CHALMERS
UNIVERSITY OF TECHNOLOGY

Digital Twin-Assisted High-Precision Massive MIMO Localization in Urban Canyons

Citation for the original published paper (version of record):
Zhou, Z., Chen, H., Steinbock, G. et al (2026). Digital Twin-Assisted High-Precision Massive MIMO Localization in Urban Canyons. 2026 IEEE 6TH INTERNATIONAL SYMPOSIUM ON JOINT COMMUNICATIONS & SENSING, JC&S. http://dx.doi.org/10.1109/JCS69321.2026.11366027

Digital Twin-Assisted High-Precision Massive MIMO Localization in Urban Canyons

Ziqin Zhou*, Hui Chen*, Gerhard Steinböck†, Henk Wymeersch*
*Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden
†Ericsson Research, Ericsson, Gothenburg, Sweden

***Abstract*—High-precision wireless localization in urban canyons is challenged by noisy measurements and severe non-line-of-sight (NLOS) propagation. This paper proposes a robust three-stage algorithm synergizing a digital twin (DT) model with the random sample consensus (RANSAC) algorithm to overcome these limitations. The method leverages the DT for geometric path association and employs RANSAC to identify reliable line-of-sight (LOS) and single-bounce NLOS paths while rejecting multi-bounce outliers. A final optimization on the resulting inlier set estimates the user's position and clock bias. Simulations validate that by effectively turning NLOS paths into valuable geometric information via the DT, the approach enables accurate localization, reduces reliance on direct LOS, and significantly lowers system deployment costs, making it suitable for practical deployment.**

***Index Terms*—Digital Twin, Massive MIMO, Localization, Map-aided Localization**

I. Introduction

A digital twin (DT) is a virtual representation that serves as the real-time digital counterpart of a physical object, process, or system [1]. The concept involves three core components: the physical entity, its virtual model, and a continuous data link that ensures the virtual model accurately reflects the state of the physical entity [2]. While originating in manufacturing, this paradigm has been widely adopted in other fields, including wireless communications, leading to the vision of a DT network [3]. A DT network creates a virtual replica of the entire communication infrastructure and, crucially, the surrounding radio environment [4]. A key function of a DT network is to act as a physics-based simulation engine that uses its detailed geometric model to simulate complex radio propagation, generating dynamic and accurate radio environment maps on demand [5]. This capability is envisioned as a cornerstone for future 6G networks, enabling proactive and predictive services by anticipating future user states and network conditions [6].

Map-aided localization has long been recognized as a powerful approach for improving the accuracy and robustness of wireless positioning systems. Early works leveraged 2D floor plans, often derived from architectural drawings, to impose crucial geometric constraints on the estimation process. For instance, these maps were instrumental in refining the solution space of probabilistic algorithms [7]. Furthermore, these maps enabled basic non-line-of-sight (NLOS) identification by checking if the direct path between a transmitter and receiver was obstructed by a structure [8]. However, the utility of such 2D representations is inherently limited in complex three-dimensional environments like urban canyons, as they fail to capture vertical signal propagation and cannot model the complex multi-path phenomena originating from reflections off building facades. The advent of high-fidelity 3D models and, more recently, DT has opened new frontiers for map-aided localization, offering a much richer description of the environment. Current research on DT-aided localization can be broadly categorized into three groups. First, some approaches treat the DT as the ultimate form of a map for simultaneous localization and mapping (SLAM), where an agent builds and refines the map using wireless signals while simultaneously determining its location within it [9], [10]. A second category of works focuses on leveraging the DT for enhanced channel characterization. By generating "massive fingerprinting" databases with virtually unlimited density through simulation [11], these methods can better distinguish between line-of-sight (LOS) and NLOS conditions, thereby improving accuracy. A third area emphasizes dynamic adaptation, where the DT is continuously synchronized with the physical world to account for changes such as moving objects [12]. Despite these advancements, a critical open challenge remains: effectively leveraging the DT's detailed geometric information to *distinguish and mitigate the effects of severe multi-bounce NLOS paths*, which are common in complex environments like indoors and urban canyons [10]. These paths act as geometric outliers that degrade performance, and existing methods often struggle to robustly identify and reject them.

In this paper, we address this challenge by proposing a novel DT-aided localization framework designed to explicitly handle multi-bounce NLOS paths as geometric outliers. By treating this as an outlier rejection problem, our framework leverages the strong geometric constraints imposed by the DT to isolate and filter these inaccurate measurements. The main contributions of this work are as follows: (i) A novel three-stage robust localization framework that explicitly formulates the multi-bounce NLOS challenge as a geometric outlier rejection problem. The framework synergizes a DT model with the random sample consensus (RANSAC) algorithm to systematically identify a consensus set of reliable LOS and single-bounce paths from noisy measurements. (ii) A probabilistic path association technique to robustly generate

This research was supported by the National Growth Fund through the Dutch 6G flagship project "Future Network Services."

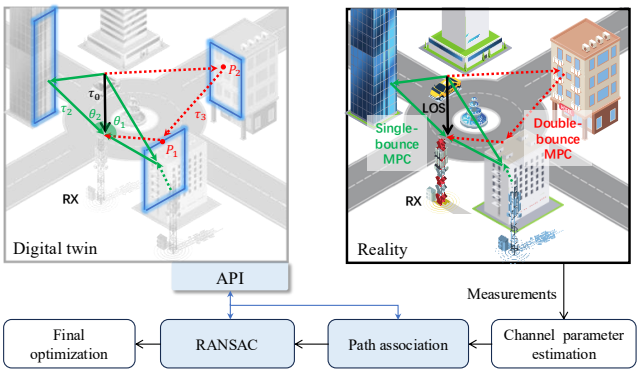

Fig. 1. Single-BS localization with aid of a DT. The DT provides an interface (API) that the localization algorithm can call at various stages (highlighted in blue).

the initial geometric hypotheses required by the framework. This method explicitly accounts for measurement uncertainty in the angles-of-arrival (AOA) by using a Monte Carlo-based ray-casting approach, ensuring reliable hypothesis generation even in noisy conditions.

II. System Model

As shown in Fig.1, our system model comprises two distinct yet complementary components: a DT model and a generative model. The DT serves as a simplified geometric representation of the environment, providing a priori knowledge that the localization algorithm uses to interpret real-world measurements. The generative model, in contrast, is a physics-based simulator designed to mimic the complexity of the real-world measurement process, creating realistic synthetic data. This allows us to use the generative model to generate synthetic data, which serves as the input for our algorithm during validation.

A. Generative Model

We consider a single-user localization system in a 3D urban environment, modeled using a cartesian coordinate system $[x, y, z]^\top$ with positions in meters. The system comprises a user equipment (UE), located at $\mathbf{u} = [u_x, u_y, u_z]^\top$, and a multi-antenna base station (BS), positioned at an elevated location $\mathbf{b} = [b_x, b_y, b_z]^\top$. The BS employs a uniform planar array (UPA) with $N = N_x \times N_y$ antennas, with an inter-antenna spacing of $d = \lambda/2$, where λ is the signal wavelength. With limited loss of generality, we consider the BS orientation to be aligned with the axes of the coordinate system, avoiding the need to carry the BS orientation in the derivations. This alignment allows the measured AOA to be directly interpreted as the global AOA, which is used throughout this paper. The generative model simulates the wireless signal propagation and subsequent measurements within the environment. The system operates in a single-input multiple-output (SIMO) configuration under an orthogonal frequency division multiplexing (OFDM) scheme.

1) Ray Generation: For any given UE location, the generative model generates L distinct propagation paths between the UE and the BS, including a potential LOS path and multiple NLOS reflective paths. Each path $\ell \in \{0, 1, \ldots, L\}$ is characterized by its propagation delay τ_ℓ and AOA at the BS, which consists of the azimuth angle θ_ℓ and the zenith angle ϕ_ℓ, denoted by the vector $\boldsymbol{\theta}_\ell = [\theta_\ell, \phi_\ell]^\top$. Signal propagation paths are constructed using the principle of specular reflection.

- *LOS path:* The direct LOS path ($\ell = 0$) is considered valid only if the straight line connecting the BS at $\mathbf{b}$ and the UE at $\mathbf{u}$ is not obstructed by any planar surface S_i. If the path is unobstructed, its ground-truth delay and AOA are calculated as
$$\tau_0 = \frac{\|\mathbf{u} - \mathbf{b}\|}{c}, \tag{1}$$
$$\boldsymbol{\theta}_0 = \left[\arctan\left(\frac{u_y - b_y}{u_x - b_x}\right), \arccos\left(\frac{u_z - b_z}{\|\mathbf{u} - \mathbf{b}\|}\right)\right]^\top, \tag{2}$$
where c is the speed of light.
- *Reflective NLOS paths:* Specular reflection paths, including both single- and multi-bounce scenarios, are generated using the image method [13]. As illustrated in Fig. 1, this technique geometrically determines the point of reflection on a surface, denoted as $\mathbf{r}_\ell$, by creating virtual images of the Receiver (RX) across reflecting surfaces. For instance, the virtual image RX' shown in the figure is used to find the reflection point for the single-bounce path by intersecting the line segment connecting the transmitter (TX) and RX' with the surface S_1. Multi-bounce paths are found by recursively applying this image creation and intersection process. Each potential path generated this way is then validated for physical feasibility, ensuring that all reflection points $\mathbf{r}_\ell$ lie within their respective surface boundaries and that all path segments remain unobstructed. For every valid path, the total path length yields the ground-truth propagation delay, while the direction of the final segment arriving at the receiver determines the global AOA.

In addition to these reflections, for simplicity, our generative model is limited to LOS and specular reflection paths. Other propagation effects such as diffuse scattering, diffraction, though supported by modern ray tracers, are not considered in this study.

2) Signal Model: Based on the geometric propagation model described previously, we formulate the uplink signal model of an OFDM system. For an OFDM system with K subcarriers, the channel frequency response (CFR) is given by (for $k \in \{0, 1, \ldots, K-1\}$) $\mathbf{H}[k] = \sum_{\ell=0}^{L} \alpha_\ell \mathbf{a}(\boldsymbol{\theta}_\ell) e^{-j2\pi f_k \tau_\ell}$, where α_ℓ is the complex channel gain, τ_ℓ is the propagation delay, $\boldsymbol{\theta}_\ell$ is the AOA, $\mathbf{a}(\boldsymbol{\theta}_\ell) \in \mathbb{C}^{N\times 1}$ is the array steering vector, and f_k is the frequency of the k-th subcarrier. Consequently, assuming the BS is equipped with a fully digital array, the received signal vector $\mathbf{y}[k] \in \mathbb{C}^{N\times 1}$ at the BS for the k-th subcarrier can be written as
$$\mathbf{y}[k] = \mathbf{H}[k]s[k] + \mathbf{n}[k], \tag{3}$$
where $s[k] \in \mathbb{C}$ is the complex-valued data symbol transmitted by the UE on the k-th subcarrier. The term $\mathbf{n}[k] \in \mathbb{C}^{N\times 1}$ represents the additive white Gaussian noise (AWGN) vector, which is modeled as a circularly-symmetric complex normal

random variable, i.e., $\mathbf{n}[k] \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$, with σ^2 being the noise variance and $\mathbf{I}_N$ being the $N \times N$ identity matrix.

B. Measurement Model

In a practical system, the BS would estimate the delay and AOA for each path by processing single-snapshot uplink reference signals from the UE using standard channel parameter estimation methods such as sparsity-based methods (e.g., OMP [14]) or spatial smoothing subspace-based methods (e.g., SS-MUSIC [15]). This work focuses on the positioning part and assumes the channel parameters are obtained from (3) using an efficient estimator, resulting in the following measurement model. The measured propagation delay for path ℓ is modeled as $\hat{\tau}_\ell = L_\ell/c + B + n_{\tau_\ell}$, where the path length L_ℓ is $\|\mathbf{u} - \mathbf{b}\|$ for the LOS path. For NLOS paths, it is the sum of the lengths of all constituent segments (e.g., $\|\mathbf{u} - \mathbf{r}_\ell\| + \|\mathbf{r}_\ell - \mathbf{b}\|$ for a single-bounce path). The model includes an unknown clock bias B common to all paths and path-specific measurement noise $n_{\tau_\ell} \sim \mathcal{N}(0, \sigma^2_{\tau_\ell})$. The measured AOA is modeled by $\hat{\boldsymbol{\theta}}_\ell = \boldsymbol{\theta}_\ell + \mathbf{n}_{\boldsymbol{\theta}_\ell}$, where $\mathbf{n}_{\boldsymbol{\theta}_\ell} = [n_{\theta_\ell}, n_{\phi_\ell}]^\top$ is the measurement noise vector. Its components are modeled as independent, zero-mean Gaussian[1] random variables with potentially different variances, i.e., $n_{\theta_\ell} \sim \mathcal{N}(0, \sigma^2_\theta)$ and $n_{\phi_\ell} \sim \mathcal{N}(0, \sigma^2_\phi)$. We assume that the estimation errors are mutually independent across different paths and parameter types.

C. Digital Twin

The DT is a digital representation of the urban environment, implemented as a functional object that encapsulates the geometry of a set of M planar surfaces, $\{S_i\}_{i=1}^M$. Each surface S_i is defined by key geometric properties, including a unit normal vector $\mathbf{n}_i = [n_{ix}, n_{iy}, n_{iz}]^\top$, a point on the surface $\mathbf{p}_i = [p_{ix}, p_{iy}, p_{iz}]^\top$, and its finite boundaries. The DT provides a well-defined interface with three primary methods for geometric queries.

The first method, $i = \texttt{FindSurface}(\mathbf{d}, \mathbf{b})$, serves as a fundamental ray-casting tool. Given a ray originating from the BS at position $\mathbf{b}$ with a direction vector $\mathbf{d}$, its purpose is to identify the first environmental surface the ray intersects. The implementation finds the intersection point $\mathbf{r} = \mathbf{b} + t\mathbf{d}$ with the minimum positive distance t for every surface in the DT. If this point lies within the surface's finite boundaries, the method returns the index of that surface.

The function $\mathbf{r} = \texttt{IncidencePoint}(\mathbf{u}, \mathbf{b}, i)$ calculates the exact 3D coordinates of a specular reflection on a given surface S_i. It takes as input the positions of the UE at $\mathbf{u}$ and the BS at $\mathbf{b}$. The method's principle is the image method, where it first computes the UE's virtual image, $\mathbf{u}'$, with respect to the plane of surface S_i

$$\mathbf{u}' = \mathbf{u} - 2((\mathbf{u} - \mathbf{p}_i)^\top \mathbf{n}_i)\mathbf{n}_i. \tag{4}$$

The reflection point $\mathbf{r}$ is then found at the intersection of the line segment connecting the BS and this virtual image with the plane of S_i. This is computed by solving for the intersection parameter of the line connecting $\mathbf{b}$ and $\mathbf{u}'$

$$\mathbf{r} = \mathbf{b} + \frac{(\mathbf{p}_i - \mathbf{b})^\top \mathbf{n}_i}{(\mathbf{u}' - \mathbf{b})^\top \mathbf{n}_i}(\mathbf{u}' - \mathbf{b}). \tag{5}$$

The function outputs the coordinates of this point, provided that $\mathbf{r}$ falls within the finite physical boundaries of the surface.

D. Problem Formulation

The objective of this work is to jointly estimate the UE position $\mathbf{u}$ and the clock bias B, from the observations $\{\mathbf{y}[k]\}_{k=0}^{K-1}$, leveraging the DT. Note that without the DT, the UE cannot be localized. Hence, the DT must be used to identify single-bounce NLOS paths and identify them to known surfaces, to render the localization problem identifiable.

III. DT-aided Localization Method

To estimate the UE position $\mathbf{u}$ and clock bias B in the presence of noisy measurements and multi-bounce NLOS paths, we propose a robust, three-stage methodology. First, we perform a candidate generation step to establish preliminary measurement-to-surface associations. In this stage, we generate a set of path hypotheses $\{h_\ell\}$, where each hypothesis h_ℓ is a tuple (ℓ, i) that pairs a measurement index ℓ with a hypothesized geometric origin index i (i.e., surface S_i) of the DT. Second, we apply the RANSAC algorithm to this expanded set of hypotheses to robustly identify a consensus set corresponding to the true LOS and single-bounce paths (inliers), while filtering out those from unmodeled multi-bounce paths (outliers). Third, we perform a refined optimization using the maximum likelihood criterion using only the final inlier set to obtain accurate estimates of $\mathbf{u}$ and B. These three stages are elaborated in the following sections.

A. Path Association

For any given path measurement $\ell \in \{0, \ldots, L\}$, we generate a hypothesis for its origin (i.e., the associated surface in the DT). Leveraging the strong prior that the LOS path exhibits the minimum travel time, we identify the path with the minimum measured delay, $\ell^* = \arg\min_\ell \hat{\tau}_\ell$, and assign it the LOS hypothesis, $h_{\ell^*} = 0$. For all other paths, the goal is to identify the most likely reflecting surface from the set $\{S_1, \ldots, S_M\}$ to assign an NLOS hypothesis, h_ℓ. The association process leverages the interface of the DT. To account for the measurement noise, for each NLOS path ℓ, we generate a set of K perturbed AOA-based directions (denoted by $\{\mathbf{d}_\ell^{(k)}\}_{k=1}^K$) by sampling from a Gaussian distribution centered at the measured AOA as $\theta_\ell^{(k)} \sim \mathcal{N}(\hat{\theta}_\ell, \sigma^2_\theta)$ and $\phi_\ell^{(k)} \sim \mathcal{N}(\hat{\phi}_\ell, \sigma^2_\phi)$, leading to $\mathbf{d}_\ell^{(k)} = [\sin\phi_\ell^{(k)} \cos\theta_\ell^{(k)}, \sin\phi_\ell^{(k)} \sin\theta_\ell^{(k)}, \cos\phi_\ell^{(k)}]^\top$. We then query the DT with this direction as $i_\ell^{(k)} = \texttt{FindSurface}(\mathbf{d}_\ell^{(k)}, \mathbf{b})$, which returns the index of the intersected surface. We then assign a score to each surface

$$\text{Score}_\ell(i) = \frac{1}{K} \sum_{k=1}^{K} \mathbb{I}(i_\ell^{(k)} = i), \tag{6}$$

where $\mathbb{I}(\cdot)$ is the indicator function. The surface with the highest score is selected as the associated surface,

$$h_\ell = \arg\max_i \text{Score}_\ell(i) \tag{7}$$

[1]More generally, a Von Mises model could be used.

$$\text{s.t. Score}_\ell(i) > \gamma,$$

where γ is a threshold, to filter out paths that lack a clear geometric origin.

B. Path Classification Using RANSAC

The second stage of our methodology employs the RANSAC algorithm to robustly classify the initial path hypotheses. It iteratively generates candidate states $(\mathbf{u}, B)$ from minimal subsets of measurements and selects the state that is consistent with the largest consensus set of inliers (LOS and single-bounce paths). Multi-bounce paths, which are inconsistent with the model, are discarded as outliers.

The core of this classification process is the residual cost function, $f_\ell(\mathbf{u}, B, h_\ell)$, which quantifies the agreement between a measurement and a given hypothesis. Assuming independent Gaussian noise, this cost is derived from the negative log-likelihood

$$f_\ell(\mathbf{u}, B, h_\ell) \triangleq \frac{1}{\sigma^2_{\tau_\ell}}\Big(\tilde{\tau}_\ell - \frac{d_\ell(\mathbf{u}, h_\ell)}{c} - B\Big)^2 + \frac{1}{\sigma^2_{\theta_\ell}}\Big(\tilde{\theta}_\ell - \theta^{\text{pred}}_\ell(\mathbf{u}, h_\ell)\Big)^2 + \frac{1}{\sigma^2_{\phi_\ell}}\Big(\tilde{\phi}_\ell - \phi^{\text{pred}}_\ell(\mathbf{u}, h_\ell)\Big)^2. \tag{8}$$

The prediction functions $d_\ell(\mathbf{u}, h_\ell)$ and $(\theta^{\text{pred}}_\ell, \phi^{\text{pred}}_\ell)$ within this cost model compute ideal path length and AOA based on the hypothesis h_ℓ. The calculation relies on the DT's `IncidencePoint` function to determine the reflection point for NLOS paths, while LOS parameters are computed directly.

The RANSAC-based classification [10] proceeds as follows:

- *Random sampling*: A minimal subset $\mathcal{S}$ from L paths is randomly selected. We set $|\mathcal{S}| = 2$ to form a solvable system for the four unknown state variables $(\mathbf{u}, B)$.
- *Model estimation*: For the selected hypothesis, a candidate state $(\mathbf{u}, B)$ is computed by minimizing the cost function over the minimal set $\mathcal{S}$:
$$\min_{\mathbf{u},B} \sum_{\ell \in \mathcal{S}} f_\ell(\mathbf{u}, B, h_\ell), \tag{9}$$
subject to geometric constraints, where the cost function f_ℓ is defined as in (8).
- *Inlier Evaluation*: The estimated state $(\mathbf{u}, B)$ is used to find the consensus set by re-evaluating all paths, $\ell \in \{0, \ldots, L\}$. Each path is individually classified as an inlier if its cost, $f_\ell(\mathbf{u}, B, h_\ell)$, does not exceed a predefined threshold T. This step ensures that unmodeled paths, such as multi-bounce reflections, are correctly rejected as outliers due to their typically large cost.
- *Iteration and selection*: The process is repeated for N iterations, where N is calculated to ensure a high probability (p) of selecting an outlier-free subset by $N = \log(1-p)/\log(1-(1-\epsilon)^s)$ [16], with ϵ as the expected outlier ratio. The model $(\mathbf{u}, B)$ that yields the largest set of inliers, denoted $\mathcal{L}_{\text{inlier}}$, is chosen.

C. Final Optimization

Given the final inlier set $\mathcal{L}_{\text{inlier}}$ and their confirmed path types (LOS or NLOS with a specific surface), we refine the estimates of $\mathbf{u}$ and B by solving the full maximum likelihood problem over all inliers:

$$\min_{\mathbf{u},B} \sum_{\ell \in \mathcal{L}_{\text{inlier}}} f_\ell(\mathbf{u}, B, h_\ell), \tag{10}$$

subject to the position bounds $x_u, y_u \in [x_{\min}, x_{\max}], z_u \in [0, z_{\max}]$. The reflection point $\mathbf{r}_\ell$ for each NLOS path is implicitly determined by the candidate position $\mathbf{u}$ and its known reflecting surface. This represents a standard non-linear least squares problem subject to box constraints; therefore, the L-BFGS-B solver is adopted in the simulation.

To ensure robust and fast convergence, a well-chosen initial point is critical. Instead of constructing a separate geometric initial point, we leverage the output from the RANSAC stage itself. The RANSAC algorithm inherently identifies the candidate state $(\mathbf{u}, B)$ that is consistent with the largest consensus set. This state, having been generated from a minimal (and likely outlier-free) sample and validated by the maximum number of inliers, serves as an excellent and robust initial estimate. Therefore, we initialize this final optimization using the state vector obtained from the winning RANSAC iteration.

IV. Simulation Results

A. Simulation Setup and Environment Configuration

The performance of the proposed algorithm is evaluated through simulations. Unless otherwise specified, the key parameters are set as follows: the receiver employs an 8×8 antenna array with an element spacing of half a wavelength, the number of OFDM subcarriers is 512, the subcarrier spacing is 30 kHz, the center frequency is 3.5 GHz, and the receiver noise figure is 7 dB. These parameters are crucial as they determine the theoretical estimation accuracy of the channel parameters. We leverage the Cramér-Rao lower bound (CRLB) to model this relationship; specifically, the measurement noise variances $(\sigma^2_{\tau_\ell}, \sigma^2_{\theta_\ell}, \sigma^2_{\phi_\ell})$ used in our simulations are set according to the CRLB, although the detailed derivation is omitted for brevity.

The simulation scenario is set within a specific urban canyon-like environment, geometrically represented by a DT model that includes three primary reflecting surfaces. This DT model serves as the a priori environmental knowledge for our localization algorithm.

Separately, to generate the ground truth for our simulations, we consider the signal propagation within this same environment. For the specific placement of the BS at $(0, 0, 15)$ m and the UE at $(-15, -15, 0)$ m, the propagation results in seven dominant multipath components. These ground-truth paths, used as the basis for generating noisy measurements, comprise one LOS path, three single-bounce NLOS paths whose reflections correspond geometrically to the surfaces defined in the DT, and three multi-bounce NLOS paths.

B. Performance Metric

We evaluate the performance of the proposed algorithm at different stages using several key metrics. First, for the initial path hypothesis generation stage, we use the correct association rate, defined as the percentage of true single-bounce

NLOS paths that are correctly associated with their ground-truth reflecting surface. Subsequently, to assess the RANSAC classification stage, we employ two standard metrics: the false alarm (FA) Rate, which measures the proportion of multi-bounce NLOS paths (true outliers) incorrectly classified as inliers, and the miss detection (MD) Rate, which measures the proportion of true LOS or single-bounce NLOS paths (true inliers) incorrectly classified as outliers. All these rates are averaged over N_{MC} Monte Carlo runs, where each run involves generating a new random realization of the measurement noise (n_{τ_ℓ} and $\mathbf{n}_{\boldsymbol{\theta}_\ell}$) added to the true path parameters for all seven paths. Finally, the overall localization accuracy is evaluated using the root mean square error (RMSE), which quantifies the Euclidean distance between the estimated UE position $\hat{\mathbf{u}}$ and the ground-truth position $\mathbf{u}$, also averaged over the runs:

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{MC}}} \sum_{n=1}^{N_{\text{MC}}} \|\hat{\mathbf{u}}^{(n)} - \mathbf{u}\|^2}, \tag{11}$$

where $\hat{\mathbf{u}}^{(n)}$ is the position estimate from the n-th run. Unless otherwise specified, all metrics presented below are averaged over 500 Monte Carlo runs.

C. Path Association Performance

Fig. 2 evaluates the performance of the path association stage by plotting the correct association rate for NLOS paths versus transmit power.

As expected, the correct association rate improves with higher transmit power. This is because increased power leads to lower measurement noise variance ($\sigma_\theta^2, \sigma_\phi^2$), which in turn causes the $K = 1000$ perturbed rays in our probabilistic casting method to be more tightly clustered around the true direction of arrival. This increases the likelihood that the score for the correct surface will be the maximum and will exceed the confidence threshold of $\gamma = 0.7$. It is important to note that this evaluation focuses exclusively on the performance for NLOS paths. The LOS path is not part of this test, as it is handled in a separate, deterministic step by identifying the path with the minimum delay.[2] The figure shows that when the transmit power exceeds -20 dBm, the correct association rate for all true NLOS paths approaches 100%, demonstrating the robustness of our proposed method.

D. RANSAC Performance

First, we evaluate the filtering performance of the RANSAC algorithm using FA and MD rates.

Fig. 3 illustrates the FA rate of the RANSAC classification stage. While FA are observed, particularly at lower transmit power levels, the FA rate decreases rapidly as power increases,

[2]We acknowledge that the minimum-delay prior can fail under very low SNR conditions, where a true NLOS path might be misidentified as the LOS candidate. Our methodology is inherently robust to this potential misclassification. The subsequent RANSAC stage employs a uniform random sampling strategy, which ensures a non-zero probability of selecting a minimal subset composed exclusively of true single-bounce NLOS paths. A model generated from such a 'pure' sample will be highly accurate and will, in turn, correctly identify the mislabeled minimum-delay path as a high-cost outlier. Therefore, the overall algorithm's success does not depend on the initial LOS hypothesis being correct.

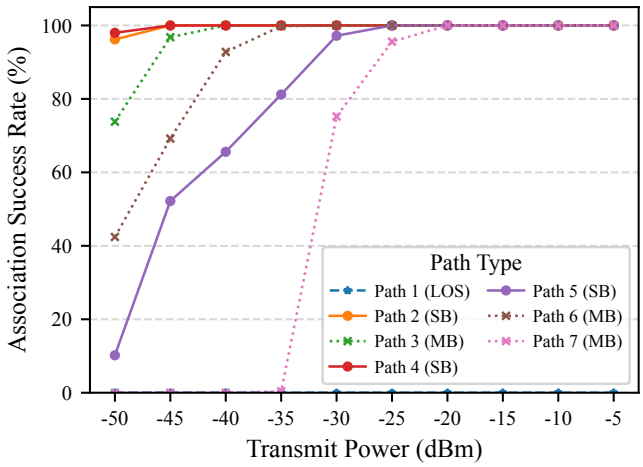

Fig. 2. Path association success rate versus transmit power (SB: Single-Bounce, MB: Multi-Bounce)

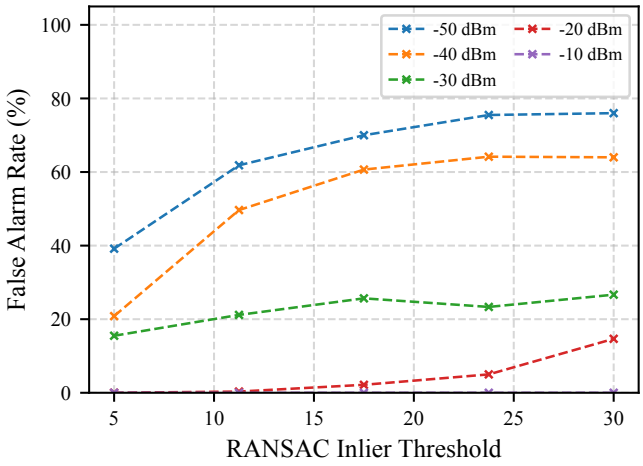

Fig. 3. FA of RANSAC with different threshold versus transmit power

demonstrating the algorithm's effectiveness in rejecting outliers under reasonable signal conditions.

This behavior stems from the interplay between geometric inconsistency and measurement uncertainty. At very low transmit power, the large measurement uncertainty (high noise variance) can obscure the inherent geometric mismatch between a multi-bounce path and the single-bounce model in our cost function. This allows an outlier to fortuitously appear consistent with a candidate model, resulting in a false alarm. However, as transmit power increases, the measurement uncertainty diminishes significantly. The underlying geometric mismatch then dominates, leading to a high residual cost for multi-bounce paths and their correct rejection as outliers. Consequently, the algorithm achieves a very low FA rate once sufficient transmit power is available.

Fig. 4 illustrates the MD rate. Counter-intuitively, the results show that the MD rate begins to increase as the transmit power becomes very high. This phenomenon can be attributed to the cost function's hypersensitivity to model inaccuracies at low noise levels.

The cost, f_ℓ, is a squared error normalized by the measurement variance ($\sigma_{\tau_\ell}^2, \sigma_{\theta_\ell}^2, \sigma_{\phi_\ell}^2$), which shrinks towards zero as

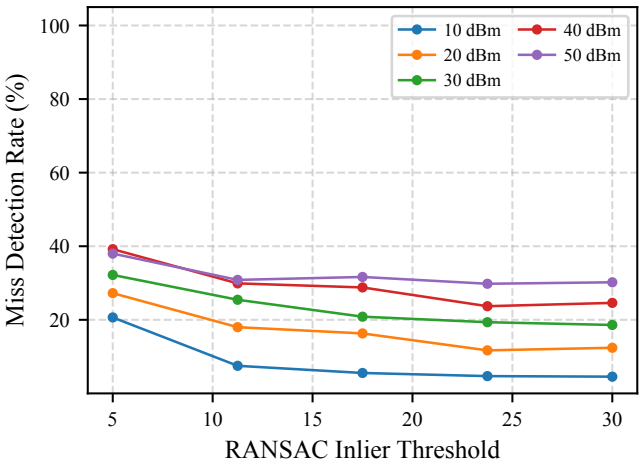

Fig. 4. MD of RANSAC with different threshold versus transmit power

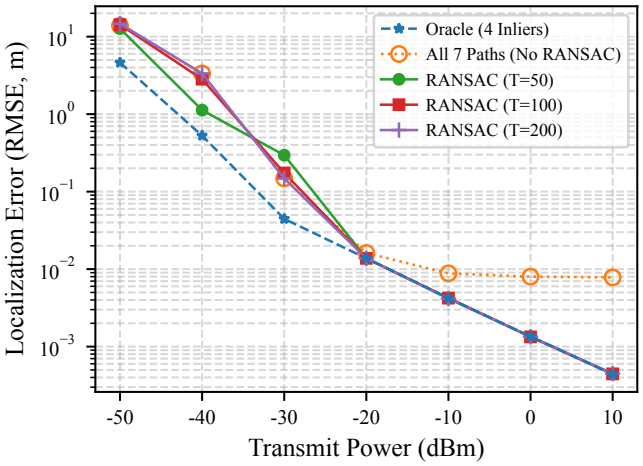

Fig. 5. RMSE with different threshold RANSAC versus transmit power

transmit power increases. In any given RANSAC iteration, the candidate state $(\mathbf{u}, B)$ is estimated from a minimal, random sample and is therefore only an approximation of the true state. Consequently, when a true inlier is evaluated against this slightly imperfect model, the small but non-zero geometric mismatch between the measurement and the prediction results in an extremely large normalized cost f_ℓ. This "cost explosion" causes the value to exceed the fixed inlier threshold T, leading to the erroneous rejection of a true inlier. As expected, the figure also confirms that a higher threshold T mitigates this effect by providing a larger acceptance region.

E. Localization Performance

Finally, we evaluate the localization performance under various RANSAC thresholds and transmission powers, using the RMSE as the metric.

As shown in Fig. 5, the localization error generally decreases as transmit power increases. Compared to the oracle performance using only true inliers ('Perfect Inlier'), simply using all available paths without outlier rejection ('All Paths') performs the worst, with its error saturating at a high value due to the bias introduced by multi-bounce outliers. In contrast, the RANSAC approach significantly improves accuracy, although performance varies with the threshold T. Notably, the localization performance is optimal across all tested thresholds within the transmit power range of -20 dBm to 10 dBm. This optimal window aligns well with the performance observed in the preceding stages: path association achieves near-perfect success, the FA rate effectively vanishes, and the MD rate remains low within this regime.

V. Conclusion

This paper presented a DT-assisted high-precision localization framework for massive MIMO systems operating in complex urban environments. By leveraging the detailed geometric knowledge provided by the DT, the proposed method effectively associates signal paths with physical surfaces and removes multi-bounce NLOS outliers. The final optimization over the inlier set enables accurate estimation of both user position and clock bias. Simulation results in a 3D scenario demonstrated the effectiveness of the RANSAC-based outlier rejection algorithm, achieving near-oracle accuracy with significantly reduced deployment complexity. Future work will extend this framework toward dynamic and large-scale environments, where the updating of DT will also be considered.

References

[1] M. Grieves, "Digital twin: Manufacturing excellence through virtual factory replication," in *White paper*, vol. 1, 2014, pp. 1–7.

[2] F. Tao *et al.*, "Digital twin in industry: State-of-the-art," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2018.

[3] L. U. Khan *et al.*, "Digital-twin-enabled 6G: Vision, architectural trends, and future directions," *IEEE Communications Magazine*, vol. 60, no. 1, pp. 70–76, 2022.

[4] P. Öhlén *et al.*, "Network digital twins – outlook and opportunities," *Ericsson Technology Review*, vol. 2022, no. 12, pp. 2–11, 2022.

[5] C. Ruah *et al.*, "Calibrating wireless ray tracing for digital twinning using local phase error estimates," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 1193–1215, 2024.

[6] H. Zhang *et al.*, "6G-enabled smart agriculture: A review and a vision," *Journal of Communications and Information Networks*, vol. 6, no. 4, pp. 363–380, 2021.

[7] D. Fox *et al.*, "Markov localization for mobile robots in dynamic environments," *J. Artif. Int. Res.*, vol. 11, no. 1, p. 391–427, Jul. 1999.

[8] V. Djaja-Josko *et al.*, "A new map based method for nlos mitigation in the UWB indoor localization system," in *2017 25th Telecommunication Forum (TELFOR)*, 2017, pp. 1–4.

[9] Y. Ge *et al.*, "5G SLAM using the clustering and assignment approach with diffuse multipath," *Sensors*, vol. 20, no. 16, p. 4656, 2020.

[10] O. Kaltiokallio *et al.*, "Robust snapshot radio SLAM," *IEEE Transactions on Vehicular Technology*, 2024.

[11] J. Morais *et al.*, "Localization in digital twin MIMO networks: A case for massive fingerprinting. arxiv 2024," *arXiv preprint arXiv:2403.09614*.

[12] H. Wang *et al.*, "Digital twin channel for 6G: Concepts, architectures and potential applications," *IEEE Communications Magazine*, vol. 63, no. 3, pp. 24–30, 2025.

[13] Z. Yun *et al.*, "Ray tracing for radio propagation modeling: Principles and applications," *IEEE Access*, vol. 3, pp. 1089–1100, 2015.

[14] Y. Pati *et al.*, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, 1993, pp. 40–44 vol.1.

[15] T.-J. Shan *et al.*, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806–811, 1985.

[16] M. A. Fischler *et al.*, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, p. 381–395, Jun. 1981.