a), b), and c), the proposed method can achieve up to 9.9%, 18.2%, and 32.9% gains in terms of the positioning error of the target UAV with the jamming power to be 10 W. These gains stem from the fact that the proposed method can select the optimal jamming attack defense method according to strength of jamming signals while baselines b) and c) use fixed jamming attack defense methods.

## VI. Conclusion

In this paper, we have proposed a novel framework that enables an active UAV and a BS cooperatively locate the target UAV under attacks from a jamming UAV. We have formulated an optimization problem whose goal is to minimize the positioning error of the target UAV while avoiding jamming attacks. To solve this problem, we proposed a GAN based jamming attack defense method and a wireless technique based jamming attack defense method. Due to the unknown jamming pattern and position of the jamming UAV, we introduced a novel distributed RL method to adaptively select the optimal jamming attack defense method from a) the GAN based method, b) the wireless technique based method, and c) the joint GAN and wireless technique based method to improve the UAV localization accuracy while avoiding jamming attacks. Compared with baseline methods, the proposed method can significantly reduce the positioning error of the target UAV.

## Appendix

### A. *Proof of Proposition 1*

Based on (23), $e_t$ can be rewritten as

$$
\begin{aligned}
e_t &= \operatorname{tr}\left(\mathbb{E}\left[\boldsymbol{\partial l}_t\left(\boldsymbol{\partial l}_t\right)^T\right]\right)\\
&= \operatorname{tr}\left(\mathbb{E}\left[\left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{\partial}\hat{\boldsymbol{d}}_t^{\mathrm{S}}\left(\left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{\partial}\hat{\boldsymbol{d}}_t^{\mathrm{S}}\right)^T\right]\right)\\
&= \left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\operatorname{tr}\left(\mathbb{E}\left[\boldsymbol{\partial}\hat{\boldsymbol{d}}_t^{\mathrm{S}}\boldsymbol{\partial}\hat{\boldsymbol{d}}_t^{\mathrm{S}T}\right]\right)\left(\left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\right)^T\\
&= \left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{J}\left(\left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\right)^T\\
&\overset{(a)}{=} \left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{J}\left(\boldsymbol{M}^T\right)^T\left(\left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\right)^T\\
&\overset{(a)}{=} \left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\boldsymbol{J}\boldsymbol{M}\left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1},
\end{aligned}
\tag{27}
$$

where equation (a) stems from the fact that $\left(\boldsymbol{M}^T\right)^T = \boldsymbol{M}$ and $\left(\left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\right)^T = \left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}$. $\boldsymbol{J} =$

$\operatorname{tr}\left(\mathbb{E}\left[\partial\hat{\boldsymbol{d}}_t^{\mathrm{S}}\left(\partial\hat{\boldsymbol{d}}_t^{\mathrm{S}}\right)^T\right]\right)$ is the variance matrix of the selected distance subset and we have

$$\begin{aligned}&\operatorname{tr}\left(\mathbb{E}\left[\partial\hat{\boldsymbol{d}}_t^{\mathrm{S}}\left(\partial\hat{\boldsymbol{d}}_t^{\mathrm{S}}\right)^T\right]\right)\\&=\begin{bmatrix}\sigma_{k_1,m_1}^2 & 0 & 0 & 0\\ 0 & \sigma_{k_2,m_2}^2 & 0 & 0\\ 0 & 0 & \sigma_{k_3,m_3}^2 & 0\\ 0 & 0 & 0 & \sigma_{k_4,m_4}^2\end{bmatrix}\end{aligned}$$

with $\sigma_{k_i,m_i}^2, i=1,\cdots,4$ being the variance of the distance measurement $\hat{d}_{k_i,m_i,t}$. Based on [19], the variance $\sigma_{k_i,m_i}^2, i=1$ depends on the SINR of the signals transmitted from active UAV $k_i$ and received by active or passive UAV $m_i$, then we have $\sigma_{k_i,m_i}^2 = k_i^1\left(j_t p^{\mathrm{J}}\left|h_{\mathrm{J},m_i,t}\right|^2\right)+k_i^2$ where $k_i^1, k_i^2$ are coefficients in [19]. This completes the proof.

### B. Proof of Lemma 1

When the measurement error variance are the same as $k$, we have $\boldsymbol{J}=k\boldsymbol{I}$ where $\boldsymbol{I}=\operatorname{diag}(1,1,1,1)$ is the identity matrix. In addition, when there are errors in the 3D coordinates of active and passive UAVs, (27) can be rewritten as

$$\begin{aligned}e_t&=\left(\widetilde{\boldsymbol{M}}^T\widetilde{\boldsymbol{M}}\right)^{-1}\widetilde{\boldsymbol{M}}^T k\boldsymbol{I}\widetilde{\boldsymbol{M}}\left(\widetilde{\boldsymbol{M}}^T\widetilde{\boldsymbol{M}}\right)^{-1}\\&=k\left(\widetilde{\boldsymbol{M}}^T\widetilde{\boldsymbol{M}}\right)^{-1}\widetilde{\boldsymbol{M}}^T\widetilde{\boldsymbol{M}}\left(\widetilde{\boldsymbol{M}}^T\widetilde{\boldsymbol{M}}\right)^{-1}\\&\overset{(a)}{=}k\left(\widetilde{\boldsymbol{M}}^T\widetilde{\boldsymbol{M}}\right)^{-1},\end{aligned}\tag{28}$$

where equation (a) is because the fact that $\left(\widetilde{\boldsymbol{M}}^T\widetilde{\boldsymbol{M}}\right)^{-1}\left(\widetilde{\boldsymbol{M}}^T\widetilde{\boldsymbol{M}}\right)=\boldsymbol{I}$. Substitute $\widetilde{\boldsymbol{M}}=\boldsymbol{M}+\Delta\boldsymbol{M}$ into (28), we have $e_t=k\left(\left(\boldsymbol{M}^T-(\Delta\boldsymbol{M})^T\right)(\boldsymbol{M}-\Delta\boldsymbol{M})\right)^{-1}$. This completes the proof.

## References


[1] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2334–2360, Mar. 2019.

[2] J. Shen, A. F. Molisch, and J. Salmi, "Accurate passive location estimation using TOA measurements," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2182–2192, 2012.

[3] P. Sinha and I. Guvenc, "Impact of antenna pattern on TOA based 3D UAV localization using a terrestrial sensor network," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 7, pp. 7703–7718, Apr. 2022.

[4] U. Bhattacherjee, E. Ozturk, O. Ozdemir, I. Guvenc, M. L. Sichitiu, and H. Dai, "Experimental study of outdoor UAV localization and tracking using passive RF sensing," *https://arxiv.org/abs/2108.07857*, Aug. 2021.

[5] M. T. Dabiri, M. Rezaee, L. Mohammadi, F. Javaherian, V. Yazdanian, M. O. Hasna, and M. Uysal, "Modulating retroreflector based free space optical link for UAV-to-ground communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 10, pp. 8631–8645, Apr. 2022.

[6] B. R. Stojkoska, J. Palikrushev, K. Trivodaliev, and S. Kalajdziski, "Indoor localization of unmanned aerial vehicles based on RSSI," *IEEE EUROCON 2017 -17th International Conference on Smart Technologies*, pp. 120–125, Aug. 2017.

[7] A. N. Bishop, B. Fidan, B. D. Anderson, K. Dogancay, and P. N. Pathirana, "Optimality analysis of sensor-target geometries in passive localization: Part 1 - bearing-only localization," in *Proc. 2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information*, pp. 7–12, Melbourne, VIC, Australia, Apr. 2007.

[8] M. Sadeghi, F. Behnia, and R. Amiri, "Optimal geometry analysis for TDOA-based localization under communication constraints," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 3096–3106, Mar. 2021.

[9] Y. Zhao, Z. Li, B. Hao, P. Wan, and L. Wang, "How to select the best sensors for TDOA and TDOA/AOA localization?" *China Communications*, vol. 16, no. 2, pp. 134–145, Feb. 2019.

[10] R. Chen, B. Yang, and W. Zhang, "Distributed and collaborative localization for swarming UAVs," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 5062–5074, Mar. 2021.

[11] I. A. Meer, M. Ozger, and C. Cavdar, "On the localization of unmanned aerial vehicles with cellular networks," in *Proc. of 2020 IEEE Wireless Communications and Networking Conference (WCNC)*, Seoul, Korea (South), May 2020, pp. 1–6.

[12] K. Gao, H. Wang, H. Lv, and P. Gao, "A DL-based high-precision positioning method in challenging urban scenarios for B5G CCUAVs," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 6, pp. 1670–1687, June 2023.

[13] R. Akter, M. Golam, V.-S. Doan, J.-M. Lee, and D.-S. Kim, "IoMT-Net: Blockchain-integrated unauthorized UAV localization using lightweight convolution neural network for internet of military things," *IEEE Internet of Things Journal*, vol. 10, no. 8, pp. 6634–6651, Apr. 2023.

[14] H. Luo, T. Chen, X. Li, S. Li, C. Zhang, G. Zhao, and X. Liu, "Keepedge: A knowledge distillation empowered edge intelligence framework for visual assisted positioning in UAV delivery," *IEEE Transactions on Mobile Computing*, vol. 22, no. 8, pp. 4729–4741, Aug. 2023.

[15] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.

[16] Y. Sun, D. Xu, D. W. K. Ng, L. Dai, and R. Schober, "Optimal 3D-trajectory design and resource allocation for solar-powered UAV communication systems," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4281–4298, Feb. 2019.

[17] M. Chen, Y. Wang, and H. V. Poor, "Performance optimization for wireless semantic communications over energy harvesting networks," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8647–8651, Singapore, May 2022.

[18] W.-F. Sun, C.-K. Lee, and C.-Y. Lee, "DFAC framework: Factorizing the value function via quantile mixture for multi-agent distributional Q-learning," in *Proc. International Conference on Machine Learning*, vol. 139, pp. 9945–9954, Dec. 2021.

[19] A. Quazi, "An overview on the time delay estimate in active and passive systems for target localization," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 3, pp. 527–533, June 1981.