Device-Free 3D Drone Localization in RIS-Assisted mmWave MIMO Networks

Jiguang He[1], Charles Vanwynsberghe[1], Hui Chen[2], Chongwen Huang[3], and Aymen Fakhreddine[1]

[1]Technology Innovation Institute, 9639 Masdar City, Abu Dhabi, United Arab Emirates

[2]Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden

[3]College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China

E-mails: {jiguang.he, charles.vanwynsberghe, aymen.fakhreddine}@tii.ae, hui.chen@chalmers.se, chongwenhuang@zju.edu.cn

Abstract—In this paper, we investigate the potential of reconfigurable intelligent surfaces (RISs) in facilitating passive/device-free three-dimensional (3D) drone localization within existing cellular infrastructure operating at millimeter-wave (mmWave) frequencies and employing multiple antennas at the transceivers. The developed localization system operates in the bi-static mode without requiring direct communication between the drone and the base station. We analyze the theoretical performance limits via Fisher information analysis and Cramér Rao lower bounds (CRLBs). Furthermore, we develop a low-complexity yet effective drone localization algorithm based on coordinate gradient descent and examine the impact of factors such as radar cross section (RCS) of the drone and training overhead on system performance. It is demonstrated that integrating RIS yields significant benefits over its RIS-free counterpart, as evidenced by both theoretical analyses and numerical simulations.

Index Terms—Cramér Rao lower bound (CRLB), millimeter wave (mmWave), passive 3D drone localization, reconfigurable intelligent surface (RIS).

I. INTRODUCTION

Radio localization presents various formats depending on the interaction between the target and the base station (BS). When the target actively communicates with the BS, it constitutes active localization, commonly utilized in fifth generation (5G) and ultra-wideband (UWB) localization [1]. Conversely, passive localization approaches, prevalent in radar sensing applications, come into play when the target does not communicate with the radar or BS. In addition, localization systems can be further categorized into mono-static or bi-static mode based on the collocation of the emitter/transmitter and receiver of pilot signals.

In the existing cellular systems, radar sensing is gaining tremendous momentum owing to the emergence of integrated communication and sensing (ISAC) paradigms [2]. This innovation enables cellular BSs, traditionally focusing on communications, to incorporate sensing functionalities such as drone detection, localization, and tracking [3]. In our latest work [4], we investigated drone detection leveraging reconfigurable intelligent surfaces (RISs) in millimeter-wave (mmWave) multiple-input multiple-output (MIMO) networks and examined the impact of various system parameters, including radar cross section (RCS) [5], on detection probability. Additional works in this domain can be found in [6], [7].

In this paper, we further extend our previous work to device-free three-dimensional (3D) drone localization, leveraging an RIS to enhance its performance within a bi-static setup. We develop a practical localization algorithm with low computational complexity based on coordinate gradient descent (CGD) while achieving performance close to the theoretical limits characterized by Fisher information analysis and the Cramér Rao lower bound (CRLB). Simulation results verify the feasibility of passive drone localization within cellular mmWave MIMO networks assisted by RISs, showcasing the advantages brought through the introduction of RISs.

Notations: A bold lowercase letter $\mathbf{a}$ denotes a vector, and a bold capital letter $\mathbf{A}$ denotes a matrix. The operators $(\cdot)^*$, $(\cdot)^\mathsf{T}$, $(\cdot)^\mathsf{H}$, and $(\cdot)^{-1}$ denote the matrix or vector conjugate, transpose, Hermitian transpose, and inverse, respectively. $[\mathbf{a}]_m$, $[\mathbf{A}]_{mn}$, and $[\mathbf{A}]_{:,m:n}$ denote the mth element of $\mathbf{a}$, the (m, n)th element of $\mathbf{A}$, and the submatrix of $\mathbf{A}$ formed by its elements in the columns $m, m + 1, \ldots, n$. $\mathrm{diag}(\mathbf{a})$ denotes a square diagonal matrix with the entries of $\mathbf{a}$ on its diagonal, $\mathrm{vec}(\mathbf{A})$ denotes the vectorization of $\mathbf{A}$ by stacking the columns of the matrix $\mathbf{A}$ on top of one another, $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$, $\mathrm{rank}(\mathbf{A})$ returns the rank of matrix $\mathbf{A}$, $\mathrm{Tr}(\mathbf{A})$ denotes the trace of $\mathbf{A}$, $\mathbb{E}[\cdot]$ is the expectation operator, $\mathbf{0}$ denotes the all-zero vector, $\mathbf{I}_M$ denotes the $M \times M$ identity matrix, and $j = \sqrt{-1}$. $\|\cdot\|_\mathrm{F}$ denotes the Frobenius norm of a matrix, $|\cdot|$ returns the absolute value of a complex number, and $\|\cdot\|_2$ denotes the Euclidean norm of a vector. Finally, $\Re\{\cdot\}$ returns the real part of its complex argument.

II. SYSTEM MODEL

Fig. 1 showcases a novel device-free drone localization system deployed on the roof of a high-rise building. It comprises a mmWave BS, a passive RIS without any radio frequency (RF) chains and baseband processing units, and a receiving user equipment (UE), all possessing a uniform planar array (UPA) structure with multiple antennas or elements. This configuration resembles a conventional bi-static radar system with spatial separation of the transmitter and receiver, but harnesses augmented capabilities through the introduction of an RIS. All the three network nodes serve as positioning reference nodes, a.k.a. anchors. The BS array aligns with the y-z plane, while the RIS and UE arrays are oriented
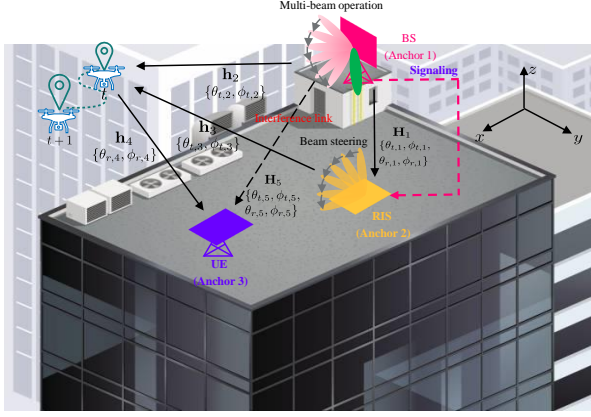
Fig. 1: Device-free drone localization system consists of a mmWave BS, an RIS, and a receiving UE, deployed on the roof of a high-rise building.

parallel to the x-y plane, facing towards the sky. To improve sounding efficiency, the BS employs a multi-beam approach with one fixed beam towards the RIS and one time-varying beam towards the sky. The primary goal of such a system is to localize a flying drone with a relatively low RCS, without requiring information exchange between the drone and the BS.

A. mmWave Channel Model

In contrast to sub-6GHz wireless channels, mmWave channels exhibit poor-scattering property. Furthermore, since we consider an open space scenario, we focus solely on the line-of-sight (LoS) path when modelling the mmWave channels. Intuitively, the RIS should be placed in the proximity of the BS, with the potential to reduce the path loss. The channel between the BS and the RIS, denoted as $\mathbf{H}_1 \in \mathbb{C}^{M_R \times M_B}$, where M_R and M_B represent the number of RIS elements and BS antennas respectively, can be effectively characterized using the Saleh-Valenzuela channel model [8], a widely adopted approach, as

$$\mathbf{H}_1 = \frac{e^{-j2\pi d_1/\lambda}}{\sqrt{\rho_1}} \boldsymbol{\alpha}_x(\theta_{r,1}, \phi_{r,1}) \otimes \boldsymbol{\alpha}_y(\theta_{r,1}, \phi_{r,1})$$
$$\left(\boldsymbol{\alpha}_y(\theta_{t,1}, \phi_{t,1}) \otimes \boldsymbol{\alpha}_z(\phi_{t,1})\right)^{\mathsf{H}}, \quad (1)$$

where ρ_1 represents the path loss, and d_1 denotes the distance between the BS and RIS. The term λ corresponds to the wavelength of the carrier frequency, while $\theta_{r,1}$ ($= \theta_{t,1}$) and $\phi_{r,1}$ ($= \phi_{t,1}$) are the azimuth and elevation angles of arrival (departure) associated with $\mathbf{H}_1$, respectively. Here, the subscripts t and r denote the transmitter and receiver, respectively. To be specific, the array response vectors $\boldsymbol{\alpha}_x(\theta_{r,1}, \phi_{r,1})$, $\boldsymbol{\alpha}_y(\theta_{r,1}, \phi_{r,1})$, $\boldsymbol{\alpha}_y(\theta_{t,1}, \phi_{t,1})$, and $\boldsymbol{\alpha}_z(\phi_{t,1})$ can be written as [9]

$$\boldsymbol{\alpha}_x(\theta_{r,1}, \phi_{r,1}) = \left[e^{-j\frac{2\pi d_{R,x}}{\lambda}\left(\frac{M_{R,x}-1}{2}\right)\cos(\theta_{r,1})\sin(\phi_{r,1})}, \right.$$
$$\left. \cdots, e^{j\frac{2\pi d_{R,x}}{\lambda}\left(\frac{M_{R,x}-1}{2}\right)\cos(\theta_{r,1})\sin(\phi_{r,1})} \right]^{\mathsf{T}}, \quad (2)$$

$$\boldsymbol{\alpha}_y(\theta_{r,1}, \phi_{r,1}) = \left[e^{-j\frac{2\pi d_{R,y}}{\lambda}\left(\frac{M_{R,y}-1}{2}\right)\sin(\theta_{r,1})\sin(\phi_{r,1})}, \right.$$
$$\left. \cdots, e^{j\frac{2\pi d_{R,y}}{\lambda}\left(\frac{M_{R,y}-1}{2}\right)\sin(\theta_{r,1})\sin(\phi_{r,1})} \right]^{\mathsf{T}}, \quad (3)$$

$$\boldsymbol{\alpha}_y(\theta_{t,1}, \phi_{t,1}) = \left[e^{-j\frac{2\pi d_{B,y}}{\lambda}\left(\frac{M_{B,y}-1}{2}\right)\sin(\theta_{t,1})\sin(\phi_{t,1})}, \right.$$
$$\left. \cdots, e^{j\frac{2\pi d_{B,y}}{\lambda}\left(\frac{M_{B,y}-1}{2}\right)\sin(\theta_{t,1})\sin(\phi_{t,1})} \right]^{\mathsf{T}}, \quad (4)$$

$$\boldsymbol{\alpha}_z(\phi_{t,1}) = \left[e^{-j\frac{2\pi d_{B,z}}{\lambda}\left(\frac{M_{B,z}-1}{2}\right)\cos(\phi_{t,1})}, \right.$$
$$\left. \cdots, e^{j\frac{2\pi d_{B,z}}{\lambda}\left(\frac{M_{B,z}-1}{2}\right)\cos(\phi_{t,1})} \right]^{\mathsf{T}}, \quad (5)$$

where $M_R = M_{R,x}M_{R,y}$ with $M_{R,x}$ and $M_{R,y}$ being the number of RIS elements along the x axis and y axis, respectively. Additionally, $d_{R,x}$ and $d_{R,y}$ represent the inter-element spacing for RIS elements along the x-axis and y-axis respectively. Similarly, $M_B = M_{B,y} \times M_{B,z}$, with $M_{B,y}$ and $M_{B,z}$ denoting the number of BS antennas along the y-axis and z-axis, respectively. Furthermore, $d_{B,y}$ and $d_{B,z}$ represent the inter-element spacing for BS antennas along the y-axis and z-axis, respectively.

Similarly, the remaining channels (marked in Fig. 1), e.g., $\mathbf{h}_2 \in \mathbb{C}^{1 \times M_B}$, $\mathbf{h}_3 \in \mathbb{C}^{1 \times M_R}$, $\mathbf{h}_4 \in \mathbb{C}^{M_U \times 1}$, and $\mathbf{H}_5 \in \mathbb{C}^{M_U \times M_B}$ can be presented in the same manner, listed below:

$$\mathbf{h}_2 = \frac{e^{-j2\pi d_2/\lambda}}{\sqrt{\rho_2}} \left(\boldsymbol{\alpha}_y(\theta_{t,2}, \phi_{t,2}) \otimes \boldsymbol{\alpha}_z(\phi_{t,2}) \right)^{\mathsf{H}}, \quad (6)$$

$$\mathbf{h}_3 = \frac{e^{-j2\pi d_3/\lambda}}{\sqrt{\rho_3}} \left(\boldsymbol{\alpha}_x(\theta_{t,3}, \phi_{t,3}) \otimes \boldsymbol{\alpha}_y(\theta_{t,3}, \phi_{t,3}) \right)^{\mathsf{H}}, \quad (7)$$

$$\mathbf{h}_4 = \frac{e^{-j2\pi d_4/\lambda}}{\sqrt{\rho_4}} \boldsymbol{\alpha}_x(\theta_{r,4}, \phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4}, \phi_{r,4}), \quad (8)$$

$$\mathbf{H}_5 = \frac{e^{-j2\pi d_5/\lambda}}{\sqrt{\rho_5}} \boldsymbol{\alpha}_x(\theta_{r,5}, \phi_{r,5}) \otimes \boldsymbol{\alpha}_y(\theta_{r,5}, \phi_{r,5})$$
$$\left(\boldsymbol{\alpha}_y(\theta_{t,5}, \phi_{t,5}) \otimes \boldsymbol{\alpha}_z(\phi_{t,5}) \right)^{\mathsf{H}}, \quad (9)$$

where M_U is the number of UE antennas, $d_2(\rho_2)$, $d_3(\rho_3)$, and $d_4(\rho_4)$ are the distance (path loss) between the drone and the BS, the RIS, and the UE, respectively, $d_5(\rho_5)$ is the distance (path loss) between the BS and the UE. $\theta_{t,i}$ and $\phi_{t,i}$ are the azimuth and elevation angles of departure associated with $\mathbf{h}_i$ for $i = 2, 3$, and $\theta_{r,4}$ ($\theta_{r,5}$) and $\phi_{r,4}$ ($\phi_{r,5}$) are the azimuth and elevation angles of arrival associated with $\mathbf{h}_4$ ($\mathbf{H}_5$). $\theta_{t,5}$ and $\phi_{t,5}$ are the azimuth and elevation angles of departure associated with $\mathbf{H}_5$. In this device-free drone localization system, $\mathbf{H}_5$ is the direct channel between the BS and the UE, which serves as an interference link.

B. Downlink Sounding Procedure

The sounding procedure, illustrated in Fig. 2, involves the BS transmitting beamformed pilot signals over the downlink to illuminate both the RIS and a specific portion of the sky simultaneously, employing a multi-beam approach. Meanwhile, the RIS adjusts its beams to cover a designated area of the sky where the drone might be located. The BS beam $\mathbf{f}_0 \in \mathbb{C}^{M_B}$
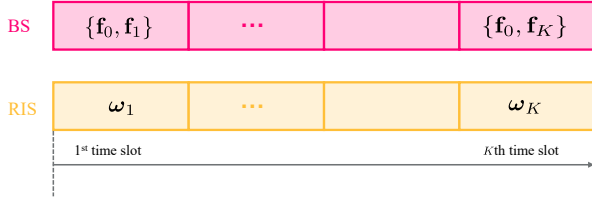
Fig. 2: The sounding framework for passive drone localization, where two beams $\mathbf{f}_0, \mathbf{f}_k \in \mathbb{C}^{M_\mathrm{B} \times 1}$ are concurrently employed at the BS, the former fixed beam towards the RIS and the latter time-varying beam towards the sky, and one time-varying beam $\boldsymbol{\omega}_k \in \mathbb{C}^{M_\mathrm{R} \times 1}$ is considered at the RIS towards the sky, during time slot k.

($\|\mathbf{f}_0\|_2 = \sqrt{M_\mathrm{B}}$) directed towards the RIS remains fixed over time, while the BS beam $\mathbf{f}_k \in \mathbb{C}^{M_\mathrm{B}}$ ($\|\mathbf{f}_k\|_2 = \sqrt{M_\mathrm{B}}$) and the RIS beam $\boldsymbol{\omega}_k \in \mathbb{C}^{M_\mathrm{R}}$ ($\|\boldsymbol{\omega}_k\|_2 = \sqrt{M_\mathrm{R}}$) directed towards the sky vary with time, for $k = 1, \cdots, K$, with K indicating the level of training overhead. We guarantee that the pairs of BS beams are mutually orthogonal, i.e., $\mathbf{f}_0^\mathsf{H} \mathbf{f}_k = 0$, $\forall k$. Each element of RIS phase profile $\boldsymbol{\omega}_k$ satisfies unit-modulus constraint, i.e., $|[\boldsymbol{\omega}_k]_i| = 1$, for $\forall i$. The pilot signals transmitted from the BS traverse various paths to reach the UE, including a direct path (BS-UE), a single-bounce path (BS-drone-UE), and a double-bounce path (BS-RIS-drone-UE).

We assume that the exact locations of the RIS and UE are known to the BS. In this sense, the BS can take full advantage of such prior information and use it for the purpose of beam design during the sounding procedure. It should be noted that during the sounding procedure, the beam design at the BS and RIS is essential for the passive 3D drone localization systems. By following the maximum ratio transmission (MRT) principle based on the known positions of the BS and RIS, we set $\mathbf{f}_0 = (\boldsymbol{\alpha}_y(\theta_{t,1}, \phi_{t,1}) \otimes \boldsymbol{\alpha}_z(\phi_{t,1}))$. To further make the BS beams $\mathbf{f}_0, \mathbf{f}_k$ orthogonal and mitigate the interference from $\mathbf{H}_5$, we define $\mathbf{g}_0 = (\boldsymbol{\alpha}_y(\theta_{t,5}, \phi_{t,5}) \otimes \boldsymbol{\alpha}_z(\phi_{t,5}))$ and constrain the remaining beams (i.e., $\mathbf{f}_k$, for $k = 1, \cdots, K$) to lie within the null space of the matrix $\mathbf{F}^\mathsf{H}$, where $\mathbf{F} = [\mathbf{f}_0, \mathbf{g}_0]$. As a result, $\mathbf{f}_k^\mathsf{H} \mathbf{f}_0 = \mathbf{f}_k^\mathsf{H} \mathbf{g}_0 = 0$, $\forall k$.

C. Signal Model

The received signal at the receiving UE during the time slot k can be expressed as

$$\mathbf{y}_k = \zeta \mathbf{h}_4 (\mathbf{h}_3 \mathrm{diag}(\boldsymbol{\omega}_k) \mathbf{H}_1 \mathbf{F}_k \mathbf{s}_k + \mathbf{h}_2 \mathbf{F}_k \mathbf{s}_k) + \mathbf{H}_5 \mathbf{F}_k \mathbf{s}_k + \mathbf{n}_k, \tag{10}$$

where ζ is the reflection coefficient of the drone, a.k.a. RCS, $\mathbf{F}_k = [\mathbf{f}_0, \mathbf{f}_k]$, $\mathbf{s}_k = [s_0, s_k]^\mathsf{T}$, and the transmitted pilot symbols during each time slot satisfy the following sum power constraint: $\mathbb{E}\{|s_0|^2\} + \mathbb{E}\{|s_k|^2\} = P$ for $\forall k$. Without loss of generality, we set $s_0 = s_k = \sqrt{P/2}$. The additive noise vector at the UE follows complex Gaussian distribution, i.e., $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{M_\mathrm{U}})$. Note that the interference term $\mathbf{H}_5 \mathbf{F}_k \mathbf{s}_k$ is equal to $\mathbf{H}_5 \mathbf{f}_0 s_0$ due to $\mathbf{H}_5 \mathbf{f}_k = \mathbf{0}$. Namely, the interference comes only from the beam $\mathbf{f}_0$. By carefully choosing the locations of the RIS and UE, we can further reduce and even

eliminate the interference term in Eq. (10). The signal vector $\mathbf{y}_k$ in (10) can be further expressed as

$$\begin{aligned}
\mathbf{y}_k &= \zeta \mathbf{h}_4 (\mathbf{h}_3 \mathrm{diag}(\boldsymbol{\alpha}_x(\theta_{r,1}, \phi_{r,1}) \otimes \boldsymbol{\alpha}_y(\theta_{r,1}, \phi_{r,1})) \boldsymbol{\omega}_k \\
&\quad \times \frac{e^{-j2\pi d_1/\lambda}}{\sqrt{\rho_1}} (\boldsymbol{\alpha}_y(\theta_{t,1}, \phi_{t,1}) \otimes \boldsymbol{\alpha}_z(\phi_{t,1}))^\mathsf{H} \mathbf{x}_k \\
&\quad + \mathbf{h}_2 \mathbf{x}_k) + \mathbf{H}_5 \mathbf{x}_k + \mathbf{n}_k \\
&= \sqrt{P/2} \frac{e^{-j2\pi d_1/\lambda}}{\sqrt{\rho_1}} M_\mathrm{B} \zeta \mathbf{h}_4 \mathbf{h}_3 \\
&\quad \times \mathrm{diag}(\boldsymbol{\alpha}_x(\theta_{r,1}, \phi_{r,1}) \otimes \boldsymbol{\alpha}_y(\theta_{r,1}, \phi_{r,1})) \boldsymbol{\omega}_k \\
&\quad + \sqrt{P/2} \zeta \mathbf{h}_4 \mathbf{h}_2 \bar{\mathbf{f}}_k + \sqrt{P/2} \mathbf{H}_5 \bar{\mathbf{f}}_k + \mathbf{n}_k, \tag{11}
\end{aligned}$$

where $\mathbf{x}_k = \mathbf{F}_k \mathbf{s}_k$ and $\bar{\mathbf{f}}_k = \mathbf{f}_0 + \mathbf{f}_k$. By stacking $\{\mathbf{y}_k\}$ column by column, the received signal can be expressed as

$$\mathbf{Y} = \sqrt{P/2} \tilde{\mathbf{H}} \bar{\boldsymbol{\Omega}} + \sqrt{P/2} \hat{\mathbf{H}} \bar{\mathbf{F}} + \sqrt{P/2} \mathbf{H}_5 \bar{\mathbf{F}} + \mathbf{N}, \tag{12}$$

where $\bar{\boldsymbol{\Omega}} = [\mathrm{diag}(\boldsymbol{\alpha}_x(\theta_{r,1}, \phi_{r,1}) \otimes \boldsymbol{\alpha}_y(\theta_{r,1}, \phi_{r,1})) \boldsymbol{\omega}_1, \cdots, \mathrm{diag}(\boldsymbol{\alpha}_x(\theta_{r,1}, \phi_{r,1}) \otimes \boldsymbol{\alpha}_y(\theta_{r,1}, \phi_{r,1})) \boldsymbol{\omega}_K]$, $\bar{\mathbf{F}} = [\bar{\mathbf{f}}_1, \cdots, \bar{\mathbf{f}}_K]$, and $\mathbf{N} = [\mathbf{n}_1, \cdots, \mathbf{n}_K]$. The two cascaded channels $\tilde{\mathbf{H}}$ and $\hat{\mathbf{H}}$ are function of ζ, $\mathbf{h}_4$, $\mathbf{h}_3$, and $\mathbf{h}_2$, as

$$\begin{aligned}
\tilde{\mathbf{H}} &= \zeta M_\mathrm{B} \frac{e^{-j2\pi d_1/\lambda}}{\sqrt{\rho_1}} \mathbf{h}_4 \mathbf{h}_3 = \tilde{\epsilon} \boldsymbol{\alpha}_x(\theta_{r,4}, \phi_{r,4}) \boldsymbol{\alpha}_x(\theta_{t,3}, \phi_{t,3})^\mathsf{H} \\
&\quad \otimes \boldsymbol{\alpha}_y(\theta_{r,4}, \phi_{r,4}) \boldsymbol{\alpha}_y(\theta_{t,3}, \phi_{t,3})^\mathsf{H}, \tag{13} \\
\hat{\mathbf{H}} &= \zeta \mathbf{h}_4 \mathbf{h}_2 = \hat{\epsilon} \boldsymbol{\alpha}_x(\theta_{r,4}, \phi_{r,4}) \boldsymbol{\alpha}_y(\theta_{t,2}, \phi_{t,2})^\mathsf{H} \\
&\quad \otimes \boldsymbol{\alpha}_y(\theta_{r,4}, \phi_{r,4}) \boldsymbol{\alpha}_z(\phi_{t,2})^\mathsf{H}, \tag{14}
\end{aligned}$$

where $\tilde{\epsilon} = \zeta M_\mathrm{B} \frac{e^{-j2\pi d_1/\lambda}}{\sqrt{\rho_1}} \frac{e^{-j2\pi d_4/\lambda}}{\sqrt{\rho_4}} \frac{e^{-j2\pi d_3/\lambda}}{\sqrt{\rho_3}}$ and $\hat{\epsilon} = \zeta \frac{e^{-j2\pi d_4/\lambda}}{\sqrt{\rho_4}} \frac{e^{-j2\pi d_2/\lambda}}{\sqrt{\rho_2}}$.

Since we assume that the angular parameters in $\mathbf{H}_1$ are known *a priori* due to the known positions of the BS and RIS, $\mathrm{diag}(\boldsymbol{\alpha}_x(\theta_{r,1}, \phi_{r,1}) \otimes \boldsymbol{\alpha}_y(\theta_{r,1}, \phi_{r,1}))$ is known in $\bar{\boldsymbol{\Omega}}$. Therefore, in Eq. (12), both $\bar{\boldsymbol{\Omega}}$ and $\bar{\mathbf{F}}$ are known to the UE. Recall that the goal of the study is to localize the drone in a passive device-free manner based on the received signal $\mathbf{Y}$ at the UE. The unknown parameters in $\tilde{\mathbf{H}}$ are $\mathbf{h}_3$, $\mathbf{h}_4$, ζ, and $\frac{e^{-j2\pi d_1/\lambda}}{\sqrt{\rho_1}}$. In $\hat{\mathbf{H}}$, all the three terms, i.e., ζ, $\mathbf{h}_4$, and $\mathbf{h}_2$, are unknown. Both $\tilde{\mathbf{H}}$ and $\hat{\mathbf{H}}$ are rank-one matrices. In the considered localization system, we have three reference nodes, i.e., the BS, the UE, and the RIS. Provided that we can estimate the angular parameters in $\mathbf{h}_2$, $\mathbf{h}_3$, and $\mathbf{h}_4$, i.e., $\theta_{t,2}$, $\phi_{t,2}$, $\theta_{t,3}$, $\phi_{t,3}$, $\theta_{r,4}$, and $\phi_{r,4}$, from the received signal $\mathbf{Y}$ in Eq. (12), we can exploit the Triangulation localization method to calculate the 3D coordinate of the drone, realizing passive 3D drone localization.

III. PERFORMANCE LIMIT ANALYSIS

In this section, we characterize the 3D drone localization via Fisher information analysis and CRLB. According to Eqs. (13) and (14), the unknown parameters in the received signal in Eq. (12) are complex-valued $\hat{\epsilon}$ and $\tilde{\epsilon}$ as well as real-valued $\theta_{t,2}$, $\phi_{t,2}$, $\theta_{t,3}$, $\phi_{t,3}$, $\theta_{r,4}$, and $\phi_{r,4}$. We stack all the unknown parameters into a vector $\boldsymbol{\eta} =$

$[\Re(\tilde{\epsilon}) \ \Im(\tilde{\epsilon}) \ \Re(\hat{\epsilon}) \ \Im(\hat{\epsilon}) \ \theta_{t,2} \ \phi_{t,2} \ \theta_{t,3} \ \phi_{t,3} \ \theta_{r,4} \ \phi_{r,4}]^{\mathsf{T}} \in \mathbb{R}^{10}$. The vectorized received signal can be expressed as

$$
\begin{aligned}
\mathbf{y} =& \sqrt{P/2}\tilde{\epsilon}(\bar{\boldsymbol{\Omega}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_x^*(\theta_{t,3}, \phi_{t,3}) \otimes \boldsymbol{\alpha}_y^*(\theta_{t,3}, \phi_{t,3}) \\
& \otimes \boldsymbol{\alpha}_x(\theta_{r,4}, \phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4}, \phi_{r,4})) \\
& + \sqrt{P/2}\hat{\epsilon}(\bar{\mathbf{F}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_y^*(\theta_{t,2}, \phi_{t,2}) \otimes \boldsymbol{\alpha}_z^*(\phi_{t,2}) \\
& \otimes \boldsymbol{\alpha}_x(\theta_{r,4}, \phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4}, \phi_{r,4})) + \sqrt{\frac{P}{2}}\mathrm{vec}(\mathbf{H}_5\bar{\mathbf{F}}) + \mathbf{n},
\end{aligned}
$$
(15)

where $\mathbf{y} = \mathrm{vec}(\mathbf{Y})$, and $\mathbf{n} = \mathrm{vec}(\mathbf{N})$. The Fisher information matrix $\mathbf{J} \in \mathbb{C}^{10 \times 10}$ for $\boldsymbol{\eta}$ is written as [10]

$$
[\mathbf{J}]_{mn} = \frac{2}{\sigma^2}\Re\left\{\frac{\partial \boldsymbol{\mu}^{\mathsf{H}}}{[\boldsymbol{\eta}]_m}\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_n}\right\},
$$
(16)

where $\boldsymbol{\mu}$ is the noise- and interference-free signal from Eq. (15). The details on $\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_n}$ are provided in Eqs. (17)–(26), where $\boldsymbol{\Phi}_1 = -j\pi\cos(\theta_{t,2})\sin(\phi_{t,2})\mathrm{diag}([-\frac{M_{\mathrm{B},y}-1}{2}, \cdots, \frac{M_{\mathrm{B},y}-1}{2}])$,
$\boldsymbol{\Phi}_2 = -j\pi\sin(\theta_{t,2})\cos(\phi_{t,2})\mathrm{diag}([-\frac{M_{\mathrm{B},y}-1}{2}, \cdots, \frac{M_{\mathrm{B},y}-1}{2}])$,
$\boldsymbol{\Phi}_3 = j\pi\sin(\phi_{t,2})\mathrm{diag}([-\frac{M_{\mathrm{B},z}-1}{2}, \cdots, \frac{M_{\mathrm{B},z}-1}{2}])$,
$\boldsymbol{\Phi}_4 = j\pi\sin(\theta_{t,3})\sin(\phi_{t,3})\mathrm{diag}([-\frac{M_{\mathrm{R},x}-1}{2}, \cdots, \frac{M_{\mathrm{R},x}-1}{2}])$,
$\boldsymbol{\Phi}_5 = -j\pi\cos(\theta_{t,3})\sin(\phi_{t,3})\mathrm{diag}([-\frac{M_{\mathrm{R},y}-1}{2}, \cdots, \frac{M_{\mathrm{R},y}-1}{2}])$,
$\boldsymbol{\Phi}_6 = -j\pi\cos(\theta_{t,3})\cos(\phi_{t,3})\mathrm{diag}([-\frac{M_{\mathrm{R},x}-1}{2}, \cdots, \frac{M_{\mathrm{R},x}-1}{2}])$,
$\boldsymbol{\Phi}_7 = -j\pi\sin(\theta_{t,3})\cos(\phi_{t,3})\mathrm{diag}([-\frac{M_{\mathrm{R},y}-1}{2}, \cdots, \frac{M_{\mathrm{R},y}-1}{2}])$,
$\boldsymbol{\Phi}_8 = -j\pi\sin(\theta_{r,4})\sin(\phi_{r,4})\mathrm{diag}([-\frac{M_{\mathrm{U},x}-1}{2}, \cdots, \frac{M_{\mathrm{U},x}-1}{2}])$,
$\boldsymbol{\Phi}_9 = j\pi\cos(\theta_{r,4})\sin(\phi_{r,4})\mathrm{diag}([-\frac{M_{\mathrm{U},y}-1}{2}, \cdots, \frac{M_{\mathrm{U},y}-1}{2}])$,
$\boldsymbol{\Phi}_{10} = j\pi\cos(\theta_{r,4})\cos(\phi_{r,4})\mathrm{diag}([-\frac{M_{\mathrm{U},x}-1}{2}, \cdots, \frac{M_{\mathrm{U},x}-1}{2}])$,
$\boldsymbol{\Phi}_{11} = j\pi\sin(\theta_{r,4})\cos(\phi_{r,4})\mathrm{diag}([-\frac{M_{\mathrm{U},y}-1}{2}, \cdots, \frac{M_{\mathrm{U},y}-1}{2}])$.

In this work, we only consider the angular parameter estimates for passive 3D drone localization. The transform matrix, a.k.a. Jacobian matrix, is calculated based on the geometrical relationship among the network nodes [10]. In its first six columns, the transform matrix $\mathbf{T} \in \mathbb{R}^{3 \times 10}$ has only non-zero elements in the following places:

$$
[\mathbf{T}]_{15} = \partial\theta_{t,2}/\partial x_{\mathrm{D}} = -\frac{\sin(\theta_{t,2})}{d_2\cos(\phi_{t,2})},
$$
(27)

$$
[\mathbf{T}]_{25} = \partial\theta_{t,2}/\partial y_{\mathrm{D}} = \frac{\cos(\theta_{t,2})}{d_2\cos(\phi_{t,2})},
$$
(28)

$$
[\mathbf{T}]_{16} = \partial\phi_{t,2}/\partial x_{\mathrm{D}} = -\frac{\cos(\theta_{t,2})\sin(\phi_{t,2})}{d_2},
$$
(29)

$$
[\mathbf{T}]_{26} = \partial\phi_{t,2}/\partial y_{\mathrm{D}} = -\frac{\sin(\theta_{t,2})\sin(\phi_{t,2})}{d_2},
$$
(30)

$$
[\mathbf{T}]_{36} = \partial\phi_{t,2}/\partial z_{\mathrm{D}} = \frac{\cos(\phi_{t,2})}{d_2},
$$
(31)

where $(x_{\mathrm{D}}, y_{\mathrm{D}}, z_{\mathrm{D}})$ is the 3D coordinate of the drone. Similarly, we can calculate the non-zero entries within $[\mathbf{T}]_{:,7:10}$. The localization error bound in terms of root mean square error (RMSE) can be characterized by

$$
\begin{aligned}
\sigma_{\hat{\mathbf{p}}_{\mathrm{D}}} &= \sqrt{\mathbb{E}\{(\mathbf{p}_{\mathrm{D}} - \hat{\mathbf{p}}_{\mathrm{D}})^{\mathsf{T}}(\mathbf{p}_{\mathrm{D}} - \hat{\mathbf{p}}_{\mathrm{D}})\}} \\
&\geq \sqrt{\mathrm{Tr}\left\{(\mathbf{TJT}^{\mathsf{T}})^{-1}\right\}},
\end{aligned}
$$
(32)

where $\mathbf{p}_{\mathrm{D}} = (x_{\mathrm{D}}, y_{\mathrm{D}}, z_{\mathrm{D}})$ and $\hat{\mathbf{p}}_{\mathrm{D}}$ is its estimate.

IV. DRONE LOCALIZATION ALGORITHM

For the sake of tractability, we develop a multi-stage localization approach: i) estimation of $\{\mathbf{h}_4, \mathbf{h}_2, \mathbf{h}_3\}$, ii) extraction of $\{\theta_{t,2}, \phi_{t,2}, \theta_{t,3}, \phi_{t,3}, \theta_{r,4}, \phi_{r,4}\}$, and iii) mapping the estimates from step ii) to the 3D drone location.

A. CGD Based Channel Estimation

In the first stage, we resort to CGD for estimating $\mathbf{h}_4$, $\mathbf{h}_2$, and $\mathbf{h}_3$ in a sequential manner. The objective is to minimize

$$
\min f(\mathbf{h}_2, \mathbf{h}_3, \mathbf{h}_4) = \left\|\mathbf{Y} - \sqrt{P/2}\tilde{\mathbf{H}}\bar{\boldsymbol{\Omega}} - \sqrt{P/2}\hat{\mathbf{H}}\bar{\mathbf{F}}\right\|_{\mathsf{F}}^2.
$$
(33)

The CGD approach provides the solutions for the three variables iteratively until convergence or reaching a certain stopping criterion. The major updating steps are summarized in the following:

$$
\mathbf{h}_4^{(k)} = \mathbf{h}_4^{(k-1)} - \eta\frac{\partial f\left(\mathbf{h}_2^{(k-1)}, \mathbf{h}_3^{(k-1)}, \mathbf{h}_4\right)}{\partial \mathbf{h}_4}\bigg|_{\mathbf{h}_4 = \mathbf{h}_4^{(k-1)}},
$$
(34)

$$
\mathbf{h}_2^{(k)} = \mathbf{h}_2^{(k-1)} - \eta\frac{\partial f\left(\mathbf{h}_2, \mathbf{h}_3^{(k-1)}, \mathbf{h}_4^{(k)}\right)}{\partial \mathbf{h}_2}\bigg|_{\mathbf{h}_2 = \mathbf{h}_2^{(k-1)}},
$$
(35)

$$
\mathbf{h}_3^{(k)} = \mathbf{h}_3^{(k-1)} - \eta\frac{\partial f\left(\mathbf{h}_2^{(k)}, \mathbf{h}_3, \mathbf{h}_4^{(k)}\right)}{\partial \mathbf{h}_3}\bigg|_{\mathbf{h}_3 = \mathbf{h}_3^{(k-1)}},
$$
(36)

where η is the step size and superscript (k) specifies the iteration index. The above gradient terms are further detailed below:

$$
\frac{\partial f\left(\mathbf{h}_2^{(k-1)}, \mathbf{h}_3^{(k-1)}, \mathbf{h}_4\right)}{\partial \mathbf{h}_4} = \mathbf{D}^{\mathsf{H}}\mathbf{D}\mathbf{h}_4 - \mathbf{D}^{\mathsf{H}}\mathrm{vec}(\mathbf{Y}),
$$

$$
\frac{\partial f\left(\mathbf{h}_2, \mathbf{h}_3^{(k-1)}, \mathbf{h}_4^{(k)}\right)}{\partial \mathbf{h}_2} = \left(\mathbf{B}^{\mathsf{H}}\mathbf{B}\mathbf{h}_2 - \mathbf{B}^{\mathsf{H}}\mathrm{vec}(\mathbf{Y}_1)\right)^{\mathsf{T}},
$$

$$
\frac{\partial f\left(\mathbf{h}_2^{(k)}, \mathbf{h}_3, \mathbf{h}_4^{(k)}\right)}{\partial \mathbf{h}_3} = \left(\mathbf{C}^{\mathsf{H}}\mathbf{C}\mathbf{h}_3 - \mathbf{C}^{\mathsf{H}}\mathrm{vec}(\mathbf{Y}_2)\right)^{\mathsf{T}},
$$

where

$$
\mathbf{Y}_1 = \mathbf{Y} - \sqrt{P/2}\mathbf{h}_4^{(k)}\mathbf{h}_3^{(k-1)}\bar{\boldsymbol{\Omega}},
$$
(37)

$$
\mathbf{Y}_2 = \mathbf{Y} - \sqrt{P/2}\mathbf{h}_4^{(k)}\mathbf{h}_2^{(k)}\bar{\mathbf{F}},
$$
(38)

$$
\mathbf{B} = \sqrt{P/2}\left(\bar{\mathbf{F}}^{\mathsf{T}} \otimes \mathbf{h}_4^{(k)}\right),
$$
(39)

$$
\mathbf{C} = \sqrt{P/2}\left(\bar{\boldsymbol{\Omega}}^{\mathsf{T}} \otimes \mathbf{h}_4^{(k)}\right),
$$
(40)

$$
\mathbf{D} = \left(\sqrt{P/2}\mathbf{h}_3^{(k-1)}\bar{\boldsymbol{\Omega}} + \sqrt{P/2}\mathbf{h}_2^{(k-1)}\bar{\mathbf{F}}\right)^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}}.
$$
(41)

Remark 1: Here, we integrate ζ into $\mathbf{h}_4$ for the sake of simplicity. Intuitively, the scaling coefficient ζ will bring a significant impact on the estimation results, verified in Section V.

Remark 2: We employ the iterative CGD method for the channel vector estimation due to its low computational complexity. Alternatively, one can also exploit nuclear norm

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_1} = \sqrt{P/2}(\bar{\boldsymbol{\Omega}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_x^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_y^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})), \tag{17}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_2} = j\sqrt{P/2}(\bar{\boldsymbol{\Omega}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_x^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_y^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})), \tag{18}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_3} = \sqrt{P/2}(\bar{\mathbf{F}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_y^*(\theta_{t,2},\phi_{t,2}) \otimes \boldsymbol{\alpha}_z^*(\phi_{t,2}) \otimes \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})), \tag{19}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_4} = j\sqrt{P/2}(\bar{\mathbf{F}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_y^*(\theta_{t,2},\phi_{t,2}) \otimes \boldsymbol{\alpha}_z^*(\phi_{t,2}) \otimes \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})), \tag{20}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_5} = \sqrt{P/2}\hat{\epsilon}(\bar{\mathbf{F}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\Phi}_1\boldsymbol{\alpha}_y^*(\theta_{t,2},\phi_{t,2}) \otimes \boldsymbol{\alpha}_z^*(\phi_{t,2}) \otimes \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})), \tag{21}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_6} = \sqrt{P/2}\hat{\epsilon}(\bar{\mathbf{F}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})((\boldsymbol{\Phi}_2\boldsymbol{\alpha}_y^*(\theta_{t,2},\phi_{t,2}) \otimes \boldsymbol{\alpha}_z^*(\phi_{t,2}) + \boldsymbol{\alpha}_y^*(\theta_{t,2},\phi_{t,2}) \otimes \boldsymbol{\Phi}_3\boldsymbol{\alpha}_z^*(\phi_{t,2})) \otimes \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})), \tag{22}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_7} = \sqrt{\frac{P}{2}}\tilde{\epsilon}(\bar{\boldsymbol{\Omega}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})((\boldsymbol{\Phi}_4\boldsymbol{\alpha}_x^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_y^*(\theta_{t,3},\phi_{t,3}) + (\boldsymbol{\alpha}_x^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\Phi}_5\boldsymbol{\alpha}_y^*(\theta_{t,3},\phi_{t,3})) \otimes \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})), \tag{23}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_8} = \sqrt{\frac{P}{2}}\tilde{\epsilon}(\bar{\boldsymbol{\Omega}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})((\boldsymbol{\Phi}_6\boldsymbol{\alpha}_x^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_y^*(\theta_{t,3},\phi_{t,3}) + (\boldsymbol{\alpha}_x^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\Phi}_7\boldsymbol{\alpha}_y^*(\theta_{t,3},\phi_{t,3})) \otimes \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})), \tag{24}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_9} = \sqrt{\frac{P}{2}}\tilde{\epsilon}(\bar{\boldsymbol{\Omega}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_x^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_y^*(\theta_{t,3},\phi_{t,3}) \otimes (\boldsymbol{\Phi}_8\boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4}) + \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\Phi}_9\boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})))$$
$$+ \sqrt{\frac{P}{2}}\hat{\epsilon}(\bar{\mathbf{F}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_y^*(\theta_{t,2},\phi_{t,2}) \otimes \boldsymbol{\alpha}_z^*(\phi_{t,2}) \otimes (\boldsymbol{\Phi}_8\boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4}) + \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\Phi}_9\boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4}))), \tag{25}$$

$$\frac{\partial \boldsymbol{\mu}}{[\boldsymbol{\eta}]_{10}} = \sqrt{\frac{P}{2}}\tilde{\epsilon}(\bar{\boldsymbol{\Omega}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_x^*(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_y^*(\theta_{t,3},\phi_{t,3}) \otimes (\boldsymbol{\Phi}_{10}\boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4}) + \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\Phi}_{11}\boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})))$$
$$+ \sqrt{\frac{P}{2}}\hat{\epsilon}(\bar{\mathbf{F}}^{\mathsf{T}} \otimes \mathbf{I}_{M_{\mathrm{U}}})(\boldsymbol{\alpha}_y^*(\theta_{t,2},\phi_{t,2}) \otimes \boldsymbol{\alpha}_z^*(\phi_{t,2}) \otimes (\boldsymbol{\Phi}_{10}\boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4}) + \boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\Phi}_{11}\boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4}))). \tag{26}$$

minimization, posing a regularized optimization problem [11]. However, this approach entails inevitably higher complexity, particularly when the numbers of RIS elements and BS antennas are large, leading to larger-sized matrices $\tilde{\mathbf{H}}$ and $\hat{\mathbf{H}}$.

B. Angular Parameter Estimation via 2D Search

In the second stage, we extract the angular parameters from the estimate of $\mathbf{h}_i$, denoted as $\hat{\mathbf{h}}_i$, for $i = 2, 3, 4$. Based on two-dimensional (2D) search, we can find the angular parameter estimates, $\hat{\theta}_{t,i}, \hat{\phi}_{t,i}$, for $i = 2, 3, 4$, as follows:

$$\hat{\theta}_{t,2}, \hat{\phi}_{t,2} = \underset{\theta_{t,2},\phi_{t,2}}{\operatorname{argmax}} \left| \hat{\mathbf{h}}_2\big(\boldsymbol{\alpha}_y(\theta_{t,2},\phi_{t,2}) \otimes \boldsymbol{\alpha}_z(\phi_{t,2})\big) \right|,$$

$$\hat{\theta}_{t,3}, \hat{\phi}_{t,3} = \underset{\theta_{t,3},\phi_{t,3}}{\operatorname{argmax}} \left| \hat{\mathbf{h}}_3\big(\boldsymbol{\alpha}_x(\theta_{t,3},\phi_{t,3}) \otimes \boldsymbol{\alpha}_y(\theta_{t,3},\phi_{t,3})\big) \right|,$$

$$\hat{\theta}_{r,4}, \hat{\phi}_{r,4} = \underset{\theta_{r,4},\phi_{r,4}}{\operatorname{argmax}} \left| \hat{\mathbf{h}}_4^{\mathsf{H}}\big(\boldsymbol{\alpha}_x(\theta_{r,4},\phi_{r,4}) \otimes \boldsymbol{\alpha}_y(\theta_{r,4},\phi_{r,4})\big) \right|.$$

C. Drone Location Mapping

Based on the estimated angular parameters $\{\hat{\theta}_{t,2}, \hat{\phi}_{t,2}, \hat{\theta}_{t,3}, \hat{\phi}_{t,3}, \hat{\theta}_{r,4}, \hat{\phi}_{r,4}\}$, we in the last stage calculate the 3D drone location by taking into consideration the geometrical relationship between the drone and the other network nodes in the spherical coordinate system. We apply the least squares (LS) principle for mapping those angular

estimates to the 3D position estimate of the drone, $\hat{\mathbf{p}}_{\mathrm{D}}$, as [4, sec. IV.A]

$$\hat{\mathbf{p}}_{\mathrm{D}} = (\mathbf{B}_{\mathrm{B}} + \mathbf{B}_{\mathrm{R}} + \mathbf{B}_{\mathrm{U}})^{-1} (\mathbf{B}_{\mathrm{B}}\mathbf{p}_{\mathrm{B}} + \mathbf{B}_{\mathrm{R}}\mathbf{p}_{\mathrm{R}} + \mathbf{B}_{\mathrm{U}}\mathbf{p}_{\mathrm{U}}), \tag{42}$$

where $\mathbf{B}_{\mathrm{B}} \triangleq \mathbf{I}_3 - \hat{\boldsymbol{\xi}}_{\mathrm{B}}\hat{\boldsymbol{\xi}}_{\mathrm{B}}^{\mathsf{T}}$ with $\hat{\boldsymbol{\xi}}_{\mathrm{B}} \triangleq [\cos(\hat{\theta}_{t,2})\cos(\hat{\phi}_{t,2}), \sin(\hat{\theta}_{t,2})\cos(\hat{\phi}_{t,2}), \sin(\hat{\phi}_{t,2})]^{\mathsf{T}}$, $\mathbf{B}_{\mathrm{R}} \triangleq \mathbf{I}_3 - \hat{\boldsymbol{\xi}}_{\mathrm{R}}\hat{\boldsymbol{\xi}}_{\mathrm{R}}^{\mathsf{T}}$ with $\hat{\boldsymbol{\xi}}_{\mathrm{R}} \triangleq [\cos(\hat{\theta}_{t,3})\cos(\hat{\phi}_{t,3}), \sin(\hat{\theta}_{t,3})\cos(\hat{\phi}_{t,3}), \sin(\hat{\phi}_{t,3})]^{\mathsf{T}}$, $\mathbf{B}_{\mathrm{U}} \triangleq \mathbf{I}_3 - \hat{\boldsymbol{\xi}}_{\mathrm{U}}\hat{\boldsymbol{\xi}}_{\mathrm{U}}^{\mathsf{T}}$ with $\hat{\boldsymbol{\xi}}_{\mathrm{U}} \triangleq [\cos(\hat{\theta}_{r,4})\cos(\hat{\phi}_{r,4}), \sin(\hat{\theta}_{r,4})\cos(\hat{\phi}_{r,4}), \sin(\hat{\phi}_{r,4})]^{\mathsf{T}}$, and $\mathbf{p}_{\mathrm{B}}, \mathbf{p}_{\mathrm{R}}$, and $\mathbf{p}_{\mathrm{U}}$ are the 3D coordinates of the BS, the RIS, and UE, respectively.

V. SIMULATION RESULTS

In this section, we evaluate the performance of the RIS-assisted drone localization systems using the following parameter configuration: $M_{\mathrm{B}} = 8 \times 8, M_{\mathrm{U}} = 4 \times 4$, and $M_{\mathrm{R}} = 6 \times 6$. The locations of the BS, RIS, drone, and UE are specified as $(0, 0, 26)$, $(0, 0.5, 25.5)$, $(3, 3, 30)$, and $(2, 2, 24)$, respectively. The path loss ρ_i is defined as $\rho_i = d_i^{\Gamma}$ with $\Gamma = 2$. Leveraging geometric relationships, we compute angular parameters for each of the five individual channels. The channel bandwidth B is fixed as 20 MHz, yielding a noise variance of $\sigma^2 = -174$ dBm$+10\log_{10}(B) = -101$ dBm. The signal-to-noise ratio (SNR) P/σ^2 is varied from -10 dB to 10
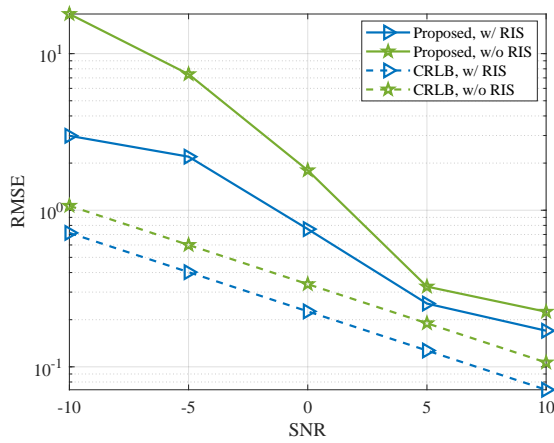
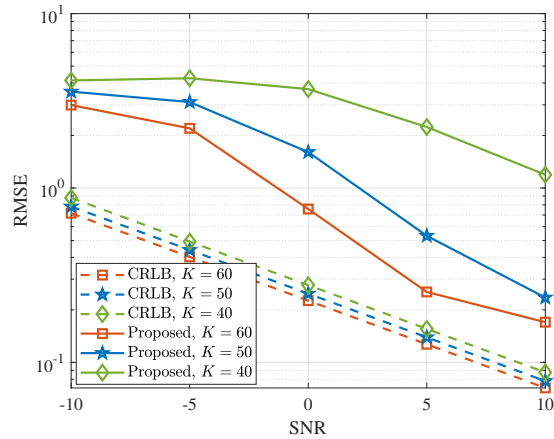Fig. 3: RMSE performance with and without RIS, where $K = 60$ and $\zeta = 1$.

Fig. 4: RMSE performance with various K values and $\zeta = 1$.

dB. A straightforward benchmark involves comparing drone localization performance without an RIS, relying solely on two reference nodes within the system. The simulation results, depicted in Fig. 3, reveal a significant improvement with the introduction of an RIS, particularly evident in the low SNR regime. In the legend, "Proposed" stands for the developed practical localization algorithm based on CGD.

We examine the impact of training overhead on performance. Various levels of training overhead, i.e., K values, are considered. The simulation results are illustrated in Fig. 4. As expected, higher training overhead leads to improved drone localization performance.

The RCS values play a pivotal role in affecting the performance. Hence, we evaluate the following three cases: $\zeta \in \{0.5, 1, 2\}$. The simulation results are depicted in Fig. 5. Notably, RCS values exhibit a substantial impact on the accuracy of angular parameter estimation (as discussed in Section IV-B), which in turn affects the passive 3D drone localization performance.

VI. Conclusions

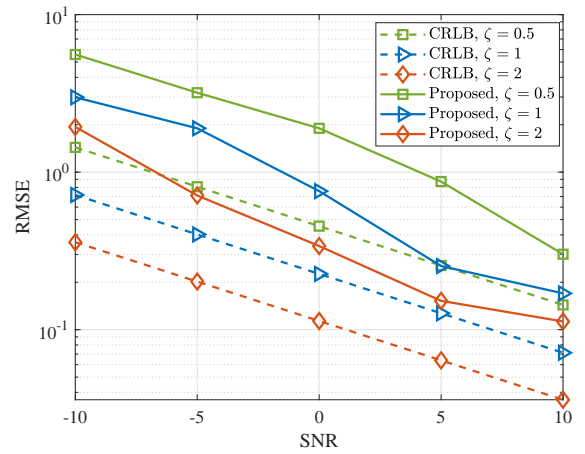In this paper, we have investigated device-free drone localization within RIS-assisted mmWave MIMO networks,

Fig. 5: RMSE performance with $K = 60$ and $\zeta \in \{0.5, 1, 2\}$.

specifically focusing on scenarios where direct communication between the drone and the BS is absent. We have theoretically analyzed the performance limits via Fisher information analysis as well as CRLB and developed a practical localization algorithm based on CGD. Furthermore, we have examined various system parameters to offer insights for the prospective deployment of such systems. The optimal power allocation among the BS beams will be considered in our future work.

References

[1] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados, and H. Wymeersch, "Error bounds for uplink and downlink 3D localization in 5G millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 4939–4954, 2018.

[2] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[3] D. Solomitckii, M. Gapeyenko, V. Semkin, S. Andreev, and Y. Koucheryavy, "Technologies for efficient amateur drone detection in 5G millimeter-wave cellular infrastructure," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 43–50, 2018.

[4] J. He, A. Fakhreddine, and G. C. Alexandropoulos, "RIS-augmented millimeter-wave MIMO systems for passive drone detection," *arXiv preprint arXiv:2402.07259*, 2024.

[5] V. Semkin, J. Haarla, T. Pairon, C. Slezak, S. Rangan, V. Viikari, and C. Oestges, "Analyzing radar cross section signatures of diverse drone models at mmWave frequencies," *IEEE Access*, vol. 8, pp. 48 958–48 969, 2020.

[6] S. Buzzi, E. Grossi, M. Lops, and L. Venturino, "Radar target detection aided by reconfigurable intelligent surfaces," *IEEE Signal Process. Lett.*, vol. 28, pp. 1315–1319, 2021.

[7] ——, "Foundations of MIMO radar detection aided by reconfigurable intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 70, pp. 1749–1763, 2022.

[8] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, 2016.

[9] Y. Tsai, L. Zheng, and X. Wang, "Millimeter-wave beamformed full-dimensional MIMO channel estimation based on atomic norm minimization," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6150–6163, 2018.

[10] J. He, A. Fakhreddine, C. Vanwynsberghe, H. Wymeersch, and G. C. Alexandropoulos, "3D localization with a single partially-connected receiving RIS: Positioning error analysis and algorithmic design," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13 190–13 202, 2023.

[11] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2862–2869.