Hybrid Near-field and Far-field Localization with Holographic MIMO

Mengyuan Cao, *Student Member, IEEE*, Haobo Zhang, *Member, IEEE*, Yonina C. Eldar, *Fellow, IEEE*, Hongliang Zhang, *Member, IEEE*.

***Abstract*—Due to its ability to precisely control wireless beams, holographic multiple-input multiple-output (HMIMO) is expected to be a promising solution to achieve high-accuracy localization. However, as the scale of HMIMO increases to improve beam control capability, the corresponding near-field (NF) region expands, indicating that users may exist in both NF and far-field (FF) regions with different electromagnetic transmission characteristics. As a result, existing methods for pure NF or FF localization are no longer applicable. We consider a hybrid NF and FF localization scenario in this paper, where a base station (BS) locates multiple users in both NF and FF regions with the aid of a reconfigurable intelligent surface (RIS), which is a low-cost implementation of HMIMO. In such a scenario, it is difficult to locate the users and optimize the RIS phase shifts because whether the location of the user is in the NF or FF region is unknown, and the channels of different users are coupled. To tackle this challenge, we propose a RIS-enabled localization method that searches the users in both NF and FF regions and tackles the coupling issue by jointly estimating all user locations. We derive the localization error bound by considering the channel coupling and propose an RIS phase shift optimization algorithm that minimizes the derived bound. Simulations show the effectiveness of the proposed method and demonstrate the performance gain compared to pure NF and FF techniques.**

***Index Terms*—Holographic multiple-input multiple-output (HMIMO), reconfigurable intelligent surface (RIS), near-field localization, far-field localization.**

I. Introduction

Future sixth-generation (6G) wireless systems will support a wide range of applications such as autonomous driving and augmented reality [2] that require not only ubiquitous communication but also user localization capabilities. Consequently, research interest in advanced localization methods has surged, among which holographic multiple-input multiple-output (HMIMO)-based localization methods are expected to play a critical role due to its potential to achieve high localization accuracy. Specifically, HMIMO is an electromagnetic surface containing a massive number of antennas or reconfigurable metamaterial elements [3] [4]. Due to the large number of tunable elements in the surface, HMIMO can precisely manipulate wireless beams to accurately locate users.

M. Cao and Hongliang Zhang are with State Key Laboratory of Photonics and Communications, School of Electronics, Peking University, Beijing 100871, China (e-mail: {caomengyuan, hongliang.zhang}@pku.edu.cn).

Haobo Zhang is with the School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, Shenzhen 518055, China (e-mail: haobo.zhang@pku.edu.cn).

Y. C. Eldar is with the Department of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 7610001, Israel (e-mail: yonina.eldar@weizmann.ac.il).

Part of this work has been published in IEEE WCNC 2024 [1].

A. Literature Review and Motivation

Traditional HMIMO-based localization works mainly fall into two categories: far-field (FF) [5]–[11] and near-field (NF) localization schemes [12]–[17]. In the FF region, the distances between the HMIMO and users are sufficiently large, and the signal channel is effectively described using a plane wave model. In [5], the fundamental bounds and localization feasibility condition are examined for FF localization in orthogonal frequency division multiplexing (OFDM) systems aided by HMIMO. The authors in [6] considered a single-input-single-output (SISO) HMIMO-aided localization system with imperfect HMIMO geometry information, and derived the misspecified Cramér-Rao bound (CRB). In [7], the positioning estimation error and orientation estimation error are investigated for a HMIMO-aided multiple-input multiple output system. In [8], the authors analyzed the CRB for a localization scenario that both the line-of-sight (LOS) path and the reflected path via the HMIMO exist simultaneously. Two HMIMO phase design methods were proposed and a maximum likelihood (ML) direction of arrival (DOA) estimator was designed in [8]. In [9], the authors considered FF user localization aided by millimeter wave (mmWave) HMIMO, and a ML location estimator and two sub-optimal estimators with lower complexity were derived. In [10], the localization method leveraged the dynamic regulation of HMIMO to cancel non-line-of-sight signals at the receiver and preserve LOS signals for precise indoor localization. HMIMO-based localization without base stations (BSs) was also investigated in [11], where the user transmitted the signal and received the signal reflected from the HMIMO.

HMIMO-based NF localization schemes have attracted much attention recently due to the increasing interest in large-scale HMIMO. Specifically, to improve beam control capability, the scale of HMIMO is enlarged, leading to the expansion of the NF region [18]. Different from the FF case, the spherical wave model has to be employed to characterize the NF channel because the signal transmission characteristics of the NF are determined by both the range and angle of the user location. The authors in [12] and [13] considered a SISO localization scenario for NF user, and the CRB is derived. In [14], a HMIMO functioning as a lens receiver was utilized to achieve NF localization. The authors in [15] developed a received signal strength-based localization method for NF users and a RIS phase shift adjustment method to enhance accuracy. In [16], the authors utilized light detection and ranging to estimate scatter locations and improve NF user

localization accuracy. A signaling and positioning method is proposed in [17], and the coverage capability under different levels of obstruction of the HMIMO is analyzed.

In practical scenarios, users may exist in both the NF and FF regions of the HMIMO. Existing algorithms for pure NF or FF users are therefore not always applicable. Specifically, FF localization methods suffer from low localization accuracy for NF users due to the significant error in describing NF channels using the plane wave model. Meanwhile, NF localization techniques are not ideal for FF users because the coupling of range and angle in the spherical model brings extra complexity, which has two main effects: first, it prolongs the running time of the algorithm; second, the more complex NF model hinders convergence of the method and leads to accuracy degradation. In addition, the localization gain for the FF users brought by the NF model is limited. Hence, a localization scheme is required to adapt to both the NF and FF users.

B. Main Contributions

In this paper, a HMIMO-based localization scheme is proposed for users in both NF and FF regions. Specifically, in the considered scenario, a base station (BS) and a reconfigurable intelligent surface (RIS) cooperate to locate multiple users in a multipath environment. The RIS is widely acknowledged as a low-cost implementation of HMIMO that can create a customized reflection beam by varying the phase shifts of its reflection elements. In the localization process, the RIS reflects the signals emitted from several single-antenna users to the BS. Then, the BS uses the received signal to estimate the locations of the users in both NF and FF regions. The phase shifts of the RIS are optimized based on the estimated locations of the users. The location estimation and RIS phase shift optimization are performed iteratively to improve localization accuracy.

Several challenges in this scenario need to be addressed. *First*, it is unknown whether each user or scatter is in the NF or the FF region, leading to difficulty in designing the localization algorithm. *Second*, the channels of different users are coupled due to shared RIS phase shifts and common scatters. To tackle the above challenges efficiently, a hybrid NF and FF localization algorithm is proposed, in which we first sample the NF and FF regions separately, and then compare the received signals with the reconstructed signals at the sampled locations to locate the users and scatters. The locations of the users and scatters are jointly estimated by considering the channel coupling effect. Then, a RIS phase shift optimization problem is formulated to minimize the sum of the Cramér-Rao bounds (CRBs) of the user localization errors. A complex circle manifold-based method is proposed to solve the optimization problem with the coupled CRBs.

Our contributions can be summarized as follows.

- We consider a hybrid NF and FF localization scenario, where the BS locates multiple users in hybrid regions with the aid of an RIS. A hybrid NF and FF localization framework is proposed, where the localization of the users and the RIS phase shifts optimization are iteratively conducted to improve localization accuracy.

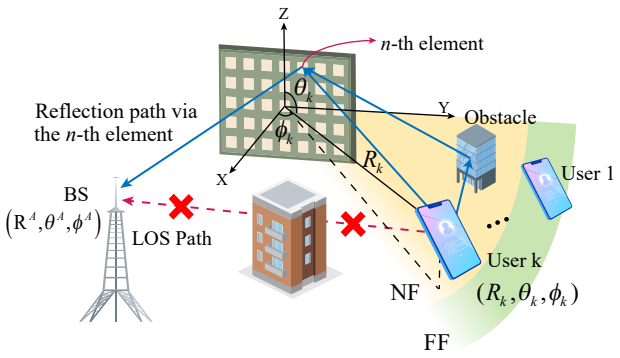

Fig. 1. RIS-enabled multi-user localization system.

- Based on the principle of minimizing localization loss, the location estimation problem is formulated and a multi-user localization algorithm is developed, which localizes users in the hybrid NF and FF region under the channel coupling effect.
- We formulate the RIS phase shift optimization problem to minimize the sum of CRBs of the location estimation errors. A complex circle manifold-based algorithm is then devised to obtain optimal RIS phase shifts.
- We compare the performance of the proposed approach and other RIS-enabled localization techniques. Simulation results show that the proposed approach outperforms other methods by more than 30% in the root mean square error (RMSE) of angle estimation for both NF and FF users and by 60% in the RMSE of range estimation for NF users given $-10 \leq \text{SNR} \leq 0$.

C. Organization

The rest of this paper is organized as follows. The localization scenario, the signal model, and the localization protocol are provided in Sec. II. In Sec. III, we propose the localization algorithm. Then, the RIS phase shift optimization algorithm is designed in Sec. IV. The performance of the proposed method is analyzed in Sec. V. In Sec. VI, we present the simulation results and discussions. Finally, conclusions are drawn in Sec. VII.

Notation: The transpose and conjugate transpose are denoted by $(\cdot)^T$ and $(\cdot)^H$, respectively. We use $\|\boldsymbol{x}\|_2$ and $\|\boldsymbol{x}\|_0$ mean the l_2 and l_0 norm of vector $\boldsymbol{x}$, and $|\cdot|$ denotes the number of elements in a set. Unbolded x, bold lower-case $\boldsymbol{x}$ and bold upper-case $\boldsymbol{X}$ denote scalar, vector and matrix, respectively. The collection of all complex numbers is denoted by $\mathbb{C}$. The Hadamard product is represented by $\odot$.

II. System Model

In this section, we first introduce the scenario of multi-user localization. Then, we describe the signal model and introduce a localization protocol for the proposed scenario.

TABLE I
MAJOR NOTATIONS

Notation	Description
N	Number of RIS elements
K	Number of users
L	Number of scatters
C	Number of localization cycles
$\boldsymbol{\beta}$	RIS phase shifts
$\boldsymbol{B}^{(c)}$	RIS phase shifts for c cycles
$\boldsymbol{h}_k$	Channel between the k-th user and the BS
$\boldsymbol{h}_k^A$	Channel between the RIS and the BS
$\boldsymbol{h}_k^t$	Channel between the RIS and the k-th user
s_k	The transmitted signal by the k-th user
$\boldsymbol{p}_k$	Location of the k-th user
$\boldsymbol{Q}$	Locations of all scatters
$\boldsymbol{g}_k^{(c)}$	The received signal of the c-th cycle
$\boldsymbol{Z}$	The sampled locations
$\boldsymbol{F}$	Atom channel in the NF and FF regions
$\boldsymbol{u}_k$	Gains of the atom channels for the k-th
$\boldsymbol{u}_k^{\boldsymbol{P}}$	The direct path component in $\boldsymbol{u}_k$
$\boldsymbol{u}_k^{\boldsymbol{Q}}$	The scattering path component in $\boldsymbol{u}_k$
$\boldsymbol{\Lambda}^k$	Atom signals for the k-th user
$\boldsymbol{D}_{NF}$, $\boldsymbol{D}_{FF}$	NF and FF region
D_R, D_θ, D_ϕ	CRBs for range and angle

A. Localization Scenario

As shown in Fig. 1, we consider a localization scenario consisting of K users, a RIS, and a BS[1]. The users are stationary or low-speed compared with the localization time. Users and the BS are each equipped with a single antenna. The RIS containing $N = N_1 \times N_2$ elements is placed on the Y-Z plane, where the center of the RIS is at the origin of the coordinate system. There also exist obstacles, which reflect the signals and create multipath channels. We assume the line-of-sight path between each user and the BS is blocked, and there only exist the reflection paths via the RIS.

During the localization process, each user sends a narrowband signal to the RIS, and the RIS reflects it to the BS. To avoid multi-user interference, frequency division multiplexing (FDM) is for different users, i.e. the k-th user sends the signal over frequency f_k, where $f_k \neq f_k'$ ($k \neq k'$). The BS then estimates the locations of the users using the received signals. We assume the locations of the RIS and the BS are known, while the locations of the users and scattering centers in the obstacles [19] are unknown. Whether the users are located in the NF or the FF regions of the RIS is also unknown.

B. Signal Model

In the uplink transmission of the k-th user, the signal y_k received by the BS is given by

$$y_k = \boldsymbol{\beta}^T \boldsymbol{h}_k(\boldsymbol{p}_k, \boldsymbol{Q}) s_k + \epsilon, \tag{1}$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, ..., \beta_N]^T \in \mathbb{C}^{N\times 1}$ is the phase shift vector of the RIS which is shared for all K users, β_n represents the phase shift of the n-th RIS element with $|\beta_n| = 1$, $s_k \in \mathbb{C}^{1\times 1}$ represents the narrowband signal transmitted by the k-th user, and ϵ represents received Gaussian noise with noise power σ^2 and mean 0. The location of the k-th user is given by $\boldsymbol{p}_k = [R_k, \theta_k, \phi_k]^T$, with R_k, θ_k, ϕ_k being the range, polar angle, and azimuth angle, respectively. Here $\boldsymbol{Q} = [\boldsymbol{q}_1, ..., \boldsymbol{q}_{L_{max}}] \in \mathbb{C}^{3\times L_{max}}$ denotes the locations of the scatters, where the number of scatters is assumed to be less than L_{max}, and the location of the l-th scatter is given by $\boldsymbol{q}_l = (r_l, \vartheta_l, \varphi_l)^T$. The cascaded channel between the k-th user and the BS $\boldsymbol{h}_k(\boldsymbol{p}_k, \boldsymbol{Q})$ is given by to $\boldsymbol{h}_k^A \odot \boldsymbol{h}_k^t(\boldsymbol{p}_k, \boldsymbol{Q})$, where $\boldsymbol{h}_k^A$, $\odot$, and $\boldsymbol{h}_k^t(\boldsymbol{p}_k, \boldsymbol{Q})$ denote the RIS-BS channel for the k-th user, the Hadamard product, and the channel between the RIS and the k-th user, respectively. The channel between the k-th user and the RIS is given by [20]

$$\boldsymbol{h}_k^t(\boldsymbol{p}_k, \boldsymbol{Q}) = \boldsymbol{h}_{k0}^t(\boldsymbol{p}_k) + \sum_{l=1}^{L} \boldsymbol{h}_{kl}^t(\boldsymbol{p}_k, \boldsymbol{q}_l), \tag{2}$$

where $\boldsymbol{h}_{kl}^t(\boldsymbol{p}_k, \boldsymbol{q}_l)$ represents the l-th path between the RIS and the k-th user, and L is the number of scatters in the scenario. We assume $l = 0$ is associated with the direct path between the k-th user and RIS, while $l \neq 0$ are the paths via scatters. Since users or scatters could be in either NF or FF regions of the RIS, the channels in these two cases are modeled separately.

1) Channel Models for the NF Region: When the user or the scatter is located in the NF region, the signal received by the RIS is described using the spherical wave model and cannot be approximated as a plane wave for the RIS. Thus, the direct user-RIS channel is modeled by [21]

$$\boldsymbol{h}_{k0}^t(\boldsymbol{p}_k) = \alpha_k \boldsymbol{b}(R_k, \theta_k, \phi_k),\ \boldsymbol{p}_k \in \boldsymbol{D}_{NF}, \tag{3}$$

where α_k is the channel gain including the effects of path loss and the directivity of the RIS elements [22], and $\boldsymbol{b}$ is the steering vector in the NF region, given as

$$\boldsymbol{b}(R_k, \theta_k, \phi_k) = \left[\exp\left(-j\frac{2\pi}{\lambda} d_{k1}^t\right), \ldots, \exp\left(-j\frac{2\pi}{\lambda} d_{kN}^t\right)\right]^T. \tag{4}$$

Here, d_{kn}^t is the distance between the k-th user and the n-th RIS element, and $\boldsymbol{D}_{NF}$ represents the NF region. Based on [18], the NF region is defined as the region where the maximum phase error between the phase shift calculated under plane wave approximation and the real phase shift is more than $\pi/8$. Mathematically, the NF region is given by $\boldsymbol{D}_{NF} = \{\boldsymbol{p} | \Delta\varphi(\boldsymbol{p}) > \pi/8\}$, where $\Delta\varphi(\boldsymbol{p})$ is the maximum phase error across all the RIS elements, given by

$$\Delta\varphi(\boldsymbol{p}) = \max_n \frac{2\pi}{\lambda}\left(d_n^t - (R - y_n \sin\theta \sin\phi - z_n \cos\theta)\right), \tag{5}$$

where d_n^t is the distance to the n-th RIS element, and $(0, y_n, z_n)$ is the coordinate of the n-th RIS element.

Similarly, we model the paths via the scatters. Specifically, when the l-th scatter is located in the NF region, the path between the l-th scatter and the RIS is given by

$$\boldsymbol{h}_{kl}^t(\boldsymbol{p}_k, \boldsymbol{q}_l) = \alpha_{kl} \boldsymbol{b}(r_l, \vartheta_l, \varphi_l),\ \boldsymbol{q}_l \in \boldsymbol{D}_{NF}, \tag{6}$$

where $\boldsymbol{b}(r_l, \vartheta_l, \varphi_l) = [\exp(-j\frac{2\pi}{\lambda} d_{l1}^t), ..., \exp(-j\frac{2\pi}{\lambda} d_{lN}^t)]^T$, and d_{ln}^t is the distance between l-th scatter and n-th element of the RIS. Here α_{kl} is the channel gain of the l-th path.

[1]In this work, we use reflective reconfigurable intelligent surface (RIS) as an example to implement HMIMO, and it can be easily replaced with other surfaces, such as transmissive RIS and reconfigurable holographic surface.

2) Channel Models for the FF Region: When the user is located in the FF region, we adopt the plane wave approximation to model the received signals. The direct user-RIS channel is modeled by [21]

$$\boldsymbol{h}_{k0}^{t}(\boldsymbol{p}_k) = \alpha_k \boldsymbol{a}(\theta_k, \phi_k),\ \boldsymbol{p}_k \in \boldsymbol{D}_{FF}, \tag{7}$$

where $\boldsymbol{a}(\theta_k, \phi_k)$ is the FF steering vector and is given as [23] [24]

$$\boldsymbol{a}(\theta_k, \phi_k) = \left[\exp\left(-j\frac{2\pi}{\lambda}(-y_1 \sin\theta_k \sin\phi_k - z_1 \cos\theta_k)\right), \right. \\ \left. ..., \exp\left(-j\frac{2\pi}{\lambda}(-y_N \sin\theta_k \sin\phi_k - z_N \cos\theta_k)\right)\right]^T. \tag{8}$$

In (7), $\boldsymbol{D}_{FF}$ represents the FF region, which is defined as the area where the maximum phase error is less than $\pi/8$, i.e., $\boldsymbol{D}_{FF} = \{\boldsymbol{p} | \Delta\varphi(\boldsymbol{p}) < \pi/8\}$. When the l-th scatter is located in the FF region, the path between the l-th scatter and the RIS can be modeled by

$$\boldsymbol{h}_{kl}^{t}(\boldsymbol{p}_k, \boldsymbol{q}_l) = \alpha_{kl} \boldsymbol{a}(\vartheta_l, \varphi_l),\ \boldsymbol{q}_l \in \boldsymbol{D}_{FF}. \tag{9}$$

C. Localization Protocol

We propose an RIS-enabled hybrid NF and FF localization protocol, which improves localization accuracy by iteratively localizing the users and optimizing the RIS phase shifts based on the estimated locations. The localization process is divided into C cycles with cycle duration being δ. Note that cycle duration δ can be adjusted for different systems. Each cycle contains three steps: transmission, localization, and optimization. The process of the localization protocol is illustrated in Fig. 2.

1) Transmission: In this step, the users send signals to the RIS, which are reflected to the BS. Let $\boldsymbol{g}^{(c)} = [g_1^{(c)}, ..., g_K^{(c)}]^T \in \mathbb{C}^{K\times 1}$ denote the signal received by the BS in the c-th cycle. The RIS phase shifts are set randomly in the first cycle, while in the following $C-1$ cycles, the phase shifts are selected based on the optimization results in the previous cycle, which will be described in the optimization step.

2) Localization: In the next step, the BS estimates users' locations using the received signals. Specifically, in the c-th cycle, based on the received signals $\boldsymbol{G}^{(c)} = [\boldsymbol{g}^{(1)}, ..., \boldsymbol{g}^{(c)}]^T \in \mathbb{C}^{K\times c}$, we determine whether the users and the scatters are located in the NF or the FF region and jointly estimate their locations $\boldsymbol{P}^{(c)} = [\boldsymbol{p}_1^{(c)}, ..., \boldsymbol{p}_K^{(c)}]$ and $\boldsymbol{Q}^{(c)} = [\boldsymbol{q}_1^{(c)}, ..., \boldsymbol{q}_L^{(c)}]$. This step focuses on the problem of accurately locating users given the received signals and the RIS phase shifts. The details of the localization algorithm are introduced in Sec. III.

3) Optimization: In the rest time of the c-th cycle, the optimal RIS phase shifts $\boldsymbol{\beta}^{(c+1)}$ are selected according to the estimated locations in the former step. Note that this step is not executed in the last cycle. This step focuses on the task of calculating the optimal RIS phase shifts for the next cycle based on the user localization results of the previous step. The details of optimization are introduced in Sec. IV.

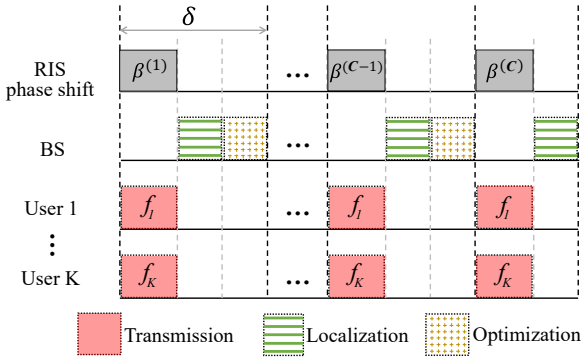

Fig. 2. RIS-enabled hybrid NF and FF source localization protocol.

III. RIS-enabled Hybrid Near-field and Far-field Localization

In this section, we first formulate the localization problem and then propose a localization algorithm to solve the formulated problem.

A. Localization Problem Formulation

We formulate the multi-user localization problem to minimize the user localization loss, which is defined as the sum of the l_2-norm of the residual between the received signals and the signals reconstructed given user locations. Thus, the localization problem is formulated as

$$\text{P1}: \min_{\boldsymbol{P}} \sum_{k=1}^{K} \left\| \boldsymbol{g}_k^{(c)} - (\boldsymbol{B}^{(c)})^T \boldsymbol{h}_k(\boldsymbol{p}_k, \overline{\boldsymbol{Q}}) s_k \right\|_2^2, \tag{10a}$$

$$s.t.\ \ \boldsymbol{p}_k \in \boldsymbol{D}_{NF} \cup \boldsymbol{D}_{FF}, \forall k, \tag{10b}$$

where $\boldsymbol{g}_k^{(c)} = [g_k^{(1)}, ..., g_k^{(c)}]^T \in \mathbb{C}^{c\times 1}$ is the received signals of the k-th user in the previous c cycles, i.e., the k-th row of $\boldsymbol{G}^{(c)}$. Here $\boldsymbol{B}^{(c)} = [\boldsymbol{\beta}^{(1)}, ..., \boldsymbol{\beta}^{(c)}] \in \mathbb{C}^{N\times c}$ is the phase shifts of the RIS in the previous c cycles, $\boldsymbol{P} = [\boldsymbol{p}_1, ..., \boldsymbol{p}_K]$ is the locations of the users, $\overline{\boldsymbol{Q}}$ is the ground truth of scatter locations. It can be observed from (10a) that the objective function is affected by the accuracy of user location estimation. By solving (P1), we can minimize the estimation error of the user location and eliminate the interference caused by the scattering paths.

Compared to single-user localization, multi-user localization is more challenging because the channels of different users are coupled with each other. Specifically, the signals transmitted by different users at different frequencies may pass through the same scatter. Thus, the localization loss of each user, i.e., a term in the summation (10a), is affected by not only the location of each user but also the scatter locations $\boldsymbol{Q}$. This is different from the single-user localization case where the locations of a user and the scatters are optimized to minimize only one localization loss term and the coupling of the localization losses among different users is not considered.

Due to the non-convex nature of the above problem (P1), conventional algorithms like gradient descent can easily fall into local minima. To solve it effectively, a location estimation method by modifying the grid search method is proposed to perform a global search. Specifically, we first sample the

search domain in NF and FF regions to create a grid map of the sampled candidate locations, denoted by $\boldsymbol{Z} = [\boldsymbol{Z}_{near}, \boldsymbol{Z}_{far}]$, where $\boldsymbol{Z}_{near}$ and $\boldsymbol{Z}_{far}$ are the sampled candidate locations in the NF and FF, respectively. Then, the most suitable sample locations are selected to minimize the localization loss.

Considering the different signal models in the NF and FF regions, the locations in these two regions are sampled in different ways, which are elaborated in the following.

NF case: In the NF region, R, θ, ϕ are uniformly sampled with sampling spacings $\Delta R, \Delta\theta = \pi/N_\theta, \Delta\phi = \pi/N_\phi$, respectively, where N_θ and N_ϕ are predetermined parameters. Thus, S sampled candidate locations in the NF are obtained.

$$\boldsymbol{Z}_{near} = \left[[R_1, \theta_1, \phi_1]^T, ..., [R_S, \theta_S, \phi_S]^T \right]. \quad (11)$$

We define the atom channels as the user-RIS channels or the scatter-RIS channels between the RIS and every possible candidate location of the users or the scatters. The atom channel in the NF region is given by [25]

$$\boldsymbol{F}_{near} = [\boldsymbol{b}(R_1, \theta_1, \phi_1), ..., \boldsymbol{b}(R_S, \theta_S, \phi_S)] \in \mathbb{C}^{N \times S}, \quad (12)$$

where $\boldsymbol{b}(R_i, \theta_i, \phi_i)$ is the NF steering vector given the location $[R_i, \theta_i, \phi_i]^T$.

FF case: In the FF region, only the angles θ, ϕ are sampled. The range R is not sampled because it does not affect the FF steering vector and it cannot be estimated similarly to NF case. The primary reason lies in the angular relationships between the user and the RIS elements. Specifically, in the NF, the angles between the user and each RIS element are different. This leads to different phase shift changes for paths from the user to each RIS element and varied received signals when the range varies, which allows for the estimation of the range R. In contrast, for FF users, since the angles between the user and each RIS element are almost the same, when the range changes, the phase shift changes for the path from the user to each RIS element are the same. This uniform phase shift leads to an overall phase change in the received signal, hence range estimation cannot be achieved.

As a result, the number of candidate locations and the algorithm complexity are significantly reduced compared with NF. We apply the same angle sampling methods in the FF as in the NF. The atom channel of the FF region is given by [26]

$$\boldsymbol{F}_{far} = [\boldsymbol{a}(\theta_1, \phi_1), ..., \boldsymbol{a}(\theta_{N_\theta N_\phi}, \phi_{N_\theta N_\phi})] \in \mathbb{C}^{N \times N_\theta N_\phi}, \quad (13)$$

where $\boldsymbol{a}(\theta_i, \phi_i)$ is the FF steering vector of the angle $[\theta_i, \phi_i]$.

We define $\boldsymbol{F} = [\boldsymbol{F}_{near}, \boldsymbol{F}_{far}] \in \mathbb{C}^{N \times M}$ as the atom channels for the hybrid field, where $M = S + N_\theta N_\phi$. For better illustration, the sampled locations at $\theta = \pi/2$ plane for both the NF and FF regions are shown in Fig. 3.

Based on the channel model (2) and the atom channels (12) and (13), each user-RIS channel can be approximated as a linear combination of multiple atom channels. We define $\boldsymbol{u}_k \in \mathbb{C}^{M \times 1}$ as the gains of the atom channels for the k-th user-RIS channels, which can be decomposed into two components: $\boldsymbol{u}_k^{\boldsymbol{P}}$ for the direct path, and $\boldsymbol{u}_k^{\boldsymbol{Q}}$ for the scattering path. Both $\boldsymbol{u}_k^{\boldsymbol{P}}$ and $\boldsymbol{u}_k^{\boldsymbol{Q}}$ are unknown, and our aim is to accurately estimate $\boldsymbol{u}_k^{\boldsymbol{P}}$, which contains user locations information. Specifically, $u_{ik}^{\boldsymbol{P}}$, the i-th element of $\boldsymbol{u}_k^{\boldsymbol{P}}$, denotes the gain of the i-th atom channel in the direct path between the k-th user and the RIS. If $u_{ik}^{\boldsymbol{P}} = 0$, it indicates that the k-th user is not at location $\boldsymbol{z}_i$, the i-th column of the candidate locations $\boldsymbol{Z}$. Otherwise, the k-th user is located at $\boldsymbol{z}_i$, and the estimated gain is $u_{ik}^{\boldsymbol{P}}$. In contrast, we do not care about the estimation accuracy of $\boldsymbol{u}_k^{\boldsymbol{Q}}$, because it is only influenced by scatter locations.

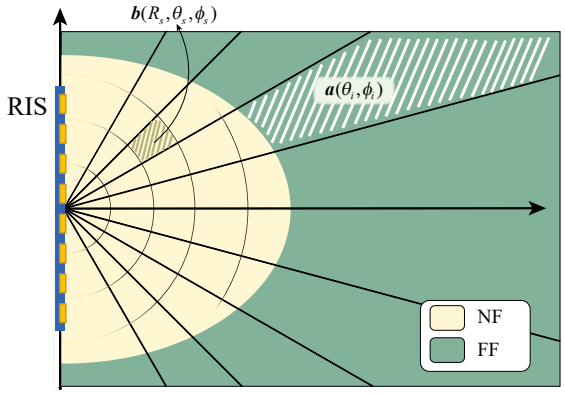

Fig. 3. Illustration of the sampled locations at $\theta = \pi/2$.

The k-th user-RIS channels are thus given as

$$\boldsymbol{h}_k^t(\boldsymbol{p}_k, \boldsymbol{Q}) = \boldsymbol{F}\boldsymbol{u}_k = \boldsymbol{F}(\boldsymbol{u}_k^{\boldsymbol{P}} + \boldsymbol{u}_k^{\boldsymbol{Q}}). \quad (14)$$

The cascaded channel of the k-th user is given by

$$\boldsymbol{h}_k(\boldsymbol{p}_k, \boldsymbol{Q}) = \boldsymbol{h}_k^A \odot \boldsymbol{h}_k^t = \text{diag}\{\boldsymbol{h}_k^A\}\boldsymbol{F}(\boldsymbol{u}_k^{\boldsymbol{P}} + \boldsymbol{u}_k^{\boldsymbol{Q}}). \quad (15)$$

Since the number of users and scatters are generally much smaller than the number of atom channels, $\boldsymbol{u}_k$ can be directly solved base on traditional sparse recovery method like orthogonal matching pursuit (OMP). However, the user location estimated in this way is not accurate enough due to interference of scattering paths. To mitigate this interference and improve user localization accuracy, (P1) is approximated as

$$\text{P1'}: \min_{\boldsymbol{u}_k^{\boldsymbol{P}}, \forall k} \sum_{k=1}^{K} \left\| \boldsymbol{g}_k^{(c)} - \boldsymbol{\Lambda}^k(\boldsymbol{u}_k^{\boldsymbol{P}} + \overline{\boldsymbol{u}_k^{\boldsymbol{Q}}}) \right\|_2^2, \quad (16a)$$

$$s.t. \ \left\| \boldsymbol{u}_k^{\boldsymbol{P}} \right\|_0 = 1, \forall k, \quad (16b)$$

where $\overline{\boldsymbol{u}_k^{\boldsymbol{Q}}}$ is the ground truth of $\boldsymbol{u}_k^{\boldsymbol{Q}}$. Here, $\boldsymbol{\Lambda}^k = (\boldsymbol{B}^{(c)})^T \text{diag}\{\boldsymbol{h}_k^A\}\boldsymbol{F}s_k \in \mathbb{C}^{c \times M}$ is the atom signals, which are the linear transformation of the atom channels $\boldsymbol{F}$ by considering the effect of RIS phase shift $\boldsymbol{B}^{(c)}$, the RIS-BS channel $\boldsymbol{h}_k^A$, and the transmit signal s_k.

B. Localization Algorithm

From (P1'), it is evident that we cannot directly estimate the user's location since the scatter locations are unknown. Therefore, we need to estimate the locations of the scatters. However, considering that scattering paths are typically weaker than direct user-RIS path, it is difficult to estimate the scatter locations directly from the received signals. To tackle this issue, we leverage the idea of successive interference cancellation (SIC) that first estimates the stronger direct user-RIS paths, and then estimates the weaker scattering paths.

Algorithm 1 The proposed localization algorithm

Input: Recieved signal $\boldsymbol{G}^{(c)}$, the atom channels matrix $\boldsymbol{F}$, max number of scatters L_{max}
Output: The estimated locations of the users $\boldsymbol{P}^{(c)}$
Initialization: the atom signals $\boldsymbol{\Lambda}^k$, the received signals energy E, the support set $\mathcal{X} = \emptyset$.
for $k = 1, ..., K$ **do**
  Obtain the coarse estimated user locations based on (20)
  Calculate residual signal by (21) and (22)
end for
Initialize the residual signal $\boldsymbol{e}_k^{ini} = \boldsymbol{e}_k$
for $l = 1, ..., L_{max}$ **do**
  Obtain the estimated l-th scatter location by (24)
  Update support set $\mathcal{X} = \mathcal{X} \cup i$
  Calculate residual signals by (25) and (27)
  Calculate the energy of the residual signal E^r
  if $E^r < \gamma E$ **then** break **end if**
end for
for $k = 1, ..., K$ **do**
  Calculate the new residual signal by (28)
  Obtain the estimated location of k-th user by (30)
end for
Return: $\boldsymbol{P}^{(c)} = [\boldsymbol{p}_1^{(c)}, ..., \boldsymbol{p}_K^{(c)}]$

Therefore, we devise a three-step algorithm to solve (P1'). Specifically, we first ignore the scattering path and roughly estimate the locations of the users based on the received signal. The residual signal is then calculated by subtracting the direct user-RIS path component from the received signal to improve the estimation accuracy of scatter locations. Next, the scatter locations are estimated using the residual signal. In the third step, we subtract the scattering path from the received signal and re-estimate the user's location to achieve higher user localization accuracy. In the following, we introduce these subproblems and the algorithm in detail. The proposed algorithm is summarized in **Algorithm 1**.

1) User Location Estimation: In this part, we estimate the location of each user using the corresponding received signal. This is a coarse estimate because we ignore the scattering path in this step. Specifically, we find the atom signal with the highest energy contribution to the received signals and estimate the gain of the atom channel [27] [28]. For example, for the k-th user, we use the $\boldsymbol{g}_k$ to estimate its location, i.e.,

$$\text{P1-a}: \min_{\boldsymbol{u}_k^{\boldsymbol{P}}, \forall k} \sum_{k=1}^{K} \left\| \boldsymbol{g}_k^{(c)} - \boldsymbol{\Lambda}^k \boldsymbol{u}_k^{\boldsymbol{P}} \right\|_2^2, \tag{17a}$$

$$s.t. \ \left\| \boldsymbol{u}_k^{\boldsymbol{P}} \right\|_0 = 1, \forall k. \tag{17b}$$

Only one element in $\boldsymbol{u}_k^{\boldsymbol{P}}$ is nonzero, therefore $\boldsymbol{u}_k^{\boldsymbol{P}}$ is a 1-sparse vector, which enables the sparse recovery method to efficiently recover it [29]. Note that the terms of summation in (17a) are independent, therefore (P1-a) can be decoupled into K independent subproblems, given by

$$\text{P1-a-k}: \min_{\boldsymbol{u}_k^{\boldsymbol{P}}} \left\| \boldsymbol{g}_k^{(c)} - \boldsymbol{\Lambda}^k \boldsymbol{u}_k^{\boldsymbol{P}} \right\|_2^2, \tag{18a}$$

$$s.t. \ \left\| \boldsymbol{u}_k^{\boldsymbol{P}} \right\|_0 = 1. \tag{18b}$$

The received signal $\boldsymbol{g}_k^{(c)}$ contains multiple paths. Since the gain of the direct path is larger than that of the scattering paths, we choose the atom signal with the strongest correlation with received signal as the direct path corresponding to the user's location. Mathematically, for the k-th user, we have

$$i^k = \arg\max_i \left| (\boldsymbol{\Lambda}_i^k)^H \boldsymbol{g}_k^{(c)} \right| \tag{19}$$

where $\boldsymbol{\Lambda}_i^k \in \mathbb{C}^{N \times 1}$ is the i-th column of $\boldsymbol{\Lambda}^k$ that corresponds to $\boldsymbol{z}_i$, the i-th location in $\boldsymbol{Z}$. Then, we have $\hat{\boldsymbol{p}}_k = \boldsymbol{z}_{i^k}$, and the estimated location for all K users is given by

$$\hat{\boldsymbol{P}} = [\boldsymbol{z}_{i^1}, ..., \boldsymbol{z}_{i^K}]. \tag{20}$$

Note that the optimal gain for any given atom signal can be given in closed form as a function i

$$\hat{u}_{ik}^{\boldsymbol{P}} = {\boldsymbol{\Lambda}_i^k}^{\dagger} \boldsymbol{g}_k^{(c)}, \tag{21}$$

where $\hat{u}_{ik}^{\boldsymbol{P}}$ is the estimated value of the i-th element in $\boldsymbol{u}_k^{\boldsymbol{P}}$, and † represents the pseudo-inverse. The residual signal for the k-th user can be given by

$$\boldsymbol{e}_k^{(c)} = \boldsymbol{g}_k^{(c)} - \hat{u}_{ik}^{\boldsymbol{P}} \boldsymbol{\Lambda}_i^k. \tag{22}$$

In this way, signals with smaller interference for scatter localization is obtained.

2) Scatter Location Estimation: In this part, the locations of the scatters are estimated using all the residual signals. This part is iteratively conducted until the energy of the new residual signal is less than a percentage of the energy of the received signals or the number of iterations reaches L_{max}. This subproblem is modeled as

$$\text{P1-b}: \min_{\boldsymbol{u}_k^{\boldsymbol{Q}}, \forall k} \sum_{k=1}^{K} \left\| \boldsymbol{e}_k^{(c)} - \boldsymbol{\Lambda}^k \boldsymbol{u}_k^{\boldsymbol{Q}} \right\|_2^2, \tag{23a}$$

$$s.t. \ |\{i | \boldsymbol{v}_i \neq \boldsymbol{0}\}| \leq L_{max}, \tag{23b}$$

where $\boldsymbol{v}_i = [u_{i1}^{\boldsymbol{Q}}, ..., u_{iK}^{\boldsymbol{Q}}] \in \mathbb{C}^{1 \times K}$. Since the numbers of scatters are generally much smaller than the number of atom channels [30], we can use the sparse recovery method [27] to estimate the sparse vector $\boldsymbol{u}_k^{\boldsymbol{Q}}$ [29]. We define the support set $\mathcal{X}$ as the indices of the atom signals corresponding to the estimated scatter locations. In each iteration, a new index is added to the support set. Note that the atom signal for each user is different, the residual signal for different users is approximated using their corresponding atom signals, and then the location with the maximum correlation with the residual signal is selected as the estimated location of a scatter. Mathematically, we have

$$i = \arg\max_i \sum_{k=1}^{K} \left| (\boldsymbol{\Lambda}_i^k)^H \boldsymbol{e}_k^{(c)} \right| \tag{24}$$

Then, we update the support set as $\mathcal{X} = \mathcal{X} \cup i$. The gains of all the support signals in the support set are calculated through orthogonal least square, which is given by

$$\hat{\boldsymbol{u}}_{k,\mathcal{X}}^{\boldsymbol{Q}} = (\boldsymbol{\Lambda}_{\mathcal{X}}^k)^{\dagger} \boldsymbol{e}_k^{ini}, \tag{25}$$

where $\boldsymbol{e}_k^{ini}$ is the initial residual signal, and $\boldsymbol{\Lambda}_{\mathcal{X}}^k$ is the matrix generated by the atom signals selected by the indices in the support set, which is given by

$$\boldsymbol{\Lambda}_{\mathcal{X}}^k = [\boldsymbol{\Lambda}_j^k], j \in \mathcal{X}. \tag{26}$$

Using the estimated gain, the residual signal is updated by

$$\boldsymbol{e}_k^{(c)} = \boldsymbol{e}_k^{ini} - \boldsymbol{\Lambda}_{\mathcal{X}}^k \hat{\boldsymbol{u}}_{k,\mathcal{X}}^{\boldsymbol{Q}}. \tag{27}$$

Then the energy of the new residual signal $\boldsymbol{e}_k$ is calculated and denoted by E^r. If $E^r < \gamma E$, the iteration terminates, else the steps (24) to (27) are repeated, where γ is a predetermined parameter and E is the energy of the original received signals.

3) User Location Refinement: In this part, we optimize the locations of users based on the estimated scatter locations, i.e., re-solve the locations of users given the locations of scatters. First, we calculate the residual signal $\boldsymbol{r}_k^{(c)}$ by subtracting the scattering path from the original received signal.

$$\boldsymbol{r}_k^{(c)} = \boldsymbol{g}_k^{(c)} - \boldsymbol{\Lambda}_{\mathcal{X}}^k {\boldsymbol{\Lambda}_{\mathcal{X}}^k}^{\dagger} \boldsymbol{g}_k^{(c)}. \tag{28}$$

Then similar to the method of the user location estimation step, by solving the following subproblem,

$$\text{P1-c}: \min_{\boldsymbol{u}_k^{\boldsymbol{P}}, \forall k} \sum_{k=1}^{K} \left\| \boldsymbol{r}_k^{(c)} - \boldsymbol{\Lambda}^k \boldsymbol{u}_k^{\boldsymbol{P}} \right\|_2^2, \tag{29a}$$

$$s.t. \ \left\| \boldsymbol{u}_k^{\boldsymbol{P}} \right\|_0 = 1, \forall k, \tag{29b}$$

the index of the refined k-th user location is given by

$$i^k = \arg\max_i \left| (\boldsymbol{\Lambda}_i^k)^H \boldsymbol{r}_k^{(c)} \right|. \tag{30}$$

Then we have $\hat{\boldsymbol{p}}_k = \boldsymbol{z}_{i^k}$ and $\hat{\boldsymbol{P}} = [\boldsymbol{z}_{i^1}, ..., \boldsymbol{z}_{i^K}]$.

IV. RIS Phase Shift Optimization

In this section, we first formulate the RIS phase shift optimization problem and propose an algorithm to solve it.

A. Phase Shift Optimization Problem Formulation

The selected optimization metric is CRB, a standard metric for assessing the parameter estimation error [31]. The CRB is derived from the model of the received signal for $c+1$ cycles, which is given by

$$\boldsymbol{y}_k = [y_k^{(1)}, ..., y_k^{(c+1)}]^T = (\boldsymbol{B}^{(c+1)})^T \boldsymbol{h}_k s + \boldsymbol{\epsilon}, \tag{31}$$

where $\boldsymbol{\epsilon} = [\epsilon^{(1)}, ..., \epsilon^{(c+1)}]^T$ is the independent and identically distributed zero-mean Gaussian noise with variance σ^2. Note that the received signal is influenced by the RIS phase shift $\boldsymbol{\beta}^{(c+1)}$, which is the $(c+1)$-th column of $\boldsymbol{B}^{(c+1)}$. Thus, we can obtain the expressions between $\boldsymbol{\beta}^{(c+1)}$ and CRB, listed in the following. By optimizing the RIS phase shifts $\boldsymbol{\beta}^{(c+1)}$, we can minimize the CRB to improve the localization accuracy.

Specifically, we formulate the optimization problem to minimize the sum of weighted CRBs of range and angle estimation errors. We introduce a weight matrix to address the different unit of angle and range CRBs [32]. Note that we estimate angles and ranges for NF users, while only angles for FF users. Hence, we give the expressions of the weighted CRBs for NF and FF users separately.

1) CRBs for NF Users: If the estimated location $\boldsymbol{p}_k^{(c)}$ is in the NF region in the c-th cycle, we expect that in the $(c+1)$-th cycle, the range and angles of the k-th user are estimated, denoted by $\boldsymbol{p}_k^{(c+1)} = [R_k, \theta_k, \phi_k]^T$. The expressions of CRBs for the NF case are provided in proposition 1.

Proposition 1: For the NF case, the CRBs of the unknown parameters $\boldsymbol{p}_k = [R_k, \theta_k, \phi_k]^T$ are given by

$$D_R = \left(\boldsymbol{J}_{NF}^{-1}\right)_{1,1}, \tag{32}$$

$$D_\theta = \left(\boldsymbol{J}_{NF}^{-1}\right)_{2,2}, \tag{33}$$

$$D_\phi = \left(\boldsymbol{J}_{NF}^{-1}\right)_{3,3}, \tag{34}$$

where $\boldsymbol{J}_{NF}$ is the 3×3 Fisher information matrix of $\boldsymbol{p}_k$ in the NF region, and the (i,j)-th element in $\boldsymbol{J}_{NF}$ is given by [33]

$$[\boldsymbol{J}_{NF}]_{i,j} = \frac{2}{\sigma^2} \sum_{m=1}^{c+1} Re\left\{ \frac{\partial (\mu^{(m)})^H}{\partial p_i} \frac{\partial \mu^{(m)}}{\partial p_j} \right\}, \tag{35}$$

where $\mu^{(m)}$ is defined as the noise-free received signal

$$\mu^{(m)} = (\boldsymbol{\beta}^{(m)})^T \boldsymbol{h}_k^{NF} s_k. \tag{36}$$

Here $\boldsymbol{h}_k^{NF}$ is the channel calculated by using $\boldsymbol{p}_k^{(c)} \in \boldsymbol{D}_{NF}$ and $\boldsymbol{Q}^{(c)}$, and p_i is the i-th element in $\boldsymbol{p}_k$.

Proof: See Appendix A. ∎

2) CRBs for FF Users: If the estimated location $\boldsymbol{p}_k^{(c)}$ is in the FF region, we expect that only the angles are estimated for the k-th user in the $(c+1)$-th cycle, and the estimated parameters are denoted by $\boldsymbol{p}_k^{(c+1)} = [\theta_k, \phi_k]^T$. The expressions of CRBs for the FF case are provided in proposition 2.

Proposition 2: For the FF case, the CRBs of the unknown parameters $\boldsymbol{p}_k$ are given by

$$D_\theta = \left(\boldsymbol{J}_{FF}^{-1}\right)_{1,1}, \tag{37}$$

$$D_\phi = \left(\boldsymbol{J}_{FF}^{-1}\right)_{2,2}, \tag{38}$$

where $\boldsymbol{J}_{FF}$ is the 2×2 Fisher information matrix of $\boldsymbol{p}_k$ in the FF region, and the (i,j)-th element is given by [33]

$$[\boldsymbol{J}_{FF}]_{i,j} = \frac{2}{\sigma^2} \sum_{m=1}^{c+1} Re\left\{ \frac{\partial (\mu^{(m)})^H}{\partial p_i} \frac{\partial \mu^{(m)}}{\partial p_j} \right\}, \tag{39}$$

where $\mu^{(m)}$ is defined as the noise-free received signal

$$\mu^{(m)} = (\boldsymbol{\beta}^{(m)})^T \boldsymbol{h}_k^{FF} s_k, \tag{40}$$

and $\boldsymbol{h}_k^{FF}$ is the channel calculated using $\boldsymbol{p}_k^{(c)}$ and $\boldsymbol{Q}^{(c)}$.

Proof: See Appendix B. ∎

3) Optimization Problem Formulation: To achieve high localization accuracy for multiple users, we optimize the phase shifts $\boldsymbol{\beta}^{(c+1)}$ by minimizing the sum of CRBs of all the users. The optimization problem can be formulated as

$$\text{P2}: \min_{\boldsymbol{\beta}^{(c+1)}} f(\boldsymbol{P}, \boldsymbol{\beta}^{(c+1)}) = \sum_{k \in \Psi_{NF}} tr\left(\boldsymbol{J}_{NF,k}^{-1} \boldsymbol{W}_{NF} \right) + \sum_{k \in \Psi_{FF}} tr\left(\boldsymbol{J}_{FF,k}^{-1} \boldsymbol{W}_{FF} \right), \tag{41a}$$

$$s.t. \ |\beta_n^{(c+1)}| = 1, \ \forall n = 1, 2, ...N, \tag{41b}$$

where $\boldsymbol{J}_{NF,k}$ or $\boldsymbol{J}_{FF,k}$ is the Fisher information matrix of $\boldsymbol{p}_k$ when the k-th user is located in the NF or FF region and is a function of the RIS phase shifts $\boldsymbol{\beta}^{(c+1)}$. Here $\boldsymbol{W}_{NF} = \text{diag}\{w_1, w_2, w_3\}$ and $\boldsymbol{W}_{FF} = \text{diag}\{w_4, w_5\}$ are the weight matrices of the CRBs in the NF and FF regions, respectively [32]. Ψ_{NF} and Ψ_{FF} are the sets of estimated NF and FF users, which are mathematically given by

$$\Psi_{NF} = \{k|\boldsymbol{p}_k^{(c)} \in \boldsymbol{D}_{NF}\}, \tag{42}$$

$$\Psi_{FF} = \{k|\boldsymbol{p}_k^{(c)} \in \boldsymbol{D}_{FF}\}. \tag{43}$$

$\beta_n^{(c+1)}$ is the RIS phase shift of the n-th element in the $(c+1)$-th cycle, satisfying the constant modulus constraint (41b).

B. RIS Phase Shift Optimization Algorithm

We design an optimization algorithm based on the complex circle manifold (CCM) method to tackle (P2). Due to the constant modulus constraint (41b), the problem (P2) is non-convex, which is numerically difficult to handle. Fortunately, the solution can be considered as lying on the CCM to satisfy the constant modulus, where the manifold is represented as

$$\mathcal{M}^N = \left\{\boldsymbol{\beta}^{(c+1)} \in \mathbb{C}^N : |\beta_1^{(c+1)}| = ... = |\beta_N^{(c+1)}| = 1\right\}. \tag{44}$$

The main idea of the CCM-based optimization method is to iteratively apply gradient descent in the manifold space. After several iterations, the algorithm terminates when the difference between two iterations of $f(\boldsymbol{P}, \boldsymbol{\beta}^{(c+1)})$ is less than a constant ζ or the number of iterations exceeds I, where ζ and I are selected to comply with the time constraints of the optimization step. The CCM-based optimization method is composed of four main steps in each iteration:

1) Compute the gradient in Euclidean space: We use the Euclidean gradient as the search direction for the minimization problem in Euclidean space. The Euclidean gradient of $f(\boldsymbol{\beta}^{(c+1)})$ is given by

$$\begin{aligned}\nabla f(\boldsymbol{\beta}^{(c+1)}) = 2\Big(&\sum_{k\in\Psi_{NF}} \left[\frac{w_1 \partial D_{R,k}}{\partial \boldsymbol{\beta}^{(c+1)^*}} + \frac{w_2 \partial D_{\theta,k}}{\partial \boldsymbol{\beta}^{(c+1)^*}} + \frac{w_3 \partial D_{\phi,k}}{\partial \boldsymbol{\beta}^{(c+1)^*}}\right] \\ &+ \sum_{k\in\Psi_{FF}} \left[\frac{w_4 \partial D_{\theta,k}}{\partial \boldsymbol{\beta}^{(c+1)^*}} + \frac{w_5 \partial D_{\phi,k}}{\partial \boldsymbol{\beta}^{(c+1)^*}}\right]\Big),\end{aligned} \tag{45}$$

where the variable $\boldsymbol{P}$ is omitted in $f(\boldsymbol{P}, \boldsymbol{\beta}^{(c+1)})$ for simplicity. Similar to the proof in [8], the specific expressions of the differentials of CRBs are given in Appendix C.

2) Compute the Riemannian gradient: The Riemannian gradient is the projection of the Euclidean gradient onto the tangent space of the complex circle manifold. The Riemannian gradient of the objective function $f(\boldsymbol{\beta}^{(c+1)})$ at the point $\boldsymbol{\beta}_j^{(c+1)}$ on the complex circle manifold $\mathcal{M}$ is given as [8]

$$\begin{aligned}\nabla_{\mathcal{M}} f(\boldsymbol{\beta}_j^{(c+1)}) = &-\nabla f(\boldsymbol{\beta}_j^{(c+1)}) \\ &- Re\left\{\left(\nabla f(\boldsymbol{\beta}_j^{(c+1)})\right)^* \odot \boldsymbol{\beta}_j^{(c+1)}\right\} \odot \boldsymbol{\beta}_j^{(c+1)}.\end{aligned} \tag{46}$$

3) Update over the Tangent Space: We choose a step size to update the current point, which is mathematically given as

$$\bar{\boldsymbol{\beta}}_j^{(c+1)} = \boldsymbol{\beta}_j^{(c+1)} + \alpha_j \nabla_{\mathcal{M}} g(\boldsymbol{\beta}_j^{(c+1)}), \tag{47}$$

where α_j is the step size in the j-th iteration.

Algorithm 2 CCM-based RIS Phase Shift Optimization Algorithm

Input: Estimated location $\hat{\boldsymbol{p}}^{(c)}$, RIS phase shift for previous c cycles $\boldsymbol{B}^{(c)}$
Output: Optimal RIS phase shift $\boldsymbol{\beta}^{(c+1)}$

```
Initialize: j = 0, β_0 ∈ M;
while |g(β_{j+1}^{(c+1)}) − g(β_j^{(c+1)})| > ζ and j < I do
    Compute the Euclidean gradient ∇f(β_j^{(c+1)}) according to (45)
    Calculate the Riemannian gradient ∇_M f(β_j^{(c+1)}) according to (46);
    Compute the RIS phase shift update on the tangent space β̄_j^{(k+1)} according to (47);
    Update RIS phase shift β_{j+1}^{(k+1)} according to (48);
    j = j + 1;
end while
```

4) Retract onto the manifold: After the update, the new point $\bar{\boldsymbol{\beta}}_j^{(c+1)}$ generally does not lie on the manifold $\mathcal{M}$. By using the retraction operator, the new point is mapped into the manifold. The retraction operator is given as

$$\boldsymbol{\beta}_{j+1}^{(c+1)} = \bar{\boldsymbol{\beta}}_j^{(c+1)} \odot \frac{1}{|\bar{\boldsymbol{\beta}}_j^{(c+1)}|}. \tag{48}$$

The proposed algorithm is summarized in **Algorithm 2**.

V. Performance Analysis

In this section, we analyze the complexity of the proposed method and discuss its localization performance.

A. Algorithm Complexity

1) Complexity of the localization algorithm: For the user location estimation step of the localization algorithm, the method is similar with the first iteration of the OMP method. Therefore, the computational cost of the user estimation step is $\mathcal{O}(cMK)$ [34]. In the scatter location estimation step, the complexity of this algorithm has been analyzed in [35]. We estimate the scatter locations for at most L_{max} rounds. Hence, the computational cost is $\mathcal{O}(cKML_{max})$. Then, similar with the user location estimation step, the computational cost of user location refinement step is $\mathcal{O}(cMK)$.

Hence, the overall cost of the localization algorithm for the c-th cycle is $\mathcal{O}(cMK) + \mathcal{O}(cKML_{max})$. The overall complexity for a C-cycle localization algorithm is $\mathcal{O}(C(C+1)MK) + \mathcal{O}(C(C+1)KML_{max})$.

2) Complexity of the RIS optimization algorithm: The complexity of the CCM algorithm has been analyzed in [36]. Let us denote the total number of iterations required to converge by T_{CCM}. Then the total complexity of the optimization algorithm is $\mathcal{O}(T_{CCM}N^2)$. Since the optimization algorithm is conducted for $C-1$ times, the overall complexity for the optimization algorithm is $\mathcal{O}((C-1)T_{CCM}N^2)$.

B. Localization Performance

Localization errors primarily arise from two sources: one is the failure to estimate the grid closest to the user, which we refer to as grid misjudgment; the other is the error between the user's continuous location in actual space and the discrete grids, known as grid mismatch, reflecting the inherent limitation of discretizing continuous space [37]. Mathematically, suppose the real location of the k-th user is $\boldsymbol{p}_k$, the estimation result is $\boldsymbol{z}_{m'}$, and the nearest grid for the k-th user is $\boldsymbol{z}_m$. Then the average localization error can be given by

$$\begin{aligned}\mathbb{E}(l_e) &= \mathbb{E}\{\|\boldsymbol{p}_k - \boldsymbol{z}_{m'}\|\} \\ &\leq \mathbb{E}\{\|\boldsymbol{p}_k - \boldsymbol{z}_m\|\} + \mathbb{E}\{\|\boldsymbol{z}_m - \boldsymbol{z}_{m'}\|\}, \end{aligned} \tag{49}$$

where the first term corresponds to grid mismatch, and the second term is grid misjudgment. In the following, we analyze these two sources of errors respectively.

1) Number of Cycles: The expectation of localization error can be given by [38]

$$\mathbb{E}(l_e) = \sum_{m=1}^{M} \gamma_m^k \int_{\boldsymbol{g}_k^{(C)} \in \mathcal{R}_{km}} \mathbb{P}(\boldsymbol{g}_k^{(C)} | \boldsymbol{B}^{(C)}, \boldsymbol{p}_k, \boldsymbol{Q}) d\boldsymbol{g}_k^{(C)}, \tag{50}$$

where $\mathcal{R}_{km}$ is the decision region for the m-th candidate location. The integral in (50) is the probability that the estimated location is the m-th candidate location, given the ground truth that the location of the k-th user is $\boldsymbol{p}_k$. Here γ_m^k is the error parameter, defined as $\gamma_m^k = \|\boldsymbol{p}_k - \boldsymbol{z}_m\|$.

We assume all locations have the same prior probabilities. Hence, the decision region can be given by

$$\begin{aligned}\mathcal{R}_{km} = \{\boldsymbol{g}_k^{(C)} :& \mathbb{P}(\boldsymbol{g}_k^{(C)} | \boldsymbol{B}^{(C)}, \boldsymbol{z}_m, \boldsymbol{Q}) \\ &\geq \mathbb{P}(\boldsymbol{g}_k^{(C)} | \boldsymbol{B}^{(C)}, \boldsymbol{z}_{m'}, \boldsymbol{Q}), \forall m' \neq m\}. \end{aligned} \tag{51}$$

Since the noise obeys Gaussian distribution, we have

$$\mathcal{R}_{km} = \{\boldsymbol{g}_k^{(C)} : |\boldsymbol{g}_k^{(C)} - \boldsymbol{\mu}_m|^2 \leq |\boldsymbol{g}_k^{(C)} - \boldsymbol{\mu}_{m'}|^2, \forall m' \neq m\}, \tag{52}$$

where $\boldsymbol{\mu}_m = [\mu_m^{(1)}, ..., \mu_m^{(C)}]^T$ is the noise-free signal for the m-th grid. Let $\boldsymbol{\xi}_k = \boldsymbol{g}_k^{(C)} - \boldsymbol{\mu}_m$, we have [38]

$$\begin{aligned}\mathcal{R}_{km} =& \{\boldsymbol{g}_k^{(C)} : |\boldsymbol{\xi}_k|^2 \leq |\boldsymbol{\xi}_k + \boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}|^2, \forall m' \neq m\}, \\ =& \{\boldsymbol{g}_k^{(C)} : |\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}|^2 + 2\boldsymbol{\xi}_k(\overline{\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}}) \geq 0, \\ & \forall m' \neq m\}. \end{aligned} \tag{53}$$

Assuming the user is located at a grid location $\boldsymbol{z}_m$, when $C \to \infty$, the first term is greater than 0, and the second term converges to 0 because $\boldsymbol{\xi}_k$ and $\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}$ are independent and $\mathbb{E}(\boldsymbol{\xi}_k) = 0$. Therefore, we have

$$\lim_{C \to \infty} \mathbb{P}(\boldsymbol{g}_k^{(C)} \in \mathcal{R}_{km}) = 1. \tag{54}$$

Hence, when ignoring grid mismatch, the grid misjudgment error $\mathbb{E}\{\|\boldsymbol{z}_m - \boldsymbol{z}_{m'}\|\}$ converges to 0.

When the user is not at a grid, if we assume the midpoint of $\boldsymbol{z}_m$ and $\boldsymbol{z}_{m'}$ also corresponds the midpoint of $\boldsymbol{\mu}_m$ and $\boldsymbol{\mu}_{m'}$, then we have $\mathbb{E}(\boldsymbol{\xi}_k) \leq |\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}|/2$. Hence, the $|\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}|^2 + 2\boldsymbol{\xi}_k(\overline{\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}}) \geq 0$ is still alwasy true, and we still have $\lim_{C \to \infty} \mathbb{P}(\boldsymbol{g}_k^{(C)} \in \mathcal{R}_{km}) = 1$. However, the number of cycles does not affect the grid mismatch error. Therefore, the localization error would converge to a non-zero number.

Remark 1: By increasing the number of estimation cycles C, the localization error first decreases and then remains fixed.

2) Sampling Spacing: First, we consider the scenario $C \to \infty$. According to the previous analysis, the localization error

$$\begin{aligned}\mathbb{E}(l_e) &\leq \mathbb{E}\{\|\boldsymbol{p}_k - \boldsymbol{z}_m\|\} + \mathbb{E}\{\|\boldsymbol{z}_m - \boldsymbol{z}_{m'}\|\}, \\ &\overset{C \to \infty}{\approx} \mathbb{E}\{\|\boldsymbol{p}_k - \boldsymbol{z}_m\|\}, \end{aligned} \tag{55}$$

where $\mathbb{E}\{\|\boldsymbol{p}_k - \boldsymbol{z}_m\|\}$ is directly related to the sampling spacing. Hence, a finer grid sampling could alleviate the grid mismatch problem [37].

When constrained by a limited number of cycle, the approximation in (55) no longer holds. Note that $\boldsymbol{\xi}_k = \boldsymbol{g}_k^{(C)} - \boldsymbol{\mu}_m$ in (53) follows Gaussian distribution. At a given SNR, decreasing the sampling spacing raises the probability that $|\boldsymbol{\xi}_k| > |\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}|/2$. This leads to a reduction in the probability of correctly estimating the nearest grid and an increase in the expected error between the estimated and nearest grid points, i.e. $\mathbb{E}\{\|\boldsymbol{z}_m - \boldsymbol{z}_{m'}\|\}$. Furthermore, the effectiveness of sparse recovery methods is constrained by the restricted isometry property (RIP) or the mutual coherence condition. Reducing the sampling spacing may breach this condition, leading to a decline in localization accuracy [39].

3) Number of RIS Elements: Similarly to [40], we characterize the average received power with respect to the number of RIS elements as $N \to \infty$. We consider two different RIS phase shifts: random phase shifts and optimal phase shifts that maximize the SNR. We consider $C = 1$ as an example.

Proposition 3: Assume $\boldsymbol{h}_k^A \sim \mathcal{CN}(\boldsymbol{0}, \rho_A^2 \boldsymbol{I})$ and $\boldsymbol{h}_k^t \sim \mathcal{CN}(\boldsymbol{0}, \rho_t^2 \boldsymbol{I})$, the average received power holds that

$$\lim_{N \to \infty} P = \begin{cases} N\rho_A^2\rho_t^2, & \boldsymbol{\beta} = [1, ..., 1]^T, \\ N^2 \dfrac{\pi^2 \rho_A^2 \rho_t^2}{16}, & \text{max SNR phase shifts.} \end{cases} \tag{56}$$

Proof: The two cases are discussed as follows:

- The channel $h = (\boldsymbol{h}_k^A)^H \text{diag}\{\boldsymbol{\beta}\} \boldsymbol{h}_k^t = (\boldsymbol{h}_k^A)^H \boldsymbol{h}_k^t$. According to Lindeberg-Lévy central limit theorem, we have $h \sim \mathcal{CN}(\boldsymbol{0}, N\rho_A^2\rho_t^2)$ as $N \to \infty$. The average user received power is given by
$$\lim_{N \to \infty} P = \mathbb{E}|h|^2 = N\rho_A^2\rho_t^2. \tag{57}$$
- For RIS phase shifts of max SNR, we have $|h| = (\boldsymbol{h}_k^A)^H \text{diag}\{\boldsymbol{\beta}\} \boldsymbol{h}_k^t = \sum_{n=1}^{N} |h_{k,n}^A||h_{k,n}^t|$, where $h_{k,n}^A$ and $h_{k,n}^t$ are the n-th element in $\boldsymbol{h}_k^A$ and $\boldsymbol{h}_k^t$, respectively. Since $h_{k,n}^A$ and $h_{k,n}^t$ are statistically independent and follow Rayleigh distribution with mean values $\sqrt{\pi}\rho_A/2$, $\sqrt{\pi}\rho_t/2$, we have $\mathbb{E}(|h_{k,n}^A||h_{k,n}^t|) = \pi\rho_A\rho_t/4$. The average user received power is given by
$$\lim_{N \to \infty} P = \lim_{N \to \infty} \left| \sum_{n=1}^{N} |h_{k,n}^A||h_{k,n}^t| \right|^2 = N^2 \frac{\pi^2 \rho_A^2 \rho_t^2}{16}. \tag{58}$$

■

Note that the power gain of order N can be achieved for fixed phase shifts, which reveals the inherent aperture gain of a larger RIS by collecting more signal power. Moreover, setting the RIS phase shifts for max SNR can also achieve another beamforming power gain of order N simultaneously.

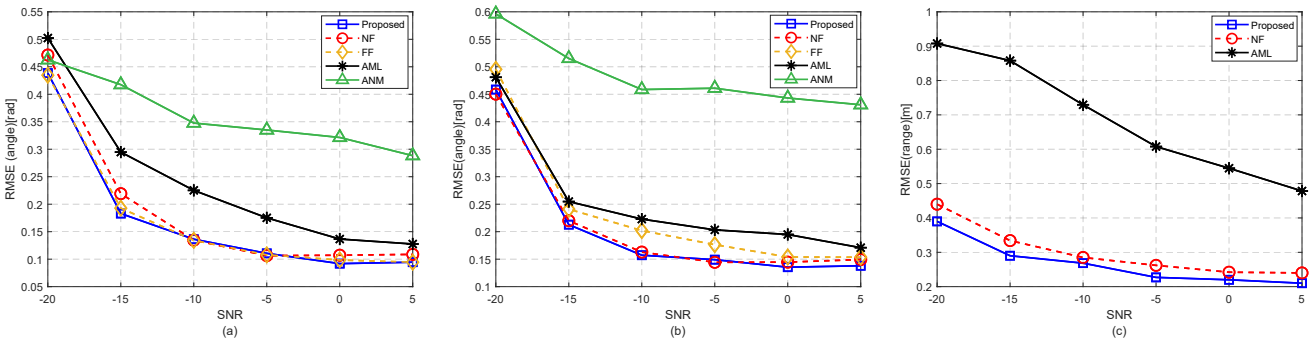

Fig. 4. (a) Angle estimation accuracy of the proposed algorithm compared with AML and ANM when the user is located in the FF region. (b) Angle estimation accuracy of the proposed algorithm compared with AML and ANM when the user is located in the NF region. (c) Range estimation accuracy of the proposed algorithm compared with AML when the user is located in the NF region.

Hence, with an increasing number of RIS elements, the SNR improves, which enlarges the difference between $\boldsymbol{\mu}_m$ and $\boldsymbol{\mu}_{m'}$. Consequently, the probability of $|\boldsymbol{\xi}_k| > |\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m'}|/2$ decreases, which lowers the likelihood of grid misjudgment. However, the impact of grid mismatch persists.

Remark 2: As the number of RIS elements increases, the localization accuracy first improves and then remains fixed.

VI. Simulation Results

In this section, we present simulation results demonstrating the performance of the proposed method. The RIS is placed on the plane $x = 0$, and its center is at the origin $(0, 0, 0)$m. We use an RIS with 10×10 elements and the element spacing is $\lambda/2 = 0.03$m. The localization range is 10m in radius [41]. The NF range for the considered RIS is approximately 4.9m, while the rest region is FF. The users are equally distributed in NF and FF regions. We consider a multipath environment with 3 randomly distributed scatters. The azimuth and elevation angles are both uniformly sampled with spacing $\pi/10$, and the range is sampled with spacing 0.25m. The center transmit frequency is 5GHz [42]. The number of cycles is set as 20. We conduct $T = 1000$ independent trials to obtain the average result. The simulation parameters are listed in Table II unless otherwise stated. The proposed method is compared with following methods.

- Approximate maximum likelihood (AML) scheme [22]: a RIS-enabled NF localization method based on maximum likelihood. This method decouples the angle and range parameters and iteratively searches the angle and range to fit best with the received signal.
- Atomic norm minimization (ANM) scheme [43]: an off-grid direction of arrival estimation method based on atomic norm that targets FF users localization.
- NF scheme: The proposed algorithm is combined with pure NF model to show the gain brought by the hybrid model, labeled as NF.
- FF scheme: The proposed algorithm is combined with pure FF model, labeled as FF.

For all compared algorithms, we employ the same system setup and localization protocol. For NF methods, including AML and the proposed method with the NF model, the whole localization range is treated as in the NF, i.e., the range of FF is also sampled, while for FF methods, we only sample the angle.

TABLE II
Simulation Parameters

Parameters	Values
Number of users K	2
Number of scatters L	3
Transmit center frequency f_c	5GHz
RIS center location	$(0, 0, 0)$m
Number of RIS element N	10×10
RIS element spacing	0.005m
Localization range	10m
Angle sampled spacing $\Delta\theta, \Delta\phi$	$\pi/10$ rad
Range sampled spacing ΔR i	0.25m
Number of cycles C	20

A. Performance Evaluation

Fig. 4 shows the RMSE of angle or range estimation versus the signal-to-noise ratio (SNR) for NF or FF users. The SNR is defined as P_s/σ^2, where $P_s = ((\boldsymbol{h}_k^A)^T \boldsymbol{I} \boldsymbol{h}_{k0}^t)^2$ is the received signal power of the direct path when the phase shifts of all the RIS elements are 1. Fig. 4 (a) and (b) show the angle estimation RMSE of the proposed algorithm and other comparison algorithms for FF and NF users, respectively. They indicate that the proposed method outperforms all other methods for FF and NF user localization. We also observe that for algorithms other than ANM, as SNR increases, the RMSE of the angle estimation first decreases and then remains fixed. The RMSE does not decrease when the SNR is very high because the estimated location must be at one of the sampled grids, while the actual location of the users is off-grid. The ANM does not suffer from the grid mismatch problem since it is an off-grid algorithm.

Fig. 4 (a) also reveals the RMSEs of the proposed algorithm with the hybrid model and FF model are similar, indicating that the hybrid model can achieve the same accuracy for FF user localization although with the interference of NF candidate locations. We also observe that the proposed algorithm performs worse in the NF model than in the hybrid model

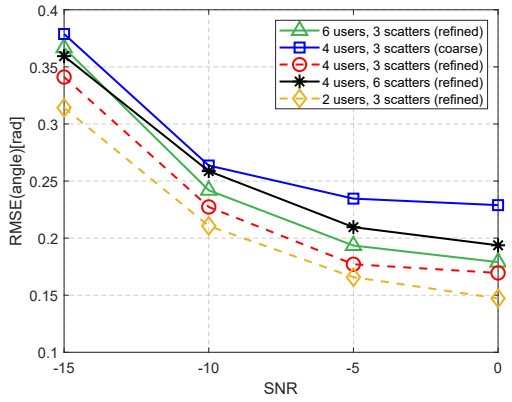

Fig. 5. Angle estimation performance for different localization scenarios.

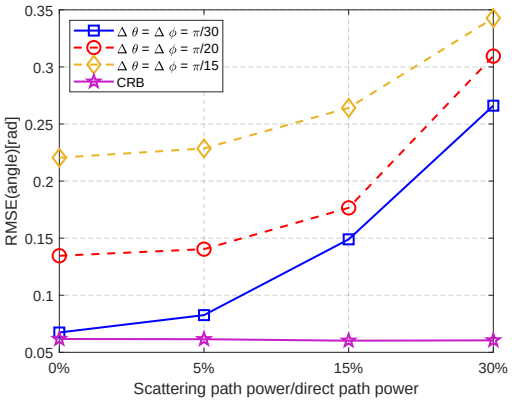

Fig. 6. Angle estimation performance for different interference power.

for FF users. This is because, for the NF model, the range of the FF region is also sampled, resulting in more candidate locations in the FF region and making the algorithm more difficult to converge. At the same time, the channels of the sampled FF candidate locations in the NF model have strong correlations, which leads to energy dispersion especially when estimating the scatters locations. In Fig. 4 (b), we observe that the FF model has a higher RMSE than the hybrid model when estimating the NF angles. This is due to the model mismatch of the FF model when describing the NF channels.

Fig. 4 (c) shows the range estimation accuracy for NF users versus SNR. For range estimation, the proposed method outperforms AML by more than 55% for all SNR conditions. It shows that the range estimation result of the NF model is not as good as those of the hybrid model for NF user localization, which reveals the superiority of the hybrid model. Additionally, the improvement of the hybrid model over the NF model is bigger in range compared to angle estimation. The reason for this phenomenon is that the NF model samples the range and angle in the whole region, i.e., the range of FF is also sampled. However, the steering vectors of the different ranges in the FF are very similar. If a NF user is misestimated as a FF point, the resulting range error becomes significantly larger, leading to an overall degradation in estimation accuracy.

In Fig. 4, we can observe that the proposed method significantly outperforms both AML and ANM. The performance gain can be attributed to the following reasons: Firstly, neither AML nor ANM effectively handles the scattering paths, and the interference of scattering paths reduces localization accuracy. Secondly, AML decouples angle and range estimation and updates them iteratively. While this approximated method reduces algorithm complexity, it also impacts performance. Moreover, ANM is a FF localization method and employs plane wave model, which inaccurately describes NF signals.

In the following, the angle sampling spacing is set as $\pi/20$ to reduce grid mismatch. Fig. 5 shows the influence of the number of users and scatters. We can observe that increasing the number of users degrades angle estimation accuracy, as the RIS beamforming becomes less precise when simultaneously serving multiple users. Moreover, increasing the number of scatters reduces localization accuracy due to interference from multipath propagation. Furthermore, by comparing the results obtained directly using the OMP algorithm, i.e., the first step of the localization algorithm (labeled as coarse) with those obtained through the three-step localization algorithm (labeled as refined), it is evident that the scattering path elimination technique can improve the localization accuracy.

Fig. 6 shows the angle estimation performance versus the scattering path power. The CRB is calculated given true user and scatter locations. It is observed that for low scattering path power and dense angle sampling, the estimation RMSE can approach the CRB. However, the effectiveness of scattering path elimination becomes limited for high scattering path interference, causing the RMSE to deviate from the CRB. Additionally, sparse angle sampling exacerbates the grid mismatch problem, further leading to accuracy degradation.

B. Localization Accuracy and Complexity Trade-off

This subsection examines the effects of hyperparameters on localization accuracy and algorithm complexity. We focus on the sampling spacing, the number of cycles, and the number of RIS elements, as they are the most important factors for both according to the analysis in Sec. V.

1) Number of Cycles: Fig. 7 (a) illustrates the angle estimation RMSE for different numbers of cycles. It shows that for the hybrid and FF model, the estimation performance converges after $C = 30$ cycles, while for the NF model, the result remains fixed after $C = 35$ cycles, proving that the NF model is harder to converge. Fig. 7 (b) reveals the CPU time versus number of cycles for three models. The results show the proposed algorithm has the lowest complexity with FF model, followed by the hybrid model, and the highest complexity with NF model. This is because the FF model has the fewest candidate locations, while the NF model has the most.

2) Sampling Spacing: Fig. 8 (a) shows the average localization RMSE for NF users of the proposed algorithm under different sampling spacing. The localization RMSE is derived by computing user positions from estimated angles and ranges, then comparing them to the ground truth. Fig. 8 (b) shows the CPU time of the proposed algorithm in three different models. It can be observed that as the sampling spacing reduces, the CPU time increases rapidly, while the estimation error first decreases and then remains fixed, which is consistent with

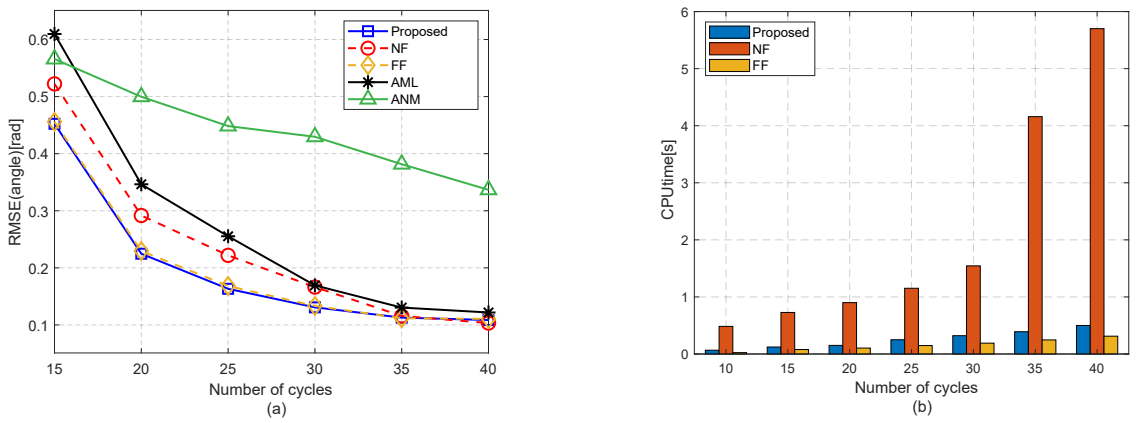

Fig. 7. (a) Angle estimation RMSE versus the number of cycles. (b) CPU time of the proposed algorithm versus number of cycles.[2]

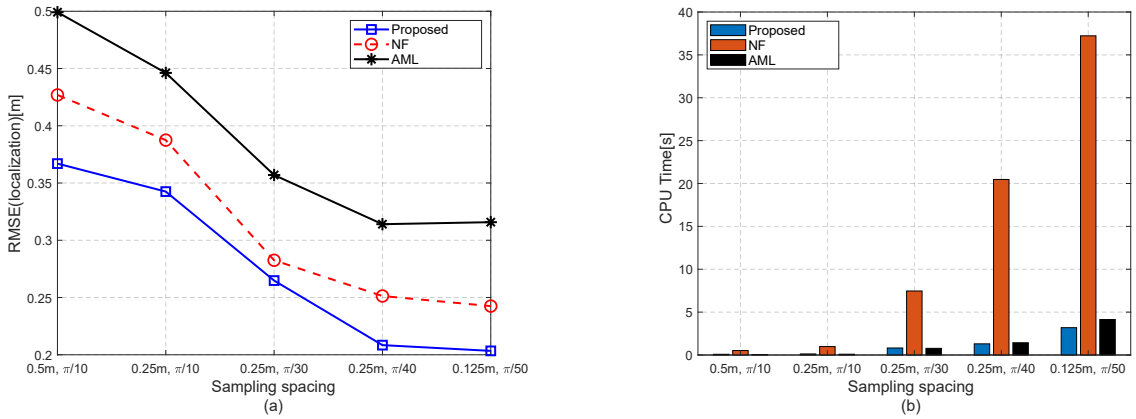

Fig. 8. (a) Localization of the NF user for different sampling spacing. (b) CPU time versus sampling spacing.

previous analysis. Note that when the spacing is less than $(0.25\text{m}, \pi/40 \text{ rad})$, the accuracy does not improve, but the complexity still increases.

3) RIS Size: We set the number of cycles as $C = 50$ to ensure convergence. Fig. 9 (a) and Fig. 9 (b) show the range estimation RMSE and the CPU time of the proposed algorithm versus the number of RIS elements. As the number of RIS elements increases, the range estimation RMSE decreases, and CPU time increases. This is because the complexities of both the localization and optimization algorithms are positively correlated with the RIS element number. Also, a centimeter-level accuracy is attainable for a RIS with 900 elements.

VII. Conclusion

In this paper, a RIS-enabled multi-user hybrid NF and FF localization method has been developed. We considered a multi-user hybrid NF and FF localization scenario and introduced a localization protocol. A localization algorithm is designed for location estimation. The CRB of the estimated location parameters was analyzed and a CCM-based optimization method is proposed to optimize RIS phase shifts. We analyzed the complexity and the localization performance of the proposed method. Simulation results have shown that: 1) the proposed method outperforms the compared algorithm and can achieve sub-meter accuracy; 2) the hybrid NF and FF model has the advantages of high accuracy and low algorithm complexity over pure NF and FF models; 3) by adjusting the hyperparameters, the trade-off between localization accuracy and complexity can be achieved.

[2]Under the specified parameters at SNR = 5 dB, our method achieves an angle estimation error of 7° and a range error of 0.2 m, with a computational latency of 0.15 s per update. This performance enables applications such as indoor navigation [44] and asset tracking [45].

Appendix A
Proof of Proposition 1

As mentioned in Section III-A, all previously received signals are used to estimate the locations, so we need to consider received signals in all previous c cycles, as modeled in (31), to obtain the expressions of CRBs. We first derive the expression of the Fisher information matrix (FIM) related to the unknown location parameters according to the received signals, denoted by $\boldsymbol{J}_{NF}$. The CRB, which gives the localization performance limit, is then derived from the FIM.

As proved in [33], FIM is given by [12]

$$[\boldsymbol{J}_{NF}]_{i,j} = E\left\{-\frac{\partial^2 \ln f(\boldsymbol{y}_k; \boldsymbol{p}_k, \boldsymbol{Q})}{\partial p_i \partial p_j}\right\}, \tag{59}$$

where $f(\boldsymbol{y}_k; \boldsymbol{p}_k, \boldsymbol{Q})$ is the likelihood function of the random vector $\boldsymbol{y}_k$ conditioned on $\boldsymbol{p}_k$ and $\boldsymbol{Q}$. Here we assume the transmitted signals are deterministic. Therefore the i-th received

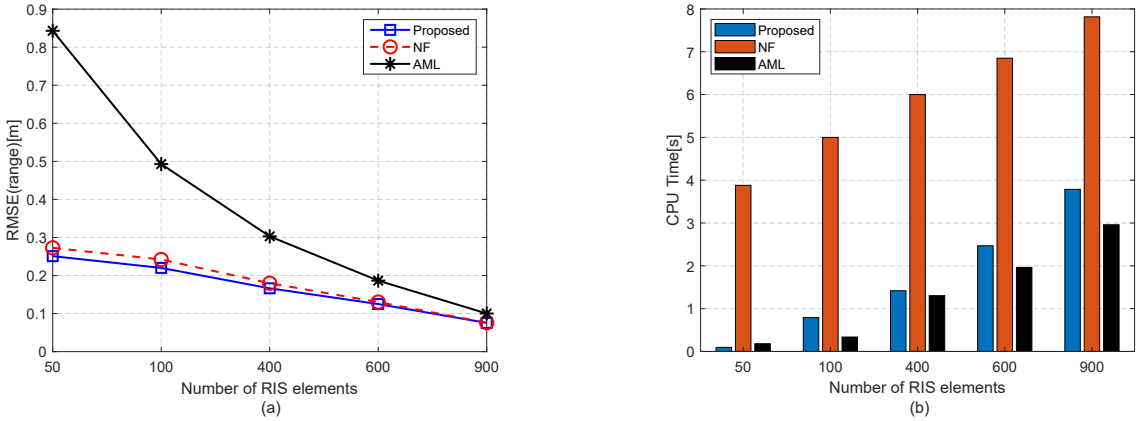

Fig. 9. (a) Range estimation performance of the NF users versus RIS size. (b) CPU time of the proposed algorithm for different RIS sizes.

signal $y_k^{(i)}$ can be assumed to be independent of each other and follow complex Gaussian distribution, i.e., $f(y_k^{(i)}; \boldsymbol{p}_k, \boldsymbol{Q}) \sim \mathcal{CN}(\mu^{(i)}, \sigma^2)$, where $\mu^{(i)} = (\boldsymbol{\beta}^{(i)})^T \boldsymbol{h}_k s$ is the noise-free received signal. Therefore, we have $f(\boldsymbol{y}_k) = \prod_{i=1}^{c+1} f(y_k^{(i)})$. Then, in the NF region, the FIM $\boldsymbol{J}_{NF}$ can be rewritten as

$$[\boldsymbol{J}_{NF}]_{i,j} = \frac{2}{\sigma^2} \sum_{m=1}^{c+1} Re\left\{ \frac{\partial (\mu^{(m)})^H}{\partial p_i} \frac{\partial \mu^{(m)}}{\partial p_j} \right\}, \tag{60}$$

where $\frac{\partial \mu^{(m)}}{\partial p_i} = (\boldsymbol{\beta}^{(m)})^T \frac{\partial \boldsymbol{h}_k^{NF}}{\partial p_i} s$. The n-th entry of $\frac{\partial \boldsymbol{h}_k^{NF}}{\partial p_i}$ can be expressed as

$$\left[\frac{\partial \boldsymbol{h}_k^{NF}}{\partial p_i} \right]_n = \alpha_k \left(-j \frac{2\pi}{\lambda} \right) \exp\left(-j \frac{2\pi}{\lambda} d_{kn}^t \right) \frac{\partial d_{kn}^t}{\partial p_i}. \tag{61}$$

Next, we take the inverse of $\boldsymbol{J}_{NF}$ and obtain the CRBs. Hence, the proof of proposition 1 is accomplished.

Appendix B
Proof of Proposition 2

Similar to the proof of proposition 2, the FIM $\boldsymbol{J}_{FF}$ in the FF region can be written as [12]

$$[\boldsymbol{J}_{FF}]_{i,j} = \frac{2}{\sigma^2} \sum_{i=1}^{c+1} Re\left\{ \frac{\partial (\mu^{(i)})^H}{\partial p_i} \frac{\partial \mu^{(i)}}{\partial p_j} \right\}, \tag{62}$$

where $\frac{\partial \mu^{(i)}}{\partial p_i} = (\boldsymbol{\beta}^{(i)})^T \frac{\partial \boldsymbol{h}_k^{FF}}{\partial p_i} s$. In the FF region, we have

$$\left[\frac{\partial \boldsymbol{h}_k^{FF}}{\partial \theta_k} \right]_n = h_{k0n}^t \left(-j \frac{2\pi}{\lambda} \right) (-y_n \cos\theta_k \sin\phi_k + z_n \sin\theta_k), \tag{63}$$

$$\left[\frac{\partial \boldsymbol{h}_k^{FF}}{\partial \phi_k} \right]_n = h_{k0n}^t \left(-j \frac{2\pi}{\lambda} \right) (-y_n \sin\theta_k \cos\phi_k), \tag{64}$$

where h_{k0n}^t is the n-th element in the direct user-RIS path.

Next, we take the inverse of $\boldsymbol{J}_{FF}$ and obtain the CRBs. Hence, the proof of proposition 2 is accomplished.

Appendix C
The Differentials of CRB

We denote the element in the i-th row, j-th column of $\boldsymbol{J}$ as J_{ij}. In order to find the differentials of CRBs, we need to find the differentials of J_{ij}. For simplicity, here we use $\boldsymbol{\beta}$ to substitute $\boldsymbol{\beta}^{(k+1)}$. J_{ij} can be considered as a function of two independent complexed-valued variables $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$. The complex differentials of J_{ij} can be expressed as

$$dJ_{ij} = \left(\frac{\partial J_{ij}}{\partial \boldsymbol{\beta}} \right) d\boldsymbol{\beta} + \left(\frac{\partial J_{ij}}{\partial \boldsymbol{\beta}^*} \right) d\boldsymbol{\beta}^*. \tag{65}$$

The complex differentials of J_{ij} are given as

$$\begin{aligned} dJ_{ij} = & \frac{1}{\sigma^2} \left[d\left(s^H \left(\frac{\partial \boldsymbol{h}}{\partial p_i} \right)^H \boldsymbol{\beta}^* \boldsymbol{\beta}^T \frac{\partial \boldsymbol{h}}{\partial p_j} s \right) \right. \\ & \left. + d\left(s^T \left(\frac{\partial \boldsymbol{h}}{\partial p_i} \right)^T \boldsymbol{\beta} \boldsymbol{\beta}^H \left(\frac{\partial \boldsymbol{h}}{\partial p_j} \right)^* s^* \right) \right]. \end{aligned} \tag{66}$$

By substitude $\frac{\partial \boldsymbol{\beta}^* \boldsymbol{\beta}^T}{\partial \boldsymbol{\beta}_n^*}$ and $\frac{\partial \boldsymbol{\beta} \boldsymbol{\beta}^H}{\partial \boldsymbol{\beta}_n^*}$ into (66), we can obtain

$$\begin{aligned} \frac{\partial J_{ij}}{\partial \beta_n^*} = & \frac{|s|^2}{\sigma^2} \left[\left(\frac{\partial h_n}{\partial p_i} \right)^* \sum_{m=1}^{N} \frac{\partial h_m}{\partial p_j} \beta_m + \right. \\ & \left. \left(\frac{\partial h_n}{\partial p_j} \right)^* \sum_{m=1}^{N} \frac{\partial h_m}{\partial p_i} \beta_m \right]. \end{aligned} \tag{67}$$

The derivatives of CRBs with respect to $\boldsymbol{\beta}^*$ are given by

$$\frac{\partial D_i}{\partial \boldsymbol{\beta}^*} = \left(\frac{\partial \boldsymbol{J}^{-1}}{\partial \boldsymbol{\beta}^*} \right)_{i,i} = \left(-\boldsymbol{J}^{-1} \frac{\partial \boldsymbol{J}}{\partial \boldsymbol{\beta}^*} \boldsymbol{J}^{-1} \right)_{i,i}. \tag{68}$$

References

[1] M. Cao, H. Zhang, B. Di, and H. Zhang, "Unified near-field and far-field localization with holographic MIMO," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2024, pp. 1–6.

[2] S. E. Trevlakis, A.-A. A. Boulogeorgos, D. Pliatsios, J. Querol, K. Ntontin, P. Sarigiannidis, S. Chatzinotas, and M. Di Renzo, "Localization as a key enabler of 6G wireless systems: A comprehensive survey and an outlook," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 2733–2801, 2023.

[3] J. An, C. Xu, D. W. K. Ng, G. C. Alexandropoulos, C. Huang, C. Yuen, and L. Hanzo, "Stacked intelligent metasurfaces for efficient holographic MIMO communications in 6G," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2380–2396, 2023.

[4] R. Deng, B. Di, H. Zhang, H. V. Poor, and L. Song, "Holographic MIMO for LEO satellite communications aided by reconfigurable holographic surfaces," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 3071–3085, 2022.

[5] D.-R. Emenonye, H. S. Dhillon, and R. M. Buehrer, "Fundamentals of RIS-aided localization in the far-field," *IEEE Trans. Wirel. Commun.*, vol. 23, no. 4, pp. 3408–3424, 2024.

[6] P. Zheng, H. Chen, T. Ballal, H. Wymeersch, and T. Y. Al-Naffouri, "Misspecified Cramér-Rao bound of RIS-aided localization under geometry mismatch," in *Proc. IEEE Int. Conf. Acoust. Speech, and Signal Process. (ICASSP)*, Rhodes Island, Greece, Jun. 2023, pp. 1–5.

[7] J. He, H. Wymeersch, L. Kong, O. Silvén, and M. Juntti, "Large intelligent surface for positioning in millimeter wave MIMO systems," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Antwerp, Belgium, May 2020, pp. 1–5.

[8] H. Chen, Y. Bai, Q. Wang, H. Chen, L. Tang, and P. Han, "DOA estimation assisted by reconfigurable intelligent surfaces," *IEEE Sens. J.*, vol. 23, no. 12, pp. 13 433–13 442, May. 2023.

[9] G. Mylonopoulos, L. Venturino, S. Buzzi, and C. D'Andrea, "Maximum-likelihood user localization in active-RIS empowered mmwave wireless networks," in *Proc. Eur. Conf. Antennas Propag. (EuCAP)*, Florence, Italy, Mar. 2023, pp. 1–5.

[10] K. Chen, "Study on indoor positioning of eliminating NLOS signals through dynamically adjusting reflected signals by RIS," in *Proc. Int. Conf. Computer and Commun. Syst. (ICCCS)*, Xi'an, China, Apr. 2024, pp. 489–493.

[11] K. Keykhosravi, G. Seco-Granados, G. C. Alexandropoulos, and H. Wymeersch, "RIS-enabled self-localization: Leveraging controllable reflections with zero access points," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Seoul, Korea, May 2022, pp. 2852–2857.

[12] 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 *Proc. Int. Symp. Wirel. Commun. Syst. (ISWCS)*, Hangzhou, China, Oct. 2022.

[13] P. Zheng, H. Chen, H. Wymeersch, and T. Y. Al-Naffouri, "Near field sidelink positioning through a single active RIS," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 2511–2516.

[14] 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 *Proc. IEEE Int. Conf. Commun. (ICC)*, Montreal, Canada, Jun. 2021, pp. 1–6.

[15] S. Huang, B. Wang, Y. Zhao, and M. Luan, "Near-field RSS-based localization algorithms using reconfigurable intelligent surface," *IEEE Sens. J.*, vol. 22, no. 4, pp. 3493–3505, 2022.

[16] O. Rinch, A. Elzanaty, and A. Alsharoa, "Enhancing near-field wireless localization with liDAR-assisted RIS in multipath environments," *IEEE Wirel. Commun. Lett.*, vol. 12, no. 12, pp. 2168–2172, 2023.

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

[18] H. Lu and Y. Zeng, "Communicating with extremely large-scale array/surface: Unified modeling and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4039–4053, Jun. 2022.

[19] H. D. Mafukidze, A. K. Mishra, J. Pidanic, and S. W. P. Francois, "Scattering centers to point clouds: A review of mmwave radars for non-radar-engineers," *IEEE Access*, vol. 10, pp. 110 992–111 021, 2022.

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

[21] X. Wei and L. Dai, "Channel estimation for extremely large-scale massive MIMO: Far-field, near-field, or hybrid-field?" *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 177–181, 2022.

[22] C. Ozturk, M. F. Keskin, H. Wymeersch, and S. Gezici, "RIS-aided near-field localization under phase-dependent amplitude variations," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 8, pp. 5550 – 5566, Apr. 2023.

[23] H. Zhang, H. Zhang, B. Di, M. D. Renzo, Z. Han, H. V. Poor, and L. Song, "Holographic integrated sensing and communication," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2114–2130, Mar. 2022.

[24] R. Deng, B. Di, H. Zhang, Y. Tan, and L. Song, "Reconfigurable holographic surface-enabled multi-user wireless communications: Amplitude-controlled holographic beamforming," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 8, pp. 6003–6017, 2022.

[25] X. Wei, L. Dai, Y. Zhao, G. Yu, and X. Duan, "Codebook design and beam training for extremely large-scale RIS: Far-field or near-field?" *China Commun.*, vol. 19, no. 6, pp. 193–204, Jun. 2022.

[26] Y. Han, S. Jin, C.-K. Wen, and X. Ma, "Channel estimation for extremely large-scale massive MIMO systems," *IEEE Wirel. Commun. Lett.*, vol. 9, no. 5, pp. 633–637, May. 2020.

[27] J. Tropp, A. Gilbert, and M. Strauss, "Simultaneous sparse approximation via greedy pursuit," in *Proc. IEEE Int. Conf. Acoust. Speech, and Signal Process. (ICASSP)*, vol. 5, Philadelphia, USA, Mar. 2005, pp. v/721–v/724 Vol. 5.

[28] M. Cui and L. Dai, "Channel estimation for extremely large-Scale MIMO: Far-field or near-field?" *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, Jan. 2022.

[29] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, Apr. 2016.

[30] Z. Wei, H. Xin, H. Jun, and Z. Nie, "Scattering centers extraction based on compressed sensing and multilevel fast multipole algorithm acceleration," in *Proc. IEEE Int. Conf. Comput. Electromagnetics (ICCEM)*, Chengdu, China, Mar. 2018, pp. 1–2.

[31] W. Lyu, S. Yang, Y. Xiu, Y. Li, H. He, C. Yuen, and Z. Zhang, "CRB minimization for RIS-aided mmwave integrated sensing and communications," *IEEE Internet Things J.*, vol. 11, no. 10, pp. 18 381–18 393, 2024.

[32] M. Zhu, L. Li, S. Xia, and T.-H. Chang, "Information and sensing beamforming optimization for multi-user multi-target MIMO ISAC systems," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, Jun. 2023, pp. 1–5.

[33] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1993.

[34] T. Wimalajeewa and P. K. Varshney, "OMP based joint sparsity pattern recovery under communication constraints," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 5059–5072, 2014.

[35] W. Zhang and T. Kim, "Successful recovery performance guarantees of SOMP under the ℓ_2-norm of noise," *IEEE Trans. Veh. Technol.*, vol. 73, no. 2, pp. 2156–2170, 2024.

[36] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 8, pp. 5218–5233, 2020.

[37] A. Xenaki and P. Gerstoft, "Grid-free compressive beamforming," *J. Acoustical Soc. of America*, vol. 137, no. 4, pp. 1923–1935, 2015.

[38] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, "Metalocalization: Reconfigurable intelligent surface aided multi-user wireless indoor localization," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 12, pp. 7743–7757, Jun. 2021.

[39] W. Wang and R. Wu, "High resolution direction of arrival (DOA) estimation based on improved orthogonal matching pursuit (OMP) algorithm by iterative local searching," *Sensors*, vol. 13, no. 9, pp. 11 167–11 183, 2013.

[40] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 11, pp. 5394–5409, 2019.

[41] R. A. Aguiar, N. Paulino, and L. M. Pessoa, "Enhancing NLoS RIS-aided localization with optimization and machine learning," in *2023 IEEE Globecom Workshops (GC Wkshps)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 1898–1903.

[42] T. Ma, Y. Xiao, X. Lei, W. Xiong, and M. Xiao, "Distributed reconfigurable intelligent surfaces assisted indoor positioning," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 1, pp. 47–58, 2023.

[43] P. Chen, Z. Yang, Z. Chen, and Z. Guo, "Reconfigurable intelligent surface aided sparse DOA estimation method with non-ULA," *IEEE Signal Process. Lett.*, vol. 28, pp. 2023–2027, 2021.

[44] S. Xu, R. Chen, G. Guo, Z. Li, L. Qian, F. Ye, Z. Liu, and L. Huang, "Bluetooth, floor-plan, and microelectromechanical systems-assisted wide-area audio indoor localization system: Apply to smartphones," *IEEE Trans. Ind. Electron.*, vol. 69, no. 11, pp. 11 744–11 754, 2022.

[45] Z. Luo, W. Li, Y. Wu, H. Dong, L. Bian, and W. Wang, "Accurate indoor localization for bluetooth low energy backscatter," *IEEE Internet Things J.*, vol. 12, no. 2, pp. 1805–1816, 2025.