GAUSSTWIN: A UNIFIED SCENE REPRESENTATION FOR WIRELESS CHANNEL PREDICTION WITH SURFACE RECOVERY

Anonymous authors
Paper under double-blind review

ABSTRACT

Given channel measurements at a set of transmitter (Tx)–receiver (Rx) links, site-specific channel prediction aims to capture the effects of the propagation environment on angular and delay power distributions and received power at unmeasured Tx–Rx links. Existing methods either take scene geometry as prior information or learn radio representations that do not recover surfaces. We propose GaussTwin, a site-specific wireless channel model that represents the scene with oriented 2D Gaussian surface primitives, predicts channel characteristics at unseen links, and recovers explicit environment geometry from channel measurements. A physical propagation operator generates these characteristics for different Tx–Rx pairs from the primitives and models both specular reflection and diffuse scattering. To the best of our knowledge, GaussTwin is the first method that predicts channel characteristics and can recover explicit environment geometry from channel measurements at the same time *without geometric supervision or prior information*. Results across five indoor scenes and one outdoor city block show that GaussTwin achieves high-quality predictions while retaining strong surface recovery capability. On public datasets with real measurements, GaussTwin outperforms state-of-the-art baselines on spatial spectrum prediction and achieves comparable performance on received power prediction.

1 INTRODUCTION

Precise prediction of wireless channel characteristics at unmeasured transmitter (Tx)–receiver (Rx) links needs a site-specific model that forms a spatially consistent, continuous wireless radiance field (WRF) within a wireless digital twin (Jiang et al., 2025). A WRF maps each Tx–Rx configuration in the scene to the channel characteristics of that link. Site-specific channel modeling therefore learns this scene-level WRF from the finite, spatially discrete Tx–Rx measurements available in practice (Böck et al., 2025).

A continuous WRF comes either from physics-based simulation or from data-driven representation learning. Ray tracing computes channels from environmental models, but its fidelity relies on detailed scene geometry and material parameters that are rarely available in practice (Hoydis et al., 2023). Neural and Gaussian radiance field methods instead fit channel observations directly by adapting visual representations (Zhao et al., 2023; Wen et al., 2026a). Their volumetric or Gaussian radiance attributes lack explicit physical surface support, so channel prediction alone does not recover the surfaces producing multipath.

Radio observations can also reveal scene geometry. Existing techniques reconstruct geometry using dedicated radar sensing waveforms, resolve individual multipath components into virtual anchors, or rely on explicit geometric supervision such as ray-traced interaction ranges and path angles (Leitinger et al., 2023; Kadambi et al., 2022; Kung et al., 2025; Moon et al., 2026). These formulations do not learn a single physical scene representation directly from position-labeled channel measurements alone. A unified representation that supports continuous WRF modeling while natively recovering propagation-relevant scene geometry remains missing.

We propose *GaussTwin*, a wireless-native physical scene representation learned directly from channel measurements without geometric or path-level supervision, as illustrated in Figure 1. GaussTwin

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

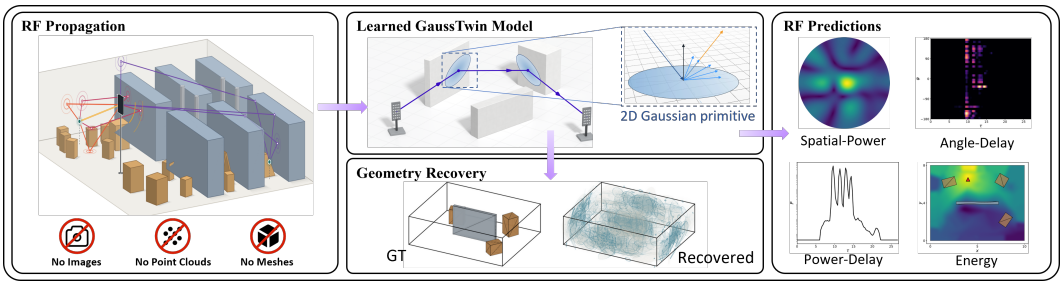


Figure 1: **Illustration of the GaussTwin framework.** From RF propagation between transmitters and receivers, without images, point clouds, or meshes (left), GaussTwin learns 2D Gaussian primitives that reflect and scatter the propagation paths (middle top). The same primitives recover the scene geometry (middle bottom) and predict channel characteristics at unmeasured links (right).

represents the environment as a set of oriented 2D Gaussian surface primitives carrying geometric and propagation attributes. A differentiable propagation forward model maps this primitive set and an arbitrary Tx–Rx configuration to joint angular, delay, and received power responses. Once its geometric and propagation attributes are fitted to the channels observed by many links, GaussTwin can predict channel characteristics at unmeasured Tx–Rx configurations to construct the continuous WRF, while its primitives natively recover propagation-relevant 3D geometry without a dedicated geometry reconstruction loss.

Our contributions are:

- To the best of our knowledge, GaussTwin is the first channel prediction method whose scene representation recovers explicit environment geometry from channel measurements alone. The propagation environment is learned from position-labeled measurements as a physical representation, instead of being supplied externally as a prior.
- We formulate a differentiable physics-based forward model on oriented 2D Gaussian surface primitives. Both geometric and electromagnetic primitive attributes are optimized through gradients of the angular, delay, and power losses.
- Across five simulated indoor scenes and one simulated outdoor city block, GaussTwin demonstrates strong channel prediction and surface recovery capabilities. On public datasets with real measurements, GaussTwin outperforms state-of-the-art baselines on spatial spectrum prediction and achieves comparable performance on received power prediction.

2 Related Work

Learning-based wireless field modeling. Neural radiance fields (NeRF) (Mildenhall et al., 2020) and 3D/2D Gaussian splatting (3DGS/2DGS) (Kerbl et al., 2023; Huang et al., 2024) represent scenes for novel view synthesis and surface reconstruction. Extending these visual representations to radio propagation, NeRF^2 (Zhao et al., 2023) and NeWRF (Lu et al., 2024) learn volumetric neural fields from position-labeled wireless observations. GRaF (Yang et al., 2026) conditions neural fields on neighboring spectra for cross-scene generalization. WRF-GS+ (Wen et al., 2026a), GSRF (Yang et al., 2025), and GSpaRC (Nukapotula et al., 2026) use explicit Gaussian primitives, while nGRF (Umer et al., 2026) treats 3D Gaussians as effective scattering sources and coherently aggregates their contributions into complex channel matrices. SwiftWRF (Liu et al., 2026) models spatial spectra with deformable 2D Gaussians, and BiSplat-WRF (Zhang et al., 2026b) captures inter-primitive dependencies through a bilinear spatial transformer. RxGS (Yang & Srivastava, 2026) learns receiver-conditioned radiance on Gaussian geometry used across receiver configurations to predict at unseen receivers. The learned primitives of these models do not correspond to physical surfaces since the models absorb multipath propagation implicitly into volumetric or Gaussian radiance attributes.

Geometry-informed wireless field modeling. Explicit geometry has also been introduced into wireless modeling through external scene observations or learned priors. RF-3DGS (Zhang et al.,

2026e) and RF-PGS (Zhang et al., 2025) reconstruct geometry optically before learning radio attributes, while CORF-GS (Zhang et al., 2026d) and URF-GS (Wen et al., 2026b) refine optical and RF variables jointly. PC-TGS (Xue et al., 2026) and Photon Splatting (Cao et al., 2026) take measured geometry as input, such as LiDAR point clouds or supplied surfaces. Wireless digital twins and geometry-conditioned predictors likewise use supplied or reconstructed scene structure to predict wireless responses (Jiang et al., 2025; Ai et al., 2026; Vaara et al., 2026; Zhang et al., 2026f; Almasan et al., 2026). Wi-GATr (Hehn et al., 2025) further couples wireless observations with a learned scene prior.

Differentiable propagation and inverse modeling. Differentiable propagation models bring explicit propagation physics into channel learning, but the channel observations in these models do not determine the scene surfaces. WiNeRT (Orekondy et al., 2023) learns interaction models along traced paths on supplied geometry, while differentiable ray tracing (Hoydis et al., 2024) optimizes material and antenna properties on a given scene model. RFIR (Wang et al., 2026) attaches a physically grounded scattering model to Gaussians reconstructed from images and freezes their geometry while learning RF attributes. BiWGS (Zhou et al., 2025) and OctCGS (Zhang et al., 2026c) model bidirectional scattering and multi-order interactions, respectively. Both learn primitive positions from channel measurements, but their primitives are virtual scatterer clusters or scattering regions with no constraint to lie on object surfaces.

Geometry recovery from radio observations. Radio observations can also support explicit scene recovery. Multipath-based RF-SLAM estimates reflecting planes, virtual anchors, or related landmarks from resolved propagation paths (Leitinger et al., 2023; 2024; Kadambi et al., 2022). Back-projection and migration imaging map walls and building layouts from through-wall radar echoes (Song et al., 2018) and from millimeter-wave communication channels (Miyake et al., 2021). Radar-based approaches such as RF4D (Zhang et al., 2026a), RadarSplat (Kung et al., 2025), GeRaF (Lu et al., 2025), and RFDT (Chen et al., 2026) recover scene geometry from dedicated radar measurements. GeRaF, for instance, reconstructs near-range geometry as a signed distance function through physics-based RF volume rendering. Channel2World (Moon et al., 2026) uses a foundation model to predict beam-domain channel responses and sparse reflector points, but its geometry branch requires ray-traced interaction ranges as supervision. Unlike these methods, GaussTwin learns an explicit physical scene representation that simultaneously serves as a part of the propagation model for site-specific channel prediction, and Table C.2 (Appendix C.8) compares representative methods on the four capabilities discussed above.

3 Problem Setting

We consider site-specific prediction of angular, delay, and received power characteristics at unmeasured Tx–Rx configurations in a fixed environment. Wireless observations are available at a finite set of configurations, while scene geometry and propagation properties are unknown.

The Tx and Rx positions of a link form its configuration $\mathcal{Q} = (\mathbf{x}_\mathrm{t}, \mathbf{x}_\mathrm{r})$, and the array geometries and frequencies are known. The available observations are $\mathcal{D} = \{(\mathcal{Q}_\ell, \mathbf{H}_\ell)\}_{\ell=1}^{N}$ from N observed links indexed by ℓ, where $\mathbf{H}_\ell \in \mathbb{C}^{K\times N_\mathrm{r}\times N_\mathrm{t}}$ is the complex channel tensor of link ℓ over subcarriers k, Rx antennas m, and Tx antennas n, K is the number of subcarriers, and $N_\mathrm{r}, N_\mathrm{t}$ are the array sizes. Each observed link provides one realization of $\mathbf{H}$, and its energy and angle–delay power estimate the mean quantities E and $\mathbf{P}$ defined below. The link index ℓ is omitted only when a single link is described. We characterize the channel by its mean energy $E = \mathbb{E}\|\mathbf{H}\|_F^2$, which is proportional to average received power under the measurement convention specified in Appendix C.1. We also describe its mean power distribution over receive directions, transmit directions, and the first D_τ delay bins of the windowed channel using the joint angle–delay power tensor $\mathbf{P} \in \mathbb{R}^{N_\mathrm{r}\times N_\mathrm{t}\times D_\tau}$, whose entry $P[\theta_\mathrm{r}, \theta_\mathrm{t}, \tau]$ is the power in receive-angle bin θ_r, transmit-angle bin θ_t, and delay bin τ. In addition, we obtain the power-delay profile (PDP) by summing $\mathbf{P}$ over both angular dimensions, $p^{\mathrm{PDP}}[\tau] = \sum_{\theta_\mathrm{r},\theta_\mathrm{t}} P[\theta_\mathrm{r}, \theta_\mathrm{t}, \tau]$, and the receive and transmit angular power spectra (APS) $\mathbf{p}^\mathrm{r}$ and $\mathbf{p}^\mathrm{t}$ by summing it over the delay and the other angular dimension.

For the configuration $\mathcal{Q}$ of an unseen link, the model predicts the joint angle–delay power distribution and channel energy via a forward operator:

$$\big(\widehat{\mathbf{P}}(\mathcal{Q}), \widehat{E}(\mathcal{Q})\big) = \mathcal{R}(\widehat{\mathcal{G}}, \mathcal{Q}), \tag{1}$$

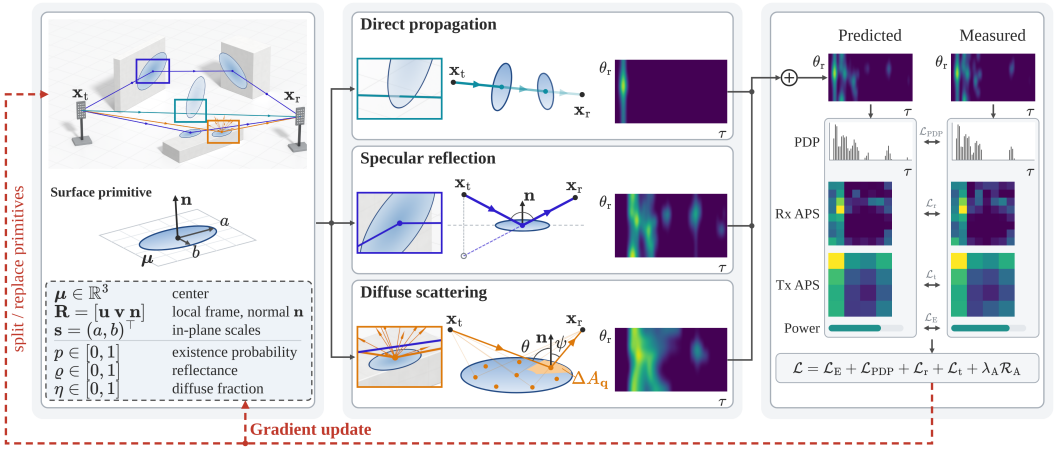


Figure 2: **Method framework.** Left: learnable 2D Gaussian primitives and the observed Tx–Rx links. Middle: the differentiable operator renders direct, specular, and diffuse angle–delay power. Right: the summed prediction is compared with the measurement through its PDP, APS, and received power, and the loss gradient updates the primitives, which are also split and replaced.

where $\widehat{(\cdot)}$ marks predicted quantities and $\widehat{\mathcal{G}}$ is the set of surface primitives learned from the observation set $\mathcal{D}$ (Section 4.1). The PDP and received power are derived from these outputs, whereas on the public datasets with real measurements, spatial spectra and received power are predicted under their respective observation models (Appendices B.3 and B.4). The same primitives $\widehat{\mathcal{G}}$ simultaneously provide explicit propagation-relevant surface geometry.

4 Method

4.1 Explicit surface representation

As illustrated in Figure 2, GaussTwin represents the scene by a sufficiently large number M of learnable surface primitives, $\mathcal{G} = \{g_i\}_{i=1}^{M}$, with $g_i = (\boldsymbol{\mu}_i, \mathbf{R}_i, \mathbf{s}_i, p_i, \varrho_i, \eta_i)$. The primitive follows the planar Gaussian surfel parameterization of 2DGS (Huang et al., 2024), which builds on 3DGS (Kerbl et al., 2023): the center $\boldsymbol{\mu}_i \in \mathbb{R}^3$, local frame $\mathbf{R}_i = [\mathbf{u}_i\ \mathbf{v}_i\ \mathbf{n}_i] \in \mathrm{SO}(3)$, and in-plane scales $\mathbf{s}_i = (a_i, b_i)^\top$ define an oriented 2D Gaussian surface primitive. Here, the normal $\mathbf{n}_i$ determines the specular reflection plane and the directional pattern of diffuse scattering. Each primitive has existence probability $p_i \in [0, 1]$, reflectance $\varrho_i \in [0, 1]$ conditional on its presence, and diffuse fraction $\eta_i \in [0, 1]$.

For a point $\mathbf{q} \in \mathbb{R}^3$ on the plane of primitive i, such as a reflection point or the crossing point of a path, the spatial kernel of the primitive is

$$w_i(\mathbf{q}) = \exp\left[-\tfrac{1}{2}\|\boldsymbol{\xi}_i(\mathbf{q})\|_2^{2\nu}\right], \tag{2}$$

where the local coordinate $\boldsymbol{\xi}_i(\mathbf{q}) \in \mathbb{R}^2$ of $\mathbf{q}$ contains the components of $\mathbf{q} - \boldsymbol{\mu}_i$ along the in-plane axes $\mathbf{u}_i$ and $\mathbf{v}_i$, normalized by the scales a_i and b_i (Appendix C.2). Here, $\|\boldsymbol{\xi}_i(\mathbf{q})\|_2^2 = (\mathbf{q} - \boldsymbol{\mu}_i)^\top \boldsymbol{\Sigma}_i^{+}(\mathbf{q} - \boldsymbol{\mu}_i)$, where $\boldsymbol{\Sigma}_i^{+}$ is the pseudo-inverse of the rank-two 2DGS covariance $\boldsymbol{\Sigma}_i = \mathbf{R}_i\,\mathrm{diag}(a_i^2, b_i^2, 0)\mathbf{R}_i^\top$. At $\nu = 1$, w_i is therefore the Gaussian of a 2DGS primitive, which weights positions on the primitive surface and does not serve as a distribution for sampling points. The exponent ν controls the kernel boundary transition, and a larger ν flattens the interior of the kernel and sharpens its boundary. The exponent is shared by all primitives and is not optimized, and it rises linearly from 2 to 8 during early training. We define the local power transmittance as $t_i(\mathbf{q}) = 1 - p_i w_i(\mathbf{q})$ and the reflection power factor as $\alpha_i = p_i \varrho_i$. At $\mathbf{q}$, specular reflection and diffuse scattering receive fractions $1 - \eta_i$ and η_i of $\alpha_i w_i(\mathbf{q})$, respectively. Since existence and reflectance are separate parameters, a weakly reflecting primitive can still block paths. The same geometry is therefore constrained by both reflection and visibility. GaussTwin uses these primitives to reflect, scatter, and block propagation between antenna arrays through the differentiable operator in Section 4.2.

4.2 DIFFERENTIABLE CHANNEL PREDICTION

Given $\mathcal{G}$ and $\mathcal{Q}$, the propagation model evaluates direct paths, first- and second-order specular reflections, and first-order diffuse scattering, and Appendix C.3 gives the geometric construction and amplitudes of these paths.

Direct propagation. Let $\mathbf{x}_\mathrm{t}$ and $\mathbf{x}_\mathrm{r}$ denote the transmitter and receiver array centers, and let $d_{mn} = \|\mathbf{x}_{\mathrm{r},m} - \mathbf{x}_{\mathrm{t},n}\|_2$ denote the distance between receive element m and transmit element n. During learning, incorrectly placed primitives may intersect the direct path and turn a line-of-sight (LoS) link into a predicted non-line-of-sight (NLoS) one. With binary occlusion, the path would be suppressed without a gradient to update their existence probabilities. We therefore scale the free-space response by a continuous amplitude transmission factor built from the local power transmittances defined in Section 4.1. At frequency f, the direct response between the two elements is

$$[\widehat{\mathbf{H}}_\mathrm{d}(f)]_{mn} = T(\mathbf{x}_\mathrm{t}, \mathbf{x}_\mathrm{r}) \frac{\lambda}{4\pi d_{mn}} e^{-\mathrm{j}2\pi f d_{mn}/c}, \qquad T(\mathbf{a}, \mathbf{b}) = \prod_{(i,\mathbf{q}) \in \mathcal{C}(\mathbf{a},\mathbf{b})} \sqrt{t_i(\mathbf{q})}, \tag{3}$$

where c is the speed of light and $\lambda = c/f$ is the wavelength. For two endpoints $\mathbf{a}$ and $\mathbf{b}$ on a propagation path, $\mathcal{C}(\mathbf{a}, \mathbf{b})$ contains the primitive intersections along the segment between them, and the square root converts each power transmittance to an amplitude transmission. We approximate the transmission factor T by its value along the array-center path and share it across all antenna pairs. The approximation holds when the array aperture is small compared with the path length. When an incorrect obstruction reduces predicted direct power below the observed value, the loss can reduce the obstructing primitive's existence probability.

Specular reflection. For a first-order path p reflecting at $\mathbf{q}$ on primitive i, the path length is $L_p = \|\mathbf{x}_\mathrm{t} - \mathbf{q}\|_2 + \|\mathbf{q} - \mathbf{x}_\mathrm{r}\|_2$, and its channel response at array centers is

$$h_p(f) = T(\mathbf{x}_\mathrm{t}, \mathbf{q}) T(\mathbf{q}, \mathbf{x}_\mathrm{r}) \frac{\lambda}{4\pi L_p} e^{-\mathrm{j}2\pi f L_p/c} \sqrt{(1-\eta_i)\alpha_i w_i(\mathbf{q})} e^{\mathrm{j}\phi_i}, \tag{4}$$

where the two transmission factors cover the incoming and outgoing segments, and $\sqrt{(1-\eta_i)\alpha_i w_i(\mathbf{q})}$ gives the local reflection amplitude. The reflection phase of a surface depends on its material and on its position at the scale of the wavelength, which lies beyond the resolution of the channel measurements. Therefore, we assign one latent reflection phase ϕ_i to each primitive and denote the phase vector by $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_M)$. The phases are not optimized, and we redraw them independently from $\mathrm{Uniform}[0, 2\pi)$ at every training iteration. For each sampled vector, the same phase ϕ_i is used at every interaction with primitive i across links, frequencies, and antenna pairs.

For a path with B bounces on primitives $i_1, \ldots, i_B$, the path length L_p is the sum of the $B+1$ segment lengths. Its response includes $B+1$ segment transmission factors and B local reflection factors, carrying phase $e^{\mathrm{j}\sum_{b=1}^{B} \phi_{i_b}}$. As each bounce introduces attenuation while B-bounce paths grow as $O(M^B)$ across M primitives, we retain first- and second-order reflections ($B \leq 2$). This truncation discards little reflected energy, as shown in Appendix D.4.

Each first- and second-order path kept by this truncation is evaluated at the antenna-pair path lengths (Appendix C.3), and the reflected channel $\widehat{\mathbf{H}}_\mathrm{r}(\boldsymbol{\phi}) = \sum_p \widehat{\mathbf{H}}_p(\boldsymbol{\phi}) \in \mathbb{C}^{K \times N_\mathrm{r} \times N_\mathrm{t}}$ for a sampled phase vector $\boldsymbol{\phi}$ is their complex sum, which keeps the interference between reflection paths.

Diffuse scattering. Each primitive also scatters diffusely over its two-dimensional region Ω_i. We model this response using the Lambertian bidirectional reflectance distribution function (BRDF) (Degli-Esposti et al., 2007), $\eta_i \alpha_i w_i(\mathbf{q})/\pi$. Under unit transmit power, the mean received power density for an antenna pair at the array centers from a surface location $\mathbf{q} \in \Omega_i$ is

$$\sigma_i(f, \mathbf{q}) = \left(\frac{\lambda}{4\pi}\right)^2 \frac{\eta_i \alpha_i w_i(\mathbf{q})}{\pi} \max(\cos\theta \cos\psi, 0) \frac{T(\mathbf{x}_\mathrm{t}, \mathbf{q})^2}{\|\mathbf{x}_\mathrm{t} - \mathbf{q}\|_2^2} \frac{T(\mathbf{q}, \mathbf{x}_\mathrm{r})^2}{\|\mathbf{q} - \mathbf{x}_\mathrm{r}\|_2^2}, \tag{5}$$

where θ and ψ are the angles between the surface normal and the directions from $\mathbf{q}$ toward the transmitter and receiver. Equation 5 is derived in detail in Appendix C.3.1.

For a point $\mathbf{q} \in \Omega_i$, let $\mathbf{G}_i(\mathbf{q}) \in \mathbb{C}^{K \times N_\mathrm{r} \times N_\mathrm{t}}$ be the complex response, per unit area, of the path from the Tx through $\mathbf{q}$ to the Rx. Its power density per antenna pair is σ_i, its delay is $\tau_i(\mathbf{q}) =$

$(\|\mathbf{x}_\mathrm{t} - \mathbf{q}\|_2 + \|\mathbf{q} - \mathbf{x}_\mathrm{r}\|_2)/c$, and its phases across the antenna elements are set by the path lengths from $\mathbf{q}$ to the individual Tx and Rx antennas (Appendix C.3.1). The angle–delay transform $\mathcal{A}$ maps antenna phase differences to angular bins and phase variation across frequencies to delay bins. Contributions from distinct surface locations are modeled as mutually uncorrelated, so their powers are summed:

$$\widehat{\mathbf{P}}_\mathrm{s} = \sum_i \int_{\Omega_i} |\mathcal{A}\mathbf{G}_i(\mathbf{q})|^2 \,\mathrm{d}A \;\approx\; \sum_i \sum_{\mathbf{q}\in\mathcal{X}_i} |\mathcal{A}\mathbf{G}_i(\mathbf{q})|^2 \Delta A_\mathbf{q}. \tag{6}$$

Here, $\mathcal{X}_i \subset \Omega_i$ contains integration points spread uniformly in area over Ω_i, each representing a patch of area $\Delta A_\mathbf{q}$ (Appendix C.7). The diffuse energy $\widehat{E}_\mathrm{s}$ is computed by integrating $\|\mathbf{G}_i(\mathbf{q})\|_F^2$ over the same integration points.

Channel characteristic prediction. The predicted mean angle–delay power is obtained by averaging the specular power over the primitive phases and adding the direct and diffuse powers,

$$\widehat{\mathbf{P}} = |\mathcal{A}\widehat{\mathbf{H}}_\mathrm{d}|^2 + \mathbb{E}_{\boldsymbol{\phi}}\left[|\mathcal{A}\widehat{\mathbf{H}}_\mathrm{r}(\boldsymbol{\phi})|^2\right] + \widehat{\mathbf{P}}_\mathrm{s}. \tag{7}$$

Each reflected path carries the independent uniform phases of its reflecting primitives, and the direct path carries none. In the phase average, the cross terms therefore vanish between the direct and reflected paths and between reflected paths whose reflection counts differ on at least one primitive, which holds for almost all path pairs, so the powers of these paths are summed. The diffuse power $\widehat{\mathbf{P}}_\mathrm{s}$ already sums the powers of uncorrelated surface locations in Eq. 6 and is modeled as uncorrelated with the direct and specular components. Appendix C.4 derives these expectations and their sampling approximation.

Since the window and delay truncation of the angle–delay transform $\mathcal{A}$ change the total power (Appendix C.1), the predicted mean energy needs to be computed from the same three components without the transform,

$$\widehat{E} = \|\widehat{\mathbf{H}}_\mathrm{d}\|_F^2 + \mathbb{E}_{\boldsymbol{\phi}}\left[\|\widehat{\mathbf{H}}_\mathrm{r}(\boldsymbol{\phi})\|_F^2\right] + \widehat{E}_\mathrm{s}. \tag{8}$$

The PDP is obtained by summing $\widehat{\mathbf{P}}$ over both angular dimensions, and $\widehat{E}$ is converted to received power in the same way as the observed energy.

4.3 Model training

We learn the primitives $\mathcal{G}$ from channel prediction errors across observed configurations. Since the propagation model is differentiable, error gradients update both the geometry and the propagation attributes of the primitives. Path delays and departure and arrival directions constrain reflection locations, and the specular reflection law constrains surface normals. For each link, we compare the measured and predicted energy E_ℓ, PDP $\mathbf{p}_\ell^\mathrm{PDP}$, and receive and transmit APS $\mathbf{p}_\ell^\mathrm{r}$ and $\mathbf{p}_\ell^\mathrm{t}$. The learning objective is formulated as

$$\mathcal{L}(\mathcal{G}) = \mathcal{L}_\mathrm{E} + \mathcal{L}_\mathrm{PDP} + \mathcal{L}_\mathrm{r} + \mathcal{L}_\mathrm{t} + \lambda_\mathrm{A}\mathcal{R}_\mathrm{A}(\mathcal{G}). \tag{9}$$

Here, $\mathcal{L}_\mathrm{E}$ measures the received power error in dB, and $\mathcal{L}_\mathrm{PDP}$, $\mathcal{L}_\mathrm{r}$, and $\mathcal{L}_\mathrm{t}$ the normalized mean squared error in dB of the PDP and the receive and transmit APS, respectively, each averaged over the N observed links (Appendix C.5). The total primitive surface area is regularized by $\lambda_\mathrm{A}\mathcal{R}_\mathrm{A}(\mathcal{G})$.

The primitives are initialized from the training channels alone by back-projecting the non-direct residual of each link by delay onto a spatial grid, as in back-projection and migration imaging (Ahmad et al., 2005; Zhuge et al., 2010; Miyake et al., 2021), and placing them at grid points supported by many links, with normals voted for under the specular reflection law. The continuous parameters are then updated by gradient descent on phase-averaged predictions over sampled batches of links and reflection phases. Since gradient descent moves primitives only locally, primitives with low existence probability are periodically relocated to surfaces the initialization missed. They become occluders where measured NLoS links remain unblocked in the prediction, or reflectors where observed delay power remains unexplained. A relocation is kept only if it improves the prediction on a subset of training links. A primitive crossed by both LoS and NLoS links has a single existence probability and cannot let the LoS links pass while blocking the NLoS links, so it is split along the in-plane axis separating them. Appendices C.6 and C.7 detail these steps, and Algorithm 1 summarizes the training sequence.

Table 1: **Channel prediction and surface recovery across simulated scenes on test links.** Precision and coverage use a tolerance of 15 cm indoors and 1.5 m for the city block, and coverage takes the surface interactions of the ground-truth paths as reference. Values report mean ± standard deviation, and the indoor mean rows average the five indoor scenes.

	Channel			Geometry		
Scene	Power MAE (dB) ↓	PDP NMSE median (dB) ↓	Joint cosine ↑	Chamfer (cm) ↓	Precision ↑	Coverage ↑
Parallel-wall	0.79 ± 0.15	-16.42 ± 0.41	0.928 ± 0.016	11.47 ± 0.60	71.60 ± 5.09	82.07 ± 1.52
Angled-wall	0.58 ± 0.02	-17.13 ± 0.36	0.959 ± 0.003	10.10 ± 0.27	73.86 ± 1.82	87.08 ± 2.78
T-shaped	0.58 ± 0.06	-18.63 ± 0.23	0.964 ± 0.010	17.96 ± 0.67	59.77 ± 2.30	90.36 ± 1.22
Cluttered	1.36 ± 0.03	-12.14 ± 0.25	0.845 ± 0.002	15.25 ± 1.66	71.58 ± 2.69	67.43 ± 2.61
Six-room	2.80 ± 0.05	-8.24 ± 0.69	0.722 ± 0.014	19.08 ± 2.35	67.77 ± 2.03	73.08 ± 2.64
Indoor mean						
MLP	3.14	-1.58	0.109	–	–	–
BiWGS	3.97	-3.49	0.520	–	–	–
RF-3DGS	2.64	-6.01	0.251	–	–	–
GaussTwin	1.22	-14.51	0.883	14.77	68.92	80.00
City block	2.04 ± 0.01	-0.57 ± 0.02	0.523 ± 0.002	87.79 ± 1.77	80.60 ± 0.70	90.53 ± 0.32

5 Experimental Results

We evaluate GaussTwin in two settings: controlled simulations on five indoor scenes and an outdoor city block, with paired ground truth for wireless propagation and 3D geometry, and two public datasets with real measurements for spatial spectrum and received power prediction. Appendices A.1 and B detail all environment layouts, array configurations, link partitions, and formal metric definitions.

5.1 Joint channel prediction and surface recovery

We test joint channel prediction and surface recovery on five Sionna RT indoor scenes (Hoydis et al., 2023) and the DeepMIMO Houston city block (Alkhateeb, 2019). Each indoor scene provides $14{,}400$ training links with 28 GHz channels over a 1 GHz band, 4×4 Tx arrays, and 8×8 Rx arrays, and $4{,}000$ test links with an unseen Tx position, Rx position, or both. The city block provides $51{,}535$ training links from three base stations and $8{,}602$ test links. Channel prediction is evaluated by the received power mean absolute error (MAE) and the median PDP normalized mean squared error (NMSE), both in dB, and by the joint cosine similarity between the predicted and observed $\mathbf{P}$, averaged over the test links. Within a distance tolerance, coverage is the fraction of propagation-relevant surface points, the surface interactions of the ground-truth paths, reached by the recovered primitive ellipses, and precision is the fraction of points sampled uniformly on these ellipses that lie on the ground-truth meshes. Chamfer-L_1 averages the two mean distances (Appendix A). To the best of our knowledge, no existing method both predicts channel characteristics and recovers geometry. With changing Tx and Rx positions, existing methods such as BiWGS (Zhou et al., 2025) predict only spatial spectra and channel power gains. Therefore, we adapt BiWGS and RF-3DGS (Zhang et al., 2026e) to this task (Appendix A.2) and also compare a multilayer perceptron (MLP) as a baseline, all trained on the same dataset with the same losses and validation.

Table 1 shows joint channel prediction and surface recovery across simulated environments. GaussTwin recovers most propagation-relevant surfaces, with coverage of up to 90% indoors and 91% outdoors under their respective tolerances. Its median PDP NMSE is -0.57 dB on the city block, compared with -8.24 to -18.63 dB indoors. The city block is observed from three base station sites across an extensive area with lower spatial measurement density, and paths outside the propagation model carry 13.3% of its received energy on average (Appendix D.4). Averaged over the indoor scenes, GaussTwin outperforms all baselines on every channel metric with a median PDP NMSE more than 8 dB lower, even though RF-3DGS starts from the ground-truth meshes.

Figure 3 visualizes recovered scene geometry across simulated indoor environments and the city block. The horizontal cross-sections show that the primitives trace outer walls, major interior parti-

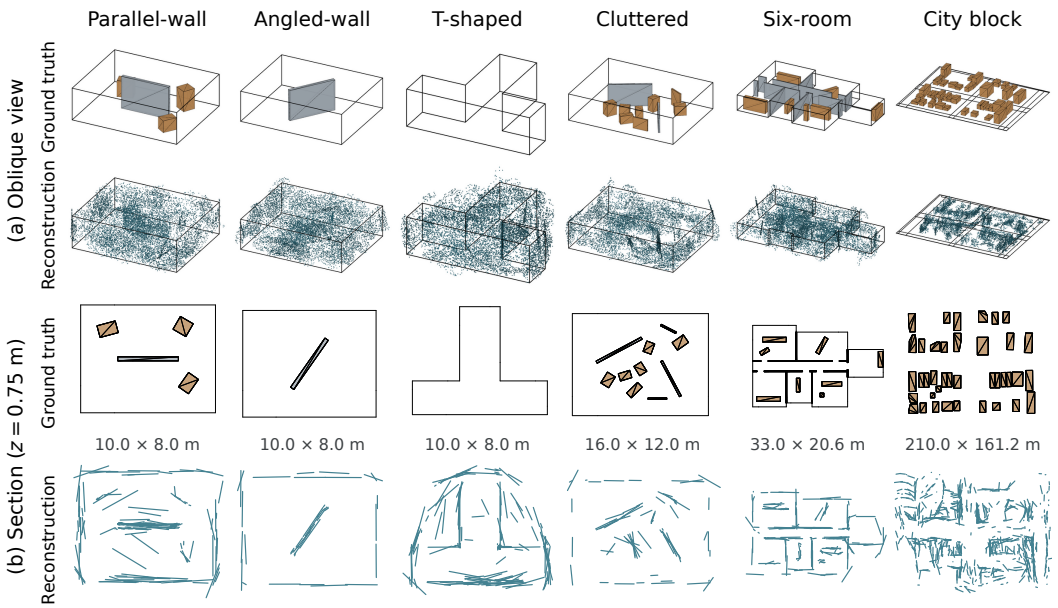


Figure 3: **Surface geometry recovered from wireless observations.** Qualitative comparison of ground-truth meshes and reconstructed point clouds across simulated scenes, shown in oblique 3D views and horizontal cross-sections at $z = 0.75$ m.

tions, and the buildings of the city block. As the layout becomes more complex, more primitives or higher-order propagation may be needed to represent it (Appendices D.2 and D.3), and the surfaces that remain missing lie in regions less constrained by the observed propagation paths. On the city block, for example, most building facades are reached by few or no ground-truth paths, as shown in Figure B.1a.

On the six-room scene, primitives kept as their initialized geometry reach a median PDP NMSE of only -1.48 dB and a Chamfer-L_1 distance of 18.5 cm, whereas the learned geometry lowers them to -8.74 dB and 16.9 cm, as shown in Table D.1 (Appendix D.2). Random, proposed, and ground-truth initialization provide increasingly accurate initial geometry, and the median PDP NMSE after training improves accordingly from -4.86 to -8.74 and -9.00 dB, indicating that more accurate geometry leads to more accurate channel prediction. Channel prediction improves only when the geometry is jointly optimized, and the process of channel prediction optimization can in turn refine the geometry as well, so channel prediction and surface recovery reinforce each other throughout training.

Appendix D further reports results by type of unseen link and with less training data.

5.2 Prediction on real-world measurements

We evaluate spatial spectrum and received power prediction on two public datasets with real measurements (Zhao et al., 2023) at unseen transmitter locations. All baselines are retrained on the same data splits and evaluated with the same metrics. For RFID, the model predicts 90×360 spatial spectra and is trained and evaluated with the structural similarity index (SSIM) (Wang et al., 2004). As shown in Table 2, GaussTwin attains the highest median SSIM, 0.935, compared to 0.912 for the strongest baseline, BiSplat-WRF, and also the highest mean SSIM (Appendix B.3). As visualized in Figure 4, GaussTwin produces less noise and avoids the abrupt seam artifacts visible in several baseline predictions. Its directional details are more accurate, and as shown in Figure B.2, it shifts the upper part of the per-transmitter SSIM distribution to the right.

For BLE, we train the model to predict received signal strength indicator (RSSI) values by minimizing absolute error in dB, a task that GSpaRC does not support. As shown in Table 2 and Figure B.3, GaussTwin achieves a median absolute error of 2.76 dB and an MAE of 3.43 dB across $8{,}400$ test

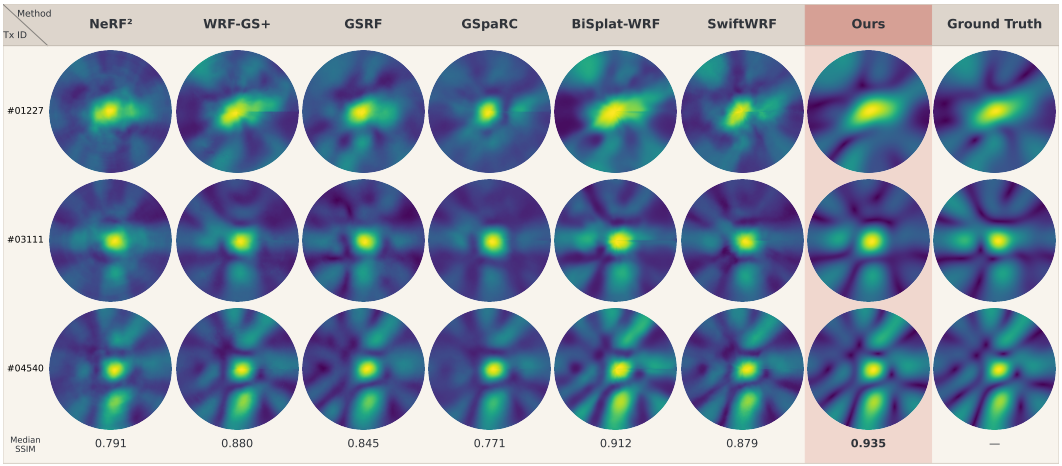


Figure 4: **Predicted spatial spectra on the NeRF^2 RFID test split.** Qualitative comparison with six baselines and ground truth across three test transmitter positions. Each panel visualizes normalized arrival power over azimuth and elevation, with overall test-split median SSIM reported at the bottom.

Table 2: **Quantitative comparison on public datasets with real measurements.** All methods are evaluated on the 1,216 RFID and 8,400 BLE test links of the NeRF^2 datasets.

Benchmark, metric	NeRF^2	WRF-GS+	GSRF	GSpaRC	BiSplat-WRF	SwiftWRF	GaussTwin
RFID, median SSIM ↑	0.791	0.880	0.845	0.771	<u>0.912</u>	0.879	**0.935**
RFID, mean SSIM ↑	0.767	0.839	0.809	0.747	<u>0.856</u>	0.836	**0.875**
BLE, median abs. error (dB) ↓	2.87	2.79	3.07	–	2.93	**2.73**	<u>2.76</u>
BLE, MAE (dB) ↓	3.45	3.48	3.89	–	3.70	**3.29**	<u>3.43</u>

links. Both rank second after SwiftWRF, at 2.73 and 3.29 dB, and are lower than those of the remaining evaluated baselines (Appendix B.4).

6 Limitations

GaussTwin recovers surfaces to the extent that the observed propagation paths constrain them. This extent depends on the paths reaching each surface and on the angular and delay resolution of the arrays. The formulation assumes synchronized multi-antenna measurements with known transceiver poses in a static scene. The propagation model truncates reflections at second order, models diffuse scattering at first order, omits diffraction, and evaluates interaction weights and blockage on the array-center path. Planar primitives represent extended surfaces compactly, whereas curved or intricate structures require more primitives. Finally, public datasets with real measurements rarely provide channel measurements together with an accurate geometric model of the scene, so surface recovery is evaluated in simulation.

7 Conclusion

We proposed GaussTwin, a wireless channel model that predicts channel characteristics from oriented 2D Gaussian surface primitives through a differentiable propagation operator and learns those primitives from channel measurements without geometric supervision. In simulation, the same primitives predict channels at unseen links and trace outer walls, major interior partitions, and building blocks. On public datasets with real measurements, GaussTwin outperforms state-of-the-art baselines on spatial spectrum prediction and ranks second in received power prediction. Future work will focus on continuously updating the learned primitives from new measurements as the scene changes, so that a wireless digital twin stays consistent with its environment.

AI USE STATEMENT

In this work, we used generative AI tools to aid and polish the writing, provide feedback on the research methodology and experiments, and assist in implementing methods. We have not used generative AI tools to generate synthetic datasets, help develop theoretical models or conceptual frameworks, propose or refine hypotheses, assist with translation, clean and reformat datasets, or interpret results, and formulating mathematical claims, providing critical ingredients for proving mathematical claims, assisting in the writing of proofs, and supporting qualitative and thematic data analysis are not applicable to this work. We have reviewed all AI-assisted work, and the authors verified the implemented methods and the experiments. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

Section 4 specifies the surface representation, prediction operator and learning procedure, and Algorithm 1 in Appendix C.7 lists the full training sequence. Appendix C gives the observation features, propagation equations, phase-marginalization derivation, optimization terms and numerical settings, and Appendices A and B describe the evaluation protocol and the datasets.

REFERENCES

Fauzia Ahmad, Moeness G. Amin, and S. A. Kassam. Synthetic aperture beamformer for imaging through a dielectric wall. *IEEE Transactions on Aerospace and Electronic Systems*, 41(1):271–283, 2005.

Junjie Ai, Shurui Xu, Yanqing Ren, Zhuoyu Liu, Weicong Chen, Wankai Tang, Xiao Li, Chao-Kai Wen, and Shi Jin. Propagation-consistent wireless environment digital twin construction under sparse measurements. *arXiv preprint arXiv:2605.22361*, 2026.

Ahmed Alkhateeb. DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications. In *Information Theory and Applications Workshop (ITA)*, pp. 1–8, 2019.

Paul Almasan, Stefanos Bakirtzis, José Suárez-Varela, and Andra Lutu. From waves to graphs: A ray-tracing-inspired neural radio propagation model. In *IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2026.

Benedikt Böck, Andreas Oeldemann, Timo Mayer, Francesco Rossetto, and Wolfgang Utschick. Physics-informed generative modeling of wireless channels. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 4602–4626, 2025.

Ge Cao, Gabriele Gradoni, and Zhen Peng. Photon splatting: A physics-guided neural surrogate for real-time channel modeling in quasi-static wireless environments. *IEEE Journal of Selected Topics in Electromagnetics, Antennas and Propagation*, 2:386–400, 2026.

Xingyu Chen, Xinyu Zhang, Kai Zheng, Xinmin Fang, Tzu-Mao Li, Chris Xiaoxuan Lu, and Zhengxiong Li. Physically accurate differentiable inverse rendering for radio frequency digital twin. In *ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2026.

Vittorio Degli-Esposti, Franco Fuschini, Enrico M. Vitucci, and Gabriele Falciasecca. Measurement and modelling of scattering from buildings. *IEEE Transactions on Antennas and Propagation*, 55 (1):143–153, 2007.

Thomas Hehn, Markus Peschl, Tribhuvanesh Orekondy, Arash Behboodi, and Johann Brehmer. Differentiable and learnable wireless simulation with geometric transformers. In *International Conference on Learning Representations (ICLR)*, 2025.

Jakob Hoydis, Fayçal Aït Aoudia, Sebastian Cammerer, Merlin Nimier-David, Nikolaus Binder, Guillermo Marcus, and Alexander Keller. Sionna RT: Differentiable ray tracing for radio propagation modeling. In *IEEE Globecom Workshops (GC Wkshps)*, pp. 317–321, 2023.

Jakob Hoydis, Fayçal Aït Aoudia, Sebastian Cammerer, Florian Euchner, Merlin Nimier-David, Stephan ten Brink, and Alexander Keller. Learning radio environments by differentiable ray tracing. *IEEE Transactions on Machine Learning in Communications and Networking*, 2:1527–1539, 2024.

Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2D Gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 Conference Papers*, pp. 32:1–32:11, 2024.

Shuaifeng Jiang, Qi Qu, Xiaqing Pan, Abhishek K. Agrawal, Richard Newcombe, and Ahmed Alkhateeb. Learnable wireless digital twins: Reconstructing electromagnetic field with neural representations. *IEEE Open Journal of the Communications Society*, 6:1568–1590, 2025.

Shreya Kadambi, Arash Behboodi, Joseph B. Soriaga, Max Welling, Roohollah Amiri, Srinivas Yerramalli, and Taesang Yoo. Neural RF SLAM for unsupervised positioning and mapping with channel state information. In *IEEE International Conference on Communications (ICC)*, pp. 3238–3244, 2022.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D Gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4):139:1–139:14, 2023.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

Pou-Chun Kung, Skanda Harisha, Ram Vasudevan, Aline Eid, and Katherine A. Skinner. RadarSplat: Radar Gaussian splatting for high-fidelity data synthesis and 3D reconstruction of autonomous driving scenes. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 27596–27606, 2025.

Erik Leitinger, Alexander Venus, Bryan Teague, and Florian Meyer. Data fusion for multipath-based SLAM: Combining information from multiple propagation paths. *IEEE Transactions on Signal Processing*, 71:4011–4028, 2023.

Erik Leitinger, Lukas Wielandner, Alexander Venus, Mingchao Liang, Florian Meyer, and Klaus Witrisal. Multipath-based SLAM with cooperation and map fusion in MIMO systems. In *Asilomar Conference on Signals, Systems, and Computers*, pp. 1316–1322, 2024.

Mufan Liu, Cixiao Zhang, Qi Yang, Yujie Cao, Yiling Xu, Yin Xu, Shu Sun, Mingzeng Dai, and Yunfeng Guan. Deformable 2D Gaussian splatting for efficient wireless radiance field rendering. *IEEE Transactions on Visualization and Computer Graphics*, 32(7):6363–6378, 2026.

Haofan Lu, Christopher Vattheuer, Baharan Mirzasoleiman, and Omid Abari. NeWRF: A deep learning framework for wireless radiation field reconstruction and channel prediction. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 33147–33159, 2024.

Jiachen Lu, Hailan Shanbhag, and Haitham Al Hassanieh. GeRaF: Neural geometry reconstruction from radio frequency signals. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, pp. 94200–94230, 2025.

Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3D reconstruction in function space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4455–4465, 2019.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *European Conference on Computer Vision (ECCV)*, pp. 405–421, 2020.

Yuto Miyake, Minseok Kim, Takeshi Tasaki, Satoshi Yamakawa, and Jun-ichi Takada. Environment mapping technique using millimeter-wave radio systems. In *International Symposium on Antennas and Propagation (ISAP)*, pp. 833–834, 2021.

Hyung-Joo Moon, Joonkyu Jang, Kwang Soon Kim, Seong-Lyun Kim, Robert W. Heath, Jr., and Chan-Byoung Chae. Channel2World: A wireless foundation model for RF environment representation. *arXiv preprint arXiv:2608.17544*, 2026.

Bhavya Sai Nukapotula, Rishabh Tripathi, Seth Pregler, Dileep Kalathil, Srinivas Shakkottai, and Theodore S. Rappaport. GSpaRC: Gaussian splatting for real-time reconstruction of RF channels. *arXiv preprint arXiv:2511.22793*, 2026.

Tribhuvanesh Orekondy, Pratik Kumar, Shreya Kadambi, Hao Ye, Joseph Soriaga, and Arash Behboodi. WiNeRT: Towards neural ray tracing for wireless channel modelling and differentiable simulations. In *International Conference on Learning Representations (ICLR)*, 2023.

Yongping Song, Jun Hu, Ning Chu, Tian Jin, Jianwen Zhang, and Zhimin Zhou. Building layout reconstruction in concealed human target sensing via UWB MIMO through-wall imaging radar. *IEEE Geoscience and Remote Sensing Letters*, 15(8):1199–1203, 2018.

Muhammad Umer, Muhammad Ahmed Mohsin, Ahsan Bilal, and John M. Cioffi. Neural Gaussian radio fields for channel estimation. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 12138–12149, 2026.

Niklas Vaara, Lam Huynh, Pekka Sangi, Miguel Bordallo López, and Janne Heikkilä. Differentiable ray tracing with Gaussians for unified radio propagation simulation and view synthesis. *arXiv preprint arXiv:2605.07781*, 2026.

Fuhai Wang, Zihan Jin, Lehang Wang, Xuehui Dong, Tiebin Mi, Robert Caiming Qiu, and Zenan Ling. Radio-frequency inverse rendering for wireless environment modeling. *arXiv preprint arXiv:2604.07086*, 2026.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004.

Chaozheng Wen, Jingwen Tong, Yingdong Hu, Zehong Lin, and Jun Zhang. Neural representation for wireless radiation field reconstruction: A 3D Gaussian splatting approach. *IEEE Transactions on Wireless Communications*, 25:7490–7504, 2026a.

Chaozheng Wen, Jingwen Tong, Zehong Lin, Chenghong Bian, and Jun Zhang. Bridging visual and wireless sensing via a unified radiation field for 3D radio map construction. *npj Wireless Technology*, 2:59, 2026b.

Ye Xue, Yiheng Wang, Xinhua Shao, Qi Yan, Shutao Zhang, and Tsung-Hui Chang. Point-cloud-assistant localized statistical channel prediction by tangent Gaussian splatting. *IEEE Transactions on Wireless Communications*, 25:17816–17830, 2026.

Kang Yang and Mani Srivastava. RxGS: Receiver-generalizable 3D Gaussian splatting for radio-frequency data synthesis. *arXiv preprint arXiv:2605.24290*, 2026.

Kang Yang, Gaofeng Dong, Sijie Ji, Wan Du, and Mani Srivastava. GSRF: Complex-valued 3D Gaussian splatting for efficient radio-frequency data synthesis. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, pp. 63076–63105, 2025.

Kang Yang, Yuning Chen, and Wan Du. Generalizable radio-frequency radiance fields for spatial spectrum synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12533–12543, 2026.

Jiarui Zhang, Zhihao Li, Chong Wang, and Bihan Wen. RF4D: Neural radar fields for novel view synthesis in outdoor dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15387–15397, 2026a.

Jinghan Zhang, Xitao Gong, Qi Wang, Richard A. Stirling-Gallacher, and Giuseppe Caire. Planar Gaussian splatting with bilinear spatial transformer for wireless radiance field reconstruction. In *IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2026b.

Jinghan Zhang, Xitao Gong, Qi Wang, Richard A. Stirling-Gallacher, and Giuseppe Caire. OctCGS: Octree-contextual Gaussian splatting with explicit multi-order propagation modeling for channel knowledge map construction. *arXiv preprint arXiv:2605.22961*, 2026c.

Jinya Zhang, Jiajia Guo, Chao-Kai Wen, and Shi Jin. CORF-GS: Real-time wireless radiance field reconstruction via coupled optical-RF Gaussian splatting. *arXiv preprint arXiv:2607.25569*, 2026d.

Lihao Zhang, Zongtan Li, and Haijian Sun. RF-PGS: Fully-structured spatial wireless channel representation with planar Gaussian splatting. *arXiv preprint arXiv:2508.16849*, 2025.

Lihao Zhang, Haijian Sun, Samuel Berweger, Camillo Gentile, and Rose Qingyang Hu. RF-3DGS: Wireless channel modeling with radio radiance field and 3D Gaussian splatting. *IEEE Transactions on Wireless Communications*, 25:10419–10433, 2026e.

Yuru Zhang, Ming Zhao, Qiang Liu, Ahmed Alkhateeb, Abhishek K. Agrawal, and Qi Qu. RadTwin: Generalizable wireless digital twin for dynamic environments. In *International Conference on Computer Communications and Networks (ICCCN)*, pp. 1–9, 2026f.

Xiaopeng Zhao, Zhenlin An, Qingrui Pan, and Lei Yang. NeRF^2: Neural radio-frequency radiance fields. In *ACM International Conference on Mobile Computing and Networking (MobiCom)*, pp. 27:1–27:15, 2023.

Juncong Zhou, Chao Hu, Guanlin Wu, Zixiang Ren, Han Hu, Juyong Zhang, Rui Zhang, and Jie Xu. 6D channel knowledge map construction via bidirectional wireless Gaussian splatting. *arXiv preprint arXiv:2510.26166*, 2025.

Xiaodong Zhuge, Alexander G. Yarovoy, Timofey Savelyev, and Leo Ligthart. Modified Kirchhoff migration for UWB MIMO array-based radar imaging. *IEEE Transactions on Geoscience and Remote Sensing*, 48(6):2692–2703, 2010.

A Evaluation Protocol

A.1 Evaluation metrics

Channel prediction metrics. NMSE and cosine similarity compare an observed quantity $\mathbf{y}$ with its prediction $\widehat{\mathbf{y}}$ on a linear scale and without normalization. Here $\mathbf{y}$ is the PDP for NMSE and the vectorized angle–delay power for cosine similarity, both obtained with the angle–delay transform $\mathcal{A}$ (Appendix C.1). For each test link, the normalized mean squared error is

$$\mathrm{NMSE} = \frac{\|\widehat{\mathbf{y}} - \mathbf{y}\|_2^2}{\|\mathbf{y}\|_2^2}. \tag{10}$$

With the subscript ℓ marking test link $\ell = 1, \ldots, L$, we report the median PDP NMSE in dB, the mean joint cosine similarity, and the mean absolute received power error,

$$\begin{aligned} \mathrm{NMSE}_{\mathrm{med,dB}} &= 10\log_{10}\big(\mathrm{median}\{\mathrm{NMSE}_1, \ldots, \mathrm{NMSE}_L\}\big), \\ \mathrm{Cosine} &= \frac{1}{L}\sum_{\ell=1}^{L} \frac{\widehat{\mathbf{y}}_\ell^{\mathsf{T}}\mathbf{y}_\ell}{\|\widehat{\mathbf{y}}_\ell\|_2\|\mathbf{y}_\ell\|_2}, \\ \mathrm{MAE}_{\mathrm{dB}} &= \frac{1}{L}\sum_{\ell=1}^{L} \left| 10\log_{10}\frac{\widehat{E}_\ell}{E_\ell} \right|, \end{aligned} \tag{11}$$

where $E_\ell = \|\mathbf{H}_\ell\|_F^2$ and $\widehat{E}_\ell$ of Eq. 8 are the observed and predicted energies, summed over all subcarriers and antenna pairs before the window and delay truncation of $\mathcal{A}$.

Surface recovery. Recovered geometry is compared with the ray-traced scene through two point sets. Ground-truth interaction points mark where training-link paths strike a surface. Each is assigned to the surface containing it, and one point is kept per 10 cm cell of that surface, without which the evaluation would be dominated by regions struck by many training paths. Sampled points lie uniformly on the reconstruction, the union of all primitive ellipses as filled planar regions regardless of existence probability, about 20,000 of them allocated by area with at least one per primitive.

For a point $\mathbf{x}$, let $d_{\mathrm{rec}}(\mathbf{x})$ be its Euclidean distance to the nearest point of the reconstruction and $d_{\mathrm{mesh}}(\mathbf{x})$ that to the nearest point of the scene mesh. Coverage is the fraction of interaction points with $d_{\mathrm{rec}}(\mathbf{x}) \leq \delta$ and precision the fraction of sampled points with $d_{\mathrm{mesh}}(\mathbf{x}) \leq \delta$. Coverage is high when the reconstruction reaches the surfaces the training paths touch, and precision when the reconstruction lies on scene geometry. We use $\delta = 0.15$ m indoors and $\delta = 1.5$ m for the city block. With N_{I} interaction points and N_{S} sampled points, the Chamfer-L_1 distance is

$$\mathrm{CD}_{L_1} = \frac{1}{2}\left[\frac{1}{N_{\mathrm{I}}}\sum_{\text{interaction } \mathbf{x}} d_{\mathrm{rec}}(\mathbf{x}) + \frac{1}{N_{\mathrm{S}}}\sum_{\text{sampled } \mathbf{x}} d_{\mathrm{mesh}}(\mathbf{x})\right]. \tag{12}$$

Following Mescheder et al. (2019), the subscript L_1 refers to averaging unsquared distances. Coverage and precision are reported in percent and CD_{L_1} in centimeters.

A.2 Adapted baselines

To the best of our knowledge, no existing method both predicts channel characteristics and recovers scene geometry. When the Tx and Rx positions both vary, existing methods such as BiWGS (Zhou et al., 2025) predict only the spatial spectrum and the channel power gain. We therefore adapt BiWGS and RF-3DGS (Zhang et al., 2026e) to predict the channel characteristics of Section 3, and both are trained on the same dataset with the same losses as GaussTwin. We modify their input and output interfaces but keep the model settings of the original papers. These results evaluate the adapted models on this task and do not represent the performance of these methods on their original tasks.

Since ground-truth scene models are available, we directly sample the initial point cloud of RF-3DGS from the scene mesh surfaces in place of its visual initialization from photographs. We also add a Tx-conditioned MLP to its radio attributes to support changes of the Tx position, which are not available in the original RF-3DGS method.

B Datasets

We use four distinct datasets: five indoor environments simulated with Sionna RT (Hoydis et al., 2023), the DeepMIMO Houston city block (Alkhateeb, 2019), and two public datasets with real measurements from NeRF^2 (Zhao et al., 2023). Table B.1 lists the radio configuration and link partition for each dataset.

B.1 Sionna indoor simulations

We use Sionna RT 2.0.1 (Hoydis et al., 2023) to generate the channel tensors without artificial noise or phase distortion, with isotropic elements in both the Tx and Rx arrays. The tensor dimensions index subcarriers, Rx antenna elements, and Tx antenna elements in that order. Each scene contains 20,800 Tx–Rx links, partitioned as shown in Table B.1.

Table B.1: **Dataset configurations.**

Dataset	Carrier	Bandwidth / subcarriers	Tx/Rx	Links (train / validation / test)
Sionna RT	28 GHz	1 GHz / 128	4×4 / 8×8, 0.5λ	14,400 / 2,400 / 4,000
DeepMIMO	28 GHz	1 GHz / 1,024	8×8 / 2×2, 0.5λ	51,535 / 8,594 / 8,602
NeRF^2 RFID	915 MHz	Not released	1×1 / 4×4, 0.16 m	4,907 / none / 1,216
NeRF^2 BLE	2.4 GHz	Not reported	1×1 / 1×1	33,600 / none / 8,400

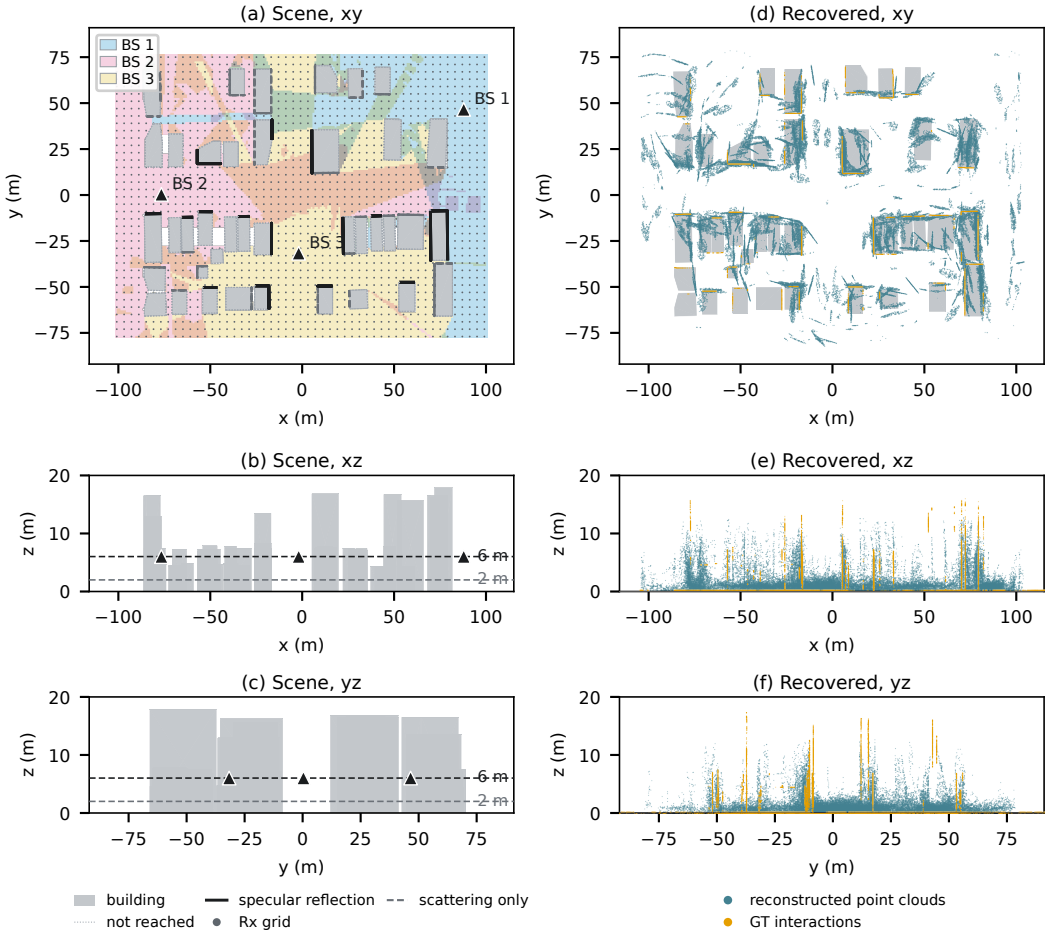


Figure B.1: **City block and recovered geometry.** (a) Plan view of 24,921 Rx positions and three BS Tx arrays. Colors indicate coverage and overlaps, and line styles mark how the ground-truth paths reach each building facade. (b), (c) Elevations along x and along y. Dashed lines mark the Rx height of 2 m and the Tx height of 6 m. (d)–(f) Projections of the reconstructed point cloud (blue) and the ground-truth interaction points (orange) onto xy, xz, and yz.

We simulate five indoor environments, with sizes given as $x \times y \times z$: *parallel-wall* ($10 \times 8 \times 3$ m, with an interior wall parallel to the outer walls and three objects), *angled-wall* ($10 \times 8 \times 3$ m, with an interior wall rotated $55°$), *T-shaped* ($10 \times 8 \times 3$ m), *cluttered* ($16 \times 12 \times 4$ m), and *six-room* ($33 \times 20.6 \times 4.2$ m, an asymmetric floor plan). Room boundaries and interior walls use ITU-R P.2040 concrete (0.3 m thickness), while interior clutter uses the wood model.

In every scene, Sionna RT models LoS paths, specular reflections, and diffuse scattering up to second order. For each Tx position, the solver traces 10^6 ray samples and retains at most 4×10^6 paths. The reference points for geometry are the interactions of the training links retraced at the array centers. They cover 191 to 2,324 m^2 of surface across the five scenes and are used for evaluation only.

B.2 DeepMIMO Houston city block

For outdoor evaluation, we take the Houston city block from DeepMIMO v4 (Alkhateeb, 2019). Following the DeepMIMO convention, the three base stations (BSs) provide the Tx arrays as vertical panels with an $8°$ downtilt, and the ground user-equipment (UE) grid provides the Rx positions. We split the 68,731 links by Rx position, with counts given in Table B.1. The reference points come from the interactions recorded with the export and reach 1.1×10^5 m^2 of ground and facade surface.

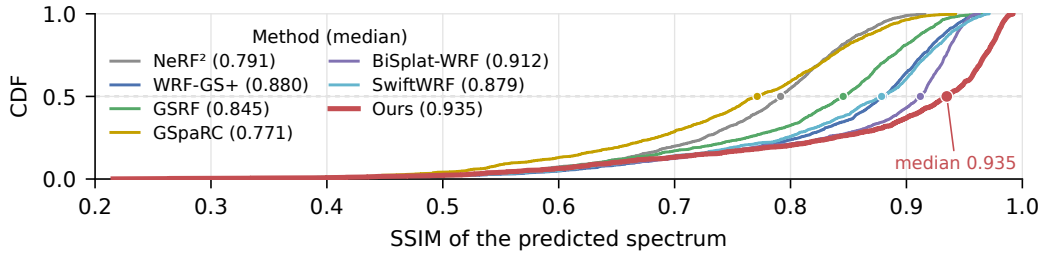


Figure B.2: **Per-Tx SSIM on the NeRF^2 RFID test split.** Empirical cumulative distribution over the 1,216 test Tx positions, with each median in the legend and on the dashed 0.5 level. GaussTwin shifts the upper part of the distribution to the right.

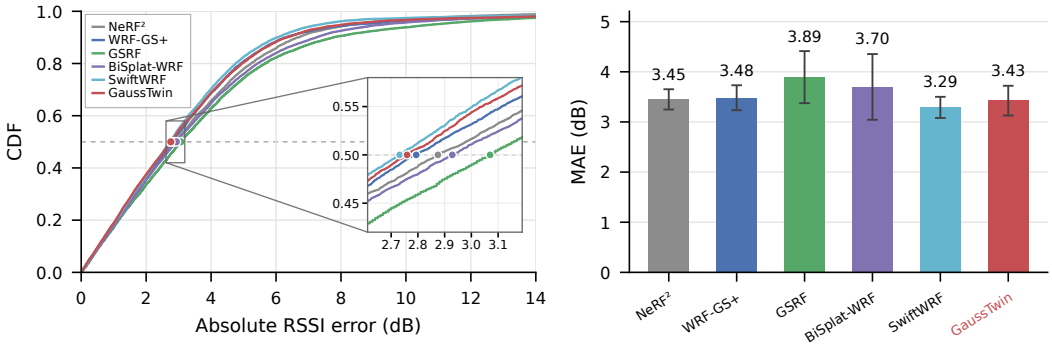


Figure B.3: **Scalar RSSI prediction on the BLE benchmark.** Left: empirical CDF of absolute error across the 8,400 test links, with inset magnifying medians around the 0.5 level. Right: mean absolute error across gateways, with error bars indicating across-gateway standard deviation.

Figure B.1 places this scene next to the recovered geometry. Panels (a)–(c) show the Rx grid, the three base stations, and where their coverage overlaps. Panels (d)–(f) project the recovered point cloud onto the ground plane and the two vertical planes, together with the reference interaction points.

B.3 Public RFID spatial spectrum benchmark

The NeRF^2 RFID benchmark records a fixed gateway Rx while the Tx occupies 6,123 positions (Zhao et al., 2023), split into training and test sets as given in Table B.1. The positions fill a $1.62 \times 1.35 \times 0.89$ m volume and lie 4.54 to 6.11 m from the gateway.

The RFID *spatial spectrum* is the normalized directional response of the gateway Rx array over arrival direction (Zhao et al., 2023; Wen et al., 2026a; Yang et al., 2026).

For Tx position $\mathbf{x}_\mathrm{t}$, let $\widehat{\mathbf{h}}(\mathbf{x}_\mathrm{t}) \in \mathbb{C}^{N_\mathrm{r}}$ be the predicted complex response across the $N_\mathrm{r} = 16$ Rx antenna elements and let $\mathbf{a}_\mathrm{Rx}(\theta_\mathrm{az}, \theta_\mathrm{el})$ be the Rx array steering vector. The predicted spatial spectrum is

$$\widehat{P}_\mathrm{spatial}(\mathbf{x}_\mathrm{t}, \theta_\mathrm{az}, \theta_\mathrm{el}) = \mathrm{scale}_{[0,1]}\left(\frac{1}{N_\mathrm{r}}\left|\mathbf{a}_\mathrm{Rx}^\mathsf{H}(\theta_\mathrm{az}, \theta_\mathrm{el})\widehat{\mathbf{h}}(\mathbf{x}_\mathrm{t})\right|\right), \tag{13}$$

where $\mathrm{scale}_{[0,1]}$ linearly maps the minimum and maximum within each spectrum to zero and one. The response is evaluated on 90 elevation angles from 1° to 90° and 360 azimuth angles from 0° to 359°, which index the rows and the columns of the 90×360 spectrum images supplied by the benchmark.

Before scoring, each predicted and measured spectrum is divided by its own maximum, with an all-zero spectrum left unchanged. We report the per-Tx median, mean, and cumulative distribution of SSIM, as shown in Figure B.2. Figure 4 provides qualitative comparisons of predicted spatial spectra with measured ground truth across all baselines.

B.4 PUBLIC BLE RECEIVED-STRENGTH BENCHMARK

To evaluate scalar received power prediction without angular or delay resolution, we use the public Bluetooth Low Energy (BLE) benchmark from NeRF^2 (Zhao et al., 2023). The measurement campaign covers a large indoor environment of approximately $1{,}400\,\mathrm{m}^2$, using 21 fixed gateways as receivers and mobile transmitters across $6{,}000$ positions. Measurements of $-100\,\mathrm{dBm}$ denote missed detections and are excluded. Each transmitter position is detected by seven gateways, yielding $42{,}000$ valid transmitter–receiver links. Following the partition of NeRF^2 and WRF-GS+, $4{,}800$ transmitter positions ($33{,}600$ links) form the training set and the remaining $1{,}200$ positions ($8{,}400$ links) form the test set.

Models take transmitter and receiver coordinates as input to predict the scalar received power in dBm. The seven gateways that hear a given transmitter are the same throughout the campaign, so the 21 receivers fall into three disjoint groups of seven. We train one GaussTwin model per group and evaluate each on its own share of the test links. SwiftWRF, WRF-GS+, GSRF and BiSplat-WRF fit one model per receiver, 21 in all. NeRF^2 fits a single model for the site and uses each gateway position as a ray origin. Predictions are evaluated by the absolute error $|\widehat{y}_\ell - y_\ell|$ dB relative to the measured RSSI on the $8{,}400$ test links. We report the overall median absolute error across all test links and the per-gateway mean with across-gateway standard deviation.

C IMPLEMENTATION AND REPRODUCTION

C.1 OBSERVATION FEATURES

The angle–delay transform $\mathcal{A}$ weights subcarrier k by the Hann window $w_k^{\mathrm{f}} = \frac{1}{2}[1-\cos(2\pi k/(K+1))]$, $k = 1, \dots, K$, applies orthonormal two-dimensional Fourier transforms over each planar array and an orthonormal inverse Fourier transform over frequency, and keeps the first D_τ delay bins. The observed angle–delay power of a link is $|\mathcal{A}\mathbf{H}|^2$, and its sum over both angle dimensions is the PDP of Section 3. Since the window and the delay truncation change the total power, the energy $\|\mathbf{H}\|_F^2$ is computed directly from the original channel tensor. When every subcarrier and transmit antenna carries equal and uncorrelated power, this energy is proportional to the total received power.

Initialization (Appendix C.6.1) uses the part of $\mathbf{H}$ that the direct path does not explain. With c the speed of light, f_k the frequency of subcarrier k, and $\lambda_k = c/f_k$, the free-space direct-path template $\mathbf{D}$ computed from the antenna positions has entries

$$D_{k,m,n} = \frac{\lambda_k}{4\pi d_{mn}} e^{-\mathrm{j}2\pi d_{mn}/\lambda_k}, \qquad d_{mn} = \|\mathbf{x}_{\mathrm{r},m} - \mathbf{x}_{\mathrm{t},n}\|_2. \tag{14}$$

The least-squares projection onto this template separates a direct component from a residual,

$$\gamma = \frac{\langle \mathbf{D}, \mathbf{H}\rangle_F}{\|\mathbf{D}\|_F^2}, \qquad \mathbf{H}^{\mathrm{d}} = \gamma \mathbf{D}, \qquad \mathbf{H}^{\mathrm{nd}} = \mathbf{H} - \mathbf{H}^{\mathrm{d}}, \tag{15}$$

where $\langle \mathbf{A}, \mathbf{C}\rangle_F = \sum A^*_{k,m,n} C_{k,m,n}$. The two parts are orthogonal, so their energies sum to $\|\mathbf{H}\|_F^2$.

C.2 SURFACE PARAMETERIZATION

Write $\mathbf{E}_i = [\mathbf{u}_i\ \mathbf{v}_i]$ and $\mathbf{S}_i = \mathrm{diag}(a_i, b_i)$. The local coordinate map and its inverse on the plane are

$$\mathbf{x}_i(\boldsymbol{\xi}) = \boldsymbol{\mu}_i + \mathbf{E}_i \mathbf{S}_i \boldsymbol{\xi}, \qquad \boldsymbol{\xi}_i(\mathbf{x}) = \mathbf{S}_i^{-1} \mathbf{E}_i^\top (\mathbf{x} - \boldsymbol{\mu}_i), \qquad \mathrm{d}A = a_i b_i \, \mathrm{d}\boldsymbol{\xi}. \tag{16}$$

In the implementation, specular amplitudes truncate the kernel to $\widetilde{w}_i(\mathbf{x}) = w_i(\mathbf{x}) \mathbf{1}[\|\boldsymbol{\xi}_i(\mathbf{x})\|_2 \le r_{\mathrm{cut}}]$, where

$$r_{\mathrm{cut}} = (2\log(1/\epsilon_w))^{1/(2\nu)}, \qquad \epsilon_w = 10^{-3}, \tag{17}$$

so $\widetilde{w}_i$ differs from w_i only where $w_i < \epsilon_w$. Diffuse integration and geometric surface readout use the ellipse $\Omega_i = \{\mathbf{x}_i(\boldsymbol{\xi}) : \|\boldsymbol{\xi}\|_2 \le 1\}$, which differs from the specular kernel support. Centers, normals and in-plane rotations parameterize geometry, positive scales are represented exponentially within the prescribed bounds, and sigmoid parameters give p_i, ϱ_i and η_i.

C.3 SURFACE PROPAGATION

Mirror geometry. The plane of primitive i satisfies $\mathbf{n}_i^\top \mathbf{x} = d_i$, where $d_i = \mathbf{n}_i^\top \boldsymbol{\mu}_i$. Reflection across this plane and intersection of a line with it are given by

$$\begin{aligned} \mathcal{M}_i(\mathbf{x}) &= \mathbf{x} - 2(\mathbf{n}_i^\top \mathbf{x} - d_i)\mathbf{n}_i, \\ \mathcal{J}_i(\mathbf{a}, \mathbf{b}) &= \mathbf{a} + \beta_i(\mathbf{a}, \mathbf{b})(\mathbf{b} - \mathbf{a}), \qquad \beta_i(\mathbf{a}, \mathbf{b}) = \frac{d_i - \mathbf{n}_i^\top \mathbf{a}}{\mathbf{n}_i^\top (\mathbf{b} - \mathbf{a})}. \end{aligned} \tag{18}$$

For array centers $\mathbf{x}_\mathrm{t}, \mathbf{x}_\mathrm{r}$, the first-order point is $\mathbf{q}_{p,1} = \mathcal{J}_i(\mathcal{M}_i(\mathbf{x}_\mathrm{t}), \mathbf{x}_\mathrm{r})$. A second-order path on the ordered pair (i, j), $i \neq j$, has points

$$\mathbf{q}_{p,2} = \mathcal{J}_j(\mathcal{M}_j(\mathcal{M}_i(\mathbf{x}_\mathrm{t})), \mathbf{x}_\mathrm{r}), \qquad \mathbf{q}_{p,1} = \mathcal{J}_i(\mathcal{M}_i(\mathbf{x}_\mathrm{t}), \mathbf{q}_{p,2}). \tag{19}$$

The validity indicator $v_p \in \{0, 1\}$ equals one when path p is feasible, that is, when both intersections lie inside their construction segments, their kernels have nonzero support, and the two physical neighbors of each reflection lie on the same side of its plane.

Each transmit element is mirrored through the same sequence of planes. For B_p reflections on $i_1, \ldots, i_{B_p}$, define

$$\mathbf{x}_{\mathrm{t},n}^{(p)} = (\mathcal{M}_{i_{B_p}} \circ \cdots \circ \mathcal{M}_{i_1})(\mathbf{x}_{\mathrm{t},n}), \qquad L_{p,m,n} = \|\mathbf{x}_{\mathrm{r},m} - \mathbf{x}_{\mathrm{t},n}^{(p)}\|_2, \qquad \tau_{p,m,n} = L_{p,m,n}/c. \tag{20}$$

For a valid center path, its unfolded length equals the sum of its physical segment lengths. Relative to center length L_p, the array phase factor is $\exp[-\mathrm{j}2\pi(L_{p,m,n} - L_p)/\lambda_k]$. Interaction weights and visibility are evaluated on the center path, while propagation phase and distance attenuation use the individual antenna-pair lengths without a far-field approximation.

Transmission and path amplitude. For each path segment, every crossing with an intervening primitive scales the amplitude by $\sqrt{1 - p_j w_j(\mathbf{x})}$ with no phase change. The reflecting primitives of a path are excluded from its occlusion calculation. With endpoints $\mathbf{q}_{p,0} = \mathbf{x}_\mathrm{t}$ and $\mathbf{q}_{p,B_p+1} = \mathbf{x}_\mathrm{r}$, the path amplitude is

$$\begin{aligned} A_{p,k,m,n} = v_p \frac{\lambda_k}{4\pi L_{p,m,n}} \prod_{a=0}^{B_p} T(\mathbf{q}_{p,a}, \mathbf{q}_{p,a+1}) \\ \times \prod_{a=1}^{B_p} \sqrt{(1 - \eta_{i_a}) \alpha_{i_a} \widetilde{w}_{i_a}(\mathbf{q}_{p,a})}. \end{aligned} \tag{21}$$

The direct response uses $B_p = 0$ and the empty reflection product, with visibility along the direct segment. Visibility searches retain crossings within a finite kernel neighborhood, using radius $\max(1.4, r_\mathrm{cut}(\nu_\mathrm{initial}))$ in normalized coordinates throughout training.

All primitives are considered for first-order reflection. For second order, the feasible ordered pairs of each link are ranked by the center-path amplitude proxy $\sqrt{(1 - \eta_i)(1 - \eta_j)\alpha_i \alpha_j \widetilde{w}_i(\mathbf{q}_{p,1}) \widetilde{w}_j(\mathbf{q}_{p,2})}/L_p$. Only the top-ranked pairs, up to the number of second-order paths per link in Table C.1, are retained before occlusion is evaluated, as a computational approximation to the full path set.

C.3.1 DIFFUSE SURFACE POWER

We derive Eq. 5 by following power from the transmitter to a surface patch and then to the receiver. The scalar model uses unit transmit power, isotropic matched antennas with unit gain, and segment transmission factors that exclude primitive i from occlusion. For $\mathbf{q} \in \Omega_i$, the signed direction cosines are

$$\cos\theta = \frac{\mathbf{n}_i^\top (\mathbf{x}_\mathrm{t} - \mathbf{q})}{\|\mathbf{x}_\mathrm{t} - \mathbf{q}\|_2}, \qquad \cos\psi = \frac{\mathbf{n}_i^\top (\mathbf{x}_\mathrm{r} - \mathbf{q})}{\|\mathbf{q} - \mathbf{x}_\mathrm{r}\|_2}. \tag{22}$$

We first assume that the transmitter and the receiver lie on the same side of the primitive and that $\mathbf{n}_i$ points toward this side, so both cosines are nonnegative.

Table C.1: **Default numerical settings for the surface model and learning procedure.**

Quantity	Value	Quantity	Value
Surface model		*Optimization*	
Surface budget M	512	Training iterations N_{it}	40,000
Initial in-plane scale	0.9 m	Minibatch links $\lvert\mathcal{I}\rvert$	4
Minimum in-plane scale	0.05 m	Area weight λ_{A}	0.01
Maximum in-plane scale	1.5 m	Position learning rate	3×10^{-4}
Initial existence probability	0.9	Orientation learning rate	10^{-3}
Initial reflectance	0.5	Scale learning rate	5×10^{-3}
Initial diffuse fraction η_i	0.0225	Existence learning rate	0.05
Initial kernel exponent ν	2	Reflectance learning rate	0.025
Final kernel exponent ν	8	Diffuse-fraction learning rate	0.025
Kernel exponent ramp	6,000 iterations	*Structure*	
Propagation		Splitting start	750
Second-order paths per link	4,096	Splitting interval	375
Phase samples S_ϕ	8	Child scale factor γ_{split}	0.6
Training integration points $\lvert\mathcal{X}_i\rvert$	16	Replacement start	750
Retained delay bins D_τ	90	Replacement interval	1,500
Initialization			
Support grid spacing	0.3 m		
Support grid margin	0.6 m		

Incident power at the surface. A unit-power isotropic transmitter produces a power flux density of $1/(4\pi\|\mathbf{x}_{\mathrm{t}}-\mathbf{q}\|_2^2)$ at distance $\|\mathbf{x}_{\mathrm{t}}-\mathbf{q}\|_2$. The projected area of a patch of area $\mathrm{d}A$ is $\cos\theta\,\mathrm{d}A$. With the power transmission along the incoming segment included, the incident irradiance is

$$E_{\mathrm{in}}(\mathbf{q}) = \frac{T(\mathbf{x}_{\mathrm{t}},\mathbf{q})^2}{4\pi\|\mathbf{x}_{\mathrm{t}}-\mathbf{q}\|_2^2}\cos\theta. \tag{23}$$

The patch receives power $E_{\mathrm{in}}(\mathbf{q})\,\mathrm{d}A$, where the incidence cosine accounts for its reduced projected area at oblique angles.

Lambertian scattering. The Lambertian BRDF is the local diffuse reflectance $\eta_i\alpha_i w_i(\mathbf{q})$ divided by π, which gives the outgoing radiance

$$L_{\mathrm{out}}(\mathbf{q}) = \frac{\eta_i\alpha_i w_i(\mathbf{q})}{\pi}E_{\mathrm{in}}(\mathbf{q}). \tag{24}$$

The normalization follows from the cosine-weighted integral over the outgoing hemisphere,

$$\int_0^{2\pi}\int_0^{\pi/2}\cos\psi\sin\psi\,\mathrm{d}\psi\,\mathrm{d}\varphi = \pi. \tag{25}$$

Consequently, the outgoing power integrated over that hemisphere equals $\eta_i\alpha_i w_i(\mathbf{q})$ times the incident power.

Power captured by the receiver. At wavelength $\lambda = c/f$, the effective area of a unit-gain matched antenna is $A_{\mathrm{e}} = \lambda^2/(4\pi)$. Toward the receiver, the patch has projected area $\cos\psi\,\mathrm{d}A$ and outgoing radiant intensity $L_{\mathrm{out}}(\mathbf{q})\cos\psi\,\mathrm{d}A$. Propagation over distance $\|\mathbf{q}-\mathbf{x}_{\mathrm{r}}\|_2$ converts this intensity to power flux density by a factor $1/\|\mathbf{q}-\mathbf{x}_{\mathrm{r}}\|_2^2$. Multiplication by the effective area and outgoing power transmission gives

$$\mathrm{d}P = L_{\mathrm{out}}(\mathbf{q})\cos\psi\,\mathrm{d}A\,\frac{A_{\mathrm{e}}}{\|\mathbf{q}-\mathbf{x}_{\mathrm{r}}\|_2^2}T(\mathbf{q},\mathbf{x}_{\mathrm{r}})^2. \tag{26}$$

Equivalently, $A_{\mathrm{e}}/\|\mathbf{q}-\mathbf{x}_{\mathrm{r}}\|_2^2$ is the effective solid angle subtended by the receiving antenna as seen from the patch.

Received power density. With the incident irradiance, the BRDF and the effective area substituted, the received power density is

$$\begin{aligned}\sigma_i(f,\mathbf{q}) = \frac{\mathrm{d}P}{\mathrm{d}A} &= \underbrace{\frac{T(\mathbf{x}_\mathrm{t},\mathbf{q})^2\cos\theta}{4\pi\|\mathbf{x}_\mathrm{t}-\mathbf{q}\|_2^2}}_{\text{incident irradiance}}\underbrace{\frac{\eta_i\alpha_i w_i(\mathbf{q})}{\pi}}_{\text{Lambertian BRDF}} \\ &\quad\times\underbrace{\frac{\cos\psi}{\|\mathbf{q}-\mathbf{x}_\mathrm{r}\|_2^2}T(\mathbf{q},\mathbf{x}_\mathrm{r})^2}_{\text{outgoing propagation}}\underbrace{\frac{\lambda^2}{4\pi}}_{\text{effective area}} \\ &= \left(\frac{\lambda}{4\pi}\right)^2\frac{\eta_i\alpha_i w_i(\mathbf{q})}{\pi}\frac{\cos\theta\cos\psi}{\|\mathbf{x}_\mathrm{t}-\mathbf{q}\|_2^2\|\mathbf{q}-\mathbf{x}_\mathrm{r}\|_2^2}T(\mathbf{x}_\mathrm{t},\mathbf{q})^2T(\mathbf{q},\mathbf{x}_\mathrm{r})^2.\end{aligned} \tag{27}$$

The two propagation segments have distinct roles: the first determines the incident irradiance, and the second delivers scattered power to the receiving antenna, with inverse-square power attenuation along each segment. The transmission factors are squared because they act on amplitude. The coefficient $(\lambda/(4\pi))^2$ contains the transmitter's spherical spreading factor $1/(4\pi)$ and the receiver's effective area $\lambda^2/(4\pi)$.

Scattering from both sides. A primitive can scatter from both sides: when the transmitter and the receiver lie on the same side and $\mathbf{n}_i$ points away from this side, the derivation above holds with $-\mathbf{n}_i$ in place of $\mathbf{n}_i$, which negates $\cos\theta$ and $\cos\psi$ and leaves $\cos\theta\cos\psi$ unchanged. When they lie on opposite sides of the primitive, the primitive scatters no power toward the receiver. The angular factor for both sides is therefore $\max(\cos\theta\cos\psi, 0)$. With the wavelength scaling factored out, the frequency-independent weight is

$$\zeta_i(\mathbf{q}) = \eta_i\alpha_i w_i(\mathbf{q})\frac{\max(\cos\theta\cos\psi,0)}{\pi\|\mathbf{x}_\mathrm{t}-\mathbf{q}\|_2^2\|\mathbf{q}-\mathbf{x}_\mathrm{r}\|_2^2}T(\mathbf{x}_\mathrm{t},\mathbf{q})^2T(\mathbf{q},\mathbf{x}_\mathrm{r})^2, \tag{28}$$

so $\sigma_i(f,\mathbf{q}) = (\lambda/(4\pi))^2\zeta_i(\mathbf{q})$ recovers Eq. 5.

Array response and surface integration. The center-path delay is $\tau_i(\mathbf{q}) = (\|\mathbf{x}_\mathrm{t}-\mathbf{q}\|_2 + \|\mathbf{q}-\mathbf{x}_\mathrm{r}\|_2)/c$. Diffuse synthesis separates a center-frequency spherical array response from the wideband center delay. At $\bar{f} = K^{-1}\sum_k f_k$, define

$$\begin{aligned}a_{\mathrm{r},m}(\mathbf{q}) &= e^{-\mathrm{j}2\pi\bar{f}(\|\mathbf{q}-\mathbf{x}_{\mathrm{r},m}\|_2-\|\mathbf{q}-\mathbf{x}_\mathrm{r}\|_2)/c}, \\ a_{\mathrm{t},n}(\mathbf{q}) &= e^{-\mathrm{j}2\pi\bar{f}(\|\mathbf{q}-\mathbf{x}_{\mathrm{t},n}\|_2-\|\mathbf{x}_\mathrm{t}-\mathbf{q}\|_2)/c}.\end{aligned} \tag{29}$$

The unit local response, including wavelength scaling, is

$$U_{i,k,m,n}(\mathbf{q}) = \frac{\lambda_k}{4\pi}a_{\mathrm{r},m}(\mathbf{q})a_{\mathrm{t},n}(\mathbf{q})e^{-\mathrm{j}2\pi(f_k-\bar{f})\tau_i(\mathbf{q})}. \tag{30}$$

The response of the patch around $\mathbf{q}$ is $\mathbf{G}_i(\mathbf{q}) = \sqrt{\zeta_i(\mathbf{q})}\,e^{\mathrm{j}\chi_i(\mathbf{q})}\mathbf{U}_i(\mathbf{q})$, with an unknown local scattering phase $\chi_i(\mathbf{q})$. This phase is the same for all subcarriers and antenna pairs, so it does not change $|\mathcal{A}\mathbf{G}_i(\mathbf{q})|^2$. Since the phases of different points are modeled as independent, the diffuse power is the integral of the local powers over the primitive surfaces,

$$\widehat{\mathbf{P}}_\mathrm{s} = \sum_i\int_{\Omega_i}\zeta_i(\mathbf{q})|\mathcal{A}\mathbf{U}_i(\mathbf{q})|^2\,\mathrm{d}A, \tag{31}$$

$$\widehat{E}_\mathrm{s} = N_\mathrm{r}N_\mathrm{t}\sum_k\left(\frac{\lambda_k}{4\pi}\right)^2\sum_i\int_{\Omega_i}\zeta_i(\mathbf{q})\,\mathrm{d}A. \tag{32}$$

Each local response is thus transformed before its power is integrated, which eliminates the phase $\chi_i(\mathbf{q})$ of every surface point $\mathbf{q}$ analytically.

C.4 Phase marginalization and characteristic prediction

In Eq. 7, the expectation is taken over the latent phases $\boldsymbol{\phi}$ alone, while the primitives, the link and the retained paths stay fixed. Let $\kappa_{p,i}$ be the number of reflections of path p on primitive i. Since each

phase ϕ_i is shared across subcarriers and antenna elements, the response $\widehat{\mathbf{H}}_p(\boldsymbol{\phi})$ of path p depends on $\boldsymbol{\phi}$ only through the scalar factor $z_p = \exp(\mathrm{j}\sum_i \kappa_{p,i}\phi_i)$. With $\widehat{\mathbf{H}}_\mathrm{r} = \sum_p \widehat{\mathbf{H}}_p$, the expansion of $|\mathcal{A}\widehat{\mathbf{H}}_\mathrm{r}(\boldsymbol{\phi})|^2$ contains the angle–delay power $|\mathcal{A}\widehat{\mathbf{H}}_p|^2$ of each path and, for every pair of paths $p \neq p'$, a cross term $2\,\mathrm{Re}[(\mathcal{A}\widehat{\mathbf{H}}_p) \odot (\mathcal{A}\widehat{\mathbf{H}}_{p'})^*]$, where $\odot$ denotes elementwise multiplication. The power of each path does not depend on $\boldsymbol{\phi}$, since $|z_p| = 1$, whereas the cross term carries the factor $z_p z_{p'}^*$, whose expectation under independent uniform phases is

$$\mathbb{E}[z_p z_{p'}^*] = \prod_{i=1}^{M} \frac{1}{2\pi} \int_0^{2\pi} e^{\mathrm{j}(\kappa_{p,i}-\kappa_{p',i})\phi}\,\mathrm{d}\phi = \mathbf{1}[\kappa_{p,i} = \kappa_{p',i} \text{ for all } i]. \tag{33}$$

Therefore, the cross term of a pair of paths is removed whenever their reflection counts differ on at least one primitive, which is the case for almost all pairs of retained paths. The cross terms of the energy $\|\widehat{\mathbf{H}}_\mathrm{r}(\boldsymbol{\phi})\|_F^2$ in Eq. 8 carry the same factors $z_p z_{p'}^*$ and are removed for the same pairs.

The implementation replaces the expectations in Eqs. 7 and 8 by averages over S_ϕ independent phase samples,

$$\widehat{\mathbf{P}} = |\mathcal{A}\widehat{\mathbf{H}}_\mathrm{d}|^2 + \frac{1}{S_\phi}\sum_{s=1}^{S_\phi} |\mathcal{A}\widehat{\mathbf{H}}_\mathrm{r}(\boldsymbol{\phi}^{(s)})|^2 + \widehat{\mathbf{P}}_\mathrm{s}, \tag{34}$$

$$\widehat{E} = \|\widehat{\mathbf{H}}_\mathrm{d}\|_F^2 + \frac{1}{S_\phi}\sum_{s=1}^{S_\phi} \|\widehat{\mathbf{H}}_\mathrm{r}(\boldsymbol{\phi}^{(s)})\|_F^2 + \widehat{E}_\mathrm{s}. \tag{35}$$

These averages are unbiased, and the cross terms that the expectations remove shrink as S_ϕ grows. Unlike training, which redraws the phases at every iteration, validation and prediction reuse a fixed set of phase samples indexed by primitive identity.

C.5 Learning objective

The four feature terms compare the predictions for a link with its observed channel. The PDP and the receive and transmit APS each sum the angle–delay power over the other two dimensions, and the predicted energy is that of Eq. 8. With N observed links, the terms of Eq. 9 are

$$\begin{aligned} \mathcal{L}_\mathrm{E} &= \frac{1}{N}\sum_{\ell=1}^{N} \left| 10\log_{10}\frac{\widehat{E}_\ell}{E_\ell} \right|, \\ \mathcal{L}_a &= \frac{1}{N}\sum_{\ell=1}^{N} 10\log_{10}\frac{\|\widehat{\mathbf{p}}_\ell^a - \mathbf{p}_\ell^a\|_2^2}{\|\mathbf{p}_\ell^a\|_2^2}, \qquad a \in \{\mathrm{PDP}, \mathrm{r}, \mathrm{t}\}. \end{aligned} \tag{36}$$

At each iteration, these terms are evaluated on a minibatch $\mathcal{I}$ of links drawn from the observation set, with $|\mathcal{I}|^{-1}$ in place of N^{-1}. The area regularizer is

$$\mathcal{R}_\mathrm{A}(\mathcal{G}) = \frac{1}{M s_\mathrm{max}^2} \sum_{i \text{ active}} a_i b_i, \tag{37}$$

which measures area relative to the surface budget M and scale bound s_max, without weighting by the learned existence probability.

C.6 Initialization and surface refinement

C.6.1 Initialization from wireless spatial support

Initialization places surfaces using only the training channels and their Tx and Rx positions, before any gradient is available. It follows the delay-and-sum back-projection used in multistatic wideband radar imaging (Ahmad et al., 2005; Zhuge et al., 2010). For every point of a regular grid it measures the share of the observed non-direct response consistent with an interaction at that point. The grid is a three-dimensional lattice over the bounding box of the observed endpoints, enlarged by a margin on every side, with the support grid spacing and margin in Table C.1.

Back-projection of one link. A training link enters the back-projection through its residual $\mathbf{H}^{\mathrm{nd}}$ of Eq. 15 since the back-projected direct component would otherwise peak along the line of sight instead of on the reflecting surfaces. Let Δf be the subcarrier spacing and $\bar{f}$ the band center. As in the diffuse response of Appendix C.3.1, the band center is factored out of the wideband phase, and the windowed residual of antenna pair (m, n) is transformed into the delay response

$$r_{\ell,m,n}(\tau) = \sum_{k=1}^{K} w_k^{\mathrm{f}} H^{\mathrm{nd}}_{\ell,k,m,n} e^{\mathrm{j}2\pi(f_k - \bar{f})\tau}. \tag{38}$$

Since the K samples are spaced Δf apart, $r_{\ell,m,n}$ is periodic in delay with period $1/\Delta f$, and delays that differ by one period are indistinguishable. The response is evaluated over $0 \le \tau < 1/\Delta f$ on a grid that oversamples its native resolution $1/(K\Delta f)$ by a factor of 8, and is read by linear interpolation between grid samples. A candidate point $\mathbf{x}$ fixes one delay per antenna pair,

$$\tau_{\ell,m,n}(\mathbf{x}) = \frac{\|\mathbf{x} - \mathbf{x}_{\mathrm{t},n}\|_2 + \|\mathbf{x} - \mathbf{x}_{\mathrm{r},m}\|_2}{c}, \tag{39}$$

the time taken by the two segments from the transmitting element to $\mathbf{x}$ and on to the receiving element. The link is back-projected onto $\mathbf{x}$ by reading each antenna pair at its own delay and restoring the band-center phase, which gives the coherent match

$$c_\ell(\mathbf{x}) = \frac{1}{N_{\mathrm{r}} N_{\mathrm{t}}} \sum_{m=1}^{N_{\mathrm{r}}} \sum_{n=1}^{N_{\mathrm{t}}} r_{\ell,m,n}\big(\tau_{\ell,m,n}(\mathbf{x})\big) e^{\mathrm{j}2\pi\bar{f}\tau_{\ell,m,n}(\mathbf{x})}. \tag{40}$$

A path of delay τ leaves the residual phase $e^{-\mathrm{j}2\pi\bar{f}\tau}$ in Eq. 38, which the exponential removes, so the antenna-pair terms are in phase when an interaction at $\mathbf{x}$ explains part of the link and have unrelated phases otherwise. Only the phases measured inside one link enter, and the support below depends on $|c_\ell|$ and is therefore invariant to a phase offset between links.

Support and orientation. Let $\mathbf{d}_{\mathrm{in}}$ and $\mathbf{d}_{\mathrm{out}}$ be the unit directions from the transmitter to $\mathbf{x}$ and from $\mathbf{x}$ to the receiver. Equal incidence and reflection angles hold at $\mathbf{x}$ only for a surface whose normal lies along $\mathbf{d}_{\mathrm{in}} - \mathbf{d}_{\mathrm{out}}$, so each link votes for that direction, $\mathbf{n}_\ell(\mathbf{x}) = (\mathbf{d}_{\mathrm{in}} - \mathbf{d}_{\mathrm{out}})/\|\mathbf{d}_{\mathrm{in}} - \mathbf{d}_{\mathrm{out}}\|_2$. Multipath-based SLAM encodes the same specular geometry through image sources, or virtual anchors, mirrored across the reflecting plane (Leitinger et al., 2023). Cross-link support and orientation consistency are summarized by

$$A(\mathbf{x}) = \frac{\sum_\ell \mathbf{1}_\ell |c_\ell|^2}{\sum_\ell \mathbf{1}_\ell}, \qquad \mathbf{V}(\mathbf{x}) = \frac{\sum_\ell \mathbf{1}_\ell |c_\ell| \mathbf{n}_\ell \mathbf{n}_\ell^\top}{\sum_\ell \mathbf{1}_\ell |c_\ell|}, \tag{41}$$

where $\mathbf{1}_\ell(\mathbf{x}) = 1$ when $\tau_{\ell,m,n}(\mathbf{x}) < 1/\Delta f$ for every antenna pair of link ℓ, $\mathbf{1}_\ell(\mathbf{x}) = 0$ otherwise, and arguments $\mathbf{x}$ are suppressed in the sums. Beyond this delay range the response of link ℓ at $\mathbf{x}$ would be read from a shorter path, so the indicator excludes that link at $\mathbf{x}$. The outer products remove the sign of each vote, left undetermined by a surface normal. A point carries support when many links back-project onto it, and their votes agree when $\mathbf{V}$ has one dominant eigenvalue. With $\mathcal{N}$ rescaling a field between its 10th and 99.5th percentiles and clipping it to $[0, 1]$, the spatial support combines cross-link support and orientation consistency as $h(\mathbf{x}) = \mathcal{N}(\mathcal{N}(A(\mathbf{x}))\lambda_{\max}(\mathbf{V}(\mathbf{x})))$.

Placement. The candidates are the grid points whose support exceeds its median over the grid, and the first center maximizes h over them. With $\mathcal{X}_{\mathrm{placed}}$ the centers already placed, each further center maximizes $(h(\mathbf{x}) + 0.25) \min_{\boldsymbol{\mu} \in \mathcal{X}_{\mathrm{placed}}} \|\mathbf{x} - \boldsymbol{\mu}\|_2$ over the same candidates, which spreads the centers across the supported region instead of crowding its strongest point. The score is a distance in meters biased by the dimensionless support, and the constant keeps weakly supported candidates eligible once the strong ones have been placed, and is the same for every scene. Placement stops at the surface budget M, or at the number of candidates when these are fewer, leaving training to start with fewer primitives. Normals follow the leading eigenvectors of $\mathbf{V}$, and the in-plane scales and radiative parameters start from their initial values in Table C.1. Both in-plane scales start at the same value, so the rotation about the normal has no effect at initialization, and the trained rotation starts from a frame completed from the normal against a fixed reference axis.

C.6.2 OBSERVATION-GUIDED REPLACEMENT

Replacement introduces reflecting or occluding surfaces where the existing representation leaves wireless evidence unexplained. The squared projection coefficient $|\gamma_\ell|^2$ of link ℓ onto the free-space direct template, with γ_ℓ from Eq. 15, serves as a proxy for its direct power transmittance, and links with $|\gamma_\ell|^2 \geq 0.9$ and $|\gamma_\ell|^2 \leq 0.01$ form the LoS and NLoS link sets $\mathcal{I}_{\mathrm{LoS}}$ and $\mathcal{I}_{\mathrm{NLoS}}$. Let $\mathcal{V}_{\mathrm{LoS}}$ contain the spatial cells crossed by LoS segments. Candidate occluders are supported by

$$d(v) = \mathbf{1}[v \notin \mathcal{V}_{\mathrm{LoS}}] \left(\sum_{\ell \in \mathcal{I}_{\mathrm{NLoS}}} \mathbf{1}[\widehat{T}_\ell^2 > \epsilon_{\mathrm{block}}] \mathbf{1}[v \cap s_\ell \neq \varnothing] + \sum_{p \in \mathcal{P}_{\mathrm{over}}} \zeta_p \, \omega_p(v) \right). \tag{42}$$

Here s_ℓ is the direct segment, $\widehat{T}_\ell^2$ is the predicted direct power transmittance, the square of $T(\mathbf{x}_{\mathrm{t}}, \mathbf{x}_{\mathrm{r}})$ in Eq. 3 evaluated with hard ellipse crossings, $\omega_p(v)$ counts the segments of path p that cross cell v, and $\mathcal{P}_{\mathrm{over}}$ contains dominant paths of NLoS links with overpredicted power. Their weights ζ_p are single-path energies divided by the corresponding predicted direct-plus- specular link energy. High-demand neighborhoods outside $\mathcal{V}_{\mathrm{LoS}}$ yield candidate centers and normals from their weighted spatial covariance. In planar occlusion proposals, this covariance uses horizontal coordinates and gives a vertical surface.

Reflecting candidates use the unexplained fraction of observed delay power,

$$u_\ell[\tau] = \max\left(0, 1 - \frac{\widehat{p}_\ell^{\mathrm{PDP}}[\tau]}{p_\ell^{\mathrm{PDP}}[\tau]}\right). \tag{43}$$

The windowed observed delay response is multiplied by $\sqrt{u_\ell[\tau]}$ and transformed back, which keeps the observed phases in the remaining response. The initialization support operator applied to this response gives candidate reflectors outside the already occupied surface support. Candidate replacements reuse low-existence slots and therefore leave the surface budget unchanged. A relocated primitive starts from the initial values of Table C.1, except for its diffuse fraction, which starts from the mean over the other active primitives, taken in logit space. Each candidate is assessed sequentially against the updated scene and kept when it lowers the mean absolute received power error on a panel of 128 training links.

C.6.3 SPLITTING CONFLICTING SURFACE SUPPORT

A primitive traversed by both LoS and NLoS links can contain regions with conflicting visibility evidence. Eligible primitives have sufficient existence and enough crossings of each type within their unit ellipse. Let $\bar{u}_{i,a}^{\mathrm{L}}$ and $\bar{u}_{i,a}^{\mathrm{N}}$ be mean crossing coordinates in meters along local in-plane axis $a \in \{1, 2\}$. The split direction and child geometry are

$$\begin{aligned} a_i^\star &= \arg\max_{a \in \{1,2\}} \frac{|\bar{u}_{i,a}^{\mathrm{L}} - \bar{u}_{i,a}^{\mathrm{N}}|}{s_{i,a}}, \\ \boldsymbol{\mu}_i^\pm &= \boldsymbol{\mu}_i \pm \tfrac{1}{2} s_{i,a_i^\star} \mathbf{e}_{i,a_i^\star}, \qquad s_{i,a_i^\star}^\pm = \gamma_{\mathrm{split}} s_{i,a_i^\star}, \end{aligned} \tag{44}$$

where $\mathbf{e}_{i,1} = \mathbf{u}_i$ and $\mathbf{e}_{i,2} = \mathbf{v}_i$. Children inherit the remaining parameters and are then optimized independently. One child occupies the parent slot and the other a low-existence slot, so splitting respects the surface budget. When splitting and replacement coincide, splitting is applied first, before the gradient update.

C.7 NUMERICAL APPROXIMATION AND OPTIMIZATION

Surface integrals over the ellipse Ω_i are computed in the local coordinates $\boldsymbol{\xi}$ of Appendix C.2, in which Ω_i is the unit disk $\|\boldsymbol{\xi}\|_2 \leq 1$. The integration points are spread uniformly in area over this disk, so each of the $|\mathcal{X}_i|$ points of primitive i represents the area $\Delta A_{\mathbf{q}} = \pi a_i b_i / |\mathcal{X}_i|$. They are drawn by random stratified sampling during training and placed on a fixed Fibonacci spiral of 128 points during evaluation. Visibility is computed at the integration points themselves during training, whereas during evaluation each integration point takes the transmission of the nearest of the 16 visibility anchors on its primitive. These integration and visibility approximations are separate from latent phase sampling.

Algorithm 1 Learning GaussTwin from wireless observations

Input: observations $\mathcal{D}$, validation links, number of iterations N_{it}.
1 Initialize $\mathcal{G} \leftarrow \mathrm{Initialize}(\mathcal{D})$ from spatial support (App. C.6.1).
2 **for** $t = 0, \ldots, N_{\mathrm{it}} - 1$ **do**
3 If $t \in \mathcal{T}_{\mathrm{val}}$, score $\mathcal{G}$ on the validation links, and keep it as $\widehat{\mathcal{G}}$ if the score is the lowest so far.
4 If $t \in \mathcal{T}_{\mathrm{split}}$, refine conflicting support with $\mathcal{G} \leftarrow \mathrm{Split}(\mathcal{G}, \mathcal{D})$ (App. C.6.3).
5 If $t \in \mathcal{T}_{\mathrm{replace}}$, relocate low-existence primitives with $\mathcal{G} \leftarrow \mathrm{Replace}(\mathcal{G}, \mathcal{D})$ (App. C.6.2).
6 Sample a minibatch $\mathcal{I}$ of links from $\mathcal{D}$ and phase vectors $\boldsymbol{\phi}^{(1)}, \ldots, \boldsymbol{\phi}^{(S_\phi)}$.
7 Compute $\widehat{\mathbf{P}}_\ell$ and $\widehat{E}_\ell$ for $\ell \in \mathcal{I}$ with Eqs. 34 and 35.
8 Update $\mathcal{G}$ with Adam on the gradient of Eq. 9 over $\mathcal{I}$.
9 **end for**, and return $\widehat{\mathcal{G}}$.

Table C.2: **Capabilities of representative wireless modeling methods.** Methods with identical profiles are grouped together. ✓: provides capability natively. ×: lacks capability or requires external geometric input/supervision. –: not applicable to setting.

Representative methods	No geom. prior	Explicit multipath	Surface recovery	Channel pred.
Neural fields: NeRF^2, NeWRF, GRaF	✓	×	×	✓
Gaussian splatting: WRF-GS+, GSRF, SwiftWRF				
RF-3DGS, RF-PGS, PC-TGS	×	×	–	✓
WiNeRT, Diff. RT, RFIR	×	✓	–	✓
BiWGS, OctCGS	✓	✓	×	✓
Multipath RF-SLAM, Neural RF SLAM	✓	✓	✓	×
GaussTwin (Ours)	✓	✓	✓	✓

We update the continuous parameters with Adam (Kingma & Ba, 2015), using a separate constant learning rate for each type of parameter. By default, the direct path passes gradients only to the existence probabilities of the primitives, whose geometry and reflectance receive no gradient from it. The iteration sets $\mathcal{T}_{\mathrm{val}}$, $\mathcal{T}_{\mathrm{split}}$ and $\mathcal{T}_{\mathrm{replace}}$ of Algorithm 1 each start at a given iteration and repeat at a fixed interval: validation starts at iteration $8{,}000$ and repeats every $1{,}000$ iterations, and splitting and replacement follow the start and interval entries of Table C.1. The validation score sums the four feature terms on the validation links, each taken as its median over links, and leaves out the area regularizer. All reported results use the selected primitives $\widehat{\mathcal{G}}$.

Table C.1 lists the reference implementation's default numerical settings. For larger scenes, initial and maximum in-plane scales are multiplied by $\max(1, L_{\mathrm{scene}}/12.2\,\mathrm{m})$, where L_{scene} is the shorter horizontal span of the endpoint bounds including their margin.

C.8 Method capability comparison

Table C.2 compares GaussTwin with representative wireless modeling methods on four capabilities: operation without geometric priors, explicit physical multipath modeling, explicit surface geometry recovery, and site-specific channel prediction at unseen links.

D Additional Experiments and Ablations

D.1 Channel prediction on unseen configurations

Figure D.1 compares channel prediction on the $2{,}400$ test links of each indoor scene with an unseen transmitter position, the 800 with an unseen receiver position, and the 800 with both unseen. The power error measures absolute power accuracy, and the joint angle–delay cosine similarity measures the shape of the predicted propagation features.

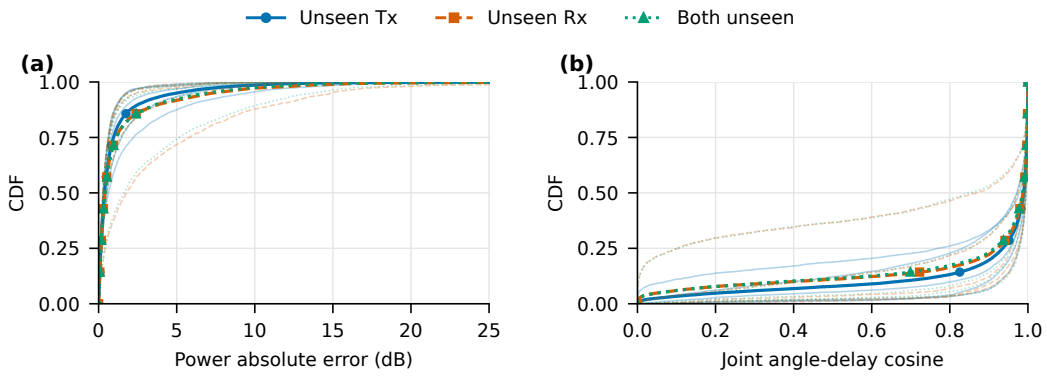


Figure D.1: **Channel prediction on unseen configurations.** CDFs of (a) received power absolute error and (b) joint angle–delay cosine similarity across unseen Tx, unseen Rx, and both unseen. Faint curves show individual scenes, and thicker curves with markers combine all five scenes.

For most links the three configurations give similar errors, with a median power error of about 0.4 dB in each. About 75% of all test links are predicted within 1 dB of the measured power. Large errors occur more often when the receiver is unseen, mostly in the six-room scene, where the median joint cosine similarity of these links drops to between 0.85 and 0.87, compared with above 0.95 in the other scenes.

D.2 Core mechanism ablation

Table D.1 varies the initialization, the geometry learning and the propagation model on the six-room scene. Random and ground-truth initialization place the primitives at random or on the ground-truth surfaces. The two fixed-geometry variants keep the primitives at their initial or ground-truth geometry, and fixed structure disables only replacement and splitting.

Fixed initial geometry and fixed structure both predict the channel less accurately than the full method and recover the surfaces with a larger Chamfer-L_1 distance, so channel prediction and geometry improve together when the geometry is learned. With fixed initial geometry, the median PDP NMSE rises from -8.74 to -1.48 dB. A good initialization also matters: from random initialization, coverage falls to 47.9% and channel prediction degrades as well. Our initialization uses only the training channels and their Tx and Rx positions, without priors from other scenes, and a finer

Table D.1: **Mechanism ablation on the six-room scene.** Each row is one training run with the same random number sequence as the first six-room run in Table 1. Rows marked with * start from the ground-truth geometry and are not ranked on the geometry metrics. Metrics follow Table 1.

	Channel			Geometry		
Variant	Power MAE (dB) ↓	PDP NMSE median (dB) ↓	Joint cosine ↑	Chamfer (cm) ↓	Precision ↑	Coverage ↑
Ours	2.75	−8.74	0.738	**16.93**	**69.67**	75.48
Initialization						
Random init.	3.24	−4.86	0.610	54.51	47.31	47.89
Ground-truth init.*	**2.72**	**−9.00**	**0.746**	13.85	73.59	81.28
Geometry learning						
Fixed init.	4.61	−1.48	0.536	18.47	62.53	71.87
Fixed structure	3.17	−5.88	0.670	20.06	50.75	**82.49**
Fixed ground truth*	3.47	−5.59	0.708	3.56	87.23	97.91
Propagation						
LoS+(1st&2nd) reflection	2.86	−8.08	0.706	23.38	65.41	69.70
LoS+1st reflection	3.85	−4.90	0.583	23.31	64.39	71.53

initialization, for example from 3D generative models conditioned on wireless observations, may further improve the recovered geometry or complete unobserved surfaces.

Ground-truth initialization followed by the same training reaches channel metrics close to those of the full method, with a median PDP NMSE of $-9.00\,\mathrm{dB}$, so ground-truth geometry adds little to the quality of channel prediction. With the ground-truth geometry fixed, the geometry metrics are close to ideal, with a coverage of 97.9%, but channel prediction is less accurate than with the learned geometry, with a median PDP NMSE of $-5.59\,\mathrm{dB}$. This indicates that part of the learned geometry compensates for propagation effects that the operator does not model, and the additional primitives placed for these effects lower the geometric accuracy. We leave the explicit modeling of these effects to future work.

Without diffuse scattering, channel prediction changes little, because diffuse scattering carries little of the received energy indoors (Appendix D.4), yet the Chamfer-L_1 distance increases, so this small share of energy still helps surface recovery. Without second-order reflection as well, channel prediction degrades while the geometry barely changes. First-order reflection possibly already constrains most planar surfaces, so second-order reflection, which follows the same mechanism, adds little geometric information, whereas it carries a considerable share of the received energy and therefore matters for channel prediction.

D.3 Primitive representation capacity

With the propagation mechanisms fixed, Figure D.2 varies representation capacity from 128 to 1,024 primitives on the six-room scene and evaluates the recovered geometry and channel predictions.

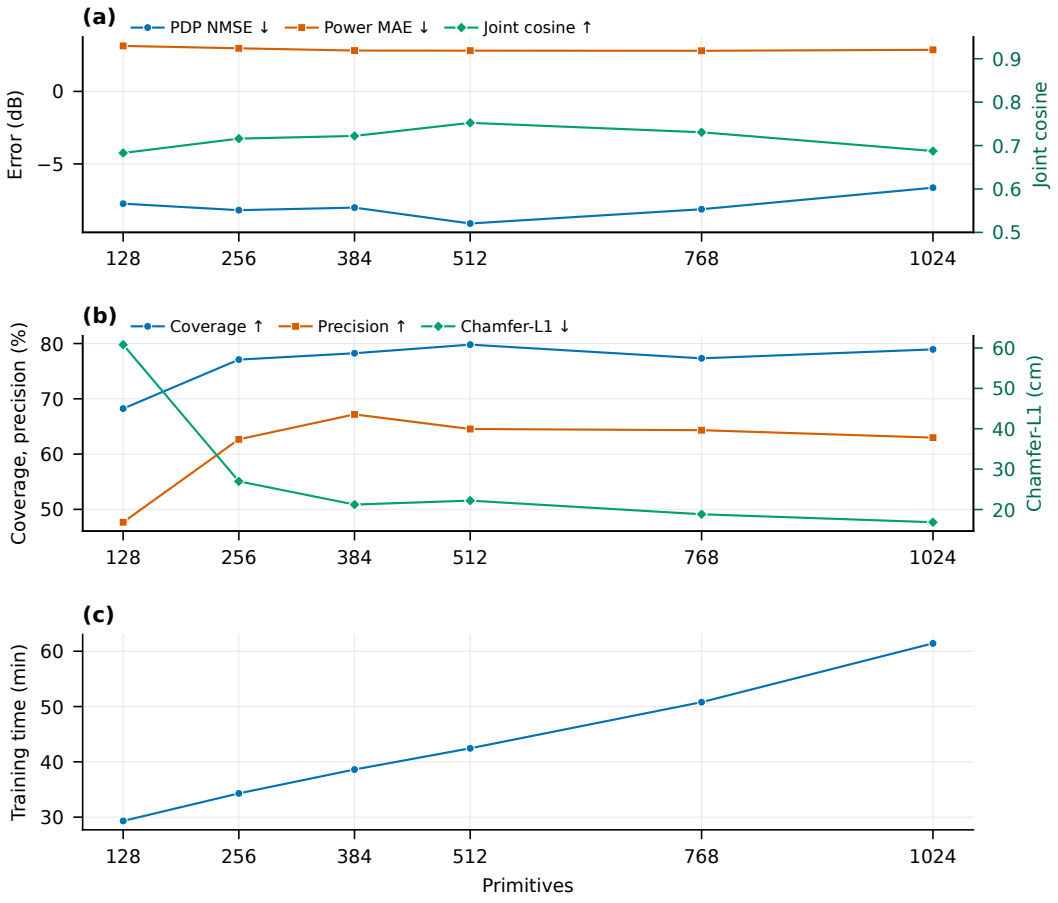


Figure D.2: **Primitive capacity on the six-room scene.** (a) Channel prediction metrics. (b) Geometry precision, coverage, and Chamfer-L_1. (c) Training time on a single GPU.

As shown in Figure D.2, more primitives mainly improve the recovered geometry. The Chamfer-L_1 distance falls as the number of primitives grows, coverage and precision level off beyond 256 primitives, and the channel metrics show no consistent trend. Training time grows with the number of primitives, by about a factor of 2.1 from 128 to 1,024.

D.4 Received-energy budget by propagation mechanism

Figure D.3 splits the received energy of each test link by propagation mechanism, following the interaction sequence of every traced path. The mechanisms GaussTwin models carry 86.7% of the received energy of a link on average on the city block and at least 99.1% indoors, where simulated paths stop after two interactions. On the city block, the remaining 13.3% travels on specular reflection followed by diffuse scattering and on third-order reflection, which carries 1.3% of the received energy. GaussTwin's diffuse energy there exceeds the ground truth by 10.1%, absorbing 76% of this remainder, while its direct and reflected energy stays within 1.8% of the ground truth. On the parallel-wall scene, the energy GaussTwin assigns to each mechanism deviates from the ground truth by less than 4%. On the six-room scene, second-order reflection carries 33.9% of the received energy and diffuse scattering 2.2%.

D.5 Measurement sparsity

We examine the effect of observation density on the parallel-wall scene by training on nested subsets containing 1%, 5%, 10%, 25%, 50%, and 100% of the training dataset, with the validation and test links unchanged.

As shown in Figure D.4, with only 5% of the training dataset, or 720 links, GaussTwin keeps channel prediction and coverage close to those of the full training set, while precision decreases gradually as

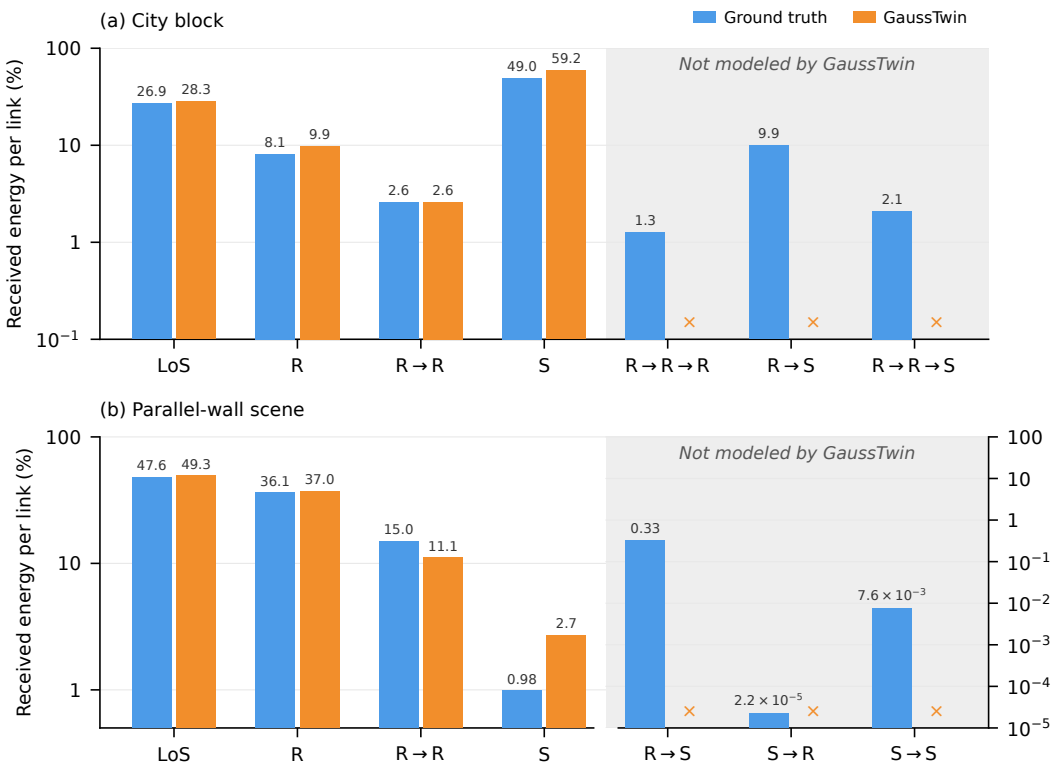


Figure D.3: **Received energy by propagation mechanism.** Percentage (with log scale) of each test link's received energy carried by each mechanism, averaged over the test links of (a) the city block and (b) the parallel-wall scene. R is specular reflection and S is diffuse scattering.

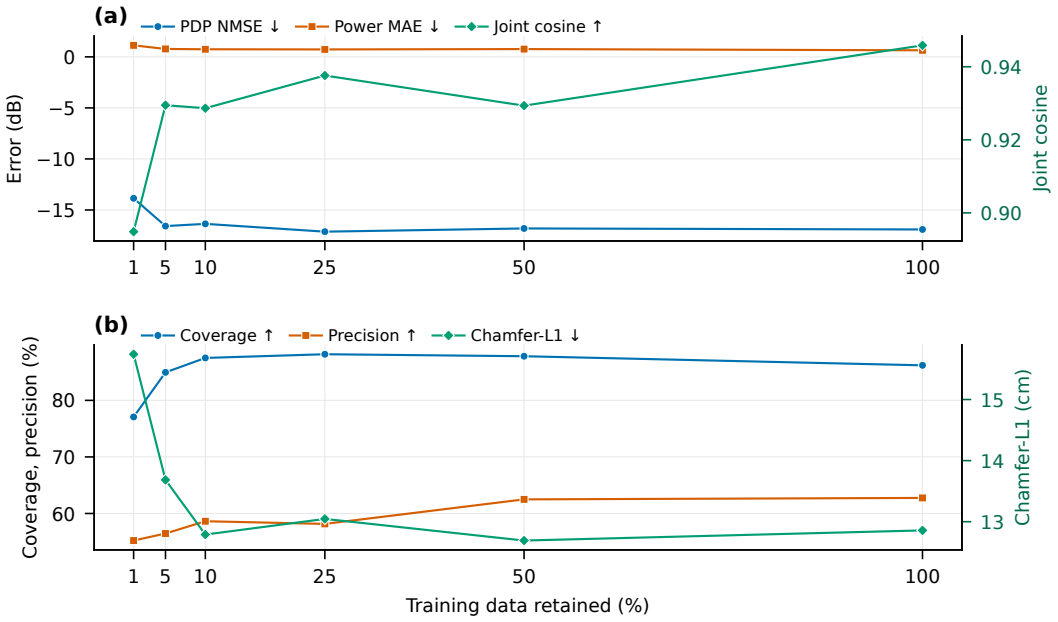


Figure D.4: **Channel prediction and surface recovery with less training data on the parallel-wall scene.** (a) Channel prediction metrics. (b) Geometry precision, coverage, and Chamfer-L_1.

data are removed. Even at 1%, or 144 links, it reaches a median PDP NMSE of -13.9 dB, although this is 3 dB above the full-data value and coverage falls by 9%.