A Two-Stage ISAC Framework for Low-Altitude Economy Based on 5G NR Signals

Haisu Wu, Hong Ren, *Member, IEEE*, Cunhua Pan, *Senior Member, IEEE*, Boshi Wang, Jun Tang,
Haoyang Weng, Feng Shu, *Member, IEEE*, Jiangzhou Wang, *Fellow, IEEE*

Abstract—The evolution of next-generation wireless networks has spurred the vigorous development of the low-altitude economy (LAE). To support this emerging field while remaining compatible with existing network architectures, integrated sensing and communication (ISAC) based on 5G New Radio (NR) signals is regarded as a promising solution. However, merely leveraging standard 5G NR signals, such as the Synchronization Signal Block (SSB), presents fundamental limitations in sensing resolution. To address the issue, this paper proposes a two-stage coarse-to-fine sensing framework that utilizes standard 5G NR initial access signals augmented by a custom-designed sparse pilot structure (SPS) for high-precision unmanned aerial vehicles (UAV) sensing. In Stage I, we first fuse information from the SSB, Type#0-PDCCH, and system information block 1 (SIB1) to ensure the initial target detection. In Stage II, a refined estimation algorithm is introduced to overcome the resolution limitations of these signals. Inspired by the sparse array theory, this stage employs a novel SPS, which is inserted into resource blocks (RBs) within the CORSET#0 bandwidth. To accurately extract the off-grid range and velocity parameters from these sparse pilots, we develop a corresponding high-resolution algorithm based on the weighted unwrapped phase (WUP) technique and the RELAX-based iterative method. Finally, the density-based spatial clustering of applications with noise (DBSCAN) algorithm is adopted to prune the redundant detections arising from beam overlap. Comprehensive simulation results demonstrate the superior estimation accuracy and computational efficiency of the proposed framework in comparison to other techniques.

Index Terms—Integrated sensing and communication (ISAC), 5G New Radio (5G NR), multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM), low-altitude economy (LAE)

I. INTRODUCTION

THE evolution of future sixth-generation (6G) wireless networks is driven by the demands of emerging applications, with the low-altitude economy (LAE) at the forefront [1]. As a new economic form, the LAE utilizes low-altitude airspace to conduct diverse flying activities, which promises to deliver revolutionary productivity advancements in sectors such as logistics delivery, environmental monitoring, and intelligent inspection [2], [3]. As this large-scale economic activity is expected to involve millions of aircraft [4], the supporting network needs to provide a dual capability: massive, real-time data exchange for communication and continuous, high-precision sensing for safe and uniform regulation [5], [6].

To address these converging requirements, integrated sensing and communication (ISAC) has emerged as a critical enabling technology [7], [8]. By multiplexing 5G-Advanced (5G-A) or future 6G communication signals, ISAC can simultaneously support high-speed data transmission for unmanned aerial vehicles (UAVs) and provide effective flight supervision. Specifically, the seamless coverage provided by the communication network is leveraged for the effective perception, localization, and surveillance of UAVs. This capability, in turn, assists in UAV navigation, path planning, and airspace management, which is essential for the prosperity of the LAE [1], [4].

Building on this motivation, considerable research has focused on the sensing of UAVs utilizing terrestrial base stations (BSs) to support the LAE. In monostatic scenarios, for example, the authors of [9] proposed a practical ISAC framework for cluttered environments and addressed the power allocation between communication and sensing beams. In addition, the authors of [10] introduced an efficient algorithm for jointly angle, velocity, and range estimation in the OFDM-based ISAC system, which includes a pruning algorithm for redundant target points. In multistatic scenarios, [11] investigated the UAV parameter estimation by employing the tensor decomposition technique and proposed a minimum spanning tree (MST)-based multi-BS data association scheme. Moreover, an extended Kalman filter (EKF) framework was introduced in [4] to predict the UAV's movement and developed a corresponding handover strategy.

Notably, the aforementioned research often operates without considering the practical constraints of standardized protocols like 5G NR [12]–[15] or Wi-Fi [16], focusing instead on theoretical aspects such as beamforming or parameter extraction. To foster practical deployment, both academia and industry have begun to focus on implementing sensing within these protocol frameworks, and the investigations generally fall into two categories: data-based and pilot-based sensing.

In monostatic data-based sensing scenarios, BS's *priori* knowledge of transmitted data can be used to estimate UAV parameter estimation via techniques such as matched filtering

H. Wu, H. Ren, C. Pan, B. Wang, J. Tang, H. Weng and J. Wang are with National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. (e-mail: wuhaisu, hren, cpan, boshi_wang, 220241016, 220241245, j.z.wang@seu.edu.cn).

Feng Shu are with the School of Information and Communication Engineering, Hainan University, Haikou 570228, China (email:shufeng@hainanu.edu.cn).

Corresponding authors: Hong Ren and Cunhua Pan.

and data cancellation, while beamforming can steer energy toward potential targets [17]. This principle can be applied within the 5G NR standard, where several downlink channels serve as potential signals of opportunity for wireless sensing. Early studies have explored the feasibility of using data-centric channels, such as the Physical Downlink Control Channel (PDCCH), the Physical Downlink Shared Channel (PDSCH), and the associated Demodulation Reference Signals (DMRS) [17]. For instance, the theoretical ambiguity functions of PDCCH and PDSCH were analyzed in [58]. However, such approaches are fundamentally dependent on active downlink data transmissions, rendering the sensing capability opportunistic rather than persistent. In addition, leveraging user-centric data channels like PDSCH for sensing poses significant privacy risks, as sensitive user information could be exposed to malicious eavesdroppers.

To overcome the reliance on user data and the associated privacy concerns, research has shifted towards dedicated reference signals that are independent of communication traffic. The Positioning Reference Signal (PRS) and the Channel State Information Reference Signal (CSI-RS) are potential candidates, as they are specifically designed for positioning and channel estimation. The work in [18] investigated the application of the PRS for radar sensing by proposing a multi-frame estimation scheme and deriving its Cramér-Rao lower bounds. In addition, [19] introduced a novel irregular resource pattern to suppress the inherent range-velocity ambiguities of the comb-like PRS structure. Despite the advantages, the dedicated signals like PRS and CSI-RS are typically transmitted only upon specific network configuration and are not always-on in most operational networks, which limits their ubiquitous applicability for persistent sensing.

In contrast, the Synchronization Signal and PBCH Block (SSB) is a fundamental component of the 5G NR frame structure that is broadcast periodically, which enables it an ideal and highly reliable signal of opportunity. Consequently, SSB-based sensing has been extensively investigated [20]–[25]. For instance, the authors of [20] leveraged the beam-sweeping nature of the SSB, utilizing RSRP measurements at the user equipment (UE) to achieve spatial environment mapping. Similarly, [21] exploited the periodicity of the SSB for passive coherent location (PCL). However, the study also highlighted that the maximum unambiguous velocity of PCL is severely constrained by the SSB's long transmission periodicity (e.g., 10 ms minimum). While the study in [22] addressed UAV detection using the SSB, its performance was limited: angular resolution is constrained by the wide beamwidth of the SSB, and velocity estimation accuracy is limited by the short duration of the per-beam OFDM transmission. To further address these limitations, the work in [26] proposed to combine the SSB with downlink control information (DCI) and system information block 1 (SIB1) signals to mitigate the radar ambiguity problem. Notably, while the 5G NR standard allocates a bandwidth part of $N_{\mathrm{RB}}^{\mathrm{DL,BWP}}$ resource blocks (RBs) for PDSCH carrying SIB1, the PDSCH itself often occupies only a fraction of this potential bandwidth [13]–[15], [26].

Since the PDSCH frequency allocation is mandated to use Downlink Resource Allocation Type 1 (i.e., a contiguous set of RBs), the requirement creates an opportunity to insert auxiliary pilots into the unused portions of the $N_{\mathrm{RB}}^{\mathrm{DL,BWP}}$ RBs [15]. These pilots can, in turn, further enhance sensing performance and assist the DMRS within the PDSCH for more effective channel estimation.[1]

The design of the aforementioned auxiliary pilots falls within the research domain of pilot placement in OFDM systems, which plays a pivotal role in optimizing ISAC performance. While initial research centered on optimizing uniformly spaced patterns [27], the equidistant approach is known to limit the maximum unambiguous velocity and range. In order to overcome the ambiguity introduced by such sparsity, subsequent works have introduced more sophisticated designs. For instance, [28] introduced a novel sensing pattern composed of multiple comb-type RS structures, where the comb sizes of the constituent structures are co-prime. Building upon a similar principle, [29] utilized co-prime intervals between the RS to eliminate ambiguity. Furthermore, [30] proposed a set of insertion criteria for OFDM RS patterns aiming at maximizing the extended ambiguity property (EAP) and applied this principle to super-resolution algorithms like 2D-MUSIC [31] and 2D-IAA [32], [33]. However, these works primarily focused on maximizing the interference-free detection region; consequently, the practical range and velocity estimation performance was not directly evaluated.

Distinct from the aforementioned pattern design methodologies, a novel approach inspired by array design theory reformulates the RS placement problem as a 2D-array design task [34]. Specifically, the proposed sparse pilot scheme achieves the same unambiguous range and peak sidelobe level (PSL) as uniform pilot placement, but with a substantial reduction in required resource elements (REs). Nevertheless, as the subsequent parameter estimation relies on a conventional 2D-FFT algorithm to generate the range-Doppler maps (RDMs), the estimation accuracy remains limited.

Motivated by the above findings, this paper proposes a novel two-stage coarse-to-fine estimation framework to leverage 5G NR broadcast signals and newly designed sparse pilots for high-precision UAV sensing. Specifically, this framework is designed to operate within the constraints of the 5G NR protocol while overcoming the inherent limitations of standard signals. The main contributions of this paper are summarized as follows:

- First, we establish a coarse estimation stage (Stage I) that fuses information from multiple 5G NR initial access signals. By combining the RDMs derived from the SSB and SIB1 with the high-resolution range profile from the Type#0-PDCCH, this stage provides robust initial target detection by utilizing the area-based combination CFAR (ABC-CFAR) algorithm.

---

[1]The use of auxiliary pilots for enhanced channel estimation is beyond the scope of this paper, which primarily focuses on their benefit to sensing performance.

- Second, to overcome the resolution limitations of standard signals, we design a novel sparse pilot structure (SPS) inspired by sparse array theory. The SPS is strategically inserted into the #CORSET0 bandwidth. We then develop a corresponding high-resolution refinement algorithm (Stage II), which employs the weighted unwrapped phase (WUP) estimator and a RELAX-based iterative method to accurately extract off-grid range and velocity parameters from these sparse pilots.
- Furthermore, we introduce the density-based spatial clustering of applications with noise (DBSCAN) based algorithm that operates in the joint range-velocity-angle domain. This method effectively identifies and prunes redundant detections of a single target that arise from the intentional overlap of adjacent beams during the SSB/SIB1 sweep.
- Finally, comprehensive simulation results are presented to validate the effectiveness of the proposed framework. The results demonstrate that our two-stage approach achieves superior estimation accuracy. Notably, the proposed WUP-based refinement algorithm significantly outperforms conventional benchmark methods, including 2D-FFT, 2D-MUSIC, and CS-AN, particularly under the challenging conditions imposed by the sparse pilot configuration.

Notations: Lowercase letters, boldface lowercase letters, boldface uppercase letters, and calligraphy letters denote the scalars, vectors, matrices, and sets, respectively, i.e., a, $\mathbf{a}$, $\mathbf{A}$, and $\mathcal{A}$. The operators $(\cdot)^T$ and $(\cdot)^H$ represent the transpose and Hermitian transpose, respectively. The operator $\odot$ denotes the Hadamard (element-wise) product. $\mathbb{R}$ and $\mathbb{C}$ denote the sets of real and complex numbers. $[\mathbf{A}]_{i,l}$ denotes the element in the i-th row and l-th column of matrix $\mathbf{A}$. $|\cdot|$ represents the absolute value of a scalar or the cardinality of a set. $\angle(\cdot)$ denotes the angle of a complex number. $\text{Diag}(\cdot)$ returns a diagonal matrix. $\text{vecl}(\cdot)$ stacks the strict lower triangular elements of a matrix into a vector. $\text{round}(\cdot)$ rounds to the nearest integer. $\cup$ denotes the set union operation.

II. SYSTEM MODEL

In this paper, we consider a mono-static MIMO-OFDM ISAC system designed for simultaneous downlink communication and radar parameter estimation. The BS is equipped with separate uniform linear arrays (ULAs), consisting of N_t transmit and N_r receive antennas, respectively.[2] As illustrated in Fig. 1, the BS first transmits SSB beams to scan the coverage area, immediately followed by Type#0-PDCCH and SIB1 signals with the same beam pattern. According to the prior studies [26], [35], these broadcast signals are utilized for a dual role. On one hand, they deliver the information necessary for uplink RACH procedures. On the flip side, they are adopted as sensing signals.

---

[2]We assume that the BS operates under an advanced full-duplex regime, where self-interference (SI) at the receive antennas is effectively canceled.

A. Transmitted Signal Model

Let $l \in \{1, \cdots, N_s\}$ and $i \in \{1, \cdots, N_c\}$ denote the indices for the OFDM symbols and subcarriers, respectively. Based on the frequency-flat assumption, the signal vector transmitted at the l-th symbol on the i-th subcarrier is modeled as [3]

$$[\mathbf{X}]_{i,l} = \mathbf{w}[l][\mathbf{S}]_{i,l}, \tag{1}$$

where $\mathbf{S} \in \mathbb{C}^{N_c \times N_s}$ denotes the transmitted matrix, and $\mathbf{w}[l] \in \mathbb{C}^{N_t}$ denotes the precoding vector for the l-th symbol. Specifically, the vector $\mathbf{w}[l]$ is selected from a predefined codebook $\mathcal{W}$ to steer the transmission towards a specific direction. This codebook, $\mathcal{W}$, is constructed from an oversampled discrete fourier transform (DFT) grid to achieve the fine-grained angular coverage. Specifically, given an integer oversampling factor O, the codebook is defined as $\mathcal{W} = \{\mathbf{w}_q\}_{q=0}^{Q-1}$, where $Q = O \cdot N_t$ and the q-th beamforming vector is given by

$$\mathbf{w}_q = \frac{1}{\sqrt{N_t}} \left[1, e^{j\frac{2\pi q}{Q}}, e^{j\frac{2\pi 2q}{Q}}, \ldots, e^{j\frac{2\pi (N_t-1)q}{Q}} \right]^T. \tag{2}$$

Then, after the inverse fast fourier transform (IFFT) operation at the transmitter, the baseband time-domain OFDM signal can be expressed as

$$\mathbf{x}(t) = \sum_{l=0}^{N_s-1} \sum_{i=0}^{N_c-1} [\mathbf{X}]_{i,l}\, e^{j2\pi i \Delta f t}\, r(t - lT) \in \mathbb{C}^{N_t}, \tag{3}$$

where Δf is the subcarrier spacing, $r(t)$ denotes the rectangular pulse, and $T = T_c + T_{cp}$ is the total duration of an OFDM symbol. Here, T_c denotes the OFDM symbol duration and T_{cp} denotes the cyclic prefix (CP) duration.

B. Sensing Signal Model

Let $\tilde{\mathbf{y}}(t) \in \mathbb{C}^{N_r \times 1}$ denote the received echoes from N_{target} unknown targets. The received signal model can be written as

$$\tilde{\mathbf{y}}(t) = \sum_{n=1}^{N_{\text{target}}} \beta_n e^{j2\pi f_{D,n} t} \mathbf{b}_r(\theta_n) \mathbf{a}_t^H(\theta_n) \mathbf{x}(t - \tau_n) + \mathbf{z}(t), \tag{4}$$

where the components are defined as follows:

- $\tilde{\mathbf{y}}(t) = [\tilde{y}(1,t), \ldots, \tilde{y}(N_r, t)]^T \in \mathbb{C}^{N_r}$ is the received signal vector at the N_r receive antennas.
- $\tau_n = 2R_n/c$ is the round-trip propagation delay for the n-th target.
- $\mathbf{x}(t) \in \mathbb{C}^{N_t \times 1}$ is the vector of transmitted baseband signals from the N_t transmit antennas.
- $\mathbf{a}_t(\theta_n) \in \mathbb{C}^{N_t \times 1}$ is the transmit array steering vector towards the direction θ_n, with $[\mathbf{a}_r(\theta_n)]_m = e^{-j2\pi(m-1)d_t \sin(\theta_n)/\lambda_c}$.
- $\mathbf{b}_r(\theta_n) \in \mathbb{C}^{N_r \times 1}$ is the receive array steering vector for the direction θ_n, with $[\mathbf{b}_r(\theta_n)]_m = e^{-j2\pi(m-1)d_r \sin(\theta_n)/\lambda_c}$.
- $\mathbf{z}(t) = [z(1,t), \ldots, z(N_r,t)]^T \in \mathbb{C}^{N_r}$ is the additive white Gaussian noise (AWGN) vector.

---

[3]Following [36], the signal bandwidth in this work is assumed to be contained within a single subband. Consequently, a common precoding vector is applied across all subcarriers for a given OFDM symbol.
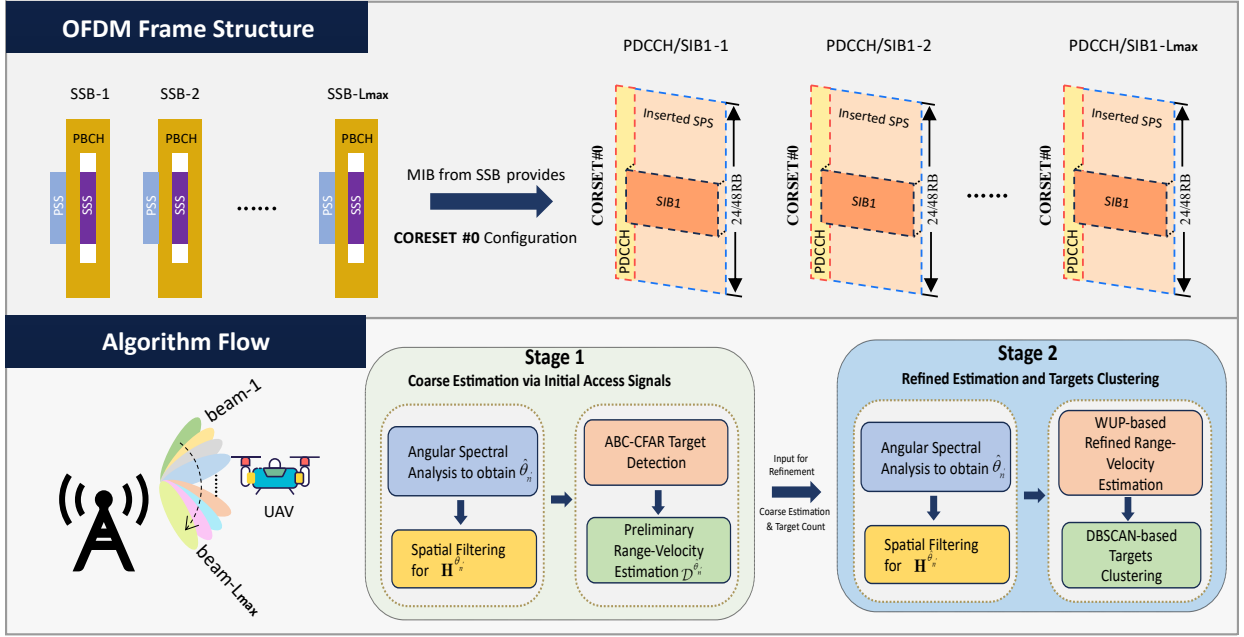
Fig. 1: Frame structure and algorithm flow for the proposed two-stage sensing framework.

After the analog-to-digital conversion (ADC) with a sampling frequency of F_s and the subsequent CP removal, the discrete echo signal vector for the l-th symbol $\hat{\mathbf{y}}(j,l) \in \mathbb{C}^{N_r \times 1}$ is obtained as

$$\hat{\mathbf{y}}(j,l) \triangleq \tilde{\mathbf{y}}(lT + j/F_s + T_{CP}). \tag{5}$$

Subsequently, following serial-to-parallel conversion, an N_s-point FFT is applied to $\hat{\mathbf{y}}(j,l)$ along the sample index dimension. As derived in Appendix A, the resulting frequency-domain OFDM symbol $\mathbf{y}(i,l)$ is given by

$$\mathbf{y}(i,l) = \sum_{n=1}^{N_{target}} \beta_n \mathbf{a}_t^H(\theta_n)[\mathbf{X}]_{i,l}\, e^{ji\omega_r(R_n)} e^{jl\omega_v(v_n)} \mathbf{b}_r(\theta_n) + \mathbf{z}(i,l), \tag{6}$$

where $\omega_r(R_n) \triangleq -4\pi\Delta f R_n/c$ and $\omega_v(v_n) \triangleq 2\pi f_{D,n}T$ are the normalized frequencies corresponding to the range and velocity of the n-th target, respectively. The notation $\mathbf{z}(i,l)$ denotes the noise term, which is approximated as AWGN [10], [37].

III. STAGE I: COARSE TARGET DETECTION AND PARAMETER ESTIMATION

In this section, we elaborate on the utilization of the SSB signal for target detection and coarse parameter estimation. Specifically, each SSB block occupies 20 resource blocks (RBs) and four consecutive OFDM symbols in the time-frequency domain and is broadcast periodically, as illustrated in Fig. 1. We exploit the inherent beam-sweeping nature of SSB to obtain coarse estimates of the target's angular, range, and velocity parameters.

A. Angular Spectral Analysis

We begin by addressing the angular estimation. To perform Direction of Arrival (DoA) estimation for targets within the k-th SSB beam centered at θ_k, we employ the high-resolution multiple signal classification (MUSIC) algorithm [10]. Firstly, we average the outer product of $\mathbf{y}(i,l)$ across all available subcarriers and OFDM symbols

$$\hat{\mathbf{R}} = \frac{1}{|\mathcal{R}_{k,SSB}|} \sum_{\{(i,l)\} \in \mathcal{R}_{k,SSB}} \mathbf{y}(i,l)\mathbf{y}(i,l)^H, \tag{7}$$

where $\mathcal{R}_{k,SSB}$ is the set of REs for the k-th SSB beam and $|\mathcal{R}_{k,SSB}|$ denotes the cardinality of this set.

Then, we perform the eigenvalue decomposition as

$$\hat{\mathbf{R}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H, \tag{8}$$

where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_{N_r}]$ is the matrix of eigenvectors and $\mathbf{\Lambda} = \mathrm{Diag}(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N_r})$ is the diagonal matrix containing the corresponding eigenvalues, sorted in descending order.

Based on the auto-correlation coefficient matrix algorithm [38], the number of potential UAVs N_{θ_k} can be effectively determined[4]. Once N_{θ_k} is estimated, the noise subspace matrix $\bar{\mathbf{U}}$ is constructed as $\bar{\mathbf{U}} = [\mathbf{u}_{N_{\theta_k}+1}, \ldots, \mathbf{u}_{N_r}]$. Then, by exploiting the orthogonality between the array steering vectors and the noise subspace, the DoA of the potential UAVs in the current beam is given by

$$\{\hat{\theta}_n\}_{n=1}^{N_{\theta_k}} = \underset{\theta}{\mathrm{argpeak}} \left(\frac{1}{\mathbf{b}_r^H(\theta)\bar{\mathbf{U}}\bar{\mathbf{U}}^H\mathbf{b}_r(\theta)}, N_{\theta_k} \right), \tag{9}$$

[4]Unlike the ideal assumption in [10], [11], the limited samples per SSB beam cause information-theoretic criteria to perform poorly; MDL [39] becomes inaccurate while AIC tends to overestimate. We therefore employ the more robust auto-correlation coefficient matrix method. The interested reader is referred to [38] for further details.

$$
\begin{aligned}
[\mathbf{H}^{\hat{\theta}_{n'}}]_{i,l} &= \frac{\mathbf{b}_{\mathrm{r}}^{H}(\hat{\theta}_{n'})\mathbf{y}(i,l)}{\mathbf{a}_{\mathrm{t}}^{H}(\hat{\theta}_{n'})[\mathbf{X}]_{i,l}} \\
&= \frac{\mathbf{b}_{\mathrm{r}}^{H}(\hat{\theta}_{n'})\left(\sum_{n=1}^{N_{\mathrm{target}}}\beta_n\mathbf{a}_{\mathrm{t}}^{H}(\theta_n)[\mathbf{X}]_{i,l}e^{ji\omega_r(R_n)}e^{jl\omega_v(v_n)}\mathbf{b}_{\mathrm{r}}(\theta_n)+\mathbf{z}(i,l)\right)}{\mathbf{a}_{\mathrm{t}}^{H}(\hat{\theta}_{n'})[\mathbf{X}]_{i,l}} \\
&= \underbrace{\sum_{k\in\mathcal{K}_{n'}}G_{n',k}\beta_k e^{ji\omega_r(R_k)}e^{jl\omega_v(v_k)}}_{\text{Desired Signals from Target Cluster}} + \underbrace{\sum_{n\notin\mathcal{K}_{n'}}\beta_n\frac{\mathbf{a}_{\mathrm{t}}^{H}(\theta_n)[\mathbf{X}]_{i,l}}{\mathbf{a}_{\mathrm{t}}^{H}(\hat{\theta}_{n'})[\mathbf{X}]_{i,l}}\left(\mathbf{b}_{\mathrm{r}}^{H}(\hat{\theta}_{n'})\mathbf{b}_{\mathrm{r}}(\theta_n)\right)e^{ji\omega_r(R_n)}e^{jl\omega_v(v_n)}}_{\text{Interference from Other Targets}} + \underbrace{\frac{\mathbf{b}_{\mathrm{r}}^{H}(\hat{\theta}_{n'})\mathbf{z}(i,l)}{\mathbf{a}_{\mathrm{t}}^{H}(\hat{\theta}_{n'})[\mathbf{X}]_{i,l}}}_{\text{Noise}}
\end{aligned}
\tag{10}
$$

where the search window $\theta \in [\theta_k - \Delta\theta, \theta_k + \Delta\theta]$, with $\Delta\theta \approx 85°/(N_t\cos(\theta_k))$ denoting the approximate half-beamwidth of the ULA at the suppression level -10 dB [40][5]. The operator argpeak$(\cdot, N)$ finds the locations of the N largest peaks.

B. Spatial Filtering

For each estimated DoA $\hat{\theta}_{n'} \in \{\hat{\theta}_n\}_{n=1}^{N_{\theta_k}}$ obtained from the preceding angular spectral analysis, we first perform the spatial filtering and remove the signal-dependent coefficients $\mathbf{a}_{\mathrm{t}}^{H}(\hat{\theta}_{n'})[\mathbf{X}]_{i,l}$ [41]. The resulting 2D channel response $[\mathbf{H}^{\hat{\theta}_{n'}}]_{i,l}$ is given in (10), which is shown at the top of this page. In this model, $\mathcal{K}_{n'}$ represents the index set of targets whose DoAs are clustered around the estimated angle $\hat{\theta}_{n'}$.

Due to the signal-to-noise ratio (SNR) gains introduced by the spatial filter, both the interference from other target clusters ($n \notin \mathcal{K}_{n'}$) and the noise term are significantly suppressed. Thus, the model in (10) can be approximated by its dominant signal components

$$
[\mathbf{H}^{\hat{\theta}_{n'}}]_{i,l} \approx \sum_{k\in\mathcal{K}_{n'}}\alpha_k e^{ji\omega_r(R_k)}e^{jl\omega_v(v_k)}, \tag{11}
$$

where $\alpha_k = G_{n',k}\beta_k$ is the effective complex amplitude of the k-th target within $\mathcal{K}_{n'}$.

Remark1: In practical systems, the half-power beamwidths of adjacent beams overlap to ensure effective user coverage. Consequently, the resource set for estimation can be expanded, e.g., $\mathcal{R}_{k,\mathrm{SSB}}^{\mathrm{ext}} = \mathcal{R}_{k-1,\mathrm{SSB}} \cup \mathcal{R}_{k,\mathrm{SSB}} \cup \mathcal{R}_{k+1,\mathrm{SSB}}$. The expansion introduces a trade-off: a larger resource set provides more samples but suffers from lower beamforming gain at the periphery. The reduced beamfoming gain $\mathbf{a}_{\mathrm{t}}^{H}(\hat{\theta}_{n'})[\mathbf{X}]_{i,l}$ increases the relative power of the interference and the noise terms in (10). Therefore, the estimation window must be judiciously configured to balance the number of samples against the degradation in effective signal to interference plus noise ratio (SINR).[6]

[5]The overlap between the half-power beamwidths of adjacent SSB beams may cause a single target to be detected multiple times. However, the clustering method is introduced in Section V to prune redundant detections and enhance the overall estimation accuracy.

[6]Due to the beam sweeping of the SSB, a target is only coherently illuminated during a subset of OFDM symbols. Processing the entire set of symbols would incorporate noise from irrelevant time instances where the beam was pointing elsewhere, thereby degrading the effective SNR.
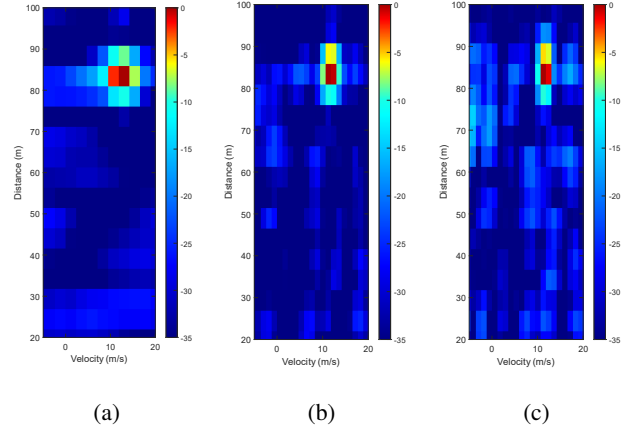

Fig. 2: RDMs with different numbers of beams: (a) 1 beams, (b) 3 beams, and (c) 5 beams.

C. Preliminary Range and Velocity Estimation

Before detailing our estimation algorithm, we first clarify two primary limitations inherent to the SSB's structure.

1) **Limited Velocity Estimation Fidelity:** The achievable velocity resolution is fundamentally limited by the short observation interval, as a UAV target is typically illuminated by only a few four-symbol SSB blocks.

2) **Limited Bandwidth and Range Resolution:** The spectral bandwidth of the SSB is confined to 240 subcarriers. This restricted bandwidth inherently limits the achievable range resolution.

To address these limitations, we further leverage the Type#0-PDCCH and SIB1 broadcasting signals, which offer a more extended observation interval and wider bandwidth than SSB, to enhance velocity estimation and range resolution. Firstly, the range-Doppler map (RDM) of the SSB is generated by performing 2D-FFT on the approximated channel matrix $\mathbf{H}^{\hat{\theta}_{n'}}$ as

$$
[\mathbf{R}_{\mathrm{SSB}}^{\hat{\theta}_{n'}}]_{p,q} = \left|\sum_{(i,l)\in\mathcal{R}_{k,\mathrm{SSB}}^{\mathrm{ext}}}[\mathbf{H}]_{i,l}^{\hat{\theta}_{n'}}e^{-j\frac{2\pi pi}{N_r}}e^{-j\frac{2\pi ql}{N_d}}\right|, \tag{12}
$$

where N_r and N_d are the FFT points for the range and doppler dimensions, respectively, with zero-padding applied

to improve resolution. Similarly, we obtain the RDM of SIB1 as $\mathbf{R}_{\text{SIB1}}^{\hat{\theta}_{n'}}$. The wider-bandwidth PDCCH, in turn, yields a high-resolution range profile, defined as the vector $\mathbf{p}_{\text{PDCCH}}^{\hat{\theta}_{n'}}$. These components are then fused to form the combined RDM as

$$\mathbf{R}_{\text{comb}}^{\hat{\theta}_{n'}} = \mathbf{R}_{\text{SSB}}^{\hat{\theta}_{n'}} \odot \mathbf{R}_{\text{SIB1}}^{\hat{\theta}_{n'}} \odot \mathbf{P}_{\text{PDCCH}}^{\hat{\theta}_{n'}}, \qquad (13)$$

where $\mathbf{P}_{\text{PDCCH}}^{\hat{\theta}_{n'}} = \mathbf{p}_{\text{PDCCH}}^{\hat{\theta}_{n'}} \mathbf{1}_{N_d}^T$, with $\mathbf{1}_{N_d}$ denoting a row vector of N_d ones. To visualize the combined RDM, Fig. 2 illustrates the combined RDM obtained with different numbers of beams used for angular estimation. It is clear that the 3-beam combination (Fig. 2b) achieves superior resolution over the single-beam case (Fig. 2a), while avoiding the noise-floor degradation apparent in the 5-beam scenario (Fig. 2c). Additionally, inaccuracies in the initial angle estimation often cause the energy of a point-based UAV target to disperse across a certain area in the RDM. Therefore, rather than the cell-average (CA) constant false-alarm rate (CFAR), we employ the area-based combination CFAR (ABC-CFAR) framework [42] for target detection. The resulting set of estimated ranges and velocities for the detected targets can be expressed as

$$\mathcal{D}^{\hat{\theta}_{n'}} = \{(\hat{R}_m, \hat{v}_m, \hat{\theta}_{n'})\}_{m=1}^M, \qquad (14)$$

where M is the total number of detected targets and $(\hat{R}_m, \hat{v}_m, \hat{\theta}_{n'})$ represents the estimated range-velocity-angle pair for the m-th target.

IV. SPARSE PILOT STRUCTURE FOR REFINED ESTIMATION

In this section, we present the necessary definitions and common structures of the SPS. The design of this SPS is directly motivated by 3GPP specifications [12]–[15] and [26], which indicate that the PDSCH carrying SIB1 typically occupies only a portion of the CORESET#0 bandwidth. This characteristic, combined with inspiration from 1D sparse array theory [43], leads to our PTRS-like SPS configuration where pilots are sparsely arranged in the frequency domain. The approach yields two key advantages: i) spanning a wide effective bandwidth to improve range resolution; ii) multiplexing communication pilots to facilitate initial access. We now introduce the necessary definitions for this SPS.

A. Pilots Signal Structure

We first introduce the structure of the pilots, which are sparsely arranged in the frequency domain as illustrated in Fig. 3.[7]

Definition 1 (Frequency-Domain Sparse Pilot Structure): For a set $\mathcal{S}_f$ of subcarrier indices where pilots are placed, let $\mathcal{D}_f$ denote the corresponding Difference Co-spectrum (DCS);
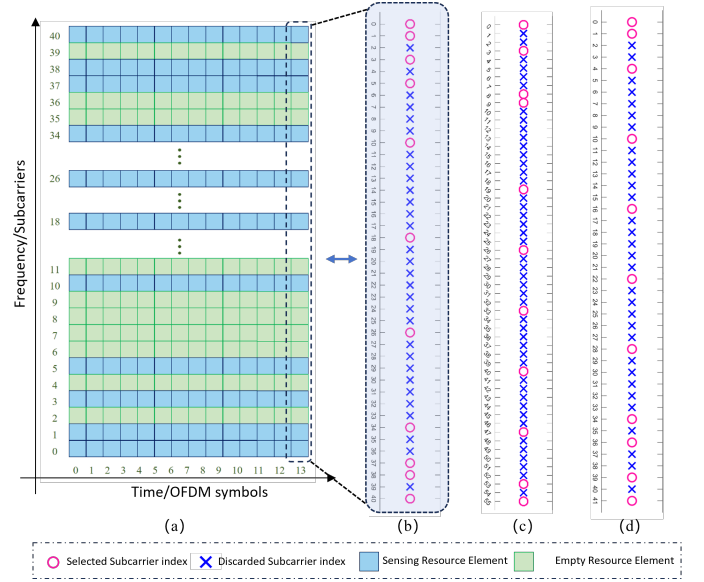
Fig. 3: Example of the Sparse Pilot Structure.

$\mathcal{U}_f$ denote the virtual uniform spectrum and $\mathcal{C}(k)$ denote all pilot pairs contributing to the frequency lag k,

$$\begin{aligned} \mathcal{S}_f &= \{p_1, p_2, \ldots, p_M\}, \\ \mathcal{D}_f &= \{p_i - p_j \mid p_i, p_j \in \mathcal{S}_f\}, \\ \mathcal{U}_f &= [-L_u, \ldots, L_u] \subseteq \mathcal{D}_f, \\ \mathcal{C}_f(k) &= \{(p_i, p_j) \mid p_i - p_j = k; p_i, p_j \in \mathcal{S}_f\}. \end{aligned} \qquad (15)$$

Here, each $p_m \in \mathcal{S}_f$ represents an individual subcarrier index, and L_u is the maximum one-sided lag of the DCS $\mathcal{U}_f$. Then, $|\mathcal{D}_f|$, $2L_u + 1$, and $|\mathcal{C}_f(k)|$ are referred to the degrees of freedom (DoF), the uniform DoF (uDoF), and the weight function $w(k)$ w.r.t. the $\mathcal{S}_f$ respectively. Similar to their roles in the DoA estimation problem, it will be shown in the subsequent analysis that the DoF, uDoF, and weight function of our sparse sensing scheme are instrumental in the performance of time-delay estimation.

Definition 2 (ANAII-2 Sparse Pilot Structure) [44]: An Augmented Nested Array with two levels and two-sided extensions (ANAII-2) for a given total of N sensors is defined by the set of sensor positions $\mathcal{S}$. First, the intermediate parameters L and M are determined from N as

$$L = \text{round}(N/6), \quad M = N - 4L. \qquad (16)$$

The set $\mathcal{S}$ is then constructed as the union of the following five sub-sets, $\mathcal{S} = \mathcal{S}_{L11} \cup \mathcal{S}_{L12} \cup \mathcal{S}_M \cup \mathcal{S}_{R21} \cup \mathcal{S}_{R22}$, where
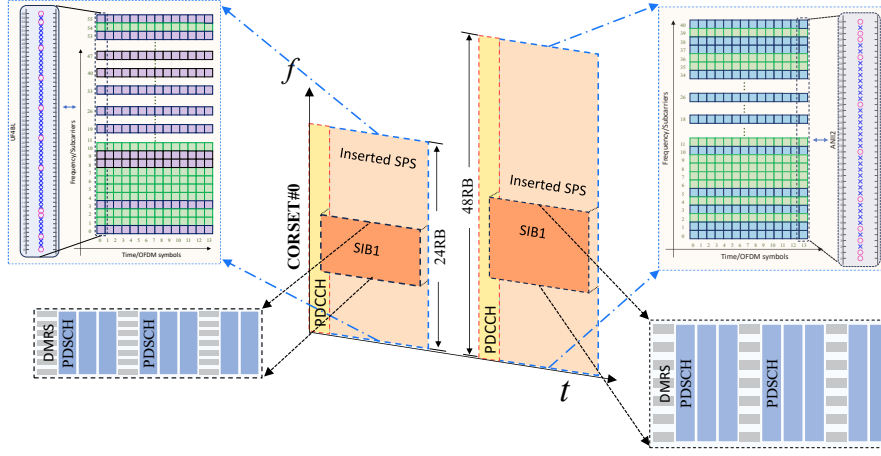
Fig. 4: Example of the SPS Configuration.

each sub-set is explicitly defined by

$$
\begin{aligned}
\mathcal{S}_M &= \left\{ 2L^2 + L + 4Ll \mid l \in \{0, \ldots, M\} \right\}, \\
\mathcal{S}_{L12} &= \left\{ 2L^2 + L - (2L+1)l \mid l \in \{0, \ldots, L-1\} \right\}, \\
\mathcal{S}_{L11} &= \left\{ 2L + 1 - l \mid l \in \{0, 2, \ldots, 2L\} \cup \{2L+1\} \right\}, \\
\mathcal{S}_{R21} &= \left\{ 2L^2 + L + 4LM + (2L-1)l \mid l \in \{0, \ldots, L-1\} \right\}, \\
\mathcal{S}_{R22} &= \left\{ \begin{array}{l} 4L^2 - 2L + 1 + 4LM + l \mid \\ l \in \{0\} \cup \{1, 3, \ldots, 2L-1\} \end{array} \right\}.
\end{aligned}
\tag{17}
$$

The resulting set $\mathcal{S}$ is normalized such that its minimum element is 0.

Definition 3 (UF-4BL Sparse Pilot Structure) [45]: A Uniform Fitting 4-Block-Level (UF-4BL) array with a total of N sensors is defined by the set of sensor positions $\mathcal{S}$. The structure is determined by two intermediate parameters, N_b and N_t, which are derived from N as:

$$
N_b = \lfloor (N - 8)/8 \rfloor, \quad N_t = N - 4N_b - 6. \tag{18}
$$

The final set of frequency positions $\mathcal{S}$ is constructed as the union of eight distinct sub-ULAs, $\mathcal{S} = \bigcup_{i=1}^{8} \mathcal{S}_i$, where each sub-array is explicitly defined by:

$\mathcal{S}_1 = \{3l \mid l = 0, 1\}$

$\mathcal{S}_2 = \{7 + 4l \mid l = 0, 1, \ldots, N_b - 1\}$

$\mathcal{S}_3 = \{(4N_b + 8) + l \mid l = 0, 1\}$

$\mathcal{S}_4 = \{(4N_b + 15) + 4l \mid l = 0, 1, \ldots, N_b - 1\}$

$\mathcal{S}_5 = \{(8N_b + 19) + (4N_b + 7)l \mid l = 0, 1, \ldots, N_t - 1\}$

$\mathcal{S}_6 = \{(4N_t N_b + 7N_t + 4N_b + 19) + 4l \mid l = 0, 1, \ldots, N_b - 1\}$

$\mathcal{S}_7 = \{(4N_t N_b + 7N_t + 8N_b + 18) + 2l \mid l = 0, 1\}$

$\mathcal{S}_8 = \{(4N_t N_b + 7N_t + 8N_b + 25) + 4l \mid l = 0, 1, \ldots, N_b - 1\}$

Based on the comparison in Table I, we recommend the UF-4BL structure for the 24RB configuration and ANAII-2 for the 48RB configuration, respectively. For the 24RB case, both structures offer comparable uDoF, making the UF-4BL preferable due to its simpler weight function and lower coupling. Conversely, for the 48RB case, the uDoF of ANAII-2

TABLE I: Comparison of Weight Functions and uDoF for Different Arrays

Array Type	$w(1)$	$w(2)$	$w(3)$	uDoF
ANAII-2 [44]	2	$N/3$	2	$\mathcal{O}(2N^2/3)$
MISC [46]	1	$2\lfloor N/4 \rfloor - 2$	$1/2$	$\mathcal{O}(N^2/2 + 3N)$
TSENA [47]	$N_1 - 6$	$N_1 - 8$	$N_1 - 9$	$\mathcal{O}(N^2/2 + 3.5N)$
UF-3BL [45]	1	1	$3N_b - 1$	$\mathcal{O}(N^2/2 + 2N)$
UF-4BL [45]	1	1	2	$\mathcal{O}(N^2/2 + 2N)$

Note: For TS-ENA array, the parameter $N_1 = 2\lfloor (N+3)/4 \rfloor$. For UF-3BL, the parameter $N_b = \lfloor (N-5)/6 \rfloor$.

significantly surpasses that of UF-4BL. This substantial uDoF gain outweighs its higher coupling, rendering it the superior choice for performance.

B. Integration of the Pilots and SIB1

In this subsection, we describe how the aforementioned SPS is incorporated into the SIB1 block. Let us denote the SIB1 data symbol matrix and the sparse sensing pilot matrix as $\mathbf{S}_{\text{SIB1}} \in \mathbb{C}^{N_{c,\text{SIB1}} \times N_{s,\text{SIB1}}}$ and $\mathbf{S}_{\text{sen}} \in \mathbb{C}^{N_{c,\text{SIB1}} \times N_{s,\text{SIB1}}}$, respectively. The placement of these signals is governed by their corresponding binary masks $\mathcal{L}_{k,\text{SIB1}}$ and $\mathcal{L}_{k,\text{sen}}$, which belong to the set $\{0,1\}^{N_{c,\text{SIB1}} \times N_{s,\text{SIB1}}}$. For the sensing pilot mask $\mathcal{L}_{k,\text{sen}}$, we adopt a configuration analogous to the PT-RS in 5G NR. Specifically, the pilots occupy all OFDM symbols within the SIB1 block but are placed only on the sparse set of subcarrier indices $\mathcal{S}_f$, as described in Section IV-A.

The final transmitted matrix $\mathbf{S}_{\text{tx}} \in \mathbb{C}^{N_{c,\text{SIB1}} \times N_{s,\text{SIB1}}}$ is formed by having the SIB1 symbols overwrite the sensing pilots on any shared REs. This is expressed as

$$
\mathbf{S}_{\text{tx}} = \mathbf{S}_{\text{SIB1}} + \mathbf{S}_{\text{sen}} \odot (\mathbf{J} - \mathcal{L}_{k,\text{SIB1}}), \tag{19}
$$

where $\mathbf{J} \in \mathbb{C}^{N_{c,\text{SIB1}} \times N_{s,\text{SIB1}}}$ is an all-ones matrix of the same size. The overwrite mechanism is illustrated in Fig. 4, where SIB1 symbols take precedence over sensing pilots. Here, we denote the set of carrier frequencies jointly occupied by both SIB1 and the sensing pilots as $\mathcal{S}_{f,\text{combined}}$.

V. STAGE II: REFINED PARAMETER ESTIMATION WITH SPARSE PILOT SYMBOLS

In this section, we refine the coarse parameter estimates obtained in Stage I by applying the RELAX-based WUP algorithm. Then, the parameter pairing is introduced to identify the most probable range-velocity pairings. Finally, we employ the DBSCAN algorithm to prune redundant targets and obtain the final estimates.

A. Weighted Unwrapped Phase Velocity Estimation

We begin by considering the channel response matrix $\mathbf{H}^{\hat{\theta}_{n'}}$ associated with a specific angle $\hat{\theta}_{n'}$, assuming it corresponds to a single target k. For $i \in \mathcal{S}_{f,\text{combined}}$, we define the Doppler vector[8]

$$\mathbf{a}_D(v_k) = [\mathbf{H}^{\hat{\theta}_{n'}}]_{i,:} = \bar{\alpha}_k[e^{j0\omega_v(v_k)}, \ldots, e^{j(N_{s,\text{SIB1}}-1)\omega_v(v_k)}]. \tag{20}$$

Then, the Doppler covariance vector $\mathbf{r}^{f_D} \in \mathbb{C}^{N_{\text{vec}} \times 1}$ is constructed as

$$\mathbf{r}^{f_D} = \text{vecl}(\mathbf{a}_D(v_k)^H \mathbf{a}_D(v_k)), \tag{21}$$

where $N_{\text{vec}} = N_{s,\text{SIB1}}(N_{s,\text{SIB1}}-1)/2$. Its n-th element is given by $[\mathbf{r}^{f_D}]_n = |\bar{\alpha}_k|^2 e^{j(n-C_n)\omega_v(v_k)}$, where $C_n = \max\{C_t | t \geq 0 \text{ and } C_t < n\}$ and $C_t = tN_{s,\text{SIB1}} - t(t+1)/2$.

Denoting the wrapped phase vector as $\mathbf{p} = \angle\mathbf{r}^{f_D} \in \mathbb{C}^{1 \times N_{s,\text{SIB1}}}$, its n-th element is $[\mathbf{p}]_n = (n - C_n)\omega_v(v_k)$ (mod 2π). Based on the coarse velocity estimate $\omega_v(v_{\text{FFT}})$ obtained in Section III, we define the reference unwrapped phase vector $\mathbf{p}_{\text{FFT}} \in \mathbb{C}^{N_{\text{vec}} \times 1}$ with its n-th element given by

$$[\mathbf{p}_{\text{FFT}}]_n = (n - C_n)\omega_v(v_{\text{FFT}}). \tag{22}$$

As such, the integer wrap vector $\boldsymbol{\kappa}$, whose elements represent the number of phase wraps, can be estimated by

$$\boldsymbol{\kappa} = \text{round}\left(\frac{\mathbf{p}_{\text{FFT}} - \mathbf{p}}{2\pi}\right), \tag{23}$$

where $\text{round}(\cdot)$ is an element-wise rounding operation. The true unwrapped phase vector $\tilde{\mathbf{p}}$ is then recovered as

$$\tilde{\mathbf{p}} = \mathbf{p} + 2\pi\boldsymbol{\kappa}. \tag{24}$$

Finally, given the unwrapped phase vector $\tilde{\mathbf{p}}$, the estimate of the Doppler frequency $\hat{\omega}_v(v_k)$ is obtained as

$$\hat{\omega}_v(v_k) = \frac{\sum_{n=1}^{N_{\text{vec}}}(n - C_n)[\tilde{\mathbf{p}}]_n}{\sum_{n=1}^{N_{\text{vec}}}(n - C_n)^2}. \tag{25}$$

Remark 2: Observing (21), we note an analogy between the phase differences captured by the cross-correlation operation and the virtual difference co-arrays exploited in DoA estimation [44], [48]. As depicted in Fig. 3, the considered PT-RS style reference signal structure, which is dense across OFDM symbols, is analogous to the ULA. Furthermore, the weighted unwrapped estimator in (25) inherently assigns higher weights to phase differences corresponding to larger effective time

[8]For notational simplicity, we omit the subcarrier index i dependence in the following derivations, as the process is identical for each $i \in \mathcal{S}_{f,\text{combined}}$.

separations (i.e., larger values of $n - C_n$), as these generally provide finer velocity resolution. This weighting strategy aligns conceptually with the objectives in the non-ULA design, which often aims for higher DoF and lower mutual coupling.

B. RELAX-Based Multi-Target Velocity Estimation

We now extend the estimation framework to the multi-target scenario, where multiple targets may share the same angle bin $\hat{\theta}_{n'}$. In this case, the i-th row of the channel response matrix is the superposition of $|\mathcal{K}_{n'}|$ signals :

$$[\mathbf{H}^{\hat{\theta}_{n'}}]_{i,:} = \sum_{k=1}^{|\mathcal{K}_{n'}|} \alpha_k[e^{j0\omega_v(v_k)}, \ldots, e^{j(N_{s,\text{SIB1}}-1)\omega_v(v_k)}]. \tag{26}$$

To estimate the parameters of each target individually, we adopt the estimate-subtract strategy inspired by the RELAX algorithm [48]. This iterative method estimates the parameters of one target at a time while treating the others as interference. Specifically, assuming that the parameters of $Q - 1$ targets have already been estimated in $\mathbf{H}^{\hat{\theta}_{n'}}$, the signal component corresponding to the k-th target ($k \leq Q$) is extracted as

$$\mathbf{s}_{f_D}^{(k)} = [\mathbf{H}^{\hat{\theta}_{n'}}]_{i,:} - \sum_{\bar{k}=1,\bar{k}\neq k}^{Q} \hat{\alpha}_{\bar{k}}[e^{j0\hat{\omega}_v(v_{\bar{k}})}, \ldots, e^{j(N_{s,\text{SIB1}}-1)\hat{\omega}_v(v_{\bar{k}})}], \tag{27}$$

where $\hat{\alpha}_{\bar{k}}$ and $\hat{\omega}_v(v_{\bar{k}})$ are the current estimates for the other targets. The Doppler frequency $\hat{\omega}_v(v_k)$ is then estimated from $\mathbf{s}_{f_D}^{(k)}$ using the weighted unwrapped phase (WUP) method described previously. Subsequently, the complex amplitude α_k can be estimated as

$$\hat{\alpha}_k = \frac{\mathbf{a}_D(\hat{\omega}_v(v_k))^* \mathbf{s}_{f_D}^{(k)}}{N_{s,\text{SIB1}}}, \tag{28}$$

where $\mathbf{a}_D(\hat{\omega}_v(v_k))$ is the estimated Doppler vector for the k-th target's Doppler frequency. This process is iterated, refining the estimates for all targets until all $Q = |\mathcal{K}_{n'}|$ targets have been identified. The detailed implementation is presented in Algorithm 1.

C. Range Estimation and Parameter Pairing

Drawing inspiration from research in the non-ULA design, we adopt the SPS for the pilots, as detailed in Section IV. Let $\mathcal{S}_{f,\text{combined}} = \{p_1, \ldots, p_{|\mathcal{S}_{f,\text{combined}}|}\}$ denote this sparse set of subcarrier indices. Assuming a single target k, we define the delay vector $\mathbf{a}_\tau(R_k) \in \mathbb{C}^{|\mathcal{S}_{f,\text{combined}}| \times 1}$ by extracting the channel response at the sparse subcarrier indices over a specific OFDM symbol l

$$\mathbf{a}_\tau(R_k) \approx \bar{\alpha}_k[e^{jp_1\omega_r(R_k)}, \ldots, e^{jp_{|\mathcal{S}_{f,\text{combined}}|}\omega_r(R_k)}]^T. \tag{29}$$

Following the weighted unwrapped phase estimator framework proposed in Sections V.A and V.B, the delay covariance vector $\mathbf{r}^\tau$ is constructed as

$$\mathbf{r}^\tau = \text{vecl}(\mathbf{a}_\tau(R_k)\mathbf{a}_\tau(R_k)^H). \tag{30}$$

Then, the corresponding unwrapped phase vector $\tilde{\mathbf{p}}^\tau$ can be expressed as

$$\tilde{\mathbf{p}}^\tau = \text{vecl}(\mathbf{R}_\tau)\omega_r(R_k), \tag{31}$$

Algorithm 1: RELAX-Based Iterative Velocity Estimation

1: **Input:** $\mathbf{s}_{f_D} = [\mathbf{H}^{\hat{\theta}_{n'}}]_{i,:}, |\mathcal{K}_{n'}|, \tau$.
2: **Initialization:** Obtain initial estimates
   $\{\hat{\omega}_v(v_k)\}_{k=1}^{|\mathcal{K}_{n'}|}, \{\hat{\alpha}_k\}_{k=1}^{|\mathcal{K}_{n'}|}, \Delta = \infty, \mathbf{y}_1 = \mathbf{s}_{f_D}$.
3: **for** $Q = 1$ to $|\mathcal{K}_{n'}|$ **do**
4:     Estimate $\hat{\omega}_v(v_Q)$ from $\mathbf{y}_Q$ according to (20)-(24).
5:     Estimate $\hat{\alpha}_Q$ from $\mathbf{y}_Q$ and $\hat{\omega}_v(v_Q)$ according to (28).
6:     Reset convergence check $\Delta = \infty$.
7:     **while** $\Delta \geq \tau$ **do**
8:         Store previous estimates $\{\omega_{\text{prev}}(v_k) = \hat{\omega}_v(v_k)\}_{k=1}^{Q}$
9:         **for** $k = 1$ to Q **do**
10:            Calculate residual signal
               $\mathbf{y}_k^{\text{refine}} = \mathbf{s}_{f_D} - \sum_{\bar{k}=1, \bar{k} \neq k}^{Q} \hat{\alpha}_{\bar{k}} \mathbf{a}_D(\hat{\omega}_v(v_{\bar{k}}))$
11:            Refine velocity $\hat{\omega}_v(v_k)$ from $\mathbf{y}_k^{\text{refine}}$ according to (20)-(24).
12:            Refine amplitude $\hat{\alpha}_k$ from $\mathbf{y}_k^{\text{refine}}$ and $\hat{\omega}_v(v_k)$ according to (28).
13:        **end for**
14:        Calculate maximum change
               $\Delta = \max_{k \in \{1..Q\}} \{||\hat{\omega}_v(v_k) - \omega_{\text{prev}}(v_k)||\}$
15:    **end while**
16:    Prepare residual for the next target's initial estimation
17:    $\mathbf{y}_{Q+1} = \mathbf{s}_{f_D} - \sum_{\bar{k}=1}^{Q} \hat{\alpha}_{\bar{k}} \mathbf{a}_D(\hat{\omega}_v(v_{\bar{k}}))$
18: **end for**
19: **Output:** Estimated velocities $\{\hat{\omega}_v(v_k)\}_{k=1}^{|\mathcal{K}_{n'}|}$.

where $[\mathbf{R}_\tau]_{i,j} = p_i - p_j$, with p_i and p_j being the i-th and j-th frequency indices from $\mathcal{S}_{f,\text{combined}}$, respectively. Notably, the vector $\text{vecl}(\mathbf{R}_\tau)$ contains the frequency index differences corresponding to the DCS, which is detailed in Definition 1. The subsequent steps for $\omega_r(R_k)$ proceed analogously to the velocity estimation.

Building upon the range and velocity estimation procedures outlined previously, we can obtain the sets of estimated normalized range frequencies $\{\hat{\omega}_r(R_k)\}_{k=1}^{|\mathcal{K}_{n'}|}$ and Doppler frequencies $\{\hat{\omega}_v(v_k)\}_{k=1}^{|\mathcal{K}_{n'}|}$. Next, parameter pairing is performed to associate the correct range and velocity estimates for each target. To this end, we first construct $\mathbf{\Gamma}$ containing all possible $|\mathcal{K}_{n'}|^2$ range-velocity pairings

$$\mathbf{\Gamma} = \{(\hat{\omega}_r(R_i), \hat{\omega}_v(v_j))\}_{i,j=1}^{|\mathcal{K}_{n'}|}, \qquad (32)$$

where for each pair $(\hat{\omega}_r(R_i), \hat{\omega}_v(v_j)) \in \mathbf{\Gamma}$, the matching score is calculated as

$$S_{\hat{\omega}_r(R_i), \hat{\omega}_v(v_j)}^{\hat{\theta}_{n'}} = \left| \mathbf{a}_\tau(\hat{R}_i)^H \mathbf{H}^{\hat{\theta}_{n'}} \mathbf{a}_D(\hat{v}_j) \right|. \qquad (33)$$

Subsequently, by identifying the $|\mathcal{K}_{n'}|$ pairs in $\mathbf{\Gamma}$ that yield the highest matching scores, the final set $\mathbf{\Xi}^{\hat{\theta}_{n'}}$ of paired range and velocity parameters for the $|\mathcal{K}_{n'}|$ targets is given by

$$\mathbf{\Xi}^{\hat{\theta}_{n'}} = \{(\hat{R}_k, \hat{v}_k)\}_{k=1}^{|\mathcal{K}_{n'}|}. \qquad (34)$$

D. DBSCAN-based Repeated Targets Pruning

In order to maintain continuous energy coverage, the half-power beamwidths of adjacent beams typically overlap during SSB and SIB1 beam sweeping. Consequently, a single UAV target may be detected by multiple beams. To address this redundancy, we employ the DBSCAN algorithm [49] to cluster the detected targets in the range-velocity-angle space. To mitigate the impact of spurious noise detections, the minimum number of points N_{min} in a cluster is set to 2.

As a prerequisite for clustering, we define the detected targets set $\mathbf{\Xi}$ across all detected angles $\{\hat{\theta}_{n'}\}$ as

$$\mathbf{\Xi} = \bigcup_{\hat{\theta}_{n'}} \{(\hat{R}_k, \hat{v}_k, \hat{\theta}_{n'}) \mid (\hat{R}_k, \hat{v}_k) \in \mathbf{\Xi}^{\hat{\theta}_{n'}}\}, \qquad (35)$$

where each element $\mathbf{s} \in \mathbf{\Xi}$ is a tuple $\mathbf{s} = (\hat{R}, \hat{v}, \hat{\theta})$. Next, we define the weighting matrix $\mathbf{W} = \text{Diag}(w_R, w_v, w_\theta)$, where the weights reflect the inverse squared resolutions: $w_R = (2N_{c,\text{SIB1}} \Delta f / c)^2$, $w_v = (2N_{s,\text{SIB1}} T_{\text{sym}} / \lambda)^2$ and $w_\theta \propto N_R^2$. Then, the weighted distance $d(\mathbf{s}_1, \mathbf{s}_2)$ between two points $\mathbf{s}_1, \mathbf{s}_2 \in \mathbf{\Xi}$ is calculated as

$$d(\mathbf{s}_1, \mathbf{s}_2) = \left((\mathbf{s}_1 - \mathbf{s}_2)^T \mathbf{W}(\mathbf{s}_1 - \mathbf{s}_2)\right)^{1/2}, \qquad (36)$$

where $(\mathbf{s}_1 - \mathbf{s}_2)$ represents the element-wise difference vector.

After applying DBSCAN with the distance metric and a suitable neighborhood radius ϵ, the algorithm groups the points in $\mathbf{\Xi}$ into clusters. Each resulting cluster ideally corresponds to multiple detections of a single, unique UAV target, while points classified as noise by DBSCAN are discarded. Specifically for each cluster $\mathcal{C}_m$ identified by DBSCAN, the final estimate $(\hat{R}_m^{\text{final}}, \hat{v}_m^{\text{final}}, \hat{\theta}_m^{\text{final}})$ for the m-th target is obtained as

$$(\hat{R}_m^{\text{final}}, \hat{v}_m^{\text{final}}, \hat{\theta}_m^{\text{final}}) = \frac{\sum_{\mathbf{s} \in \mathcal{C}_m} S_{\hat{\omega}_r(R), \hat{\omega}_v(v)}^{\hat{\theta}} \cdot \mathbf{s}}{\sum_{\mathbf{s} \in \mathcal{C}_m} S_{\hat{\omega}_r(R), \hat{\omega}_v(v)}^{\hat{\theta}}}, \qquad (37)$$

where the summation and division are performed element-wise for the tuple $\mathbf{s} = (\hat{R}, \hat{v}, \hat{\theta})$. This process yields the final pruned set of detected UAV parameters.

VI. NUMERICAL RESULTS

In this section, simulation results are presented to evaluate the performance of the proposed sensing framework. Firstly, we provide the simulation parameter settings. Then, we compare the performance of our proposed RELAX-based WUP algorithm with conventional techniques in a single-target scenario. Furthermore, we also evaluate the impact of SPS configurations. Finally, we compare our DBSCAN-based clustering algorithm with the benchmark pruning method in the multi-target scenario.

A. Simulation Settings

Unless stated otherwise, the simulation parameters are set as follows: We consider a DFRC BS operating at a carrier frequency of $f_c = 30$ GHz with $\Delta f = 120$ kHz. The BS is equipped with seperate half-wavelength ULAs with $N_t = 32$ and $N_r = 64$ antennas, respectively. Assuming

the normal CP, the 0th and 7th OFDM symbols have a CP duration of $T_{\rm cp} \approx 1.107$ μs with $T \approx 9.44$ μs, while the remaining 12 symbols have $T_{\rm cp} \approx 0.586$ μs with $T \approx 8.92$ μs. The RCS of each UAV is set to $\sigma = 0.05$ m^2. Moreover, the sensing procedure is sequential, comprising two distinct stages. Firstly, 64 SSB beams are transmitted to perform initial detection across the $[-45°, 45°]$ angular range. Secondly, SPS-augmented PDCCH/SIB1 signals are transmitted to provide the essential information for UE system access, while simultaneously enabling refined parameter estimation. All results are averaged over 200 Monte Carlo trials.

B. Estimation Accuracy and Computational Cost for Single Target

We first evaluate the performance of our proposed algorithm in the single-target scenario by comparing it against three benchmark methods[9]: 2D-FFT [10], 2D-MUSIC [31], 2D CS-AN [50]. These benchmarks represent the periodogram, subspace-based and compressed sensing methods, respectively. For this simulation, the range, radial velocity, and angle of the single target are set to (50 m, 25 m/s, 15°).
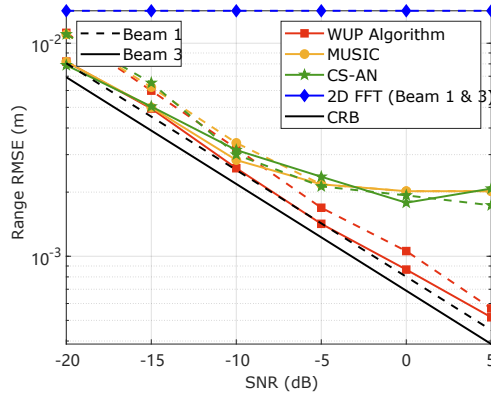

Fig. 5: RMSE of range estimation versus SNR.

Fig. 5 compares the range RMSE performance versus SNR for different algorithms. We can observe that the 2D-FFT method exhibits a almost fixed RMSE, limited by bandwidth resolution and unaffected by the number of beams. In contrast, both 2D-MUSIC and CS-AN show performance gains as SNR increases; however, their grid-based search leads to eventual error floors. Among these, CS-AN provides better range estimation than 2D-MUSIC. The proposed WUP algorithm significantly outperforms all benchmarks, achieving the lowest RMSE and highlighting its effective off-grid estimation capability under challenging sparse pilot conditions.

Fig. 6 presents the velocity RMSE versus SNR for different algorithms. A key observation is that utilizing multiple beams significantly enhances velocity estimation accuracy for all algorithms by extending the effective observation interval and thus improving velocity resolution. Notably, this performance

[9]To ensure reproducibility, the source code for all benchmark algorithms is provided at https://github.com/OnePieceofCakeforYou.
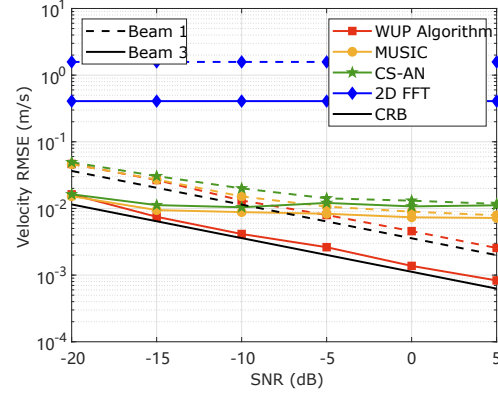

Fig. 6: RMSE of velocity estimation versus SNR.

gain is markedly stronger than that observed for range estimation. In contrast to the range estimation results, it is noted that the 2D-MUSIC slightly outperforms CS-AN in velocity estimation under the sparse frequency configuration. Furthermore, the proposed WUP algorithm demonstrates superior performance, consistently achieving the lowest RMSE across the entire simulated SNR range.
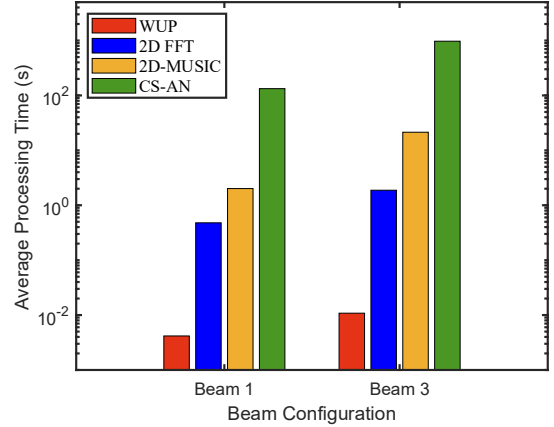

Fig. 7: Computational time comparison .

Fig. 7 demonstrates the average processing time of the proposed WUP algorithm against benchmark methods under different beam configurations. The results reveal a clear computational hierarchy, with WUP algorithm being the most efficient, followed in order by 2D-FFT, 2D-MUSIC, and the substantially more costly CS-AN. While the computational burden increases for all algorithms with more beams, the WUP algorithm maintains a marked efficiency edge. This advantage demonstrates its superior scalability compared to the other methods, particularly the highly complex CS-AN.

Fig. 8 illustrates the performance of range estimation versus transmit power under different array layouts, where the Fig. 8(a) and 8(b) correspond to the 24RB and 48RB CORSET#0 configurations, respectively. Since the frequency-domain configuration schemes primarily affect range estimation, only the ranging RMSE is evaluated. As shown in Fig. 8(a), for the
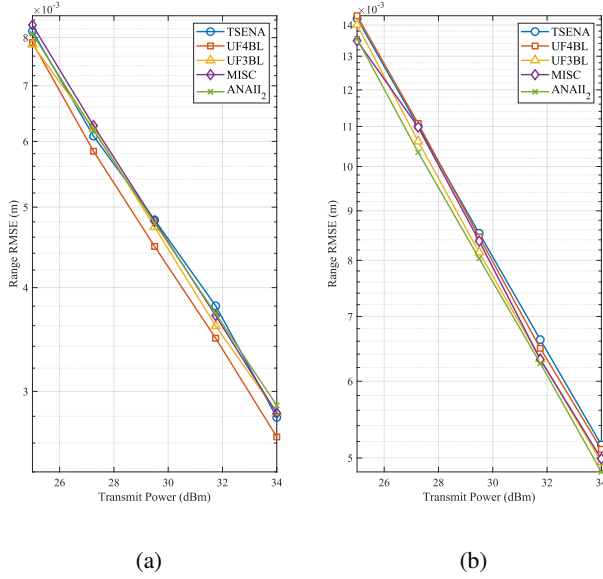
(a)   (b)

Fig. 8: RMSE of range estimation versus transmit power with different CORSET#0 Configuration: (a) 24RBs, (b) 48RBs

24RB configuration, the UF4BL scheme achieves the best performance, which is attributed to its lower mutual coupling. Conversely, for the 48RB configuration in Fig. 8(b), the ANAII-2 scheme provides the highest accuracy due to the increased uDoF it offers. Furthermore, we observe that for a given transmit power, the increased bandwidth from 24RBs to 48RBs does not improve sensing performance. This can be attributed to the decrease in the SNR per RE as the total power is distributed over a larger number of REs.

C. Estimation Accuracy for Multiple Targets

To evaluate the multi-target performance, we employ the optimal sub-pattern assignment (OSPA) metric [10], which provides a unified performance assessment by jointly capturing both localization and cardinality errors (i.e., the difference between the true and estimated number of targets) in a single value.

Given the ground truth set of L target positions $P = \{p_1, \ldots, p_L\}$, and the estimated set of $\hat{L}$ positions $\hat{P} = \{\hat{p}_1, \ldots, \hat{p}_{\hat{L}}\}$, the distance between any true position p and estimated position $\hat{p}$ is defined as

$$d(p, \hat{p}; \bar{c}) = \min\{\bar{c}, ||p - \hat{p}||_2\}, \tag{38}$$

where $\bar{c} > 0$ is the cutoff parameter that determines how the metric penalizes cardinality error with respect to the localization error. For the case $L \leq \hat{L}$, the OSPA metric of order q is defined as [51]

$$\overline{d}_q(P, \hat{P}; \bar{c})$$
$$= \left(\frac{1}{\hat{L}} \left(\min_{\pi \in \Pi_{\hat{L}}} \sum_{l=1}^{L} \left(d(p_l, \hat{p}_{\pi(l)}; \bar{c}) \right)^q + \bar{c}^q (\hat{L} - L) \right) \right)^{1/q}, \tag{39}$$

where $\Pi_{\hat{L}}$ is the set of permutations of $\{1, 2, \ldots, \hat{L}\}$ and $\pi(l)$ is the index of the estimated point in $\hat{P}$ assigned to the l-th true point p_l under a specific permutation $\pi \in \Pi_{\hat{L}}$. In particular, the metric is computed symmetrically as $\overline{d}_q(P, \hat{P}; \bar{c}) = \overline{d}_q(\hat{P}, P; \bar{c})$ if $L > \hat{L}$. Moreover, the order q determines the sensitivity to outliers, while the cutoff $\bar{c}$ balances the localization and cardinality error penalties. In our simulations, we follow the recommendations in [51] and set $q = 2$ and $\bar{c} = 10$ m to ensure a proper balance between these error components. Specifically, we consider $L = 5$ targets, whose parameters are randomly drawn from uniform distributions: ranges in $[5, 85]$ m, velocities in $[-35, 35]$ m/s, and angles in $[-35°, 35°]$. We employ this setup to validate the performance of our proposed two-stage (coarse-to-fine) detection scheme in a multi-target environment.
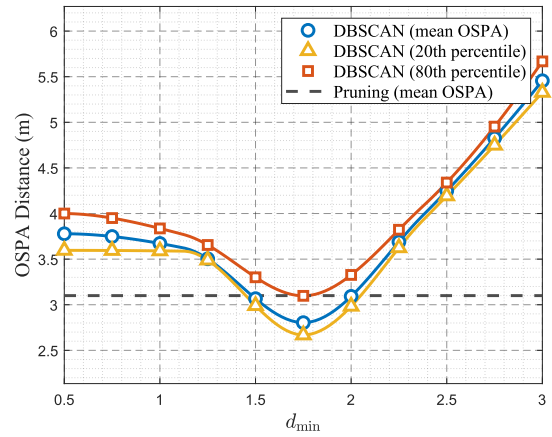

Fig. 9: OSPA distance versus the minimum distance threshold $d_{\min}$ of DBSCAN.

Fig. 9 illustrates the OSPA performance of the proposed DBSCAN-based algorithm versus the minimum distance threshold $d_{\min}$, compared against the Pruning algorithm from [10]. For the Pruning algorithm, parameters are set to $\epsilon_r = \Delta r$ and $\epsilon_v = 3\Delta v$ as recommended in [10] to achieve its optimal performance, which is plotted as the black dashed line (mean OSPA). For our proposed algorithm, the mean, 20th, and 80th percentile curves are presented.

The OSPA distance for the DBSCAN-based fusion algorithm exhibits a U-shaped trend, highlighting a critical trade-off. Specifically, when the threshold $d_{\min}$ is too low, the algorithm fail to associate detections from the same target (e.g., across different beams), resulting in an overestimation of the target count and increasing the cardinality error. Conversely, a large $d_{\min}$ risks merging detections from distinct targets that are closely spaced in the range-Doppler-angle domain, which introduces both localization and cardinality errors. As shown, the DBACAN-based fusion algorithm achieves an optimal balance between both error types around $d_{\min} = 1.7$. With this proper hyperparameter selection, our method demonstrates superior performance over the benchmark.
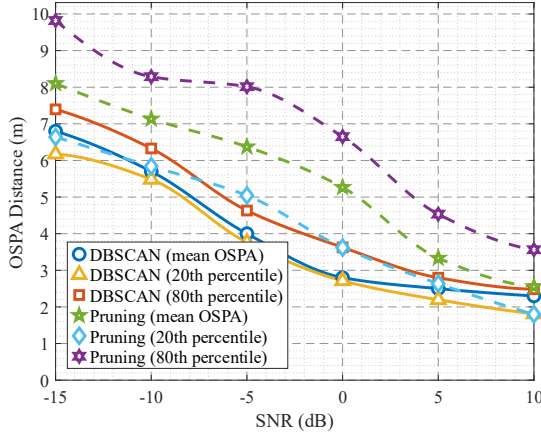
Fig. 10: OSPA distance versus SNR.

Fig. 10 depicts the OSPA distance versus SNR for the proposed DBSCAN-based fusion algorithm and the Pruning benchmark [10]. The proposed algorithm consistently achieves better accuracy and greater robustness across the entire SNR range. This improved robustness is reflected in the narrower spread between its 20th and 80th percentile curves around the mean. In contrast, the Pruning algorithm exhibits larger performance fluctuations, which can be attributed to its selection mechanism: it retains only the single target candidate with the maximum matching score. The clustering nature of DBSCAN, however, inherently suppresses such variability by aggregating the regional detections, resulting in more stable and reliable performance.

VII. CONCLUSION

This paper proposed a two-stage ISAC framework designed to support the emerging LAE by leveraging 5G NR networks. Specifically, the framework addressed the fundamental sensing resolution limitations of standard broadcast signals. In Stage I, we first fused RDMs from the SSB, Type#0-PDCCH, and SIB1 for robust initial detection. In Stage II, we employed the novel SPS, inspired by sparse array theory and inserted into unused SIB1 resources, to overcome the resolution limitations. To this end, a corresponding high-resolution algorithm based on the WUP estimator and the RELAX-based iterative method was developed to accurately extract off-grid range and velocity parameters from these sparse pilots. Finally, the DBSCAN-based clustering scheme was employed to prune redundant detections. Comprehensive simulation results demonstrated that the proposed framework achieves superior estimation accuracy at a significantly lower computational cost compared to conventional benchmark methods. Future work will focus on investigating the framework's performance in more complex, dense multi-target scenarios and exploring further optimization of the SPS patterns in 5G NR.

APPENDIX A

In this appendix, we derive the received frequency-domain signal model for a MIMO-OFDM system as presented in the main text. We begin with the continuous-time baseband signal models and proceed through the steps of sampling, CP-removal, and DFT.

Substituting (5) into (3), and assuming a sufficiently long CP such that inter-symbol interference (ISI) is eliminated, the sampled signal from the l-th transmitted symbol is

$$\hat{\mathbf{y}}(j,l) \approx \sum_{n=1}^{N_{\text{target}}} \beta_n e^{j2\pi f_{D,n} t_{j,l}} \mathbf{b}_{\text{r}}(\theta_n) \mathbf{a}^H(\theta_n)$$
$$\times \sum_{i=0}^{N_c-1} [\mathbf{X}]_{i,l} e^{j2\pi i \Delta f (t_{j,l} - \tau_n)}, \quad (40)$$

where $t_{j,l} = lT + j/F_{\text{s}} + T_{\text{CP}}$ and the noise term is omitted for clarity.

An N_s-point Discrete Fourier Transform (DFT) is then applied to $\hat{\mathbf{y}}(j,l)$ along the index j to obtain the frequency-domain signal $\mathbf{y}(k,l)$:

$$\mathbf{y}(k,l) = \sum_{j=0}^{N_s-1} \hat{\mathbf{y}}(j,l) e^{-j2\pi kj/N_s}$$
$$\approx \sum_{j=0}^{N_s-1} \sum_{n=1}^{N_{\text{target}}} \beta_n e^{j2\pi f_{D,n} t_{j,l}} \mathbf{b}_{\text{r}}(\theta_n) \mathbf{a}^H(\theta_n)$$
$$\times \sum_{i=0}^{N_c-1} [\mathbf{X}]_{i,l} e^{j2\pi i \Delta f (t_{j,l} - \tau_n)} e^{-j2\pi kj/N_s}. \quad (41)$$

By rearranging the summations and substituting the definition of $t_{j,l}$, (41) can be written as

$$\mathbf{y}(k,l) \approx \sum_{n=1}^{N_{\text{target}}} \beta_n \mathbf{b}_{\text{r}}(\theta_n) \mathbf{a}^H(\theta_n) \sum_{i=0}^{N_c-1} [\mathbf{X}]_{i,l} e^{-j2\pi i \Delta f \tau_n}$$
$$\times \left(\sum_{j=0}^{N_s-1} e^{j2\pi (f_{D,n} + i\Delta f) t_{j,l}} e^{-j2\pi kj/N_s} \right). \quad (42)$$

The summation over j can be simplified. First, we expand $t_{j,l}$ and factor out the term that is constant with respect to j:

$$\sum_{j=0}^{N_s-1} e^{j2\pi (f_{D,n} + i\Delta f)(lT + T_{CP} + j/F_s)} e^{-j2\pi kj/N_s}$$
$$= e^{j2\pi (f_{D,n} + i\Delta f)(lT + T_{CP})}$$
$$\times \sum_{j=0}^{N_s-1} e^{j2\pi (f_{D,n} + i\Delta f)(j/F_s)} e^{-j2\pi kj/N_s}. \quad (43)$$

Using the standard OFDM relation $F_s = N_s \Delta f$, the summation term in (43) becomes

$$\sum_{j=0}^{N_s-1} e^{j2\pi \frac{j}{N_s}(i-k+f_{D,n}/\Delta f)}. \quad (44)$$

For scenarios where the Doppler shift is much smaller than the subcarrier spacing ($f_{D,n} \ll \Delta f$), we can approximate $f_{D,n}/\Delta f \approx 0$. This approximation neglects the inter-carrier

interference (ICI). Under this condition, (44) simplifies to the delta function

$$\sum_{j=0}^{N_s-1} e^{j2\pi \frac{j}{N_s}(i-k)} = N_s \cdot \delta_{i,k}. \tag{45}$$

Substituting (45) back into (42) causes the summation over i to collapse to a single term where $i = k$:

$$\mathbf{y}(k,l) \approx \sum_{n=1}^{N_{\text{target}}} \beta_n \mathbf{b}_\text{r}(\theta_n)\mathbf{a}^H(\theta_n)[\mathbf{X}]_{k,l} e^{-j2\pi k\Delta f \tau_n}$$
$$\times\, e^{j2\pi(f_{D,n}+k\Delta f)(lT+T_{CP})} \cdot N_s. \tag{46}$$

After compensating for the known phase component $e^{j2\pi k\Delta f(lT+T_{CP})}$ and neglecting the negligible term $e^{j2\pi f_{D,n}T_{CP}}$ and the constant N_s, (46) simplifies to

$$\mathbf{y}(k,l) = \sum_{n=1}^{N_{\text{target}}} \beta_n \mathbf{a}_\text{t}^H(\theta_n)[\mathbf{X}]_{k,l}\, e^{jk\omega_r(R_n)}e^{jl\omega_v(v_n)}\mathbf{b}_\text{r}(\theta_n)+\mathbf{z}(k,l) \tag{47}$$

where $\omega_r(R_n) \triangleq -4\pi\Delta f R_n/c$ and $\omega_v(v_n) \triangleq 2\pi f_{D,n}T$ respectively.

APPENDIX B

A. Stacking Order Convention

To facilitate the joint estimation of all target parameters, we aggregate the observations from all receive antennas. First, for each receive antenna p, we form a data vector $\mathbf{b}_p = \text{vec}(\mathbf{Y}_p) \in \mathbb{C}^{MN\times 1}$, where $[\mathbf{Y}_p]_{i,l} = [\mathbf{y}(i,l)]_p$. Then the total observation vector $\mathbf{b} \in \mathbb{C}^{N_r MN\times 1}$ is expressed as:

$$\mathbf{b} = \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{N_r-1} \end{bmatrix}. \tag{48}$$

B. Contribution of the n-th Target

For the n-th target, we define the time-frequency phase vector $\mathbf{s}_n \in \mathbb{C}^{MN\times 1}$ with elements arranged as

$$\mathbf{s}_n = \begin{bmatrix} \mathbf{a}_\text{t}^H(\theta_n)[\mathbf{X}]_{0,0}e^{j0\cdot\omega_r(R_n)}e^{j0\cdot\omega_v(v_n)} \\ \mathbf{a}_\text{t}^H(\theta_n)[\mathbf{X}]_{1,0}e^{j1\cdot\omega_r(R_n)}e^{j0\cdot\omega_v(v_n)} \\ \vdots \\ \mathbf{a}_\text{t}^H(\theta_n)[\mathbf{X}]_{M-1,0}e^{j(M-1)\cdot\omega_r(R_n)}e^{j0\cdot\omega_v(v_n)} \\ \mathbf{a}_\text{t}^H(\theta_n)[\mathbf{X}]_{0,1}e^{j0\cdot\omega_r(R_n)}e^{j1\cdot\omega_v(v_n)} \\ \vdots \\ \mathbf{a}_\text{t}^H(\theta_n)[\mathbf{X}]_{M-1,N-1}e^{j(M-1)\cdot\omega_r(R_n)}e^{j(N-1)\cdot\omega_v(v_n)} \end{bmatrix}. \tag{49}$$

The contribution of the n-th target to the complete received vector is given by

$$\mathbf{y}_n = \beta_n \mathbf{b}_\text{r}(\theta_n) \otimes \mathbf{s}_n,$$
$$= \beta_n \mathbf{b}_\text{r}(\theta_n) \otimes \left(\mathbf{T}^T\mathbf{a}_\text{t}^*(\theta_n) \odot (\mathbf{a}_v(v_n) \otimes \mathbf{a}_r(R_n))\right) \tag{50}$$

where $\mathbf{a}_r(R_n) = [1, e^{j\omega_r(R_n)}, \ldots, e^{j(M-1)\omega_r(R_n)}]^T \in \mathbb{C}^{M\times 1}$, $\mathbf{a}_v(v_n) = [1, e^{j\omega_v(v_n)}, \ldots, e^{j(N-1)\omega_v(v_n)}]^T \in \mathbb{C}^{N\times 1}$ and $\mathbf{T} \in \mathbb{C}^{N_t\times MN}$ is formed by stacking $[\mathbf{X}]_{i,l}$ in column-major order.

C. Complete Signal Model

The full received signal is the superposition of all target contributions plus noise

$$\mathbf{y} = \sum_{n=1}^{N_{\text{target}}} \beta_n \left(\mathbf{b}_\text{r}(\theta_n) \otimes \left[(\mathbf{T}^T\mathbf{a}_\text{t}^*(\theta_n)) \odot (\mathbf{a}_v(v_n) \otimes \mathbf{a}_r(R_n))\right]\right)+\mathbf{z}. \tag{51}$$

Denoting the joint steering vector for the n-th target as

$$\mathbf{a}_n = \mathbf{b}_\text{r}(\theta_n) \otimes \left[(\mathbf{T}^T\mathbf{a}_\text{t}^*(\theta_n)) \odot (\mathbf{a}_v(v_n) \otimes \mathbf{a}_r(R_n))\right], \tag{52}$$

the model can then be written in the matrix form

$$\mathbf{y} = \mathbf{A}\boldsymbol{\beta} + \mathbf{z}, \tag{53}$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{N_{\text{target}}}]$ and $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_{N_{\text{target}}}]^T$.

Substituting these results back, the final expression for the partial derivative of the joint steering vector with respect to the angle θ_n is:

$$\frac{\partial \mathbf{a}_n}{\partial \theta_n} = (\mathbf{J}_\text{br}\mathbf{b}_\text{r}(\theta_n)) \otimes \mathbf{s}_n$$
$$- \mathbf{b}_\text{r}(\theta_n) \otimes \left[(\mathbf{T}^T\mathbf{J}_\text{at}\mathbf{a}_\text{t}^*(\theta_n)) \odot (\mathbf{a}_v(v_n) \otimes \mathbf{a}_r(R_n))\right]. \tag{54}$$

$$\frac{\partial \mathbf{a}_n}{\partial \omega_r(R_n)} = \mathbf{b}_\text{r}(\theta_n)\otimes\left[(\mathbf{T}^T\mathbf{a}_\text{t}^*(\theta_n)) \odot (\mathbf{a}_v(v_n) \otimes (\mathbf{J}_r\mathbf{a}_r(R_n)))\right] \tag{55}$$

$$\frac{\partial \mathbf{a}_n}{\partial \omega_v(v_n)} = \mathbf{b}_\text{r}(\theta_n)\otimes\left[(\mathbf{T}^T\mathbf{a}_\text{t}^*(\theta_n)) \odot ((\mathbf{J}_v\mathbf{a}_v(v_n)) \otimes \mathbf{a}_r(R_n))\right] \tag{56}$$

where $\mathbf{J}_\text{br}$, $\mathbf{J}_\text{at}$, $\mathbf{J}_r$ and $\mathbf{J}_v$ are diagonal matrixs. $[\mathbf{J}_\text{br}]_{i,i} = -j2\pi(i-1)d_r\cos\theta_n/\lambda_c$ and $[\mathbf{J}_\text{at}]_{i,i} = -j2\pi(i-1)d_t\cos\theta_n/\lambda_c$, $\mathbf{J}_r = j\cdot\text{diag}(0,1,2,\ldots,M-1)$, $\mathbf{J}_v = j\cdot\text{diag}(0,1,2,\ldots,N-1)$

Based on the Theorem 1 in [52], the CRLB w.r.t $\boldsymbol{\Theta} = \left[\boldsymbol{\theta}^T, \boldsymbol{\omega}_r^T, \boldsymbol{\omega}_v^T\right]^T$ can be given as

$$\text{CRLB}(\boldsymbol{\Theta}) = \frac{\sigma_v^2}{2}\left[\text{Re}\left\{\tilde{\mathbf{B}}^H\left(\bar{\mathbf{A}}^H\mathbf{P}_\mathbf{A}^\perp\bar{\mathbf{A}}\right)\tilde{\mathbf{B}}\right\}\right]^{-1}, \tag{57}$$

where

$$\bar{\mathbf{A}} = \left[\frac{\partial\mathbf{a}_1}{\partial\theta_1}, \frac{\partial\mathbf{a}_2}{\partial\theta_2}, \ldots, \frac{\partial\mathbf{a}_{N_{\text{target}}}}{\partial\theta_{N_{\text{target}}}},\right.$$
$$\frac{\partial\mathbf{a}_1}{\partial\omega_r(R_1)}, \frac{\partial\mathbf{a}_2}{\partial\omega_r(R_2)}, \ldots, \frac{\partial\mathbf{a}_{N_{\text{target}}}}{\partial\omega_r(R_{N_{\text{target}}})}, \tag{58}$$
$$\left.\frac{\partial\mathbf{a}_1}{\partial\omega_v(v_1)}, \frac{\partial\mathbf{a}_2}{\partial\omega_v(v_2)}, \ldots, \frac{\partial\mathbf{a}_{N_{\text{target}}}}{\partial\omega_v(v_{N_{\text{target}}})}\right]$$

$$\mathbf{P}_\mathbf{A}^\perp = \mathbf{I} - \mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}, \tag{59}$$

$$\tilde{\mathbf{B}} = \mathbf{I}_3 \otimes \text{diag}(\boldsymbol{\beta}) \tag{60}$$

REFERENCES

[1] Y. Jiang, X. Li, G. Zhu, H. Li, J. Deng, K. Han, C. Shen, Q. Shi, and R. Zhang, "Integrated sensing and communication for low altitude economy: Opportunities and challenges," *IEEE Commun. Mag.*, 2025.

[2] Y. Wang, G. Sun, Z. Sun, J. Wang, J. Li, C. Zhao, J. Wu, S. Liang, M. Yin, P. Wang, *et al.*, "Toward realization of low-altitude economy networks: Core architecture, integrated technologies, and future directions," *arXiv preprint arXiv:2504.21583*, 2025.

[3] M. Song, Y. Lin, J. Wang, G. Sun, C. Dong, N. Ma, D. Niyato, and P. Zhang, "Trustworthy intelligent networks for low-altitude economy," *IEEE Commun. Mag.*, 2025.

[4] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in uav communication networks," *IEEE Commun. Surv. Tutor.*, vol. 18, no. 2, pp. 1123–1152, 2015.

[5] M. A. Hoque, M. Hossain, S. Noor, S. R. Islam, and R. Hasan, "IoTaaS: Drone-based internet of things as a service framework for smart cities," *IEEE Internet Things J.*, vol. 9, no. 14, pp. 12425–12439, 2021.

[6] Q. Wu, J. Xu, Y. Zeng, D. W. K. Ng, N. Al-Dhahir, R. Schober, and A. L. Swindlehurst, "A comprehensive overview on 5G-and-beyond networks with uavs: From communications to sensing and intelligence," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 2912–2945, 2021.

[7] 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.*, vol. 68, no. 6, pp. 3834–3862, 2020.

[8] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[9] H. Luo, Y. Wang, D. Luo, J. Zhao, H. Wu, S. Ma, and F. Gao, "Integrated sensing and communications in clutter environment," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 10941–10956, 2024.

[10] L. Pucci, E. Paolini, and A. Giorgetti, "System-level analysis of joint sensing and communication based on 5G new radio," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2043–2055, 2022.

[11] J. Tang, Y. Yu, C. Pan, H. Ren, D. Wang, J. Wang, and X. You, "Cooperative ISAC-empowered low-altitude economy," *IEEE Trans. Wireless Commun.*, 2025.

[12] 3GPP, "NR; physical channels and modulation," Technical Specification (TS) 38.211, 3GPP, Jul. 2023. Version 17.5.0.

[13] 3GPP, "NR; Multiplexing and channel coding," Technical Specification (TS) 38.212, 3GPP, Sep. 2025. Version 19.1.0.

[14] 3GPP, "NR; physical layer procedures for control," Technical Specification (TS) 38.213, 3GPP, Oct. 2023. Version 17.7.0.

[15] 3GPP, "NR; Physical layer procedures for data," Technical Specification (TS) 38.214, 3GPP, Sep. 2025. Version 19.1.0.

[16] IEEE, "IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 1: Enhancements for High-Efficiency WLAN," 2021. (Amendment to IEEE Std 802.11-2020).

[17] M. F. Keskin, M. M. Mojahedian, J. O. Lacruz, C. Marcus, O. Eriksson, A. Giorgetti, J. Widmer, and H. Wymeersch, "Fundamental trade-offs in monostatic ISAC: A holistic investigation towards 6G," *IEEE Trans. Wireless Commun.*, 2025.

[18] Z. Wei, Y. Wang, L. Ma, S. Yang, Z. Feng, C. Pan, Q. Zhang, Y. Wang, H. Wu, and P. Zhang, "5G PRS-based sensing: A sensing reference signal approach for joint sensing and communication system," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 3250–3263, 2022.

[19] M. Golzadeh, E. Tiirola, J. Talvitie, L. Anttila, K. Hooli, O. Tervo, and M. Valkama, "Joint sensing and ue positioning in 5G-6G: Prs range estimation with suppressed ambiguity," in *2024 IEEE Radar Conference (RadarConf24)*, pp. 1–6, IEEE, 2024.

[20] J. Yang, C.-K. Wen, J. Xu, H. Que, H. Wei, and S. Jin, "Angle-based SLAM on 5G mmWave systems: Design, implementation, and measurement," *IEEE Internet Things J.*, vol. 10, no. 20, pp. 17755–17771, 2023.

[21] K. Abratkiewicz, A. Księżyk, M. Płotka, P. Samczyński, J. Wszołek, and T. P. Zieliński, "SSB-based signal processing for passive radar using a 5G network," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 16, pp. 3469–3484, 2023.

[22] Y. Awad, A. Celik, and A. Eltawil, "SSB-based multi-target detection and joint velocity, angle, and range estimation in 5G-NR," *IEEE Commun. Lett.*, 2025.

[23] L. Chen, X. Zhou, F. Chen, L.-L. Yang, and R. Chen, "Carrier phase ranging for indoor positioning with 5G NR signals," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 10908–10919, 2021.

[24] Z. Liu, L. Chen, X. Zhou, Z. Jiao, G. Guo, and R. Chen, "Machine learning for time-of-arrival estimation with 5G signals in indoor positioning," *IEEE Internet Things J.*, vol. 10, no. 11, pp. 9782–9795, 2023.

[25] H. Yang, L. Chen, H. Liu, and G. Zhu, "Dynamic feature-fused localization with smartphones exploiting 5G NR SSB and Wi-Fi for indoor environments," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–14, 2024.

[26] M. Golzadeh, E. Tiirola, L. Anttila, J. Talvitie, K. Hooli, O. Tervo, I. Peruga, S. Hakola, and M. Valkama, "Downlink sensing in 5G-advanced and 6G: SIB1-assisted SSB approach," in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, pp. 1–7, IEEE, 2023.

[27] Q. Zhao, A. Tang, and X. Wang, "Reference signal design and power optimization for energy-efficient 5G V2X integrated sensing and communications," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 1, pp. 379–392, 2023.

[28] W. Liu, X. Hou, J. Liu, and L. Chen, "Low-overhead sensing rs design for integrated sensing and communication," in *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2025.

[29] D. Mei, Z. Wei, X. Chen, L. Wang, and Z. Feng, "A coprime and periodic pilot design for ISAC system," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2024.

[30] R. Zhang, S. Tsai, T.-H. Chou, J. Ren, W. Qu, and O. Sun, "OFDM reference signal pattern design criteria for integrated communication and sensing," *IEEE Internet Things J.*, 2024.

[31] R. Xie, D. Hu, K. Luo, and T. Jiang, "Performance analysis of joint range-velocity estimator with 2D-MUSIC in OFDM radar," *IEEE Trans. Signal Process.*, vol. 69, pp. 4787–4800, 2021.

[32] P. Stoica, J. Li, J. Ling, and Y. Cheng, "Missing data recovery via a nonparametric iterative adaptive approach," in *2009 ieee international conference on acoustics, speech and signal processing*, pp. 3369–3372, IEEE, 2009.

[33] W. Roberts, P. Stoica, J. Li, T. Yardibi, and F. A. Sadjadi, "Iterative adaptive approaches to MIMO radar imaging," *IEEE J. Sel. Top. Signal Process.*, vol. 4, no. 1, pp. 5–20, 2010.

[34] J. Ren and S. Tsai, "Multi-tier structured array for sensing pilot design in integrated sensing and communication," in *2024 IEEE 13rd Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 1–5, IEEE, 2024.

[35] A. Księżyk, M. Płotka, K. Abratkiewicz, R. Maksymiuk, J. Wszołek, P. Samczyński, and T. P. Zieliński, "Opportunities and limitations in radar sensing based on 5G broadband cellular networks," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 38, no. 9, pp. 4–21, 2023.

[36] X. Fu, D. Le Ruyet, R. Visoz, V. Ramireddy, M. Grossmann, M. Landmann, and W. Quiroga, "A tutorial on downlink precoder selection strategies for 3GPP MIMO codebooks," *IEEE Access*, vol. 11, pp. 138897–138922, 2023.

[37] Z. Zhang, H. Ren, C. Pan, S. Hong, D. Wang, J. Wang, and X. You, "Target localization in cooperative ISAC systems: A scheme based on 5G NR OFDM signals," *IEEE Trans. Commun.*, 2024.

[38] T. Salman, A. Badawy, T. M. Elfouly, A. Mohamed, and T. Khattab, "Estimating the number of sources: An efficient maximization approach," in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 199–204, IEEE, 2015.

[39] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 33, no. 2, pp. 387–392, 2003.

[40] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory.* John Wiley & Sons, 2002.

[41] Z. Xiao, R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst, "A novel joint angle-range-velocity estimation method for MIMO-OFDM ISAC systems," *IEEE Trans. Signal Process.*, 2024.

[42] Z. Wei, B. Li, T. Feng, Y. Tao, and C. Zhao, "Area-based CFAR target detection for automotive millimeter-wave radar," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 2891–2906, 2022.

[43] X. Wang, Z. Yang, F. Liu, P. Chu, and J. Zheng, "Coprime-based frame-level sensing OFDM waveform for 5G NR integrated radar sensing and

communication via downlink sensing," *IEEE Sens. J.l*, vol. 24, no. 15, pp. 24421–24437, 2024.

[44] J. Liu, Y. Zhang, Y. Lu, S. Ren, and S. Cao, "Augmented nested arrays with enhanced DOF and reduced mutual coupling," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5549–5563, 2017.

[45] W. Shi, S. A. Vorobyov, and Y. Li, "ULA fitting for sparse array design," *IEEE Trans. Signal Process.*, vol. 69, pp. 6431–6447, 2021.

[46] Z. Zheng, W.-Q. Wang, Y. Kong, and Y. D. Zhang, "MISC array: A new sparse array design achieving increased degrees of freedom and reduced mutual coupling effect," *IEEE Trans. Signal Process.*, vol. 67, no. 7, pp. 1728–1741, 2019.

[47] S. Ren, W. Dong, X. Li, W. Wang, and X. Li, "Extended nested arrays for consecutive virtual aperture enhancement," *IEEE Signal Process. Lett.*, vol. 27, pp. 575–579, 2020.

[48] Y. Ma, X. Cao, X. Wang, M. S. Greco, and F. Gini, "Multi-source off-grid DOA estimation with single snapshot using non-uniform linear arrays," *Signal Processing*, vol. 189, p. 108238, 2021.

[49] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *KDD*, vol. 96, pp. 226–231, 1996.

[50] L. Zheng and X. Wang, "Super-resolution delay-doppler estimation for OFDM passive radar," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2197–2210, 2017.

[51] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, 2008.

[52] Z. Hu, Q. Ye, Y. Huang, S. Hu, and G. Yang, "Joint range-velocity-azimuth estimation for OFDM-based integrated sensing and communication," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 12933–12948, 2024.