

Blind Radio Mapping via Spatially Regularized Bayesian Trajectory Inference

Zheng Xing and Juntao Chen

School of Science and Engineering (SSE) and Shenzhen Future Network of Intelligence Institute (FNii-Shenzhen)
The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China

Abstract—Radio maps enable intelligent wireless applications by capturing the spatial distribution of channel characteristics. However, conventional construction methods demand extensive location-labeled data, which are costly and impractical in many real-world scenarios. This paper presents a blind radio map construction framework that infers user trajectories from indoor multiple-input multiple-output (MIMO)-Orthogonal Frequency-Division Multiplexing (OFDM) channel measurements without relying on location labels. It first proves that channel state information (CSI) under non-line-of-sight (NLOS) exhibits spatial continuity under a quasi-specular environmental model, allowing the derivation of a CSI-distance metric that is proportional to the corresponding physical distance. For rectilinear trajectories in Poisson-distributed access point (AP) deployments, it is shown that the Cramer-Rao Lower Bound (CRLB) of localization error vanishes asymptotically, even under poor angular resolution. Building on these theoretical results, a spatially regularized Bayesian inference framework is developed that jointly estimates channel features, distinguishes line-of-sight (LOS)/NLOS conditions and recovers user trajectories. Experiments on a ray-tracing dataset demonstrate an average localization error of 0.68 m and a beam map reconstruction error of 3.3%, validating the effectiveness of the proposed blind mapping method.

Index Terms—Radio map, trajectory inference, spatial continuity, LOS/NLOS discrimination, indoor localization

I. INTRODUCTION

Radio maps link physical locations with channel characteristics, enabling new methodologies for channel state information (CSI) acquisition, tracking, and prediction [1]–[3]. Conventional methods for constructing radio maps predominantly rely on labeled datasets, where the CSI measurements are associated with accurate location information [4]–[6]. However, obtaining such location-labeled CSI data is costly and labor-intensive, often requiring dedicated positioning infrastructure, or manual calibration [7]–[9]. These limitations become particularly pronounced in scenarios demanding rapid deployment or frequent updates, such as dynamic urban environments, shopping malls, and exhibition halls that require frequent reconfiguration. Consequently, developing radio map construction methods that do not rely on location labels is crucial for enhancing the scalability, flexibility, and practicality of wireless network intelligence.

There were some attempts on reducing the reliance on location-labeled measurements. For instance, the work [10] employed Kriging-based space–frequency interpolation to construct radio maps from a limited set of labeled data. Similarly, the work [11] developed a geometry-driven matrix completion approach, leveraging virtual anchor modeling and spatial

consistency to extrapolate multipath angles and delays. Other methods have incorporated Bayesian inference; for example, the work [12] investigated sparse sampling combined with Bayesian learning to optimize sampling locations and account for shadow fading in spectrum recovery. Note that these methods are based on the methodologies of interpolation and supervised learning, and thus, they still necessitate a sufficient amount of location-labeled data, which requires labor cost in radio map construction and update.

Another line of research aims to reduce the dependence on location-labeled data by exploiting additional sensors or side information from the environment. For example, the work [13] investigated ultrawideband (UWB)-assisted positioning, where synchronized angle and time measurements enabled accurate localization. The work [14] employed Light Detection and Ranging (LiDAR)-based scanning to achieve fine-grained environmental perception and self-localization, while the work [15] utilized Inertial Measurement Unit (IMU) measurements to estimate relative positions from accelerometer data. However, these approaches require specialized hardware deployments, incur additional costs, and are often impractical in large-scale or resource-constrained scenarios. Recently, channel charting has emerged as a promising paradigm for radio map construction, wherein high-dimensional CSI is embedded into low-dimensional manifolds using neural architectures such as triplet-based networks [16], Siamese networks [17], and bilateration-based networks [18]. Despite their potential, these methods typically require a small number of labeled points to align the latent and geometric spaces. Moreover, such alignment may fail to faithfully reflect the underlying physical geometry, especially in dense multipath non-line-of-sight (NLOS) environments.

This paper focuses on recovering the location along an unknown trajectory that measures the multiple-input multiple-output (MIMO)-Orthogonal Frequency-Division Multiplexing (OFDM) channels in an indoor environment where there could be NLOS regions. With the recovered location labels, a radio map can be constructed by associating the recovered locations with the channel measurements. The two main challenges addressed in this paper are: (i) *How to infer the location information from a sequence of channel measurements*, where there are multi-paths and fading that give randomness to the channel, and (ii) *How to improve the location inference in NLOS by exploiting the spatial correlation of the channel*.

To tackle these challenges, we develop a spatially regularized Bayesian framework for trajectory inference from

unlabeled channel measurements. The following two principles are investigated. First, we establish a Bayesian model to identify line-of-sight (LOS) measurements, where the angle information can be extracted more accurately compared to the NLOS counterpart. Second, we develop a CSI distance metric for NLOS measurements, where the metric is shown to be proportional to the physical distance of the measurements, and hence, it leads to a spatial regularization for Bayesian trajectory inference.

Specifically, the following technical contributions are made:

- We establish a *quasi-specular environment model* to understand whether there exists a radio signature that is statistically and locally *continuous* in the physical space given the randomness nature of the wireless channel. We derive a spatial continuity theorem and find a theoretical CSI distance metric that is proven to be proportional to the physical distance scaled by the bandwidth if the measurements are obtained in a fully scattered NLOS scenario.
- We investigate whether it is theoretically possible to recover a rectilinear trajectory using a sufficient number of measurements despite an arbitrarily poor accuracy in angle of arrival (AoA) estimation under NLOS. We theoretically show that this is possible under certain conditions on access point (AP) topology, and the corresponding localization error decays as $\mathcal{O}(1/T)$ for T measurements.
- We formulate a spatially regularized Bayesian framework for trajectory inference, in which the user trajectory is recovered jointly with LOS/NLOS assignment and channel feature estimation.
- We evaluate the proposed method on a ray-tracing dataset with uniform linear antenna array (ULA) configuration. The proposed method achieves an average localization error of 0.68 m (1.07 m in NLOS), a LOS/NLOS identification error of 2%, and a relative error of 3.3% in constructing the MIMO beam map based on the estimated trajectory.

The remainder of this paper is organized as follows. Section II introduces the propagation model in MIMO-OFDM systems, the user mobility model, and the Bayesian framework for trajectory inference. Section III develops the quasi-specular environment model, establishes the spatial continuity property of the wireless channel in NLOS region, and derives the Cramer-Rao Lower Bound (CRLB) of the localization error under limited and unlimited regions. Section IV presents the formulation of the spatially regularized Bayesian problem and the design of the trajectory inference algorithm. Section V reports experimental evaluations, and Section VI concludes the paper.

II. SYSTEM MODEL

A. Propagation Model in MIMO-OFDM Networks

Consider a single-antenna mobile user moving in an indoor environment. There are Q APs located at known positions $\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_Q \in \mathbb{R}^2$. Each AP is equipped with a ULA consisting of N_t antennas. For the ease of elaboration, we

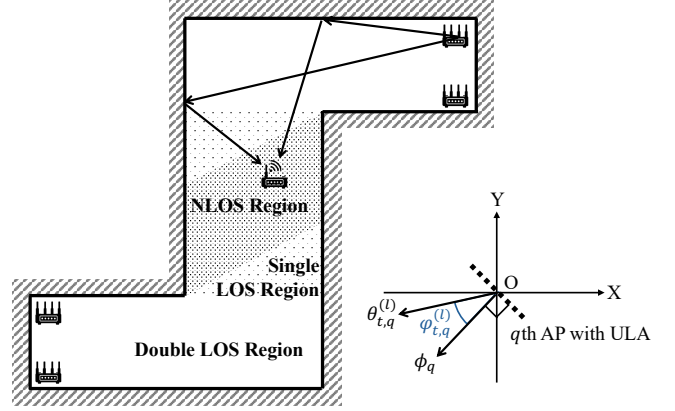


Figure 1. An example indoor environment with NLOS regions.

assume no angular ambiguity in the antenna geometry. For example, assume that the APs are installed on a wall or the antenna arrays are installed with a rear panel, such that there is approximately no signal arriving from the back of the AP.

Consider the XOY coordinate system as shown in Figure 1. For each AP q , denote ϕ_q as the angle of the reference direction that is orthogonal to the array axis of the antenna. For each user position $\mathbf{x}_t$ at time t , denote $\theta_{t,q}^{(l)}$ as the angle of departure (AOD) of the l th propagation path from the q th AP at position $\mathbf{o}_q$ to the user at position $\mathbf{x}_t$ as shown in Figure 1. As a result, the *relative* AOD $\varphi_{t,q}^{(l)} \in (-\pi/2, \pi/2)$ with respect to (w.r.t.) the reference direction ϕ_q satisfies $\theta_{t,q}^{(l)} = (\varphi_{t,q}^{(l)} + \phi_q) \bmod 2\pi$.

Assuming far-field propagation, the steering vector at the q th AP for a propagation path with a relative AOD φ is defined as

$$\mathbf{a}(\varphi) = [1, e^{-j\frac{2\pi}{\lambda}\Delta \sin \varphi}, \dots, e^{-j\frac{2\pi}{\lambda}(N_t-1)\Delta \sin \varphi}]^T \quad (1)$$

where Δ denotes the inter-element spacing, $\lambda = \frac{c}{f_c}$ is the wavelength at the carrier frequency f_c , and $c = 3 \times 10^8$ m/s is the speed of light.

Consider an OFDM system with M subcarriers ($M > N_t$) and bandwidth B . At time slot t , the baseband equivalent MIMO channel $\mathbf{h}_{t,q}^{(m)} \in \mathbb{C}^{N_t}$ of the m th subcarrier at the q th AP is given by¹

$$\mathbf{h}_{t,q}^{(m)} = \sum_{\ell=1}^L \kappa_{t,q}^{(\ell)} e^{-j2\pi \frac{m}{M} B \tau_{t,q}^{(\ell)}} \mathbf{a}(\varphi_{t,q}^{(\ell)}) \quad (2)$$

where L is the number of paths from the q th AP to the user at position $\mathbf{x}_t$. For the path l between AP q and location $\mathbf{x}_t$, $\kappa_{t,q}^{(\ell)} \in \mathbb{C}$ denotes the complex gain, and $\tau_{t,q}^{(\ell)} \in \mathbb{R}_+$ represents the propagation delay.

By stacking the vectors (2), the OFDM channel measured at the q th AP at time slot t is represented by

$$\mathbf{H}_{t,q} = [\mathbf{h}_{t,q}^{(1)}, \mathbf{h}_{t,q}^{(2)}, \dots, \mathbf{h}_{t,q}^{(M)}] \in \mathbb{C}^{N_t \times M}.$$

¹Here, we have implicitly assumed that the delays $\tau_{t,q}^{(\ell)}$ are discretized following a tap-delay model.

B. Mobility Model

Consider that the slot duration is even. The mobility of the user may follow a distribution where a big jump from $\mathbf{x}_{t-1}$ to a far away position $\mathbf{x}_t$ is rare. To capture such a prior information, a widely adopted model is the Gauss-Markov model [19]

$$\mathbf{x}_t - \mathbf{x}_{t-1} = \gamma(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}) + (1 - \gamma)\delta\bar{\mathbf{v}} + \sqrt{1 - \gamma^2}\delta\epsilon \quad (3)$$

which captures the temporal correlation in both the position $\mathbf{x}_t$ and the speed $\mathbf{x}_t - \mathbf{x}_{t-1}$, where $0 < \gamma \leq 1$ is the velocity correlation coefficient, δ is the time slot duration, $\bar{\mathbf{v}}$ is the average velocity, and $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I})$ models the randomness.

We adopt a graph-based approach to convert (3) into a probability model. Specifically, the indoor environment is represented as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes the set of possible discretized positions $\mathbf{p}_i \in \mathbb{R}^2$ placed uniformly over the area of the indoor environment. Denote D_m as the maximum travel distance within a single time slot. The set $\mathcal{E}$ consists of edges, such that there is an edge between $\mathbf{p}_i$ and $\mathbf{p}_j$ if $\|\mathbf{p}_j - \mathbf{p}_i\|_2 \leq D_m$. Each node is also considered adjacent to itself, i.e., $(\mathbf{p}_i, \mathbf{p}_i) \in \mathcal{E}$.

Let $\mathcal{N}_i = \{\mathbf{p}_j : \|\mathbf{p}_j - \mathbf{p}_i\|_2 \leq D_m\}$ denote the set of neighbors of node $\mathbf{v}_i$ (including itself). The user trajectory is modeled as a sequence of positions constrained to the graph $\mathcal{G}$.

To construct the transition probabilities on the discrete graph, we evaluate the continuous-state transition density at each neighbor and normalize across all feasible neighbors. Specifically, the probability that the user moves from node $\mathbf{p}_n$ at time $t-2$ and node $\mathbf{p}_i$ at time $t-1$ to a neighbor node $\mathbf{p}_j \in \mathcal{N}_i$ at the next time t is given by

$$\mathbb{P}(\mathbf{p}_j | \mathbf{p}_i, \mathbf{p}_n) = \frac{a_{nij}}{\sum_{k \in \mathcal{N}_i} a_{nik}} \quad (4)$$

where the normalization $1/\sum_{k \in \mathcal{N}_i} a_{nik}$ guarantees that for any given $\mathbf{p}_i$ and $\mathbf{p}_n$, $\sum_{j \in \mathcal{N}_i} \mathbb{P}(\mathbf{p}_j | \mathbf{p}_i, \mathbf{p}_n) = 1$, the probability $\mathbb{P}(\mathbf{p}_j | \mathbf{p}_i, \mathbf{p}_n) = 0$ for $\mathbf{p}_j \notin \mathcal{N}_i$, and the factor

$$\begin{aligned} a_{nij} = & \frac{1}{2\pi(1-\gamma^2)\delta^2\sigma_v^2} \exp\left(-\frac{1}{2(1-\gamma^2)\delta^2\sigma_v^2}\right. \\ & \times (\mathbf{p}_j + \gamma\mathbf{p}_n - (1+\gamma)\mathbf{p}_i - (1-\gamma)\delta\bar{\mathbf{v}})^T \\ & \left. \times (\mathbf{p}_j + \gamma\mathbf{p}_n - (1+\gamma)\mathbf{p}_i - (1-\gamma)\delta\bar{\mathbf{v}})\right) \end{aligned}$$

is given from the Gauss-Markov model (3). Based on (4), one can evaluate the transition probability $\mathbb{P}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2})$.

C. A Bayesian Framework for Trajectory Inference

Denote $\mathcal{H}$ as a mapping from the MIMO-OFDM channel $\mathbf{H}_{t,q}$ to the observed radio signature $\mathbf{y}_{t,q} = \mathcal{H}(\mathbf{H}_{t,q})$ that will be exploited for trajectory inference. The development of a specific mapping $\mathcal{H}$ will be discussed in Section IV. Assuming that the observations $\mathbf{y}_{t,q}$ are independent across t and q , a Bayesian model that describes the evolution of the observation

$\mathbf{y}_t = \{\mathbf{y}_{t,q}\}_q$ as a function of the trajectory can be formulated as

$$p(\mathcal{Y}_t, \mathcal{X}_t) = p(\mathbf{y}_t | \mathbf{x}_t) \mathbb{P}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2}) p(\mathcal{Y}_{t-1}, \mathcal{X}_{t-1}) \quad (5)$$

$$= \prod_{t=1}^T \prod_{q=1}^Q p(\mathbf{y}_{t,q} | \mathbf{x}_t) \prod_{t=3}^T \mathbb{P}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2}), \quad (6)$$

where $\mathcal{X}_t = (\mathbf{x}_1, \dots, \mathbf{x}_t)$ and $\mathcal{Y}_t = (\mathbf{y}_1, \dots, \mathbf{y}_t)$ are the trajectory of the mobile user and the accumulated observations up to time t , respectively and (6) is obtained by repeatedly applying the chain rule in (5). While the mobility model $\mathbb{P}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2})$ is given in (4), the main challenge is to develop the conditional probability $p(\mathbf{y}_t | \mathbf{x}_t)$ that will be discussed in Section IV.

In the rest of the paper, we propose to extract the trajectory $\mathcal{X}_t$ by maximizing the parameterized likelihood (6) with a proper design of the radio signature mapping $\mathcal{H}$ that extracts spatial information from the MIMO-OFDM channels and a cost function that exploits the property of the propagation.

III. SPATIAL CONTINUITY AND TRAJECTORY IDENTIFIABILITY

The key challenge of trajectory inference without any location labels in an indoor environment is that the propagation is probably NLOS where the path may arrive at any angle, and in this case, the AOD may provide very little information on direction of the user. Moreover, the path amplitude also barely contains any information on the propagation distance due to the fading in indoor with rich scattering.

In this section, we try to understand some intuitive theoretical principle of blind trajectory inference by studying the properties of the channel and the trajectory estimation in simplified and special scenarios. First, for NLOS, we try to establish the spatial continuity of the channel, where under the same propagation environment as to be specified later, there exists a distance metric D such that the distance between the CSI is consistent with the physical distance d between the corresponding physical locations, i.e., $D(\mathbf{H}_1, \mathbf{H}_2) \propto d(\mathbf{x}_1, \mathbf{x}_2)$. Second, we try to establish a simple scenario with theoretical guarantee where user trajectories are identifiable, i.e., to perfectly recover a simple trajectory under infinite amount of independent measurements.

A. Quasi-Specular Environment Model

We first specify the environment model for mathematical tractability. Consider the 2D indoor environment to be surrounded by a finite number of *quasi-specular surfaces* and a finite number of *diffractive scatters*. A *quasi-specular surface* is defined as a patch that roughly holds a specular property, where the patch absorbs some energy of an incident wave, and reflects the wave with a majority energy towards the direction with the emergence angle equal to the incident angle and with a minor energy towards the directions dispersed around the major reflected path as shown in Figure 2(a). While this is also known as scattering in some literature [20, 21], our model limits the range of scattering around the major reflected path, whose emergence angle equal to the incident angle. As

as a result, when a receiver (RX) is located at a position on the major reflected path of a transmitter (TX), it may also receive a number of scattered paths arriving from a similar angle as shown in Figure 2(b). Based on the geometry, we can model the reflection with scattering due to this patch using a mirror TX located at the symmetric position of the true TX about the patch and a cluster of virtual TXs surrounding the mirror TX, as shown in Figure 2(c). Note that the scattering paths may arrive constructively or destructively at the RX, modeling the fading phenomenon. Therefore, the amplitude and phase of the received signal at the RX cannot be computed in a deterministic way, but the major propagation delay and angle can still be computed geometrically based on our model.

A *diffractive scatterer* models the diffraction phenomenon, where a radio propagation path may bend when passing over the edge of an obstacle as shown in Figure 2(d). Geometrically, one can place a virtual TX with the same distance away from the diffractive scatterer as that of the true TX such that the RX, diffractive scatterer, and the virtual TX are on the same line as shown in Figure 2(e). In our model, we assume diffraction exists only for a limited angle range, meaning a diffractive path can only bend for a certain angle. As a result, given a TX, each diffractive scatterer corresponds to a series of TXs located on an arc segment.

We only consider a finite number of reflections and diffractions. Note that each mirror TX and the associated cluster of virtual TXs can be reflected and diffracted again, resulting in double reflections, reflection-then-diffraction, triple reflections, and so on. Note that, the more reflections and diffractions, the further away a mirror TX from the original TX is created due to the geometry relation as shown in Figure 2(f), forming a *lattice* of mirror or virtual TX clusters. Thus, it is a natural assumption to consider that the clusters of virtual TXs do not overlap in a simple environment that has not so many patches and diffractive scatters. A typical example that likely matches with the quasi-specular environment model is a polygonal indoor office with several solid walls.

B. Spatial Continuity

The quasi-specular environment model inspires the construction of radio signature S_i for identifying position $\mathbf{p}_i$. Specifically, consider to place a TX at position $\mathbf{p}_i$. One can construct a lattice of mirror images of $\mathbf{p}_i$. Note that if one moves the TX, its images also move. Thus, in general, the lattice of mirror images differs from $\mathbf{p}_i$ to $\mathbf{p}_j$ unless there is perfect symmetry of the two distinct positions $\mathbf{p}_i$ and $\mathbf{p}_j$ in the environment, which is rare. As a result, one can select the L mirror images that are the nearest from the AP as the spatial signature. More precisely, a *conceptual* radio signature S_i for position $\mathbf{p}_i$ can be constructed as $S_i = \{(\tau_i^{(l)} - \tau_i^{(1)}, \theta_i^{(l)}), l = 1, 2, \dots, L\}$, where $\tau_i^{(l)} - \tau_i^{(1)}$ is the relative delay of the l th shortest path to the delay of the first arrival path and $\theta_i^{(l)}$ is the AoA of the l th shortest path at the RX. The reason that absolute delays are excluded in S_i because time synchronization is not assumed between TX and RX. In addition, the amplitude information is also excluded in S_i because the amplitude can be random due to

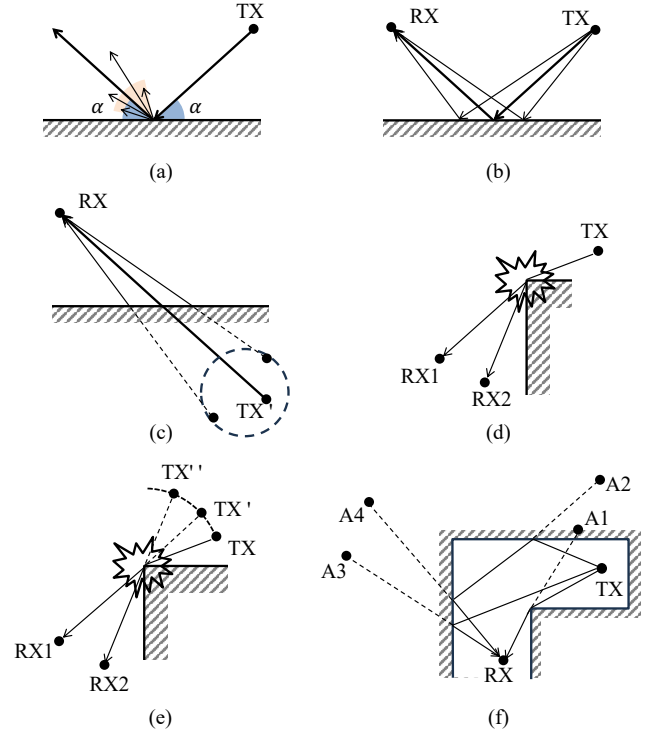


Figure 2. Illustration of quasi-specular environment model. (a) The scattered paths spread over a limited range of angles. (b) The RX receives a number of scattered paths arriving from a similar angle. (c) It is equivalent to view that the RX receives signals from a cluster of virtual TXs surrounding the mirror TX. (d) Diffraction. (e) A series of virtual TXs on an arc segment. (f) A lattice of mirror TXs A_1 , A_2 , A_3 , and A_4 .

various effects in practice, including fading, antenna pattern, and body shadowing.

The essential problem to investigate here is *whether the radio signature S is locally continuous in the physical space*. More precisely, can we find a distance metric on S , where the distance between S_i and S_j is consistent with the physical distance between the two corresponding positions $\mathbf{p}_i$ and $\mathbf{p}_j$, at least when the two positions are nearby? This problem is essential for radio map construction, because it asserts whether one can abstract the radio signature S in a continuous physical area to a *finite* number of *representations* $\{S_i\}$, such that for any continuous position $\mathbf{p}$, there is a discrete position $\mathbf{p}_i$ nearby whose radio signature S_i is close to the signature at position $\mathbf{p}$. In addition, if spatial continuity exists, it helps infer the latent physical topology from the CSI and enables a blind construction of radio map from CSI measurements without location labels.

We first specify the scenario where two nearby positions share the same propagation environment. Denote the AP position as point A . Following the quasi-specular environment model, a lattice of mirror positions $\mathcal{A} = \{A_1, A_2, \dots, A_L\}$ can be formed based on sequential specular operations, where we only consider a finite number of reflections and diffractions. Recall that, for each mirror position $A_l \in \mathcal{A}$, it represents a cluster of virtual mirror APs due to possible scattering and diffraction as discussed in Section III-A. For a position $\mathbf{p}_i$, only a subset $\mathcal{A}^i \subset \mathcal{A}$ of mirror images are visible from $\mathbf{p}_i$

because some paths may be blocked as shown in Figure 2. Two positions $\mathbf{p}_i$ and $\mathbf{p}_j$ are said to share the same propagation environment if the corresponding visible subsets are identical $\mathcal{A}^i = \mathcal{A}^j$.

We then formulate the problem starting from a single antenna case for the ease of elaboration. Consider two nearby positions $\mathbf{p}_1$ and $\mathbf{p}_2$ that share the same propagation environment, *i.e.*, $\mathcal{A}^1 = \mathcal{A}^2$. Following the model in (2), the m th subcarrier of the OFDM channel $\mathbf{h}_1$ at position $\mathbf{p}_1$ is given by

$$h_1^{(m)} = \sum_{l=1}^L \kappa_1^{(l)} e^{-jm \frac{2\pi}{M} B \tau_1^{(l)}} \quad (7)$$

where, following our quasi-specular environment model, $\kappa_1^{(l)}$ and $\tau_1^{(l)}$ are, respectively and conceptually, the complex path gain and delay for the path that associated with the l th mirror AP at $A_l \in \mathcal{A}^1$ in the visible subset. Likewise, we can express the channel $h_1^{(m)}$ for position $\mathbf{p}_2$ that share the same subset of visible mirror images $\mathcal{A}^2 = \mathcal{A}^1$.

Define

$$R(u) = \frac{1}{M} \sum_{m=0}^{M-1} \mathbb{E}\{h_1^{(m)} h_2^{(m)*}\} e^{jm \frac{2\pi}{M} u} \quad (8)$$

where the expectation is taken over the randomness of the complex path gain $\kappa_1^{(l)}$, $\kappa_2^{(l)}$, and direction θ' from $\mathbf{p}_1$ to $\mathbf{p}_2$. It follows that $R(u)$ is the correlation function in the delay domain, because $h_i^{(m)}$ represent the frequency domain channel, the product $h_1^{(m)} h_2^{(m)*}$ computes the correlation, and the sum with the exponential term is just the Inverse Discrete Fourier Transform (IDFT) formula, *i.e.*, transforming the correlation to the time domain.

Consider the following theoretical metric

$$\hat{u}(\mathbf{h}_1, \mathbf{h}_2) = \arg \max_u |R(u)|. \quad (9)$$

Denote $d = \|\mathbf{p}_1 - \mathbf{p}_2\|_2$ as the physical distance between the two positions, which have been assumed to share the same propagation environment. Suppose that the complex path gains $\kappa_i^{(l)}$ are zero-mean and uncorrelated among l , but there is correlation among $i = 1, 2$, *i.e.*, $\mathbb{E}\{\kappa_1^{(l)} \kappa_2^{(l')*}\} = 0$ for $l \neq l'$, and $\mathbb{E}\{\kappa_1^{(l)} \kappa_2^{(l)*}\} = C(d)$ for all l . Typically, one may expect that the correlation $C(d)$ decreases in distance d . Then, we have the following result.

Theorem 1. (*Spatial Continuity*) Under the multipath model (7) and a large M approximation, we have $\hat{u}(\mathbf{h}_1, \mathbf{h}_2) \approx \frac{B}{c} d$.

Proof. See Appendix A. $\square$

Theorem 1 confirms that within the same propagation environment, there exists a metric where the CSI distance is consistent with the physical distance of the corresponding positions. In a special case of using metric (9), the two distances are related by a factor of B/c .

The implication of Theorem 1 is two-fold: First, one can approximate the CSI sampled randomly in a neighborhood using the CSI sampled from a deterministic representative location in a set $\mathcal{V}$. While this result may sound intuitive, Theorem 1 finds a metric to quantify the CSI approximation

error in terms of distance d away from the representative location. Second, Theorem 1 provides a perspective to convert the trajectory inference problem to a multidimensional scaling (MDS) problem [22] that embeds high-dimensional CSI observation $\mathbf{y}_t$ onto a 2D space while preserving the pairwise distances under a certain scaling. Such a philosophy will be exploited for algorithm development in Section IV.

Remark 1 (A Practical CSI Distance Metric). The metric $R(u)$ defined in (8) is a theoretical one because it needs to compute the expectation. A practical metric can be defined as

$$\hat{R}(u) = \frac{1}{M} \sum_{m=0}^{M-1} h_1^{(m)} h_2^{(m)*} e^{jm \frac{2\pi}{M} u}. \quad (10)$$

The two metrics (8) and (10) are related under a rich scattering environment where there are a large number of independent paths. Then, the law of large numbers comes into play to connect the expectation with the sum of paths (See (28) in Appendix A). Therefore, Theorem 1 is only meaningful in NLOS case.

Remark 2 (Multiple Antenna Extension). A rigorous metric that exploits the angular domain information for MIMO-OFDM channel with analytical justification similar to (9) remains unknown. The challenge is that the angular discrepancy is inversely proportional to the total propagation distance, which is relatively large compared to d . By contrast, the frequency domain discrepancy that leads to (9) does not depend on the absolute propagation distance. A naive extension to the multiple antenna case without exploiting the angular domain discrepancy is to apply (8) to each antenna and yield $R_n(u)$ for the n th antenna. Then, defining $R(u) = \frac{1}{N_t} \sum_{n=1}^{N_t} R_n(u)$, and the same metric $\hat{u}(\mathbf{H}_1, \mathbf{H}_2)$ in (9) follows and the same spatial continuity result applies.

C. Trajectory Identifiability

We now examine the problem of trajectory inference from another perspective. Suppose that the AOD $\theta_{t,q}$ of the dominant path from AP q to the user at location $\mathbf{x}_t$ follows a Gaussian distribution

$$\theta_{t,q} \sim \mathcal{N}(\phi(\mathbf{x}_t, \mathbf{o}_q), \sigma_\theta^2) \quad (11)$$

where $\phi(\mathbf{x}_t, \mathbf{o}_q)$ denotes the geometric azimuth angle between the user location $\mathbf{x}_t$ and the AP position $\mathbf{o}_q$. In an LOS scenario, σ_θ^2 tends to be small, whereas, for NLOS, σ_θ^2 can be large. While this model might be a bit artificial for an indoor case, the problem we investigate is highly non-trivial: *for an arbitrarily large angular variance σ_θ^2 , is it theoretically possible to recover a rectilinear trajectory using a sufficient number of measurements?*

The prior work [23] attempted a similar problem in an outdoor case exploiting the power law that relates the distance information with the signal strength. However, in the indoor case, we do *not* assume a power law of the signal strength due to the more complicated propagation. Instead, we only rely on the noisy and unbiased AOD information from (11) to recover the trajectory, which leads to non-straight-forward answers.

Specifically, consider the Gauss-Markov mobility model in (3) in the continuous space with $\gamma = 1$, which degenerates to a constant speed rectilinear mobility

$$\mathbf{x}_t = \mathbf{x} + t\mathbf{v} \quad (12)$$

where we only need to estimate the start position and the speed $(\mathbf{x}, \mathbf{v}) \in \mathbb{R}^4$ in a continuous space. Consequently, the distance $d(\mathbf{x}_t, \mathbf{o}_q)$ between the mobile location $\mathbf{x}_t$ and the q th AP location $\mathbf{o}_q$ is simplified as $d_{t,q}(\mathbf{x}, \mathbf{v}) \triangleq \|\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}\|_2$, where $\mathbf{l}_q(\mathbf{x}) = \mathbf{x} - \mathbf{o}_q$ is the direction from the q th AP to the initial position $\mathbf{x}$ of the trajectory.

Thus, the log-likelihood function $\log p(\mathcal{Y}_T, \mathcal{X}_T)$ in (6) simplifies to

$$f(\psi) = \sum_{t=1}^T \sum_{q=1}^Q \left[-\ln(2\pi\sigma_\theta^2) - \frac{1}{2\sigma_\theta^2} (\theta_{t,q} - \phi(\mathbf{x} + t\mathbf{v}, \mathbf{o}_q))^2 \right] \quad (13)$$

where $\psi = (\mathbf{x}, \mathbf{v})$ represents the mobility parameters, and the term $\log \mathbb{P}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2})$ vanishes under the constant-speed mobility model.

The focus here is to understand the fundamental limit of estimating the mobility parameter $\psi = (\mathbf{x}, \mathbf{v})$ using only AOD measurements.

1) *The Fisher Information Matrix:* The Fisher information matrix (FIM) $\mathbf{F}_{T,\psi}$ of $\psi = (\mathbf{x}, \mathbf{v}) \in \mathbb{R}^4$ from the measurements over a duration T can be computed as

$$\begin{aligned} \mathbf{F}_{T,\psi} &\triangleq \mathbb{E}\{-\nabla_{\psi}^2 f(\psi)\} \\ &= \sum_{t,q} \frac{1}{\sigma_\theta^2} \nabla_{\psi} \phi(\mathbf{x} + t\mathbf{v}, \mathbf{o}_q) (\nabla_{\psi} \phi(\mathbf{x} + t\mathbf{v}, \mathbf{o}_q))^T. \end{aligned}$$

The derivative $\nabla_{\psi} \phi(\mathbf{x} + t\mathbf{v}, \mathbf{o}_q)$ is derived as

$$\begin{aligned} \nabla_{\psi} \phi(\mathbf{x} + t\mathbf{v}, \mathbf{o}_q) &= \frac{1}{d_{t,q}^2(\mathbf{x}, \mathbf{v})} \begin{bmatrix} 1 \\ t \end{bmatrix} \\ &\quad \otimes \left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} (\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}) \right). \end{aligned}$$

Thus, the FIM can be expressed as

$$\begin{aligned} \mathbf{F}_{T,\psi} &= \sum_{t,q} \frac{1}{\sigma_\theta^2 d_{t,q}^4(\mathbf{x}, \mathbf{v})} \begin{bmatrix} 1 & t \\ t & t^2 \end{bmatrix} \\ &\quad \otimes ((\|\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}\|^2 \mathbf{I} - (\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})(\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})^T) \end{aligned} \quad (14)$$

in which, $\otimes$ is the Kronecker product.

For an unbiased estimator $\hat{\psi}$, the mean squared error (MSE) is lower bounded by $\mathbb{E}\{\|\hat{\psi} - \psi\|^2\} \geq \text{tr}\{\mathbf{F}_{T,\psi}^{-1}\}$, where $\text{tr}\{\mathbf{F}_{T,\psi}^{-1}\}$ is the CRLB of estimating $\psi = (\mathbf{x}, \mathbf{v})$. Similarly, we define the FIMs $\mathbf{F}_{T,x} = \mathbb{E}\{-\nabla_{\mathbf{x}\mathbf{x}}^2 f(\psi)\}$ and $\mathbf{F}_{T,v} = \mathbb{E}\{-\nabla_{\mathbf{v}\mathbf{v}}^2 f(\psi)\}$, which are the diagonal blocks of $\mathbf{F}_{T,\psi}$ and are associated with the CRLB $B(\mathbf{x}) = \text{tr}\{\mathbf{F}_{T,x}^{-1}\}$ and CRLB $B(\mathbf{v}) = \text{tr}\{\mathbf{F}_{T,v}^{-1}\}$ for the parameters $\mathbf{x}$ and $\mathbf{v}$, respectively.

2) *AP Deployed in a Limited Region:* We investigate the case where the APs are deployed in a limited region, but the measurement trajectory is allowed to go unbounded as T goes to infinity. Signals can always be collected by the APs regardless of the distance. As a result, an infinite amount of measurements can be collected as $T \rightarrow \infty$.

It is observed that $\mathbf{F}_{T,\psi} \prec \mathbf{F}_{T+1,\psi}$, indicating that the Fisher information is strictly increasing from Lemma 4 in Appendix B, provided that $\mathbf{l}_q(\mathbf{x})$ and $\mathbf{v}$ are linear independent for at least one q .

However, it is somewhat surprising that the CRLB for $\mathbf{x}$ and $\mathbf{v}$ does not decrease to zero as $T \rightarrow \infty$, despite the infinitely increasing amount of independent data.

Specifically, assume that the trajectory $\mathbf{x}_t$ does not pass any of the AP location $\mathbf{o}_q$, and hence, $d_{\min} = \min_{t,q} \{d_{t,q}(\mathbf{x}, \mathbf{v})\} > 0$ for all t, q .

Proposition 1. *The CRLB of $\mathbf{x}$ satisfies $B(\mathbf{x}) = \text{tr}\{\mathbf{F}_{T,x}^{-1}\} \geq \bar{\Delta}_{T,x}$ with equality achieved when $d_{t,q} = d_{\min}$ and $\phi_{t,q} = 0$ for all t, q . In addition, $\bar{\Delta}_{T,x}$ is strictly decreasing in T , provided that at least two vectors in $\{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_Q, \mathbf{v}\}$ are linear independent, but $\bar{\Delta}_{T,x}$ converges to a strictly positive number as $T \rightarrow \infty$.*

Proof. See Appendix B. $\square$

Proposition 1 suggests that the CRLB of $\mathbf{x}$ cannot decrease to zero even when we estimate only two parameters for the initial location $\mathbf{x} \in \mathbb{R}^2$ based on *infinite* AOD measurements collected over an infinite geographical horizon as $T \rightarrow \infty$.

Through the development of the proof, a physical interpretation of Proposition 1 can be given as follows. As T increases, the distances $d_{t,q}(\mathbf{x}, \mathbf{v}) = \|\mathbf{x}_t - \mathbf{o}_q\|_2$ grow larger because the user moves away from the APs. For a position $\mathbf{x}_t$ at a sufficiently large distance, the term $\mathbf{x}_t - \mathbf{o}_q$ approximates $\mathbf{x}_t$ since $\|\mathbf{x}_t\| \gg \|\mathbf{o}_q\|$. Consequently, the angle measurement $\phi(\mathbf{x}_t, \mathbf{o}_q)$ changes very little as $t \rightarrow \infty$, making successive AOD observations almost indistinguishable regardless of the AP locations $\mathbf{o}_q$. Although the Fisher information matrix $\mathbf{F}_{T,x}$ increases with T , the increment $\mathbf{F}_{T+1,x} - \mathbf{F}_{T,x}$ decays rapidly. Hence, while the CRLB decreases monotonically, it converges to a strictly positive lower bound.

We obtain a similar conclusion for estimating the velocity variable $\mathbf{v}$.

Proposition 2. *The CRLB of $\mathbf{v}$ satisfies $B(\mathbf{v}) = \text{tr}\{\mathbf{F}_{T,v}^{-1}\} \geq \bar{\Delta}_{T,v}$, with equality achieved when $d_{t,q} = d_{\min}$ and $\phi_{t,q} = 0$ for all t, q . In addition,*

$$\bar{\Delta}_{T,v} \rightarrow C_v = \frac{d_{\min}^2 \sigma_n^2}{G_1 N_t (N_t^2 - 1)} \left(\sum_{q=1}^Q s_{\infty,q}^{(2)} \|\mathbf{P}_v^\perp \mathbf{l}_q(\mathbf{x})\|^2 \right)^{-1}$$

as $T \rightarrow \infty$, where G_1 depends on the antenna configuration, σ_n^2 is the signal noise, and

$$s_{\infty,q}^{(2)} = \lim_{T \rightarrow \infty} \sum_{t=1}^T \frac{t^2}{d_{t,q}^4(\mathbf{x}, \mathbf{v})}, \quad \mathbf{P}_v^\perp = \mathbf{I} - \mathbf{v}\mathbf{v}^T / \|\mathbf{v}\|^2$$

in which, the parameter $s_{\infty,q}^{(2)}$ is upper bounded by $1/\rho^4 \lim_{T \rightarrow \infty} \sum_{t=1}^T 1/t^2 \approx \frac{\pi^2}{6\rho^4}$, where $\rho > 0$ is sufficiently small such that $d_{t,q}(\mathbf{x}, \mathbf{v}) > \rho t$ for all $t \geq 1$.

Proof. See Appendix C. $\square$

Propositions 1 and 2 suggest that, under a finite number of APs in a limited region, one cannot perfectly recover a trajectory even for a simple constant speed rectilinear mobility under infinite measurements.

Proposition 2 quantifies the fundamental limit to the estimation accuracy that is affected by the spatial distribution of the APs and the nature of the AOD measurements. Specifically, the term $\|\mathbf{P}_v^\perp \mathbf{l}_q(\mathbf{x})\|$ is the length of the q th AP position vector component orthogonal to the velocity vector $\mathbf{v}$. If all APs lie nearly collinear with $\mathbf{v}$, then $\mathbf{P}_v^\perp \mathbf{l}_q \approx \mathbf{0}$ and the sum $\sum_q \|\mathbf{P}_v^\perp \mathbf{l}_q\|^2$ is small, providing little information in directions orthogonal to motion. By contrast, a wider angular spread of APs (APs surrounding the agent or spanning different angles relative to $\mathbf{v}$) increases $\sum_q \|\mathbf{P}_v^\perp \mathbf{l}_q\|^2$, improving geometric conditioning and reducing the CRLB. The equality case $\phi_{t,q} = 0$ corresponds to each AP being *broadside* to $\mathbf{v}$, which maximizes information gain. Limited region deployments exhibit geometric dilution of precision, causing a nonzero lower bound on estimation error. Moreover, the lower bound C_v decreases with higher effective signal-to-noise ratio (SNR) (G_1/σ_n^2) and scales as $\mathcal{O}(1/N_t^3)$ with increasing N_t , as larger ULA apertures improve angular resolution and array gain—practically reducing the CRLB by a factor of eight when doubling the number of transmit antennas.

3) *AP Deployed in an Unlimited Region:* We now consider a theoretical scenario in which the APs are distributed over an unbounded region according to a homogeneous Poisson Point Process (PPP) with density κ . Despite the infinite spatial domain, the user is restricted to connecting only with those APs located within a fixed connectivity radius R . Consequently, the number of connected APs at any time remains finite, with the average number per time slot given by $\bar{Q} = \kappa\pi R^2$. We are interested in the asymptotic behaviour of the CRLBs, as $T \rightarrow \infty$. This is to understand the error decrease rate, i.e., how fast the error may decrease as we increase the number of observations T .

Remarkably, even though the set of active APs at any instant is always limited, the estimation lower bound for user state parameters vanishes in the limit of long observation duration.

Theorem 2. Assume that the minimum distance to the nearest AP is greater than r_0 along the trajectory.² The CRLB of $\mathbf{x}$ satisfies $B(\mathbf{x}) = \text{tr}\{\mathbf{F}_{T,x}^{-1}\} \leq \tilde{\Delta}_{T,x}$ and as $T \rightarrow \infty$

$$T\tilde{\Delta}_{T,x} \rightarrow \frac{16\sigma_n^2}{\kappa\pi(r_0^{-2} - R^{-2})G_1N_t(N_t^2 - 1)}.$$

The CRLB of $\mathbf{v}$ satisfies $B(\mathbf{v}) = \text{tr}\{\mathbf{F}_{T,v}^{-1}\} \leq \tilde{\Delta}_{T,v}$ and as $T \rightarrow \infty$

$$T(T+1)(2T+1)\tilde{\Delta}_{T,v} \rightarrow \frac{96\sigma_n^2}{\kappa\pi(r_0^{-2} - R^{-2})G_1N_t(N_t^2 - 1)}.$$

Proof. See Appendix D. $\square$

²In practice, the parameter r_0 can be understood as the height of the antenna. More rigorously, we should employ a 3D model to compute the distance $d_{t,q}$, but the asymptotic result would be the same.

From the above theorem, it is evident that the CRLB for the initial position $\mathbf{x}$ decays as $\mathcal{O}(1/T)$, while the CRLB for the velocity $\mathbf{v}$ decreases at the faster rate of $\mathcal{O}(1/T^3)$. This demonstrates that velocity estimation becomes asymptotically much more accurate than position estimation as the observation window grows. Furthermore, both a larger connectivity radius R and a higher AP density κ within the coverage region enhance the achievable estimation accuracy. Additionally, $\tilde{\Delta}_{T,v}$ exhibits a scaling relationship with the number of antennas N_t as $\mathcal{O}(1/N_t^3)$, indicating that increasing the antenna count at each AP provides substantial improvements in velocity estimation precision.

It is worth highlighting that the results do *not* assume LOS or NLOS conditions. Specifically, regardless of a possibly large angular variance σ_θ^2 in NLOS, we can achieve arbitrarily high accuracy in estimating the parameters of the trajectory given a sufficient number of independent measurements. Moreover, a larger G_1/σ_n^2 , which results from richer array geometries or higher SNR, accelerates the rate at which the CRLB for both $\mathbf{x}$ and $\mathbf{v}$ decreases as T increases.

IV. ALGORITHM DESIGN

The previous section delivers two important messages: First, from Theorem 2, accumulating enough AOD measurements can recover the trajectory even when the AOD is highly noisy under NLOS. While the result was developed under an artificial model for a simplified mobility, we can expect to recover at least a partial trajectory given sufficient data. Second, from Theorem 1, the pairwise CSI distance is consistent with the physical distance of the corresponding positions in a small neighborhood under NLOS. As a result, we can regulate the estimated trajectory by relating the CSI distance with the physical distance, hoping to enhance the estimation in the NLOS area. As indicated by Theorem 1, we also need to identify the propagation conditions and only pair measurements in the same propagation condition. Finally, we can exploit the mobility model (4) and the Bayesian formulation (6) so that the estimated trajectory in the NLOS area also benefit from the information collected in the LOS area via the Bayesian chain rule.

In this section, we develop an algorithm framework to incorporate all these design philosophy using Bayesian approaches.

A. Feature Engineering

1) *RSS Feature for LOS/NLOS Discrimination:* While we try not to rely on received signal strength (RSS) for location signature, RSS is still a good indicator for LOS and NLOS discrimination especially around the boundary of the two propagation regions. The RSS $s_{t,q}$ at the q th AP at time slot t is simply extracted as the channel power in logarithm scale:

$$s_{t,q} = 10 \log_{10} \|\mathbf{H}_{t,q}\|_F^2. \quad (15)$$

The RSS measurements are to be fitted to a conditional path loss model that accounts for location and AP dependent propagation conditions:

$$s_{t,q} = \beta_q^{(k)} - \alpha_q^{(k)} \log_{10} d(\mathbf{x}_t, \mathbf{o}_q) + \xi_q^{(k)}, \quad k \in \{0, 1\} \quad (16)$$

where k indicates the propagation condition, with $k = 0$ for LOS and $k = 1$ for NLOS. Here, $\beta_q^{(k)}$ represents the AP-dependent reference path loss, $\alpha_q^{(k)}$ is the AP-dependent path loss exponent; and $d(\mathbf{x}_t, \mathbf{o}_q) = \|\mathbf{o}_q - \mathbf{x}_t\|_2$ denotes the Euclidean distance between the AP at $\mathbf{o}_q$ and the mobile user at $\mathbf{x}_t$. The term $\xi_q^{(k)} \sim \mathcal{N}(0, \sigma_{s,q,k}^2)$ is to model the randomness due to multipath fading, body shadowing, and antenna pattern. The parameters $\alpha_q^{(k)}$, $\beta_q^{(k)}$, $\sigma_{s,q,k}^2$ are to be jointly estimated from the data.

2) *AoD Feature*: We extract the AOD of the *dominant* path from the MIMO-OFDM channel $\mathbf{H}_{t,q}$ using a subspace approach just as the MUSIC algorithm [24]. Specifically, perform eigen-decomposition of the sample covariance matrix

$$\mathbf{R}_{t,q} = \frac{1}{M} \mathbf{H}_{t,q} \mathbf{H}_{t,q}^H$$

and obtain the eigenvectors $\tilde{\mathbf{u}}_{t,q}^{(1)}, \tilde{\mathbf{u}}_{t,q}^{(2)}, \dots, \tilde{\mathbf{u}}_{t,q}^{(N_t)}$ arranged in a decreasing order of the corresponding eigenvalues. Construct the noise subspace matrix

$$\mathbf{U}_{t,q} = [\tilde{\mathbf{u}}_{t,q}^{(2)}, \tilde{\mathbf{u}}_{t,q}^{(3)}, \dots, \tilde{\mathbf{u}}_{t,q}^{(N_t)}] \in \mathbb{C}^{N_t \times (N_t - 1)}$$

by skipping the dominant eigenvector $\tilde{\mathbf{u}}_{t,q}^{(1)}$.

Using the same principle of MUSIC algorithm, the relative AOD $\varphi_{t,q}$ of the dominant path for position $\mathbf{x}_t$ w.r.t. the reference direction ϕ_q of the q th AP can be obtained by maximizing the following pseudo-spectrum

$$\hat{\varphi}_{t,q} = \underset{\varphi \in (-\pi/2, \pi/2)}{\operatorname{argmax}} \frac{1}{\mathbf{a}^H(\varphi) \mathbf{U}_{t,q} \mathbf{U}_{t,q}^H \mathbf{a}(\varphi)}$$

where $\mathbf{a}(\varphi)$ is the steering vector given by (1). Thus, the AOD in the XOY coordinate system of the dominant path from the AP at $\mathbf{o}_q$ to position $\mathbf{x}_t$ is estimated as

$$\hat{\theta}_{t,q} = (\hat{\varphi}_{t,q} + \phi_q) \bmod 2\pi. \quad (17)$$

The estimated AOD $\hat{\theta}_{t,q}$ is to be fitted to a conditional Gaussian model

$$\hat{\theta}_{t,q} \sim \mathcal{N}(\phi(\mathbf{x}_t, \mathbf{o}_q), \sigma_{\theta,k}^2), \quad k \in \{0, 1\} \quad (18)$$

where $\phi(\mathbf{x}_t, \mathbf{o}_q)$ defines the geometric azimuth angle from the AP location $\mathbf{o}_q$ to the position $\mathbf{x}_t$. Likewise, $k = 0$ stands for the LOS condition and $k = 1$ stands for the NLOS condition, and the parameters $\sigma_{\theta,k}^2$ are to be jointly fitted from the data. Experimental results in Figure 3 (b) verify that the variance in the LOS region is significantly smaller than the variance in the NLOS region.

3) *Delay Spread Feature for LOS/NLOS Discrimination*: From the multipath channel model (2), if there is only one dominant path, the magnitude of the frequency domain channel is roughly constant, because the Fourier transform of a delta function is constant. By contrast, if there are several significant multipath components, the magnitude of the frequency domain channel fluctuates. We thus construct a feature $\nu_{t,q}$ as follows

$$\nu_{t,q} = 10 \log_{10} \left(\operatorname{Var} \left(\left| \frac{\mathbf{H}_{t,q}}{\|\mathbf{H}_{t,q}\|_2} \right| \right) \right) \quad (19)$$

to empirically capture how the channel energy spreads over the multipaths, where $|\cdot|$ denotes the element-wise absolute

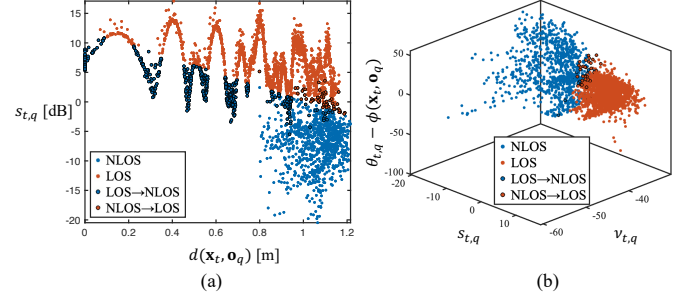


Figure 3. (a) With known user location, the RSS (path-loss) model is fitted to perform LOS/NLOS identification, yielding an error rate of 20.5%. (b) By further leveraging the relationship among relative delay spread, AOD, and RSS, the proposed method reduces the identification error to 1.6%.

value operation and $\operatorname{Var}(\cdot)$ denotes variance of the elements of the matrix or vector.

The feature pair $(s_{t,q}, \nu_{t,q})$ can be modeled using the following Gaussian mixture model:

$$[s_{t,q}, \nu_{t,q}]^T \sim \mathcal{N}(\mathbf{m}_k, \mathbf{\Upsilon}_k), \quad k \in \{0, 1\} \quad (20)$$

where $k = 0$ corresponds to LOS case and $k = 1$ to the NLOS case. The parameters $\mathbf{m}_k$ and $\mathbf{\Upsilon}_k$ are jointly estimated from empirical data. This Gaussian mixture model is motivated by experimental observations, as illustrated in Figure 3(b). Compared to relying solely on RSS for LOS/NLOS identification as shown in Figure 3(a), the proposed feature demonstrates a much better clustering behavior that may simplify the LOS/NLOS discrimination.

4) *CSI Distance Feature in NLOS Condition*: Define the CSI distance of two channels $\mathbf{H}_{i,q}$ and $\mathbf{H}_{j,q}$ based on (10) and Remarks 1 and 2 as

$$\hat{u}(\mathbf{H}_{i,q}, \mathbf{H}_{j,q}) = \underset{u \in \{0, 1, \dots, M-1\}}{\operatorname{argmax}} \left| \frac{1}{N_t M} \sum_{n=1}^{N_t} \sum_{m=0}^{M-1} \left[h_{i,q}^{(m,n)} \times \left(h_{j,q}^{(m,n)} \right)^* \cdot e^{j \frac{2\pi m u}{M}} \right] \right|$$

where $h_{t,q}^{(m,n)}$ is the channel of the m th subcarrier at the n th antenna.

Denote $u_{t,q}^{(k)} \in \{0, 1\}$ to indicate whether $\mathbf{H}_{t,q}$ is LOS, i.e., $u_{t,q}^{(0)} = 1$ or NLOS, i.e., $u_{t,q}^{(1)} = 1$. Note that $u_{t,q}^{(0)} + u_{t,q}^{(1)} \equiv 1$, and either $u_{t,q}^{(k)} = u_{\tau,q}^{(k)}$ for both $k = 1, 2$ or $u_{t,q}^{(k)} \neq u_{\tau,q}^{(k)}$ for both $k = 1, 2$.

As an implication of Theorem 1, the feature $\hat{u}(\mathbf{H}_{i,q}, \mathbf{H}_{j,q})$ is to be fitted to the following regression model for every pair of measurements i, j such that they share the same NLOS propagation condition $u_{i,q}^{(1)} = u_{j,q}^{(1)} = 1$,

$$\hat{u}(\mathbf{H}_{i,q}, \mathbf{H}_{j,q}) = \frac{B}{c} d(\mathbf{x}_i, \mathbf{x}_j) + \mathcal{N}(0, \sigma_u^2), \quad k \in \{0, 1\} \quad (21)$$

where the variance σ_u^2 quantifies the uncertainty. The model (21) describes the phenomenon that the larger the physical distance d , the larger the CSI distance $\hat{u}$, following the scaling $\hat{u} = \frac{B}{c} d$ inspired from Theorem 1.

The model (21) was found to be consistent with the data from the empirical studies in NLOS condition as shown in Figure 6(a).

B. A Bayesian Formulation

1) *Gaussian Mixture Model*: Based on the feature derived in (15), (17), and (19), the observed radio signature can be constructed as

$$\mathbf{y}_{t,q} = \mathcal{H}(\mathbf{H}_{t,q}) = [s_{t,q}, \hat{\theta}_{t,q}, s_{t,q}, \nu_{t,q}]^T.$$

According to the conditional models (16), (18), and (20), the observed radio signature vector $\mathbf{y}_{t,q}$ follows a Gaussian mixture distribution

$$p(\mathbf{y}_{t,q}|\mathbf{x}_t) = \sum_{k=0}^1 \pi_{t,q}^{(k)} \mathcal{N}(\boldsymbol{\mu}^{(k)}(\mathbf{x}_t), \boldsymbol{\Sigma}^{(k)}) \quad (22)$$

where $\pi_{t,q}^{(k)} = \mathbb{P}\{u_{t,q}^{(k)} = 1\}$ is the probability that position $\mathbf{x}_t$ is at propagation condition k for the q th AP, the mean vector is given by

$$\boldsymbol{\mu}^{(k)}(\mathbf{x}_t) = \begin{bmatrix} \beta_q^{(k)} + \alpha_q^{(k)} \log_{10} d(\mathbf{o}_q, \mathbf{x}_t) \\ \phi(\mathbf{x}_t, \mathbf{o}_q) \\ \mathbf{m}_k \end{bmatrix}$$

and the covariance matrix

$$\boldsymbol{\Sigma}^{(k)} = \begin{bmatrix} \sigma_{s,q,k}^2 & 0 & 0 \\ 0 & \sigma_{\theta,k}^2 & 0 \\ 0 & 0 & \boldsymbol{\Upsilon}_k \end{bmatrix}.$$

For the CSI distance model in (21), considering that the spatial continuity is only valid in a small neighborhood and under NLOS conditions $k = 1$. For each AP q , if the location $\mathbf{x}_t$ is in the NLOS region of AP q , the NLOS neighborhood of $\mathbf{x}_t$, with respect to all other position variables in $\mathcal{X}_T$, is defined as

$$\tilde{\mathcal{N}}_q(\mathbf{x}_t) = \left\{ \tau : d(\mathbf{x}_t, \mathbf{x}_\tau) < \tilde{\delta}, u_{t,q}^{(1)} = 1, u_{\tau,q}^{(1)} = 1 \right\},$$

where $\tilde{\delta}$ is a predefined distance threshold. Define $\mathcal{H}_t$ as a collection of the channel $\mathbf{H}_{t,q}$ up to time t for all APs q . We can construct a Gaussian log-likelihood function as

$$f(\mathbf{H}_{t,q}|\mathcal{H}_T, \mathcal{X}_T) = \frac{1}{|\tilde{\mathcal{N}}_q(\mathbf{x}_t)|} \sum_{\tau \in \tilde{\mathcal{N}}_q(\mathbf{x}_t)} \log \left[\frac{1}{\sqrt{2\pi\sigma_u^2}} \right] \quad (23)$$

$$\times \exp \left(- \frac{[\hat{u}(\mathbf{H}_{t,q}, \mathbf{H}_{\tau,q}) - \frac{B}{c} d(\mathbf{x}_t, \mathbf{x}_\tau)]^2}{2\sigma_u^2} \right),$$

where $|\tilde{\mathcal{N}}_q(\mathbf{x}_t)|$ denotes the cardinality of the set $\tilde{\mathcal{N}}_q(\mathbf{x}_t)$. Moreover, $f(\mathbf{H}_{t,q}|\mathcal{H}_T, \mathcal{X}_T)$ is set to zero if $\mathbf{x}_t$ is located in the LOS region of AP q .

2) *Spatially Regularized Likelihood Formulation*: Recall the joint probability $p(\mathcal{Y}_T, \mathcal{X}_T)$ in (6). Denote $\boldsymbol{\Theta}_m$ as the collection of all the parameters in the mobility model (4) and $\boldsymbol{\Theta}_p$ as the collection of all the remaining parameters which are related to the radio signature, we construct a regularized likelihood as follows

$$\mathcal{L}(\mathcal{X}_T, \boldsymbol{\Theta}_p, \boldsymbol{\Theta}_m) = \log p(\mathcal{Y}_T, \mathcal{X}_T) + \eta \sum_{t=1}^T \sum_{q=1}^Q f(\mathbf{H}_{t,q}|\mathcal{H}_T, \mathcal{X}_T)$$

where $\eta > 0$ is some weighting factor and the last term applies regularization only to locations under the NLOS condition.

The trajectory inference problem can be formulated as

$$\begin{aligned} & \underset{\mathcal{X}_T, \boldsymbol{\Theta}_p, \boldsymbol{\Theta}_m}{\text{maximize}} && \mathcal{L}(\mathcal{X}_T, \boldsymbol{\Theta}_p, \boldsymbol{\Theta}_m) \\ & \text{subject to} && u_{t,q}^{(k)} \in \{0, 1\}, u_{t,q}^{(0)} + u_{t,q}^{(1)} = 1 \\ & && \mathbf{x}_t \in \mathcal{V}, \quad t = 1, 2, \dots, T \\ & && (\mathbf{x}_t, \mathbf{x}_{t-1}) \in \mathcal{E}, \quad t = 2, \dots, T. \end{aligned} \quad (24)$$

C. Algorithm Design

To solve the joint trajectory inference and parameter estimation problem in equation (24), we observe that, given $\mathcal{X}_T$, the variables $\boldsymbol{\Theta}_p$ and $\boldsymbol{\Theta}_m$ are decoupled. This is because the term $p(\mathbf{y}_{t,q}|\mathbf{x}_t)$ in equation (24) only depends on $\boldsymbol{\Theta}_p$, while the term $\mathbb{P}(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_{t-2})$ only depends on $\boldsymbol{\Theta}_m$. Consequently, $\boldsymbol{\Theta}_p$ and $\boldsymbol{\Theta}_m$ can be solved through two parallel subproblems derived from equation (24), as follows:

$$\begin{aligned} \text{(P1)} : & \underset{\boldsymbol{\Theta}_m}{\text{maximize}} && \sum_{t=3}^T \log \mathbb{P}(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_{t-2}; \boldsymbol{\Theta}_m) \\ \text{(P2)} : & \underset{\boldsymbol{\Theta}_p}{\text{maximize}} && \sum_{t=1}^T \sum_{q=1}^Q \log \sum_{k=0}^1 \pi_{t,q}^{(k)} \mathcal{N}(\boldsymbol{\mu}^{(k)}(\mathbf{x}_t), \boldsymbol{\Sigma}^{(k)}) \\ & && + \eta \sum_{t=1}^T \sum_{q=1}^Q f(\mathbf{H}_{t,q}|\mathcal{H}_T, \mathcal{X}_T) \\ & \text{subject to} && u_{t,q}^{(k)} \in \{0, 1\}, u_{t,q}^{(0)} + u_{t,q}^{(1)} = 1 \end{aligned}$$

On the other hand, given the variables $\hat{\boldsymbol{\Theta}}_p$ and $\hat{\boldsymbol{\Theta}}_m$ as the solutions to (P1) and (P2), respectively, the trajectory $\mathcal{X}_T$ can be solved by:

$$\begin{aligned} \text{(P3)} : & \underset{\mathcal{X}_T}{\text{maximize}} && \mathcal{L}(\mathcal{X}_T, \hat{\boldsymbol{\Theta}}_p, \hat{\boldsymbol{\Theta}}_m) \\ & \text{subject to} && \mathbf{x}_t \in \mathcal{V}, \quad t = 1, 2, \dots, T \\ & && (\mathbf{x}_t, \mathbf{x}_{t-1}) \in \mathcal{E}, \quad t = 2, \dots, T. \end{aligned}$$

This naturally leads to an alternating optimization strategy. In this strategy, we solve for $\mathcal{X}_T$ from problem (P3), and then for $\hat{\boldsymbol{\Theta}}_p$ and $\hat{\boldsymbol{\Theta}}_m$ from problems (P1) and (P2) iteratively. Since the corresponding iterations never decrease the objective function in equation (24), which is bounded above, the iterations are guaranteed to converge.

1) *Solution to (P1) for the Mobility Model*: With given $\boldsymbol{\Theta}_p$ and $\mathcal{X}_T$, according to the mobility model in (3), setting the derivative of $p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_{t-2}; \boldsymbol{\Theta}_m)$ w.r.t. $\boldsymbol{\Theta}_m = \{\bar{\mathbf{v}}, \sigma_v^2\}$ to zero, we find that the corresponding solution

$$\bar{\mathbf{v}} = \frac{\sum_{t=3}^T (\mathbf{x}_t - (1 + \gamma)\mathbf{x}_{t-1} + \gamma\mathbf{x}_{t-2})}{(T-2)(1-\gamma)\delta} \quad (25)$$

$$\sigma_v^2 = \frac{\sum_{t=3}^T \|\mathbf{x}_t - (1 + \gamma)\mathbf{x}_{t-1} + \gamma\mathbf{x}_{t-2} - (1 - \gamma)\delta\bar{\mathbf{v}}\|_2^2}{2(T-2)\delta^2} \quad (26)$$

is unique. Since (P1) is an unconstrained optimization problem, (25)–(26) give the optimal solution to (P1).

2) Solution to (P2) via Expectation Maximization:

Given the known parameters Θ_m and the trajectory $\mathcal{X}_T$, we can solve (P2) for the set of parameters $\Theta_p = \{\beta_q^{(k)}, \alpha_q^{(k)}, \mathbf{m}_k, \sigma_{s,q,k}^2, \sigma_{\theta,k}^2, \Upsilon_k, u_{t,q}^{(k)}, \sigma_u^2\}$. This is a typical maximum likelihood estimation problem with latent variables, which can be efficiently solved using the Expectation-Maximization (EM) algorithm [25], where one may iteratively estimate the posterior probability $c_{t,q}^{(k)}$ and the model parameter $\Xi = \{\pi_{t,q}^{(k)}, \boldsymbol{\mu}^{(k)}(\mathbf{x}_t), \Sigma^{(k)}\}$.

Specifically, given a random initialization Ξ , the posterior probability $c_{t,q}^{(k)}$ is calculated by

$$c_{t,q}^{(k)} = \frac{\pi_{t,q}^{(k)} \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}^{(k)}(\mathbf{x}_t), \Sigma^{(k)})}{\sum_{j=0}^1 \pi_{t,q}^{(j)} \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}^{(j)}(\mathbf{x}_t), \Sigma^{(j)})}. \quad (27)$$

With the updated $c_{t,q}^{(k)}$, we update Ξ as follows. For updating the parameters $\{\beta_q^{(k)}, \alpha_q^{(k)}, \mathbf{m}_k\}$ in $\boldsymbol{\mu}^{(k)}(\mathbf{x}_t)$, let $\mathbf{D}_q \in \mathbb{R}^{T \times 2}$ be a matrix with the t th row $[1, \log_{10} d(\mathbf{o}_q, \mathbf{x}_t)]$ and $\mathbf{w}_q^{(k)} = [c_{1,q}^{(k)}, \dots, c_{T,q}^{(k)}]^T$ be a vector. Then, the mean parameters $\mathbf{m}_k$ is updated by $\mathbf{m}_k = \sum_{t,q} c_{t,q}^{(k)} [s_{t,q}, \nu_{t,q}]^T / (\sum_{t,q} c_{t,q}^{(k)})$, and $[\beta_q^{(k)}, \alpha_q^{(k)}]^T$ are estimated using weighted least squares

$$[\beta_q^{(k)}, \alpha_q^{(k)}]^T = (\mathbf{D}_q^T \text{Diag}(\mathbf{w}_q^{(k)}) \mathbf{D}_q)^{-1} \mathbf{D}_q^T \text{Diag}(\mathbf{w}_q^{(k)}) \mathbf{s}_q$$

where $\text{Diag}(\mathbf{w}_q^{(k)})$ denotes the diagonal matrix whose diagonal entries are given by the elements of the vector $\mathbf{w}_q^{(k)}$.

To update the variances $\{\sigma_{s,q,k}^2, \sigma_{\theta,k}^2, \Upsilon_k\}$ in $\Sigma^{(k)}$, we leverage its diagonal structure and estimate each component independently

$$\begin{aligned} \sigma_{s,q,k}^2 &= \frac{\sum_{t,q} c_{t,q}^{(k)} \left(s_{t,q} - \beta_q^{(k)} - \alpha_q^{(k)} \log_{10} d(\mathbf{o}_q, \mathbf{x}_t) \right)^2}{\sum_{t,q} c_{t,q}^{(k)}} \\ \sigma_{\theta,k}^2 &= \frac{\sum_{t,q} c_{t,q}^{(k)} \left(\hat{\theta}_{t,q} - \phi(\mathbf{x}_t, \mathbf{o}_q) \right)^2}{\sum_{t,q} c_{t,q}^{(k)}} \\ \Upsilon_k &= \frac{\sum_{t,q} c_{t,q}^{(k)} ([s_{t,q}, \nu_{t,q}]^T - \mathbf{m}_k) ([s_{t,q}, \nu_{t,q}]^T - \mathbf{m}_k)^T}{\sum_{t,q} c_{t,q}^{(k)}}. \end{aligned}$$

Then, the mixture weights are updated as $\pi_{t,q}^{(k)} = \frac{1}{TQ} \sum_{t=1}^T \sum_{q=1}^Q c_{t,q}^{(k)}$.

With the updated Ξ , we compute $c_{t,q}^{(k)}$ using (27) again. Thereafter, a new round of updating for Ξ can be started. Let $\mathcal{L}^{(i)}$ denote the log-likelihood at iteration i . The EM algorithm is terminated once the convergence condition is satisfied $|\mathcal{L}^{(i+1)} - \mathcal{L}^{(i)}| < 10^{-6}$. Note that the spatial regularization is only meaningful in the NLOS case and is therefore not considered in the LOS/NLOS discrimination procedure.

Finally, the LOS/NLOS assignment is given by

$$u_{t,q}^{(k)} = \begin{cases} 1, & \text{if } k = \arg \max_j c_{t,q}^{(j)} \\ 0, & \text{otherwise.} \end{cases}$$

Algorithm 1 An alternating optimization algorithm for trajectory inference.

Initialize the parameter $\Theta_p^{(0)}, \Theta_m^{(0)}$ randomly.

Loop for the $(i+1)$ th iteration:

- Update $\mathcal{X}_T^{(i+1)}$ using the method in Section (IV-C3).
- Update $\Theta_p^{(i+1)}$ using method in Section (IV-C2).
- Update $\Theta_m^{(i+1)}$ using (25)-(26).

Until $\mathcal{X}_T^{(i+1)} = \mathcal{X}_T^{(i)}$.

and the regularization variance σ_u^2 is calculated as

$$\sigma_u^2 = \frac{1}{\sum_{t=1}^T \sum_{q=1}^Q |\tilde{\mathcal{N}}_q(\mathbf{x}_t)|} \sum_{t=1}^T \sum_{q=1}^Q \sum_{\tau \in \tilde{\mathcal{N}}_q(\mathbf{x}_t)} \left[\hat{u}(\mathbf{H}_{t,q}, \mathbf{H}_{\tau,q}) - \frac{B}{c} d(\mathbf{x}_t, \mathbf{x}_\tau) \right]^2$$

where $|\tilde{\mathcal{N}}_q(\mathbf{x}_t)|$ denotes the cardinality of the set $\tilde{\mathcal{N}}_q(\mathbf{x}_t)$.

3) *Solution to (P3) for Trajectory Optimization:* Problem (P3) searches for a trajectory in a discrete space that maximizes the log-likelihood $\mathcal{L}(\mathcal{X}_T, \Theta_p, \Theta_m)$ given the signal propagation parameters $\hat{\Theta}_p$ and mobility model parameters $\hat{\Theta}_m$. Problem (P3) follows a classical hidden Markov model (HMM) optimization form, and can be efficiently solved using a modified version of the Viterbi algorithm with globally optimal guarantee.

At each step, there are $|\mathcal{V}|$ candidate locations considered, but states with very low probabilities $\prod_{q=1}^Q p(\mathbf{y}_{t,q} | \mathbf{x}_t; \hat{\Theta}_p)$ are highly unlikely to contribute to the optimal path. To improve efficiency, states with probabilities below a threshold ζ are pruned. Mathematically, this corresponds to retaining only the top $n_t(\zeta)$ most probable locations at time slot t , where $n_t(\zeta)$ is the number of elements in the set $\{\mathbf{x}_t \mid \prod_{q=1}^Q p(\mathbf{y}_{t,q} | \mathbf{x}_t; \hat{\Theta}_p) > \zeta, \mathbf{x}_t \in \mathcal{V}\}$. Denote the maximum number of element in the set $n_{\max}(\zeta) = \max_t \{n_t(\zeta)\}$.

Considering the number of candidate previous states for the current state, which is constrained by the graph structure, it is of the order $\mathcal{O}(\varrho^2(D_m))$ for a square region, where the max hop $\varrho(D_m)$ at each step is determined by D_m . Thus, the computational complexity of solving problem (P3) is upper bounded by $\mathcal{O}(T n_{\max}(\zeta) \varrho^2(D_m))$.

The overall algorithm is summarized in Algorithm 1. We first initialize the propagation parameter Θ_p and the mobility parameter Θ_m randomly and then begin the alternating update of $\mathcal{X}_T$, Θ_p and Θ_m alternatively until convergence. Since each iteration of this procedure never decreases the objective function, which is bounded above, the iterative process is therefore guaranteed to converge.

V. NUMERICAL EXPERIMENTS

In this section, we first present the experimental setup and scenarios in Section V-A, followed by a numerical validation of the theoretical results in Section V-B. Finally, we evaluate the accuracy of trajectory inference and the constructed radio map in Section V-C.

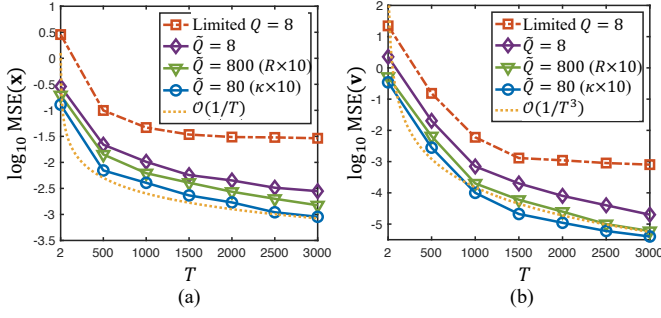


Figure 4. MSE of (a) $\mathbf{x}$ and (b) $\mathbf{v}$ with different sample number T , the number of AP Q , radius R , and density κ .

A. Environmental Setup and Scenarios

This paper validates the proposed algorithm using two datasets:

Synthetic Dataset I: We simulate a trajectory of length 100 meters using the mobility model defined in (3), parameterized by $\gamma = 1$, $\mathbf{v} = [1, 0]^T$ m/s, $\mathbf{x} = [0, 0]^T$ m, and $\delta = 0.1$ s. We consider AP at a height of 3 meters and a mobile user at a height of 1.5 meters equipped with N_t antennas. Two scenarios are considered: in Scenario 1 (AP deployed in a limited region, c.f., Section IV-B), the number of APs surrounding the trajectory is fixed at $Q = 4, 8, 12, 16, 20$; in Scenario 2 (AP deployed in an unlimited region, c.f., Section III-C), the APs in the target area follow a PPP with densities $\kappa = 0.64 \times 10^{-2}, 2.55 \times 10^{-2}, 3.02 \times 10^{-2}, 5.02 \times 10^{-2}, 7.02 \times 10^{-2}, 9.02 \times 10^{-2}, 1.02 \times 10^{-1}$, and 2.55×10^{-1} units per m^2 . The mobile user can only detect APs within a radius of $R = 10, 20, 30, 40, 50, 60, 100$ meters. The number of antennas in each AP is $N_t = 2, 4, 8, 16, 32, 64, 128$.

Synthetic Dataset II: We utilized Wireless Insite[®] to simulate a $26 \text{ m} \times 24 \text{ m}$ indoor environment with a 264 m^2 area. As illustrated in Figure 1, four APs with a height of 3 meters were manually deployed at the corners of the room. Each AP is equipped with an 8-antenna omnidirectional ULA array and configured with $M = 64$ subcarriers using a MIMO-OFDM model. We recorded the CSI at receivers positioned at a height of 1.5 meters along a trajectory with length of 167 m. The sampling interval is set to $\delta = 0.2$ s

B. Numerical Validation of the Theoretical Results

Figure 4 illustrates the MSE defined as $\text{MSE}(\mathbf{x}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 \text{ m}^2$ and $\text{MSE}(\mathbf{v}) = \|\mathbf{v} - \hat{\mathbf{v}}\|_2^2 \text{ m}^2/\text{s}^2$ on the synthetic dataset I with the parameter $\sigma_\theta = 0.1$, where $\hat{\mathbf{x}}$ and $\hat{\mathbf{v}}$ are the outputs of the proposed algorithm.

In Scenario 1 of the synthetic dataset I with $Q = 8$, the MSE of $\mathbf{x}$ and $\mathbf{v}$ decreases as T increases within a limited region but does not converge to zero even when $T = 20000$ in our experiments. This behavior is consistent with Proposition 1 and Proposition 2.

In Scenario 2 of the synthetic dataset I, we set $R = 10$ m and $\kappa = 2.55 \times 10^{-2}$ units per m^2 , resulting in $\tilde{Q} \approx 8$. As T increases, the rate at which $\text{MSE}(\mathbf{x})$ decreases follows $\mathcal{O}(1/T)$, and the rate at which $\text{MSE}(\mathbf{v})$ decreases follows $\mathcal{O}(1/T^3)$, which is consistent with Theorem 2. Scenario 2 with $\tilde{Q} \approx 8$ achieves a lower MSE compared to Scenario 1

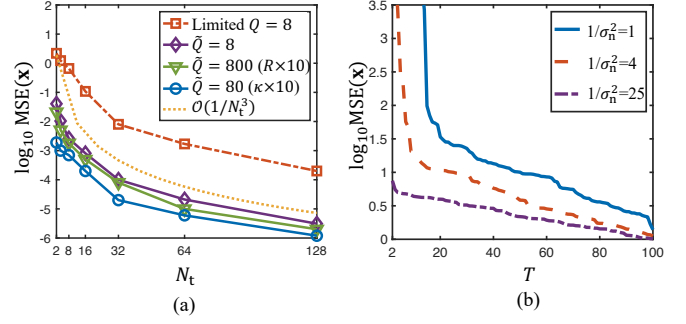


Figure 5. (a) The relationship between $\text{MSE}(\mathbf{x})$ and the number of antennas N_t . (b) $\text{MSE}(\mathbf{x})$ under different noise σ_n^2 .

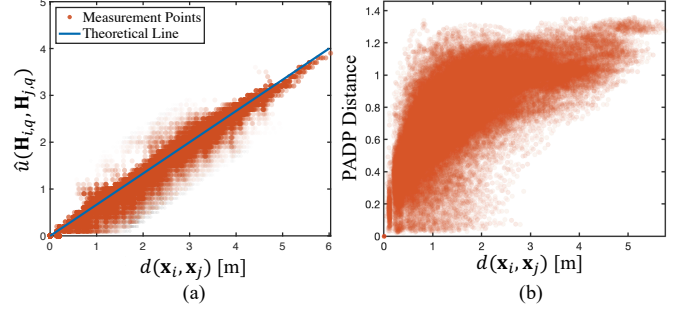


Figure 6. (a) The relationship between $\hat{u}(\mathbf{H}_{i,q}, \mathbf{H}_{j,q})$ and $d(\mathbf{x}_i, \mathbf{x}_j)$ in NLOS region. (b) The relationship between PADP distance and $d(\mathbf{x}_i, \mathbf{x}_j)$ in NLOS region.

with $Q = 8$. Furthermore, The $\text{MSE}(\mathbf{x})$ and $\text{MSE}(\mathbf{v})$ for the curves $\tilde{Q} = 8$ and $\tilde{Q} = 800$ in Figure 4 both reach zero when $T > 3200$. We found that increasing R results in a lower MSE than increasing κ under the same number of APs. This is because $\tilde{\Delta}_{T,x}$ and $\tilde{\Delta}_{T,v}$ in Theorem 2 is related to $\mathcal{O}(1/(\kappa\pi(r_0^{-2} - R^{-2})))$. In addition, we found that increasing the radius R from 50 to 500 meters results in a lower MSE, and increasing the density κ from 2.55×10^{-2} to 2.55×10^{-1} also yields a lower MSE.

We investigate the effect of the number of antenna N_t on $\text{MSE}(\mathbf{x})$. The trajectory length is set to 100 m. As shown in Figure 5 (a), as N_t increases, the rate at which $\text{MSE}(\mathbf{x})$ decreases follows $\mathcal{O}(1/N_t^3)$, which is consistent with Theorem 2.

We also investigate the effect of the noise variance σ_n^2 under the unlimited scenario, with $N_t = 2$, $R = 10$ m, $\kappa = 2.55 \times 10^{-2}$ per m^2 , and a trajectory length of 500 m. We consider $1/\sigma_n^2 = 1, 4, 25$ for all APs. As shown in Figure 5(b), a larger σ_n^2 results in a faster convergence rate. Recall that $1/\sigma_n^2$ is proportional to the SNR. Thus, a smaller $1/\sigma_n^2$ leads to a slower decrease in the CRLB of $\mathbf{x}$ and $\mathbf{v}$ as T increases, as stated in Theorem 2.

Figure 6(a) illustrates the data point pairs $(d(\mathbf{x}_i, \mathbf{x}_j), \hat{u}(\mathbf{H}_{i,q}, \mathbf{H}_{j,q}))$ along with the reference line $y = \frac{B}{c}x$. The data points exhibit a good fit to the line, with a variance of 0.2, thereby verifying the accuracy of Theorem 1. As for a benchmark, we also analyze the power-angular-delay profile (PADP) distance to capture spatial correlations between locations $\mathbf{x}_t$. Let $\mathbf{H}_{t,q} \in \mathbb{C}^{N_t \times M}$ denote the channel matrix. A dictionary $\mathbf{D} \in \mathbb{C}^{N_t \times N}$ with $N = 8N_t$ columns is constructed, and a normalized DFT matrix $\mathbf{F}_M \in \mathbb{C}^{M \times M}$

Table I
COMPARISON OF AVERAGE LOCALIZATION ERROR (E_{loc}) ON THE DATASET II.

	WCL [26]	AoDL [27]	TDoAL [28]	HAT [29]	CC [18]
NLOS	4.95	4.31	6.74	3.98	3.91
Single LOS	4.24	3.82	4.81	3.20	3.14
Double LOS	3.01	2.95	4.14	2.63	1.75
All	3.55	3.38	4.72	2.97	2.19
	HRE [30]	Proposed ($\eta = 0$)	Proposed	GMA (Ideal)	
NLOS	3.31	1.56	1.07	1.02	
Single LOS	2.51	1.21	0.82	0.79	
Double LOS	1.26	0.91	0.59	0.57	
All	1.64	1.02	0.68	0.66	

is defined with unit-norm columns. The normalized PADP is computed as $\mathbf{G}_{t,q} = |\mathbf{D}^H \mathbf{H}_{t,q} \mathbf{F}_M^H| / \|\mathbf{D}^H \mathbf{H}_{t,q} \mathbf{F}_M^H\|_F$, projecting $\mathbf{H}_{t,q}$ onto the spatial-delay domain and normalizing by its Frobenius norm. Figure 6(b) shows the PADP distance $\|\mathbf{G}_i - \mathbf{G}_j\|_F$ and the physical distance between two samples. It is evident that the simple PADP distance does not exhibit a clear relationship with physical distance.

C. Trajectory Inference Performance

We evaluate the trajectory inference performance of the proposed method using the average localization error, defined as $E_{\text{loc}} = \frac{1}{T} \sum_{t=1}^T \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2$, where $\mathbf{x}_t$ denotes the true data collection location at time slot t , and $\hat{\mathbf{x}}_t$ is the corresponding estimated location. The proposed approach is compared against six baseline methods and two variants. The baselines include: (i) Weighted Centroid Localization (WCL) [26], which estimates $\hat{\mathbf{p}}_t = \sum_{q=1}^Q w_{t,q} \mathbf{o}_q$ with weights $w_{t,q} = 10^{s_{t,q}/20} / \sum_{l=1}^Q 10^{s_{t,l}/20}$ derived from the received signal power $s_{t,q}$; (ii) AoD-Based Localization (AoDL) [27], which applies geometric triangulation using AOD measurements and known AP coordinates; (iii) TDoA-Based Localization (TDoAL) [28], which estimates position through the intersection of hyperboloids formed by time difference of arrival (TDOA) measurements; (iv) Hybrid AoD-TDoA (HAT) [29], which combines AoD and TDoA measurements to improve robustness; (v) Channel Charting (CC) [18], which maps CSI features to physical space via a bilateration loss and line-of-sight bounding-box regularization; and (vi) HMM-based RSS Embedding (HRE) [30], which employs a graph-based hidden Markov model to infer trajectories from RSS. In addition, we consider two variants of our approach: the proposed method without the spatial continuity constraint ($\eta = 0$), and a Genius-aided Map-Assisted (GMA) variant that assumes perfect knowledge of propagation parameters and alternately updates the mobility model and trajectory, serving as an upper performance bound. For the proposed method, we set the mobility model trade-off parameter to $\gamma = 0.5$, the regularization parameter to $\eta = 3000$, and the location space resolution in the graph $\mathcal{G}$ to 0.2 m. The parameter D_m is set to 12.48 m, where 12.4 m/s corresponds to the maximum human walking speed. We set $\tilde{\delta} = 2$ m.

As shown in Table I, the proposed method achieves the best overall performance, with an average localization error of just 0.68 meters, significantly outperforming all baselines.

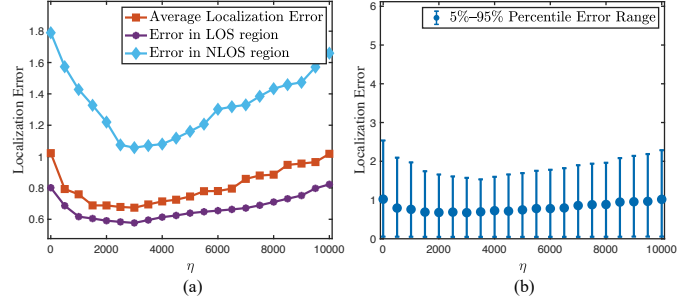


Figure 7. (a) The relationship between average localization error E_{loc} and the spatial regularization parameter η . (b) Average localization error along with minimum and maximum errors as the parameter η .

Among the benchmarks, the time-delay-based method TDoAL performs poorly across all scenarios, with an average error of 4.72 meters. The RSS-based WCL method also yields unsatisfactory results, exhibiting an average error of 3.55 meters, primarily due to the strong fluctuations in signal strength caused by environmental factors. The angle-based AoDL approach delivers slightly better results but still fails to provide reliable localization, as angle information in NLOS regions cannot ensure accurate positioning. The hybrid AoD-TDoA method HAT demonstrates improved performance over both AoDL and TDoAL by combining angular and delay information, achieving an average error of 2.97 meters. However, it still suffers in NLOS environments, where multipath-induced angular deviations significantly degrade its accuracy, resulting in an error of 3.98 meters in the NLOS region. While CC demonstrates moderately improved performance compared to HAT, it remains fundamentally constrained in NLOS-dominated environments due to unresolved multipath interference limitations. In contrast, the superior accuracy of our proposed method stems from its integrated exploitation of power, angle, and delay information, enabling a more holistic characterization of the multipath propagation process. This joint modeling allows the system to better distinguish between LOS and NLOS scenarios and enhances localization robustness in complex environments. In addition, the proposed method with $\eta = 0$ demonstrates that incorporating spatial regularization constraints from the trajectory optimization framework results in a 33.3% reduction in error (decreasing from 1.02 to 0.68 meters), thereby quantitatively validating the necessity of geometric consistency enforcement in trajectory inference.

Figure 7 presents the average localization error as a function of the regularization parameter η . When $\eta = 0$, i.e., without using the regularization term, the error reaches its maximum. As η increases, the error gradually decreases, reaching its minimum value of 0.68 when $\eta = 3000$. However, further increasing η leads to a rise in error. Nonetheless, any $\eta > 0$ leads to an improvement in localization accuracy due to the regularization term. The effectiveness of the regularization term is particularly pronounced in the NLOS regions, where it significantly reduces the estimation error. This is also evidenced by the reduction in the maximum error as η increases (for $\eta < 3000$), since the maximum error mainly originates from NLOS regions.

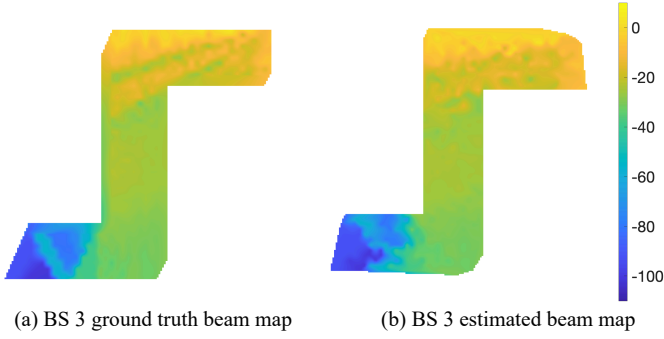


Figure 8. (a) The ground truth beam map in dB (relative value) and (b) the estimated map.

Figure 8 shows the beam map in dB for a particular beam. The proposed method is able to accurately reconstruct the beam map. We define the relative error as $E_{\text{map}} = \frac{1}{NQ} \sum |e_{t,q} - \hat{e}_{t,q}|$, where $\hat{e}_{t,q}$ denotes the estimated beam energy at the ground truth location, based on the beam energy map constructed from the estimated locations and their corresponding beam energy values. The average relative reconstruction error is 3.27 %.

VI. CONCLUSION

This paper presented a blind radio map construction framework that infers user trajectories from indoor MIMO-OFDM channel measurements without requiring explicit location information. Theoretically, a quasi-specular environment model was developed, and a spatial continuity theorem under NLOS conditions was established, leading to a CSI-distance metric proportional to the corresponding physical distance. For rectilinear trajectories under Poisson-distributed AP deployments, it was shown that the CRLB of localization error vanishes asymptotically, even with poor angular resolution—demonstrating the theoretical feasibility of blind localization. Building on these results, a spatially regularized Bayesian inference framework was formulated to jointly estimate channel features, LOS/NLOS conditions and user trajectories. Experiments on a ray-tracing dataset validated the practical effectiveness of the approach, achieving an average localization error of 0.68 m and abeam map reconstruction error of 3.3%.

REFERENCES

- [1] Y. Zeng, J. Chen, J. Xu, D. Wu, X. Xu, S. Jin, X. Gao, D. Gesbert, S. Cui, and R. Zhang, "A tutorial on environment-aware communications via channel knowledge map for 6G," *IEEE Commun. Surv. Tutorials*, vol. 26, no. 3, pp. 1478–1519, 2024.
- [2] D. Romero and S.-J. Kim, "Radio map estimation: A data-driven approach to spectrum cartography," *IEEE Signal Process. Mag.*, vol. 39, no. 6, pp. 53–72, 2022.
- [3] Y. Zeng and X. Xu, "Toward environment-aware 6G communications via channel knowledge map," *IEEE Wireless Commun.*, vol. 28, no. 3, pp. 84–91, 2021.
- [4] S. Timilsina, S. Shrestha, and X. Fu, "Quantized radio map estimation using tensor and deep generative models," *IEEE Trans. Signal Process.*, vol. 72, no. 0, pp. 173–189, 2024.
- [5] S. Shrestha, X. Fu, and M. Hong, "Deep spectrum cartography: Completing radio map tensors using learned neural models," *IEEE Trans. Signal Process.*, vol. 70, pp. 1170–1184, 2022.
- [6] Ç. Yapar, R. Levie, G. Kutyniok, and G. Caire, "Real-time outdoor localization using radio maps: A deep learning approach," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9703–9717, 2023.
- [7] Y. Teganya and D. Romero, "Deep completion autoencoders for radio map estimation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1710–1724, 2021.
- [8] R. Levie, Ç. Yapar, G. Kutyniok, and G. Caire, "Radiounet: Fast radio map estimation with convolutional neural networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 4001–4015, 2021.
- [9] Z. Xing and J. Chen, "Constructing indoor region-based radio map without location labels," *IEEE Trans. Signal Process.*, vol. 72, 2024.
- [10] K. Sato, K. Suto, K. Inage, K. Adachi, and T. Fujii, "Space-frequency-interpolated radio map," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 714–725, 2021.
- [11] Q. Li, X. Liao, A. Li, and S. Valaee, "Automatic indoor radio map construction and localization via multipath fingerprint extrapolation," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 5814–5827, 2023.
- [12] J. Wang, Q. Zhu, Z. Lin, J. Chen, G. Ding, Q. Wu, G. Gu, and Q. Gao, "Sparse Bayesian learning-based hierarchical construction for 3D radio environment maps incorporating channel shadowing," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 14 560–14 574, 2024.
- [13] S. Wang, G. Mao, and J. A. Zhang, "Joint time-of-arrival estimation for coherent UWB ranging in multipath environment with multi-user interference," *IEEE Trans. Signal Process.*, vol. 67, no. 14, pp. 3743–3755, 2019.
- [14] C. Shi, X. Chen, J. Xiao, B. Dai, and H. Lu, "Fast and accurate deep loop closing and relocalization for reliable LiDAR SLAM," *IEEE Trans. Robot.*, vol. 40, pp. 2620–2640, 2024.
- [15] J. Choi, "Sensor-aided learning for Wi-Fi positioning with beacon channel state information," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5251–5264, 2022.
- [16] S. Taner, V. Palhares, and C. Studer, "Channel charting in real-world coordinates," in *Proc. IEEE Global Commun. Conf. (GlobeCom)*. IEEE, 2023, pp. 3940–3946.
- [17] M. Stahlke, G. Yammine, T. Feigl, B. M. Eskofier, and C. Mutschler, "Indoor localization with robust global channel charting: A time-distance-based approach," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 1, pp. 3–17, 2023.
- [18] S. Taner, V. Palhares, and C. Studer, "Channel charting in real-world coordinates with distributed MIMO," *IEEE Trans. Wireless Commun.*, 2025.
- [19] R. He, B. Ai, G. L. Stüber, and Z. Zhong, "Mobility model-based non-stationary mobile-to-mobile channel modeling," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4388–4400, 2018.
- [20] C. Oestges, V. Erceg, and A. J. Paulraj, "A physical scattering model for MIMO macrocellular broadband wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 721–729, 2003.
- [21] D. Astley and B. Ottersten, "The effects of local scattering on direction of arrival estimation with MUSIC," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3220–3234, 2002.
- [22] L. Kong, C. Qi, and H.-D. Qi, "Classical multidimensional scaling: A subspace perspective, over-denoising, and outlier detection," *IEEE Trans. Signal Process.*, vol. 67, no. 14, pp. 3842–3857, 2019.
- [23] Z. Xing and J. Chen, "Blind construction of angular power maps in massive MIMO networks," *Submitted to IEEE Trans. Signal Process.*, 2025.
- [24] Q. Zhang, "Probability of resolution of the MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 43, no. 4, pp. 978–987, 1995.
- [25] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 47–60, 1996.
- [26] J. Wang, P. Urriza, Y. Han, and D. Cabric, "Weighted centroid localization algorithm: Theoretical analysis and distributed implementation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3403–3413, 2011.
- [27] N. Garcia, H. Wymeersch, and D. T. Slock, "Optimal precoders for tracking the AoD and AoA of a mmwave path," *IEEE Trans. Signal Process.*, vol. 66, no. 21, pp. 5718–5729, 2018.
- [28] Y. Sun, K. Ho, and Q. Wan, "Solution and analysis of TDOA localization of a near or distant source in closed form," *IEEE Trans. Signal Process.*, vol. 67, no. 2, pp. 320–335, 2018.
- [29] M. Ahadi and F. Kaltenberger, "5G NR indoor positioning by joint DL-TDOA and DL-AoD," in *IEEE Wirel. Commun. Netw. Conf. (WCNC)*, 2023, pp. 1–6.
- [30] Z. Xing and J. Chen, "HMM-based CSI embedding for trajectory recovery from RSS measurements of non-cooperative devices," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2024, pp. 7060–7064.
- [31] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Courier Corporation, 1965, vol. 55.
- [32] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Academic press, 2014.

APPENDIX A
PROOF OF THEOREM 1

From (7), we have

$$\begin{aligned}\mathbb{E}\{h_1^{(m)}h_2^{(m)*}\} &= \mathbb{E}\left\{\sum_{l=1}^L \kappa_1^{(l)}\kappa_2^{(l)*} e^{-jm\frac{2\pi}{M}B(\tau_1^{(l)}-\tau_2^{(l)})}\right. \\ &\quad \left. + \sum_{l=1}^L \sum_{l' \neq l} \kappa_1^{(l)}\kappa_2^{(l')*} e^{-jm\frac{2\pi}{M}B(\tau_1^{(l)}-\tau_2^{(l')})}\right\} \\ &= C(d) \cdot L \cdot \mathbb{E}\left\{e^{-jm\frac{2\pi}{M}B(\tau_1^{(l)}-\tau_2^{(l)})}\right\}\end{aligned}\quad (28)$$

where the second equality is due to the zero-mean independent and identically distributed (i.i.d.) assumption of the path gains, i.e., $\mathbb{E}\{\kappa_1^{(l)}\kappa_2^{(l')*}\} = 0$ for $l \neq l'$, and the correlation assumption $\mathbb{E}\{\kappa_1^{(l)}\kappa_2^{(l)*}\} = C(d)$ for the small movement.

For the l th path that comes from the mirror image $A_l \in \mathcal{A}^1 = \mathcal{A}^2$, denote the movement direction $\mathbf{p}_2 - \mathbf{p}_1$ relative to the l th arriving at $\mathbf{p}_1$ as θ'_l . It is natural to consider that θ'_l is uniformly distributed in $[-\pi, \pi)$. As a result, for the l th path, the delays satisfy $\tau_2^{(l)} - \tau_1^{(l)} = \frac{d}{c} \cos(\theta'_l)$. Denoting $\omega = B \frac{d}{c}$ for simplification, (28) becomes

$$\begin{aligned}\mathbb{E}\{h_1^{(m)}h_2^{(m)*}\} &= C(d) \cdot L \cdot \mathbb{E}\left\{e^{-jm\frac{2\pi}{M}\omega \cos(\theta'_l)}\right\} \\ &= C(d)L \int_{-\pi}^{\pi} e^{-jm\frac{2\pi}{M}\omega \cos(\theta'_l)} \frac{1}{2\pi} d\theta'_l \\ &= C(d)L \cdot J_0\left(m\frac{2\pi}{M}\omega\right)\end{aligned}\quad (29)$$

where $J_0(x)$ is the zero-th order Bessel function of the first kind [Formula 9.1.21, [31]].

Substituting (29) to (8) for computing the IDFT of (29) yields

$$R(u) = C(d)L \cdot \frac{1}{M} \sum_{m=0}^{M-1} J_0\left(m\frac{2\pi}{M}\omega\right) e^{jm\frac{2\pi}{M}u} \quad (30)$$

$$\approx C(d)L \cdot \int_0^1 J_0(2\pi f\omega) e^{j2\pi f u} df \quad (31)$$

for large M , where (30) becomes a Riemann sum for “frequency bin” $f_m = (\frac{m}{M}, \frac{m+1}{M})$ with interval $\frac{1}{M}$, and hence, asymptotically for $M \rightarrow \infty$, f_m becomes $f \in (0, 1)$, $\frac{1}{M}$ becomes df , and the sum is approximated by the integral.

The integral can be converted into an inverse continuous-time Fourier transform for a function $J_0(2\pi f\omega)\text{rect}(f)$, where $\text{rect}(f)$ is a rectangle function that acts as a window with $\text{rect}(f) = 1$ for $f \in [0, 1]$ and $\text{rect}(f) = 0$ otherwise. If we denote $R_J(u) = \mathcal{F}^{-1}\{J_0(2\pi f\omega)\}$ as the inverse Fourier transform of $J_0(2\pi f\omega)$ and $R_r(u) = \mathcal{F}^{-1}\{\text{rect}(f)\}$, then according to the convolution property of Fourier transform, we have $\mathcal{F}\{R_J(u) * R_r(u)\} = J_0(2\pi f\omega)\text{rect}(f)$, where $\mathcal{F}\{\cdot\}$ denotes the Fourier transform and the convolution is defined as

$$R_J(u) * R_r(u) = \int_{-\infty}^{\infty} R_J(u-t)R_r(t)dt.$$

It follows that $R(u) \approx C(d)L \cdot R_J(u) * R_r(u)$.

The inverse Fourier transform can be computed as

$$\begin{aligned}R_J(u) &= \int_{-\infty}^{\infty} J_0(2\pi f\omega) e^{j2\pi f u} df \\ &= 2 \int_0^{\infty} J_0(2\pi f\omega) \cos(2\pi f u) df \\ &= \begin{cases} \frac{1}{\pi\sqrt{\omega^2 - u^2}} & |u| < \omega \\ 0 & |u| \geq \omega \end{cases}\end{aligned}$$

where the second equality is due to the fact that the Bessel function $J_0(2\pi f\omega)$ is real and even, and the third equality follows [Formula 6.671, [32]]. In addition,

$$R_r(u) = \int_0^1 e^{j2\pi f u} df = e^{j\pi u} \frac{\sin(\pi u)}{\pi u} = e^{j\pi u} \text{sinc}(u).$$

Therefore, we have

$$R(u) \approx C(d)L \left(\frac{\mathbb{I}\{|u| < \omega\}}{\pi\sqrt{\omega^2 - u^2}} \right) * \left(e^{j\pi u} \text{sinc}(u) \right).$$

It is observed that $R_J(u) = \frac{\mathbb{I}\{|u| < \omega\}}{\pi\sqrt{\omega^2 - u^2}}$ is a U-shape function that has its peak at $u \rightarrow \omega_-$ where the $R_J(u)$ approaches ∞ . In addition, $\text{sinc}(u)$ is a pulse with ripples and peaked at $u = 0$. Thus, the convolution roughly shifts the pulse $\text{sinc}(u)$ to $u \approx \omega$, and thus, $|R(u)|$ is peaked at $u \approx \omega = \frac{B}{c}d$, which can be easily verified by numerical plots.

APPENDIX B
PROOF OF PROPOSITION 1

From (14), the FIM $\mathbf{F}_{T,x}$, as the upper diagonal block in $\mathbf{F}_{T,\psi}$, can be expressed as

$$\begin{aligned}\mathbf{F}_{T,x} &= \sum_{t,q} \frac{1}{\sigma_{\theta}^2 d_{t,q}^4(\mathbf{x}, \mathbf{v})} \left(\|\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}\|^2 \mathbf{I} \right. \\ &\quad \left. - (\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})(\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})^T \right)\end{aligned}\quad (32)$$

In the following text, we simplify the notation by writing $\mathbf{l}_q(\mathbf{x})$ as $\mathbf{l}_q$, $d_{t,q}(\mathbf{x}, \mathbf{v})$ as $d_{t,q}$, and $\phi(\mathbf{x} + t\mathbf{v}, \mathbf{o}_q)$ as $\phi_{t,q}$.

Lemma 3. *For a narrow-band, far-field source illuminating an N_t -element uniform linear array, the single-snapshot CRLB for the angle of departure θ is*

$$\text{CRLB}(\hat{\theta}) = \frac{d^2}{G_1 N_t (N_t^2 - 1) \cos^2 \phi}$$

where G_1 is a constant that depends on the antenna configuration with $G_1/d^2 \propto \text{SNR}$, d denotes the distance between the transmitter and the receiver, and ϕ is the angle between the transmitter and the receiver.

Proof. See Appendix E. $\square$

For each q , we have $\cos^2 \phi_{t,q} \leq 1$ for all t with equality achieved when $\phi_{t,q} = 0$. Denote $d_{\min} = \min_q \{d_{\min,q}\}$. Thus, using Lemma 3, the angle variance between the AP and the user location is bounded by

$$\frac{d_{t,q}^2 \sigma_n^2}{G_1 N_t (N_t^2 - 1) \cos^2 \phi_{t,q}} \geq \frac{d_{\min}^2 \sigma_n^2}{G_1 N_t (N_t^2 - 1)}$$

for any t, q .

Thus, we have

$$\sigma_\theta^2 \geq \frac{d_{\min}^2 \sigma_n^2}{G_1 N_t (N_t^2 - 1)}$$

for any t, q .

Thus, $\mathbf{F}_{T,x}$ can be written as

$$\mathbf{F}_{T,x} \preceq C_0 \mathbf{A}_{T,x} \quad (33)$$

where

$$\begin{aligned} \mathbf{A}_{T,x} = \sum_{q=1}^Q & [s_{T,q}^{(0)} (\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T) + s_{T,q}^{(1)} (2\mathbf{v}^T \mathbf{l}_q \mathbf{I} \\ & - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T) + s_{T,q}^{(2)} (\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v} \mathbf{v}^T)] \end{aligned}$$

$$\text{and } C_0 = \frac{N_t(N_t^2-1)\sigma_n^2}{G_1 d_{\min}^2}, \quad s_{T,q}^{(n)} = \sum_{t=1}^T \frac{t^n}{d_{t,q}^4(\mathbf{x}, \mathbf{v})}.$$

Lemma 4. Assume that the trajectory $\mathbf{x}_t$ does not pass any of the AP location $\mathbf{o}_q$. Then, $\mathbf{A}_{T,x} \prec \mathbf{A}_{T+1,x}$ if at least two vectors in $\{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_Q, \mathbf{v}\}$ are linear independent.

Proof. Denote $\mathbf{d}_{t,q} \triangleq \mathbf{l}_q(\mathbf{x}) + t\mathbf{v} = [d_{t,q,1}, d_{t,q,2}]^T$ as the direction from the q th AP to the user position at time slot t . Assume that the trajectory $\mathbf{x}_t$ does not pass any of the AP location $\mathbf{o}_q$. Thus, $\mathbf{l}_q = \mathbf{x} - \mathbf{o}_q \neq \mathbf{0}$ for all q , $d_{t,q} > 0$ for all t, q .

The matrix $\mathbf{A}_{T,x}$ can be expressed as

$$\mathbf{A}_{T,x} = \sum_{q=1}^Q \sum_{t=1}^T \frac{1}{\|\mathbf{d}_{t,q}\|_2^4} (\|\mathbf{d}_{t,q}\|_2^2 \mathbf{I} - \mathbf{d}_{t,q} \mathbf{d}_{t,q}^T).$$

The incremental matrix is defined as:

$$\Delta \mathbf{A}_T = \mathbf{A}_{T+1,x} - \mathbf{A}_{T,x} = \sum_{q=1}^Q \mathbf{P}_q$$

where $\mathbf{d}_{T+1,q} = \mathbf{l}_q + (T+1)\mathbf{v}$ and

$$\mathbf{P}_q = \frac{1}{\|\mathbf{d}_{T+1,q}\|^2} \left(\mathbf{I} - \frac{\mathbf{d}_{T+1,q} \mathbf{d}_{T+1,q}^T}{\|\mathbf{d}_{T+1,q}\|^2} \right).$$

The matrix $\mathbf{P}_q$ is P.S.D. because of the Cauchy-Schwarz inequality. For any non-zero vector $\mathbf{u} \in \mathbb{R}^2$:

$$\mathbf{u}^T \mathbf{P}_q \mathbf{u} = \frac{1}{\|\mathbf{d}_{T+1,q}\|^2} \left(\|\mathbf{u}\|^2 - \frac{(\mathbf{u}^T \mathbf{d}_{T+1,q})^2}{\|\mathbf{d}_{T+1,q}\|^2} \right) \geq 0. \quad (34)$$

Equality holds if and only if $\mathbf{u} \parallel \mathbf{d}_{T+1,q}$, where $\parallel$ represents parallelism. Thus, $\Delta \mathbf{A}_T$ is P.S.D..

Therefore, $\mathbf{A}_{T+1,x} = \mathbf{A}_{T,x}$ holds if and only if $\mathbf{u} \parallel \mathbf{d}_{T+1,q}$ for all q . Recall $\mathbf{d}_{T+1,q} = \mathbf{l}_q + (T+1)\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$, and $\frac{1}{d_{t,q}^4} > 0$ for all q . If $\mathbf{u} \parallel \mathbf{d}_{T+1,q}$ for all q , then $\mathbf{l}_q = k_q \mathbf{u} - (T+1)\mathbf{v}$ for any q , where all vectors $\mathbf{l}_q$ and $\mathbf{v}$ are linear combinations of $\mathbf{u}$. This implies that the set $\{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_Q, \mathbf{v}\}$ spans a subspace of rank at most 1, as all vectors are collinear with $\mathbf{u}$. Thus, if at least two vectors in $\{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_Q, \mathbf{v}\}$ are linearly independent, there exists some q such that $\mathbf{u} \nparallel \mathbf{d}_{T+1,q}$. This ensures that $\mathbf{A}_{T+1,x} \neq \mathbf{A}_{T,x}$. Consequently, we have shown that $\Delta \mathbf{A}_T$ is positive definite when at least two vectors from $\{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_Q, \mathbf{v}\}$ are linearly independent. $\square$

Lemma 4 proves that $\mathbf{A}_{T,x} \prec \mathbf{A}_{T+1,x}$ if at least two vectors in $\{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_Q, \mathbf{v}\}$ are linear independent. Therefore, $\text{tr}\{\mathbf{F}_{T,x}^{-1}\} \geq \bar{\Delta}_{T,x} \triangleq \text{tr}\{(C_0 \mathbf{A}_{T,x})^{-1}\}$.

Similar to , we have

Lemma 5. (Lemma 8 in [23]) Suppose $d_{\min,q} > 0$. The sequence $s_{T,q}^{(n)}$ is bounded for $n < 3$ and divergent as $s_{T,q}^{(n)} \rightarrow \infty$ as $T \rightarrow \infty$ for $n \geq 3$. In addition, $s_{T,q}^{(n+1)}/s_{T,q}^{(n)} \rightarrow \infty$ as $T \rightarrow \infty$ for $n > 3$.

Using Lemma 5, since $s_{T,q}^{(n)}$ are bounded for $n < 3$ and Q is finite, we have $\mathbf{A}_{T,x}$ bounded. Thus, $\bar{\Delta}_{T,x}$ converges to a strictly positive number as $T \rightarrow \infty$.

APPENDIX C

PROOF OF PROPOSITION 2

From (14), the FIM $\mathbf{F}_{T,v}$, as the lower diagonal block in $\mathbf{F}_{T,\psi}$ can be expressed as

$$\begin{aligned} \mathbf{F}_{T,v} = \sum_{t,q} \frac{t^2}{\sigma_\theta^2 d_{t,q}^4(\mathbf{x}, \mathbf{v})} & \left(\|\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}\|^2 \mathbf{I} \right. \\ & \left. - (\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})(\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})^T \right) \end{aligned}$$

Similar to (33), we have,

$$\mathbf{F}_{T,v} \preceq C_0 \mathbf{A}_{T,v} \quad (35)$$

where

$$\begin{aligned} \mathbf{A}_{T,v} = \sum_{q=1}^Q & \left[s_{T,q}^{(2)} (\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T) + \sum_{q=1}^Q s_{T,q}^{(3)} (2\mathbf{v}^T \mathbf{l}_q \mathbf{I} \right. \\ & \left. - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T) + \sum_{q=1}^Q s_{T,q}^{(4)} (\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v} \mathbf{v}^T) \right] \end{aligned}$$

Lemma 6. The eigenvalues of $\mathbf{A}_{T,v}$ satisfies

$$\lambda_{\min}(\mathbf{A}_{T,v}) \rightarrow \sum_{q=1}^Q s_{T,q}^{(2)} \|\mathbf{P}_v^\perp \mathbf{l}_q\|^2$$

and $\lambda_{\max}(\mathbf{A}_{T,v}) \rightarrow \sum_{q=1}^Q s_{T,q}^{(4)} \|\mathbf{v}\|^2$ as $T \rightarrow \infty$, where $\mathbf{P}_v^\perp = \mathbf{I} - \mathbf{v} \mathbf{v}^T / \|\mathbf{v}\|^2$.

Proof. Since the term $\sum_{q=1}^Q s_{T,q}^{(4)} \mathbf{v} \mathbf{v}^T$ dominates $\mathbf{A}_{T,v}$ for a sufficiently large T , for a sufficiently large T , the larger eigenvalue satisfies

$$\begin{aligned} \lambda_{\max}(\mathbf{A}_{T,v}) &= \max_{\|\mathbf{u}\|=1} \mathbf{u}^T \mathbf{A}_{T,v} \mathbf{u} \\ &= \max_{\|\mathbf{u}\|=1} \mathbf{u}^T \left[\sum_{q=1}^Q s_{T,q}^{(2)} (\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T) \right. \\ &\quad \left. + \sum_{q=1}^Q s_{T,q}^{(3)} (2\mathbf{v}^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T) \right. \\ &\quad \left. + \sum_{q=1}^Q s_{T,q}^{(4)} (\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v} \mathbf{v}^T) \right] \mathbf{u} \\ &\approx \max_{\|\mathbf{u}\|=1} \sum_{q=1}^Q s_{T,q}^{(4)} \cdot \mathbf{u}^T (\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v} \mathbf{v}^T) \mathbf{u} \quad (36) \end{aligned}$$

where when $\mathbf{u}$ is orthogonal to $\mathbf{v}$ (i.e., $\mathbf{u}^T \mathbf{v} = 0$), the expression $\|\mathbf{v}\|^2 - (\mathbf{u}^T \mathbf{v})^2 = \|\mathbf{v}\|^2$ attains its maximum value, that is $\mathbf{u} = \mathbf{v}_\perp / \|\mathbf{v}_\perp\|$, we have $\mathbf{v}_\perp^T \mathbf{v}_\perp = 1$, $\mathbf{v}^T \mathbf{v}_\perp = 0$, $\mathbf{v}_\perp^T \mathbf{v} = 0$ and $\lambda_{\max}(\mathbf{A}_{T,v}) \rightarrow \sum_{q=1}^Q s_{T,q}^{(4)} \|\mathbf{v}\|^2$. As a result, asymptotically, the larger eigenvector is $\mathbf{v}_\perp / \|\mathbf{v}_\perp\| \in \mathbb{R}^2$, and hence, the smaller eigenvector is denoted as $\tilde{\mathbf{u}} = \frac{\mathbf{v}}{\|\mathbf{v}\|_2}$.

Consequently, we have

$$\begin{aligned} & \lambda_{\min}(\mathbf{A}_{T,v}) \\ &= \tilde{\mathbf{u}}^T \left[\sum_{q=1}^Q s_{T,q}^{(2)} (\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T) + \sum_{q=1}^Q s_{T,q}^{(3)} (2\mathbf{v}^T \mathbf{l}_q \mathbf{I} \right. \\ & \quad \left. - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T) + \sum_{q=1}^Q s_{T,q}^{(4)} (\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v} \mathbf{v}^T) \right] \tilde{\mathbf{u}} \\ &= \sum_{q=1}^Q s_{T,q}^{(2)} \|\mathbf{P}_v^\perp \mathbf{l}_q\|^2. \end{aligned}$$

where $\mathbf{P}_v^\perp = \mathbf{I} - \mathbf{v} \mathbf{v}^T / \|\mathbf{v}\|^2$ is orthogonal projector, and $\mathbf{P}_v^\perp \mathbf{l}_q$ is to project the vector $\mathbf{l}_q$ onto the null space spanned by $\mathbf{v}_\perp$ of $\mathbf{v}$. $\square$

From $\mathbf{F}_{T,v} \preceq C_0 \mathbf{A}_{T,v}$, since both $\mathbf{F}_{T,v}$ and $\mathbf{A}_{T,v}$ are P.S.D., we have

$$\lambda_{\min}(\mathbf{F}_{T,v}) \leq C_0 \lambda_{\min}(\mathbf{A}_{T,v}), \lambda_{\max}(\mathbf{F}_{T,v}) \leq C_0 \lambda_{\max}(\mathbf{A}_{T,v}). \quad (37)$$

Denoting the Eigen Value Decomposition (EVD) of $\mathbf{F}_{T,v}$ as $\mathbf{F}_{T,v} = \mathbf{u}_{T,v} \mathbf{A}_{T,v} \mathbf{u}_{T,v}^{-1}$, where

$$\mathbf{A}_{T,v} = \begin{bmatrix} \lambda_{\max}(\mathbf{A}_{T,v}) & 0 \\ 0 & \lambda_{\min}(\mathbf{A}_{T,v}) \end{bmatrix},$$

we have

$$\begin{aligned} \text{tr}\{\mathbf{F}_{T,v}^{-1}\} &= \text{tr}\{(\mathbf{u}_{T,v} \mathbf{A}_{T,v} \mathbf{u}_{T,v}^{-1})^{-1}\} = \text{tr}\{\mathbf{u}_{T,v} \mathbf{A}_{T,v}^{-1} \mathbf{u}_{T,v}^{-1}\} \\ &= \lambda_{\max}^{-1}(\mathbf{F}_{T,v}) + \lambda_{\min}^{-1}(\mathbf{F}_{T,v}) \\ &\geq \frac{1}{C_0} \lambda_{\max}^{-1}(\mathbf{A}_{T,v}) + \frac{1}{C_0} \lambda_{\min}^{-1}(\mathbf{A}_{T,v}) \end{aligned} \quad (38)$$

$$\geq \frac{1}{C_0} \lambda_{\min}^{-1}(\mathbf{A}_{T,v}) \triangleq \bar{\Delta}_{T,v} \quad (39)$$

where (38) is due to (37) with equality achieved when $d_{t,q}^2 = d_{\min}^2$, $\cos^2 \phi_{t,q} = 1$ and (39) is due to the fact that $C_0 \lambda_{\max}(\mathbf{A}_{T,v}) > 0$ and equality can be asymptotically achieved at large T as $\lambda_{\max}^{-1}(\mathbf{A}_{T,v}) \rightarrow 1/(\sum_{q=1}^Q s_{T,q}^{(4)} \|\mathbf{v}\|^2)$ which converges to zero.

Using Lemma 6, as $T \rightarrow \infty$, we have

$$\bar{\Delta}_{T,v} \rightarrow C_v = \left(C_0 \sum_{q=1}^Q s_{T,q}^{(2)} \|\mathbf{P}_v^\perp \mathbf{l}_q\|^2 \right)^{-1},$$

which is strictly positive, where $\mathbf{P}_v^\perp = \mathbf{I} - \mathbf{v} \mathbf{v}^T / \|\mathbf{v}\|^2$ is orthogonal projector, and $\mathbf{P}_v^\perp \mathbf{l}_q$ is to project the vector $\mathbf{l}_q$ onto the null space spanned by $\mathbf{v}_\perp$ of $\mathbf{v}$. $s_{T,q}^{(2)}$ is bounded as stated

in Lemma 5. Suppose $\rho > 0$ is sufficiently small such that $d_{t,q} > \rho t$ for all $t \geq 1$, we have

$$\begin{aligned} s_{\infty,q}^{(2)} &= \lim_{T \rightarrow \infty} \sum_{t=1}^T \frac{t^2}{d_{t,q}^4} < \lim_{T \rightarrow \infty} \sum_{t=1}^T \frac{t^2}{(\rho t)^4} \\ &= \frac{1}{\rho^4} \lim_{T \rightarrow \infty} \sum_{t=1}^T \frac{1}{t^2} \approx \frac{\pi^2}{6\rho^4}. \end{aligned}$$

Thus, the element $s_{\infty,q}^{(2)}$ is upper bounded by $\frac{\pi^2}{6\rho^4}$.

APPENDIX D PROOF OF THEOREM 2

Consider the CRLB $B(\mathbf{x})$. Denote $\mathcal{Q}_t = \{q | d_{t,q} \leq R\}$ as the set of APs that are within a range of R from the mobile user at time slot t . Based on the FIM $\mathbf{F}_{T,\psi}$ in (32), we have

$$\begin{aligned} \mathbf{F}_{T,x} &= \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} \frac{1}{\sigma_\theta^2 d_{t,q}^4(\mathbf{x}, \mathbf{v})} (\|\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}\|^2 \mathbf{I} \right. \\ & \quad \left. - (\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})(\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})^T) \right\} \end{aligned} \quad (40)$$

For each term in the sum, the geometry-dependent Fisher information per sample is less than or equal to that from the uniform minimum-variance case. Thus, using Lemma 3, we have

$$\begin{aligned} \mathbf{F}_{T,x} &\succeq \frac{G_1 N_t (N_t^2 - 1)}{\sigma_n^2} \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} \frac{d_{t,q,1}^2}{d_{t,q}^8} (\|\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}\|^2 \mathbf{I} \right. \\ & \quad \left. - (\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})(\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})^T) \right\} \\ &= \tilde{C}_0 \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 \left[\frac{\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T}{d_{t,q}^8} \right. \right. \\ & \quad \left. \left. + t \frac{2\mathbf{v}^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T}{d_{t,q}^8} + t^2 \frac{\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v} \mathbf{v}^T}{d_{t,q}^8} \right] \right\} \\ &= \tilde{C}_0 \tilde{\mathbf{A}}_{T,x} \end{aligned}$$

where $\tilde{C}_0 = G_1 N_t (N_t^2 - 1) / \sigma_n^2$ and

$$\begin{aligned} \tilde{\mathbf{A}}_{T,x} &= \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 \frac{\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T}{d_{t,q}^8} \right\} \\ &+ \sum_{t=1}^T t \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 \frac{2\mathbf{v}^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T}{d_{t,q}^8} \right\} \\ &+ \sum_{t=1}^T t^2 \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 \frac{\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v} \mathbf{v}^T}{d_{t,q}^8} \right\}. \end{aligned} \quad (41)$$

Since $\mathbf{F}_{T,x}$ and $\tilde{\mathbf{A}}_{T,x}$ are 2×2 symmetric and positive semi-definite, their eigenvalues are real and non-negative. From $\mathbf{F}_{T,x} \succeq \tilde{C}_0 \tilde{\mathbf{A}}_{T,x}$, we have

$$\lambda_{\min}(\mathbf{F}_{T,x}) \geq \tilde{C}_0 \lambda_{\min}(\tilde{\mathbf{A}}_{T,x}), \lambda_{\max}(\mathbf{F}_{T,x}) \geq \tilde{C}_0 \lambda_{\max}(\tilde{\mathbf{A}}_{T,x}).$$

Since $\text{tr}\{\mathbf{F}_{T,x}^{-1}\} = \lambda_{\max}^{-1}(\mathbf{F}_{T,x}) + \lambda_{\min}^{-1}(\mathbf{F}_{T,x})$, we have

$$\text{tr}\{\mathbf{F}_{T,x}^{-1}\} \leq 2\lambda_{\min}^{-1}(\mathbf{F}_{T,x}) \leq 2\left(\tilde{C}_0\lambda_{\min}(\tilde{\mathbf{A}}_{T,x})\right)^{-1} \triangleq \tilde{\Delta}_{T,x}. \quad (42)$$

Lemma 7. Assume that $d_{t,q} \geq r_0$ for all t and q . The eigenvalue of $\tilde{\mathbf{A}}_{T,x}$ satisfies

$$\frac{1}{T}\lambda_{\min}(\tilde{\mathbf{A}}_{T,x}) \rightarrow \frac{1}{8}\pi\kappa\left(\frac{1}{r_0^2} - \frac{1}{R^2}\right)$$

as $T \rightarrow \infty$.

Proof. The term $\sum_{t=1}^T t^2 \mathbb{E}\{\sum_{q \in \mathcal{Q}_t} d_{t,q,2}^2 (\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v}\mathbf{v}^T) / d_{t,q}^8\}$ in (41) dominates $\tilde{\mathbf{A}}_{T,x}$ for a sufficiently large T , because t^2 increases quadratically. Thus, as $T \rightarrow \infty$, the larger eigenvalue satisfies:

$$\begin{aligned} \frac{1}{T}\lambda_{\max}(\tilde{\mathbf{A}}_{T,x}) &= \frac{1}{T} \max_{\|\tilde{\mathbf{u}}\|=1} \tilde{\mathbf{u}}^T \tilde{\mathbf{A}}_{T,x} \tilde{\mathbf{u}} \\ &\rightarrow \max_{\|\tilde{\mathbf{u}}\|=1} \frac{1}{T} \sum_{t=1}^T t^2 \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 \frac{1}{d_{t,q}^8} \right. \\ &\quad \left. \times \tilde{\mathbf{u}}^T (\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v}\mathbf{v}^T) \tilde{\mathbf{u}} \right\} \end{aligned} \quad (43)$$

where the solution to (43) is $\tilde{\mathbf{u}} = \mathbf{v}_\perp / \|\mathbf{v}_\perp\|_2$, which satisfies $\mathbf{v}^T \mathbf{v}_\perp = 0$.

As a result, asymptotically, the larger eigenvector is $\mathbf{v}_\perp / \|\mathbf{v}_\perp\|_2 \in \mathbb{R}^2$ and hence, the smaller eigenvector is $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|_2}$. Consequently, from (41), as $T \rightarrow \infty$, we have:

$$\begin{aligned} \frac{1}{T}\lambda_{\min}(\tilde{\mathbf{A}}_{T,x}) & \quad (44) \\ &\rightarrow \mathbf{u}^T \left[\frac{1}{T} \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 \frac{\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T}{d_{t,q}^8} \right\} \right. \\ &\quad + \frac{1}{T} \sum_{t=1}^T t \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 \frac{1}{d_{t,q}^8} (2\mathbf{v}^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T) \right\} \\ &\quad \left. + \frac{1}{T} \sum_{t=1}^T t^2 \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 \frac{\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v}\mathbf{v}^T}{d_{t,q}^8} \right\} \right] \mathbf{u} \\ &= \frac{1}{T} \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} \frac{d_{t,q,1}^2}{d_{t,q}^8} \|\mathbf{P}_v^\perp \mathbf{l}_q\|^2 \right\}. \end{aligned} \quad (45)$$

To compute the expectation in (45), we note that as the APs follow a Poisson distribution within a radius of R from the user location $\mathbf{x}_t$, the expected number of the APs is $\kappa\pi R^2$. In addition, given the number of the APs, the APs are independently and uniformly distributed. As a result, consider a coordinate system with the initial position $\mathbf{x}$ as the origin and the direction $\mathbf{v}$ as the x -axis as stated in Lemma 10 in [23], and then, $\mathbf{P}_v^\perp \mathbf{l}_q$ is simply to project the vector $\mathbf{l}_q = \mathbf{x} - \mathbf{o}_q$ onto the y -axis. Denote $\mathbf{l}_q = (l_{q,x}, l_{q,y})$ and it follows that $\mathbf{P}_v^\perp \mathbf{l}_q = l_{q,y}$ and $d_{t,q,1}^2 = l_{q,y}^2$.

We have

$$\begin{aligned} &\mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} \frac{d_{t,q,1}^2}{d_{t,q}^8} \|\mathbf{P}_v^\perp \mathbf{l}_q\|^2 \right\} \\ &= \mathbb{E} \left\{ \frac{l_{q,y}^2 l_{q,x}^2}{(l_{q,x}^2 + l_{q,y}^2)^4} \right\} \kappa\pi R^2 \\ &= \kappa\pi R^2 \frac{1}{\pi R^2} \int_{-R}^R \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{x^2 y^2}{(x^2 + y^2)^4} dy dx \quad (46) \\ &= \kappa \frac{\pi}{8} \left(\frac{1}{r_0^2} - \frac{1}{R^2} \right). \end{aligned} \quad (47)$$

where (46) is due to the prior condition that $d_{t,q} > r_0$ for all t and q .

As a result, from (45) and (47), we have $\frac{1}{T}\lambda_{\min}(\tilde{\mathbf{A}}_{T,x}) \rightarrow \frac{1}{8}\pi\kappa\left(\frac{1}{r_0^2} - \frac{1}{R^2}\right)$ as $T \rightarrow \infty$. $\square$

According Lemma 7 and from (42), we have

$$T\tilde{\Delta}_{T,x} \rightarrow \frac{16}{\kappa\pi(r_0^{-2} - R^{-2})G_1 N_t(N_t^2 - 1)}$$

as $T \rightarrow \infty$.

Consider the CRLB $B(\mathbf{v})$. Based on the FIM $\mathbf{F}_{T,v}$ in (32), we have

$$\begin{aligned} \mathbf{F}_{T,v} &= \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} \frac{t^2}{\sigma_{\theta,t,q}^2 d_{t,q}^4 (\mathbf{x}, \mathbf{v})} [\|\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}\|^2 \mathbf{I} \right. \\ &\quad \left. - (\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})(\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})^T] \right\} \end{aligned} \quad (48)$$

Using Lemma 3, we have

$$\begin{aligned} \mathbf{F}_{T,x} &\succeq \frac{G_1 N_t(N_t^2 - 1)}{\sigma_n^2} \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} \frac{d_{t,q,1}^2 t^2}{d_{t,q}^8} (\|\mathbf{l}_q(\mathbf{x}) + t\mathbf{v}\|^2 \mathbf{I} \right. \\ &\quad \left. - (\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})(\mathbf{l}_q(\mathbf{x}) + t\mathbf{v})^T) \right\} \\ &= \tilde{C}_0 \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 t^2 \left[\frac{\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T}{d_{t,q}^8} \right. \right. \\ &\quad \left. \left. + t \frac{2\mathbf{v}^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T}{d_{t,q}^8} + t^2 \frac{\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v}\mathbf{v}^T}{d_{t,q}^8} \right] \right\} \\ &= \tilde{C}_0 \tilde{\mathbf{A}}_{T,v} \end{aligned}$$

where

$$\begin{aligned} \tilde{\mathbf{A}}_{T,v} &= \sum_{t=1}^T \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 t^2 \frac{\mathbf{l}_q^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{l}_q^T}{d_{t,q}^8} \right\} \\ &\quad + \sum_{t=1}^T t \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 t^2 \frac{2\mathbf{v}^T \mathbf{l}_q \mathbf{I} - \mathbf{l}_q \mathbf{v}^T - \mathbf{v} \mathbf{l}_q^T}{d_{t,q}^8} \right\} \\ &\quad + \sum_{t=1}^T t^2 \mathbb{E} \left\{ \sum_{q \in \mathcal{Q}_t} d_{t,q,1}^2 t^2 \frac{\|\mathbf{v}\|^2 \mathbf{I} - \mathbf{v}\mathbf{v}^T}{d_{t,q}^8} \right\}. \end{aligned} \quad (49)$$

Since $\mathbf{F}_{T,v}$ is symmetric and positive semi-definite, its eigenvalues are real and non-negative. Similar to (42), we have

$$\text{tr}\{\mathbf{F}_{T,v}^{-1}\} \leq 2\lambda_{\min}^{-1}(\mathbf{F}_{T,v}) \leq 2[\tilde{C}_0\lambda_{\min}(\tilde{\mathbf{A}}_{T,v})]^{-1} \triangleq \tilde{\Delta}_{T,v}. \quad (50)$$

Lemma 8. Under the same condition in Lemma 7, the eigenvalue of $\tilde{\mathbf{A}}_{T,v}$ satisfies

$$\frac{\lambda_{\min}(\tilde{\mathbf{A}}_{T,x})}{T(T+1)(2T+1)} \rightarrow \frac{\pi}{48} \kappa \left(\frac{1}{r_0^2} - \frac{1}{R^2} \right)$$

as $T \rightarrow \infty$.

Proof. Similar to the derivation of Lemma 7, the asymptotic larger eigenvector of $\tilde{\mathbf{A}}_{T,v}$ is $\mathbf{u} = \mathbf{v}_\perp / \|\mathbf{v}_\perp\|_2$, because the last term in (49) dominates when T is large.

As a result, asymptotically, the smaller eigenvector is $\mathbf{v} / \|\mathbf{v}\|_2$. Consequently, from (49), as $T \rightarrow \infty$, we have

$$\frac{\lambda_{\min}(\tilde{\mathbf{A}}_{T,v})}{T(T+1)(2T+1)} \rightarrow \frac{\kappa\pi}{48} \left(\frac{1}{r_0^2} - \frac{1}{R^2} \right)$$

as $T \rightarrow \infty$. $\square$

According Lemma 8 and from (50), we have

$$T(T+1)(2T+1)\tilde{\Delta}_{T,v} \rightarrow \frac{96}{\kappa\pi(r_0^{-2} - R^{-2})G_1N_t(N_t^2 - 1)}$$

as $T \rightarrow \infty$.

APPENDIX E PROOF OF LEMMA 3

Consider a narrowband MIMO system where the received signal model is given by

$$\mathbf{y} = \mathbf{a}(\theta)s + \mathbf{n},$$

where $\mathbf{a}(\theta)$ is the array steering vector in (1), s is the transmitted signal (a scalar), and $\mathbf{n} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$ represents additive complex Gaussian noise. The SNR is defined as

$$\text{SNR} = \frac{|s|^2}{\sigma_n^2} = \frac{G_0}{d^2 \sigma_n^2}.$$

where the received signal power $|s|^2$ decays inversely with the square of the distance d^2 , as given by the Friis free-space transmission equation, d denotes the distance between the transmitter and the receiver, and G_0 is a constant that

aggregates transmit power, receive power gain, antenna gains, etc.

Taking the derivative of $\mathbf{a}(\theta)$ with respect to θ , we obtain

$$\frac{\partial \mathbf{a}(\theta)}{\partial \theta} = -j2\pi \frac{\Delta}{\lambda} \cos \theta \mathbf{b}^T(\theta)$$

where

$$\mathbf{b}(\theta) = [0, e^{-j2\pi \frac{\Delta}{\lambda} \sin \theta}, 2e^{-j2\pi \frac{\Delta}{\lambda} 2 \sin \theta}, \dots, (N_t - 1)e^{-j2\pi \frac{\Delta}{\lambda} (N_t - 1) \sin \theta}].$$

According to the standard FIM formula for complex Gaussian models, we have

$$\begin{aligned} F_{\theta\theta} &= 2 \operatorname{Re} \left\{ \frac{1}{\sigma_n^2} \left(\frac{\partial(\mathbf{a}(\theta)s)}{\partial \theta} \right)^H \frac{\partial(\mathbf{a}(\theta)s)}{\partial \theta} \right\} \\ &= \frac{2|s|^2}{\sigma_n^2} \operatorname{Re} \left\{ \left(\frac{\partial \mathbf{a}^H(\theta)}{\partial \theta} \right) \left(\frac{\partial \mathbf{a}(\theta)}{\partial \theta} \right) \right\}. \end{aligned}$$

Substituting the previous derivative, we have

$$\frac{\partial \mathbf{a}^H(\theta)}{\partial \theta} = j2\pi \frac{\Delta}{\lambda} \cos \theta \mathbf{b}^H(\theta),$$

thus

$$\left(\frac{\partial \mathbf{a}^H(\theta)}{\partial \theta} \right) \left(\frac{\partial \mathbf{a}(\theta)}{\partial \theta} \right) = \left(2\pi \frac{\Delta}{\lambda} \cos \theta \right)^2 \mathbf{b}^H(\theta) \mathbf{b}(\theta).$$

Therefore, the FIM becomes

$$F_{\theta\theta} = 2 \text{SNR} \cdot \left(2\pi \frac{\Delta}{\lambda} \cos \theta \right)^2 \mathbf{b}^H(\theta) \mathbf{b}(\theta).$$

By definition, $\mathbf{b}^H(\theta) \mathbf{b}(\theta) = \sum_{n=1}^{N_t-1} n^2$, and using the formula for the sum of squares, $\sum_{n=1}^{N_t-1} n^2 = \frac{(N_t-1)N_t(2N_t-1)}{6}$. Substituting back, we get

$$F_{\theta\theta} = \frac{3}{2} \text{SNR} \left(\pi \frac{\Delta}{\lambda} \cos \theta \right)^2 N_t(N_t^2 - 1).$$

Thus, the CRLB for angle estimation is:

$$\text{CRLB}(\theta) = \frac{1}{F_{\theta\theta}} = \frac{d^2 \sigma_n^2}{G_1 N_t (N_t^2 - 1) \cos^2 \theta}$$

where $G_1 = \frac{3G_0(\pi\Delta/\lambda)^2}{2}$ and $\frac{G_1}{d^2 \sigma_n^2} \propto \text{SNR}$.