CHALMERS
UNIVERSITY OF TECHNOLOGY

3D Cooperative Positioning Via RIS and Sidelink Communications With Zero Access Points

Citation for the original published paper (version of record):
Ammous, M., Chen, H., Wymeersch, H. et al (2025). 3D Cooperative Positioning Via RIS and Sidelink Communications With Zero Access Points. IEEE Transactions on Mobile Computing, 24(7): 6119-6136. http://dx.doi.org/10.1109/TMC.2025.3541575

(article starts on next page)

3D Cooperative Positioning via RIS and Sidelink Communications with Zero Access Points

Mustafa Ammous, *Student Member, IEEE*, Hui Chen, *Member, IEEE*, Henk Wymeersch, *Fellow, IEEE*, and Shahrokh Valaee, *Fellow, IEEE*

Abstract—Reconfigurable intelligent surfaces (RISs) are expected to be a main component of future 6G networks due to their capability to create a controllable wireless environment, achieve extended coverage, and improve localization accuracy. In this paper, we present a novel cooperative positioning use case of the RIS in mmWave frequencies and show that in the presence of RIS, together with sidelink communications, localization with zero access points (APs) is possible. We show that multiple (at least three) half-duplex single-antenna user equipments (UEs) can cooperatively estimate their positions through device-to-device communications with a single RIS as an anchor without the need for any APs. We start by formulating a three-dimensional positioning problem with Cramér-Rao lower bound (CRLB) derived for performance analysis. After that, we discuss the RIS profile design and the power allocation strategy between the UEs. Then, we propose low-complexity estimators for estimating the channel parameters and UEs' positions. Finally, we evaluate the performance of the proposed estimators and RIS profiles in the considered scenario via extensive simulations and show that sub-meter level positioning accuracy can be achieved under multi-path propagation.

Index Terms—Cooperative positioning; device-to-device (D2D); millimeter wave (mmWave); reconfigurable intelligent surface (RIS).

1 Introduction

1.1 Motivation

Accurate positioning[1] is crucial for various services and technologies such as emergency services, Internet-of-Things, intelligent transportation systems, and smart cities [1], [2]. While the global navigation satellite system (GNSS) is commonly used, its accuracy diminishes in urban and indoor settings due to signal blockage [3]. Radio positioning using communication infrastructure offers a viable alternative by leveraging power–, time–, or angle–based measurements for position estimation [4]. 5G and beyond 5G networks are expected to enhance positioning accuracy by utilizing higher bandwidths, higher frequencies and massive multiple-input multiple-output (mMIMO) [5]. One of the challenging issues for systems operating at high frequency (e.g., mmWave and THz bands) is the high penetration and path losses, limiting the coverage of communication and positioning services. Network densification, by adding more access points (APs), helps to improve network connectivity and coverage, but will drastically increase hardware and deployment costs, and energy consumption [6]. Instead, adding passive nodes to the network will achieve a more optimized system cost and lower energy consumption.

Recently, the concept of *reconfigurable intelligent surfaces* (RISs), also referred to as *intelligent reflective surfaces* (IRSs), has drawn attention from both academia and industry due to their relatively low cost and ability to enhance coverage while reducing energy consumption [7]. RIS is an array of digitally-controlled metamaterial elements that reflect incident signals [8]. RISs can improve communications by creating an indirect path between an AP and a user equipment (UE) when the line-of-sight (LoS) is not available, or to boost the communication capability in the presence of LoS [9]. Recent studies have shown that RIS can also assist localization services [10]. Although one AP is sufficient for communication purposes, more anchors (APs) are usually needed for location estimation [2]. For example, three time-difference-of-arrival (TDOA) measurements from four APs are needed for localization in 4G systems, while angle measurements from two APs (equipped with antenna arrays) are sufficient for localization in 5G networks [9]. By acting as an additional anchor, RIS enables location estimation with only one AP [9]. Recently, it has been shown that with the assistance of RIS, a full-duplex UE can achieve self-localization without the need for APs [11].

Furthermore, the introduction of sidelink or device-to-device (D2D)[2] communications in 5G-NR opens the road for a slew of opportunities to achieve more accurate UE positioning due to the availability of a larger number of radio measurements via cooperation between different UEs [12], [13]. Without the need for a coordinating AP, the UEs can autonomously select sidelink resources from a pre-configured sidelink resource pool(s) [14]. Thus, the

- *M. Ammous and S. Valaee are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada. E-mail:mustafa.ammous@mail.utoronto.ca; valaee@ece.utoronto.ca.*
- *H. Chen and H. Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, 412 58 Gothenburg, Sweden. E-mail: {hui.chen; henkw}@chalmers.se.*

The work of M. Ammous and S. Valaee was financially supported by Huawei Technologies Canada Inc.. The work of H. Chen and H. Wymeersch was supported by the SNS JU project 6G-DISAC under the EU's Horizon Europe research and innovation programme under Grant Agreement No 101139130.

1. We use the terms localization and positioning interchangeably in this paper, as a core function of sensing, indicating the estimation of UE location (or position).

2. Sidelink or D2D communications refer to direct communications between UEs without the data passing through the AP.

introduced sidelink communication can make localization services available in partial-coverage and out-of-coverage scenarios. This is particularly important in environments with severe AP and GNSS signal blockages, such as indoors, in tunnels, or in dense urban areas [2]. Moreover, when a large number of nodes are present within the AP's coverage area, the central localizer on the AP would suffer from significant overhead in managing the positioning of all UEs. In contrast, sidelink communications alleviate this burden by enabling positioning through cooperation among UEs and thus reducing the strain on the AP. In this paper, we argue that with D2D communications and cooperation between multiple UEs, a single RIS anchor can enable AP-free positioning.

1.2 Related Work

In this section, we review related works on RIS-aided/enabled UE positioning, categorizing them into three main groups. First, we consider studies that explore RIS for positioning in the presence of APs. Next, we discuss cooperative positioning systems involving APs. Finally, we review works focused on AP-free localization with the assistance of RISs.

Several recent works have studied the importance of RISs in improving positioning accuracy [15], [16], [17], [18], [19]. In [15], [16], the authors derive the position and orientation error bounds in a RIS-assisted system with one AP and one UE using the Fisher information matrix (FIM). They also show that the positioning error decreases by finding the optimal configuration of the RIS elements. In [17], the authors optimize the configuration of RIS elements by minimizing a localization loss function with received signal strength (RSS) measurements in a multi-user scenario. It has also been shown that the RIS can be used to build a radio map of RSS measurements for localization using fingerprinting and supervised learning [18]. In [19], the authors consider a positioning scenario with multiple RISs and non-line-of-sight (NLoS) links. They propose the design of orthogonal phase shifts at the RISs to differentiate the signals arriving at the UE from different RISs for location estimation. Other studies utilize RISs for localizing and tracking objects, and creating the image of the scene [20], [21], [22]. However, none of these works considers AP-free localization or tracking.

Other works have considered cooperative positioning (CP) with RIS-assisted systems. In [23], [24], the authors have considered positioning systems with one BS, one RIS and multiple UEs with D2D capability. Measurements collected from D2D communications in addition to downlink measurements have been shown to improve positioning performance [23], [24]. However, both works only consider direct D2D links and ignore the D2D links in the UE-RIS-UE paths. On the other hand, the authors in [25] derive the performance limits of a CP system with multiple APs, one RIS and one UE. Again, their results show improvement in positioning accuracy in comparison to non-cooperative methods, and all the mentioned works assume the availability of AP(s).

More recently, it has been argued that RISs can enable localization and mapping without the need for APs [2]. In [11], the authors propose a new use case for the RIS where a full-duplex single-antenna UE can localize itself by leveraging controllable reflections with zero APs (i.e., acting like a radar). Again, in [26], the authors show that a full-duplex multiple-antenna UE can perform simultaneous localization and mapping of the environment with the aid of RIS and zero APs. In [13], the authors propose a CP use case in a scenario with one RIS and two single-antenna UEs. They assume that one UE acts like a radar and transmits pilot signals to the other UE via a direct path and an indirect path through the RIS. The radar-like UE first localizes itself and then localizes the other UE. However, these works on RIS-enabled zero APs localization assume the presence of at least one UE with full-duplex capability, which requires additional hardware costs compared to half-duplex UEs. In addition, the full-duplex UE in these works can determine its position by finding the angle of departure at the RIS and the round-trip time of the reflected signals, which is not possible with a half-duplex UE.

1.3 Contribution

In this paper, we study a novel CP use case of the RIS with D2D communications without the need for APs. We consider a positioning scenario with one RIS and K single-antenna UEs. We build the system model based on a uniform planar array (UPA) architecture at the RIS and orthogonal frequency-division multiplexing (OFDM) signals. The system model proposed in this paper is an extension of the work in [13], where the authors only consider two-dimensional (2D) CP with two UEs, one of them being full-duplex. The main contributions of this work are:

- **RIS-enabled CP problem formulation:** In contrast to [13], we expand the system model to consider a three-dimensional (3D) scenario involving K half-duplex single-antenna UEs with unknown locations and one RIS acting as an anchor. Hence, we leverage D2D communications between the UEs to localize them. Our feasibility analysis reveals that at least three UEs are needed to make the location identifiable, which differs from [11] where only one full-duplex UE is needed (i.e., a hardware upgrade is required). In addition, we study the performance limits of the CP system by deriving the Cramér-Rao lower bound (CRLB) on the channel parameters and the UEs' positions. The derived CRLB will be used to evaluate the proposed estimators as well as the proposed RIS profile designs.
- **RIS profile design:** To assist the estimation of the channel parameters, we decompose the received signal into two components, a LoS path and a RIS path, via orthogonal RIS profiles. To design the phase shifts at the RIS, we propose the use of random phase and directional codebooks. The choice of the codebook depends on whether prior information about the UEs' positions is available or not. Our proposed directional codebook depends on spatial frequencies rather than angles due to the unknown positions of the UEs. In addition, a generalization of the directional codebook involving a power allocation between the UEs is proposed and evaluated.

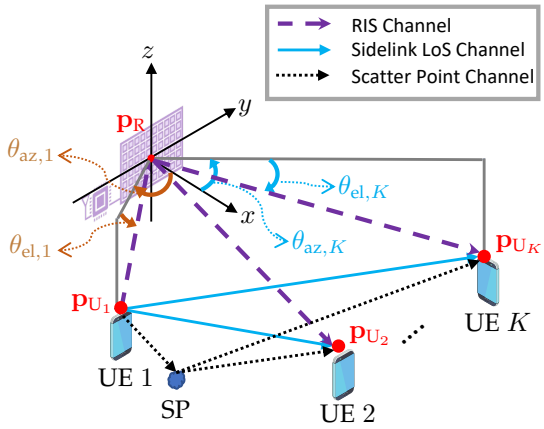


Figure 1. The positioning system with one RIS and K single-antenna UEs showing UE 1 as the transmitter.

- **Positioning algorithm development:** To localize the UEs, we use a two-stage approach. First, we start by estimating the channel parameters for each pair of UEs, including delay estimations (LoS and reflected path through RIS), and spatial frequency estimations. Then, we formulate a simple one-dimensional (1D) coarse search localization algorithm, followed by maximum likelihood estimation (MLE), to obtain the positions of all the UEs.
- **Performance evaluation:** We evaluate the performance of the proposed localization algorithm via extensive simulations. We verify that the proposed estimators attain the CRLB at sufficient transmitting power. We show that a sufficiently large RIS can replace an AP without any performance degradation, and the proposed method can achieve sub-meter level positioning even under multi-path propagation.

The rest of the paper is organized as follows. Section 2 discusses the system setup and the channel model. Then, we derive the CRLB in Section 3. Next, we propose different RIS profile designs in Section 4. We then discuss the estimation of the channel parameters and the proposed localization algorithm in Section 5. Finally, we evaluate the performance of the proposed method via numerical studies in Section 6, followed by Section 7 that concludes this work.

2 Problem Formulation

2.1 System Setup

Consider the 3D scenario shown in Figure 1. The system consists of one RIS and $K \geq 3$ single-antenna UEs. The RIS consists of M passive[3] reflecting elements placed in a UPA facing the x-axis. We assume that the location and orientation of the RIS are known[4] where $\mathbf{p}_{\mathrm{R}} = [x_{\mathrm{R}}, y_{\mathrm{R}}, z_{\mathrm{R}}]^\top$ represents the position of the RIS center as well as the origin of the Cartesian coordinate system. We denote the unknown position of the k-th UE where $k \in \{1, 2, \ldots, K\}$ as $\mathbf{p}_{\mathrm{U}_k} = [x_{\mathrm{U}_k}, y_{\mathrm{U}_k}, z_{\mathrm{U}_k}]^\top$. In addition, we assume that there are S_j scatter points (SPs) for each receiving UE j and the location of SP s seen by the j-th receiving UE is denoted by $\mathbf{p}_{\mathrm{SP}_{j,s}} = [x_{\mathrm{SP}_{j,s}}, y_{\mathrm{SP}_{j,s}}, z_{\mathrm{SP}_{j,s}}]^\top$ where $s \in \{1, 2, \ldots, S_j\}$.

We assume that the UEs communicate with each other directly by utilizing D2D communications at the mmWave band (i.e., frequencies above 24 GHz [30]). As seen in Figure 1, we assume that only one UE is transmitting at a time to reduce interference[5]. We assume that the K UEs are located within a small geographical area, and thus we have a fully connected D2D network. Each receiving UE receives the transmitted signal from three paths: a direct path, a reflected path through the RIS, and reflections from SPs in the environment. These assumptions are consistent with other works on RIS-aided and D2D localization and D2D communications, such as [19], [24], [33], [34], [35], [36], [37].

2.2 Channel Model

As mentioned above, we assume that only one UE is transmitting at a time and that the other UEs are in the receiving mode. Without loss of generality, let us assume that UE i is transmitting OFDM pilot signals with N subcarriers for T consecutive time slots. We also assume that the channel coefficients are fixed during the T transmissions. Furthermore, it is assumed that all pilot symbols transmitted by UE i are equal to $\sqrt{E_i}$, where E_i is the energy of the symbol. Similar to [28], [38], we adopt the geometric channel model to represent the received signal. Then, the signal received at UE j from the transmitting UE i at the discrete times $t = 1, 2, \ldots, T$ can be defined as

$$\mathbf{y}_{ij,t} = \mathbf{y}_{ij,t}^{(\mathrm{LoS})} + \mathbf{y}_{ij,t}^{(\mathrm{RIS})} + \mathbf{y}_{ij,t}^{(\mathrm{MP})} + \mathbf{n}_{j,t}, \tag{1}$$

where $\mathbf{y}_{ij,t}^{(\mathrm{LoS})}$, $\mathbf{y}_{ij,t}^{(\mathrm{RIS})}$ and $\mathbf{y}_{ij,t}^{(\mathrm{MP})}$ represent the received signal at UE j from the transmitting UE i through the direct path, the reflected path through the RIS and the reflected signals from SPs in the environment, respectively. Since the path degradation in mmWave is strong, we ignore the double-bounced paths (e.g., UE i-RIS-SP-UE j channel) and only consider the NLoS components from the UE i-SP-UE j channel [25]. In (1), the vector $\mathbf{n}_{j,t} \in \mathbb{C}^{N\times 1}$ is an additive white Gaussian noise[6] (AWGN) at the j-th receiving UE where $\mathbf{n}_{j,t} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_N)$ with σ^2 being the noise variance and $\mathbf{I}_N$ is the identity matrix. The LoS signal received at the j-th UE from the i-th UE, at time t, can be defined as

$$\mathbf{y}_{ij,t}^{(\mathrm{LoS})} = \sqrt{E_i}\beta_{ij}\mathbf{d}\left(\tau_{ij}\right), \tag{2}$$

3. We focus only on passive RISs because they offer benefits like low power usage and cost-effective deployment. Nevertheless, there are other types of RISs, including active RISs, hybrid RISs, and simultaneously transmitting and reflecting (STAR) RISs [27], which will be explored in future work.

4. The RIS location and orientation can be predetermined via a calibration procedure as discussed in [28], [29], or by using the available map of the environment.

5. Any form of orthogonal waveforms, such as positioning reference signal (PRS) comb signals, could mitigate interference and simplify processing. These waveforms are widely used in positioning systems and are considered the standard due to their ability to separate the signals from different UEs [31]. Alternatively, non-orthogonal waveforms could be employed for pilot transmissions, which theoretically improve positioning performance due to the increased transmission power, given fixed frequency and time resources. However, non-orthogonal waveforms would require power adaptation to enable the receiver to perform sequential interference cancellation [32], which introduces additional complexity that is beyond the scope of this work.

6. The noise here includes both the thermal noise at the receiver and interferences, which can be assumed to vary slowly, with their power estimated as part of the background noise power.

where $\beta_{ij} = \alpha_{ij}e^{j\rho_{ij}}$ is the complex channel gain and τ_{ij} is the transmission delay. The delay steering vector, $\mathbf{d}(\tau)$, is defined as $\mathbf{d}(\tau) = [1, e^{-j2\pi\tau\Delta_f}, \ldots, e^{-j2\pi\tau(N-1)\Delta_f}]^\top$ where Δ_f is the subcarrier spacing and $(\cdot)^\top$ is the transpose operation. In (2), $\tau_{ij} = d_{\mathrm{U}_i\mathrm{U}_j}/c + \Delta_{t_j} - \Delta_{t_i}$ where c is the speed of light, $d_{\mathrm{U}_i\mathrm{U}_j} = \|\mathbf{p}_{\mathrm{U}_i} - \mathbf{p}_{\mathrm{U}_j}\|_2$ represents the Euclidean distance between UEs i and j, and Δ_{t_j} is the clock offset at UE j with respect to a reference UE. Furthermore, the signal arriving from the uncontrolled multi-path propagation, $\mathbf{y}_{ij,t}^{(\mathrm{MP})}$, can be defined as

$$\mathbf{y}_{ij,t}^{(\mathrm{MP})} = \sqrt{E_i}\sum_{s=1}^{S_j}\beta_{ij,s}\mathbf{d}\left(\tau_{ij,s}\right), \tag{3}$$

where S_j represents the total number of SPs seen by the receiving UE j. Here, $\beta_{ij,s} = \alpha_{ij,s}e^{j\rho_{ij,s}}$ and $\tau_{ij,s}$ are the complex channel gain and transmission delay of the uncontrolled path UE i- SP s-UE j, respectively. Again, $\tau_{ij,s} = \left(d_{\mathrm{U}_i\mathrm{SP}_{j,s}} + d_{\mathrm{U}_j\mathrm{SP}_{j,s}}\right)/c + \Delta_{t_j} - \Delta_{t_i}$, where $d_{\mathrm{U}_i\mathrm{SP}_{j,s}} = \|\mathbf{p}_{\mathrm{U}_i} - \mathbf{p}_{\mathrm{SP}_{j,s}}\|_2$ denotes the distance between UE i and SP s seen by the j-th receiving UE.

Finally, the signal received at UE j from UE i through the RIS can be modeled as [9], [38]

$$\mathbf{y}_{ij,t}^{(\mathrm{RIS})} = \sqrt{E_i}\beta_{ij,r}\mathbf{d}\left(\tau_{ij,r}\right)\left(\mathbf{a}\left(\boldsymbol{\vartheta}_i\right)\odot\mathbf{a}\left(\boldsymbol{\vartheta}_j\right)\right)^\top\boldsymbol{\omega}_{i,t}, \tag{4}$$

where $\beta_{ij,r} = \alpha_{ij,r}e^{\rho_{ij,r}}$ is the complex channel gain of the path UE i-RIS-UE j, $\tau_{ij,r}$ is the transmission delay of the same path and $\odot$ represents the Hadamard product. Here, $\tau_{ij,r} = \left(d_{\mathrm{RU}_i} + d_{\mathrm{RU}_j}\right)/c + \Delta_{t_j} - \Delta_{t_i}$, where $d_{\mathrm{RU}_i} = \|\mathbf{p}_{\mathrm{R}} - \mathbf{p}_{\mathrm{U}_i}\|_2$ represents the distance between UE i and the RIS center. The azimuth and elevation angles measured at the RIS center with respect to UE i are represented by the vector $\boldsymbol{\vartheta}_i = [\theta_{az,i}, \theta_{\mathrm{el},i}]^\top$. The angles $\theta_{az,i}$ and $\theta_{el,i}$ are related to the position vector $\mathbf{p}_{\mathrm{U}_i}$ by:

$$\theta_{az,i} = \arctan\left(\frac{y_{\mathrm{U}_i}}{x_{\mathrm{U}_i}}\right), \tag{5}$$

$$\theta_{el,i} = \arcsin\left(\frac{z_{\mathrm{U}_i}}{\sqrt{x_{\mathrm{U}_i}^2 + y_{\mathrm{U}_i}^2 + z_{\mathrm{U}_i}^2}}\right). \tag{6}$$

In (4), the vector $\mathbf{a}(\boldsymbol{\vartheta}_i) \in \mathbb{C}^{M\times 1}$ represents the RIS response vector with respect to the angle vector $\boldsymbol{\vartheta}_i$ and the m-th element of $\mathbf{a}(\boldsymbol{\vartheta}_i)$ is defined as [39]

$$\left[\mathbf{a}\left(\boldsymbol{\vartheta}_i\right)\right]_m = e^{j\frac{2\pi}{\lambda}\mathbf{p}_{\mathrm{R},m}^\top\mathbf{t}(\boldsymbol{\vartheta}_i)}, \tag{7}$$

where λ is the wavelength of the transmitted signal and $\mathbf{p}_{\mathrm{R},m}$ represents the position of the m-th RIS element. In (7), $\mathbf{t}(\boldsymbol{\vartheta}_i)$ is the direction vector of UE i with respect to the RIS center defined as [38], [39]

$$\mathbf{t}(\boldsymbol{\vartheta}_i) = \begin{bmatrix}\cos(\theta_{\mathrm{az},i})\cos(\theta_{\mathrm{el},i})\\ \sin(\theta_{\mathrm{az},i})\cos(\theta_{\mathrm{el},i})\\ \sin(\theta_{\mathrm{el},i})\end{bmatrix}. \tag{8}$$

The coefficient vector at the RIS at time t, when UE i is the transmitter, is defined as

$$\boldsymbol{\omega}_{i,t} = \left[e^{j\omega_{i,t,1}}, \ \ldots, \ e^{j\omega_{i,t,M}}\right]^\top \in \mathbb{C}^{M\times 1}, \tag{9}$$

where $\omega_{i,t,m}$ is the phase shift of the m-th RIS element at time t when UE i is the transmitter. We can rewrite (4) as

$$\mathbf{y}_{ij,t}^{(\mathrm{RIS})} = \sqrt{E_i}\beta_{ij,r}\mathbf{d}\left(\tau_{ij,r}\right)\mathbf{c}^\top\left(\boldsymbol{\gamma}_{ij}\right)\boldsymbol{\omega}_{i,t}, \tag{10}$$

where the m-th element of the intermediate steering vector at the RIS, $\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right) = \mathbf{a}\left(\boldsymbol{\vartheta}_i\right)\odot\mathbf{a}\left(\boldsymbol{\vartheta}_j\right)$, can be computed as

$$\left[\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right]_m = e^{j\frac{2\pi}{\lambda}\mathbf{p}_{\mathrm{R},m}^\top\boldsymbol{\gamma}_{ij}}. \tag{11}$$

In (11), $\boldsymbol{\gamma}_{ij}$ is the spatial frequency vector defined as $\boldsymbol{\gamma}_{ij} = [\kappa_{ij}, \xi_{ij}, \zeta_{ij}]^\top$. Since the RIS is placed in the yz-plane, the first entry of $\mathbf{p}_{\mathrm{R},m}$ equals zero, and hence κ_{ij} is never utilized and will be ignored in the rest of the paper. To avoid angle ambiguity at the RIS, we set the RIS element spacing to be $\lambda/4$, similar to [11]. Hence, we can define the spatial frequencies as [28]

$$\begin{aligned}\xi_{ij} &= \sin\left(\theta_{\mathrm{az},i}\right)\cos\left(\theta_{\mathrm{el},i}\right) + \sin\left(\theta_{\mathrm{az},j}\right)\cos\left(\theta_{\mathrm{el},j}\right),\\ \zeta_{ij} &= \sin\left(\theta_{\mathrm{el},i}\right) + \sin\left(\theta_{\mathrm{el},j}\right).\end{aligned} \tag{12}$$

Consequently, only two spatial frequencies, ξ_{ij} and ζ_{ij}, can be estimated and thus we cannot estimate the angles $\theta_{\mathrm{az},i}$, $\theta_{\mathrm{el},i}$, $\theta_{\mathrm{az},j}$ and $\theta_{\mathrm{el},j}$ directly [28]. From (12), we observe that both ξ_{ij} and ζ_{ij} are in the interval $[-2,\ 2]$.

Given that our system model comprises K single-antenna UEs, we operate within a single-input single-output (SISO) framework. In SISO systems, estimating the transmission delay of multi-path components does not offer additional information for UE positioning. This is because a multiple-antenna array is needed at the transmitter or receiver to determine the SP's location (for angle estimation). Although the same SP may affect multiple UE pairs and make delay-based SP positioning possible, this requires the assumption of a point target and correct data association. Hence, multi-path components are ignored in developing algorithms in this work by dropping $\mathbf{y}_{ij,t}^{(\mathrm{MP})}$ from (1). We evaluate the robustness of our proposed estimators to multi-path propagation in Sections 6.8 and 6.9, where we generate a rich scattering environment based on real-world channel measurements. Consequently, the received signal is simplified to

$$\mathbf{y}_{ij,t} = \mathbf{y}_{ij,t}^{(\mathrm{LoS})} + \mathbf{y}_{ij,t}^{(\mathrm{RIS})} + \mathbf{n}_{j,t}. \tag{13}$$

2.3 Problem Statement

Based on the channel model described above, we define an unknown channel parameter vector, $\boldsymbol{\eta}$, as

$$\boldsymbol{\eta} = \left[\boldsymbol{\eta}_{12}^\top, \boldsymbol{\eta}_{13}^\top, \ldots, \boldsymbol{\eta}_{1K}^\top, \boldsymbol{\eta}_{21}^\top, \boldsymbol{\eta}_{23}^\top, \ldots, \boldsymbol{\eta}_{K,K-1}^\top\right]^\top, \tag{14}$$

which contains $2\binom{K}{2} = (K-1)K$ subvectors. Each subvector $\boldsymbol{\eta}_{ij} \in \mathbb{R}^8$ contains the channel parameters of the channel between the i-th UE (as the transmitter) and the j-th UE (as the receiver) given by

$$\boldsymbol{\eta}_{ij} = \ [\tau_{ij}, \tau_{ij,r}, \xi_{ij}, \zeta_{ij}, \alpha_{ij}, \rho_{ij}, \alpha_{ij,r}, \rho_{ij,r}]^\top. \tag{15}$$

We further define an unknown state vector, $\mathbf{s} \in \mathbb{R}^{4K-1}$, as

$$\mathbf{s} = [\mathbf{p}_{\mathrm{U}_1}^\top, \ldots, \mathbf{p}_{\mathrm{U}_K}^\top, \Delta_{t_1}, \Delta_{t_2}, \ldots, \Delta_{t_{i-1}}, \Delta_{t_{i+1}}, \ldots, \Delta_{t_K}]^\top, \tag{16}$$

where we take UE i as a reference resulting in $\Delta_{t_i} = 0$. By concatenating all the observed signal symbols into the vector $\mathbf{y} = [\mathbf{y}_{12}^\top, \ldots, \mathbf{y}_{1K}^\top, \mathbf{y}_{21}^\top, \ldots, \mathbf{y}_{K,K-1}^\top]^\top$ with $\mathbf{y}_{ij} = [\mathbf{y}_{ij,1}^\top, \ldots, \mathbf{y}_{ij,T}^\top]^\top$, the localization problem can be formulated as follows:

1) Design the RIS coefficient vectors, which depend on whether or not prior information about the UEs' positions is available;
2) Extract the channel parameters from the observed signals;
3) Localization of all the UEs.

2.4 Feasibility Analysis

The complex channel gains will not contribute to estimating the locations of the UEs. By removing the nuisance parameters (i.e., channel gain parameters), we define a channel parameter vector, $\bar{\boldsymbol{\eta}}_{ij} \in \mathbb{R}^4$, as

$$\bar{\boldsymbol{\eta}}_{ij} = [\tau_{ij}, \tau_{ij,r}, \xi_{ij}, \zeta_{ij}]^\top, \tag{17}$$

and $\bar{\boldsymbol{\eta}}$ can be obtained similarly based on (14). Note that $\bar{\boldsymbol{\eta}}_{ji}$ provides the same geometrical information as $\bar{\boldsymbol{\eta}}_{ij}$, which can be treated as multiple measurements (more details in Section 5). Consequently, $\bar{\boldsymbol{\eta}}$ (without the complex channel gains) can provide $4\binom{K}{2} = 2K(K-1)$ geometrical equations to solve $4K-1$ state unknowns. To solve the localization problem, we need to have at least as many geometrical equations as there are state unknowns. This means that we need to satisfy the condition $2K(K-1) \geq 4K-1$, which in turn means that $K=3$ is the minimum requirement for localizability. For more challenging scenarios (e.g., blockage between UE-UE or UE-RIS), more UEs can enable localization by providing more geometrical information (more details in Section 5.5).

2.5 Cooperation Strategy

We consider a localization scenario with at least three out-of-coverage UEs where one UE (a target UE) requires its location information. The UEs communicate with each other via sidelink communications under mode 2, which does not require a coordinating AP [14]. We design an AP-free coordinated architecture as follows, where steps 1, 3, and 5 are similar to [2]:

1) The target UE uses sidelink communication to discover nearby devices such as RIS and other UEs. It then obtains the location information of the RIS.
2) The target UE selects the localization coordinator, which could be itself or one of its neighboring UEs. Due to the unknown positions of all the UEs, the coordinator is chosen randomly, and the target UE shares the RIS location information with the coordinator and other UEs.
3) The coordinator receives a localization request from the target UE.
4) The coordinator sends pilot signals to the RIS controller and other UEs for communication-level synchronization (i.e., to align the phase profile changes at the RIS with the transmitted OFDM symbols).
5) The coordinator assigns the same frequency resources to all the UEs and sets them to transmit in a sequential order, for example, UE 1 is the first transmitter, then UE 2, and so on to avoid interference. The coordinator also triggers the RIS controller to configure the phase shifts of the RIS elements and shares the phase shift information with the other UEs.
6) Each UE transmits the positioning pilot signals for T consecutive time slots within a coherence time block. The total time needed for K UEs to send all pilot signals is KT/Δ_f. This time needs to be short enough to ensure that any UE movement is negligible during the localization epoch. To impose only a small overhead, these signals can be sent over relatively large periods (e.g., every 100 ms).
7) Each receiving UE estimates the channel parameters locally based on the received signals and then shares the channel parameter information with the coordinator.
8) The coordinator estimates its position and the other UEs' positions, and distributes the position information to the target UE and other UEs, if needed.

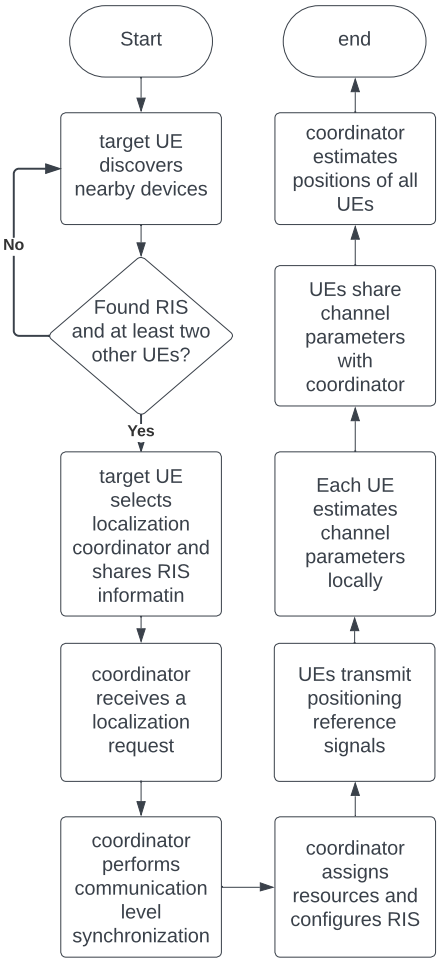


Figure 2. Flow chart of the localization system.

Figure 2 summarizes the steps needed for localization.

3 Derivation of Cramér-Rao Bound

In this section, we start by deriving the FIM for the unknown channel parameter vector $\boldsymbol{\eta}$ defined in (14). Then, the FIM for the UEs' positions can be derived by computing a transformation matrix that links the two matrices. Finally, the CRLB for the estimation error of the UEs' positions is

derived, which serves as a benchmark for the proposed positioning algorithm.

We define the FIM for $\boldsymbol{\eta}$ as $\mathbf{J}_{\boldsymbol{\eta}}$. Similar to [13], [24], the matrix $\mathbf{J}_{\boldsymbol{\eta}}$ can be computed as

$$\mathbf{J}_{\boldsymbol{\eta}} = \sum_{i=1}^{K} \sum_{j \neq i}^{K} \mathbf{J}_{\boldsymbol{\eta}}^{ij}. \tag{18}$$

The (b, v)-th entry of $\mathbf{J}_{\boldsymbol{\eta}}^{ij}$ can be obtained as [40], [41]

$$\left[\mathbf{J}_{\boldsymbol{\eta}}^{ij}\right]_{b,v} = \frac{2}{\sigma^2} \sum_{t=1}^{T} \Re \left\{ \left(\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \eta_b} \right)^H \left(\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \eta_v} \right) \right\}, \tag{19}$$

where the b-th entry of $\boldsymbol{\eta}$ is represented by η_b, $\Re\{\cdot\}$ represents the real part of a complex quantity, $(\cdot)^H$ is the conjugate transpose (Hermitian) operation, and $\boldsymbol{\mu}_{ij,t} = \mathbf{y}_{ij,t}^{(\text{LoS})} + \mathbf{y}_{ij,t}^{(\text{RIS})}$. The partial derivatives needed for computing (19) are derived in Appendix A.

Let us define $\mathbf{J}$ as the FIM associated with the state vector $\mathbf{s}$ (16) and the complex channel gains (i.e., $\alpha_{ij}, \alpha_{ij,r}, \rho_{ij}$ and $\rho_{ij,r}$ $\forall$ $i \neq j$). The matrix $\mathbf{J}$ can be calculated by applying the transformation matrix $\mathbf{T}$ as follows:

$$\mathbf{J} = \mathbf{T} \mathbf{J}_{\boldsymbol{\eta}} \mathbf{T}^\top, \tag{20}$$

where $\mathbf{T} \in \mathbb{R}^{g \times f}$, for $g = 4K - 1 + 4K(K-1)$ and $f = 8K(K-1)$, is defined as $\mathbf{T} = \frac{\partial \boldsymbol{\eta}^\top}{\partial \mathbf{h}}$. The number of columns in $\mathbf{T}$ (i.e., f) equals the length of the vector $\boldsymbol{\eta}$ (14), and $\mathbf{h}$ can be written as

$$\mathbf{h} = \left[\mathbf{s}^\top, \alpha_{ij}, \ldots, \alpha_{ij,r}, \ldots, \rho_{ij}, \ldots, \rho_{ij,r}, \ldots\right]^\top. \tag{21}$$

The length of $\mathbf{h}$ (i.e., g) equals the length of $\mathbf{s}$ (i.e., $4K-1$) in addition to $4K(K-1)$ elements, which are the total number of parameters related to the channel gains. The partial derivatives of the channel parameters with respect to $\mathbf{h}$ are derived in Appendix B. The first $4K-1$ rows (or columns) of $\mathbf{J}$ correspond to the UEs' positions and clock offsets while the next $4K(K-1)$ rows (or columns) correspond to the channel gains.

Finally, the positioning error bound (PEB) of UE i can be calculated as

$$\text{PEB}_i = \sqrt{\text{tr}\left(\left[\mathbf{J}^{-1}\right]_{3i-2:3i, 3i-2:3i} \right)}. \tag{22}$$

The clock offset error bound (CEB) at UE i can be computed as

$$\text{CEB}_i = \sqrt{\left(\left[\mathbf{J}^{-1}\right]_{3K+i, 3K+i} \right)}. \tag{23}$$

4 RIS Profile Design

In this section, we discuss orthogonal RIS profiles and how to tune the phase shifts of the RIS elements via different codebook designs. In addition, we discuss the power allocation between the UEs for a fixed codebook.

4.1 Orthogonal RIS Profiles

To assist the channel parameters estimation in Section 5, we decompose the received signal into two components via orthogonal RIS profiles. Similar to [26], [38], we design the RIS coefficient vectors for half of the transmission time and represent them by $\tilde{\boldsymbol{\omega}}_{i,\tilde{t}}$, where $\tilde{t} = 1, 2, \ldots, \frac{T}{2}$ (T is even). We discuss the choice of the RIS element coefficients in $\tilde{\boldsymbol{\omega}}_{i,\tilde{t}}$ in the following subsections. Without loss of generality, let us define $\boldsymbol{\omega}_{i,2\tilde{t}-1} = \tilde{\boldsymbol{\omega}}_{i,\tilde{t}}$ and $\boldsymbol{\omega}_{i,2\tilde{t}} = -\tilde{\boldsymbol{\omega}}_{i,\tilde{t}}$. By subtracting and adding two consecutive received signals, we obtain the following two quantities assuming that UE i is the transmitter and UE j is the receiver:

$$\begin{aligned} \tilde{\mathbf{y}}_{ij,\tilde{t}}^{(\text{RIS})} &= \frac{1}{2}\left(\mathbf{y}_{ij,2\tilde{t}-1} - \mathbf{y}_{ij,2\tilde{t}}\right) \\ &= \sqrt{E_i}\beta_{ij,r}\mathbf{d}\left(\tau_{ij,r}\right)\mathbf{c}^\top\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\boldsymbol{\omega}}_{i,\tilde{t}} + \tilde{\mathbf{n}}_{j,\tilde{t}}, \end{aligned} \tag{24}$$

$$\begin{aligned} \tilde{\mathbf{y}}_{ij,\tilde{t}}^{(\text{LoS})} &= \frac{1}{2}\left(\mathbf{y}_{ij,2\tilde{t}-1} + \mathbf{y}_{ij,2\tilde{t}}\right) \\ &= \sqrt{E_i}\beta_{ij}\mathbf{d}\left(\tau_{ij}\right) + \tilde{\mathbf{n}}_{j,\tilde{t}}, \end{aligned} \tag{25}$$

where (24) follows from (10) and (13), and (25) follows from (2) and (13). The vector $\tilde{\mathbf{n}}_{j,\tilde{t}} \in \mathbb{C}^{N\times 1}$ is an AWGN at UE j where $\tilde{\mathbf{n}}_{j,\tilde{t}} \sim \mathcal{CN}(0, \frac{\sigma^2}{2}\mathbf{I}_N)$. We can see that we are able to decompose the received signal into two parts. The first part represents the signal arriving through the RIS path, while the second part represents the LoS path. We can concatenate the received signals (24) and (25) over different transmissions into the matrices $\mathbf{Y}_{ij}^{(\text{RIS})}$ and $\mathbf{Y}_{ij}^{(\text{LoS})}$, respectively, which can be defined as

$$\begin{aligned} \mathbf{Y}_{ij}^{(\text{RIS})} &= \left[\tilde{\mathbf{y}}_{ij,1}^{(\text{RIS})},\ \tilde{\mathbf{y}}_{ij,2}^{(\text{RIS})},\ \ldots,\ \tilde{\mathbf{y}}_{ij,\frac{T}{2}}^{(\text{RIS})}\right] \in \mathbb{C}^{N\times\frac{T}{2}}, \\ \mathbf{Y}_{ij}^{(\text{LoS})} &= \left[\tilde{\mathbf{y}}_{ij,1}^{(\text{LoS})},\ \tilde{\mathbf{y}}_{ij,2}^{(\text{LoS})},\ \ldots,\ \tilde{\mathbf{y}}_{ij,\frac{T}{2}}^{(\text{LoS})}\right] \in \mathbb{C}^{N\times\frac{T}{2}}. \end{aligned} \tag{26}$$

We concatenate the RIS coefficient vectors over time for each transmitting UE i into $\tilde{\boldsymbol{\Omega}}_i$, which can be expressed as

$$\tilde{\boldsymbol{\Omega}}_i = \left[\tilde{\boldsymbol{\omega}}_{i,1},\ \tilde{\boldsymbol{\omega}}_{i,2},\ \ldots,\ \tilde{\boldsymbol{\omega}}_{i,\frac{T}{2}}\right] \in \mathbb{C}^{M\times\frac{T}{2}}. \tag{27}$$

In the following subsections, we introduce two practical RIS profile codebooks and a power allocation optimization problem.

4.2 Random Phase Codebook

If no prior information on the UEs' positions is available, a natural way is to set the RIS coefficients as random phases. Since the RIS coefficient vector for the i-th transmitting UE and time slot $\tilde{t}$ is defined as $\tilde{\boldsymbol{\omega}}_{i,t} = [e^{j\tilde{\omega}_{i,\tilde{t},1}}, \ldots, e^{j\tilde{\omega}_{i,\tilde{t},M}}]^\top$, we can choose the phase shift value at each RIS element from a uniform distribution over the interval $[0, 2\pi)$ (i.e., $\tilde{\omega}_{i,\tilde{t},m} \sim \mathcal{U}[0, 2\pi)$, $\forall\ m,\ i,\ \tilde{t}$).

4.3 Directional Codebook

If we have any prior information about the UEs' positions (e.g., from previous estimates with random RIS phases and/or information from extra sources such as GNSS), we can utilize this information to improve the positioning performance [38], [42]. A simple design is to reflect the incident signal on the RIS from the transmitting UE towards the direction of the receiving UE. However, the locations of the transmitting and receiving UEs are unknown, and there

are $K-1$ receiving UEs for each transmitting UE. To create the codebook, we choose the elements ξ, ζ in the vector $\boldsymbol{\gamma}$ based on some prior information, and then the phase shift of each RIS element can be set as the conjugate of (11) (i.e., $e^{-j\frac{2\pi}{\lambda}\mathbf{p}_{\mathrm{R},m}^{\top}\boldsymbol{\gamma}_{ij}}$), named as **directional codebook**.

Let us assume that the locations of the UEs follow a certain distribution based on the prior information (e.g., $\mathbf{p}_{\mathrm{U}_i} \sim \mathcal{N}(\tilde{\mathbf{p}}_{\mathrm{U}_i}, \tilde{\boldsymbol{\Sigma}}_{\mathbf{p}_{\mathrm{U}_i}})$) [43]. Then, for each time slot $\tilde{t}$, we first sample the location of the transmitting UE i. Next, we uniformly sample the index of the receiving UE j from $\{1,2,\ldots,K\}-\{i\}$ (since there are multiple receiving UEs), and then we sample the location of the selected UE j from the prior distribution. Based on the sampled locations of UEs i and j, we can calculate $\boldsymbol{\gamma}_{ij}$ (i.e., ξ_{ij} and ζ_{ij}) using (5), (6) and (12). The RIS profile can be written as

$$[\tilde{\boldsymbol{\omega}}_{i,\tilde{t}}]_m = e^{-j\frac{2\pi}{\lambda}\mathbf{p}_{\mathrm{R},m}^{\top}\boldsymbol{\gamma}_{ij}} \ \forall m, \tag{28}$$

where $[\tilde{\boldsymbol{\omega}}_{i,\tilde{t}}]_m$ is the phase shift at the m-th RIS element for the i-th transmitting UE and time slot $\tilde{t}$. All RIS elements use the same value of $\boldsymbol{\gamma}_{ij}$ during a transmission time slot to maximize the signal energy towards a certain direction.

4.4 Power Allocation Optimization

The positioning performance can be improved by optimizing the power allocation between the UEs. Let us define $\boldsymbol{P} = [P_1, P_2, \ldots, P_K]^{\top}$ as the power allocation vector, where $P_i = NE_i$ is the transmitted power of UE i. We assume that there is a constraint on the total transmitted power in the system such that $\sum_{i=1}^{K} P_i = P_{\text{tot}}$, where P_{tot} represents the total power transmitted in the system during each time slot. The relationship between the positioning error and the transmitted power is apparent since the symbol energy E_i appears in all derivatives needed to calculate the FIM (19). For a fixed RIS codebook, we minimize the average squared PEB of the involved UEs by formulating the following optimization problem[7]:

$$\begin{aligned} \min_{\boldsymbol{P}} \ & \frac{1}{K}\sum_{i=1}^{K}(\mathrm{PEB}_i)^2 \\ \text{s.t.} \ & \sum_{i=1}^{K} P_i = P_{\text{tot}}, \\ & 0 < P_i, \end{aligned} \tag{29}$$

where PEB_i is defined in (22). The first constraint in (29) ensures that the total transmitted power equals P_{tot} and the second constraint ensures that each UE has a non-zero transmission power. The optimization problem (29) is convex as discussed in [44] and thus can be solved using any convex optimization toolbox. One limitation with the optimization problem (29) is that the objective function (i.e., $\mathrm{PEB}_i \ \forall \ i$) requires the positions of all the UEs, which are unknown. Assuming that we have the prior information of all the UEs following multi-variable Gaussian distributions similar to Section 4.3, we can then calculate PEB_i as $\mathrm{PEB}_i(\tilde{\mathbf{p}}_{\mathrm{U}_1}, \ldots, \tilde{\mathbf{p}}_{\mathrm{U}_i}, \ldots, \tilde{\mathbf{p}}_{\mathrm{U}_K}, \tilde{\boldsymbol{\Omega}}_1, \ldots, \tilde{\boldsymbol{\Omega}}_K)$, where $\tilde{\mathbf{p}}_{\mathrm{U}_i}$ is the mean of the prior information of UE i and $\tilde{\boldsymbol{\Omega}}_i$ is obtained from the directional codebook. This is not guaranteed to be optimal, but it is a practical solution for the RIS profile design.

7. Note that this is not the only way of formulating the power allocation problem. The objective function can also be expressed as minimizing the worst PEB.

5 Proposed Estimators

In this section, we start by estimating the delays and spatial frequencies based on the received signals. Then, we develop a localization algorithm to estimate the UEs' positions.

5.1 Delay Estimation

We start by estimating the transmission delays of the signals reflected by the RIS (i.e., $\tau_{ij,r} \ \forall \ i, \ j, \ i \neq j$). Coarse estimation of the delays can be computed by applying the inverse fast Fourier transform (IFFT) on the matrix $\mathbf{Y}_{ij}^{(\mathrm{RIS})}$. Let us define $\mathbf{F} \in \mathbb{C}^{N_f \times N}$ as the IFFT matrix where N_f represents the length of the IFFT vector, which is a design parameter. The (l,h)-th entry of the IFFT matrix can be computed as

$$[\mathbf{F}]_{l,h} = \frac{1}{N_f} e^{j2\pi lh/N_f}. \tag{30}$$

Similar to [11], [23], the coarse estimates of the transmission delays can be calculated as

$$\bar{\tau}_{ij,r} = \frac{l_{ij,r}^{*}}{N_f \Delta f}, \tag{31}$$

where $l_{ij,r}^{*}$ is an integer that can be computed as follows [11], [23]:

$$l_{ij,r}^{*} = \arg\max_{l} \left\| \mathbf{e}_l^{\top} \left(\mathbf{F}\mathbf{Y}_{ij}^{(\mathrm{RIS})} \right) \right\|_2^2. \tag{32}$$

In (32), $\mathbf{e}_l$ is a column vector of length N_f with all elements equal to zero except the l-th entry that has a value of 1. In (31), we assume that the delay estimate ($\bar{\tau}_{ij,r}$) is an integer multiple of the time resolution $\frac{1}{N_f \Delta f}$. To refine the delay estimate, we first define an intermediate variable $\delta_{ij,r} \in [0, \frac{1}{N_f \Delta f})$. Then, the optimal value of the intermediate variable can be calculated by [38]:

$$\bar{\delta}_{ij,r} = \arg\max_{\delta_{ij,r}} \left\| \mathbf{e}_{l_{ij,r}^{*}}^{\top} \mathbf{F} \left(\mathbf{Y}_{ij}^{(\mathrm{RIS})} \odot \mathbf{D}(\delta_{ij,r}) \right) \right\|_2^2, \tag{33}$$

where $\mathbf{D}(\delta_{ij,r})$ is defined as

$$\mathbf{D}(\delta_{ij,r}) = \underbrace{\left[\mathbf{d}(\delta_{ij,r}), \ \ldots, \ \mathbf{d}(\delta_{ij,r}) \right]}_{\frac{T}{2} \text{ times}} \in \mathbb{C}^{N \times \frac{T}{2}}. \tag{34}$$

The optimization problem (33) can be solved using a quasi-Newton method by setting the starting point to 0. Next, the refined delay estimate can be obtained as

$$\tilde{\tau}_{ij,r} = \bar{\tau}_{ij,r} - \bar{\delta}_{ij,r}. \tag{35}$$

We observe that the delays $\tau_{ij,r}$ and $\tau_{ji,r}$ provide the same geometrical information. Due to the channel reciprocity and coherence assumption, we can simply average the two estimates as

$$\hat{\tau}_{ij,r} = \frac{\tilde{\tau}_{ij,r} + \tilde{\tau}_{ji,r}}{2}. \tag{36}$$

Since $\tau_{ij,r} = \left(d_{\mathrm{RU}_i} + d_{\mathrm{RU}_j}\right)/c + \Delta_{t_j} - \Delta_{t_i}$ and $\tau_{ji,r} = \left(d_{\mathrm{RU}_i} + d_{\mathrm{RU}_j}\right)/c + \Delta_{t_i} - \Delta_{t_j}$, the clock offsets cancel out when we take the average of the two estimates. Thus, we

do not need to estimate the clock offsets anymore (we can think of this as two-way communication).

Similarly, we can estimate the delay of the LoS path (i.e., $\hat{\tau}_{ij}\ \forall\ i,\ j,\ i \neq j$) using (31)–(36) by replacing $\mathbf{Y}_{ij}^{(\text{RIS})}$, $\bar{\tau}_{ij,r}$, $\tilde{\tau}_{ij,r}$, $\hat{\tau}_{ij,r}$, $l_{ij,r}^{*}$ and $\bar{\delta}_{ij,r}$, with $\mathbf{Y}_{ij}^{(\text{LoS})}$, $\bar{\tau}_{ij}$, $\tilde{\tau}_{ij}$, $\hat{\tau}_{ij}$, l_{ij}^{*} and $\bar{\delta}_{ij}$, respectively.

5.2 Spatial Frequency Estimation

In this section, we estimate the spatial frequencies ξ_{ij} and $\zeta_{ij}\ \forall\ i,\ j,\ i \neq j$. We start by eliminating the effect of the transmission delay on the received signal through the RIS path by computing:

$$\mathbf{Y}_{ij}^{(\text{a})} = \mathbf{Y}_{ij}^{(\text{RIS})} \odot \mathbf{D}\left(-\tilde{\tau}_{ij,r}\right). \tag{37}$$

We then take the sum of all rows of the matrix in (37) (i.e., taking the sum over all subcarriers) and we get the following quantity:

$$\tilde{\mathbf{y}}_{ij}^{(\text{r})} = {\mathbf{Y}_{ij}^{(\text{a})}}^{\top}\mathbf{1}_N = N\sqrt{E_i}\beta_{ij,r}\tilde{\mathbf{\Omega}}_i^{\top}\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right) + \tilde{\mathbf{q}}, \tag{38}$$

where $\mathbf{1}_N$ is a column vector of length N with all entries equal to 1, $\tilde{\mathbf{q}} \in \mathbb{C}^{\frac{T}{2}\times 1}$ represents the noise component and $\tilde{\mathbf{\Omega}}_i$ is defined in (27). By ignoring the noise component in (38), we can evaluate $\beta_{ij,r}$ as a function of $\boldsymbol{\gamma}_{ij}$ and obtain:

$$\tilde{\beta}_{ij,r}\left(\boldsymbol{\gamma}_{ij}\right) = \frac{\mathbf{c}^H\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\mathbf{\Omega}}_i^{*}\tilde{\mathbf{y}}_{ij}^{(\text{r})}}{N\sqrt{E_i}\left\|\mathbf{c}^H\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\mathbf{\Omega}}_i^{*}\right\|_2^2}, \tag{39}$$

where $(\cdot)^*$ is the conjugate operation[8]. Finally, we can estimate ξ_{ij} and ζ_{ij} by solving the following optimization problem:

$$\begin{aligned}\tilde{\xi}_{ij},\ \tilde{\zeta}_{ij} = \underset{\xi_{ij},\ \zeta_{ij}}{\arg\min}&\left\|\tilde{\mathbf{y}}_{ij}^{(\text{r})} - N\sqrt{E_i}\tilde{\beta}_{ij,r}\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\mathbf{\Omega}}_i^{\top}\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right\|_2^2\\ \text{s.t.}\quad &-2 \leq \xi_{ij},\ \zeta_{ij} \leq 2,\end{aligned} \tag{40}$$

where the constraints follow from (12). Coarse estimates of ξ_{ij} and ζ_{ij} can be obtained by solving the optimization problem above by searching over a 2D grid, whose area can also be reduced with prior UE position information. Once coarse estimates are computed, we can refine the estimates by utilizing a quasi-Newton method and using the coarse estimates as the starting point. By looking at (12) and due to the reciprocity of the channel, we observe that ξ_{ij} and ξ_{ji} are equal. This equality also applies to ζ_{ij} and ζ_{ji}. Hence, we can average the two estimates as

$$\begin{aligned}\hat{\xi}_{ij} &= \frac{\tilde{\xi}_{ij} + \tilde{\xi}_{ji}}{2},\\ \hat{\zeta}_{ij} &= \frac{\tilde{\zeta}_{ij} + \tilde{\zeta}_{ji}}{2}.\end{aligned} \tag{41}$$

To summarize, for every transmitting UE i and receiving UE j, we estimate four channel parameters (two delays and two spatial frequencies).

8. We derive (39) in Appendix C.

5.3 Localization Algorithm

To estimate the UEs' positions, we first need to compute the angles at the RIS center with respect to each UE using the estimated spatial frequencies. Next, we obtain coarse position estimates, which are then refined through a maximum likelihood estimation (MLE) process, with the coarse estimates serving as initial values[9].

5.3.1 Angle Estimation

The estimated spatial frequencies can be concatenated as

$$\begin{aligned}\hat{\boldsymbol{\xi}} &= \left[\hat{\xi}_{12}, \ldots, \hat{\xi}_{ij}, \ldots\right]^{\top},\\ \hat{\boldsymbol{\zeta}} &= \left[\hat{\zeta}_{12}, \ldots, \hat{\zeta}_{ij}, \ldots\right]^{\top},\end{aligned} \tag{42}$$

where each of the vectors $\hat{\boldsymbol{\xi}}$ and $\hat{\boldsymbol{\zeta}}$ has $K(K-1)/2$ elements. For example, if $\hat{\boldsymbol{\xi}}$ contains $\hat{\xi}_{12}$, it will not contain $\hat{\xi}_{21}$ since both $\hat{\xi}_{12}$ and $\hat{\xi}_{21}$ are equal due to the averaging in (41). We can rewrite (12) as

$$\begin{aligned}\xi_{ij} &= w_{1,i} + w_{1,j},\\ \zeta_{ij} &= w_{2,i} + w_{2,j},\end{aligned} \tag{43}$$

where $w_{1,i} = \sin\left(\theta_{\text{az},i}\right)\cos\left(\theta_{\text{el},i}\right)$ and $w_{2,i} = \sin\left(\theta_{\text{el},i}\right)$ which can be concatenated into the following two vectors:

$$\begin{aligned}\mathbf{w}_1 &= \left[w_{1,1}, w_{1,2}, \ldots, w_{1,K}\right]^{\top},\\ \mathbf{w}_2 &= \left[w_{2,1}, w_{2,2}, \ldots, w_{2,K}\right]^{\top}.\end{aligned} \tag{44}$$

Based on (43), we can relate $\boldsymbol{\xi}$ and $\mathbf{w}_1$, and $\boldsymbol{\zeta}$ and $\mathbf{w}_2$ as follows:

$$\begin{aligned}\boldsymbol{\xi} &= \mathbf{G}\mathbf{w}_1,\\ \boldsymbol{\zeta} &= \mathbf{G}\mathbf{w}_2,\end{aligned} \tag{45}$$

where $\mathbf{G} \in \mathbb{R}^{\frac{(K-1)K}{2}\times K}$, with the k-th row corresponding to the k-th element in $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$. For example, for the k-th element ξ_{ij}, the entries i and j in the k-th row of the matrix $\mathbf{G}$ are set as 1 with all other entries as zeros. Based on the estimated spatial frequency vectors $\hat{\boldsymbol{\xi}}$ and $\hat{\boldsymbol{\zeta}}$, we can estimate $\hat{\mathbf{w}}_1$ and $\hat{\mathbf{w}}_2$ using the least squares solution, which requires $K \geq 3$, as [45]

$$\begin{aligned}\hat{\mathbf{w}}_1 &= \left(\mathbf{G}^{\top}\mathbf{G}\right)^{-1}\mathbf{G}^{\top}\hat{\boldsymbol{\xi}},\\ \hat{\mathbf{w}}_2 &= \left(\mathbf{G}^{\top}\mathbf{G}\right)^{-1}\mathbf{G}^{\top}\hat{\boldsymbol{\zeta}}.\end{aligned} \tag{46}$$

Consequently, the angle information of UE i can be estimated as

$$\hat{\theta}_{\text{az},i} = \arcsin\left(\hat{w}_{1,i}/\sqrt{1-\hat{w}_{2,i}^2}\right), \tag{47}$$

$$\hat{\theta}_{\text{el},i} = \arcsin\left(\hat{w}_{2,i}\right), \tag{48}$$

which can be concatenated into the vector $\hat{\boldsymbol{\vartheta}}_i = [\hat{\theta}_{\text{az},i}, \hat{\theta}_{\text{el},i}]^{\top}$.

9. This is a standard approach in positioning literature (e.g., [38]). While position estimation can be performed directly through MLE without coarse estimates, the use of coarse position estimates as initial values significantly enhances the efficiency and robustness of the MLE process. This is because MLE for positioning typically involves solving a non-linear, non-convex optimization problem, which can suffer from slow convergence or convergence to local optima if the initial guess is far from the true position. Coarse estimates, often obtained using simpler search methods, provide a reasonable starting point that reduces the search space for MLE and improves convergence speed.

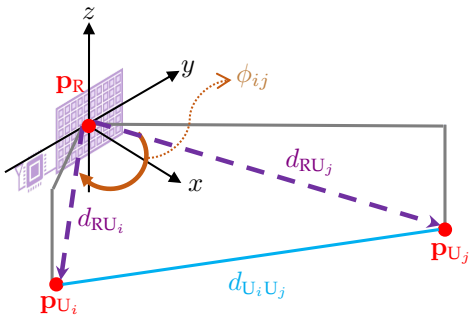


Figure 3. Every two UEs i and j along with the RIS center forms a triangle. ϕ_{ij} represents the angle between the two directional vectors pointing towards each of UE i and UE j.

5.3.2 Coarse Estimation of UEs' Positions

Based on the angle and delay estimates, we design a 1D search algorithm for UE positioning along the direction of the coordinating UE. By taking UE i as our coordinator, for a given candidate distance between UE i and the RIS (denoted as $\check{d}_{\mathrm{RU}_i}$), a candidate location for the UE will be:

$$\check{\mathbf{p}}_{\mathrm{U}_i} = \check{d}_{\mathrm{RU}_i}\mathbf{t}(\hat{\boldsymbol{\vartheta}}_i), \tag{49}$$

where $\mathbf{t}(\hat{\boldsymbol{\vartheta}}_i)$ is the direction vector as defined in (8). Each pair of UEs (e.g., i and j), together with the RIS center, forms a triangle as shown in Figure 3. Based on the law of cosines, the relationship between the i-th UE, the j-th UE ($j \neq i$), and the RIS is given by [46]:

$$d^2_{\mathrm{RU}_i} + d^2_{\mathrm{RU}_j} - 2d_{\mathrm{RU}_i}d_{\mathrm{RU}_j}\cos(\phi_{ij}) = d^2_{\mathrm{U}_i\mathrm{U}_j}, \tag{50}$$

where d_{RU_i}, d_{RU_j} and $d_{\mathrm{U}_i\mathrm{U}_j}$ denote the distances between UE i and the RIS, between UE j and the RIS, and between UE i and UE j, respectively. The angle between the directional vectors $\mathbf{t}(\boldsymbol{\vartheta}_i)$ and $\mathbf{t}(\boldsymbol{\vartheta}_j)$ is represented by ϕ_{ij}. Since $\cos(\phi_{ij}) = \mathbf{t}^\top(\boldsymbol{\vartheta}_i)\mathbf{t}(\boldsymbol{\vartheta}_j)$, we rewrite (50) as

$$d^2_{\mathrm{RU}_i} + d^2_{\mathrm{RU}_j} - 2d_{\mathrm{RU}_i}d_{\mathrm{RU}_j}\mathbf{t}^\top(\boldsymbol{\vartheta}_i)\mathbf{t}(\boldsymbol{\vartheta}_j) = d^2_{\mathrm{U}_i\mathrm{U}_j}. \tag{51}$$

In (50) and (51), $d_{\mathrm{U}_i\mathrm{U}_j}$ can be calculated using the delay of the LoS path as $c\tau_{ij}$. However, d_{RU_i} and d_{RU_j} cannot be calculated directly from the delay estimates. Using the transmission delay of the RIS path, we can calculate $d_{\mathrm{RU}_i} + d_{\mathrm{RU}_j}$ as $c\tau_{ij,r}$. In (51), we can show that $d^2_{\mathrm{U}_i\mathrm{U}_j}$ is equivalent to $(c\tau_{ij,r} - c\tau_{ij} - d_{\mathrm{RU}_i} - d_{\mathrm{RU}_j})^2$. For a given $\check{d}_{\mathrm{RU}_i}$ and by replacing $d^2_{\mathrm{U}_i\mathrm{U}_j}$ with its equivalent expression, we can solve (51) for $\check{d}_{\mathrm{RU}_j}$ as

$$\check{d}_{\mathrm{RU}_j} = \frac{2(c\hat{\tau}_{ij,r} - c\hat{\tau}_{ij})\check{d}_{\mathrm{RU}_i} - (c\hat{\tau}_{ij,r} - c\hat{\tau}_{ij})^2}{2\check{d}_{\mathrm{RU}_i}(1 + \mathbf{t}^\top(\hat{\boldsymbol{\vartheta}}_i)\mathbf{t}(\hat{\boldsymbol{\vartheta}}_j)) + 2(c\hat{\tau}_{ij} - c\hat{\tau}_{ij,r})}. \tag{52}$$

Then, the candidate position of all the other UEs can be obtained as

$$\check{\mathbf{p}}_{\mathrm{U}_j} = \check{d}_{\mathrm{RU}_j}\mathbf{t}(\hat{\boldsymbol{\vartheta}}_j), \quad j \neq i. \tag{53}$$

The final estimation of the distance between the i-th UE and the RIS can be calculated by formulating a cost function as

$$d^*_{\mathrm{RU}_i} = \underset{d_{\mathrm{RU}_i}}{\arg\min} \sum_{\substack{j,k \\ j\neq k\neq i}} \left(\check{\psi}_{jk} - c(\hat{\tau}_{jk,r} - \hat{\tau}_{jk})\right)^2, \tag{54}$$

where $\check{\psi}_{jk} = \|\check{\mathbf{p}}_{\mathrm{U}_j} - \mathbf{p}_{\mathrm{R}}\|_2 + \|\check{\mathbf{p}}_{\mathrm{U}_k} - \mathbf{p}_{\mathrm{R}}\|_2 - \|\check{\mathbf{p}}_{\mathrm{U}_j} - \check{\mathbf{p}}_{\mathrm{U}_k}\|_2$. The cost function (54) calculates the mismatch between the distance measurements based on the location estimates and the distance measurements based on the delay estimates for each pair of UEs not including the coordinating UE i. Once the optimal $d^*_{\mathrm{RU}_i}$ is found, the rest of the state parameters can be calculated based on (52) and (53). The 1D search along the direction of the coordinating UE i provides coarse estimates of the UEs' positions, which will be refined in the next section. We can do the 1D search using distance values given by the vector $\mathbf{d}_{\text{search}}$ which has a total of L elements. The steps of the coarse estimation of UEs' positions are summarized in Algorithm 1. We can then concatenate the outputs of Algorithm 1 into the vector $\bar{\mathbf{s}}$ defined as[10]

$$\bar{\mathbf{s}} = \left[\bar{\mathbf{p}}^\top_{\mathrm{U}_1}, \ldots, \bar{\mathbf{p}}^\top_{\mathrm{U}_K}\right]^\top. \tag{55}$$

Algorithm 1 Coarse Estimation of UEs' positions

Input: $\hat{\boldsymbol{\vartheta}}_j, \hat{\tau}_{jk}, \hat{\tau}_{jk,r}\ \forall\ j,\ k,\ j \neq k, i$ (index of coordinating UE) and $\mathbf{d}_{\text{search}}$

Output: $\bar{\mathbf{p}}_{\mathrm{U}_j}\ \forall\ j$

$d^*_{\mathrm{RU}_i} \leftarrow \underset{d_{\mathrm{RU}_i}\in\mathbf{d}_{\text{search}}}{\arg\min} \sum_{\substack{j,k \\ j\neq k\neq i}} \left(\check{\psi}_{jk} - c(\hat{\tau}_{jk,r} - \hat{\tau}_{jk})\right)^2$

$\bar{\mathbf{p}}_{\mathrm{U}_i} \leftarrow d^*_{\mathrm{RU}_i}\mathbf{t}(\hat{\boldsymbol{\vartheta}}_i)$

for $j = 1$ to K **do**

  if $(j \neq i)$ **then**

    $\bar{d}_{\mathrm{RU}_j} \leftarrow$ calculate using (52) and $d^*_{\mathrm{RU}_i}$

    $\bar{\mathbf{p}}_{\mathrm{U}_j} \leftarrow \bar{d}_{\mathrm{RU}_j}\mathbf{t}(\hat{\boldsymbol{\vartheta}}_j)$

  end if

end for

return $\bar{\mathbf{p}}_{\mathrm{U}_j}\ \forall\ j$

5.3.3 Maximum Likelihood Estimation

With the coarse estimation of the state parameters (i.e., $\bar{\mathbf{s}}$), refinement procedures can be performed by solving an MLE problem. The objective function can be formulated as

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\arg\min}\ (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}(\mathbf{s}))^\top \boldsymbol{\Sigma}^{-1}_{\boldsymbol{\eta}} (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}(\mathbf{s})), \tag{56}$$

where $\hat{\boldsymbol{\eta}}$ is the estimated nuisance-free channel parameter vector defined as $\hat{\boldsymbol{\eta}} = [\hat{\boldsymbol{\eta}}^\top_{12}, \ldots, \hat{\boldsymbol{\eta}}^\top_{ij}, \ldots]^\top$ and $\hat{\boldsymbol{\eta}}_{ij} = [\hat{\tau}_{ij}, \hat{\tau}_{ij,r}, \hat{\xi}_{ij}, \hat{\zeta}_{ij}]^\top$ (i.e., two delay estimates and two spatial frequency estimates for each pair of UEs i and j obtained from Sections 5.1 and 5.2). The vector $\boldsymbol{\eta}(\mathbf{s})$, defined similar to $\hat{\boldsymbol{\eta}}$, is the mapping from the state parameters to the channel parameters where for example, we calculate $\xi_{ij}(\mathbf{s})$ and $\zeta_{ij}(\mathbf{s})$ by substituting the relevant information from $\mathbf{s}$ in (5), (6) and (12). In (56), $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$ is the covariance matrix of the estimated channel parameters. If the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$ is not available, we replace it with the identity matrix and thus (56) becomes a least squares optimization problem. In general, the optimization problem (56) can be solved by standard multi-variable optimization methods, such as the trust-region method [47] and using the coarse estimate of the state parameters (i.e., $\bar{\mathbf{s}}$ in (55)) as an initial point.

10. Note that the clock offsets are not part of the state vector anymore since they cancel out when we average the delay estimates and they do not affect the positioning algorithm.

5.4 Complexity Analysis

In this section, we study the computational complexity of the proposed estimators discussed in Sections 5.1–5.3. The system model, as described in Section 2.3, encompasses a total of $K(K-1)$ pairs of transmitters and receivers. For each transmitter-receiver pair, we perform two 1D N_f-point FFTs to perform coarse delay estimation, resulting in a complexity of $\mathcal{O}(K^2TN_f \log N_f)$. The refinement process of coarse delay estimates, evaluated by (33), has a complexity of $\mathcal{O}(K^2I_1N_fNT)$, where I_1 denotes the number of iterations required. Furthermore, for every transmitter-receiver pair, a 2D search is conducted to determine the coarse estimates of ξ_{ij} and ζ_{ij}, leading to a complexity of $\mathcal{O}(K^2I_2I_3TM)$, where I_2 and I_3 represent the search dimensions. To refine spatial frequency estimates, a complexity of $\mathcal{O}(K^2I_4TM)$ is incurred, where I_4 is the number of iterations. Furthermore, the angle estimation process exhibits a complexity of $\mathcal{O}(K^3)$ due to matrix inversion in (46). The coarse estimation of UEs' positions carries a complexity of $\mathcal{O}(LK^2)$, where L denotes the size of the search vector $\mathbf{d}_{\text{search}}$. Finally, the MLE refinement process, employing an identity covariance matrix, entails a complexity of $\mathcal{O}(K^2I_5)$, where I_5 is the number of iterations. Thus, the overall complexity of the proposed estimators is given by

$$\begin{aligned}\mathcal{O}_{\text{total}} &= \mathcal{O}(K^2TN_f \log N_f) + \mathcal{O}(K^2I_1N_fNT) + \mathcal{O}(K^3) \\ &+ \mathcal{O}(K^2I_2I_3TM) + \mathcal{O}(K^2I_4TM) + \mathcal{O}(LK^2) + \mathcal{O}(K^2I_5)\end{aligned} \tag{57}$$

This expression highlights that system parameters such as the number of subcarriers (N), the number of transmissions per UE (T), and the number of RIS elements (M) contribute to the computational complexity of the channel parameter estimators linearly, while the number of UEs involved affects most complexity terms quadratically, except for the matrix inversion term, which is affected cubically. With more UEs involved in the positioning process and possibly more frequency and/or time resources, the positioning performance will improve, as will be shown in Figure 9. However, the computational complexity will increase.

5.5 Extension to Non-fully Connected Network

Although the proposed localization algorithm relies on a fully connected network, it can be extended to non-fully connected scenarios. In a non-fully connected scenario with disconnected UEs, additional functionalities are required, such as an algorithm to identify blockages or limitations preventing direct communication between the UEs. This blockage information is crucial for determining whether the remaining network is sufficient for localization. A non-fully connected network of UEs is localizable as long as the number of geometric equations provided by the channel parameter estimates is greater than or equal to the number of state unknowns. The following steps can be applied to a localizable non-fully connected network of K UEs:

1) Identify which links are blocked using a blockage identification tool (e.g., a deep neural network).
2) Estimate the transmission delays and spatial frequencies of the unblocked links using the proposed channel estimators in Sections 5.1 and 5.2.
3) Divide the non-fully connected network of K UEs into fully connected subnetworks, provided that each subnetwork consists of at least three UEs.
4) Apply the coarse localization algorithm proposed in Sections 5.3.1 and 5.3.2 to each of the fully connected subnetworks of UEs. The results of the subnetworks can then be averaged.
5) If some UEs do not belong to any fully connected subnetwork, they can be localized using the estimated positions of the UEs within the fully connected subnetworks as anchors.
6) The results can be further refined using the MLE in (56).

Table 1
Simulation Parameters

Parameter	Value
Carrier Frequency	28 GHz
Wavelength	$\lambda = 1$ cm
RIS array size	$M = 11 \times 11$ elements
Speed of light	$c = 3 \times 10^8$ m/s
Number of UEs	$K = 3$
Number of transmissions per UE	$T = 40$
Total number of transmission	$KT = 120$
RIS element spacing	$\lambda/4 = 0.25$ cm
Number of subcarriers	$N = 3000$
subcarrier spacing	$\Delta_f = 120$ kHz
Symbol duration	8.33 μs
Noise figure	$n_f = 8$ dB
Noise Power Spectral Density	$N_0 = -174$ dBm/Hz
Noise variance	$\sigma^2 = n_f N_0 \Delta_f = -115.2$ dBm
Transmission power per UE	$P_i = NE_i = 23$ dBm
Total transmission power	$P_{\text{tot}} = 27.8$ dBm
Length of IFFT	$N_f = 10N$

6 Simulation Results

In this section, we evaluate the performance of the proposed RIS-enabled CP algorithm, benchmarked by the derived CRLB. We start by describing the simulation environment and then we discuss the numerical results generated over different system parameters.

6.1 Simulation Setup

We consider the scenario shown in Figure 1 as our simulation environment. We assume that there are three UEs that need to be localized. There is also one RIS placed along the yz-plane. The center of the RIS (i.e., $\mathbf{p}_{\text{R}}$) is located at the origin $[0,0,0]^\top$ m. The dimensions of the simulation environment are $[0,7]$ m, $[-3.5,3.5]$ m and $[-2,2]$ m on the x-axis, y-axis and z-axis, respectively. These dimensions align with those commonly found in indoor RIS-aided localization scenarios documented in the literature, such as [9], [17], [28]. The three UEs: UE 1, UE 2 and UE 3 are located at $\mathbf{p}_{\text{UE}_1} = [4,3,-1]^\top$ m, $\mathbf{p}_{\text{UE}_2} = [4.5,1,-0.5]^\top$ m and $\mathbf{p}_{\text{UE}_3} = [5,-3,-1]^\top$ m, respectively. The distances between the RIS and each of the UEs are $d_{\text{RU}_1} \approx 5$ m, $d_{\text{RU}_2} \approx 4.6$ m and $d_{\text{RU}_3} \approx 6$ m. In addition, the distances between each pair of UEs are $d_{\text{U}_1\text{U}_2} \approx 2$ m, $d_{\text{U}_1\text{U}_3} \approx 6$ m and $d_{\text{U}_2\text{U}_3} \approx 4$ m. The communication channel parameters are summarized in Table 1 which are consistent with the simulation parameters in [23], [38]. The total transmission power in the system P_{tot}

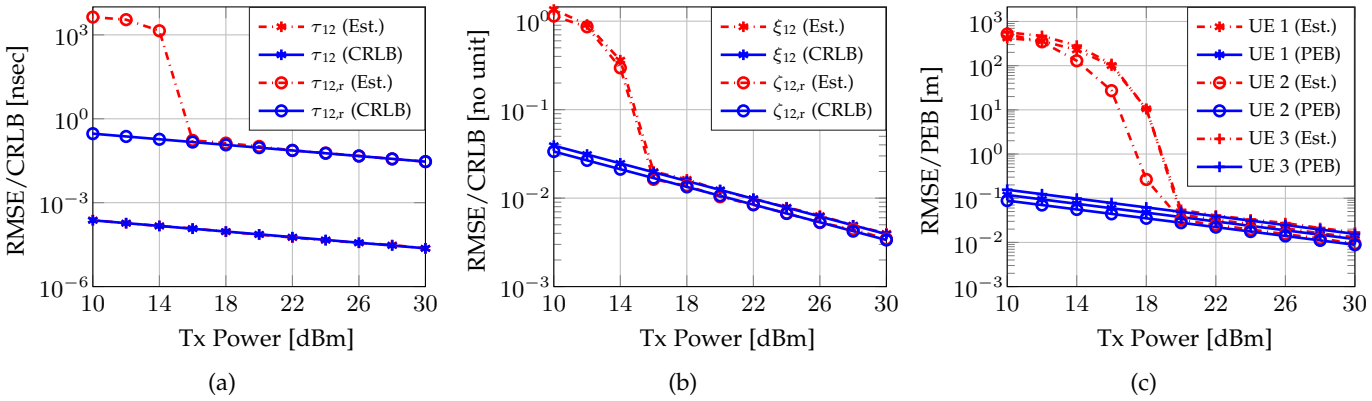


Figure 4. Comparison of the RMSE and CRLB/PEB for (a) transmission delays, (b) spatial frequencies, and (c) position estimates, as a function of the transmission power.

is set to 600 mW or equivalently 27.8 dBm. We assume that P_{tot} is evenly distributed between the three UEs (i.e., $P_i = 200$ mW or 23 dBm $\forall\ i$). The number of SPs is set to zero (i.e., $S_j = 0\ \forall j$) unless otherwise stated. The channel gain of each path in the simulations is calculated as [21], [48]

$$\beta_{ij} = \alpha_{ij} e^{j\rho_{ij}} = \frac{\lambda}{4\pi d_{\text{U}_i\text{U}_j}} e^{-j\frac{2\pi}{\lambda} d_{\text{U}_i\text{U}_j}}, \tag{58}$$

$$\beta_{ij,r} = \alpha_{ij,r} e^{j\rho_{ij,r}} = \frac{\lambda^2}{16\pi^2 d_{\text{RU}_i} d_{\text{RU}_j}} e^{-j\frac{2\pi}{\lambda}\left(d_{\text{RU}_i} + d_{\text{RU}_j}\right)}, \tag{59}$$

$$\begin{aligned}\beta_{ij,s} &= \alpha_{ij,s} e^{j\rho_{ij,s}} \\ &= \sqrt{\frac{\sigma_{\text{rcs}}^2}{4\pi}} \frac{\lambda}{4\pi d_{\text{U}_i\text{SP}_{j,s}} d_{\text{U}_j\text{SP}_{j,s}}} e^{-j\frac{2\pi}{\lambda}\left(d_{\text{U}_i\text{SP}_{j,s}} + d_{\text{U}_j\text{SP}_{j,s}}\right)},\end{aligned} \tag{60}$$

where (58)–(60) are the channel gains for the LoS channel, the RIS channel, and the multi-path channel, respectively. The radar cross-section (RCS) σ_{rcs}^2 indicates the reflection capability of the NLoS path in terms of m^2. Note that $\beta_{i,j}$, $\beta_{ij,r}$ and $\beta_{ij,s}$ are unknown to be estimated. Finally, we use the root-mean-squared error (RMSE) as a metric for evaluation.

6.2 Effect of Transmission Power on Positioning Performance

In this section, we evaluate the performance of the proposed estimators by comparing their performance to the CRLB as a function of the transmission power. We vary the transmission power from 6 dBm to 30 dBm for the three UEs. The RIS coefficient vectors are designed using the random phase codebook. The lower bounds are computed using one realization of the RIS profile. We average the results of the estimators over 200 noise realizations.

For ease of illustration, we only focus on the case where UE 1 is the transmitter and UE 2 is the receiver to evaluate the delay and spatial frequency estimators since there are many channel parameters involved and the results are similar. Figure 4(a) shows the performance of the transmission delay estimator for the LoS path and the RIS path (i.e., τ_{12} and $\tau_{12,r}$) in comparison to the CRLB. As seen in the figure, the LoS delay estimation error attains the CRLB for all transmission power values due to the large channel gain value as observed in (58). However, in the case of the RIS path delay, the channel gain value (59) is much lower and thus there is a high estimation error when the power is below 16 dBm since the RIS signal has a low signal-to-noise ratio (SNR). The RMSE of the RIS path delay gets very close to the CRLB when the transmission power is 16 dBm or higher. Figure 4(b) compares the estimator RMSE values of the spatial frequencies ξ_{12} and ζ_{12} with the CRLB. Since the spatial frequencies are only included in the RIS path, we see a similar performance to the RIS path delay, where the error is relatively large if the power is below 16 dBm and small otherwise.

Figure 4(c) shows the RMSE of the positioning algorithm for all three UEs in comparison to the PEBs. Again, due to the large estimation error of the channel parameters when the power is below 16 dBm, the positioning errors are also large. The positioning errors are in the range of hundreds of meters with low transmission power since we do not account for the geometry of the environment in our proposed localization algorithm. On the other hand, the positioning error gets close to the PEB for all the UEs when the transmission power is 18 dBm or higher. We can see from Figure 4(c) that the proposed algorithm can provide centimeter-level accuracy at 20 dBm transmission power. As depicted in Figure 4, the effectiveness of the estimators proposed in Sections 5.1–5.3 is evident, as their estimation errors consistently reach the lower bounds when the transmission power is larger than 20 dBm.

6.3 Comparison with Other Systems

In this section, we compare the performance of our proposed localization system with other systems. We would like to highlight that, to the best of our knowledge, this paper is the first to consider RIS-enabled CP with only RIS as an anchor. As a comparison scenario, we replace the RIS with an AP at the same location and refer to this scenario as AP-aided CP. In the AP-aided CP scenario, each UE estimates the LoS transmission delays from its neighboring UEs. Furthermore, the AP estimates the delay and angle of arrival (AoA) from each UE. As another comparison scenario, we consider the RIS-enabled self-localization system proposed in [11] where each UE localizes itself via reflections from RIS with zero APs. However, in [11], each UE must have a full-duplex capability, which is a different hardware requirement compared to our proposed system. For the RIS-enabled CP and RIS-enabled self-localization

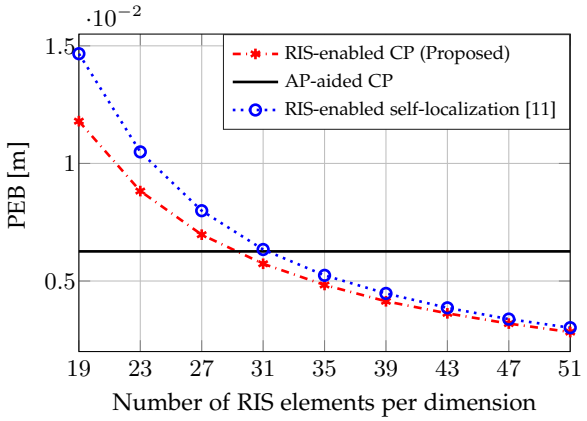


Figure 5. The average PEB of the three UEs under different localization systems for different RIS sizes.

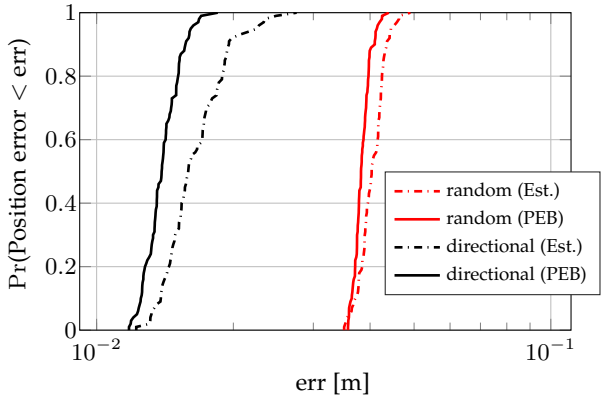


Figure 6. ECDF of the positioning error of UE 3 for the lower bound and the estimator for two RIS profiles.

scenarios, we consider the same parameters described in Section 6.1. In the AP-aided CP scenario, the AP parameters are 10 dBm transmission power, 3×3 array, and 12 transmissions. In all scenarios, we fix the total number of transmissions to 120. The number of transmissions per UE is 40 in the two scenarios involving RIS, while it is 36 per UE in the AP scenario. In the AP-aided CP case, we assume that the UEs communicate with each other directly using the same parameters as in the RIS-enabled CP scenario. We use the random phase codebook for the RIS profile design.

Figure 5 shows the average PEB of the three UEs under different scenarios as a function of the number of RIS elements per dimension. Here, we vary the RIS size from 19×19 elements to 51×51 elements. For accurate representation, each value is averaged over 200 Monte-Carlo simulations. We can see that the PEB of the AP scenario is a constant line since there is no RIS involved. We observe that our proposed system outperforms the RIS-enabled self-localization system. It is also clear that the PEBs of the RIS-enabled scenarios decrease as the RIS size increases. With a sufficiently large RIS (31×31 array), we can replace the AP with a RIS without performance degradation. The far-field condition is satisfied when the distance between the RIS center and the closest UE is much greater than the maximum distance between the RIS center and one RIS element. In other words, when

$$\min_i \|\mathbf{p}_{\mathrm{R}} - \mathbf{p}_{\mathrm{U}_i}\|_2 \gg \max_m \|\mathbf{p}_{\mathrm{R}} - \mathbf{p}_{\mathrm{R},m}\|_2. \tag{61}$$

When the RIS size is 31×31, the far-field condition is met. When the RIS size is very large (e.g., 500×500), the far-field condition is not satisfied anymore, and near-field localization needs to be considered which is beyond the scope of this paper.

6.4 Effect of RIS Profile and Power Allocation on Positioning Performance

In this section, we evaluate the positioning performance under different RIS profiles and different power allocation strategies. We consider two cases for the phase shifts at the RIS: random phase codebook and directional codebook.

Figure 6 shows the empirical cumulative distribution function (ECDF) for both the estimator and the PEB for two RIS profiles. For ease of illustration, we only show the performance of UE 3. Here, we set the number of RIS elements to $M = 121$ and the total transmission power is evenly distributed between the UEs. Prior information about the UEs' positions is needed in the case of the directional codebook. We assume that the distribution of the prior information of UE i is normal with $\mathcal{N}(\mathbf{p}_{\mathrm{U}_i}, \mathbf{\Sigma}_{\mathbf{p}_{\mathrm{U}_i}})$. We set $\mathbf{\Sigma}_{\mathbf{p}_{\mathrm{U}_i}} = 1.5\mathbf{I}_3$ in Figure 6. To plot the ECDF, we generate 100 different codebooks in each case. To accurately evaluate the estimator performance, we take the average of 200 noise realizations for each codebook. It is clear that the directional codebook outperforms the random phase codebook since the energy is directed toward the UEs' positions.

To evaluate the power allocation strategy on the positioning performance, we define $P_2 = \epsilon P_1$ and $P_3 = \upsilon P_1$ where ϵ, $\upsilon > 0$ are scaling factors. Then, $P_1 + P_2 + P_3 = (1+\epsilon+\upsilon)P_1 = P_{\mathrm{tot}}$ and thus $P_1 = 1/(1+\epsilon+\upsilon)P_{\mathrm{tot}}$. Hence, the transmission power of each UE can be defined in terms of ϵ, υ, and P_{tot}. We observe that P_{tot} is evenly distributed between the UEs when $\epsilon = \upsilon = 1$. Here, we consider two cases for the UEs' positions. In the first case, we use the positions defined in Section 6.1. In the second case, we move UE 3 from $\mathbf{p}_{\mathrm{UE}_3} = [5, -3, -1]^\top$ m to $\mathbf{p}_{\mathrm{UE}_3} = [1, -1, -1]^\top$ m (i.e., closer to the RIS) while we keep the positions of UEs 1 and 2 the same as in the first case. Figures 7(a) and 7(b) show the average PEB of the three UEs as a function of the power scaling factors ϵ and υ. Here, we set $P_{\mathrm{tot}} = 600$ mW and use one realization of the directional profile with $\mathbf{\Sigma}_{\mathbf{p}_{\mathrm{U}_i}} = 0.5\mathbf{I}_3$. As seen in both figures, it is clear that there are three corners with high average PEB: bottom left (power allocated to UE 1 only), bottom right (power allocated to UE 2 only) and top left (power allocated to UE 3 only). In Figure 7(a), we observe a small area centered around $\epsilon = 1$ and $\upsilon = 1$ where the PEB value is small, suggesting that a uniform power allocation between the UEs is near-optimal in the first case. However, in Figure 7(b), the values of both ϵ and υ are less than 1 in the optimal PEB area. This suggests that uniform power allocation is not optimal in the second case, where we observe that UE 1 is allocated more power because it is the farthest from the RIS.

Next, we study how the uncertainty of the prior information of the UEs' positions affects the average PEB. To study the effect of the uncertainty, we define the covariance matrix of the distribution of the prior information as $\mathbf{\Sigma}_{\mathbf{p}_{\mathrm{U}_i}} = \sigma^2_{\mathrm{unc}}\mathbf{I}_3$ and we vary σ^2_{unc} from 10^{-3} to 100 m^2. Here, there are

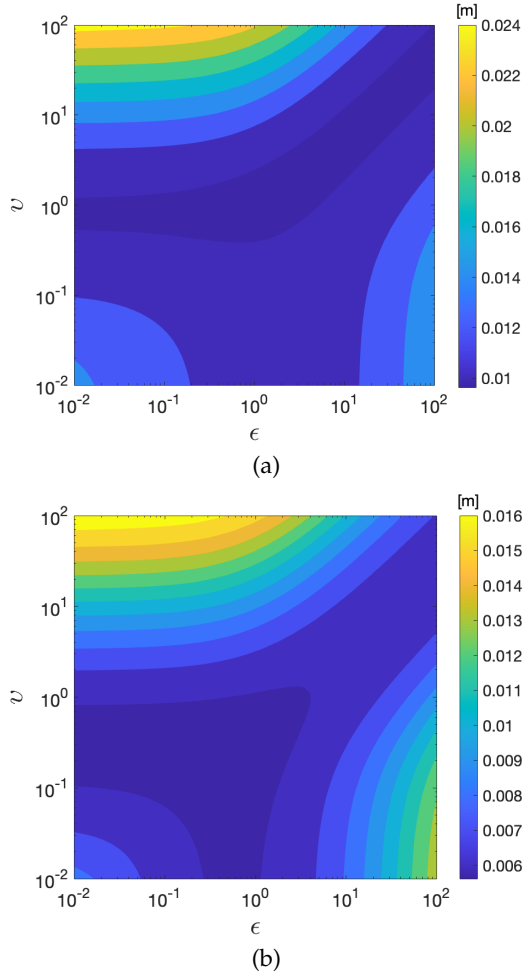


Figure 7. The average PEB of the three UEs as a function of the power scaling factors ϵ and υ for a fixed RIS profile for two cases of UEs' positions. In both cases, $\mathbf{p}_{\text{UE}_1} = [4, 3, -1]^\top$ m and $\mathbf{p}_{\text{UE}_2} = [4.5, 1, -0.5]^\top$ m. In (a), $\mathbf{p}_{\text{UE}_3} = [5, -3, -1]^\top$ m. In (b), $\mathbf{p}_{\text{UE}_3} = [1, -1, -1]^\top$ m.

$M = 121$ reflecting elements at the RIS and we use the UEs' positions defined in Section 6.1. Figure 8 shows the average PEB for random and directional RIS profiles as the uncertainty of the prior information increases. Again, the results are averaged over 100 different RIS profiles. We use uniform power allocation for the random codebook, while we adopt different power allocation strategies for the directional codebook. We observe that the random codebook is a constant line since it does not depend on the prior information. We can see from the figure that the directional profile performance is relatively poor when the uncertainty value is large. The positioning performance of the directional profile improves when the uncertainty level is low. However, when the uncertainty value is too small, the performance of the directional profile is worse than that of the random profile. This happens because when the prior error is too small, the RIS reflects the beams to small areas and hence there is less spatial diversity. This could be mitigated by finding an optimal RIS profile by jointly optimizing (29) over transmission powers and RIS coefficient vectors. However, this joint optimization problem is non-convex and will be left for future work. Figure 8 illustrates that the positioning performance is notably influenced by the selected power allocation strategy. It is evident that the optimal power allocation strategy defined in (29) yields the best positioning performance, with uniform power allocation trailing closely behind.

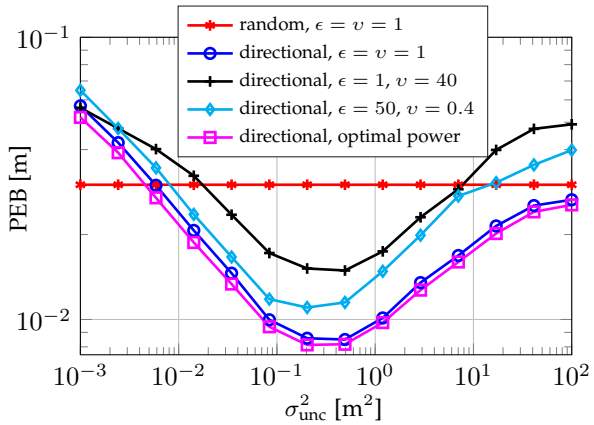


Figure 8. The average PEB of the three UEs for random and directional profiles as a function of the uncertainty of the prior information with different power allocation strategies.

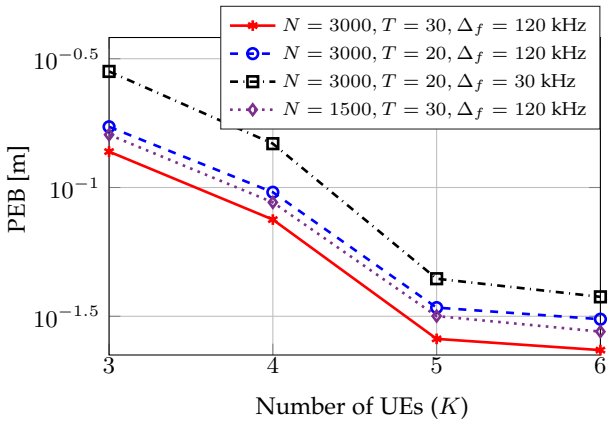
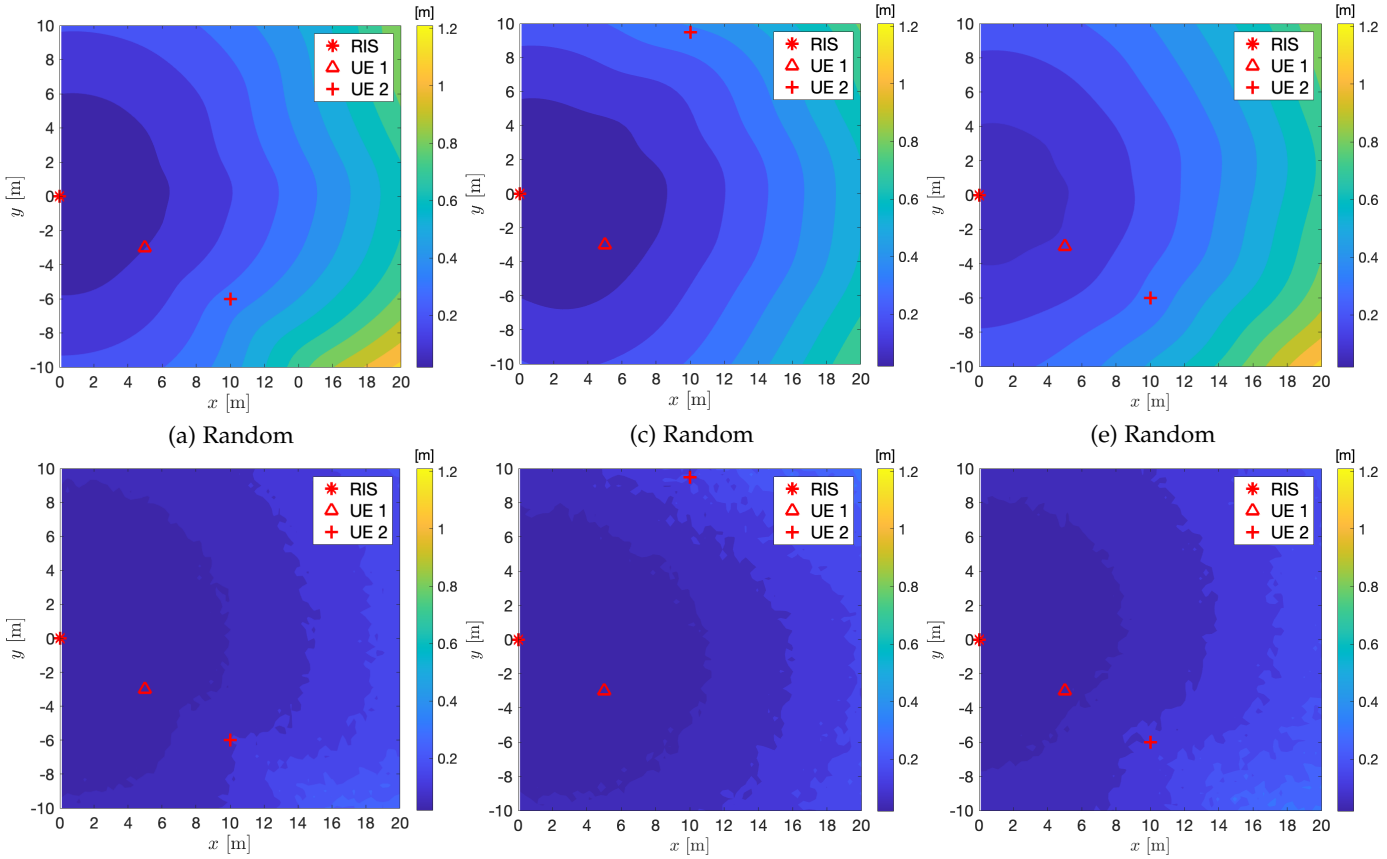


Figure 9. The PEB of UE 1 as a function of the number of UEs involved in the localization under different parameter settings.

6.5 Effect of Number of UEs on Positioning Performance

In this section, we study the effect of the number of UEs involved in the localization algorithm on the positioning performance for the random RIS profile under different system parameters. Here, there are 121 reflecting elements at the RIS, and the results are averaged over 200 different codebook realizations.

Figure 9 shows the PEB of UE 1 as the number of involved UEs increases for various parameter settings. The parameters considered in this scenario are the number of subcarriers (N), the number of transmissions per UE (T), and the subcarrier spacing (Δ_f). The positions of the UEs are randomly generated over the area $[1, 11]$ m, $[-6, 6]$ m, and $[-2, 2]$ m on the x-axis, y-axis, and z-axis, respectively. We observe that the localization accuracy improves with the involvement of more UEs across all cases. However, this improvement comes at the expense of increased computational complexity, as evident from Section 5.4. For example, with $T = 30$ and $\Delta_f = 120$ kHz, the total number of transmissions is 90 for three UEs (equivalent to 0.75 ms in duration), while it doubles to 180 (equivalent to 1.5 ms



(a) Random (c) Random (e) Random

(b) Directional (d) Directional (f) Directional

Figure 10. The PEB of UE 3 under different placements of UEs 1 and 2 for both random and directional profiles with the RIS located at $\mathbf{p}_{\mathrm{R}} = [0, 0, 0]^\top$ m. In (a) and (b), $\mathbf{p}_{\mathrm{UE}_1} = [5, -3, -1.5]^\top$ m, $\mathbf{p}_{\mathrm{UE}_2} = [10, -6, -1]^\top$ m and $\mathbf{p}_{\mathrm{UE}_3} = [x, y, 1]^\top$ m. In (c) and (d), $\mathbf{p}_{\mathrm{UE}_1} = [5, -3, -1.5]^\top$ m, $\mathbf{p}_{\mathrm{UE}_2} = [10, 9.5, -1]^\top$ m and $\mathbf{p}_{\mathrm{UE}_3} = [x, y, 1]^\top$ m. In (e) and (f), $\mathbf{p}_{\mathrm{UE}_1} = [5, -3, -1.5]^\top$ m, $\mathbf{p}_{\mathrm{UE}_2} = [10, -6, -1]^\top$ m and $\mathbf{p}_{\mathrm{UE}_3} = [x, y, 4]^\top$ m.

in duration) for six UEs. Furthermore, for a fixed number of UEs, increasing either N or T enhances the positioning performance due to the larger number of observations available. Finally, with a fixed N, increasing Δ_f improves the positioning performance due to better range resolution.

6.6 Effect of UEs' Positions on Positioning Performance

To study the effect of the UEs' positions on the positioning performance, we fix the RIS location at $\mathbf{p}_{\mathrm{R}} = [0, 0, 0]^\top$ m. We also fix the locations of UEs 1 and 2 while we vary the location of UE 3 and calculate its PEB. In this section, we consider three scenarios where we vary the location of UE 3 in the intervals $[0.1, 20]$ m and $[-10, 10]$ m along the x-axis and y-axis, respectively. In the first scenario, we place UE 1, UE 2 and UE 3 at $\mathbf{p}_{\mathrm{UE}_1} = [5, -3, -1.5]^\top$ m, $\mathbf{p}_{\mathrm{UE}_2} = [10, -6, -1]^\top$ m and $\mathbf{p}_{\mathrm{UE}_3} = [x, y, 1]^\top$ m, respectively. In this section, the first scenario is used as a benchmark. In the second scenario, we move UE 2 to $\mathbf{p}_{\mathrm{UE}_2} = [10, 9.5, -1]^\top$ (i.e., farther away from UE 1) and we keep the same parameters for UEs 1 and 3 as in the first scenario. In the third scenario, we increase the height of UE 3 and place it at $\mathbf{p}_{\mathrm{UE}_3} = [x, y, 4]^\top$ m, while we keep the locations of UEs 1 and 2 the same as in the first scenario. For each scenario, we consider two RIS profiles: random phase codebook and directional codebook. Finally, for the directional profile, we use $\mathbf{\Sigma}_{\mathbf{p}_{\mathrm{U}_i}} = 0.5\mathbf{I}_3$.

Figures 10(a) and 10(b) show the PEB for UE 3 under the first scenario for random and directional profiles, respectively. We observe that the positioning performance is mainly affected by the distance between UE 3 and the RIS in both figures. The results of the second scenario are shown in Figures 10(c) and 10(d) for both profiles. Compared to Figures 10(a) and 10(b), Figures 10(c) and 10(d) show slightly better performance in terms of the PEB, which is due to the placement of UE 2 farther away from UE 1 resulting in better coverage. Figures 10(e) and 10(f) show the results of the third scenario. In comparison to Figures 10(a) and 10(b), Figures 10(e) and 10(f) show slightly poorer positioning performance, which happens because UE 3 is placed at a higher height (i.e., larger distances from the RIS and both UEs). As expected, the directional profile outperforms the random one in the three scenarios. RIS-aided localization systems might suffer from blind areas [28], [29]. However, Figure 10 shows that the proposed use case does not suffer from blind areas and the localization error is mainly affected by the distance from the RIS and the other UEs. Finally, the study depicted in Figure 10 can be extended to cover a larger localization area. However, as UE 3 moves away from the RIS and other UEs, the positioning performance may degrade. In such scenarios, improving performance

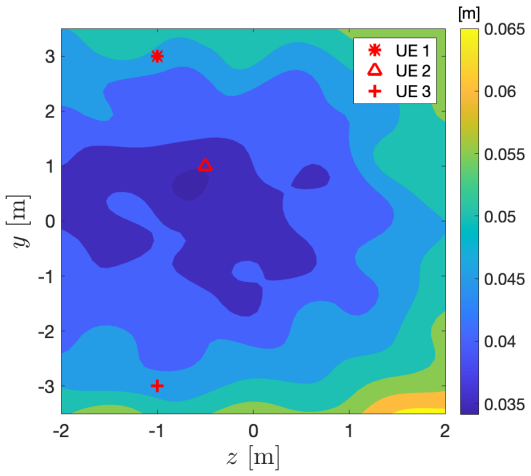


Figure 11. The PEB of UE 3 as the RIS location varies while the UEs' positions are fixed.

is possible by increasing the transmission power and/or increasing the RIS size.

6.7 Effect of RIS Position on Positioning Performance

To study the effect of the RIS position on the positioning performance, we fix the locations of the three UEs at $\mathbf{p}_{\text{UE}_1} = [4, 3, -1]^\top$ m, $\mathbf{p}_{\text{UE}_2} = [4.5, 1, -0.5]^\top$ m and $\mathbf{p}_{\text{UE}_3} = [5, -3, -1]^\top$ m. We place the RIS along the yz-plane at $\mathbf{p}_{\text{R}} = [0, y, z]^\top$ m and we vary its y and z values. Figure 11 shows the PEB of UE 3 as the location of the RIS varies on the plane. We can see that the optimal PEB area is centered at $y \approx 0.5$ m and $z \approx -0.5$ m. This center is approximately the average of the UE's positions on the y-axis and z-axis. The performance degrades when the $|z|$ value is high (i.e., high or low elevation) or when $|y|$ is relatively large.

6.8 Effect of Multi-path on Positioning Performance

In this section, we study the effect of the multi-path on the positioning performance. We consider the same setup as in Section 6.1 and randomly generate four SPs for each receiving UE (i.e., $S_j = 4\ \forall j$). We also consider three different values for the RCS which are $\sigma^2_{\text{rcs}} = 1,\ 10$ and 30 m^2. Here, we consider 100 realizations of the random codebook and we take the average of 200 noise realizations for each codebook. For ease of illustration, we only consider the positioning error of UE 3.

In Figure 12, the solid red curve represents the ECDF of the PEB while the dashed red curve represents the ECDF of the RMSE of the estimator with zero SPs involved. The positioning performance degrades if there are SPs in the environment and worsens as the RCS increases. However, even with $\sigma^2_{\text{rcs}} = 30$ m^2, sub-meter level positioning can still be achieved.

6.9 Evaluation on Real-World Multi-path Channel Statistics

In this section, we study the effect of the multi-path on the channel estimation using channel statistics obtained from a real-world experiment. We use the indoor scenario described in [37], which involves a transmitter, a receiver, and a 20×20 RIS located at $[6.1, 11.7, 0]^\top$ m, $[7.1, 10, 0]^\top$ m and $[4, 8.5, 0]^\top$ m, respectively. The dataset obtained from [37] contains the angle, delay and channel gain information of the 20 strongest multi-path components for the transmitter-receiver and transmitter-RIS-receiver channels. We use the channel statistics obtained from the dataset, together with our channel model described in (1)–(4), to generate data in a more realistic and rich scattering environment. We then estimate the LoS and RIS delays, denoted as τ_{LoS} and τ_{RIS}, and we also estimate the spatial frequencies ξ and ζ. Here, we consider 100 realizations of the random codebook each averaged over 200 noise realizations.

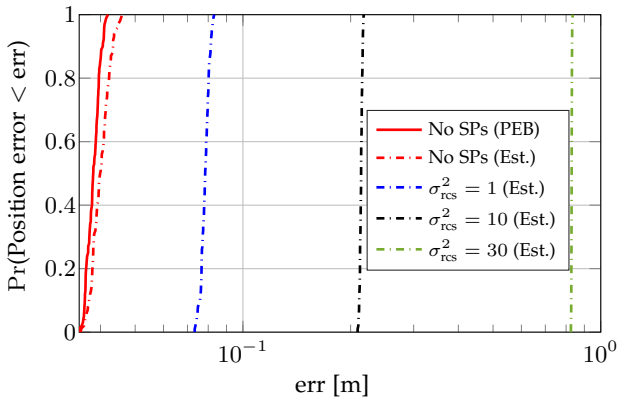
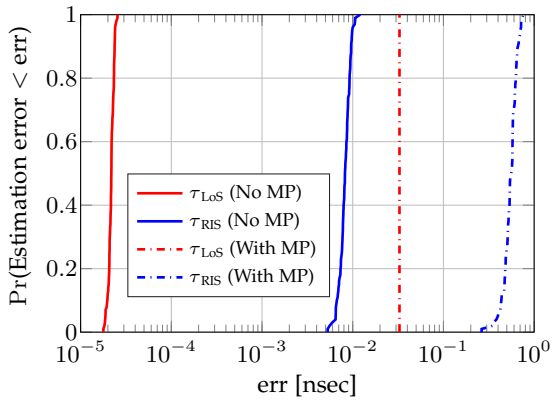


Figure 12. ECDF of the positioning error of UE 3 under multi-path scenario with $S_j = 4\ \forall j$ and different RCS values.

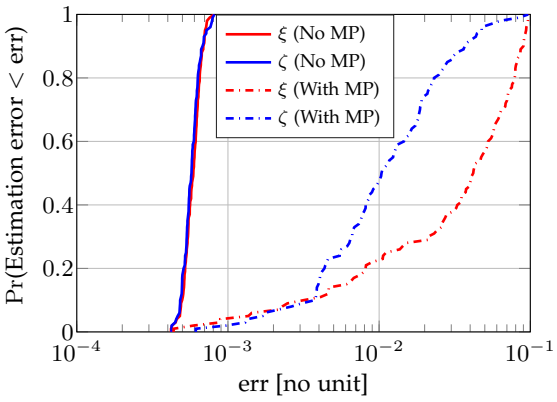
Figure 13(a) shows the ECDF of the estimation error for both τ_{LoS} and τ_{RIS}. In the figure, the solid curves consider the scenario of no scattering, while the dashed curves consider a rich scattering environment. We observe that the estimation accuracy degrades with scattering in the environment. However, the delay estimation errors are still below 1 nsec or 30 cm in distance. We observe a similar trend with the spatial frequencies, as shown in Figure 13(b), where the maximum estimation error is around 0.1 (at a level comparable to the results of Figure 4). Since the proposed localization algorithm in Section 5.3 needs at least three transmitters/receivers, we cannot apply it directly to this scenario. We can compute the positioning error in this scenario by assuming that the transmitter location is known. The azimuth and elevation angles of the receiver with respect to the RIS center can be computed directly from (12) using the estimated ξ and ζ and the known azimuth and elevation angles of the transmitter with respect to the RIS. The distance between the RIS and the receiver can be calculated directly from τ_{RIS} since the distance between the transmitter and the RIS is known. The 90-th percentile error values of the channel parameters in Figures 13(a) and 13(b) result in approximately 70 cm of positioning error, which is also consistent with the results shown in Figure 12 with σ^2_{rcs} between 10 m^2 and 30 m^2.

7 Conclusion

In this paper, we studied a 3D CP scenario with one RIS and K UEs to estimate the locations of the UEs without using an AP via sidelink communications. We started by introducing the channel model and deriving the lower bounds of the channel parameters and position estimates. We then introduced two practical RIS profiles and formulated a power



(a)



(b)

Figure 13. ECDF of the estimation error for (a) delays and (b) spatial frequencies.

allocation optimization problem for a fixed RIS profile to minimize the average PEB. We also developed three low-complexity estimators to estimate the transmission delays, spatial frequencies, and positions of the UEs. Through numerical studies, we showed that a sufficiently large RIS could replace an AP without any performance degradation. The proposed localization algorithm can achieve sub-meter level accuracy even under multi-path scenarios. We also evaluated the positioning performance under different RIS profiles and showed that directional profiles resulted in the best performance when the uncertainty of the prior information was low and the power allocation was optimal. Further extension to this work can consider synchronization protocols, overhead calculation and optimization, simultaneous pilot transmissions, robustness improvement to non-standard noises, and partial blockage between the UEs. Finally, a critical step forward would be the development and testing of a real-world prototype to validate the effectiveness of the proposed positioning system.

REFERENCES

[1] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1124–1148, 2018.

[2] H. Chen, H. Kim, M. Ammous, G. Seco-Granados, G. C. Alexandropoulos, S. Valaee, and H. Wymeersch, "RISs and sidelink communications in smart cities: The key to seamless localization and sensing," *IEEE Communications Magazine*, vol. 61, no. 8, pp. 140–146, 2023.

[3] Z. Z. M. Kassas, J. Khalife, K. Shamaei, and J. Morales, "I hear, therefore I know where I am: Compensating for GNSS limitations with cellular signals," *IEEE Signal Processing Magazine*, vol. 34, no. 5, pp. 111–124, 2017.

[4] J. A. Zhang, M. L. Rahman, K. Wu, X. Huang, Y. J. Guo, S. Chen, and J. Yuan, "Enabling joint communication and radar sensing in mobile networks—A survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 306–345, 2022.

[5] A. Behravan, V. Yajnanarayana, M. F. Keskin, H. Chen, D. Shrestha, T. E. Abrudan, T. Svensson, K. Schindhelm, A. Wolfgang, S. Lindberg *et al.*, "Positioning and sensing in 6G: Gaps, challenges, and opportunities," *IEEE Vehicular Technology Magazine*, vol. 18, no. 1, pp. 40–48, 2023.

[6] Z. Chu, W. Hao, P. Xiao, and J. Shi, "Intelligent reflecting surface aided multi-antenna secure transmission," *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 108–112, 2020.

[7] M. Javad-Kalbasi, M. S. Al-Abiad, and S. Valaee, "Energy efficient communications in RIS-assisted UAV networks based on genetic algorithm," in *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, 2023, pp. 5901–5906.

[8] X. Gan, C. Zhong, C. Huang, and Z. Zhang, "RIS-assisted multi-user MISO communications exploiting statistical CSI," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6781–6792, 2021.

[9] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications," *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 135–158, 2022.

[10] X. Gan, C. Huang, Z. Yang, C. Zhong, and Z. Zhang, "Near-field localization for holographic RIS assisted mmwave systems," *IEEE Communications Letters*, vol. 27, no. 1, pp. 140–144, 2023.

[11] K. Keykhosravi, G. Seco-Granados, G. C. Alexandropoulos, and H. Wymeersch, "RIS-enabled self-localization: Leveraging controllable reflections with zero access points," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 2852–2857.

[12] S.-Y. Lien, D.-J. Deng, C.-C. Lin, H.-L. Tsai, T. Chen, C. Guo, and S.-M. Cheng, "3GPP NR sidelink transmissions toward 5G V2X," *IEEE Access*, vol. 8, pp. 35 368–35 382, 2020.

[13] M. Ammous and S. Valaee, "Cooperative positioning with the aid of reconfigurable intelligent surfaces and zero access points," in *2022 IEEE 96th Vehicular Technology Conference: (VTC2022-Fall)*, 2022, pp. 1–5.

[14] M. Harounabadi, D. M. Soleymani, S. Bhadauria, M. Leyh, and E. Roth-Mandutz, "V2X in 3GPP standardization: NR sidelink in release-16 and beyond," *IEEE Communications Standards Magazine*, vol. 5, no. 1, pp. 12–21, 2021.

[15] J. He, H. Wymeersch, L. Kong, O. Silvén, and M. Juntti, "Large intelligent surface for positioning in millimeter wave MIMO systems," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1–5.

[16] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5386–5402, 2021.

[17] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, "MetaLocalization: Reconfigurable intelligent surface aided multi-user wireless indoor localization," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7743–7757, 2021.

[18] C. L. Nguyen, O. Georgiou, G. Gradoni, and M. Di Renzo, "Wireless fingerprinting localization in smart environments using reconfigurable intelligent surfaces," *IEEE Access*, vol. 9, pp. 135 526–135 541, 2021.

[19] D. Dardari, N. Decarli, A. Guerra, and F. Guidi, "LOS/NLOS near-field localization with a large reconfigurable intelligent surface," *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 4282–4294, 2022.

[20] M. Ammous and S. Valaee, "Positioning and tracking using reconfigurable intelligent surfaces and extended kalman filter," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–6.

[21] A. Parchekani and S. Valaee, "Sensing and localization using reconfigurable intelligent surfaces and the swendsen-wang algorithm," in *2022 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2022, pp. 981–986.

[22] A. Parchekani and S. Valaee, "Sensing via orthogonal time frequency space signalling and reconfigurable intelligent surface," in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2022, pp. 1–6.

[23] M. Ammous and S. Valaee, "Cooperative positioning with the aid of reconfigurable intelligent surfaces and device-to-device communications in mmwave," in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2022, pp. 683–688.

[24] Q. Cheng, L. Li, M.-M. Zhao, and M.-J. Zhao, "Cooperative localization for reconfigurable intelligent surface-aided mmwave systems," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 1051–1056.

[25] P. Gao, L. Lian, and J. Yu, "Optimal passive beamforming for cooperative localization with RIS-assisted mmwave systems," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 222–227.

[26] H. Kim, H. Chen, M. F. Keskin, Y. Ge, K. Keykhosravi, G. C. Alexandropoulos, S. Kim, and H. Wymeersch, "RIS-enabled and access-point-free simultaneous radio localization and mapping," *IEEE Transactions on Wireless Communications*, vol. 23, no. 4, pp. 3344–3360, 2024.

[27] E. Basar and H. V. Poor, "Present and future of reconfigurable intelligent surface-empowered communications [perspectives]," *IEEE Signal Processing Magazine*, vol. 38, no. 6, pp. 146–152, 2021.

[28] P. Zheng, H. Chen, T. Ballal, M. Valkama, H. Wymeersch, and T. Y. Al-Naffouri, "JrCUP: Joint RIS calibration and user positioning for 6G wireless systems," *IEEE Transactions on Wireless Communications*, pp. 1–16, 2023.

[29] R. Ghazalian, K. Keykhosravi, H. Chen, H. Wymeersch, and R. Jäntti, "Bi-static sensing for near-field RIS localization," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 6457–6462.

[30] "3GPP TS 38.322 V16.0.0: NR; Physical channels and modulation (Release 16)," Jan. 2020. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3213

[31] Z. Wei, Y. Wang, L. Ma, S. Yang, Z. Feng, C. Pan, Q. Zhang, Y. Wang, H. Wu, and P. Zhang, "5G PRS-based sensing: A sensing reference signal approach for joint sensing and communication system," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 3, pp. 3250–3263, 2023.

[32] Y. Sun, W. Cao, M. Zhou, and Z. Ding, "Hybrid successive interference cancellation and power adaptation: A win-win strategy for robust uplink NOMA transmission," *IEEE Transactions on Communications*, vol. 72, no. 2, pp. 771–785, 2024.

[33] A. Douik, S. Sorour, H. Tembine, T. Y. Al-Naffouri, and M.-S. Alouini, "A game-theoretic framework for network coding based device-to-device communications," *IEEE Transactions on Mobile Computing*, vol. 16, no. 4, pp. 901–917, 2017.

[34] J.-W. Qiu, C. C. Lo, C.-K. Lin, and Y.-C. Tseng, "A D2D relative positioning system on smart devices," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, 2014, pp. 2168–2172.

[35] S. Fujii, A. Fujita, T. Umedu, S. Kaneda, H. Yamaguchi, T. Higashino, and M. Takai, "Cooperative vehicle positioning via V2V communications and onboard sensors," in *2011 IEEE Vehicular Technology Conference (VTC Fall)*, 2011, pp. 1–5.

[36] F. Zhang, M.-M. Zhao, M. Lei, and M. Zhao, "Joint power allocation and phase-shift design for RIS-aided cooperative near-field localization," in *2022 International Symposium on Wireless Communication Systems (ISWCS)*, 2022, pp. 1–6.

[37] T. Mazloum, L. Santamaria, F. Munoz, A. Clemente, J.-B. Gros, Y. Nasser, M. Odit, G. Lerosey, and R. D'Errico, "Impact of multiple RIS on channel characteristics: An experimental validation in ka band," in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2023, pp. 13–18.

[38] K. Keykhosravi, M. F. Keskin, G. Seco-Granados, P. Popovski, and H. Wymeersch, "RIS-enabled SISO localization under user mobility and spatial-wideband effects," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 1125–1140, 2022.

[39] E. Björnson and L. Sanguinetti, "Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces," *IEEE Wireless Communications Letters*, vol. 10, no. 4, pp. 830–834, 2021.

[40] M. S. Kay, *Fundamentals of Statistical Signal Processing. estimation theory*. Prentice-Hall, 1993.

[41] L. Scharf and C. Demeure, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley Publishing Company, 1991.

[42] A. Albanese, P. Mursia, V. Sciancalepore, and X. Costa-Pérez, "PAPIR: Practical RIS-aided localization via statistical user information," in *2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2021, pp. 531–535.

[43] Z. Abu-Shaban, K. Keykhosravi, M. F. Keskin, G. C. Alexandropoulos, G. Seco-Granados, and H. Wymeersch, "Near-field localization with a reconfigurable intelligent surface acting as lens," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.

[44] T. Zhang, A. F. Molisch, Y. Shen, Q. Zhang, H. Feng, and M. Z. Win, "Joint power and bandwidth allocation in wireless cooperative localization networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6527–6540, 2016.

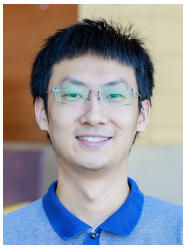
[45] L. Hogben, *Handbook of linear algebra*. Chapman & Hall/CRC, 2013.

[46] M. A. Murray-Lasso, "On the generalization of the law of cosines of triangles to three and more dimensions," *Ingeniería Investigación y Tecnología*, vol. 3, no. 4, p. 169–183, 2002.

[47] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," *Journal of Machine Learning Research*, vol. 15, no. 42, pp. 1455–1459, 2014. [Online]. Available: https://www.manopt.org

[48] 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 Transactions on Wireless Communications*, vol. 17, no. 8, pp. 4939–4954, 2018.

Mustafa Ammous (S'17) received his B.Sc. degree in Electrical Engineering from King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia in 2016 and his M.Sc. degree in Electrical and Computer Engineering at the University of Idaho, Moscow, Idaho, United States in 2018. He is currently pursuing his PhD degree at the University of Toronto, Toronto, Ontario, Canada. His research interests include radio localization and sensing using reconfigurable intelligent surfaces.



Hui Chen (S'18, M'21) obtained the Ph.D. degree in electrical and computer engineering from the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, in 2021. He is currently a Research Specialist with the Department of Electrical Engineering at Chalmers University of Technology, Gothenburg, Sweden. His research interests include 5G/6G localization, mismatch analysis and calibration, and machine learning for localization and sensing.

Henk Wymeersch (S'01, M'05, SM'19, F'24) obtained the Ph.D. degree in Electrical Engineering/Applied Sciences in 2005 from Ghent University, Belgium. He is currently a Professor of Communication Systems with the Department of Electrical Engineering at Chalmers University of Technology, Sweden. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 until 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. Prof. Wymeersch served as Associate Editor for IEEE Communication Letters (2009-2013), IEEE Transactions on Wireless Communications (since 2013), and IEEE Transactions on Communications (2016-2018) and is currently Senior Member of the IEEE Signal Processing Magazine Editorial Board. During 2019-2021, he was an IEEE Distinguished Lecturer with the Vehicular Technology Society. His current research interests include the convergence of communication and sensing, in a 5G and Beyond 5G context.

Shahrokh Valaee is a Professor with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, and the holder of Nortel Chair of Network Architectures and Services. He is the Founder and the Director of the Wireless Innovation Research Laboratory (WIRLab) at the University of Toronto. Professor Valaee was the TPC Co-Chair and the Local Organization Chair of the IEEE Personal Mobile Indoor Radio Communication (PIMRC) Symposium 2011. He was the TPC Co-Chair of ICT 2015, and PIMRC 2017, and the Track Co-Chair of WCNC 2014, PIMRC 2020, VTC Fall 2020. He was the co-chair of the organizing committee for PIMRC 2023. He is currently a member of the Steering Committee of IEEE PIMRC. From December 2010 to December 2012, he was the Associate Editor of the IEEE Signal Processing Letters. From 2010 to 2015, he served as an Editor of IEEE Transactions on Wireless Communications. Currently, he is an Editor of the Journal of Computer and System Science and also an Editor of IEEE Transactions on Wireless Communications. From 2021 to 2023, he was a Distinguished Lecturer of the IEEE Communications Society. Currently, he serves as a Distinguished Lecturer for the IEEE Vehicular Technology Society. He was the co-recipient of the best paper award in the IEEE Machine Learning for Signal Processing (MLSP) 2020 workshop. Professor Valaee is a Fellow of the Engineering Institute of Canada and a Fellow of IEEE.

APPENDIX A
DERIVATIVES OF $\boldsymbol{\mu}_{ij,t}$

The partial derivatives of $\boldsymbol{\mu}_{ij,t}$ with respect to η_b, $b = 1, 2, \ldots, 8(K-1)K$, can be computed as

$$\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \tau_{ij}} = -j2\pi\Delta f\sqrt{E_i}\beta_{ij}\left(\mathbf{n}\odot\mathbf{d}\left(\tau_{ij}\right)\right), \tag{A.1}$$

$$\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \tau_{ij,r}} = -j2\pi\Delta f\sqrt{E_i}\beta_{ij,r}\left(\mathbf{n}\odot\mathbf{d}\left(\tau_{ij,r}\right)\right)\mathbf{c}^\top\left(\boldsymbol{\gamma}_{ij}\right)\boldsymbol{\omega}_{i,t}, \tag{A.2}$$

$$\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \xi_{ij}} = j\frac{2\pi}{\lambda}\sqrt{E_i}\beta_{ij,r}\mathbf{d}\left(\tau_{ij,r}\right)\left(\mathbf{p}_y\odot\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right)^\top\boldsymbol{\omega}_{i,t}, \tag{A.3}$$

$$\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \zeta_{ij}} = j\frac{2\pi}{\lambda}\sqrt{E_i}\beta_{ij,r}\mathbf{d}\left(\tau_{ij,r}\right)\left(\mathbf{p}_z\odot\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right)^\top\boldsymbol{\omega}_{i,t}, \tag{A.4}$$

$$\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \alpha_{ij}} = \sqrt{E_i}e^{j\rho_{ij}}\mathbf{d}\left(\tau_{ij}\right), \tag{A.5}$$

$$\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \rho_{ij}} = j\sqrt{E_i}\beta_{ij}\mathbf{d}\left(\tau_{ij}\right), \tag{A.6}$$

$$\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \alpha_{ij,r}} = \sqrt{E_i}e^{j\rho_{ij,r}}\mathbf{d}\left(\tau_{ij,r}\right)\mathbf{c}^\top\left(\boldsymbol{\gamma}_{ij}\right)\boldsymbol{\omega}_{i,t}, \tag{A.7}$$

$$\frac{\partial \boldsymbol{\mu}_{ij,t}}{\partial \rho_{ij,r}} = j\sqrt{E_i}\beta_{ij,r}\mathbf{d}\left(\tau_{ij,r}\right)\mathbf{c}^\top\left(\boldsymbol{\gamma}_{ij}\right)\boldsymbol{\omega}_{i,t}, \tag{A.8}$$

where $\mathbf{n} = \left[0, 1, \ldots, N-1\right]^\top$. $\mathbf{p}_y = \left[p_{1,y}, p_{2,y}, \ldots, p_{M,y}\right]^\top$ where $p_{m,y}$ represents the location of the m-th RIS element along the y-axis. Similarly, $\mathbf{p}_z = \left[p_{1,z}, p_{2,z}, \ldots, p_{M,z}\right]^\top$ denotes the RIS elements positions' along the z-axis. Please note that the partial derivatives in (A.1)–(A.8) will be equal to the zero vector if the indices on the numerator and denominator are different. For example, $\partial\boldsymbol{\mu}_{ij,t}/\partial\tau_{bv} = \mathbf{0}_N$ if $i \neq b$ or $j \neq v$ where $\mathbf{0}_N$ is a zero column vector of length N.

APPENDIX B
DERIVATIVES OF $\boldsymbol{\eta}$

The partial derivatives of the delays with respect to $\mathbf{p}_{\mathrm{U}_k}$ (i.e., UE k) can be computed as

$$\frac{\partial \tau_{ij}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \frac{\partial \tau_{ji}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \begin{cases} \frac{1}{cd_{\mathrm{U}_k\mathrm{U}_j}}\left(\mathbf{p}_{\mathrm{U}_k} - \mathbf{p}_{\mathrm{U}_j}\right) & \text{if } k = i \\ \frac{1}{cd_{\mathrm{U}_i\mathrm{U}_k}}\left(\mathbf{p}_{\mathrm{U}_k} - \mathbf{p}_{\mathrm{U}_i}\right) & \text{if } k = j \\ \left[0,0,0\right]^\top & \text{otherwise,} \end{cases} \tag{B.1}$$

$$\frac{\partial \tau_{ij,r}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \frac{\partial \tau_{ji,r}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \begin{cases} \frac{1}{cd_{\mathrm{RU}_k}}\mathbf{p}_{\mathrm{U}_k} & \text{if } k = i \text{ or } k = j \\ \left[0,0,0\right]^\top & \text{otherwise.} \end{cases} \tag{B.2}$$

By substituting (5) and (6) in (12), we get:

$$\begin{aligned} \xi_{ij} &= \frac{y_{\mathrm{U}_i}}{\sqrt{x_{\mathrm{U}_i}^2 + y_{\mathrm{U}_i}^2 + z_{\mathrm{U}_i}^2}} + \frac{y_{\mathrm{U}_j}}{\sqrt{x_{\mathrm{U}_j}^2 + y_{\mathrm{U}_j}^2 + z_{\mathrm{U}_j}^2}}, \\ \zeta_{ij} &= \frac{z_{\mathrm{U}_i}}{\sqrt{x_{\mathrm{U}_i}^2 + y_{\mathrm{U}_i}^2 + z_{\mathrm{U}_i}^2}} + \frac{z_{\mathrm{U}_j}}{\sqrt{x_{\mathrm{U}_j}^2 + y_{\mathrm{U}_j}^2 + z_{\mathrm{U}_j}^2}}. \end{aligned} \tag{B.3}$$

The derivatives of the spatial frequencies with respect to $\mathbf{p}_{\mathrm{U}_k}$ can be computed as

$$\frac{\partial \xi_{ij}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \begin{cases} \left[-\frac{x_{\mathrm{U}_k}y_{\mathrm{U}_k}}{d_{\mathrm{RU}_k}^3}, \frac{1}{d_{\mathrm{RU}_k}} - \frac{y_{\mathrm{U}_k}^2}{d_{\mathrm{RU}_k}^3}, -\frac{z_{\mathrm{U}_k}y_{\mathrm{U}_k}}{d_{\mathrm{RU}_k}^3}\right]^\top & \text{if } k = i \text{ or } k = j \\ \left[0,0,0\right]^\top & \text{otherwise,} \end{cases} \tag{B.4}$$

$$\frac{\partial \zeta_{ij}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \begin{cases} \left[-\frac{x_{\mathrm{U}_k}z_{\mathrm{U}_k}}{d_{\mathrm{RU}_k}^3}, -\frac{y_{\mathrm{U}_k}z_{\mathrm{U}_k}}{d_{\mathrm{RU}_k}^3}, \frac{1}{d_{\mathrm{RU}_k}} - \frac{z_{\mathrm{U}_k}^2}{d_{\mathrm{RU}_k}^3}\right]^\top & \text{if } k = i \text{ or } k = j \\ \left[0,0,0\right]^\top & \text{otherwise.} \end{cases} \tag{B.5}$$

The derivatives of the channel gains with respect to $\mathbf{p}_{\mathrm{U}_k}$ is the zero vector. In other words,

$$\frac{\partial \alpha_{ij}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \frac{\partial \alpha_{ij,r}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \frac{\partial \rho_{ij}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \frac{\partial \rho_{ij,r}}{\partial \mathbf{p}_{\mathrm{U}_k}} = \left[0,0,0\right]^\top \forall\ i,\ j,\ k. \tag{B.6}$$

Next, the derivatives of the channel parameters with respect to the clock offset Δ_{t_k} can be computed as

$$\frac{\partial \tau_{ij}}{\partial \Delta_{t_k}} = \frac{\partial \tau_{ij,r}}{\partial \Delta_{t_k}} = \begin{cases} 1 & \text{if } k = j \\ -1 & \text{if } k = i \\ 0 & \text{otherwise,} \end{cases} \tag{B.7}$$

$$\begin{aligned} \frac{\partial \xi_{ij}}{\partial \Delta_{t_k}} &= \frac{\partial \zeta_{ij}}{\partial \Delta_{t_k}} = \frac{\partial \alpha_{ij}}{\partial \Delta_{t_k}} = \frac{\partial \alpha_{ij,r}}{\partial \Delta_{t_k}} \\ &= \frac{\partial \rho_{ij}}{\partial \Delta_{t_k}} = \frac{\partial \rho_{ij,r}}{\partial \Delta_{t_k}} = 0\ \forall\ i,\ j,\ k. \end{aligned} \tag{B.8}$$

APPENDIX C
DERIVATION OF (39)

We can evaluate $\beta_{ij,r}$ as a function of $\boldsymbol{\gamma}_{ij}$, denoted as $\tilde{\beta}_{ij,r}\left(\boldsymbol{\gamma}_{ij}\right)$, by ignoring the noise component in (38) and consequently we have:

$$\tilde{\mathbf{y}}_{ij}^{(\mathrm{r})} = N\sqrt{E_i}\tilde{\beta}_{ij,r}\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\boldsymbol{\Omega}}_i^\top\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right). \tag{C.1}$$

By dividing both sides of (C.1) by $N\sqrt{E_i}$, we obtain:

$$\frac{\tilde{\mathbf{y}}_{ij}^{(\mathrm{r})}}{N\sqrt{E_i}} = \tilde{\beta}_{ij,r}\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\boldsymbol{\Omega}}_i^\top\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right). \tag{C.2}$$

Since the term $\tilde{\boldsymbol{\Omega}}_i^\top\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)$ is a vector, we multiply both sides of (C.2) by its Hermitian (i.e., $\left(\tilde{\boldsymbol{\Omega}}_i^\top\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right)^H$) and get:

$$\frac{\left(\tilde{\boldsymbol{\Omega}}_i^\top\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right)^H\tilde{\mathbf{y}}_{ij}^{(\mathrm{r})}}{N\sqrt{E_i}} = \tilde{\beta}_{ij,r}\left(\boldsymbol{\gamma}_{ij}\right)\left(\tilde{\boldsymbol{\Omega}}_i^\top\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right)^H\tilde{\boldsymbol{\Omega}}_i^\top\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right). \tag{C.3}$$

The term $\left(\tilde{\boldsymbol{\Omega}}_i^\top\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right)^H$ can be computed as $\mathbf{c}^H\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\boldsymbol{\Omega}}_i^*$. In addition, the multiplication of a vector $\boldsymbol{v}$ with its Hermitian results in $\|\boldsymbol{v}\|_2^2$. Hence,

the term $\left(\tilde{\boldsymbol{\Omega}}_i^{\top}\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)\right)^{H}\tilde{\boldsymbol{\Omega}}_i^{\top}\mathbf{c}\left(\boldsymbol{\gamma}_{ij}\right)$ can be rewritten as $\left\|\mathbf{c}^{H}\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\boldsymbol{\Omega}}_i^{*}\right\|_2^2$. Then, we simplify (C.3) to:

$$\frac{\mathbf{c}^{H}\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\boldsymbol{\Omega}}_i^{*}\tilde{\mathbf{y}}_{ij}^{(\mathrm{r})}}{N\sqrt{E_i}}=\tilde{\beta}_{ij,r}\left(\boldsymbol{\gamma}_{ij}\right)\left\|\mathbf{c}^{H}\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\boldsymbol{\Omega}}_i^{*}\right\|_2^2. \tag{C.4}$$

By dividing both sides of (C.4) by $\left\|\mathbf{c}^{H}\left(\boldsymbol{\gamma}_{ij}\right)\tilde{\boldsymbol{\Omega}}_i^{*}\right\|_2^2$, we obtain (39) as required.