



## **Pilot-Based End-to-End Radio Positioning and Mapping for ISAC: Beyond Point-Based Landmarks**

Downloaded from: <https://research.chalmers.se>, 2026-04-16 06:21 UTC

Citation for the original published paper (version of record):

Ge, Y., Keskin, M., Chen, H. et al (2025). Pilot-Based End-to-End Radio Positioning and Mapping for ISAC: Beyond Point-Based Landmarks. Proceedings - IEEE Global Communications Conference, GLOBECOM

N.B. When citing this work, cite the original published paper.

© 2025 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

(article starts on next page)

# Pilot-Based End-to-End Radio Positioning and Mapping for ISAC: Beyond Point-Based Landmarks

Yu Ge<sup>\*</sup>, Musa Furkan Keskin<sup>\*</sup>, Hui Chen<sup>\*</sup>, Ossi Kaltiokallio<sup>†</sup>, Mengting Li<sup>\*‡</sup>,  
Mikko Valkama<sup>†</sup>, Christos Masouros<sup>§</sup>, Henk Wymeersch<sup>\*</sup>

<sup>\*</sup>Chalmers University of Technology, Sweden, <sup>†</sup>Tampere University, Finland,

<sup>‡</sup>Aalborg University, Denmark, <sup>§</sup>University College London, The United Kingdom

**Abstract**—Integrated sensing and communication enables simultaneous communication and sensing tasks, including precise radio positioning and mapping, essential for future 6G networks. Current methods typically model environmental landmarks as isolated incidence points or small reflection areas, lacking detailed attributes essential for advanced environmental interpretation. This paper addresses these limitations by developing an end-to-end cooperative uplink framework involving multiple base stations and users. Our method uniquely estimates extended landmark objects and incorporates obstruction-based outlier removal to mitigate multi-bounce signal effects. Validation using realistic ray-tracing data demonstrates substantial improvements in the richness of the estimated environmental map.

**Index Terms**—ISAC, 5G/6G, positioning and mapping, extended object, outlier removal, raytracing.

## I. INTRODUCTION

Integrated sensing and communication (ISAC), which enables the radio system to simultaneously provide communication and sensing abilities, has gained significant attention in the evolution of 5G mobile radio systems and has become a pivotal pillar of 6G networks [1]–[3]. Sensing using radio signals is at the heart of ISAC, which encompasses the estimation of the states of both connected devices and passive objects in the propagation environment, also termed as radio *positioning* and *mapping*, respectively [4], [5].

Intense research activities in radio positioning and mapping have been taking place, due to its wide range of foreseen applications [5], [6]. High time and angular resolutions are enabled by the large available bandwidth and antenna arrays in 5G and 6G systems. By exploiting geometric information from the propagation paths, including both line-of-sight (LoS) and non-line-of-sight (NLoS) paths [7], [8], the user equipment (UE) trajectory and the map of reflecting or scattering objects in the environment can be jointly estimated. Real-world experiments in [9], [10] demonstrate the feasibility of radio positioning and mapping, which rely solely on transmissions between a single base station (BS) and a single UE. However, all these works either model the landmarks in the environment as single points and only estimate the positions of points, or recover the incidence points (IPs) of the transmitted signals with the landmarks. However, such a way of modeling cannot convey richer attributes of landmarks such as object size, shape, types, and others, all of which are crucial to higher-level semantic understanding and digital radio twins [11].

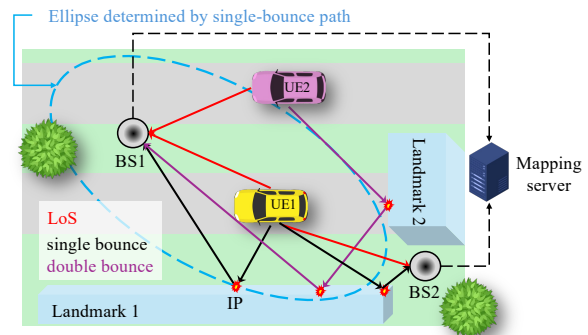


Fig. 1. An example of the considered scenario with 2 BSs and 2 UEs. It also provides an illustrative demonstration of the closed-form solution for estimating the IP of a single-bounce path.

Radio positioning and mapping considering an extended object model, which contains richer attributes of landmarks than only position and considers multiple measurements from the same landmark, has also been explored, for example, in [12]–[14]. However, [12], [13] support simultaneous UE localization and landmark mapping, and can distinguish different types of landmarks, but other features of the landmarks, such as size or shape, are not considered. The approach in [13] associates a small reflection/scattering area to each estimated IP, but landmarks are still represented as collections of mapped IPs. The sizes and shapes of landmarks are considered in [14] by incorporating a probabilistic framework; however, the problem is not solved in an end-to-end manner. In addition, this approach assumes only single-bounce paths and considers a single UE without any cooperation.

In this paper, we study the radio positioning and mapping problem in the multi-UE-multi-BS uplink system, as illustrated in Fig 1, where each BS localizes its connected UEs on its own and all BSs cooperate with each other by sharing its own estimated IPs to a fusion center to create a more refined map. The problem is solved in an end-to-end manner, which includes a channel modeling, a high-resolution channel estimator, a low-complexity positioning and mapping algorithm for each BS, and a post-processing mapping refinement process for the fused map. The main contributions of this paper are as follows: (i) We propose a novel method that can estimate the extended objects, not only separate points, in the field of radio mapping; (ii) We develop an outlier removal method which is based on obstruction

detection to remove the artificial landmarks caused by multi-bounce paths; (iii) We validate the effectiveness of the end-to-end framework using realistic ray-tracing data, demonstrating the ability of the proposed post-processing algorithm to create more subtle map.

*Notations:* Scalars (e.g.,  $x$ ) are denoted in italic, vectors (e.g.,  $\mathbf{x}$ ) in bold lower-case letters, matrices (e.g.,  $\mathbf{X}$ ) in bold capital letters, sets (e.g.,  $\mathcal{X}$ ) in calligraphic. Transpose is denoted by  $(\cdot)^\top$ , the Hermitian transpose is denoted by  $(\cdot)^H$ , and the outer product is denoted by  $\circ$ .

## II. SIGNAL MODEL

In this paper, we consider a multi-BS and multi-UE scenario, where each BS performs bistatic sensing using uplink signals to localize UEs and map the surrounding environment. After that, multiple BSs share sensing results to a fusion center to create more refined results. In this section, the state models and the received signal model are introduced.

### A. State Model

In the considered environment, there are multiple UEs moving over time, and each UE is equipped with a single antenna. The state of the  $i$ -th UE at the time step  $k$  is described by the 3D position of its antenna in the global reference coordinate system as  $\mathbf{x}_{\text{UE},k}^i = [x_{\text{UE},k}^i, y_{\text{UE},k}^i, z_{\text{UE},k}^i]^\top$ . Neither the total number of UEs nor the specific UE states are known to the BSs. Unlike UEs, all BSs are fixed over time, and each BS is equipped with a uniform rectangular array (URA). The state of the  $p$ -th BS contains a 3D location  $\mathbf{x}_{\text{BS}}^p = [x_{\text{BS}}^p, y_{\text{BS}}^p, z_{\text{BS}}^p]^\top$ , which denotes the location of the center of its URA in the global reference coordinate system, and a 3D orientation  $\boldsymbol{\psi}_{\text{BS}}^p = [\varepsilon_{\text{BS}}^p, \nu_{\text{BS}}^p, \gamma_{\text{BS}}^p]^\top$ , which is ordered as roll, pitch, and yaw, respectively, and denotes the Euler angles of the orientation of the URA with respect to the global reference coordinate system. The total number of BSs and specific BS states are known. Each BS has a limited coverage, and only provide positioning services for UEs in its field-of-view (FoV). Apart from the BSs, there are other unknown landmarks in the considered environment. The state of the  $m$ -th landmark can be parameterized as a convex shape, denoted as  $\mathcal{H}^m$ , which contains a list of points that describe the convex hull of the landmark.

### B. Received Signal Model

In this paper, we follow the round-trip-time (RTT) protocol, where a UE responds directly to the corresponding BS when it receives downlink signals from that BS [15]. Suppose that at time step  $k$ , the  $p$ -th BS receives uplink signals from the  $i$ -th UE, and the signal can arrive via a LoS path and/or via NLoS paths resulting from reflections, diffractions, or diffusions off landmarks. Each landmark can generate multiple NLoS paths, and the signals can interact with several landmarks before reaching the BS. Considering a short transmission interval (neglecting Doppler effects) and Orthogonal Frequency Division Multiplexing (OFDM) transmissions with  $N_{\text{OFDM}}$  symbols and  $S$  subcarriers, the received signal at the  $p$ -th BS from the  $i$ -th

UE for the  $g$ -th OFDM symbol at the  $\kappa$ -th subcarrier and time step  $k$  can be expressed as [16]<sup>1</sup>

$$\mathbf{y}_{\kappa,g,k} = \sum_{m=0}^{M_k} \sum_{l=1}^{L_k^m} \rho_k^{m,l} \mathbf{a}(\boldsymbol{\theta}_k^{m,l}) e^{-j2\pi\kappa\Delta f\tau_k^{m,l}} s_{\kappa,g} + \boldsymbol{\omega}_{\kappa,g,k}, \quad (1)$$

where  $\mathbf{y}_{\kappa,g,k}$  is the received signal across the BS antenna array at the subcarrier  $\kappa$ ,  $s_{\kappa,g}$  is the precoded pilot signal at the subcarrier  $\kappa$  at the UE side,  $\boldsymbol{\omega}_{\kappa,g,k}$  is the additive white Gaussian noise,  $\mathbf{a}(\cdot)$  is the steering vector of the BS antenna arrays, and  $\Delta f$  is the subcarrier spacing. There are  $M_k + 1$  visible landmarks at time step  $k$ , with  $m = 0$  representing the BS, and  $L_k^m$  is the number of paths from each landmark. Each path can be characterized by a complex gain  $\rho_k^{m,l}$ , a time of arrival (TOA)  $\tau_k^{m,l}$ , and an angles of arrival (AOA) pair  $\boldsymbol{\theta}_k^{m,l} = [\theta_{\text{az},k}^{m,l}, \theta_{\text{el},k}^{m,l}]^\top$  in azimuth and elevation. The AOA pair  $\boldsymbol{\theta}_k^{m,l}$  is determined by the direction of the signals that arrive at the receiver for the  $(m,l)$ -th path, and the TOA  $\tau_k^{m,l}$  is determined by the propagation distance of the  $(m,l)$ -th path (the clock offset between the transmitter and the receiver is assumed to be cancelled out through the RTT protocol).

## III. END-TO-END PROCESS

To address the positioning and mapping problem, the system exploits the received signals described in (1) to localize the UEs and map the surrounding environment. This section introduces two key components: the channel parameter estimator, which converts received signals into geometric information, and the snapshot positioning and mapping algorithm, which estimates the positions of UEs and the locations of IPs.

### A. Channel Parameter Estimation

1) *3-D Tensor Observations:* We present an ESPRIT-based high-resolution channel estimator to extract geometric path parameters. In this model, the constant modulus pilots  $s_{\kappa,g}$  are eliminated through conjugate multiplication. The signal expression in (1) is reformulated into a three-dimensional tensor  $\mathcal{Y}_k \in \mathbb{C}^{N_z \times N_x \times S}$ , given by

$$\mathcal{Y}_k = \sum_{t=1}^{T_k} \rho_k^t \mathbf{a}_z(\boldsymbol{\theta}_k^t) \circ \mathbf{a}_x(\boldsymbol{\theta}_k^t) \circ \mathbf{d}(\tau_k^t) + \mathcal{N}_k, \quad (2)$$

where  $T_k = \sum_{m=0}^{M_k} L_k^m$  denotes the number of all paths at time step  $k$ , and  $N_x$  and  $N_z$  are the numbers of elements along the horizontal and vertical axes of the URA, respectively, and the integration gains over multiple symbols are encapsulated within  $\rho_k^t$ . The spatial steering vectors in the horizontal and vertical domains are defined, respectively, as  $\mathbf{a}_x(\boldsymbol{\theta}) \in \mathbb{C}^{N_x \times 1}$  and  $\mathbf{a}_z(\boldsymbol{\theta}) \in \mathbb{C}^{N_z \times 1}$ , with their elements given by  $[\mathbf{a}_x(\boldsymbol{\theta})]_n = e^{j\frac{2\pi}{\lambda} d_x n \cos(\theta_{\text{el}}) \sin(\theta_{\text{az}})}$  and  $[\mathbf{a}_z(\boldsymbol{\theta})]_n = e^{j\frac{2\pi}{\lambda} d_z n \sin(\theta_{\text{el}})}$ , respectively. Here,  $d_x$  and  $d_z$  represent the spacing between array elements. The complete spatial-domain steering vector is obtained through the Kronecker product:  $\mathbf{a}(\boldsymbol{\theta}) = \mathbf{a}_x(\boldsymbol{\theta}) \otimes \mathbf{a}_z(\boldsymbol{\theta})$ . Additionally, the frequency-domain steering vector  $\mathbf{d}(\tau) \in \mathbb{C}^{S \times 1}$  is defined element-wise as

<sup>1</sup>For notational simplicity, we omit the indices of the BS and UE. Unless otherwise stated, all descriptions in the remainder of this section refer specifically to the transmission between the  $p$ -th BS and the  $i$ -th UE.

$[\mathbf{d}(\tau)]_\kappa = e^{-j2\pi\kappa\Delta f\tau}$ , while  $\mathcal{N}_k$  accounts for the noise component in the tensor representation.

2) *Tensor Decomposition With Spatial Augmentation*: To extract the path parameters from the tensor expression in (2), we utilize a high-resolution channel estimation method that is based on tensor decomposition, as proposed in [15]. The key idea is to overcome the rank deficiency issue inherent in conventional CP decomposition (CPD) approaches. To achieve this, we reshape the original observation tensor into a new form,  $\tilde{\mathbf{Y}}_k \in \mathbb{C}^{N_z(n_z+1) \times N_x(n_x+1) \times V}$ , where  $V = S - n_z - n_x$ . This transformation is achieved by augmenting the spatial dimensions using frequency-domain data, following the spatial augmentation (SA) framework in [15, Eq. (15)]. After applying SA, the signal tensor can be expressed as

$$\tilde{\mathbf{Y}}_k = \sum_{t=1}^{T_k} \rho_k^t \tilde{\mathbf{a}}_z(\boldsymbol{\theta}_k^t, \tau_k^t) \circ \tilde{\mathbf{a}}_x(\boldsymbol{\theta}_k^t, \tau_k^t) \circ \tilde{\mathbf{d}}(\tau_k^t) + \tilde{\mathcal{N}}_k, \quad (3)$$

where  $\tilde{\mathcal{N}}_k$  denotes the reshaped form of  $\mathcal{N}_k$ , and

$$\tilde{\mathbf{a}}_z(\boldsymbol{\theta}, \tau) = \mathbf{a}_z(\boldsymbol{\theta}) \otimes [\mathbf{d}(\tau)]_{1:n_z+1} \in \mathbb{C}^{N_z(n_z+1) \times 1}, \quad (4a)$$

$$\tilde{\mathbf{a}}_x(\boldsymbol{\theta}, \tau) = \mathbf{a}_x(\boldsymbol{\theta}) \otimes [\mathbf{d}(\tau)]_{1:n_x+1} \in \mathbb{C}^{N_x(n_x+1) \times 1}, \quad (4b)$$

$$\tilde{\mathbf{d}}(\tau) = [\mathbf{d}(\tau)]_{1:V} \in \mathbb{C}^{V \times 1}. \quad (4c)$$

In this formulation, the augmented spatial-frequency domain representations  $\tilde{\mathbf{a}}_z$  and  $\tilde{\mathbf{a}}_x$  allow for improved identifiability of the path parameters by effectively increasing the rank of the observed tensor, thereby enabling more accurate decomposition and estimation. Following this, we proceed by applying CPD to the augmented tensor  $\tilde{\mathbf{Y}}_k$  as expressed in (3). The decomposition problem is formulated as

$$\min_{\{\mathbf{a}_z^t, \mathbf{a}_x^t, \mathbf{d}^t\}_{t=1}^{T_k}} \|\tilde{\mathbf{Y}} - \tilde{\mathbf{Y}}_k\|_F^2 \quad (5a)$$

$$\text{s.t. } \hat{\mathbf{Y}} = \sum_{t=1}^{T_k} \mathbf{a}_z^t \circ \mathbf{a}_x^t \circ \mathbf{d}^t. \quad (5b)$$

Using the structure of the augmented steering vectors introduced in (4), the path-specific parameters for all components at time step  $k$  can be retrieved from the decomposition results in (5) as follows:

$$\hat{\tau}_k^t = -\frac{\angle([\mathbf{J}_1 \mathbf{d}^t]^H [\mathbf{J}_2 \mathbf{d}^t])}{2\pi\Delta f}, \quad (6a)$$

$$\hat{\theta}_{\text{el},k}^t = \arg \max_{\theta_{\text{el}}} |(\mathbf{a}_z^t)^H \tilde{\mathbf{a}}_z(\boldsymbol{\theta}, \hat{\tau}_k^t)|^2, \quad (6b)$$

$$\hat{\theta}_{\text{az},k}^t = \arg \max_{\theta_{\text{az}}} |(\mathbf{a}_x^t)^H \tilde{\mathbf{a}}_x(\theta_{\text{az}}, \hat{\theta}_{\text{el},k}^t, \hat{\tau}_k^t)|^2. \quad (6c)$$

Here,  $\mathbf{J}_1$  and  $\mathbf{J}_2$  refer to selection matrices that extract the first and final  $V - 1$  entries of any right-multiplied vector, respectively. Once the delay and angular parameters have been estimated, the path gains are inferred via a least-squares approach, i.e.,  $\hat{\rho}_k = \mathbf{B}_k^\dagger \tilde{\mathbf{y}}_k$ . In this expression,  $\mathbf{B}_k \in \mathbb{C}^{N_z N_x S \times T_k}$  is a matrix whose  $t$ -th column is defined as

$$[\mathbf{B}_k]_{:,t} = \mathbf{d}(\hat{\tau}_k^t) \otimes \mathbf{a}_x(\hat{\theta}_{\text{az},k}^t, \hat{\theta}_{\text{el},k}^t) \otimes \mathbf{a}_z(\hat{\theta}_{\text{az},k}^t, \hat{\theta}_{\text{el},k}^t), \quad (7)$$

$\tilde{\mathbf{y}}_k = \text{vec}(\tilde{\mathbf{Y}}_k) \in \mathbb{C}^{N_z N_x S \times 1}$ ,  $\hat{\rho}_k \in \mathbb{C}^{T_k \times 1}$  with  $[\hat{\rho}_k]_t = \hat{\rho}_k^t$  denoting the estimate of the path gain  $\rho_k^t$  in (2), and  $(\cdot)^\dagger$  denotes Moore-Penrose pseudo-inverse.

## B. User Positioning and Landmark Mapping

1) *UE Positioning*: The ESPRIT channel parameter estimator outputs a set of estimated paths for each UE-BS pair, denoted by  $\{\hat{\tau}_k^t, \hat{\boldsymbol{\theta}}_k^t\}_{t \in \{1, \dots, \hat{T}_k\}}$ , where  $\hat{T}_k$  is the total number of estimated paths. These paths may include both LoS and/or NLoS components. For UE positioning, only the LoS paths are used, since the IPs of the NLoS paths are unknown. We further assume that the UEs move on a flat ground plane in the considered scenarios, their heights are known, and thus only the horizontal coordinates,  $x_{\text{UE},k}$  and  $y_{\text{UE},k}$ , need to be estimated. Given the measurement  $\mathbf{z}_k = [\tau_k^t, \boldsymbol{\theta}_k^t, \boldsymbol{\theta}_{k,\text{LoS}}^t]$  for the LoS path, which is assumed to be the shortest estimated path, and its associated covariance matrix  $\mathbf{R}_k$ , the unknown UE coordinates  $(x_{\text{UE},k}, y_{\text{UE},k})$  can be estimated using a maximum likelihood (ML) estimator, defined as

$$\hat{x}_{\text{UE},k}, \hat{y}_{\text{UE},k} = \arg \min_{x_{\text{UE},k}, y_{\text{UE},k}} (\mathbf{z}_k - \mathbf{h}(x_{\text{UE},k}, y_{\text{UE},k}, \mathbf{x}_{\text{BS}}))^T \mathbf{R}_k^{-1} (\mathbf{z}_k - \mathbf{h}(x_{\text{UE},k}, y_{\text{UE},k}, \mathbf{x}_{\text{BS}})), \quad (8)$$

where  $\mathbf{h}(x_{\text{UE},k}, y_{\text{UE},k}, \mathbf{x}_{\text{BS}})$  denotes the measurement function that maps the positions of the UE and the BS to the TOA and AOA of the LoS path. The estimation problem can be solved using gradient descent, initialized with a closed-form least-squares (LS) solution as described in [17].

2) *Landmark Mapping*: After estimating the UE position, we proceed to map the environment by estimating the IPs associated with the remaining NLoS paths. Since each resolved path provides only three observable parameters, we can estimate at most one unknown IP. Determining multiple IPs for multi-bounce paths is infeasible without additional information, making the problem underdetermined. In addition, it is not known whether each NLoS path corresponds to a single-bounce or a multi-bounce path. Therefore, we conservatively treat all remaining NLoS paths as single-bounce paths and estimate the corresponding IP for each path accordingly, even though some, in fact, are multi-bounce paths.

For a NLoS path with measurements  $\{\tau_k^t, \boldsymbol{\theta}_k^t\}$ , the signal is modeled as propagating in the order UE  $\rightarrow$  IP  $\rightarrow$  BS, where  $\mathbf{x}_{\text{IP},k}^t$  denotes the position of the corresponding IP. The IP lies along the line originating from the BS in the direction of  $\boldsymbol{\theta}_k^t$ , and the total path length from the UE to the BS via the IP is constrained by the measured delay  $\tau_k^t$ . Based on these geometric constraints, the position  $\mathbf{x}_{\text{IP},k}^t$  can be determined by

$$\begin{cases} \|\mathbf{x}_{\text{IP},k}^t - \mathbf{x}_{\text{UE},k}^t\| + \|\mathbf{x}_{\text{IP},k}^t - \mathbf{x}_{\text{BS}}\| = \tau_k^t, \\ (\mathbf{x}_{\text{IP},k}^t - \mathbf{x}_{\text{BS}})^T \mathbf{u}_k^t = \|\mathbf{x}_{\text{IP},k}^t - \mathbf{x}_{\text{BS}}\|, \end{cases} \quad (9)$$

where  $\mathbf{u}_k^t$  denotes the unit direction vector determined by  $\boldsymbol{\theta}_k^t$

$$\mathbf{u}_k^t = \mathbf{R}(\boldsymbol{\psi}_{\text{BS}})^T \begin{bmatrix} \cos(\theta_{\text{az},k}^t) \cos(\theta_{\text{el},k}^t) \\ \sin(\theta_{\text{az},k}^t) \cos(\theta_{\text{el},k}^t) \\ \sin(\theta_{\text{el},k}^t) \end{bmatrix}, \quad (10)$$

with  $\mathbf{R}(\boldsymbol{\psi}_{\text{BS}})$  denoting the rotation matrix that transforms coordinates from the global reference frame to the local coordinate system of the BS URA. The first entry in (9) ensures that  $\mathbf{x}_{\text{IP},k}^t$  lies on an ellipse with foci at  $\mathbf{x}_{\text{BS}}$  and  $\mathbf{x}_{\text{UE},k}$ . The second entry specifies that the IP lies on a half-line

originating from  $\mathbf{x}_{\text{BS}}$  in the direction of  $\mathbf{u}_k^t$ . Consequently,  $\mathbf{x}_{\text{IP},k}^t$  is determined as the intersection point of this half-line and the ellipse. An illustrative representation of this geometric interpretation is shown in Fig. 1. While a closed-form solution to (9) exists, the derivation is omitted for brevity. This process is repeated for all resolved NLoS paths to estimate the corresponding IPs.

#### IV. POST-PROCESSING

At each time step, each BS sends a set of estimated IPs to a central fusion center, forming a global set  $\mathcal{D}$ . This section introduces a three-phase method to estimate extended landmarks and detect outliers—i.e., estimated IPs of multi-bounce paths mistakenly treated as single-bounce reflections.

##### A. Phase I: DBSCAN for Clustering

An IP represents the point where a signal hits a landmark. As such, IPs corresponding to paths bounced from the same landmark should lie on that landmark. This allows us to estimate a landmark that best fits the set of IPs originating from the same source. However, the origin source of each NLoS path is initially unknown, making it infeasible to directly estimate landmarks from the complete set of collected paths. To address this, we first cluster all IPs that likely originate from the same source, based on the assumption that such points are spatially close to one another. For this clustering, we employ the density-based spatial clustering of applications with noise (DBSCAN) algorithm [18], which groups points based on density-reachability and classifies them either as members of a cluster or as outliers.

The DBSCAN algorithm is controlled by two parameters  $\epsilon$  and  $N_{\text{min}}$ . A core point is defined as a point with at least  $N_{\text{min}}$  points (including itself) within the distance of  $\epsilon$ , where we use the Euclidean distance in this paper. A point is directly reachable from a core point if the distance between two points is within  $\epsilon$ , and points are only directly reachable from core points. Moreover, a point is reachable from a given core point if it can be reached through a series of core points starting from the given core point, where the current core point is directly reachable from the last core point. Please note that all the points on the path from the given core point to the target point should be core points, with the possible exception of the target point. Any points that are not reachable from any other point are defined as outlier points. The DBSCAN algorithm clusters a core point with all its reachable points, no matter core or non-core points, to create a cluster, and forms different clusters and picks up all outlier points. If we input  $\mathcal{D}$ , the output of the DBSCAN algorithm will be a list of point clusters  $\mathcal{L}_1, \dots, \mathcal{L}_C$ , and set of outlier points  $\mathcal{P}$ .

##### B. Phase II: Convex Hull Extraction

Applying DBSCAN to  $\mathcal{D}$  yields clusters  $\mathcal{L}_c = \{\mathbf{x}_{\text{IP}}^1, \dots, \mathbf{x}_{\text{IP}}^{|\mathcal{L}_c|}\} \subset \mathcal{D}$ , where each cluster represents a group of estimated IPs that are spatially close and likely to originate from the same physical landmark. To extract a geometric boundary for subsequent tasks, such as occlusion

detection or shape analysis, we compute the convex hull of each cluster. The convex hull, denoted as  $\text{CH}(\mathcal{L}_c)$  for the defined cluster  $\mathcal{L}_c$ , is the smallest convex set that encloses all points in  $\mathcal{L}_c$ , defined as

$$\text{CH}(\mathcal{L}_c) = \left\{ \mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x} = \sum_{n=1}^{|\mathcal{L}_c|} \alpha_n \mathbf{x}_{\text{IP}}^n, \alpha_n \geq 0, \sum_{n=1}^{|\mathcal{L}_c|} \alpha_n = 1 \right\}. \quad (11)$$

Any point  $\mathbf{x}$  within the convex hull is represented as a convex combination of the points in the cluster. Efficient algorithms such as Quickhull [19] are commonly used to compute  $\text{CH}(\mathcal{L}_c)$ , yielding a compact and computationally efficient boundary representation for each cluster, even in the presence of internal irregularities. While the convex hull may not capture concavities in the original shape, its simplicity and efficiency make it well-suited for many real-time 3D applications. Alternative approaches, such as concave hulls, can provide more accurate representations of complex geometries, but their exploration lies outside the scope of this paper.

##### C. Phase III: Outlier Removal

After extraction, a convex hull can be recovered for each cluster, capturing geometric properties of the associated landmarks, such as location, size, and shape. However, not all clusters produced by DBSCAN necessarily correspond to real physical landmarks. Many multi-bounce paths may be misinterpreted as single-bounce reflections, resulting in the formation of false landmarks. Distinguishing between single- and multi-bounce paths based solely on the measurements—angles, delays, or the 3D positions of the estimated IPs—is inherently challenging, as this information does not directly reveal the nature of the underlying bounces.

A key geometric cue for identifying false landmarks is occlusion: IPs generated by multi-bounce paths often appear *behind* real landmarks due to longer propagation distances. Although such paths may yield similar AOA compared to single-bounce paths, they typically exhibit greater TOA, placing the estimated IPs farther from both the BS and UE. These false IPs are often occluded by real landmarks. To eliminate such false landmarks, we introduce a visibility-based test for each estimated IP  $\mathbf{x}_{\text{IP}}^n$ . Let  $\mathbf{x}_{\text{BS}}$  and  $\mathbf{x}_{\text{UE},k}$  denote the corresponding BS and UE positions, respectively, and we define the BS-IP and UE-IP line segments as

$$\mathcal{S}_{\text{BS}}^n = \{\mathbf{x}(\alpha) = \mathbf{x}_{\text{BS}} + \alpha(\mathbf{x}_{\text{IP}}^n - \mathbf{x}_{\text{BS}}) \mid 0 < \alpha < 1\}, \quad (12)$$

$$\mathcal{S}_{\text{UE}}^n = \{\mathbf{x}(\alpha) = \mathbf{x}_{\text{UE}} + \alpha(\mathbf{x}_{\text{IP}}^n - \mathbf{x}_{\text{UE},k}) \mid 0 < \alpha < 1\}. \quad (13)$$

We check whether  $\mathcal{S}_{\text{BS}}^n$  or  $\mathcal{S}_{\text{UE}}^n$  intersects with any other convex hull  $\text{CH}(\mathcal{L}_c)$  formed from another IP cluster  $\mathcal{L}_c$ . An IP is retained only if both lines are unobstructed:

$$\mathcal{S}_{\text{BS}}^n \cap \text{CH}(\mathcal{L}_c) = \emptyset \text{ and } \mathcal{S}_{\text{UE}}^n \cap \text{CH}(\mathcal{L}_c) = \emptyset, \quad \forall \mathcal{L}_c, \mathbf{x}_{\text{IP}}^n \notin \mathcal{L}_c. \quad (14)$$

The remaining visible IPs are used to construct convex hulls that represent the estimated landmarks.

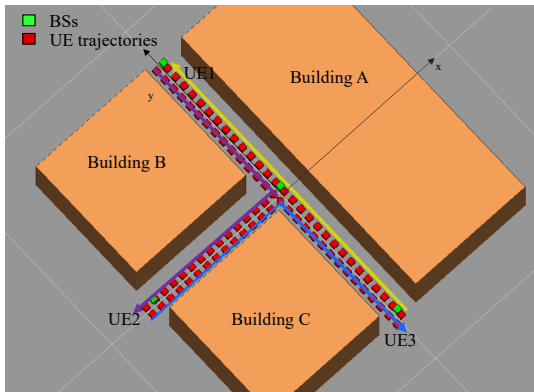


Fig. 2. The ray-tracing simulation is conducted in a T-junction environment, featuring four fixed BSs and three UEs, each following its own trajectory over time. Building A is centered at  $[40\text{ m}, 0\text{ m}, 15\text{ m}]^T$  and measures 140 m in length, 50 m in width, and 30 m in height. Buildings B and C are centered at  $[40\text{ m}, 40\text{ m}, 15\text{ m}]^T$  and  $[40\text{ m}, -40\text{ m}, 15\text{ m}]^T$ , respectively, each measuring 60 m in length, 60 m in width, and 30 m in height.

## V. RESULTS

### A. Simulation Environment

We consider an urban vehicular scenario situated at a T-junction, as illustrated in Fig. 2. The environment consists of three surrounding buildings with brick surfaces. Four BSs and three vehicles that act as UEs are deployed. The BSs are located at  $[0\text{ m}, 0\text{ m}, 15\text{ m}]^T$ ,  $[0\text{ m}, -70\text{ m}, 15\text{ m}]^T$ ,  $[0\text{ m}, 70\text{ m}, 15\text{ m}]^T$ , and  $[-70\text{ m}, 0\text{ m}, 15\text{ m}]^T$ , respectively. The UEs follow distinct trajectories over time: UE1 moves from  $[2\text{ m}, -70\text{ m}]^T$  to  $[2\text{ m}, 70\text{ m}]^T$ . UE2 starts at  $[-70\text{ m}, -2\text{ m}]^T$ , turns right at  $[-2\text{ m}, -2\text{ m}]^T$ , and ends at  $[-2\text{ m}, -70\text{ m}]^T$ . UE3 starts at  $[-2\text{ m}, 70\text{ m}]^T$ , turns right at  $[-2\text{ m}, 2\text{ m}]^T$ , and ends at  $[-70\text{ m}, 2\text{ m}]^T$ . Along each trajectory, samples are taken at 5 m intervals. Each BS is equipped with an  $8 \times 8$  multi-panel antenna array, while each UE has a single antenna mounted at a fixed height of 1.5 m, which is assumed to be known. The antenna radiation patterns follow the specifications in [15]. The coverage radius of each BS is set to 70 m. A UE communicates only with the BSs within the coverage ranges of BSs, ensuring LoS availability for all active BS-UE pairs.

The carrier frequency is set to 27.2 GHz. The system employs OFDM pilot signals composed of 12 symbols, each with constant amplitude. The total active bandwidth is 400 MHz, with a subcarrier spacing of 120 kHz. The transmit power is configured to 10 dBm, while the noise spectral density is set to  $-174$  dBm/Hz and the receiver noise figure to 8 dB. Ground-truth channel data is obtained using the REMCOM Wireless InSite@ray-tracing simulator [20]. At each UE location and time step, 50 propagation paths are retained per BS. On average, each snapshot contains approximately 9 single-bounce paths and 40 double-bounce paths per BS. For clustering using the DBSCAN algorithm,  $\epsilon$  is set to 3 m, and  $N_{\min}$  is set to 5. 100 Monte Carlo (MC) simulations are performed.

Each simulation starts with generating propagation paths using ray-tracing data, followed by channel parameter estimation described in Section III-A. The positioning and mapping method in Section III-B is then applied to estimate

UE positions and the corresponding IPs. Post-processing in Section IV clusters IPs, reconstructs convex hulls, and removes outliers. Positioning performance is evaluated using the sub-meter accuracy rate and the mean absolute error (MAE) across all UEs over time and 100 MC simulations. Mapping performance is assessed qualitatively through visual comparison with the ground-truth map, and quantitatively by measuring the outlier removal rate.

### B. Results and Discussion

We begin by evaluating the positioning performance of all UEs, using the channel parameter estimates corresponding to the LoS path. The LoS path is taken to be the shortest among the resolved paths provided by the channel parameter estimator at each time step. The results indicate that sub-meter accuracy is achieved in 95.9% of the cases, with an overall MAE of 0.21 m. This high positioning accuracy is primarily attributed to the wide signal bandwidth and large antenna array, which offer fine angular and delay resolution, as well as the use of a high-precision channel parameter estimator.

Next, we evaluate the mapping performance and demonstrate how the proposed post-processing methods enhance landmark reconstruction and outlier removal. Fig. 3 illustrates the mapping results *without* applying our post-processing scheme, where the estimated IPs for all resolved paths are shown. All first-order IPs corresponding to single-bounce paths, and ground-truth map are also shown. From the figure, we observe that some estimated IPs align well with the first-order IPs and ground truth map, indicating that certain estimated paths do correspond to true single-bounce paths. In these cases, our mapping method performs effectively, producing IPs that are spatially close to the true IPs. Specifically, out of 870 estimated IPs, 379 lie within 2 m of the ground-truth facades. However, we also notice a significant number of estimated IPs located far from the real first-order IPs and building facades. These errors arise from paths that are either clutter-induced or involve multiple bounces. Since these paths do not conform to the single-bounce assumption, their corresponding IPs are misleading and should be excluded from the mapping process.

Fig. 4 shows the mapping performance with the proposed post-processing scheme. It is evident that most of the spurious IPs have been successfully removed. In total, 43.1% of the estimated IPs are discarded, and the remaining IPs align well with the building facades, demonstrating the effectiveness of our outlier removal method. Furthermore, Fig. 4 also shows the reconstructed convex hulls, which closely match the actual building facades. These convex hulls provide richer geometric information beyond individual IPs, enabling estimation of the approximate size and shape of the facades. Some discontinuities are observed compared to the ground-truth, primarily due to missing or sparse IP estimates in those regions. In such cases, DBSCAN may classify the sparse IPs as outliers. Additionally, larger channel parameter estimation errors occur when the azimuth angle is near zero, leading to increased UE positioning and IP mapping errors, and resulting

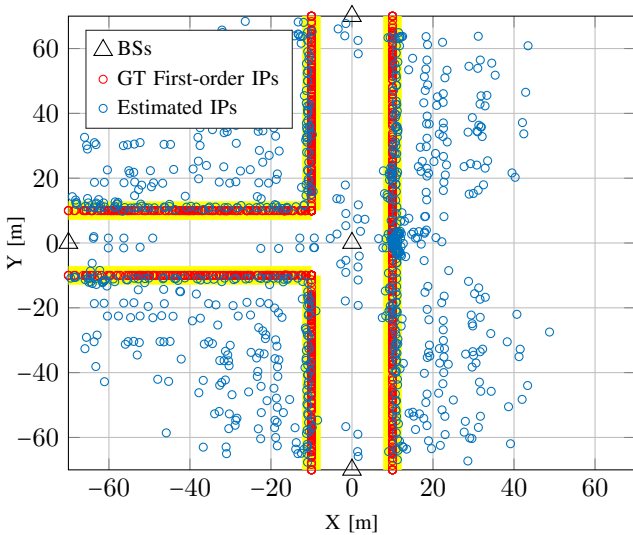


Fig. 3. 2D visualization of the mapping results, showing the estimated IPs corresponding to all resolved channel paths for all UE-BS pairs. The ground truth (denoted as “GT” in the figure) consists of the true IPs from all single-bounce paths across all UE-BS pairs and time steps. All ground-truth IPs lie on the building facades, which are indicated by yellow lines. The results correspond to a single MC simulation.

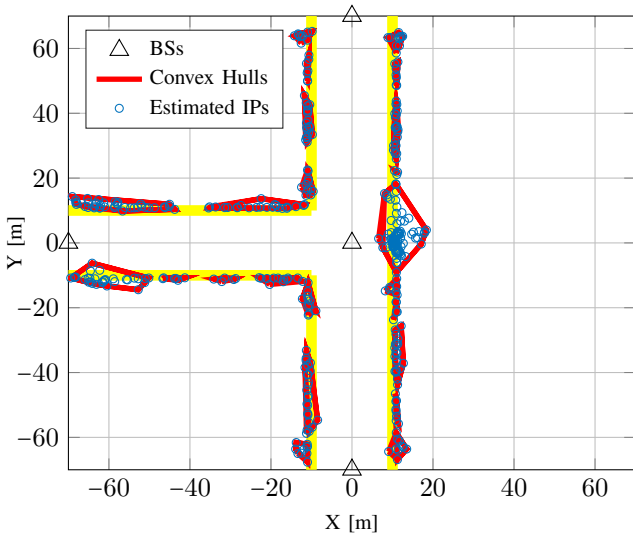


Fig. 4. 2D visualization of the mapping results after applying the proposed post-processing scheme. The results correspond to a single MC simulation.

in a distorted cluster, for example around  $[0 \text{ m}, 10 \text{ m}]^T$ . The mapping performance is expected to improve further with more UEs and/or BSs in the environment, as additional radio paths offer more opportunities for signal interaction with the environment and thus provide more spatial information.

## VI. CONCLUSIONS

In this paper, we address the limitations of existing environmental modeling approaches in ISAC, which typically focus on isolated IPs. Our goal is to enable rich, radio-based environmental mapping for future 6G networks by moving beyond simplistic point-based representations and leveraging more expressive features to enhance sensing performance. To achieve this, we propose an end-to-end cooperative uplink framework involving multiple BSs and UEs.

Within this framework, UEs are localized, extended landmarks are estimated, and a visibility-based outlier removal method is introduced to eliminate clutter caused by multi-bounce paths. Simulation validation using realistic ray-tracing data demonstrates the effectiveness of the proposed method for high-precision positioning and environmental reconstruction in complex wireless scenarios.

## ACKNOWLEDGMENT

This work has been supported by the SNS JU project 6G-DISAC under the EU’s Horizon Europe research and innovation Program under Grant Agreement No. 101139130, SNS JU project 6G-MUSICAL under Grant Agreement No. 101139176 and the Swedish Research Council (VR) through the project 6G-PERCEF under Grant 2024-04390. The authors wish to thank Remcom for providing Wireless InSite@ray-tracer.

## REFERENCES

- [1] A. Liu, *et al.*, “A survey on fundamental limits of integrated sensing and communication,” *IEEE Commun. Surv. Tutor.*, vol. 24, no. 2, pp. 994–1034, 2022.
- [2] F. Liu, *et al.*, “Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond,” *IEEE JSAC*, 2022.
- [3] N. González-Prelcic, *et al.*, “The integrated sensing and communication revolution for 6G: Vision, techniques, and applications,” *Proceedings of the IEEE*, 2024.
- [4] C. Chaccour, *et al.*, “Seven defining features of terahertz (THz) wireless systems: A fellowship of communication and sensing,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 967–993, 2022.
- [5] Y. Ge, *et al.*, “Mmwave mapping and SLAM for 5G and beyond,” in *Integrated Sensing and Communications*. Springer, 2023, pp. 445–475.
- [6] B. Amjad, *et al.*, “Radio SLAM: A review on radio-based simultaneous localization and mapping,” *IEEE Access*, vol. 11, pp. 9260–9278, 2023.
- [7] E. Leitinger, *et al.*, “A belief propagation algorithm for multipath-based SLAM,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5613–5629, 2019.
- [8] Y. Ge, *et al.*, “A computationally efficient EK-PMBM filter for bistatic mmwave radio SLAM,” *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2179–2192, 2022.
- [9] Y. Ge, *et al.*, “Experimental validation of single BS 5G mmwave positioning and mapping for intelligent transport,” *IEEE Transactions on Vehicular Technology*, 2024.
- [10] E. Rastorgueva-Foi, *et al.*, “Millimeter-wave radio SLAM: End-to-end processing methods and experimental validation,” *IEEE Journal on Selected Areas in Communications*, 2024.
- [11] L. U. Khan, *et al.*, “Digital twin of wireless systems: Overview, taxonomy, challenges, and opportunities,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2230–2254, 2022.
- [12] Y. Ge, *et al.*, “5G SLAM using the clustering and assignment approach with diffuse multipath,” *Sensors (Basel, Switzerland)*, vol. 20, no. 16, August 2020. [Online]. Available: <https://doi.org/10.3390/s20164656>
- [13] A. Karttunen, *et al.*, “Towards semantic radio SLAM with landmark feature extraction in mmWave networks,” in *IEEE JC&S*, 2025, pp. 1–6.
- [14] S. Zhai, *et al.*, “Multipath-based SLAM exploiting extended object estimation and classification,” *IEEE Transactions on Wireless Communications*, 2025.
- [15] Y. Ge, *et al.*, “V2X sidelink positioning in FR1: Scenarios, algorithms, and performance evaluation,” *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 10, pp. 2608–2624, 2024.
- [16] R. W. Heath, *et al.*, “An overview of signal processing techniques for millimeter wave MIMO systems,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.
- [17] S. Zhu *et al.*, “A simple approach of range-based positioning with low computational complexity,” *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5832–5836, 2009.
- [18] M. Ester, *et al.*, “A density-based algorithm for discovering clusters in large spatial databases with noise,” in *International Conference on Knowledge Discovery and Data Mining*, 1996, pp. 226–231.
- [19] C. B. Barber, *et al.*, “The quickhull algorithm for convex hulls,” *ACM TOMS*, vol. 22, no. 4, pp. 469–483, 1996.
- [20] Remcom. Wireless InSite - 3D Wireless Prediction Software. Accessed: Oct 9, 2022. [Online]. Available: <https://www.remcom.com/wireless-insite-em-propagation-software>