

# RadioDiff- $k^2$ : Helmholtz Equation Informed Generative Diffusion Model for Multi-Path Aware Radio Map Construction

Xiucheng Wang<sup>1</sup>, Graduate Student Member, IEEE, Qiming Zhang<sup>2</sup>, Nan Cheng<sup>1</sup>, Senior Member, IEEE, Ruijin Sun<sup>1</sup>, Member, IEEE, Zan Li<sup>1</sup>, Fellow, IEEE, Shuguang Cui<sup>1</sup>, Fellow, IEEE, and Xuemin Shen<sup>3</sup>, Fellow, IEEE

**Abstract**—In this paper, we propose a novel physics-informed generative learning approach, named RadioDiff- $k^2$ , for accurate and efficient multipath-aware radio map (RM) construction. As future wireless communication evolves towards environment-aware paradigms, the accurate construction of RMs becomes crucial yet highly challenging. Conventional electromagnetic (EM)-based methods, such as full-wave solvers and ray-tracing approaches, exhibit substantial computational overhead and limited adaptability to dynamic scenarios. Although existing neural network (NN) approaches have efficient inferring speed, they lack sufficient consideration of the underlying physics of EM wave propagation, limiting their effectiveness in accurately modeling critical EM singularities induced by complex multipath environments. To address these fundamental limitations, we propose a novel physics-inspired RM construction method guided explicitly by the Helmholtz equation, which inherently governs EM wave propagation. Specifically, based on the analysis of partial differential equations (PDEs), we theoretically establish a direct correspondence between EM singularities, which correspond to the critical spatial features influencing wireless propagation, and regions defined by negative wave numbers in the Helmholtz equation. We then design an innovative dual diffusion model (DM)-based large artificial intelligence framework comprising one DM dedicated to accurately inferring EM singularities and another DM responsible for reconstructing the complete RM using these singularities along with environmental contextual information. Experimental results demonstrate that the proposed RadioDiff- $k^2$  framework achieves state-of-the-art (SOTA) performance in both image-level RM construction and localization tasks, while maintaining inference latency within a

few hundred milliseconds. Code is available at <https://github.com/UNIC-Lab/RadioDiff-k>

**Index Terms**—Radio map, Helmholtz equation, partial differential equations, physics-informed neural network, large artificial intelligence, diffusion model.

## I. INTRODUCTION

THE paradigm shift in wireless communication from environment-independent operation to an environment-aware framework marks a fundamental transformation in the design and functionality of future 6G [1], [2]. This transition is driven by the increasing demand for intelligent, adaptive, and context-aware communication systems, particularly in the era of 6G networks, where network nodes must not only operate within but also dynamically interact with complex and time-varying environments [3], [4]. A fundamental enabler of this transformation is the radio map (RM), which provides a spatial representation of critical wireless channel characteristics, including path loss, angle of arrival (AoA), and interference distribution [5]. By integrating location-specific channel state information (CSI), RMs facilitate proactive network optimization, intelligent resource allocation, and real-time adaptation to environmental dynamics, offering a significant advantage over conventional reactive network management strategies [6]. One of the most impactful applications of RM lies in low-overhead channel estimation, where it enables infrastructure components, such as intelligent reflecting surfaces (IRS) and massive multiple-input multiple-output (MIMO) systems, to obtain high-fidelity CSI with minimal reliance on pilot signals [7]. This reduces training overhead, mitigates pilot contamination, and enhances spectral efficiency, making RM a cornerstone for sustainable and scalable 6G wireless communication. Furthermore, RM plays a pivotal role in the trajectory planning, coverage optimization, and interference mitigation of mobile network elements, such as autonomous aerial vehicles (AAVs) and satellites in space-air-ground integrated networks (SAGIN) [8], [9]. By precomputing the spatial variations of wireless channel properties, RM allows mobile nodes to anticipate and mitigate signal blockages, optimize handovers, and enhance connectivity resilience in highly dynamic environments.

Received 30 April 2025; revised 9 October 2025; accepted 29 November 2025. Date of publication 5 December 2025; date of current version 20 February 2026. This work was supported by the National Key Research and Development Program of China under Grant 2024YFB907500. (Corresponding author: Nan Cheng.)

Xiucheng Wang, Nan Cheng, Ruijin Sun, and Zan Li are with the State Key Laboratory of ISN and the School of Telecommunications Engineering, Xidian University, Xi'an 710071, China (e-mail: xcwang\_1@stu.xidian.edu.cn; dr.nan.cheng@ieee.org; sunruijin@xidian.edu.cn; zanli@xidian.edu.cn).

Qiming Zhang is with the School of Artificial Intelligence, Xidian University, Xi'an 710071, China (e-mail: 23009200991@stu.xidian.edu.cn).

Shuguang Cui is with the School of Science and Engineering (SSE), Shenzhen Future Network of Intelligence Institute (FNii-Shenzhen), and Guangdong Provincial Key Laboratory of Future Networks of Intelligence, The Chinese University of Hong Kong, Shenzhen Campus, Shenzhen 511453, China (e-mail: shuguangcui@cuhk.edu.cn).

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Digital Object Identifier 10.1109/JSAC.2025.3641105

1558-0008 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

Authorized licensed use limited to: University of Waterloo. Downloaded on March 27, 2026 at 14:07:03 UTC from IEEE Xplore. Restrictions apply.

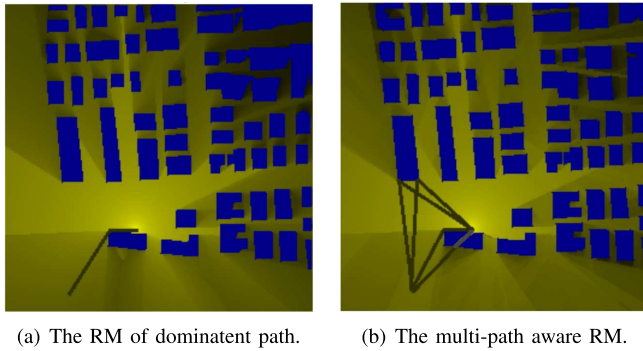


Fig. 1. Illustration of the RM, where (a) is the RM only considering the dominating path, (b) is the RM considering the multi-path effect, and the propagating path is shown for a specific location.

However, the construction of RM is inherently challenging due to the simultaneous demands for efficiency and accuracy, particularly within highly dynamic 6G environments [5]. Broadly, RM construction methods can be categorized into two distinct classes: electromagnetic (EM)-based high-precision methods [14] and data-driven approaches emphasizing computational efficiency [12]. Traditional high-precision EM-based methods rely on solving Maxwell's equations via full-wave numerical simulations, providing highly accurate spatial distributions of wireless channel characteristics [15]. However, their computational complexity is substantial, typically requiring hours to simulate EM wave propagation even within meter-scale scenarios [14]. To mitigate computational demands, approximate methods such as ray-tracing techniques have emerged, significantly reducing the required time to tens of minutes for kilometer-scale RM generation by simplifying EM wave propagation laws [16]. However, these EM-based approaches inherently assume a static environment during the computation period; this assumption is incompatible with the temporally and spatially dynamic characteristics of 6G networks, which include mobile network elements such as AAVs and satellites [17]. Consequently, recent attention has shifted toward neural network (NN)-based data-driven methods, offering efficiency through their rapid inference capabilities [10], [12], [13]. Nonetheless, existing NN-based approaches encounter significant limitations, notably their inability to accurately capture electromagnetic singularities resulting from intricate EM-environment interactions, as is shown in Fig. 2. These singularities, characterized by abrupt spatial variations in channel properties like pathloss, critically impact network performance by influencing factors such as user signal quality, AAV trajectory design, and beamforming alignment [9]. Moreover, traditional NN architectures, originally designed for tasks such as image processing or pattern recognition, inadequately address the physical nature and complexity of EM singularities, as these singularities are intrinsically tied to the physics of wave propagation rather than visual patterns or frequency domain characteristics [12]. An additional fundamental shortcoming of existing data-driven methods is their predominant focus on modeling the main propagation path, largely neglecting the complex yet critical multipath phenomena inherent in real-world wireless

communication environments. Multipath propagation leads to numerous secondary signals arriving at the receiver from various directions and with diverse delays, significantly complicating the spatial distribution of EM singularities. Moreover, as shown in Fig. 3, even classical digital filtering techniques fail to recover these fine-scale multipath-induced patterns [18], further underscoring the limitations of existing approaches. These observations highlight the urgent need for a new class of RM construction methods that not only infer the spatial distribution of channel characteristics but also incorporate a deep understanding of the physical mechanisms, particularly the multipath-driven singularities underpinning wave propagation.

However, achieving both efficiency and accuracy in RM construction remains challenging, motivating the development of novel methods to bridge the gap between the high precision of physics-based approaches and the computational efficiency of data-driven techniques. To address this challenge, this paper proposes a physics-inspired theoretical framework for effectively modeling the spatial distribution features of EM wave propagation. By explicitly incorporating the physics embedded in the EM wave equation, particularly the Helmholtz equation [15], into the NN training process, our method enables the NN to capture spatial distribution characteristics intrinsic to EM wave propagation. Specifically, through rigorous theoretical analysis of the Helmholtz equation, we discover that electromagnetic singularities, which are key regions exhibiting abrupt changes in wireless channel features due to multipath effects, directly correspond to regions characterized by a complex wave number where  $k^2 < 0$ . Leveraging this novel insight, we devise a specialized NN architecture and training strategy that significantly enhances the NN's capability to identify and model abrupt variations in wireless channel characteristics caused by multipath propagation. Moreover, inspired by the success of the generative diffusion model (DM) [19] for RM construction, where the DM-based method, such as RadioDiff [13], achieves much better performance than pixel-to-pixel methods, such as RadioUNet [12] and RME-GAN [10], within hundreds of milliseconds latency. The generative DM is used as the backbone to achieve high performance. Consequently, our physics-informed approach bridges the critical gap between traditional data-driven models and the underlying physical principles governing EM propagation, significantly advancing the accuracy and reliability of RM construction in complex 6G communication scenarios. The main contributions in this paper are summarized as follows.

- 1) Theoretically, a physics-inspired RM construction theory is proposed in this paper, which explicitly integrates the underlying EM propagation characteristics governed by the Helmholtz equation into NN learning frameworks. Through an in-depth theoretical analysis of the Helmholtz equation, we reveal that electromagnetic singularities, corresponding to abrupt spatial variations in the wireless channel caused by multipath propagation, are characterized by regions where the  $k^2 < 0$ . This insight fundamentally enhances the ability of NNs to accurately capture and predict intricate spatial features, significantly improving the accuracy of RM-

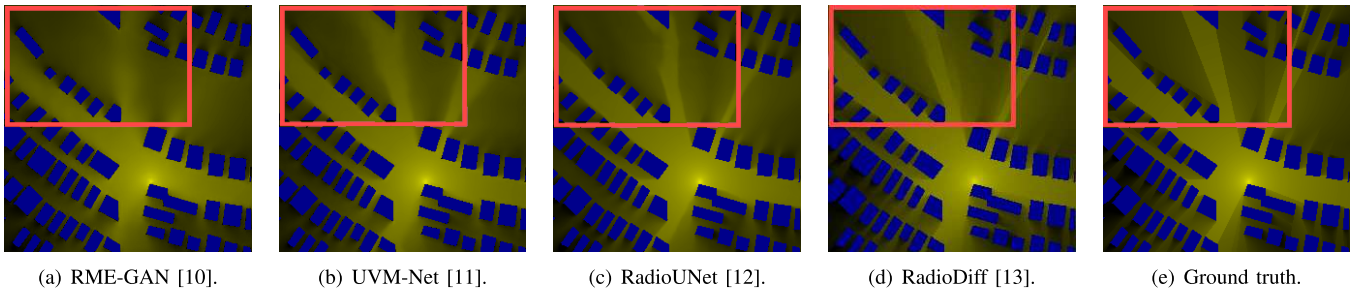


Fig. 2. Illustration of RMs constructed by various NN-based baseline methods, where RadioDiff is the SOTA NN-based RM construction method. The results indicate that existing NN-based approaches exhibit limited capability in capturing fine-grained texture features of the RM, particularly those associated with abrupt spatial variations in wireless channel characteristics. This deficiency hinders the ability of these methods to reliably detect electromagnetic singularities, thereby limiting their effectiveness in downstream wireless network optimization tasks.

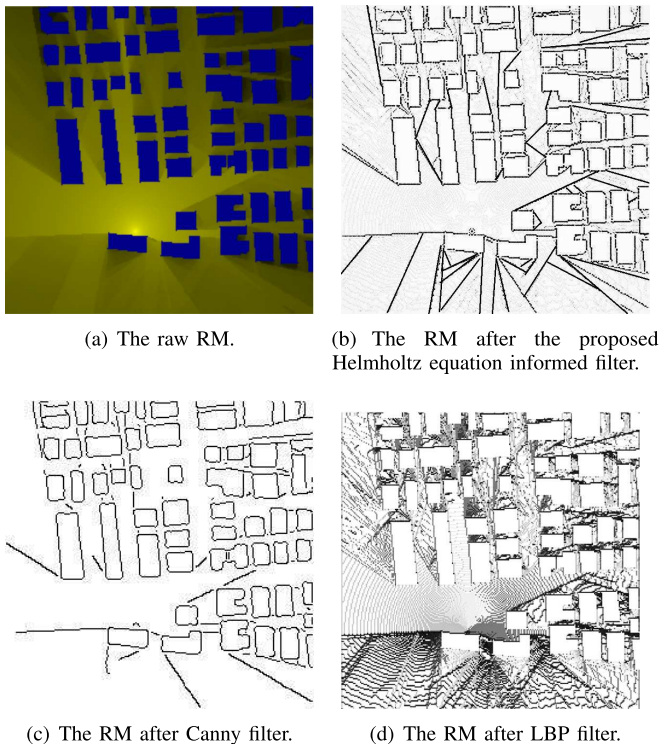


Fig. 3. Illustration of the different texture extraction methods.

based wireless environmental awareness required by next-generation communication systems.

- 2) Based on the physics-informed theory and partial differential equations (PDEs), we propose an innovative dual DM-based RM construction method specifically tailored for effectively modeling the complex multipath propagation environment. In our architecture, the first DM is explicitly designed and trained to predict the spatial distribution of EM singularities. The second NN leverages these learned singularities together with environmental context to accurately reconstruct the RM. This novel dual-stage DM approach ensures that intricate propagation characteristics induced by multipath effects are meticulously captured and effectively translated into the final RM.

- 3) Different from traditional NN-based RM methods that primarily model the main path propagation features, our proposed method explicitly addresses multipath effects. By comprehensively incorporating multipath-induced singularities into RM construction, our method is particularly advantageous in highly dynamic and complex 6G scenarios.
- 4) Extensive experimental results demonstrate that the proposed method achieves remarkable improvements in widely adopted evaluation metrics, such as normalized mean square error (NMSE), root mean square error (RMSE), structural similarity (SSIM), and peak signal-to-noise ratio (PSNR). Moreover, it establishes a new SOTA baseline in downstream localization tasks, achieving sub-5-meter accuracy under both static and dynamic scenarios in heterogeneous line-of-the-sight (LoS) and non-LoS scenarios. Despite its dual-stage diffusion architecture, the model maintains practical inference latency within a few hundred milliseconds, facilitating its deployment in near-real-time applications.

## II. RELATED WORKS AND PRELIMINARIES

### A. Related Works

RM construction methods can generally be classified into sampling-based and sampling-free approaches. Sampling-based methods rely on sparse pathloss measurements (SPM) collected at specific locations, which are then interpolated to estimate the RM over the entire area. These methods do not require prior knowledge of the environment or base station (BS) locations. Classic techniques include K-nearest neighbors (KNN) interpolation, which estimates unknown values by a weighted average of nearby measurements [20], and local multinomial regression, which fits a local linear model via least-squares minimization using nearby data points [21]. More sophisticated techniques, such as Kriging treat the interpolation as a stochastic process governed by spatial correlation, allowing for more accurate pathloss prediction by modeling the covariance structure of measurements [22].

Despite their simplicity, these methods suffer from two core limitations: a strong dependence on the availability of SPM and limited reconstruction accuracy, particularly in regions with sparse measurements or complex propagation environments [12], [23]. As a result, sampling-free RM construction

has gained increasing attention. These approaches eliminate the need for measurement data in the target area and instead leverage environmental features—such as obstacle positions, heights, and BS coordinates—to infer location-specific channel information. Representative efforts include RadioUNet, which adapts the U-Net architecture from image-to-image translation tasks to RM generation using MSE-based supervision [12], and RadioNet, which incorporates transformer-based attention mechanisms to capture global spatial dependencies in RM construction [23]. Further, graph neural networks (GNNs) have been applied to exploit the relational structure of spatial layouts for more expressive modeling [24]. While these methods offer improvements, most treat RM construction as a purely discriminative learning task and may struggle to capture the full distributional complexity of pathloss patterns in dynamic wireless environments. Some attempts, such as RME-GAN, introduced generative frameworks based on generative adversarial networks (GANs) to enhance output realism [10]. However, RME-GAN remains dependent on sparse measurements, thus not truly sampling-free. In contrast, this work frames RM construction as a conditional generative modeling task, proposing a diffusion-based method that generates high-fidelity RMs solely from environmental and BS input features—without relying on sampling within the target region. This fundamentally shifts the RM construction paradigm, enabling stronger generalization and distribution modeling capabilities.

---

**Algorithm 1** Pipeline of RadioDiff- $k^2$ 


---

- 1: **Input:** environment maps  $\mathbf{H}_s$ , dynamic maps  $\mathbf{H}_d$ , BS location  $r$ , ground-truth RM  $\mathbf{P}$ , initialized DM1 for map outline and initialized DM2 for RM generation.
  - 2: # **Training Pipeline** #
  - 3: **Stage 0: Data preprocessing**
  - 4: For each RM, compute  $k_{\text{eff}}^2(x)$  using (33).
  - 5: Define the outline mask  $\mathbf{1}\{k_{\text{eff}}^2(x) < 0\}$ , which keep locations where  $k_{\text{eff}}^2$  is negative.
  - 6: **Stage 1: Train DM1 for map-outline generation**
  - 7: Use  $(\mathbf{H}_s, \mathbf{H}_d, r)$  as inputs and train DM1 to predict the outline map.
  - 8: **Stage 2: Train DM2 for RM reconstruction**
  - 9: Use  $(\mathbf{H}_s, \mathbf{H}_d, r)$ , and the outline map generated by DM1 as inputs and train DM2 to generate the RM  $\hat{\mathbf{P}}$ .
  - 10: # **Inferencing Pipeline** #
  - 11: Input  $(\mathbf{H}_s, \mathbf{H}_d, r)$  to the DM1 to get the outline map.
  - 12: Input  $(\mathbf{H}_s, \mathbf{H}_d, r)$  and outline map generated by the DM1 to the DM2 to predict the RM.
- 

### B. Physics Informed Dual DM Framework

#### C. Score-Based Denoising Diffusion Model

DMs have recently emerged as powerful generative frameworks capable of capturing complex data distributions and synthesizing high-fidelity samples [19]. Among these, score-based diffusion models define a generative process grounded in

stochastic differential equations (SDEs), wherein data undergoes progressive corruption and is later reconstructed by learning to approximate the gradient of the data log-likelihood, known as the score function [25]. Unlike classical DMs that adopt discrete Markov chains for the forward and reverse noise processes, score-based models perturb data in continuous time via SDEs, making them well-suited for solving inverse problems such as radio map reconstruction through Bayesian sampling principles. The forward process is formulated by the SDE as follows.

$$d\mathbf{x} = f(\mathbf{x}, t)dt + g(t)d\mathbf{w}, \quad (1)$$

where  $f(\mathbf{x}, t)$  denotes the drift component,  $g(t)$  is a time-dependent diffusion coefficient, and  $d\mathbf{w}$  is a Wiener process. As time  $t$  increases, the data distribution  $p_t(\mathbf{x})$  transitions toward an isotropic Gaussian. Reconstructing the original data entails solving the reverse-time SDE as follows.

$$d\mathbf{x} = [f(\mathbf{x}, t) - g^2(t)\nabla_{\mathbf{x}} \log p_t(\mathbf{x})] dt + g(t)d\bar{\mathbf{w}}, \quad (2)$$

where  $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$  is the score function, typically approximated by a neural network  $s_{\theta}(\mathbf{x}, t)$ . The objective is to train  $s_{\theta}$  as follows.

$$s_{\theta}(\mathbf{x}, t) \approx \nabla_{\mathbf{x}} \log p_t(\mathbf{x}). \quad (3)$$

For deterministic sampling, an equivalent probability flow ODE is derived:

$$d\mathbf{x} = \left[ f(\mathbf{x}, t) - \frac{1}{2}g^2(t)\nabla_{\mathbf{x}} \log p_t(\mathbf{x}) \right] dt, \quad (4)$$

which allows sample generation without stochasticity, analogous to denoising diffusion probabilistic models (DDPMs). In DDPMs, the forward noise process is discretized as follows.

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \alpha_t \mathbf{x}_{t-1}, \beta_t \mathbf{I}), \quad (5)$$

and the score is estimated via a learned denoiser  $\epsilon_{\theta}$  as follows.

$$s_{\theta}(\mathbf{x}, t) = -\frac{\epsilon_{\theta}(\mathbf{x}, t)}{\sqrt{1 - \bar{\alpha}_t}}. \quad (6)$$

This formulation highlights DDPM as a discrete approximation of continuous score-based models using a variance-preserving scheme.

In the context of RM construction, decoupled diffusion models (DDMs) have been proposed to improve generative stability, particularly in the RadioDiff framework [26]. Unlike standard models that inject noise directly, DDM first attenuates the original input to a zero baseline, followed by noise injection. The forward transition from  $\mathbf{x}_0$  to  $\mathbf{x}_t$  is governed as follows.

$$q(\mathbf{x}_t | \mathbf{x}_0) = \mathcal{N}(\gamma_t \mathbf{x}_0, \delta_t^2 \mathbf{I}), \quad (7)$$

with time-dependent coefficients  $\gamma_t$  and  $\delta_t$  controlling signal decay and noise variance. The process is further described by the SDE as follows.

$$d\mathbf{x}_t = \mathbf{f}_t \mathbf{x}_t dt + g_t d\epsilon_t, \quad (8)$$

$$\mathbf{f}_t = \frac{d \log \gamma_t}{dt}, \quad (9)$$

$$\int_0^t \mathbf{f}_t dt = -\mathbf{x}_0 \quad (10)$$

$$g_t^2 = \frac{d\delta_t^2}{dt} - 2\mathbf{f}_t \delta_t^2, \quad (11)$$

where  $\mathbf{f}_t$  dictates attenuation rate and  $g_t^2$  captures noise accumulation over time.

The reverse-time generation process reconstructs  $\mathbf{x}_0$  from  $\mathbf{x}_t$  by solving the following equation.

$$d\mathbf{x}_t = [\mathbf{f}_t \mathbf{x}_t - g_t^2 \nabla_{\mathbf{x}} \log q(\mathbf{x}_t)] dt + g_t d\bar{\epsilon}_t. \quad (12)$$

Thanks to its two-phase design, DDM stabilizes training and improves sampling quality by decoupling data transformation and noise injection. The deterministic transformation toward the zero vector simplifies forward sampling as follows.

$$q(\mathbf{x}_t | \mathbf{x}_0) = \mathcal{N}\left(\mathbf{x}_0 + \int_0^t \mathbf{f}_\tau d\tau, t\mathbf{I}\right), \quad (13)$$

and enables an efficient reverse update as follows.

$$q(\mathbf{z}_{t-\Delta t} | \mathbf{z}_t, \mathbf{z}_0) = \mathcal{N}\left(\mathbf{z}_t + \int_t^{t-\Delta t} \mathbf{f}_t dt - \frac{\Delta t}{\sqrt{t}} \boldsymbol{\epsilon}, \frac{\Delta t(t-\Delta t)}{t} \mathbf{I}\right). \quad (14)$$

Through this structured perturbation mechanism, DDM enhances generation quality and computational efficiency, making it especially suitable for high-fidelity RM reconstruction in next-generation wireless networks.

#### D. Helmholtz Wave Equation

The derivation of the Helmholtz equation begins with Maxwell's equations in a source-free, linear, isotropic, and time-invariant medium. Consider the time-harmonic forms of Maxwell's equations with an  $e^{j\omega t}$  dependence, where  $\omega$  is the angular frequency. The two curl equations are as follows.

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{G}, \quad \nabla \times \mathbf{G} = j\omega\epsilon\mathbf{E}, \quad (15)$$

where  $\mathbf{E}$  and  $\mathbf{G}$  denote the electric and magnetic fields, respectively, and  $\epsilon$  and  $\mu$  represent the permittivity and permeability of the medium. Taking the curl of Faraday's law,  $\nabla \times (\nabla \times \mathbf{E}) = -j\omega\mu\nabla \times \mathbf{G}$ , and substituting Ampère's law into it as follows.

$$\nabla \times (\nabla \times \mathbf{E}) = -j\omega\mu(j\omega\epsilon\mathbf{E}) = -\omega^2\mu\epsilon\mathbf{E}. \quad (16)$$

Utilizing the vector identity  $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ , and assuming a source-free medium where  $\nabla \cdot \mathbf{E} = 0$ , the equation reduces to following form.

$$\nabla^2 \mathbf{E} + \omega^2\mu\epsilon\mathbf{E} = 0. \quad (17)$$

By defining the wavenumber  $k = \omega\sqrt{\mu\epsilon}$ , we obtain the vector Helmholtz equation as follows.

$$\nabla^2 \mathbf{E} + k^2 \mathbf{E} = 0. \quad (18)$$

For fields with a radiation source, the Helmholtz equation can be expressed as follows.

$$\nabla^2 \mathbf{E} + k^2 \mathbf{E} + \mathbf{e} = 0. \quad (19)$$

where  $\mathbf{e}$  is the source term.

### III. SYSTEM MODEL AND PROBLEM FORMULATION

In this paper, we consider an RM construction scenario over a discrete spatial region modeled as an  $N \times N$  grid. Each grid cell is assumed to be sufficiently small such that the pathloss within a cell remains approximately constant. Consequently, the RM can be represented by a matrix  $\mathbf{P} \in \mathbb{R}^{N \times N}$ , where each entry  $P(i, j)$  denotes the pathloss value at the corresponding grid location. Similar to [12] and [13], a single BS equipped with a dipole antenna serves as the sole radiation source within this region. The antenna emits EM energy as a spherical wave, and its position is specified by the tuple  $r = \langle d_x, d_y, d_z \rangle$ , where  $d_z$  denotes the BS height and  $(d_x, d_y)$  indicates its horizontal coordinates in the grid. Since the channel characteristics between different antennas in a massive MIMO system are largely independent, the single-antenna case generalizes to multi-antenna settings by simply varying the BS position across antennas [15].

The environment comprises both static and dynamic obstacles. Static obstacles such as buildings are composed of homogeneous materials and exhibit consistent EM reflection and diffraction behavior. Following established assumptions [10], [12], [23], the interior of a static obstacle is modeled as a total EM shield, resulting in infinite pathloss, i.e., no signal propagation. The presence of static obstacles is described by a matrix  $\mathbf{H}_s \in \mathbb{R}^{N \times N}$ , where  $H_s(i, j) = 0$  indicates the absence of a static obstacle at location  $(i, j)$ . In contrast, dynamic obstacles such as vehicles induce partial attenuation and scattering of EM waves due to their smaller physical size and lower elevation. Unlike static structures, dynamic objects do not completely block wave propagation. Their spatial distribution is encoded by the matrix  $\mathbf{H}_d \in \mathbb{R}^{N \times N}$ , with  $H_d(i, j) = 0$  denoting the absence of dynamic obstacles at grid position  $(i, j)$ .

The goal of this work is to develop a NN  $\mu_\theta(\cdot)$ , parameterized by  $\theta$ , to predict the pathloss distribution  $\hat{\mathbf{P}} \in \mathbb{R}^{N \times N}$  based on the environmental context and BS configuration. The network is trained to minimize the discrepancy between the predicted RM  $\hat{\mathbf{P}}$  and the ground truth RM  $\mathbf{P}$ , quantified by a loss function  $\mathcal{L}(\hat{\mathbf{P}}, \mathbf{P})$ . The overall RM construction task can thus be formulated as the following optimization problem:

*Problem 1:*

$$\min_{\theta} \mathcal{L}(\hat{\mathbf{P}}, \mathbf{P}), \quad (20)$$

$$\text{s.t. } \hat{\mathbf{P}} = \mu_\theta(\mathbf{H}_s, \mathbf{H}_d, r). \quad (20a)$$

### IV. HELMHOLTZ EQUATION INFORMED DM FOR RM CONSTRUCTION

#### A. Feature Analysis of Helmholtz Alignment

To enable NNs to more effectively learn the spatial characteristics of EM-wave propagation, we ground the analysis in

the time-harmonic Helmholtz equation. In a locally isotropic free-space region, the scalar field  $\mathbf{u}(r)$ <sup>1</sup> satisfies

$$\nabla^2 \mathbf{u} + k^2 \mathbf{u} = 0, \quad (21)$$

$$\begin{aligned} \nabla^2 \mathbf{u} &= \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\mathbf{u}}{dr} \right) \\ &= \frac{2}{r} \frac{d\mathbf{u}}{dr} + \frac{d^2 \mathbf{u}}{dr^2}. \end{aligned} \quad (22)$$

The resulting radial equation

$$\frac{d^2 \mathbf{u}}{dr^2} + \frac{2}{r} \frac{d\mathbf{u}}{dr} + k^2 \mathbf{u} = 0. \quad (23)$$

admits the outward spherical-wave solution  $\mathbf{u}(r) = Ae^{ikr}/r$ .

Write  $\mathbf{u} = Ae^{i\phi}$  with amplitude  $A > 0$  and phase  $\phi$ . Expanding  $\nabla^2(Ae^{i\phi})$  and separating real/imaginary parts as the eikonal–transport form [15], gives

$$\nabla^2 A - A \|\nabla \phi\|^2 + k^2 A = 0, \quad (24)$$

$$2\nabla A \cdot \nabla \phi + A \nabla^2 \phi = 0. \quad (25)$$

Thus,

$$k^2 = \|\nabla \phi\|^2 - \frac{\nabla^2 A}{A} \quad (26)$$

$$k_{\text{eff}}^2 \triangleq -\frac{\nabla^2 A}{A} = k^2 - \|\nabla \phi\|^2 \leq k^2. \quad (27)$$

If  $k_{\text{eff}}^2(\mathbf{x}) < 0$ , then  $\|\nabla \phi(\mathbf{x})\|^2 > k^2$ , i.e., local spatial frequency/curvature dominates the medium wavenumber. This is a sufficient indicator for shadowing/diffraction/evanescent-like mutation zones in RM [16], [27]. Therefore, the field magnitude decays as  $|\mathbf{u}(r)| \propto 1/r$ , and the power obeys the inverse-square law, when the  $k^2 > 0$

$$I(r) \propto |\mathbf{u}(r)|^2 = \frac{|A|^2}{r^2}, \quad (28)$$

which is consistent with the Friis transmission relation in free space [28]. In strongly confining or cutoff-like situations, the relevant component of the propagation constant becomes imaginary, giving rise to an evanescent behavior that can be captured by an effective imaginary wavenumber  $k = i\gamma$ , and  $\gamma > 0$ , where  $k^2 < 0$ :

$$\mathbf{u}(r) = \frac{Ae^{-\gamma r}}{r}, \quad (29)$$

$$I(r) \propto \frac{|A|^2 e^{-2\gamma r}}{r^2}. \quad (30)$$

Compared with the gradual  $1/r$  decay of radiating waves, such evanescent regions exhibit rapid spatial attenuation and sharp texture transitions in radio maps.

Motivated by this contrast, we introduce a local, curvature-based effective wavenumber

$$k_{\text{eff}}^2(\mathbf{x}) \triangleq -\frac{\nabla^2 \mathbf{u}(\mathbf{x})}{\mathbf{u}(\mathbf{x})}, \quad (31)$$

<sup>1</sup>We define the scalar field  $\mathbf{u}(r)$  to represent the time-harmonic electric field in a radially symmetric free-space region, which simplifies the Helmholtz equation into a one-dimensional radial form. This representation enables tractable analysis of wave attenuation behaviors—distinguishing between radiative  $k^2 > 0$  and evanescent  $k^2 < 0$  regimes—and supports the derivation of curvature-based indicators that guide singularity detection in radio maps.

which serves as a physics-informed indicator of high-variation zones: negative values of  $k_{\text{eff}}^2$  are empirically aligned with abrupt changes due to shadowing, diffraction, or strong multipath. Unlike the physical  $k^2 = \omega^2 \mu \epsilon > 0$  in homogeneous media,  $k_{\text{eff}}^2$  is a data-derived quantity that summarizes local wavefront curvature and attenuation; it is therefore well suited to guide NNs toward EM-consistent discontinuities and singular structures that are often underrepresented by purely discriminative RM learners.

Given a RM to represent the EM power  $I = |u|^2$ , we use  $A = \sqrt{I}$  and compute the follows.

$$k_{\log}(\mathbf{x}) \triangleq -\nabla^2 \log(A(\mathbf{x}) + \varepsilon), \quad (32)$$

$$k_{\text{eff}}^2(\mathbf{x}) \triangleq -\frac{\nabla^2 A(\mathbf{x})}{A(\mathbf{x}) + \varepsilon}. \quad (33)$$

We primarily rely on the sign, such as  $k_{\log} < 0$ , which is gain-invariant and robust under light smoothing and cross-scale persistence, and operationally actionable: negative belts co-locate with sharp power transitions and guide coverage enhancement, active sampling allocation, and high order stability tuning along user corridors. Importantly, we do not interpret  $k_{\text{eff}}^2$  as the material wavenumber; it is a curvature-based, physics-aligned prior for RMs.

### B. Discretization for Helmholtz-Aligned Indicators

All finite-difference operators are applied to  $A = \sqrt{I}$  (or  $\log(A + \varepsilon)$ ) rather than to the complex field:

$$\left. \frac{\partial^2 A}{\partial x^2} \right|_{(i,j)} \approx \frac{A(i+1, j) - 2A(i, j) + A(i-1, j)}{h^2}, \quad (34)$$

$$\left. \frac{\partial^2 A}{\partial y^2} \right|_{(i,j)} \approx \frac{A(i, j+1) - 2A(i, j) + A(i, j-1)}{h^2}, \quad (35)$$

$$\nabla_h^2 A(i, j) \approx (A(i+1, j) + A(i-1, j) + A(i, j+1) \quad (36)$$

$$+ A(i, j-1) - 4A(i, j))/h^2, \quad (37)$$

and the following.

$$k_{\text{eff}}^2(i, j) = -\frac{\nabla_h^2 A(i, j)}{A(i, j) + \varepsilon}, \quad (38)$$

$$k_{\log}(i, j) = -\nabla_h^2 \log(A(i, j) + \varepsilon). \quad (39)$$

A single-scale 5-point Laplacian is  $O(N)$  for  $N$  grid points. With  $s$  Gaussian scales, the cost is  $O(sN)$  and remains negligible. We use mild Gaussian smoothing before differentiation and a small  $\varepsilon$  to avoid division by near-zero  $A$ . Boundary conditions follow the RM acquisition setup, and anisotropic grids use  $(h_x, h_y)$  accordingly.

We refer to the raw EM power distribution, which is the RM, as the “Field Strength Map”, while the binary or grayscale mask extracted from the Helmholtz-informed curvature indicator  $k_{\text{eff}}^2 < 0$  is denoted as the “Map Outline”. This outline highlights potential electromagnetic singularities and serves as an intermediate structural prior during training.

### C. Framework of the RadioDiff- $k^2$

Building upon the theoretical analysis of the Helmholtz equation, we introduce Radiodiff- $k^2$ , a novel dual-stage DM

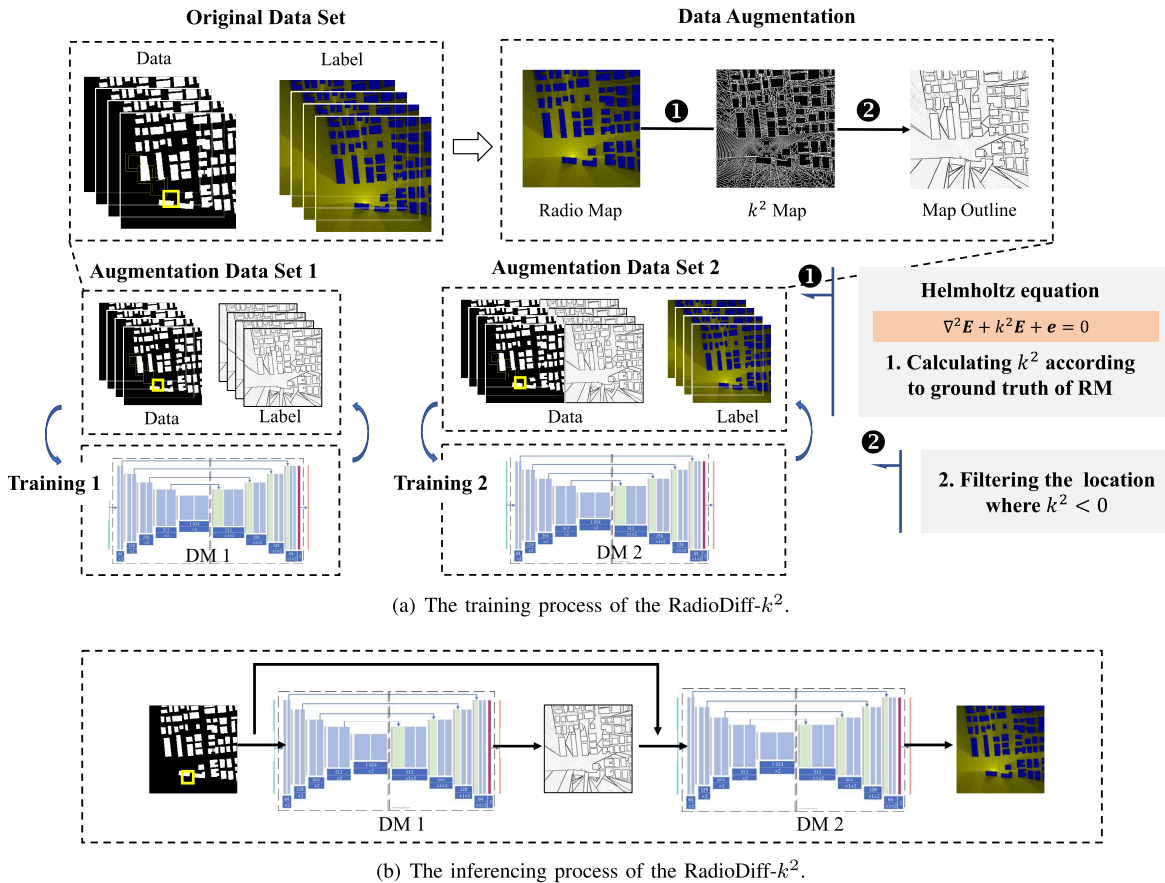


Fig. 4. Illustration of the proposed two-stage diffusion framework. (a) Training: The  $k^2$  map is derived from the Helmholtz equation and GT, and points with  $k^2 < 0$  form the Map Outline, refer to the Helmholtz-informed singularity map, used to train a dedicated DM. (b) Inference: Without solving the Helmholtz equation, the first DM predicts the Map Outline, which, together with environmental information, guides a second DM to generate the RM.

framework designed to address the complexities of RM construction, which is shown as Fig. 4. This framework draws inspiration from existing SOTA RM generation architectures but incorporates significant enhancements to leverage the underlying physics of EM wave propagation for improved performance. The architecture consists of two integral components: a Variational Autoencoder (VAE) and a Denoising UNet. The VAE's primary role is to encode the radio map from image space into a latent hidden space, which allows the denoising UNet to operate more efficiently in this compressed representation. The denoising UNet, in turn, is responsible for predicting the  $\epsilon_t$  and  $f_t$  terms required by the diffusion model during the denoising process with the condition of environment information and BS location, embedded by the cross attention, where the condition is used as the key and value vector of the cross attention [13]. Moreover, the model adopts a conditional DM architecture, similar to the approach used in RadioDiff, enabling controllable RM generation based on various input conditions such as environmental features and the locations of buildings and base stations. Considering the drawbacks of directly enforcing a Helmholtz PDE loss inside the diffusion backbone, we deliberately avoid embedding PDE residuals into the training objective. In latent diffusion, the denoiser operates in a non-physical latent space where differential operators lack meaning; projecting a Helmholtz residual to

image space  $x_0$  yields gradients that are typically much weaker than the diffusion objective and becomes ill-posed across noisy intermediates  $x_t$ , while real-world RM measurements are power-only and phase-free, with uncertain boundaries and material parameters, further complicating faithful PDE enforcement [29]. Instead, we translate the physics prior into a measurement-faithful spatial feature: a Helmholtz-aligned curvature outline  $\mathcal{M}$  derived from the observable envelope  $A = \sqrt{I}$  via  $k_{\log} = -\nabla^2 \log(A + \epsilon)$ , or  $k_{\text{eff}}^2 = -\nabla^2 A / (A + \epsilon)$ , and we rely on its sign, such as  $k_{\log} < 0$  to obtain a gain-invariant, smoothing-robust indicator of shadow/diffraction belts. This outline is computed once in preprocessing, learned by DM1 as structured supervision, and injected into DM2 as conditional guidance through cross-attention together with environmental/BS inputs. The design preserves the efficiency of latent diffusion while delivering consistent gains over a single-stage diffusion baseline, as RadioDiff.

A key innovation of RadioDiff- $k^2$  lies in its integration of electromagnetic singularities, derived from the Helmholtz equation, to enhance RM generation. These singularities, corresponding to regions where  $k^2 < 0$ , are particularly important for accurately modeling areas of rapid change in wireless channel characteristics, such as sudden shifts in pathloss. To effectively capture these features, we designed a two-stage dual framework. In the first stage, a diffusion model

is trained to learn and generate a feature map of regions where  $k^2 < 0$ , using environmental distributions and the base station location as input conditions. Once this model is trained, a second diffusion model is employed, which takes the predicted electric field singularity distribution map from the first model, along with the environmental features and base station location, to predict the final RM. It is important to note that, as illustrated in Fig. 4 and Algorithm 1, computations involving the Helmholtz equation are required only during the data preprocessing stage of training. In the inference stage, the framework merely performs two neural network forward passes. Since NN inference is highly efficient, the overall computational complexity is dominated solely by the network size.

This framework is inspired by curriculum learning [30], where the complexity of the data distribution is tackled progressively. The first diffusion model focuses on capturing regions where wireless channel features undergo abrupt changes due to electromagnetic singularities. This reduces the complexity faced by the second model, which can then use the information on the distribution of the electric field singularities to generate a more accurate radio map. This stepwise decomposition of the RM generation problem improves both the efficiency and the effectiveness of the process, ensuring that each model specializes in a distinct aspect of the radio map's spatial characteristics. Consequently, this method not only enhances the learning of fine-grained details but also provides better generalization and scalability in dynamic and complex environments.

#### D. Training Method of RadioDiff- $k^2$

In this work, the VAE is trained by minimizing the negative evidence lower bound (ELBO) of the RM, which consists of a reconstruction loss and a regularization term. Given an input sample  $\mathbf{x}$ , the encoder approximates the posterior distribution  $q_\phi(\mathbf{z}|\mathbf{x})$  over latent variables  $\mathbf{z}$ , while the decoder seeks to reconstruct  $\mathbf{x}$  from  $\mathbf{z}$  via  $p_\theta(\mathbf{x}|\mathbf{z})$ . The overall objective is given by

$$\mathcal{L}_{\text{VAE}}(\theta, \phi; \mathbf{x}) = \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})} [-\log p_\theta(\mathbf{x}|\mathbf{z})] + D_{\text{KL}}(q_\phi(\mathbf{z}|\mathbf{x}) \| p(\mathbf{z})). \quad (40)$$

The reconstruction loss is implemented as an  $\ell_2$ -norm between the input and output, assuming a Gaussian likelihood [31]. As for the denoising framework, Radiodiff- $k^2$  builds upon the DDM framework, where the inference process is governed by the reverse-time stochastic differential equation defined in (14). This formulation necessitates the generative model to estimate two critical components: the deterministic drift term  $\mathbf{f}_t$ , which characterizes the latent trajectory of the denoising process, and the noise term  $\hat{\epsilon}_t$ , which compensates for the stochastic perturbations introduced during the forward diffusion phase. In alignment with the design paradigm of RadioDiff, the ground truth for the drift term is defined as  $\mathbf{f}_t = -\mathbf{z}_0$ , where  $\mathbf{z}_0$  denotes the clean latent representation sampled from the variational encoder.

Accordingly, we define the drift loss as the MSE between the predicted drift term  $\hat{\mathbf{f}}_t$  and its target:

$$\mathcal{L}_{\text{drift}} = \mathbb{E}_t \left[ \left\| \hat{\mathbf{f}}_t + \mathbf{z}_0 \right\|_2^2 \right]. \quad (41)$$

In parallel, the noise loss is constructed as the MSE between the predicted noise component  $\hat{\epsilon}_t$  and the actual injected noise:

$$\mathcal{L}_{\text{noise}} = \mathbb{E}_t \left[ \left\| \hat{\epsilon}_t - \epsilon_t \right\|_2^2 \right]. \quad (42)$$

Meanwhile, these losses ensure that the model captures both the deterministic and stochastic components of the generative process. Although this dual-objective training strategy fosters convergence to an accurate generative path, further analysis of the reverse-time formulation offers a critical enhancement. By setting the denoising step size  $\Delta t = t$ , the model can reconstruct the clean latent variable  $\hat{\mathbf{z}}_0$  in a single step as follows.

$$\hat{\mathbf{z}}_0 = \mathbf{z}_t + \int_0^t \hat{\mathbf{f}}_\tau d\tau - \sqrt{t} \hat{\epsilon}_t. \quad (43)$$

This observation enables the incorporation of an auxiliary reconstruction loss, defined as follows.

$$\mathcal{L}_{\text{recon}} = \left\| \hat{\mathbf{z}}_0 - \mathbf{z}_0 \right\|_2^2, \quad (44)$$

which provides a global constraint on the learned denoising trajectory. This auxiliary term enhances training stability and promotes more faithful recovery of the latent representation, particularly under constrained denoising budgets. Combining these components, the overall loss function guiding the optimization of Radiodiff- $k^2$  is given as follows.

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{drift}} + \lambda_2 \mathcal{L}_{\text{noise}} + \lambda_3 \mathcal{L}_{\text{recon}}, \quad (45)$$

where  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  are scalar weights that balance the influence of each component. These coefficients can be tuned based on empirical performance and task-specific constraints.

Furthermore, the continuous-time formulation of the decoupled diffusion framework grants Radiodiff- $k^2$  the ability to perform adaptive-step inference. By modifying  $\Delta t$ , the model dynamically adjusts the granularity of the denoising trajectory, offering a practical trade-off between generation fidelity and computational efficiency. In real-time applications with stringent latency requirements, coarse inference with fewer steps can be adopted, while more refined RM generation is achievable via finer denoising schedules. This operational flexibility ensures that Radiodiff- $k^2$  remains both accurate and efficient in dynamic wireless communication environments.

## V. EXPERIMENTAL RESULTS

### A. Datasets and Metrics

To rigorously assess the performance of the proposed method, we conduct experiments on three distinct datasets, each representing varying degrees of environmental complexity and EM propagation fidelity. These datasets are designed to evaluate the effectiveness of RM reconstruction under both static and dynamic wireless scenarios, including the presence or absence of multipath effects. The foundation of our evaluation is the RadioMapSeer dataset, introduced in

the Pathloss RM Construction Challenge [32]. It contains 700 uniquely structured urban maps, each incorporating a varying number of buildings derived from OpenStreetMap data across cities, urban areas, and villages. For each map, 80 transmitter positions are specified along with the corresponding pathloss ground truth (GT). All maps are converted into  $256 \times 256$  binary morphological images, where each pixel denotes one square meter, with ‘1’ representing building interiors and ‘0’ denoting open areas. Transmitter and receiver heights are both set to 1.5 meters, with buildings uniformly modeled at 25 meters. The transmit power is fixed at 23 dBm, and the carrier frequency is 5.9 GHz. In particular, we consider three RM variants as follows. The GTs of these methods are all generated by electromagnetic ray tracing methods, so they can also be regarded as a performance comparison between neural network methods and computational electromagnetic methods.

- Static Radio Maps (SRM), generated using the dominant path model (DPM) [33] that only accounts for the main propagation path and the influence of large-scale static buildings;
- Dynamic Radio Maps (DRM), also based on the dominant path model but additionally incorporating small-scale dynamic obstacles (e.g., vehicles) randomly distributed along roadways, thereby simulating temporal urban dynamics;
- Multipath-Aware Radio Maps (MRM), produced using an intelligent ray tracing (IRT) [17] engine that includes multipath propagation with up to four environmental interactions per signal path. Due to computational constraints, the MRM dataset excludes dynamic obstacles and focuses solely on static structures.

To enhance the realism of our data, the DRM generation pipeline explicitly models dynamic reflectors and heterogeneous electromagnetic properties. Mobile objects such as vehicles are injected into scenes at traffic-relevant spatial and temporal scales, and different dielectric/conductive parameters are assigned to vehicles and to buildings so as to realistically reproduce vehicle-induced multipath fluctuations, partial occlusions, and localized attenuation differences. The training corpus comprises 600 distinct environments with 80 base-station placements per environment, while evaluation is performed on 100 disjoint, unseen environments with 80 base-station placements each. This setup constitutes a zero-shot, cross-environment generalization test that measures transferability to novel layouts and varying building densities.

Our framework is implemented in PyTorch and trained in two sequential stages using the AdamW optimizer with a cosine-decayed learning rate, initially set to  $5 \times 10^{-5}$  and gradually annealed to  $5 \times 10^{-6}$ . In the first stage, we train a VAE on the radio map images from the full training split. The VAE employs a latent dimensionality of 128 with 3 input channels, and its training is conducted on four NVIDIA RTX Pro6000 GPUs with batch size of 6, requiring approximately 30 hours. The second stage trains the denoising U-Net component of the diffusion model using the fixed latent representations produced by the VAE. This step is performed on four NVIDIA RTX Pro6000 GPUs with a batch size of 64 and completes in around 24 hours. Throughout, all images are resized to

$256 \times 256$  resolution. The diffusion model adopts a time horizon of  $T = 1000$ , uses a Gaussian starting distribution, and minimizes an  $\ell_2$ -based loss targeting the prediction of noise. Consistent with [34], the VAE encodes each image into a latent tensor of shape  $64 \times 64 \times 4$ , enabling efficient denoising in a compressed representation space.

To comprehensively evaluate the performance of RM reconstruction, we adopt a combination of classical error-based and perceptual quality metrics. Following previous works such as [12], we employ normalized mean squared error (NMSE) and root mean squared error (RMSE) to quantify overall prediction accuracy. However, since RM reconstruction aims not only to minimize global error but also to preserve fine structural and textural features critical to communication performance, we further incorporate the structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR) to assess perceptual and spatial fidelity.

NMSE measures the relative energy of the reconstruction error with respect to the ground truth, while RMSE reflects the square root of the averaged pixel-wise error. These metrics are defined as:

$$\text{NMSE} = \frac{\sum_{m=1}^M \sum_{n=1}^N (I_b(m, n) - I(m, n))^2}{\sum_{m=1}^M \sum_{n=1}^N I^2(m, n)}, \quad (46)$$

$$\text{RMSE} = \sqrt{\frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (I_b(m, n) - I(m, n))^2}, \quad (47)$$

where  $I_b$  and  $I$  denote the predicted and ground truth RM values, respectively, and  $M, N$  are the image dimensions.

To better evaluate structural fidelity, we employ SSIM, which captures perceptual similarity by considering luminance, contrast, and structural correlation between the predicted and ground truth maps. SSIM is especially relevant in RM tasks due to the presence of high-frequency textures and edge features that reflect electromagnetic wave behavior. It is computed as:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (48)$$

where  $\mu_x, \mu_y$  are the local means,  $\sigma_x, \sigma_y$  are variances, and  $\sigma_{xy}$  is the covariance between inputs  $x$  and  $y$ . Constants  $C_1$  and  $C_2$  stabilize the division and are derived from the dynamic range  $L$  of the image.

Lastly, we include PSNR to measure the reconstruction fidelity relative to the maximum possible signal intensity, with a particular focus on edge preservation, which is crucial in reflecting signal boundaries in RMs. It is defined as:

$$\text{PSNR} = 10 \log_{10} \left( \frac{r^2}{\text{MSE}} \right), \quad (49)$$

where  $r$  denotes the maximum pixel value in the image. Higher PSNR values correspond to lower reconstruction error and better visual quality. These four metrics provide a well-rounded evaluation framework, enabling both quantitative accuracy and perceptual quality assessment of radio map reconstruction performance.

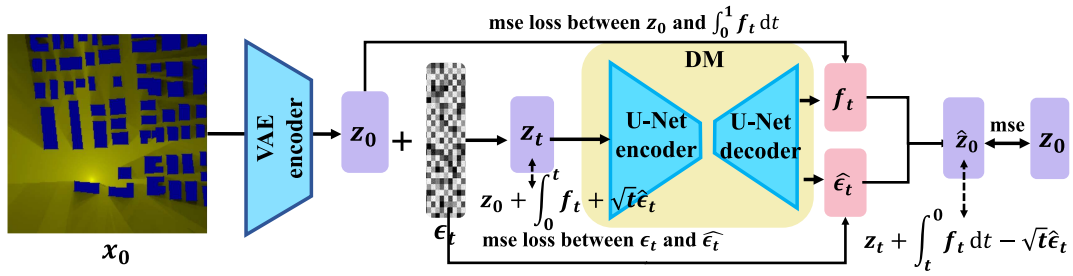


Fig. 5. Illustration of the training details.

### B. Comparisons With SOTA Methods

To provide a fair and informative benchmark, we compare iRadioDiff against four representative deep models that span the major architectural families; all baselines are trained and evaluated under the same data and experimental settings.

- **RadioUNet [12]:** A widely adopted sampling-free convolutional neural network (CNN)-based method for RM reconstruction. It leverages a U-Net architecture trained with supervised learning to infer RMs directly from environmental information. Due to its simplicity and effectiveness, RadioUNet has become a foundational benchmark in RM construction research.
- **RME-GAN [10]:** A generative adversarial network (GAN)-based model that originally incorporates both environmental features and sparse pathloss measurements (SPM) for RM generation. To ensure a fair comparison under the sampling-free setting, we adapt RME-GAN in our experiments to utilize only environmental features. While RME-GAN demonstrates the potential of adversarial training in generative RM tasks, it is not considered SOTA due to its reliance on sampled measurements in its original form.
- **RadioDiff [13]:** The current state-of-the-art in sampling-free RM construction. RadioDiff formulates the task as a conditional generative problem based on a DDM. It combines a VAE and a denoising UNet to model reverse-time EM propagation dynamics in latent space. This architecture enables fine-grained reconstruction of pathloss textures and structural features, achieving superior performance in accuracy and perceptual quality.
- **RMDM [35]:** A conditional diffusion framework that synthesizes high-resolution RM in the image domain from environmental geometry and base-station metadata. Training incorporates a Helmholtz-equation PDE residual as a physics-alignment loss, fusing probabilistic generative priors with electromagnetic consistency for RM reconstruction.

Since RadioDiff employs the same backbone architecture as our proposed method and directly generates the RM from environmental and base station inputs, it effectively serves as an ablation baseline corresponding to the case without the map-outline guidance derived from the  $k^2$ -based analysis.

### C. Evaluation on DPM

Based on the quantitative results in Table I and the qualitative comparisons in Fig. 6 and Fig. 7, the superiority of

the proposed RadioDiff- $k^2$  method is clearly demonstrated across all evaluation metrics and RM types. Our method achieves substantial improvements over baseline approaches, particularly in its ability to capture fine-grained structural details and abrupt spatial variations caused by EM singularities. In the case of SRM, RadioDiff- $k^2$  reduces NMSE by over 40% compared to RME-GAN and outperforms the strong diffusion-based baseline RadioDiff by a margin of 43%. It also achieves the highest SSIM score of 0.9773, reflecting superior structural similarity, and the highest PSNR of 34.46 dB, indicating enhanced fidelity in reconstructing signal boundaries and texture features. For DRM, which are more challenging due to additional scattering and dynamic obstacles, RadioDiff- $k^2$  again leads with an NMSE of 0.0054, a 40% reduction from the best-performing baseline, and shows consistent gains in RMSE, SSIM, and PSNR. This indicates that our method generalizes effectively to environments with more complex propagation characteristics.

The visual comparisons further support these conclusions. As shown in Fig. 6 and Fig. 7, the RMs generated by RadioDiff- $k^2$  display significantly sharper textures, clearer signal gradients, and finer spatial transitions compared to other methods. Unlike baseline models, which often blur high-frequency structures and fail to localize abrupt changes, our approach precisely reconstructs EM singularities—regions corresponding to rapid pathloss changes induced by multipath and diffraction effects. These detailed reconstructions are crucial for real-world applications such as UAV trajectory planning and low-pilot CSI estimation, where local signal variation must be modeled accurately. In summary, both the numerical metrics and visual results validate that RadioDiff- $k^2$  achieves state-of-the-art performance in RM reconstruction, with particularly strong advantages in modeling complex multipath and singularity-dominated scenarios.

### D. Evaluation on IRT

Based on the results shown in Fig. 8 and the accompanying performance in Table II, our proposed RadioDiff- $k^2$  method demonstrates clear superiority over all baselines, particularly in the challenging MRM reconstruction scenario modeled using IRT4. While IRT-based datasets reflect the complex multipath propagation effects typically encountered in real-world wireless environments, they also pose significant challenges for accurate reconstruction, especially in modeling abrupt pathloss variations resulting from EM singularities. Quantitatively, RadioDiff- $k^2$  achieves an NMSE of 0.0066, reducing the error by 45.5% compared to the best-performing baseline

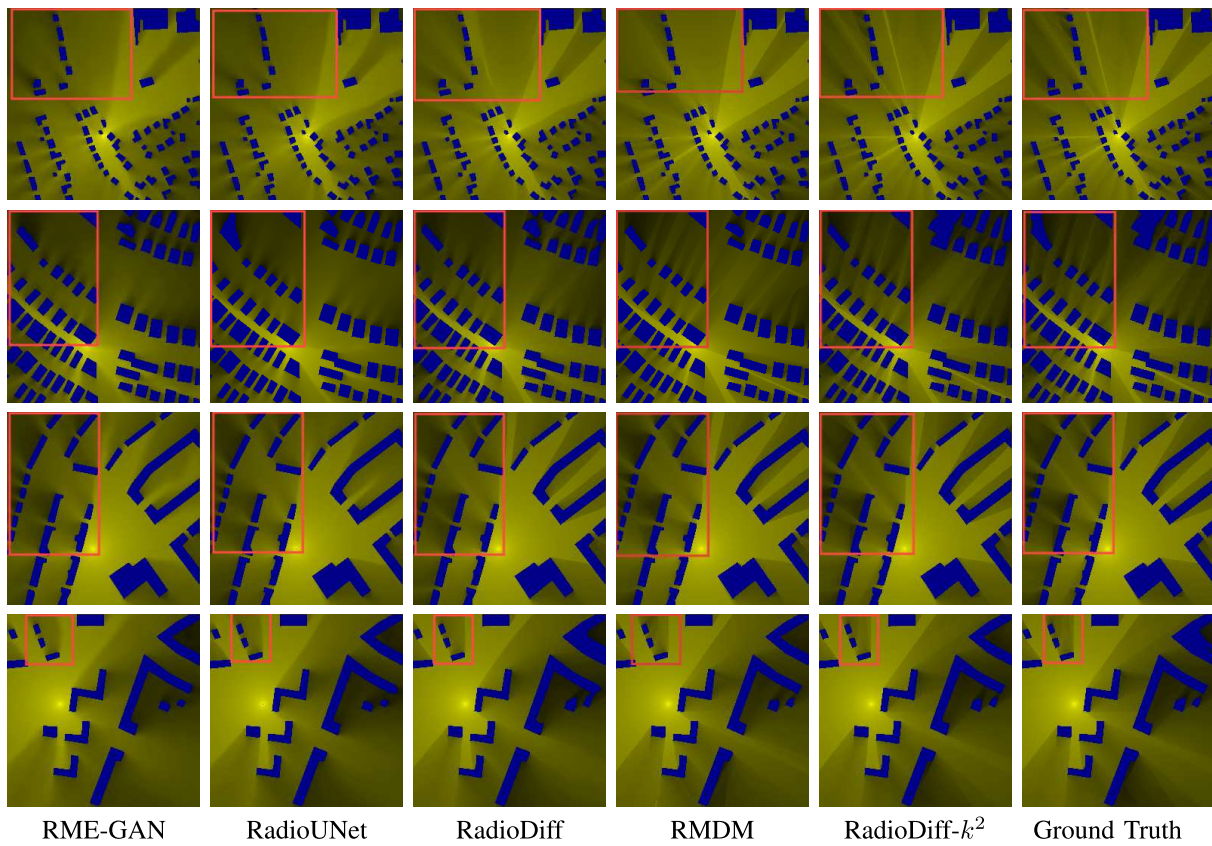


Fig. 6. The comparisons of constructed SRM on different methods.

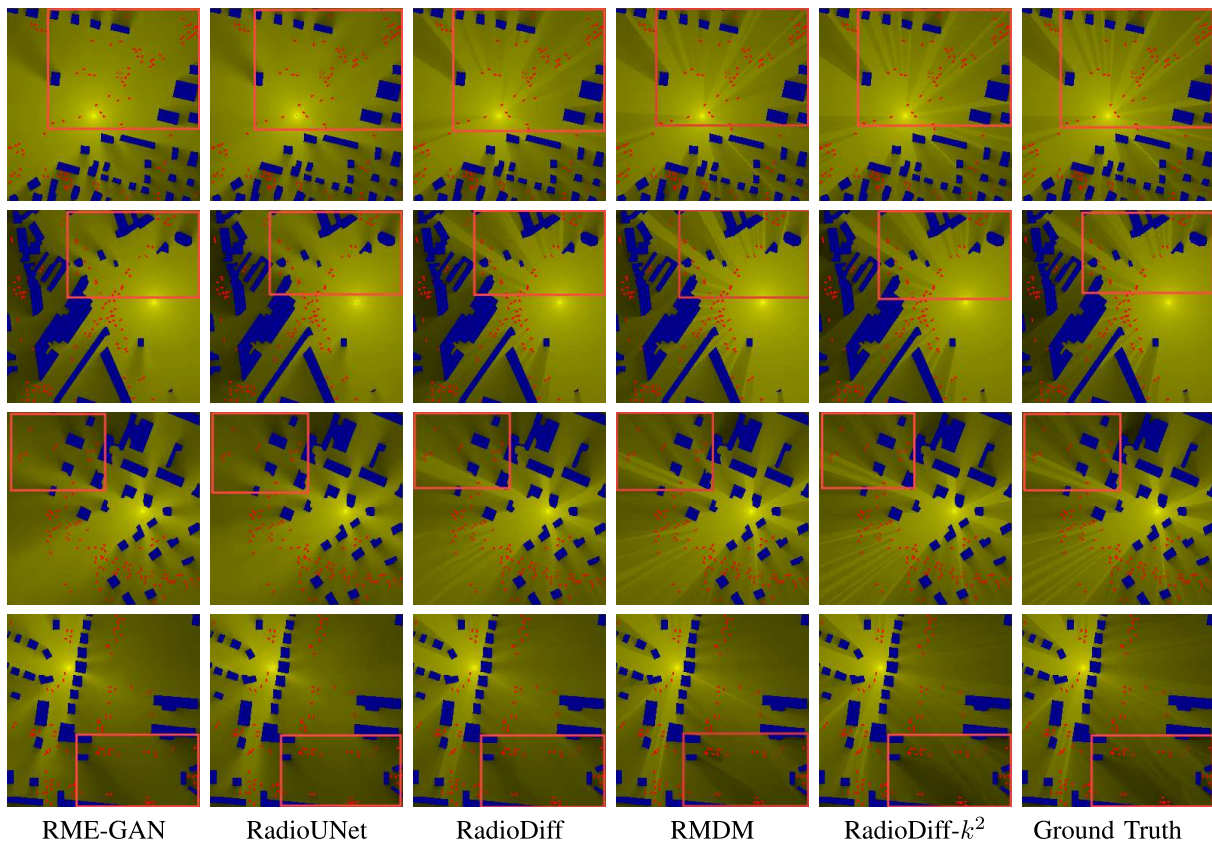


Fig. 7. The comparisons of constructed DRM on different methods. The buildings are colored blue, while the cars are colored red.

TABLE I

QUANTITATIVE COMPARISON ON DPM. RESULTS IN BOLD RED AND UNDERLINED BLUE HIGHLIGHT THE HIGHEST AND SECOND HIGHEST, RESPECTIVELY. THE  $\uparrow$  INDICATES METRICS WHEREBY HIGHER VALUES CONSTITUTE IMPROVED OUTCOMES, WITH HIGHER VALUES PREFERRED FOR ALL OTHER METRICS. "RATE w/  $k^2$  MAP (%)" DENOTES THE PERCENTAGE IMPROVEMENT OF THE PROPOSED METHOD WHEN CONDITIONED ON THE  $k^2$  MAP VERSUS THE BASELINE RADIODIFF WITHOUT THE  $k^2$  MAP

Methods		RME-GAN	RadioUNet	RadioDiff	RMDM	RadioDiff- $k^2$	Rate w/ $k^2$ Map (%)
SRM	NMSE	0.0096	0.0088	0.0072	<u>0.0059</u>	<b>0.0043</b>	$\downarrow$ 40.28%
	RMSE	0.0279	0.0266	0.0240	<u>0.0214</u>	<b>0.0193</b>	$\downarrow$ 19.58%
	SSIM $\uparrow$	0.9431	0.9466	0.9560	<u>0.9591</u>	<b>0.9773</b>	$\uparrow$ 2.232%
	PSNR $\uparrow$	31.35	31.77	32.67	<u>33.67</u>	<b>34.46</b>	$\uparrow$ 5.483%
DRM	NMSE	0.0115	0.0107	<u>0.0090</u>	0.0098	<b>0.0054</b>	$\downarrow$ 40.00%
	RMSE	0.0306	0.0291	<u>0.0266</u>	0.0283	<b>0.0208</b>	$\downarrow$ 21.80%
	SSIM $\uparrow$	0.9276	0.9291	<u>0.9432</u>	0.9363	<b>0.9704</b>	$\uparrow$ 2.884%
	PSNR $\uparrow$	30.42	30.89	<u>31.71</u>	31.13	<b>33.79</b>	$\uparrow$ 6.569%

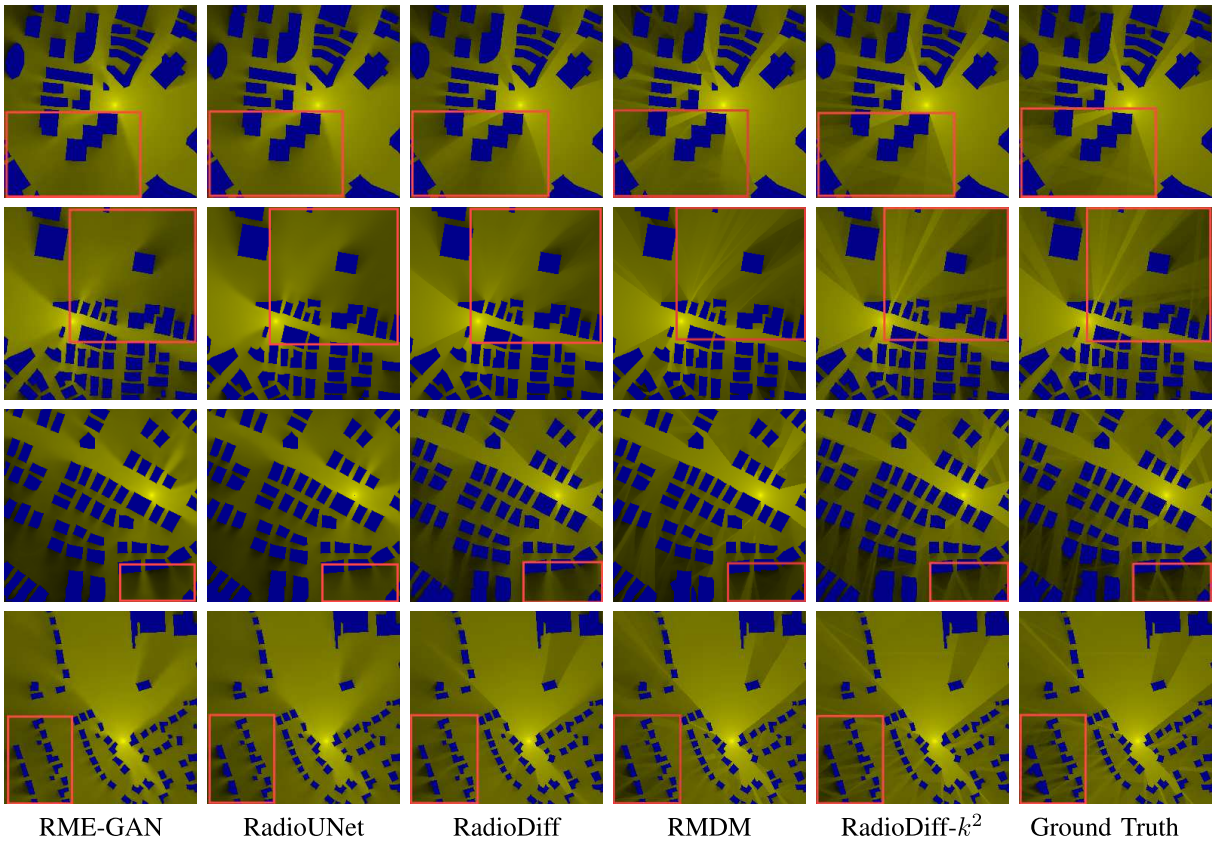


Fig. 8. The comparisons of constructed MRM on different methods.

TABLE II

QUANTITATIVE COMPARISON ON IRT. RESULTS IN BOLD RED AND UNDERLINED BLUE HIGHLIGHT THE HIGHEST AND SECOND HIGHEST, RESPECTIVELY. "RATE w/  $k^2$  MAP (%)" DENOTES THE PERCENTAGE IMPROVEMENT OF THE PROPOSED METHOD WHEN CONDITIONED ON THE  $k^2$  MAP VERSUS THE BASELINE RADIODIFF WITHOUT THE  $k^2$  MAP

Methods		RME-GAN	RadioUNet	RadioDiff	RMDM	RadioDiff- $k^2$	Rate w/ $k^2$ Map (%)
MRM	NMSE	0.0155	0.0159	0.0121	<u>0.0100</u>	<b>0.0066</b>	$\downarrow$ 45.45%
	RMSE	0.0340	0.0344	0.0309	<u>0.0278</u>	<b>0.0236</b>	$\downarrow$ 23.62%
	SSIM $\uparrow$	0.9123	0.9102	<u>0.9268</u>	0.9180	<b>0.9674</b>	$\uparrow$ 4.380%
	PSNR $\uparrow$	29.74	29.64	30.44	<u>31.40</u>	<b>32.68</b>	$\uparrow$ 7.352%

and significantly outperforming the standard RadioDiff. It also attains the lowest RMSE of 0.0236, and the highest SSIM of 0.9674 and PSNR of 32.68 dB scores, reflecting exceptional performance in preserving spatial structure and edge fidelity.

TABLE III  
LOCALIZATION ERROR COMPARISON ON DIFFERENT  
METHODS IN METERS

Method	SRM	DRM	MRM
RadioDiff- $k^2$	<b>3.32</b>	<b>3.87</b>	<b>4.72</b>
RadioDiff	<u>5.90</u>	<u>7.18</u>	<u>8.87</u>
RMDM	7.10	7.63	9.70
RadioUNet	8.04	9.86	11.80
RME-GAN	8.53	10.53	11.04

TABLE IV  
INFERENCE TIME AND MEMORY COMPARISON OF DIFFERENT MODELS

Model	RME-GAN	RadioUNet	RadioDiff	RMDM	RadioDiff- $k^2$
Time (s)	0.0057	0.0056	0.2400	19.30	0.7600
Memory (GB)	0.819	0.865	1.520	1.653	3.020

These gains are critical in capturing high-frequency variations in EM intensity that result from reflection, diffraction, and scattering phenomena directly linked to the distribution of EM singularities.

Visually, as shown in Fig. 8, RadioDiff- $k^2$  produces significantly sharper RM images with clearer boundaries, richer textures, and finer detail compared to existing methods. Competing baselines tend to oversmooth or blur critical signal transitions, failing to capture the multipath-induced fluctuations observable in the ground truth. In contrast, our method successfully preserves localized high-gradient regions, making it particularly well-suited for applications like trajectory planning and beamforming in dense and dynamic 6G environments. Overall, the results validate the effectiveness of integrating physics-based EM singularity modeling into a generative diffusion framework, establishing RadioDiff- $k^2$  as a robust and high-fidelity solution for realistic, multipath-aware RM construction.

### E. Localization Performance Comparison

To evaluate the practical utility of the generated radio maps, we performed localization experiments using a K-Nearest Neighbors (KNN) [36] algorithm with  $K = 5$ , randomly selecting 3,000 test points per map. These test positions include both Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) conditions, thereby capturing realistic localization challenges. As shown in Table III, across all three evaluation SRM, DRM, and MRM, RadioDiff- $k^2$  consistently achieves the lowest localization error. In particular, the average localization error remains below 5 meters in all cases, significantly outperforming prior methods such as RadioDiff, RMDM, and RadioUNet. These results demonstrate that the physics-informed singularity-aware structure of RadioDiff- $k^2$  enhances not only radio map reconstruction quality but also downstream localization accuracy under complex propagation conditions.

TABLE V  
PERFORMANCE COMPARISON USING DIFFERENT OUTLINE MAPS AS  
INPUT FOR RADIODIFF- $k^2$

Method	Dataset	NMSE	RMSE	SSIM	PSNR
Canny	SRM	0.0414	0.0596	0.8831	24.85
	DRM	0.0316	0.0495	0.8949	26.62
	MRM	0.6288	0.2614	0.3090	11.97
LBP	SRM	0.8865	0.3096	0.1544	10.34
	DRM	0.7529	0.2767	0.2448	11.38
	MRM	0.6297	0.2530	0.3085	12.06

### F. Efficiency Comparison

We evaluated the inference latency and memory consumption of all baseline and proposed models on an NVIDIA RTX Pro 6000 GPU. As shown in Table IV, traditional models such as RME-GAN and RadioUNet exhibit extremely low inference times and modest memory footprints, owing to their relatively shallow architectures. In contrast, models based on diffusion processes, including RadioDiff, RMDM, and RadioDiff- $k^2$ , incur approximately 3 to 4 $\times$  higher memory usage due to the iterative nature and larger model capacities inherent in diffusion-based frameworks. Nevertheless, both RadioDiff and RadioDiff- $k^2$  leverage the LDM architecture, allowing them to maintain inference times under 1 second—significantly faster than RMDM’s 19.3 seconds. This demonstrates that even for more realistic dynamic RM scenarios, such as those captured by DRM with moving vehicles and material heterogeneity, the inference latency of RadioDiff- $k^2$  remains practically acceptable for near real-time deployment.

### G. Ablation Study

To further assess the importance of physically informed guidance, we replaced the Helmholtz-derived map outline in our framework with two classical texture extraction methods, Canny and LBP, and used them as condition inputs to the diffusion model. As shown in Table V, both alternatives result in significantly degraded performance across all datasets, particularly in terms of NMSE, SSIM, and PSNR. Compared to the results in Table V to Tables I and Table II, the Canny and LBP based variants underperform the baseline RadioDiff model, which uses no outline guidance at all. This sharp contrast underscores the importance of using a physics-grounded indicator like  $k_{\text{eff}}^2$  to extract meaningful singularity structures. Unlike heuristic edge detectors, the Helmholtz-informed outline preserves EM-consistent transitions, leading to more effective conditioning and improved radio map reconstruction.

To investigate the impact of network architecture on the accuracy of  $k^2$  map prediction, we compare a standard U-Net with our proposed diffusion-based model across three RM datasets. As shown in Table VI, the diffusion model significantly outperforms the U-Net across all evaluation metrics, including the distance-tolerant coverage (DTC), distance-tolerant intersection over union (DTIoU), and learned perceptual image patch similarity (LPIPS) [37]. These

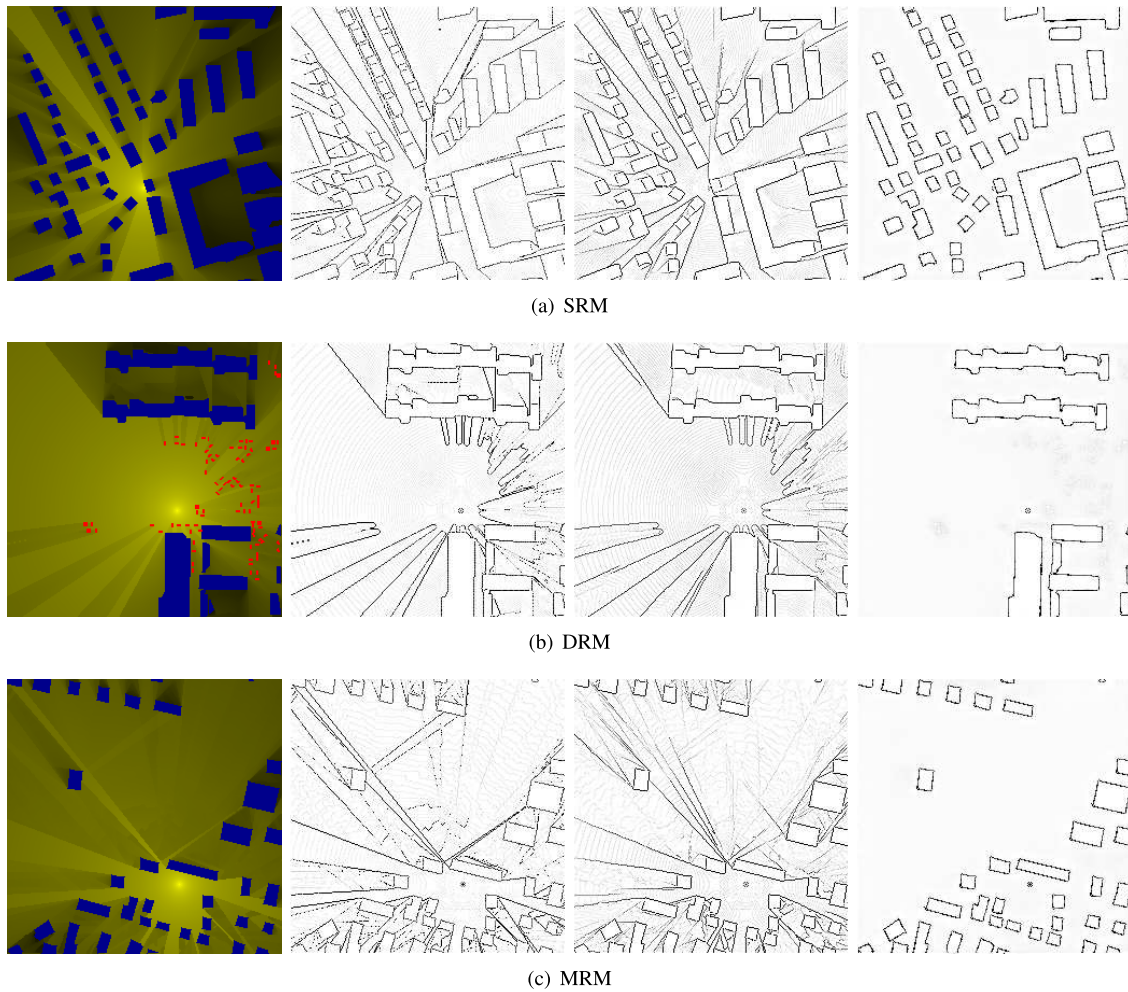


Fig. 9. Qualitative comparison of  $k^2$  maps. The first column shows the ground-truth RM, the second column shows the ground-truth  $k^2$  map derived from the RM, the third column presents the  $k^2$  map predicted by the DM, and the fourth column displays the  $k^2$  map predicted by a standard U-Net.

TABLE VI  
QUANTITATIVE COMPARISON ON  $k^2$  MAP

Metric	Model	SRM	DRM	MRM
DTC $\uparrow$	UNet	0.8141	0.6857	0.7433
	Diffusion	0.9703	0.9782	0.9577
DTIoU $\uparrow$	UNet	0.8283	0.6997	0.7483
	Diffusion	0.9023	0.8403	0.8615
LPIPS	UNet	0.2042	0.2449	0.2590
	Diffusion	0.0933	0.1107	0.1313

improvements underscore the superior capacity of the diffusion backbone in capturing the fine-grained singularity structures that are often missed by conventional architectures. Furthermore, the qualitative results in Fig. 9 provide intuitive visual evidence of this performance gap. The  $k^2$  maps generated by the diffusion model closely resemble the ground-truth singularity patterns and exhibit sharper texture boundaries, effectively highlighting abrupt transitions in the radio field.

In contrast, the U-Net predictions are severely smoothed, lacking clear structural detail, which aligns with the blurred edges observed in its RM outputs in Fig. 6 to Fig. 8. These results demonstrate that the use of diffusion models not only benefits full RM reconstruction but also enhances the fidelity of intermediate physics-informed guidance signals such as the  $k^2$  map.

## VI. CONCLUSION

In this work, we have proposed RadioDiff- $k^2$ , a physics-informed, diffusion-based framework for multipath-aware RM construction. By incorporating a Helmholtz-informed curvature indicator, the method first localizes electromagnetic singularities and then conditions a second-stage generative model on these features, improving interpretability and reconstruction fidelity. Extensive experiments on static and dynamic multipath environments show substantial gains over state-of-the-art baselines. The approach is lightweight at inference and has immediate applicability to real-time wireless tasks such as UAV trajectory planning and IRS beam selection; future work will extend robustness to material heterogeneity via conditional dielectric estimates. In future work, regarding material-mismatch realism, large-scale datasets with per-object

heterogeneous dielectric/conductive annotations are currently unavailable; we will incorporate estimated environmental electromagnetic parameters as additional conditional inputs to the diffusion model to improve robustness to material heterogeneity scenarios.

## REFERENCES

- [1] X. Shen et al., "Toward immersive communications in 6G," *Frontiers Comput. Sci.*, vol. 4, Jan. 2023, Art. no. 1068478.
- [2] N. Cheng et al., "6G omni-scenario on-demand services provisioning: Vision, technology and prospect (in chinese)," *Scientia Sinica Informationis*, vol. 54, pp. 1025–1054, 2024.
- [3] Y. Zeng and X. Xu, "Toward environment-aware 6G communications via channel knowledge map," *IEEE Wireless Commun.*, vol. 28, no. 3, pp. 84–91, Jun. 2021.
- [4] Y. Yang et al., "6G network AI architecture for everyone-centric customized services," *IEEE Netw.*, vol. 37, no. 5, pp. 71–80, Sep. 2023.
- [5] Y. Zeng et al., "A tutorial on environment-aware communications via channel knowledge map for 6G," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1478–1519, 3rd Quart., 2024.
- [6] H. Yilmaz, T. Tugcu, F. Alagöz, and S. Bayhan, "Radio environment map as enabler for practical cognitive radio networks," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 162–169, Dec. 2013.
- [7] Z. Wang et al., "A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1560–1605, 3rd Quart., 2024.
- [8] N. Cheng et al., "Space/aerial-assisted computing offloading for IoT applications: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117–1129, May 2019.
- [9] X. Wang et al., "Joint flying relay location and routing optimization for 6G UAV–IoT networks: A graph neural network-based approach," *Remote Sens.*, vol. 14, no. 17, p. 4377, Sep. 2022.
- [10] S. Zhang, A. Wijesinghe, and Z. Ding, "RME-GAN: A learning framework for radio map estimation based on conditional generative adversarial network," *IEEE Internet Things J.*, vol. 10, no. 20, pp. 18016–18027, Oct. 2023.
- [11] Z. Zheng and C. Wu, "U-shaped vision mamba for single image dehazing," 2024, *arXiv:2402.04139*.
- [12] R. Levie, Ç. Yapar, G. Kutyniok, and G. Caire, "RadioUNet: Fast radio map estimation with convolutional neural networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 4001–4015, Jun. 2021.
- [13] X. Wang et al., "RadioDiff: An effective generative diffusion model for sampling-free dynamic radio map construction," *IEEE Trans. Cognit. Commun. Netw.*, vol. 11, no. 2, pp. 738–750, Apr. 2024.
- [14] S.-H. Oh and N.-H. Myung, "MIMO channel estimation method using ray-tracing propagation model," *Electron. Lett.*, vol. 40, no. 21, pp. 1350–1352, Oct. 2004.
- [15] D. S. Jones, *The Theory of Electromagnetism*. Amsterdam, The Netherlands: Elsevier, 2013.
- [16] G. A. Deschamps, "Ray techniques in electromagnetics," *Proc. IEEE*, vol. 60, no. 9, pp. 1022–1035, 1972.
- [17] T. Rautiainen, G. Wölfle, and R. Hoppe, "Verifying path loss and delay spread predictions of a 3D ray tracing propagation model in urban environment," in *Proc. IEEE 56th Veh. Technol. Conf.*, vol. 4, Sep. 2002, pp. 2470–2474.
- [18] B. Jähne, *Digital Image Processing*. Cham, Switzerland: Springer, 2005.
- [19] J. Ho, A. N. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, Dec. 2020, pp. 6840–6851.
- [20] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inform. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.
- [21] F. J. Breidt and J. D. Opsomer, "Local polynomial regression estimators in survey sampling," *Ann. Statist.*, vol. 28, no. 4, pp. 1026–1053, Aug. 2000.
- [22] Y. Qiu et al., "Channel knowledge map construction based on a UAV-assisted channel measurement system," *Drones*, vol. 8, no. 5, p. 191, May 2024.
- [23] H. Li, K. Gupta, C. Wang, N. Ghose, and B. Wang, "RadioNet: Robust deep-learning based radio fingerprinting," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Oct. 2022, pp. 190–198.
- [24] G. Chen, Y. Liu, T. Zhang, J. Zhang, X. Guo, and J. Yang, "A graph neural network based radio map construction method for urban environment," *IEEE Commun. Lett.*, vol. 27, no. 5, pp. 1327–1331, May 2023.
- [25] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," 2020, *arXiv:2011.13456*.
- [26] Y. Huang, Z. Qin, X. Liu, and K. Xu, "Decoupled diffusion models: Simultaneous image to zero and zero to noise," 2024, *arXiv:2306.13720*.
- [27] C. A. Balanis, *Antenna Theory: Analysis and Design*. Hoboken, NJ, USA: Wiley, 2016.
- [28] E. A. Lee and D. G. Messerschmitt, *Digital Communication*. Cham, Switzerland: Springer, 2012.
- [29] Y. Ye, K. Xu, Y. Huang, R. Yi, and Z. Cai, "DiffusionEdge: Diffusion probabilistic model for crisp edge detection," 2024, *arXiv:2401.02032*.
- [30] X. Wang, Y. Chen, and W. Zhu, "A survey on curriculum learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 4555–4576, Sep. 2022.
- [31] D. P. Kingma and M. Welling, "An introduction to variational autoencoders," *Found. Trends Mach. Learn.*, vol. 12, no. 4, pp. 307–392, 2019.
- [32] Ç. Yapar, F. Jaensch, R. Levie, G. Kutyniok, and G. Caire, "The first pathloss radio map prediction challenge," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–2.
- [33] R. Wahl, G. Wölfle, P. Wertz, P. Wildbolz, and F. Landstorfer, "Dominant path prediction model for urban scenarios," in *Proc. 14th IST Mobile Wireless Commun.*, 2005, pp. 1–5.
- [34] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2022, pp. 10684–10695.
- [35] H. Jia et al., "Rmdm: Radio map diffusion model with physics informed," in *Proc. 33rd ACM Int. Conf. Multimedia (MM)*, 2025, pp. 1–9.
- [36] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM Conf. Comput. Commun. 19th Annu. Joint Conf. IEEE Comput. Commun. Soc.*, vol. 2, Mar. 2000, pp. 775–784.
- [37] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 586–595.



**Xiucheng Wang** (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree with Xidian University. His research interests include machine learning in wireless networks.



**Qiming Zhang** is currently pursuing the bachelor's degree with Xidian University, Xi'an, China. His research interests include diffusion models and radio maps.



**Nan Cheng** (Senior Member, IEEE) received the B.E. and M.S. degrees from the Department of Electronics and Information Engineering, Tongji University, Shanghai, China, in 2009 and 2012, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, in 2016.

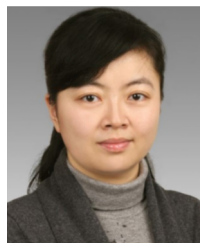
He was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, from 2017 to 2019. He is currently a Professor with the State Key Laboratory of ISN and the School of Telecommunications Engineering, Xidian University, Xi'an, Shaanxi, China. He has published over 90 journal articles in IEEE TRANSACTIONS and other top journals. His current research interests include B5G/6G, AI-driven future networks, and space-air-ground integrated networks.

Prof. Cheng serves as an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, and *Peer-to-Peer Networking and Applications*, and serves/served as a guest editor for several journals.



**Ruijin Sun** (Member, IEEE) received the Ph.D. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 2019. From September 2017 to September 2018, she was a Visiting Student with the University of Waterloo, Waterloo, ON, Canada. From 2019 to 2021, she was a joint Post-Doctoral Fellow with the Peng Cheng Laboratory, Shenzhen, China, and Tsinghua University, Beijing. She is currently a Lecturer with the School of Telecommunications Engineering and the State Key Laboratory of Integrated Services

Networks, Xidian University, Xi'an, China. Her research interests include knowledge-driven wireless resource allocation.



**Zan Li** (Fellow, IEEE) received the B.S. degree in communications engineering and the M.S. and Ph.D. degrees in communication and information systems from Xidian University, Xi'an, China, in 1998, 2001, and 2006, respectively.

She is currently a Professor with the State Key Laboratory of Integrated Services Networks, School of Telecommunications Engineering, Xidian University. Her research interests include wireless communications and signal processing, such as covert communication, spectrum sensing, and cooperative communications. She is a fellow of the Institution of Engineering and Technology, China Institute of Electronics, and China Institute of Communications. She received the National Science Fund for Distinguished Young Scholars. She serves as an Associate Editor for IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING and *China Communications*.



**Shuguang Cui** (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2005.

He was an Assistant Professor, an Associate Professor, a Full Professor, and the Chair Professor of electrical and computer engineering with The University of Arizona, Texas A&M University, UC Davis, and CUHK-Shenzhen, Shenzhen, China, respectively. He was the Executive Dean of the School of Science and Engineering, CUHK-Shenzhen, and the Director of the Future Network of Intelligence Institute. His current research interests include the merging of AI and communication networks. He is a fellow of Canadian Academy of Engineering and the Royal Society of Canada. He is a member of the Steering Committee for IEEE TRANSACTIONS ON BIG DATA. He is the Vice Chair of the IEEE VT Fellow Evaluation Committee and a member of the IEEE ComSoc Award Committee. He was an Elected Member of the IEEE Signal Processing Society SPCOM Technical Committee from 2009 to 2014 and the Elected Chair of the IEEE ComSoc Wireless Technical Committee from 2017 to 2018. He was selected as a Thomson Reuters Highly Cited Researcher and listed in the World's Most Influential Scientific Minds by ScienceWatch in 2014. He was elected as an IEEE ComSoc Distinguished Lecturer in 2014 and the IEEE VT Society Distinguished Lecturer in 2019. He is the Chair of the Steering Committee for IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING. He is the Editor-in-Chief of IEEE TRANSACTIONS ON MOBILE COMPUTING, an Area Editor of *IEEE Signal Processing Magazine*, and an Associate Editor of IEEE TRANSACTIONS ON BIG DATA, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.



**Xuemin (Sherman) Shen** (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990.

He is currently a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include network resource management, wireless network security, the Internet of Things, 5G and beyond, and vehicular ad hoc and sensor networks.

He is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, Chinese Academy of Engineering Foreign Member, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and the Communications Society. He received the R. A. Fessenden Award from IEEE Canada in 2019, the Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) in 2019, the James Evans Avant Garde Award from the IEEE Vehicular Technology Society in 2018, the Joseph LoCicero Award in 2015 and Education Award in 2017 from the IEEE Communications Society, and the Technical Recognition Award from the Wireless Communications Technical Committee in 2019 and the AHSN Technical Committee in 2013. He has also received the Excellent Graduate Supervision Award from the University of Waterloo and the Premier's Research Excellence Award (PREA) from the Province of Ontario, Canada, in 2003. He served as the Technical Program Committee Chair/Co-Chair for IEEE Globecom'16, IEEE Infocom'14, IEEE VTC'10 Fall, and IEEE Globecom'07, and the Chair for the IEEE Communications Society Technical Committee on Wireless Communications. He is the President-Elect of the IEEE Communications Society. He was the Vice President for Technical and Educational Activities, the Vice President for Publications, the Member-at-Large on the Board of Governors, the Chair of the Distinguished Lecturer Selection Committee, and a member of the IEEE Fellow Selection Committee of the ComSoc. He served as the Editor-in-Chief for IEEE INTERNET OF THINGS JOURNAL, *IEEE Network*, and *IET Communications*.