CHALMERS
UNIVERSITY OF TECHNOLOGY

Joint RIS Calibration and Multi-User Positioning

Joint RIS Calibration and Multi-User Positioning

Yi Lu*, Hui Chen†, Jukka Talvitie*, Henk Wymeersch†, and Mikko Valkama*
*Electrical Engineering, Tampere University, Finland
†Chalmers University of Technology, Sweden
e-mail: yi.lu@tuni.fi, hui.chen@chalmers.se

***Abstract*—Reconfigurable intelligent surfaces (RISs) are expected to be a key component enabling the mobile network evolution towards a flexible and intelligent 6G wireless platform. In most of the research works so far, RIS has been treated as a passive base station (BS) with a known state, in terms of its location and orientation, to boost the communication and/or terminal positioning performance. However, such performance gains cannot be guaranteed anymore when the RIS state is not perfectly known. In this paper, by taking the RIS state uncertainty into account, we formulate and study the performance of a joint RIS calibration and user positioning (JrCUP) scheme. From the Fisher information perspective, we formulate the JrCUP problem in a network-centric single-input multiple-output (SIMO) scenario with a single BS, and derive the analytical lower bound for the states of both user and RIS. We also demonstrate the geometric impact of different user locations on the JrCUP performance while also characterizing the performance under different RIS sizes. Finally, the study is extended to a multi-user scenario, shown to further improve the state estimation performance.**

***Index Terms*—5G New Radio, 6G, Fisher information, joint calibration and positioning, reconfigurable intelligent surfaces**

I. Introduction

Evolving from 5G to 6G, the wireless networks are transforming into a ubiquitous, intelligent and multi-function service platform with the support of several key technical enablers, such as artificial intelligence, cognitive slicing, proactive channel coding, and reconfigurable intelligent surface (RIS) [1], [2]. Among them, the RIS technology is seen as a promising transformative component to realize smart connectivity, which also ties the three fundamental wireless applications, i.e., communications, localization and sensing/mapping together [3]. Besides the potential benefits in communications [4]–[7], RIS can also be effectively deployed to construct a controllable and reconfigurable channel for improved positioning performance with lower costs than, e.g., the ultra-dense deployments of 5G New Radio (NR) BSs [8].

Earlier works on RIS-enabled positioning include [9]–[21], covering fingerprinting approaches as well as methods based on geometric channel parameters (i.e., angles and delays). For fingerprinting-based approaches, the utilization of RIS has provided reduced computational complexity and improved accuracy, as shown in [9], [10]. By employing the delay and angle measurements obtained from radio signals, the positioning performance with RISs as reflectors or scatters has been studied in different situations, e.g., in the line-of-sight (LoS) (with respect to the BS) [11]–[13], the corresponding non-line-of-sight (NLoS) [14]–[17], in the downlink side [11] or the uplink side [15], [18], and in indoor environments [19], [20]. Finally, a comparison between 5G and 6G systems is provided in [21].

This work was financially supported by the Academy of Finland under grants #319994, #338224, #328214, #323244, and through the H2020 project Hexa-X (Grant Agreement no. 101015956).

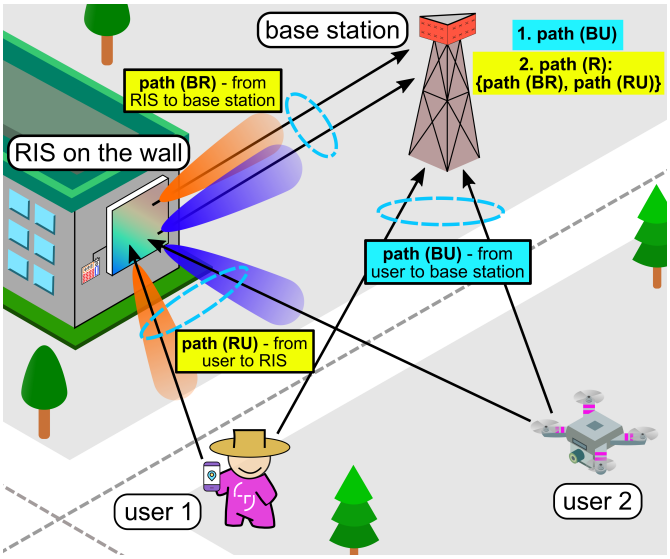


Fig. 1: A graphical illustration of the considered RIS-integrated network for joint RIS calibration and user positioning (JrCUP). Direct user-BS paths are marked in cyan, while paths including the RIS are marked in yellow.

An important limitation in all the above works is that the RIS location and orientation were assumed to be precisely known. With such assumption, RISs can therefore be treated as secondary BSs or anchor points to improve the positioning and/or communication performance. However, in reality, due to potential deployment faults, external disturbances, and/or improper installation, the state of RIS may not be perfectly known, thus, needing further calibration. Towards this end, we formulate and investigate the feasibility and performance of a joint RIS calibration and user positioning (JrCUP) scheme in this paper, where the state of RIS is not assumed to be perfectly known. The contributions of this work are as follows:

- We formulate the JrCUP problem in the 3D scenario, where the states of both users and the RIS need to be jointly estimated;
- We carry out the Fisher information matrix (FIM) analysis, and compute the lower bound of the estimated channel parameters and states of user and RIS, while also address their dependence on the problem geometry;

- We propose a low-cost 2D searching-based initialization method and devise an efficient snapshot-like joint state estimation method in the general multi-user scenario;
- We provide an extensive set of numerical results, showing that the lower bound can be achieved by the proposed estimation method, while also demonstrating that the multi-user scenario provides improved state estimation performance compared to the single-user case.

II. System Model

In this work, a 3D outdoor positioning scenario is considered and illustrated in Fig. 1, where there exists several users, one RIS deployed on the surfaces of a building, and a BS. The state of BS is known and used as the reference point in the considered coordinate system, while the states of users and RIS are to be estimated. Facilitating a network-centric wireless system, the BS observes and processes the uplink signals received from indirect paths via RIS and the direct path from user(s) to BS.

A. Geometry Model

Without loss of generality, we start the system descriptions by considering one arbitrary user and the RIS, while the extension towards multi-user and multi-RIS will be treated later. Specifically, a single-antenna user, an N_{R}-element RIS and an N_{B}-antenna BS located at $\mathbf{p}_{\mathrm{U}} = [x_{\mathrm{U}}, y_{\mathrm{U}}, z_{\mathrm{U}}]^\top$, $\mathbf{p}_{\mathrm{R}} = [x_{\mathrm{R}}, y_{\mathrm{R}}, z_{\mathrm{R}}]^\top$ and $\mathbf{p}_{\mathrm{B}} = [x_{\mathrm{B}}, y_{\mathrm{B}}, z_{\mathrm{B}}]^\top$, respectively, are considered where $\mathbf{p}_{\mathrm{B}}$ is the center of the BS array defined as the origin of the coordinate system (the BS is facing x-axis by default). The direction vector from the BS to the user can be expressed as

$$\mathbf{t}_{\mathrm{BU}} = -\mathbf{t}_{\mathrm{UB}} = \begin{bmatrix} t_{\mathrm{BU},x} \\ t_{\mathrm{BU},y} \\ t_{\mathrm{BU},z} \end{bmatrix} = \frac{\mathbf{p}_{\mathrm{U}} - \mathbf{p}_{\mathrm{B}}}{d_{\mathrm{BU}}}, \tag{1}$$

where $d_{\mathrm{BU}} = \|\mathbf{p}_{\mathrm{U}} - \mathbf{p}_{\mathrm{B}}\|$ is the distance between BS and user. Similar definitions apply to the RIS-user (RU) and BS-RIS (BR) paths. Furthermore, the orientation of RIS is defined as a 3D Euler angle vector $\mathbf{o}_{\mathrm{R}} = [o_1, o_2, o_3]^\top$ (pitch, roll, yaw) while the transformation from Euler angle vector $\mathbf{o}_{\mathrm{R}}$ to the rotation matrix $\mathbf{R}_{\mathrm{R}}$ can be found in [22].

In terms of the angle of arrival (AoA) and angle of departure (AoD), we denote $\boldsymbol{\varphi}_{\mathrm{BU}}$ as the AoA from user to BS, $\boldsymbol{\varphi}_{\mathrm{RU}}$ as the AoA from user to RIS, $\boldsymbol{\varphi}_{\mathrm{RB}}$ as the AoD from RIS to BS, and $\boldsymbol{\varphi}_{\mathrm{BR}}$ as the AoA from RIS to BS, where each $\boldsymbol{\varphi} = [\phi, \theta]^\top$ consists of azimuth and elevation angle and is measured in the local coordinate system. Since the BS is assumed to be the coordinate origin, its local and global direction vectors are identical. For the sake of clarity, we define the local direction vector $\tilde{\mathbf{t}}$ of the channel from BS to RIS as

$$\tilde{\mathbf{t}}_{\mathrm{RB}}(\boldsymbol{\varphi}_{\mathrm{RB}}) = \mathbf{R}_{\mathrm{R}}^\top \mathbf{t} = \begin{bmatrix} \cos(\phi_{\mathrm{RB}})\cos(\theta_{\mathrm{RB}}) \\ \sin(\phi_{\mathrm{RB}})\cos(\theta_{\mathrm{RB}}) \\ \sin(\theta_{\mathrm{RB}}) \end{bmatrix}. \tag{2}$$

B. Signal Model

We consider an OFDM-based uplink communication system, where two radio paths are observed and processed at the BS. The received frequency domain signal vector at the gth transmission (one transmission means one OFDM symbol transmission) and kth subcarrier can be formulated as [21]

$$y_{g,k} = \mathbf{w}_{\mathrm{B},g}^\top \left[\mathbf{h}_{g,k} x_{g,k} + \mathbf{n}_{g,k}\right], \tag{3}$$

where $\mathbf{w}_{\mathrm{B},g} \in \mathbb{C}^{N_{\mathrm{B}} \times 1}$ is the combiner vector at the BS, and $\mathbf{n}_{g,k} \in \mathbb{C}^{N_{\mathrm{B}} \times 1}$ refers to the additive white Gaussian noise vector with a complex normal distribution $\mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{N_B})$. The channel $\mathbf{H}_{g,k}$ comprises two parts: the direct channel from user to BS and the reflected channel via the RIS, i.e., $\mathbf{h}_{g,k} = \mathbf{h}_{\mathrm{BU},k} + \mathbf{h}_{\mathrm{R},g,k}$.

The channel from user to BS is given by

$$\mathbf{h}_{\mathrm{BU},k} = \alpha_{\mathrm{BU}} \mathbf{a}_{\mathrm{BU}}(\boldsymbol{\varphi}_{\mathrm{BU}}) e^{-j2\pi\Delta_f k \tau_{\mathrm{BU}}} \tag{4}$$

where Δ_f is the subcarrier spacing, $\tau_{\mathrm{BU}} = d_{\mathrm{BU}}/c + \beta$ represents the delay of BU path (including the synchronization offset β). The speed of light is denoted as c. Additionally, $\alpha_{\mathrm{BU}} = \rho_{\mathrm{BU}} + j\xi_{\mathrm{BU}}$ is the complex gain of the LOS channel, $\mathbf{a}_{\mathrm{BU}}(\boldsymbol{\varphi}_{\mathrm{BU}})$ is the steering vector of the BS-user channel with the bth element ($1 \le b \le N_{\mathrm{B}}$), expressed as

$$\mathbf{a}_{\mathrm{BU},b}(\boldsymbol{\varphi}_{\mathrm{BU}}) = e^{j\frac{2\pi}{\lambda_c}\mathbf{p}_b^\top \tilde{\mathbf{t}}_{\mathrm{BU}}}, \tag{5}$$

and $\mathbf{p}_b$ is the local position of the bth element with respect to the array center. We note that $\mathbf{p}_{b,x}$ is equal to $\mathbf{0}$ since a uniform planar array (UPA) is assumed.

The RIS-relayed channel $\mathbf{h}_{\mathrm{R},g,k}$ is defined as

$$\mathbf{h}_{\mathrm{R},g,k} = \alpha_{\mathrm{R}} \mathbf{a}_{\mathrm{BR}}(\boldsymbol{\varphi}_{\mathrm{BR}}) (\mathbf{a}_{\mathrm{RB}}(\boldsymbol{\varphi}_{\mathrm{RB}}))^\top \boldsymbol{\Omega}_g \mathbf{a}_{\mathrm{RU}}(\boldsymbol{\varphi}_{\mathrm{RU}}) e^{-j2\pi\Delta_f k \tau_{\mathrm{R}}} \tag{6}$$

where $\alpha_{\mathrm{R}} = \rho_{\mathrm{R}} + j\xi_{\mathrm{R}}$ is the complex channel gain of the RIS path, $\tau_{\mathrm{R}} = (d_{\mathrm{BR}} + d_{\mathrm{RU}})/c + \beta$ refers to the delay of the path R, and $\boldsymbol{\Omega}_g = \mathrm{diag}[\omega_{1,g}, \ldots, \omega_{N_{\mathrm{R}},g}]$ is the time-varying RIS configuration, with $|\omega_{i,g}| = 1$, $\forall i, g$. Furthermore, $\mathbf{a}_{\mathrm{RB}}(\boldsymbol{\varphi}_{\mathrm{RB}})$ and $\mathbf{a}_{\mathrm{RU}}(\boldsymbol{\varphi}_{\mathrm{RU}})$ are the steering vectors that can be obtained from (5). We further define an intermediate steering vector at the RIS as $\mathbf{a}_{\mathrm{R}}(\boldsymbol{\varphi}_{\mathrm{RB}}, \boldsymbol{\varphi}_{\mathrm{RU}}) = \mathbf{a}_{\mathrm{RB}}(\boldsymbol{\varphi}_{\mathrm{RB}}) \odot \mathbf{a}_{\mathrm{RU}}(\boldsymbol{\varphi}_{\mathrm{RU}})$ and its rth element can be obtained as

$$\mathbf{a}_{\mathrm{R},r}(\boldsymbol{\varphi}_{\mathrm{RB}}, \boldsymbol{\varphi}_{\mathrm{RU}}) = e^{j\frac{2\pi}{\lambda_c}\mathbf{p}_r^\top \tilde{\mathbf{t}}_{\mathrm{R}}} = e^{j\frac{2\pi}{\lambda_c}\mathbf{p}_r^\top (\tilde{\mathbf{t}}_{\mathrm{RB}} + \tilde{\mathbf{t}}_{\mathrm{RU}})}, \tag{7}$$

where $\tilde{\mathbf{t}}_{\mathrm{R}} = [\vartheta_1, \vartheta_2, \vartheta_3]^\top$ is the intermediate direction vector.

Remark: Considering that the first entry of the $\mathbf{p}_b$ is 0 and the fact that RIS does not perform any signal processing, the following intermediate AoA measurements can be obtained:

$$\begin{aligned} \vartheta_2 &= \sin(\phi_{\mathrm{RU}})\cos(\theta_{\mathrm{RU}}) + \sin(\phi_{\mathrm{RB}})\cos(\theta_{\mathrm{RB}}) \\ \vartheta_3 &= \sin(\theta_{\mathrm{RU}}) + \sin(\theta_{\mathrm{RB}}). \end{aligned} \tag{8}$$

In other words, there are overall four angle measurements $(\phi_{\mathrm{RB}}, \theta_{\mathrm{RB}}, \phi_{\mathrm{RU}}, \theta_{\mathrm{RU}})$ involved and correlated with the RIS-relayed channel. However, only two intermediate angles (8) can be acquired and estimated, resulting in 8 measurements. The fundamental purpose of this work is to estimate the RIS location, the RIS orientation, the user location and the clock offset simultaneously, based on these 8 measurements.

C. RIS Profile and BS Combiner Design

In this work, we assume that no prior information about the RIS and UE states is available. In such a case, we use random RIS profiles for each transmission, i.e., the coefficient of the ith RIS element in the gth transmission, $\omega_{i,g}$, is assigned with a unit amplitude $|\omega_{i,g}| = 1$ and a random phase following a uniform distribution as $\angle\omega_{i,g} \sim \mathcal{U}(-\pi, \pi)$. Similar random coefficients are also considered for BS combiner vectors across different transmissions. Such a method does not require any prior information of the RIS or the user(s), nor any specifically designed codebook, but necessitates a large number of transmissions, thus yielding increased latency. The optimization of the RIS profile and the BS combiner with prior information (e.g., iteratively obtained with localization feedback) will be considered in our future work.

III. Performance Bound and Estimation Algorithm

In this section, we outline and derive the lower bound of the parameters of interest using the ubiquitous FIM. Thereafter, we present the proposed initialization and snapshot estimation methods for JrCUP.

A. From FIM to Lower Bound

In general, the analytical lower bound of the parameters of interest can be obtained by computing the corresponding FIM based on the observed measurements. Herein, the observed measurements are the received signals at the BS, obtained by collecting (3) over the considered transmissions and active subcarriers, yielding $\boldsymbol{Y} \in \mathbb{C}^{G\times K}$ where $\boldsymbol{Y} = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_G]$ and $\boldsymbol{y}_g \in \mathbb{C}^{K\times 1}$, in which G is the total number of OFDM transmission and K refers to the number of subcarriers. Thereafter, the FIM of the channel parameters can be computed as follows [23]

$$\mathbf{I}(\boldsymbol{\eta}) = \frac{2}{\sigma^2}\sum_{g=1}^{\mathcal{G}}\sum_{k=1}^{K}\mathrm{Re}\left\{\left(\frac{\partial\mu_{g,k}}{\partial\boldsymbol{\eta}}\right)^H\left(\frac{\partial\mu_{g,k}}{\partial\boldsymbol{\eta}}\right)\right\} \tag{9}$$

where $\mu_{g,k} = \mathbf{w}_{\mathrm{B},g}^\top\mathbf{h}_{g,k}x_{g,k}$ is the noise-free version of the observed symbol $y_{g,k}$ in (3). Moreover, the channel parameter vector is denoted as

$$\boldsymbol{\eta} = [\boldsymbol{\varphi}_{\mathrm{BU}}, \boldsymbol{\varphi}_{\mathrm{BR}}, \vartheta_2, \vartheta_3, \tau_{\mathrm{BU}}, \tau_{\mathrm{R}}, \rho_{\mathrm{BU}}, \rho_{\mathrm{R}}, \xi_{\mathrm{BU}}, \xi_{\mathrm{R}}]^\top \tag{10}$$

where the last four are the nuisance parameters, in which ρ and ξ refer to the real and imaginary parts of the channel gain. The sub-indices $(\cdot)_{\mathrm{BU}}$ and $(\cdot)_{\mathrm{R}}$ refer to the gains of the user to BS (BU) and the RIS-relayed (R) paths, respectively. More importantly, the first 8 parameters[1] are the geometry-related measurements from which we extract the information of both the user and the RIS states.

By performing Schur complement [24], an effective FIM of (9) can be obtained with the nuisance parameters (i.e., the channel gains) being removed. Thereafter, the FIM of the state parameter vector can be calculated as

$$\mathbf{I}(\mathbf{s}) = \mathbf{J}_{\mathrm{S}}^\top\acute{\mathbf{I}}(\boldsymbol{\eta})\mathbf{J}_{\mathrm{S}}, \tag{11}$$

[1]Note that $\boldsymbol{\varphi} = [\phi, \theta]^\top$ defined in Sec. II-A contains both azimuth and elevation angles, therefore, the overall number of available measurements is 8.

Algorithm 1: Proposed initialization and snapshot estimation methods

1: Based on the direction induced by $\hat{\boldsymbol{\varphi}}_{\mathrm{BU}}$, find the intersection with the area of user $\mathcal{A}_{\mathrm{U}}$. This leads to a distance range of $[d_{\min}, d_{\max}]$ between BS and user.
2: **For** every candidate LoS distance $\check{d}_0 \in [d_{\min}, d_{\max}]$
3: Estimate the clock offset $\check{\beta}$ as $\check{\beta} = \hat{\tau}_{\mathrm{BU}} - \check{d}_0/c$
4: Estimate the user location as $\check{\mathbf{p}}_{\mathrm{U}} = \mathbf{p}_{\mathrm{B}} + \hat{\mathbf{t}}_{\mathrm{BU}}\check{d}_0$
5: Compute the distance of path (R) $\check{d}_{\mathrm{R}} = (\hat{\tau}_{\mathrm{R}} - \check{\beta})c$
6: Determine an ellipsoid (denoted by E_{R}) with focal points $\mathbf{p}_{\mathrm{B}}$ and $\check{\mathbf{p}}_{\mathrm{U}}$ and focal distance $\check{d}_{\mathrm{R}}$
7: Intersect the ellipsoid E_{R} and the line formed by $\hat{\mathbf{t}}_{\mathrm{BU}}(\hat{\boldsymbol{\varphi}}_{\mathrm{BR}})$, to determine the RIS location $\check{\mathbf{p}}_{\mathrm{R}}$
8: **For** every $\check{o}_3$ in the orientation prior $\mathcal{O}_{\mathrm{R}}$
9: Predict the intermediate AoA measurements $\check{\vartheta}_2$, $\check{\vartheta}_3$ (8) based upon $\check{o}_3$, $\check{\mathbf{p}}_{\mathrm{U}}$, $\check{\mathbf{p}}_{\mathrm{R}}$ and $\mathbf{p}_{\mathrm{B}}$
10: Compute the metric: $\Delta\left(\check{d}_0, \check{o}_3\right) = \|[\check{\vartheta}_2, \check{\vartheta}_3]^\top - [\hat{\vartheta}_2, \hat{\vartheta}_3]^\top\|$
11: **End**
12: **End**
13: Determine the RIS orientation $\hat{o}_3$ and the LoS distance $\hat{d}_0$ by solving

$$[\hat{d}_0, \hat{o}_3] = \underset{\check{d}_0, \check{o}_3}{\arg\min}\, \boldsymbol{\Delta}\left(\check{d}_0, \check{o}_3\right)$$

14: Based $\hat{d}_0$, determine $\hat{\mathbf{p}}_{\mathrm{U}}$, $\hat{\beta}$ and $\hat{\mathbf{p}}_{\mathrm{R}}$
15: Refine user location $\hat{\mathbf{p}}_{\mathrm{U}}$, clock offset $\hat{\beta}$, RIS location $\hat{\mathbf{p}}_{\mathrm{R}}$ and orientation $\hat{o}_3$ with a Gauss-Newton (GN) method in (13)

where $\acute{\mathbf{I}}(\boldsymbol{\eta})$ is the effective FIM computed from (9), and $\mathbf{J}_{\mathrm{S}} \triangleq \partial\boldsymbol{\eta}/\partial\mathbf{s}$ represents the Jacobian matrix, essentially the derivative of the channel parameters with respect to the state parameters. In particular, the state parameter vector $\mathbf{s} \in \mathbb{R}^{8\times 1}$ is defined as

$$\mathbf{s} = [\mathbf{s}_{\mathrm{R}}^\top, \mathbf{s}_{\mathrm{U}}^\top]^\top = [\mathbf{p}_{\mathrm{R}}^\top, o_3, \mathbf{p}_{\mathrm{U}}^\top, \beta]^\top. \tag{12}$$

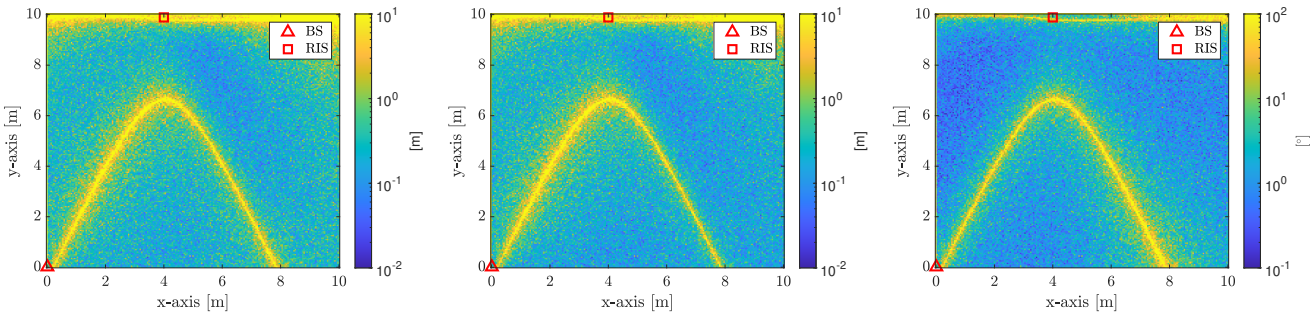
The derivation principles of $\mathbf{I}(\boldsymbol{\eta})$, $\mathbf{I}(\mathbf{s})$ and $\mathbf{J}_{\mathrm{S}}$ can be found for example in [13].

The lower bound of parameters of the state vector can be computed by taking the inverse of the FIM $\mathbf{I}(\mathbf{s})$, from which the lower bound of RIS and user locations can be calculated as $\sqrt{\mathrm{trace}\left(\mathbf{I}^{-1}(\mathbf{s})\right)_{[1:3]}}$ and $\sqrt{\mathrm{trace}\left(\mathbf{I}^{-1}(\mathbf{s})\right)_{[5:7]}}$, respectively. Similarly, the lower bound of RIS orientation and clock offset can be acquired as $\sqrt{\mathbf{I}^{-1}(\mathbf{s})_{[4]}}$ and $\sqrt{\mathbf{I}^{-1}(\mathbf{s})_{[8]}}$. Corresponding numerical examples will be provided in Section IV.

B. Proposed Initialization and Estimation Methods

In this subsection, we present an initialization method and a snapshot estimation method for JrCUP with one or multiple users in the network. For the consistency of notation and formulations, we present first the single-user case, followed by the extension to the multi-user case.

Specifically, the step-wise initialization method is outlined from step 1 to step 8 in Algorithm 1, where the outputs are



(a) User location. (b) RIS location. (c) RIS orientation.

Fig. 2: Visualization of different estimation lower bounds with different user locations: (a) User location error bound, (b) RIS location error bound, (c) RIS orientation error bound. The BS is facing the positive direction of x-axis and RIS is facing the negative direction of y-axis.

employed as the initial state of user (location and clock offset) and RIS (location and orientation) before the final estimation. In addition to the channel parameters $\boldsymbol{\eta}_{[1:8]}$ from (10), the inputs for the initialization contain the prior search area of user, $\mathcal{A}_{\mathrm{U}}$, within the considered environment and that of the RIS orientation $\mathcal{O}_{\mathrm{R}}$. Thereafter, with the searched user and RIS state as the initial guess, the last step of JrCUP scheme can be carried out using the iterative Gauss-Newton (GN) method [25, Ch. 4], expressed as

$$\hat{\mathbf{s}}_i = \hat{\mathbf{s}}_{i-1} + \mathbf{I}^{-1}(\hat{\mathbf{s}}_{i-1})\mathbf{J}_{\mathrm{S}}^{\top}\acute{\mathbf{I}}(\boldsymbol{\eta})\left(\hat{\boldsymbol{\eta}}_{[1:8]} - \boldsymbol{\eta}(\hat{\mathbf{s}}_{i-1})_{[1:8]}\right), \quad (13)$$

in which i denotes the iteration index, while $\mathbf{I}^{-1}(\hat{\mathbf{s}}_{i-1})$, $\mathbf{J}_{\mathrm{S}}$ and $\acute{\mathbf{I}}(\boldsymbol{\eta})$ were given in (11). Moreover, the input measurements $\hat{\boldsymbol{\eta}}_{[1:8]}$ are generated using the lower bound of channel parameter, i.e., $\acute{\mathbf{I}}(\boldsymbol{\eta})$ that approximates a distribution $\mathcal{N}(\boldsymbol{\eta}_{[1:8]}; \hat{\boldsymbol{\eta}}_{[1:8]}, (\acute{\mathbf{I}}(\boldsymbol{\eta}))^{-1})$. Finally, the predicted measurements $\boldsymbol{\eta}(\hat{\mathbf{s}}_{i-1})_{[1:8]}$ are evaluated using the estimated state of the $i-1$ iteration.

In the case of the multi-user scenario, the overall state vector $\mathbf{s}$ can be redefined as

$$\mathbf{s} = \left[\mathbf{s}_{\mathrm{R}}^{\top}, \mathbf{s}_1^{\top}, \cdots, \mathbf{s}_M^{\top}\right]^{\top}, \quad (14)$$

where $\mathbf{s}_m = [\mathbf{p}_{\mathrm{U},m}^{\top}, \beta_m]^{\top}$ is the state vector of the mth user. The corresponding Jacobian matrix, measurement covariance and measurement vector can all be extended accordingly, and thereon applied in the GN algorithm in (13).

IV. Numerical Results

In this section, we present the achievable accuracy of the proposed JrCUP scheme in a concrete example scenario at the 28 GHz mmWave band, with the evaluation parameters listed in Table I. In particular, the BS is located at $[0, 0, 0]^{\top}$ of the applied coordinate system, with antenna array facing the positive of x-axis. The combiner matrix $\mathbf{w}_{\mathrm{B},g}$ at BS is chosen according to the employed codebooks which we will describe in the next subsection. The whole area under consideration is 10 m $\times$ 10 m with a RIS located at $[4, 10, 0]^{\top}$. The user height is set to 5 m below the BS to reflect a terrestrial user scenario. Finally, the distance and orientation intervals are set to 0.1 m and 0.1°, respectively, which are adopted in both Algorithm 1 and the numerical simulations.

TABLE I: Evaluation Assumptions and Numerology

Parameter	Value
Carrier frequency, f_c	28 GHz
Bandwidth, W	400 MHz
Number of transmissions, G	500
Number of subcarriers, K	128
Transmit power	$+30$ dBm
Noise power spectral density, N_0	-173.8 dBm/Hz
Noise figure of RX	10 dB
BS array size, N_{B}	16×16
RIS array size, N_{R}	20×20
Number of RF chains at BS/user	1
Number of GN iterations	30

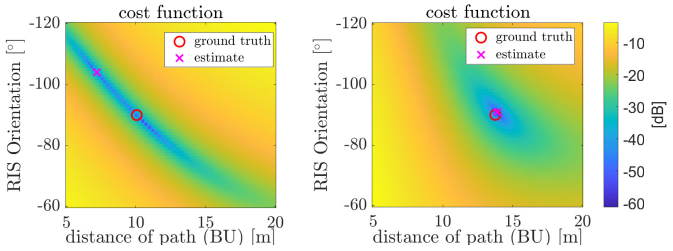


Fig. 3: The cost function for user at the blind area (left plot, user located at $[5, 6, -5]$) and non-blind area (right plot, user located at $[9, 8, -5]$) of Fig. 2.

A. Performance Bounds at Different User Locations

We start with the heat map of different lower bounds, obtained through the derivations in Section III-A, when user is located throughout the whole map. In terms of the RIS profile, we assume no prior information is available and apply random RIS coefficients (unit amplitude and random phases) for each transmission as discussed in Section II-C. The pattern for clock offsets is approximately the same as that of user location, therefore, is omitted herein for presentation brevity.

From the obtained pattern in Fig. 2, we see that except for the two blind areas (observed in yellow), both the location and orientation estimations achieve good performance over the considered area as shown in Fig. 2, where the location lower bound is in general under 1 m and the orientation lower bound is under 2°. In terms of the blind area, one is around the

locations that are in line with the RIS or BS array planes, where accurate angle estimates cannot be obtained. Another blind area is the parabola-style zone across the BS and is symmetric to the norm of the RIS array. The reason for such a parabolic blind area lies in the fact that there exist other candidate state vectors that generate the same channel parameters, resulting in ambiguous solutions.

We further evaluate the reason for the blind area in Fig. 3 from the perspective of the cost function that has been described in Algorithm 1. It is interesting to see that a unique solution exists on the right subplot when the user is located at $[9, 8, -5]$ (at the non-blind area), while ambiguous solutions can be found on the left subplot when the user is located at $[5, 6, -5]$ (at the blind area). This finding indicates that there exists an optimal area for solving the JrCUP problem.

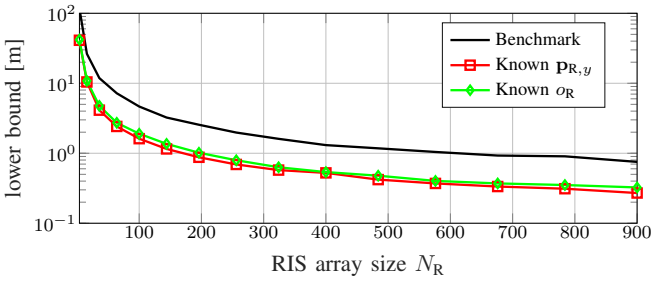
B. Impacts of RIS Size and Known States

We continue by evaluating the performance bounds for the different state parameters as functions of RIS sizes (i.e., the overall number of RIS elements N_{R}), while also considering different special cases of known states. The state of BS and RIS remain the same as in Fig. 2, while the user is set at an example location of $[8, 8, -5]^\top$ in the non-blind area. The lower bounds for user location, RIS location, and RIS orientation are respectively shown in Fig. 4, in which the benchmark scenario (black curve) assumes the user location, clock offset, RIS location, and RIS orientation are all unknown, yielding overall 8 unknowns. Compared to the benchmark scenario, we find that when one coordinate of RIS location (i.e., $\mathbf{p}_{\mathrm{R},y}$) is known, the performance of all the state parameters is improved as demonstrated by the red curves. In particular, the location estimation performance shown in Fig. 4a and Fig. 4b benefit more from this scenario than the orientation estimates, shown in Fig. 4c. As of the green curve scenario when the RIS orientation is assumed to be accurately known, the achieved performance is on a similar level as the red curve scenario because there are overall 7 unknowns in these two cases.

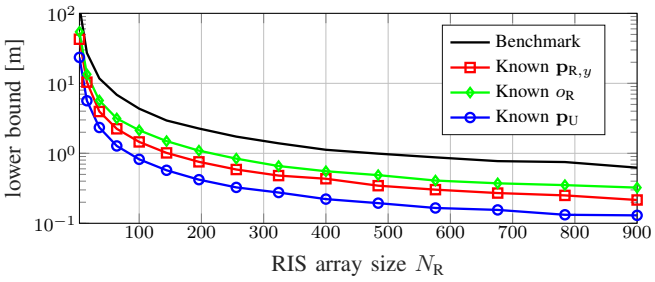
Moving next towards the blue curve scenario, when the user acts as a calibration agent within the system, i.e., the user location $\mathbf{p}_{\mathrm{U}}$ is known, the lower bound of the RIS state is vastly improved. In this case, the location lower bound and orientation lower bound drop to around 0.3 m and 0.5° individually with overall 400 RIS elements. It can also be clearly observed that the performance in all different scenarios becomes better as the RIS size N_{R} increases. This is intuitive since a large RIS size directs the signal power more efficiently and provides a finer angular resolution. These observations suggest that reliable and accurate information on the state of either the user or the RIS as well as a sufficiently large RIS size undoubtedly improve the performance.

C. Proposed Estimator Performance vs. Bounds

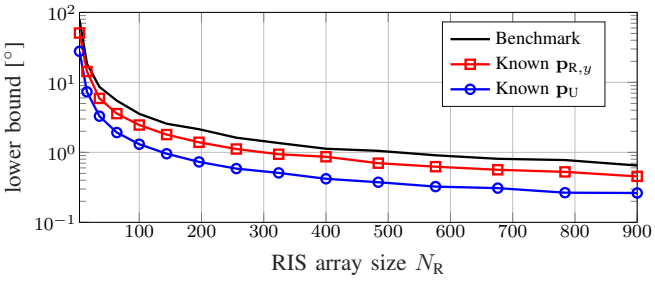
Finally, we assess and compare the performance of the proposed estimator against the corresponding lower bounds, while also varying the number of the involved users. The obtained results are shown in Fig. 5 where the dashed lines with



(a) User location lower bound



(b) RIS location lower bound



(c) RIS orientation lower bound

Fig. 4: Lower bounds as functions of RIS size N_{R}.

different colors represent the accuracy of initial state, which is calculated using the searching method described in Algorithm 1. Such accuracy can be considered as a reference benchmark. Moreover, the solid lines are the lower bounds of different state parameters, whereas the star markers represent the proposed estimator RMSE over 100 trials. The user locations are randomly generated in each trial with an orthogonal frequency division multiple access (OFDMA) resource allocation scheme where each user possesses equal bandwidth and the same physical height. Learning from the performance pattern in Fig. 2, we assume that the users are uniformly sampled in a 3×3 m^2 area $\mathcal{A}_{\mathrm{U}}$ with the start point located at $[6.5, 5.5, -5]^\top$. In such a way, the blind area in Fig. 2 where the positioning solution cannot be uniquely identified is avoided. From the numerical results presented in Fig. 5, we can see that with more users in the network, the joint estimation performance improves due to stronger geometric restraint and more information obtained from more measurements. In other words, the stronger geometric constraint formed by more users enhances the accuracy of RIS state, which in turn helps the user states after the iterations via GN approach. It can be observed that

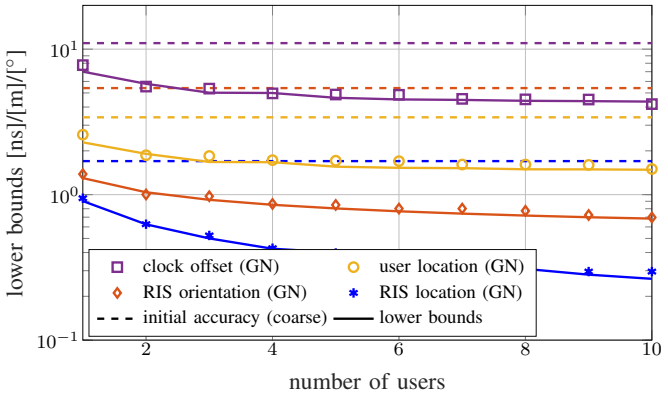


Fig. 5: The estimator RMSE and corresponding lower bounds as functions of the overall number of users.

the performance of the iterative GN method can approach the corresponding analytical lower bound. The proposed method can thus be considered as an efficient estimation solution for JrCUP.

V. Conclusion

In this paper, we investigated the problem of joint RIS calibration and user positioning, called JrCUP, towards intelligent 6G wireless communication systems. The ultimate objective was to jointly estimate the state parameters of both the users, in terms of clock offsets and 3D positions, and the RIS, in terms of array orientation and 3D position. To this end, we first expressed and computed the lower bound for all the state parameters. We then formulated signal processing methods for state initialization and iterative estimation. Our numerical results showed that the geometric impact can be detrimental and therefore needs extra attention and evaluation in the network planning phase. More importantly, we have found and shown that multi-user scenario in general outperforms the single-user case, demonstrating the potential benefits of deploying multiple users in the system. Moreover, the accuracy of the proposed estimation methods was shown to approach the lower bound, indicating that the proposed methods are efficient, and that the JrCUP problem is efficiently solvable. Our future research topics include the development of computationally efficient joint state tracking methods with moving users as well as the optimization strategy of RIS phase profiles and BS combiners.

References

[1] H. Sarieddeen, N. Saeed, T. Y. Al-Naffouri, and M.-S. Alouini, "Next Generation Terahertz Communications: A Rendezvous of Sensing, Imaging, and Localization," *IEEE Commun. Mag.*, vol. 58, no. 5, pp. 69–75, Jun. 2020.

[2] E. C. Strinati, G. C. Alexandropoulos, H. Wymeersch, B. Denis, V. Sciancalepore, R. D'Errico, A. Clemente, D.-T. Phan-Huy, E. De Carvalho, and P. Popovski, "Reconfigurable, Intelligent, and Sustainable Wireless Environments for 6G Smart Connectivity," *IEEE Commun. Mag.*, vol. 59, no. 10, pp. 99–105, 2021.

[3] H. Wymeersch, J. He, B. Denis, A. Clemente, and M. Juntti, "Radio Localization and Mapping With Reconfigurable Intelligent Surfaces: Challenges, Opportunities, and Research Directions," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 52–61, 2020.

[4] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless Communications Through Reconfigurable Intelligent Surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[5] S. Hu, F. Rusek, and O. Edfors, "Beyond Massive MIMO: The Potential of Data Transmission With Large Intelligent Surfaces," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2746–2758, 2018.

[6] E. Björnson, O. Özdogan, and E. G. Larsson, "Reconfigurable Intelligent Surfaces: Three Myths and Two Critical Questions," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 90–96, 2020.

[7] D. Dardari, "Communicating with large intelligent surfaces: Fundamental limits and models," *IEEE J. Sel. Areas in Commun.*, vol. 38, no. 11, pp. 2526–2537, Nov. 2020.

[8] H. Zhang, J. Lee, and T. Quek, *Ultra-dense Networks: Principles and Applications*. Shaftesbury Road, Cambridge, UK: Cambridge University Press, 2020.

[9] 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.

[10] P. del Hougne, "RIS-Based Radio Localization in Rich Scattering Environments: Harnessing Multi-Path with ANN Decoders," in *Proc. IEEE SPAWC*, 2021, pp. 511–515.

[11] H. Wymeersch and B. Denis, "Beyond 5G Wireless Localization with Reconfigurable Intelligent Surfaces," in *Proc. IEEE ICC*, 2020, pp. 1–6.

[12] K. Keykhosravi, M. F. Keskin, S. Dwivedi, G. Seco-Granados, and H. Wymeersch, "Semi-Passive 3D Positioning of Multiple RIS-Enabled Users," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 11 073–11 077, 2021.

[13] K. Keykhosravi, M. F. Keskin, G. Seco-Granados, and H. Wymeersch, "SISO RIS-Enabled Joint 3D Downlink Localization and Synchronization," in *Proc. IEEE ICC*, 2021, pp. 1–6.

[14] J. He, H. Wymeersch, T. Sanguanpuak, O. Silven, and M. Juntti, "Adaptive Beamforming Design for mmWave RIS-Aided Joint Localization and Communication," in *Proc. IEEE WCNC Workshops*, 2020, pp. 1–6.

[15] A. Albanese, P. Mursia, V. Sciancalepore, and X. Costa-Pérez, "PAPIR: Practical RIS-aided Localization via Statistical User Information," in *Proc. IEEE SPAWC*, 2021, pp. 531–535.

[16] D. Dardari, N. Decarli, A. Guerra, and F. Guidi, "Localization in NLOS Conditions using Large Reconfigurable Intelligent Surfaces," in *Proc. IEEE SPAWC*, 2021, pp. 551–555.

[17] M. Luan, B. Wang, Y. Zhao, Z. Feng, and F. Hu, "Phase Design and Near-Field Target Localization for RIS-Assisted Regional Localization System," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 1766–1777, 2022.

[18] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable Intelligent Surfaces for Localization: Position and Orientation Error Bounds," *IEEE Trans. Signal Process.*, vol. 69, pp. 5386–5402, 2021.

[19] T. Ma, Y. Xiao, X. Lei, W. Xiong, and Y. Ding, "Indoor Localization With Reconfigurable Intelligent Surface," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 161–165, 2021.

[20] Z. Yang, H. Zhang, B. Di, H. Zhang, K. Bian, and L. Song, "Wireless Indoor Simultaneous Localization and Mapping Using Reconfigurable Intelligent Surface," in *Proc. IEEE GLOBECOM*, 2021, pp. 1–6.

[21] H. Chen, H. Sarieddeen, T. Ballal, H. Wymeersch, M.-S. Alouini, and T. Y. Al-Naffouri, "A tutorial on terahertz-band localization for 6G communication systems," *Accepted for publication in IEEE Commun. Surveys Tuts. arXiv preprint arXiv:2110.08581*, 2022.

[22] J. Lasenby, H. Lasenby, and A. Lasenby, "Calculating the Rotor Between Conformal Objects," *Adv. Appl. Clifford Algebras*, vol. 29, no. 102, Oct. 2019.

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

[24] F. Zhang, *The Schur complement and its applications*. Springer Science & Business Media, 2005, vol. 04.

[25] S. Sand, A. Dammann, and C. Mensing, *Positioning in Wireless Communication Systems*. John Wiley & Sons Ltd., June 2014.