

Analysis and Algorithm for Multi-IRS Collaborative Localization via Hybrid Time–Angle Estimation

Ziheng Zhang, Wen Chen, Qingqing Wu, Haoran Qin, Zhendong Li, and Qiong Wu

Abstract—This paper proposes a novel multiple intelligent reflecting surfaces (IRSs) collaborative hybrid localization system, which involves deploying multiple IRSs near the target area and achieving target localization through joint time delay and angle estimation. Specifically, echo signals from all reflective elements are received by each sensor and jointly processed to estimate the time delay and angle parameters. Based on the above model, we derive the Fisher Information Matrix (FIM) for cascaded delay, Angle of Arrival (AOA), and Angle of Departure (AOD) estimation in semi-passive passive models, along with the corresponding Cramér-Rao Bound (CRB). To achieve precise estimation close to the CRB, we design efficient algorithms for angle and location estimation. For angle estimation, reflective signals are categorized into three cases based on their rank, with different signal preprocessing. By constructing an atomic norm set and minimizing the atomic norm, the joint angle estimation problem is transformed into a convex optimization problem, and low-complexity estimation of multiple AOA and AOD pairs is achieved using the Alternating Direction Method of Multipliers (ADMM). For location estimation, we propose a three-stage localization algorithm that combines weighted least squares, total least squares, and quadratic correction to handle errors in the coefficient matrix and observation vector, thus improving accuracy. Numerical simulations validate the superiority of the proposed system, demonstrating that the system's collaboration, hybrid localization, and distributed deployment provide substantial benefits, as well as the accuracy of the proposed estimation algorithms, particularly in low signal-to-noise ratio (SNR) condition.

Index Terms—Intelligent reflecting surface, collaborative localization, hybrid localization, Cramér-Rao bound.

I. INTRODUCTION

The integration of sensing technologies into 6G networks is expected to be a defining characteristic, enhancing the capabilities of wireless communication systems in unprecedented ways [1]. This integration primarily involves two key aspects: one is the emergence of environment-aware applications, such as autonomous vehicles [2], drone delivery systems [3], and advanced satellite communications [4], and the other is dual role of communication signals for both data transmission and sensing, allowing for target localization and environmental imaging [5]. Additionally, environmental data

gathered through sensing can improve channel estimation and reduce the need for pilot signals, enhancing overall network efficiency [6]. However, delivering high-precision sensing capabilities across diverse and complex electromagnetic environments presents several challenges. Firstly, high-frequency sensing signals, required for precise localization, suffer from significant propagation loss due to free-space path loss and atmospheric attenuation, limiting their effective range [7]. Secondly, achieving reliable line-of-sight (LoS) links between base stations (BS) and targets is particularly difficult in urban environments, where obstacles obstruct signal paths. Finally, the computational demands for processing large-scale real-time data, such as radar imaging or MIMO measurements, pose significant challenges in terms of both computational power and energy consumption. Overcoming these challenges is essential for improving the performance of sensing applications in 6G networks.

Intelligent Reflecting Surface (IRS) has emerged as a transformative technology in the development of 6G wireless networks, offering the ability to manipulate the electromagnetic environment in a cost-effective and energy-efficient manner. By dynamically adjusting the phase shifts of reflected signals, IRS enhances communication performance, mitigates signal blockage, and improves spectral efficiency. This capability extends beyond conventional communication systems and has led to growing interest in IRS-assisted wireless sensing, where it facilitates the creation of additional LoS links and enhances environmental sensing [8]. In the context of 6G, which is characterized by its focus on ubiquitous intelligence and the integration of sensing and communication (ISAC), IRS plays a crucial role in advancing sensing capabilities. Specifically, IRS can significantly improve target detection, localization, and environmental mapping by optimizing signal reflection paths, particularly in non-line-of-sight (NLoS) environments [9]. Recent research has proposed several IRS-assisted sensing architectures, including fully passive IRS sensing [10], semi-passive IRS sensing [10], and IRS-based self-sensing systems [11]. These architectures differ in terms of the control and feedback mechanisms employed.

Under these IRS-assisted sensing architectures—including fully passive, semi-passive, and IRS-based self-sensing systems—that effectively exploit IRS's ability to dynamically reconfigure electromagnetic environments and establish additional LoS links [12]–[14], recent research has developed innovative approaches for wireless localization, target tracking, and detection applications [15]–[24]. In wireless localization, extensive studies have shown that IRS can overcome inherent limitations of conventional locationing methods in complex

Z. Zhang, W. Chen, Q. Wu and H. Qin are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: zhangziheng@sjtu.edu.cn; wenchen@sjtu.edu.cn; qingqingwu@sjtu.edu.cn; haoranqin@sjtu.edu.cn).

Z. Li is with the School of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an 710049, China (email: lizhendong@xjtu.edu.cn).

Q. Wu is with the School of Internet of Things Engineering, Jiangnan University, Wuxi, China (email: qiongwu@jiangnan.edu.cn).

(Corresponding author: Wen Chen.)

environments [15], [16], [25]. Specifically, IRS establishes virtual line-of-sight (VLoS) links by manipulating reflected signal paths, thus compensating for direct-path blockage and multipath interference [15], [26]. This capability not only mitigates challenges posed by obstacles but also strengthens signal availability, particularly beneficial in massive MIMO systems [16], [17]. Consequently, IRS-aided localization exhibits improved accuracy and stability compared to traditional schemes, [15]–[17]. Similarly, IRS demonstrates substantial potential in target tracking, addressing typical accuracy deterioration arising from multipath fading and signal attenuation in dynamic multi-target scenarios [18]–[21]. By adaptively optimizing reflection phases and beam patterns, IRS provides stable and reliable tracking paths, reducing tracking error and ensuring continuous signal availability [18], [19], [21]. Such adaptability enables effective real-time optimization of propagation paths, significantly improving multi-target tracking accuracy under severe NLoS conditions [20], [21]. Furthermore, IRS has enabled new capabilities in target detection, especially where traditional systems suffer from limited coverage and insufficient sensitivity [22]–[24]. By adaptively controlling reflected signals, IRS enhances echo signals from targets, thereby improving detection sensitivity in cluttered and obstructed environments [23]. This functionality is particularly advantageous for real-time monitoring in scenarios such as intelligent transportation and autonomous driving, as IRS effectively compensates for direct-path losses and substantially boosts system detection capabilities [22]–[24]. Looking forward, further research is anticipated to focus on collaborative multi-IRS strategies, intelligent beam management schemes, and comprehensive optimization in more complex scenarios, thereby achieving superior sensing performance and precision in future wireless sensing systems.

First, while IRS technology has shown great promise, the study of multi-IRS collaboration remains insufficient. Although prior works have explored the benefits of multi-IRS systems in extending coverage, they primarily focus on non-cooperative schemes where each IRS serves a separate zone [27], or employ time-division protocols to avoid inter-IRS interference [28], [29]. A framework enabling fully collaborative localization at the physical layer, where multiple semi-passive IRSs jointly process all signals, including inter-IRS reflection paths, remains to be thoroughly investigated. Secondly, there is a noticeable gap regarding the joint use of sensing information. Existing IRS-assisted localization methods predominantly rely on a single type of sensing information. For instance, a significant body of research focuses on Angle-of-Arrival (AOA)-based localization [27]–[30], employing advanced signal processing techniques like MUSIC, ESPRIT [28], or Atomic Norm Minimization (ANM) [31]–[33] to achieve high-resolution AOA estimation. However, these angle-only methods neglect the valuable range information embedded in time-delay measurements. While some works have separately studied Time-of-Arrival (TOA)-based localization [34], the potential of a hybrid approach that jointly processes both angle and delay information within a multi-IRS collaborative framework to maximize localization accuracy has not yet been fully explored. Finally, the research on specific signal processing

techniques for estimating these parameters remains limited. Most existing research on IRS-assisted sensing focuses primarily on system-level optimization, such as joint beamforming design to minimize the Cramér-Rao Bound (CRB) [27], [29]. There is insufficient investigation into the signal processing algorithms for extracting essential information (such as angle and delay) from the reflected signals as a prerequisite for localization. Although efficient algorithms leveraging angular domain sparsity have been proposed for channel estimation [35], the design of a complete set of parameter estimation algorithms specifically for the hybrid localization scenario would not only enhance accuracy but also provide valuable insights for system and protocol design, ultimately contributing to the overall performance improvement of IRS-assisted sensing systems.

Motivated by the above considerations, we study a multiple semi-passive IRS deployed near the target area to enable collaborative localization of potential targets, as shown in Fig. 1. Our objective is to analyze the CRB of the proposed system for target localization and design efficient algorithms to estimate the cascaded delay, angle, and the target's location. It is worth noting that collaborative localization refers to the unified signal processing of the reflections received by the sensors of all semi-passive IRS, while hybrid localization involves jointly estimating the location based on both delay and angle information. **The analysis and algorithms in this paper are specifically developed for a single-target localization scenario, which lays a crucial foundation for more complex multi-target systems.** The main contributions of the paper are summarized as follows.

- **Proposed a multi-IRS collaborative localization system model and derived its performance bounds.** We propose a novel multi-IRS collaborative hybrid localization system. A key assumption of this model is the mutual orthogonality of signals reflected by different IRSs, which establishes a tractable theoretical framework for designing our core estimation algorithms. The feasibility of this idealized assumption is grounded in established techniques, such as zero-forcing (ZF) beamforming at the base station or the strategic orthogonal deployment of IRSs. Based on this model, we derive the Fisher Information Matrix (FIM) for cascaded delay, AOA, and Angle of Departure (AOD), and further obtain the CRB for target localization.
- **Designed a joint angle estimation algorithm adaptable to various signal conditions.** For the angle estimation problem, we classify the IRS reflected signal into three distinct cases based on its matrix rank and design unique signal preprocessing techniques for each. By constructing an atomic norm set, we transform the joint angle estimation problem into a convex optimization problem. We then derive angle estimation algorithms for different cases using the Alternating Direction Method of Multipliers (ADMM), achieving low-complexity estimation of multiple AOA and AOD pairs.
- **Designed a robust three-stage hybrid localization algorithm.** For the location estimation problem, we propose

a three-stage algorithm to perform hybrid localization that is resilient to errors in delay and angle estimates. By combining Weighted Least Squares (WLS), Total Least Squares (TLS) to handle errors in the coefficient matrix, and a final correction step using second-order information, the algorithm systematically addresses error propagation and accumulation, yielding an accurate location estimate.

- **Validated system gains and algorithm performance under the ideal model via simulations.** Finally, the superiority of the proposed system is validated through numerical simulations. The results clearly demonstrate the significant performance gains from multi-IRS collaboration, the use of hybrid information, and the distributed deployment strategy under the ideal condition of signal orthogonality. The simulations also confirm the effectiveness and accuracy of the proposed joint angle estimation and three-stage localization algorithms, particularly in low signal-to-noise ratio (SNR) environments.

The rest of this paper is as follows. In Section II, we introduce the system model of multiple IRSs collaborative hybrid localization. Then, Section III the CRB for delay, angle, and location estimation for semi-passive systems. Then, in Sections IV and V, we elaborate on the algorithms for angle and location estimation, respectively. Section VI reveals the performance superiority of the proposed system and algorithm compared to other benchmarks and some useful insights. Finally, Section VII concludes this paper.

Notations: Matrices, vectors and scalars are represented by bold uppercase, bold lowercase and standard lowercase letters, respectively. For a complex-valued scalar x , $|x|$ denotes its absolute value. For a complex-valued vector $\mathbf{x}$, $\|\mathbf{x}\|_p$, $[\mathbf{x}]_i$ represents the its p -norm and i -th element, respectively. For a general matrix $\mathbf{A}$, $\text{rank}(\mathbf{A})$, $\mathbf{A}^H$, $\mathbf{A}^\dagger$ and $[\mathbf{A}]_{i,j}$ denote its rank, conjugate transpose, Moore-Penrose pseudo-inverse and (i,j) -th element, respectively. For a square matrix $\mathbf{X}$, and $\text{tr}(\mathbf{X})$ denote its trace. $\mathbf{X} \succeq 0$ denotes that $\mathbf{X}$ is a positive semidefinite matrix. $\mathbb{C}^{M \times N}$ represents the $M \times N$ dimensional complex matrix space. $\mathbb{E}(\cdot)$ denotes the expectation operation. $\sim$ represents “distributed as” and $\mathcal{CN}(\mathbf{x}, \mathbf{R})$ represents the distribution of a circularly symmetric complex Gaussian random vector with mean vector $\mathbf{x}$ and covariance matrix $\mathbf{R}$. $\otimes$ represents the Kronecker product.

II. SYSTEM MODEL

As shown in Fig. 1, the model of a multi-IRSs collaborative hybrid localization system employing a semi-passive architecture is illustrated. This semi-passive architecture for IRS-assisted sensing consists of a BS with N_t transmit antennas for transmitting sensing signals, and K IRSs. Each IRS is equipped with N reflective elements to construct VLoS links. In this semi-passive configuration, the sensors are co-located with the IRSs, meaning each of the K IRSs is equipped with M sensors. The location of BS, k -th IRS, and target are denoted by $\mathbf{l}_{\text{BS}} = (x_{\text{BS}}, y_{\text{BS}})$, $\mathbf{l}_{l,k} = (x_{l,k}, y_{l,k})$ and $\mathbf{x} = (x_{\text{T}}, y_{\text{T}})$, respectively. The distance and angle between the target and the k -th IRS can be expressed

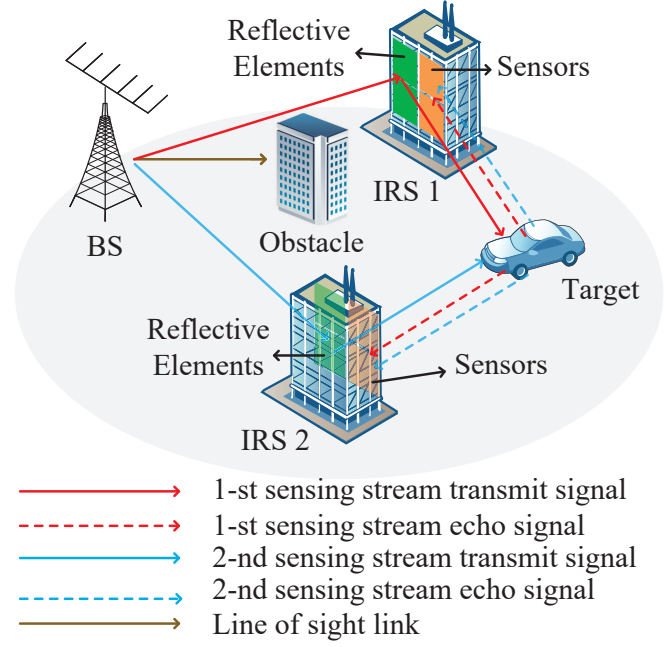


Fig. 1. Multiple IRSs collaborative hybrid localization system.

as $d_k = c\tau_k = \sqrt{(y_{\text{T}} - y_{l,k})^2 + (x_{\text{T}} - x_{l,k})^2}$ and $\theta_k = \arctan(y_{\text{T}} - y_{l,k} / x_{\text{T}} - x_{l,k})$, respectively where c and τ_k represent the speed of light and corresponding propagation delay, respectively. Under the assumption of the far-field condition, the AOD from the reflective elements on IRS to target is the same as the AOA, i.e., $\theta_k = \theta_{k,\text{AOA}} = \theta_{k,\text{AOD}}$.

A. Channel Model

The channel between the BS and the k -th IRS $\mathbf{H}_k$ consists of R_k multipath components, which can be expressed as

$$\mathbf{H}_k = \sum_{r=1}^{R_k} \alpha_{k,r} \mathbf{a}(\varphi_{k,r}, N) \mathbf{a}^H(\theta_{k,r}, N_t), \quad (1)$$

where r_k represents the number of paths between the base station and the k -th IRS, and $\alpha_{k,r}$ represent the path gain and $\theta_{k,r}$, $\varphi_{k,r}$ represents the AOD and AOA of the r -th path, respectively; $\mathbf{a}(\theta_0, N_0)$ represents the steering vector where θ_0 represents the phase difference between two adjacent elements or antennas, and N_0 represents the number of antennas transmitted or received, which is defined as

$$\mathbf{a}(\theta_0, N_0) = \left[1, e^{j2\pi \sin \theta_0}, \dots, e^{j2\pi(N_0-1) \sin \theta_0} \right]^T \in \mathbb{C}^{N_0}. \quad (2)$$

Because localization related information can only be extracted from LoS, the channel from k -th IRS to l -th IRS via target of semi-passive system can be modeled as

$$\mathbf{H}_{l,k} = \alpha_{l,k} \mathbf{a}(\theta_{l,\text{AOA}}, M) \mathbf{a}^H(\theta_{k,\text{AOD}}, N), \quad (3)$$

where $\alpha_{l,k} = \sqrt{\lambda^2 \kappa / 64 \pi^3 d_k^2 d_l^2}$ denotes the path loss, where κ denotes the radar cross section, λ denotes the wavelength of the signal.

B. Performance Analysis

Let $\mathbf{X} \in \mathbb{C}^{N_t \times L}$ be the sensing signal matrix, with $L \gg N_t$ being the length of the radar pulse frame and $\mathbf{X}$ can be modeled as

$$\mathbf{X} = \sum_{k=1}^K \mathbf{W}_k \mathbf{X}_k, \quad (4)$$

where $\mathbf{X}_k, \mathbf{W}_k$ represent the k -th sensing stream and its beam-forming matrix. To ensure that the receiver can distinguish the signals corresponding to different paths reflected by the K IRSs, the signals reflected by the K IRSs are designed to be mutually orthogonal, i.e.,

$$\mathbf{H}_k \mathbf{W}_{k'} = 0, \forall k \neq k', \quad (5)$$

which can be achieved through the orthogonal deployment of IRSs or zero-forcing beamforming. The sensing streams corresponding to the same IRS are assumed to be independent in different pulse frame while the sensing streams corresponding to the different IRSs are assumed to be independent in any frame, i.e.,

$$\mathbf{X}_k \mathbf{X}_k^H = \mathbf{I}_{N_t}, \mathbf{X}_k \mathbf{X}_{k'}^H = \mathbf{0}_{N_t}, \forall k \neq k'. \quad (6)$$

We adopt a semi-passive IRS architecture. While this increases hardware complexity, it is a well-established approach in sensing literature that significantly improves the signal-to-noise ratio (SNR) by reducing round-trip path loss. For tractability, our model focuses on the primary reflection paths, assuming sensor data is collected via an ideal backhaul. Higher-order reflections, such as inter-IRS paths, are considered negligible due to higher path loss but can be explored in future work. For the semi-passive architecture, the signal received by the sensors on the l -th IRS can be expressed as

$$\mathbf{R}_l(t) = \sum_{k=1}^K \mathbf{H}_{l,k} \mathbf{\Theta}_k \mathbf{H}_k \mathbf{X}(t - \dot{\tau}_{l,k}) + \mathbf{N}_l, \quad (7)$$

where $\mathbf{N}_l$ represents the Gaussian white noise at the sensors in l -th IRS and $\text{vec}(\mathbf{N}_l) \sim \mathcal{CN}(0, \sigma_l^2 \mathbf{I}_{ML})$, $\dot{\tau}_{l,k} \triangleq \tau_{\text{B2I},k} + \tau_{l,k}$, $\tau_{\text{B2I},k}$ represents the signal propagation delay between the BS and the k -th IRS and $\tau_{l,k} \triangleq \tau_k + \tau_l$ represent the signal cascade propagation delay from k -th IRS to the l -th IRS via target. Let $\mathbf{\Theta}_k = \text{diag}(\alpha_{k,1} e^{j\beta_{k,1}}, \dots, \alpha_{k,N} e^{j\beta_{k,N}})$ denotes the reflection coefficient matrix of k -th IRS, where $\alpha_{k,n}$ and $\beta_{k,n}$ denotes the amplitude and phase reflection coefficient of the n -th element of k -th IRS, respectively. The phase shift matrix $\mathbf{\Theta}_k$ of each IRS can be dynamically configured to enhance system performance. It is important to note that the performance analysis and estimation algorithms developed in this work are general and applicable to any given IRS phase configuration.

Since the location of the BS and IRS are fixed, $\tau_{\text{B2I},k}$ can be treated as known. Our hybrid localization model first extracts unknown angle $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_K] \in \mathbb{R}^{K \times 1}$ and cascade delay $\boldsymbol{\tau} = [\tau_{1,1}, \tau_{1,2}, \dots, \tau_{K,K}] \in \mathbb{R}^{K^2 \times 1}$ information from received signal $\mathbf{R} = [\mathbf{R}_1, \dots, \mathbf{R}_K] \in \mathbb{C}^{M \times KL}$, and then utilizes above hybrid information to estimate the target's location $\mathbf{x}$.

III. PERFORMANCE ANALYSIS OF MULTI-IRS COLLABORATIVE HYBRID LOCALIZATION SYSTEM

CRB provides a lower bound of variance of any unbiased estimator, therefore, we first derive the CRB for the location related estimation in the proposed system to characterize the system's performance and validate the effectiveness of the subsequent estimation algorithms. For any variable to be estimated $\mathbf{u}$, the variance of the unbiased estimator $\hat{\mathbf{u}}$ satisfies

$$\text{var}(\hat{\mathbf{u}}) \geq \text{CRB}(\mathbf{u}) = \mathbf{F}^{-1}(\mathbf{u}), \quad (8)$$

where $\text{CRB}(\mathbf{u})$ and $\mathbf{F}^{-1}(\mathbf{u})$ denote the CRB matrix and FIM of $\mathbf{u}$ to be estimated, respectively.

A. CRB for Cascade Delay Estimation

In this subsection, we analyze the CRB for $\boldsymbol{\tau}$ estimation involved in the localization system. The Fisher matrix for $\boldsymbol{\tau}$ can be modeled as

$$\mathbf{F}(\boldsymbol{\tau}) = \begin{bmatrix} F(\tau_{1,1}, \tau_{1,1}) & F(\tau_{1,1}, \tau_{1,2}) & \dots & F(\tau_{1,1}, \tau_{K,K}) \\ F(\tau_{1,2}, \tau_{1,1}) & F(\tau_{1,2}, \tau_{1,2}) & \dots & F(\tau_{1,2}, \tau_{K,K}) \\ \dots & \dots & \dots & \dots \\ F(\tau_{K,K}, \tau_{1,1}) & F(\tau_{K,K}, \tau_{1,2}) & \dots & F(\tau_{K,K}, \tau_{K,K}) \end{bmatrix}, \quad (9)$$

where the element of $\mathbf{F}(\boldsymbol{\tau})$ is given by

$$F(\tau_{l,k}, \tau_{l',k'}) = \mathbb{E}_{\mathbf{R}|\boldsymbol{\tau}, \boldsymbol{\theta}} \left[\frac{\partial^2 \log p(\mathbf{R}|\boldsymbol{\tau}, \boldsymbol{\theta})}{\partial \tau_{l,k} \partial \tau_{l',k'}} \right], \quad (10)$$

where $p(\mathbf{R}|\boldsymbol{\tau}, \boldsymbol{\theta})$ is the joint probability density function of $\mathbf{R}$ conditioned on $\boldsymbol{\tau}, \boldsymbol{\theta}$ and can be given as

$$p(\mathbf{R}|\boldsymbol{\tau}, \boldsymbol{\theta}) = \exp \left(\sum_{l=1}^K \int_T -\frac{1}{\sigma_l^2} |\mathbf{R}_l(t) - \overline{\mathbf{R}}_l(t)|^2 dt \right), \quad (11)$$

where $\overline{\mathbf{R}}_l(t)$ denotes the deterministic component of the received echo signal [36]. By substituting (7) into (11) and ignoring terms that are unrelated to the derivative, $F(\tau_{l,k}, \tau_{l',k'})$ can be expressed as

$$F(\tau_{l,k}, \tau_{l',k'}) = M \alpha_{l,k}^2 \left\| \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{\Theta}_k \mathbf{H}_k \mathbf{W}_k \dot{\mathbf{X}}_k(t) \right\|_2^2 \quad (12)$$

when $l = l', k = k'$ and 0 otherwise, where $\dot{\mathbf{X}}_k(t) = \partial \mathbf{X}_k(t) / \partial t$ and the CRB for the cascade delay estimation can be expressed as $\text{CRB}(\boldsymbol{\tau}) = \mathbf{F}^{-1}(\boldsymbol{\tau})$.

B. CRB for Angle Estimation

In this subsection, we analyze the CRB for angle estimation involved in the localization system, which includes two types of angles: AOA and AOD. Similar to the Fisher matrix for delay estimation, the Fisher matrix for AOA and AOD angle estimations can also be represented as a diagonal matrix as

$$\mathbf{F}(\boldsymbol{\theta}_{\text{AOA}}) = \text{diag}(F(\theta_{1,\text{AOA}}), \dots, F(\theta_{K,\text{AOA}})), \quad (13)$$

$$\mathbf{F}(\boldsymbol{\theta}_{\text{AOD}}) = \text{diag}(F(\theta_{1,\text{AOD}}), \dots, F(\theta_{K,\text{AOD}})), \quad (14)$$

where $F(\theta_{l,\text{AOA}})$ and $F(\theta_{k,\text{AOD}})$ can be obtained by computing the second-order derivatives of $p(\mathbf{R}|\boldsymbol{\tau}, \boldsymbol{\theta})$ with respect to $\theta_{l,\text{AOA}}$ and $\theta_{k,\text{AOD}}$, respectively, i.e.,

$$F(\theta_{l,\text{AOA}}) = \sum_{k=1}^K \left\| \dot{\mathbf{a}}(\theta_{l,\text{AOA}}, N) \mathbf{b}_k^H \right\|_2^2, \quad (15)$$

$$F(\theta_{k,\text{AOD}}) = \sum_{l=1}^K \left\| \mathbf{a}(\theta_{l,\text{AOA}}, N) \mathbf{c}_k^H \right\|_2^2, \quad (16)$$

where $\dot{\mathbf{a}}(\theta, N) \triangleq \partial \mathbf{a}(\theta, N) / \partial \theta$, $\mathbf{b}_k^H = \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{\Theta}_k \mathbf{H}_k \mathbf{W}_k \mathbf{X}_k$ and $\mathbf{c}_k^H = \dot{\mathbf{a}}^H(\theta_{k,\text{AOD}}, N) \mathbf{\Theta}_k \mathbf{H}_k \mathbf{X}_k$. Here, we omit the derivative calculation process, as it is similar to the process in the previous subsection.

C. CRB for Location Estimation

Since the location cannot be directly estimated from the signal, the Fisher Information Matrix (FIM) of the location can be derived using the chain rule, i.e.,

$$\mathbf{F}(\mathbf{x}) = \frac{\partial [\boldsymbol{\tau}, \boldsymbol{\theta}]}{\partial \mathbf{x}} \mathbf{F}(\boldsymbol{\tau}, \boldsymbol{\theta}) \left(\frac{\partial [\boldsymbol{\tau}, \boldsymbol{\theta}]}{\partial \mathbf{x}} \right)^T \quad (17)$$

where $\mathbf{F}(\boldsymbol{\tau}, \boldsymbol{\theta}) = \text{diag}(\mathbf{F}(\boldsymbol{\theta}_{\text{AOA}}) + \mathbf{F}(\boldsymbol{\theta}_{\text{AOD}}), \mathbf{F}(\boldsymbol{\tau}))$ and $\partial [\boldsymbol{\tau}, \boldsymbol{\theta}] / \partial \mathbf{x}$ represents the geometric relationship, which can be expressed as

$$\frac{\partial [\boldsymbol{\tau}, \boldsymbol{\theta}]}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \boldsymbol{\tau}}{\partial x_T} & \frac{\partial \boldsymbol{\theta}}{\partial x_T} \\ \frac{\partial \boldsymbol{\tau}}{\partial y_T} & \frac{\partial \boldsymbol{\theta}}{\partial y_T} \end{bmatrix}, \quad (18)$$

where $\partial \boldsymbol{\tau} / \partial x_T = [\cos \theta_1 + \cos \theta_1, \dots, \cos \theta_K + \cos \theta_K] / c \in \mathbb{R}^{K^2}$, $\partial \boldsymbol{\tau} / \partial y_T = [\sin \theta_1 + \sin \theta_1, \dots, \sin \theta_K + \sin \theta_K] / c \in \mathbb{R}^{K^2}$, $\partial \boldsymbol{\theta} / \partial x_T = -[\sin \theta_1 / d_1, \dots, \sin \theta_K / d_K] \in \mathbb{R}^K$ and $\partial \boldsymbol{\theta} / \partial y_T = [\cos \theta_1 / d_1, \dots, \cos \theta_K / d_K] \in \mathbb{R}^K$. Then, the CRB for the estimation of $\mathbf{x}$ can be expressed as

$$\text{CRB}(\mathbf{x}) = \text{tr}(\mathbf{F}^{-1}(\mathbf{x})) \quad (19)$$

IV. DELAY AND ANGLE ESTIMATION ALGORITHMS

In this section, we propose angle and cascade delay estimation algorithms. The estimation of the cascade delay is relatively straightforward. For sensors on each semi-passive IRS, the received signal $\mathbf{R}_l$ is used to search for the time delay $\tau_{l,k}$ through the following matched filtering.

$$\hat{\tau}_{l,k} = \arg \max_{\tau} \|\mathbf{R}_l(t) \mathbf{X}_k^H(t - \tau - \tau_{\text{B2I},k})\|_2, \quad (20)$$

and $\mathbf{R}_{l,k}$ is obtained through the following matched filtering

$$\begin{aligned} \mathbf{R}_{l,k} &= \mathbf{R}_l(t) \mathbf{X}_k^H(t - \hat{\tau}_{l,k}) \\ &= \alpha_{l,k} \mathbf{a}(\theta_{l,\text{AOA}}, M) \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{S}_k + \mathbf{N}_l \mathbf{X}_k^H(t - \hat{\tau}_{l,k}), \end{aligned} \quad (21)$$

where $\mathbf{S}_k = \mathbf{\Theta}_k \mathbf{H}_k \mathbf{W}_k (\mathbf{X}_k \mathbf{X}_k^H) = \mathbf{\Theta}_k \mathbf{H}_k \mathbf{W}_k$. Since path gain is related to distance and the target's location is unknown in the angle estimation problem, the path gain is unknown. As a result, the traditional angle estimation problem can be expressed as

$$\min_{\theta_{l,\text{AOA}}, \theta_{k,\text{AOD}}, \alpha_{l,k}} \|\mathbf{R}_{l,k} - \alpha_{l,k} \mathbf{a}(\theta_{l,\text{AOA}}) \mathbf{a}^H(\theta_{k,\text{AOD}}) \mathbf{S}_k\|_2^2. \quad (22)$$

Because of the non-convexity of the steering vector and the coupling between variables, this problem cannot be directly solved using convex optimization methods.

The traditional MUSIC, compressive sensing, and ESPRIT algorithms, due to their complexity, discretization, and sensitivity to noise, may not achieve optimal performance. Furthermore, these algorithms cannot simultaneously utilize observation samples from different angles for joint angle estimation, which limits the full potential of the collaborative localization system.

A. Atomic Norm Minimization Formulation

First, we introduce the concept of an atomic set. A 2-D atomic norm set $\mathcal{A}$ can be represented as

$$\mathcal{A} = \{\mathbf{A}(\theta_{l,\text{AOA}}, \theta_{k,\text{AOD}}) | \theta_{l,\text{AOA}}, \theta_{k,\text{AOD}} \in [-\pi/2, \pi/2]\}, \quad (23)$$

with $\mathbf{A}(\theta_{l,\text{AOA}}, \theta_{k,\text{AOD}}) = \mathbf{a}(\theta_{l,\text{AOA}}, M) \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \in \mathbb{C}^{M \times N}$. The atomic norm of a matrix $\mathbf{X}$ is defined as:

$$\|\mathbf{X}\|_{\mathcal{A}} = \inf \left\{ \sum_{l,k} \alpha_{l,k} | \mathbf{X} = \sum_{l,k} \beta_{l,k} \mathbf{A}(\theta_{l,\text{AOA}}, \theta_{k,\text{AOD}}) \right\}. \quad (24)$$

with $\beta_{l,k} > 0$ represents the weight coefficient of each atomic term. By minimizing the atomic norm of $\mathbf{X}$, we can decompose $\mathbf{X}$ within the atomic set, thereby obtaining the steering vectors corresponding to the components. It is noted that the signal model in (21) does not directly fit the standard atomic norm minimization framework, which typically decomposes a signal into a linear combination of atoms with scalar coefficients. To make the problem compatible with the ANM framework, a critical signal pre-processing step is performed first. Specifically, we eliminate the effect of the known matrix $\mathbf{S}_k$ by right-multiplying the received signal matrix $\mathbf{R}_{l,k}$ by the pseudo-inverse of $\mathbf{S}_k$, $\mathbf{S}_k^\dagger$. Let the pre-processed signal be $\mathbf{R}'_{l,k} = \mathbf{R}_{l,k} \mathbf{S}_k^\dagger \approx \alpha_{l,k} \mathbf{a}(\theta_{l,\text{AOA}}, M) \mathbf{a}^H(\theta_{k,\text{AOD}}, N)$. However, due to the presence of noise, the received signal is often difficult to accurately decompose within the atomic set. Therefore, we consider minimizing the atomic norm within the F -norm domain of the received signal, i.e.,

$$(P1): \min_{\mathbf{X}} \|\mathbf{X}\|_{\mathcal{A}} \quad (25)$$

$$\text{s.t.} \quad \|\mathbf{X} - \mathbf{R}'_{l,k}\|_F \leq \eta, \quad (26)$$

where η represents the allowable margin of error. By utilizing the properties of the Vandermonde decomposition, problem (P1) can be equivalently transformed into problem (P2), which can be expressed as

$$(P2): \min_{\mathbf{X}, \mathbf{Z}, \mathbf{v}, \mathbf{u}} \text{tr}(\mathbf{T}(\mathbf{v})) + \text{tr}(\mathbf{T}(\mathbf{u})) \quad (27)$$

$$\text{s.t.} \quad \|\mathbf{X} - \mathbf{R}'_{l,k}\|_F \leq \eta, \quad (28)$$

$$\begin{bmatrix} \mathbf{T}(\mathbf{v}) & \mathbf{Z} \\ \mathbf{Z}^H & \mathbf{T}(\mathbf{u}) \end{bmatrix} \succeq 0, \quad (29)$$

where $\mathbf{T}(\mathbf{u})$ denotes a Hermitian Toeplitz matrix with $\mathbf{u}$ as its main diagonal element [37]. After the above equivalence, the problem becomes easier to solve. The previous subsection only introduced the basic one-dimensional atomic norm. In the subsequent solution process, we will apply the corresponding transformations to the one-dimensional atomic norm for the proposed model.

B. Joint Angle Estimation Algorithm Based on ADMM

In this subsection, we present a detailed description of the angle estimation algorithm. First, we formulate a joint estimation framework where the objective function is a weighted

sum over all K collaborative paths, enabling the simultaneous processing of multiple observation matrices to estimate a common angle. This is a significant departure from traditional methods that handle a single observation matrix. Second, we propose a comprehensive multi-case framework, tailoring the signal processing and algorithm design based on the rank of the IRS-reflected signal matrix $\mathbf{S}_k$.

Specifically, for any angle θ_i , there are K observation samples containing information about $\theta_{i,\text{AOA}}$ (i.e., $\mathbf{R}_{i,k}, \forall k$) and K observation samples containing information about $\theta_{i,\text{AOD}}$ (i.e., $\mathbf{R}_{l,i}, \forall l$). Joint estimation using all relevant observation samples achieves better performance compared to estimation based on single observation sample. Therefore, we first estimate $\theta_{i,\text{AOA}}$ and $\theta_{i,\text{AOD}}$ based on $\mathbf{R}_{i,k}, \forall k$ and $\mathbf{R}_{l,i}, \forall l$, respectively, and obtain θ_i through weighted combination of $\theta_{i,\text{AOA}}$ and $\theta_{i,\text{AOD}}$. Through the equivalence transformation in previous section, the estimation of $\theta_{i,\text{AOA}}$ can be expressed as

$$(P3) \quad \min_{\mathbf{v}_i, \{\mathbf{u}_k\}, \{\mathbf{Z}_{i,k}\}} \sum_{k=1}^K e_{i,k} [\text{tr}(\mathbf{T}(\mathbf{u}_k)) + \text{tr}(\mathbf{T}(\mathbf{v}_i))] \quad (30a)$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{T}(\mathbf{v}_i) & \mathbf{Z}_{i,k} \\ \mathbf{Z}_{i,k}^H & \mathbf{T}(\mathbf{u}_k) \end{bmatrix} \succeq 0, \forall k, \quad (30b)$$

where $e_{i,k} = \|\mathbf{R}_{i,k}\|_2^2$ is a weighted coefficient representing the energy of the received signal, which characterizes the confidence level for each observation sample. Problem (P3) is non-convex due to $\|\mathbf{Z}_{i,k}\mathbf{S}_k - \mathbf{R}_{i,k}\|_2^2$ is an indefinite quadratic form and the semi-definite constraints. To address this, we choose to project the corresponding matrix onto the semi-definite cone and use $\mathbf{Z}_{i,k} - \mathbf{R}_{i,k}\mathbf{S}_k^\dagger$ to replace $\mathbf{Z}_{i,k}\mathbf{S}_k - \mathbf{R}_{i,k}$ for characterizing the reconstruction error of the signal. However, the Moore-Penrose inverse of $\mathbf{S}_k$ depends on its rank, which will affect the subsequent solution process. Therefore, we categorize the discussion based on the rank of $\mathbf{S}_k$ as follows:

1) $\text{rank}(\mathbf{S}_k) = N$: Since the rank of $\mathbf{S}_k$ is N , the Moore-Penrose inverse of $\mathbf{S}_k$ can be denoted as $\mathbf{S}_k^\dagger = \mathbf{S}_k^H (\mathbf{S}_k \mathbf{S}_k^H)^{-1}$. Then, the $\mathbf{R}_{i,k}$ can be multiplied on the right by $\mathbf{S}_k$ and $\|\mathbf{Z}_{i,k} - \mathbf{R}_{i,k}\mathbf{S}_k^\dagger\|_2^2$ is a convex function with respect to $\mathbf{Z}_{i,k}$. Then the Lagrangian function of (P3) can be formulated as (32), where $\mathbf{G}_{i,k}$ represents the Lagrange multiplier, and ρ is the penalty factor. For ease of subsequent representation, the matrix $\mathbf{G}_{i,k}, \mathbf{Z}_{i,k}$ is expressed as the following block matrix:

$$\mathbf{G}_{i,k} = \begin{bmatrix} \mathbf{G}_{i,k,0} & \mathbf{G}_{i,k,1} \\ \mathbf{G}_{i,k,1}^H & \mathbf{G}_{i,k,2} \end{bmatrix}, \mathbf{Z}_{i,k} = \begin{bmatrix} \mathbf{Z}_{i,k,0} & \mathbf{Z}_{i,k,1} \\ \mathbf{Z}_{i,k,1}^H & \mathbf{Z}_{i,k,2} \end{bmatrix}, \quad (31)$$

where $\mathbf{G}_{i,k,0}, \mathbf{Z}_{i,k,0} \in \mathbb{C}^{M \times M}$, $\mathbf{G}_{i,k,1}, \mathbf{Z}_{i,k,1} \in \mathbb{C}^{M \times N}$ and $\mathbf{G}_{i,k,2}, \mathbf{Z}_{i,k,2} \in \mathbb{C}^{N \times N}$.

These variables can be divided into three parts and updated sequentially by solving the corresponding subproblems. Specifically, the subproblems and their solutions are elaborated as follows.

- Subproblem with respect to $\mathbf{v}_i, \mathbf{u}_k, \mathbf{X}_{i,k}$

By computing the derivative of $\mathcal{L}_1(\mathbf{v}_i, \mathbf{u}_k, \mathbf{X}_{i,k}, \mathbf{Z}_{i,k}, \mathbf{G}_{i,k})$, we have

$$\begin{aligned} \nabla_{\mathbf{v}_i} \mathcal{L}_1 &= \sum_{k=1}^K e_{i,k} \{M \mathbf{e}_{M,1} - f(\mathbf{G}_{i,k,0}) \\ &\quad + 2\rho [f(\mathbf{Z}_{i,k,0}) - f(\mathbf{T}(\mathbf{u}_k))]\}, \end{aligned} \quad (33)$$

$$\nabla_{\mathbf{u}_k} \mathcal{L}_1 = N \mathbf{e}_{N,1} - f(\mathbf{G}_{i,k,2}) + 2\rho [f(\mathbf{T}(\mathbf{u}_k)) - f(\mathbf{Z}_{i,k,2})], \quad (34)$$

$$\nabla_{\mathbf{X}_{i,k}} \mathcal{L}_1 = (\mathbf{X}_{i,k} - \mathbf{R}_{i,k} \mathbf{S}_k^\dagger) - \mathbf{G}_{i,k,1} + \rho (2\mathbf{X}_{i,k} - 2\mathbf{Z}_{i,k,1}) \quad (35)$$

where $f(\mathbf{X}, N)$ is a mapping from a matrix $\mathbf{X} \in \mathbb{C}^{N \times N}$ to a vector, which is defined as

$$[f(\mathbf{X}, N)]_n = \begin{cases} \sum_{i=1}^N [\mathbf{X}]_{i,i}, n = 1, \\ 2 \sum_{j=i+1}^N [\mathbf{X}]_{i,j}, n > 1. \end{cases} \quad (36)$$

By setting these derivatives to 0, $\mathbf{v}_i, \mathbf{u}_k, \mathbf{X}_{i,k}$ are updated by

$$[\mathbf{u}_k]_n = \frac{1}{g(n, N)} \left[\frac{f(\mathbf{G}_{i,k,2}) + 2\rho f(\mathbf{Z}_{i,k,2}) - N \mathbf{e}_{N,1}}{2\rho} \right]_n, \quad (37)$$

$$[\mathbf{v}_i]_m = \frac{1}{g(m, M)} \left[\frac{\sum_{l=1}^K e_{i,k} [f(\mathbf{G}_{i,k,0}, M) + 2\rho f(\mathbf{Z}_{i,k,0}, M) - M \mathbf{e}_{M,1}]}{\sum_{l=1}^K e_{i,k} 2\rho} \right]_m, \quad (38)$$

$$\mathbf{X}_{i,k} = (\lambda \mathbf{R}_{i,k} \mathbf{S}_k^\dagger + \mathbf{G}_{i,k,1} + 2\rho \mathbf{Z}_{i,k,1}) (\lambda + 2\rho)^{-1}, \quad (39)$$

where t is the iteration index and function $g(\cdot)$ is defined as

$$g(n, N) = \begin{cases} N, n = 1, \\ 2(N - n + 1), n > 1, \end{cases} \quad (40)$$

- Subproblem with respect to $\mathbf{Z}_{i,k}$

The update of $\mathbf{Z}_{i,k}$ can be expressed as solving the following problem:

$$\mathbf{Z}_{i,k}^r = \arg \min_{\mathbf{Z}_{i,k} \succeq 0} \mathcal{L}_1(\mathbf{v}_i^r, \mathbf{u}_k^r, \mathbf{X}_{i,k}^r, \mathbf{Z}_{i,k}^r, \mathbf{G}_{i,k}^{r-1}). \quad (41)$$

Because the positive semi-definite constraint is a non-convex constraint, we first ignore this constraint to solve the problem, leading to the following result

$$\tilde{\mathbf{Z}}_{i,k}^r = \begin{bmatrix} \mathbf{T}(\mathbf{v}_i^r) & \mathbf{X}_{i,k}^r \\ (\mathbf{X}_{i,k}^r)^H & \mathbf{T}(\mathbf{u}_k^r) \end{bmatrix} - \frac{\mathbf{G}_{i,k}^{r-1}}{2\rho}. \quad (42)$$

To satisfy the positive semi-definite constraint, we perform eigenvalue decomposition on $\tilde{\mathbf{Z}}_{i,k}^{r+1}$, resulting in $\tilde{\mathbf{Z}}_{i,k}^r = \sum_{q=1}^{R_{i,k}} \lambda_q \mathbf{v}_q \mathbf{v}_q^H$, where $R_{i,k}$ is the rank of $\tilde{\mathbf{Z}}_{i,k}^r$ and

$$\mathbf{Z}_{i,k}^r = \sum_{q=1}^{R_{i,k}} \max(\lambda_q, 0) \mathbf{v}_q \mathbf{v}_q^H. \quad (43)$$

$$\begin{aligned} \mathcal{L}_1(\mathbf{v}_i, \mathbf{u}_k, \mathbf{X}_{i,k}, \mathbf{G}_{i,k}, \mathbf{Z}_{i,k}) &= \sum_{k=1}^K e_{i,k} \left[\text{tr}(\mathbf{T}(\mathbf{u}_k) + \mathbf{T}(\mathbf{v}_i)) + \lambda \left\| \mathbf{Z}_{i,k} - \mathbf{R}_{i,k} \mathbf{S}_k^\dagger \right\|_2^2 \right. \\ &\quad \left. + \left\langle \mathbf{G}_{i,k}, \left(\mathbf{X}_{i,k} - \begin{bmatrix} \mathbf{T}(\mathbf{v}_i) & \mathbf{Z}_{i,k} \\ \mathbf{Z}_{i,k}^H & \mathbf{T}(\mathbf{u}_k) \end{bmatrix} \right) \right\rangle + \rho \left\| \mathbf{X}_{i,k} - \begin{pmatrix} \mathbf{T}(\mathbf{v}_i) & \mathbf{Z}_{i,k} \\ \mathbf{Z}_{i,k}^H & \mathbf{T}(\mathbf{u}_k) \end{pmatrix} \right\|_2^2 \right], \end{aligned} \quad (32)$$

Algorithm 1 Proposed Alternating Optimization Algorithm.

- 1: **Input:** Initialize $\rho, \lambda, \mathbf{u}_k^0, \mathbf{v}_i^0, \mathbf{X}_{i,k}^0, \mathbf{Z}_{i,k}^0, \mathbf{G}_{i,k}^0$, threshold $\varepsilon_1, \varepsilon_2$ and iteration index $r = 1$
- 2: **repeat:**
- 3: Update $\mathbf{u}_k^r, \mathbf{v}_i^r, \mathbf{X}_{i,k}^r$ according to (34), (35) and (36), respectively.
- 4: Update $\mathbf{Z}_{i,k}^r$ by solving according to (40).
- 5: Update $\mathbf{G}_{i,k}^r$ by solving according to (41).
- 6: $r = r + 1$.
- 7: **until** $\left\| \mathbf{X}_{i,k}^r - \mathbf{X}_{i,k}^{r-1} \right\|_2^2 \leq \varepsilon_1$ and $\left\| \mathbf{X}_{i,k}^r - \mathbf{X}_{i,k}^{r-1} \right\|_2^2 \leq \varepsilon_2$.
- 8: **Output:** $\mathbf{v}_i$

- Subproblem with respect to $\mathbf{G}_{i,k}$

The update for $\mathbf{G}_{i,k}$ is based on the multiplier update rule, i.e.,

$$\mathbf{G}_{i,k}^r = \mathbf{G}_{i,k}^{r-1} + \rho \left(\mathbf{Z}_{i,k}^r - \begin{bmatrix} \mathbf{T}(\mathbf{v}_i^r) & \mathbf{X}_{i,k}^r \\ (\mathbf{X}_{i,k}^r)^H & \mathbf{T}(\mathbf{u}_k^r) \end{bmatrix} \right). \quad (44)$$

At this point, we have derived the closed-form solutions for updating each variable. By alternately updating these variables until convergence, the entire algorithm is summarized in **Algorithm 1**.

After solving for $\mathbf{T}(\mathbf{v}_i)$, we perform eigenvalue decomposition, and the normalised eigenvector corresponding to the largest eigenvalue is taken as the estimation of $\mathbf{a}(\theta_{i,\text{AOA}}, M)$. The estimation of $\theta_{i,\text{AOD}}$ is similar to that of $\theta_{i,\text{AOA}}$. Specifically, replace Lagrange function $\mathcal{L}_1(\mathbf{v}_i, \mathbf{u}_k, \mathbf{X}_{i,k}, \mathbf{G}_{i,k}, \mathbf{Z}_{i,k})$ with $\mathcal{L}_1(\mathbf{v}_k, \mathbf{u}_i, \mathbf{X}_{k,i}, \mathbf{G}_{k,i}, \mathbf{Z}_{k,i})$ and the subsequent variable update process can follow **Algorithm 1**.

2) $1 < \text{rank}(\mathbf{S}_k) < N$: The singular value decomposition (SVD) of matrix $\mathbf{S}_k = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H$. Based on the result of the singular value decomposition and substituting it into (21), $\mathbf{R}_{i,k}$ can be expressed as

$$\mathbf{R}_{i,k} = \alpha_{i,k} \mathbf{a}(\theta_{i,\text{AOA}}, M) \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H + \mathbf{N}_{i,k} \quad (45)$$

where $\mathbf{U}_k \in \mathbb{C}^{N \times \text{rank}(\mathbf{S}_k)}$, $\mathbf{\Sigma}_k \in \mathbb{C}^{\text{rank}(\mathbf{S}_k) \times \text{rank}(\mathbf{S}_k)}$, $\mathbf{V}_k \in \mathbb{C}^{L \times \text{rank}(\mathbf{S}_k)}$, $\mathbf{N}_{i,k} = \mathbf{N}_i \mathbf{X}_k^H(t - \hat{\tau}_{i,k})$ represents the Gaussian white noise after matched filtering. It is noted that when the vector $\mathbf{a}^H(\theta_{k,\text{AOD}}, N)$ containing the information of $\theta_{k,\text{AOD}}$ is multiplied by the left singular matrix $\mathbf{U}_k$, and $\mathbf{a}^H(\theta_{k,\text{AOD}}, N)$ is compressed into a lower-dimensional vector, which leads to the destruction of the structure of the steering vector. Therefore, we reconstruct the atomic set as

$$\mathcal{A} = \left\{ \mathbf{A}(\theta_{l,\text{AOA}}, \mathbf{b}) \in \mathbb{C}^{M \times \text{rank}(\mathbf{S}_k)} \mid \theta_{l,\text{AOA}} \in [-\pi/2, \pi/2] \right\}, \quad (46)$$

with $\mathbf{A}(\theta_{l,\text{AOA}}, \mathbf{b}) = \mathbf{a}(\theta_{l,\text{AOA}}, M) \mathbf{b}^H$ and $\|\mathbf{b}\|_2 = 1$. By multiplying matrix $(\mathbf{\Sigma}_k \mathbf{V}_k)^\dagger$ on the right, the received signal $\mathbf{R}_{i,k}$ can be re-expressed as the sum of the atoms and noise, i.e.,

$$\mathbf{R}_{i,k} (\mathbf{\Sigma}_k \mathbf{V}_k)^\dagger = \beta_{i,k} \mathbf{A}(\theta_{l,\text{AOA}}, \mathbf{b}) + \mathbf{N}_{i,k} (\mathbf{\Sigma}_k \mathbf{V}_k)^\dagger, \quad (47)$$

where $\beta_{i,k} \mathbf{b}^H$ equivalently represents $\alpha_{i,k} \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{U}_k$ in $\mathbf{R}_{i,k}$, and $\beta_{i,k} \mathbf{b}^H$ is more consistent with the definition of atomic norm compared to $\alpha_{i,k} \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{U}_k$. Then, we compute the atomic norm of matrix $\mathbf{R}_{i,k} (\mathbf{\Sigma}_k \mathbf{V}_k)^\dagger$ using a method similar to **Algorithm 1**, and the corresponding Lagrangian function is modified as shown in (48). Specifically, these variables can be divided into three parts $(\mathbf{v}_i, \mathbf{P}_k, \mathbf{X}_{i,k})$, $\mathbf{G}_{i,k}$ and $\mathbf{Z}_{i,k}$ and updated sequentially. Since the specific update method is similar to **Algorithm 1**. Once the estimation of angle $\theta_{l,\text{AOA}}$ is completed, the next step is to estimate angle $\theta_{k,\text{AOD}}$. The corresponding problem can be expressed as

$$\min_{\alpha_{i,k}, \theta_{k,\text{AOD}}} \left\| \alpha_{i,k} \hat{\mathbf{a}}(\theta_{i,\text{AOA}}, M) \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{S}_k - \mathbf{R}_{l,k} \right\|_2^2. \quad (49)$$

To simplify the problem, by performing the left multiplication of vector $\hat{\mathbf{a}}^\dagger(\theta_{i,\text{AOA}}, M)$ by matrix $\mathbf{R}_{l,k}$, the problem can be simplified to

$$\min_{\alpha_{i,k}, \theta_{k,\text{AOD}}} h(\alpha_{i,k}, \theta_{k,\text{AOD}}) = \left\| \alpha_{i,k} \mathbf{S}_k \mathbf{a}(\theta_{k,\text{AOD}}, N) - \mathbf{r}_{i,k} \right\|_2^2, \quad (50)$$

where $\mathbf{r}_{i,k}^H = \hat{\mathbf{a}}^\dagger(\theta_{i,\text{AOA}}, M) \mathbf{R}_{i,k}$. Due to the non-convexity of $\mathbf{a}(\theta_{k,\text{AOD}}, N)$, we employ a gradient-based trust-region algorithm for estimation. Specifically, we use a second-order Taylor expansion $\tilde{h}$ to approximate h , i.e.,

$$\begin{aligned} \tilde{h}(\mathbf{y}^r) &= h(\mathbf{y}^{r-1}) + (\mathbf{g}^r)^H (\mathbf{y}^r - \mathbf{y}^{r-1}) \\ &\quad + (\mathbf{y}^r - \mathbf{y}^{r-1})^H \mathbf{G}^r (\mathbf{y}^r - \mathbf{y}^{r-1}), \end{aligned} \quad (51)$$

where $\mathbf{y}^r = [\alpha_{i,k}^r, \theta_{k,\text{AOD}}^r]$, $\mathbf{g}^r$ and $\mathbf{G}^r$ represent the value of the optimization variable, gradient and Hessian matrix at the r -th iteration, respectively. For the r -th iteration, the subproblem to be solved can be expressed as

$$(\text{P4}) \quad \min_{\mathbf{y}^r} \tilde{h}(\mathbf{y}^r) \quad \text{s.t.} \quad \left\| \mathbf{y}^r - \mathbf{y}^{r-1} \right\|_2^2 \leq l^r, \quad (52)$$

where l^r represents the upper bound of the search step size at the r -th iteration. This problem is a typical and common quadratic optimization problem, and here we omit the detailed solution process. After solving the subproblem, the approximation degree of the Taylor expansion s^r is computed to

$$\begin{aligned} \mathcal{L}_2(\mathbf{v}_i, \mathbf{P}_k, \mathbf{X}_{i,k}, \mathbf{G}_{i,k}, \mathbf{Z}_{i,k}) &= \sum_{k=1}^K e_{i,k} [\text{tr}(\mathbf{T}(\mathbf{v}_i)) + \text{tr}(\mathbf{P}_k) + \lambda \|\mathbf{Z}_{i,k} - \mathbf{R}_{i,k} \mathbf{S}_k^\dagger\|_2^2] \\ &\quad \left\langle \mathbf{G}_{i,k}, \left(\mathbf{X}_{i,k} - \begin{bmatrix} \mathbf{T}(\mathbf{v}_i) & \mathbf{Z}_{i,k} \\ \mathbf{Z}_{i,k}^H & \mathbf{P}_k \end{bmatrix} \right) \right\rangle + \rho \left\| \mathbf{X}_{i,k} - \begin{bmatrix} \mathbf{T}(\mathbf{v}_i) & \mathbf{Z}_{i,k} \\ \mathbf{Z}_{i,k}^H & \mathbf{P}_k \end{bmatrix} \right\|_2^2, \end{aligned} \quad (48)$$

Algorithm 2 Proposed AOD estimation Algorithm.

1: **Input:** Initialize $\mathbf{y}^0$, l^1 , threshold ε_3 and iteration index $r = 1$ for outer loop
2: **repeat:** inner loop
3: Update $\mathbf{y}^r$ by solving (P4).
4: Update $\mathbf{s}^r$ according to (47).
5: **if** $s^r > 0.75$ **then**
6: $l^r = 2l^r$.
7: **else**
8: $l^r = \frac{1}{4} \|\mathbf{y}^r - \mathbf{y}^{r-1}\|_2^2$.
9: **end if**
10: $r = r + 1$.
11: **until** $|h(\mathbf{y}^r) - h(\mathbf{y}^{r-1})|/h(\mathbf{y}^{r-1}) \leq \varepsilon_3$.
12: **Output:** $\hat{\theta}_{k,\text{AOD}} = [\mathbf{y}^r]_2$

determine the range of the search step size upper bound for the next iteration, which can be expressed as

$$s^r = \frac{h(\mathbf{y}^r) - h(\mathbf{y}^{r-1})}{\bar{h}(\mathbf{y}^r) - \bar{h}(\mathbf{y}^{r-1})}. \quad (53)$$

The updating and solving process is repeated until convergence, and the specific algorithm flow is shown in **Algorithm 2**.

3) $\text{rank}(\mathbf{S}_k) = 1$: The singular value decomposition (SVD) of matrix $\mathbf{S}_k$ is given by $\mathbf{S}_k = \mathbf{u}_k \Sigma_k \mathbf{v}_k^H$. Based on the result of the singular value decomposition and substituting it into (21), $\mathbf{R}_{i,k}$ can be expressed as

$$\mathbf{R}_{i,k} = \alpha_{i,k} \mathbf{a}(\theta_{i,\text{AOA}}, M) \mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{u}_k \sum_k \mathbf{v}_k^H + \mathbf{N}_{i,k} \quad (54)$$

where $\mathbf{N}_{i,k} = \mathbf{N}_i \mathbf{X}_k^H(t - \hat{\tau}_{i,k})$ represents the Gaussian white noise after matched filtering.

It is noted that when the vector $\mathbf{a}^H(\theta_{k,\text{AOD}}, N)$ containing the information of $\theta_{k,\text{AOD}}$ is multiplied by the left singular vector $\mathbf{u}_k$, the angular information is compressed into a scalar. At the same time, $\alpha_{i,k}$ is unknown, which results in the coupling of $\alpha_{i,k}$ and $\mathbf{a}^H(\theta_{k,\text{AOD}}, N) \mathbf{u}_k$, making $\theta_{k,\text{AOD}}$ unestimable. Therefore, we use the one-dimensional ADMM to estimate a , and the corresponding Lagrangian function is modified as shown in (55).

V. LOCATION ESTIMATION ALGORITHMS

In the previous section, the estimation of all time delay information $\boldsymbol{\tau} = [\tau_{1,1}, \dots, \tau_{K,K}] \in \mathbb{R}^{K^2 \times 1}$ and angle information $\boldsymbol{\theta} = [\theta_1, \dots, \theta_K] \in \mathbb{R}^{K \times 1}$ has been completed. Based on the geometric relationship shown in Fig. 2, the

relationship between the time delay, angle, and the coordinates to be estimated can be expressed as

$$-2x_{1,k}x_T - 2y_{1,k}y_T + x_T^2 + y_T^2 = \tau_k^2 c^2 - x_{1,k}^2 - y_{1,k}^2, \forall k, \quad (56)$$

$$x_T \tan \theta_k - y_T = x_{1,k} \tan \theta_k - y_{1,k}, \forall k. \quad (57)$$

We need to estimate (x_T, y_T) based on the cascade delay and angle estimates $\hat{\tau}$ and $\hat{\theta}$ according to (56) and (57). The estimation of (x_T, y_T) faces three challenges: first, the directly estimated data is the cascade delay rather than the segment delay while segment delays are needed for localization. Then (56) contains quadratic information (i.e., $x_T^2 + y_T^2$) about the location, and second, There exists estimation error in the coefficient matrix.

To address the above challenges, we propose a three-stage algorithm. In the first stage, we estimate the segment delays based on the cascade delay. Then, in the second stage, we ignore the information in the quadratic terms and use weighted least squares to process error in the coefficient matrix to obtain a rough estimate. Finally, we use the information in the quadratic terms to refine the rough estimate.

A. Stage 1

First, the segment delays τ_k and τ_l can be estimated based on the cascade delay estimate $\tau_{l,k}$. Specifically, the error in the delay estimate $\tau_{l,k}$ can be modeled as $\hat{\tau}_{l,k} = \tau_{l,k} + \varepsilon_{l,k}$, where $\tau_{l,k}$, $\hat{\tau}_{l,k}$, and $\varepsilon_{l,k}$ represent the true value, the true estimated value, and the estimation error of the cascade delay, respectively. Then, we construct the following equation to perform a weighted least squares estimation of τ_k .

$$\underbrace{\begin{bmatrix} 2 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 2 \end{bmatrix}}_{\mathbf{A}_1} \underbrace{\begin{bmatrix} \tau_1 \\ \tau_2 \\ \cdots \\ \tau_K \end{bmatrix}}_{\mathbf{x}_1} = \underbrace{\begin{bmatrix} \hat{\tau}_{1,1} \\ \hat{\tau}_{1,2} \\ \cdots \\ \hat{\tau}_{K,K} \end{bmatrix}}_{\mathbf{b}_1} + \underbrace{\begin{bmatrix} \varepsilon_{1,1} \\ \varepsilon_{1,2} \\ \cdots \\ \varepsilon_{K,K} \end{bmatrix}}_{\Delta \mathbf{b}_1}. \quad (58)$$

Thus, the estimate of $\mathbf{x}_1$ can be expressed as

$$\hat{\mathbf{x}}_1 = (\mathbf{A}_1^T \mathbf{Q}_1^{-1} \mathbf{A}_1)^{-1} \mathbf{A}_1^T \mathbf{Q}_1^{-1} \mathbf{b}_1, \quad (59)$$

Where $\mathbf{Q}_1$ is the autocorrelation matrix of the estimation error $\mathbf{b}_1$, which can be approximated through statistical averaging.

B. Stage 2

The relationship between the estimated and true values of the segment delays and angles can be modeled as

$$\tau_k = \hat{\tau}_k + \varepsilon_{\tau,k}, \forall k, \quad (60)$$

$$\theta_k = \hat{\theta}_k + \varepsilon_{\theta,k}, \forall k, \quad (61)$$

$$\begin{aligned} \mathcal{L}_3(\mathbf{v}_i, p_k, \mathbf{X}_{i,k}, \mathbf{G}_{i,k}, \mathbf{z}_{i,k}) = & \sum_{k=1}^K e_{i,k} [\text{tr}(\mathbf{T}(\mathbf{v}_i)) + p_k + \lambda \|\mathbf{z}_{i,k} - \mathbf{R}_{i,k} \mathbf{S}_k^\dagger\|_2^2] \\ & \left\langle \mathbf{G}_{i,k}, \left(\mathbf{X}_{i,k} - \begin{bmatrix} \mathbf{T}(\mathbf{v}_i) & \mathbf{z}_{i,k} \\ \mathbf{z}_{i,k}^H & p_k \end{bmatrix} \right) \right\rangle + \rho \left\| \mathbf{X}_{i,k} - \begin{bmatrix} \mathbf{T}(\mathbf{v}_i) & \mathbf{z}_{i,k} \\ \mathbf{z}_{i,k}^H & p_k \end{bmatrix} \right\|_2^2, \end{aligned} \quad (55)$$

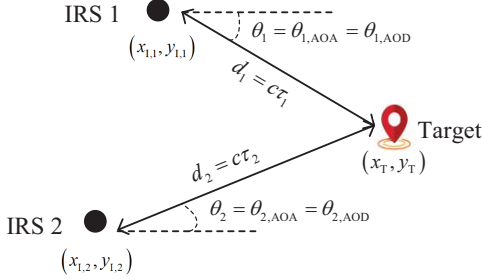


Fig. 2. The geometric relationship between delay, angle, and location.

where ε_k and ε_k represent the estimation errors of the distance for the k -th path and the angle for the k -th IRS, respectively. By substituting equations (60) and (61) into equations (56) and (57) and defining the variables in equation (74), we can establish the following system of equations:

$$\underbrace{\begin{bmatrix} \mathbf{A}_\tau \\ \mathbf{A}_\theta + \Delta \mathbf{A}_\theta \end{bmatrix}}_{\mathbf{A}_2} \underbrace{\begin{bmatrix} x_T \\ y_T \\ x_T^2 + y_T^2 \end{bmatrix}}_{\mathbf{x}_2} = \underbrace{\begin{bmatrix} \mathbf{b}_\tau \\ \mathbf{b}_\theta \end{bmatrix}}_{\mathbf{b}_2} + \underbrace{\begin{bmatrix} \Delta \mathbf{b}_\tau \\ \Delta \mathbf{b}_\theta \end{bmatrix}}_{\Delta \mathbf{b}_2}, \quad (62)$$

where $\Delta \mathbf{b}_\tau \sim \mathcal{CN}(\mathbf{0}_K, \text{diag}(c^2 \hat{\tau}_1^2 \sigma_1^2, \dots, c^2 \hat{\tau}_K^2 \sigma_K^2))$, $\Delta \mathbf{b}_\theta \sim \mathcal{CN}(\mathbf{0}_K, \text{diag}(x_{1,1}^2 \sigma_1^2 \sec^4 \hat{\theta}_1, \dots, x_{1,K}^2 \sigma_K^2 \sec^4 \hat{\theta}_K))$ and $\text{vec}(\Delta \mathbf{A}_\theta) \sim \mathcal{CN}(\mathbf{0}_{3K}, \text{diag}(x_{1,1}^2 \sigma_1^2 \sec^4 \hat{\theta}_1, \dots, x_{1,K}^2 \sigma_K^2 \sec^4 \hat{\theta}_K, \mathbf{0}_{2K}))$.

First, we perform a variable substitution by defining $s \triangleq x_T^2 + y_T^2$ and treating s as a new variable independent of x_T and y_T , so that the unknown terms become linear and the least squares estimation for $\mathbf{x}_2$ can be modeled as

$$\begin{aligned} \text{(P4)} \quad & \min_{\Delta \mathbf{a}_\theta, \Delta \mathbf{b}_\theta, \Delta \mathbf{b}_\tau} \Delta \mathbf{a}_\theta^T \mathbf{R}_{\Delta \mathbf{a}_\theta}^{-1} \Delta \mathbf{a}_\theta + \Delta \mathbf{b}_\theta^T \mathbf{R}_{\Delta \mathbf{b}_\theta}^{-1} \Delta \mathbf{b}_\theta \\ & + \Delta \mathbf{b}_\tau^T \mathbf{R}_{\Delta \mathbf{b}_\tau}^{-1} \Delta \mathbf{b}_\tau \\ \text{s.t.} \quad & \mathbf{A}_2 \mathbf{x}_2 = \mathbf{b}_2 + \Delta \mathbf{b}_2 \end{aligned} \quad (63)$$

The Lagrange function for P4 can be expressed as

$$\begin{aligned} \mathcal{L}_4(\boldsymbol{\lambda}_\theta^T, \boldsymbol{\lambda}_\tau^T, \Delta \mathbf{a}_\theta, \Delta \mathbf{b}_\theta, \Delta \mathbf{b}_\tau, \mathbf{x}_2) = & \boldsymbol{\lambda}_\tau^T (\mathbf{A}_\tau \mathbf{x}_2 - \mathbf{b}_\tau - \Delta \mathbf{b}_\tau) \\ & + \boldsymbol{\lambda}_\theta^T ((\mathbf{A}_\theta + [\Delta \mathbf{a}_\theta \quad \mathbf{0}_{K \times 2}]) \mathbf{x}_2 - \mathbf{b}_\theta - \Delta \mathbf{b}_\theta) \\ & + \Delta \mathbf{a}_\theta^T \mathbf{R}_{\Delta \mathbf{a}_\theta}^{-1} \Delta \mathbf{a}_\theta + \Delta \mathbf{b}_\theta^T \mathbf{R}_{\Delta \mathbf{b}_\theta}^{-1} \Delta \mathbf{b}_\theta + \Delta \mathbf{b}_\tau^T \mathbf{R}_{\Delta \mathbf{b}_\tau}^{-1} \Delta \mathbf{b}_\tau, \end{aligned} \quad (64)$$

where $\mathbf{R}_{\Delta \mathbf{a}_\theta}$ and $\mathbf{R}_{\Delta \mathbf{b}}$ and represent the variances of $\text{vec}(\Delta \mathbf{A}_\theta)$ and $\Delta \mathbf{b}$ respectively. By taking the partial deriva-

tives with respect to each variable and setting the results to zero, we can obtain the following results.

$$\Delta \mathbf{b}_\tau = \mathbf{R}_{\Delta \mathbf{b}_\tau} \boldsymbol{\lambda}_\tau / 2, \quad (65)$$

$$\Delta \mathbf{b}_\theta = \mathbf{R}_{\Delta \mathbf{b}_\theta} \boldsymbol{\lambda}_\theta / 2, \quad (66)$$

$$\Delta \mathbf{a}_\theta = -\mathbf{e}_3^T \mathbf{x}_2 \mathbf{R}_{\Delta \mathbf{a}_\theta} \boldsymbol{\lambda}_\theta / 2, \quad (67)$$

$$\mathbf{A}_\tau \mathbf{x}_2 - \mathbf{b}_\tau - \Delta \mathbf{b}_\tau = 0, \quad (68)$$

$$\mathbf{A}_\theta \mathbf{x}_2 + \Delta \mathbf{a}_\theta \mathbf{e}_3^T \mathbf{x}_2 - \mathbf{b}_\theta - \Delta \mathbf{b}_\theta = 0, \quad (69)$$

$$\mathbf{A}_\theta \boldsymbol{\lambda}_\theta + \boldsymbol{\lambda}_\theta^T \Delta \mathbf{a}_\theta \mathbf{e}_3 + \mathbf{A}_\tau \boldsymbol{\lambda}_\tau = 0. \quad (70)$$

By substituting equation (65) into (68) and equations (66) and (67) into (69), we can obtain

$$\boldsymbol{\lambda}_\tau = 2 \mathbf{R}_{\Delta \mathbf{b}_\tau}^{-1} (\mathbf{A}_\tau \mathbf{x} - \mathbf{b}_\tau), \quad (71)$$

$$\boldsymbol{\lambda}_\theta = 2 (\mathbf{e}_3^T \mathbf{x}_2 \mathbf{e}_3^T \mathbf{x}_2 \mathbf{R}_{\Delta \mathbf{a}_\theta} + \mathbf{R}_{\Delta \mathbf{b}_\theta})^{-1} (\mathbf{A}_\theta \mathbf{x}_2 - \mathbf{b}_\theta). \quad (72)$$

For the convenience of subsequent processing, we define $\tilde{x} \triangleq \mathbf{e}_3^T \mathbf{x}_2 \mathbf{e}_3^T \mathbf{x}_2$. By substituting equation (71) and (72) into (70), the estimate of $\mathbf{x}_2$ can be expressed as

$$\hat{\mathbf{x}}_2 = (\mathbf{B}_\theta \mathbf{A}_\theta + \mathbf{A}_\tau \mathbf{R}_{\Delta \mathbf{b}_\tau}^{-1} \mathbf{A}_\tau)^{-1} (\mathbf{B}_\theta \mathbf{b}_\theta + \mathbf{A}_\tau \mathbf{R}_{\Delta \mathbf{b}_\tau}^{-1} \mathbf{b}_\tau), \quad (73)$$

where $\mathbf{B}_\theta = (\mathbf{A}_\theta - [\Delta \mathbf{a}_\theta \quad \mathbf{0}_{K \times 2}]) (\tilde{x} \mathbf{R}_{\Delta \mathbf{a}_\theta} + \mathbf{R}_{\Delta \mathbf{b}_\theta})^{-1}$. We update the corresponding variables using an iterative method until the estimation results converge. The initial estimate can be obtained using least squares estimation, i.e., $\hat{\mathbf{x}}_{\text{LS}} = (\mathbf{A}_2^T \mathbf{A}_2)^{-1} \mathbf{A}_2^T \mathbf{b}_2$.

C. Stage 3

After completing the second stage of estimation, the error between the estimated values and the true values can be modeled as

$$x_T = \hat{x}_{T,2} + \varepsilon_x, \quad (75)$$

$$y_T = \hat{y}_{T,2} + \varepsilon_y, \quad (76)$$

$$x_T^2 + y_T^2 = \hat{s} + \varepsilon_s, \quad (77)$$

where $\hat{x}_{T,2}$, $\hat{y}_{T,2}$, and $\hat{s}$ represent the estimations of x_T , y_T , and $x_T^2 + y_T^2$ based on the second stage of estimation and ε_x , ε_y , and ε_s represent the corresponding estimation errors.

In the third stage, we utilize the location information contained in s to refine the solution obtained in the first stage. This leads to the following system of equations

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{A}_3} \underbrace{\begin{bmatrix} x_T^2 \\ y_T^2 \\ s \end{bmatrix}}_{\mathbf{x}_3} = \underbrace{\begin{bmatrix} \hat{x}_{T,2}^2 \\ \hat{y}_{T,2}^2 \\ \hat{s} \end{bmatrix}}_{\mathbf{b}_3} + \underbrace{\begin{bmatrix} 2x_T \varepsilon_x \\ 2y_T \varepsilon_y \\ \varepsilon_s \end{bmatrix}}_{\Delta \mathbf{b}_3}. \quad (78)$$

$$\begin{aligned}
\mathbf{A}_\tau &\triangleq \begin{bmatrix} -2x_{1,1} & -2y_{1,1} & 1 \\ \cdots & \cdots & \cdots \\ -2x_{1,K} & -2y_{1,K} & 1 \end{bmatrix}, \mathbf{A}_\theta \triangleq \begin{bmatrix} \tan \hat{\theta}_1 & -1 & 0 \\ \cdots & \cdots & \cdots \\ \tan \hat{\theta}_K & -1 & 0 \end{bmatrix}, \Delta \mathbf{A}_\theta \triangleq [\Delta \mathbf{a}_\theta \quad \mathbf{0}_{K \times 2}] = \begin{bmatrix} \tan \varepsilon_1 & 0 & 0 \\ \cdots & \cdots & \cdots \\ \tan \varepsilon_K & 0 & 0 \end{bmatrix}, \\
\mathbf{b}_\tau &\triangleq [\hat{\tau}_1^2 c^2 - (x_{1,1}^2 + y_{1,1}^2), \cdots, \hat{\tau}_K^2 c^2 - (x_{1,K}^2 + y_{1,K}^2)], \Delta \mathbf{b}_\tau \triangleq [2\tau_1 \varepsilon_1 c^2, \cdots, 2\tau_K \varepsilon_K c^2], \\
\mathbf{b}_\theta &\triangleq [x_{1,1} \tan \hat{\theta}_1 - y_{1,1}, \cdots, x_{1,K} \tan \hat{\theta}_K - y_{1,K}], \Delta \mathbf{b}_\theta \triangleq [x_{1,1} \tan \varepsilon_1, \cdots, x_{1,K} \tan \varepsilon_K],
\end{aligned} \tag{74}$$

Algorithm 3 Proposed Three-Stage Localization Algorithm.

- 1: **Input:** Initialize $\hat{\tau}$ and $\hat{\theta}$ threshold $\varepsilon_2, \varepsilon_3$ and iteration index $r = 1$
 - 2: **Stage 1:**
 - 3: Update $\hat{\mathbf{x}}_1$ according to (59).
 - 4: **Stage 2:**
 - 5: **repeat:**
 - 6: Update $\tilde{x}$ according to its definition.
 - 7: Update $\hat{\mathbf{x}}_2$ according to (73).
 - 8: **until** $\|\hat{\mathbf{x}}_2^{r+1} - \hat{\mathbf{x}}_2^r\|_2^2 \leq \varepsilon_3$.
 - 9: **Stage 3:**
 - 10: $\hat{x}_{T,2} = [\hat{\mathbf{x}}_2]_1, \hat{y}_{T,2} = [\hat{\mathbf{x}}_2]_2, \hat{s} = [\hat{\mathbf{x}}_2]_3$.
 - 11: Calculate $\hat{\mathbf{x}}_3$ according to (79).
 - 12: $\hat{x}_{T,3} = \pm \sqrt{[\hat{\mathbf{x}}_3]_1}, \hat{y}_{T,3} = \pm \sqrt{[\hat{\mathbf{x}}_3]_2}$ and the signs are consistent with $\hat{x}_{T,2}, \hat{y}_{T,2}$.
 - 13: **Output:** $\hat{\mathbf{x}} = [\hat{x}_{T,3}, \hat{y}_{T,3}]$.
-

and the least squares estimation for $\mathbf{x}_3$ can be expressed as

$$\hat{\mathbf{x}}_3 = (\mathbf{A}_3^T \mathbf{Q}_3^{-1} \mathbf{A}_3)^{-1} \mathbf{A}_3^T \mathbf{Q}_3^{-1} \mathbf{b}_3 \tag{79}$$

where $\mathbf{Q}_3 = \text{diag}(4\hat{x}_T^2, 4\hat{y}_T^2, \hat{s}) \text{conv}(\hat{\mathbf{x}}_2) \text{diag}(4\hat{x}_T^2, 4\hat{y}_T^2, \hat{s})$. After obtaining the estimation $\hat{\mathbf{x}}_3$ from the second stage, we take the square root of $\hat{x}_T^2$ and $\hat{y}_T^2$ in $\hat{\mathbf{x}}_3$, ensuring that the signs are consistent with those in the first stage, and then complete the final location estimation. and the entire three-stage localization algorithm is summarized in **Algorithm 3**.

VI. SIMULATION RESULTS

In this section, we validate the superiority of the proposed multi-IRS collaborative localization model and the accuracy of the angle and location estimation algorithms through numerical simulations. We consider a two-dimensional coordinate system in this section. The BS is located at (0, 0) m while $K = 3$ IRSs are located at (10, 50) m, (10, -50) m, and (50, 0) m. A single target is randomly distributed within the rectangular area defined by $0 < x < 10$ m and $0 < y < 10$ m. A key aspect of our simulation setup is the configuration of the IRS elements. In our simulations, we adopt a random phase configuration, where the phase shift of each element is chosen independently and uniformly from $[0, 2\pi)$. In the setup, each IRS is a Uniform Linear Array (ULA) placed vertically with its broadside oriented towards the origin. The target's Radar Cross Section (κ), listed in Table I, is modeled as a constant value independent of the viewing angle. Other main system parameters are listed in Table I.

TABLE I
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value	Parameter	Value
N_t	50	N	10	M	10
W	50 MHz	L	100	κ	7 dBsm
$P_{\max}$	50 dBmW	σ_l^2	-100 dBm	λ	0.3m

A. CRB Performance for Localization System

The accuracy of the localization system within the region is characterized by calculating the average CRB of localization estimation of multiple random target, i.e.,

$$\mathbf{C}_{\text{ave}} = \frac{1}{N_{\text{trials}}} \sum_{i=1}^{N_{\text{trials}}} \mathbf{C}(\mathbf{I}_{T,i}), \tag{80}$$

where we set N_{trials} is the total number of Monte Carlo trials, set to 100; $\mathbf{C}(\mathbf{x}_{T,i})$ is the CRB for a single target at a random location $\mathbf{I}_{T,i}$. We compare the proposed system design with the following schemes:

- **Angle-based:** The system only estimates the angle and performs localization using the angle information.
- **Delay-based:** The system only estimates the delay and performs localization using the delay information.
- **No collaboration:** All semi-passive IRSs employ frequency-division or code-division techniques for the independent reflection of incident signals and reception of echo signals, thereby facilitating independent localization.
- **Single IRS:** All the reflective elements and sensors are deployed at the location of the first IRS.

1) *CRB Versus Number of Reflective Elements:* In Fig. 3, we compare the average CRB for location estimation versus the number of reflective elements. First, the performance of all schemes improves as the number of reflective elements increases, which can be attributed to two main reasons. One reason is that the transmitter units provide beamforming gain of $\mathcal{O}(N)$ (in this model, we do not consider the optimal IRS coefficient design, which could provide an $\mathcal{O}(N^2)$ gain. Our model assumes random IRS coefficients, resulting in an $\mathcal{O}(N)$ gain), which positively affects the estimation of delay, AOA, and AOD. The CRB scales down with N approximately in the the order of $1/N$. Another reason is that for AOD estimation, the increase in reflective elements leads to a higher dimensionality of the corresponding steering vector, thereby enhancing spatial resolution.

2) *CRB Versus Number of Sensors:* In Fig. 4, we compare the CRB for location estimation of the single target versus the number of transmitting antennas, Based on Fig. 3 and Fig.

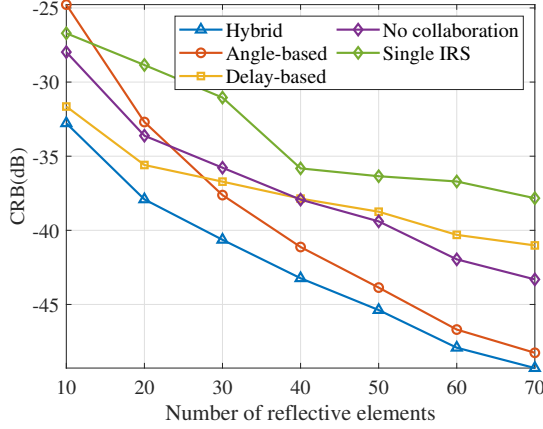


Fig. 3. CRB versus the number of reflective elements.

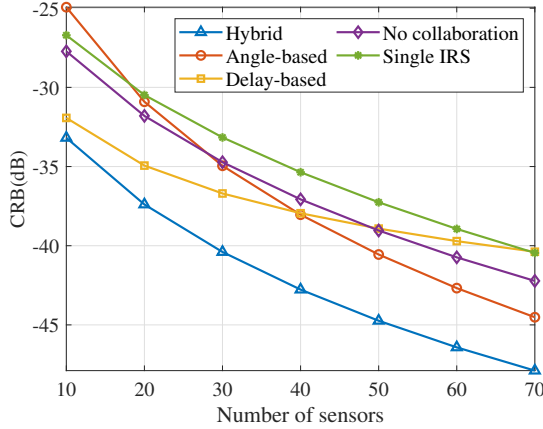


Fig. 4. CRB versus the number of sensors.

4, we observe that multiple semi-passive IRS collaborative localization can improve accuracy by approximately 7 dB, and this value remains unchanged regardless of the number of reflective elements or sensors. We also find that when the number of reflective elements or sensors is low, the delay information plays a dominant role in localization, while at higher numbers of reflective elements or sensors, the angle information becomes dominant. This is because, in the case of random IRS coefficients, the Fisher information provided by delay information is approximately $\mathcal{O}(MN)$, while the Fisher information provided by AOA and AOD is approximately $\mathcal{O}(M^3N)$ and $\mathcal{O}(MN^3)$, respectively. Therefore, for the design of IRS-assisted localization systems, the choice of direct estimators can be determined based on the number of reflective elements, the number of sensors, and the bandwidth W of the sensing signal. This choice will influence the complexity of signal processing. Finally, we observe a significant performance difference between traditional single-IRS and multi-IRS systems, which remains unchanged regardless of the number of reflective elements or sensors. To investigate the main reasons behind this difference, we will explore the impact of distance in the next subsection.

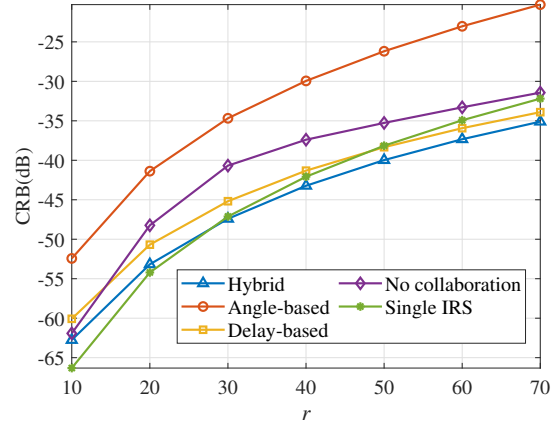


Fig. 5. CRB versus the coverage area.

3) *CRB Versus Coverage area:* In Fig. 5, we change the target area for localization to a circle with the first semi-passive IRS as the center and a radius of r and compare the average CRB for location estimation versus r . It is observed that When r is less than 30m, the performance of the single-IRS system is the best. However, as r increases, the performance advantage shifts to the multi-IRS system. This is because at larger distances, the benefit of mitigating path loss through distributed deployment outweighs the drawback of splitting system resources. This observation reveals the distinct application scenarios for each deployment strategy: a single-IRS deployment is more suitable for localization tasks in "hotspot" areas where the target is relatively static and at a close range, whereas a multi-IRS collaborative deployment is better adapted for localization over larger areas, significantly enhancing the system's coverage and robustness. Furthermore, it is interesting to note that the collaborative gain varies with r and does not remain the same as in the previous two simulations. This is because, as r changes, the degree to which multiple IRSs are involved in the system also changes. Specifically, when the distance between a certain IRS and the target is much smaller than the distances of the other IRSs to the target, that IRS dominates the localization, resulting in a lower collaborative gain. On the other hand, when the distances from all IRSs to the target are roughly the same, the collaborative gain is maximized. Furthermore, the collaborative gain increases as the number of IRSs increases, because cooperative signal processing can expand the number of observation samples from K to K^2 .

B. Performance of Angle Estimation Algorithm

In this section, we select the reflection signal from the IRS with $\mathbf{S}_k = N$ for simulation. The algorithm proposed in Sections IV is compared with several classical angle estimation algorithms, and the MSE of K angles is calculated to verify the accuracy of the angle estimation.

1) *MSE Versus transmitting power at the BS:* Fig. 6 shows the average MSE versus transmitting power at the BS. It can be observed that the estimation error decreases linearly with increasing power. The proposed estimation algorithm achieves

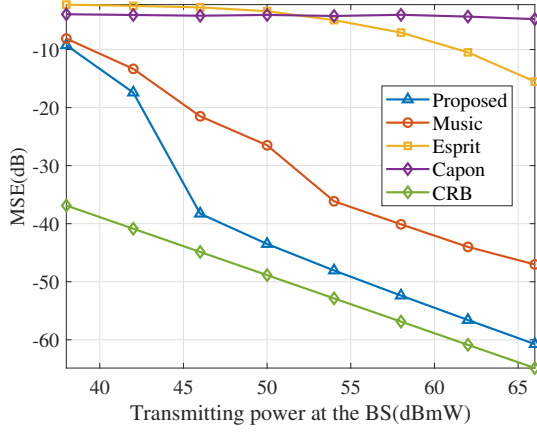


Fig. 6. MSE versus the transmitting power at the BS.

the best performance compared to other benchmarks, with only a 4 dB gap from the CRB. The Capon algorithm is sensitive to noise and thus fails to accurately estimate angles under the SNR conditions set in the simulation. The ESPRIT algorithm can perform angle estimation only at relatively high SNR levels. Although the MUSIC algorithm is capable of estimation under lower SNR, it exhibits a gap of more than 10 dB from the CRB. Since this paper does not consider the design of optimal IRS coefficients, the required base station transmission power is relatively high. In practice, with proper beamforming design, the required transmission power can be significantly reduced.

C. Performance of Localization Estimation Algorithm

We generate error-prone estimates of $\hat{\tau}$ and $\hat{\theta}$ with variances corresponding to the derived CRB of the cascaded delay and angle. Then, the estimation of based on $\hat{\tau}$ and $\hat{\theta}$ is completed using the three-stage algorithm proposed in Section V, and compared with the following schemes.

- **Two-stage:** Only the first two stages of the three-stage algorithm are used, meaning the information in the second-order terms is ignored.
- **WLS:** After estimating the segmented delay in the first stage, WLS estimation is applied to estimate the location.
- **LS:** After estimating the segmented delay in the first stage, LS estimation is applied to estimate the location.
- **CRB:** The CRB derived in Section III.

1) *MSE Versus transmitting power at the BS:* Fig. 7 shows the average MSE versus transmitting power at the BS. The MSE differences among different estimation schemes are significant. Specifically, the LS estimation has the highest MSE because it completely ignores the variance information of the errors in the coefficient matrix $\mathbf{A}_2$ and the observation sample $\mathbf{b}_2$, as well as the quadratic information related to the location contained in the observation samples. Compared to LS, WLS takes into account the variance information of the errors in the observation sample $\mathbf{b}_2$, leading to an approximately 4 dB reduction in the estimated MSE. Building upon this, our proposed algorithm considers the errors in the coefficient

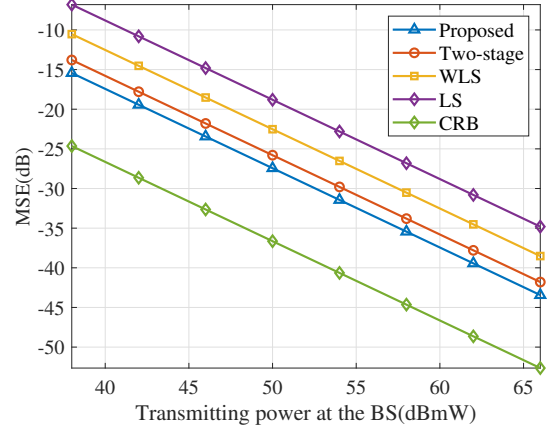


Fig. 7. MSE versus the transmitting power at the BS.

matrix $\mathbf{A}_2$ in the second stage and the quadratic information of the location included in $\mathbf{b}_2$ in the third stage, further reducing the estimation MSE by 4 dB and 2 dB, respectively. Finally, we found that the performance improvement gained by addressing the three major challenges not considered in LS estimation remains unaffected by changes in the signal-SNR. Therefore, in practical locating algorithms, these challenges should be addressed in both high and low SNR scenarios.

2) *MSE Versus the distance:* In this subsection, we explore the impact of IRS deployment on estimation performance by changing the coordinates of the first two IRS location from $(10, -50)$ and $(10, 50)$ to $(d, -50)$ and $(d, 50)$, respectively. Fig. 8 shows the average MSE versus d . It can be observed that only the estimation error of the proposed scheme experiences minimal fluctuations, similar to the CRB, as the IRS deployment changes within the given range. When d is small, the error of the WLS algorithm increases because the path loss of the system signal increases as d decreases, particularly when $d < 10$. Specifically, the horizontal distance between the first two IRSs and the base station is d , and the horizontal distance between the first two IRSs and the target is $x - d$. The echoed signal travels through the latter distance twice. The necessity of the third-stage algorithm increases as d increases. This is because the variance of the quadratic term estimation related to the location decreases as the absolute value of the IRS coordinates increases. When the variance of the quadratic term estimation becomes smaller, the correction of the coarse estimate by the third-stage algorithm becomes more accurate.

VII. CONCLUSIONS

This paper presents a novel collaborative hybrid localization system using multi-IRS to estimate target location through joint time delay and angle estimation. Echo signals from IRS reflective elements are processed by sensors to estimate time delay and angle parameters. The FIM and CRB are derived for delay, AOA, and AOD estimation in. Efficient algorithms for angle and location estimation are designed, including a convex optimization approach using ADMM for AOA/AOD estimation and a three-stage algorithm combining least squares methods for location estimation. Simulations show that the

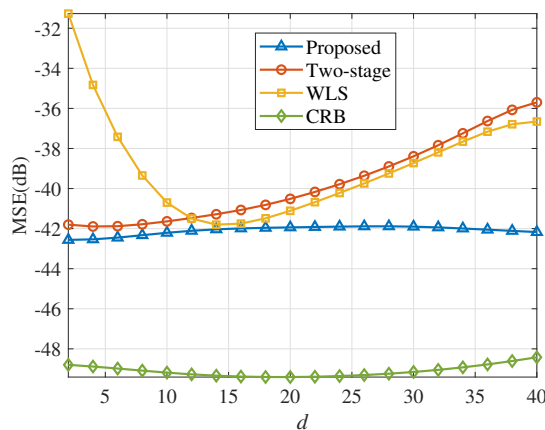


Fig. 8. CRB versus the distance.

system provides significant benefits, especially in low signal-to-noise ratio conditions.

REFERENCES

- [1] Z. Wei, H. Qu, Y. Wang, X. Yuan, H. Wu, Y. Du, K. Han, N. Zhang, and Z. Feng, "Integrated sensing and communication signals toward 5g-a and 6g: A survey," *IEEE Internet Things J.*, vol. 10, no. 13, pp. 11 068–11 092, Jan. 2023.
- [2] X. Cheng, H. Zhang, J. Zhang, S. Gao, S. Li, Z. Huang, L. Bai, Z. Yang, X. Zheng, and L. Yang, "Intelligent multi-modal sensing-communication integration: Synesthesia of machines," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 1, pp. 258–301, Nov. 2024.
- [3] K. Meng, Q. Wu, J. Xu, W. Chen, Z. Feng, R. Schober, and A. L. Swindlehurst, "Uav-enabled integrated sensing and communication: Opportunities and challenges," *IEEE Wireless Communications*, vol. 31, no. 2, pp. 97–104, Apr. 2024.
- [4] S. Pala, K. Singh, C.-P. Li, and O. A. Dobre, "Empowering isac systems with federated learning: A focus on satellite and ris-enhanced terrestrial integrated networks," *IEEE Trans. Wireless Commun.*, vol. 24, no. 1, pp. 810–824, Nov. 2025.
- [5] F. Dong, F. Liu, Y. Cui, W. Wang, K. Han, and Z. Wang, "Sensing as a service in 6g perceptive networks: A unified framework for isac resource allocation," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3522–3536, Nov. 2023.
- [6] M. Hua, Q. Wu, W. Chen, A. Jamalipour, C. Wu, and O. A. Dobre, "Integrated sensing and communication: Joint pilot and transmission design," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 16 017–16 032, Aug. 2024.
- [7] H. Qin, W. Chen, Z. Li, Q. Wu, N. Cheng, and F. Chen, "Antenna positioning and beamforming design for fluid antenna-assisted multi-user downlink communications," *IEEE Wireless Commun.Lett.*, vol. 13, no. 4, pp. 1073–1077, Jan. 2024.
- [8] Q. Wu, B. Zheng, C. You, L. Zhu, K. Shen, X. Shao, W. Mei, B. Di, H. Zhang, E. Basar, L. Song, M. Di Renzo, Z.-Q. Luo, and R. Zhang, "Intelligent surfaces empowered wireless network: Recent advances and the road to 6g," *Proc. IEEE*, vol. 112, no. 7, pp. 724–763, 2024.
- [9] X. Shao, C. You, W. Ma, X. Chen, and R. Zhang, "Target sensing with intelligent reflecting surface: Architecture and performance," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2070–2084, Mar. 2022.
- [10] Z. Wang, X. Hu, C. Liu, and M. Peng, "RIS-enabled multi-target sensing: Performance analysis and space-time beamforming design," *IEEE Trans. Wireless Commun.*, pp. 1–1, Jun. 2024.
- [11] Y. Lin, S. Jin, M. Matthaiou, and X. You, "Tensor-based algebraic channel estimation for hybrid IRS-assisted MIMO-OFDM," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3770–3784, Jan. 2021.
- [12] H. Yan, H. Chen, W. Liu, S. Yang, G. Wang, and C. Yuen, "Ris-enabled joint near-field 3d localization and synchronization in siso multipath environments," *IEEE Trans. Green Commun. Netw.*, vol. 9, no. 1, pp. 367–379, 2025.
- [13] X. Mu, L. Zhou, H. Fu, and J. Dai, "Target localization for irs-assisted massive mimo systems," *IEEE Sens. J.*, vol. 23, no. 23, pp. 29 260–29 270, Jul. 2023.
- [14] P. Gao, L. Lian, and J. Yu, "Efficiency of spatial correlation for multi-ris-assisted multi-user direct localization," in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 2741–2746.
- [15] Y. Han, S. Jin, C.-K. Wen, and T. Q. S. Quek, "Localization and channel reconstruction for extra large ris-assisted massive mimo systems," *IEEE J. Sel. Top. Sign. Process.*, vol. 16, no. 5, pp. 1011–1025, Aug. 2022.
- [16] T. Ji, M. Hua, X. Yan, C. Li, Y. Huang, and L. Yang, "Intelligent reflecting surface aided target localization with unknown transceiver-irs channel state information," *IEEE Trans. Commun.*, pp. 1–1, Apr. 2024.
- [17] S. Zhao, Y. Liu, L. Wu, J. Rodríguez-Piñeiro, X. Yin, and J. Hong, "Recursive UE localization for a multi-ris-assisted wireless system in an obstacle-dense environment," in *2023 17th European Conference on Antennas and Propagation (EuCAP)*, 2023, pp. 1–5.
- [18] Z. Esmaeilbeig, K. V. Mishra, A. Eamaz, and M. Soltanalian, "Cramér-rao lower bound optimization for hidden moving target sensing via multi-IRS-aided radar," *IEEE Signal Processing Lett.*, vol. 29, pp. 2422–2426, Nov. 2022.
- [19] X. Shao and R. Zhang, "Target-mounted intelligent reflecting surface for secure wireless sensing," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9745–9758, Feb. 2024.
- [20] H. Wang, B. Zheng, X. Shao, and R. Zhang, "Intelligent reflecting surface-aided radar spoofing," *IEEE Wireless Commun.Lett.*, vol. 13, no. 10, pp. 2722–2726, Aug. 2024.
- [21] B. Zheng, X. Xiong, J. Tang, and R. Zhang, "Intelligent reflecting surface-aided electromagnetic stealth against radar detection," *IEEE IEEE Trans. Signal Process.*, vol. 72, pp. 3438–3452, Jun. 2024.
- [22] Y. Zhang, X. Shao, H. Li, B. Clerckx, and R. Zhang, "Full-space wireless sensing enabled by multi-sector intelligent surfaces," *arXiv preprint arXiv:2406.15945*, 2024.
- [23] X. Song, X. Li, X. Qin, and J. Xu, "Fully-passive versus semi-passive irs-enabled sensing: Snr analysis," in *2023 IEEE Globecom Workshops (GC Wkshps)*, 2023, pp. 1874–1879.
- [24] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, "Metaradar: Multi-target detection for reconfigurable intelligent surface aided radar systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 6994–7010, Sept. 2022.
- [25] G. Zhan, B. Zheng, X. Xiong, and M. Wen, "Multi-irs-assisted wireless localization," *IEEE Wireless Commun.Lett.*, pp. 1–1, Jan. 2025.
- [26] Z. Zhang, W. Chen, Q. Wu, Z. Li, X. Zhu, and J. Yuan, "Intelligent omni surfaces assisted integrated multi-target sensing and multi-user MIMO communications," *IEEE Trans. Commun.*, vol. 72, no. 8, pp. 4591–4606, Aug. 2024.
- [27] Y. Fang, S. Zhang, X. Li, X. Yu, J. Xu, and S. Cui, "Multi-IRS-enabled integrated sensing and communications," *IEEE Trans. Commun.*, pp. 1–1, Apr. 2024.
- [28] S. Lyu, X. Hu, C. Liu, and M. Peng, "IRS-aided location sensing and beamforming in ISAC systems: Distributed or not?" *IEEE Trans. Veh. Technol.*, pp. 1–15, Apr. 2024.
- [29] Y. Fang, X. Yu, J. Xu, and Y.-J. A. Zhang, "Multi-active-irs-assisted cooperative sensing: Cramér-Rao bound and joint beamforming design," *IEEE Trans. Wireless Commun.*, pp. 1–1, Jul. 2025.
- [30] Q. Li, K. Huang, Z. Deng, Z. Wan, and M. Yi, "A multi-RIS-assisted static environment sensing method," *IEEE Commun. Lett.*, vol. 28, no. 6, pp. 1317–1321, Jun. 2024.
- [31] R. Li, S. Sun, and M. Tao, "Atomic norm minimization-based DoA estimation for IRS-assisted sensing systems," *IEEE Wireless Commun.Lett.*, vol. 13, no. 10, pp. 2672–2676, Oct. 2024.
- [32] P. Chen, Z. Chen, P. Miao, and Y. Chen, "RIS-ADMM: A RIS and ADMM-based passive and sparse sensing method with interference removal," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 867–871, Apr. 2024.
- [33] K. Huang, Q. Li, S. Ren, Z. Wan, and M. Yi, "A RIS-assisted MUSIC-like single-channel blind direction-of-arrival estimation method," *IEEE Wireless Commun.Lett.*, vol. 13, no. 8, pp. 2190–2194, May 2024.
- [34] Z. Zhang, W. Chen, Q. Wu, Z. Li, X. Zhu, J. Chen, and N. Cheng, "Multiple intelligent reflecting surfaces collaborative wireless localization system," *IEEE Trans. Wireless Commun.*, vol. 24, no. 1, pp. 134–148, Nov. 2025.
- [35] T. Lin, X. Yu, Y. Zhu, and R. Schober, "Channel estimation for IRS-assisted millimeter-wave MIMO systems: Sparsity-inspired approaches," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4078–4092, Jun. 2022.
- [36] H. Godrich, A. M. Haimovich, and R. S. Blum, "Target localization accuracy gain in MIMO radar-based systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2783–2803, May 2010.
- [37] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," in *Academic Press Library in Signal Processing, Volume 7*. Elsevier, 2018, pp. 509–581.