RadioDiff-Inverse: Diffusion Enhanced Bayesian Inverse Estimation for ISAC Radio Map Construction

Xiucheng Wang, *Student Member, IEEE*, Zhongsheng Fang, *Student Member, IEEE*, Nan Cheng, *Senior Member, IEEE*, Ruijin Sun, *Member, IEEE*, Haibo Zhou, *Fellow, IEEE*, Zhou Su, *Senior Member, IEEE*, Zan Li, *Fellow, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*—Radio maps (RMs) are essential for environment-aware communication and sensing, providing location-specific wireless channel information. Existing RM construction methods often rely on precise environmental data and base station (BS) locations, which are not always available in dynamic or privacy-sensitive environments. While sparse measurement techniques reduce data collection, the impact of noise in sparse data on RM accuracy is not well understood. This paper addresses these challenges by formulating RM construction as a Bayesian inverse problem under coarse environmental knowledge and noisy sparse measurements. Although maximum a posteriori (MAP) filtering offers an optimal solution, it requires a precise prior distribution of the RM, which is typically unavailable. To solve this, we propose RadioDiff-Inverse, a diffusion-enhanced Bayesian inverse estimation framework that uses an unconditional generative diffusion model to learn the RM prior. This approach not only reconstructs the spatial distribution of wireless channel features but also enables environmental building outlines perception, just relying on pathloss, through integrated sensing and communication (ISAC). The proposed method operates on routine communication measurements, without new waveforms, specialized feedback, or protocol changes, thereby enabling a plug-and-play ISAC capability. Remarkably, RadioDiff-Inverse is training-free, leveraging a pre-trained model from Imagenet without task-specific fine-tuning, which significantly reduces the training cost of using a generative large model in wireless networks. Experimental results demonstrate that RadioDiff-Inverse achieves state-of-the-art performance in accuracy of RM construction and environmental reconstruction, and robustness against noisy sparse sampling.**

***Index Terms*—Radio map, diffusion model, Bayesian inverse problem, maximum a posteriori, integrated sensing and communication.**

Received 21 May 2025; revised 29 October 2025; accepted 21 March 2026. Date of current version 1 April 2026. This work was supported in part by the National Key Research and Development Program of China under Grant 2024YFB2907500 and in part by the National Natural Science Foundation of China under Grant 62571402. The associate editor coordinating the review of this article and approving it for publication was Y. Zhou. *(Corresponding author: Nan Cheng.)*

Xiucheng Wang, Zhongsheng Fang, 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; zsfang@stu.xidian.edu.cn; dr.nan.cheng@ieee.org; sunruijin@xidian.edu.cn; zanli@xidian.edu.cn).

Haibo Zhou is with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China (e-mail: haibozhou@nju.edu.cn).

Zhou Su is with the School of Cyber Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: zhousu@ieee.org).

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/TWC.2026.3677479

I. Introduction

THE evolution of wireless communication from an environment-independent paradigm to an environment-aware paradigm marks a significant shift in the design and operation of next-generation networks [1]. A key enabler of this transition is the radio map (RM), which provides a graphical representation for the spatial distribution of wireless channel characteristics, such as pathloss, time of arrival (ToA), and angle of arrival (AoA) [2]. By encoding location-specific channel information, RM enables proactive network optimization and real-time adaptation to environmental changes [3]. One of the primary applications of RM lies in low pilot-consumption communication, where it allows infrastructure components such as intelligent reflecting surfaces (IRS) and massive multiple input multiple output (MIMO) systems to acquire high-fidelity channel state information (CSI) with no/low need for pilot signals, significantly reducing training overhead and enhancing spectral efficiency [4], [5]. Furthermore, RM plays a crucial role in trajectory planning and interference mitigation for dynamic network elements, such as autonomous aerial vehicles (AAVs) and satellites in space-air-ground integrated networks (SAGIN) [6], [7]. Recent RM-related works also model the scatters by explicitly parameterizing scatterers' angular responses, reflection coefficients, and path contributions to capture multipath structure in the environment [8].

Although numerous advantages of the RM, constructing an accurate and efficient RM remains a formidable challenge, particularly in dynamic and complex environments [2]. From a taxonomic perspective, RM construction can be divided into two primary categories: known-scenario RM construction [9], [10], [11], [12], [13], [14], [15] and unknown or semi-aware scenario RM construction [16], [17], [18], as is shown in Fig. 1. In known-scenario RM construction, all relevant environmental details, such as base station (BS) locations, the outlines and locations of environment entities are fully available, making RM construction a deterministic problem [11]. In theory, given the source radiation characteristics,

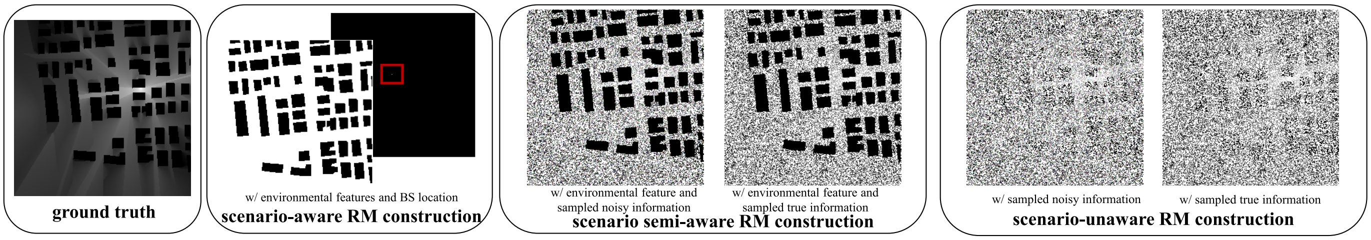


Fig. 1. Illustration of RM construction under different information.

dielectric constants, and magnetic permeability of the environment, the propagation of electromagnetic (EM) waves at any point in space can be derived by solving Maxwell's equations [19]. However, the computational complexity of directly solving Maxwell's equations is often prohibitive, requiring high-dimensional numerical solutions [20]. To mitigate this, approximate methods such as electromagnetic ray tracing (ERT) have been widely adopted, providing sufficiently accurate EM wave spatial distribution estimations at the cost of limited but extensive computational time [21]. To further enhance efficiency, neural network (NN)-based approaches have emerged as an alternative, leveraging data-driven learning to approximate RM distributions. Techniques such as UNet based RadioUNet [9] and Transformer-based RadioNet [10] enable fast inference by training deep models on environmental features like building layouts and BS locations. More recently, generative artificial intelligence (GAI) methods have been introduced to model RM as a stochastic process, where random Gaussian noise is incorporated during training to capture environmental uncertainty, such as generative adversarial network (GAN) based RME-GAN [22], [23], [24] and diffusion model (DM) based RadioDiff [11] approaches.

Despite the advances in scenario-aware RM construction, its practical deployment remains severely constrained by its reliance on high-precision environmental modeling and the known of the location of BS [2], [25]. These methods assume access to detailed maps of environmental structures, which define the propagation characteristics of EM waves. However, in privacy-sensitive areas, highly dynamic urban landscapes, or adversarial environments with non-cooperative radiation sources, obtaining such high-precision environmental data is often impractical, infeasible, or restricted [26]. This fundamental limitation significantly reduces the applicability of scenario-aware RM methods, particularly in real-world deployments where environmental conditions are constantly evolving. In such cases, traditional interpolation-based approaches, such as linear interpolation and Kriging, are commonly employed to estimate missing wireless channel characteristics from sparse measurements [16], [17]. However, these methods fail to generalize in complex environments where signal propagation is highly nonlinear, due to factors such as obstructions, multipath fading, and small-scale spatial variations. Furthermore, prior works have largely overlooked the impact of noise in sparsely sampled data on RM construction, assuming that sparse measurements remain noise-free [16]. In reality, noise in sparse sampling significantly degrades RM reconstruction accuracy, particularly in environments with severe multipath effects, dynamic interference, and varying signal propagation conditions [9]. The presence of noise in sparse measurements leads to highly unstable RM estimations, as traditional interpolation and machine learning-based methods often fail to recover structured channel characteristics under such uncertainties. Addressing scenario-unaware RM construction is particularly critical in emerging wireless applications, such as BS deployment optimization, interference management, and spectrum allocation in satellite and aerial networks [2]. In these scenarios, the ability to infer the RM of non-cooperative transmitters is essential for interference avoidance, coexistence strategies, and adaptive resource allocation [27], [28]. However, the absence of detailed environmental models and precise BS positioning information, coupled with the challenge of noise in sparse measurements, fundamentally limits the effectiveness of traditional RM construction techniques. This limitation underscores the urgent need for alternative approaches that can both leverage sparse measurements and mitigate noise effects to accurately reconstruct RMs in complex, dynamic, and unstructured settings. To overcome these challenges, scenario-unaware and scenario semi-aware RM construction approaches offer a potential solution by bypassing the need for detailed environmental maps. Instead, these methods rely on coarse environmental knowledge and sparse wireless channel measurements to approximate the spatial distribution of EM waves. Unlike precomputed scenario-aware models, these methods provide greater adaptability in dynamic, unstructured, and privacy-sensitive environments, where environmental features may be unavailable, unreliable, or constantly evolving. However, achieving accurate RM reconstruction under sparse noisy sampling remains a fundamental challenge, necessitating the development of new techniques that can extract and infer radio propagation characteristics efficiently under uncertainty.

Moreover, an often-overlooked yet fundamental characteristic of RMs is their bidirectional relationship with the surrounding environment. While traditionally employed to represent the spatial distribution of wireless channel characteristics, RMs inherently encode structural information about their environment. Variations in pathloss, ToA, and AoA across space provide implicit cues about the geometry, material properties, and layout of surrounding objects. This intrinsic property extends RM's utility beyond mere communication optimization, making it a key enabler of integrated sensing and communication (ISAC), where wireless communication and environmental perception are seamlessly integrated [29]. However, leveraging RMs for environment inference and environment unware RM construction poses a significant challenge, as it necessitates a robust framework capable of extracting meaningful prior knowledge from coarse environmental information and sparse, noisy wireless measurements. To address these challenges, we formulate RM construction as a Bayesian inverse problem, where the objective is to infer the underlying RM distribution from limited and noisy observations. Bayesian inference provides a principled framework for handling uncertainty and incorporating prior knowledge to improve estimation accuracy. Although maximum a posteriori (MAP) filtering offers a theoretically optimal solution to the Bayesian inverse problem, it requires an accurate prior distribution of the RM, which is often unavailable or difficult to model explicitly [30]. To circumvent this limitation, DMs—which have demonstrated remarkable generative capabilities—offer a promising alternative by learning the prior distribution of RMs across diverse environments. By integrating diffusion models into the Bayesian inference framework, we can construct RMs even when precise BS locations and detailed environmental maps are missing, while simultaneously mitigating the adverse effects of noise in sparse measurements. To this end, we propose RadioDiff-Inverse, a diffusion-enhanced Bayesian inverse framework designed for robust RM reconstruction under coarse environmental knowledge and sparse, noisy measurements. Unlike traditional approaches that rely on deterministic interpolation or predefined environmental models, RadioDiff-Inverse leverages diffusion-based generative modeling to learn the underlying statistical structure of RMs, enabling adaptive RM reconstruction in highly dynamic and uncertain environments. By integrating Bayesian inverse estimation with diffusion-based generative modeling, RadioDiff-Inverse represents a paradigm shift in RM construction, moving beyond static scenario-aware methods to a data-driven, uncertainty-aware framework. The resulting RM not only reconstructs the spatial distribution of wireless channel features but also enables accurate environmental inference—such as detecting the outlines and locations of buildings and obstacles—through the ISAC framework. As a result, RadioDiff-Inverse provides a powerful, scalable solution for next-generation wireless communication systems, particularly in dynamic, unstructured, and privacy-sensitive environments where traditional RM construction methods fail. The main contributions of this paper are summarized as follows.

1) To address scenario-unaware/semi-aware RM construction, we cast RM estimation as a Bayesian inverse problem and develop RadioDiff-Inverse, which performs MAP filtering under sparse and noisy sampling. Treating the task as an inverse problem enables the inference of missing environmental and base-station information and yields robust reconstructions in dynamic, unstructured, and privacy-sensitive settings, beyond deterministic interpolation or predefined environmental models.
2) We instantiate an unconditional generative diffusion model, pretrained on ImageNet and used without task-specific fine-tuning or post-training, to capture a data-driven prior over radio maps across diverse environments. Leveraging diffusion modeling provides a strong, training-free prior that preserves fidelity even when precise environmental details are unavailable.
3) We precisely define the RM and standardize key parameters, such as carrier frequency, bandwidth, transmit power, antenna heights, and articulate how the method serves ISAC using only sparse position–power reports gathered by routine communication measurements. The pipeline reconstructs the map and extracts building outlines via a noise-aware shadow/edge logic, operates as a passive, drop-in ISAC module without bespoke waveforms or specialized feedback, and maintains low reporting overhead that scales with the number of samples while remaining forward-compatible with multi-band and multi-height extensions.
4) Extensive experiments demonstrate state-of-the-art accuracy in radio-map reconstruction, environment perception, and base-station localization from path-loss alone, with strong robustness to sparsity and measurement noise. The proposed method consistently outperforms interpolation baselines and deterministic generative models in the absence of high-precision environmental information, underscoring its applicability to next-generation wireless networks, autonomous systems, and satellite communication infrastructures.

II. Preliminary

A. Bayesian Inverse Problems

Bayesian inverse problems arise in numerous scientific and engineering applications where one seeks to recover an unknown signal or parameter from indirect and often noisy observations [30]. These problems are particularly prevalent in fields such as medical imaging, geophysics, computer vision, and wireless signal processing, where direct measurements of the desired quantities are either infeasible or incomplete. Mathematically, an inverse problem is typically formulated as a system of equations that relate the unknown variable $\boldsymbol{x}$ to the observed data $\boldsymbol{y}$ through a known forward operator A, often corrupted by noise as follows.

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{n} \sim \mathcal{N}(0, \sigma^2 I). \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^D$ represents the unknown quantity to be estimated, $\boldsymbol{y} \in \mathbb{R}^d$ denotes the observed measurements, and $A \in \mathbb{R}^{d\times D}$ is a known transformation matrix encoding the mapping between the two. The term $\boldsymbol{n}$ represents measurement noise, which is commonly assumed to follow a Gaussian distribution with variance σ^2. A fundamental challenge in inverse problems arises when the number of observations is smaller than the number of unknowns (i.e., $d < D$), making the system underdetermined and the solution non-unique. Such problems are often classified as ill-posed, necessitating additional constraints or regularization techniques to obtain a meaningful estimate of $\boldsymbol{x}$.

A principled approach to addressing the ill-posed nature of inverse problems is to adopt a Bayesian inference framework, which incorporates prior knowledge about the unknown variable $\boldsymbol{x}$. Instead of seeking a single deterministic solution, the Bayesian approach formulates the inverse problem as an inference task, where the goal is to estimate the posterior distribution of $\boldsymbol{x}$ given the observations $\boldsymbol{y}$. Using Bayes' theorem, the posterior distribution is expressed as follows.

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x}). \tag{2}$$

where $p(\boldsymbol{x})$ represents the prior distribution, which encodes prior knowledge about the likely values of $\boldsymbol{x}$ based on domain-specific assumptions. The term $p(\boldsymbol{y}|\boldsymbol{x})$ is the likelihood function, which models the probability of observing $\boldsymbol{y}$ given a specific realization of $\boldsymbol{x}$. Under the assumption of Gaussian noise in the measurement process, the likelihood follows a normal distribution as follows.

$$p(\boldsymbol{y}|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{y}|A\boldsymbol{x}, \sigma^2 I). \tag{3}$$

By combining the prior with the likelihood, Bayesian inference provides a posterior distribution that quantifies the uncertainty in the recovered solution. This formulation naturally leads to several inference strategies for estimating $\boldsymbol{x}$. One common approach is Maximum a Posteriori (MAP) estimation, which seeks the most probable value of $\boldsymbol{x}$ by solving:

$$\hat{\boldsymbol{x}}_{\text{MAP}} = \arg\max_{\boldsymbol{x}} p(\boldsymbol{x}|\boldsymbol{y}) = \arg\max_{\boldsymbol{x}} p(\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x}). \tag{4}$$

In cases where the prior is Gaussian, i.e., $p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|0, \Sigma)$, the MAP estimate reduces to a regularized least-squares problem as follows.

$$\hat{\boldsymbol{x}}_{\text{MAP}} = \left(A^T A + \sigma^2 \Sigma^{-1}\right)^{-1} A^T \boldsymbol{y}, \tag{5}$$

which corresponds to Tikhonov regularization in classical inverse problem literature.

B. Score-Based Denoising Diffusion Model

DM have emerged as a powerful class of generative models capable of producing high-quality samples from complex distributions [31]. Among them, score-based diffusion models provide a probabilistic framework in which data is gradually perturbed using a stochastic process, and the generative model learns to reverse this process through score function estimation [32]. Unlike standard DMs, which rely on discrete Markov chains for noise injection and denoising, score-based diffusion models utilize stochastic differential equations (SDEs) to continuously perturb and reconstruct data distributions. This formulation naturally aligns with Bayesian inference, enabling efficient posterior sampling for inverse problems, such as radio map reconstruction. Formally, the forward diffusion process is defined as a stochastic perturbation process, where an initial data distribution $p_0(\boldsymbol{x})$ is gradually transformed into a known prior, typically as Gaussian, via a controlled SDE as follows.

$$d\boldsymbol{x} = f(\boldsymbol{x}, t)dt + g(t)d\boldsymbol{w}, \tag{6}$$

where $f(\boldsymbol{x}, t)$ is the drift term that governs deterministic transformation, $g(t)$ is a time-dependent diffusion coefficient, and $d\boldsymbol{w}$ represents a standard Wiener process. As t progresses from 0 to T, the distribution $p_t(\boldsymbol{x})$ transitions from the data distribution to an isotropic Gaussian distribution, ensuring tractable sampling. The generative process in score-based diffusion models relies on reversing the forward stochastic process, which involves solving a corresponding reverse-time SDE as follows.

$$d\boldsymbol{x} = \left[f(\boldsymbol{x}, t) - g^2(t)\nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x})\right] dt + g(t)d\bar{\boldsymbol{w}}, \tag{7}$$

where $\nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x})$ is the score function, representing the gradient of the log probability density at time t, and $d\bar{\boldsymbol{w}}$ is a standard Wiener process in reverse time. In practice, $p_t(\boldsymbol{x})$ is unknown, so a neural network $s_\theta(\boldsymbol{x}, t)$ is trained to approximate the score function as follows.

$$s_\theta(\boldsymbol{x}, t) \approx \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x}). \tag{8}$$

By leveraging score-matching objectives, the model learns to recover data from noise, effectively inverting the stochastic diffusion process.

Although the reverse-time SDE formulation provides a stochastic sampling procedure, an equivalent probability flow ordinary differential equation (ODE) can be derived as follows.

$$d\boldsymbol{x} = \left[f(\boldsymbol{x}, t) - \frac{1}{2}g^2(t)\nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x})\right] dt. \tag{9}$$

This ODE formulation eliminates stochasticity in sampling, enabling deterministic inference paths akin to denoising diffusion probabilistic models (DDPMs) [31]. Specifically, DDPMs discretize the forward diffusion process into a finite sequence of steps as follows.

$$q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \mathcal{N}(\boldsymbol{x}_t; \alpha_t \boldsymbol{x}_{t-1}, \beta_t \boldsymbol{I}), \tag{10}$$

where α_t and β_t are step-wise coefficients controlling noise injection. The generative process then reconstructs data using a learned noise estimator $\epsilon_\theta(\boldsymbol{x}_t, t)$, which corresponds to the score function as follows.

$$s_\theta(\boldsymbol{x}, t) = -\frac{\epsilon_\theta(\boldsymbol{x}, t)}{\sqrt{1-\bar{\alpha}_t}}. \tag{11}$$

Thus, DDPM can be interpreted as a discrete implementation of score-based diffusion using a variance-preserving diffusion process.

The decoupled diffusion model (DDM) [33], which is employed in the SOTA scenario-aware RM construction method RadioDiff, introduces a structured diffusion framework by decoupling data attenuation and noise injection into two distinct stages. Unlike conventional diffusion models that directly introduce Gaussian noise to the original data, DDM first attenuates the initial state $\boldsymbol{n}_0$ to a zero vector before adding stochastic noise. This forward process is defined as a continuous Markov process, where the transition from $\boldsymbol{n}_0$ to $\boldsymbol{n}_t$ follows a Gaussian distribution as follows.

$$q\left(\boldsymbol{x}_t \mid \boldsymbol{x}_0\right) = \mathcal{N}\left(\gamma_t \boldsymbol{x}_0, \delta_t^2 \boldsymbol{I}\right), \tag{12}$$

where γ_t and δ_t as time-dependent coefficients controlling the attenuation and noise variance, respectively. The differential form of this process follows.

$$d\boldsymbol{x}_t = f_t \boldsymbol{x}_t dt + g_t d\boldsymbol{\epsilon}_t, \tag{13}$$

$$f_t = \frac{d \log \gamma_t}{dt}, \tag{14}$$

$$g_t^2 = \frac{d\delta_t^2}{dt} - 2\ f_t \delta_t^2 \tag{15}$$

where f_t governs the rate of decay and g_t^2 determines the noise accumulation. This structured perturbation stabilizes the diffusion process, mitigating early-stage variance and ensuring more controlled generative dynamics. The reverse process reconstructs $\boldsymbol{x}_0$ from $\boldsymbol{x}_t$ by solving the corresponding stochastic differential equation as follows.

$$d\boldsymbol{x}_t = \left[f_t\boldsymbol{x}_t - g_t^2\nabla_{\boldsymbol{x}}\log q\left(\boldsymbol{x}_t\right)\right]dt + g_t d\bar{\boldsymbol{\epsilon}}_t. \tag{16}$$

By leveraging the decoupling property, DDM enhances stability in both training and inference, ensuring improved sample quality and reduced computational overhead. The transformation from $\boldsymbol{x}_0$ to $\mathbf{0}$ is deterministic, simplifying the forward process as follows.

$$q(\boldsymbol{x}_t|\boldsymbol{x}_0) = \mathcal{N}\left(\boldsymbol{x}_0 + \int_0^t \boldsymbol{f}_t \mathrm{d}t, t\boldsymbol{I}\right), \tag{17}$$

which leads to an efficient reverse sampling scheme as follows.

$$q\left(\boldsymbol{x}_{t-\Delta t} \mid \boldsymbol{x}_t, \boldsymbol{x}_0\right) = \mathcal{N}\left(\boldsymbol{x}_t + \int_t^{t-\Delta t} \boldsymbol{f}_t \ \mathrm{d}t - \frac{\Delta t}{\sqrt{t}}\boldsymbol{\epsilon}_t, \frac{\Delta t(t-\Delta t)}{t}\boldsymbol{I}\right). \tag{18}$$

III. System Model and Problem Formulation

A. System Model

In this work, we consider the task where the RM needs to be constructed in a given area, represented as a grid of size $N \times N$. Each grid cell is assumed to have a constant pathloss value, allowing the RM to be expressed as a pathloss matrix $\boldsymbol{P} \in \mathbb{R}^{N\times N}$. Unlike traditional scenario-aware RM construction, we operate under scenario-unaware conditions, meaning that precise BS location and environmental information are unavailable, while only coarse environmental knowledge and sparse, noisy measurements are provided. Within this region, there exists an unknown BS, whose location, denoted as $\boldsymbol{R}$, is not explicitly given. However, the environment contains some obstacles, which affect the propagation of EM waves. The environmetal obstacles, such as buildings and walls, have fixed positions, varying sizes and shapes, and are composed of materials that reflect, absorb, and diffract EM waves. Similar to prior works [9], [10], [22], the pathloss within their interiors is assumed to be zero. The presence of static obstacles is represented by a matrix $\boldsymbol{h}_s \in \mathbb{R}^{N\times N}$, where $h^s_{i,j} = 0, \forall h^s_{i,j} \in \boldsymbol{h}_s$ indicates the absence of a static obstacle at grid cell (i, j). Unlike fully scenario-aware RM construction, where large-scale environmental features are available, we assume only sparse and noisy pathloss measurements can be obtained. Let $\boldsymbol{y}$ represent the sparse sampled measurements of the true pathloss matrix $\boldsymbol{p}$, with noise $\boldsymbol{n}$ modeled as an additive term as follows.

$$\boldsymbol{y} = \mathcal{S}(\boldsymbol{p}) + \boldsymbol{n}, \tag{19}$$

where $\mathcal{S}(\cdot)$ is a sampling operator that selects a small subset of the total RM points, and $\boldsymbol{n} \sim \mathcal{N}(0, \sigma^2)$ represents measurement noise. The objective is to train a NN $\boldsymbol{\mu}_{\boldsymbol{\theta}}(\cdot)$ with parameters $\boldsymbol{\theta}$ to predict the pathloss matrix $\hat{\boldsymbol{p}}$ using coarse environmental information and sparse, noisy pathloss measurements while not relying on BS location information. The reconstruction accuracy is measured using a criterion function $\mathcal{L}(\hat{\boldsymbol{p}}, \boldsymbol{p})$, which quantifies the difference between the predicted $\hat{\boldsymbol{p}}$ and the ground truth $\boldsymbol{p}$.

By reconstructing the radio map, ISAC is achieved in a passive and plug-and-play manner, converting routine communication measurements into environmental awareness without the need for bespoke waveforms, specialized feedback, or protocol modifications. To extract environmental structure, we estimate a noise-floor threshold based on low-variance regions or repeated samples. Any reconstructed grid cell whose value falls below this threshold is identified as a radio shadow, indicating the likely presence of a signal-blocking obstacle such as a building or wall. This allows the system to deliver both communication-aware pathloss reconstruction and sensing-based building structure inference, all from the same sparse and noisy wireless measurements.

Problem 1:

$$\min_{\boldsymbol{\theta}} \ \mathcal{L}\left(\boldsymbol{p}, \hat{\boldsymbol{p}}\right) + \alpha\mathcal{L}\left(\boldsymbol{h}_s, \mathbb{I}\left(\hat{\boldsymbol{p}} \le \tau\right)\right), \tag{20}$$

$$s.t. \ \ \hat{\boldsymbol{p}} = \boldsymbol{\mu}_{\boldsymbol{\theta}}(\boldsymbol{y}), \tag{20a}$$

where α is a weighting factor, $\mathbb{I}(\cdot)$ is indicating function whose value is 1 if the input is true, else is 0, and τ is the threshold factor. This formulation captures the key challenges of scenario-unaware RM construction, where no explicit BS location information is available, only coarse environmental features are provided, and sparse, noisy measurements must be leveraged to accurately reconstruct the RM and sense the buildings.

B. Problem Analysis

In the context of RM construction under the condition of known partial environmental information and noisy sparse sampling, we aim to reconstruct an RM based on limited, noisy observations. The RM is a grid representation of wireless channel characteristics, which can be flattened into a 1-dimensional vector $\boldsymbol{x} \in \mathbb{R}^{N^2}$, where N represents the grid size. Each element of $\boldsymbol{x}$ corresponds to a pathloss value or another characteristic (e.g., time of arrival, angle of arrival) at a specific location in the grid. However, due to practical constraints, only partial environmental information is available for constructing the RM. Additionally, the data we receive is sparse, meaning that only a subset of the grid cells are measured, and these measurements come with noise. This situation leads to a problem that can be formulated as a Bayesian inverse problem of the form as follows.

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{n}, \tag{21}$$

where $\boldsymbol{y} \in \mathbb{R}^d$ is the observed data vector, which includes both the partial environmental information (e.g., coarse grid features) and noisy sparse sampling information. The data $\boldsymbol{y}$ is a sparse and noisy version of the true RM. $\boldsymbol{n} \in \mathbb{R}^d$ is the noise vector in the sparse sampling measurements, assumed to follow a Gaussian distribution $\boldsymbol{n} \sim \mathcal{N}(0, \sigma^2 I)$. $A \in \mathbb{R}^{d\times N^2}$ is the mask matrix, where only the elements on the diagonal are 1, and all other elements are 0. The mask matrix represents the sampling pattern. If the diagonal element $A_{i,i} = 1$, it indicates that the corresponding element x_i of the flattened RM has been observed and is not masked; otherwise, if $A_{i,i} = 0$, the corresponding element x_i has been masked and is unavailable in the measurement. In this setup, we have a sparse and noisy observation $\boldsymbol{y}$, and our goal is to infer the full RM vector $\boldsymbol{x}$ from these incomplete measurements. The Bayesian inverse problem is the task of inferring $\boldsymbol{x}$ given the observed data $\boldsymbol{y}$, the known mask matrix A, and the noise model. We express this as a posterior distribution as follows.

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x}), \tag{22}$$

where $p(\boldsymbol{x})$ is the prior distribution that encodes the statistical properties or assumptions about the RM structure (e.g., smoothness or spatial correlation in wireless channels). $p(\boldsymbol{y}|\boldsymbol{x})$ is the likelihood function, which models the likelihood of observing $\boldsymbol{y}$ given a particular realization of $\boldsymbol{x}$, governed by the noise model $\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{n}$. Since the noise is assumed to be Gaussian, the likelihood is expressed as follows.

$$p(\boldsymbol{y}|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{y}|A\boldsymbol{x}, \sigma^2 I). \tag{23}$$

This problem formulation illustrates how RM construction under coarse environmental knowledge, sparse sampling, and measurement noise can be framed as a Bayesian inverse problem. The goal is to estimate the full RM vector $\boldsymbol{x}$ while accounting for sparse, noisy data and incomplete information, thereby enabling a more robust and accurate RM reconstruction in challenging environments where traditional methods may fail due to data sparsity or noise.

IV. Diffusion Enhanced Bayesian Inverse Estimation

In this section, we give a detailed introduction for the framework of the proposed RadioDiff-Inverse method.

A. Bayesian Filtering

Bayesian filtering provides a framework for recursively updating the belief about a latent variable $\boldsymbol{x}_t$ based on new observations $\boldsymbol{y}_t$. The core of Bayesian filtering consists of two main steps: prediction and update. These steps, when applied iteratively over time, allow for the estimation of the posterior distribution $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:k})$ at each time step k. The prediction step leverages the prior knowledge of how the state evolves over time, while the update step incorporates new observations to adjust the prediction and refine the estimate of the state. The prediction step is mathematically formulated as follows.

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:k-1}) = \int p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})p(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:k-1})d\boldsymbol{x}_{t-1}, \tag{24}$$

where $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ represents the transition model that governs how the state evolves from $\boldsymbol{x}_{t-1}$ to $\boldsymbol{x}_t$, and $p(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:k-1})$ is the prior distribution of the state at time $k-1$. The update step incorporates the new observation $\boldsymbol{y}_t$ to adjust the predicted distribution using Bayes' rule:

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:k}) = \frac{p(\boldsymbol{y}_t|\boldsymbol{x}_t)p(\boldsymbol{x}_t|\boldsymbol{y}_{1:k-1})}{\int p(\boldsymbol{y}_t|\boldsymbol{x}_t)p(\boldsymbol{y}_t|\boldsymbol{y}_{1:k-1})d\boldsymbol{x}_t}, \tag{25}$$

where $p(\boldsymbol{y}_t|\boldsymbol{x}_t)$ is the likelihood, representing the probability of observing $\boldsymbol{y}_t$ given the state $\boldsymbol{x}_t$, and $p(\boldsymbol{y}_t|\boldsymbol{y}_{1:k-1})$ is the normalization term to ensure that the posterior is a valid probability distribution.

Repeatedly executing these two steps allows for Bayesian filtering to solve the Bayesian inverse problem. For systems where both the dynamics and measurements follow a linear Gaussian distribution, the Kalman filter can effectively solve this problem. However, for more complex, nonlinear problems, a more sophisticated method is required to estimate the posterior distribution. In this work, we propose utilizing DMs to extract prior distribution characteristics, which enables the use of Bayesian posterior sampling in a nonlinear setting.

B. Diffusion Models for Bayesian Filtering

DMs provide a natural method for solving Bayesian filtering problems. These models involve a forward diffusion process that gradually adds noise to the data and a backward sampling process that generates posterior samples. In the context of Bayesian filtering, the forward diffusion process corresponds to the prediction step, while the backward sampling process corresponds to the update step. The forward diffusion process can be expressed as follows.

$$\boldsymbol{x}_t = a_t\boldsymbol{x}_{t-1} + b_t\boldsymbol{n}_t, \tag{26}$$

where $\boldsymbol{n}_t$ is independent standard normal noise, and a_t and b_t are scaling factors for the state and the noise. According to [34], [35], the noise sharing method is used as the diffusion process in the sequence of $\boldsymbol{y}$ as follows.

$$\boldsymbol{y}_t = a_k\boldsymbol{y}_{t-1} + b_t\boldsymbol{A}\boldsymbol{n}_t. \tag{27}$$

Therefore since the $\boldsymbol{y}_0 = \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{A}\boldsymbol{x}_0, \sigma^2\boldsymbol{I})$, the observation model is similarly affected by noise, leading to the noisy observations as follows.

$$\boldsymbol{y}_t \sim \mathcal{N}(\boldsymbol{A}\boldsymbol{x}_t, c_t^2\sigma_t^2\boldsymbol{I}), \tag{28}$$

where $c_t = a_1a_2\cdots a_t$, and σ^2 represents the noise variance. This formulation establishes the Bayesian posterior sampling problem as a reverse-time Bayesian filtering problem, which is well-studied with existing algorithms and solutions. By following the prediction and update steps in Bayesian filtering, we compute the marginal posterior distribution $p_\theta(\boldsymbol{x}_t|\boldsymbol{y}_{t:N})$, which enables us to recover the initial latent state $\boldsymbol{x}_0$ from the noisy measurements.

1) Generation Sequence: To generate the sequence of states $\boldsymbol{x}_t$, we need to obtain the sequence $\boldsymbol{y}_{t:N}$. Although we have full access to the initial noisy measurement $\boldsymbol{y}_0$, which represents the coarse environmental features and noisy sparse measurements, this alone is not enough for posterior sampling according to Bayesian filtering principles. Each observation $\boldsymbol{y}_t$ should be sampled from a Gaussian distribution with the mean defined by $\boldsymbol{A}\boldsymbol{x}_t$.

Given this, we use the backward sampling process to sample $\boldsymbol{y}_t$ from the sequence $\boldsymbol{y}_{t+1}$. This backward process allows us to sample from the posterior distribution $p_\theta(\boldsymbol{y}_t|\boldsymbol{x}_t)$. As stated in [35], the equation for sampling $\boldsymbol{y}_t$ is as follows.

$$\boldsymbol{y}_t = u_t\boldsymbol{y}_0 + v_t\boldsymbol{y}_t + w_t \cdot \boldsymbol{A}\boldsymbol{n}_t, \tag{29}$$

where u_t, v_t, and w_t are determined by the diffusion parameters. This process is computed recursively, where the initial

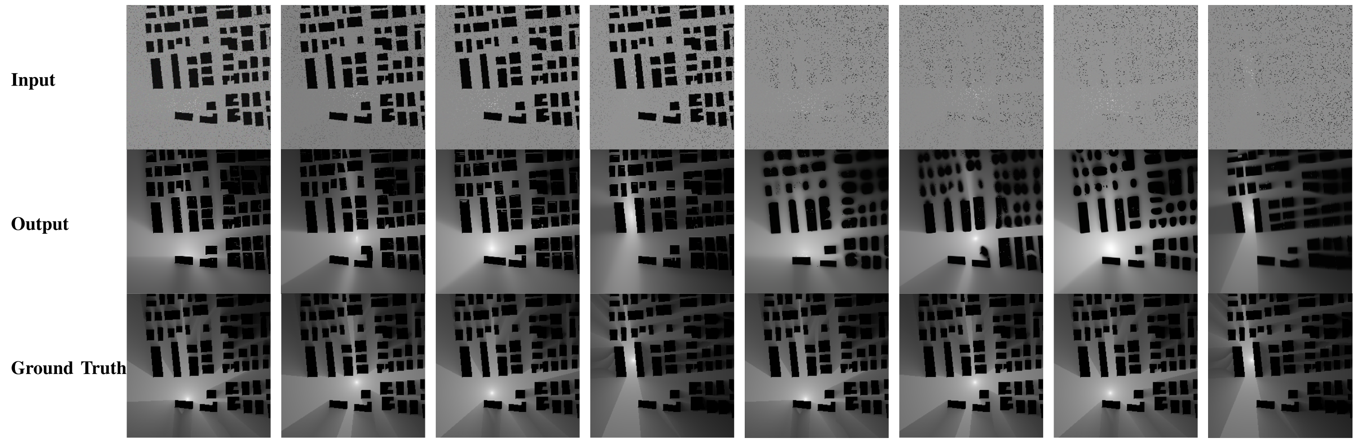


Fig. 2. Comparison of different scenarios. Row 1: Input measurements; Row 2: Output RMs; Row 3: Ground truth.

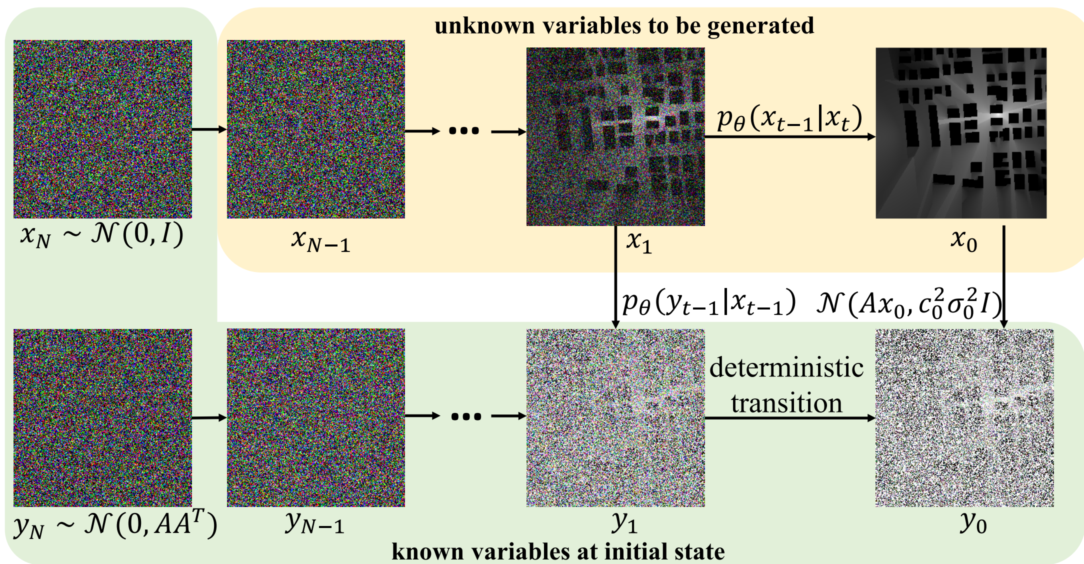


Fig. 3. The illustration of inferencing sequence.

point $\boldsymbol{y}_N$ is sampled from $\mathcal{N}(0, \boldsymbol{A}\boldsymbol{A}^T)$, and subsequent values are generated based on the prior $\boldsymbol{y}_0$.

2) Backward Sequence: The backward sampling sequence is crucial for recursively sampling $\boldsymbol{x}_t$ from the noisy observations. Given that $\boldsymbol{x}_N$ is approximately a standard Gaussian, we can express the posterior distribution of $\boldsymbol{x}_N$ given the observations $\boldsymbol{y}_N$ as follows.

$$\boldsymbol{x}_N \sim p_\theta(\boldsymbol{x}_N|\boldsymbol{y}_N) \propto p_\theta(\boldsymbol{x}_N) \cdot p_\theta(\boldsymbol{y}_N|\boldsymbol{x}_N). \tag{30}$$

The posterior distribution $p_{\boldsymbol{\theta}}(\boldsymbol{x}_N|\boldsymbol{y}_N)$ can be derived from the likelihood $p_{\boldsymbol{\theta}}(\boldsymbol{y}_N|\boldsymbol{x}_N)$, which is expressed as follows.

$$p_{\boldsymbol{\theta}}(\boldsymbol{y}_N|\boldsymbol{x}_N) = q(\boldsymbol{y}_N|\boldsymbol{x}_N)\mathcal{N}(\boldsymbol{A}\boldsymbol{x}_N, c_N^2\sigma^2\boldsymbol{I}), \tag{31}$$

where $c_N = a_1a_2\cdots a_N$. The posterior distribution $p_{\boldsymbol{\theta}}(\boldsymbol{x}_N|\boldsymbol{y}_N)$ is Gaussian and can therefore be expressed in closed form. Next, we recursively sample $\boldsymbol{x}_{t-1}$ conditioned on $\boldsymbol{x}_t$ and $\boldsymbol{y}_{t-1}$. The likelihood $p_{\boldsymbol{\theta}}(\boldsymbol{y}_{t-1}|\boldsymbol{x}_{t-1})$ is given as follows.

$$p_{\boldsymbol{\theta}}(\boldsymbol{y}_{t-1}|\boldsymbol{x}_{t-1}) = q(\boldsymbol{y}_{t-1}|\boldsymbol{x}_{t-1}) = \mathcal{N}(\boldsymbol{A}\boldsymbol{x}_{t-1}, c_{t-1}^2\sigma^2\boldsymbol{I}), \tag{32}$$

where $c_{t-1} = a_1a_2\cdots a_{t-1}$. Additionally, the posterior $p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)$ is determined by the score function as follows.

$$p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) = \mathcal{N}\left(u_t\hat{\boldsymbol{x}}_0(\boldsymbol{x}_t) + v_t\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t_t), w_t^2\boldsymbol{I}\right), \tag{33}$$

where $\hat{\boldsymbol{x}}_0(\boldsymbol{x}_t) := \frac{\boldsymbol{x}_t + d_t^2\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t_t)}{c_t}$ is the conditional expectation of $\boldsymbol{x}_0$ given $\boldsymbol{x}_t$, computed using Tweedie's formula. The parameters c_t, d_t, u_t, v_t, and w_t are determined through the process of unconditional diffusion sampling. Specific values for these parameters in the DDPM and DDM frameworks are provided in Section II-B. Now, we can compute the posterior distribution as follows.

$$\begin{aligned} p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{y}_{t-1}) &= \frac{p_{\boldsymbol{\theta}}\left(\boldsymbol{x}_{t-1}, \boldsymbol{y}_{t-1}|\boldsymbol{x}_t\right)}{p_{\boldsymbol{\theta}}\left(\boldsymbol{y}_{t-1}|\boldsymbol{x}_t\right)}, \\ &\propto p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) \cdot p_{\boldsymbol{\theta}}(\boldsymbol{y}_{t-1}|\boldsymbol{x}_{t-1}). \end{aligned} \tag{34}$$

The process of sampling $\boldsymbol{x}_{t-1}$ based on $\boldsymbol{x}_t$ and $\boldsymbol{y}_{t-1}$ follows the probability distribution described in (34). As outlined in [30], Monte Carlo sampling is employed to estimate the desired quantity, as is shown in Fig. 4. In Section II-B, it is

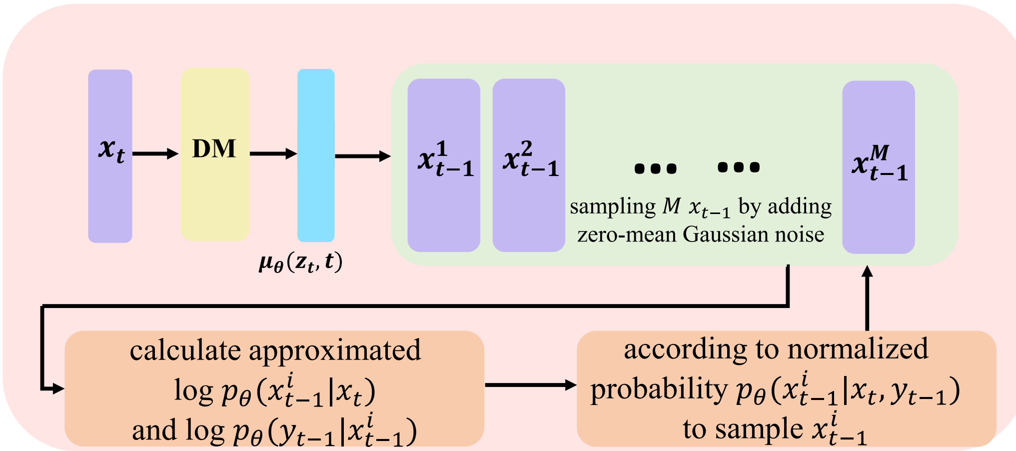


Fig. 4. The illustration of denoising procedure of RadioDiff-Inverse.

described that a trained DM $\mu_{\boldsymbol{\theta}}$ can sample $\boldsymbol{x}_{t-1}$ from $\boldsymbol{x}_t$ using the following distribution.

$$\mathcal{N}(\boldsymbol{x}_{t-1}|\mu_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t), \sigma_t^2). \tag{35}$$

The procedure begins by sampling M instances of $\boldsymbol{x}_{t-1}^i$, where $i \in \{1, 2, \cdots, M\}$, according to the distribution given by (35). Next, the log-probability $p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}^i|\boldsymbol{x}_t)$ is approximated using the following expression.

$$\log p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}^i|\boldsymbol{x}_t) \propto -\frac{\|\boldsymbol{x}_{t-1}^i - \mu_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t)\|^2}{2\sigma_t^2}. \tag{36}$$

Since $\boldsymbol{x}_t$ is given, there is no need to include the superscript on $\boldsymbol{x}_t$. Given $\boldsymbol{y}_{t-1}$, which is generated as described in Section IV-B1, and the sequence $\{\boldsymbol{x}_{t-1}^i\}_{i=1}^M$, the log-probability $p_{\boldsymbol{\theta}}(\boldsymbol{y}_{t-1}|\boldsymbol{x}_{t-1}^i)$ is expressed as follows.

$$\log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{t-1}|\boldsymbol{x}_{t-1}^i) \propto -\frac{\|\boldsymbol{y}_{t-1} - \boldsymbol{A}\boldsymbol{x}_{t-1}^i\|^2}{2c_t^2\sigma_t^2}. \tag{37}$$

Thus, the combined log-probability is as follows.

$$\begin{aligned}\log p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}^i|\boldsymbol{x}_t, \boldsymbol{y}_{t-1}) &\propto \log p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}^i|\boldsymbol{x}_t) \\ &\quad + \log p_{\boldsymbol{\theta}}(\boldsymbol{y}_{t-1}|\boldsymbol{x}_{t-1}^i), \qquad (38)\\ &\propto -\left(\frac{\|\boldsymbol{x}_{t-1}^i - \mu_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t)\|^2}{2\sigma_t^2}\right. \\ &\quad \left. + \frac{\|\boldsymbol{y}_{t-1} - \boldsymbol{A}\boldsymbol{x}_{t-1}^i\|^2}{2c_t^2\sigma_t^2}\right). \qquad (39)\end{aligned}$$

To finalize the sampling process, each of $\boldsymbol{x}_{t-1}^i$ is sampled by probability $p_{\boldsymbol{\theta}}^{sample}(\boldsymbol{x}_{t-1}^i|\boldsymbol{x}_t, \boldsymbol{y}_{t-1})$ as follows.

$$p_{\boldsymbol{\theta}}^{sample}(\boldsymbol{x}_{t-1}^i|\boldsymbol{x}_t, \boldsymbol{y}_{t-1}) = \frac{p_{\boldsymbol{\theta}}^{wgt}(\boldsymbol{x}_{t-1}^i|\boldsymbol{x}_t, \boldsymbol{y}_{t-1})}{\sum_{j=1}^M p_{\boldsymbol{\theta}}^{wgt}(\boldsymbol{x}_{t-1}^j|\boldsymbol{x}_t, \boldsymbol{y}_{t-1})}, \tag{40}$$

$$p_{\boldsymbol{\theta}}^{wgt}(\boldsymbol{x}_{t-1}^i|\boldsymbol{x}_t, \boldsymbol{y}_{t-1}) = e^{-\left(\frac{\|\boldsymbol{x}_{t-1}^i - \mu_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t)\|^2}{2\sigma_t^2} + \frac{\|\boldsymbol{y}_{t-1} - \boldsymbol{A}\boldsymbol{x}_{t-1}^i\|^2}{2c_t^2\sigma_t^2}\right)}. \tag{41}$$

The total sampling procedure from the $\boldsymbol{y}_0$ to obtain the $\boldsymbol{x}_0$ is shown in Fig. 3.

Remark 1: The RadioDiff-Inverse is a training-free framework, where the DM pre-trained on Imagenet without any task-specific fine-tune or post-training is used as the backbone NN of $\mu_{\boldsymbol{\theta}}$. The detail of how to train a DM can be found in [36].

V. Experiments

A. Dataset

This study employs the DPM (Dynamic Pathloss Map) dataset from the RM pathloss construction challenge to evaluate our proposed method. The dataset encompasses 700 unique radio maps, each containing distinct geographic information and building configurations. Each map incorporates 80 transmitter locations with corresponding ground truth measurements and features between 50 and 150 buildings. We designate 200 maps for testing purposes, ensuring no overlap in terrain information between training and test sets to maintain evaluation integrity.

The dataset represents diverse urban environments, sourced from OpenStreetMap, including major metropolitan areas such as Ankara, Berlin, Glasgow, Ljubljana, London, and Tel Aviv. Environmental parameters are standardized across all maps: both transmitter and receiver heights are fixed at 1.5 meters, while building heights are uniformly set at 25 meters. Each map is discretized into a 256×256 pixel binary morphological image, where each pixel represents a $1m^2$ area. The binary values denote building presence (1) or absence (0), providing a precise representation of the urban landscape.

Transmitter positions are encoded using two-dimensional numerical coordinates and represented in morphological images through binary indicators, where the transmitter's location is marked as 1 and all other positions as 0. The transmission parameters are standardized with 23 dBm transmitter power and 5.9 GHz carrier frequency. Ground truth radio maps (RMs) are generated using Maxwell's equations, accounting for electromagnetic ray reflection and diffraction phenomena. The dataset provides two variants of ground truth maps: Static Radio Maps (SRM), which consider only the electromagnetic interactions with static building structures, and Dynamic Radio Maps (DRM), which incorporate both static building effects and the influence of randomly distributed vehicles along roadways.

While this dataset provides high-fidelity signal strength (pathloss) maps, it does not include sensing-related parameters such as Angle of Arrival (AoA), Angle of Departure (AoD), or delay spread, which would be part of a comprehensive ISAC radio map. This focus on signal strength is a limitation imposed by the lack of publicly available, large-scale datasets that jointly map both communication and sensing parameters. Our proposed framework is designed to be extensible, and its application to such richer, multi-parameter ISAC maps remains a critical direction for future work, pending the availability of such data.

This comprehensive dataset enables rigorous evaluation of radio map reconstruction methods under diverse urban scenarios and propagation conditions, providing a robust foundation for assessing the performance of our proposed approach.

B. Evaluation Framework

We establish a comprehensive evaluation framework to assess the performance of our proposed diffusion model, focusing on two primary aspects: the fidelity of the reconstructed radio map and the accuracy of environmental perception (specifically, building detection). This framework enables a systematic assessment of our method's effectiveness across various operational scenarios relevant to ISAC applications.

1) Reconstruction Quality Metrics and Building Detection Accuracy: To quantify the fidelity of reconstructed radio maps and the accuracy of building detection, we employ four key metrics.

For assessing the overall reconstruction quality, we use: The Peak Signal-to-Noise Ratio (PSNR) measures the logarithmic ratio between maximum possible signal power and noise power, defined as:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right) \tag{42}$$

where MSE denotes mean squared error and MAX represents the maximum signal value.

The Structural Similarity Index (SSIM) evaluates perceptual similarity. Its design rationale does not stem from optimization principles but is rather a principled perceptual model designed to empirically mimic the Human Visual System (HVS) by comparing three components: luminance, contrast, and structure [37]. The index is defined as:

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{43}$$

where μ_x and μ_y are the means, σ_x^2 and σ_y^2 are the variances of x and y, σ_{xy} is the covariance of x and y, and C_1, C_2 are stabilization constants.

The Root Mean Square Error (RMSE) quantifies the absolute reconstruction error in the original signal units (e.g., dBm):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \tag{44}$$

where x_i represents the ground truth pixel value, $\hat{x}_i$ is the reconstructed pixel value, and N is the total number of pixels in the map.

To evaluate the accuracy of building outline perception derived from the reconstructed radio map, we utilize the Intersection over Union (IoU) metric. This metric is commonly used in object detection and segmentation tasks. In our context, we first apply a threshold τ (as described in Section III and Equation (1)) to the reconstructed map $\hat{\boldsymbol{p}}$ to obtain a binary prediction mask $\hat{\boldsymbol{B}} = \mathbb{I}(\hat{\boldsymbol{p}} \leq \tau)$, where pixels below the threshold are predicted as building shadows (value 1) and others as non-shadow (value 0). We compare this prediction mask $\hat{\boldsymbol{B}}$ with the ground truth building mask $\boldsymbol{B}$. The IoU is then calculated as the ratio of the area of overlap to the area of union between these two binary masks:

$$\mathrm{IoU} = \frac{\mathrm{Area}(\hat{\boldsymbol{B}} \cap \boldsymbol{B})}{\mathrm{Area}(\hat{\boldsymbol{B}} \cup \boldsymbol{B})} = \frac{\sum_{i=1}^{N}(\hat{B}_i \cdot B_i)}{\sum_{i=1}^{N}\max(\hat{B}_i, B_i)} \tag{45}$$

where $\hat{B}_i$ and B_i are the binary values (0 or 1) of the i-th pixel in the predicted and ground truth masks, respectively. A higher IoU value indicates better accuracy in detecting the building footprints from the radio map.

Our evaluation framework systematically assesses performance using these four metrics (PSNR, SSIM, RMSE, and IoU) across diverse operational scenarios, including varying building presence conditions ('Known' vs. 'Unknown'), different sampling strategies ('1 × 1' vs. '2 × 3'), mask ratios (0.50 to 0.95), and noise levels (0.01 to 0.09). This comprehensive approach ensures thorough performance characterization across scenarios representative of real-world ISAC applications.

C. Implementation Details

We implement our radio map reconstruction framework using PyTorch, leveraging a pre-trained diffusion model initially developed for ImageNet classification. The implementation balances computational efficiency with reconstruction fidelity, making it suitable for both offline analysis and semi-real-time applications in electromagnetic environment assessment.

1) Model Architecture: The core architecture employs a modified U-Net backbone designed for optimal radio map reconstruction. The network is configured with a base channel dimension of 256 and incorporates two residual blocks per resolution level. Operating at an input resolution of 256×256 pixels, the architecture integrates multi-head attention mechanisms at 32×32, 16×16, and 8×8 resolutions, utilizing 4 attention heads with 64 channels per head. Key architectural enhancements include scale-shift normalization with residual connections, adaptive feature fusion across resolution levels, and progressive upsampling with skip connections.

2) Diffusion Process Configuration: The diffusion process employs a linear noise schedule with DDIM sampling method [38], configured for 1000 timesteps. The process utilizes ϵ-prediction with learned range variance for optimal noise prediction. Training optimizations incorporate gradient clipping during denoising, preservation of original timestep scaling, and

TABLE I

PERFORMANCE COMPARISON WITH RANDOM PIXEL-WISE MASKING (NOISE LEVEL = 0.01)

Sampling Setting	Method			
	Ours	RadioUNet	RME-GAN	Interp.
PSNR (↑)				
Aware (70% Missing)	**32.34**	26.68	23.37	22.47
Aware (80% Missing)	**30.29**	25.43	25.51	22.25
Aware (90% Missing)	**26.33**	24.16	25.63	22.47
Unaware (70% Missing)	**32.37**	22.28	21.76	19.76
Unaware (80% Missing)	**30.10**	21.28	20.03	19.49
Unaware (90% Missing)	**26.58**	20.05	19.82	19.67
SSIM (↑)				
Aware (70% Missing)	**0.9680**	0.9195	0.7485	0.7906
Aware (80% Missing)	**0.9495**	0.8988	0.7885	0.7767
Aware (90% Missing)	**0.8913**	0.8687	0.7580	0.8127
Unaware (70% Missing)	**0.9682**	0.7684	0.7257	0.6068
Unaware (80% Missing)	**0.9495**	0.7538	0.7157	0.5453
Unaware (90% Missing)	**0.8941**	0.7275	0.7622	0.6013
RMSE (↓)				
Aware (70% Missing)	**0.0493**	0.0872	0.1425	0.1539
Aware (80% Missing)	**0.0624**	0.0872	0.1264	0.1578
Aware (90% Missing)	**0.0979**	0.1009	0.1091	0.1539
Unaware (70% Missing)	**0.0492**	0.1245	0.1654	0.2081
Unaware (80% Missing)	**0.0637**	0.1218	0.2039	0.2147
Unaware (90% Missing)	**0.0956**	0.1221	0.2080	0.2104
IoU (↑)				
Unaware (70% Missing)	**0.9052**	0.1823	0.5486	0.1268
Unaware (80% Missing)	**0.8944**	0.1756	0.2656	0.1413
Unaware (90% Missing)	**0.8469**	0.1558	0.5495	0.1536

TABLE II

PERFORMANCE COMPARISON WITH RANDOM PIXEL-WISE MASKING (NOISE LEVEL = 0.05)

Sampling Setting	Method			
	Ours	RadioUNet	RME-GAN	Interp.
PSNR (↑)				
Aware (70% Missing)	**35.29**	22.19	23.20	22.26
Aware (80% Missing)	**32.16**	22.22	25.34	22.35
Aware (90% Missing)	**27.44**	22.04	26.17	22.19
Unaware (70% Missing)	**34.65**	19.54	22.02	19.47
Unaware (80% Missing)	**31.52**	19.76	20.67	19.22
Unaware (90% Missing)	**26.80**	19.73	18.99	19.57
SSIM (↑)				
Aware (70% Missing)	**0.9840**	0.9187	0.7743	0.7950
Aware (80% Missing)	**0.9748**	0.9183	0.7889	0.7967
Aware (90% Missing)	**0.9376**	0.9167	0.7628	0.7923
Unaware (70% Missing)	**0.9829**	0.8635	0.7219	0.5784
Unaware (80% Missing)	**0.9737**	0.8808	0.7189	0.5565
Unaware (90% Missing)	**0.9365**	0.8885	0.5616	0.5895
RMSE (↓)				
Aware (70% Missing)	**0.0499**	0.0872	0.1402	0.1581
Aware (80% Missing)	**0.0630**	0.0869	0.1118	0.1565
Aware (90% Missing)	**0.0952**	0.1017	0.1009	0.1600
Unaware (70% Missing)	**0.0494**	0.1246	0.1601	0.2155
Unaware (80% Missing)	**0.0630**	0.1226	0.1884	0.2217
Unaware (90% Missing)	**0.0945**	0.1226	0.2285	0.2126
IoU (↑)				
Unaware (70% Missing)	**0.9041**	0.1972	0.6388	0.1192
Unaware (80% Missing)	**0.8950**	0.1597	0.4517	0.1196
Unaware (90% Missing)	**0.8472**	0.1404	0.3115	0.1150

adaptive learning rate scheduling. This configuration ensures stable training and efficient sampling while maintaining high reconstruction quality.

3) Computational Environment: Our experiments are conducted on NVIDIA H100 GPUs with 80GB HBM3 memory per unit, supported by 512GB DDR5 system memory. The implementation achieves an average reconstruction time of 100 seconds per 256×256 map, with the capability to process 16 maps simultaneously. Peak memory usage is maintained at 45GB, demonstrating efficient resource utilization. The modular architecture design facilitates easy adaptation to different input resolutions and environmental configurations while maintaining consistent performance characteristics.

D. Comparative Analysis

We evaluate our approach against state-of-the-art approaches in radio map reconstruction, including both deep learning-based and traditional methods. For fairness, all compared methods use identical training and testing datasets.

1) Comparison Methods:

- **RadioUNet** [9]: A convolutional U-Net architecture that learns environmental characteristics for radio map reconstruction. As a supervised learning approach, RadioUNet represents the current standard in non-generative neural network solutions.
- **RME-GAN** [22]: A conditional GAN architecture that generates radio maps from environmental features. While the original implementation incorporates pathloss measurements, our comparison uses only environmental features to maintain consistency with our method's inputs.
- **Kriging Interpolation**: A geostatistical technique that estimates unknown values based on spatial correlation of sampled points. We implement ordinary Kriging with exponential variogram models.

2) Evaluation Protocol: All neural network models were trained until convergence with early stopping based on validation loss. We maintain consistent hyperparameter optimization procedures across all learning-based methods. The evaluation metrics include:

- Peak Signal-to-Noise Ratio (PSNR)
- Structural Similarity Index (SSIM)
- Root Mean Square Error (RMSE)
- Intersection over Union (IoU)

For statistical significance, we report mean values across 5 independent runs with different random seeds, along with 95% confidence intervals.

E. Experimental Results and Analysis

We conduct comprehensive experiments to evaluate our proposed inverse method for radio map reconstruction. The evaluation encompasses four key aspects: 1) the impact of building information on reconstruction accuracy, 2) the effectiveness of different sampling strategies, 3) the influence of mask ratios, and 4) the robustness against noise. Tables I–III present the quantitative results across these different experimental configurations.

We conduct comprehensive experiments to evaluate our proposed inverse method for radio map reconstruction. The

TABLE III
PERFORMANCE COMPARISON WITH RANDOM PIXEL-WISE MASKING (NOISE LEVEL = 0.09)

Sampling Setting	Method			
	Ours	RadioUNet	RME-GAN	Interp.
PSNR (↑)				
Aware (70% Missing)	**33.35**	21.99	22.04	21.83
Aware (80% Missing)	**31.82**	21.84	23.41	22.07
Aware (90% Missing)	**28.17**	21.82	25.57	22.04
Unaware (70% Missing)	**32.71**	19.20	20.68	18.87
Unaware (80% Missing)	**31.18**	19.40	19.37	18.93
Unaware (90% Missing)	**27.53**	19.45	19.27	18.94
SSIM (↑)				
Aware (70% Missing)	0.8483	**0.9159**	0.7625	0.7703
Aware (80% Missing)	0.8771	**0.9145**	0.7701	0.7754
Aware (90% Missing)	0.8501	**0.9144**	0.7464	0.7723
Unaware (70% Missing)	0.8472	**0.8551**	0.6982	0.5425
Unaware (80% Missing)	0.8760	**0.8740**	0.6937	0.5453
Unaware (90% Missing)	0.8490	**0.8843**	0.4760	0.5505
RMSE (↓)				
Aware (70% Missing)	**0.0534**	0.0892	0.1601	0.1661
Aware (80% Missing)	**0.0649**	0.0898	0.1368	0.1614
Aware (90% Missing)	**0.0956**	0.1030	0.1064	0.1620
Unaware (70% Missing)	**0.0548**	0.1300	0.1878	0.2310
Unaware (80% Missing)	**0.0664**	0.1274	0.2185	0.2291
Unaware (90% Missing)	**0.0967**	0.1265	0.2223	0.2290
IoU (↑)				
Unaware (70% Missing)	**0.9011**	0.1104	0.5980	0.0573
Unaware (80% Missing)	**0.8933**	0.1328	0.1705	0.0688
Unaware (90% Missing)	**0.8401**	0.1234	0.4085	0.0714

evaluation encompasses four key aspects: 1) the impact of building information on reconstruction accuracy, 2) the effectiveness of different sampling strategies, 3) the influence of mask ratios, and 4) the robustness against noise. Tables I–III present the quantitative results across these different experimental configurations.

1) Impact of Building Information: The experimental results demonstrate that building information significantly enhances radio map reconstruction accuracy. This is visually evident in Figure 6 and 7, where the solid lines (‘Known’ scenario, equivalent to ‘Aware’ in the tables) consistently outperform the dashed lines (‘Unknown’ scenario, equivalent to ‘Unaware’). Our analysis reveals a consistent performance advantage of scenario-aware generation (with building information) over scenario-unaware approaches across all test configurations. At low mask ratios (70%) and low noise (0.01, see Table I), scenario-aware generation achieves superior performance with a PSNR of 34.55dB and SSIM of 0.9847, compared to scenario-unaware generation’s 33.91dB PSNR and 0.9836 SSIM. This performance gap becomes more pronounced as the reconstruction task grows more challenging with higher mask ratios. At 80% masking (Noise Level 0.01), the scenario-aware approach maintains a PSNR of 32.57dB, outperforming its unaware counterpart by 0.64dB (31.93dB). The advantage extends to structural preservation metrics, with scenario-aware generation consistently achieving higher SSIM values (0.9765 vs. 0.9754 at 80% masking, Noise Level 0.01), indicating better preservation of critical radio map features. The IoU metric, evaluated only in the ‘Unaware’ scenario (as shown in Figure 9), further highlights the challenge of environmental perception without prior building knowledge, achieving 0.9052 at 70% masking (Noise Level 0.01, Table I).

2) Effect of Sampling Strategy: The comparison between random pixel-wise masking (represented by 1×1 sampling in our plots) and structured sampling (represented by 2×3 sampling) reveals distinct characteristics that influence reconstruction quality. Visual comparisons, such as contrasting Figure 6 with its 2×3 counterpart, illustrate these differences. Our method demonstrates remarkable resilience under random pixel-wise masking, maintaining high PSNR values even at challenging mask ratios. The performance advantage of our method over baselines is particularly evident in scenario-aware settings (’Known’, solid lines).

3) Mask Ratio Analysis: The impact of mask ratio reveals important insights into the robustness and scalability of different reconstruction approaches. As mask ratios increase from 70% to 90%, all methods exhibit performance degradation, but with varying degrees of resilience, as clearly visualized by the downward trend in all generated plots (e.g., Figures 6–9). Our proposed method demonstrates remarkable stability, maintaining PSNR above 27dB even at 90% masking in scenario-aware settings (’Known’, solid lines). The performance degradation follows a consistent pattern, with PSNR decreasing from 34.55dB at 70% to 27.83dB at 90% masking under noise level 0.01 (Table I). Notably, while RME-GAN shows competitive performance compared to other baselines at high mask ratios (achieving 25.63dB PSNR at 90% masking in the ‘Known’ scenario, Noise Level 0.01), our method maintains superior structural preservation as evidenced by consistently higher SSIM values (0.9420 vs. 0.7580 for RME-GAN under these conditions).

4) Noise Level Impact: The analysis of noise impact, visualized in Figures 10 through 13, reveals crucial insights into the robustness of different reconstruction methods under a challenging 80% mask ratio (Mask Ratio=0.80). Across all metrics, our proposed method demonstrates exceptional resilience to varying noise levels (0.01-0.09). Its performance curves are consistently superior and significantly flatter than all competing methods, indicating that increasing noise has minimal impact on its accuracy.Specifically, in the PSNR analysis (Figure 10), our method maintains the highest performance. In the ‘Known’ scenario (solid blue line), the PSNR only drops trivially from approximately 30.2dB at 0.01 noise to 29.8dB at 0.09 noise. In sharp contrast, the next best method, RadioUNet (‘Known’, solid green line), starts lower at $\approx$ 26.8dB and degrades more significantly to $\approx$ 24.2dB under high noise. The trend is identical in the ‘Unknown’ scenario (dashed lines), where our method remains stable while competitors’ performance collapses.The SSIM metrics (Figure 11) mirror this robustness. Our method’s ‘Known’ (solid blue) and ‘Unknown’ (dashed blue) SSIM values remain high, staying above 0.93 even at the highest noise level. Conversely, other methods in the ‘Unknown’ setting (e.g., RadioUNet, interpolation) show a steep decline, dropping to values near 0.5.The RMSE plot (Figure 12) provides the most dramatic evidence of robustness. Our method’s RMSE (blue lines) is exceptionally low (at $\approx$ 0.06-0.065) and remains almost perfectly flat across all noise levels in both ‘Known’ and ‘Unknown’ scenarios.

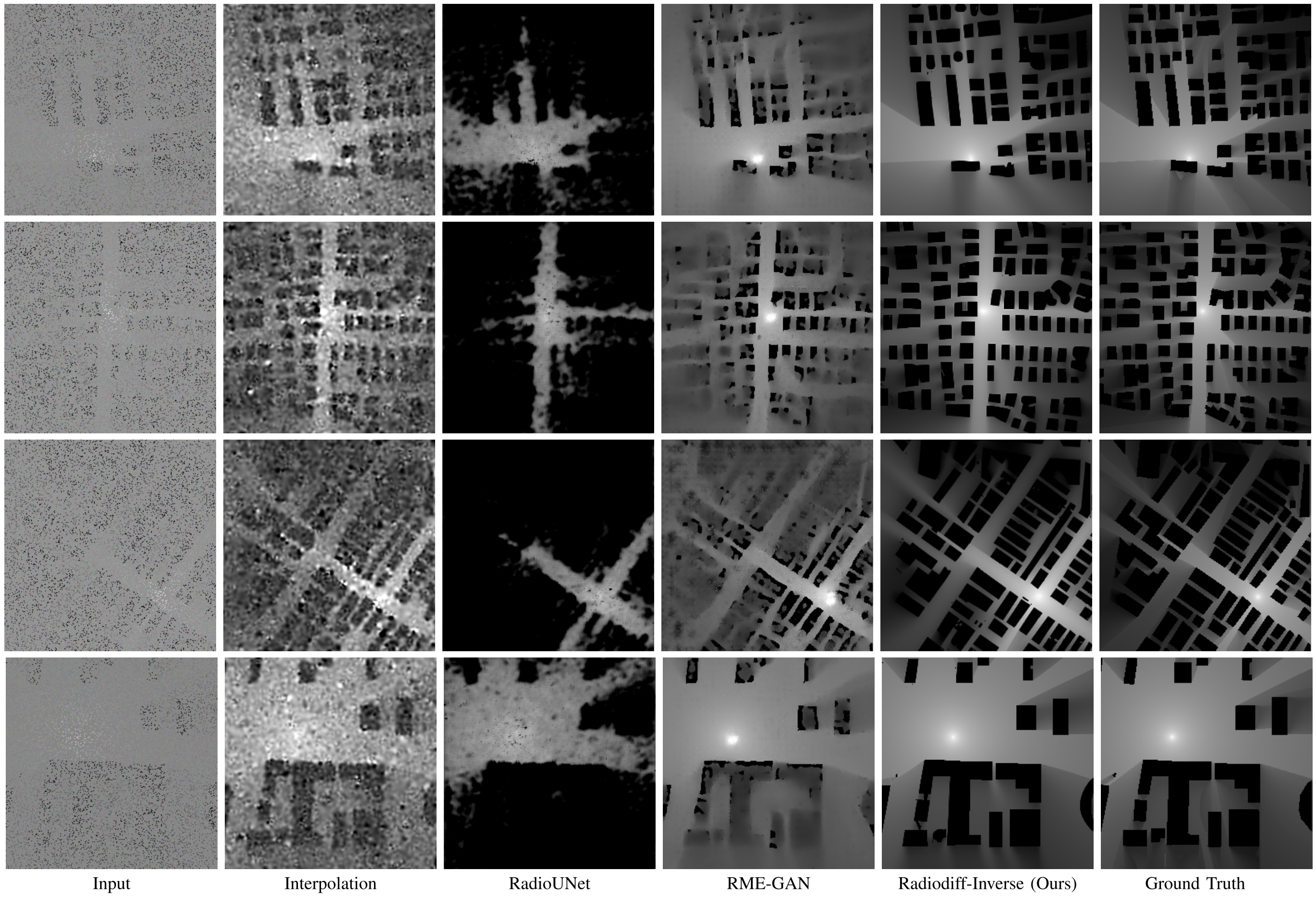


Fig. 5. The comparision of constructed RM between different methods.

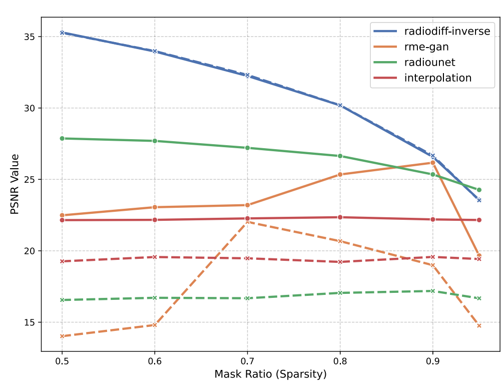


Fig. 6. Performance comparison for PSNR under 1×1 sampling with a noise level of 0.05. Solid lines represent the 'Known' scenario, and dashed lines represent the 'Unknown' scenario.

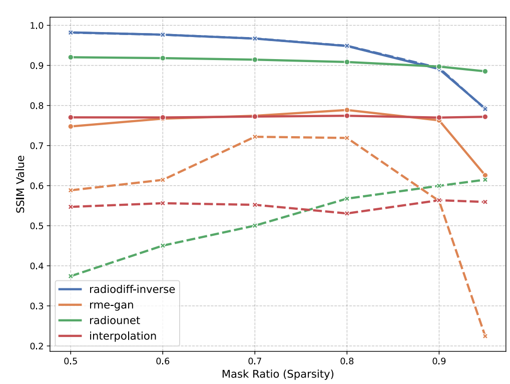


Fig. 7. Performance comparison for SSIM under 1×1 sampling with a noise level of 0.05. Solid lines represent the 'Known' scenario, and dashed lines represent the 'Unknown' scenario.

This demonstrates a highly controlled error propagation. All other methods not only start with higher error but also show a distinct upward trend, indicating that their reconstruction error increases significantly with noise.Finally, the IoU analysis for the 'Unknown' scenario (Figure 13) confirms our method's superiority in identifying unknown transmitter locations. Our method (blue line) achieves a near-constant IoU of ≈ 0.9, unaffected by noise. Competitors perform poorly, with IoU values below 0.2 (e.g., RadioUNet) or exhibiting erratic behavior (e.g., rme-gan).In summary, under high sparsity (80% masking), our proposed method is the only one to consistently deliver high-fidelity reconstructions (high PSNR/SSIM/IoU) and low error (low RMSE), proving its exceptional robustness against varying noise levels.

F. Discussion

A critical consideration for practical deployment is the computational complexity and inference latency, particularly in dynamic ISAC environments. Our proposed method necessitates a full reconstruction of the map when new

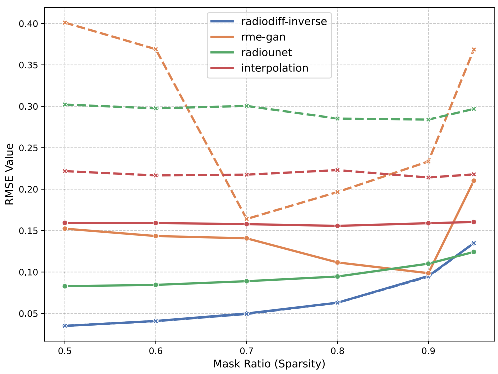


Fig. 8. Performance comparison for RMSE under 1×1 sampling with a noise level of 0.05. Solid lines represent the 'Known' scenario, and dashed lines represent the 'Unknown' scenario.

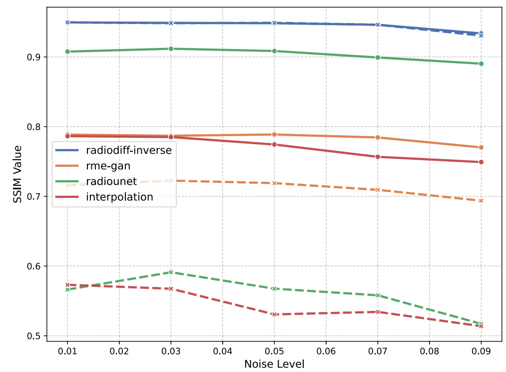


Fig. 11. Performance comparison for SSIM under 1×1 sampling with mask ratio of 0.80. Solid lines represent the 'Known' scenario, and dashed lines represent the 'Unknown' scenario.

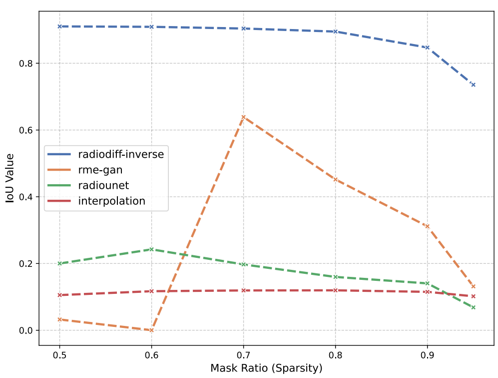


Fig. 9. Performance comparison for IoU under 1×1 sampling with a noise level of 0.05. Note that IoU is evaluated only for the 'Unknown' scenario (dashed lines), as it is not applicable to the 'Known' scenario.

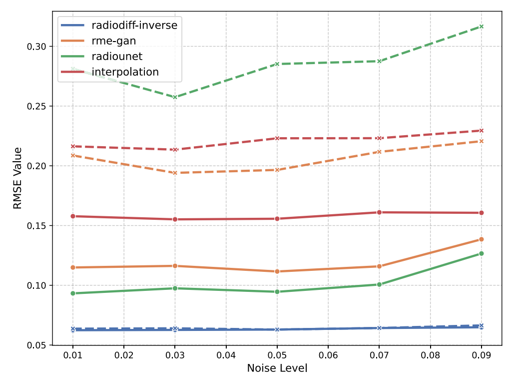


Fig. 12. Performance comparison for RMSE under 1×1 sampling with mask ratio of 0.80. Solid lines represent the 'Known' scenario, and dashed lines represent the 'Unknown' scenario.

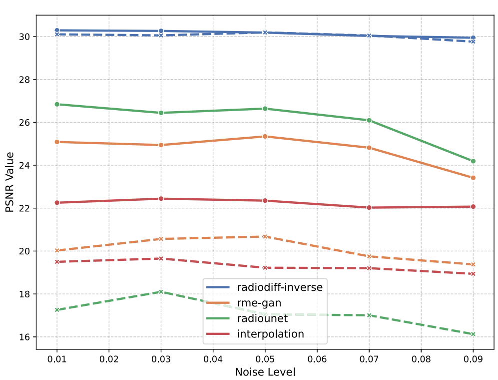


Fig. 10. Performance comparison for PSNR under 1×1 sampling with mask ratio of 0.80. Solid lines represent the 'Known' scenario, and dashed lines represent the 'Unknown' scenario.

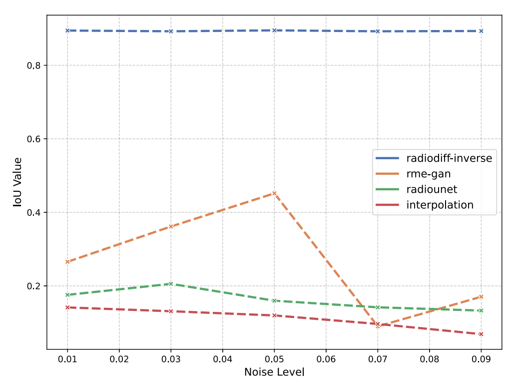


Fig. 13. Performance comparison for IoU under 1×1 sampling with mask ratio of 0.80. Note that IoU is evaluated only for the 'Unknown' scenario, as it is not applicable to the 'Known' scenario.

sparse data is provided and does not, in its current implementation, support lightweight incremental updates. While this full-reconstruction paradigm implies an inherent latency, the computational complexity of our approach scales more favorably than traditional methods, particularly as the map resolution (dimension M) and the number of sparse samples (N) increase.

Specifically, traditional interpolation methods such as Kriging are computationally burdensome. Kriging requires the inversion of an $N \times N$ covariance matrix, incurring $O(N^3)$ complexity. This cost becomes prohibitive as the number of sensing devices (N) grows large. Likewise, conventional optimization methods, such as iterative MAP estimation using gradient descent, also present scalability challenges. Their computational cost per iteration typically scales with the map dimension M (the number of pixels), and achieving convergence for high-resolution maps can be protracted.

The proposed method also differs fundamentally from Physics-based Prediction (e.g., Ray-Tracing), which operates on an entirely different paradigm. Instead of interpolating from sparse measurements, ray-tracing attempts to deterministically predict the map from first principles. This simulation-based approach faces two prohibitive barriers to real-time operation. First, its Extreme Environmental Data Requirement is its primary bottleneck. Accurate prediction demands a complete, high-fidelity, real-time 3D model of the environment, encompassing precise building geometry, electromagnetic material properties like permittivity, and the instantaneous state of all dynamic objects. Acquiring and maintaining such an extraordinarily complex dataset is logistically infeasible for most practical ISAC applications. Second, even if such data were available, the Enormous Computational Load of simulating the complex physics of ray propagation, which involves millions of reflections and diffractions, for every pixel (M) is immense.

Conversely, our data-driven, diffusion-based approach circumvents these challenges. This framework employs a U-Net, a type of Convolutional Neural Network (CNN), as its core backbone network to execute the fundamental denoising step (i.e., noise prediction). This model does not require explicit 3D environmental geometry; instead, it relies solely on the N sparse measurements and leverages physical properties implicitly learned during offline training. Consequently, once this U-Net backbone is trained, its single forward pass exhibits a computational complexity that scales linearly with the map dimension M, resulting in $O(M)$ complexity. Crucially, the cost of this single pass is independent of the number of sparse input samples N. Therefore, as the problem scales, the fundamental computational kernel of our deep learning approach is significantly more efficient than methods dependent on $O(N^3)$ complexity or slow iterative convergence.

We acknowledge, however, that the primary latency bottleneck of the overall method is not the $O(M)$ complexity of a single U-Net pass, but the iterative nature of the diffusion sampling process. This process requires T sequential passes of the U-Net backbone to generate the final map. Addressing this cumulative latency (which scales with $T \times O(M)$) is the key pathway to achieving (near) real-time viability. This remains an active area of research, and as part of our future work, we will investigate state-of-the-art acceleration techniques. These include: (1) Advanced Samplers, such as DPM-Solver and DDIM, to drastically reduce the number of required sampling steps T (for instance, from 1000 down to 10–20) with minimal quality loss, and (2) Model Compression/Distillation to create a more compact and faster U-Net backbone.

VI. Conclusion

This study has presented RadioDiff-Inverse, a diffusion-enhanced Bayesian framework for radio map reconstruction under environmental uncertainty and sparse noisy measurements. Two principal contributions have been established: (1) A Bayesian inverse problem formulation systematically addressing measurement and environmental uncertainties has been developed; (2) A training-free methodology leveraging pre-trained diffusion models for prior distribution learning has demonstrated effective cross-domain knowledge transfer. The framework's capability to infer environmental features from limited measurements has advanced integrated sensing and communication in semi-aware scenarios. Future research directions have been identified, including dynamic environment adaptation and integration with emerging wireless technologies. This work has established a new paradigm for addressing inverse problems in wireless communications through synergistic integration of generative models and Bayesian inference.

References

[1] X. Shen et al., "Toward immersive communications in 6G," *Frontiers Comput. Sci.*, vol. 4, Jan. 2023, Art. no. 1068478.

[2] Y. Zeng et al., "A tutorial on environment-aware communications via channel knowledge map for 6G," *IEEE Commun. Surv. Tut.*, vol. 26, no. 3, pp. 1478–1519, 3rd Quart., 2024.

[3] X. Wang, X. Wang, S. Mao, J. Zhang, S. C. G. Periaswamy, and J. Patton, "Adversarial deep learning for indoor localization with channel state information tensors," *IEEE Internet Things J.*, vol. 9, no. 19, pp. 18182–18194, Oct. 2022.

[4] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?," *Nature Electron.*, vol. 3, no. 1, pp. 20–29, Jan. 2020.

[5] S. Liu, X. Yu, Z. Gao, J. Xu, D. W. K. Ng, and S. Cui, "Sensing-enhanced channel estimation for near-field XL-MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 3, pp. 628–643, Mar. 2025.

[6] 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.

[7] 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.

[8] H. Sun, L. Zhu, and R. Zhang, "Channel gain map estimation for wireless networks based on scatterer model," *IEEE Trans. Wireless Commun.*, vol. 24, no. 8, pp. 7012–7028, Aug. 2025.

[9] 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.

[10] 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.

[11] 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. 2025.

[12] X. Wang et al., "RadioDiff-k2: Helmholtz equation informed generative diffusion model for multi-path aware radio map construction," *IEEE J. Sel. Areas Commun.*, vol. 44, pp. 2318–2333, 2026.

[13] K. D. Polyzos et al., "Bayesian active learning for sample efficient 5G radio map reconstruction," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 19382–19396, Dec. 2024.

[14] J. Wang et al., "Sparse Bayesian learning-based hierarchical construction for 3D radio environment maps incorporating channel shadowing," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 14560–14574, Oct. 2024.

[15] X. Gong et al., "Digital twin of channel: Diffusion model for sensing-assisted statistical channel state information generation," *IEEE Trans. Wireless Commun.*, vol. 24, no. 5, pp. 3805–3821, May 2025.

[16] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.

[17] F. J. Breidt and J. D. Opsomer, "Local polynomial regresssion estimators in survey sampling," *Ann. Statist.*, vol. 28, no. 4, pp. 1026–1053, Aug. 2000.

[18] Q. Li, X. Liao, A. Li, and S. Valaee, "Automatic indoor radio map construction and localization via multipath fingerprint extrapolation," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 5814–5827, Sep. 2023.

[19] C. A. Balanis, "Antenna theory: A review," *Proc. IEEE*, vol. 80, no. 1, pp. 7–23, 2002.

[20] A. Bondeson, T. Rylander, and P. Ingelström, *Computational Electromagnetics*. Cham, Switzerland: Springer, 2012.

[21] 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.

[22] 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.
[23] Z. Zhang, G. Zhu, J. Chen, and S. Cui, “Fast and accurate cooperative radio map estimation enabled by GAN,” in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2024, pp. 1641–1646.
[24] P. Liao, X. Wang, L. An, S. Mao, T. Zhao, and C. Yang, “TFSemantic: A time-frequency semantic GAN framework for imbalanced classification using radio signals,” *ACM Trans. Sensor Netw.*, vol. 20, no. 4, pp. 1–22, Jul. 2024.
[25] W. Chen and J. Chen, “Diffraction and scattering aware radio map and environment reconstruction using geometry model-assisted deep learning,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 19804–19819, Dec. 2024.
[26] S. Zhang, T. Yu, B. Choi, F. Ouyang, and Z. Ding, “Radiomap inpainting for restricted areas based on propagation priority and depth map,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9330–9344, Aug. 2024.
[27] C. Rose, S. Ulukus, and R. D. Yates, “Wireless systems and interference avoidance,” *IEEE Trans. Wireless Commun.*, vol. 1, no. 3, pp. 415–428, Jul. 2002.
[28] X. Wang, Z. Yu, S. Mao, J. Zhang, S. C. G. Periaswamy, and J. Patton, “MapLoc: LSTM-based location estimation using uncertainty radio maps,” *IEEE Internet Things J.*, vol. 10, no. 15, pp. 13474–13488, Aug. 2023.
[29] A. Liu et al., “A survey on fundamental limits of integrated sensing and communication,” *IEEE Commun. Surv. Tut.*, vol. 24, no. 2, pp. 994–1034, 2nd Quart., 2022.
[30] G. E. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*. Hoboken, NJ, USA: Wiley, 2011.
[31] J. Ho, “Denoising diffusion probabilistic models,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2024, pp. 6840–6851.
[32] 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*.
[33] Y. Huang, Z. Qin, X. Liu, and K. Xu, “Decoupled diffusion models: Simultaneous image to zero and zero to noise,” 2024, *arXiv:2306.13720*.
[34] Y. Song, L. Shen, L. Xing, and S. Ermon, “Solving inverse problems in medical imaging with score-based generative models,” in *Proc. Int. Conf. Learn. Represent.*, 2021, pp. 1–18.
[35] Z. Dou and Y. Song, “Diffusion posterior sampling for linear inverse problem solving: A filtering perspective,” in *Proc. 12th Int. Conf. Learn. Represent.*, 2024, pp. 1–13.
[36] 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.*, Sep. 2022, pp. 10684–10695.
[37] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[38] J. Song, C. Meng, and S. Ermon, “Denoising diffusion implicit models,” in *Proc. Int. Conf. Learn. Represent.*, 2021, pp. 1–20.

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

Zhongsheng Fang (Student Member, IEEE) is currently pursuing the Ph.D. degree in engineering with Xidian University, specializing in the intersection of generative models and satellite communications. His research interests include electromagnetic digital twin technology, with the goal of enabling advanced applications, such as satellite radio map estimation and 3D channel reconstruction.

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.

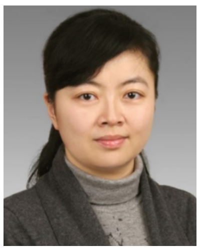
He was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, from 2017 to 2019. He is currently a Professor with the State Key Laboratory of ISN and the School of Telecommunications Engineering, Xidian University, Xi’an, Shaanxi, 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.

Haibo Zhou (Fellow, IEEE) received the Ph.D. degree in information and communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2014. From 2014 to 2017, he was a Post-Doctoral Fellow with the Broadband Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo. He is currently a Full Professor with the School of Electronic Science and Engineering, Nanjing University, Nanjing, China. His research interests include resource management and protocol design in B5G/6G networks, vehicular ad hoc networks, and space-air-ground integrated networks. He was a recipient of the 2019 IEEE ComSoc Asia–Pacific Outstanding Young Researcher Award. He served as the Track/Symposium Co-Chair for IEEE/CIC ICCC 2019, IEEE VTC-Fall 2020, IEEE VTC-Fall 2021, and IEEE Globecom 2022. He is also an Associate Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, *IEEE Network Magazine*, and IEEE WIRELESS COMMUNICATIONS LETTERS.

Zhou Su (Senior Member, IEEE) is currently a Professor with Xi'an Jiaotong University. He has published technical papers, including top journals and top conferences, such as IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE Infocom. His research interests include multimedia communication, wireless communication, network security, and network traffic. He received the Best Paper Award of International Conference, including IEEE AIoT2024, IEEE WCNC2023, IEEE VTC-Fall2023, and IEEE ICC2020. He is the Chair of the IEEE VTS Xi'an Chapter Section. He is an Associate Editor of IEEE INTERNET OF THINGS JOURNAL and IEEE OPEN JOURNAL OF THE COMPUTER SOCIETY.



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*.

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, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, a Chinese Academy of Engineering Foreign Member, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society. He received the R. A. Fessenden Award in 2019 from IEEE, Canada, Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) in 2019, the James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, 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 (2019) and the AHSN Technical Committee (2013). He has also received the Excellent Graduate Supervision Award in 2006 from the University of Waterloo and the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada. 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, a Member-at-Large on the Board of Governors, the Chair of the Distinguished Lecturer Selection Committee, and a Member of IEEE Fellow Selection Committee of the ComSoc. He served as the Editor-in-Chief for the IEEE INTERNET OF THINGS JOURNAL, IEEE NETWORK, and *IET Communications*.