

Physics-informed line-of-sight learning for scalable deterministic channel modeling

Xiucheng Wang

Xidian University

Junxi Huang

Xidian University

Conghao Zhou

Xidian University

Xuemin Shen

University of Waterloo

Nan Cheng

`dr.nan.cheng@ieee.org`

Xidian University

Article

Keywords: Deterministic channel modeling, radio map, line-of-sight, 6G native-AI

Posted Date: May 26th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9261533/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Physics-informed line-of-sight learning for scalable deterministic channel modeling

Xiucheng Wang¹, Junxi Huang¹, Conghao Zhou^{1*},
Xuemin Shen², Nan Cheng^{1*}

^{1*}School of Telecommunications Engineering, Xidian University, No. 2
South Taibai Road, Xi'an, 710071, Shaanxi, China.

²Department of Electrical and Computer Engineering, University of
Waterloo, 200 University Avenue West, Waterloo, N2L 3G1, Ontario,
Canada.

*Corresponding author(s). E-mail(s): conghao.zhou@ieee.org;
dr.nan.cheng@ieee.org;

Contributing authors: xwang_1@stu.xidian.edu.cn;
24012100067@stu.xidian.edu.cn; sshen@uwaterloo.ca;

Abstract

Deterministic channel modeling maps a physical environment to its site-specific electromagnetic response. Ray tracing produces complete multi-dimensional channel information but remains prohibitively expensive for area-wide deployment. We identify line-of-sight (LoS) region determination as the dominant bottleneck. To address this, we propose D²LoS, a physics-informed neural network that reformulates dense pixel-level LoS prediction into sparse vertex-level visibility classification and projection point regression, avoiding the spectral bias at sharp boundaries. A geometric post-processing step enforces hard physical constraints, yielding exact piecewise-linear boundaries. Because LoS computation depends only on building geometry, cross-band channel information is obtained by updating material parameters without retraining. We also construct RayVerse-100, a ray-level dataset spanning 100 urban scenarios with per-ray complex gain, angle, delay, and geometric trajectory. Evaluated against rigorous ray tracing ground truth, D²LoS achieves 3.28 dB mean absolute error in received power, 4.65° angular spread error, and 20.64 ns delay spread error, while accelerating visibility computation by over 25×.

Keywords: Deterministic channel modeling, radio map, line-of-sight, 6G native-AI.

1 Introduction

Deterministic channel modeling maps a physical environment to its exact electromagnetic response. It provides site-specific channel state information (CSI) for 6G network optimization [1], digital twin construction [2, 3], and physical-layer AI applications [4–6]. Unlike stochastic models that discard spatial structure by design [7–9], deterministic approaches preserve the gain, angle, delay, and geometric trajectory of every propagation path [10]. This level of detail is essential for coherent beamforming [11, 12], MIMO channel matrix construction [13], and beam management in dense urban deployments [14, 15]. However, generating such complete channel information at area-wide scale remains an open challenge. Ray tracing (RT) is the only method that produces multi-dimensional channel parameters in a single computation [10, 16], yet its cost is prohibitive for large-scale deployment [17]. This tension between information completeness and computational scalability motivates the search for efficient alternatives.

Existing learning-based approaches address only fragments of this problem [18–21]. Link-level methods predict point-to-point channel parameters from local features but cannot produce spatial coverage maps [22]. Area-level radio map (RM) construction methods provide coverage predictions, yet most focus on received power alone [23]. Recent efforts have explored angular and delay profile estimation, but each dimension requires a separately trained model at considerable cost. Meanwhile, available datasets offer either aggregate path-loss maps or dominant-path summaries [21, 24]. A few datasets provide per-path angle and delay parameters, but explicit geometric ray trajectories remain absent [24, 25]. This gap between what AI-driven network optimization demands and what current methods deliver remains unresolved.

We make two key observations that guide our approach. First, the computational bottleneck of RT lies not in the multi-bounce path search but in the preceding line-of-sight (LoS) preprocessing step [26, 27]. For a scenario with m building vertices, the classical rotational sweep algorithm requires $O(m \log m)$ operations per transmitter [28]. The subsequent path search operates on precomputed visibility adjacency via bounded-depth graph traversal, whose per-query cost is substantially lower [29, 30]. Second, in a 2D environment of polygonal obstacles, the LoS boundary seen from any point source is piecewise-linear. Each boundary segment is defined by a visible building vertex and its projection onto a farther building edge. These two observations suggest a natural decomposition: the network predicts sparse vertex-level attributes, and the sharp boundary is reconstructed analytically.

Here we introduce D²LoS¹, a physics-informed neural network that addresses the LoS bottleneck through vertex-level geometric decomposition. We also release RayVerse-100², a ray-level dataset spanning 100 diverse urban scenarios with per-ray complex gain, angle, delay, and complete geometric trajectory. The main contributions are as follows.

1. We reformulate LoS map prediction from dense pixel-level classification into sparse vertex-level visibility classification and projection point regression. This formulation

¹Code is available at <https://github.com/UNIC-Lab/D2LoS>.

²Dataset is available at <https://github.com/UNIC-Lab/RayVerse>.

avoids the spectral bias of neural networks at sharp LoS boundaries. Combined with a geometric post-processing step that enforces hard physical constraints, D²LoS reduces per-transmitter preprocessing complexity from $O(m \log m)$ to $O(M \log M + n)$, where M is the number of boundary vertices with $M \ll m$ and n is the number of evaluation points. In typical urban scenarios, $M \log M \ll n$, yielding an effective cost of $O(n)$.

2. Because LoS computation depends only on building geometry, it is inherently frequency-independent. Once geometric ray paths are determined, cross-band channel information is obtained by updating material parameters through uniform theory of diffraction (UTD) [31] formulations without retraining. This enables unified multi-dimensional radio map construction covering received signal strength (RSS), angular power spectrum (APS), and power delay profile (PDP) from a single pipeline.
3. We construct RayVerse-100, a ray-level dataset spanning 100 diverse urban scenarios. Each record captures per-ray complex gain, angle of arrival (AoA), angle of departure (AoD), propagation delay, and complete geometric trajectory with 3D reflection-point coordinates. To our knowledge, this is the first open dataset providing explicit per-ray geometric trajectories at this scale. We also provide post-processing interfaces for coherent superposition, MIMO channel matrix synthesis, and cross-band CSI generation.
4. Evaluated against rigorous RT ground truth on unseen test scenarios, D²LoS achieves 3.28 dB mean absolute error in RSS, 4.65° angular spread error, and 20.64 ns delay spread error, while accelerating visibility computation by over 25×.

2 Results

We evaluate D²LoS on the RayVerse-100 dataset against three baselines. No-Geom removes the geometric post-processing from D²LoS and uses raw neural network predictions directly. RadioUNet is a U-Net architecture that predicts pixel-level LoS maps. RMTransformer is a vision transformer variant designed for radio map estimation. All methods receive the same input and are trained on the same LoS corpus. The ground truth is produced by exact rotational sweep ray tracing with UTD field computation. For each evaluation point, the top-8 rays ranked by path gain are retained.

D²LoS achieves high-fidelity multi-dimensional radio maps

D²LoS consistently outperforms all baselines across three channel dimensions. In received signal strength (RSS) prediction, D²LoS achieves a mean absolute error (MAE) of 3.28 dB and a Pearson correlation of 0.954 averaged over 100 test scenarios (Table 1). The closest baseline, No-Geom, yields an MAE of 12.65 dB and a correlation of 0.647. RadioUNet and RMTransformer produce MAEs exceeding 32 dB with correlations below 0.18, indicating near-complete failure in spatial power prediction. The advantage extends to angular and delay domains. For the angular power spectrum (APS), D²LoS achieves an angular spread error of 4.65° and a shape cosine similarity of 0.892 (Table 2). For the power delay profile (PDP), D²LoS achieves a delay spread error of 20.64 ns and a K -factor error of 2.23 dB (Table 3). In both domains,

Table 1: Received signal strength prediction accuracy across 100 test scenarios. Six metrics are reported as mean $\pm$ standard deviation over all scenarios in RayVerse-100. D²LoS achieves the lowest error and highest correlation across all metrics. No-Geom denotes D²LoS without geometric post-processing. RadioUNet and RMTransformer are pixel-level LoS prediction baselines. Bold values indicate the best performance.

Metric	D ² LoS	No-Geom	RadioUNet	RMTransformer
Bias (dB)	-1.69 $\pm$ 3.82	-2.96 $\pm$ 11.76	-6.78 $\pm$ 26.30	-26.60 $\pm$ 28.39
MAE (dB)	3.28 $\pm$ 3.82	12.65 $\pm$ 10.01	32.80 $\pm$ 11.87	40.85 $\pm$ 15.98
RMSE (dB)	7.28 $\pm$ 4.57	19.92 $\pm$ 11.27	40.12 $\pm$ 11.38	48.11 $\pm$ 15.42
MSE (dB ²)	73.93 $\pm$ 174.32	523.75 $\pm$ 725.86	1738.76 $\pm$ 1120.59	2552.54 $\pm$ 1616.15
NMSE	0.0088 $\pm$ 0.0297	0.0826 $\pm$ 0.1481	0.2695 $\pm$ 0.2756	0.4789 $\pm$ 0.4912
Correlation	0.9537 $\pm$ 0.0603	0.6473 $\pm$ 0.3076	0.1350 $\pm$ 0.2824	0.1708 $\pm$ 0.3084

Table 2: Angular power spectrum prediction accuracy. Four metrics are reported as the mean over all evaluation points across 100 test scenarios. Angular spread (AS) absolute error and mean direction of arrival (MDoA) absolute error measure first-order angular statistics. Shape cosine and shape RMSE quantify the similarity between predicted and ground-truth APS profiles.

Metric	D ² LoS	No-Geom	RadioUNet	RMTransformer
AS Abs. Err. (deg)	4.65	11.80	22.91	23.37
MDoA Abs. Err. (deg)	17.07	37.72	70.09	74.64
Shape Cosine	0.8924	0.6667	0.3158	0.2902
Shape RMSE	0.0035	0.0090	0.0170	0.0177

RadioUNet and RMTransformer yield shape cosine values below 0.32. These low values indicate that their predicted angular and delay profiles bear little resemblance to the ground truth.

Fig. 1 provides a qualitative comparison across four representative scenarios. D²LoS reproduces both the LoS regions near the transmitter and the shadow boundaries behind buildings. No-Geom captures the coarse spatial pattern but produces blurred transitions at building edges. RadioUNet and RMTransformer fail to reconstruct meaningful spatial structures. RadioUNet generates near-uniform power distributions that lose all shadow boundary detail. RMTransformer produces narrow beam-like artifacts that do not correspond to any physical propagation pattern. These qualitative observations are consistent with the quantitative metrics in Tables 1–3.

Table 3: Power delay profile prediction accuracy. Six metrics are reported as the mean over all evaluation points across 100 test scenarios. Delay spread (DS) and median delay capture temporal dispersion. K -factor error reflects the dominant-to-scattered path ratio accuracy. Effective count measures the number of resolvable multipath components.

Metric	D^2 LoS	No-Geom	RadioUNet	RMTransformer
DS Abs. Err. (ns)	20.64	46.11	90.93	97.94
Median Delay Abs. Err. (ns)	47.91	126.67	245.72	237.56
K -factor Abs. Err. (dB)	2.23	4.61	8.77	9.41
Effective Count Abs. Err.	4.88	11.82	26.78	26.48
Shape Cosine	0.6434	0.4610	0.1743	0.1407
Shape RMSE	0.0327	0.0496	0.0743	0.0763

Geometric post-processing is critical for boundary-sensitive metrics

To isolate the contribution of geometric post-processing, we compare D^2 LoS with its ablated variant No-Geom. This variant uses raw neural network predictions without edge snapping or ray-edge intersection refinement. The impact of removing this step is substantial across all three channel dimensions. In RSS, the MAE increases from 3.28 to 12.65 dB and the correlation drops from 0.954 to 0.647 (Table 1). In the angular domain, the APS shape cosine decreases from 0.892 to 0.667, and the mean direction of arrival error grows from 17.07° to 37.72° (Table 2). In the delay domain, the degradation is even more pronounced. The PDP shape cosine drops from 0.643 to 0.461, and the median delay error increases from 47.91 to 126.67 ns (Table 3). These results confirm that geometric post-processing is not a marginal refinement but a critical component of the pipeline.

The qualitative comparisons reinforce these findings. In the APS curves (Fig. 2), No-Geom preserves the dominant arrival directions but introduces noticeable errors in secondary peaks. In the PDP curves (Fig. 3), No-Geom exhibits shifted peak positions and attenuated late-arriving components. These shifts indicate that boundary inaccuracies propagate into systematic path length errors in the downstream ray tracing. The delay domain shows a larger relative degradation than the angular domain. This asymmetry arises because a small boundary shift may preserve the approximate arrival angle of a reflected path while directly altering the total propagation distance. The resulting delay errors accumulate across multiple bounces. This pattern is consistent across all 100 test scenarios, as confirmed by the violin distributions in Fig. 5 and Fig. 6.

Pixel-level baselines fail at LoS boundary reconstruction

RadioUNet and RMTransformer both formulate LoS prediction as dense pixel-level image-to-image translation. Their performance collapses across all channel dimensions. In RSS prediction, RadioUNet achieves a correlation of only 0.135 and RMTransformer

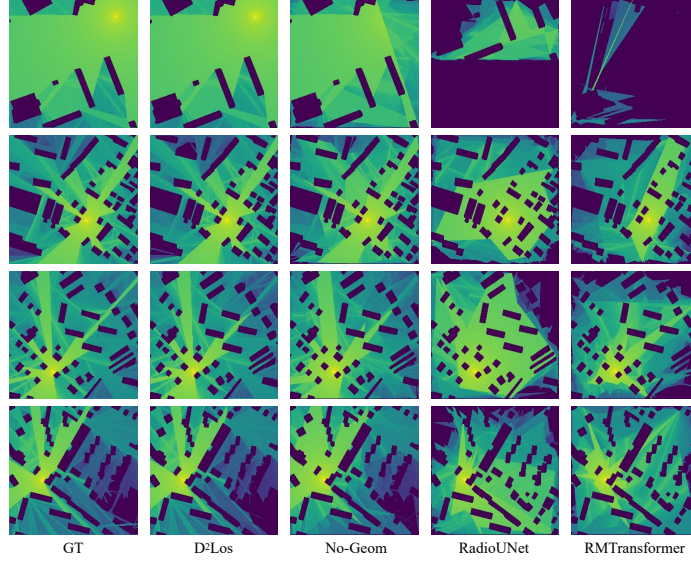


Fig. 1: Qualitative comparison of RSS radio maps across four representative scenarios. Each row corresponds to one scenario from the RayVerse-100 test set. Columns from left to right: ground truth (GT), D²LoS, No-Geom, RadioUNet [23], and RMTransformer [32]. D²LoS reproduces the spatial power distribution with high fidelity. No-Geom captures the coarse pattern but exhibits blurred boundaries. RadioUNet and RMTransformer fail to reconstruct meaningful spatial structures.

achieves 0.171 (Table 1). These values indicate that the predicted power maps are nearly uncorrelated with the ground truth. The violin distributions in Fig. 4 further reveal heavy-tailed error distributions with large inter-scenario variance. Some scenarios produce biases exceeding -50 dB, corresponding to order-of-magnitude power prediction errors. The failure is equally severe in the angular and delay domains. Both baselines yield APS shape cosine values below 0.32 (Table 2) and PDP shape cosine values below 0.18 (Table 3). These shape cosine values indicate that the predicted profiles are essentially random with respect to the ground truth.

The qualitative results reveal the underlying cause. The APS curves in Fig. 2 show noisy angular profiles with incorrect peak positions and power levels for both baselines. The PDP curves in Fig. 3 exhibit spurious peaks and incorrect power floors, particularly at large delays where reflected and diffracted paths dominate. The RSS maps in Fig. 1 further illustrate this failure mode. RadioUNet produces heavily smoothed maps that erase all shadow boundary detail. RMTransformer generates narrow beam-like artifacts unrelated to physical propagation. These patterns are consistent with the spectral bias of neural networks, which favors low-frequency functions during training. The resulting blurred LoS boundaries remove the sharp spatial transitions that define shadow regions in urban environments. This spectral bias is the fundamental reason why pixel-level formulations cannot reconstruct the fine-grained LoS structure required for accurate channel modeling.

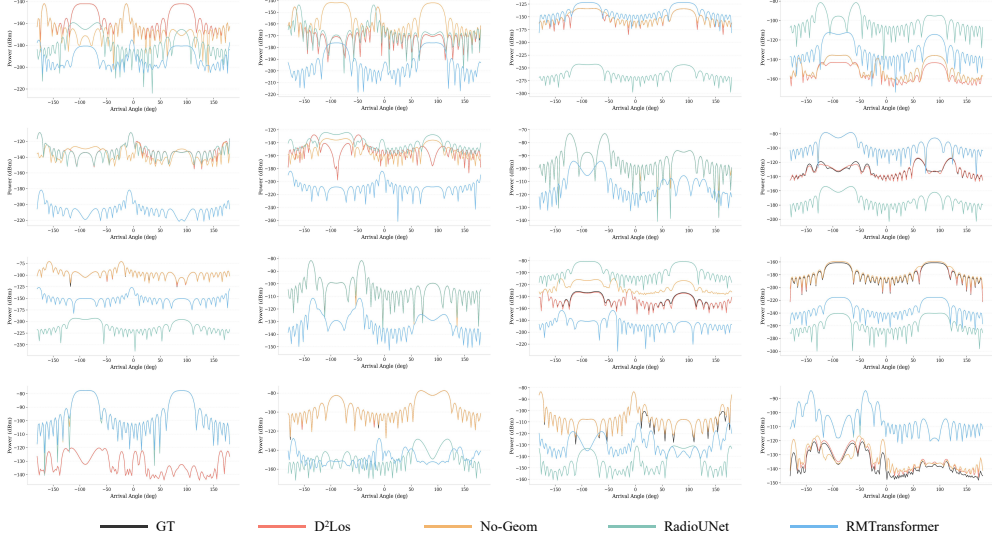


Fig. 2: Qualitative comparison of angular power spectra at selected receiver locations. Each row corresponds to one receiver position from a different test scenario. The horizontal axis is the azimuth arrival angle and the vertical axis is the received power in dBm. D²LoS closely tracks the ground truth. No-Geom preserves dominant arrival directions but introduces errors in secondary peaks. RadioUNet and RMTransformer produce noisy angular profiles with incorrect peak positions.

Per-ray angular information enables post-hoc beam pattern integration

A key advantage of ray-level output is that arbitrary transmit antenna patterns can be applied after the propagation simulation without re-execution. For each retained ray, the framework stores the AoD in both azimuth and elevation. A directional antenna gain $G_{\text{ant}}(\phi_{\text{AoD}}, \theta_{\text{AoD}})$ is then multiplied onto each ray's complex field before power summation. This produces beam-specific RSS, APS, and PDP maps from the same underlying ray data. We demonstrate this capability with a 15° half-power beamwidth pattern. Fig. 7 shows the resulting RSS maps for the same four scenarios as Fig. 1. The directional pattern concentrates energy along the boresight and suppresses off-axis multipath components. D²LoS accurately captures this directional filtering effect. It preserves both the beam footprint and the shadow structure behind buildings. No-Geom produces blurred beam edges due to inaccurate LoS boundaries. RadioUNet and RMTransformer fail to generate physically meaningful directional power distributions.

The quantitative evaluation in Fig. 8 confirms these observations across all three channel dimensions. D²LoS maintains compact, low-error distributions consistent with the omnidirectional results. The performance ranking among methods is preserved after beam integration. This indicates that applying a directional pattern does not amplify the relative errors introduced by LoS prediction inaccuracy. Additional results with 30° , 45° , and 60° beamwidths are provided in the Supplementary Information

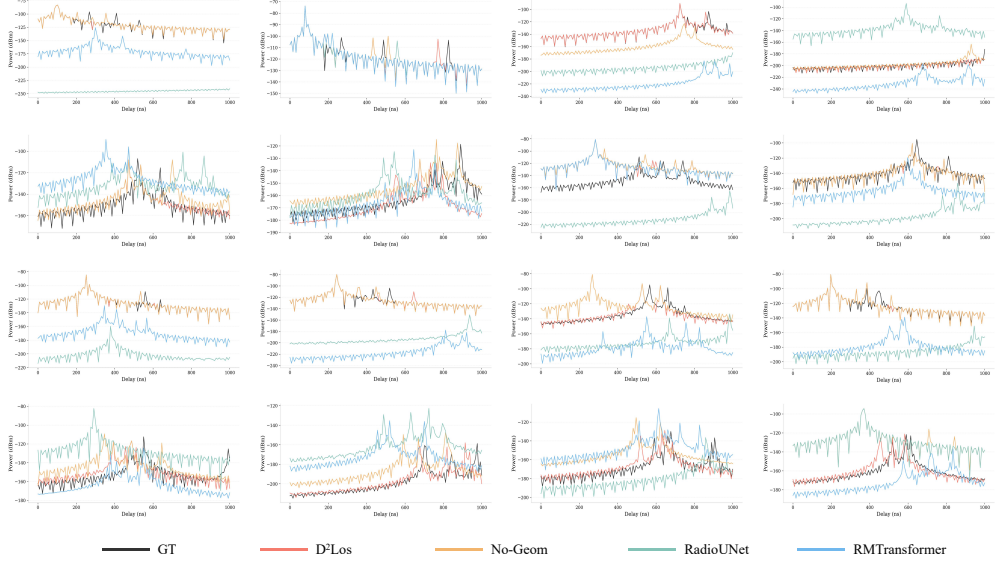


Fig. 3: Qualitative comparison of power delay profiles at selected receiver locations. Each row corresponds to one receiver position from a different test scenario. The horizontal axis is the propagation delay in nanoseconds and the vertical axis is the received power in dBm. D²LoS accurately reproduces the delay-domain structure. No-Geom exhibits shifted peak positions. RadioUNet and RMTransformer produce delay profiles with spurious peaks at large delays.

and show the same trend. This post-hoc integration capability is unavailable to conventional radio map methods that output only aggregate received power. Such methods do not preserve per-ray angular information and therefore cannot apply antenna patterns or evaluate beam-specific coverage after the fact. The ray-level output of D²LoS thus offers a practical advantage for beam management and antenna deployment planning in 6G systems.

D²LoS accelerates visibility computation

We compare the end-to-end processing time of D²LoS against the traditional RT pipeline. The traditional pipeline uses the rotational sweep algorithm for LoS preprocessing. Both pipelines share the same breadth-first search (BFS) path search [33] and UTD field computation backend. All experiments are conducted on a single NVIDIA RTX Pro 6000 GPU. Fig. 9a shows the per-scenario processing time for ten test maps. The traditional pipeline requires 15 to 48 min per scenario, while D²LoS completes the same computation in under 1 min. The speedup factors range from 25 $\times$ to 71 $\times$, with the variation attributable to differences in scenario complexity. Fig. 9b shows the cumulative distribution across all test scenarios. The median processing time is 0.5 min for D²LoS and 24.9 min for the traditional pipeline, representing a median speedup of approximately 50 $\times$.

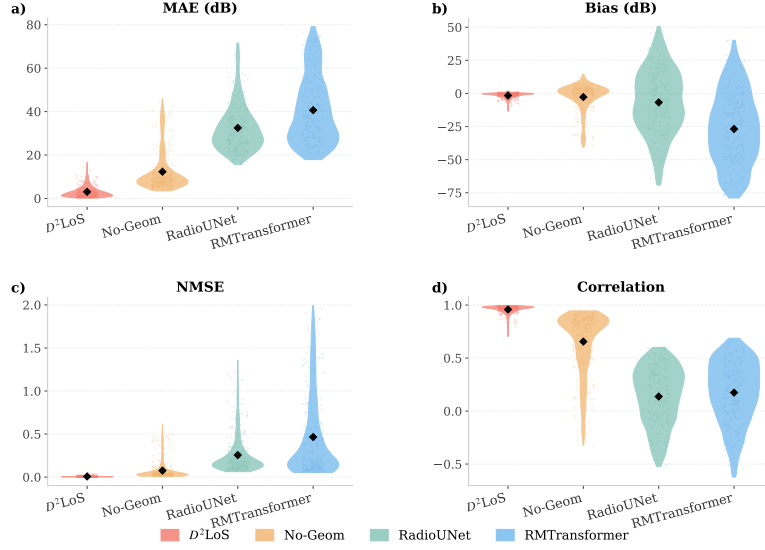


Fig. 4: Distribution of RSS prediction errors across 100 test scenarios. Violin plots show the per-scenario distribution of six metrics: **a)** MAE, **b)** RMSE, **c)** bias, **d)** MSE, **e)** NMSE, and **f)** Pearson correlation. Black diamonds indicate the mean. D²LoS exhibits both the lowest mean error and the smallest variance. RadioUNet and RMTransformer show wide distributions with heavy tails.

This acceleration is achieved without sacrificing output completeness. The D²LoS pipeline produces the same ray-level output format as the traditional pipeline. Each ray stores complex gain, AoA, AoD, propagation delay, and complete geometric trajectory with 3D reflection-point coordinates. Fig. 9c shows the cumulative distribution of the total ray count per scenario. The D²LoS distribution closely matches the ground truth, confirming that the learned LoS prediction preserves the ray-level completeness of the full RT pipeline. Each ray is fully characterized for downstream applications, including coherent superposition, MIMO channel matrix synthesis, and cross-band CSI generation. The combination of over 25 $\times$ speedup and full ray-level output makes D²LoS a practical drop-in replacement for the LoS preprocessing stage in existing RT workflows.

3 Discussion

The results demonstrate that accurate LoS prediction is sufficient to produce high-fidelity multi-dimensional radio maps. In the ray tracing pipeline, LoS determination serves as the gating step for all downstream computations. A correct LoS map ensures that both the direct path and the first-level vertices of the BFS visibility tree are correctly identified. When LoS boundaries are inaccurate, errors propagate multiplicatively through the multi-bounce path search, corrupting all channel dimensions simultaneously. This mechanism explains why geometric post-processing has a

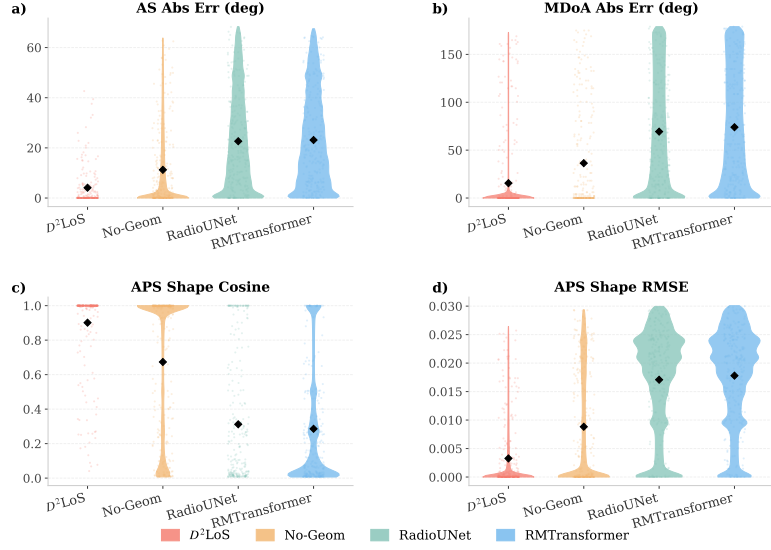


Fig. 5: Distribution of APS prediction errors across 100 test scenarios. Violin plots show: **a)** angular spread error, **b)** MDoA error, **c)** APS shape cosine, and **d)** APS shape RMSE. D²LoS achieves concentrated low-error distributions. The No-Geom variant shows moderate degradation. RadioUNet and RMTransformer produce near-random angular predictions.

stronger impact on delay-domain metrics than on angular-domain metrics: a small boundary shift may preserve the approximate arrival angle of a reflected path, but it directly alters the total propagation distance, producing systematic delay errors that accumulate across multiple bounces.

The complete failure of RadioUNet and RMTransformer highlights a fundamental limitation of pixel-level LoS prediction. Neural networks exhibit spectral bias, favoring low-frequency functions during training. The resulting blurred LoS boundaries remove the sharp shadow edges that define urban multipath structure. D²LoS circumvents this by predicting sparse vertex attributes instead of dense pixel maps. Binary visibility labels and approximate projection coordinates are both low-frequency targets. The sharp LoS boundary is then reconstructed analytically, preserving the geometric precision that downstream ray tracing requires.

A key advantage of the LoS-based decomposition is its inherent frequency independence. The LoS computation depends only on building geometry, not on carrier frequency. Frequency enters solely through the UTD interaction coefficients, which can be updated without re-executing visibility computation or path search. This distinguishes D²LoS from existing radio map methods that require retraining for each frequency band. Furthermore, the per-ray angular output enables post-hoc antenna pattern integration, as demonstrated by the beam-specific results across 15°, 30°, 45°, and 60° beamwidths. The ability to evaluate arbitrary beam patterns from a single

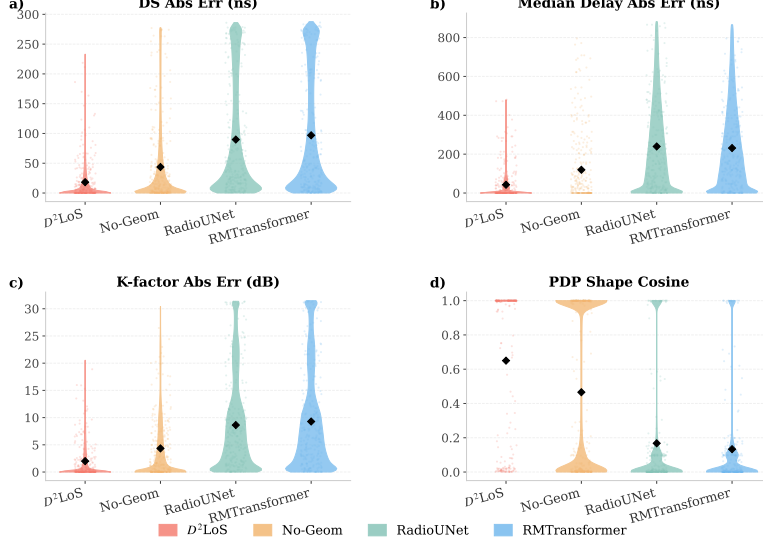


Fig. 6: Distribution of PDP prediction errors across 100 test scenarios. Violin plots show: a) delay spread error, b) median delay error, c) K -factor error, d) effective count error, e) PDP shape cosine, and f) PDP shape RMSE. D²LoS maintains compact error distributions. The gap between D²LoS and No-Geom is larger in the delay domain than in the angular domain.

ray tracing computation provides substantial time savings for antenna deployment planning in multi-band 6G systems.

The RayVerse-100 dataset provides ray-level ground truth that existing datasets do not offer. Neither RadioMapSeer nor DeepMIMO exposes explicit per-ray geometric trajectories with 3D reflection-point coordinates. This trajectory information is essential for environment-aware beam tracking, RIS placement optimization, and radar-communication co-design.

Several limitations should be noted. The 2.5D building model introduces errors for complex rooftop geometries. The 2D $\times$ 2D decomposition excludes over-rooftop diffraction paths. The evaluation is conducted at street-level height (1.5 m) and retains only the top-8 rays per point. The geometric post-processing uses a fixed search radius that may benefit from adaptive scaling. Future work will extend the building model, incorporate over-rooftop diffraction, and integrate measured channel data through transfer learning to bridge the gap between simulation and real-world deployment.

4 Methods

Environment representation and problem formulation

The propagation environment is encoded as a top-down binary occupancy map $E \in \{0, 1\}^{H \times W}$ with $H = W = 257$ pixels at a spatial resolution of $\Delta r = 1$ m.

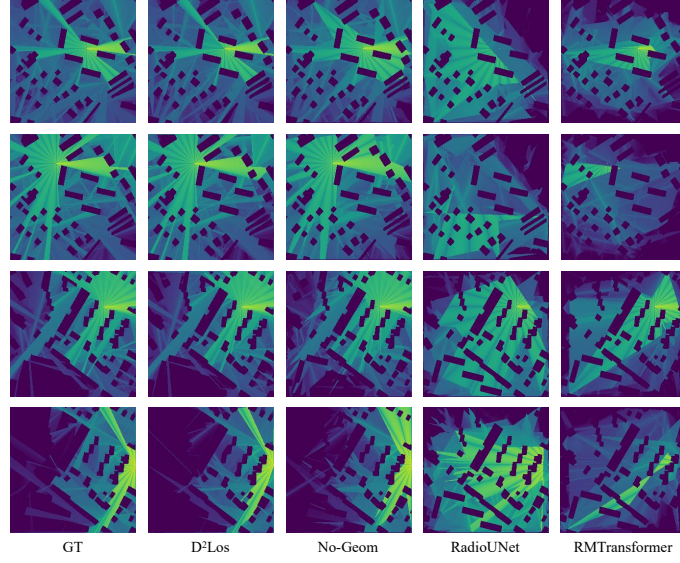


Fig. 7: RSS radio maps with a 15° directional beam pattern. The same scenarios as Fig. 1 are shown after post-hoc beam integration. D²LoS accurately captures the directional filtering effect, preserving both the beam footprint and shadow structure. RadioUNet and RMTransformer fail to produce meaningful directional patterns.

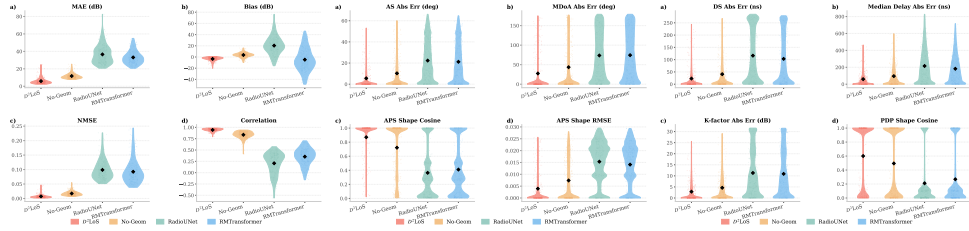


Fig. 8: Prediction accuracy with a 15° half-power beamwidth antenna pattern. Violin plots show per-scenario error distributions after beam pattern integration using per-ray AoD. Left: RSS metrics. Center: APS metrics. Right: PDP metrics. D²LoS maintains compact, low-error distributions consistent with the omnidirectional results. Additional beamwidths (30° , 45° , 60°) are provided in the Supplementary Information.

Building footprints are extracted from OpenStreetMap (OSM) and approximated as quadrilateral prisms. The vertices of each footprint polygon serve as candidate reflection and diffraction points. Per-building heights are obtained from OSM attributes when available; a default value of 20 m is assigned otherwise.

Given a transmitter at $\mathbf{p}_s = (x_s, y_s)$, the goal is to determine a binary LoS map $L \in \{0, 1\}^{H \times W}$, where $L(x_r, y_r) = 1$ indicates that pixel (x_r, y_r) is visible from $\mathbf{p}_s$. The network input is a $257 \times 257 \times 4$ tensor comprising: the binary occupancy map

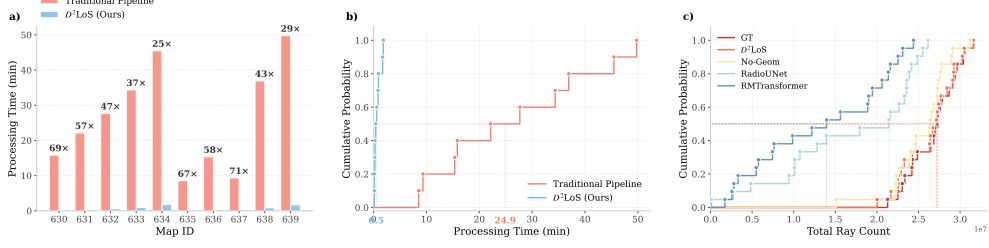


Fig. 9: Computational efficiency comparison between D²LoS and the traditional RT pipeline. **a)** Per-scenario processing time for ten test maps. Speedup factors range from 25 $\times$ to 71 $\times$. **b)** Cumulative distribution of processing time. The median is 0.5 min for D²LoS and 24.9 min for the traditional pipeline. **c)** Cumulative distribution of the total ray count per scenario. All experiments use a single NVIDIA RTX Pro 6000 GPU.

E , a Gaussian-blurred transmitter location heatmap H_{Tx} with $\sigma = 3$ pixels, and two spatial coordinate grids X_{coord} and Y_{coord} .

The ray tracing pipeline is decomposed into two stages (Fig. 10): LoS preprocessing, which determines mutual visibility among all geometric primitives, and BFS multi-bounce path search, which enumerates valid propagation paths. D²LoS replaces the LoS preprocessing stage while preserving the standard path search and UTD field computation backend.

Vertex-level LoS reformulation

In a 2D environment composed of polygonal obstacles, the LoS region boundary as seen from a point source $\mathbf{p}_s$ is piecewise-linear. Each linear segment is defined by a building vertex $\mathbf{v}_i$ visible from $\mathbf{p}_s$ and its projection point $\mathbf{v}_i^{proj}$ on a more distant building edge. The projection point is computed via ray casting: a ray from $\mathbf{p}_s$ passes through $\mathbf{v}_i$ and extends until it first intersects a building edge e_j that is farther from $\mathbf{p}_s$ and not parallel to the ray:

$$\mathbf{v}_i^{proj} = \mathbf{p}_s + t_i^* \cdot (\mathbf{v}_i - \mathbf{p}_s), \quad t_i^* = \min_j \{t > 1 \mid (\mathbf{p}_s + t \cdot \mathbf{d}_i) \in e_j, \mathbf{d}_i \times \mathbf{n}_{e_j} \neq 0\}, \quad (1)$$

where $\mathbf{d}_i = \mathbf{v}_i - \mathbf{p}_s$ and $\mathbf{n}_{e_j}$ is the direction vector of edge e_j .

Two prediction targets are defined for each building vertex $\mathbf{v}_i$: a binary visibility label $c_i \in \{0, 1\}$ indicating whether $\mathbf{v}_i$ is visible from $\mathbf{p}_s$, and the projection point coordinates $\hat{\mathbf{v}}_i^{proj} = (\hat{x}_i^{proj}, \hat{y}_i^{proj})$, defaulting to $\mathbf{v}_i$ for non-boundary vertices. All M building vertices are predicted simultaneously, avoiding the need to pre-select which vertices determine the LoS boundary. The network predicts low-frequency attributes rather than a dense high-frequency boundary map. The sharp LoS boundary is reconstructed analytically from the predicted vertex attributes in the post-processing stage.

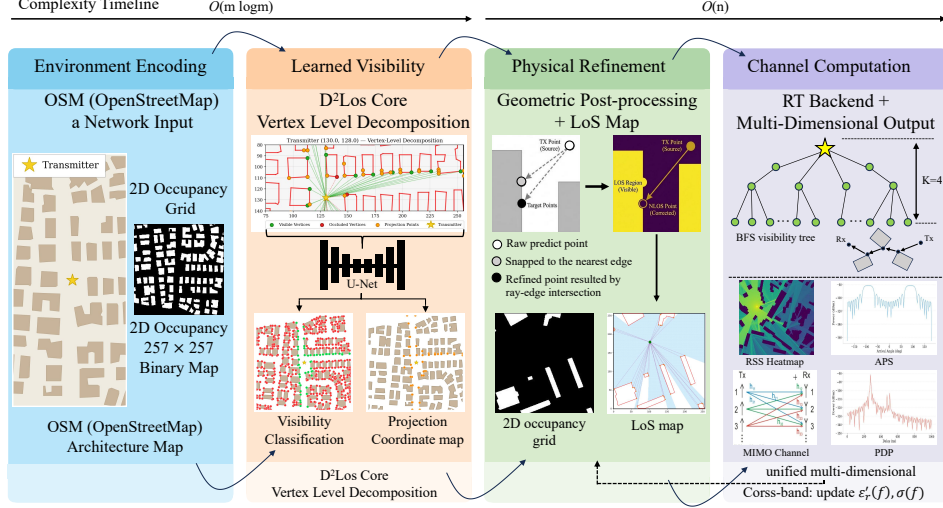


Fig. 10: Overview of the D²LoS framework. The pipeline takes a binary occupancy map and transmitter position as input. The D²LoS network predicts vertex-level visibility and projection coordinates. Geometric post-processing enforces physical constraints to produce an exact LoS map. A BFS path search and UTD field computation yield per-ray channel parameters. Cross-band information is obtained by updating material parameters without re-executing visibility computation.

Network architecture and loss function

We employ a U-Net encoder-decoder with two modifications at the bottleneck layer. Atrous spatial pyramid pooling (ASPP) applies dilated convolutions with rates $\{1, 6, 12, 18\}$ to capture occlusion effects at multiple spatial scales. Spectral feature modulation follows the Fourier neural operator (FNO) paradigm: spatial features are transformed to the frequency domain via real FFT, modulated by learned complex-valued weights, and transformed back via inverse FFT. This provides a global receptive field in a single layer and captures long-range shadow connectivity that local convolutions cannot efficiently model.

The decoder produces two output heads via sub-pixel upsampling with coordinate attention: a visibility classification map $\hat{M} \in [0, 1]^{H \times W}$ and normalised projection coordinates $\hat{P} \in [0, 1]^{H \times W \times 2}$. During inference, predictions are extracted at known building vertex pixel locations, effectively implementing sparse vertex-level prediction through a dense network architecture.

The composite loss function is

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{Focal}} + \lambda_1 \mathcal{L}_{\text{Dice}} + \lambda_2 \mathcal{L}_{\text{MaskedL1}}. \quad (2)$$

Focal loss with $\gamma = 2.0$ addresses class imbalance between visible and occluded vertices. Dice loss with $\lambda_1 = 1.0$ enforces spatial connectivity of predicted LoS regions. The

coordinate regression head uses a masked L1 loss computed only at building-vertex pixels:

$$\mathcal{L}_{\text{MaskedL1}} = \frac{\sum \|\hat{P} - P^{\text{GT}}\|_1 \cdot \mathbf{1}(M^{\text{GT}} > 0.5)}{\sum \mathbf{1}(M^{\text{GT}} > 0.5)}, \quad (3)$$

ensuring that representational capacity concentrates on geometrically meaningful boundary regions. We set $\lambda_2 = 20.0$ based on hyperparameter search.

Geometric post-processing and 3D extension

Raw neural network predictions contain residual regression errors in the projection coordinates. We exploit the physical constraint that projection points must lie on building edges through a two-step refinement.

In Step 1, for each predicted projection point $\hat{\mathbf{v}}_i^{\text{proj}}$, the nearest building edge e_j^* within a search radius $R_{\text{search}} = 5$ m is identified. In Step 2, the intersection of the ray from $\mathbf{p}_s$ through vertex $\mathbf{v}_i$ with edge e_j^* is computed:

$$\mathbf{v}_i^{\text{refined}} = \mathbf{p}_s + t^*(\mathbf{v}_i - \mathbf{p}_s), \quad t^* = \frac{(\mathbf{a} - \mathbf{p}_s) \times (\mathbf{b} - \mathbf{a})}{(\mathbf{v}_i - \mathbf{p}_s) \times (\mathbf{b} - \mathbf{a})}, \quad (4)$$

where $\mathbf{a}$ and $\mathbf{b}$ are the endpoints of e_j^* . The refined point is accepted if $t^* > 1$ and the intersection lies within the edge segment; otherwise, the raw prediction is retained. A global path-clearance check verifies that the refined ray does not intersect other buildings.

The complete LoS map is reconstructed by selecting boundary vertices $V_{\text{bnd}} = \{\mathbf{v}_i \mid c_i = 1, \mathbf{v}_i^{\text{refined}} \neq \mathbf{v}_i\}$, constructing segments $[\mathbf{v}_i, \mathbf{v}_i^{\text{refined}}]$, sorting them angularly with respect to $\mathbf{p}_s$, taking the union of unobstructed angular sectors, and rasterising onto the grid to produce $\hat{L} \in \{0, 1\}^{H \times W}$.

To extend to 3D visibility, we adopt the 2D $\times$ 2D decomposition:

$$\text{LoS}_{3\text{D}}(\mathbf{p}_s, \mathbf{p}_r) \approx \text{LoS}_{2\text{D-H}}(\mathbf{p}_s, \mathbf{p}_r) \wedge \text{LoS}_{2\text{D-V}}(\mathbf{p}_s, \mathbf{p}_r). \quad (5)$$

The horizontal check $\text{LoS}_{2\text{D-H}}$ is computed by D^2LoS . The vertical check $\text{LoS}_{2\text{D-V}}$ evaluates whether the elevation angle of the direct path clears the heights of all intervening buildings. Under the 2.5D extruded prism model, this decomposition is mathematically exact (Supplementary Theorem S3). The transmitter and receiver heights are input parameters; the current evaluation uses $h_{\text{tx}} = h_{\text{rx}} = 1.5$ m. Over-rooftop diffraction paths are excluded from the current framework, as discussed in the limitations.

Ray tracing backend

With LoS relationships determined, an N -ary visibility tree is constructed rooted at the transmitter. Vertex-to-vertex LoS adjacency is precomputed once per scenario using the exact rotational sweep algorithm. D^2LoS provides transmitter-to-vertex visibility, forming the first level of the tree. Multi-bounce paths are enumerated via breadth-first search (BFS) with a maximum depth of $K = 4$, covering up to four combined reflections and diffractions. The search is executed with GPU-parallel graph traversal.

For each valid path, the complex electric field is computed as the product of interaction coefficients and free-space propagation factors along all segments. Interaction coefficients include Fresnel reflection coefficients and UTD diffraction coefficients following the Luebbers lossy-dielectric extension. Two numerical stability mechanisms are applied: an envelope clamp ensuring that diffracted field magnitude does not exceed half the incident field, and a forward-scatter smoothing for deflection angles below 30° (Supplementary Theorem S2 and Supplementary Remark 1). Default material parameters at the reference frequency of 3.5 GHz are $\epsilon'_r = 5.31$ and $\sigma = 0.0326$ S/m.

At each evaluation point, all arriving rays are ranked by path gain and the top $N_{\text{ray}} = 8$ rays are retained, capturing over 98% of the total received energy across all scenarios. Each retained ray stores: complex gain, AoA and AoD in azimuth and elevation, propagation delay, complete geometric trajectory with 3D interaction-point coordinates, interaction type at each point, and material index of each interacting surface. The RSS output is computed via incoherent power summation over all retained rays. Because per-ray phase information is fully preserved, coherent superposition is available through the post-processing interface.

The geometric ray paths are frequency-independent. To obtain channel information at a different carrier frequency, users specify the target frequency and corresponding material parameters $\epsilon'_r(f)$ and $\sigma(f)$. The framework re-evaluates all interaction coefficients along the stored geometric trajectories without re-executing visibility computation or path search. Similar to [34], the narrowband MIMO channel matrix for an $N_t \times N_r$ array system is

$$\mathbf{H}(f) = \sum_{p=1}^{N_{\text{ray}}} \alpha_p \mathbf{a}_R(\theta_{R,p}, \phi_{R,p}) \mathbf{a}_T^H(\theta_{T,p}, \phi_{T,p}) e^{-j2\pi f \tau_p}, \quad (6)$$

where α_p is the complex path gain, τ_p is the delay, and $\mathbf{a}_R$, $\mathbf{a}_T$ are the receive and transmit array steering vectors.

Dataset construction and training configuration

To train D²LoS, we extracted 701 urban topographies from OSM covering diverse building densities and street layouts. Each environment is rasterised into a 257×257 grid. For each scenario, multiple transmitter positions are sampled and the ground-truth LoS map and projection point coordinates are computed using the exact rotational sweep algorithm. The 701 scenarios are strictly partitioned: scenarios 0–630 for training and validation, and scenarios 631–700 as a completely unseen test set.

To provide ray-level ground truth for end-to-end evaluation, we constructed RayVerse-100 by selecting 100 scenarios from geographically distinct regions. For each scenario, the full RT pipeline is executed: exact LoS computation, BFS path enumeration with $K = 4$, and UTD field computation at every grid point. The output contains the top-8 per-ray multipath parameters. The RayVerse-100 scenarios are drawn from the test partition to ensure no overlap with D²LoS training data. Built on UTD, the dataset exposes open material parameter interfaces for cross-band CSI generation.

Post-processing interfaces for coherent superposition, MIMO channel matrix synthesis, and antenna pattern integration are provided.

D²LoS is trained on NVIDIA RTX 4090 and A100 GPUs using AdamW with a cosine annealing learning rate schedule. The initial learning rate is 1.8×10^{-3} , the batch size is 64, and training runs for 120 epochs. All metrics reported in Tables 1–3 are computed as mean $\pm$ standard deviation over the 100 independent test scenarios. No scenario in the test set overlaps with the training data geographically or topologically.

Declarations

Founding Information: This work was supported by the National Key Research and Development Program of China (2024YFB907500), and the National Natural Science Foundation of China (62571402).

Data availability: The RayVerse-100 dataset generated and analyzed during this study is publicly available at <https://github.com/UNIC-Lab/RayVerse>. The urban building footprints used for environmental construction were obtained from OpenStreetMap (<https://www.openstreetmap.org>) under the Open Database License. All data needed to evaluate the conclusions in this paper are present in the paper and the Supplementary Information.

Code availability: Data is available at <https://github.com/UNIC-Lab/D2LoS>.

Author contribution: X.W. conceived the idea, designed the D²LoS framework, implemented the algorithm, conducted all experiments, constructed the RayVerse-100 dataset, and wrote the main manuscript text. J.H. assisted with dataset construction, baseline implementation, and figure preparation. C.Z. co-supervised the research, contributed to the methodology design and theoretical analysis, and revised the manuscript. X.S. provided guidance on the system design and revised the manuscript. N.C. supervised the overall research, contributed to the problem formulation and research direction, and revised the manuscript. All authors reviewed the manuscript.

References

- [1] Zheng, Y., Chen, J.: A radio map approach for reduced pilot csi tracking in massive mimo networks. *IEEE Transactions on Signal Processing* (2025)
- [2] Wang, X., Zheng, P., Cheng, N., Sun, R., Chen, J., Tao, K., Yin, Z., Liu, Z., Zeng, Y.: Radiodiff-turbo: Lightweight generative large electromagnetic model for wireless digital twin construction. In: *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6 (2025). <https://doi.org/10.1109/INFOCOMWKSHPS65812.2025.11152929>
- [3] Wang, X., Zhang, Q., Cheng, N., Sun, R., Li, Z., Cui, S., Shen, X.: Radiodiff- k^2 : Helmholtz equation informed generative diffusion model for multi-path aware radio map construction. *IEEE J. Sel. Areas Commun.* **44**, 2318–2333 (2026) <https://doi.org/10.1109/JSAC.2025.3641105>

- [4] Zeng, Y., Chen, J., Xu, J., Wu, D., Xu, X., Jin, S., Gao, X., Gesbert, D., Cui, S., Zhang, R.: A tutorial on environment-aware communications via channel knowledge map for 6G. *IEEE Commun. Surveys Tuts.* **26**(3), 1478–1519 (2024)
- [5] Wang, Z., Zhang, J., Du, H., Niyato, D., Cui, S., Ai, B., Debbah, M., Letaief, K.B., Poor, H.V.: A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications. *IEEE Commun. Surveys Tuts.* **26**(3), 1560–1605 (2024)
- [6] Liu, B., Gao, S., Liu, X., Cheng, X., Yang, L.: Wifo: Wireless foundation model for channel prediction. *Science China Information Sciences* **68**(6), 162302 (2025)
- [7] Eyceoz, T., Duel-Hallen, A., Hallen, H.: Deterministic channel modeling and long range prediction of fast fading mobile radio channels. *IEEE Communications Letters* **2**(9), 254–256 (2002)
- [8] Khuwaja, A.A., Chen, Y., Zhao, N., Alouini, M.-S., Dobbins, P.: A survey of channel modeling for uav communications. *IEEE Communications Surveys & Tutorials* **20**(4), 2804–2821 (2018)
- [9] Kammoun, A., Khanfir, H., Altman, Z., Debbah, M., Kamoun, M.: Preliminary results on 3d channel modeling: From theory to standardization. *IEEE Journal on Selected Areas in Communications* **32**(6), 1219–1229 (2014)
- [10] Deschamps, G.A.: Ray techniques in electromagnetics. *Proc. IEEE* **60**(9), 1022–1035 (1972)
- [11] Tang, J., Bian, D., Li, G., Hu, J., Cheng, J.: Optimization Method of Dynamic Beam Position for LEO Beam-Hopping Satellite Communication Systems. *IEEE Access* **9**, 57578–57588 (2021)
- [12] Ha, T.N., Romero, D., López-Valcarce, R.: Radio maps for beam alignment in mmwave communications with location uncertainty. In: 2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring), pp. 1–7 (2024). IEEE
- [13] Oh, S., Myung, N.-H.: MIMO channel estimation method using ray-tracing propagation model. *Electronics Lett.* **40**(21), 1 (2004)
- [14] Chen, W., Chen, J.: Learning blockage and reflection geometry for mimo beam map construction. In: ICC 2025-IEEE International Conference on Communications, pp. 2321–2326 (2025). IEEE
- [15] Hoffmann, M., Kryszkiewicz, P.: Beam management driven by radio environment maps in o-ran architecture. In: 2023 IEEE International Conference on Communications Workshops (ICC Workshops), pp. 54–59 (2023). IEEE
- [16] Hoydis, J., Aït Aoudia, F., Cammerer, S., Nimier-David, M., Binder, N., Marcus,

- G., Keller, A.: Sionna rt: Differentiable ray tracing for radio propagation modeling. In: 2023 IEEE Globecom Workshops (GC Wkshps), pp. 317–321 (2023). IEEE
- [17] He, D., Ai, B., Guan, K., Wang, L., Zhong, Z., Kürner, T.: The design and applications of high-performance ray-tracing simulation platform for 5g and beyond wireless communications: A tutorial. *IEEE communications surveys & tutorials* **21**(1), 10–27 (2018)
 - [18] Wang, X., Tao, K., Cheng, N., Yin, Z., Li, Z., Zhang, Y., Shen, X.: RadioDiff: An effective generative diffusion model for sampling-free dynamic radio map construction. *IEEE Trans. Cognit. Commun. Networking*, Early Access, 1–13 (2024)
 - [19] Zhang, S., Wijesinghe, A., Ding, Z.: Rme-gan: A learning framework for radio map estimation based on conditional generative adversarial network. *IEEE Internet Things J.* **10**(20), 18016–18027 (2023)
 - [20] Wang, X., Yuan, T., Cao, Y., Cheng, N., Sun, R., Zhuang, W.: iradiodiff: Physics-informed diffusion model for indoor radio map construction and localization. In: ICC 2026 - IEEE International Conference on Communications, pp. 1–6 (2026)
 - [21] Xiucheng Wang, N.C. Yuhao Pan: A tutorial on learning-based radio map construction: Data, paradigms, and physics-awareness. *arXiv preprint arXiv:2603.17499* (2026)
 - [22] Guo, J., Wen, C.-K., Jin, S., Li, G.Y.: Overview of deep learning-based csi feedback in massive mimo systems. *IEEE Transactions on Communications* **70**(12), 8017–8045 (2022)
 - [23] Levie, R., Yapar, Ç., Kutyniok, G., Caire, G.: RadioUNet: Fast radio map estimation with convolutional neural networks. *IEEE Trans. Wireless Commun.* **20**(6), 4001–4015 (2021)
 - [24] Wang, X., Zhang, Q., Cheng, N., Chen, J., Zhang, Z., Li, Z., Cui, S., Shen, X.: Radiodiff-3d: A 3d×3d radio map dataset and generative diffusion based benchmark for 6g environment-aware communication. *IEEE Trans. Netw. Sci. Eng.* **13**, 3773–3789 (2026) <https://doi.org/10.1109/TNSE.2025.3590545>
 - [25] Yapar, Ç., Levie, R., Kutyniok, G., Caire, G.: Real-time outdoor localization using radio maps: A deep learning approach. *IEEE Transactions on Wireless Communications* **22**(12), 9703–9717 (2023) <https://doi.org/10.1109/TWC.2023.3273202>
 - [26] Wolffe, G., Gschwendtner, B.E., Landstorfer, F.M.: Intelligent ray tracing—a new approach for field strength prediction in microcells. In: 1997 IEEE 47th Vehicular Technology Conference. Technology in Motion, vol. 2, pp. 790–7942 (1997). <https://doi.org/10.1109/VTC97.583942>

- [27] Long, K., Wang, W., Wu, Y., Liu, Y.: A fast ray tracing algorithm for urban areas. In: 2022 IEEE 10th Asia-Pacific Conference on Antennas and Propagation (APCAP), pp. 1–2 (2022). <https://doi.org/10.1109/APCAP56600.2022.10069838>
- [28] Choi, H., Oh, J., Chung, J., Alexandropoulos, G.C., Choi, J.: Withray: A versatile ray-tracing simulator for smart wireless environments. *IEEE Access* **11**, 56822–56845 (2023)
- [29] Valenzuela, R.: A ray tracing approach to predicting indoor wireless transmission. In: IEEE 43rd Vehicular Technology Conference, pp. 214–218 (1993). IEEE
- [30] Yang, C.-F., Wu, B.-C., Ko, C.-J.: A ray-tracing method for modeling indoor wave propagation and penetration. *IEEE transactions on Antennas and Propagation* **46**(6), 907–919 (1998)
- [31] Tsingos, N., Funkhouser, T., Ngan, A., Carlbom, I.: Modeling acoustics in virtual environments using the uniform theory of diffraction. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, pp. 545–552 (2001)
- [32] Li, Y., Zhang, C., Wang, W., Huang, Y.: Rmtransformer: Accurate radio map construction and coverage prediction. In: 2025 IEEE 101st Vehicular Technology Conference (VTC2025-Spring), pp. 1–5 (2025). IEEE
- [33] Beamer, S., Asanović, K., Patterson, D.: Direction-optimizing breadth-first search. *Scientific Programming* **21**(3-4), 137–148 (2013)
- [34] Alkhateeb, A.: Deepmimo: A generic deep learning dataset for millimeter wave and massive mimo applications. arXiv preprint arXiv:1902.06435 (2019)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [RayVerseAppendix.pdf](#)