

Point Cloud for Wireless Communications: A Comprehensive Survey on Fundamentals, Techniques, and Applications

Yu Zheng, Kezhi Wang, *Senior Member, IEEE*, Zhicheng Du, Wenji Xi, Gang Yu, Jinbo Hou, Haonan Hu, *Member, IEEE*, Jie Zhang, *Senior Member, IEEE*

Abstract—Point cloud has recently emerged as a promising technology for next-generation wireless communication systems, offering precise geometric representations of complex environments. Despite growing interest, a comprehensive survey on its integration with wireless communications remains lacking. To bridge this gap, this survey provides a thorough review of state-of-the-art methods and applications of the point cloud across various wireless communication tasks over the last six years. We first introduce the fundamentals of point cloud for wireless communications, including acquisition techniques, preprocessing and representation methods, as well as open datasets and simulation tools. Subsequently, we investigate how point cloud effectively enhances key wireless communication tasks such as propagation channel modeling, blockage management, beam management, and proactive handoff. Additionally, emerging applications are extensively reviewed, including point cloud-enabled semantic wireless communications, wireless localization, metasurface coverage optimization, and digital-twin systems. Alongside these opportunities, we identify key challenges such as real-time processing of large-scale point clouds, robustness to environmental variations, multimodal fusion and calibration, generalization across diverse scenarios, the extension from single-user to multi-user wireless settings, the need for standardized benchmarking and open datasets, and associated privacy and security risks. We also outline promising research directions to guide the development of point cloud-enabled wireless networks.

Index Terms—Point cloud, wireless communications, propagation channel modeling, blockage management, beam management, proactive handoff, integrated sensing and communication, semantic communications, digital twin.

Yu Zheng, Wenji Xi, Gang Yu, and Jinbo Hou are with the School of Electrical and Electronic Engineering, The University of Sheffield, Sheffield S10 2TN, U.K. (e-mail: {yzheng120, wxi4, gyu2, jhou9}@sheffield.ac.uk).

Kezhi Wang is with the Department of Computer Science, Brunel University of London, UB8 3PH, U.K. (e-mail: kezhi.wang@brunel.ac.uk).

Zhicheng Du is with Tsinghua Shenzhen International Graduate School, Shenzhen 518055, China (e-mail: duzc21@mails.tsinghua.edu.cn).

Haonan Hu is with the School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China. (email: huhn@cqupt.edu.cn).

Jie Zhang is with the R&D Department, Ranplan Wireless Network Design Ltd., Cambridge CB23 3UY, U.K. (e-mail: Jie.Zhang@ranplanwireless.com) and also with the School of Information Science and Technology, Harbin Institute of Technology, Shenzhen 518055, China.

This work was partly supported by Innovate UK COMET (10099265) and i2D-MSW (10071278) projects, the Horizon Europe COVER project (No. 101086228 with funding from UKRI grant EP/Y028031/1), and the National Key R&D Program of China (with funding from Grant no. 2024YFE0200101). K. Wang would like to acknowledge the support in part by the Royal Society Industry Fellowship (IF/R2/23200104).

Kezhi Wang and Jie Zhang are corresponding authors.

I. INTRODUCTION

Wireless communication underpins almost every aspect of modern society, from video streaming and online collaboration to autonomous driving and industrial automation [1]. Driven by surging data traffic, massive device connectivity, and increasingly complex application scenarios, modern wireless networks need to continuously evolve towards higher data rates, lower latency, and improved reliability [2]. Meeting these multifaceted demands introduces numerous technical challenges, including accurately characterizing dynamic propagation environments, efficiently managing real-time channel variations and blockage effects, performing precise beam management, and executing proactive handoffs.

To address these challenges, the three-dimensional (3D) point cloud has emerged as a powerful sensing modality for accurately capturing the geometric structure of physical environments. It represents physical environments as collections of discrete 3D points, with each point specified by a coordinate triplet (x, y, z) of finite precision and dynamic range [3]. By encoding fine-grained geometric features of objects and surfaces, point clouds allow accurate reconstruction of structural elements, such as walls, vehicles, and other scatterers, that critically impact electromagnetic wave propagation. Typically acquired through advanced sensing techniques such as photogrammetry, light detection and ranging (LiDAR), structured light scanner, or radar, point clouds support high-resolution mapping in complex scenarios. Their capability to model scene geometry, detect obstructions, and capture dynamic spatial changes has drawn growing interest from the wireless communications community.

In this survey, we define point-cloud-aided wireless communication as the integration of point cloud information into wireless communication systems, where such information is used as environment-aware spatial context to assist communication-related tasks. It encompasses applications such as propagation channel modeling, blockage detection and prediction, beam management, proactive handoff, and other emerging wireless functions. With access to detailed geometric context, wireless communication systems can better predict channel variations, adapt beam directions, and maintain stable links under various conditions. Furthermore, advanced applications, such as semantic wireless communication, wireless localization, metasurface coverage optimization, and digital-twin systems, can all significantly benefit from rich spatial representations

enabled by point clouds.

Motivated by the substantial potential and growing interest in point cloud technology, this article presents a comprehensive survey of point cloud-aided wireless communication over the last six years. We first introduce the fundamentals of point cloud for wireless communications, including acquisition techniques, preprocessing, and representation methods, as well as open datasets and simulation tools. We then examine how point clouds enhance key wireless communication tasks, including propagation channel modeling, blockage management, beam management, and proactive handoff. Furthermore, emerging use cases beyond traditional paradigms are also discussed. Finally, we outline key challenges and future research directions, aiming to provide a clear, structured, and comprehensive foundation for further research in point-cloud-aided wireless communications.

A. Related Work

Despite growing interest in sensing-aided wireless communication systems, there is currently no dedicated survey that systematically examines the comprehensive role of the point cloud in wireless communications. While several recent surveys [4]–[8] in integrated sensing and communication (ISAC) and environment awareness have discussed the integration of LiDAR sensing within wireless communications, they only occasionally refer to point cloud sensing as one possible modality, without a focused analysis of its unique capabilities or broader impact on key wireless communication tasks. Similarly, surveys [9]–[11] on digital twin systems emphasize the role of the point cloud in the context of creating accurate virtual models for wireless network optimization and simulation. However, they fall short of detailing how point clouds contribute to wireless performance enhancement throughout the digital twin lifecycle. In addition, channel modeling surveys [12], [13] have acknowledged the point cloud as a valuable source of environmental geometry, but do not extend their scope to key wireless communication tasks such as blockage management, beam management, or proactive network handoffs.

Distinct from these works, this survey provides an explicit and comprehensive overview of the application of the point cloud in wireless communications. It covers key acquisition techniques, their integration into key wireless communication tasks, and their emerging applications. Furthermore, this survey identifies key technical challenges and proposes future research directions, offering a systematic roadmap for advancing point cloud-aided wireless systems. A detailed comparison of related surveys is presented in Table II.

B. Contribution

This survey makes four key contributions to the emerging field of point cloud-aided wireless communications. First, it provides a systematic overview of point cloud fundamentals for wireless communications, including acquisition techniques, preprocessing and representation methods, as well as open-source datasets and simulation tools. Second, it establishes a unified task-oriented framework for integrating

Acronyms	Full Form
2D	Two Dimensional
3D	Three Dimensional
ANN	Artificial Neural Networks
AoD	Angle of Departure
BS	Base Station
CAD	Computer-Aided Design
CNN	Convolutional Neural Network
ConvLSTM	Convolutional Long Short-Term Memory
CSI	Channel State Information
dToF	direct Time-of-Flight
FMCW	Frequency-Modulated Continuous Wave
GBDT	Gradient Boosting Decision Trees
GNN	Graph Neural Network
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
iToF	indirect Time-of-Flight
ISAC	Integrated Sensing and Communication
JPDA	Joint Probabilistic Data Association
LiDAR	Light Detection And Ranging
LOS	Line-of-Sight
LSTM	Long Short-Term Memory
mmWave	millimeter-wave
NLOS	Non-Line-of-Sight
RANSAC	Random Sample Consensus
RAT	Radio Access Technology
RF	Radio Frequency
RGB	Red Green Blue
RIS	Reconfigurable Intelligent Surface
RNN	Recurrent Neural Network
RMSE	Root Mean Square Errors
RoV	Region of View
RSS	Received Signal Strength
RSSI	Received Signal Strength Indication
RSU	Road-Side Unit
Rx	Receive
SFM	Structure From Motion
SNR	Signal-to-Noise Ratio
THz	Terahertz
ToF	Time of Flight
Tx	Transmit
UAV	Unmanned Aerial Vehicle
UE	User Equipment
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle

TABLE I: List of acronyms

point cloud into key wireless communication tasks. Specifically, it classifies existing studies on propagation channel modeling into deterministic, stochastic, and machine-learning-aided approaches, and further organizes point-cloud-aided methods for blockage management, beam management, and proactive handoff under a coherent taxonomy. In this way, the survey clarifies the distinct roles of point cloud across different wireless tasks and provides a unified view

Work	Type	Year	Covered Content of Point Cloud for Wireless Communication				
			Propagation Channel Modeling	Blockage Management	Beam Management	Proactive Handoff	Digital Twin
[4]	Survey	2024	✓				
[5]	Journal	2022			✓		
[6]	Survey	2023			✓		
[7]	Survey	2024	✓		✓		
[8]	Tutorial	2024	✓		✓		
[9]	Magazine	2025	✓				✓
[10]	Magazine	2023					✓
[11]	Journal	2025					✓
[12]	Journal	2024	✓				
[13]	Survey	2025	✓				
Our Survey	Survey	2025	✓	✓	✓	✓	✓

TABLE II: Comparison of existing related works

of geometry-aware wireless system design. Third, it offers a comprehensive review of emerging applications beyond these tasks, such as point cloud-driven semantic wireless communications, wireless localization, metasurface coverage optimization, and digital-twin systems, thereby illuminating new research frontiers. Fourth, it offers a structured and forward-looking discussion of open challenges and future research directions. In particular, it identifies the main bottlenecks that currently limit the practical deployment of point-cloud-aided wireless communications, including real-time processing of large-scale point clouds, robustness to environmental variations, multimodal fusion and calibration, generalization across diverse scenarios, the extension from single-user to multi-user wireless settings, the lack of standardized benchmarking and open datasets, and associated privacy and security risks. It further outlines corresponding directions for advancing more reliable, scalable, and practically deployable environment-aware wireless systems. Together, these contributions offer a structured roadmap for harnessing point cloud in the design and optimization of next-generation wireless communication systems.

C. Structure of the Survey

This survey is organized as follows. Section II introduces the fundamentals of the point cloud, including point cloud acquisition methods, point cloud preprocessing and representation, as well as open-source datasets and simulation tools. Section III reviews how point cloud is used for propagation channel modeling, covering deterministic, stochastic, and machine learning-aided approaches. Section IV details the use of the point cloud to manage blockage in wireless communication, with a focus on detection and prediction strategies. Section V presents point-cloud-aided beam management, including current beam selection, future beam prediction, and beamforming optimization. Section VI explores proactive handoff schemes leveraging point cloud for frequency, base-station, and multi-radio access technology (RAT) transitions. Section VII surveys other applications, including semantic communication, wireless localization, metasurface coverage optimization, and digital twin. Section VIII outlines key challenges and proposes

future research directions. Finally, Section IX concludes this survey. The main structure of the survey is shown in Fig. 1. Some of the acronyms used in this paper are summarized in Table I.

II. POINT CLOUD FUNDAMENTALS

In this section, we introduce the fundamental aspects of point clouds for wireless communication, including acquisition methods, preprocessing and representation techniques, as well as available datasets and simulation tools.

A. Point Cloud Acquisition Methods

In this section, we discuss and classify the methods to acquire the point cloud for general measurement, including photogrammetry, LiDAR, structured light scanner, and radar, based on the sensor type and working principle.

- Photogrammetry generates the point cloud through the analysis of overlapping photographs captured from multiple viewpoints. It relies primarily on structure from motion or stereo triangulation, requiring adequate illumination, texture-rich environments, and computationally intensive post-processing. Photogrammetry provides high-resolution spatial data and is cost-effective, making it suitable for large-scale static scene reconstruction, such as urban terrain modeling [14].
- LiDAR actively measures distances by emitting laser pulses and detecting their reflections, with point coordinates precisely calculated based on the time-of-flight (ToF) principle. It is categorized into direct ToF (dToF), measuring direct pulse reflection time, and indirect ToF (iToF), measuring phase delays of modulated laser signals. LiDAR systems offer high spatial resolution and robust performance under varying lighting conditions [15]. However, they may face limitations due to environmental conditions such as fog or heavy rain [16] and involve higher equipment costs. LiDAR is particularly effective in autonomous driving, urban mapping, and detailed indoor environment scanning.

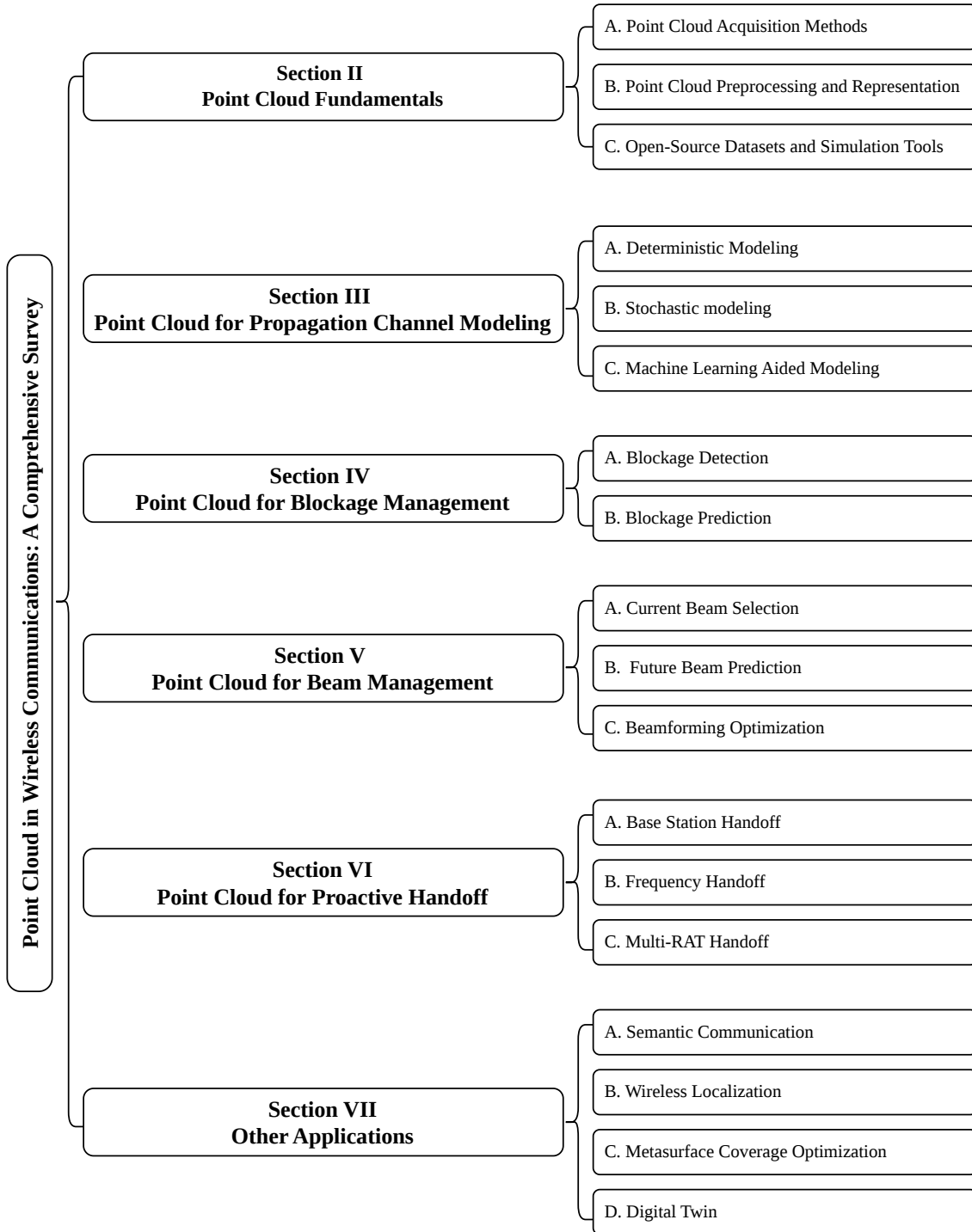


Fig. 1: Structure of this survey

- Structured light scanner projects a known infrared pattern onto a scene and captures its deformation with a calibrated camera. It offers several advantages for point cloud acquisition, including high spatial resolution (typically centimeter to millimeter scale), relatively low hardware cost, and suitability for real-time depth imaging at frame rates exceeding 30 Hz. However, structured light scanner exhibits limited effectiveness under strong ambient lighting and highly reflective or translucent surfaces, making it most suitable for indoor and controlled-environment scenarios such as robotics navigation, human-computer interaction, and precise indoor 3D reconstruction tasks [17].
- Radar generates point clouds by analyzing the time delay and frequency characteristics of reflected radio frequency (RF) signals. Based on its ranging principle, radar can be categorized into pulsed radar, which directly measures the round-trip travel time of a transmitted pulse, and frequency-modulated continuous-wave (FMCW) radar, which indirectly derives range from the instantaneous frequency difference between the transmitted chirp and the received echo. Radar is characterized by its superior capability to penetrate adverse weather conditions such as rain and fog, maintaining effective performance at considerable distances [18]. It is ideal for applications such as vehicular sensing, obstacle detection in poor visibility conditions, and geospatial remote sensing. However, radar has relatively lower spatial resolution compared to optical methods, limiting its precision in detailed object characterization [19].

The concise comparison of these methods is provided in Table III.

1) *Photogrammetry*: Photogrammetry leverages multiple photographs or digital images taken from different viewpoints to identify matching points and perform 3D reconstruction, allowing precise extraction of an object's geometry, displacement, and deformation [20]. In the following section, we explore how to apply photogrammetric methods to produce a high-precision point cloud.

a) *Monocular Camera Vision System*: A monocular camera can only capture 2D RGB images. However, the point cloud is a 3D representation of the environment, which means each point has a 3D coordinate. Therefore, the generation of a point cloud from a monocular camera involves the conversion of 2D image data into a 3D spatial representation. The process begins with depth estimation. For instance, Xu *et al.* utilized the technique of structure from motion (SFM) to recover depth information and camera poses by analyzing the motion between images captured from different angles in [21]. The methodology outlined includes feature point extraction and matching from these images, forming a sparse reconstruction of the scene by identifying key points. These matched feature points are then used to estimate the 3D coordinates, resulting in an initial sparse point cloud. The camera's position and orientation are determined through geometric relations, such as vanishing points in the images. To enhance the accuracy of the 3D reconstruction and form a denser point cloud, digital image processing methods like edge detection, Hough transform, and

random sample consensus (RANSAC) algorithms are applied to identify and classify lines and points in the image. These geometric features guide the calculation of the camera's pose and enable the generation of a dense point cloud.

In addition to structure and motion-based methods, a deep learning model, typically a convolutional neural network (CNN), is employed to estimate depth directly from the 2D image and form the point cloud as demonstrated in [22] and [23]. The network predicts the depth for each pixel in the image, providing the necessary spatial information. Once the depth map is obtained, the 2D-pixel coordinates and their depth values are converted into 3D coordinates using the camera's intrinsic parameters, such as focal length and the principal point. Specifically, given a pixel's coordinates (u, v) and its estimated depth d , the 3D coordinates (x, y, z) are calculated using the equations:

$$\begin{cases} x = \frac{(u - P_x)z}{f} \\ y = \frac{(v - P_y)z}{f} \\ z = d \end{cases}, \quad (1)$$

where P_x and P_y represent the principal point, and f is the focal length. This transformation projects the pixels from the 2D image space into the 3D camera coordinate space, resulting in a point cloud that represents the scene in 3D. To enhance the point cloud's discriminative power, RGB information from the image can also be incorporated, enriching the 3D representation with colour features. However, as depth information is derived either through deep learning techniques or multi-angle approaches, the accuracy is contingent upon the specific machine learning model and algorithm employed. The characteristics of the camera, like resolution, would greatly influence the accuracy and limit the measurement distance.

b) *Multi-Camera Vision System*: Unlike the monocular camera vision system, which only uses a single camera to capture information and form the point cloud by deep learning or SFM, a multi-camera vision system, such as the binocular camera vision system in [24], the trinocular camera vision system in [25], and the quadrinocular camera vision system in [26], consists of multiple cameras to simultaneously capture two or more slightly different images of the same scene. The depth information can be directly computed based on the disparity between them, which is called triangulation. For instance, the vision system proposed in [27] employs a binocular setup to create the point cloud. The process begins with local and global calibration. Local calibration determines camera-specific intrinsic and extrinsic parameters, while global calibration establishes a unified coordinate frame to enable coherent multi-camera data processing. Once calibrated, the system uses triangulation to generate 3D point cloud from the disparity between corresponding pixels in the stereo images, resulting in a raw local point cloud. In detail, the process of triangulation typically starts with stereo rectification, which places corresponding points on a common horizontal plane to facilitate correspondence matching. Next, the disparity in the position of a point between the left and right images is

calculated, which is crucial for inferring the depth of the point relative to the cameras.

The homogeneous 3D coordinates are then calculated by the general triangulation formula [27]:

$$[x, y, z, w]^T = Q[u, v, d, 1]^T \quad (2)$$

where x, y, z, w are the homogeneous coordinates of the physical point relative to the camera coordinate system, Q is the projection matrix derived from the camera's intrinsic parameters (such as focal length and principal point) and extrinsic parameters (such as the translation and rotation between the left and right cameras), u, v are the image coordinates of the pixel, d is the disparity between the left and right images.

These homogeneous coordinates are normalized to convert the coordinates into real-world 3D points, resulting in a dense 3D point cloud that represents the scene. These raw point clouds are further subjected to a series of filtering techniques to improve the overall quality of the data, such as voxel, statistical, and radius filters, to remove noise and outliers. The filtered local point cloud from different cameras is then stitched together using transformation matrices derived during the global calibration process, aligning the points in a global coordinate system.

2) *LiDAR*: A LiDAR system typically operates by actively emitting electromagnetic radiation, like a laser, and subsequently detecting the reflected signals. By analyzing the characteristics of the reflected light, the system calculates the distance between the sensor and the object to form the 3D point cloud. Based on its working principle, it can be broadly categorized into dToF and iToF.

a) *Direct Time of Flight*: A dToF sensor determines the distance to a target by measuring the time interval between the emission of a light pulse and the detection of its reflected echo. As demonstrated in [28], a vertical-cavity surface-emitting laser with a wavelength of 940 nm is used as the light source. The emitted pulses reflect off a target and are detected by an array of single-photon avalanche diodes, which are highly sensitive photodetectors to receive the returning photons. By measuring the round-trip time of an emitted laser pulse, the ranging system computes the distance to the target with high temporal resolution. The distance d can be calculated as [29]:

$$d = \frac{cT_{\text{ToF}}}{2}, \quad (3)$$

where c denotes the speed of emitted light and T_{ToF} is the round-trip time of photons. To further form the point cloud of the measured scene, an m beams system with n measurements taken per scan cycle is used in [30]. The point cloud from a single scan is organized into a spherical coordinate system where each point corresponds to a shared azimuth angle θ and a shared inclination angle ϕ with the distance d_{dToF} . These angles represent the relative horizontal and vertical positions of each detected point concerning the sensor's origin. The

commonly used point cloud with Cartesian coordinates is transformed as [30]:

$$\begin{cases} x = d_{\text{dToF}} \cos(\phi) \cos(\theta) \\ y = d_{\text{dToF}} \cos(\phi) \sin(\theta) \\ z = d_{\text{dToF}} \sin(\phi) \end{cases} \quad (4)$$

where x, y, z denotes the Cartesian coordinates of the point cloud. The process is illustrated in Fig. 2.

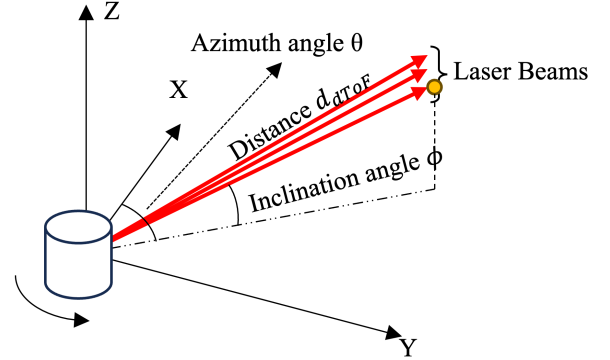


Fig. 2: Transform LiDAR coordinates to Cartesian coordinates

b) *Indirect Time of Flight*: An iToF sensor calculates distance by measuring the phase delay of the reflected light signal, rather than directly measuring the time interval between emission and reflection. As a representative example, consider the widely adopted four-phase clock demodulation scheme, in which the receiver employs four reference clocks with phase offsets of 0° , 90° , 180° , and 270° relative to the transmitted signal. These four clock signals are coherently combined with the received reflected light, with the results denoted as A_1 , A_2 , A_3 , and A_4 . The phase difference $\Delta\phi$ between emitted and reflected light can be estimated by the Four-Sampling-Bucket algorithm [31] [32]. The phase difference $\Delta\phi$, distance d_{iToF} is given by [33] respectively:

$$\Delta\phi = \arctan \left[\frac{A_1 - A_3}{A_0 - A_2} \right], \quad (5)$$

$$d_{\text{iToF}} = \frac{c \times \Delta\phi}{4\pi f_m}, \quad (6)$$

where f_m stands for the modulation frequency of the light. The point cloud Cartesian coordinates (x, y, z) are then computed using Eq. 4 by replacing d_{dToF} with d_{iToF} .

3) *Structured light scanner*: Structured Light is another active optical method for generating a point cloud, which obtains depth information by projecting known patterns of light onto the environment and analyzing the resulting distortions observed by one or more calibrated cameras [34]. This approach typically utilizes infrared light to avoid interference with visible illumination and achieves robust performance in controlled indoor scenarios.

The working principle of structured-light systems involves three key steps [35], [36]:

- 1) **Pattern Projection:** A pre-defined infrared pattern, commonly a grid of dots, stripes, or encoded light structures, is projected onto the target surfaces. This pattern is specifically designed to facilitate accurate and efficient depth extraction through subsequent decoding processes.
- 2) **Pattern Capture and Distortion Analysis:** An infrared-sensitive camera captures the distorted projection of the structured pattern from a viewpoint slightly offset from the projector. Due to variations in surface geometry and distances, the originally uniform pattern appears distorted, deformed, or displaced, allowing depth cues to be encoded within these observed deviations.
- 3) **Triangulation-based Depth Estimation:** Depth measurements are computed through triangulation, using the calibrated baseline distance between the projector and the camera, along with measured displacements of the structured pattern features. Specifically, the depth d_s at each pixel can be calculated from the known camera focal length f , projector-camera baseline distance B , and the measured disparity or pattern displacement Δd

$$d_s = \frac{f \cdot B}{\Delta d}. \quad (7)$$

Through pixel-wise triangulation, structured-light systems produce a dense point cloud that accurately represents the geometry of the scanned surfaces in three-dimensional space.

4) **Radar:** A radar system actively transmits an electromagnetic signal and detects the echo reflected from surrounding targets. By analyzing the time delay and frequency characteristics of the returned signal, the range and angular position of each target are estimated and a 3D point cloud is assembled. Following the same ranging-principle criterion used to classify LiDAR systems (dToF vs. iToF), radar can be broadly categorized into pulsed radar and Frequency-Modulated Continuous Wave(FMCW) radar.

a) **Pulsed Radar:** A pulsed radar estimates target range by directly measuring the round-trip travel time Δt of a short, high-power electromagnetic pulse, analogous to the dToF principle in LiDAR. The range R is given by [37]:

$$R = \frac{c \Delta t}{2}, \quad (8)$$

where the range resolution $\Delta R = c\tau/2$ is determined by the pulse width τ . To recover the angular position of each echo, a mechanically or electronically steered beam scans the spatial domain; each received pulse is associated with the beam's pointing direction (θ, ϕ) at the time of transmission. The 3D Cartesian coordinates are then computed via Eq. (4) by replacing d_{dToF} with R , yielding a full 3D point cloud. Owing to its high transmitted peak power, pulsed radar is capable of long-range detection and is widely employed in air-traffic control, weather sensing, and long-range geospatial mapping.

b) **Frequency-Modulated Continuous Wave Radar:** A FMCW radar continuously transmits a linearly frequency-swept (chirp) signal of bandwidth B over a sweep period T_c .

When this signal reflects off a target at range R , the received echo is delayed by $\tau = 2R/c$, producing a beat frequency $\Delta f = K\tau$ where $K = B/T_c$ is the chirp slope. The range is therefore derived as [38], [39]:

$$R = \frac{c \Delta f}{2K}. \quad (9)$$

Unlike pulsed radar, FMCW radar infers range indirectly from a frequency measurement rather than a direct time interval, directly analogous to the phase-delay principle of iToF LiDAR.

To resolve targets in the angular domain and generate a full 3D point cloud, FMCW radar systems typically employ a MIMO antenna configuration with N_T transmit and N_R receive elements [40]. The transmitters emit mutually orthogonal waveforms, which the receivers process jointly, synthesizing a virtual aperture of $N_T \times N_R$ elements. For a uniform linear virtual array with inter-element spacing d , the phase shift between adjacent virtual elements encodes the angle of arrival as $\psi = \frac{2\pi d}{\lambda} \sin \theta$. A spatial Discrete Fourier Transform (DFT) applied across the virtual aperture resolves the azimuth angle θ with angular resolution [41]:

$$\Delta \theta \approx \frac{\lambda}{N_T N_R d \cos \theta}. \quad (10)$$

Elevation angle ϕ is estimated analogously from the vertical dimension of the virtual array; high-resolution algorithms such as MUSIC or ESPRIT may further refine both estimates beyond the DFT resolution limit [40]. Combining the estimated (R, θ, ϕ) via Eq. (4) yields the 3D radar point cloud.

B. Point Cloud Preprocessing and Representation

Raw point clouds acquired from different sensing modalities are rarely used directly in wireless communication tasks. Due to measurement noise, varying point density, and redundancy, preprocessing is typically required to improve data quality and convert raw observations into task-suitable representations. In point cloud processing, such operations are commonly performed before downstream modeling or learning in order to balance geometric fidelity and computational efficiency [42], [43].

a) **Denoising and Outlier Removal:** Denoising and outlier removal are commonly applied to suppress measurement noise and eliminate isolated or spurious points. Techniques such as statistical filtering and radius-based filtering are widely used to improve the reliability of geometric structures, which is particularly important for LiDAR- and radar-based point clouds under adverse sensing conditions [42], [44]. In wireless applications, effective noise removal helps avoid false obstacles or reflectors in subsequent modeling. For example, prior studies on point-cloud-aided ray tracing and propagation simulation have shown that preprocessing noisy point clouds improves the agreement between simulations and measurements [45]–[47].

b) **Down-sampling and Structural Simplification:** Down-sampling and structural simplification are adopted to reduce computational complexity while preserving the essential geometry of the environment. Representative methods include voxel grid filtering and farthest point sampling [42], [48]. In

Feature	Photogrammetry	LiDAR (dToF)	LiDAR (iToF)	Structured Light Scanner	Radar (Pulsed)	Radar (FMCW)
Measurement Principle	Image triangulation	Direct pulse timing	Phase-delay measurement	Project coded IR and triangulation	Direct pulse timing	Beat frequency measurement
Spatial Resolution	High (cm-level)	Very high (mm–cm level)	High (cm-level)	Very High (mm–cm level)	Medium (dm–m level)	Medium (dm–m level)
Range	Far	Far	Medium	Near	Very far	Far
Environmental Robustness	Poor (light dependent)	Moderate (weather effects)	Moderate (weather effects)	Poor (light dependent)	Excellent (all-weather)	Excellent (all-weather)
Typical Application	Large-area mapping	Autonomous sensing	Indoor	Indoor	Surveillance and weather sensing	Automotive and UAV sensing
Cost	Low	High	Medium	Low	High	Medium

TABLE III: Comparative overview of point cloud acquisition methods

addition, surface fitting, plane extraction, and mesh reconstruction are often employed to obtain more compact and structured representations for wireless modeling [49]–[51]. However, excessive simplification may remove fine-grained geometric details that are critical for accurately characterizing blockage, reflection, and diffraction effects. This trade-off is also evident in propagation studies, where different levels of point cloud preprocessing lead to different levels of radio-map accuracy [46], [52].

c) Task-oriented Representation: Beyond basic preprocessing, point clouds are often transformed into task-oriented representations. Depending on the downstream application, raw point sets may be converted into voxel grids, range images, polygonal meshes, or semantically annotated structures. For example, geometry-aware channel modeling usually relies on surface-based or mesh representations with material or semantic information [50], [53], [54], whereas learning-based beam management and localization often adopt tensorized representations such as voxelized inputs, projected maps, or range images for efficient neural processing [55]–[57]. Therefore, the choice of representation directly affects how spatial and structural information is encoded and ultimately influences downstream wireless performance.

C. Open-Source Datasets and Simulation Tools

To facilitate reproducible research on point cloud-aided wireless communications, several open-source datasets and simulation tools have been developed. This subsection provides a consolidated overview of these resources, summarizing their available modalities, communication data types, target tasks, and scenarios.

1) Simulation Tools: Several simulation tools enable researchers to generate point cloud-aided wireless communication data for specific scenarios and configurations. These tools can be broadly categorized into ray-tracing engines for wireless channel simulation, 3D environment simulators for point cloud and sensor data generation, and mobility simulators for dynamic scenario construction.

For ray-tracing-based channel simulation, Wireless InSite [58] is a widely used commercial platform that accepts 3D geometric models, including those reconstructed from point clouds, as input for deterministic channel prediction. It has been adopted as the channel simulation backend in multiple open-source datasets, such as Raymobtime [59], M3SC [60], and SynthSoM [61]. Ranplan Professional [62], developed by Ranplan Wireless, is another commercial-grade ray-tracing and ray-launching simulator with a true 3D propagation engine. It supports both indoor and outdoor environments and enables the import of 3D building models from BIM/IFC files and LiDAR scans, making it particularly relevant for point cloud-aided indoor network planning and channel simulation. Sionna RT [63], developed by NVIDIA, is an open-source, GPU-accelerated, differentiable ray-tracing framework. It supports direct import of 3D scene representations for path loss and channel matrix computation, and its differentiable nature enables end-to-end optimization of communication system parameters jointly with scene geometry.

For 3D environment and sensor simulation, AirSim [64] provides high-fidelity LiDAR point cloud generation through physics-based rendering in Unreal Engine, and has been adopted by M3SC and SynthSoM for multi-modal sensory data collection. CARLA [65] is another open-source simulator originally designed for autonomous driving, which provides configurable LiDAR sensors and has been integrated with channel simulators such as MATLAB and Sionna for joint sensing–communication data generation. WaveFarer [66], also developed by Remcom, is a high-fidelity radar simulator for mmWave frequencies that has been used alongside AirSim and Wireless InSite to generate radar sensory data in M3SC and SynthSoM. For mobility simulation, SUMO [67] (Simulation of Urban Mobility) is an open-source microscopic traffic simulator that provides realistic vehicular trajectories, and has been adopted by Raymobtime to generate dynamic vehicular scenarios.

Together, these simulation tools form a flexible ecosystem for researching point cloud-aided wireless communication.

However, the fidelity of the generated point clouds and channels depends critically on the accuracy of the underlying 3D scene models and propagation algorithms, which remains a challenge for ensuring transferability of simulation-trained models to real-world deployments.

2) *Open-Source Datasets*: Table IV summarizes representative datasets that include point cloud or LiDAR modalities alongside wireless communication ground truth. Existing datasets can be broadly divided into measurement-based and simulation-based categories.

Among measurement-based datasets, FLASH [68] and e-FLASH [69] are among the earliest real-world multimodal benchmarks for mmWave beam management. They were collected using autonomous vehicles equipped with Velodyne LiDAR (16-channel and 32-channel, respectively), cameras, and GPS, paired with 60 GHz mmWave beam labels. The LiDAR data is quantized into a $20 \times 20 \times 20$ 3D occupancy grid to provide a compact spatial representation for beam selection. DeepSense6G [70] provides a larger-scale benchmark comprising over 40 scenarios, with a notable difference that the LiDAR and other sensors are deployed at the roadside base station rather than on the vehicle. Its wireless ground truth consists of 64×1 power vectors from beam training with a 64-beam codebook. For indoor environments, the indoor radio mapping dataset [71] pairs high-resolution 3D LiDAR scans with Wi-Fi RSSI measurements across 20 access point configurations in a multi-room space, supporting point cloud-aided channel modeling.

On the simulation side, Raymobtime [59] provides paired ray-tracing channels and LiDAR point cloud data generated using Wireless InSite [58] and SUMO [67] across 10 vehicular scenario variants, and has served as one of the most frequently referenced benchmarks for point cloud-aided beam selection. ViWi [72] is a parametric simulation framework that generates LiDAR, depth, RGB, and CIR data for dynamic outdoor mmWave scenarios, supporting beam management, blockage prediction, and proactive handoff. More recently, M3SC [60] and SynthSoM [61] have significantly expanded the data scale and scenario diversity. M3SC integrates AirSim [64], WaveFarer [66], and Wireless InSite to provide precisely aligned LiDAR point clouds, CIR matrices, and multi-modal sensory data under varying weather and frequency conditions. SynthSoM further extends the coverage to air-ground cooperative scenarios including UAV environments, providing 79K LiDAR point cloud sets alongside 140K channel matrices and 18K path loss measurements across five scene types, with real-world measurement validation.

As shown in Table IV, current open-source datasets are still skewed toward vehicular scenarios, whereas indoor point-cloud-aided wireless benchmarks remain limited. Moreover, although both measurement-based and simulation-based datasets are available, the latter tend to offer broader scenario diversity and larger-scale data collection, leaving a non-negligible sim-to-real gap. The implications of these dataset limitations for model generalization are further discussed in Section VIII.

III. POINT CLOUD FOR PROPAGATION CHANNEL MODELING

In wireless communications, accurately characterizing the propagation channel is essential for applications such as base station (BS) deployment and evaluating building wireless performance [73]. Based on the adopted methodology, channel modeling approaches can be broadly categorized into deterministic, stochastic, and machine learning-aided methods.

Traditional channel models, which rely on idealized or simplified environment assumptions, cannot capture the complex propagation characteristics encountered in realistic environments, particularly at millimeter-wave (mmWave) and sub-terahertz (THz) frequencies [74]. To address this limitation, a detailed and geometry-aware model of the communication environment is essential for accurately characterizing signal propagation under practical conditions. However, conventional approaches to constructing such models are both time-consuming and labor-intensive, often requiring manual development of 3D computer-aided design (CAD) representations. Recent advancements in sensing technology offer an efficient alternative by enabling the precise acquisition of a 3D point cloud. These technologies enable the construction of highly detailed environmental models that reflect fine-grained variations in terrain, building structures, and vegetation that can severely affect signal propagation.

A. Deterministic modeling

Deterministic modeling explicitly computes signal propagation for every location in the environment [75]. The electromagnetic field is precisely characterized by solving Maxwell's equations under specific boundary conditions within a given environment [12]. Methods like ray tracing [76] and ray launching [77] are the most widely used techniques in this category. These methods simulate the physical propagation of electromagnetic waves by capturing interactions like reflection, diffraction, and scattering based on the geometrical and material characteristics of the environment. Therefore, they demand extensive environmental modeling to achieve high accuracy. Point cloud enables the generation of site-specific models that can capture fine-grained spatial features, serving as the foundation for deterministic channel models. The modeling strategies differ depending on the physical characteristics of the scene, such as openness, object density, material diversity, and human activity. This section reviews recent works that utilize point cloud for deterministic modeling in various scenarios.

1) *Indoor scenarios*: Indoor scenarios typically contain a variety of furniture such as desks, chairs, shelves, and electronic devices. These objects are distributed within a space that is enclosed by walls, a ceiling, and a floor. They often cause signal scattering, reflection, and blockage, which significantly affect wireless signal propagation and reduce the accuracy of channel modeling [78]. Modeling these effects requires detailed and geometrically faithful representations of the environment. Recent research efforts have focused on utilizing point clouds captured by LiDAR scanners, RGB-D cameras, or other depth sensors to automate the reconstruction

Dataset	Modalities					Wireless Data	Task	Scenario	Data Source
	RGB	Depth	LiDAR	Radar	GPS				
FLASH [68]	✓		✓		✓	Beam Index	Beam Management	Vehicular	Measurement
e-FLASH [69]	✓		✓		✓	Beam Index	Beam Management	Vehicular	Measurement
DeepSense6G [70]	✓		✓	✓	✓	RSSI, Beam Index	Beam Management, Blockage Management	Multi-scenario	Measurement
Indoor Radio Mapping Dataset [71]			✓			RSSI	Channel Modeling	Indoor	Measurement
Raymobtime [59]	✓		✓			Beam Index	Beam Management	Vehicular	Simulation
ViWi [72]	✓	✓	✓			CIR, Beam Index	Beam Management, Blockage Management, Proactive Handoff	Vehicular	Simulation
M3SC [60]	✓	✓	✓	✓		CIR	Channel Modeling, Beam Management	Vehicular	Simulation
SynthSoM [61]	✓	✓	✓	✓		RSSI, CIR	Channel Modeling, Beam Management	Vehicular, UAV	Simulation

TABLE IV: Summary of existing open-source datasets for wireless communications

of these indoor scenes. For instance, Okamura *et al.* developed a pipeline that extracts room structure from LiDAR-generated point cloud and validates the model with 11 GHz channel sounding measurements in [45]. Similarly, Vaara *et al.* proposed a GPU-based ray launching framework that computes exact propagation paths from a noisy dense point cloud, using voxel-based environment representation and gradient-based path refinement in [47], [79]. These works highlight the growing shift from traditional CAD-based modeling to sensor-driven scene reconstruction, which offers increased realism and scalability. Although indoor environments share common structural elements such as walls, ceilings, and floors, they differ substantially in scatterer density, dominant propagation mechanisms, and material composition. These differences impose distinct requirements on point-cloud-based reconstruction and channel modeling. Accordingly, we classify indoor scenarios into three representative categories. Industrial halls are densely populated with large metallic machinery and equipment [80], which gives rise to severe multipath propagation with high-order reflections and diffuse scattering. Faithful point cloud modeling in such environments must therefore capture the complex geometry of cluttered metallic scatterers to ensure simulation accuracy [81]. Entrance halls, in contrast, are characterized by high ceilings, spacious layouts, and relatively sparse obstructions. Their propagation is more often dominated by sparse but strong specular reflections from large structural elements such as walls, pillars, and staircases [82]. Point cloud reconstruction in these spaces therefore prioritizes preserving large-scale architectural features and subtle structural details [83]. Other functional indoor rooms, including standard offices, lecture halls, cafeterias, and multi-room floor plans, are grouped together because they are enclosed rooms with moderate levels of furniture and functional equipment [49], [84], [85]. At the same time, they still differ in room geometry, occupancy patterns, and material composition. Some of them may also involve dynamic elements, such as moving occupants, which further affect propagation characteristics [86], [87]. Point cloud pipelines for these environments must therefore preserve both fine-grained

static geometry and, when relevant, temporal variations. In the following subsections, we discuss recent studies that apply point clouds to these representative indoor scenarios for propagation channel modeling.

a) *Industrial Hall*: In the context of Industry 4.0, the deployment of dense wireless networks to support the Industrial Internet of Things demands accurate propagation modeling in environments dominated by metallic machinery and structural obstructions. To model such environments accurately, Niu *et al.* proposed a comprehensive ray tracing multi-band simulation workflow for highly cluttered industrial hall scenarios [81]. The workflow encompasses 3D point cloud acquisition via laser scanning. This point cloud is then used to reconstruct a detailed CAD model that preserves key structural and material features essential for propagation modeling. Ray tracing simulations at 6.75, 30, and 60 GHz are conducted on the generated CAD models. The study showed improved agreement with measurements at higher frequencies, highlighting the importance of high-fidelity geometry in dense and cluttered industrial spaces.

b) *Entrance Hall*: The open layout and high ceilings typical of entrance halls result in propagation channels shaped by a small number of strong specular reflections, placing particular emphasis on geometrically faithful reconstruction of large-scale architectural features. Several studies have leveraged the point cloud to reconstruct entrance hall environments for deterministic modeling. For instance, Okamura *et al.* employed LiDAR to capture a dense 3D point cloud of a university entrance hall, and proposed a geometric reconstruction pipeline based on 2D height histograms and contour extraction in [83]. The resulting mesh was validated through 11 GHz channel sounding, showing good agreement in delay and angular profiles. At higher frequencies, De Guzman *et al.* conducted channel measurements at 140 GHz in the entrance hall of Aalto University and developed a measurement-driven ray launching framework in [88]. This approach utilized point cloud-derived geometry to recover double-directional multipath characteristics, thereby addressing the limitations of single-sided angular measurements. In

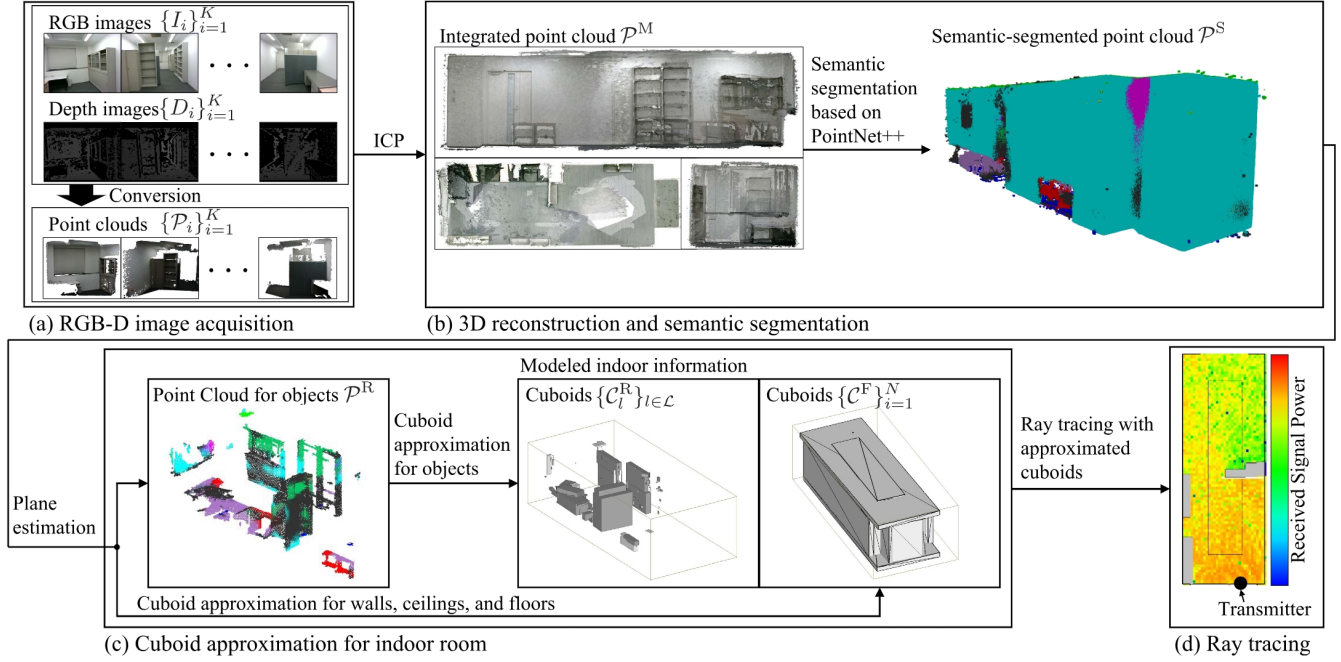


Fig. 3: Proposed point cloud assisted radio map construction framework [50]. RGB-D multi-view depth fusion and PointNet++-based semantic segmentation convert raw point clouds into material-labeled cuboid models for ray tracing.

follow-up work [82], the authors combined a dense indoor point cloud with annotated object segmentation to analyze wave-object interactions at 142 GHz. Their results revealed the dominant role of structural components such as staircases and pillars in shaping reflection clusters and contributing to propagation path diversity.

c) Other Functional Indoor Rooms: The following subparagraphs review point cloud-based channel modeling studies across several representative functional indoor environments, organized by their distinct room geometries and occupancy characteristics.

Office: Offices are typically characterized by dense arrangements of partition walls, furniture, and electronic equipment, which introduce complex multipath conditions and frequent signal obstruction. Wang *et al.* developed a LiDAR-based workflow in [49] that reconstructs full indoor geometry, including walls, doors, windows, and furniture, using voxel grid filtering, RANSAC segmentation, and corner detection. The resulting model is used in custom ray tracing simulations at 1, 3, and 5 GHz, allowing precise modeling of non-line-of-sight (NLOS) paths. Nevertheless, Suga *et al.* proposed an RGB-D-based approach in [50]. They utilize multi-view depth fusion and pointnet++ enabled semantic segmentation to transform the point cloud into annotated cuboid models shown in Fig. 3. Experiments at 2.452 GHz and 5.18 GHz demonstrated that these models are particularly useful for wireless local area network system simulations where object classification, like desks and shelves, improves the prediction of signal blockage and reflection. Further, by integrating an image-based CNN for pixel-wise material classification into the pipeline, Suga *et al.* achieved up to 2.14 dB root mean square errors (RMSE) improvement and over 16% reduction

in transmission rate estimation error in [53].

Dynamic Office Room: Beyond static modeling, point cloud methods have also been extended to dynamic office scenarios. In a realistic dynamic office room, the movement of occupants significantly alters propagation characteristics [86], introducing additional complexity for accurate modeling. To address these temporal dynamics, studies by Saito *et al.* [87], [89] utilized continuous LiDAR scanning to capture both static structures and moving occupants. Changes in point density over time are used to classify dynamic elements, enabling accurate simulation of time-varying channels. The resulting ray tracing predictions at 28 GHz showed high consistency with empirical measurements, validating the approach for dynamic environments.

Lecture and Conference Room: Lecture and conference rooms feature symmetric layouts, fixed seating, and presentation equipment, resulting in relatively uniform propagation conditions compared to standard offices. Recent work [84] by Mi *et al.* integrated LiDAR-based 3D reconstruction with synchronized 60 GHz channel measurements in a lecture room. A screened Poisson surface reconstruction was applied to the merged point cloud. Ray tracing simulations with up to three reflections were conducted, and a cluster centroid distance metric was introduced to evaluate spatial consistency between measured and simulated multipaths. Similarly, in [90], Kitao *et al.* constructed polygonal models from static LiDAR scans of a conference room to simulate throughput performance at 4.7, 28, and 100 GHz in a 6G ray tracing framework. Their findings demonstrated the utility of point cloud-derived geometry in predicting accurate channel responses for high-frequency indoor wireless systems.

Cafeteria: Office cafeterias, present open layouts with regu-

larly arranged furniture and fewer obstructions. In a representative study, Koivumäki *et al.* performed point cloud-assisted ray launching simulations at 60 GHz in a university cafeteria in [85]. The environment was reconstructed using high-resolution LiDAR scans without mesh conversion, treating each point as a reflective micro-element. The method allowed modeling up to six reflections with GPU acceleration. The results showed strong agreement with real-world measurements, validating the accuracy of point cloud-based environment modeling for high-frequency channel prediction.

Multi-Room Office Floor: Multi-room office floor plan consists of interconnected individual offices, meeting rooms, and corridors separated by partition walls and arranged furniture, creating a varied propagation environment with both direct line of sight paths and obstructed signal paths. In [51], Cui *et al.* utilized a mobile-laser-scanning-generated point cloud to automatically reconstruct the 3D indoor environment for signal simulation. By combining 2D floorplan line extraction and 3D surface geometry and applying multi-label graph cuts for room segmentation and model optimization, the method generates semantically rich and geometrically precise models. The detailed flowchart is shown in Fig. 4. The structured models are further used to simulate 5G signal propagation at 100 GHz with ray tracing, enabling accurate coverage prediction and optimal placement of small cells in complex indoor environments.

2) **Outdoor Scenarios:** Outdoor wireless propagation is governed by large-scale geometry and environmental clutter. Compared to indoor settings, outdoor scenarios cover wider areas with more diverse and irregular obstacles, such as buildings, trees, and vehicles. High-fidelity 3D representations are essential for accurate modeling, making the point cloud increasingly indispensable. Based on spatial layout and obstruction density, outdoor environments are typically categorized into street canyons, open squares, and urban macro scenarios.

a) **Street Canyon:** Street canyons, formed by rows of buildings along narrow urban roads, create confined propagation environments with rich multipath and frequent line-of-sight (LOS) blockages. High-accuracy modeling of building facades, signs, and vehicles is essential for realistic simulation. In [91], Koivumäki *et al.* evaluated ray-based simulations at 28 and 60 GHz using LiDAR scans of multiple Helsinki streets. They proposed an adaptive-density point cloud representation, assigning higher density to critical objects like poles and signs. This strategy improved delay and angular domain accuracy with minimal computational overhead. De Guzman *et al.* studied object-level interactions at 142 GHz in a residential street scene in [82]. By isolating structures such as fences and utility poles, they quantified their impact on reflection clusters and delay profiles, showing the importance of fine-grained point cloud modeling.

b) **Open Square:** Open squares are large, low-clutter urban areas typically enclosed by buildings, offering mostly LOS propagation paths. However, edge diffraction and building-induced reflections remain significant. In [46], Koivumäki *et al.* investigated a 100m×100m open square at 28 GHz through synchronized measurements and ray tracing. They compared

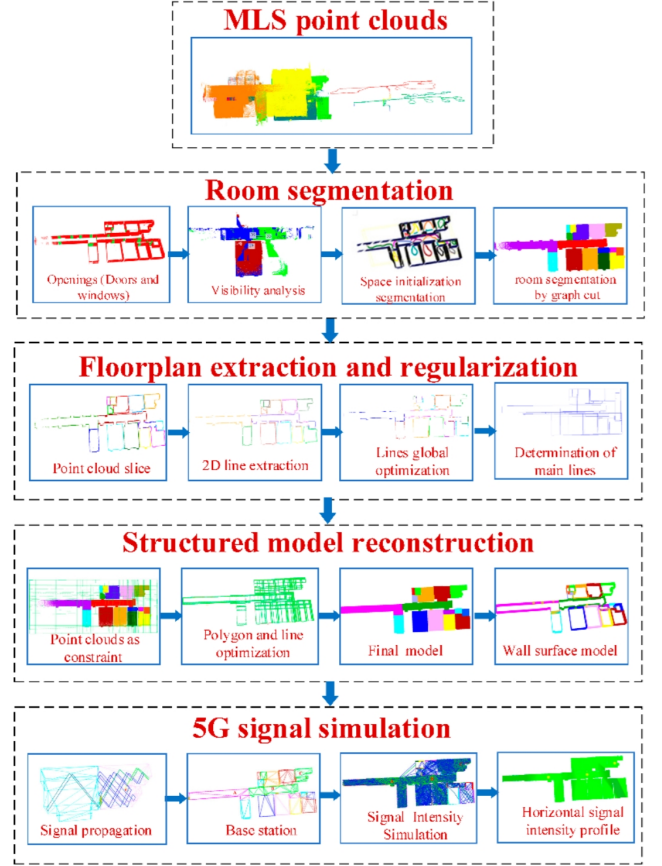


Fig. 4: Flowchart of multi-room reconstruction and ray tracing simulation [51]. A mobile-laser-scanning point cloud is processed through 2D floorplan extraction, 3D surface fitting, and graph cut-based room segmentation to produce structured models for ray tracing.

three levels of point cloud preprocessing: raw, denoised, and surface-fitted. Preprocessed point cloud significantly improved agreement with measurements, especially in weak multipath components and diffracted paths. The detection rate of physical paths increased by over 15% with better alignment in delay and angular spreads. This underscores the importance of detailed and noise-reduced geometry for accurate modeling in open square scenarios.

c) **Urban Macro Scenario:** Urban macro scenario poses significant challenges for propagation channel modeling due to its large spatial extent, heterogeneous structures, and propagation complexity. High-resolution point cloud enables the precise reconstruction of such areas, allowing precise simulation of multipath, diffraction, and foliage-induced loss. In [92], Xia *et al.* reconstructed a 900m×800m city district using aerial photogrammetry, producing a dense segmented point cloud comprising buildings, terrain, vegetation, and man-made objects. The model was converted into surface meshes for accurate Sionna-ray-tracing-based path loss prediction at 2.8 GHz. Their results showed that the model containing buildings, ground and vegetation achieved an overall RMSE of approximately 6.4-6.5 dB. Similarly, Zhang *et al.* applied

unmanned aerial vehicle (UAV)-based oblique imaging and semantic segmentation to reconstruct a 2km² mixed urban area in [93]. Their point cloud captured fine-scale building facades and terrain details, enabling realistic ray tracing at 1.8 and 1.99 GHz. Using the accurate 3D model, the mean absolute errors were 5.5 and 7.4 dB for two measurement routes, demonstrating the benefit of high-fidelity 3D point cloud reconstruction, particularly in regions dominated by diffraction and vegetation attenuation.

To address the computational burden of simulating such dense urban scenarios, Pang *et al.* proposed a progressive processing pipeline for radio wave propagation on GPU platforms in [52]. Their method decomposes the raw point cloud into planar surfaces, diffraction wedges, and nonplanar residuals, generating simplified triangle-based geometries optimized for parallel processing. This representation achieved up to 49.8× acceleration in ray tracing at 3.5 GHz while preserving key propagation effects such as specular reflection and edge diffraction.

3) *Other Scenarios*: Beyond standard indoor and outdoor scenarios, the point cloud has also been employed in specialized and complex scenarios where traditional modeling approaches face inherent limitations. This subsection highlights two representative applications: the railway tunnel and the outdoor-to-indoor scenario.

a) *Railway Tunnel*: Railway tunnel presents a distinct propagation environment characterized by long, narrow, and enclosed geometries with highly reflective surfaces and limited LOS diversity. Signal propagation in such environments is strongly guided by the tunnel structure, exhibiting waveguide-like behavior, multi-order reflections, and limited diffraction due to sparse openings. Accurate modeling of these channels is critical for railway communication systems, including train control and safety applications [94]. Qi *et al.* used the LiDAR-generated point cloud to reconstruct detailed 3D models of railway tunnels using Poisson surface reconstruction in [54]. The resulting mesh was used to perform ray tracing simulations at 1.8 GHz. Their model captured the tunnel's curved and straight segments with high geometric fidelity. When compared to the measurement data, the simulations achieved high accuracy with an RMSE of 4.0 dB in the straight tunnel and 4.8 dB in the curved tunnel, demonstrating the effectiveness of point cloud-based modeling in such constrained environments.

b) *Outdoor-to-Indoor Scenario*: Accurate modeling of outdoor-to-indoor propagation is essential for delivering reliable wireless connectivity to indoor users, especially in high-frequency bands where building penetration loss is significant. Koivumäki *et al.* investigated this scenario using detailed 3D models reconstructed from the laser-scanned point cloud in [95]. Their method includes both exterior and interior structural elements, including walls, windows, and facades. Feature extraction algorithms are applied to identify critical surfaces, which are then assigned material properties based on permittivity and conductivity values to simulate their impact on signal behaviour. Ray tracing simulations were conducted to capture direct transmission, reflection, diffraction, and window penetration effects as signals transitioned from outdoor

environments into indoor spaces. The results demonstrated strong agreement with measurement data, achieving mean path loss errors below 6 dB and delay/angular spread deviations within 10%, validating the utility of point cloud-based modeling for outdoor-to-indoor scenarios.

B. Stochastic Modeling

Stochastic channel models typically rely on statistical or geometry-based stochastic models, with their parameters usually derived from extensive measurement campaigns. Although such models are efficient and scalable, they often simplify local propagation environments into generic statistical assumptions, which limits their ability to capture site-specific variations. In this context, point clouds mainly serve as a source of site-specific geometric information that enhances stochastic modeling in two aspects: environment-aware statistical parameterization, where statistical assumptions are grounded in real-world geometry, and geometry-informed attenuation estimation, where propagation loss is characterized based on environment-dependent geometric features.

First, point clouds enable environment-aware statistical parameterization. In conventional stochastic models, the spatial distribution of scatterers or blockers is often described by simplified assumptions or idealized probability distributions. By incorporating point cloud data, these parameters can be related more closely to the actual morphology of the environment. For example, in [96], Huang *et al.* used semantic segmentation on LiDAR point clouds to distinguish dynamic vehicles from static structures, allowing multipath-related parameters, such as scatterer density, power, and angular spread, to vary with the actual vehicular traffic density instead of fixed generic assumptions. Similarly, in [97], Koivumäki *et al.* employed point clouds to characterize the effective blocking regions of small-scale cluttered objects, such as street signs and pedestrians, thereby providing more realistic density parameters for Poisson-point-process-based LOS probability modeling.

Second, point clouds enable geometry-informed attenuation estimation in scenarios where coarse statistical assumptions are insufficient. This is particularly relevant in GNSS-related propagation studies involving vegetation and other irregular obstacles. In [98], Kubelka *et al.* extracted local geometric features from dense point clouds to distinguish occlusions caused by buildings from attenuation caused by forest canopies, enabling more accurate attenuation estimation under different obstruction types. In [99], Kou *et al.* further used slice-based analysis on airborne laser scanning point clouds to estimate the signal path length through tree canopies, allowing attenuation prediction to be associated with vegetation thickness rather than relying on a coarse binary visible/blocked classification.

C. Machine Learning Aided Modeling

The integration of machine learning techniques with the point cloud has emerged as a powerful approach for enhancing the accuracy of propagation channel modeling, especially in complex environments. Point cloud provides detailed 3D representations of physical environments, capturing essential

REF	Point Cloud Acquisition	Scenario	Freq. (GHz)	Method	Evaluation Metric	Reported Performance
[83]	LiDAR	Entrance Hall	11	RT	Channel gain; delay/angular profiles	Median CDF difference 1.7 dB; gain error <5 dB
[88]	LiDAR	Entrance Hall	140	RL	Path mapping success rate	70% paths and 84% power mapped
[82]	LiDAR	Entrance Hall / Street Canyon	142	RL	Path mapping rate; reflection loss	85% paths / 88% power mapped; per-object reflection loss quantified
[49]	LiDAR	Office	1, 3, 5	RT	Std. deviation of received power	2.60 dBm (1 GHz); 6.38 dBm (3 GHz); 6.75 dBm (5 GHz)
[50]	RGB-D	Office	2.452, 5.18	RT	RMSE of RSS radio map	3.27-4.30 dB
[53]	RGB-D	Office	2.452, 5.18	RT	RMSE; transmission rate error	2.14 dB RMSE improvement; 16% rate error reduction
[90]	LiDAR	Conference Room	4.7, 28, 100	RT	Throughput evaluation	>100 Gbps feasibility validated at sub-THz
[85]	LiDAR	Cafeteria	60	RL	Delay spread relative error	<10% with path grouping
[46]	LiDAR	Open Square	28	RT	Path detection rate; delay/angular spread	>15% detection rate improvement
[92]	Photogrammetry	Urban Macro	2.8	RT	RMSE (path loss)	6.4-6.5 dB
[93]	Photogrammetry	Urban Macro	1.8, 1.99	RT	MAE (received power)	5.5/7.4 dB for two routes
[54]	LiDAR	Railway Tunnel	1.8	RT	RMSE (path loss)	4.0 dB (straight); 4.8 dB (curved)
[95]	LiDAR	Outdoor-to-Indoor	4, 14	RT	Mean PL error; delay/angular spread deviation	<6 dB PL error; <10% spread deviation

TABLE V: Performance summary of point cloud-assisted deterministic channel modeling studies.

geometric information that enables machine learning models to better predict path loss and detailed channel characteristics.

1) *RSS/Path Loss Prediction*: The point cloud has been extensively utilized to predict received signal strength (RSS) and path loss in wireless communications. By encoding rich 3D environmental features, the point cloud enables machine learning models to accurately infer signal attenuation patterns.

a) *Outdoor Scenarios*: Outdoor scenarios present unique challenges for propagation channel modeling, due to complex geometries, sparse measurements, and the presence of large-scale scatterers like buildings and vegetation. Point cloud offers a detailed spatial representation that empower data-driven models to predict RSS and path loss more accurately.

For instance, Li *et al.* proposed NeRA, a neural radiative architecture trained on 3D grids constructed by voxelization of UAV-acquired LiDAR point cloud in [100]. Three separate networks are trained to model attenuation, reflections, and radiative field strength, reconstructing the full RSS distribution at 4.6 GHz. Nevertheless, Wang *et al.* extended the 3D Gaussian Splatting framework to wireless communication applications by initializing spatial Gaussians from point cloud geometry in [101]. Through sparse RSS supervision and relaxed-mean reparameterization, it enables efficient and scalable radio map inference with sub-second latency. In [102], Egi *et al.* used the airborne-LiDAR-generated point cloud to model foliage-induced signal attenuation on a university campus. An artificial neural network (ANN) is trained to regress path loss based on canopy density and geometry. Gupta *et al.* [103] applied regression models such as SVR and Random Forests to predict urban canyon path loss using public LiDAR-generated point cloud of Manhattan street.

b) *Vehicular Scenarios*: Vehicular communication scenarios, comprising both vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) links, are characterized by inherently dynamic and heterogeneous radio propagation environments.

The variations in surrounding geometries, the presence of mobile scatterers, and the susceptibility to ambient factors, including meteorological conditions and traffic density, pose substantial impediments to conventional channel modeling methodologies. To mitigate these challenges, contemporary research has harnessed the point cloud to facilitate enhanced characterization of path loss attributes.

In [104], Wei *et al.* introduced a GAN-pretrained architecture integrating LiDAR-generated point cloud projection, RGB imagery, and depth maps to predict probabilistic path loss distributions at 5.9 and 28 GHz under diverse environmental and traffic conditions. The framework leverages Vision Transformers and FasterNet for modality-specific feature extraction, fused via multi-head attention and regressed through MLP layers.

For V2I applications, Wang *et al.* proposed MFEF-Net, a modular architecture integrating LiDAR-generated point cloud, RGB, and global positioning system (GPS) data to predict path loss at dense urban intersections in [105]. The system employs PointNet for spatial feature extraction, ResNet for vision encoding, and a GRU for trajectory-aware positional embedding, achieving robustness across varied lighting and topologies.

In the V2V domain, Peng *et al.* [106] introduced an end-to-end multimodal framework that estimates the distribution of path loss values using LiDAR-generated point cloud, RGB, and depth inputs. The architecture combines PointNet++ and ResNet34 for spatial and visual feature extraction, with channel-wise attention fusion, and classifies signal strength into discrete bins.

c) *Autonomous Robotic Scenarios*: Point cloud-assisted propagation channel modeling has also been extended to robotic platforms operating in constrained or GPS-denied environments such as mines, tunnels, and caves. In [107], Clark *et al.* proposed PropEM, a voxel-based learning framework

for real-time RSS prediction. The system transforms 3D point clouds into voxel occupancy grids, which are processed by neural networks to estimate signal strength for arbitrary robot poses. The framework supports communication-aware exploration by enabling trajectory planning that jointly considers exploration utility and link quality.

2) *Channel Estimation*: Beyond path loss prediction, point cloud has been leveraged to enhance the estimation of detailed channel characteristics. Channel estimation recovers the propagation channel either as the time-domain CIR or its frequency-domain CSI, which are related through the Fourier transform [108]. By incorporating 3D spatial context from point clouds, learning-based methods can infer these channel representations with reduced pilot overhead and improved spatial awareness.

In the frequency domain, Zhang *et al.* proposed a multi-branch CNN that jointly processes sparse pilot signals and LiDAR-derived point cloud to estimate full-band, multi-user CSI at 28 GHz in [57]. The fusion of pilot and geometric modalities enabled the model to reconstruct CSI with high accuracy, while significantly reducing the need for dense pilot transmissions.

In the time domain, Wang *et al.* proposed a direct regression framework for CIR vector $\mathbf{h}[n]$ across delay taps using a PointNet-based architecture in [109]. A spatial selection module extracts local regions from the environment's point cloud, and aggregated point-wise features are used to model geometry-induced multipath without requiring ray tracing or CAD-based scene reconstruction. Nevertheless, [110] employed point cloud classification to identify candidate reflection points, from which geometric path parameters such as delays, angles of arrival, and departure are estimated. This path-wise reconstruction approach proves especially effective in vehicular scenarios, where channel variation closely follows dynamic scene geometry.

The point cloud has been further extended to scatterer-level analysis, offering a more granular understanding of CIR formation in [111]. Huang *et al.* employed a clustering-based LiDAR-generated point cloud processing pipeline to identify individual scatterers in the environment and classify them as static or dynamic using MLP. This classification enables the generation of structurally-aware CIRs that account for non-stationarity and mobility in the environment.

Furthermore, point cloud has also been used to predict statistical channel descriptors. Mi *et al.* [112] estimated delay spread, angular spread, and the Rician K -factor directly from high-resolution indoor LiDAR-generated point cloud using a PointNet-based regressor trained on 60.5 GHz measurement data. The model demonstrates robust prediction of key CIR metrics across multiple LOS settings.

D. Discussion

Point cloud plays different roles across the three modeling paradigms. In deterministic modeling, it serves as the basis for environmental reconstruction, enabling site-specific propagation simulation. In stochastic modeling, it provides geometric priors that constrain statistical parameterization, such as anchoring scatterer distributions to measured geometry. In machine learning-aided modeling, it acts as an

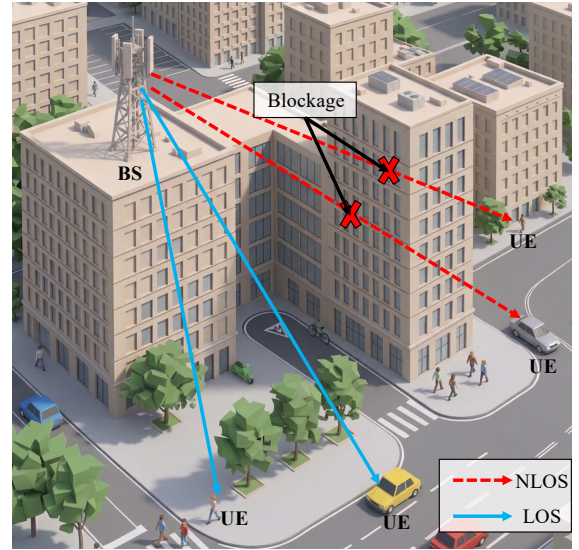


Fig. 5: Illustration of a point-cloud-aided blockage management scenario in an urban wireless environment.

input representation that encodes spatial context for channel prediction. Representative quantitative results of deterministic studies are summarized in Table V, and machine learning-aided studies are summarized in Table VI.

A common factor across all three paradigms is that the level of geometric detail extracted from the point cloud directly influences modeling accuracy. Within deterministic modeling, raw point cloud representations [85] preserve maximum surface detail for propagation simulation, while cuboid approximations [50] trade geometric fidelity for computational efficiency. Mesh-based reconstructions [84], [90] offer an intermediate balance. In addition, higher carrier frequencies generally require finer point cloud representations to capture the dominant scattering mechanisms. In machine learning-aided modeling, a similar trade-off exists in how the point cloud is encoded as model input. Projection-based and voxel-based encodings [100], [104] enable standard CNN architectures but introduce quantization or dimensional loss. Point-based architectures [109], [112] process the raw geometry directly, better preserving the spatial relationships that determine multipath characteristics.

IV. POINT CLOUD FOR BLOCKAGE MANAGEMENT

Emerging applications such as virtual and augmented reality, Internet of Things, and autonomous vehicles have created unprecedented demands for ultra-high throughput, ultra-low latency, and massive connectivity. These requirements have led to a dramatic surge in data traffic, thereby driving the shift toward higher frequency bands to meet the growing need for enhanced communication capacity [113], [114]. As demonstrated in Section III, signal blockage due to obstacles, including buildings, vehicles, vegetation, and pedestrians, severely impacts wireless link stability, propagation reliability, and overall system performance. This phenomenon becomes even more pronounced as communication systems migrate to higher frequency bands, where signals are inherently more susceptible

REF	Sensor	Input Parameters	Output Parameters	Testing Scenarios	Freq. (GHz)	DataSet	Model	Application
[102]	Airborne LiDAR	Height and Width of detected tree canopies	Path loss	Campus	2.121	Measured	ANN	RSS/path loss prediction
[103]	LiDAR	Street clutter features, building features	Path loss	Street canyon	28	Measured	RF, SVR, Lasso, Elastic-net	RSS/path loss prediction
[104]	LiDAR	RGB and Depth images, LiDAR top-down projection images	Path loss	V2X (weather, traffic density)	5.9, 28	Simulated	MLP, GAN, ViT, FasterNet	RSS/path loss prediction
[105]	LiDAR	RGB images, point cloud, GPS	Path loss	V2I	NA	Measured	PointNet, ResNet, GRU	RSS/path loss prediction
[106]	LiDAR	LiDAR point cloud, RGB images and depth images	Path loss	V2V	NA	M3SC	PointNet++, ResNet34, MLP	RSS/path loss prediction
[107]	LiDAR	Sparse point cloud, Tx/Rx position, power and frequency	RSS	Mine, Cave (Robotics)	NA	Measured	5 FCNNs	RSS/path loss prediction
[100]	UAV LiDAR	Voxelized point cloud, Tx location and Ray direction	RSS	General Outdoor	4.6	Measured	Three MLPs	RRSS/path loss prediction
[101]	LiDAR	Point cloud, BS location	RSS	City-scale outdoor environment	NA	Simulated and Measured	RadSplatter framework	RSS/path loss prediction
[57]	LiDAR	Sparse pilot obs., Processed LiDAR range images	CSI	V2V	28	M3SC	CNNs	CSI Estimation
[111]	LiDAR	LiDAR features, Tx/Rx positions	CIR, scatter (type and number)	Urban crossroad	28	Simulated	MLP, DBSCAN	CIR Estimation
[112]	LiDAR	Fused point cloud, RGB, Rx position, semantics	Path Loss, Delay Spread, Angular Spread, K-Factor	Lecture Room	60.5	Measured	Regression PointNet	CIR Estimation
[110]	NA	Point cloud feature, BS pos., MS pos.	Channel characteristics	V2I	28	Simulated	PointNet, MLP	CIR Estimation
[109]	UAV DJI Terra	Local point subset (spatial)	CIR	Campus	4.6	Measured	PointNet++	CIR Estimation

TABLE VI: Summary of studies on point cloud and machine learning aided channel modeling

to attenuation and diffraction loss. Accurate management of these blockage events thus becomes critical for ensuring robust wireless communication, especially at millimeter-wave and sub-terahertz frequencies.

To effectively mitigate the challenges associated with blockages, wireless communication systems incorporate advanced environmental awareness into their operational framework [115]. A point cloud derived from LiDAR or similar sensors provides detailed geometric insights into the surrounding environment, facilitating proactive blockage management. Fig. 5 illustrates a representative urban blockage management scenario, where surrounding buildings block the direct BS–UE propagation paths and create NLOS conditions for some links. Based on such geometric awareness, the system can not only detect whether an existing link is currently blocked, but also predict potential future blockage events, for example when moving pedestrians or vehicles enter the blockage region of a propagation path. Such environmental awareness enables both the timely detection of existing obstructions and the prediction of potential blockages, laying the foundation for the approaches reviewed in the following sections.

A. Blockage Detection

Point clouds have become a valuable resource for accurately capturing environmental geometry and providing detailed insights into the physical surroundings for wireless communication systems. This rich 3D information enables the identification of impediments and obstructions within wireless communication links, encompassing both static and dynamic entities. By leveraging these spatial and structural details, communication systems can better model the impact of blockages and implement mitigation strategies to enhance link reliability in complex environments.

Various studies have investigated the versatility of the point cloud for blockage detection across diverse scenarios, ranging from cluttered indoor environments with dense obstacles to outdoor settings characterized by large-scale variability and mobility. These approaches employ diverse sensing technologies, including LiDAR, radar, and imagery-based systems, and integrate advanced machine learning models for real-time blockage detection. Table VII summarizes representative works in this area, highlighting their sensing methods, application domains, and performance characteristics.

1) *Indoor Scenarios*: Indoor scenarios present unique challenges for wireless communication systems due to dense obstacles and frequent environmental changes. Point cloud provides precise 3D representations of both static elements, such as walls and cabinets, and dynamic factors, including people and movable furniture, supporting advanced blockage detection and enabling accurate classification of LOS and NLOS conditions.

a) *Static Blockage*: Static indoor obstacles such as walls and large furniture can consistently obstruct LOS paths between BS and UE. For instance, Woodford *et al.* proposed a real-time blockage detection method in [116]. It applies a ray-tracing algorithm to LiDAR-generated 3D models, using path-averaging and pruning to improve robustness while maintaining accurate identification of blockage surfaces. Similarly, Sharma *et al.* utilized a projection-based approach to evaluate LOS conditions in [117]. By projecting LiDAR-generated 3D point cloud onto 2D planes, their algorithm identifies obstacles that obstruct LOS by analyzing their placement relative to the direct sensor-device line.

b) *Dynamic Blockage*: Dynamic indoor environments feature both immobile structures and mobile objects, leading to complex, time-varying blockages. For instance, Kang *et al.* modeled human body shadowing at 58 GHz using point cloud and propagation measurements in [118]. They precisely simulate time-varying blockage effects of shadowing distances and shadowing gains by employing geometric models such as vertical screens and elliptic cylinders. Expanding further, Marasinghe *et al.* combined ray-casting and deep learning to identify human-induced blockages in dynamic settings using LiDAR-generated point cloud in [119].

Sensor fusion further improves blockage detection performance. Zhang *et al.* combined LiDAR-based point cloud with a stereo-camera to identify blockages to adapt lighting conditions in [120]. Point cloud is preprocessed by object clustering and overlap analysis to identify obstacles, excelling in low-light scenarios. Stereo camera leverages a YOLOX-based RNN to detect blockages in well-lit environments. Kamari *et al.* leveraged RGB-D images and point cloud reconstruction to detect both static and dynamic obstructions via machine learning in [121]. Similarly, Nishio *et al.* combined depth images and point cloud with convolutional long short-term memory (ConvLSTM) networks in [122] to model the spatiotemporal patterns of signal degradation in multi-user environments, enabling real-time blockage detection.

Industrial scenarios introduce additional complexity due to large-scale machinery and continuous equipment movement. In [123], Feng *et al.* utilized LiDAR-generated point cloud combined with transfer learning-based models to detect static and semi-dynamic obstacles with minimal retraining, enabling robust blockage detection in dynamic industrial spaces.

2) *Vehicular Scenarios*: Vehicular communication faces even greater blockage challenges due to highly dynamic surroundings, including moving vehicles, pedestrians, and complex urban layouts. Static obstacles such as buildings and terrain remain significant, and rapidly changing LOS conditions introduce the demand for real-time adaptation. To meet the challenges, point cloud-based methods have emerged

as indispensable tools for blockage detection in vehicular communication.

a) *Static blockage*: Nerini *et al.* addressed static blockage detection using LiDAR-generated point cloud with a physics-based graph neural network (GNN) in [124]. This method utilizes k-nearest neighbor graph construction and message-passing algorithms to detect obstacles, while ray tracing on reconstructed 3D surfaces further improves LOS and NLOS classification. In [55], Klautau *et al.* applied CNNs to LiDAR-generated point cloud for static obstacle detection, while Dias *et al.* explored both distributed and centralized CNN-based architectures using binary point cloud arrays in [125]. Okada *et al.* enhanced detection performance by fusing radar point clouds from multiple units using a Kalman filter and joint probabilistic data association (JPDA) in [126]. This multi-sensor fusion approach extends coverage and improves the accuracy of identifying static blockages, such as parked vehicles and road infrastructure, across large urban areas.

b) *Dynamic blockage*: Asano *et al.* [127] incorporated radar point cloud, DBSCAN clustering, and Kalman filters to detect and track moving obstacles, enabling detection for LOS/NLOS transitions in [127]. Nevertheless, Pradhan *et al.* proposed a cooperative perception method [128] that quantizes LiDAR point clouds from connected autonomous vehicles into structured grids and applies clustering with deep attention networks to detect both static and dynamic obstacles. This cooperative perception method enhances blockage detection by integrating multiple perspectives, reducing blind spots in highly dynamic environments. In addition, Kamari *et al.* used street-view imagery to generate 3D point cloud enriched with material properties for improved detection of transient obstacles in urban settings in [129].

B. Blockage Prediction

Expanding on the foundation of point cloud-assisted blockage detection, blockage prediction focuses on anticipating future obstructions by analyzing spatial and temporal patterns in the environment. This approach leverages point clouds to model dynamic changes, such as the movement of vehicles or pedestrians, and predict potential LOS or NLOS transitions before they occur. By utilizing machine learning techniques and sophisticated algorithms, blockage prediction enables communication systems to proactively adapt to environmental variability, thereby improving reliability in dynamic communication scenarios.

As summarized in Table VIII, recent studies present a diverse set of point cloud-driven methodologies for blockage prediction. These methods integrate advanced sensing technologies such as LiDAR, stereo cameras, and depth sensors with machine learning and deterministic modeling to address the challenges posed by dynamic blockage prediction. Techniques range from deep learning models, such as CNNs and RNNs, to reinforcement learning and sequence-to-sequence frameworks, enabling precise predictions of disruptions caused by human movement and vehicular motion. These methods demonstrate the potential of point cloud-assisted prediction in improving wireless communication reliability in dynamic and complex scenarios.

Study	Sensor	Methodology	Blockages Type	Scenarios
[116]	LiDAR	LiDAR-generated 3D models with ray-tracing algorithm	Static	Indoor
[117]	LiDAR	Projection-based approach with 3D-to-2D point cloud conversion and rectangular RoV analysis	Static	Indoor
[118]	LiDAR	Dynamic deterministic models (e.g., vertical screens, elliptic cylinders) for time-varying human blockage simulation	Dynamic	Indoor
[119]	LiDAR	Ray-casting algorithms combined with deep learning models using LiDAR point cloud data	Dynamic	Indoor
[120]	LiDAR and Stereo Cameras	LiDAR preprocessing (ground removal, clustering) and YOLOX CNN for stereo camera data	Static and dynamic	Indoor
[121]	RGB-D	RGB-D images and point cloud data with machine learning models for inter-user occlusions and 3D environment reconstruction	Static and dynamic	Indoor
[122]	RGB-D	ConvLSTM networks with depth images and point cloud data to model spatiotemporal signal degradation patterns	Dynamic	Indoor
[123]	LiDAR	LiDAR-based point cloud data with machine learning models, incorporating transfer learning for static and semi-dynamic obstacles	Static and dynamic	Indoor
[124]	LiDAR	Physics-based GNN with ray tracing for LOS/NLOS classification	Static	Vehicular
[55]	LiDAR	Convolutional neural networks for LOS/NLOS classification	Static	Vehicular
[125]	LiDAR	Binary array conversion of point cloud analyzed via CNNs	Static	Vehicular
[126]	Radar	Kalman filter and JPDA for multi-radar data fusion	Static	Vehicular
[127]	Radar	DBSCAN clustering and Kalman filters with route prediction	Static and dynamic	Vehicular
[128]	LiDAR	Quantization into structured 3D grids, clustering, and attention-based deep learning	Static and dynamic	Vehicular
[129]	Street view	Segmentation and material identification models on street view images	Static and dynamic	Vehicular

TABLE VII: Summary of Studies on point cloud-assisted blockage detection methods

1) *Pedestrian Scenarios*: Pedestrian-induced obstructions are highly dynamic and unpredictable, making them a critical focus of blockage prediction. Variability in human size, speed, and behavior poses unique challenges to reliable link maintenance. The point cloud assisted approaches address this issue by combining advanced sensing modalities such as LiDAR and depth cameras with machine learning to anticipate disruptions caused by human motion. The reviewed studies adopt diverse methodologies, including deterministic modeling, deep learning, and multimodal sensing, to accurately forecast pedestrian blockages.

For instance, Marasinghe *et al.* proposed an end-to-end system in [119] that leverages infrastructure-mounted LiDAR sensors to capture the dynamics of indoor environments. The system processes the collected point cloud using ray-casting and deep learning models, achieving an impressive 87% accuracy in predicting human blockages within a 300 ms time window. Nevertheless, dynamic deterministic modeling has also proven effective. Kang *et al.* introduced a method in [118] that simulates human body shadowing at 58 GHz by representing human movement geometry with vertical screens and elliptic cylinders. By synchronizing the point cloud with

channel measurements, their approach accurately estimates the shadowing gain and forecasts LOS conditions.

2) *Vehicular Scenarios*: Vehicular blockages introduce additional complexities compared to pedestrian scenarios due to the larger size, higher speed, and unpredictable trajectories of vehicles. These characteristics require advanced prediction methods that leverage point cloud to anticipate the impact of vehicles on LOS/NLOS conditions in communication networks. Various studies have developed innovative methods to predict vehicular blockages effectively.

In [130], Wu *et al.* proposed a LiDAR-based method that denoises point clouds to extract dynamic object traces, which are then analyzed by a CNN to learn vehicular motion patterns. The method achieves a prediction accuracy of 95% for blockages occurring within 100 ms and 80% accuracy within 1 second, showcasing its effectiveness in real-time applications. In a follow-up study, Wu *et al.* extended this point cloud-only pipeline by fusing LiDAR-generated point cloud with mmWave signatures, applying static cluster removal to isolate dynamic objects and machine learning models to analyze spatial density shifts and trajectories in [131]. This cross-modal fusion enables a richer prediction output of

the occurrence, timing, and severity of vehicular blockages, facilitating proactive link management.

Radar-based systems offer an alternative point cloud source with distinct trade-offs. Unlike the dense LiDAR point clouds used in [130], [131], radar-generated point clouds are sparser and lower in spatial resolution, but operate robustly under adverse weather and lighting conditions. In [127], Asano *et al.* processed radar-generated point clouds using DB-SCAN clustering and Kalman filters to predict vehicular trajectories and blockages in urban environments, enabling route-aware resource allocation. The object-level clustering strategy extracts discrete targets from the sparse radar point cloud, compensating for its lower density through structured trajectory modeling. Okada *et al.* extended this concept in [126] by fusing point cloud from multiple radar units with Kalman filters and JPDA, enhancing coverage and prediction accuracy in highway and large urban areas.

Cooperative frameworks further extend the sensing coverage area. In [128], Pradhan *et al.* proposed a cooperative design in which LiDAR-generated point cloud from multiple connected vehicles are fused through an attention-based mechanism. This system identifies transient vehicular blockages, mitigates blind spots, and determines optimal road-side units (RSUs) to maintain connectivity, enabling real-time blockage prediction in complex, high-mobility environments.

C. Discussion

Point cloud serves as the primary source of environmental geometry for blockage management. By providing explicit 3D spatial information about obstacles, it enables communication systems to identify and anticipate link obstructions without relying solely on signal-level indicators.

The reviewed methods differ in how they exploit this geometric information. Some approaches perform explicit geometric reasoning on the point cloud through ray-casting or ray tracing [116], [124], directly evaluating whether physical objects obstruct the propagation path. Others use the point cloud as a learned feature input, compressing it into 2D projections or binary arrays for neural network-based classification [55], [125]. The former preserves the physical meaning of the point cloud but is computationally heavier. The latter achieves faster inference but relies on training data and may lose spatial details during compression. For blockage prediction, the point cloud further provides temporal motion cues through sequential frames, enabling recurrent models to forecast future obstructions [122], [130].

V. POINT CLOUD FOR BEAM MANAGEMENT

As discussed in the previous sections, Section III revealed that link blockages caused by obstacles such as buildings, vehicles, and pedestrians can significantly degrade wireless communication performance. To address this issue, approaches in Section IV focused on detecting and predicting potential obstructions in wireless communications. Building upon these efforts, beam management serves as a complementary solution that actively mitigates the impact of blockages and

alleviates performance degradation in wireless communications. However, conventional beam management approaches typically rely on exhaustive beam sweeping, repeated beam measurements, and feedback exchange, which incur substantial training overhead and latency.

Point clouds offer a fundamentally different source of information by explicitly capturing the 3D structure of the propagation environment and therefore supply a physically meaningful prior for beam management. Unlike purely signal-driven approaches, they make it possible to reason about why a beam is likely to be effective before extensive beam training is performed. Specifically, point cloud can reveal whether the direct path between transmitter and receiver is available or obstructed, identify dominant reflecting or scattering structures that may support strong indirect paths, and characterize the spatial evolution of users, vehicles, and blockers over time. These geometry-derived cues can then be used to prune infeasible beam directions, rank candidate beams according to likely propagation support, predict future beam transitions, and guide beamforming or codebook design under environment-dependent constraints. Therefore, the benefit of point clouds in beam management lies not only in richer sensing, but in enabling a shift from blind beam exploration to geometry-aware beam inference and optimization.

Beam management encompasses a set of Layer 1 (physical layer) protocols [132] and Layer 2 (medium access control) protocols [133] designed to acquire and maintain beam pair links [134]. These links are established by aligning beams at the Tx/Rx points on the BS side with corresponding beams at the user equipment (UE) side, ensuring robust and efficient communication through precise beam management operations [135]. According to Section 6.1.6.1 in the 3GPP technical report [136], the operation of beam management includes at least four aspects:

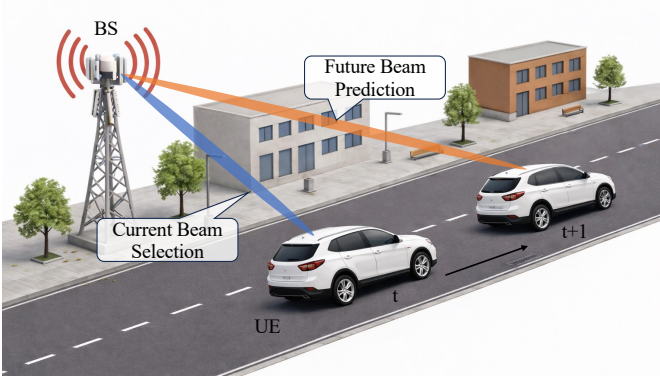
- **Beam Determination:** The process by which Tx/Rx points or UE select their respective Tx or Rx beams
- **Beam Measurement:** The evaluation of characteristics of received beamformed signals by Tx/Rx points or UE to assess signal quality and performance.
- **Beam Reporting:** The transmission of information by the UE regarding beamformed signals, based on the results of beam measurements.
- **Beam Sweeping:** An operation that systematically covers a spatial area by Tx and/or Rx beams in a predetermined manner over a specific time interval.

However, in practice, the boundaries between these stages are increasingly blurred. Many learning-based approaches integrate multiple functions, such as directly predicting the optimal beam index without explicit beam sweeping or measurement. Others leverage 3D environmental modeling to implicitly guide beam determination, thereby reducing or bypassing conventional procedures.

To better reflect the modeling goals, data usage, and prediction targets in point cloud-based beam management, we adopt a task-oriented classification strategy. Specifically, existing works can be categorized into three groups based on their primary functional objective:

Study	Sensor	Methodology	Key Contribution	Scenarios
[119]	LiDAR	Ray-casting and deep learning models	Achieves 87% accuracy within 300 ms time window	Pedestrian
[118]	LiDAR	Dynamic deterministic modeling with vertical screens and elliptic cylinders	Simulates shadowing gain and its impact on LOS conditions	Pedestrian
[130]	LiDAR	Denoising algorithm for LiDAR data and CNN-based object analysis	Achieves 95% accuracy for blockages within 100 ms	Vehicular
[131]	LiDAR	Static cluster removal and machine learning to detect pre-blockage indicators	Proactively predicts blockage severity with LiDAR and mmWave integration	Vehicular
[127]	Radar	DBSCAN clustering and Kalman filters	Enables precise route prediction and resource allocation	Vehicular
[126]	Radar	Kalman filters and JPDA for multi-radar data fusion	Enhances spatial range and accuracy with multi-radar fusion	Vehicular
[128]	LiDAR	Attention-based fusion of LiDAR data from multiple vehicles	Extends coverage and mitigates blind spots through cooperative perception	Vehicular

TABLE VIII: Summary of studies on point cloud-assisted blockage prediction methods

Fig. 6: Illustration of a vehicular beam management scenario, showing current beam selection at time t and future beam prediction at time $t + 1$.

- Current Beam Selection (relates to beam determination)
- Future Beam Prediction (relates to beam measurement and determination)
- Beamforming Optimization (relates to beam determination and beam sweeping)

This perspective allows a clearer differentiation of methods based on the specific roles that point clouds play within the beam management pipeline, whether it is enabling accurate current beam alignment, predicting future beam transitions, or accelerating optimization processes. Fig. 6 illustrates a representative vehicular beam management scenario. With point-cloud-based geometric awareness of the road environment and vehicle motion, the system can determine the beam aligned to the UE at the current time step t and further predict the beam required at the future time step $t + 1$.

A. Current Beam Selection

Building on the role of beam management, one major research direction focuses on current beam selection, which aims to identify an appropriate beam pair for the current channel state from sensing observations. Based on how point cloud data is exploited during the beam selection process,

existing methods can be broadly classified into two categories: direct beam selection and search-efficient beam selection. Direct beam selection infers the current optimal beam from current sensing inputs, whereas search-efficient beam selection improves the efficiency of beam acquisition by reducing the beam search space or sweeping overhead during the selection process.

1) *Direct Beam Selection*: In direct beam selection, the optimal beam is inferred directly from current sensor inputs, typically without explicit sweeping over the full beam codebook. This task is typically formulated as a multi-class classification problem, where inputs such as point clouds, images, and GPS are mapped to the optimal transmit or receive beam index. By leveraging point cloud-assisted inference, these methods aim to replace traditional exhaustive sweeping and measurement procedures with fast, one-shot predictions, thereby reducing latency and improving link robustness.

a) *Indoor Scenarios*: In indoor environments, where the geometry is relatively static but highly reflective and complex, point cloud-based beam selection can achieve high precision without requiring channel measurements. A representative work is [116], which leverages RGB-D scans to construct a 3D mesh of the environment and applies a custom RF ray-tracing engine to compute optimal beam paths. This system precomputes a lookup table of beam indices for all feasible Tx/Rx positions, enabling real-time one-shot selection without in-band sweeping. Evaluations across multiple indoor scenarios, including empty rooms, cluttered environments, and garages, demonstrate over 80% latency reduction and 2× throughput gain compared to standard 802.11ad beam training.

Indoor industrial environments such as smart factories exhibit unique wireless propagation challenges due to dense metallic structures, dynamic machinery, and rapidly changing layouts. Beam selection needs to account for both static and dynamic components that significantly impact the channel. Feng *et al.* proposed a point cloud-assisted framework combining RGB camera images and LiDAR scans to jointly model static surroundings and dynamic scatterers in [123]. Spatial and temporal features are extracted using 2D and 3D

Study	Sensor	Methodology	Scenario	Task
[116]	RGB-D	Ray-tracing precomputed lookup table	Indoor	Current Beam Selection
[123]	LiDAR, Camera	LSTM	Indoor	Current Beam Selection
[137]	LiDAR, Camera	Deep neural network	V2V	Current Beam Selection
[138]	Camera	Deep neural network	V2I	Current Beam Selection
[139]	LiDAR, Camera, Radar, GPS	GRU, Power-based accuracy	V2I	Current Beam Selection
[140]	LiDAR, Camera, Radar, GPS	Transformer	V2I	Current Beam Selection
[55], [56]	LiDAR, GPS	CNN	V2I	Current Beam Selection
[141]	LiDAR, Camera, IMU, mmWave	LSTM	V2I	Current Beam Selection
[142]	LiDAR, GPS, Camera	CNN with dynamic input fusion	V2I	Current Beam Selection
[143]	LiDAR, GPS, Camera	MLP	V2I	Current Beam Selection
[144]	LiDAR,Radar	Cross-modal knowledge distillation	V2I	Current Beam Selection
[145]	LiDAR, GPS, Camera	Cooperative edge inference	V2I	Current Beam Selection
[146]	LiDAR	Federated learning	V2I	Current Beam Selection
[147], [148]	LiDAR, GPS, Camera	Federated learning	V2I	Current Beam Selection
[149]	LiDAR, GPS	Split learning	V2I	Current Beam Selection
[129]	Camera	Ray-tracing	Urban	Current Beam Selection
[150]	Camera	FusionNet	Urban	Current Beam Selection
[151]	LiDAR	RNN	Indoor	Future Beam Prediction
[152]	LiDAR	GRU	V2I	Future Beam Prediction
[153]	LiDAR	LSTM	V2I	Future Beam Prediction
[154]	LiDAR, GPS	GRU with PointNet	V2V	Future Beam Prediction
[155]	LiDAR, Radar, Camera, GPS	LLM	V2I	Future Beam Prediction
[121]	RGB-D	Ray-tracing	Multi-user	Beamforming Optimization
[140]	LiDAR, Camera, GPS	Transformer	Multi-user	Beamforming Optimization
[156]	LiDAR, Camera, GPS	Deep neural network	Multi-user	Beamforming Optimization
[124]	LiDAR	GNN with ray-tracing	Single-user	Beamforming Optimization

TABLE IX: Summary of studies on point cloud aided beam management

CNNs, fused with user identity inputs, and fed into an LSTM-based beam classifier. Additionally, a transfer learning strategy enables the pre-trained model to adapt to a new factory with only 30% labeled data, achieving 94% Top-10 beam prediction accuracy and reducing training cost by 75%.

b) Vehicular Scenarios: Vehicular scenarios encompass both V2V and V2I paradigms, which are essential components of intelligent transportation systems. Beam selection in these settings is particularly challenging due to rapidly changing geometries, dynamic occlusions, and mobility.

V2V: In V2V communication, both Tx and Rx are mobile, necessitating low-latency beam selection strategies. Point cloud-based methods offer an opportunity to dynamically capture surrounding vehicles and environments. In [137], Mollah *et al.* integrated LiDAR, RGB images, and vehicle kinematics to train a unified deep neural network capable of predicting top-M beam candidates for various V2V relative positions. In addition to improving beam selection accuracy, this method reduces beam sweeping space and time by 79.67% and 91.89%, respectively.

V2I: In V2I communication, the Tx or Rx is typically part of

a fixed infrastructure, such as an RSU or BS, while the other resides on a moving vehicle. Recent studies have explored various paradigms, including centralized, decentralized, and cooperative frameworks, that leverage point cloud sensing for beam selection in V2I systems.

- *Infrastructure-Located Sensing:* For infrastructure-located sensing, beam selection can be performed centrally by leveraging complete environmental awareness. For example, Xu *et al.* utilized panoramic 3D scene reconstructions generated from camera images and a deep neural network to predict beam indices at the BS side in [138]. This method avoids the high cost of LiDAR by leveraging offline photogrammetry to create the point cloud, enabling a scalable and low-cost solution for static infrastructure. The work [139] by Vuckovic *et al.* presents a fusion framework with GRU that jointly processes camera, LiDAR, radar, and GPS data collected from a BS to classify beam indices. Importantly, it also revises evaluation metrics by proposing power-based accuracy (Power Ratio and Top-K), which better reflects communication performance. Similarly, Cui *et*

al. introduced a transformer-based approach in [140] that processes multimodal point cloud derived features from BS sensors. Their method uses 3D-ResNet for spatial encoding and merged-attention transformers for final beam selection, achieving 91.59% top-3 accuracy in high-mobility environments.

- **Vehicle-Based Sensing:** In vehicle-based sensing, beam selection is carried out using onboard sensors, enabling decentralized deployment. Dias *et al.* further demonstrated in [125] that purely vehicle-side sensing is particularly effective under partial V2X deployment, where infrastructure-based perception is unavailable. Klautau *et al.* [55] and Zecchin *et al.* [56] transformed raw LiDAR-generated point cloud into 2D histograms or voxelized tensors for CNN-based top-k beam classification with non-local attention modules. Their models rely exclusively on vehicle-mounted LiDAR and GPS, enabling decentralized deployment. Zhu *et al.* presented a practical, low-cost mmWave testbed that integrates LiDAR point cloud with multiple sensors, including camera, mmWave, and inertial measurement unit (IMU), on a remote-controlled car to collect real-world data for LSTM-based beam alignment in [141]. The Omni-CNN model [142] extends this idea by supporting dynamic combinations of inputs, including LiDAR, camera, and location, through a shared encoder design that adapts to available modalities while maintaining prediction robustness. To enhance efficiency, Niu *et al.* proposed a principled approach to reduce redundancy in multi-modal data, including vehicle-mounted LiDAR point clouds in [143]. The model extracts LiDAR features using CNNs, projects them into lower-dimensional representations via PCA, and fuses them with image and GPS data for efficient beam classification using a lightweight MLP. Park *et al.* further enhanced efficiency through cross-modal knowledge distillation, training radar-only lightweight models under the supervision of LiDAR-rich teacher networks in [144]. High-resolution point clouds guide feature learning during training, whereas inference relies solely on radar data, significantly reducing sensor cost, model size, and computation while maintaining accuracy. This approach enables near real-time beam prediction on resource-constrained edge platforms while implicitly leveraging point cloud structure.
- **Joint Infrastructure-Vehicle Sensing:** Joint sensing frameworks combine the infrastructure's global environmental awareness with the vehicles' high-resolution local perception. The comparison between centralized and distributed inference paradigms in [145] shows that a cooperative edge inference model, trained on GPS, LiDAR, and camera data, achieves high accuracy even under incomplete sensor inputs. They also develop a fusion-based distributed optimal beam index inference model tailored for V2I edge systems, supporting flexible deployment either on vehicles or MEC servers. Considering federated learning, [146] introduces a multi-vehicle federated learning framework in which vehicles locally train beam selection models using onboard LiDAR point clouds

and periodically synchronize with a central server. This architecture eliminates the need to transmit raw sensor data, thereby preserving privacy and reducing bandwidth requirements. In [147], [148], Salehi *et al.* extended this to federated sector selection with model pruning, achieving a 52% beam alignment latency reduction compared with IEEE 802.11ad sweeping. Nevertheless, [149] adopts a split learning architecture where the model is partitioned between the vehicle and the BS. The vehicle processes raw LiDAR and positional data through the encoder, while the BS completes the forward pass, ensuring both data privacy and joint model training. This method also offers strong resilience to non-IID data across users.

2) *Search-efficient beam selection:* In contrast, search-efficient beam selection does not necessarily determine the final beam index in a single step. Instead, it improves the efficiency of current beam acquisition by reducing the search space or sweeping overhead. These methods typically exploit point-cloud-derived scene geometry, environment-aware beam probing, or adaptive codebook design to narrow the candidate beam set before or during beam selection.

Recent advances have explored point cloud-aided approaches to accelerate this beam selection process. Kamari *et al.* proposed a method that generates reflection-aware 3D maps of urban environments from street-view images in [129]. By performing material segmentation and ray-based projection, the system constructs a point cloud-derived environment model offline. At runtime, vehicles perform environment-driven hierarchical ray tracing and restricted beam probing to select high-SNR beams with minimal channel measurements. This approach reduces the number of probed beams from 64 to 6, achieving less than 1 dB SNR loss while improving throughput and reducing latency by over 70%.

Point cloud data can also enhance codebook-based beam selection by enabling environment-adaptive beam codebooks tailored to site-specific propagation. Chen *et al.* proposed a computer vision-aided codebook design framework in [150] where different strategies are used for LOS and NLOS users. For NLOS users, the authors explicitly incorporate 3D environmental features obtained from point clouds, including architectural structure and terrain geometry. These features are processed by a deep learning model (FusionNet) to construct codebooks tailored to the surrounding environment without requiring measured CSI. This approach reduces system overhead while significantly improving beam coverage in NLOS conditions.

B. Future Beam Prediction

Future beam prediction aims to anticipate the optimal beam indices several time steps ahead by leveraging temporal sequences of point clouds. By modeling the spatiotemporal evolution of motion and environmental dynamics, these methods reduce reliance on reactive beam sweeping and improve link robustness. Most approaches employ RNNs, such as GRUs or LSTMs, to capture temporal dependencies across sequential frames.

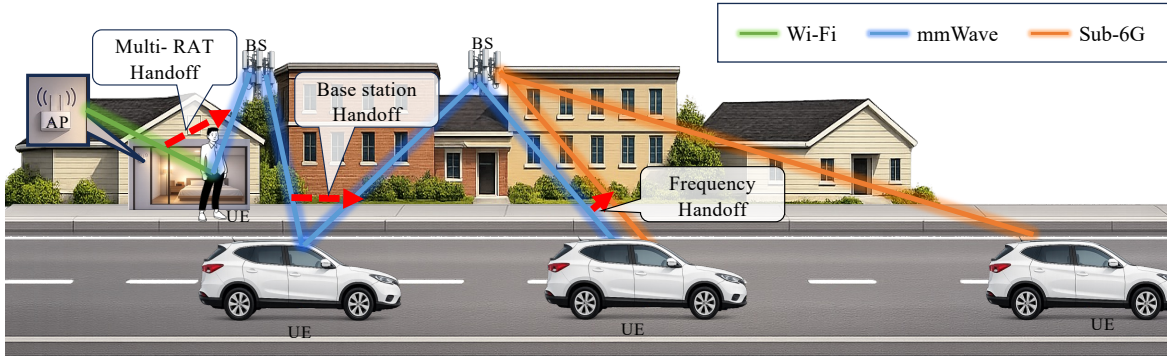


Fig. 7: Illustration of representative proactive handoff scenarios in dynamic wireless environments, including multi-RAT handoff, base station handoff, and frequency handoff.

In [151], Marasinghe *et al.* proposed an RNN-based approach to capture user movement from LiDAR-generated point cloud. This enables proactive inference of user beam transitions in indoor environments, achieving comparable accuracy (88.7%).

In vehicular scenarios, Jiang *et al.* [152] and Rinchi *et al.* [153] utilized sequential LiDAR frames as the primary input to GRU- and LSTM-based models, respectively. These models learn to map environmental motion cues into beam index trajectories without requiring CSI. While Jiang *et al.* focused on a fully LiDAR-only input pipeline and demonstrated 95% top-1 accuracy on real-world V2I scenarios, Rinchi *et al.* targeted high-signal-to-noise ratio (SNR) beam prediction using codebook constraints, achieving over 84% accuracy in simulated environments. Both confirm that temporal sequences of point cloud frames can support highly accurate multi-step beam forecasting. For the V2V scenario, Bian *et al.* utilized 3D LiDAR point cloud and GPS data collected from vehicles to perform beam prediction in [154]. The method employs PointNet to process spatial features, combined with a GRU network to capture temporal dependencies, achieving over 92% top-20 prediction accuracy on DeepSense 6G highway datasets.

Recent work extends future beam prediction to multimodal and large language model (LLM)-assisted frameworks. Zheng *et al.* proposed M2BeamLLM [155], which fused LiDAR point clouds, radar, RGB images, and GPS data to predict multi-step beam indices in V2I scenarios. Heterogeneous features are aligned through a multimodal encoder and fine-tuned on a GPT-2 backbone, enabling the model to reason over spatiotemporal dynamics and generate future beam sequences without explicit sweeping. Experiments on the DeepSense 6G dataset show that M2BeamLLM outperforms RNN- and transformer-based baselines, achieving up to 68.9% Top-1 accuracy while maintaining robustness in few-shot and sensor-missing conditions.

C. Beamforming Optimization

Point clouds enable geometry-aware enhancements of beamforming strategies by providing accurate spatial information about obstacles, reflectors, and user distribution. By leveraging

these spatial cues, beamforming strategies can be optimized to enhance link reliability and reduce interference.

In [121], Kamari *et al.* constructed a material-aware 3D map by combining RGB-D derived point cloud and object segmentation, enabling accelerated ray-tracing-based beamforming in multi-user urban scenarios. Cui *et al.* [140] developed a transformer-based deep fusion network that aggregates features from point cloud, image, and GPS data to support high-reliability beamforming under high-speed mobility. Patel *et al.* also demonstrated that predicting beamspace AoDs from point cloud-based multimodal sensing enhances overall beamforming and minimizes inter-user interference in multi-user settings in [156]. In their framework, LiDAR point clouds fused with GPS and camera data are processed by a deep neural network to estimate each user's beamspace channel representation.

In addition to digital beamforming, analog or hybrid precoding can also benefit from point cloud-assisted environmental modeling, reducing reliance on conventional CSI acquisition and exhaustive beam training. Nerini *et al.* [124] proposed a physics-based framework that employs LiDAR point clouds to reconstruct the environment, performs localized ray tracing to generate channel estimates, and applies GNN for precoder selection.

D. Discussion

Current beam selection, future beam prediction, and beamforming optimization address different objectives but all use point cloud as a source of geometry-aware prior information. Current beam selection infers the best beam for the present channel state. Future beam prediction extends this to the temporal domain for proactive adaptation. Beamforming optimization incorporates point cloud into beam design under environment-dependent propagation conditions. Representative studies are summarized in Table IX. Overall, point cloud enables beam management to shift from blind beam exploration toward geometry-aware beam inference.

Across these sub-tasks, the reviewed methods also differ in how point cloud information is acquired and represented. Infrastructure-mounted sensors [138], [140] capture the scene from a fixed and stable viewpoint, providing consistent spatial

context for beam decisions. Vehicle-mounted sensors [55], [56] capture the local surroundings from the user side, offering higher resolution of nearby objects but with a continuously changing perspective. The way point cloud is encoded also affects how effectively its geometric information is preserved. 2D projection [55] discards depth but enables efficient processing. Voxelization [56] retains 3D structure at moderate cost. Direct point set processing [154] preserves full geometry but scales less efficiently. These choices determine how much of the spatial prior provided by the point cloud is actually available to the beam management model.

VI. POINT CLOUD FOR PROACTIVE HANDOFF

While Section IV and Section V focus on mitigating the impact of obstacles through detection, prediction, and dynamic beam adaptation, these strategies do not always prevent link degradation when the existing path becomes untenable. Proactive handoff provides a complementary approach, maintaining seamless connectivity by preemptively switching users to alternative communication resources, such as different frequencies, BSs, or radio interfaces, before the current link fails.

By leveraging the spatial and temporal fidelity of the point cloud, proactive handoff can anticipate mobility-induced link disruptions and environmental blockages with high precision. This enables networks to trigger handoff events in advance, reducing latency and maintaining reliable connectivity in highly dynamic scenarios.

This section examines point cloud-enabled proactive handoff in three key domains:

- **Base Station Handoff:** Anticipatory reassignment to neighboring base stations guided by user trajectory and blockage prediction.
- **Frequency Handoff:** Switching to alternative carrier frequencies before signal degradation occurs.
- **Multi-RAT Handoff:** Coordinated transitions across heterogeneous radio access technologies to sustain robust links.

Fig. 7 provides a visual overview of the representative proactive handoff scenarios discussed in this section. Specifically, a user leaving the house may experience a multi-RAT handoff from indoor Wi-Fi to outdoor mmWave access, while a moving vehicle may first undergo a proactive base station handoff between neighboring BSs and later a frequency handoff from mmWave to sub-6 GHz as the link quality deteriorates. A summary of representative point cloud-assisted proactive handoff methods is provided in Table X.

A. Base Station Handoff

Base station handoff is essential in dynamic and dense environments, ensuring that users maintain seamless connectivity as they transition between serving BSs. By leveraging point cloud, proactive handoff mechanisms can predict and mitigate link degradation, enabling timely transitions and sustaining robust communication links. Existing studies have explored methodologies tailored to vehicular, urban, and indoor scenarios, demonstrating the adaptability of point cloud-assisted solutions across diverse environments.

1) *Vehicular Scenarios:* High-mobility vehicular environments pose significant challenges for handoff due to rapid topology changes and frequent blockages. Koda *et al.* integrated depth cameras and deep reinforcement learning to process high-dimensional sensory data in [157]. This method enables precise handover decisions by real-time blockage prediction that address blind spots during fast transitions. Similarly, the COPILLOT framework proposed by Pradhan *et al.* in [128] fuses LiDAR point clouds from connected vehicles using attention-based mechanisms to predict transient blockages. This facilitates proactive V2I handoffs between RSUs, improving coverage and link reliability in highly dynamic road networks.

2) *Urban Scenarios:* Urban environments are characterized by complex geometries, frequent LOS interruptions, and high pedestrian and vehicle activity, demanding more sophisticated handoff strategies. Wu *et al.* employed infrastructure-mounted LiDAR systems to monitor urban intersections in [130], using GBDT to process point cloud and predict future blockages, thereby enabling timely BS transitions in crowded conditions. Building on this, Charan *et al.* combined RGB imagery with LiDAR point clouds to identify disruptions caused by moving pedestrians and vehicles in [115]. Neural network-based fusion of these modalities supports proactive handoff decisions in dense urban areas with frequent LOS changes.

3) *Industrial Scenarios:* In indoor industrial settings such as smart factories, frequent layout reconfigurations and moving equipment can disrupt wireless connectivity. Asano *et al.* [127] applied CNNs to LiDAR-generated point clouds to predict blockages and trigger proactive BS handoffs, maintaining reliable communication in industrial environments with rapidly changing obstacle configurations. Similarly, Ohta *et al.* [158] employed voxelized point cloud modeling to capture LOS conditions and predict link degradation caused by human or robotic movement. Their system enables timely BS transitions, ensuring uninterrupted connectivity in smart factory deployments.

B. Frequency Handoff

Frequency handoff is essential for maintaining reliable connectivity as networks adapt to varying propagation conditions and dynamic blockages. By leveraging point cloud, frequency handoff mechanisms can proactively trigger transitions between frequency bands, ensuring optimal connectivity in challenging scenarios.

Zhang *et al.* [120] proposed a dual-platform framework that employs both stereo cameras and LiDAR to support handoffs between mmWave and sub-6 GHz bands. The camera-based subsystem utilizes YOLOX object detection and RNNs to predict LOS blockages and initiate frequency transitions, while the LiDAR-based platform operates robustly in low-light conditions. This multi-sensor design enables continuous communication by dynamically switching frequencies based on predicted link degradation, demonstrating effectiveness in dense urban deployments.

C. Multi-RAT Handoff

Multi-Radio Access Technology (Multi-RAT) handoff enhances connectivity by dynamically switching between heterogeneous networks [159], such as mmWave, Wi-Fi, and cellular links. Point cloud augments these strategies by providing fine-grained spatial awareness, enabling networks to anticipate link degradation and trigger seamless transitions across RATs.

Zhang *et al.* introduced a framework that combines full-body pose tracking with point cloud to predict link disruptions and facilitate multipath networking in [160]. By leveraging stereo cameras and VR headset tracking systems, the framework captures dynamic spatial correlations in user motion, enabling adaptive handoffs between mmWave and Wi-Fi networks. A Transformer-based sequence-to-sequence model is employed to predict link quality and guide proactive handoff decisions, achieving up to a 29% improvement in throughput compared to baseline models. This approach proves particularly effective in immersive content delivery scenarios, where seamless transitions between heterogeneous radios are essential to maintaining high Quality of Experience.

In vehicular networks, Okada *et al.* developed a multi-radar system that fuses point clouds via a Kalman filter to expand sensing range and predict vehicle trajectories [126]. The system proactively initiates handoffs by transitioning mmWave communication to macro BSs before vehicles enter predicted blockage zones, significantly reducing service disruption in high-mobility environments.

D. Discussion

Point cloud enables proactive handoff by providing spatial awareness of user mobility and environmental obstructions before link degradation occurs. This allows handoff decisions to shift from reactive, signal-triggered switching to predictive, geometry-informed transitions.

The three handoff types exploit point cloud information at different granularities. Base station handoff uses point cloud primarily for object-level trajectory prediction, requiring moderate spatial resolution to track major blockers such as vehicles and buildings [128], [130]. Frequency handoff relies on point cloud more for robust environmental awareness under varying sensing conditions, where sensing reliability matters more than spatial resolution [120]. Multi-RAT handoff places the highest demand on point cloud fidelity, as fine-grained body pose tracking at millimeter scale is needed to predict link quality across heterogeneous radio interfaces [160].

VII. OTHER APPLICATIONS

Beyond channel modeling, blockage management, beam management, and proactive handoff, point cloud also supports a broad range of wireless communication applications. These include semantic communication, localization, metasurface optimization, and digital twin construction, where 3D spatial awareness significantly enhances network intelligence and robustness.

A. Semantic Communication

In semantic wireless communication, the focus shifts from bit-level accuracy to preserving task-relevant meaning and end-to-end performance [161]. Zhang *et al.* proposed a point cloud-augmented channel semantic model that couples environmental perception with wireless propagation across three hierarchical semantic layers in [162]:

- **Status Semantics:** Multipath components are semantically annotated by segmenting the point cloud and clustering delay–power profiles, associating paths with physical scatterers (“tree at 36.9 m”) and instantaneous channel features.
- **Behavior Semantics:** Temporal evolution of scatterer clusters (appearance, disappearance, acceleration, deceleration) is tracked to model dynamic channel behaviors.
- **Event Semantics:** Sequences of behaviors are aggregated into high-level events (e.g., “turning at intersection” or “crossing roadway”), providing contextual reasoning for channel variation.

By leveraging 3D point clouds to localize and classify scatterers, it enables:

- **Adaptive Semantic Compression:** Selecting task-critical features to reduce redundant transmission.
- **Robust Inference:** Using temporal consistency in behavior semantics to mitigate transient occlusions and blockages.
- **Event-Driven Resource Allocation:** Dynamically adjusting bandwidth and sensing rates in response to detected events.

This semantic approach integrates environmental understanding into wireless communication, paving the way for context-aware 6G systems.

B. Wireless Localization

Wireless localization aims to determine the position of mobile devices or targets by exploiting radio signal properties. In multipath-rich environments, traditional time-of-arrival, angle-of-arrival, or RSS methods are prone to ambiguity and NLOS-induced errors [163]. To mitigate these ambiguities, point clouds provide explicit geometric information about the propagation environment. It can be derived either from externally acquired point clouds, such as LiDAR- or vision-derived point clouds, or from radio-generated point clouds directly reconstructed by the BS through mmWave/radar sensing.

When point clouds are acquired by external sensors, they mainly serve as environmental priors for localization. For instance, in [164], LiDAR is co-located with a mmWave uplink to a reconfigurable intelligent surface (RIS). After filtering out irrelevant ground and building points, scatterer clusters are extracted, and their centroids inform a compressive-sensing localization algorithm. In this way, the point cloud does not directly output the user position, but improves UE localization by reducing the number of geometric variables to be estimated in multipath environments. The reported results show a 65% reduction in localization error compared with the standard compressed sensing scheme. Externally acquired point clouds can also enhance fingerprint-based

Paper	Sensor	Method	Key Contribution	Scenarios	Issue
[157]	Depth Cameras	Deep reinforcement learning with cooperative sensing to map sensory data to handover decisions	Compensates for blind spots; ensures optimal handoffs in vehicular scenarios	Vehicular	Base station handoff
[128]	LiDAR	LiDAR fusion with attention mechanisms for transient blockage prediction	Enhances coverage and reliability in high-speed V2I networks	Vehicular	Base station handoff
[130]	LiDAR	GBDT applied to LiDAR data to predict future blockages	Predicts blockages in crowded urban settings; enables seamless handoffs	Urban	Base station handoff
[115]	Camera and LiDAR	Neural networks for data fusion from RGB images and point cloud data	Proactively identifies disruption caused by urban pedestrians and vehicles	Urban	Base station handoff
[127]	LiDAR	CNNs applied to LiDAR-generated point clouds for proactive handoff	Ensures uninterrupted communication in reconfigurable industrial layouts	Industrial	Base station handoff
[158]	LiDAR	Voxelized LiDAR data for link quality prediction in LOS	Facilitates timely transition in factory setups	Industrial	Base station handoff
[120]	LiDAR and Stereo Cameras	Dual-platform approach with YOLOX-based camera detection and LiDAR-supported low-light decisions	Maintains continuous communication by dynamically switching frequencies based on blockage prediction	Urban	Frequency handoff
[160]	Stereo Camera	Full-body pose tracking with sequence-to-sequence learning models for proactive handoff decisions	Improves throughput by up to 20%; enables seamless transitions between mmWave and Wi-Fi networks in immersive content delivery	VR	Multi-RAT Handoff
[126]	Radar	Multi-radar system with Kalman filter for trajectory prediction and preemptive frequency transitions	Predicts vehicle trajectories and ensures seamless handoffs to macro BS in high-mobility scenarios	Vehicular	Multi-RAT Handoff

TABLE X: Summary of Proactive Handoff Techniques in Different Applications

localization. In [165], features extracted from LiDAR point clouds within the first Fresnel zone between the transmitter and receiver are used to characterize obstacle distribution in the environment and condition a diffusion-based fingerprint generation model. Rather than directly inferring user position from point clouds, this approach exploits point-cloud-aware environmental descriptors to synthesize dense CIR fingerprint databases for new environments with minimal real data, thereby improving the sample efficiency and robustness of fingerprint-based localization.

When point clouds are reconstructed by the wireless infrastructure itself, localization no longer relies on external sensing modalities. Instead, the BS or ISAC transceiver acts as both a communication node and a sensing node, generating point-cloud-like environmental representations from mmWave/radar measurements. For example, Rastorgueva-Foi *et al.* [166] developed an end-to-end mmWave radio SLAM framework that jointly estimates user equipment positions and environmental landmarks from 60 GHz cellular measurements. Du *et al.* [167] further showed that active mmWave sensing in 5G NR can generate point clouds in a manner analogous to laser scanning, and then use point cloud registration, loop closure detection, and pose graph optimization for terminal localization. In this way, radio-generated point clouds support joint environment reconstruction and user localization within the wireless system

itself.

C. Metasurface Coverage Optimization

In mmWave networks, passive metasurfaces offer a cost-effective solution to mitigate coverage blind spots by redirecting and phase-tuning incident signals. However, their effectiveness depends critically on precise placement and orientation relative to the surrounding environment [168]. Point cloud data, captured via LiDAR, provides high-resolution geometric information about walls, furniture, and other dominant surfaces. This geometric context enables accurate extraction of surface planes and normal vectors, and when combined with material databases or minimal calibration, allows estimation of per-surface mmWave reflection coefficients.

In the AUTOMS framework [169], point clouds play a central role in guiding the optimization process. The processed 3D scene is input to a GPU-accelerated ray-tracing engine that simulates realistic multipath propagation, including delay, angle of arrival, and attenuation. Based on this physically grounded model, AUTOMS jointly optimizes beam-forming codebooks, metasurface placement, and per-element phase configurations to maximize signal strength in predefined blind-spot regions. Experimental results show that this point-cloud-assisted approach achieves 9–12 dB improvements in

RSS in typical indoor environments, while enabling ultra-low-cost (\$1–3 per panel) and precise metasurface deployment without manual surveying or iterative measurement campaigns.

D. Digital Twin

Digital twins for 6G wireless networks rely on a continuously updated, high-fidelity replica of the physical environment where accurate geometry and material properties are essential for reliable channel prediction and network control. In this context, point clouds serve as the geometric backbone from which virtual environment meshes are reconstructed and annotated with location, size, and material parameters. These point-cloud-based virtual replicas then feed into predictive channel-modeling engines where per-scatterer features such as position and surface normals enable real-time forecasting of channel impulse responses, power-delay profiles, and angular spectra across sub-6 GHz to THz bands [170]. Crucially, the digital twin remains synchronized with the live network via periodic or event-triggered point-cloud updates, capturing dynamic changes like moving objects or new obstructions; this continuous loop of sensing, model update, and action (adaptive beam selection or RIS reconfiguration) underpins sub-millisecond decision support and autonomous, context-aware operation in future 5G/6G systems [171], [172].

VIII. OPEN CHALLENGES AND FUTURE RESEARCH DIRECTIONS

Despite the promising progress reviewed in the previous sections, point-cloud-aided wireless communications are still far from practical and large-scale deployment. Most existing studies have demonstrated encouraging gains in specific scenarios, but several fundamental challenges remain unresolved. These challenges include the real-time processing of large-scale point clouds, robustness to environmental variations, multimodal fusion and calibration, generalization across diverse scenarios, the extension from single-user to multi-user wireless settings, the lack of standardized benchmarking and open datasets, and associated privacy and security risks. Addressing these issues will be essential for transforming point clouds from an auxiliary sensing input into a reliable and scalable foundation for future environment-aware wireless systems. In the following, we discuss these open problems and highlight corresponding research directions.

A. Real-Time Processing of Large-Scale Point Clouds

One of the most fundamental challenges lies in the efficient processing of large-scale point clouds for time-sensitive wireless tasks. High-resolution LiDAR, RGB-D sensors, and radar-based 3D sensing can provide detailed geometric information, but they also generate large volumes of raw data. These data require processing before being used for blockage prediction, beam management, or proactive handoff, and the associated processing burden may directly affect communication performance.

Several studies have quantified this issue. In [145], sensor-data delivery, processing, and control signaling require approximately 3.662 ms, while the subsequent beam sweeping requires 0.94–2.04 ms. The resulting end-to-end beam-selection delay is 4.6–5.71 ms and retains 95.37–97.95% of the relative throughput, compared with 145 ms for exhaustive 5G-NR beam selection. In [148], model inference requires approximately 0.6 ms, and the end-to-end sector-selection delay is up to 3.6 ms after including sector reporting and baseband processing. This delay exceeds the 2.65 ms exhaustive-search baseline for small 5G-NR codebooks, but remains below the 64.914 ms delay reported for larger 5G-NR codebooks. These results indicate that the same point-cloud-aided processing pipeline may provide clear latency reduction when the avoided communication overhead is large, but little or no net benefit when the baseline search cost is already low.

A complete real-time evaluation should therefore compare the full sensing-to-control delay with the communication gain that the point-cloud-aided method is expected to provide. This delay includes sensor acquisition, point-cloud preprocessing, data or feature transmission, model inference, and communication-control execution. Future research should develop task- and time-aware point-cloud pipelines that balance the required environmental information with the available sensing and processing budget. Relevant directions include adaptive sensing rates and resolutions, communication-oriented processing, sparse or hierarchical representations, and hardware-efficient inference. Future evaluations should report end-to-end latency, computational complexity, and hardware configuration, along with communication metrics such as SNR, throughput, beam selection accuracy, and beam training delay.

B. Robustness to Environmental Variations

A major challenge for point-cloud-aided wireless communications is that the quality of sensed 3D geometry can degrade significantly under adverse environments. In practice, rain, fog, snow, poor illumination, reflective surfaces, occlusion, and fast motion may introduce missing points, false reflections, sparse returns, or geometric distortion into raw point cloud data [15], [16], [173]. Since many existing methods rely on accurate spatial structure to support propagation modeling, blockage prediction, beam selection, and proactive handoff, degradation at the sensing stage can directly propagate to downstream wireless tasks and reduce their reliability. Therefore, robustness in this context is not only a sensing issue, but also a system-level challenge concerning how environmental uncertainty is transferred from perception to communication decisions.

Future research should address this problem along the full pipeline from sensing to wireless control. First, more robust point cloud acquisition and preprocessing methods are needed to suppress weather-induced noise, recover incomplete geometry, and preserve wireless-relevant structures under adverse conditions. Second, downstream wireless algorithms should explicitly account for sensing uncertainty, rather than assuming point cloud inputs are always accurate. For example, blockage prediction, beam management, and handoff control may

need to adapt their decisions according to the confidence of perceived geometry. Third, future studies should move beyond reporting performance under nominal conditions and establish robustness-oriented evaluations under weather variation, partial occlusion, sparse sensing, and sensor degradation. Such efforts would make point-cloud-aided wireless methods more reliable in realistic and continuously changing environments.

C. Multimodal Fusion and Calibration

Point-cloud-aided wireless communication is inherently related to multimodal sensing. As discussed in Section II, point clouds themselves can be acquired from different sensing modalities, including photogrammetry, LiDAR, structured light scanners, and radar. Although these sensing sources may ultimately produce a common geometric representation, they differ substantially in sampling density, spatial resolution, sensing range, noise characteristics, and environmental sensitivity. In addition, many wireless communication tasks do not rely on point clouds alone, but also combine them with other information sources such as RGB images, depth maps, GPS, IMU data, radar features, or communication-side observations [105], [140], [142], [155], [156]. Therefore, multimodal fusion in this context should be understood more broadly as the joint utilization of heterogeneous sensing sources, including both different sources of point clouds and point clouds combined with non-point-cloud modalities.

Future research should therefore move beyond simple feature concatenation and develop more principled multimodal fusion frameworks for point-cloud-aided wireless systems. First, systematic methods are needed to align and fuse heterogeneous point cloud sources, such as LiDAR point clouds and radar point clouds, as well as to connect point cloud representations with image-based, location-based, and communication-side information. Second, future studies should clarify which fusion stage is most appropriate for different wireless tasks, including raw-data fusion, geometric-level fusion, intermediate feature fusion, and decision-level fusion. Third, the benefit of multimodal fusion should be evaluated together with its system cost, since richer sensing may improve performance but also increase calibration burden, processing latency, and deployment complexity. Addressing these issues will be important for building multimodal and practically scalable point-cloud-aided wireless communication systems.

D. Generalization Across Diverse Scenarios

The literature also shows that many point-cloud-aided wireless methods are evaluated in specific indoor layouts, selected outdoor routes, or custom simulation environments. Although such evaluations are useful for demonstrating feasibility, they also reveal that model generalization remains limited. A method trained for a particular sensing setup, propagation environment, or carrier frequency may not transfer well to unseen scenarios [81], [95], [123], [174]. This issue becomes even more important across frequency bands, since the relevance of environmental details may differ substantially from sub-6 GHz to mmWave and further to THz systems.

This challenge is important because the long-term value of point-cloud-aided wireless communication depends not only on site-specific optimization, but also on its ability to generalize across heterogeneous environments. Several directions therefore deserve further investigation, including self-supervised representation learning, synthetic-to-real transfer, few-shot adaptation, and distributed or federated updating across sites [123], [147]–[149]. More broadly, future studies should pay greater attention to identifying which geometric features are task-specific and which may remain informative across indoor, outdoor, mmWave, and THz settings. Improving such transferability will be essential if point-cloud-aided methods are to move beyond carefully curated demonstration scenarios.

E. Extension to Multi-User Wireless Settings

A further observation from the surveyed literature is that many existing frameworks are formulated in a single-user or single-link manner. In many beam management, blockage prediction, and handoff studies, point clouds are used to infer the communication state of one target user at a time. By contrast, comparatively fewer works explicitly consider multi-user selection, multi-user beamforming, or multi-user CSI learning [57], [156], [175]. Recent work has further identified user identification as a key prerequisite for extending vision- or sensing-aided wireless communication from single-user settings to crowded multi-user scenarios [176]. This suggests that the transition from single-user case studies to dense multi-user systems remains relatively underexplored.

This gap is important because practical wireless networks must serve multiple users simultaneously under interference, competition for resources, and partially overlapping mobility patterns. Future work may therefore move beyond per-user prediction and investigate user association in shared sensing scenes, interference-aware beamforming, coordinated beam switching, and load-aware mobility management. Extending point-cloud-aided wireless communication from single-user to multi-user demonstrations would substantially improve the practical relevance of this research direction.

F. Benchmarking and Open Datasets

Another issue in this field is that a commonly adopted benchmarking framework is still lacking. Although some public datasets have become available for point-cloud-aided wireless communications, existing studies still differ substantially in baseline settings, data splits and evaluation metrics. Moreover, most methods are evaluated only on a single dataset, while cross-dataset comparison is still very limited. As a result, results reported in different works are still difficult to compare under a unified evaluation framework.

This gap is even more evident for point-cloud-aided channel modeling. Open datasets for propagation channel modeling remain much more limited, especially in indoor and measurement-driven settings. Many existing studies on indoor channel modeling are still based on scene-specific measurements, proprietary reconstructions, or custom simulation pipelines, rather than widely shared open datasets.

Future work should therefore proceed in two directions. First, more standardized benchmarking practices are needed, especially in terms of data splits, evaluation metrics, and cross-dataset evaluation protocols. Second, greater effort should be devoted to building open datasets for point-cloud-aided channel modeling, particularly for indoor environments. Ideally, such datasets should cover diverse room topologies, material compositions, human blockage conditions and transceiver locations, together with aligned point clouds and channel measurements, so as to better support reproducible evaluation and cross-scenario generalization.

G. Privacy and Security

The use of point clouds introduces two concerns for practical point-cloud-aided wireless communication systems. The first is privacy. Detailed geometric sensing data may contain sensitive information, such as indoor layouts, user locations, and vehicle trajectories. Previous studies have demonstrated that point clouds should not be regarded as inherently privacy-preserving data [177], [178]. The second concern is security, as the sensing-to-communication pipeline itself may become a potential attack target. Existing LiDAR security studies have demonstrated both false-object injection and real-object removal through the manipulation of LiDAR returns [179], [180]. Manipulated or spoofed point clouds can provide an incorrect representation of surrounding objects and subsequently degrade wireless communication performance. Therefore, practical deployment requires ensuring both the privacy and security of point clouds used by wireless communication systems.

One future research direction is privacy protection by reducing the exposure of raw point clouds. At the representation stage, instead of sharing complete point clouds, point cloud anonymization, geometric obfuscation, or semantic abstraction can be combined with task-oriented processing so that only information required by a particular wireless task is retained. For example, the privacy-preserving localization method in [181] lifts 3D map points to parallel planes, such that the shared map provides only partial geometric constraints rather than complete scene information. A similar principle could be applied to wireless systems by sharing only communication-relevant geometric information. At the model training and inference stages, federated learning can keep locally collected point cloud data on the user side while exchanging model updates [146], whereas split learning partitions the model between the user and the network to avoid uploading raw sensing data during training and inference [149].

Another research direction is to improve the security of the sensing-to-communication pipeline. Future systems should verify the authenticity and integrity of sensing information before it is used for wireless communication tasks. Existing defenses provide useful starting points. For example, ADoPT detects point cloud spoofing by examining point-level temporal consistency across consecutive frames [182]. Cooperative observations may further help identify inconsistencies when the same environment is observed by different devices. Combined with secure transmission and authentication [2], such mechanisms

can help prevent corrupted or unauthorized sensing information from influencing wireless decisions.

Finally, privacy and security should become explicit dimensions of evaluation for point-cloud-aided wireless communications. In addition to conventional communication metrics, future benchmarks should assess how much sensitive information can be recovered from shared sensing representations and how wireless performance degrades under spoofed or manipulated sensing inputs. Protection mechanisms should also be evaluated across the complete sensing-to-communication pipeline. Establishing such evaluation procedures, together with appropriate data-governance and compliance requirements, will be important for the practical deployment of point-cloud-aided communications in future 6G networks.

IX. CONCLUSION

In this survey, we have reviewed the growing role of point cloud technology in wireless communications. We first introduced the fundamentals of point cloud for wireless communications, including acquisition techniques, preprocessing and representation methods, as well as open datasets and simulation tools. We then examined how the point cloud can enhance core wireless functionalities such as propagation channel modeling, blockage management, beam management, and proactive handoff. Additionally, we explored emerging applications extending beyond traditional communication paradigms, including semantic communications, wireless localization, metasurface coverage optimization, and digital-twin systems. Finally, we discussed several technical challenges, including the real-time processing of large-scale point clouds, robustness to environmental variations and sensing imperfections, multimodal fusion and calibration, generalization across diverse scenarios, the extension from single-user to multi-user wireless settings, the lack of standardized benchmarking and open datasets, and associated privacy and security risks together with possible directions for future research.

Looking ahead, point clouds are poised to become a key enabler of environment-aware and adaptive wireless systems. By providing fine-grained spatial awareness, they can help networks evolve from reactive to predictive, and from rigid to intelligent. We hope this survey provides a clear foundation for future work and inspires further exploration of point cloud-aided wireless communications.

REFERENCES

- [1] B. S. Khan, S. Jangsher, A. Ahmed, and A. Al-Dweik, "URLLC and eMBB in 5G industrial iot: A survey," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 1134–1163, 2022.
- [2] V.-L. Nguyen, P.-C. Lin, B.-C. Cheng, R.-H. Hwang, and Y.-D. Lin, "Security and privacy for 6G: A survey on prospective technologies and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2384–2428, 2021.
- [3] G. Qualcomm, "On representation format– dynamic point cloud representation," Tech. Rep. S4-241620, 3GPP, Aug. 2024. TSG-SA WG4 #129-e meeting draft.
- [4] D. Wen, Y. Zhou, X. Li, Y. Shi, K. Huang, and K. B. Letaief, "A survey on integrated sensing, communication, and computation," *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2024.
- [5] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE journal on selected areas in communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

- [6] X. Cheng, H. Zhang, J. Zhang, S. Gao, S. Li, Z. Huang, L. Bai, Z. Yang, X. Zheng, and L. Yang, "Intelligent multi-modal sensing-communication integration: Synesthesia of machines," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 1, pp. 258–301, 2023.
- [7] W. Jiang, Q. Zhou, J. He, M. A. Habibi, S. Melnyk, M. El-Absi, B. Han, M. D. Renzo, H. D. Schotten, F.-L. Luo, T. S. El-Bawab, M. Juntti, M. Debbah, and V. C. M. Leung, "Terahertz communications and sensing for 6g and beyond: A comprehensive review," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 4, pp. 2326–2381, 2024.
- [8] 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 Communications Surveys & Tutorials*, vol. 26, no. 3, pp. 1478–1519, 2024.
- [9] J. Li, C.-X. Wang, C. Huang, T. Qi, and T. Wu, "Digital twin online channel modeling: Challenges, principles, and applications," *IEEE Vehicular Technology Magazine*, vol. 20, no. 1, pp. 94–103, 2025.
- [10] A. Alkhateeb, S. Jiang, and G. Charan, "Real-time digital twins: Vision and research directions for 6G and beyond," *IEEE Communications Magazine*, vol. 61, no. 11, pp. 128–134, 2023.
- [11] H. Zhang, Z. Zhang, X. Liu, W. Li, H. Li, and C. Sun, "Integrated sensing and communication for 6G holographic digital twins," *IEEE Wireless Communications*, vol. 32, no. 2, pp. 104–112, 2025.
- [12] E. Zhu, H. Sun, and M. Ji, "Physics-informed generalizable wireless channel modeling with segmentation and deep learning: Fundamentals, methodologies, and challenges," *IEEE Wireless Communications*, vol. 31, no. 6, pp. 170–177, 2024.
- [13] L. Bai, Z. Huang, M. Sun, X. Cheng, and L. Cui, "Multi-modal intelligent channel modeling: A new modeling paradigm via synesthesia of machines," *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2025.
- [14] Q. Li, H. Huang, W. Yu, and S. Jiang, "Optimized views photogrammetry: Precision analysis and a large-scale case study in qingdao," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 1144–1159, 2023.
- [15] A. Haider, M. Pigniczki, S. Koyama, M. H. Köhler, L. Haas, M. Fink, M. Scharadt, K. Nagase, T. Zeh, A. Eryildirim, *et al.*, "A methodology to model the rain and fog effect on the performance of automotive lidar sensors," *Sensors*, vol. 23, no. 15, p. 6891, 2023.
- [16] M. Dreissig, D. Scheuble, F. Piwak, and J. Boedecker, "Survey on lidar perception in adverse weather conditions," in *2023 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1–8, IEEE, 2023.
- [17] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "Kinectfusion: Real-time dense surface mapping and tracking," in *2011 10th IEEE international symposium on mixed and augmented reality*, pp. 127–136, IEEE, 2011.
- [18] M. Sheeny, E. De Pellegrin, S. Mukherjee, A. Ahrabian, S. Wang, and A. Wallace, "Radiate: A radar dataset for automotive perception in bad weather," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1–7, IEEE, 2021.
- [19] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE transactions on signal processing*, vol. 57, no. 6, pp. 2275–2284, 2009.
- [20] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile, "Photogrammetry and optical methods in structural dynamics—a review," *Mechanical Systems and Signal Processing*, vol. 86, pp. 17–34, 2017.
- [21] B. Xu and C. Liu, "A 3D reconstruction method for buildings based on monocular vision," *Computer-Aided Civil and Infrastructure Engineering*, vol. 37, no. 3, pp. 354–369, 2022.
- [22] X. Ma, Z. Wang, H. Li, P. Zhang, W. Ouyang, and X. Fan, "Accurate monocular 3D object detection via color-embedded 3D reconstruction for autonomous driving," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 6851–6860, 2019.
- [23] G. Wang, X. Tian, R. Ding, and H. Wang, "Unsupervised learning of 3D scene flow from monocular camera," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4325–4331, IEEE, 2021.
- [24] J. Tian, B. Ma, S. Lu, B. Yang, S. Liu, and Z. Yin, "Three-dimensional point cloud reconstruction method of cardiac soft tissue based on binocular endoscopic images," *Electronics*, vol. 12, no. 18, p. 3799, 2023.
- [25] M. Ahrnbom, M. Nilsson, H. Ardö, K. Åström, O. Yastremska-Kravchenko, and A. Lareshyn, "Calibration and absolute pose estimation of trinocular linear camera array for smart city applications," in *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 103–110, IEEE, 2021.
- [26] L. Huang, F. Da, and S. Gai, "Research on multi-camera calibration and point cloud correction method based on three-dimensional calibration object," *Optics and Lasers in Engineering*, vol. 115, pp. 32–41, 2019.
- [27] M. Chen, Y. Tang, X. Zou, K. Huang, L. Li, and Y. He, "High-accuracy multi-camera reconstruction enhanced by adaptive point cloud correction algorithm," *Optics and Lasers in Engineering*, vol. 122, pp. 170–183, 2019.
- [28] H. Yu, L. Wang, J. Xu, and P. Y. Chiang, "A dToF ranging sensor with accurate photon detector measurements for lidar applications," *Sensors*, vol. 23, no. 6, p. 3011, 2023.
- [29] J. Ma, S. Zhuo, L. Qiu, Y. Gao, Y. Wu, M. Zhong, R. Bai, M. Sun, and P. Y. Chiang, "A review of tof-based LiDAR," *Journal of Semiconductors*, vol. 45, no. 10, p. 101201, 2024.
- [30] L. Fan, X. Xiong, F. Wang, N. Wang, and Z. Zhang, "Rangedet: In defense of range view for lidar-based 3D object detection," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 2918–2927, 2021.
- [31] R. Lange and P. Seitz, "Solid-state time-of-flight range camera," *IEEE Journal of quantum electronics*, vol. 37, no. 3, pp. 390–397, 2001.
- [32] M. Perenzoni and D. Stoppa, "Figures of merit for indirect time-of-flight 3D cameras: Definition and experimental evaluation," *Remote Sensing*, vol. 3, no. 11, pp. 2461–2472, 2011.
- [33] F. Chen, R. Ying, J. Xue, F. Wen, and P. Liu, "A configurable and real-time multi-frequency 3D image signal processor for indirect time-of-flight sensors," *IEEE Sensors Journal*, vol. 22, no. 8, pp. 7834–7845, 2022.
- [34] B. Cui, W. Tao, and H. Zhao, "High-precision 3d reconstruction for small-to-medium-sized objects utilizing line-structured light scanning: A review," *Remote Sensing*, vol. 13, no. 21, p. 4457, 2021.
- [35] J. Salvi, J. Pages, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern recognition*, vol. 37, no. 4, pp. 827–849, 2004.
- [36] Z. Zhang, "Microsoft kinect sensor and its effect," *IEEE multimedia*, vol. 19, no. 2, pp. 4–10, 2012.
- [37] M. I. Skolnik, *Radar Handbook*. McGraw-Hill, 3rd ed., 2008.
- [38] A. Stelzer, M. Jahn, and S. Scheiblhofer, "Precise distance measurement with cooperative FMCW radar units," in *2008 IEEE Radio and Wireless Symposium*, pp. 771–774, IEEE, 2008.
- [39] Y. Cheng, J. Su, M. Jiang, and Y. Liu, "A novel radar point cloud generation method for robot environment perception," *IEEE Transactions on Robotics*, vol. 38, no. 6, pp. 3754–3773, 2022.
- [40] G. Li *et al.*, "Comprehensive mPoint: A method for 3D point cloud generation of human bodies utilizing FMCW MIMO mm-wave radar," *Sensors*, vol. 21, no. 19, p. 6455, 2021.
- [41] X. Gao, S. Roy, and G. Xing, "MIMO-SAR: A hierarchical high-resolution imaging algorithm for mmWave FMCW radar in autonomous driving," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 7322–7334, 2021.
- [42] R. B. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," in *2011 IEEE International Conference on Robotics and Automation*, pp. 1–4, 2011.
- [43] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu, and M. Bennamoun, "Deep learning for 3d point clouds: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 12, pp. 4338–4364, 2021.
- [44] M.-J. Rakotosaona, V. La Barbera, P. Guerrero, N. J. Mitra, and M. Ovsjanikov, "Pointcleannet: Learning to denoise and remove outliers from dense point clouds," *Computer Graphics Forum*, vol. 39, no. 1, pp. 185–203, 2020.
- [45] W. Okamura, R. Lukita, G. Ching, Y. Matsuyama, Y. Kishiki, Z. Chen, K. Saito, and J.-i. Takada, "Simplification method of 3D point cloud data for ray trace simulation in indoor environment," *IEICE Communications Express*, vol. 9, no. 6, pp. 182–187, 2020.
- [46] P. Koivumäki, G. Steinböck, and K. Haneda, "Impacts of point cloud modeling on the accuracy of ray-based multipath propagation simulations," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 8, pp. 4737–4747, 2021.
- [47] N. Vaara, P. Sangi, M. B. López, and J. Heikkilä, "A ray launching approach for computing exact paths with point clouds," pp. 800–804, 2024.
- [48] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," in *Advances in Neural Information Processing Systems* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds.), vol. 30, Curran Associates, Inc., 2017.

- [49] Y.-J. Wang, L.-X. Guo, Z.-Y. Liu, and X.-J. Li, "Ray-tracing-based indoor channel prediction using point cloud geometrical modeling," in *2021 13th International Symposium on Antennas, Propagation and EM Theory (ISAPE)*, pp. 1–3, IEEE, 2021.
- [50] N. Suga, Y. Maeda, and K. Sato, "Indoor radio map construction via ray tracing with RGB-D sensor-based 3D reconstruction: concept and experiments in wlan systems," *IEEE Access*, vol. 11, pp. 24863–24874, 2023.
- [51] Y. Cui, Q. Li, and Z. Dong, "Structural 3D reconstruction of indoor space for 5G signal simulation with mobile laser scanning point clouds," *Remote Sensing*, vol. 11, no. 19, p. 2262, 2019.
- [52] M. Pang, H. Wang, K. Lin, and H. Lin, "A gpu-based radio wave propagation prediction with progressive processing on point cloud," *IEEE Antennas and Wireless Propagation Letters*, vol. 20, no. 6, pp. 1078–1082, 2021.
- [53] N. Suga, N. Yoshida, R. Gozono, Y. Maeda, and K. Sato, "RGB-D sensor-aided radio map estimation using materials classification," *IEEE Wireless Communications Letters*, vol. 14, no. 6, pp. 1811–1815, 2025.
- [54] Z. Qi, R. He, M. Yang, X. Zhang, B. Ai, W. Li, S. Wang, W. Sun, and L. Li, "Point cloud-based environment reconstruction and ray tracing simulations for railway tunnel channels," *High-speed Railway*, vol. 1, no. 4, pp. 241–247, 2023.
- [55] A. Klautau, N. González-Prelcic, and R. W. Heath, "Lidar data for deep learning-based mmwave beam-selection," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 909–912, 2019.
- [56] M. Zecchin, M. B. Mashhadi, M. Jankowski, D. Gündüz, M. Kountouris, and D. Gesbert, "LIDAR and position-aided mmwave beam selection with non-local CNNs and curriculum training," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 2979–2990, 2022.
- [57] H. Zhang, S. Gao, X. Cheng, and L. Yang, "Synesthesia of machines (SoM)-enhanced wideband Multi-User CSI learning with lidar sensing," *IEEE Transactions on Vehicular Technology*, pp. 1–6, 2025.
- [58] Remcom Inc., "Wireless InSite Propagation Software." <https://www.remcom.com/wireless-insite-em-propagation-software>. Accessed: 2026.
- [59] G. Reus-Muns, B. Salehi, D. Roy, T. Jian, Z. Wang, J. Dy, S. Ioannidis, and K. Chowdhury, "Deep Learning on Visual and Location Data for V2I mmWave Beamforming," in *IEEE International Conference on Mobility, Sensing and Networking*, Dec 2021.
- [60] X. Cheng, Z. Huang, L. Bai, H. Zhang, M. Sun, B. Liu, S. Li, J. Zhang, and M. Lee, "M³sc: A generic dataset for mixed multi-modal (mmm) sensing and communication integration," *China Communications*, vol. 20, no. 11, pp. 13–29, 2023.
- [61] X. Cheng, Z. Huang, Y. Yu, L. Bai, M. Sun, Z. Han, R. Zhang, and S. Li, "Synthsom: A synthetic intelligent multi-modal sensing-communication dataset for synesthesia of machines (som)," *Scientific Data*, vol. 12, no. 1, p. 819, 2025.
- [62] Ranplan Wireless, "Ranplan professional." [Online]. Available: <https://ranplanwireless.com/>, 2026.
- [63] J. Hoydis, F. A. Aoudia, S. Cammerer, M. Nimier-David, N. Binder, G. Marcus, and A. Keller, "Sionna RT: Differentiable Ray Tracing for Radio Propagation Modeling," in *2023 IEEE Globecom Workshops (GC Wkshps)*, pp. 317–322, 2023.
- [64] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles," in *Field and Service Robotics: Results of the 11th International Conference*, pp. 621–635, Springer, 2018.
- [65] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An Open Urban Driving Simulator," in *Proceedings of the 1st Annual Conference on Robot Learning (CoRL)*, pp. 1–16, 2017.
- [66] Remcom Inc., "WaveFarer Radar Sensing Simulation Software." <https://www.remcom.com/wavefarer-radar-sensing-simulation-software>. Accessed: 2026.
- [67] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wiessner, "Microscopic traffic simulation using sumo," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2575–2582, 2018.
- [68] B. Salehi, J. Gu, D. Roy, and K. Chowdhury, "FLASH: Federated Learning for Automated Selection of High-band mmWave Sectors," in *IEEE INFOCOM*, 2022.
- [69] J. Gu, B. Salehi, D. Roy, and K. R. Chowdhury, "Multimodality in mmwave mimo beam selection using deep learning: Datasets and challenges," *IEEE Communications Magazine*, vol. 60, no. 11, pp. 36–41, 2022.
- [70] A. Alkhateeb, G. Charan, T. Osman, A. Hredzak, J. Morais, U. Demirhan, and N. Srinivas, "Deepsense 6g: A large-scale real-world multi-modal sensing and communication dataset," *IEEE Communications Magazine*, vol. 61, no. 9, pp. 122–128, 2023.
- [71] L. Milosheski, K. Akiyama, B. Bertalanici, J. Hribar, and R. Shinkuma, "A multimodal dataset for indoor radio mapping with 3d point clouds and rss," *arXiv preprint arXiv:2511.00494*, 2025.
- [72] M. Alrabeiah, A. Hredzak, Z. Liu, and A. Alkhateeb, "Viwi: A deep learning dataset framework for vision-aided wireless communications," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pp. 1–5, 2020.
- [73] J. Zhang, A. A. Glazunov, W. Yang, and J. Zhang, "Fundamental wireless performance of a building," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 186–193, 2021.
- [74] S. Ju, Y. Xing, O. Kanhere, and T. S. Rappaport, "Millimeter wave and sub-terahertz spatial statistical channel model for an indoor office building," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1561–1575, 2021.
- [75] H. Obeidat, A. Alabdullah, E. Elkhazmi, W. Suhaib, O. Obeidat, M. Alkhambashi, M. Mosleh, N. Ali, Y. Dama, Z. Abidin, *et al.*, "Indoor environment propagation review," *Computer Science Review*, vol. 37, p. 100272, 2020.
- [76] J. W. McKown and R. L. Hamilton, "Ray tracing as a design tool for radio networks," *IEEE Network*, vol. 5, no. 6, pp. 27–30, 1991.
- [77] G. Yu, L. Zhou, J. Zhang, and J. Zhang, "A ray-launching algorithm for polarized wireless channel prediction," in *2023 IEEE Globecom Workshops (GC Wkshps)*, pp. 1928–1933, 2023.
- [78] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, "On mm-wave multipath clustering and channel modeling," *IEEE transactions on antennas and propagation*, vol. 62, no. 3, pp. 1445–1455, 2013.
- [79] N. Vaara, P. Sangi, M. Boddalo López, and J. Heikkilä, "Ray launching-based computation of exact paths with noisy dense point clouds," *IEEE Transactions on Antennas and Propagation*, vol. 73, no. 5, pp. 3270–3283, 2025.
- [80] L. Zhang, C.-X. Wang, Z. Zhou, Y. Li, J. Huang, L. Xin, C. Pan, D. Zheng, and X. Wu, "Wireless channel measurements and characterization in industrial iot scenarios," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 2, pp. 2292–2307, 2025.
- [81] H. Niu, D. Dupleich, Y. Völker-Schöneberg, A. Ebert, R. Müller, J. Eichinger, A. Artemenko, G. Del Galdo, and R. S. Thomä, "From 3D point cloud data to ray-tracing multi-band simulations in industrial scenario," in *2022 IEEE 95th Vehicular Technology Conference (VTC2022-Spring)*, pp. 1–5, IEEE, 2022.
- [82] M. F. De Guzman and K. Haneda, "Analysis of wave-interacting objects in indoor and outdoor environments at 142 GHz," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 12, pp. 9838–9848, 2023.
- [83] W. Okamura, K. Sugiyama, G. S. Ching, Y. Kishiki, K. Saito, and J.-i. Takada, "Indoor model reconstruction using 3D point cloud data for ray tracing simulation," in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–5, IEEE, 2022.
- [84] H. Mi, B. Ai, R. He, R. Caromi, J. Wang, A. Bodi, C. Gentile, and Y. Miao, "Cluster association for 3D environment based on 60 GHz indoor channel measurements," in *2023 17th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, IEEE, 2023.
- [85] P. Koivumäki and K. Haneda, "Point cloud ray-launching simulations of indoor multipath channels at 60 GHz," in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 01–07, IEEE, 2022.
- [86] F. Quattresooz, S. Demey, and C. Oestges, "Tracking of interaction points for improved dynamic ray tracing," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 7, pp. 6291–6301, 2021.
- [87] K. Saito, N. Keerativoranan, and J.-i. Takada, "Dynamic propagation simulation method from lidar point cloud data for smart office scenario," in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–6, IEEE, 2022.
- [88] M. F. De Guzman, P. Koivumäki, and K. Haneda, "Double-directional multipath data at 140 GHz derived from measurement-based ray-launcher," in *2022 IEEE 95th Vehicular Technology Conference (VTC2022-Spring)*, pp. 1–6, IEEE, 2022.
- [89] K. Saito, N. Keerativoranan, and J.-i. Takada, "Scenario-specific radio propagation simulation from LiDAR point cloud data for smart office scenario," *IEICE Communications Express*, vol. 12, no. 7, pp. 379–384, 2023.

- [90] K. Kitao, S. Suyama, T. Tomie, N. Kuno, K. Tateishi, and H. Jiang, "6g simulator using real environment model based on point cloud data," in *2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)*, pp. 1–5, IEEE, 2024.
- [91] P. Koivumäki, K. Haneda, and A. F. Molisch, "On the use of adaptive-density point cloud for site-specific ray-optics simulations," in *2024 18th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, IEEE, 2024.
- [92] G. Xia, C. Zhou, F. Zhang, Z. Cui, C. Liu, H. Ji, X. Zhang, Z. Zhao, and Y. Xiao, "Path loss prediction in urban environments with sionnar based on accurate propagation scene models at 2.8 ghz," *IEEE Transactions on Antennas and Propagation*, vol. 72, no. 10, pp. 7986–7997, 2024.
- [93] F. Zhang, C. Zhou, C. Brennan, R. Wang, Y. Li, G. Xia, Z. Zhao, and Y. Xiao, "A radio wave propagation modeling method based on high-precision 3-D mapping in urban scenarios," *IEEE Transactions on Antennas and Propagation*, vol. 72, no. 3, pp. 2712–2722, 2024.
- [94] R. He, B. Ai, G. Wang, K. Guan, Z. Zhong, A. F. Molisch, C. Briso-Rodriguez, and C. P. Oestges, "High-speed railway communications: From gsm-r to lte-r," *IEEE Vehicular Technology Magazine*, vol. 11, no. 3, pp. 49–58, 2016.
- [95] P. Koivumäki, A. Karttunen, and K. Haneda, "Ray-optics simulations of outdoor-to-indoor multipath channels at 4 and 14 GHz," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 7, pp. 6046–6059, 2023.
- [96] Z. Huang, L. Bai, M. Sun, and X. Cheng, "A lidar-aided channel model for vehicular intelligent sensing-communication integration," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 12, pp. 20105–20119, 2024.
- [97] P. Koivumäki, A. F. Molisch, and K. Haneda, "Line-of-sight probability in cluttered urban microcells: Analyses using poisson point process and point cloud," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 3, pp. 2161–2173, 2021.
- [98] V. Kubelka, P. Dandurand, P. Babin, P. Giguère, and F. Pomerleau, "Radio propagation models for differential GNSS based on dense point clouds," *Journal of Field Robotics*, vol. 37, no. 8, pp. 1347–1362, 2020.
- [99] R. Kou, R. Tan, S. Wang, B. Yang, Z. Dong, S. Yang, and F. Liang, "Satellite visibility analysis considering signal attenuation by trees using airborne laser scanning point cloud," *GPS Solutions*, vol. 27, no. 2, p. 64, 2023.
- [100] Y. Li, Y. Wang, and C. Huang, "Nera: Neural reflectance and attenuation fields for radio map reconstruction," in *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, pp. 1–5, IEEE, 2024.
- [101] Y. Wang, Y. Xue, S. Zhang, and T.-H. Chang, "Radsplatter: Extending 3D gaussian splatting to radio frequencies for wireless radiomap extrapolation," *arXiv preprint arXiv:2502.12686*, 2025.
- [102] Y. Egi and C. E. Otero, "Machine-learning and 3D point-cloud based signal power path loss model for the deployment of wireless communication systems," *IEEE Access*, vol. 7, pp. 42507–42517, 2019.
- [103] A. Gupta, J. Du, D. Chizhik, R. A. Valenzuela, and M. Sellathurai, "Machine learning-based urban canyon path loss prediction using 28 GHz manhattan measurements," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 6, pp. 4096–4111, 2022.
- [104] Z. Wei, B. Mao, H. Guo, Y. Xun, J. Liu, and N. Kato, "An intelligent path loss prediction approach based on integrated sensing and communications for future vehicular networks," *IEEE Open Journal of the Computer Society*, vol. 5, pp. 170–180, 2024.
- [105] K. Wang, L. Yu, J. Zhang, Y. Tian, E. Guo, and G. Liu, "Multi-modal environmental sensing based path loss prediction for V2I communications," *arXiv preprint arXiv:2412.07681*, 2024.
- [106] B. Peng, N. Ma, K. Zhang, K. Peng, and J. Chen, "Multimodal sensory data fusion-aided channel modelling: An end-to-end approach," in *2024 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1222–1226, IEEE, 2024.
- [107] L. Clark, J. A. Edlund, M. S. Net, T. S. Vaquero, and A. Aghamohammadi, "Propem-l: Radio propagation environment modeling and learning for communication-aware multi-robot exploration," *Robotics: Science and Systems (RSS)*, June 2022.
- [108] A. F. Molisch, *Wireless Communications: From Fundamentals to Beyond 5G*. Wiley-IEEE Press, 3rd ed., 2022.
- [109] Y. Wang, Y. Li, G. Chen, and C. Huang, "Channel knowledge maps construction based on point cloud environment information," in *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, pp. 1–5, IEEE, 2024.
- [110] P. Qi, Y. Zhang, Z. Yuan, L. Yu, P. Tang, and J. Zhang, "Channel modeling based on 3d scenario information for V2I communications," in *2021 15th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, IEEE, 2021.
- [111] Z. Huang, L. Bai, Z. Han, and X. Cheng, "Scatterer recognition for multi-modal intelligent vehicular channel modeling via synesthesia of machines," *IEEE Wireless Communications Letters*, pp. 1–1, 2025.
- [112] H. Mi, B. Ai, R. He, A. Bodi, R. Caromi, J. Wang, J. Senic, C. Gentile, and Y. Miao, "Measurement-based prediction of mmwave channel parameters using deep learning and point cloud," *IEEE Open Journal of Vehicular Technology*, vol. 5, pp. 1059–1072, 2024.
- [113] A. N. Uwaechia and N. M. Mahyuddin, "A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges," *IEEE Access*, vol. 8, pp. 62367–62414, 2020.
- [114] Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, "5g network capacity: Key elements and technologies," *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 71–78, 2014.
- [115] G. Charan, M. Alrabeiah, and A. Alkhateeb, "Vision-aided dynamic blockage prediction for 6G wireless communication networks," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, IEEE, 2021.
- [116] T. Woodford, X. Zhang, E. Chai, K. Sundaresan, and A. Khojastepour, "Spacebeam: Lidar-driven one-shot mmwave beam management," in *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 389–401, 2021.
- [117] R. Sharma, V. Badarla, and V. Sharma, "Pcoc: A fast sensor-device line of sight detection algorithm for point cloud representations of indoor environments," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1258–1261, 2020.
- [118] C. Kang, X. Du, and J.-I. Takada, "Point cloud-based prediction models of dynamic human body shadowing at 58 GHz," *IEEE Open Journal of Antennas and Propagation*, vol. 5, no. 4, pp. 1013–1025, 2024.
- [119] D. Marasinghe, N. Rajatheva, and M. Latva-aho, "Lidar aided human blockage prediction for 6G," in *2021 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, IEEE, 2021.
- [120] T. Zhang, J. Liu, and F. Gao, "Vision aided beam tracking and frequency handoff for mmwave communications," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–2, IEEE, 2022.
- [121] A. Kamari, N. Wu, H. K. Narsani, B. Han, and P. Pathak, "Environment-driven mmwave beamforming for multi-user immersive applications," in *Proceedings of the 1st ACM Workshop on Mobile Immersive Computing, Networking, and Systems*, pp. 268–275, 2023.
- [122] T. Nishio, H. Okamoto, K. Nakashima, Y. Koda, K. Yamamoto, M. Morikura, Y. Asai, and R. Miyatake, "Proactive received power prediction using machine learning and depth images for mmwave networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, pp. 2413–2427, Aug. 2019.
- [123] C. Zhao, Y. Feng, F. Gao, Y. Zhang, S. Ma, and H. V. Poor, "Environment sensing-aided beam prediction with transfer learning for smart factory," pp. 2077–2082, 2024.
- [124] M. Nerini and B. Clerckx, "Overhead-free blockage detection and precoding through physics-based graph neural networks: LIDAR data meets ray tracing," *IEEE Wireless Communications Letters*, vol. 12, no. 3, pp. 565–569, 2023.
- [125] M. Dias, A. Klautau, N. González-Prelcic, and R. W. Heath, "Position and lidar-aided mmwave beam selection using deep learning," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, IEEE, 2019.
- [126] T. Okada, T. Ueta, K. Ono, and M. Uesugi, "Wide area route prediction scheme by integrating multiple radars," in *2022 Asia-Pacific Microwave Conference (APMC)*, pp. 910–912, IEEE, 2022.
- [127] H. Asano, H. Noguchi, N. Shimizu, T. Asanuma, M. Yasugi, T. Ueta, K. Ohno, and R. Honda, "High power efficiency millimeter-wave network with communication quality prediction technology," in *2021 IEEE VTS 17th Asia Pacific Wireless Communications Symposium (APWCS)*, pp. 1–5, IEEE, 2021.
- [128] S. Pradhan, D. Roy, B. Salehi, and K. Chowdhury, "Copilot: Cooperative perception using lidar for handoffs between road side units," in *IEEE INFOCOM 2024-IEEE Conference on Computer Communications*, pp. 1301–1310, IEEE, 2024.
- [129] A. Kamari, Y. Chae, and P. Pathak, "mmsv: mmwave vehicular networking using street view imagery in urban environments," in *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*, pp. 1–16, 2023.

- [130] S. Wu, C. Chakrabarti, and A. Alkhateeb, "Lidar-aided mobile blockage prediction in real-world millimeter wave systems," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2631–2636, IEEE, 2022.
- [131] S. Wu, C. Chakrabarti, and A. Alkhateeb, "Proactively predicting dynamic 6G link blockages using lidar and in-band signatures," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 392–412, 2023.
- [132] 3GPP, "3rd generation partnership project; technical specification group radio access network; NR; physical layer procedures for data (release 18)," 2024.
- [133] 3GPP, "3rd generation partnership project; technical specification group radio access network; NR; medium access control (MAC) protocol specification (release 18)," 2024.
- [134] Y.-N. R. Li, B. Gao, X. Zhang, and K. Huang, "Beam management in millimeter-wave communications for 5G and beyond," *IEEE Access*, vol. 8, pp. 13282–13293, 2020.
- [135] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmwave frequencies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 173–196, 2018.
- [136] 3GPP, "3rd generation partnership project; technical specification group radio access network; study on new radio access technology physical layer aspects (release 14)," Tech. Rep. TR 38.802 V14.2.0 (2017-09), 3GPP, 2017.
- [137] M. B. Mollah, H. Wang, M. A. Karim, and H. Fang, "Multi-modality sensing in mmwave beamforming for connected vehicles using deep learning," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2025.
- [138] W. Xu, F. Gao, S. Jin, and A. Alkhateeb, "3d scene-based beam selection for mmwave communications," *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 1850–1854, 2020.
- [139] K. Vuckovic, S. M. Hosseini, and N. Rahnavard, "Revisiting performance metrics for multimodal mmwave beam prediction using deep learning," in *MILCOM 2024-2024 IEEE Military Communications Conference (MILCOM)*, pp. 881–887, IEEE, 2024.
- [140] Y. Cui, J. Nie, X. Cao, T. Yu, J. Zou, J. Mu, and X. Jing, "Sensing-assisted high reliable communication: A transformer-based beamforming approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 18, no. 5, pp. 782–795, 2024.
- [141] E. Zhu, H. Sun, and M. Ji, "Hawkrover: An autonomous mmwave vehicular communication testbed with multi-sensor fusion and deep learning," in *International Symposium on Intelligent Computing and Networking*, pp. 61–72, Springer, 2024.
- [142] B. Salehi, D. Roy, T. Jian, C. Dick, S. Ioannidis, and K. Chowdhury, "Omni-CNN: A modality-agnostic neural network for mmwave beam selection," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 6, pp. 8169–8183, 2024.
- [143] X. Niu, Y. Fu, M. Dong, N. Cheng, and C. Li, "Multi-source low redundancy data-aided beam prediction for V2I communication," in *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)*, pp. 1–5, IEEE, 2023.
- [144] Y. M. Park, Y. K. Tun, W. Saad, and C. S. Hong, "Resource-efficient beam prediction in mmwave communications with multimodal realistic simulation framework," *arXiv preprint arXiv:2504.05187*, 2025.
- [145] B. Salehi, G. Reus-Muns, D. Roy, Z. Wang, T. Jian, J. Dy, S. Ioannidis, and K. Chowdhury, "Deep learning on multimodal sensor data at the wireless edge for vehicular network," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 7, pp. 7639–7655, 2022.
- [146] M. B. Mashhadi, M. Jankowski, T.-Y. Tung, S. Kobus, and D. Gündüz, "Federated mmwave beam selection utilizing lidar data," *IEEE Wireless Communications Letters*, vol. 10, no. 10, pp. 2269–2273, 2021.
- [147] B. Salehi, J. Gu, D. Roy, and K. Chowdhury, "Flash: Federated learning for automated selection of high-band mmwave sectors," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*, pp. 1719–1728, IEEE, 2022.
- [148] B. Salehi, D. Roy, J. Gu, C. Dick, and K. Chowdhury, "Flash-and-prune: Federated learning for automated selection of high-band mmwave sectors using model pruning," *IEEE Transactions on Mobile Computing*, vol. 23, no. 12, pp. 11655–11669, 2024.
- [149] M. Tian, Z. Zhang, Q. Xu, and L. Yang, "A privacy-preserved split learning solution for deep learning-based mmwave beam selection," *IEEE Communications Letters*, vol. 26, no. 7, pp. 1474–1478, 2022.
- [150] J. Chen, F. Gao, X. Tao, G. Liu, C. Pan, and A. Alkhateeb, "Computer vision aided codebook design for MIMO communications systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3341–3354, 2022.
- [151] D. Marasinghe, N. Jayaweera, N. Rajatheva, S. Hakola, T. Koskela, O. Tervo, J. Karjalainen, E. Tirola, and J. Hultkonen, "Lidar aided wireless networks-beam prediction for 5g," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, pp. 1–7, IEEE, 2022.
- [152] S. Jiang, G. Charan, and A. Alkhateeb, "LiDAR aided future beam prediction in real-world millimeter wave V2I communications," *IEEE Wireless Communications Letters*, vol. 12, no. 2, pp. 212–216, 2022.
- [153] O. Rinchi, A. Alsharoa, and I. Shatnawi, "Deep-learning-based accurate beamforming prediction using lidar-assisted network," in *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–5, IEEE, 2023.
- [154] Y. Bian, J. Yang, S. Xia, and S. Jin, "3-D LiDAR and GPS aided beam tracking in millimeter wave vehicular communications," *IEEE Wireless Communications Letters*, vol. 13, no. 12, pp. 3290–3294, 2024.
- [155] C. Zheng, J. He, C. G. Kang, G. Cai, Z. Yu, and M. Debbah, "M2beamllm: Multimodal sensing-empowered mmwave beam prediction with large language models," *arXiv preprint arXiv:2506.14532*, 2025.
- [156] K. Patel and R. W. Heath, "Harnessing multimodal sensing for multi-user beamforming in mmwave systems," *IEEE Transactions on Wireless Communications*, vol. 23, no. 12, pp. 18725–18739, 2024.
- [157] Y. Koda, K. Nakashima, K. Yamamoto, T. Nishio, and M. Morikura, "Cooperative sensing in deep rl-based image-to-decision proactive handover for mmwave networks," in *2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1–6, IEEE, 2020.
- [158] S. Ohta, T. Nishio, R. Kudo, K. Takahashi, and H. Nagata, "Point cloud-based proactive link quality prediction for millimeter-wave communications," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 1, pp. 258–276, Sep. 2023.
- [159] J. Zhang and G. de la Roche, *Femtocells: Technologies and Deployment*. Wiley Publishing, 2010.
- [160] A. Zhang, C. Wang, Y. Hu, A. Hassan, Z. Zhang, B. Han, F. Qian, and S. Xu, "Habitat: Boosting mobile immersive content delivery through full-body pose tracking and multipath networking," in *21st USENIX Symposium on Networked Systems Design and Implementation (NSDI 24)*, pp. 1677–1695, 2024.
- [161] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE transactions on signal processing*, vol. 69, pp. 2663–2675, 2021.
- [162] Z. Zhang, R. He, M. Yang, X. Zhang, Z. Qi, Y. Yuan, and B. Ai, "Characterization of wireless channel semantics: A new paradigm," pp. 1–5, 2024.
- [163] L. Xie, Y. Zhang, C. Zhao, C. Zhang, M. Zhou, and X. Yang, "Positioning under multipath environments in wireless network: Survey, design, and opportunities," *IEEE Network*, vol. 38, no. 6, pp. 476–484, 2024.
- [164] O. Rinchi, A. Elzanaty, and A. Alsharoa, "Enhancing near-field wireless localization with LiDAR-Assisted RIS in multipath environments," *IEEE Wireless Communications Letters*, vol. 12, no. 12, pp. 2168–2172, 2023.
- [165] J. Gao, F. Yin, W. Yan, Q. Kong, L. Xu, and S. Cui, "Genmetaloc: Learning to learn environment-aware fingerprint generation for sample efficient wireless localization," *arXiv preprint arXiv:2503.18078*, 2025.
- [166] E. Rastorgueva-Foi, O. Kaltiokallio, Y. Ge, M. Turunen, J. Talvitie, B. Tan, M. Furkan Keskin, H. Wymeersch, and M. Valkama, "Millimeter-wave radio slam: End-to-end processing methods and experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 9, pp. 2550–2567, 2024.
- [167] T. Du, J. Yang, F. Liu, J. Guo, S. Xia, C.-K. Wen, and S. Jin, "Simultaneous localization and mapping using active mmwave sensing in 5g nr," in *ICC 2025 - IEEE International Conference on Communications*, pp. 6716–6722, 2025.
- [168] J. Zhang and D. M. Blough, "Optimizing coverage with intelligent surfaces for indoor mmwave networks," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*, pp. 830–839, IEEE, 2022.
- [169] R. Ma, S. Zheng, H. Pan, L. Qiu, X. Chen, L. Liu, Y. Liu, W. Hu, and J. Ren, "Automs: Automated service for mmwave coverage optimization using low-cost metasurfaces," in *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, pp. 62–76, 2024.
- [170] Y. Miao, Y. Zhang, J. Zhang, Y. Sun, Y. Tian, L. Yu, and G. Liu, "Demo abstract: Predictive radio environment for digital twin communication platform via enhanced sensing," in *IEEE INFOCOM 2023-IEEE*

- Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–2, IEEE, 2023.
- [171] G. Nie, J. Zhang, Y. Zhang, L. Yu, Z. Zhang, Y. Sun, L. Tian, Q. Wang, and L. Xia, “A predictive 6g network with environment sensing enhancement: From radio wave propagation perspective,” *China Communications*, vol. 19, no. 6, pp. 105–122, 2022.
 - [172] H. Wang, J. Zhang, G. Nie, L. Yu, Z. Yuan, T. Li, J. Wang, and G. Liu, “Digital twin channel for 6G: Concepts, architectures and potential applications,” *IEEE Communications Magazine*, vol. 63, no. 3, pp. 24–30, 2025.
 - [173] C. Linnhoff, K. Hofrichter, L. Elster, P. Rosenberger, and H. Winner, “Measuring the influence of environmental conditions on automotive lidar sensors,” *Sensors*, vol. 22, no. 14, p. 5266, 2022.
 - [174] X. Xia, N. P. Bhatt, A. Khajepour, and E. Hashemi, “Integrated inertial-lidar-based map matching localization for varying environments,” *IEEE transactions on intelligent vehicles*, 2023.
 - [175] B. Lin, C. Zhao, F. Gao, G. Ye Li, J. Qian, and H. Wang, “Environment reconstruction based on multi-user selection and multi-modal fusion in ISAC,” *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 15083–15095, 2024.
 - [176] G. Charan and A. Alkhateeb, “User identification: A key enabler for multi-user vision-aided communications,” *IEEE Open Journal of the Communications Society*, vol. 5, pp. 472–488, 2024.
 - [177] F. Pittaluga, S. J. Koppal, S. B. Kang, and S. N. Sinha, “Revealing scenes by inverting structure from motion reconstructions,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 145–154, IEEE, 2019.
 - [178] K. Chelani, F. Kahl, and T. Sattler, “How privacy-preserving are line clouds? recovering scene details from 3d lines,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15668–15678, June 2021.
 - [179] Z. Hau, K. T. Co, S. Demetriou, and E. C. Lupu, “Object removal attacks on LiDAR-based 3d object detectors,” in *Third International Workshop on Automotive and Autonomous Vehicle Security (AutoSec)*, pp. 17–23, 2021.
 - [180] T. Sato, Y. Hayakawa, R. Suzuki, Y. Shiiki, K. Yoshioka, and Q. A. Chen, “LiDAR spoofing meets the new-gen: Capability improvements, broken assumptions, and new attack strategies,” in *Network and Distributed System Security Symposium (NDSS)*, 2024.
 - [181] M. Geppert, V. Larsson, J. L. Schönberger, and M. Pollefeys, “Privacy preserving partial localization,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 17316–17326, 2022.
 - [182] M. Cho, Y. Cao, Z. Zhou, and Z. M. Mao, “ADoPT: LiDAR spoofing attack detection based on point-level temporal consistency,” in *34th British Machine Vision Conference 2023, BMVC 2023, Aberdeen, UK, November 20-24, 2023*, BMVA, 2023.