Near-Field Low-Complexity Sensing for Reconfigurable Holographic Surfaces with Hardware Constraints

Yinuo Dong, Qingchao Li (Member, IEEE), Soon Xin Ng, (Senior Member, IEEE), Mohammed El-Hajjar, (Senior Member, IEEE)

***Abstract*—Reconfigurable holographic surfaces (RHS) have emerged as a promising enablers for Integrated Sensing and Communication (ISAC) owing to their ultra-dense electromagnetic control capability. In this work, we consider near-field multi-user localization as a key ISAC function. However, hardware-induced phase-shift constraints and complex-plane coupling effects severely degrade the performance of classical subspace-based estimators such as multiple signal classification (MUSIC). In this work, we analytically characterize the impact of practical phase-shift constraints on the MUSIC spectrum and reveal the resulting pseudo-peak and angular bias phenomena. To address these challenges, we propose a hierarchical sensing framework that integrates subspace estimation with sparsity-aware recovery. Specifically, we introduce a regularized orthogonal matching pursuit (ROMP)-assisted multi-sector sensing mechanism that exploits directional sparsity and sector-level structural partitioning of RHS, enabling robust multi-user angle-of-arrival (AoA) estimation under limited-resolution hardware. Furthermore, to alleviate near-field distance estimation bias, we design a nonlinear fusion estimator that aggregates angle observations from multiple local far-field sectors using confidence-weighted integration based on sector-specific priors. A confidence-aware sector selection scheme is also developed to suppress unreliable observations without relying on data-driven learning models. Extensive Monte Carlo simulations demonstrate that our method achieves over 80% reduction in root mean square error (RMSE) localization compared to conventional MUSIC and least-squares schemes, and maintains meter-level accuracy even under severe hardware distortions and sector limitations.**

***Index Terms*—Near-field, Hierarchical Codebook, Beam Training, Reconfigurable Holographic Surfaces.**

I. Introduction

Integrated sensing and communication (ISAC) has emerged as a key application scenario for forthcoming sixth-generation (6G) communications [1]. Future 6G systems aim to deliver not only enhanced communication services, but also advanced sensing capabilities [2]. ISAC is expected to deliver advanced sensing capabilities such as positioning, tracking, and imaging [3]. However, incorporating these sensing functions into networks originally optimized for communication remains an unresolved challenge. In ISAC frameworks, base stations are expected to serve as reference nodes for localization, leveraging their known positions and the ability to transmit and process radio signals to estimate target range and angle of arrival (AoA) [4] [5].

In parallel, reconfigurable holographic surfaces (RHS), featuring continuous apertures and elements capable of dynamically controlling phase and amplitude, have been proposed to meet the growing demand for high-speed, cost-effective, and highly directional communication solutions [6]. The RHS can synthesize flexible electromagnetic wavefronts over their apertures [7], enabling near-field beamforming and holographic manipulation of propagation environments. These capabilities position RHS as a promising enabler for next-generation ISAC systems [8], where it can jointly support high-capacity transmission and fine-grained sensing and imaging.

However, existing RHS-assisted ISAC designs typically assume ideal and fully controllable hardware [9]–[12], where each surface element can independently and continuously adjust both phase and amplitude without any constraints. In reality, such assumptions are not attainable due to hardware imperfections, limited quantization resolution, mutual coupling among dense elements, and power restrictions that reduce the dynamic range of reflection coefficients. In practical RHS, these limitations introduce quantization errors and nonlinear distortions that can degrade both communication and sensing performance.

While works explicitly addressing ISAC performance with hardware constrained RHS remain scarce, particularly in near-field scenarios where spherical wavefront propagation characteristics deviate from conventional far-field assumptions [13] [14]. This gap is further accentuated by the fact that near-field sensing requires finer spatial resolution and more precise control, which are more susceptible to hardware imperfections. Existing studies most closely related to this work can be broadly classified into three categories. First, RHS-/RIS-assisted ISAC and holographic sensing works mainly assume ideal or fully controllable hardware responses, and therefore do not capture the amplitude-phase coupling and feasible-response constraints of practical meta-atoms [9]–[12]. Second, recent near-field localization and beam-training methods reduce the complexity of joint angle-range estimation, but they are typically developed for ideal arrays or ideal RIS/RHS responses and do not address the degradation of subspace sensing under hardware-constrained metasurface implementations [15]–[17]. Third, multi-sector or segmented-

The authors are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, United Kingdom. (e-mail: yd4u21@soton.ac.uk; sxn@ecs.soton.ac.uk; meh@ecs.soton.ac.uk).

Mohammed El-Hajjar would like to acknowledge the support of the Future Telecoms Research Hub, Platform for Driving Ultimate Connectivity (TITAN), sponsored by the Department of Science Innovation and Technology (DSIT) and the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/Y037243/1.

surface architectures have been shown to improve spatial coverage or enable local approximations [18]–[21], but the joint problem of hardware-constrained near-field localization, sector-wise confidence evaluation, and geometric fusion has not been explicitly studied. These limitations motivate the present work, which combines hardware-aware subspace analysis, hierarchical sector-wise sensing, and confidence-aware multi-sector fusion in a unified near-field RHS localization framework. Table I summarizes the main methodological differences between the proposed framework and the representative prior works. In contrast to the existing studies, our work jointly considers hardware-constrained subspace sensing, hierarchical sector-wise AoA estimation, and confidence-aware multi-sector fusion for near-field RHS localization.

Under high-frequency operation and large apertures, the conventional far-field assumptions become increasingly invalid [22], highlighting the need for accurate modeling of near-field communication and sensing behaviors. Precise AoA estimation are particularly critical in such systems, as they enable effective spatial signal processing and robust beamforming. Consequently, system performance is heavily dependent on the accuracy of channel estimation.

The estimation of AoA is a fundamental technique in multi-user positioning and sensing [23]. Classical high-resolution methods, including Multiple Signal Classification (MUSIC) [16], [17], [24], ESPRIT [25], and ℓ_1-SVD [26], have demonstrated strong performance when signal and noise subspaces can be reliably extracted from the covariance matrix of received signals. However, these methods incur substantial computational complexity, especially in multi-dimensional grid searches for jointly estimating angle and distance parameters [16].

To address these challenges, subsequent studies have developed AoA estimation techniques tailored for near-field scenarios, such as decoupling angle and distance estimation to reduce complexity [15]–[17], or employing deep learning approaches like the DCNN-based classifier proposed in [27]. Nevertheless, these methods predominantly assume ideal hardware conditions. In holographic multi-input multi-output (MIMO) systems, metasurface-based architectures introduce significant hardware impairments, including quantization errors and mutual coupling, which inevitably degrade the fidelity of the measured angle information [28]. Most existing near-field localization schemes rely on multi-stage procedures that first estimate coarse angular directions and subsequently refine distance estimation [29], [30], since exhaustive three-dimensional searches are computationally prohibitive. However, under a spherical wavefront model, the gain differences between beams corresponding to the same direction but different distances are inherently limited, making the distance estimation highly sensitive to errors in the initial angle estimation.

In addition to algorithmic adaptations, another promising strategy is to decompose large-scale RHS into multiple spatial sectors or blocks. In [18] demonstrated that such segmentation can locally approximate the near-field propagation within each block as an effective 'piecewise-far-field region'. *Subsequently, applying architectures in our previous work [19] with geometric multi-sectors metasurface, we show that performing angle estimation exclusively within each spatial sector under a quasi-far-field approximation, without attempting direct distance estimation, could fundamentally improve robustness.* This perspective suggests that decomposing large-scale RHS or massive arrays into multiple sectors [20], [21], where each exhibits an approximately planar wavefront over moderate ranges, enables the application of conventional far-field models to achieve more reliable angular sensing. Subsequently, geometric relationships among sectors can be exploited to reconstruct accurate three-dimensional user positions via multi-sector fusion. This gap motivates the development of new frameworks capable of combining sector-wise estimation, confidence-aware fusion, and compensation for hardware-induced distortions within the ISAC paradigm.

Specifically, in reconfigurable metasurface-based architectures, practical hardware limitations introduce amplitude and phase coupling effects, which significantly degrade the accuracy of multi-user localization and exacerbate computational complexity. *Therefore, it is critical to develop practical, efficient, and robust localization schemes that explicitly account for these hardware impairments while maintaining low complexity.*

In our work, the preliminary empirical observations reveal that classical MUSIC algorithms still exhibit spectrum peaks near true directions, yet accompanied by pseudo-peaks and estimation bias that have not been rigorously characterized. Moreover, due to the estimation bias introduced by these impairments, traditional near-field distance estimation methods, such as those based on MUSIC or fourth-order cumulants, exhibit significant localization errors. These observations motivate the development of more robust estimation and compensation techniques.

In this paper, we focus on near-field multi-user localization within the broader ISAC framework. We consider a near-field localization system assisted by a RHS deployed at the transmitter. The RHS is partitioned into multiple spatial sectors, each capable of quasi-far-field directional sensing, thereby enabling sector-wise estimation of incident wavefronts. A key challenge arises from practical hardware constraints and inter-element mutual coupling, which fundamentally impair the orthogonality of array steering vectors and severely distort classical subspace-based estimators such as MUSIC. In particular, even modest angular estimation errors under such hardware impairments can be significantly magnified in the distance domain, resulting in large localization errors when conventional near-field distance estimation techniques are employed. This observation motivates the adoption of a multi-sector RHS design, whereby each sector operates approximately in a local far-field regime and provides independent angle observations. By systematically fusing these sector-wise directional estimates, the proposed approach circumvents the need for explicit distance estimation, thus enhancing robustness against hardware-induced distortions and maintaining reliable localization performance under practical near-field conditions. Compared with distributed RHS deployments [37], the proposed multi-sector RHS integrates multiple directional views into a single BS-side platform, avoiding inter-node syn-

TABLE I
NOVELTY COMPARISON WITH THE LITERATURE

	Our paper	[5]	[16]	[17]	[15]	[31]	[30]	[32]	[33]	[34]	[35]	[36]
3D AoA Estimation	✓	✓	✓	✓		✓	✓	✓			✓	
Subspace Algorithm	✓	✓	✓	✓	✓	✓						
Compress Sensing	✓						✓	✓			✓	
Hardware Constraints	✓								✓	✓		✓
Hierarchical Algorithm	✓						✓	✓				
Multi-sectors RHS	✓											
Multi-sector Fusion	✓											

chronization, backhaul exchange, and network-wide geometric calibration. Although distributed RHS can provide large-baseline diversity, achieving comparable near-field coverage over the same local service region may require additional panels and denser site planning, which increases deployment and coordination overhead.

Motivated by these practical considerations, our contributions in this paper are summarized as follows:

- We provide a theoretical characterization of the classical MUSIC algorithm under hardware-constrained complex-plane phase responses. In particular, we show that for a broad class of amplitude-phase-coupled hardware mappings represented by an affine shift of the ideal phase-only steering response on the complex plane, the MUSIC spectrum can still preserve significant peaks near the true directions. At the same time, such hardware-constrained responses introduce deterministic pseudo-peaks and angular bias. This analysis clarifies both the residual effectiveness and the fundamental limitation of directly applying conventional subspace sensing to practical RHS-assisted near-field localization.
- We propose a ROMP-assisted hierarchical sector-wise sensing framework for low-complexity AoA estimation. The key motivation is that ROMP and MUSIC are affected differently by hardware constraints: over a large global search space, MUSIC is more vulnerable to hardware-induced pseudo-peaks and mismatch accumulation, whereas pure ROMP suffers from reduced resolution and degraded support recovery on dense local grids. By combining ROMP-based coarse support detection with MUSIC-based fine refinement, the proposed method progressively narrows the angular search space, avoids exhaustive global scanning, and achieves a more favorable balance among robustness, resolution, and complexity.
- We further develop a confidence-aware multi-sector geometric fusion scheme by exploiting the joint global near-field and local quasi-far-field property of the proposed multi-sector RHS. Specifically, users remain in the near field of the full RHS aperture, while each active sector admits a locally quasi-far-field approximation due to its smaller aperture. This allows the localization task to be transformed from explicit angle-range estimation into sector-wise angular sensing followed by geometric fusion. As a result, the proposed method avoids direct distance-domain sampling, reduces the sensitivity of range estimation to angular bias, and improves localization robustness under hardware constraints through reliability screening, consistency checking, and weighted nonlinear refinement.
- Extensive Monte Carlo simulations validate the effectiveness of the proposed framework. The proposed hierarchical sensing module achieves an average angular estimation error below 1° over a wide central region, while the overall localization framework reduces the RMSE by more than 80% compared with conventional LS- and MUSIC-based schemes. Even under limited sector availability and non-ideal hardware conditions, the proposed method maintains meter-level localization accuracy without relying on explicit distance-domain sampling.

In order to highlight the contributions of this paper relative to other existing works, Table I provides a comparative analysis of key features in various studies. The rest of this paper is organized as follows: we firstly introduce the system model, including the near-field channel model and the hardware constraint model for RHS in Section II. Then, the proposed hierarchical sensing framework, which combines sparse recovery and subspace refinement is detailed in Section III. Section IV introduces the multi-sector fusion strategy and the confidence-aware estimation algorithm. The computational complexity of the proposed schemes is analyzed in Section V. Section VI presents extensive simulation results to evaluate the performance of the proposed approach under various hardware configurations and impairments. Finally, Section VII concludes the paper and discusses potential future research directions.

II. SYSTEM MODEL

In this section, we describe the considered RHS architecture and the associated near-field channel model that underpins our sensing and communication framework.

In this work, we consider an RHS-assisted ISAC system where the base station equipped with a multi-sector RHS receives signals transmitted by the users. These signals are exploited for localization, in particular for estimating the AoA of each user. To establish the propagation model, we denote a global coordinate system centered at the base station, while each RHS sector is associated with a local coordinate frame aligned with its bore-sight. The overall sensing model is then formulated by combining the local steering vectors within each sector with the global spatial geometry. A schematic illustration of the multi-sector RHS configuration and its sensing operation is provided in Fig. 1. The RHS is mounted

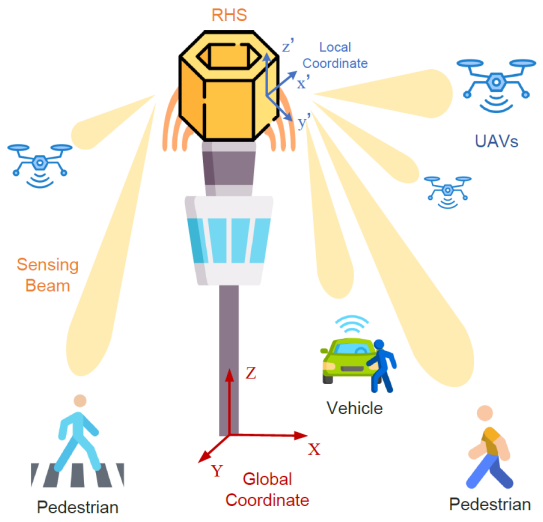


Fig. 1. System model using the multi-sector holographic surface at the base station.

at the BS and partitioned into L planar sectors, each oriented towards a distinct angular segment of space. As depicted, user signals impinge on different sectors under spherical wavefront propagation in the global coordinate system. However, within the limited aperture of each sector, the wavefront can be approximated as planar, thereby enabling quasi-far-field modeling. The local coordinate frames aligned with each sector's boresight are also shown in the figure, highlighting the geometric relation between global user positions and sector-wise observations.

A. RHS Architecture

We consider a multisector RHS architecture, which is physically implemented as a prism-shaped metasurface mounted on a base station. The RHS comprises multiple planar sectors, each oriented to cover a distinct angular segment of the surrounding three-dimensional space. This configuration allows the system to achieve quasi-omnidirectional coverage by combining the sensing capability of all sectors.

For each sector, the electromagnetic propagation between the RHS and users is modeled under a near-field spherical wavefront assumption when considering the global geometry. However, due to the limited angular extent and aperture of each individual sector, the wavefront can be locally approximated as a planar model (quasi-far-field) within the sector's field of view. For clarity, we specify the propagation assumptions adopted in this work. At the scale of the entire RHS aperture and operating frequency, the users fall within the near-field region, and spherical wavefront propagation must be taken into account. In contrast, when focusing on a single sector with a smaller aperture, the user distance typically exceeds the local Rayleigh threshold. Under this condition, the wavefront can be accurately approximated as planar, enabling the use of conventional far-field steering vectors. This joint near-field and quasi-far-field characterization forms the basis of the proposed multi-sector framework. This piecewise approximation simplifies the formulation of the steering vectors and enables the use of classical angle estimation techniques that rely on planar wave models. Specifically, for each user k, only $L/2$ contiguous sectors are considered as active sectors for the channel response. Formally, let $\mathcal{M}_k \subset \{1, 2, \ldots, L\}$ denote the serving sector index set for user k, satisfying $|\mathcal{M}_k| = L/2$.

This modeling assumption follows from the directional coverage geometry of the multi-sector RHS. Each sector is oriented towards its own forward-facing angular region, while the opposite half-space is not effectively covered. Therefore, for a user at any given location, the number of sectors with direct visibility is upper-bounded by $L/2$. Moreover, these visible sectors are adjacent in angle, which motivates the contiguous-sector assumption adopted in this work. The sector-wise quasi-far-field assumption is enforced through the user-region definition. Specifically, users are generated in the annular range $r_k \in [\max\{R_{\min}^{(\mathrm{RHS})}, R_{\mathrm{FF}}^{(\mathrm{sec})}\}, R_{\max}^{(\mathrm{RHS})}]$, where $R_{\min}^{(\mathrm{RHS})}$ and $R_{\max}^{(\mathrm{RHS})}$ characterize the near-field region of the whole RHS, while $R_{\mathrm{FF}}^{(\mathrm{sec})}$ denotes the far-field threshold of a single sector. Since each sector has a much smaller aperture than the entire RHS, i.e. $D_{\mathrm{sec}} \ll D_{\mathrm{RHS}}$, this construction ensures that the global channel remains near-field whereas each active sector admits a locally quasi-far-field approximation.

To capture the angularly limited contribution of each sector, we define a sector selection matrix $\mathbf{M}_k \in \mathbb{C}^{L\times L}$, expressed as

$$M_k = \mathrm{Diag}(\underbrace{1, \ldots, 1}_{\ell\in\mathcal{M}_k}, \underbrace{0, \ldots, 0}_{\ell\notin\mathcal{M}_k}). \tag{1}$$

B. Channel Model

Overall, we consider an RHS-aided near-field scenario in which the base station (BS) employs a $N = N_x \times N_y \times L$ uniform planar array holographic element surface to serve K single antenna users. The RHS surface with L sectors located at the origin, with N_x elements along the x-axis and N_y elements along the y-axis in each sectors. Inspired by our prior study on multi-sector metasurface architectures [19], where a large surface was divided into several angular sectors to extend the coverage, we adopt a similar multi-sector configuration for the RHS. In that design, multiple sectors can simultaneously contribute to the service of a single user if their coverage regions overlap. Building on this principle, in this paper we focus on RHS-assisted sensing, where the continuous aperture enables near-field modeling and quasi-far-field approximations within each sector under practical hardware constraints. To support multi-user connectivity, each user is assumed to be effectively served by a subset of sectors whose spatial coverage overlaps with the user's position. The general multi-path channel from the lth RHS sector to the kth user can be modeled as

$$\mathbf{h}_{l,k} = \sqrt{N}\bar{h}_k \mathbf{b}(\theta_k, \phi_k, r_k) + \sum_{i}^{I} \sqrt{\frac{N}{I}} \tilde{h}_{k,i} \mathbf{b}(\bar{\theta}_k, \bar{\phi}_k, \bar{r}_k), \tag{2}$$

where the complex-valued channel gains are the line-of-sight (LoS) channel $\bar{h}_k = \sqrt{\frac{\kappa}{1+\kappa}} \frac{\sqrt{\rho_0}}{r_k} e^{-j\frac{2\pi}{\lambda_c} r_k}$ and the non-line of

sight channel (NLoS) $\tilde{h}_k = \sqrt{\frac{1}{1+\kappa}}\frac{\sqrt{\rho_0}}{\bar{r}_k}e^{-j\frac{2\pi}{\lambda_c}r_k}$, while κ represents the Rician factor and ρ_0 denotes the channel gain at the reference distance of 1 m and I denotes the number of NLoS paths, respectively.

From an overview perspective, $\mathbf{H} = [\mathbf{h}_1; \ldots; \mathbf{h}_L]^{\mathrm{T}} \in \mathbb{C}^{N\times K}$ denotes the channel matrix between the transmission side in each sector, where in each sector $\mathbf{h}_l = [\mathbf{h}_{l,1}, \ldots, \mathbf{h}_{l,K}]^{\mathrm{T}} \in \mathbb{C}^{N/L\times K}$ denotes the channel matrix between the lth sector and each user.

This sector-wise sparsification is a principled approach reflecting the quasi-far-field approximations adopted locally within each sector, while preserving the inherent near-field spherical wavefront of the global geometry. Compared with multi-sector RIS [19], the RHS exhibits enhanced spatial resolution and continuous aperture characteristics, which further accentuate the necessity to enforce sector-dependent support constraints in practical deployments.

In this paper, we focus on RHS assisted near-field communication scenarios within high-frequency bands. From an overview perspective, the power of NLoS channel paths is significantly attenuated due to severe path loss and shadowing effects [38], [39]. As a result, our analysis is restricted to the main path, that is, the LoS channels, where the BS-user channel can be closely approximated by the LoS path alone, using the uniform spherical wave (USW) model [22]. The near-field LoS channel from the BS to the user can be characterized as

$$\mathbf{h}_{l,k} = \sqrt{N}\bar{h}_{l,k}\mathbf{b}_{l,k}(\theta_{l,k}, \phi_{l,k}, r_{l,k}), \tag{3}$$

where $\theta_{l,k}$, $\phi_{l,k}$, and $r_{l,k}$ represent the azimuth angle and elevation angle, and the distance from the transmitter to user k, with reference to the center of the RHS surface, respectively. Consequently, for any sector $\ell \notin \mathcal{M}_k$, the corresponding channel components are explicitly set to zero

$$\mathbf{h}_{\ell,k} = \mathbf{0}, \quad \forall \ell \notin \mathcal{M}_k. \tag{4}$$

Moreover, $\mathbf{b}_{l,k}(\cdot)$ is the steering vector of the kth user, which is derived based on the spherical wave model, where it samples the angles and the distance as follows [22]

$$\mathbf{b}(\theta,\phi,r) = \frac{1}{\sqrt{N}}\Big[e^{-j\frac{2\pi}{\lambda_c}(r^{(1,1)}-r)}, \ldots, e^{-j\frac{2\pi}{\lambda_c}(r^{(1,N_y)}-r)}, \ldots, e^{-j\frac{2\pi}{\lambda_c}(r^{(N_x,1)}-r)}, \ldots, e^{-j\frac{2\pi}{\lambda_c}(r^{(N_x,N_y)}-r)}\Big]^T, \tag{5}$$

where $r^{(i,j)}$ indicates the distance between the (i,j)th element of the RHS and the user.

For each sector's local coordinate, the location of the (n_x, n_y)th element can be expressed as $U_{\text{element}}^{(n_x,n_y)} = [\delta_{n_x}d, \delta_{n_y}d, 0]$, where $\delta_{n_x} = \frac{2n_x - N_x + 1}{2}$ and $\delta_{n_y} = \frac{2n_y - N_y + 1}{2}$, for $n_x = 1, \ldots, N_x$ and $n_y = 1, \ldots, N_y$. While the location of the kth user can be represented as $L_k = [r_k\cos\theta_k\sin\phi_k - \delta_{n_x}, r_k\cos\theta_k\sin\phi_k - \delta_{n_x}, r_k\cos\phi_k]$. The distance from the kth user to the (n_x, n_y)th element of the surface is then given by

$$r_k^{(n_x,n_y)} = \sqrt{(L_k^{(x)} - \delta_{n_x}d)^2 + (L_k^{(y)} - \delta_{n_y}d)^2 + (L_k^{(z)})^2}. \tag{6}$$

In this paper, we assume that the user is located in the Fresnel region and the distance does not exceed the Rayleigh distance from the aspect of the total RHS planar, i.e. $r > r_{\min} = \max\left\{\frac{1}{2}\sqrt{\frac{D^3}{\lambda_c}}, 1.2D\right\}$ and $r < r_{\max} = \frac{D^2}{\lambda_c}$, where D and λ_c denote the aperture size of the array and the carrier wavelength, respectively. In this region, the amplitude variations between the antennas are considered negligible [22] [40].

As illustrated in the previous example, while a user at a distance of a few meters remains in the near field of the global RHS, it can simultaneously lie in the quasi-far field of a single sector with a smaller aperture. This motivates the sector-wise approximation adopted in the channel formulation below. That means $r_k^{(n_x,n_y)} \approx r_k - (\delta_{n_x}\cos\theta_k\sin\phi_k - \delta_{n_y}\sin\theta_k\sin\phi_k)d$, where $\delta_{n_x} = \frac{2n_x - N_x - 1}{2}$ and $\delta_{n_y} = \frac{2n_y - N_y - 1}{2}$. Then, in this case $\mathbf{b}(\theta,\phi,r)$ is equivalent to the conventional far-field steering vector

$$\mathbf{a}(\theta,\phi) = \frac{1}{\sqrt{N}}\Big[1, \cdots, e^{j\pi(n_x\sin\phi\cos\theta + n_y\sin\phi\sin\theta)} \cdots, e^{j\pi(N_x\sin\phi\cos\theta + N_y\sin\phi\sin\theta)}\Big]^T. \tag{7}$$

The aggregated near-field channel vector $\mathbf{H}_k$ incorporates the sector-wise sparsification and is formed as

$$\mathbf{H}_k = \left[\mathbf{h}_{1,k}^\top, \ldots, \mathbf{h}_{L,k}^\top\right]^\top. \tag{8}$$

The signal received in the RHS from the kth user is given by

$$\mathbf{y}_k = \mathbf{H}_k x_k + \mathbf{n}_k, \tag{9}$$

where x_k denotes the symbol transmitted from the kth user to the RHS, $\mathbf{H}_k \in \mathbb{C}^{N\times 1}$ represents the near-field channel vector between the kth user and the RHS with N elements, and $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$ is the additive white Gaussian noise (AWGN) vector at the RHS, with variance σ^2.

C. Hardware Constraint Model

In this work, we consider a practical RHS model, which is subject to inherent hardware constraints that limit the achievable complex-valued response of each element. Specifically, we adopt the Lorentzian phase constraint model [41] [42] [43] [44],which characterizes the feasible set of element response coefficients as a function of a tunable control phase. It is worth emphasizing that the adopted model corresponds to a continuous-phase hardware-constrained response region, rather than a discrete phase-quantized architecture. Its role is to capture the amplitude-phase coupling of practical resonant meta-atoms through a normalized Lorentzian-type feasible locus in the complex plane. More concretely, the complex-valued response coefficient of each receiving element satisfies

$$\psi \in \left\{\frac{j + e^{j\zeta}}{2}\right\}, \quad \zeta \in [0, 2\pi), \tag{10}$$

where ζ denotes the control phase of the meta-atom[1]. Accordingly, the ideal steering vector $\mathbf{b}_{l,k}(\theta_{l,k}, \phi_{l,k}, r_{l,k})$ is elementally weighted by complex-valued coefficients ψ_n. Therefore,

[1]This model reflects the physical Lorentzian-type electromagnetic response of practical metasurfaces and implies that the effective amplitude and phase cannot be independently adjusted. In particular, the amplitude response is inherently coupled to the phase state, leading to a partial loss of amplitude information in the received signal.

the effective near-field channel under hardware constraints can be expressed as

$$\mathbf{h}_{l,k}^{(\mathrm{hw})} = \sqrt{N}\,\bar{h}_{l,k}\,\mathbf{\Psi}\,\mathbf{b}_{l,k}(\theta_{l,k},\phi_{l,k},r_{l,k}), \tag{11}$$

where $\mathbf{\Psi} = \mathrm{Diag}\big(\psi_1,\psi_2,\ldots,\psi_N\big)$ denotes the diagonal weighting matrix determined by the hardware constraint parameters. This hardware constraint modifies the effective steering vector of each sector and introduces a non-uniform weighting across the aperture. As a result, the composite channel exhibits both phase-dependent attenuation and distortion, which must be explicitly considered in the algorithm design and performance analysis.

In this section, we described the multi-sector RHS architecture with the near-field channel and hardware constraint models. Based on this formulation of the system, Section III develops the proposed AOA estimation framework to enable robust detection under practical impairments.

III. AoA Estimation Scheme

In this section, we develop the proposed AoA estimation scheme. We first analyze the limitations of classical subspace methods under hardware constraints, and then introduce a hierarchical framework that combines sparse recovery with refined subspace processing.

MUSIC relies on the theory of self-adjoint matrices. After collecting the T snapshots of measurements received as $\boldsymbol{Y} = [\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_T]$, we can rewrite the received signal as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n} \tag{12}$$

where $\boldsymbol{X} = [\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_T]$ and $\boldsymbol{N} = [\boldsymbol{n}_1,\boldsymbol{n}_2,\cdots,\boldsymbol{n}_T]$. The covariance matrix of the received signal can be written as

$$\boldsymbol{R}_{\boldsymbol{Y}} = \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{X}}\boldsymbol{A}^{\mathrm{H}} + N_0\boldsymbol{I}. \tag{13}$$

Here, $\boldsymbol{R}_{\boldsymbol{X}}$ is the covariance matrix of the transmitted signals, and it is a diagonal matrix under the assumption of independent transmission from the K sources. The k^{th} column of $\boldsymbol{A}$ corresponds to the array response due to the k^{th} source.

A. MUSIC Scheme under Hardware Constrained Scenario in Fine-resolution Layers

1) Pseudo-Peak in Mismatch Model: In this section, we provide a rigorous mathematical proof to elucidate the emergence of pseudo-peaks in the MUSIC algorithm under model mismatch conditions, where the sensing signal is subject to hardware constraints while the codebook employs ideal steering vectors.

From the above receive signal model Eq. (11), it can be observed that an ideal receive codeword should align with the near-field channel response [30]. Specifically, the ideal steering vector for a direction (θ,ϕ) is expressed in Eq. (5). When hardware constraints are imposed, the sensing signal's direction matrix becomes

$$\mathbf{A}_{\mathrm{hw}} = [f(\mathbf{a}(\theta_1,\phi_1)),\ldots,f(\mathbf{a}(\theta_K,\phi_K))], \tag{14}$$

where $f(\mathbf{a}) = 0.5\cdot(1j + \exp(1j\cdot\angle(\mathbf{a})))$ retains the phase information while normalizing the amplitude.

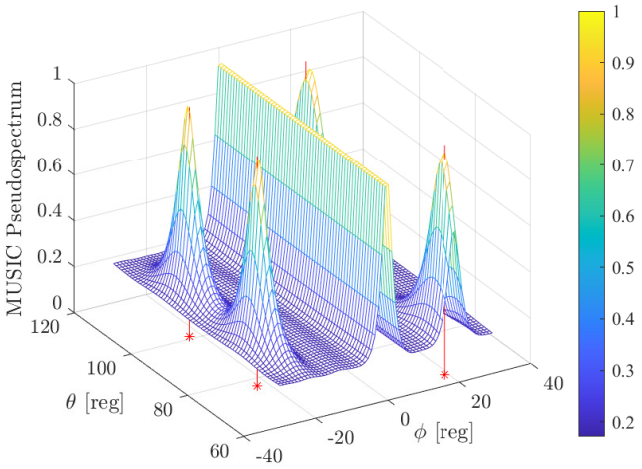


Fig. 2. MUSIC pseudo-spectrum versus azimuth angle and elevation angle.

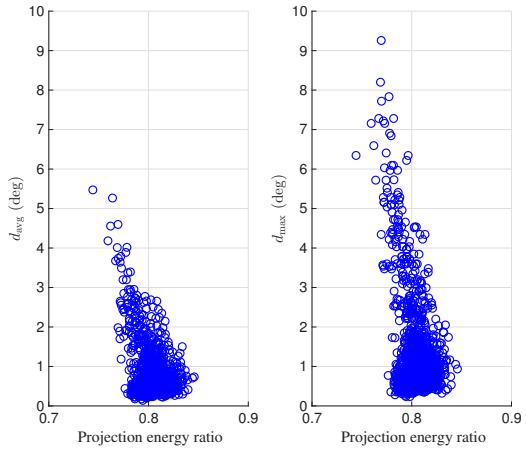


Fig. 3. Projection-energy ratio of the constant component versus the angular displacement of the nearest MUSIC local peak under hardware-constrained sensing: average displacement over users (left) and maximum displacement over users (right).

Therefore, the hardware-constrained steering vector becomes

$$\mathbf{a}_{\mathrm{hw}}(\theta,\phi) = 0.5\cdot 1j\cdot\mathbf{1} + 0.5\cdot\mathbf{a}(\theta,\phi), \tag{15}$$

with $\mathbf{1}$ being the all-one vector. Then, the covariance matrix of the received signal is given by

$$\mathbf{R}_{\boldsymbol{Y}}^{\mathrm{hw}} = \mathbf{A}_{\mathrm{hw}}\mathbf{R}_s\mathbf{A}_{\mathrm{hw}}^H + \sigma^2\mathbf{I}, \tag{16}$$

and its eigendecomposition provides the noise subspace $\mathbf{U}_n^{\mathrm{hw}}$, satisfying

$$\mathbf{A}_{\mathrm{hw}}^H\mathbf{U}_n^{\mathrm{hw}} = 0. \tag{17}$$

The MUSIC spectrum function [5], [15] is given by

$$P_{\mathrm{mismatch}}(\theta,\phi) = \frac{1}{\mathbf{a}(\theta,\phi)^H\mathbf{U}_n^{\mathrm{hw}}(\mathbf{U}_n^{\mathrm{hw}})^H\mathbf{a}(\theta,\phi)}, \tag{18}$$

which exhibits a pseudo-peak when $\mathbf{a}(\theta_k,\phi_k)^H\mathbf{U}_n^{\mathrm{hw}} \approx 0$ at the true direction (θ_k,ϕ_k).

From Eq. (15) and the orthogonality condition of Eq. (17), we can derive that

$$\mathbf{a}(\theta_k,\phi_k)^H\mathbf{U}_n^{\text{hw}} = 1j\cdot\mathbf{1}^H\mathbf{U}_n^{\text{hw}}. \tag{19}$$

To avoid relying on the heuristic statement that the all-one vector is always close to the hardware-constrained steering subspace, we rewrite the pseudo-peak mechanism in terms of an explicit projection relation. Let $S = \text{span}(A_{\text{hw}})$ and let P_S denote the orthogonal projector onto S. Under the affine hardware mapping $a_{\text{hw}}(\theta,\phi) = \alpha\bar{\mathbf{1}} + \beta a(\theta,\phi)$, where $\bar{\mathbf{1}} = \mathbf{1}/\sqrt{N}$ and $\beta \neq 0$, for any true direction $a_k = a(\theta_k,\phi_k)$, one has $a_{\text{hw}}(\theta_k,\phi_k) \in S$. Therefore, $(I-P_S)(\alpha\bar{\mathbf{1}}+\beta a_k) = 0$, which yields $(I-P_S)a_k = -(\alpha/\beta)(I-P_S)\bar{\mathbf{1}}$. This relation shows that the leakage of the true ideal steering vector outside the hardware-constrained signal subspace is directly controlled by the projection residual of the constant component. Hence, pseudo-peak preservation should be interpreted as a conditional phenomenon governed by subspace alignment, rather than as a universal consequence of $\mathbf{1}^H U_n^{\text{hw}} \approx 0$.

Attributed to the high dimensionality and the specific spatial sampling structure inherent to reconfigurable holographic surfaces, thereby explaining the accurate alignment of pseudo-peaks with the true directions. This analytical derivation confirms that the MUSIC pseudo-spectrum retains prominent peaks near the true directions, despite the presence of hardware-induced distortions and amplitude-phase coupling. To visually validate this property, Fig. 2 illustrates the pseudo-spectrum obtained under the hardware-constrained mismatch model, with $N = 256$ elements at 30 GHz, where four users are located in different directions within the near-field region at $r = 5$ m. The channel follows the LoS spherical wave model, and the hardware response is generated according to the Lorentzian phase constraint in Eq. (10). The result demonstrates that significant peaks indeed occur at the correct angular positions, thus corroborating the theoretical results. Accordingly, we define the projection residual as

$$\frac{\|(I-P_S)\mathbf{1}\|_2}{\|\mathbf{1}\|_2},$$

and the corresponding projection-energy ratio as shown in Fig. 3 is defined by

$$\frac{\|P_S\mathbf{1}\|_2^2}{\|\mathbf{1}\|_2^2}.$$

A larger projection-energy ratio equal to a small residual, indicates a stronger alignment between the hardware-induced constant component and the hardware-constrained steering subspace.

Furthermore, this property can be generalized to any hardware mapping defined by an affine combination of the constant vector and the original phase-only steering vector, as summarized in Corollary 1.

Corollary 1. *Consider a hardware mapping of the form $f(a) = \alpha\bar{\mathbf{1}} + \beta\exp(j\angle a)$, where $\alpha,\beta \in \mathbb{C}$ and $\beta \neq 0$. Then, for each true direction, the leakage of the corresponding ideal steering vector outside the hardware-constrained signal subspace is proportional to the projection residual of $\bar{\mathbf{1}}$ onto* $\text{span}(A_{\text{hw}})$. *Consequently, when this residual is small and the hardware-constrained steering vectors remain sufficiently separable, the MUSIC pseudo-spectrum tends to retain a prominent local peak near the true direction. Otherwise, competing pseudo-peaks and larger peak displacement may occur.*

2) Match Sensing Signal and Codebook in Hardware-Constrained Scheme: This subsection investigates how hardware constraints affect the alignment between the received sensing signals and the codebook, and analyzes their impact on the resulting MUSIC spectrum and angle estimation accuracy. In this case, the MUSIC spectrum function is defined as

$$P_{\text{hw}}(\theta,\phi) = \frac{1}{\mathbf{a}_{\text{hw}}(\theta,\phi)^H\mathbf{U}_n^{\text{hw}}(\mathbf{U}_n^{\text{hw}})^H\mathbf{a}_{\text{hw}}(\theta,\phi)}, \tag{20}$$

where $\mathbf{U}_n^{\text{hw}}$ is the noise subspace from the covariance matrix $\mathbf{R}_{xx}$. Ideally, the spectrum peaks in the true direction (θ_k,ϕ_k) since $\mathbf{a}_{\text{hw}}(\theta_k,\phi_k)^H\mathbf{U}_n^{\text{hw}} = 0$, but hardware constraints impair the directional distinguishability.

For two distinct directions (θ_1,ϕ_1) and (θ_2,ϕ_2), ideal steering vectors satisfy

$$\mathbf{a}(\theta_1,\phi_1)^H\mathbf{a}(\theta_2,\phi_2) \approx 0 \tag{21}$$

when sufficiently separated, whereas under hardware constraints, the inner product

$$\begin{aligned}\mathbf{a}_{\text{hw}}(\theta_1,\phi_1)^H\mathbf{a}_{\text{hw}}(\theta_2,\phi_2) = \\ 0.25\sum_{n_x,n_y}\Big[1+(-j)e^{jr_2^{(n_x,n_y)}} + je^{-jr_1^{(n_x,n_y)}} \\ +e^{j(r_2^{(n_x,n_y)}-r_1^{(n_x,n_y)})}\Big],\end{aligned} \tag{22}$$

with $r_{(n_x,n_y),i} = \frac{2\pi d}{\lambda}(m\sin\theta_i\cos\phi_i + n\sin\theta_i\sin\phi_i)$, reveals non-zero cross-interference terms like $e^{jr_{(n_x,n_y),2}}$ and $e^{-jr_{(n_x,n_y),1}}$, indicating that hardware-constrained steering vectors lack orthogonality. This reduced distinguishability alters the noise subspace $\mathbf{U}_n^{\text{hw}}$, orthogonal to $\mathbf{A}_{\text{hw}}$, such that $\mathbf{a}_{\text{hw}}(\theta,\phi)^H\mathbf{U}_n^{\text{hw}} = 0$ may hold not only at (θ_k,ϕ_k) but also at adjacent directions, shifting the spectrum peak and manifesting as a direction offset in the estimated angle, a phenomenon consistently observed in simulations results.

Building on these observations, the next subsection introduces sparse recovery methods to further achieve robust coarse-resolution angle estimation under hardware constraints.

B. OMP/ROMP in Coarse-resolution Layers

In multi-source signal processing applications, accurately estimating the signal covariance matrix $\mathbf{R}_s$ is crucial for identifying the power and correlations of individual sources. This work details an estimation method for $\mathbf{R}_s$ using sparse optimization, leveraging the sparse structure of sources in the angle-distance space. The complete process is outlined in this section, including the signal model, sparse matrix optimization, and covariance matrix construction.

Consider an antenna array with N elements receiving signals from M sources. The received signal vector $\mathbf{x}(t) \in \mathbb{C}^{N\times 1}$ at time t can be modeled as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{23}$$

where $\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_M)] \in \mathbb{C}^{N\times M}$ is the array steering matrix, with $\mathbf{a}(\theta_i)$ representing the steering vector for the i-th source. $\mathbf{s}(t) \in \mathbb{C}^{M\times 1}$ is the source signal vector. $\mathbf{n}(t) \sim \mathbb{CN}(0, \sigma_n^2\mathbf{I})$ is the additive white Gaussian noise with variance σ_n^2.

The covariance matrix of the received signal in Eq. (23) can then be expressed as

$$\mathbf{R}_{xx} = \mathbb{E}[\mathbf{x}(t)\mathbf{x}^H(t)] = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma_n^2\mathbf{I}, \tag{24}$$

where $\mathbf{R}_s = \mathbb{E}[\mathbf{s}(t)\mathbf{s}^H(t)]$ is the signal covariance matrix of interest.

In the angular domain, the sources are assumed to be sparsely distributed. Let K^2 represent the number of possible grid points in this space, where $K^2 \gg M$. We introduce a diagonal sparse matrix $\mathbf{S} \in \mathbb{C}^{K^2\times K^2}$ to represent the signal power distribution across this grid, with non-zero entries corresponding to the actual source locations.

To model the received signal's sparse structure, we define an overcomplete dictionary matrix $\mathbf{\Phi} \in \mathbb{C}^{N\times K^2}$, where each column represents a potential steering vector for a grid point in the angle-distance space. We can approximate $\mathbf{R}_{xx}$ as

$$\mathbf{R}_{xx} \approx \mathbf{\Phi}\mathbf{S}\mathbf{\Phi}^H + \sigma_n^2\mathbf{I}. \tag{25}$$

To estimate $\mathbf{S}$, we solve a sparse optimization problem that promotes sparsity in $\mathbf{S}$ while fitting the observed covariance $\mathbf{R}_{xx}$

$$\text{(P1):} \quad \min_{\mathbf{S}} \|\mathbf{S}\|_1 \quad \text{subject to} \quad \|\mathbf{R}_{xx} - \mathbf{\Phi}\mathbf{S}\mathbf{\Phi}^H - \sigma_n^2\mathbf{I}\|_F^2 \le \epsilon, \tag{26}$$

where $\|\cdot\|_1$ is the L_1 norm, encouraging sparsity in $\mathbf{S}$, and ϵ controls the fitting error.

Once $\mathbf{S}$ is estimated, we identify the non-zero elements in $\mathbf{S}$, which indicate the locations and power levels of the signal sources. Let $\{i_1, i_2, \ldots, i_M\}$ be the indices of these non-zero elements, representing the detected source positions in the angle-distance grid.

Define $\mathbf{S}_{\text{selected}} \in \mathbb{C}^{M\times M}$ as a diagonal matrix containing the power of the detected sources, with

$$\mathbf{S}_{\text{selected}} = \text{diag}(s_{i_1}, s_{i_2}, \ldots, s_{i_M}), \tag{27}$$

where s_{i_j} is the power at the j-th source position in $\mathbf{S}$.

The corresponding steering vectors for these positions are extracted from $\mathbf{\Phi}$, forming the matrix $\mathbf{A}_{\text{selected}} \in \mathbb{C}^{N\times M}$:

$$\mathbf{A}_{\text{selected}} = [\phi_{i_1}, \phi_{i_2}, \ldots, \phi_{i_M}]. \tag{28}$$

The estimated signal covariance matrix $\mathbf{R}_s$ can now be constructed as

$$\mathbf{R}_s = \mathbf{A}_{\text{selected}}^H \mathbf{S}_{\text{selected}} \mathbf{A}_{\text{selected}}. \tag{29}$$

To use the OMP [45] [46], we change the constraint in (P1) into the linear representation:

$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{subject to} \quad \|y - \mathbf{U}\mathbf{s}\|_F \le \sqrt{\epsilon}, \tag{30}$$

where $\mathbf{U} = \mathbf{\Phi} \otimes \mathbf{\Phi}$ and $y = \text{vec}(\mathbf{R}_s) - \text{vec}(\sigma_n^2\mathbf{I})$.

To further improve the stability and robustness of sparse support detection under hardware constraints, we employ ROMP as an enhanced alternative to conventional OMP. Unlike OMP method, which selects the single index with the largest correlation in each iteration, ROMP identifies a set of indices with comparable magnitudes and selects the subset that captures the maximum residual energy, thereby providing improved noise resilience and more reliable support recovery in near-field dense sampling scenarios [47] [48].

Specifically, ROMP operates as an iterative greedy algorithm to approximate the sparse representation vector $\mathbf{s}$ in Eq. (30). At each iteration t, ROMP first constructs a candidate set Λ_t combined by all indices with significant correlation in the codebook $\mathbf{U}$, represented as

$$\Lambda_t = \left\{ i \;\middle|\; \left|\langle u_i, e^{(t)}\rangle\right| \ge \alpha \cdot \max_j \left|\langle u_j, e^{(t)}\rangle\right| \right\}, \tag{31}$$

where $e^{(t)}$ is the current residual vector and $\alpha \in (0, 1)$ is a fixed threshold parameter. Here, we denote the ROMP threshold by α_{R} to distinguish it from the confidence thresholds used in the fusion stage. In this work, α_{R} is used as a conservative candidate-retention parameter in the coarse screening stage, rather than as a final support-size control parameter. A smaller α_{R} retains more candidate atoms and reduces the risk of excluding the true angular region under hardware-induced pseudo-peaks, while the final number of retained atoms is further bounded by the ROMP sparsity budget, set proportional to the number of users as $2K$ in our simulations.

Then, Λ_t is partitioned into groups of comparable magnitude into set $\mathbf{G}$, where the selected group $\mathbf{G}_t$ maximizes the ℓ_2-norm of correlation among all groups as

$$\mathbf{G}_t = \arg\max_{\mathbf{G}\subset\Lambda_t} \left\| \left[\langle u_j, e^{(t)}\rangle\right]_{j\in\mathbf{G}} \right\|_2. \tag{32}$$

The indices in G_t are then added to the cumulative support set $\mathcal{S}_t$, defined as

$$\mathcal{S}_t = \mathcal{S}_{t-1} \cup \mathbf{G}_t, \quad \mathcal{S}_0 = \emptyset \tag{33}$$

At each iteration, the selective vector $\mathbf{s}$ is updated by solving the least squares projection onto the subspace spanned by the columns of $\mathbf{U}$ indexed by $\mathcal{S}_t$

$$\mathbf{s}_{\mathcal{S}_t} = \arg\min_z \|y - \mathbf{U}_{\mathcal{S}_t} z\|_2, \tag{34}$$

where the full coefficient vector $\mathbf{s}$ is updated by assigning the components of z by the indices in $\mathcal{S}_t$, and setting all other components to zero. The residual is recomputed by

$$e^{(t+1)} = y - \mathbf{U}\mathbf{s}. \tag{35}$$

This process continues until the residual norm falls below a specified threshold ϵ or the predefined sparsity level is reached.

Unlike OMP, which greedily selects single atoms in each iteration, ROMP leverages the structured grouping criterion to enhance robustness against noise and coherence among dictionary atoms. In this context, the estimated sparse vector s directly corresponds to the nonzero entries associated with the detected support indices across all iterations.

While in this section we focused on developing robust sector-wise AoA estimation under hardware constraints, Section IV extends this framework by introducing a multi-sector fusion strategy to further improve localization performance.

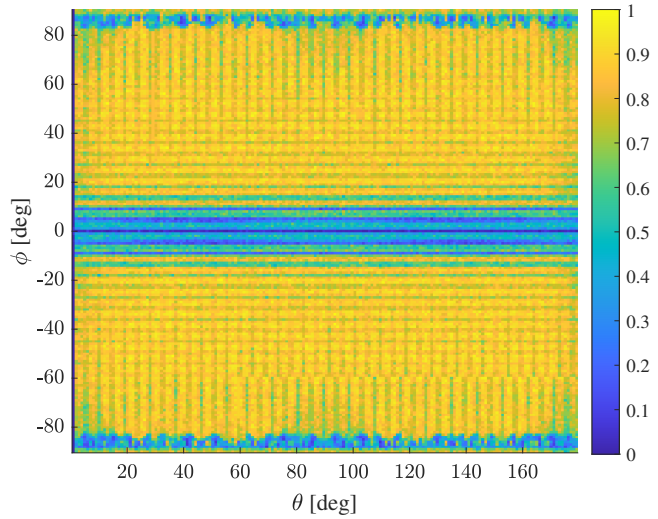


Fig. 4. RMSE confidence versus azimuth angle and elevation angle.

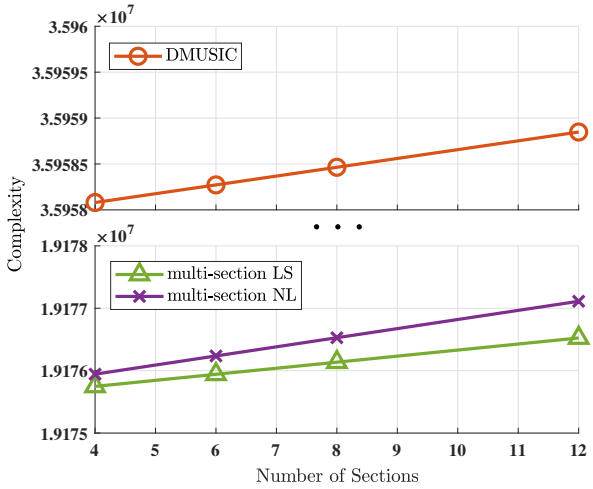


Fig. 5. Comparison of the complexity versus the number of sectors, when varying the estimate schemes.

IV. Multi-Sector Fusion Strategy

As conventional near-field localization approaches typically rely on a single planar sensing aperture and assume ideal hardware capabilities. In practical case, limited phase quantization, element coupling, and amplitude-phase response constraints introduce non-negligible distortions that degrade the accuracy of angle estimation. These hardware-induced errors result in partial loss of angular information, leading to systematic estimation bias that cannot be fully compensated by increasing the snapshot count or applying conventional calibration techniques.

To address these challenges, we adopt a multi-sector fusion strategy that exploits the spatial diversity inherent in independently oriented sectors of the reconfigurable surface. The key motivation is that different sectors observe the same target from distinct perspectives, thereby providing complementary angular information. This redundancy mitigates the impact of hardware impairments that may affect specific angular regions. Experimental observations further support this approach. As illustrated in Fig. 4, the confidence level is defined as the empirical probability that the angular estimation error remains within a threshold of 1°. To obtain this measure, a uniform grid is generated over the azimuth and elevation angles, and at each point of the grid, the system performs 10^3 Monte Carlo trials under randomized noise and hardware distortion. The confidence level is then computed as the ratio of successful trials to the total number of realizations, and the resulting two-dimensional confidence map is illustrated in Fig. 4.

Regions with systematically lower confidence often correspond to directions where the array geometry and phase quantization jointly reduce estimation resolution. Unlike conventional near-field estimators that perform explicit angle-range joint search, the proposed method estimates sector-wise AoAs first and then recovers the user position through multi-sector geometric fusion. Therefore, the distance does not enter the first-stage search explicitly, which reduces the online complexity. By combining observations from multiple sectors, the system can compensate for local reliability deficits and improve the robustness of the final position estimate.

In the proposed multi-sector fusion scheme, for the kth user, each sector provides an azimuth-elevation estimate $(\hat{\theta}_{l,k}, \hat{\phi}_{l,k})$ along with a confidence score $c_{l,k} \in [0,1]$, reflecting the reliability of the estimate under noise and hardware distortion. The array is partitioned into L sectors, each defined by a central azimuth orientation β_l globally. In this formulation, the sector-specific estimation procedures are performed independently within each local coordinate system. By aligning the resulting directional estimates with the global reference frame and incorporating the known positions of the sectors, the estimates can be consistently integrated to infer the user location in the global coordinate system. The transformation from the sector's local coordinate system to the global reference frame is represented by the rotation matrix

$$\mathbf{R}_l = \begin{bmatrix} -\sin(\beta_l) & \cos(\beta_l) & 0 \\ 0 & 0 & 1 \\ \cos(\beta_l) & \sin(\beta_l) & 0 \end{bmatrix}. \tag{36}$$

Given the local estimated direction vector as

$$\mathbf{d}_{\text{local}}^{(l,k)} = \begin{bmatrix} \sin(\hat{\theta}_{l,k})\sin(\hat{\phi}_{l,k}) \\ \cos(\hat{\phi}_{l,k}) \\ \sin(\hat{\phi}_{l,k})\cos(\hat{\theta}_{l,k}) \end{bmatrix}, \tag{37}$$

the global unit vector is computed as

$$\mathbf{d}_{l,k} = \frac{\mathbf{R}_l\, \mathbf{d}_{\text{local}}^{(l,k)}}{\left\| \mathbf{R}_l\, \mathbf{d}_{\text{local}}^{(l,k)} \right\|}. \tag{38}$$

After transforming the local estimates of each sector into the global coordinate system, an initial estimate of the position of the user $\mathbf{u} \in \mathbb{R}^3$ can be obtained by solving an unweighted LS problem. In this formulation, each sector contributes to an observation of the direction vector pointing towards the user, and the position is estimated by minimizing the aggregate squared projection errors relative to the sector phase centers. Geometrically, each sector provides a directional ray starting

from its phase center. In the absence of estimation error, these rays intersect at the true user position. Under hardware-induced bias and noise, however, the rays are generally inconsistent and do not meet at a single point. The confidence-aware fusion therefore seeks a point that best matches the most reliable and mutually consistent rays, while suppressing the impact of biased sector observations.

Specifically, the reference point $\mathbf{r}_i$ denotes the phase center of sector i, and the LS estimate is given by

$$\hat{\mathbf{u}}_{\mathrm{LS}} = \arg\min_{\mathbf{u}} \sum_{i\in\mathcal{T}} \left\| \left(\mathbf{I}_3 - \mathbf{d}_i\mathbf{d}_i^\top\right)\left(\mathbf{u} - \mathbf{r}_i\right)\right\|_2^2, \tag{39}$$

where $\mathbf{d}_i$ represents the normalized global direction vector estimated by sector i. This problem has a closed-form solution expressed as

$$\hat{\mathbf{u}}_{\mathrm{LS}} = \mathbf{A}^{-1}\mathbf{b}, \tag{40}$$

where

$$\mathbf{A} = \sum_{i\in\mathcal{T}} \left(\mathbf{I}_3 - \mathbf{d}_i\mathbf{d}_i^\top\right), \quad \mathbf{b} = \sum_{i\in\mathcal{T}} \left(\mathbf{I}_3 - \mathbf{d}_i\mathbf{d}_i^\top\right)\mathbf{r}_i. \tag{41}$$

This unweighted estimate provides a computationally efficient initial solution. However, it does not incorporate any information about the relative confidence or mutual consistency of the sector observations. Consequently, its performance can degrade if certain sectors are subject to severe hardware impairments or produce unreliable estimates. To address these limitations, a confidence-based weighting and directional consistency evaluation are introduced in the following stage, ensuring that sectors with higher estimation quality contribute proportionally more to the final position estimate.

While the LS estimate provides an efficient initial approximation of the user position, it assumes that all sector observations are equally reliable and does not account for potential hardware-induced biases or local estimation uncertainties. To improve robustness, it is necessary to assess the reliability of each directional estimate and assign appropriate confidence weights before incorporating them into the final fusion process. For this purpose, each sector is assigned a confidence value c_l derived from offline simulations. Specifically, during the training phase, a uniform grid of azimuth and elevation angles (θ, ϕ) is sampled within the coverage area of each sector. For each grid point, T_{MC} Monte Carlo trials are conducted by generating noisy observations under randomized hardware distortion parameters. The estimation error in each trial is measured as the angular deviation relative to the ground truth direction. The confidence score $c_l(\theta, \phi)$ is then computed as the empirical probability that the estimation error remains below a predefined threshold τ_c. During online inference, the confidence corresponding to the estimated direction is interpolated from this lookup table. Sectors with confidence below a minimum threshold α_1 are excluded from further processing

$$c_l \geq \alpha_1, \tag{42}$$

where α_1 denotes the minimum acceptable confidence level.

In addition to confidence evaluation, a second criterion is applied to assess the consistency of each sector's estimated direction relative to the remaining sectors. For each sector l, the mean direction of all other sectors is computed as

$$\bar{\mathbf{d}}^{(-l)} = \frac{1}{L-1}\sum_{k\neq l} \mathbf{d}_k, \tag{43}$$

and normalized by

$$\hat{\mathbf{d}}^{(-l)} = \frac{\bar{\mathbf{d}}^{(-l)}}{\|\bar{\mathbf{d}}^{(-l)}\|}. \tag{44}$$

The angular deviation is then defined as

$$\alpha_l = \arccos\left(\mathbf{d}_l \cdot \hat{\mathbf{d}}^{(-l)}\right). \tag{45}$$

A sector is considered consistent if this deviation is below a specified threshold α_2

$$\alpha_l \leq \alpha_2. \tag{46}$$

Combining both criteria, the set of selected sectors is defined as

$$\mathcal{T} = \left\{ l : \, c_l \geq \alpha_1 \text{ and } \alpha_l \leq \alpha_2 \right\}, \tag{47}$$

with $|\mathcal{T}|$ sectors retained for fusion. The confidence function c_l is generated offline by Monte Carlo simulation. For each sector l, a uniform grid of azimuth-elevation pairs (θ, ϕ) within the sector coverage is sampled. For each point, T_{MC} independent realizations are simulated under randomized noise and hardware distortion. The estimation error in each trial is measured as the angular deviation $\Delta\alpha^{(t)}$. The confidence score is computed as the empirical probability that the estimation error remains below a threshold α_1

$$c_l(\theta, \phi) = \frac{1}{T_{\mathrm{MC}}} \sum_{t=1}^{T_{\mathrm{MC}}} \mathbb{I}\left[\Delta\alpha^{(t)} \leq \alpha_1\right], \tag{48}$$

where $\mathbb{I}[\cdot]$ denotes the indicator function. This two-dimensional confidence map is stored in a lookup table and interpolated during runtime.

Finally, the selected sectors contribute to the localization estimate through a nonlinear fusion procedure that incorporates the confidence weights derived from the previous stage. For each sector $i \in \mathcal{T}$, the confidence value c_i denotes the interpolated reliability score corresponding to the estimated azimuth and elevation. To assign higher importance to more reliable observations, the weight w_i can be defined as

$$w_i = \sqrt{c_i},$$

which provides a smoother scaling compared to using c_i directly. The solution is refined by minimizing the weighted discrepancy between the estimated and observed directions. The cost function is expressed as

$$J(\mathbf{u}) = \sum_{i\in\mathcal{T}} w_i \left\| \frac{\mathbf{u} - \mathbf{r}_i}{\|\mathbf{u} - \mathbf{r}_i\|} - \mathbf{d}_i \right\|_2^2, \tag{49}$$

where $\mathbf{r}_i$ denotes the phase center of sector i and $\mathbf{d}_i$ is its normalized global direction vector. The final position is obtained by solving a confidence-weighted nonlinear least-squares problem

$$\hat{\mathbf{u}}_{\mathrm{NLS}} = \arg\min_{\mathbf{u}} J(\mathbf{u}), \tag{50}$$

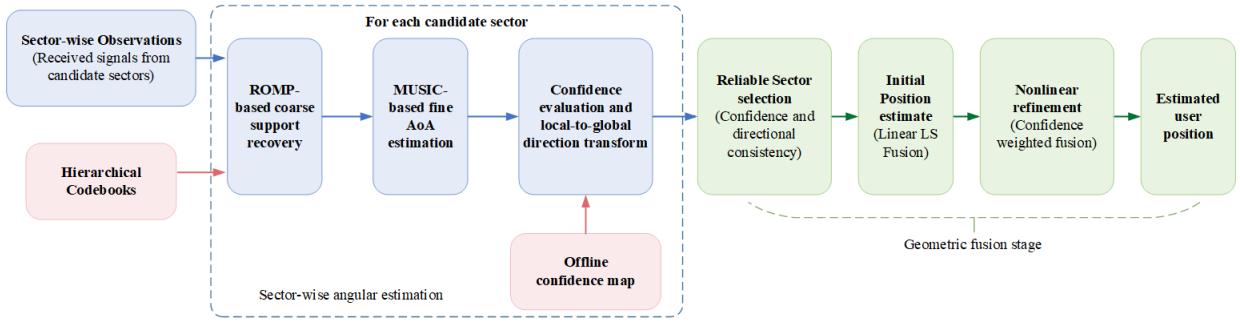


Fig. 6. The block diagram for the overall workflow of the proposed algorithm.

where the cost function $J(\mathbf{u})$ is nonlinear in the user position $\mathbf{u}$. We therefore refer to this refinement as nonlinear least-squares (NL) estimation, which is solved iteratively using the Levenberg–Marquardt algorithm (LMA) [49] initialized at the linear LS solution.

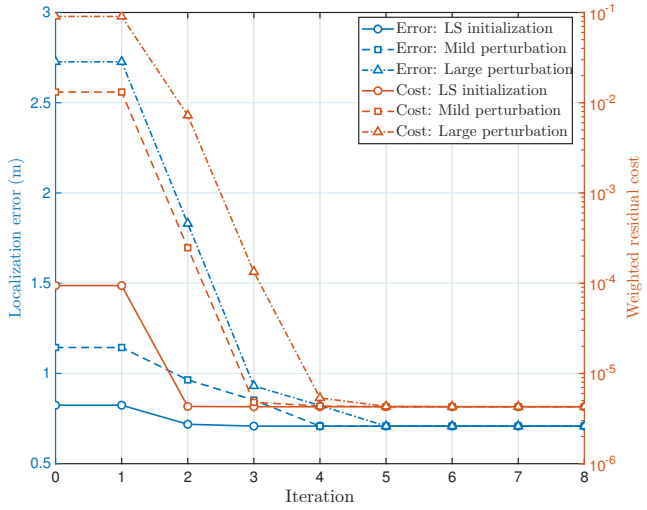


Fig. 7. Convergence behavior of the proposed nonlinear fusion estimator under different initializations for $L = 12$ and $K = 10$.

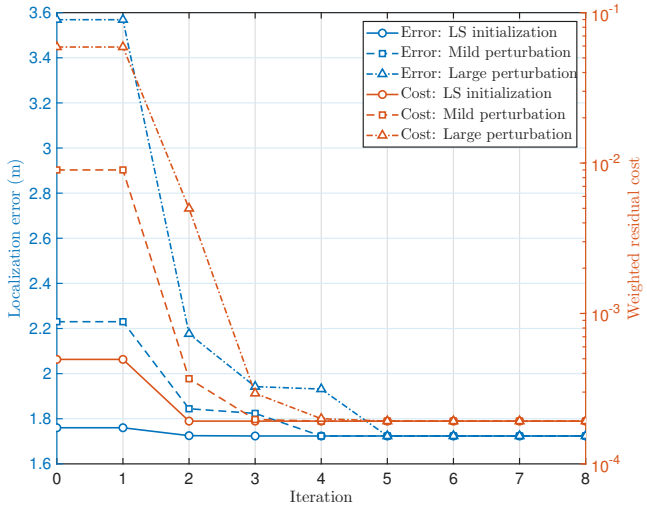


Fig. 8. Convergence behavior of the proposed nonlinear fusion estimator under different initializations for $L = 8$ and $K = 5$.

To further examine the convergence behavior of the proposed nonlinear fusion estimator, we tested the Levenberg–Marquardt procedure under three different initializations: the LS initialization, a mildly perturbed point, and a largely perturbed point. Specifically, the perturbed initial points are generated by adding a random three-dimensional displacement to the LS initialization, where the displacement direction is uniformly random and the perturbation magnitude is set to 0.5 m for the mild case and 1.5 m for the large case. As shown in Fig. 7 and Fig. 8, the weighted residual cost decreases rapidly within the first few iterations, and the localization error converges to nearly the same final value in all tested cases. This indicates that the nonlinear stage mainly performs local refinement around a geometry-based coarse estimate, rather than a global search from arbitrary initial points. Since the LS initialization is constructed from the selected sector-wise directional intersections after confidence and consistency filtering, it is usually already close to the feasible region of the true solution. In addition, the confidence-weighted cost function downweights unreliable sector observations, which improves robustness against angular bias and reduces the sensitivity to imperfect initialization.

The overall processing flow of the proposed hierarchical multi-sector localization framework is illustrated in Fig. 6, showing the sector-wise angular estimation and confidence-aware multi-sector fusion strategy.

V. Complexity Analysis

In this section, we analyze the complexity of the multi-sector fusion estimation scheme using different sensing methods. To evaluate the computational efficiency of the proposed localization scheme, we analyze the asymptotic complexity of each processing stage. The system adopts N_x horizontal antennas and N_y vertical antennas, and L is sector divisions, resulting in a total antenna count of $N = N_x \times N_y \times L$. For each sector, the hierarchical estimation is conducted via an initial ROMP-based sparse direction recovery, followed by layer-wise MUSIC refinement and optional nonlinear optimization based on each sector's angular stage results.

In Stage I, sparse support detection is performed by ROMP on the first three hierarchical layers. The complexity can be approximated as

$$C_{\mathrm{ROMP}} = \mathcal{O}\big(N_{\mathrm{active}}^{(1)} \cdot s_1 \cdot L\big), \tag{51}$$

where $N_{\text{active}}^{(1)}$ denotes the number of retained atoms in the first layer, and s_1 represents the per-iteration sparsity level.

Subsequent layer-wise MUSIC refinement across the hierarchical levels incurs an overall complexity expressed as

$$C_{\text{MUSIC}} = \sum_{\ell=2}^{L_{\max}} \Big[\mathcal{O}\big(N_{\text{active}}^{(\ell)\,3}\big) + \mathcal{O}\big(N_{\text{active}}^{(\ell)} \cdot s_2\big) \Big], \tag{52}$$

where the cubic term is due to covariance eigendecomposition, and s_2 denotes the number of angular grid points evaluated.

In Stage II, the final position estimation procedure varies depending on the selected method. Applying the MUSIC algorithm on the distance estimation (DMUSIC) relies on a complexity of

$$C_{\text{DMUSIC}} = \mathcal{O}\big(N_{\text{active}}^3\big) + \mathcal{O}\big(N_{\text{active}} \cdot s_3\big), \tag{53}$$

where s_3 is the sampling grid size.

In the multi-sector fusion scheme, the linear LS estimation scales as $\mathcal{O}(L)$, while the nonlinear optimization incurs

$$C_{\text{NL}} = \mathcal{O}(T \cdot L), \tag{54}$$

with T denoting the number of iterations required for convergence.

To further illustrate the offline computational burden, the system employs Monte Carlo training over $T_{\text{mc}} = 1000$ simulations and a sampling grid of $S_t = 31 \times 31$ angular pairs. The total offline training complexity scales as

$$C_{\text{training}} = T_{\text{MC}} \cdot S_t \cdot \big(C_{\text{ROMP}} + C_{\text{MUSIC}}\big). \tag{55}$$

This cost, however, is amortized during deployment and does not impact online inference latency.

The complexity results for DMUSIC, multi-sector LS, and multi-sector NL schemes are shown in Fig. 5. It can be observed that the proposed multi-sector approaches achieve significantly lower complexity than the DMUSIC baselines while maintaining superior localization accuracy. It is worth distinguishing the offline and online computational cost. The confidence map is generated offline before sensing, and therefore it does not increase the online localization complexity. During online inference, the dominant operations are sector-wise ROMP support detection, covariance eigendecomposition in the retained local MUSIC regions, and the final fusion step. The fusion stage itself is lightweight: LS solves only a small three-dimensional linear problem, while nonlinear fusion optimizes the same three-dimensional position variable over a limited number of Levenberg–Marquardt iterations. The main memory usage is dominated by candidate dictionaries and covariance-related matrices. Since the proposed hierarchical scheme retains only reduced local candidate regions and avoids global distance-domain search, it has a lower practical memory footprint than exhaustive or DMUSIC-based searches.

The simulation-based evaluation of these schemes is presented in Section VI.

VI. Simulation Result

In this section, we present the simulation results of the achievable rate to validate the effectiveness of our proposed beam training schemes under the proposed hybrid RHS structure, considering different modes of RHS surface hardware constraints. The Monte Carlo method is used with 10^5 iterations. Specifically, we consider an RHS-aided multi-user communication system with carrier frequency $f_c = 30$ GHz to represent a typical mmWave ISAC scenario, where the short wavelength leads to compact RHS deployment and near-field effects. The fixed total number of RHS elements $N = 3072$, $N_x = 16$, and the N_y in each sectors is determined by the number of sectors . The reference SNR of the system is defined as SNR $= \frac{P\beta N}{r^2\sigma^2}$ [32], where the noise power is set to $\sigma^2 = -70$ dBm. Unless otherwise specified, user locations are randomly generated within the considered global near-field coverage region of the RHS. Hence, the reported localization results are averaged over users with different distances and angles, rather than evaluated at a single fixed range.

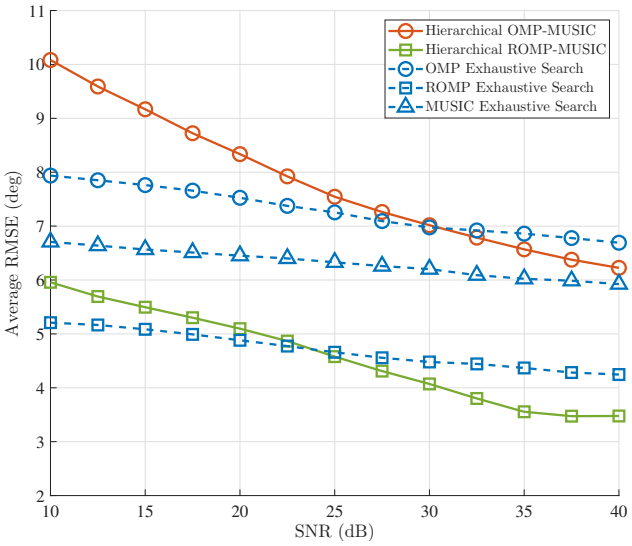


Fig. 9. Comparison of the RMSE versus SNR in single sector, when varying the number of sectors and estimation scheme.

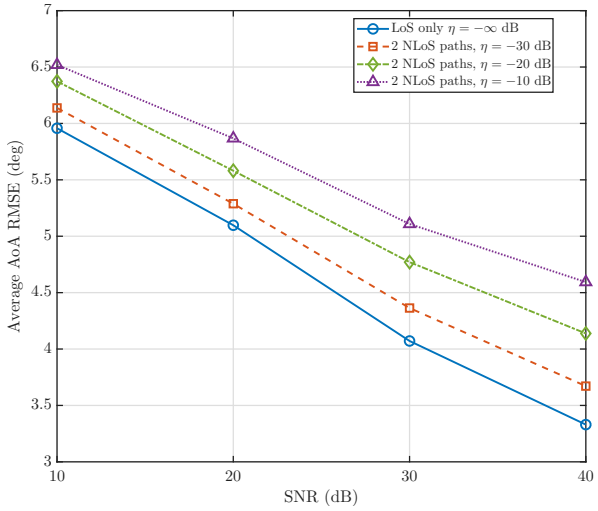


Fig. 10. Average AoA RMSE versus SNR under different NLoS-to-LoS power ratios for the proposed ROMP-MUSIC sensing module in a single-sector 16×16 RHS configuration.

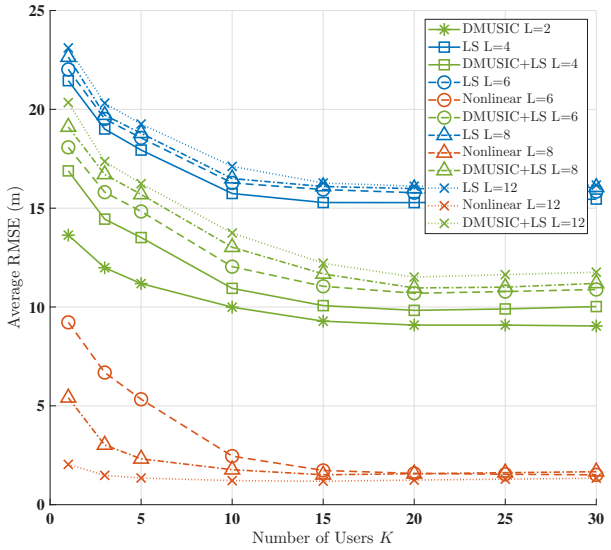


Fig. 11. Comparison of the RMSE versus the number of users, when varying the number of sectors and estimation scheme.

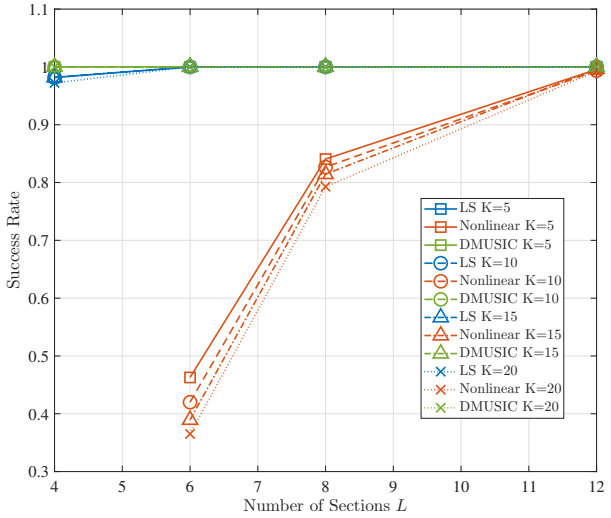


Fig. 12. Comparison of the success rate versus the number of sectors, when varying the number of users and estimation scheme.

In Fig. 9, we compare the localization accuracy of different sparse recovery methods in a single-sector setting, including OMP-MUSIC and ROMP-MUSIC with hierarchical algorithm in single surface, as well as exhaustive search baselines. The RMSE here is evaluated for the estimated user location, defined as RMSE $= \sqrt{\frac{1}{N_{\mathrm{MC}}}\sum_{mc=1}^{N_{\mathrm{MC}}}\|\hat{\mathbf{u}}^{(mc)}-\mathbf{u}\|^2}$, where $\hat{\mathbf{u}}^{(mc)}$ denotes the estimated position in the mc-th Monte Carlo trial, $\mathbf{u}$ is the ground-truth location, and MC is the total number of Monte Carlo trials. It is observed that hierarchical OMP-MUSIC exhibits the highest RMSE, particularly in low SNR regimes, because OMP tends to select incorrect support in noisy measurements. In contrast, Hierarchical ROMP-MUSIC demonstrates significantly lower error across all SNRs, highlighting its robustness to noise and grid mismatch. Moreover, the performance of pure MUSIC with exhaustive search remains moderate, as it lacks explicit sparsity constraints and can converge to fake pseudo-peaks cause by hardware constraints. The results motivate our subsequent adoption of ROMP-MUSIC as the fundamental estimation module in multi-sector configurations to achieve a more accurate sensing result. The slightly inferior performance of hierarchical ROMP-MUSIC compared with ROMP exhaustive search in the low-SNR regime is mainly due to error propagation in the coarse screening stage. Since the proposed scheme progressively reduces the angular search space to lower the computational complexity, an incorrect coarse support decision may exclude the true direction before the MUSIC refinement stage. By contrast, exhaustive search preserves global candidates and is therefore more robust under severe noise, but at a substantially higher computational cost. This behavior reflects the intended complexity–robustness trade-off of the proposed hierarchical design.

In Fig. 10, we consider a single-sector 16×16 RHS configuration with $K = 8$ users. For each user, the channel is modeled as the superposition of one dominant LoS component and $I_{\mathrm{NLoS}} = 2$ weak reflected NLoS components, modeled as

$$\mathbf{h} = \sqrt{1-\rho_{\mathrm{NLoS}}}\mathbf{h}_{\mathrm{LoS}} + \sqrt{\rho_{\mathrm{NLoS}}}\mathbf{h}_{\mathrm{NLoS}},$$

where $\mathbf{h}_{\mathrm{LoS}}$ denotes the normalized LoS channel response and $\mathbf{h}_{\mathrm{NLoS}}$ denotes the normalized aggregate NLoS response. The relative NLoS strength is controlled by $\rho_{\mathrm{NLoS}} = 10^{\eta/10}/(1+10^{\eta/10})$, where $\eta = 10\log_{10}(P_{\mathrm{NLoS}}/P_{\mathrm{LoS}})$. We consider $\eta = -30$, -20, and -10 dB, together with the LoS-only case.

The results show that the average AoA RMSE decreases monotonically with SNR for all tested multipath levels. As expected, stronger NLoS components lead to higher RMSE because the reflected paths perturb the dominant LoS angular signature. Nevertheless, the degradation is smooth and no abrupt failure is observed even for $\eta = -10$ dB, indicating that the proposed ROMP-MUSIC sensing module remains stable under weak-to-moderate NLoS contamination.

In Fig. 11, we illustrate the average positioning RMSE achieved by LS, DMUSIC and the proposed NL methods under varying numbers of users and different number of RHS sectors L. The simulations are conducted by user positions that are randomly generated within the global near-field coverage area to emulate realistic spatial diversity. It is observed that in both the LS and DMUSIC methods, the RMSE increases as the number of sectors grows, indicating that simply incorporating more independent directional observations without an error-based filtering mechanism may introduce additional inaccurate inputs due to hardware limitations, ultimately degrading the estimation performance. Additionally, when the total number of antenna elements is fixed and evenly distributed among sectors, configurations with fewer sectors allocate a denser antenna array per sector, which enhances local angle estimation resolution and helps mitigate the impact of random hardware impairments.

Notably, the proposed nonlinear estimation scheme, which incorporates confidence-aware weighting and bias compensation, effectively mitigates the issue that introduce additional inaccurate and consistently outperforms the baseline LS and DMUSIC approaches across all tested configurations. For example, the nonlinear method reduces the RMSE by approximately 80% compared to LS when $L = 12$, demonstrating

superior robustness against hardware-induced angle estimation errors.

Another key observation is that the RMSE remains relatively stable as the number of users increases beyond $K = 15$. This can be explained by the fact that ROMP-based angle recovery benefits from sparsity in the angular domain. Although extremely sparse scenarios facilitate more distinctive angular activations, they also tend to suffer from higher sensitivity to noise and estimation ambiguities, leading to an increased RMSE in the low-K regime. As the number of user K increases, the diversity of active directions improves the reliability of sparse recovery, thereby stabilizing positioning accuracy.

To further provide a fair comparison using the same total number of elements, we include the $L = 2$ case in Fig. 11 while keeping the total number of RHS elements fixed at $N = 3072$. Under the considered sector-visibility rule, each user in the $L = 2$ configuration is effectively observed by only one sector over the near-field region, so this case serves as a single-sector baseline with DMUSIC adopted. The results show that the proposed multi-sector framework achieves lower localization RMSE than the $L = 2$ baseline for the same number of elements. Additionally, we observe that simply fusing more DMUSIC-based sector estimates does not necessarily improve performance, since explicit distance-domain estimation is highly sensitive to angular bias under hardware constraints. By contrast, the proposed method relies on sector-wise angular estimation and confidence-aware geometric fusion, which better exploits multi-sector diversity and improves robustness as L increases.

It is worth emphasizing that maintaining meter-level localization accuracy, as observed in Fig. 11, already constitutes practically useful performance for near-field ISAC systems. Prior studies have highlighted that meter-level positioning is sufficient to support applications such as beam management, user clustering, and IoT device localization in 5G and beyond networks [5], [35], [48]. Achieving sub-meter accuracy typically requires significantly more complex system architectures or idealized hardware assumptions, which are often impractical under real hardware constraints.

In Fig. 12, we present the success rate, defined as the proportion of realizations in which each method produces a valid position estimate. For the proposed non-linear method, a valid estimate requires that, after applying confidence-aware weighting and bias compensation, at least two sectors remain with mutually consistent directional estimates within a 3° global angular deviation. Due to this strict requirement, the nonlinear method exhibits lower success rates when there are a few number of sectors, as the probability of retaining a sufficient number of high-confidence, globally consistent observations decreases. As the number of sector increases, the available directional estimations improve, leading to a increase in the success rate. In contrast, the LS and DMUSIC methods do not incorporate any filtering, they only require that the final estimated position is within the global near-field radius. Consequently, their success rates remain consistently high in all tested configurations. These results indicate that the number of sectors should be selected by balancing fusion reliability and implementation cost. With a fixed total aperture, a small number of sectors provides insufficient angular diversity for reliable fusion, whereas an excessively large number of sectors reduces the per-sector aperture and increases sector-level control and offline training overhead.

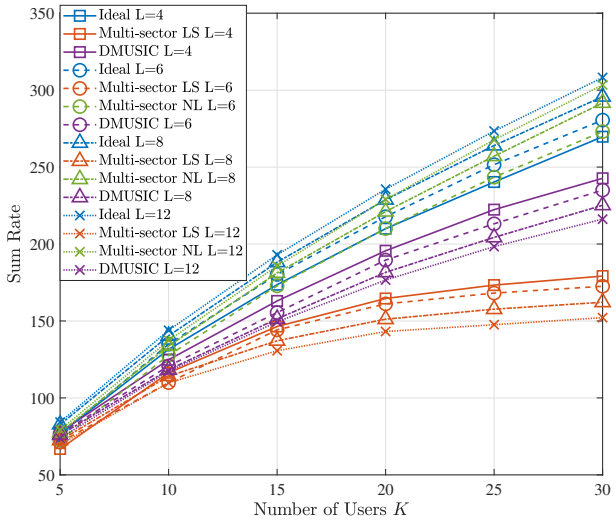


Fig. 13. Comparison of the sum rate versus the number of users, when varying the number of sectors and estimation scheme.

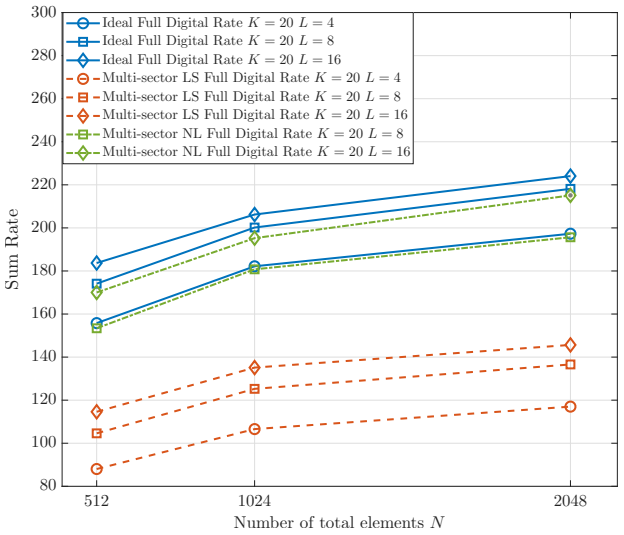


Fig. 14. Comparison of the achievable sum rate versus the total number of RHS elements for $K = 20$, under different sector numbers and localization schemes.

Furthermore, in Fig. 13 we calculate the full digital sum rate to further confirm that more accurate localization results in a higher achievable sum rate. The achievable sum rate is given by

$$R = \sum_{k=1}^{K} \log_2 \left(1 + \frac{\Gamma_k P \beta N |\mathbf{h}_k^H(\theta_k, \phi_k, r_k)\mathbf{w}_k|^2}{r_k^2 \sigma^2} \right), \quad (56)$$

where the $\beta = \left(\frac{\lambda}{4\pi}\right)^2 = -53$ dB, Γ_k is a power allocation parameter, which can be calculated based on the water-filling

method [50], where $\mathbf{w}_k$ is the fully digital beamforming vector calculated by the SVD method.

From the Fig. 13 we can observe that the proposed non-linear method consistently outperforms LS and DMUSIC, achieving approximately 90% of the ideal sum rate. In contrast, the LS approach suffers from persistent performance gaps due to uncorrected misalignment. Furthermore, while increasing L improves angular diversity, it yields diminishing returns beyond $L = 8$, since hardware constraints and limited SNR per sector restrict further improvements. These results jointly demonstrate the trade-off between sector partitioning and estimation reliability, and underscore the importance of incorporating prior knowledge and confidence weighting in near-field multi-user localization for supporting high-rate communication.

To supplement the fixed-N results, we further examine the effect of the total RHS element budget on the achievable sum rate. Here, N_x is fixed to 16, and the total number of elements are selected as $N = 512, 1024, 2048$ so that $N_y = N/(N_x L)$ remains an integer for all compared sector numbers. Accordingly, we consider the representative sector settings $L = 4, 8, 16$. The results in Fig. 14 show that, for all considered sector numbers, the achievable sum rate increases monotonically with N, since a larger RHS provides higher beamforming gain and improves the quality of the sector-wise localization input. Moreover, the proposed nonlinear fusion scheme consistently remains much closer to the ideal fully digital rate than the LS baseline, which confirms that the benefit of the proposed framework persists under different total number of elements.

Overall, all our simulation results highlight several important insights. First, increasing the number of sectors helps to mitigate hardware constraints and estimation inconsistencies, especially when employing fusion strategies that are sensitive to confidence and bias. Second, the proposed non-linear method achieves a lower RMSE and higher success rates especially in multi-sector architecture. Then, the proposed algorithm demonstrates stable performance under varying user densities, indicating its potential applicability in dense IoT and multi-user wireless sensing scenarios.

VII. Conclusion

In this paper, we proposed a low-complexity near-field localization framework for RHS under practical hardware constraints. We analytically characterized how limited phase quantization induces pseudo-peaks and angular bias in subspace estimators such as MUSIC. To overcome these effects, we developed a hierarchical ROMP-MUSIC scheme combined with confidence-aware multi-sector fusion, which is capable of achieving robust angle estimation without relying on the real-time training. Simulations demonstrated that our method reduces localization RMSE by over 80% compared to conventional approaches and maintains meter-level accuracy under severe distortions. Future work will explore broadband extensions and cooperative multi-panel configurations.

References

[1] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo. Joint radar and communication design: Applications, state-of-the-art, and the road ahead. *IEEE Trans. Commun.*, 68(6):3834–3862, 2020.

[2] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu. An overview of signal processing techniques for joint communication and radar sensing. *IEEE J. Sel. Topics Signal Process.*, 15(6):1295–1315, 2021.

[3] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu, Y. Shen, F. Colone, and K. Chetty. A survey on fundamental limits of integrated sensing and communication. *IEEE Commun. Surveys Tuts.*, 24(2):994–1034, 2022.

[4] G. Chen, X. Zeng, S. Jiao, A. Yu, and Q. Luo. High accuracy near-field localization algorithm at low snr using fourth-order cumulant. *IEEE Commun. Lett.*, 24(3):553–557, 2020.

[5] X. Zhang, W. Chen, W. Zheng, Z. Xia, and Y. Wang. Localization of near-field sources: A reduced-dimension music algorithm. *IEEE Commun. Lett.*, 22(7):1422–1425, 2018.

[6] S. Zeng et al. Reconfigurable intelligent surfaces in 6G: Reflective, transmissive, or both? *IEEE Commun. Lett.*, 25(6):2063–2067, June 2021.

[7] A. Pizzo, T. L. Marzetta, and L. Sanguinetti. Spatially-stationary model for holographic MIMO small-scale fading. *IEEE J. Sel. Areas Commun.*, 38(9):1964–1979, September 2020.

[8] H. Zhang, H. Zhang, B. Di, M. Di Renzo, Z. Han, H. V. Poor, and L. Song. Holographic integrated sensing and communication. *IEEE J. Sel. Areas Commun.*, 40(7):2114–2130, 2022.

[9] J. Hu, Z. Chen, and J. Luo. Multi-band reconfigurable holographic surface based isac systems: Design and optimization. In *Proc. IEEE Int. Conf. Commun. (ICC)*, pages 2927–2932, 2023.

[10] H. Zhang, H. Zhang, B. Di, and L. Song. Holographic integrated sensing and communications: Principles, technology, and implementation. *IEEE Commun. Mag.*, 61(5):83–89, 2023.

[11] P. Zhu, W. Ni, and X. Wang. Integrated sensing and communication with reconfigurable holographic surface. *IEEE Trans. Commun.*, 2025. To be published.

[12] S. Li, T. Mao, G. Liu, F. Zhang, R. Liu, M. Hua, Z. Gao, Q. Wu, and G. K. Karagiannidis. Intelligent metasurface-enabled integrated sensing and communication: Unified framework and key technologies. *arXiv*, 2025. arXiv:2506.13713, submitted to IEEE Wireless Commun.

[13] X. Shi, J. Wang, Z. Sun, and J. Song. Spatial-chirp codebook-based hierarchical beam training for extremely large-scale massive MIMO. *IEEE Trans. Wireless Commun.*, 23(4):2824–2838, April 2024.

[14] J. An, C. Yuen, L. Dai, M. Di Renzo, M. Debbah, and L. Hanzo. Near-field communications: Research advances, potential, and challenges. *IEEE Wireless Commun.*, 31(3):100–107, June 2024.

[15] J. He, M. N. S. Swamy, and M. O. Ahmad. Efficient application of music algorithm under the coexistence of far-field and near-field sources. *IEEE Trans. Signal Process.*, 60(4):2066–2070, 2012.

[16] M. Haghshenas, A. Mahmood, and M. Gidlund. Efficient multi-source localization in near-field using only angular domain music. *arXiv preprint arXiv:2501.11460*, 2025.

[17] C. Huang, J. Xu, W. Xu, X. You, C. Yuen, and Y. Chen. Low-complexity channel estimation for extremely large-scale mimo in near field. *IEEE Wireless Commun. Lett.*, 13(3):671–675, 2024.

[18] M. Cui and L. Dai. Near-field wideband beamforming for extremely large antenna arrays. *IEEE Trans. Wireless Commun.*, 23(10):13110–13124, 2024.

[19] Y. Dong, Q. Li, S. X. Ng, and M. El-Hajjar. Reconfigurable intelligent surface relying on low-complexity joint sector non-diagonal structure. *IEEE Open J. Veh. Technol.*, 5:1106–1123, 2024.

[20] Hongyu Li, Shanpu Shen, and Bruno Clerckx. Beyond diagonal reconfigurable intelligent surfaces: A multi-sector mode enabling highly directional full-space wireless coverage. *IEEE Journal on Selected Areas in Communications*, 41(8):2446–2460, 2023.

[21] Hongyu Li, Shanpu Shen, and Bruno Clerckx. Beyond diagonal reconfigurable intelligent surfaces: A multi-sector mode enabling highly directional full-space wireless coverage. *IEEE Journal on Selected Areas in Communications*, 41(8):2446–2460, 2023.

[22] Y. Liu, Z. Wang, J. Xu, C. Ouyang, X. Mu, and R. Schober. Near-field communications: A tutorial review. *IEEE Open J. Commun. Soc.*, 4:1999–2049, 2023.

[23] M. A. Islam, G. C. Alexandropoulos, and B. Smida. Simultaneous multi-user mimo communications and multi-target tracking with full duplex radios. In *Proc. IEEE Globecom Workshops*, pages 19–24, 2022.

[24] Y.-D. Huang and M. Barkat. Near-field multiple source localization by passive sensor array. *IEEE Trans. Antennas Propag.*, 39(7):968–975, 1991.

[25] F. Gao and A. B. Gershman. A generalized esprit approach to direction-of-arrival estimation. *IEEE Signal Process. Lett.*, 12(3):254–257, 2005.

[26] D. Malioutov, M. Cetin, and A. S. Willsky. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.*, 53(8):3010–3022, 2005.

[27] C. M. Mylonakis and Z. D. Zaharis. A novel three-dimensional direction-of-arrival estimation approach using a deep convolutional neural network. *IEEE Open J. Veh. Technol.*, 5:643–657, 2024.

[28] Nikolaos Kolomvakis and Emil Björnson. Exploiting mutual coupling characteristics for channel estimation in holographic mimo. In *GLOBECOM 2024 - 2024 IEEE Global Communications Conference*, pages 3570–3575, 2024.

[29] K. Qu, S. Guo, J. Ye, and H. Zhao. Two-stage beamspace music-based near-field channel estimation for hybrid xl-mimo. *IEEE Commun. Lett.*, 28(8):1949–1953, 2024.

[30] C. Wu, C. You, Y. Liu, L. Chen, and S. Shi. Two-stage hierarchical beam training for near-field communications. *IEEE Trans. Veh. Technol.*, 73(2):2032–2044, February 2024.

[31] Y. Du, Y. Hu, F. Zhang, V. G. Menon, and B. Ji. Three-dimensional channel estimation for near-field xl-ris: A polar-domain approach. *IEEE Trans. Consum. Electron.*, pages 1–1, 2025.

[32] Y. Zhang, X. Wu, and C. You. Fast near-field beam training for extremely large-scale array. *IEEE Wireless Commun. Lett.*, 11(12):2625–2629, December 2022.

[33] R. Deng, B. Di, H. Zhang, Y. Tan, and L. Song. Reconfigurable holographic surface: Holographic beamforming for metasurface-aided wireless communications. *IEEE Trans. Veh. Technol.*, 70(6):6255–6259, June 2021.

[34] 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. Wireless Commun.*, 21(8):6003–6017, August 2022.

[35] S. Yue, S. Zeng, L. Liu, Y. C. Eldar, and B. Di. Hybrid near-far field channel estimation for holographic mimo communications. *IEEE Trans. Wireless Commun.*, 23(11):15798–15813, 2024.

[36] Q. Li, M. El-Hajjar, Y. Sun, I. Hemadeh, A. Shojaeifard, and L. Hanzo. Energy-efficient reconfigurable holographic surfaces operating in the presence of realistic hardware impairments. *IEEE Trans. Commun.*, pages 1–1, 2024.

[37] George C. Alexandropoulos, Ioanna Vinieratou, and Henk Wymeersch. Localization via multiple reconfigurable intelligent surfaces equipped with single receive rf chains. *IEEE Wireless Communications Letters*, 11(5):1072–1076, 2022.

[38] W. Liu, C. Pan, H. Ren, F. Shu, S. Jin, and J. Wang. Low-overhead beam training scheme for extremely large-scale ris in near field. *IEEE Trans. Commun.*, 71(8):4924–4940, August 2023.

[39] Ibrahim A. Hemadeh, Katla Satyanarayana, Mohammed El-Hajjar, and Lajos Hanzo. Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget. *IEEE Communications Surveys & Tutorials*, 20(2):870–913, 2018.

[40] E. Björnson, Ö. T. Demir, and L. Sanguinetti. A primer on near-field beamforming for arrays and reconfigurable intelligent surfaces. In *Proc. Asilomar Conf. Signals, Syst., Comput.*, pages 105–112, 2021.

[41] J. Xu, L. You, G. C. Alexandropoulos, X. Yi, W. Wang, and X. Gao. Near-field wideband extremely large-scale MIMO transmissions with holographic metasurface-based antenna arrays. *IEEE Trans. Wireless Commun.*, pages 1–1, 2024.

[42] T. Sleasman, M. F. Imani, W. Xu, J. Hunt, T. Driscoll, M. S. Reynolds, and D. R. Smith. Waveguide-fed tunable metamaterial element for dynamic apertures. *IEEE Antennas Wireless Propag. Lett.*, 15:606–609, 2016.

[43] Seraphin F. Kimaryo and Kyungchun Lee. Downlink beamforming for dynamic metasurface antennas. *IEEE Transactions on Wireless Communications*, 22(7):4745–4755, 2023.

[44] Amirhossein Azarbahram, Onel L. A. Lopez, Richard D. Souza, Rui Zhang, and Matti Latva-Aho. Energy beamforming for rf wireless power transfer with dynamic metasurface antennas. *IEEE Wireless Communications Letters*, 13(3):781–785, 2024.

[45] 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.*, 64(6):2370–2386, 2016.

[46] N.-S. Duong, Q.-T. Nguyen, and T.-M. Dinh-Thi. Omp-based channel estimation with dynamic grid for mmwave mimo positioning systems. *IEEE Commun. Lett.*, 27(10):2623–2627, 2023.

[47] D. Needell and R. Vershynin. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. *Found. Comput. Math.*, 9:317–334, 2009.

[48] L. Mei, F. Gao, H. Pan, and Y. Xue. An improved romp sparse channel estimation algorithm in ofdm system. In *Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC)*, pages 1–4, 2015.

[49] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch. Position and orientation estimation through millimeter-wave mimo in 5g systems. *IEEE Trans. Wireless Commun.*, 17(3):1822–1835, 2018.

[50] A. Goldsmith. *Wireless Communications*. Cambridge University Press, Cambridge, 2005.



Yinuo Dong received the B.Eng. degree in Electronic Information Engineering from Changzhou University, Changzhou, China, in 2021, and the M.Sc. degree in Mobile Communications and Smart Networking from the University of Southampton, Southampton, U.K., in 2022. She is currently pursuing the Ph.D. degree with the Next Generation Wireless Communications Group, School of Electronics and Computer Science, University of Southampton. Her research interests include reconfigurable intelligent surfaces, near-field communications, holographic MIMO, and signal processing for advanced wireless systems.



Qingchao Li (Member, IEEE) received the B.Eng. degree from the University of Science and Technology Beijing, Beijing, China, in 2017, the M.S. degree from the University of Science and Technology of China, Heifei, China, in 2020, and the Ph.D. degree in wireless communications from the University of Southampton, Southampton, U.K., in 2024 under the University of Southampton-Interdigital Communications Inc. Fully Funded Scholarship. His research interests include reconfigurable intelligent surfaces, holographic MIMO and non-terrestrial wireless networks. He was the recipient of 2022 Chinese Government Award for Outstanding Self-Financed Student Abroad and 2023 Doctoral College Research Award, University of Southampton.



Soon Xin Ng (S'99-M'03-SM'08) received the B.Eng. degree (First class) in electronic engineering and the Ph.D. degree in telecommunications from the University of Southampton, U.K., in 1999 and 2002, respectively. He is currently a Professor of Next Generation Communications at the University of Southampton. His research interests include adaptive coded modulation, coded modulation, channel coding, space-time coding, joint source and channel coding, iterative detection, OFDM, MIMO, cooperative communications, distributed coding, quantum communications, quantum error correction codes, joint wireless-and-optical-fibre communications, game theory, artificial intelligence and machine learning. He has published over 300 papers and co-authored two John Wiley/IEEE Press books in this field.



Mohammed El-Hajjar (M'02, SM'14) is a Professor of Signal Processing for Wireless Communications in the School of Electronics and Computer Science in the University of Southampton. He is the recipient of several academic awards and has published a Wiley-IEEE book and more than 150 IEEE journal and conference papers and in excess of 10 patents. Mohammed's research interests include the design of intelligent and energy-efficient transceivers, MIMOs, millimeter wave communications, non-terrestrial networks and machine learning for wireless communications. Mohammed's research is funded by the Engineering and Physical Sciences Research Council, the Royal Academy of Engineering and many industrial partners.