Learning-Based Delay, Doppler, and Angle Estimation Using Quantized DDAE for OFDM Bistatic ISAC

Shiv Shankar[†*], Radhakrishna Ganti[†] and J. Klutto Milleth[*†]
[*]*Center of Excellence in Wireless Technology (CEWiT)*
[†]*Department of Electrical Engineering, IIT Madras*, Chennai, India
{shivshankar, klutto}@cewit.org.in
rganti@ee.iitm.ac.in

***Abstract*—Integrated sensing and communication (ISAC) is emerging as a key capability for 5G-Advanced and 6G systems, motivating sensing pipelines that remain compatible with practical OFDM-based deployments. In this work, we consider bistatic estimation of azimuth, elevation, Doppler, and delay in a 5G-style OFDM sensing setup using PRS-like pilot resources. Instead of feeding the full complex channel frequency response (CFR) tensor to a learning model, the receiver first transforms the measurements into a delay–Doppler–azimuth–elevation (DDAE) representation, retains the delay–Doppler region of interest, and uses its 8-bit quantized log-magnitude as input. This yields a compact sensing-oriented representation with reduced feedback and model-input complexity. A lightweight coarse-to-fine multi-task estimator is then designed to exploit the DDAE structure for continuous-valued parameter prediction. The method is motivated by the fact that direct peak selection on a discretized DDAE map is fundamentally grid-limited. Simulation results show that the learned estimator consistently improves azimuth, elevation, Doppler, and delay estimation accuracy over direct DDAE peak-based estimation, demonstrating a practical sub-grid refinement mechanism for communication-compatible sensing.**

***Index Terms*—ISAC, OFDM sensing, bistatic sensing, DDAE, delay–Doppler–angle estimation, learning-based sensing**

I. Introduction

Sixth-generation (6G) wireless systems are expected to provide not only high-rate connectivity, but also native intelligence and environmental awareness. In the IMT-2030 framework, the International Telecommunication Union (ITU) identifies *Ubiquitous Connectivity*, *AI and Communication*, and *Integrated Sensing and Communication* (ISAC) as major usage scenarios beyond IMT-2020 [1]. In parallel, 3GPP has initiated sensing-related studies through Rel-19 ISAC work and subsequent Rel-20 architectural and service-layer studies, indicating that sensing is evolving from a research concept toward a practical capability for 5G-Advanced and beyond [2], [3].

This trend naturally motivates ISAC solutions that remain close to existing cellular deployments. In particular, OFDM-based signaling and reference-signal-driven processing are attractive because they are already embedded in current wireless systems and can support communication-compatible sensing with shared spectrum and hardware [4]. Recent IEEE works have explored PRS-based sensing for range and velocity estimation, 5G-NR synchronization-signal-based sensing, and joint angle–range–velocity estimation for MIMO-OFDM ISAC systems [5]–[7]. These studies demonstrate the feasibility of practical OFDM sensing, but they also highlight the importance of efficient receiver-side estimation and compact sensing representations.

A key difficulty is that many practical sensing pipelines still rely on discretized response maps, where delay, Doppler, and angular parameters are extracted by direct peak selection. While simple and robust, such estimators are inherently limited by the underlying delay, Doppler, and angular grid. Classical off-grid or super-resolution approaches, such as MUSIC or likelihood-based refinement, can improve accuracy, but often at the cost of increased complexity, sensitivity to model mismatch, or higher search burden [8], [9]. At the same time, recent IEEE literature has shown growing interest in machine-learning-enhanced ISAC, including survey-level overviews and end-to-end learning-based designs under practical impairments [10]–[12]. Directly operating on the full complex CFR/CSI tensor across subcarriers, OFDM symbols, and antenna elements therefore leads to both high-dimensional sensing measurements and heavy model inputs, which is undesirable for lightweight learning-based sensing.

This issue becomes even more relevant in architectures where sensing measurements are processed beyond the local radio node, as illustrated in Fig. 1. A straightforward centralized high-resolution pipeline would require forwarding or storing the full complex CFR tensor, whose size scales with the product of the number of subcarriers, OFDM symbols, and receive-antenna elements, with separate real and imaginary components. In contrast, if the measurements are first transformed locally into a DDAE representation, restricted to the delay–Doppler region of interest, and then converted to 8-bit quantized log-magnitude values, the sensing representation becomes substantially more compact. This reduces both measurement-transfer overhead and network-input complexity, while preserving the structure needed for sub-grid estimation

This work was funded in part by the Ministry of Electronics and Information Technology (MeitY), Government of India.

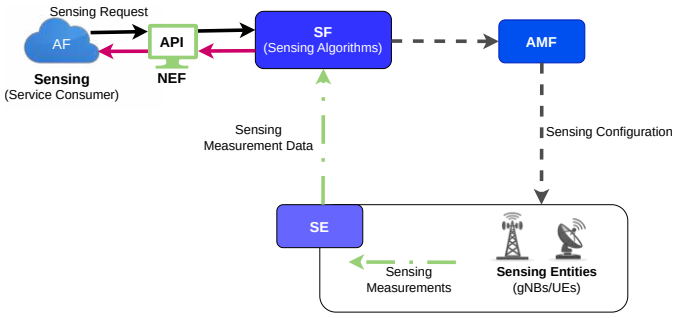

Fig. 1. Illustrative ISAC architecture with local sensing entities and a sensing function (SF).

at the sensing function (SF).

Motivated by the above gap, this work considers bistatic estimation of azimuth, elevation, Doppler, and delay in a 5G-style OFDM sensing setup. Instead of directly selecting the strongest DDAE bin, we use a lightweight learning-based estimator that operates on a compact quantized DDAE tensor and predicts continuous-valued target parameters. The resulting approach acts as a practical sub-grid alternative to direct grid-based estimation while remaining significantly lighter than pipelines that rely on the full complex CFR tensor.

The main contributions of this paper are as follows:

- We formulate bistatic estimation of azimuth, elevation, Doppler, and delay as a supervised regression problem using a DDAE-domain sensing representation derived from OFDM pilot measurements.
- We employ a compact sensing input obtained from the log-magnitude DDAE tensor in 8-bit quantized form, after retaining only the delay–Doppler region of interest, thereby avoiding the direct use of the high-dimensional complex CFR tensor and reducing sensing-input complexity.
- We design a lightweight coarse-to-fine multi-task estimator with dedicated delay/Doppler processing and a decoupled angular branch for continuous-valued parameter prediction, and show that it consistently improves upon direct DDAE peak-based estimation as a practical sub-grid refinement mechanism.

II. System Model and DDAE Representation

We consider a bistatic sensing link extracted from a multi-static ISAC deployment. Let $\mathbf{p}_{\mathrm{T}}, \mathbf{p}_{\mathrm{R}} \in \mathbb{R}^3$ denote the transmitter (TX) and receiver (RX) positions, and let a point target be located at $\mathbf{p}_{\mathrm{o}} \in \mathbb{R}^3$ with velocity $\mathbf{v}_{\mathrm{o}} \in \mathbb{R}^3$. The TX and RX employ uniform planar arrays (UPAs) with inter-element spacing $d = \lambda/2$, where λ is the carrier wavelength. The dominant direct/LOS component and strong static background components are assumed to be removed or sufficiently suppressed by preprocessing, so that the retained measurements are dominated by the target response plus additive noise. The target direction at the RX is parameterized by azimuth ϕ_o and elevation θ_o.

The bistatic propagation delay is

$$\tau_o = \frac{\|\mathbf{p}_{\mathrm{o}} - \mathbf{p}_{\mathrm{T}}\| + \|\mathbf{p}_{\mathrm{R}} - \mathbf{p}_{\mathrm{o}}\|}{c}, \tag{1}$$

where c is the speed of light. Defining the unit vectors

$$\hat{\mathbf{r}}_{\mathrm{to}} = \frac{\mathbf{p}_{\mathrm{o}} - \mathbf{p}_{\mathrm{T}}}{\|\mathbf{p}_{\mathrm{o}} - \mathbf{p}_{\mathrm{T}}\|}, \qquad \hat{\mathbf{r}}_{\mathrm{or}} = \frac{\mathbf{p}_{\mathrm{R}} - \mathbf{p}_{\mathrm{o}}}{\|\mathbf{p}_{\mathrm{R}} - \mathbf{p}_{\mathrm{o}}\|}, \tag{2}$$

the bistatic Doppler shift is

$$f_{d,o} = \frac{\mathbf{v}_{\mathrm{o}}^T \hat{\mathbf{r}}_{\mathrm{to}} + \mathbf{v}_{\mathrm{o}}^T \hat{\mathbf{r}}_{\mathrm{or}}}{\lambda}. \tag{3}$$

We consider an OFDM waveform with K subcarriers, subcarrier spacing Δf_{sc}, and L sensing symbols. Let $T_{\mathrm{sym}} \approx 1/\Delta f_{\mathrm{sc}}$ denote the OFDM-symbol spacing used in the sensing grid. During sensing symbols, PRS-like pilots are transmitted on a resource-element set Ω. The received grid is

$$\mathbf{Y} = \mathbf{H} \odot \mathbf{X} + \mathbf{N}, \tag{4}$$

where $\mathbf{X}$ is the transmitted resource grid, the pilot symbols are normalized to unit average power, $\mathbf{N}$ is modeled as additive white Gaussian noise (AWGN), and $\odot$ denotes elementwise multiplication over (k, l). For the target component, the CFR at RX element (n_y, n_z) is

$$\mathbf{H}_o(n_y, n_z, k, l) = \alpha_o e^{-j2\pi k \Delta f_{\mathrm{sc}} \tau_o} e^{j2\pi f_{d,o} l T_{\mathrm{sym}}} a_{\mathrm{RX}}(n_y, n_z; \theta_o, \phi_o), \tag{5}$$

where α_o absorbs the target reflectivity, path gain, and TX beamforming factor. On pilot REs, the least-squares CFR estimate is

$$\widetilde{\mathbf{H}}(n_y, n_z, k, l) = \frac{\mathbf{Y}(n_y, n_z, k, l)}{\mathbf{X}_p(k, l)}, \qquad (k, l) \in \Omega. \tag{6}$$

The estimated CFR is then transformed into a delay–Doppler–azimuth–elevation (DDAE) tensor through a separable IFFT/FFT chain. First, an IFFT over subcarriers gives delay bins,

$$\mathbf{H}_\tau(n_y, n_z, q, l) = \frac{1}{K} \sum_{k=0}^{K-1} \widetilde{\mathbf{H}}(n_y, n_z, k, l) e^{j2\pi kq/K}. \tag{7}$$

Next, an FFT over OFDM symbols gives Doppler bins,

$$\mathbf{H}_{f_d}(n_y, n_z, q, p) = \sum_{l=0}^{L-1} \mathbf{H}_\tau(n_y, n_z, q, l) e^{-j2\pi lp/L}. \tag{8}$$

Finally, spatial FFTs across the RX UPA yield the DDAE tensor $\mathbf{H}_{\mathrm{DDAE}}(a, e, p, q)$, whose axes correspond to azimuth, elevation, Doppler, and delay, respectively.

The native delay and Doppler bin spacings are

$$\Delta\tau = \frac{1}{K\Delta f_{\mathrm{sc}}} = \frac{1}{B_{\mathrm{eff}}}, \qquad \Delta f_d = \frac{1}{T_{\mathrm{eff}}}, \tag{9}$$

where B_{eff} is the effective occupied bandwidth and T_{eff} is the coherent observation time of the retained sensing snapshot. Instead of operating on the full complex CFR tensor, we use

the DDAE representation as a compact sensing-oriented front-end. The complex DDAE cube is converted to a real-valued log-magnitude tensor

$$\mathbf{P}(a,e,p,q) = 10\log_{10}\big(|\mathbf{H}_{\mathrm{DDAE}}(a,e,p,q)|^2 + \epsilon\big), \quad (10)$$

and uniformly quantized to 8-bit unsigned integers,

$$\mathbf{Q} = \mathrm{uint8}\left(\mathrm{round}\left(255 \cdot \frac{\mathrm{clip}(\mathbf{P}, \mathbf{P}_{\min}, \mathbf{P}_{\max}) - \mathbf{P}_{\min}}{\mathbf{P}_{\max} - \mathbf{P}_{\min}}\right)\right). \quad (11)$$

In the implementation, the retained tensor is arranged as

$$\mathbf{Q} \in \mathbb{R}^{N_a \times N_e \times N_{f_d} \times N_\tau} = \mathbb{R}^{8\times 8\times 128\times 1200}, \quad (12)$$

where $N_a = 8$ and $N_e = 8$ are the angular bins induced by the 8×8 RX UPA, while $N_{f_d} = 128$ and $N_\tau = 1200$ are the retained Doppler and delay bins selected according to the expected spread over the considered target region. For the angular dimensions, the native spatial-FFT grid is uniform in direction-cosine space, and the corresponding angular discretization is therefore characterized by effective local bin widths $\Delta_\phi^{\mathrm{eff}}$ and $\Delta_\theta^{\mathrm{eff}}$.

For each bistatic-link sample, the parameter vector of interest is

$$\boldsymbol{\xi} = [\phi_o,\ \theta_o,\ f_{d,o},\ \tau_o]^T, \quad (13)$$

and the learning objective is to estimate

$$f_\Theta : \mathbf{Q} \mapsto \hat{\boldsymbol{\xi}}, \quad (14)$$

where $\hat{\boldsymbol{\xi}} = [\hat{\phi}, \hat{\theta}, \hat{f}_d, \hat{\tau}]^T$ denotes the predicted azimuth, elevation, Doppler, and delay of the target.

For the direct DDAE peak-based baseline, the discretized grid itself induces a fundamental off-grid error. In the ideal noiseless case,

$$\begin{aligned} |\hat{\tau} - \tau_o| &\le \frac{\Delta_\tau}{2}, \qquad |\hat{f}_d - f_{d,o}| \le \frac{\Delta f_d}{2}, \\ |\hat{\phi} - \phi_o| &\le \frac{\Delta_\phi^{\mathrm{eff}}}{2}, \qquad |\hat{\theta} - \theta_o| \le \frac{\Delta_\theta^{\mathrm{eff}}}{2}. \end{aligned} \quad (15)$$

For delay and Doppler, whose grids are uniform, these correspond to the usual half-bin quantization limits of direct grid-based estimation.

III. Proposed Learning-Based Estimator

The proposed network predicts $[\hat{\phi}, \hat{\theta}, \hat{f}_d, \hat{\tau}]^T$ from the quantized DDAE tensor. For each sample, the input has size $\mathbf{Q} \in \mathbb{R}^{8\times 8\times 128\times 1200}$. The two angular dimensions are merged into the channel dimension to obtain $\widetilde{\mathbf{Q}} \in \mathbb{R}^{64\times 128\times 1200}$. The reshaped tensor is first processed by a lightweight shared convolutional stem,

$$\mathbf{F} = f_{\mathrm{stem}}(\widetilde{\mathbf{Q}}), \quad (16)$$

followed by separate task-specific adapters for delay and Doppler:

$$\mathbf{F}_\tau = g_\tau(\mathbf{F}), \qquad \mathbf{F}_{f_d} = g_{f_d}(\mathbf{F}). \quad (17)$$

The delay and Doppler branches exploit the anisotropic structure of the DDAE tensor by aggregating features along one axis while preserving the other. This produces coarse delay and Doppler logits, denoted by $\mathbf{z}_\tau^{(c)}$ and $\mathbf{z}_{f_d}^{(c)}$, from which coarse bin locations are obtained using soft-argmax:

$$\hat{b}_\tau^{(c)} = \mathrm{SoftArgMax}\Big(\mathbf{z}_\tau^{(c)}\Big), \qquad \hat{b}_{f_d}^{(c)} = \mathrm{SoftArgMax}\Big(\mathbf{z}_{f_d}^{(c)}\Big). \quad (18)$$

A local refinement step is then applied around the detached coarse center in each branch, yielding refined bin estimates $\hat{b}_\tau$ and $\hat{b}_{f_d}$. This two-stage design stabilizes sub-bin estimation by first identifying a reliable coarse neighborhood and then performing local refinement within that region. The final delay and Doppler estimates are

$$\hat{\tau} = (\hat{b}_\tau + d_0)\Delta\tau, \qquad \hat{f}_d = (\hat{b}_{f_d} - v_0)\Delta f_d, \quad (19)$$

where d_0 and v_0 denote the delay offset and Doppler-center bin, respectively.

The angular branch is decoupled from delay/Doppler learning. Specifically, the coarse delay and Doppler logits are converted to detached attention weights and used to form an attention-focused angular map from the input tensor. This map is combined with mean and max angular projections and coordinate channels, and the resulting feature map is processed by a lightweight CNN:

$$\mathbf{u} = f_{\mathrm{ang}}(\mathbf{A}). \quad (20)$$

Azimuth is predicted using a sine–cosine representation,

$$[\hat{s}_\phi,\ \hat{c}_\phi] = h_\phi(\mathbf{u}), \quad (21)$$

while elevation is predicted as a bounded scalar,

$$r_\theta = h_\theta(\mathbf{u}), \qquad \hat{\theta} = \theta_{\min} + (\theta_{\max} - \theta_{\min})\sigma(r_\theta), \quad (22)$$

where $\sigma(\cdot)$ denotes the sigmoid function. Azimuth is represented through sine and cosine to avoid angular wrap-around, whereas elevation is regressed directly as a bounded scalar over its physically valid interval. The azimuth estimate is recovered by

$$\hat{\phi} = \mathrm{atan2}(\hat{s}_\phi, \hat{c}_\phi). \quad (23)$$

Training uses a composite objective consisting of: i) smooth-L1 losses for refined delay and Doppler regression, ii) a unit-vector loss for azimuth, iii) a normalized regression loss for elevation, iv) auxiliary coarse delay/Doppler regression losses, and v) coarse-logit supervision using soft-bin targets, with an additional hard-bin cross-entropy term for Doppler. The overall loss is written compactly as

$$\begin{aligned} L = \lambda_\tau L_\tau + \lambda_{f_d} L_{f_d} + \lambda_\phi L_\phi + \lambda_\theta L_\theta + \lambda_\tau^{(c)} L_\tau^{(c)} + \lambda_{f_d}^{(c)} L_{f_d}^{(c)} \\ + \lambda_{\tau,\mathrm{bin}} L_{\tau,\mathrm{bin}} + \lambda_{f_d,\mathrm{bin}} L_{f_d,\mathrm{bin}} + \lambda_{f_d,\mathrm{CE}} L_{f_d,\mathrm{CE}}. \end{aligned} \quad (24)$$

During inference, the final reported outputs are the refined delay and Doppler estimates together with the angular-branch outputs, while the coarse terms are used only to guide training and refinement.

The resulting estimator remains lightweight by operating on the compact quantized DDAE input, while improving upon direct grid-based peak selection through task-specific feature adaptation, coarse-to-fine sub-bin refinement, and decoupled angular estimation.

TABLE I
SIMULATION AND DATASET PARAMETERS.

Parameter	Value
TX locations	$(250, 250, 25)$ m, $(-250, -250, 25)$ m
RX locations	$(250, -250, 25)$ m, $(-250, 250, 25)$ m
RX panel azimuths	$135^\circ, 45^\circ$
Target region	$x, y \in [-240, 240]$ m, $z \in [20, 300]$ m
Link usage	4 bistatic TX–RX links, treated independently
RX array	8×8 UPA, $d = \lambda/2$
Waveform	$f_c = 13$ GHz, $\Delta f_{\text{sc}} = 120$ kHz, $K = 3276$
PRS	PRS every 8 OFDM symbols
Noise model	AWGN, receiver noise figure = 5 dB
Retained DDAE size	$8 \times 8 \times 128 \times 1200$
Samples / split	Single target; train/val/test = 19200/2400/2400

TABLE II
P50/P90/P99 ABSOLUTE ESTIMATION ERRORS FOR THE PROPOSED MACHINE LEARNING (ML) BASED ESTIMATOR AND DIRECT GRID-BASED DDAE PEAK ESTIMATION. RATIO DENOTES GRID/ML, SO LARGER VALUES INDICATE GREATER IMPROVEMENT.

	P50			P90			P99		
Parameter	ML	Grid	Ratio	ML	Grid	Ratio	ML	Grid	Ratio
e_ϕ (deg)	0.26	4.46	17.28	0.86	8.52	9.95	2.16	15.40	7.13
e_θ (deg)	0.23	3.68	15.73	0.78	7.06	9.03	1.83	20.87	11.40
e_{f_d} (Hz)	15.22	30.54	2.01	25.86	53.14	2.06	29.67	60.06	2.02
e_τ (ns)	0.095	0.546	5.76	0.339	0.922	2.72	1.004	1.070	1.07

IV. SIMULATION SETUP

The main simulation parameters are summarized in Table I, and Fig. 2 illustrates the deployment geometry. The dataset is generated from a deployment with two transmitters and two receivers, yielding four bistatic TX–RX links. Orthogonal PRS allocation isolates per-link CFR estimates, and each extracted link is treated as an independent single-target sample. Because targets are distributed over a bounded target region, the bistatic path loss and hence the effective received SNR vary across samples, so a single fixed SNR value is not fully representative of the dataset. Treating the four TX–RX links independently increases geometric diversity over the target region without requiring multi-link fusion. The model is trained using AdamW with learning rate 10^{-4}, weight decay 10^{-4}, batch size 12, and a maximum of 28 epochs. A 3-epoch warmup followed by cosine learning-rate decay is used, mixed-precision training is enabled, gradients are clipped to unit norm, and early stopping with patience 8 is applied based on validation improvement. To avoid leakage across correlated samples, the train/validation/test partition 80/10/10 is performed at the drop level rather than at the individual sample level.

V. RESULTS AND DISCUSSION

The performance is evaluated using the absolute estimation errors

$$\begin{aligned} e_\tau &= |\hat{\tau} - \tau_o|, \qquad e_{f_d} = |\hat{f}_d - f_{d,o}|, \\ e_\phi &= |\hat{\phi} - \phi_o|, \qquad e_\theta = |\hat{\theta} - \theta_o|. \end{aligned} \tag{25}$$

We summarize the error distribution using the 50th, 90th, and 99th percentiles, denoted by $P_{50}(\cdot)$, $P_{90}(\cdot)$, and $P_{99}(\cdot)$, respectively. In addition to the absolute error values, Table II reports the improvement ratio defined as the grid-based error divided by the corresponding ML error.

Table II shows that the proposed estimator consistently outperforms direct grid-based DDAE peak estimation across all four parameters. For delay, the proposed method reduces the median and 90th-percentile errors by factors of $5.76\times$ and $2.72\times$, respectively, while the gain at the 99th percentile is smaller at $1.07\times$. For Doppler, the improvement is consistently around $2\times$ across P_{50}, P_{90}, and P_{99}. The largest gains are observed for angular estimation: azimuth improves by $17.28\times$, $9.95\times$, and $7.13\times$ at P_{50}, P_{90}, and P_{99}, respectively, while elevation improves by $15.73\times$, $9.03\times$, and $11.40\times$. These results indicate that the learned estimator provides not only lower median error, but also substantially improved tail behavior relative to direct grid-based estimation.

Fig. 3 further illustrates this behavior through the full CDFs of the absolute errors. In all four panels, the ML curves are shifted to the left of the grid-based curves, confirming that the proposed model reduces errors over a broad portion of the test distribution rather than at a single operating point. The separation is particularly pronounced for azimuth and elevation, where direct peak selection is strongly constrained by the discretized angular map, whereas the learned model is able to infer more accurate continuous-valued angular parameters from the compact DDAE structure. For Doppler, the improvement is more moderate but remains consistent across percentiles, indicating stable sub-grid refinement. For delay, the gain is strongest in the median and upper-middle region of the CDF, while the extreme tail remains comparatively closer to the direct grid-based baseline. The reported errors are obtained over a test set with geometry-dependent path loss and hence nonuniform effective SNR across samples.

Overall, the results show that the proposed coarse-to-fine estimator can extract meaningful sub-grid information from the selected DDAE representation even after 8-bit log-magnitude quantization. This supports the main claim of the paper: compact DDAE-domain learning can significantly improve azimuth, elevation, Doppler, and delay estimation over direct peak-based sensing without requiring the full complex CFR tensor.

VI. CONCLUSION

This paper presented a compact DDAE-domain learning framework for bistatic OFDM sensing. Using an 8-bit DDAE input and a lightweight coarse-to-fine estimator, the proposed method consistently improves azimuth, elevation, Doppler, and delay estimation over direct grid-based DDAE peak selection while reducing feedback overhead and enabling sensing-function processing without relying on full-CFR super-resolution methods. This also provides a path toward future extensions based on multi-link fusion, more realistic channel and hardware impairments, and monostatic operation. The same framework can also be extended to the monostatic case ($\mathbf{p}_{\text{T}} = \mathbf{p}_{\text{R}}$) after adequate direct-path/self-interference suppression and hardware-impairment mitigation,

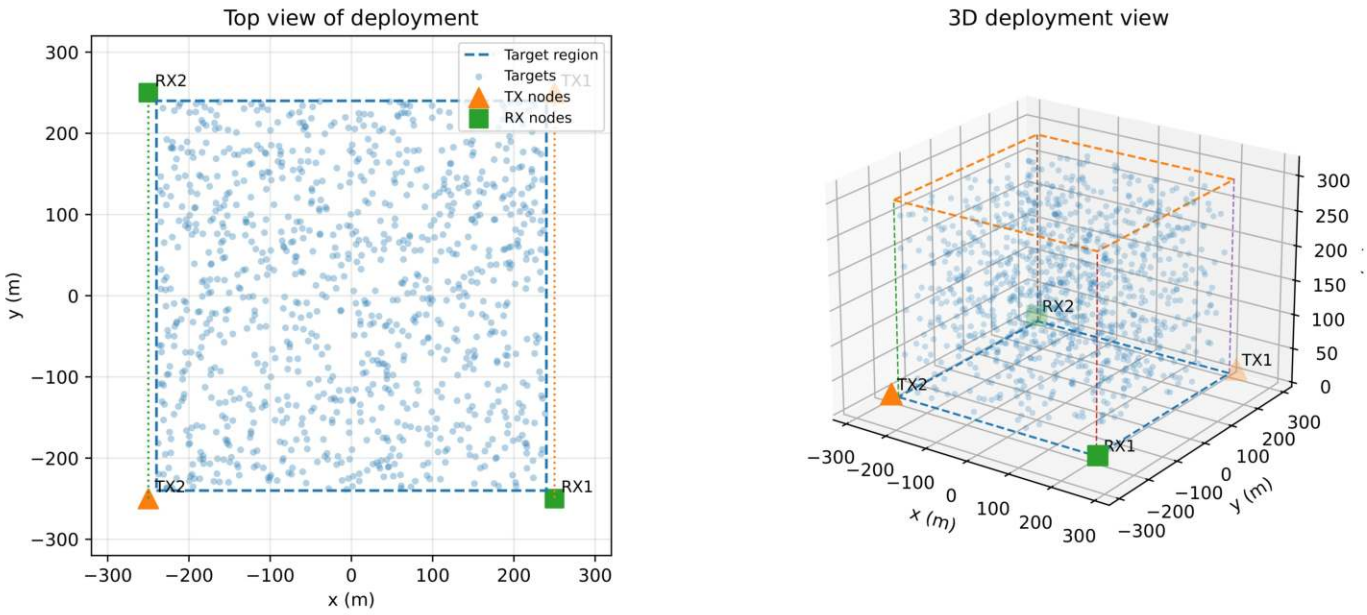

Fig. 2. Deployment geometry used for dataset generation. Two TXs and two RXs are placed on a rectangular grid with inter-site distance 500 m, yielding four bistatic TX–RX combinations. Targets are uniformly distributed over the region used for training and evaluation.

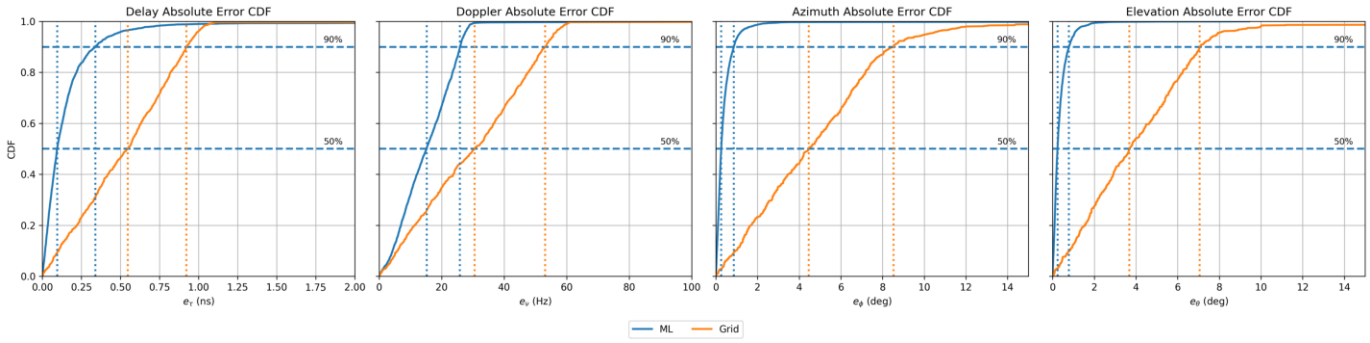

Fig. 3. CDF of absolute estimation errors for delay, Doppler, azimuth, and elevation for the proposed ML estimator and the direct grid-based baseline. Dashed horizontal lines indicate the 50th and 90th percentiles, and dotted vertical lines mark the corresponding error values.

with the bistatic geometry reducing to the corresponding monostatic model.

REFERENCES

[1] ITU-R, "Framework and overall objectives of the future development of imt for 2030 and beyond," Recommendation ITU-R M.2160-0, 2023.

[2] 3GPP, "Study on integrated sensing and communication," TR 22.837, Release 19, 2024.

[3] ——, "Study on stage 2 for integrated sensing and communication," TR 23.700-14, Release 20, 2025.

[4] W. Zhou, R. Zhang, G. Chen, and W. Wu, "Integrated sensing and communication waveform design: A survey," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 1930–1949, 2022.

[5] 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 Transactions on Vehicular Technology*, vol. 72, no. 3, pp. 3250–3263, 2023.

[6] Y. Awad, A. Celik, and A. Eltawil, "Ssb-based multi-target detection and joint velocity, angle, and range estimation in 5g-nr," *IEEE Communications Letters*, vol. 29, no. 6, pp. 1340–1344, 2025.

[7] 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 Transactions on Signal Processing*, vol. 72, pp. 3805–3818, 2024.

[8] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Transactions on Information Theory*, vol. 20, no. 5, pp. 591–598, 1974.

[9] P. Stoica and A. Nehorai, "Music, maximum likelihood, and cramer-rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.

[10] M. Ade Krisna Respati and B. M. Lee, "A survey on machine learning enhanced integrated sensing and communication systems: Architectures, algorithms, and applications," *IEEE Access*, vol. 12, pp. 170 946–170 964, 2024.

[11] J. M. Mateos-Ramos, B. Chatelier, C. Häger, M. F. Keskin, L. Le Magoarou, and H. Wymeersch, "Semi-supervised end-to-end learning for integrated sensing and communications," in *2024 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, 2024, pp. 132–138.

[12] J. Miguel Mateos-Ramos, C. Häger, M. Furkan Keskin, L. Le Magoarou, and H. Wymeersch, "Model-based end-to-end learning for multi-target integrated sensing and communication under hardware impairments," *IEEE Transactions on Wireless Communications*, vol. 24, no. 3, pp. 2574–2589, 2025.