

WiNeRF: Measurement Constrained Radiance Fields for Actionable Wireless Channel Modeling

Saif Ur Rahman¹, Rafid Umayer Murshed¹, Anton Dmitriev¹, Cagri Tanriover²,

Rahul C. Shah², Elahé Soltanaghai¹

¹University of Illinois Urbana-Champaign, ²Intel

{saifu2, rum3, antond2, elahé}@illinois.edu, {cagri.tanriover, rahul.c.shah}@intel.com

Abstract

Wireless embedded systems increasingly rely on wireless channel information for decision making, yet practical platforms operate under severe constraints, including few antennas, narrow bandwidth, and sparse, noisy measurements. While neural field based approaches inspired by Neural Radiance Field (NeRF) have recently been explored for continuous wireless channel modeling, existing approaches depend on dense measurements or external priors such as known geometry, visual context, or angle-of-arrival (AoA) information, limiting their practicality in real-world deployments. We present WiNeRF, a neural field framework that learns a spatially continuous, complex-valued wireless channel representation directly from sparse channel state information (CSI) collected by commodity WiFi devices. WiNeRF embeds intrinsic system constraints, such as antenna geometry, limited spatial resolution, and phase uncertainty, as inductive biases through a 3D conical wave sampling model, a multi-resolution implicit scene representation, and a differentiable optimization framework for complex-valued channel learning. Across diverse indoor environments with non-line-of-sight (NLoS) regions, WiNeRF achieves a median prediction SNR of 5.3 dB, outperforming prior neural baselines by 4.9 dB on average (approximately 3 $\times$ higher prediction SNR), and produces a task-agnostic channel representation that can be directly reused in standard signal-processing pipelines, including beamforming, AoA estimation, and RSSI coverage mapping, without modifying existing hardware or wireless protocols.

CCS Concepts

• **Networks** $\rightarrow$ **Wireless access networks**; *Network performance modeling*; • **Computing methodologies** $\rightarrow$ *Machine learning*.

Keywords

Wireless Channel Modeling, Channel State Information (CSI), Neural Fields, Neural Radiance Fields, Commodity WiFi, Beamforming, Angle-of-Arrival Estimation

1 Introduction

Modern embedded wireless systems increasingly seek to exploit wireless channel information to support environment-aware decision making under tight sensing and communication constraints. In these systems, communication signals can be leveraged to sense and map the surrounding environment, enabling proactive decision-making that accounts for wireless conditions alongside physical factors [7, 19, 38]. For example, a mobile robot can map and avoid communication dead zones to plan connectivity-aware paths, while an embedded node can predict coverage gaps to maintain reliable links under mobility. Achieving this intelligence requires modeling

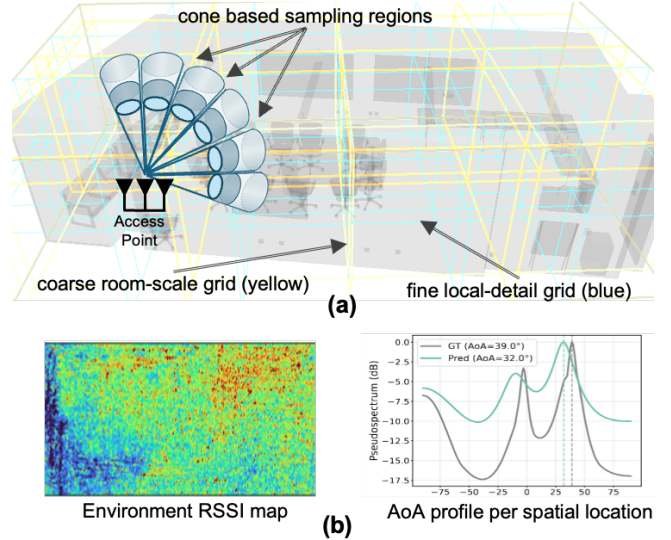


Figure 1: (a) WiNeRF combines cone-based sampling regions with a multi-resolution implicit scene representation of coarse room-scale grids (yellow) and fine local-detail grids (blue) to learn a continuous wireless field from sparse CSI measurements. (b) The learned field supports downstream tasks such as RSSI coverage mapping and AoA profiling.

the channel under severe measurement constraints, a task central to building high-fidelity Digital Twins and site-planning tools where scene-specific accuracy is the primary objective. In this context, CSI provides a rich representation of how wireless signals propagate and interact with surrounding objects [7]. Specifically, the phase variations of CSI across frequency subcarriers and antenna elements capture intrinsic multipath characteristics that define the environment's spatial structure [18]. As such, accurately capturing these multidimensional relationships is essential for understanding and predicting the static spatial behavior of the wireless channel in resource-constrained Joint Sensing and Communication (JSAC) systems.

Prior works on wireless channel modeling have explored NeRF-based approaches to represent implicit volumetric representations of wireless channels, capturing spatial signal spectra or coverage patterns [8, 16, 21, 40]. However, all of these methods make strong assumptions such as dense training data [40] or rely on priors such as known 3D geometry [29], visual context [8], or precise AoA information [21] to guide learning and stabilize training. This reliance makes them impractical for real-world deployments where such external data is often unavailable. A key challenge is that wireless channel

modeling is an inherently difficult reconstruction task, with more unknown variables (e.g. multipath characteristics) than available observations due to the small number of antennas and aggregate representation of multipath in the form of CSI matrices. Due to these reasons, prior works focus on task-specific objectives, rather than learning general-purpose channel representations that can be reused across downstream signal-processing tasks.

To address these limitations, we ask a different question: *Can we design a neural wireless channel model that uses measurement measurement constraints as natural guides to produce actionable and consistent channel representations?* These intrinsic measurement characteristics, such as spatial resolution and antenna array geometry, fundamentally shape how the channel is observed at the wireless node. Rather than treating them as noise to be compensated for, we use them as inductive biases that regularize what the model is allowed to represent.

We present WiNeRF, a site-specific neural field framework for learning spatially continuous, complex-valued wireless channel representations subject to measurement constraints of commodity hardware. An overview of WiNeRF is presented in Figure. 1. Prior NeRF-based wireless channel models often idealize propagation using infinitesimally thin rays or fixed-resolution voxel grids, implicitly assuming arbitrarily fine angular resolution at the receiver. In contrast, WiNeRF replaces line-ray sampling with conical primitives whose angular extent reflects the finite AoA resolvability of practical antenna arrays, as illustrated in Figure. 2. Importantly, these cones do not model directional radiation; rather, they encode the set of propagation directions that are indistinguishable given the hardware’s angular observability, mitigating AoA aliasing while preserving multipath-related structure.

As a result, WiNeRF learns a site-specific channel field that can be queried at unseen receiver locations within the same environment, even under lower measurement density than prior approaches. These estimates are immediately actionable for offline tasks such as beamforming, AoA estimation, and spatial received signal strength (RSSI) coverage mapping. The framework consists of three interconnected measurement-driven modules: (i) *Cone Sampling Wave Model*, (ii) *Implicit Scene Representation*, and (iii) *CSI Reconstruction Module*.

Cone Sampling Wave Model : Learning a continuous neural representation directly from sparse CSI is difficult because many different signal paths can produce the same measurement at the receiver. Instead of relying on extra data like 3D maps or images, we guide the representation using the actual measurement characteristics of the antenna arrays themselves. Specifically, we represent each propagation direction using a volumetric cone that reflects the array’s natural angular resolution, and approximate each cone with a multivariate Gaussian to enable stable, differentiable CSI rendering. This replaces idealized line-ray sampling with a representation that explicitly accounts for the finite angular resolution of practical antenna arrays. The cone width encodes antenna-dependent angular ambiguities [11, 23], such as left–right aliasing in linear arrays or elevation ambiguity in horizontal arrays. By aligning the sampling process with the receiver’s angular observability, the learned channel representation captures only multipath structure that is distinguishable under the available measurements.

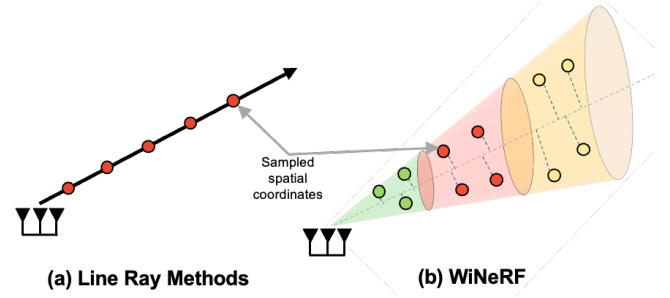


Figure 2: (a) Prior NeRF-based wireless models sample along infinitesimally thin line rays, implicitly assuming arbitrarily fine angular resolution. In contrast, (b) WiNeRF replaces line rays with cone-based support regions that reflect finite antenna-array observability, enabling hardware-aware spatial aggregation.

Implicit Scene Representation : Unlike camera images where each pixel corresponds to a localized point in space [33], a CSI measurement aggregates contributions from many multipath signals with different path lengths and angles; nearby objects can dominate certain measurements while distant reflectors may only appear at coarse spatial scales [1, 11]. As a result, a continuous field model must capture meaningful world structure without over-smoothing or overfitting. This need is amplified because our cone formulation represents distributions of feasible arrival directions rather than single paths. To address this, WiNeRF adopts a Multi-Scale Scene Representation using grids of various resolutions to represent the wireless environment. This compact and adaptive feature field is capable of encoding large room-scale boundaries (e.g., walls) and fine local reflectors (e.g., furniture) within the same framework. This multi-scale encoding preserves the spatial detail needed to separate complex signal paths while still generalizing from sparse CSI.

CSI Reconstruction Module : This module combines the Cone Sampling Wave Model and Implicit Scene Representation to synthesize the CSI at receiver (Rx) locations via differentiable rendering over cone-based primitives, producing complex-valued CSI across antennas and subcarriers. A key challenge is that CSI phase from commodity hardware is heavily corrupted by noise and impairments such as phase wrapping, CFO, and SFO [20, 34], making direct prediction of absolute phase unreliable. We therefore decouple hardware impairment compensation from learning by applying magnitude normalization, removing global CFO/SFO, and training with a relative phase loss that captures phase relationships across antennas and subcarriers. This preserves multipath cues such as AoA and relative time-of-flight while providing a stable and physically meaningful learning objective from sparse CSI.

We evaluate WiNeRF as a high-fidelity spatial planning tool using CSI collected at various positions using RT-AC86U commodity WiFi routers with four antennas and 50 OFDM subcarriers (20MHz bandwidth channel) in the 5 GHz band. Experiments span indoor environments with different room scales and furniture configurations, enabling evaluation under diverse multipath conditions, including both line-of-sight (LoS) and NLoS regions. We compare

Table 1: Evolution of Neural Field Methods for Wireless Applications

Method	Wave Representation	Priors Need	Experimental Data	Channel Output
NeRF ²	Voxelized Rays	Dense Training Data	✓	Task-dependent (e.g. RSSI, spatial spectrum)
WiNeRT	Ray segments	Path parameters (AoA, ToF)	✗	CIR
RF-Canvas	Tensorial Field	3D Geometry	✓	RSSI
NeWRF	Voxelized Ray	Multipath AoA	✗	Complex CSI
WiNeRF	Hash-based Cones	–	✓	Complex CSI

WiNeRF against existing neural channel modeling baselines, including NeRF² [40], NeWRF [21], VAE [17], and GAN [12]. WiNeRF estimates complex-valued CSI matrix at held-out receiver locations within the same environment with high fidelity, achieving a median SNR of 5.30 dB, **outperforming state-of-the-art baselines by 3.0–6.7 dB (131–368% improvement)**, while using only 4 samples per cubic feet of training data. Beyond channel estimation accuracy, we demonstrate the offline actionability of the estimated CSI and its preserved phase coherence through three representative applications: beamforming gain evaluation, spatial signal coverage mapping for wireless-aware path planning, and LoS AoA estimation using the classical MUSIC algorithm [18, 31]. Importantly, all evaluations are performed using standard signal processing methods directly on the predicted CSI, rather than training downstream estimators. This ensures that observed performance reflects true physical fidelity and learned multipath structure rather than task-specific overfitting.

As such, our contributions are as follows:

- We introduce WiNeRF, a neural field framework for wireless channel modeling that directly incorporates the measurement limits of commodity devices as inductive constraints, enabling structured learning of complex-valued CSI from sparse data.
- We propose a cone-based wave sampling formulation that models hardware-imposed angular resolution and ambiguity, limiting the learned representation to multipath components that commodity antenna arrays can actually resolve.
- We incorporate a multi-resolution implicit scene representation that jointly captures large-scale environmental structure and fine-grained scattering effects, allowing accurate channel prediction at unseen receiver locations within the same environment.
- We present a rendering-based training pipeline with relative phase learning objective that enables stable learning from noisy commodity CSI, and show that the learned channel representations can be directly used for downstream tasks such as beamforming and AoA estimation in real indoor environments.

2 Background and Related Work

2.1 Wireless Channel Estimation

Wireless channel modeling has been a central topic in wireless communication [11, 41], with a more recent trend on using data-driven models [37]. The main paradigm is to use machine learning to learn complex channel distributions and model the channel input-output relationships as a conditional probability distribution [36], generative adversarial networks (GANs) [39], or variational auto-encoders

(VAEs) [9]. However, these methods have seen limited practical adoption due to their reliance on large training datasets and their limited ability to generalize to scenarios not observed during training. Another category of wireless channel modeling approach focuses on wireless simulation and ray-tracing [15] with more recent techniques leveraging deep learning for optimizing the ray tracing [14]. However, they rely on accurate 3D model of the environment and large training data.

2.2 Neural Radiance Fields for Wireless

Recent advances in neural radiance fields have transformed visual computing by representing physical scene properties as continuous functions in space. The key strength of this approach is the ability to construct spatial models from sparse observations and predict signal characteristics at unseen locations within a scene. Viewed through the same lens as the original definition of physical fields in physics [10], a wireless propagation environment can also be parameterized as a field that maps spatial coordinates to electromagnetic quantities such as complex attenuation, phase, or frequency response. Motivated by this idea, recent work has proposed NeRF-inspired models to synthesize wireless channels. WiNeRT [29] and NeWRF [21], for example, aim to reconstruct complete Channel Impulse Response or Channel Frequency Response fields that capture multipath amplitude, phase, and delay. However, these models rely exclusively on synthetic data because they require per-path information such as Angle of Arrival and Time of Flight for training or evaluation, and such annotations are not practical to acquire from real-world measurements.

Another line of work focuses on synthesizing RF measurements for downstream tasks. Examples include NeRF² [40] and RF-Canvas [8], which primarily estimate RSSI or spatial spectrum rather than full complex propagation behavior and therefore do not demonstrate the phase accuracy required for coherent MIMO processing. Furthermore, these methods rely on unrealistic assumptions such as extremely dense measurements or a known coarse geometry of the scene. These assumptions reflect the fundamentally ill-posed nature of wireless channel reconstruction, where many different scattering configurations can produce similar sparse CSI. Instead, our method incorporates the intrinsic measurement limits of commercial WiFi hardware as constraints in the model, enabling physically observable structure to guide learning and allowing complex wireless channels to be estimated from sparse, noisy experimental data.

Architectural Limits of NeRF-based Wireless Models: These limitations point to two recurring architectural mismatches in prior NeRF-based wireless models: how propagation is sampled and how scene structure is represented across scales. The existing papers primarily apply the classical NeRF architecture directly to the wireless

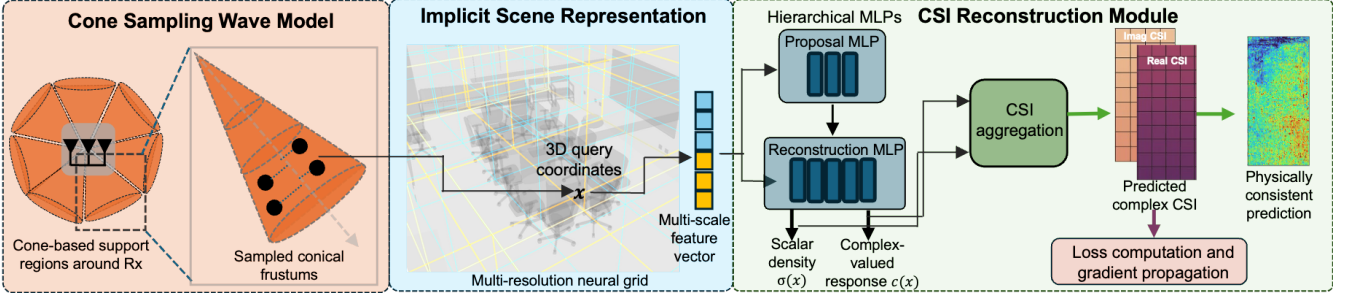


Figure 3: Given sparse spatial CSI measurements, WiNeRF reconstructs a continuous complex channel field through three modules: (1) Cone Sampling Wave Model, which defines cone-based support regions and sampled conical frustums around the receiver; (2) Implicit Scene Representation, which encodes the (x, y, z) samples using a multi-resolution hash grid; and (3) CSI Reconstruction Module, which predicts and aggregates complex CSI.

domain, suffering from several architectural limitations that our method addresses:

- **Inefficient Sampling:** NeRF assumes dense ray sampling along infinitesimally thin rays, which works for a pinhole camera where each pixel corresponds to a unique direction and megapixel images yield millions of angular samples [24]. In wireless, the receiver antenna array is the sensor, and its AoA resolution is fundamentally limited by the number of elements, aperture size, and SNR [30, 35]. Therefore, directly porting NeRF produces two inefficient extremes: dense ray casting with huge training sets as in NeRF² [40], which is wasteful because most rays do not correspond to physical paths, or sparse ray casting guided only by a few AoAs as in NeWRF [21], which is highly susceptible to aliasing and hardware-limited angular resolution, resulting in inaccurate field reconstruction when real-world data is used.
- **Scale Ambiguity:** Classical NeRF uses a single MLP to represent the entire scene, and most NeRF-based wireless methods inherit this same architecture. However, wireless propagation involves both large, dominant reflectors (such as walls) and fine, localized scattering features (such as furniture edges). These different scattering mechanisms operate at different spatial scales and impose very different structure on the signal. Prior works such as WiNeRT [29] and RF-Canvas [8] attempt to cope with this by injecting explicit geometry priors, but this is often impractical because it requires either perfect 3D scene models from simulation or additional sensing modalities that are not available in real deployments. Without such priors, a single MLP tends to oversmooth (spectral-bias) and fails to capture the sharp multipath variations.

Recent NeRF progress in computer vision offers architectural ideas that can help address these structural limitations for wireless: (i) Anti-aliasing: Mip-NeRF [3] replaces infinitely thin rays with cones and integrates features over their finite footprint to explicitly suppress aliasing artifacts. (ii) Multi-resolution encoding: Instant-NGP [27] introduces a multi-resolution hash-grid positional encoding so that coarse cells capture large structures while fine cells capture small details. More recently, Zip-NeRF [4] combines both concepts, using cone footprints to select the appropriate hash-grid scale at each query location. However, Zip-NeRF assumes a dense camera pixel grid that preserves angular information, which wireless systems do not provide because RF measurements are sparse, array-limited, and

do not directly yield pixel-level direction vectors. Inspired by Zip-NeRF, we therefore propose a new wave-and-world field formulation that addresses the inefficient sampling and scale ambiguity discussed above, but repurposes these ideas for wireless by explicitly embedding hardware-dependent observability into the representation.

3 System Design

WiNeRF introduces a neural field framework that replaces the idealized line-ray abstraction in classical NeRF formulations with a cone-based volumetric wave model. This design explicitly reflects the finite angular resolution of commodity antenna arrays, which fundamentally limits how precisely incoming propagation directions can be distinguished [35]. As a result, commodity WiFi devices observe propagation only up to their finite angular resolvability, collapsing multiple incoming directions into indistinguishable angular sectors rather than infinitesimally precise rays. In contrast, prior NeRF-based wireless models [21, 40] assume arbitrarily fine angular resolution by representing propagation with infinitesimally thin rays. This abstraction is mismatched with real antenna observability and leads to AoA aliasing: thin rays represent directions that are indistinguishable at the hardware level and therefore unlikely to consistently intersect resolvable multipath structure in the scene. This mismatch often results in physically implausible channel estimates and weaker spatial prediction at unseen receiver locations within the same environment.

An overview of the end-to-end workflow of WiNeRF is shown in Figure 3. Without loss of generality, we focus on a Uniform Linear Array (ULA), though the formulation naturally extends to other array geometries. The framework consists of three tightly coupled components: (1) a **Cone Sampling Wave Model**, (2) a **Implicit Scene Representation**, and (3) a **CSI Reconstruction Module**. We detail each component in the following subsections.

3.1 Cone Sampling Wave Model

The goal of WiNeRF’s wave model is to incorporate antenna-induced observability limits as inductive biases that constrain what the neural field can represent. Rather than attempting to compensate for hardware imperfections through post-processing or learning arbitrary corrections, the model is explicitly restricted to represent only those angular structures that are distinguishable given the finite resolution of the antenna array. By construction, multipath components that

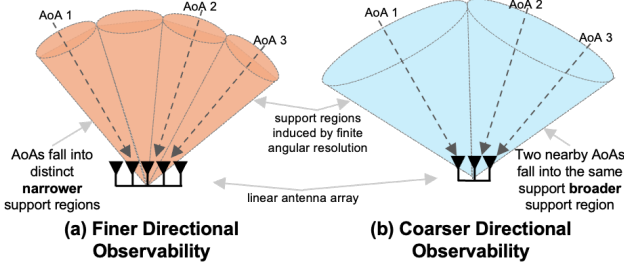


Figure 4: Array-dependent angular support in WiNeRF. The figure schematically compares two ULA configurations with different numbers of antennas: (a) higher-element ULA with finer directional observability and (b) lower-element ULA with coarser directional observability. WiNeRF adjusts cone-based support width accordingly. With finer observability, three example AoAs occupy distinct conical support regions; with coarser observability, the broader support regions are grouped into two nearby AoAs.

are closer in angle than the array’s resolvable limit are treated as indistinguishable and are not artificially separated during learning. This alignment between the sampling model and the measurement process encourages physically consistent channel representations that remain usable for downstream tasks such as beamforming and AoA estimation.

Cone Geometry: In WiNeRF, angular sampling is performed using volumetric cones rather than infinitesimally thin rays. Importantly, these cones are not intended to approximate the full beamforming process, directional radiation pattern, or exact array response function. Instead, they represent the set of arrival directions that are indistinguishable under the antenna array’s finite angular resolution. Each cone is parameterized by a central direction vector $\mathbf{u}$ and two angular widths: the azimuthal width α_{az} in the horizontal plane and the elevation width α_{el} in the vertical plane.

For a cone centered at the antenna array, the cross-sectional radii at distance d are given by

$$r_{az}(d) = d \tan\left(\frac{\alpha_{az}}{2}\right), \quad r_{el}(d) = d \tan\left(\frac{\alpha_{el}}{2}\right). \quad (1)$$

This formulation defines an anisotropic (elliptical) cone whose shape reflects direction-dependent angular resolution imposed by the array geometry. The linear expansion with distance preserves constant angular uncertainty, consistent with array resolution being determined by aperture and antenna element placement. As a result, the cone geometry serves as a measurement-constrained inductive bias that regularizes the inverse problem under AoA aliasing, rather than learned smoothing or a claim of full electromagnetic fidelity.

More generally, WiNeRF sets cone width according to the directional observability of the receiving array: arrays that can resolve directions more finely are assigned narrower angular support regions, while arrays with coarser directional observability are assigned broader ones. Figure 4 schematically illustrates this idea using three example AoAs. In the finer-observability case, the three AoAs occupy distinct support regions and therefore remain distinguishable at the level of the measurement model. In the coarser-observability case,

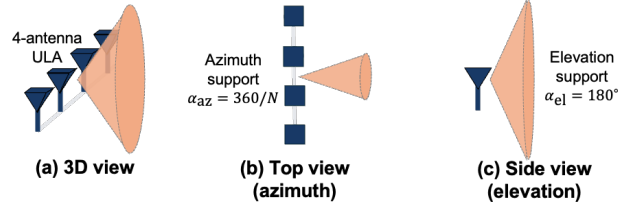


Figure 5: Angular support region for a 4-antenna ULA.

two nearby AoAs fall within the same broader support region, indicating that they are not separately resolved by the array under the assumed observability. The figure is intended as a conceptual illustration of support-region width, not as a depiction of beamforming patterns or exact array response.

As one concrete example, consider a ULA with four antenna elements spaced at half-wavelength intervals. Such an array provides directional discrimination primarily in azimuth, while exhibiting strong ambiguity in elevation due to the absence of vertically separated elements. To reflect this asymmetry, WiNeRF assigns the azimuthal cone width as $\alpha_{az} = \frac{360}{N}^\circ$, where N denotes the number of angular partitions around the receiver, and sets $\alpha_{el} = 180^\circ$ to capture near-complete elevation ambiguity. Figure 5 illustrates the resulting anisotropic support region used in WiNeRF for a 4-antenna ULA, shown in 3D, top, and side views. The figure depicts one representative support region; in practice, WiNeRF defines multiple such regions around the receiver by partitioning the azimuth plane into discrete sectors. In our implementation, this observability model is instantiated using 12 azimuthal sectors of 30° each, which are used consistently during both training and inference.

Each cone is discretized into K segments, or conical frustums, along the propagation path to enable volumetric integration. Let $\mathbf{o}$ denote the ray origin (antenna position), $\mathbf{u}$ the unit direction vector, and d_k the midpoint distance of the k -th segment from the origin. The center of each frustum is given by $\boldsymbol{\mu}_k = \mathbf{o} + d_k \mathbf{u}$. To represent the spatial extent of each frustum, WiNeRF employs the supersampling strategy from Zip-NeRF [5]. In this formulation, each sample within a frustum is modeled as a Gaussian distribution parameterized by its mean and variance. The variance reflects the local volume of the frustum as smaller frustums correspond to narrower Gaussians. This allows the representation to adapt continuously to the geometric scale of the segment. By sampling multiple Gaussian-distributed points within each frustum, the model captures both the mean and variance (first and second spatial moments) of the enclosed volume, ensuring that the aggregated features accurately approximate its volumetric contribution. This approach provides a differentiable alternative to explicit Gaussian volume modeling, as the sampling pattern and segment length $\Delta d_k = d_{k+1} - d_k$ implicitly encode the axial extent of each segment without requiring a separate axial variance term.

3.2 Implicit Scene Representation

The cone-based wave model in WiNeRF induces spatial queries with varying footprints and scales. Narrow conical frustums emphasize localized spatial variation, while wider frustums aggregate information over larger regions. Moreover, the spatial extent of each frustum

increases with distance from the antenna array, reflecting the fact that array-induced angular uncertainty corresponds to larger spatial ambiguity in the far field. As a result, a single frustum may simultaneously span coarse structures (e.g., walls or floors) and finer scattering features (e.g., edges or furniture). Representing such mixed-scale interactions using a single fixed-scale representation would blur fine spatial variation into coarse background structure and limit expressivity.

To address this, WiNeRF employs a multi-resolution implicit scene representation that supports spatial feature encoding across scales. Rather than explicitly reconstructing geometry or material properties, this representation learns a latent, spatially indexed feature field that is queried exclusively through cone-based sampling. This design aligns with the measurement-driven nature of the system and avoids imposing external geometric or visual priors. Figure 2 illustrates this distinction: unlike classical line-ray formulations that rely on a single-scale MLP, WiNeRF uses a multi-resolution hash grid [28] to efficiently support coarse-to-fine spatial variation while maintaining memory efficiency and training stability.

Multi-Resolution Hash Grid Architecture: The implicit scene representation maps each 3D location $\mathbf{x} = (x, y, z)$ to a feature vector constructed from L resolution levels. At each level $l \in [1, L]$, space is discretized into a regular grid, with features stored in a fixed-size hash table to enable high effective resolution without incurring the memory cost of dense voxel grids. The coarse grid levels capture room-scale structures while fine levels resolve small scatterers and detailed geometry.

For each resolution level l , the feature embedding at location $\mathbf{x}$ is computed via trilinear interpolation over the eight nearest grid vertices:

$$f^{(l)}(\mathbf{x}) = \sum_{v \in \mathcal{N}(\mathbf{x}, l)} w_v(\mathbf{x}) \cdot g^{(l)}(h(v)), \quad (2)$$

where $\mathcal{N}(\mathbf{x}, l)$ denotes the neighboring grid vertices at level l , $w_v(\mathbf{x})$ are the trilinear interpolation weights, $h(\cdot)$ is the spatial hash function, and $g^{(l)}(\cdot)$ are trainable feature vectors stored in the hash table. Importantly, this interpolation is applied only to latent multi-scale scene features, not directly to absolute RF phase or to a full electromagnetic field representation. This allows the model to preserve coarse-to-fine spatial continuity in the learned representation without claiming smoothness of the underlying electromagnetic field itself. Features from all resolution levels are concatenated to form a multi-scale embedding:

$$\mathbf{f}(\mathbf{x}) = \bigoplus_{l=1}^L f^{(l)}(\mathbf{x}), \quad (3)$$

which is then passed to downstream neural networks for channel synthesis.

This representation provides three key benefits: (i) *Compactness*: Hash-based table reuses a small pool of trainable features instead of storing full voxel grids at each resolution, greatly reducing memory. (ii) *Multi-scale support*: Coarse grid levels capture large structures such as walls and ceilings, while fine levels resolve small scatterers like furniture and clutter. (iii) *Continuity*: Trilinear interpolation ensures smooth feature variation in the latent scene features with respect to spatial location, which is important for stable volumetric

aggregation without directly imposing smoothness on absolute RF phase.

WiNeRF Neural Network Architecture: Conditioned on the multi-resolution scene features, WiNeRF predicts intermediate latent quantities that govern how samples along each cone contribute to channel synthesis. Specifically, for each sampled location $\mathbf{x}$, the network estimates: (1) a scalar density $\sigma(\mathbf{x})$, which acts as a soft weighting function that modulates the relative contribution of the location during volumetric aggregation along the cone, and (2) a complex-valued response $\mathbf{c}(\mathbf{x}) \in \mathbb{C}$, which encodes amplitude scaling and phase rotation applied to the propagating signal. These quantities are not interpreted as explicit physical material properties or scattering parameters. Instead, they are learned latent variables whose sole role is to support accurate reconstruction of observed CSI under the cone-based sampling and aggregation process.

To improve computational efficiency, WiNeRF adopts a hierarchical two-stage architecture inspired by Zip-NeRF [5]. A lightweight proposal network first evaluates coarse samples along each conical ray to estimate a rough volumetric density distribution, identifying spatial regions with non-negligible contributions to the received signal. The input to the proposal network is the multi-resolution hash grid encoding corresponding to each cone frustum sample, enabling efficient assessment of where meaningful interactions are likely to occur.

Guided by the proposal network’s weights, a second network then processes a refined set of samples, focusing computational resources on regions that contribute most strongly to the channel. This refined network predicts both the scalar density $\sigma(\mathbf{x})$ and the complex-valued response $\mathbf{c}(\mathbf{x})$ used for final channel synthesis. To ensure consistency between stages, the proposal network is trained with auxiliary consistency losses that encourage its density estimates to align with the refined network’s outputs, preventing divergence while significantly reducing the number of expensive network evaluations.

3.3 CSI Reconstruction Module

WiNeRF reconstructs channel state information by aggregating learned complex responses along cone-shaped sampling volumes defined by the wave model. This module combines cone-based sampling with the implicit scene representation to synthesize CSI in a manner consistent with the angular resolution and ambiguity of the antenna array. A key challenge in this aggregation is aligning the continuous spatial support of conical frustums with the discrete multi-resolution hash grid used to encode the scene. WiNeRF addresses this using a level-of-detail (LOD) weighting strategy adapted from Zip-NeRF [5], where the spatial extent of each frustum determines how features from different grid resolutions contribute to the final response.

The channel reconstruction process begins by partitioning the angular space around each receiver into P non-overlapping bimodal wedges, where each wedge corresponds to a cone of arrival directions defined by the array’s intrinsic angular resolution. Within each cone, contributions are accumulated along the propagation direction by discretizing the cone into N conical frustums. For a given cone, the complex channel contribution is computed as a weighted sum over frustums:

$$\mathbf{H}_{\text{cone}} = \sum_{i=1}^N w_i \mathbf{c}_i, \quad (4)$$

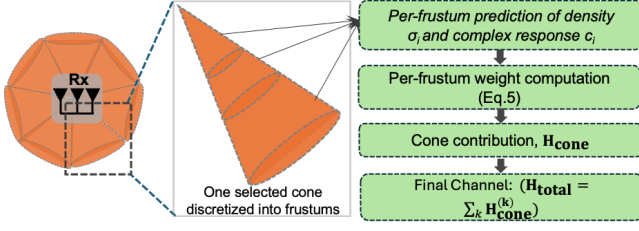


Figure 6: The full pipeline of CSI estimation via per-frustum prediction and aggregation.

where c_i denotes the complex-valued response predicted at frustum i , and w_i is a scalar weight that modulates the relative contribution of that location during aggregation. The weights are computed using an exponential attenuation formulation:

$$w_i = T_i \cdot \alpha_i, \quad \alpha_i = 1 - \exp(-\sigma_i \delta_i), \quad T_i = \exp\left(-\sum_{j=1}^{i-1} \sigma_j \delta_j\right), \quad (5)$$

where σ_i is the learned density at frustum i , δ_i is its axial length, α_i controls the local contribution strength, and T_i accumulates the residual contribution from preceding frustums along the cone. These quantities are treated as learned aggregation variables rather than explicit physical interaction probabilities. Figure 6 depicts the channel reconstruction process. The full channel response at a receiver is obtained by summing contributions from all cones:

$$\mathbf{H}_{\text{total}} = \sum_{k=1}^P \mathbf{H}_{\text{cone}}^{(k)}, \quad (6)$$

where $\mathbf{H}_{\text{cone}}^{(k)}$ denotes the aggregated response from the k -th cone. To support multi-antenna and multi-subcarrier CSI, the network predicts complex responses with dimensions $N_{\text{rx}} \times N_{\text{sc}}$, enabling simultaneous reconstruction across spatial and frequency domains within a unified framework.

It is worth noting that the entire reconstruction pipeline, that progresses from cone initialization and frustum sampling to weighted aggregation and CSI synthesis, is fully differentiable. This allows end-to-end optimization of both the implicit scene representation and the neural networks using gradient-based learning.

3.4 Training Objective

WiNeRF is trained by aligning synthesized channel responses from the forward rendering pipeline with measured CSI through end-to-end differentiable optimization. Rather than performing inverse rendering or recovering explicit physical scene parameters, training constrains the latent neural representation so that its forward-rendered CSI matches observations collected by commodity hardware. Mini-batches of receiver locations are sampled from the training set, and the corresponding measured CSI matrices serve as supervision. Gradients propagate from the CSI reconstruction loss through the channel aggregation, implicit scene representation, and cone-based sampling process.

A central challenge when training on experimental CSI is that the absolute phase measured by commodity devices is unreliable. Hardware-induced effects such as carrier frequency offset (CFO),

sampling frequency offset (SFO), and packet detection delay introduce packet-dependent global phase drifts that are unrelated to propagation [18]. As a result, directly regressing absolute phase forces the model to explain device-specific artifacts rather than multipath structure.

Following common practice in WiFi CSI processing, we first apply standard packet-level phase sanitization to mitigate global CFO and SFO components that are shared across antennas and subcarriers within each snapshot [18, 20]. However, residual phase offsets and snapshot-to-snapshot drift remain unavoidable due to hardware instability and measurement noise, making absolute phase values inconsistent across packets [20, 34].

To ensure that learning focuses on stable and physically informative structure, we propose a relative phase loss that operates on unwrapped, antenna-referenced phase differences rather than absolute phase. This removes remaining snapshot-level global phase offsets while preserving spatial phase differences across antennas and subcarriers, which encode relative angle-of-arrival and delay information under the array geometry. [11, 13]. For each CSI snapshot, we normalize the phase by subtracting the phase of a reference antenna element from all antenna channels. This removes snapshot-level global phase drift while preserving spatial phase differences across antennas and subcarriers, which encode relative AoA and delay information. Both the predicted channel $\hat{\mathbf{H}}$ and the measured channel $\mathbf{H}$ are normalized in this manner prior to loss evaluation.

The training objective minimizes the normalized mean squared error between the relative-phase CSI representations. Let $\mathbf{H} \in \mathbb{C}^{N_r \times N_f}$ denote the measured CSI at a receiver with N_r antennas and N_f subcarriers. We define the relative-phase CSI $\mathbf{H}_{\text{rel}}$ by normalizing each CSI entry with respect to a reference antenna-subcarrier pair:

$$\mathcal{L}_{\text{rel-phase}} = \frac{\|\hat{\mathbf{H}}_{\text{rel}} - \mathbf{H}_{\text{rel}}\|^2}{\|\mathbf{H}_{\text{rel}}\|^2} \quad (7)$$

This loss does not aim to recover absolute phase or correct hardware impairments. Instead, it encourages the learned neural field to reproduce the relative phase structure supported by the measurements under the cone-based sampling and rendering process, rather than arbitrary sub-wavelength absolute field oscillations, enabling accurate CSI reconstruction for downstream multi-antenna tasks.

4 Implementation

4.1 Data Collection

Mobile Platform: As shown in Figure 7a, we employ a mobile platform based on iROBOT Create 3 to collect the spatial CSI data. The robot is equipped with an ASUS RT-AX86U router, configured as a receiver (Rx). The router’s firmware was modified using Nexmon CSI [32] tool to enable CSI extraction, and WiROS [2] framework is used to collect CSI and RSSI. We also use an Intel RealSense depth D435i camera for simultaneous localization and mapping (SLAM) based on ORBSLAM3 [6] to track the 6-DoF pose of the robot. The robot was moved manually throughout the space using random trajectories.

Wireless Devices: We collected CSI data on a 20 MHz channel in the 5 GHz band (Channel 157), resulting in 50 active subcarriers captured by 4 Rx antennas from a single spatial stream. To address the inherent phase noise in commodity WiFi hardware, we implemented

a two-step calibration process. First, we modified the firmware to disable automatic receive-chain switching. Second, we performed a dynamic phase-compensation procedure [2] prior to data collection to ensure that the CSI matrices retain a consistent and correct phase. Maintaining phase coherence across antennas and subcarriers is essential because the relative phase encodes the channel’s underlying multipath structure. Without proper alignment, the signatures of individual paths become distorted, which in turn affects any downstream processing that relies on stable spatial or frequency characteristics. With the compensated phase, the resulting CSI accurately reflects the true multipath profile of the environment.

Experimental Environments: We collected data in three distinct indoor environments: a Meditation Room ($4\text{ m} \times 2.25\text{ m} \times 2.4\text{ m}$), an Office Space ($5\text{ m} \times 3\text{ m} \times 2.7\text{ m}$), and a Conference Room ($12\text{ m} \times 6\text{ m} \times 2.4\text{ m}$), resulting in a total of 30k synchronized CSI and pose samples, with an average data sampling density of 4 samples per ft^3 (equivalent of 150 samples per meter^3). All environments contain various furniture configurations and exhibit rich multipath propagation conditions. The Conference Room dataset includes numerous non-line-of-sight (NLoS) cases caused by a large table positioned at the same height as the robotic platform, which blocks the line-of-sight (LoS) path between the transmitter and receiver across most locations. The Office Space, on the other hand, contains several metal cabinets that enhance multipath propagation and increase signal reflections. A snapshot of all environments is shown in Figs. 7b, 7c, and 7d.

4.2 Model Training and Configuration

We implement WiNeRF in PyTorch (v2.3) with CUDA 12 and employ the AdamW optimizer for end-to-end training. All experiments are conducted on our server equipped with two NVIDIA RTX A6000 GPUs. Training uses a batch size of 128, an initial learning rate of 10^{-3} with cosine decay to 10^{-4} , and a maximum of 50,000 epochs. For each environment, WiNeRF is trained separately to predict the full complex CSI matrix for 4 receive antennas across 50 OFDM subcarriers, and the current system does not include a cross-environment transfer or rapid adaptation mechanism. Evaluations are conducted using 5-fold cross-validation, where each fold uses an 80%/20% split between training and testing samples. The per-fold wall-clock training time is approximately 15–30 minutes on our RTX A6000 server, depending on the dataset split and scene complexity.

Overfitting is mitigated by the broad diversity of CSI measurements across spatial positions, LoS/NLoS regions, and multipath patterns induced by metal reflectors. Combined with the strong inductive biases of our NeRF-style architecture [5, 25, 28] (e.g., volumetric density fields and hash-grid feature sharing), the model is encouraged to learn physically consistent propagation behavior instead of memorizing individual CSI samples.

WiNeRF uses one proposal network and one reconstruction network. The proposal network employs a multiresolution hash grid with a base resolution of 16^3 and a finest resolution of 256^3 , corresponding to 5 resolution levels, and contains 0.86M learnable parameters. This proposal stage efficiently identifies regions of significant multipath energy. The reconstruction network uses a larger multiresolution hash grid with the same base resolution of 16^3 and a finest

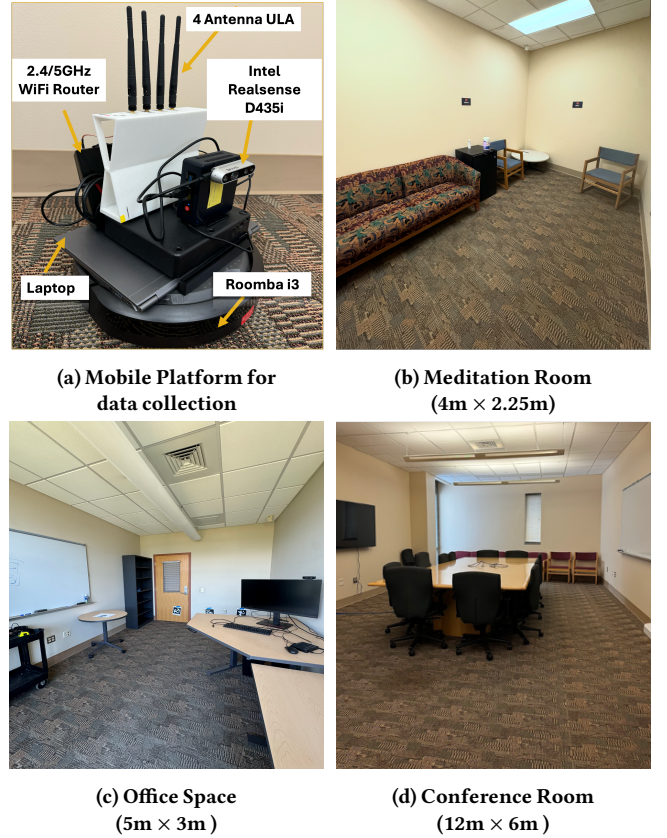


Figure 7: (a) Mobile platform used for CSI data collection with a 4-antenna WiFi router, Intel Realsense depth camera, and laptop mounted on a Roomba. (b-d) WiNeRF experimental environments.

resolution of 2048^3 , corresponding to 8 resolution levels, and contains 1.87M learnable parameters. Across both networks, WiNeRF contains 2.73M learnable parameters in total.

The reconstruction network processes hash-encoded features through two multi-layer perceptrons (MLPs). A density MLP (2 layers, 64 units per layer) outputs volumetric density along with a 64-dimensional bottleneck vector. A view-conditioned response branch (2 layers, 256 units per layer) combines this bottleneck with directionally encoded viewing cones to predict the complex-valued response for all subcarriers and receiver antennas. Under our setup with 4 receive antennas, 1 transmit antenna, and 50 subcarriers, this branch predicts the real and imaginary parts of the full complex CSI tensor. The final complex CSI is computed via volumetric integration, where the predicted responses are weighted by the volumetric density and integrated along each cone to synthesize the channel response.

Cone Quantization: To operationalize the observability-driven cone model in our 4-antenna ULA setting, we divide the azimuth plane into 12 equal 30° sectors. During both training and inference, we emit *bimodal cone-wedge* support regions from the receiver for each discretized azimuth bin. Each bimodal cone-wedge has two properties: (i) its elevation span covers the entire elevation range,

since our receiver array does not provide elevation resolution, and (ii) its azimuth span is restricted to the corresponding 30° bin, reflecting the angular discretization used in our model. This implementation directly mirrors the finite angular observability assumed in the wave model, allowing the representation to remain physically meaningful while staying aligned with the hardware’s actual directional resolution.

5 Evaluation

We evaluate WiNeRF on the task of estimating the full 4×50 complex CSI channel across the three indoor environments shown in Figure 7, selected to span different room sizes, geometries, and multipath conditions and thus assess site-specific reconstruction under diverse static layouts. Here, 4 denotes the number of receive antennas and 50 denotes the number of OFDM subcarriers per packet. This setup allows us to assess both spatial (across antennas) and frequency (across subcarriers) reconstruction fidelity, which is critical for practical MIMO applications such as beamforming, localization, and coverage mapping.

5.1 Baselines

We compare our system’s performance against the following state-of-the-art and classical baselines:

- **NeRF²** [40], which models the spatial behavior of radio signals using a continuous volumetric representation parameterized by MLPs. The model can estimate complex CSI by accumulating contributions along the signal propagation path. We employ the publicly available implementation and train it on our datasets. Since the framework is designed for scenarios with moving transmitters, we prepare compatible datasets by fixing device locations and interchanging the transmitter and receiver roles.
- **NeWRF** [21], which models wireless propagation as a volumetric field guided by direction-of-arrival (DoA) priors. As NeWRF requires DoA input, we estimate the top three arrival directions per receiver using the MUSIC algorithm on the ground truth CSI matrix and use them to train the model on our datasets.
- **VAE** [17], a generative model. Variational Autoencoders have previously been used to generate wireless channel representations, albeit in different contexts [20].
- **DC-GAN** [12], a deep convolutional GAN chosen as a baseline because it represents a generative modeling approach capable of learning complex, high-dimensional data distributions. This allows a direct comparison between a generative neural network and WiNeRF.

5.2 Evaluation Metrics

We evaluate WiNeRF using three complementary metrics that assess different aspects of channel estimation quality: overall channel fidelity, signal power estimation accuracy, and phase coherence preservation. These metrics collectively provide a comprehensive assessment of the model’s ability to reconstruct physically meaningful and actionable channel representations.

Channel Prediction SNR: This metric measures the overall fidelity of the complex channel estimate by computing the signal-to-noise ratio between predicted and ground-truth CSI matrices, and is employed by prior works to assess overall channel quality [20, 22, 40].

The SNR is defined as:

$$\text{SNR}_{\text{pred}} = -10 \log_{10} \left(\frac{\|\mathbf{H}_{\text{pred}} - \mathbf{H}_{\text{true}}\|_2}{\|\mathbf{H}_{\text{true}}\|_2} \right) \quad [\text{dB}] \quad (8)$$

where $\mathbf{H}_{\text{pred}}$ and $\mathbf{H}_{\text{true}}$ are the predicted and ground-truth channel matrices, respectively. Higher SNR values indicate better reconstruction quality, and, by construction, this metric captures both amplitude and phase accuracy.

RSSI Estimation Error: This metric evaluates the accuracy of received signal strength by comparing the average power of predicted and measured channels. RSSI is computed from the CSI magnitude as:

$$\text{RSSI} = \frac{1}{N_{\text{ant}} N_{\text{sc}}} \sum_{i=1}^{N_{\text{ant}}} \sum_{j=1}^{N_{\text{sc}}} |H[i, j]|^2 \quad [\text{dBm}], \quad (9)$$

where N_{ant} and N_{sc} denote the number of antennas and subcarriers. The RSSI error is the absolute difference between predicted and ground-truth values; lower errors indicate better power estimation, which is critical for coverage mapping and link budget analysis [26].

Relative Phase Similarity: This metric evaluates how well the model preserves spatial phase relationships across antenna elements, which is critical for applications such as beamforming and angle-of-arrival estimation. For each subcarrier, we normalize the phase by the first antenna to remove global phase offsets, then compute the cosine similarity between predicted and ground-truth phase vectors:

$$\text{Similarity} = \frac{\phi_{\text{pred}} \cdot \phi_{\text{true}}}{\|\phi_{\text{pred}}\| \|\phi_{\text{true}}\|}, \quad (10)$$

where ϕ_{pred} and ϕ_{true} are the phase vectors across antennas for a given subcarrier. Values range from -1 to 1 , with 1 indicating perfect alignment. This metric quantifies the model’s ability to capture the relative phase structure across antennas and subcarriers.

5.3 Channel Estimation Accuracy

SNR performance: As shown in Figure 8a, WiNeRF achieves a median prediction SNR of 5.30 dB and 90th percentile of 11.29 dB, outperforming all baselines by at least 130%. This is mainly because all baselines heavily rely on dense training data or priors. NeRF² requires a very dense sampling of ~ 178 measurements/ ft^3 , while our experimental data has an average density of ~ 4 measurements/ ft^3 . On the other hand, NeWRF heavily depends on precise angle-of-arrival (AoA) priors, which are not available in real-world measurements. Figure 8b presents the prediction SNR across different environments, showing that WiNeRF consistently outperforms all baselines regardless of environment size or multipath complexity. In the smallest environment, the Meditation Room, WiNeRF achieves a median SNR of 9.61 dB, compared to 3.96 dB for NeRF² and 3.05 dB for NeWRF, yielding gains of 5.6–8.8 dB. While performance slightly decreases in larger environments due to increased spatial extent and richer multipath, WiNeRF still surpasses all baselines by a clear margin. Even in these more challenging settings, the resulting SNR remains sufficient to produce actionable and spatially consistent estimations across space and antenna elements, as we demonstrate in the following results.

RSSI Estimation Performance: While the previous analysis focused on SNR, which captures both the magnitude and phase similarity of the predicted complex channel, we now examine the role of CSI magnitude in RSSI estimation. RSSI serves as a key indicator

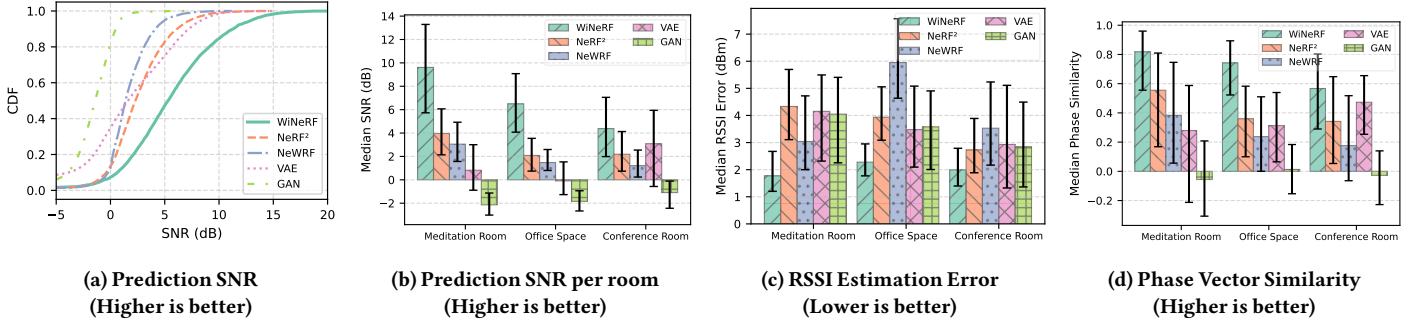


Figure 8: WiNeRF outperforms all baselines in channel estimation accuracy at unseen WiFi receiver locations, while accurately preserving both magnitude and phase characteristics.

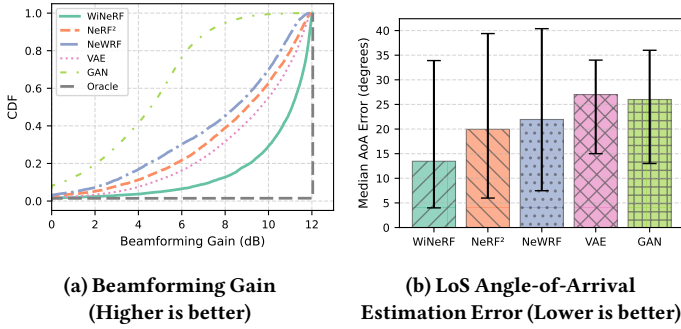


Figure 9: WiNeRF preserves phase coherence for practical MIMO tasks: (a) Beamforming and (b) AoA estimation, using predicted CSI with classical signal processing methods.

for coverage mapping and MCS index selection in WiFi systems. Figure 8c presents the per-room RSSI error, showing that WiNeRF achieves a median RSSI error between 1.8 and 2.3 dB, consistently outperforming all baselines by 35–50% and delivering more accurate power estimations across diverse environments.

Channel Phase Similarity Performance: Preserving relative phase across antennas and subcarriers is essential for maintaining multipath information. To evaluate this capability, we compute the similarity between the estimated and ground-truth relative phase vectors. Figure 8d shows the per-room phase similarity, where WiNeRF achieves a median phase similarity between 0.5 and 0.8, outperforming all baselines by 40–300%. WiNeRF’s performance in relative phase estimation slightly drops in the largest room due to numerous NLoS cases. We next show how these similarity factors are reflected in AoA estimation.

5.4 Beamforming Gain Performance

We validate that WiNeRF’s channel estimates are not only numerically accurate but also physically actionable, preserving the phase coherence necessary for beamforming. To test this, we perform equal-gain combining (EGC), a stringent benchmark that relies solely on CSI phase information using the estimated CSI for test data. In EGC, the combining weights are defined as $w_{m,s} = e^{-j\angle h_{m,s}}$, meaning any

phase error directly causes destructive interference and loss of array gain. We use our test data and the estimated CSI across 4 antennas in all three environments and compute the beamforming gain as

$$G_s = 10 \log_{10} \left(\frac{P_{\text{EGC},s}}{P_{\text{avg},s}} \right), \quad (11)$$

aggregated over 50 subcarriers per packet. Importantly, the combining weights are derived from the predicted CSI but applied to the measured channels, isolating phase accuracy from other confounding factors. Figure 9a presents the CDFs of beamforming gains. WiNeRF achieves a median gain of 11.4 dB, closely tracking the oracle’s 12 dB (only a 0.6 dB or 5% gap), while outperforming all baselines by large margins, up to 145% higher than NeRF² and even larger improvements over NeWRF, VAE, and GAN. Unlike these models, WiNeRF’s gain distribution is consistently high, with fewer than 10% of packets falling below 8 dB, compared to about 40% for baselines.

In practical terms, this 3 dB improvement in beamforming translates to a 19% increase in theoretical link capacity for a 20 MHz Wi-Fi channel at 15 dB baseline SNR, equivalent to a doubling of range or a fourfold reduction in transmit power. These results demonstrate that WiNeRF’s CSI estimates retain the correct relative phase structure across antennas and subcarriers, making them immediately usable for beamforming and other downstream PHY-layer operations without requiring training of a new downstream task.

5.5 LoS Angle-of-Arrival Estimation Accuracy

WiNeRF is designed to preserve the underlying physics of wireless propagation, particularly multipath structure such as the AoA, which is captured in the relative phase across antennas. The AoA of the direct path between the WiFi transmitter and receiver is especially important for applications such as localization, tracking, and robotic path planning. To validate that WiNeRF’s channel estimates retain this physical information, we apply the classical MUSIC algorithm [18, 31] to the predicted CSI and evaluate its ability to recover the LoS AoA at unseen receiver locations within the same environment.

Figure 9b shows the LoS AoA estimation error across methods. WiNeRF achieves a median error of 13.0°, consistently outperforming all baselines. In practical terms, this corresponds to an angular bias of roughly 11% within a 120° sector, translating to a localization error of about 23 cm at a 1m distance that is sufficient for coarse, room-scale tracking without additional model training. As illustrated in

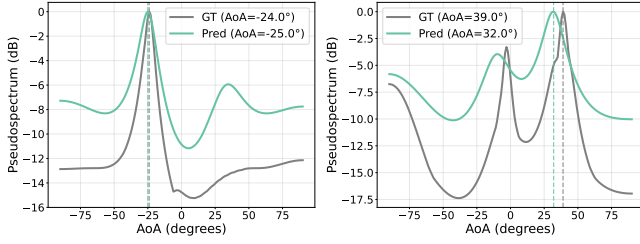


Figure 10: Multipath resolution results using estimated CSI; WiNeRF preserves accurate multipath AoA characterization

Figure 10, WiNeRF’s predicted CSI produces sharp, well-aligned LoS and multipath peaks. This ability to maintain spatial structure makes WiNeRF’s CSI actionable for beam tracking and interference localization. It is worth noting that unlike prior work such as NeRF², which trains a separate downstream network to infer spatial spectra from synthetic wireless data generated by a base NeRF model, WiNeRF directly enables classical signal processing methods (e.g., MUSIC) to recover multipath characteristics from its predicted CSI. This distinction demonstrates that WiNeRF truly captures the underlying physical multipath structure rather than merely memorizing patterns inside images.

5.6 Radio Coverage Prediction via RSSI Mapping

Next, we demonstrate WiNeRF’s ability to preserve spatial CSI patterns across an entire room. This is an essential capability for applications such as coverage mapping, access point (AP) placement, link budgeting, and trajectory planning. Using pretrained CSI estimation models from WiNeRF and the baselines, we query each model over a dense 2D spatial grid in the conference room and convert the predicted CSI into RSSI values. Figure 11 compares the inferred coverage maps with sparse ground-truth measurements, showing that WiNeRF accurately captures the expected spatial variations, including RSSI drops in NLoS regions and constructive or destructive multipath interference patterns. Quantitatively, WiNeRF achieves lower median RSSI error than the strongest learned baseline and also reduces worst-case error, indicating that its spatial predictions are not only visually plausible but also more reliable for deployment-facing tasks such as dead-zone identification and path planning. This results in actionable, physically consistent predictions suitable for proactive, context-aware decision-making.

For comparison, Figure 11 also includes a ray-tracing-based coverage map generated using SionnaRT and a detailed 3D CAD model of the same environment. While ray tracing can yield precise results when surface materials and geometry are modeled accurately, its performance degraded around the conference table due to incomplete geometry and missing details beneath the table captured during scanning. This comparison further highlights that WiNeRF can recover room-scale coverage structure directly from sparse measurements even when complete geometric fidelity is unavailable.

6 Sensitivity Analysis

Next, we evaluate WiNeRF’s sensitivity to key design factors, including cone geometry, hash-grid parameters, and training data density.

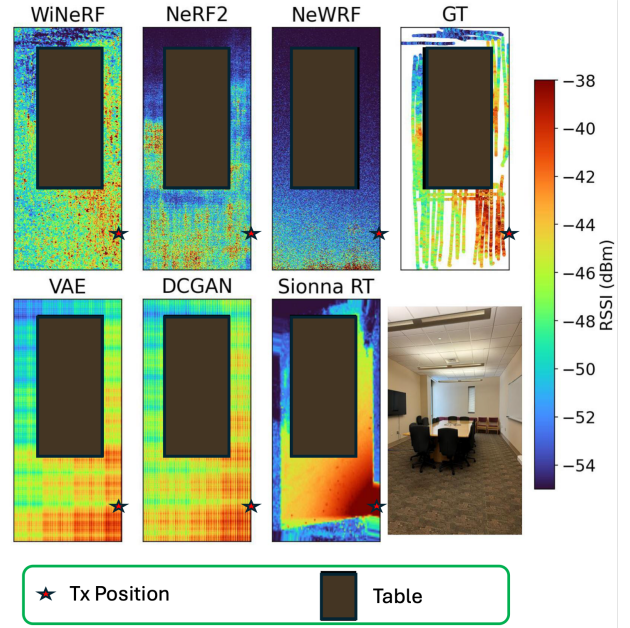


Figure 11: RSSI spatial coverage prediction in Conference Room. WiNeRF accurately reproduces room-scale coverage patterns with fine spatial detail matching ground truth, capturing both the LoS hotspot and NLoS attenuation at edges.

We further conduct an ablation study to quantify the contribution of the proposed cone-based ray structure.

Impact of Cone Sizes: We evaluate how the angular discretization in our wave model influences channel estimation accuracy by varying the number of conical rays around each receiver. This parameter defines the trade-off between spatial resolution and computational efficiency in the cone-based representation. We test configurations with 2–16 cones, corresponding to azimuthal bin widths from 180° to 22.5°, while keeping all other model parameters fixed. As shown in Fig. 12a, performance improves with finer angular sampling up to 12 cones, which yields the highest median prediction SNR of 5.3 dB. This aligns with the theoretical angular resolution limit imposed by our 4-antenna array hardware.

Impact of Hash Grid Levels: We analyze how the number of multi-resolution grid levels influences channel estimation accuracy (Fig. 12b). The grids span spatial resolutions from coarse (meter-scale) to fine (centimeter-scale), and we vary the number of levels used to represent this range. Results show a clear upward trend in performance as grid depth increases, confirming that distributing the representation across multiple exponentially spaced levels enables the model to capture both large-scale spatial trends and fine-grained multipath variations. This finding underscores the value of hierarchical feature encoding for accurately modeling complex wireless propagation environments.

Impact of Training Data Density: We evaluate how training data density influences channel estimation accuracy across all environments (Fig. 12c). As expected, higher sampling density improves performance for all models. However, WiNeRF consistently achieves

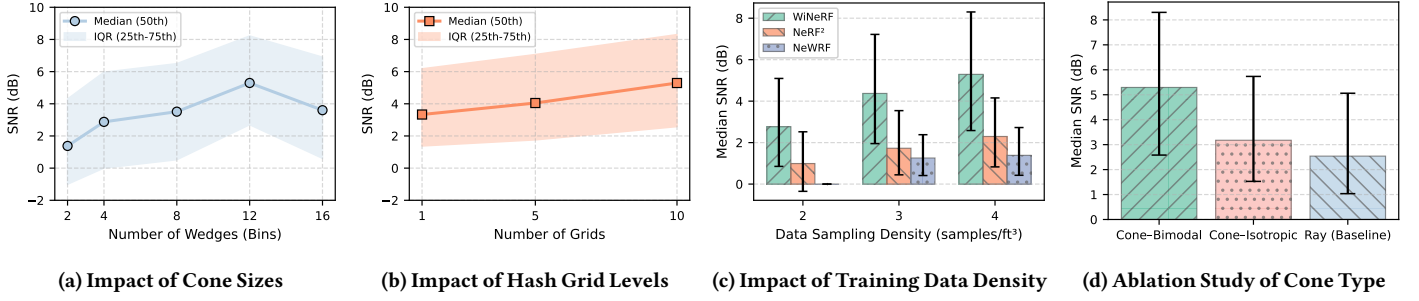


Figure 12: WiNeRF Sensitivity Analysis

higher SNR and degrades more gracefully under sparse data conditions, whereas baseline methods show sharp performance drops at lower densities. This robustness highlights WiNeRF’s efficiency in learning spatial channel structure from limited training data and predicting at held-out receiver locations within the same environment..

Ablation Study: We compare three spatial propagation modeling strategies to assess their impact on channel estimation performance (Fig. 12d): (1) our proposed anisotropic wedge cone, which uses wedge-shaped sampling regions to reflect the directional angular observability constraints of the antenna array; (2) an isotropic cone, which radiates uniformly in all directions; and (3) a simple ray baseline, which uses idealized line rays without spatial spreading, similar to NeRF² [40] and NeWRF [21]. All configurations employ 12 angular bins (30° azimuth resolution) and are evaluated across three environments, with the ray baseline using 12 rays centered in each bin. As shown in Fig. 12d, the anisotropic wedge cone achieves the highest median SNR of 5.29 dB, demonstrating that incorporating angular anisotropy consistent with the array’s measurement constraints improves ULA-based channel estimation over both omnidirectional and idealized ray models.

7 Discussion and Future Works

WiNeRF effectively learns both phase coherence and multipath characteristics, enabling physically meaningful and actionable channel representations. By embedding hardware-informed inductive biases directly into the neural field, the framework successfully reconstructs complex CSI even under the tight sensing and communication constraints typical of embedded platforms. However, the current paper evaluates WiNeRF only in static environments and therefore does not establish performance under moving scatterers and changing layouts. While the current framework focuses on linear antenna arrays, it is readily extendable to other antenna geometries.

We intentionally designed WiNeRF as a site-specific model to support high-fidelity Digital Twins and spatial planning tools, where capturing the unique multipath “fingerprint” of an environment is the primary objective. This specialization allows the model to achieve high prediction accuracy in complex, static scenes from sparse measurements, but it also means that WiNeRF is currently trained separately for each environment and does not include a cross-environment transfer or rapid adaptation mechanism. At present, WiNeRF is evaluated as a server-side system for digital twinning applications, and embedded deployment remains an important direction for future work.

Several extensions remain open. First, dynamic environments with moving objects are not validated by the current experiments. A possible future direction is to incorporate object-aware neural representations together with external physical priors such as camera or radar feeds to help disentangle static background structure from dynamic components. Second, transfer learning or related adaptation strategies could reduce retraining cost and improve portability across environments.

8 Conclusion

We presented WiNeRF, a neural wireless channel modeling framework that learns a spatially continuous channel representation constrained by what is observable from commodity wireless hardware, without relying on idealized thin-ray propagation or external priors such as known geometry or explicit AoA information. By incorporating typical constraints of embedded platforms such as antenna geometry and finite angular resolution into its modeling formulation, WiNeRF focuses on learning the subset of channel variations that can be reliably inferred from sparse and noisy measurements. This constrained representation enables stable and reusable channel estimates that can be directly queried for downstream tasks. Importantly, the learned CSI can be used with standard signal-processing pipelines—such as beamforming, LoS AoA estimation, and signal coverage mapping—without task-specific retraining or changes to existing hardware or protocols. Experimental evaluation across three indoor environments demonstrates a median prediction SNR of 5.30 dB (3.0–6.7 dB improvement over prior methods), near-oracle beamforming performance with only 0.6 dB loss, and a median LoS AoA error of 13°. Together, these results show that WiNeRF provides a practical and task-agnostic channel representation that supports proactive, wireless-aware decision making in embedded JSAC systems.

Acknowledgments

We thank Mr. Jakob Link and Dr. Matthias Hollick from the Secure Mobile Networking Lab at TU Darmstadt for their help with updating the router firmware files used in our data collection pipeline. We also thank Navid Tajkhorshid, an undergraduate student at the University of Illinois Urbana-Champaign, for his help with setting up the data collection system. This work was supported in part by the National Science Foundation (NSF) under Grants #2414227, and by Intel.

References

- [1] Fadel Adib and Dina Katabi. 2013. See through walls with WiFi!. In *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*. 75–86.
- [2] Aditya Arun, William Hunter, and Dinesh Bharadia. 2023. Demo Abstract: Accessible WiFi sensing leveraging Robot Operating System. In *Proceedings of the 22nd International Conference on Information Processing in Sensor Networks*. 356–357.
- [3] Jonathan T. Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P. Srinivasan. 2021. Mip-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 5855–5864.
- [4] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. 2023. Zip-nerf: Anti-aliased grid-based neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 19697–19705.
- [5] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. 2023. Zip-nerf: Anti-aliased grid-based neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 19697–19705.
- [6] Carlos Campos, Richard Elvira, Juan J Gómez Rodríguez, José MM Montiel, and Juan D Tardós. 2021. Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam. *IEEE transactions on robotics* 37, 6 (2021), 1874–1890.
- [7] Chen Chen, Gang Zhou, and Youfang Lin. 2023. Cross-Domain WiFi Sensing with Channel State Information: A Survey. *ACM Comput. Surv.* 55, 11, Article 231 (Feb. 2023), 37 pages. doi:10.1145/3570325
- [8] Xingyu Chen, Zihao Feng, Ke Sun, Kun Qian, and Xinyu Zhang. 2024. Rfcanvas: Modeling rf channel by fusing visual priors and few-shot rf measurements. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems*. 464–477.
- [9] Sebastian Dörner, Marcus Henninger, Sebastian Cammerer, and Stephan ten Brink. 2020. WGAN-based Autoencoder Training Over-the-air. In *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 1–5.
- [10] Richard P Feynman, Robert B Leighton, and Matthew Sands. 1965. The feynman lectures on physics; vol. i. *American Journal of Physics* 33, 9 (1965), 750–752.
- [11] Andrea Goldsmith. 2005. *Wireless communications*. Cambridge university press.
- [12] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. *Advances in neural information processing systems* 27 (2014).
- [13] Monson H. Hayes. 1996. *Statistical Digital Signal Processing and Modeling* (1st ed.). John Wiley & Sons, Inc., New York, NY, USA.
- [14] Jakob Hoydis, Fayçal Ait Aoudia, Sebastian Cammerer, Merlin Nimier-David, Nikolaus Binder, Guillermo Marcus, and Alexander Keller. 2023. Sionna RT: Differentiable ray tracing for radio propagation modeling. In *2023 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 317–321.
- [15] Zhong Ji, Bin-Hong Li, Hao-Xing Wang, Hsing-Yi Chen, and Tapan K Sarkar. 2001. Efficient ray-tracing methods for propagation prediction for indoor wireless communications. *IEEE Antennas and propagation Magazine* 43, 2 (2001), 41–49.
- [16] Shuaifeng Jiang, Qi Qu, Xiqing Pan, Abhishek Agrawal, Richard Newcombe, and Ahmed Alkhateeb. 2025. Learnable wireless digital twins: Reconstructing electromagnetic field with neural representations. *IEEE Open Journal of the Communications Society* (2025).
- [17] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).
- [18] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. Spotfi: Decimeter level localization using wifi. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*. 269–282.
- [19] Fan Liu, Yuanhao Cui, Christos Masouros, Jie Xu, Tony Xiao Han, Yonina C. Eldar, and Stefano Buzzi. 2022. Integrated Sensing and Communications: Toward Dual-Functional Wireless Networks for 6G and Beyond. *IEEE Journal on Selected Areas in Communications* 40, 6 (2022), 1728–1767. doi:10.1109/JSAC.2022.3156632
- [20] Zikun Liu, Gagandeep Singh, Chenren Xu, and Deepak Vasishth. 2021. FIRE: enabling reciprocity for FDD MIMO systems. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 628–641.
- [21] Haofan Lu, Christopher Vaththuer, Baharan Mirzasoleiman, and Omid Abari. 2024. A Deep Learning Framework for Wireless Radiation Field Reconstruction and Channel Prediction. *arXiv preprint arXiv:2403.03241* (2024).
- [22] Andrew Luo, Yilun Du, Michael Tarr, Josh Tenenbaum, Antonio Torralba, and Chuhan Gan. 2022. Learning neural acoustic fields. *Advances in Neural Information Processing Systems* 35 (2022), 3165–3177.
- [23] Frederic Matter, Tobias Fischer, Marius Pesavento, and Marc E Pfetsch. 2022. Ambiguities in DoA estimation with linear arrays. *IEEE Transactions on Signal Processing* 70 (2022), 4395–4407.
- [24] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Commun. ACM* 65, 1 (2021), 99–106.
- [25] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Commun. ACM* 65, 1 (2021), 99–106.
- [26] Andreas F Molisch. 2012. *Wireless communications*. Vol. 34. John Wiley & Sons.
- [27] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. 2022. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (ToG)* 41, 4 (2022), 1–15.
- [28] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. 2022. Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)* 41, 4 (2022), 1–15.
- [29] Tribhuvanesh Orekondy, Pratik Kumar, Shreya Kadambi, Hao Ye, Joseph Soriaga, and Arash Behboodi. 2023. Winert: Towards neural ray tracing for wireless channel modelling and differentiable simulations. In *The Eleventh International Conference on Learning Representations*.
- [30] Arogyaswami Paulraj, Rohit Nabar, and Dhananjay Gore. 2003. *Introduction to space-time wireless communications*. Cambridge university press.
- [31] Ralph Schmidt. 1986. Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation* 34, 3 (1986), 276–280.
- [32] Matthias Schulz, Daniel Wegemer, and Matthias Hollick. 2017. Nexmon: The C-based Firmware Patching Framework. <https://nexmon.org>
- [33] Richard Szeliski. 2022. *Computer vision: algorithms and applications*. Springer Nature.
- [34] Jit Ken Tan. 2006. *An adaptive orthogonal frequency division multiplexing baseband modem for wideband wireless channels*. Ph. D. Dissertation. Master's thesis, MIT.
- [35] Harry L Van Trees. 2002. *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons.
- [36] Joram Walfisch and Henry L Bertoni. 1988. A theoretical model of UHF propagation in urban environments. *IEEE Transactions on antennas and propagation* 36, 12 (1988), 1788–1796.
- [37] Jie Wang, Ying Ding, Shujie Bian, Yang Peng, Miao Liu, and Guan Gui. 2019. UL-CSI data driven deep learning for predicting DL-CSI in cellular FDD systems. *IEEE Access* 7 (2019), 96105–96112.
- [38] Jian Wang, Neeraj Varshney, Camillo Gentile, Steve Blandino, Jack Chuang, and Nada Golmie. 2022. Integrated Sensing and Communication: Enabling Techniques, Applications, Tools and Data Sets, Standardization, and Future Directions. *IEEE Internet of Things Journal* 9, 23 (2022), 23416–23440. doi:10.1109/IJOT.2022.3190845
- [39] Hao Ye, Le Liang, Geoffrey Ye Li, and Bing-Hwang Juang. 2020. Deep learning-based end-to-end wireless communication systems with conditional GANs as unknown channels. *IEEE Transactions on Wireless Communications* 19, 5 (2020), 3133–3143.
- [40] Xiaopeng Zhao, Zhenlin An, Qingrui Pan, and Lei Yang. 2023. Nerf2: Neural radio-frequency radiance fields. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*. 1–15.
- [41] Beixiong Zheng, Changsheng You, and Rui Zhang. 2020. Fast channel estimation for IRS-assisted OFDM. *IEEE Wireless Communications Letters* 10, 3 (2020), 580–584.