

Radiomap Blind Prediction under Incomplete Observation: Error Characterization and Correctable Propagation-Prior Learning

Xiaojie Li, *Student Member, IEEE*, Yu Han, *Member, IEEE*, Han Fang, Shangqing Liu, Guangxu Zhu, *Member, IEEE*, Shi Jin, *Fellow, IEEE*, and Chao-Kai Wen, *Fellow, IEEE*

Abstract—Radiomap blind prediction aims to infer radiomaps from observable representations of the propagation environment and base station configuration without field measurements. In practice, the observable representations are inherently incomplete. Thus, the target radiomap is not fully determined by the inputs when generalizing to unseen configurations or environments. Under incomplete observation, we establish a population-level theory of deterministic radiomap blind prediction that identifies the conditional mean as its optimal target and separates prediction error into reducible predictor approximation and irreducible uncertainty caused by missing physical information. The framework further characterizes the train-test risk gap and the uncertainty reduction enabled by observation enrichment. Building on it, we reveal the dual role of propagation priors: they provide physically grounded guidance, yet their implementable forms may bias the attainable predictor. This motivates RadioDecomp, which treats a prior-guided predictor as a correctable base and learns its remaining predictable discrepancy through residual refinement. To evaluate RadioDecomp across distinct propagation-prior designs, we instantiate it with a feature-guided monolithic base and a LoS-Shadow structured base, yielding RadioFR and RadioLSR, respectively. Across random, cross-configuration, and cross-environment settings, experiments confirm the benefit of propagation-related representations and show that both instantiations improve upon their respective bases. Further controlled studies on base capacity, training-support coverage, and observation coarsening corroborate the proposed analysis.

Index Terms—radiomap, incomplete observation, generalization, residual learning, wireless propagation model

I. INTRODUCTION

A preliminary version of this work, entitled “Rethinking Radiomap Blind Prediction with Limited Environment and Configuration Representations,” has been accepted for presentation at the 2026 IEEE Global Communications Conference (GLOBECOM).

Xiaojie Li, Yu Han, and Shi Jin are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China, and also with National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: {xiaojieli, hanyu, and jinshi}@seu.edu.cn).

Han Fang is with the School of Cyber Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: h_fang@seu.edu.cn).

Shangqing Liu is with the State Key Laboratory of Novel Software Technology, Nanjing University, Nanjing 210023, China (e-mail: shangqingliu@nju.edu.cn).

Guangxu Zhu is with the Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong-Shenzhen, Guangdong 518172, China (e-mail: gxzhu@sribd.cn).

Chao-Kai Wen is with the Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan (e-mail: chaokai.wen@mail.nsysu.edu.tw).

FUTURE 6G networks are expected to rely on environment-aware radio intelligence, creating strong demand for radiomaps that characterize the spatial distribution of radio attributes such as received power and angle of arrival [1]. Such radiomaps are valuable for wireless network planning [2], optimization [3], predictive radio management [4], [5], and covert communications [6], [7]. In practice, obtaining radiomaps through exhaustive measurements is prohibitively costly, motivating extensive research on radiomap reconstruction [8]–[10]. However, many applications require reliable radiomaps for areas or network configurations without available radiomap measurements, such as coverage prediction for base station (BS) deployment, reconfiguration, or decommissioning. This gives rise to *radiomap blind prediction*, which predicts radiomaps solely from observable representations of the propagation environment and BS configuration [11], [12].

Traditional radiomap blind prediction mainly relied on statistical propagation models [13], [14] or ray tracing [15]. Statistical models infer radiomaps from empirical propagation laws calibrated for specific scenario classes or from stochastic models of network geometry, enabling efficient estimation but limited site-specific characterization. Ray tracing instead simulates propagation paths within a locally modeled scene, offering finer radiomap characterization but requiring detailed environment and BS modeling and substantial computation. Enabled by advances in deep learning, learning-based methods have emerged as a critical alternative approach. Such methods directly learn the mapping from observable environment and BS-configuration representations to their corresponding radiomaps. Early studies mainly investigated radiomap blind prediction across varying environments under fixed or simplified BS configurations [16]–[18]. Subsequent studies extended this setting to jointly varying environments and configurations, explicitly accounting for variations in carrier frequency, antenna patterns, array configurations, and beamforming [12], [19]–[23]. However, only partial environment and configuration representations are available in practical blind prediction.

Such incomplete representations omit propagation-relevant factors, such as material properties, fine-grained geometry, small scatterers, indoor structures, and hardware imperfections. Consequently, the same observable representation may be compatible with multiple complete physical states that produce different radiomaps. Nevertheless, radiomap blind prediction is commonly formulated using deterministic predictors that return a single estimate for each observable input.

This contrast raises fundamental questions about the statistical target approximated by such a predictor and about how its prediction error and generalization should be understood under incomplete observation.

Understanding these issues requires distinguishing radiomap prediction under incomplete observation from radiomap generation under a complete physical state. Theoretically, the observable and hidden factors together define a complete physical state. Under such state, the radiomap is deterministically generated through electromagnetic propagation ultimately governed by Maxwell's equations. Thus, the physical regularities make propagation knowledge a natural and principled source of structural guidance for radiomap blind prediction, particularly under finite training data and unseen test conditions.

Existing methods incorporate such propagation knowledge at the feature, model, and loss levels. Feature-level methods transform propagation knowledge into model inputs [24]–[28], model-level methods encode propagation processes into network structures [29]–[32], and loss-level methods constrain model training using propagation rules [33]–[35]. These mechanisms utilise useful propagation regularities and reduce the difficulty of learning from finite samples [36]. However, under incomplete observation, the complete propagation process cannot be fully characterized from the observable inputs, while implementable propagation priors typically capture selected or simplified physical regularities. Consequently, such priors may be inconsistent with the statistically optimal prediction relation, thereby introducing prior-induced bias. Existing studies have primarily emphasized their empirical benefits, leaving this benefit-bias tradeoff under incomplete observation insufficiently understood.

To address these issues, we develop a population-level framework for radiomap blind prediction under incomplete observation. We formulate radiomap generation using observable and hidden physical factors and identify the conditional mean given the observable representation as the optimal deterministic predictor under squared loss, providing a statistical basis for understanding prediction error and train-test generalization. We further examine propagation-prior injection relative to this conditional mean target and show that, although propagation priors can facilitate finite-sample learning, their partial or simplified physical assumptions may introduce prior-induced bias. This observation motivates *RadioDecomp*, which retains a propagation-prior-guided predictor as its base and introduces a residual predictor to correct its deviation from the conditional target. The main contributions of this paper are summarized as follows:

- **Incomplete-Observation Error Analysis Framework.** We establish a population-level framework to distinguish radiomap blind prediction errors addressable by predictor improvement from those caused by incomplete observation. Under squared loss, we identify the conditional-mean radiomap as the optimal prediction target and decompose the expected error of any deterministic predictor into its approximation error relative to this target and irreducible uncertainty. The framework further characterizes how these two terms contribute to the train-test risk gap and proves that additional observations strictly reduce the

uncertainty whenever they change the conditional-mean radiomap with positive probability.

- **Propagation-Prior Benefit-Bias Analysis.** Building on this framework, a unified analysis of feature-, model-, and loss-level prior injection characterizes information-compression, structural-approximation, and loss-induced biases. Together, these results show how propagation priors can facilitate finite-sample learning while their implementable forms may bias the attainable predictor.
- **RadioDecomp: Correctable Propagation-Prior Learning.** We propose RadioDecomp, which retains a prior-guided radiomap predictor as a base and learns its remaining predictable discrepancy through a residual branch. Two instantiations examine complementary correction needs: RadioFR addresses prediction errors that remain under information-preserving feature augmentation, while RadioLSR addresses discrepancies under a more restrictive LoS-Shadow base with region-specific inputs. Experiments under random, cross-configuration, cross-environment, and limited-data settings demonstrate the benefits of prior guidance and residual correction, while studies of training-support coverage and observation coarsening support the theoretical analysis.

II. RELATED WORK

A. Learning-Based Radiomap Prediction

Early blind-prediction methods mainly modeled environmental variations under fixed or simplified transmitter configurations (isotropic antenna pattern). RadioUNet [16] established an image-to-image prediction paradigm using building maps and transmitter locations, while RME-GAN [17] employed a conditional generative adversarial network for radiomap prediction. RadioDiff [18] subsequently introduced a conditional diffusion model for measurement-free prediction.

More recent studies have extended radiomap learning toward configuration awareness. Directional antenna patterns were incorporated through spatial gain projections [19], while UniRM [20] studied adaptation across frequency bands and receiver heights with sparse target measurements. BeamCKM [21] and BeamCKMDiff [22] modeled discrete and continuous beamforming configurations, respectively. Configuration-dependent coverage was also investigated through disentangled representations using real measurements [23] and explicit beam-map representations for U6G XL-MIMO systems [12]. These studies demonstrate a progression from environment-aware to joint environment- and configuration-aware prediction, but the statistical target of deterministic blind prediction under incomplete observation remains insufficiently understood.

B. Propagation Priors for Radiomap Prediction

Propagation knowledge has been incorporated at the feature, model, and loss levels. Feature-level methods provide models with propagation-related representations derived from geometry, materials, attenuation, obstruction, or wave equations [24]–[28]. Such features can expose useful propagation patterns to a

radiomap predictor but may be inaccurate because they depend on simplified models and incomplete observation.

Model-level methods embed propagation assumptions into predictor structures. Existing examples include ray-tracing-informed large-model modules [29], geometry-aware decomposition of propagation components [30], reciprocity-constrained bidirectional Gaussian splatting [31], physics-guided diffusion architectures [32], and disentangled modeling of coverage factors [23]. These structures can improve data efficiency, but structural mismatch may restrict the hypothesis class and limit its approximation of the optimal predictor.

Loss-level methods incorporate propagation knowledge through physical constraints or propagation-aware error metrics during training. Representative approaches use volume-integral equations [33], LoS power constraints and NLoS propagation-delay lower bounds [34], Helmholtz-equation and boundary residuals [35], or logarithmic-power-domain prediction losses [31]. Approximate constraints and their weights may, however, shift the population optimum of the original prediction task.

Totally, existing studies primarily demonstrate the empirical benefits of task-specific propagation priors. A unified understanding of their finite-sample benefits and prior-induced biases under incomplete observation, together with a general mechanism for correcting such biases, is still lacking.

III. ERROR CHARACTERIZATION UNDER INCOMPLETE OBSERVATION

A. From Physical Generation to Blind Prediction

Let E denote the complete physical environment and C the complete BS configuration. The former includes propagation-relevant scene information, such as geometry and material properties, while the latter specifies the BS location and transmission setup, including the antenna array, carrier frequency, and beamforming configuration. A radiomap is a discretized spatial representation of a radio attribute, such as received power, angle of arrival, or angle of departure. In the 2D setting considered here, a radiomap realization is represented by $\mathbf{y} \in \mathbb{R}^{H \times W}$. The same formulation can be extended to voxelized 3D radiomaps [37], [38]. When E and C are fully specified, radiomap generation can be expressed as

$$\mathbf{y} = F(E, C), \quad (1)$$

where F denotes the deterministic propagation mechanism. Therefore, full-information radiomap generation is a deterministic forward problem, which can in principle be numerically solved using full-wave simulation [39] or high-fidelity ray tracing.

Radiomap blind prediction differs because the predictor receives only incomplete observation of the environment and BS configuration. The observable environment may omit fine-scale geometry, electromagnetic material properties, indoor structures, or small scatterers. Likewise, recorded BS configurations may contain only coarse parameters, such as the beam direction, 3-dB beamwidth, and nominal transmit power, while the exact beam pattern, hardware-induced power offsets, and other site-specific calibration parameters remain unrecorded.

Let $\mathbf{Z} = (E_{\text{obs}}, C_{\text{obs}})$ denote the observable representation, taking values in the observable-input space $\mathcal{Z}$. Elements of $\mathcal{Z}$ may combine spatial environment representations with discrete and continuous BS-configuration variables. Let $\mathbf{H}$ collect the remaining hidden propagation factors. The complete physical state can then be partitioned into $(\mathbf{Z}, \mathbf{H})$. Accordingly,

$$\mathbf{Y} = F(\mathbf{Z}, \mathbf{H}). \quad (2)$$

For a given observable representation $\mathbf{z}$, multiple hidden states may be compatible with it and produce different radiomaps. For example, scenes with the same observable geometry may differ in material properties or unobserved scatterers. Accordingly,¹

$$\mathbf{Y} \mid (\mathbf{Z} = \mathbf{z}) = F(\mathbf{z}, \mathbf{H}), \quad \mathbf{H} \sim p(\cdot \mid \mathbf{z}). \quad (3)$$

Although $\mathbf{H}$ is unobserved, a deterministic blind predictor must return a single radiomap for each observable input $\mathbf{z}$. This single-valued mapping is learned from paired observations of $\mathbf{Z}$ and $\mathbf{Y}$. Under squared loss, its population-optimal output is

$$\begin{aligned} m^*(\mathbf{z}) &:= \arg \min_{\hat{\mathbf{y}}} \mathbb{E} \left[\|\mathbf{Y} - \hat{\mathbf{y}}\|_F^2 \mid \mathbf{Z} = \mathbf{z} \right] \\ &= \mathbb{E}[\mathbf{Y} \mid \mathbf{Z} = \mathbf{z}] = \mathbb{E}_{\mathbf{H} \sim p(\cdot \mid \mathbf{z})} [F(\mathbf{z}, \mathbf{H})]. \end{aligned} \quad (4)$$

Remark 1 (Learning target and error-analysis perspective). For a given data population, deterministic radiomap blind prediction aims to learn a mapping $f : \mathcal{Z} \rightarrow \mathbb{R}^{H \times W}$ from the available observation to a single radiomap estimate. **Under squared loss, the ideal predictor is the conditional-mean mapping m^* , and the learning objective is to approximate this mapping as closely as possible.** Even if $f(\mathbf{Z}) = m^*(\mathbf{Z})$ almost surely, its prediction may still differ from the realized radiomap $\mathbf{Y}$, because hidden propagation factors remain unresolved by the observation. The following analysis therefore uses m^* as the reference predictor to distinguish the error due to imperfect approximation by f from the irreducible uncertainty under incomplete observation, and examines how these two terms vary between training and test domains.

B. Domain-Wise Risk Decomposition and Train-Test Gap

A radiomap predictor is often trained on available environments and BS configurations and then applied to unseen ones. Its prediction error may change because both the observable environment/configuration distribution and the unrecorded propagation factors differ between training and test deployment. Let $d \in \{\text{tr}, \text{te}\}$ index the training and test domains, respectively. Each domain is characterized by a joint distribution $p_d(\mathbf{z}, \mathbf{h})$ over the observable input and the unobserved propagation factors. Accordingly, the population-optimal deterministic predictor under squared loss is domain-dependent:

$$m_d^*(\mathbf{z}) := \mathbb{E}_d[\mathbf{Y} \mid \mathbf{Z} = \mathbf{z}] = \mathbb{E}_{\mathbf{H} \sim p_d(\cdot \mid \mathbf{z})} [F(\mathbf{z}, \mathbf{H})]. \quad (5)$$

The corresponding conditional uncertainty is defined as

$$v_d(\mathbf{z}) := \mathbb{E}_d \left[\|\mathbf{Y} - m_d^*(\mathbf{z})\|_F^2 \mid \mathbf{Z} = \mathbf{z} \right]. \quad (6)$$

¹Throughout this paper, bold uppercase symbols denote random variables, and their lowercase counterparts denote realizations. We use $p(\cdot)$ generically for probability laws, including discrete, continuous, and mixed cases.

The conditional target m_d^* describes the deterministic prediction target m^* in domain d , whereas v_d measures the remaining radiomap variation induced by the unobserved factors. These functions may differ across domains when the conditional hidden-factor distributions differ. Independently, different observable-input distributions determine how the corresponding pointwise quantities are aggregated within each domain. The following theorem shows that these two quantities yield an exact decomposition of the domain risk.

Theorem 1 (Domain-wise risk decomposition). *For any deterministic predictor f with finite second moment and any $d \in \{\text{tr}, \text{te}\}$, the population risk satisfies*

$$\begin{aligned} R_d(f) &:= \mathbb{E}_d \left[\|\mathbf{Y} - f(\mathbf{Z})\|_F^2 \right] \\ &= \underbrace{\mathbb{E}_d \left[\|f(\mathbf{Z}) - m_d^*(\mathbf{Z})\|_F^2 \right]}_{\mathcal{A}_d(f)} + \underbrace{\mathbb{E}_d [v_d(\mathbf{Z})]}_{\mathcal{U}_d}, \end{aligned} \quad (7)$$

where $\mathcal{A}_d(f)$ is the domain-specific approximation error and $\mathcal{U}_d$ is the irreducible uncertainty induced by incomplete observation in domain d .

Proof. Adding and subtracting $m_d^*(\mathbf{Z})$ gives

$$\begin{aligned} R_d(f) &= \mathbb{E}_d \left[\|f(\mathbf{Z}) - m_d^*(\mathbf{Z})\|_F^2 \right] + \mathbb{E}_d \left[\|\mathbf{Y} - m_d^*(\mathbf{Z})\|_F^2 \right] \\ &\quad + 2\mathbb{E}_d [\langle f(\mathbf{Z}) - m_d^*(\mathbf{Z}), m_d^*(\mathbf{Z}) - \mathbf{Y} \rangle_F]. \end{aligned} \quad (8)$$

The cross term is zero because $\mathbb{E}_d [\mathbf{Y} - m_d^*(\mathbf{Z}) \mid \mathbf{Z}] = 0$. The second term equals $\mathbb{E}_d[v_d(\mathbf{Z})]$ by the law of total expectation, which proves (7). $\square$

Theorem 1 separates two sources of radiomap prediction error with different physical implications. The approximation term $\mathcal{A}_d(f)$ measures how effectively the predictor exploits the recorded environment and BS configuration, whereas $\mathcal{U}_d$ captures the remaining radiomap variability associated with hidden propagation factors. For a fixed observation and domain, $R_d(f) \geq \mathcal{U}_d$, with equality when the predictor recovers m_d^* almost surely. Hence, increasing model capacity or improving training can bring the prediction closer to the conditional-mean radiomap, but cannot resolve propagation differences that the input does not distinguish. For example, a geometry-only predictor may learn typical attenuation and blockage patterns while retaining uncertainty associated with unrecorded material properties. The decomposition therefore distinguishes the benefit attainable through predictor improvement from that requiring richer physical observations. Since both terms depend on the deployment domain, their changes jointly determine the train-test risk gap.

Corollary 1 (Train-test risk gap). *For any fixed deterministic predictor f ,*

$$\begin{aligned} R_{\text{te}}(f) - R_{\text{tr}}(f) &= \underbrace{\mathcal{A}_{\text{te}}(f) - \mathcal{A}_{\text{tr}}(f)}_{\text{domain-specific approximation gap}} \\ &\quad + \underbrace{\mathcal{U}_{\text{te}} - \mathcal{U}_{\text{tr}}}_{\text{conditional-uncertainty gap}}. \end{aligned} \quad (9)$$

Corollary 1 explains why radiomap prediction performance can change substantially when a model is transferred to new

environments or BS configurations. The approximation gap reflects how well the learned mapping captures propagation behavior in the test domain. It can increase when training data provide limited coverage of the relevant layouts or configurations, or when different hidden propagation conditions change the conditional-mean radiomap. The uncertainty gap captures a complementary effect: the same observable representation may leave different amounts of propagation variability unresolved across domains. For example, a building-geometry map may provide a more complete description of propagation in a domain with relatively uniform materials than in one with substantial unrecorded material variation. The train-test gap therefore reflects both the transferability of the learned propagation relation and the adequacy of the available physical description in each domain.

In practice, learning is primarily supported by the available training radiomaps. Broader coverage of training data (relevant environments and BS configurations) helps the predictor learn propagation relations in the training domain, while effective predictor design determines how well these relations are extracted from finite samples. The uncertainty term points to a separate question: can additional environment or configuration information make the target radiomap more predictable? This motivates the following analysis of observation enrichment.

C. Observation Enrichment and Conditional Uncertainty

For radiomap blind prediction, observation enrichment means providing additional propagation-relevant information, such as adding material properties to a building-geometry map or supplying a detailed antenna pattern alongside a coarse beam descriptor. Such information may distinguish propagation conditions that were indistinguishable under the original input, thereby lowering the minimum achievable radiomap prediction error.

Let $\tilde{\mathbf{Z}}$ denote an enriched observation from which the original observation can be recovered:

$$\mathbf{Z} = \psi(\tilde{\mathbf{Z}}). \quad (10)$$

Thus, $\tilde{\mathbf{Z}}$ retains all information contained in $\mathbf{Z}$ and may additionally reveal propagation-relevant factors.

For each domain $d \in \{\text{tr}, \text{te}\}$, define the conditional mean and conditional uncertainty under the enriched observation as

$$\tilde{m}_d^*(\tilde{\mathbf{z}}) := \mathbb{E}_d [\mathbf{Y} \mid \tilde{\mathbf{Z}} = \tilde{\mathbf{z}}] \quad (11)$$

and

$$\tilde{v}_d(\tilde{\mathbf{z}}) := \mathbb{E}_d \left[\|\mathbf{Y} - \tilde{m}_d^*(\tilde{\mathbf{z}})\|_F^2 \mid \tilde{\mathbf{Z}} = \tilde{\mathbf{z}} \right]. \quad (12)$$

Let

$$\tilde{\mathcal{U}}_d := \mathbb{E}_d [\tilde{v}_d(\tilde{\mathbf{Z}})]. \quad (13)$$

Proposition 1 (Monotonicity under observation enrichment). *Suppose the involved random variables have finite second moments and $\mathbf{Z} = \psi(\tilde{\mathbf{Z}})$. Then, for every domain $d \in \{\text{tr}, \text{te}\}$,*

$$\begin{aligned} v_d(\mathbf{Z}) - \mathbb{E}_d [\tilde{v}_d(\tilde{\mathbf{Z}}) \mid \mathbf{Z}] \\ = \mathbb{E}_d \left[\left\| \tilde{m}_d^*(\tilde{\mathbf{Z}}) - m_d^*(\mathbf{Z}) \right\|_F^2 \mid \mathbf{Z} \right] \geq 0 \end{aligned} \quad (14)$$

almost surely. Consequently,

$$\mathcal{U}_d - \tilde{\mathcal{U}}_d = \mathbb{E}_d \left[\left\| \tilde{m}_d^*(\tilde{\mathbf{Z}}) - m_d^*(\mathbf{Z}) \right\|_F^2 \right] \geq 0. \quad (15)$$

Proof. Because $\mathbf{Z} = \psi(\tilde{\mathbf{Z}})$, $m_d^*(\mathbf{Z})$ is also determined by $\tilde{\mathbf{Z}}$. Expanding

$$\mathbf{Y} - m_d^*(\mathbf{Z}) = \mathbf{Y} - \tilde{m}_d^*(\tilde{\mathbf{Z}}) + \tilde{m}_d^*(\tilde{\mathbf{Z}}) - m_d^*(\mathbf{Z})$$

and conditioning first on $\tilde{\mathbf{Z}}$ eliminates the cross term because

$$\mathbb{E}_d \left[\mathbf{Y} - \tilde{m}_d^*(\tilde{\mathbf{Z}}) \mid \tilde{\mathbf{Z}} \right] = 0.$$

Conditioning subsequently on $\mathbf{Z}$ proves (14). Taking expectations proves (15). $\square$

Proposition 1 gives a precise criterion for the value of additional physical observations: the uncertainty reduction equals the expected squared change in the conditional-mean radiomap. Thus, additional information is useful for lowering the prediction limit when it distinguishes propagation conditions that were previously averaged together. For example, recording material properties can distinguish attenuation behavior among scenes with the same geometry, while a detailed antenna pattern can distinguish coverage variations hidden by a coarse beam descriptor. The relevant consideration is therefore the additional propagation information revealed by the representation.

Together, these results distinguish the roles of training-data acquisition, observation design, and predictor learning. More training radiomaps and broader coverage of environment and configuration variations support learning the predictable propagation relation; richer observations can make previously unresolved radiomap variations predictable. With the available observations fixed, the remaining task is to exploit their propagation information effectively from finite training data. Propagation priors offer a natural source of guidance by making selected physical relations explicit, motivating the following analysis of their benefits and potential biases.

IV. PROPAGATION-PRIOR GUIDANCE AND BIAS

A. Propagation Guidance for Predictor Approximation

With the available environment and BS information fixed, radiomap predictor design aims to reduce the approximation term $\mathcal{A}_d(f)$. Given training data, minimizing the training-domain population risk is equivalent to reducing

$$\mathcal{A}_{\text{tr}}(f) = \mathbb{E}_{\text{tr}} \left[\left\| f(\mathbf{Z}) - m_{\text{tr}}^*(\mathbf{Z}) \right\|_F^2 \right], \quad (16)$$

that is, learning a predictor that approximates the training-domain conditional target m_{tr}^* . In practice, empirical training error can often be substantially reduced, for example, by increasing model capacity or improving optimization. However, a small empirical training error does not necessarily imply that f closely approximates m_{tr}^* at the population level, much less that it approximates m_{te}^* well in the test domain. To identify a possible source of test-domain guidance, recall that the conditional targets in both domains can be written as

$$m_d^*(\mathbf{z}) = \mathbb{E}_{\mathbf{H} \sim p_d(\cdot|\mathbf{z})} [F(\mathbf{z}, \mathbf{H})], \quad d \in \{\text{tr}, \text{te}\}. \quad (17)$$

Although the conditional distributions of the hidden factors may differ between the two domains, both targets are governed by the same physical propagation mechanism F . Partial knowledge of F can therefore provide domain-relevant guidance for learning relations that may transfer from the training samples to unseen test inputs. This makes propagation knowledge a natural inductive prior for improving test-domain approximation.

In this paper, Φ denotes a *propagation prior*, namely, partial structural knowledge of F , such as distance-dependent attenuation, antenna radiation, LoS propagation, blockage, reflection, diffraction, or scattering. It may be incorporated through input features, model structures, training losses, or their combinations. The following analysis examines how these different prior-injection mechanisms affect the training-domain learning problem, including their population solutions and possible induced biases.

B. Prior-Injection Forms and Induced Biases

1) *Feature-level prior:* For radiomap blind prediction, feature-level injection converts available environment and BS information into propagation-related inputs, such as geometry-derived blockage indicators, distance-based attenuation maps, or beam-gain maps computed from recorded antenna and beamforming parameters. For a fixed prior Φ , this construction is represented as

$$\mathbf{U} = T_\Phi(\mathbf{Z}), \quad (18)$$

where T_Φ makes selected propagation relations explicit to the radiomap predictor. The resulting features may replace the original observations or be supplied alongside them, yielding *prior-only* and *prior-augmented* inputs, respectively.

1) *Prior-only input:* Consider a predictor $g(\mathbf{U})$ that uses only $\mathbf{U} = T_\Phi(\mathbf{Z})$. Under training-domain squared loss, its population conditional target is

$$\begin{aligned} m_{\Phi, \text{tr}}^*(\mathbf{u}) &:= \mathbb{E}_{\text{tr}} [\mathbf{Y} \mid \mathbf{U} = \mathbf{u}] = \mathbb{E}_{\text{tr}} [\mathbb{E}_{\text{tr}} [\mathbf{Y} \mid \mathbf{Z}] \mid \mathbf{U} = \mathbf{u}] \\ &= \mathbb{E}_{\text{tr}} [m_{\text{tr}}^*(\mathbf{Z}) \mid T_\Phi(\mathbf{Z}) = \mathbf{u}]. \end{aligned} \quad (19)$$

The second equality follows because $\mathbf{U} = T_\Phi(\mathbf{Z})$ is measurable with respect to $\mathbf{Z}$.

To compare this target with $m_{\text{tr}}^*(\mathbf{z})$ on the original observable space, define its pullback to the $\mathbf{Z}$ -space as

$$\bar{m}_{\Phi, \text{tr}}^*(\mathbf{z}) := m_{\Phi, \text{tr}}^*(T_\Phi(\mathbf{z})) = m_{\Phi, \text{tr}}^*(\mathbf{u}). \quad (20)$$

The information-compression bias induced by replacing $\mathbf{Z}$ with $T_\Phi(\mathbf{Z})$ is then

$$\mathcal{B}_{\Phi, \text{tr}}^{\text{feat}} := \mathbb{E}_{\text{tr}} \left[\left\| m_{\text{tr}}^*(\mathbf{Z}) - \bar{m}_{\Phi, \text{tr}}^*(\mathbf{Z}) \right\|_F^2 \right]. \quad (21)$$

This bias is zero only if the information retained by $T_\Phi(\mathbf{Z})$ is sufficient to determine the original conditional target. Equivalently, there must exist a measurable function q_Φ such that

$$m_{\text{tr}}^*(\mathbf{Z}) = q_\Phi(T_\Phi(\mathbf{Z})) \quad (22)$$

almost surely under the training distribution. Otherwise, $\mathcal{B}_{\Phi, \text{tr}}^{\text{feat}} > 0$.

For any predictor $g(\mathbf{U})$, the prediction error can be expanded as

$$\mathbf{Y} - g(\mathbf{U}) = \mathbf{Y} - m_{\text{tr}}^*(\mathbf{Z}) + m_{\text{tr}}^*(\mathbf{Z}) - m_{\Phi, \text{tr}}^*(\mathbf{U}) + m_{\Phi, \text{tr}}^*(\mathbf{U}) - g(\mathbf{U}). \quad (23)$$

The first component is orthogonal in expectation to the remaining two because

$$\mathbb{E}_{\text{tr}} [\mathbf{Y} - m_{\text{tr}}^*(\mathbf{Z}) \mid \mathbf{Z}] = 0. \quad (24)$$

Moreover,

$$\begin{aligned} \mathbb{E}_{\text{tr}} [m_{\text{tr}}^*(\mathbf{Z}) - m_{\Phi, \text{tr}}^*(\mathbf{U}) \mid \mathbf{U}] \\ = \mathbb{E}_{\text{tr}} [m_{\text{tr}}^*(\mathbf{Z}) \mid \mathbf{U}] - m_{\Phi, \text{tr}}^*(\mathbf{U}) = 0, \end{aligned} \quad (25)$$

so the second component is also orthogonal in expectation to the third. Consequently,

$$\begin{aligned} \mathbb{E}_{\text{tr}} [\|\mathbf{Y} - g(\mathbf{U})\|_{\text{F}}^2] &= \mathcal{U}_{\text{tr}} + \mathcal{B}_{\Phi, \text{tr}}^{\text{feat}} \\ &+ \mathbb{E}_{\text{tr}} [\|m_{\Phi, \text{tr}}^*(\mathbf{U}) - g(\mathbf{U})\|_{\text{F}}^2]. \end{aligned} \quad (26)$$

The three terms respectively represent the irreducible conditional uncertainty under the original observation, the additional error caused by compressing $\mathbf{Z}$ into $T_{\Phi}(\mathbf{Z})$, and the predictor-dependent error relative to the prior-optimal conditional target. In particular, even the population-optimal predictor based only on $\mathbf{U}$ has minimum risk $\mathcal{U}_{\text{tr}} + \mathcal{B}_{\Phi, \text{tr}}^{\text{feat}}$.

2) *Prior-augmented input*: Consider instead the augmented representation

$$\mathbf{Z}_{\Phi} = [\mathbf{Z}, T_{\Phi}(\mathbf{Z})]. \quad (27)$$

Because the original observation is retained and $T_{\Phi}(\mathbf{Z})$ is deterministically generated from it, $\mathbf{Z}_{\Phi}$ and $\mathbf{Z}$ contain the same population-level information. Therefore,

$$\mathbb{E}_{\text{tr}} [\mathbf{Y} \mid \mathbf{Z}, T_{\Phi}(\mathbf{Z})] = \mathbb{E}_{\text{tr}} [\mathbf{Y} \mid \mathbf{Z}] = m_{\text{tr}}^*(\mathbf{Z}). \quad (28)$$

Thus, retaining the original observations preserves the optimal radiomap target and $\mathcal{U}_{\text{tr}}$. Compared with the prior-only input, however, it presents the model with a higher-dimensional input space, so identifying a mapping close to m_{tr}^* may require a more demanding finite-sample search.

2) *Model-level prior*: Model-level injection embeds propagation knowledge F into how a radiomap predictor processes and combines information. One example is the adoption of CNN-based architectures [16], which reflects the prior belief that propagation patterns exhibit local spatial dependence and approximate translation equivariance. Another is to decompose signal propagation into distinct components or regimes, learn them through dedicated branches, and aggregate their contributions to form the final prediction [23].

Let $\mathcal{H}_{\Phi}$ denote the effective hypothesis class induced by the model-level prior. Because the predictor still operates on $\mathbf{Z}$, its training-domain conditional target remains $m_{\text{tr}}^*(\mathbf{Z})$. The prior instead restricts the functions available for approximating this target. The resulting structural approximation bias is

$$\mathcal{B}_{\Phi, \text{tr}}^{\text{model}} := \inf_{f \in \mathcal{H}_{\Phi}} \mathbb{E}_{\text{tr}} [\|f(\mathbf{Z}) - m_{\text{tr}}^*(\mathbf{Z})\|_{\text{F}}^2]. \quad (29)$$

Accordingly, the minimum attainable training-domain population risk is

$$\inf_{f \in \mathcal{H}_{\Phi}} R_{\text{tr}}(f) = \mathcal{U}_{\text{tr}} + \mathcal{B}_{\Phi, \text{tr}}^{\text{model}}. \quad (30)$$

The bias is zero when functions in $\mathcal{H}_{\Phi}$ can approximate the conditional-mean radiomap mapping m_{tr}^* arbitrarily well under the training distribution. Otherwise, the imposed structure produces an approximation error that remains even with unlimited data and ideal optimization.

3) *Loss-level prior*: Without loss of generality, we consider radiomap predictors trained with an additional penalty that encourages consistency with physical propagation mechanisms for loss-level injection. Let

$$\Omega_{\Phi, \text{tr}}(f) := \mathbb{E}_{\text{tr}} [\omega_{\Phi}(f; \mathbf{Z}, \mathbf{Y})] \quad (31)$$

denote its training-domain population form. Examples include physical-consistency penalties, propagation-guided auxiliary supervision, and smoothness or boundary constraints. The resulting population objective is

$$\mathcal{J}_{\lambda, \text{tr}}^{\text{pop}}(f) = R_{\text{tr}}(f) + \lambda \Omega_{\Phi, \text{tr}}(f), \quad f \in \mathcal{H}, \quad (32)$$

where $\lambda \geq 0$ controls the strength of the prior.

The loss-level prior introduces a soft preference when selecting a predictor from $\mathcal{H}$. Because $\mathcal{U}_{\text{tr}}$ is independent of f , the corresponding population solution satisfies

$$f_{\lambda, \text{tr}}^* \in \arg \min_{f \in \mathcal{H}} \left\{ \mathbb{E}_{\text{tr}} [\|f(\mathbf{Z}) - m_{\text{tr}}^*(\mathbf{Z})\|_{\text{F}}^2] + \lambda \Omega_{\Phi, \text{tr}}(f) \right\}. \quad (33)$$

To isolate the bias induced by this preference from the intrinsic approximation limitation of $\mathcal{H}$, define

$$\begin{aligned} \mathcal{B}_{\lambda, \text{tr}}^{\text{loss}} &:= \mathbb{E}_{\text{tr}} [\|f_{\lambda, \text{tr}}^*(\mathbf{Z}) - m_{\text{tr}}^*(\mathbf{Z})\|_{\text{F}}^2] \\ &- \inf_{f \in \mathcal{H}} \mathbb{E}_{\text{tr}} [\|f(\mathbf{Z}) - m_{\text{tr}}^*(\mathbf{Z})\|_{\text{F}}^2] \geq 0. \end{aligned} \quad (34)$$

This bias is zero when the loss-guided solution remains a best approximation to m_{tr}^* within $\mathcal{H}$. In particular, if $m_{\text{tr}}^* \in \mathcal{H}$ and also minimizes $\Omega_{\Phi, \text{tr}}$, the original population optimum is preserved. Otherwise, an incompatible penalty with fixed nonzero λ may shift the selected solution away from the conditional target, even with unlimited data and ideal optimization.

Discussion: For radiomap blind prediction, propagation priors can facilitate learning by exposing radiation-related features, organizing model components of the predictor, or designing a loss to guide physical consistency during training. Their potential biases arise through different mechanisms: feature replacement may discard informative observation details, architectural restrictions may limit the representable mapping, and incompatible loss penalties may favor inaccurate propagation relations. Although feature augmentation preserves the conditional-mean target, the higher input dimension may make identifying an effective radiomap mapping more demanding under finite data and limited optimization. In principle, each injection mechanism can avoid population-level bias if its design preserves the information, representational capacity, or loss compatibility needed to recover the conditional-mean radiomap. In practice, however, these designs are chosen using incomplete observations, finite training data, and selected propagation assumptions; their compatibility with the optimal mapping under incomplete observation is generally difficult to ensure. This motivates a predictor design that retains propagation guidance while allowing data-driven correction of its remaining predictable discrepancies.

V. RADIODECOMP: CORRECTABLE PROPAGATION-PRIOR LEARNING

A. RadioDecomp Formulation

For radiomap blind prediction, let $b_\Phi(\mathbf{z})$ denote a prior-guided estimate of the radiomap generated from the observable environment and BS configuration through the injection mechanisms in Section IV. The target remains the conditional-mean radiomap $m^*(\mathbf{z})$. The base estimate may depart from this target because of information compression, structural restrictions, incompatible prior losses, or finite-sample learning. To correct these potential deviations while retaining the encoded propagation guidance, we propose *RadioDecomp*, which augments the prior-guided base with a data-driven residual correction:

$$\hat{\mathbf{y}} = b_\Phi(\mathbf{z}) + r(\mathbf{z}, b_\Phi(\mathbf{z})), \quad (35)$$

The residual branch receives $\mathbf{z}$ to exploit observable cues not captured by the base and $b_\Phi(\mathbf{z})$ to condition the correction on the prediction being corrected. Although $b_\Phi(\mathbf{z})$ adds no new observable information because it is deterministic given $\mathbf{z}$, it serves as a propagation-guided prediction anchor. The base thus provides prior guidance, while the residual branch corrects its data-supported discrepancy.

To formalize this correctability, we consider a fixed prior-guided base predictor. Let $\mathcal{H}_R$ denote the residual hypothesis class and define

$$\mathcal{H}_{RD}(b_\Phi) := \{f : f(\mathbf{z}) = b_\Phi(\mathbf{z}) + r(\mathbf{z}, b_\Phi(\mathbf{z})), \quad r \in \mathcal{H}_R\}.$$

Proposition 2 (Residual correctability). *If the zero function belongs to the residual class, $0 \in \mathcal{H}_R$, then, for any $d \in \{\text{tr}, \text{te}\}$,*

$$\begin{aligned} & \inf_{f \in \mathcal{H}_{RD}(b_\Phi)} \mathbb{E}_d \left[\|f(\mathbf{Z}) - m_d^*(\mathbf{Z})\|_F^2 \right] \\ & \leq \mathbb{E}_d \left[\|b_\Phi(\mathbf{Z}) - m_d^*(\mathbf{Z})\|_F^2 \right]. \end{aligned} \quad (36)$$

Moreover, under squared loss, the pointwise population target of the residual branch in domain d is

$$r_{\Phi,d}^*(\mathbf{z}, b_\Phi(\mathbf{z})) = m_d^*(\mathbf{z}) - b_\Phi(\mathbf{z}). \quad (37)$$

Proof. Because $0 \in \mathcal{H}_R$, choosing $r \equiv 0$ recovers b_Φ . Hence, $b_\Phi \in \mathcal{H}_{RD}(b_\Phi)$, which proves (36).

For a fixed base predictor, $b_\Phi(\mathbf{Z})$ is deterministic given $\mathbf{Z}$. Therefore, for any $d \in \{\text{tr}, \text{te}\}$,

$$\begin{aligned} \mathbb{E}_d [\mathbf{Y} - b_\Phi(\mathbf{Z}) \mid \mathbf{Z} = \mathbf{z}] &= \mathbb{E}_d [\mathbf{Y} \mid \mathbf{Z} = \mathbf{z}] - b_\Phi(\mathbf{z}) \\ &= m_d^*(\mathbf{z}) - b_\Phi(\mathbf{z}). \end{aligned} \quad (38)$$

Because the conditional mean minimizes conditional squared loss, (37) follows. $\square$

Proposition 2 shows that the zero residual recovers b_Φ , while the population-optimal residual in domain d targets $m_d^* - b_\Phi$. RadioDecomp thus preserves the base prediction while making its predictable domain-specific discrepancy correctable.

B. Representative Instantiations

RadioDecomp is not tied to a specific propagation prior or base architecture. To evaluate its effectiveness under different prior-guided predictors, we develop two instantiations that examine complementary correction needs in radiomap blind prediction. RadioFR examines residual correction under feature augmentation, whereas RadioLSR examines correction under more restrictive propagation-guided prediction. RadioFR jointly processes the available representations to assess correction when feature augmentation does not introduce information-compression bias. RadioLSR instead uses a LoS-Shadow partition and restricted branch inputs, providing a stricter propagation prior setting. These instantiations demonstrate how the general base-plus-residual framework can be applied to distinct forms of prior injection. We first introduce the observable and propagation-prior representations shared by the two instantiations.

1) *Common Representations and Two-Stage Training:* For each radiomap sample, the observable representation contains an environment description and a BS configuration. The environment is represented by a height map $\mathbf{G} \in \mathbb{R}^{H \times W}$, while the BS configuration specifies the antenna array, carrier frequency, transmit power, and beamforming parameters. Based on the observable environment and configuration, several deterministic propagation-related representations are constructed. These representations expose physically meaningful relations that would otherwise need to be inferred entirely from finite training samples.

First, a beam map $\mathbf{B} \in \mathbb{R}^{H \times W}$ represents the analytically generated LoS beamforming power. Let $k \in \{1, \dots, HW\}$ denote a spatial grid point. The beam-map value is defined as [12]

$$B_k = \frac{\lambda^2}{(4\pi)^2} P_t |\mathbf{w}^H \mathbf{h}_k^{\text{LoS}}|^2, \quad (39)$$

where λ is the wavelength, P_t is the transmit power, $\mathbf{w}$ is the beamforming vector, and $\mathbf{h}_k^{\text{LoS}}$ denotes the analytically generated LoS channel at grid point k . The beam map therefore exposes the configuration-dependent direct-coverage pattern.

Second, two directional edge maps are extracted from the height map:

$$\begin{aligned} \mathbf{E}^{(x)} &= \mathcal{D}_x(\mathbf{G}), \\ \mathbf{E}^{(y)} &= \mathcal{D}_y(\mathbf{G}), \end{aligned} \quad (40)$$

where $\mathcal{D}_x(\cdot)$ and $\mathcal{D}_y(\cdot)$ denote directional edge operators along the two spatial axes. The edge maps expose geometric boundaries that may affect blockage, reflection, and diffraction.

Third, a blockage score map $\mathbf{S} \in \mathbb{R}^{H \times W}$ is constructed from the observable geometry and BS position. For grid point k , the BS-grid line segment is sampled at M intermediate locations:

$$S_k = \frac{1}{M} \sum_{m=1}^M [h_{\text{bld}}(\mathbf{p}_{\text{BS}} + t_m(\mathbf{p}_k - \mathbf{p}_{\text{BS}})) - h_{\text{ray}}(t_m)]_+, \quad (41)$$

where $\mathbf{p}_{\text{BS}}$ and $\mathbf{p}_k$ denote the horizontal positions of the BS and grid point k , respectively, $\{t_m\}_{m=1}^M \subset (0, 1)$ are interpolation factors, $h_{\text{bld}}(\cdot)$ denotes the building height, $h_{\text{ray}}(t_m)$ denotes the height of the direct BS-grid segment,

and $[a]_+ = \max(a, 0)$. The blockage score is clipped and normalized before being supplied to the network.

Prediction is performed only over the valid non-building region. Let $\mathbf{M}_{\text{valid}} \in \{0, 1\}^{H \times W}$ denote the corresponding binary mask. The propagation-feature representation is written as

$$\mathbf{X}_\Phi = [\mathbf{B}, \mathbf{G}, \mathbf{E}^{(x)}, \mathbf{E}^{(y)}, \mathbf{S}]. \quad (42)$$

The beam map, directional edge maps, and blockage score are deterministic transformations of the available environment and configuration information. Therefore, they do not constitute genuine observation enrichment in the sense of Proposition 1. Instead, they reorganize the original observation into propagation-related representations and serve as feature-level propagation priors. RadioFR and RadioLSR are constructed from the same underlying observable environment and configuration information. However, they organize and supply the above representations differently according to their respective prior-guided base structures.

For subsequent training, let $\mathbf{y}$ denote the normalized ground-truth radiomap. For a prediction $\mathbf{a}$, target $\mathbf{b}$, and binary mask $\mathbf{M}$, define the masked squared L_2 loss as

$$\ell_{\text{mask}}^{(2)}(\mathbf{a}, \mathbf{b}; \mathbf{M}) = \frac{\|(\mathbf{a} - \mathbf{b}) \odot \mathbf{M}\|_{\text{F}}^2}{\max(\|\mathbf{M}\|_1, 1)}. \quad (43)$$

Both instantiations follow a two-stage training procedure. The prior-guided base is first trained using its model-specific objective and then frozen. Only the residual branch is subsequently optimized against the difference between the training radiomap and the frozen base prediction under the masked squared L_2 loss. All prediction branches in RadioFR and RadioLSR use the common U-Net backbone summarized in Table I, where C_b denotes the initial channel width of each branch.

2) *RadioFR (Feature-Residual Prediction)*: RadioFR realizes RadioDecomp with a monolithic network that jointly uses all propagation-related features to predict the radiomap. These feature representation in (42) is directly supplied to a U-Net:

$$b_{\text{F}}(\mathbf{z}) = f_{\text{F}}(\mathbf{X}_\Phi) \odot \mathbf{M}_{\text{valid}}, \quad (44)$$

Here, f_{F} , a monolithic U-Net, jointly processes all input channels. Relative to the height-map and beam-map inputs, adding edge and blockage features preserves the available information and does not introduce compression bias. Nevertheless, finite capacity, limited training data, and imperfect optimization may leave predictable radiomap errors. RadioFR therefore tests whether residual correction remains useful in this general feature-augmentation setting. The residual branch receives the same features together with the base radiomap:

$$\mathbf{X}_{\text{R},\text{F}} = [\mathbf{B}, \mathbf{G}, \mathbf{E}^{(x)}, \mathbf{E}^{(y)}, \mathbf{S}, b_{\text{F}}(\mathbf{z})]. \quad (45)$$

The residual prediction is

$$\hat{\mathbf{r}}_{\text{F}} = g_{\text{F}}(\mathbf{X}_{\text{R},\text{F}}) \odot \mathbf{M}_{\text{valid}}, \quad (46)$$

where g_{F} is implemented using another U-Net.

The final RadioFR prediction is

$$\hat{\mathbf{y}}_{\text{F}} = b_{\text{F}}(\mathbf{z}) + \hat{\mathbf{r}}_{\text{F}}. \quad (47)$$

In the first training stage, the feature-guided base is optimized using

$$\mathcal{L}_{\text{base},\text{F}} = \ell_{\text{mask}}^{(2)}(b_{\text{F}}(\mathbf{z}), \mathbf{y}; \mathbf{M}_{\text{valid}}). \quad (48)$$

In the second stage, the trained base is frozen, and the residual target is defined as

$$\mathbf{y}_{\text{res},\text{F}} = \mathbf{y} - b_{\text{F}}(\mathbf{z}). \quad (49)$$

Only the residual branch is then optimized using

$$\mathcal{L}_{\text{res},\text{F}} = \ell_{\text{mask}}^{(2)}(\hat{\mathbf{r}}_{\text{F}}, \mathbf{y}_{\text{res},\text{F}}; \mathbf{M}_{\text{valid}}). \quad (50)$$

3) *RadioLSR: LoS-Shadow-Residual Prediction*: RadioLSR examines a more restrictive propagation-guided base for radiomap blind prediction. It partitions the prediction region using the blockage score and assigns different inputs to the two branches: the beam map for the LoS region, and the beam map together with the blockage score for the Shadow region. This design emphasizes direct-coverage and blockage trends while limiting the geometric information available to the base prediction. Based on the blockage score, the LoS and Shadow masks are defined as

$$\begin{aligned} \mathbf{M}_{\text{LoS}} &= \mathbb{I}(\mathbf{S} \leq \epsilon) \odot \mathbf{M}_{\text{valid}}, \\ \mathbf{M}_{\text{Shd}} &= \mathbb{I}(\mathbf{S} > \epsilon) \odot \mathbf{M}_{\text{valid}}, \end{aligned} \quad (51)$$

where ϵ is the blockage-score threshold. The two masks satisfy

$$\mathbf{M}_{\text{LoS}} + \mathbf{M}_{\text{Shd}} = \mathbf{M}_{\text{valid}}. \quad (52)$$

The LoS branch uses the beam map to estimate the direct-coverage trend:

$$\hat{\mathbf{y}}_{\text{LoS}} = f_{\text{LoS}}(\mathbf{B}) \odot \mathbf{M}_{\text{LoS}}, \quad (53)$$

where f_{LoS} is implemented using a U-Net.

The Shadow branch jointly uses the beam map and blockage score:

$$\hat{\mathbf{y}}_{\text{Shd}} = f_{\text{Shd}}([\mathbf{B}, \mathbf{S}]) \odot \mathbf{M}_{\text{Shd}}, \quad (54)$$

where f_{Shd} is another U-Net.

The LoS-Shadow structured base is then given by

$$\begin{aligned} b_{\text{LS}}(\mathbf{z}) &= \hat{\mathbf{y}}_{\text{LoS}} + \hat{\mathbf{y}}_{\text{Shd}} \\ &= f_{\text{LoS}}(\mathbf{B}) \odot \mathbf{M}_{\text{LoS}} + f_{\text{Shd}}([\mathbf{B}, \mathbf{S}]) \odot \mathbf{M}_{\text{Shd}}. \end{aligned} \quad (55)$$

The two branches are supervised by the total radiomap within their respective regions. However, their restricted inputs may miss geometry-dependent coverage variations, including those associated with reflection and diffraction. The residual branch therefore receives the height and edge maps alongside the beam map and base prediction to recover predictable spatial details unavailable to the base:

$$\mathbf{X}_{\text{R},\text{LS}} = [\mathbf{B}, \mathbf{G}, \mathbf{E}^{(x)}, \mathbf{E}^{(y)}, b_{\text{LS}}(\mathbf{z})]. \quad (56)$$

The residual prediction is

$$\hat{\mathbf{r}}_{\text{LS}} = g_{\text{LS}}(\mathbf{X}_{\text{R},\text{LS}}) \odot \mathbf{M}_{\text{valid}}, \quad (57)$$

where g_{LS} is implemented using a U-Net.

The final RadioLSR prediction is

$$\hat{\mathbf{y}}_{\text{LS}} = b_{\text{LS}}(\mathbf{z}) + \hat{\mathbf{r}}_{\text{LS}}. \quad (58)$$

TABLE I

Unified U-Net backbone used by the prediction branches.

Stage	Specification
Encoder	DoubleConv (C_b), then $3 \times [\text{MaxPool} + \text{DoubleConv}]$ Channels: $2C_b, 4C_b, 8C_b$
Bottleneck	DoubleConv ($8C_b$)
Decoder	$3 \times [\text{Up} + \text{skip concat} + \text{DoubleConv}]$ Channels: $4C_b, 2C_b, C_b$
Output	1×1 convolution with one output channel

In the first training stage, the LoS and Shadow branches are optimized over their corresponding regions using

$$\mathcal{L}_{\text{base,LS}} = \lambda_{\text{LoS}} \ell_{\text{mask}}^{(2)}(\hat{\mathbf{y}}_{\text{LoS}}, \mathbf{y}; \mathbf{M}_{\text{LoS}}) + \lambda_{\text{Shd}} \ell_{\text{mask}}^{(2)}(\hat{\mathbf{y}}_{\text{Shd}}, \mathbf{y}; \mathbf{M}_{\text{Shd}}). \quad (59)$$

where $\lambda_{\text{LoS}} \geq 0$ and $\lambda_{\text{Shd}} \geq 0$ weight the LoS- and Shadow-region objectives, respectively.

In the second stage, the trained base branches are frozen, and only g_{LS} is optimized using

$$\mathcal{L}_{\text{res,LS}} = \ell_{\text{mask}}^{(2)}(\hat{\mathbf{r}}_{\text{LS}}, \mathbf{y} - b_{\text{LS}}(\mathbf{z}); \mathbf{M}_{\text{valid}}). \quad (60)$$

VI. EXPERIMENTS

In the experiments, we validate the preceding analysis and evaluate the effectiveness of RadioDecomp under different training and evaluation conditions.

A. Experimental Setup

1) *Dataset and Evaluation Protocols*: Experiments are conducted on the U6G XL-MIMO Radiomap dataset [12], which contains 78,400 radiomaps generated from 800 urban scenes and 98 BS configurations. Each sample contains an observable environment representation, a BS configuration representation, and a 128×128 received-power radiomap. Unless otherwise specified, the dataset is divided using a 6/1/3 training/validation/test ratio with random seed 42. Three evaluation protocols are considered:

- *Random*: all radiomap samples are randomly divided into training, validation, and test sets.
- *Cross-config*: the BS configurations used for testing are disjoint from those used for training.
- *Cross-env*: the urban scenes used for testing are disjoint from those used for training.

2) *Compared Predictors*: The following controlled predictors are considered:

- *Vanilla*: a standalone U-Net that directly predicts the radiomap from the observable environment and BS configuration. The environment is represented by the height map, while the BS configuration is represented by a binary 3-dB beam mask.
- *BaseFR*: the feature-guided monolithic base defined in Section V-B2, without residual correction.
- *RadioFR*: the feature-guided monolithic instantiation defined in Section V-B2.
- *BaseLSR*: the LoS-Shadow structured base defined in Section V-B3, without residual correction.

- *RadioLSR*: the complete LoS-Shadow-Residual instantiation defined in Section V-B3.

Unless otherwise specified, all experiments use the following initial channel widths. BaseFR uses $C_b = 32$, while the base and residual branches of RadioFR both use $C_b = 32$. For BaseLSR and RadioLSR, the LoS and Shadow branches each use $C_b = 16$; the residual branch of RadioLSR uses $C_b = 32$.

3) *Training and Metrics*: All predictors are optimized using AdamW with a batch size of 32 and an initial learning rate of 10^{-3} . All training objectives use the masked squared L_2 loss defined in (43). RadioFR and RadioLSR are trained in two stages: the base predictor is trained first and then frozen, after which only the residual branch is optimized. Unless otherwise specified, each training stage uses 60 epochs, and the checkpoint with the lowest validation loss is retained. For the RadioLSR base, $\lambda_{\text{LoS}} = \lambda_{\text{Shd}} = 0.5$.

Performance is evaluated over the valid non-building region. Training and test Root Mean Squared Error (RMSE) are reported to measure squared-error performance, while test Mean Absolute Error (MAE) is additionally provided as a complementary measure of average prediction error that is less sensitive to large local errors.

B. Propagation-Prior Guidance and Residual Correction

We first evaluate the effects of propagation-prior injection and residual correctability under the three evaluation protocols. Table II compares the considered predictors.

Under the random protocol, the training and test RMSE values remain close for all predictors, with gaps below 0.24 dB. This confirms that random splitting produces only a small train-test gap. In contrast, the gaps become substantially larger under the cross-config and cross-env protocols. Vanilla is consistently outperformed by BaseFR, RadioFR, and RadioLSR in both test RMSE and MAE under all three protocols. In particular, compared with Vanilla, BaseFR reduces test RMSE by 2.3339, 4.8974, and 3.4660 dB under the random, cross-config, and cross-env protocols, respectively. Since BaseFR does not include residual correction, these improvements demonstrate the benefit of injecting propagation-related representations. BaseLSR does not provide uniformly better standalone performance because its LoS-Shadow structure deliberately models only selected propagation effects, leaving the remaining discrepancy for residual correction.

Consistent with this design, residual correction improves both prior-guided bases under all three protocols. Relative to BaseFR, RadioFR reduces test RMSE by 0.0487, 0.4684, and 0.3178 dB. The correction is more pronounced for BaseLSR: RadioLSR reduces test RMSE by 3.8386, 1.8635, and 2.0240 dB. The consistent improvements over two structurally different bases support the effectiveness of RadioDecomp in correcting their remaining predictable discrepancies.

Fig. 1 provides representative spatial comparisons under the two shifted protocols. RadioFR refines the local coverage details captured by BaseFR and reduces the corresponding localized errors. The correction is particularly pronounced from BaseLSR to RadioLSR in the cross-config setting: the residual branch captures additional spatial variations not represented by

TABLE II

Training RMSE and test RMSE/MAE (dB) under the 6/1/3 random, cross-config, and cross-env protocols.

Method	Random			Cross-config			Cross-env		
	Train RMSE	Test RMSE	Test MAE	Train RMSE	Test RMSE	Test MAE	Train RMSE	Test RMSE	Test MAE
Vanilla	5.7382	5.9414	4.0935	5.7474	11.3364	8.7591	8.5897	11.6593	8.3522
BaseFR	3.4333	3.6075	2.1264	3.7145	6.4390	4.6388	6.1419	8.1933	4.7766
RadioFR	3.3841	3.5588	2.0856	3.0630	5.9706	4.2466	4.8478	7.8755	4.4046
BaseLSR	7.2615	7.2742	3.7781	7.2774	11.0581	8.2632	8.4982	10.0741	5.3493
RadioLSR	3.2009	3.4356	1.8905	3.0369	9.1946	7.1526	5.9953	8.0501	4.4138

TABLE III

Prediction performance with only 10 training radiomaps under the random protocol.

Method	Train RMSE	Test RMSE	Test MAE
Vanilla	14.9871	17.9310	13.6009
BaseFR	12.3616	15.6376	11.0481
BaseLSR	11.6288	20.8432	11.8843
RadioFR	11.2497	14.4285	10.4248
RadioLSR	7.9732	19.6369	9.3178

the LoS-Shadow base, which may arise from reflection and other complex propagation interactions. These results visually support the corrective role of RadioDecomp.

C. Propagation-Prior Learning under Extremely Limited Supervision

In practical deployments, acquiring sufficient radiomaps for a target area can be costly. We therefore restrict the training set to only 10 radiomaps under the random protocol to evaluate RadioDecomp under extremely limited supervision. Each training stage uses 200 epochs in this experiment. The results are reported in Table III.

Compared with the standard random protocol, the ten-sample setting produces a substantially larger train-test gap. The test RMSE exceeds the training RMSE by at least 2.9439 dB for all predictors, whereas the corresponding gaps (random) in Table II are below 0.24 dB. Vanilla exhibits the smallest train-test RMSE gap of 2.9439 dB. This may partly result from its simpler input representation and weaker training fit, as it also has the highest training RMSE. Moreover, BaseFR reduces the test RMSE from 17.9310 to 15.6376 dB and the test MAE from 13.6009 to 11.0481 dB, demonstrating that propagation-related representations remain beneficial when labeled radiomaps are scarce.

Residual correction improves both prior-guided bases. Relative to BaseFR, RadioFR reduces the training RMSE by 1.1119 dB and the test RMSE by 1.2091 dB, achieving the lowest test RMSE of 14.4285 dB. Relative to BaseLSR, RadioLSR reduces the training RMSE by 3.6556 dB, the test RMSE by 1.2063 dB, and the test MAE by 2.5665 dB. These consistent reductions support the effectiveness of RadioDecomp under extremely limited supervision.

RadioLSR achieves the lowest test MAE of 9.3178 dB, although its test RMSE remains relatively high. Its LoS-Shadow partition may make the dominant propagation patterns easier to fit over most locations, thereby reducing average

absolute error, while leaving a small number of large errors in complex or imperfectly partitioned regions. These localized errors receive substantially greater weight in RMSE, which explains the divergence between the two metrics.

D. Residual Correction across Feature-Guided Base Capacities

Although feature concatenation does not alter the population conditional target, a base predictor learned from finite samples may still retain an optimisation discrepancy. We examine this effect by varying the initial channel width of BaseFR over $C_b \in \{32, 48, 64, 96\}$. For each width, RadioFR uses the corresponding frozen BaseFR and a residual branch with a fixed width of 32.² Table IV reports the results under the three evaluation protocols.

Under the random and cross-config protocols, residual correction is more beneficial at smaller or intermediate base widths. The largest RMSE reductions are 0.1707 dB under the random protocol and 0.4684 dB under cross-config. At $C_b = 96$, the absolute RMSE differences are only 0.0031 and 0.0038 dB under the random and cross-config protocols, respectively, with RadioFR being marginally worse in the latter case. These results suggest that a sufficiently expressive BaseFR can capture most of the predictable relation in these settings. In particular, the analytically constructed beam map directly represents configuration-dependent radiation, which may limit the remaining correctable discrepancy under cross-config evaluation.

A different pattern appears under cross-env, where RadioFR reduces test RMSE at every width from 0.2245 to 0.3315 dB. Cross-environment prediction involves unseen spatial layouts, for which the relation between the observable geometric features and the radiomap is less completely covered by the training samples. The residual branch therefore remains useful for correcting finite-sample errors relative to the feature-guided base. RadioFR-64 also achieves a lower cross-env RMSE than BaseFR-96 (7.6678 versus 8.0248 dB) despite using fewer parameters, indicating that the gain is not attributable solely to increased model capacity.

E. Training-Support Coverage under a Fixed Test Domain

Corollary 1 motivates a practical data-acquisition question: when the objective is to minimize prediction error over a pre-scribed target domain, should training data prioritize relevant

²Fixing the residual capacity provides a controlled comparison across base widths; a wider residual branch may provide additional correction but is not considered here.

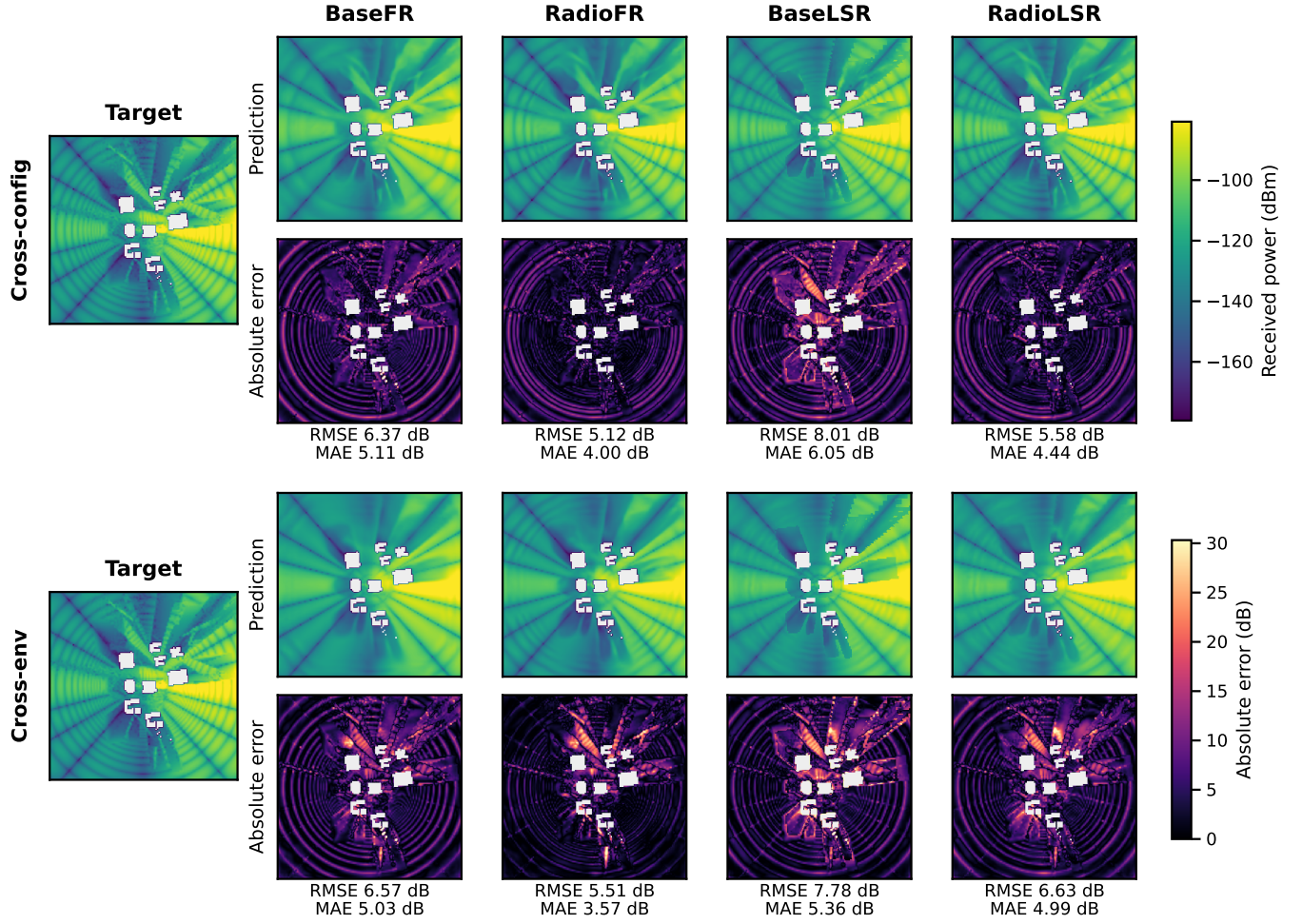


Fig. 1: Qualitative comparison on representative cross-config and cross-env test samples, with radiomaps and error maps using common received-power and error scales, respectively, white regions denoting masked buildings, and per-sample RMSE and MAE reported below each error map.

TABLE IV

Residual correction across feature-guided base capacities under the three evaluation protocols, with all RadioFR residual branches using $C_b = 32$ and bold indicating the lower error within each BaseFR-RadioFR pair.

Method	Params	Random			Cross-config			Cross-env		
		Train RMSE	Test RMSE	Test MAE	Train RMSE	Test RMSE	Test MAE	Train RMSE	Test RMSE	Test MAE
BaseFR-32	3.1295M	3.4333	3.6075	2.1264	3.7145	6.4390	4.6388	6.1419	8.1933	4.7766
BaseFR-48	7.0375M	2.9885	3.2183	1.9032	2.7451	5.5745	3.9754	4.6841	8.0554	4.4492
BaseFR-64	12.5075M	2.5282	2.8246	1.6136	2.5515	5.6783	4.0489	4.9321	7.9993	4.4656
BaseFR-96	28.1340M	2.5231	2.8096	1.6132	2.1523	5.7090	4.0976	4.1787	8.0248	4.2922
RadioFR-32	6.2594M	3.3841	3.5588	2.0856	3.0630	5.9706	4.2466	4.8478	7.8755	4.4046
RadioFR-48	10.1673M	2.7781	3.0476	1.7520	2.7298	5.5734	3.9512	4.1774	7.7778	4.2241
RadioFR-64	15.6373M	2.4831	2.7865	1.5862	2.5248	5.5063	3.8939	4.2651	7.6678	4.1429
RadioFR-96	31.2638M	2.5185	2.8065	1.6042	2.1504	5.7128	4.0915	3.9909	7.8003	4.2199

environments, relevant BS configurations, or both? To isolate how this choice affects the train-test gap, we keep the test set fixed. Consequently, the test observable distribution and conditional-uncertainty term remain unchanged, and performance differences mainly reflect how the selected training data affect the learned predictor over the same target domain.

We partition the 800 environments into $\mathcal{E}_{te}$ and $\mathcal{E}_{other}$, each containing 400 environments, and the 98 configurations into $\mathcal{C}_{te}$ and $\mathcal{C}_{other}$, each containing 49 configurations. The fixed

test set is $\mathcal{D}_{te} = \mathcal{E}_{te} \times \mathcal{C}_{te}$, which contains 19,600 radiomaps. We consider four fixed-budget training policies:

- *Both Covered*: $(\mathcal{E}_{te} \times \mathcal{C}_{other}) \cup (\mathcal{E}_{other} \times \mathcal{C}_{te})$;
- *Environment Only*: $\mathcal{E}_{te} \times \mathcal{C}_{other}$;
- *Configuration Only*: $\mathcal{E}_{other} \times \mathcal{C}_{te}$;
- *Neither Covered*: $\mathcal{E}_{other} \times \mathcal{C}_{other}$.

Each policy uses 17,600 training and 2,000 validation radiomaps. We additionally include *Both Covered (Full)*, which uses the complete leakage-free Both Covered pool,

TABLE V

Fixed-test performance of RadioLSR under different policies.

Training Policy	Train RMSE	Test RMSE	Test MAE
Both Covered	3.4032	4.6924	2.5305
Environment Only	3.1738	5.7332	3.8413
Configuration Only	3.0217	9.3269	4.4443
Neither Covered	3.0554	10.4874	6.1990
Both Covered (Full)	3.1288	4.4540	2.3567

with 37,200 training and 2,000 validation samples. No exact environment-configuration pair in the fixed test set is used for training or validation. The RadioLSR architecture and optimization settings are held constant.

As shown in Table V, the four policies obtain similar training RMSE values, ranging only from 3.0217 to 3.4032 dB under the common 17,600-sample budget, but their fixed-test performance differs substantially. Neither Covered produces the highest test RMSE and MAE of 10.4874 and 6.1990 dB. Relative to this policy, Configuration Only reduces the two errors by 1.1605 and 1.7547 dB, whereas Environment Only reduces them by 4.7542 and 2.3577 dB. Thus, when only one factor can be covered, acquiring data from test-relevant environments is more beneficial than acquiring data with test-relevant configurations in the current dataset utilising RadioLSR. Covering both factors yields complementary gains and achieves the best fixed-budget results, with a test RMSE of 4.6924 dB and MAE of 2.5305 dB. These differences, with comparable training errors and no pair-level leakage, show that the train-test gap depends strongly on where the training samples are collected.

Expanding Both Covered to the complete candidate pool more than doubles the training-set size, from 17,600 to 37,200, but provides smaller additional reductions of 0.2384 dB in RMSE and 0.1738 dB in MAE. The results suggest a practical acquisition priority: training data should first cover both the target-relevant environments and configurations, with particular attention to environment coverage; increasing sampling density within the covered support provides a further but smaller benefit.

F. Effect of Observable-Representation Coarsening

Proposition 1 shows that enriching the observation with additional predictive information can reduce irreducible conditional uncertainty. We examine this effect in the reverse direction by progressively coarsening the observable environment and BS-configuration representations. The same simple monolithic U-Net, optimization strategy, and random 6/1/3 split across all settings are used. Training RMSE is treated as the primary indicator of information loss, while test RMSE is reported as a practical reference because input coarsening may also change the train-test distribution mismatch.

We consider two BS-configuration representations: the continuous beam map in Section V-B1 and its binary 3-dB high-gain-region mask. The environment is represented by either the complete height map or a coarsened version with 10% of the height observations removed. Table VI reports the four resulting input combinations.

TABLE VI

RMSE under different beam and height-map representations.

Beam Input	Height Input	Train RMSE	Test RMSE
Continuous	Complete	3.4488	3.6185
Continuous	10% missing	3.5844	3.7440
3-dB mask	Complete	5.7382	5.9414
3-dB mask	10% missing	5.7254	5.9323

The continuous beam representation provides the dominant improvement. Relative to the 3-dB mask, it reduces training RMSE by 2.2894 dB with the complete height map and by 2.1410 dB under 10% height missingness. In comparison, removing 10% of the height observations increases training RMSE by only 0.1356 dB with the full beam map and changes it by a negligible 0.0128 dB in the opposite direction with the 3-dB mask. Under the controlled training setting, the substantial reductions obtained from the continuous beam map are consistent with the lower conditional uncertainty enabled by richer observable information. The test results largely follow the same dominant trend. The only exception is that the fourth setting outperforms the third by 0.0091 dB. This minor reversal may result from changes in the induced train-test distribution mismatch or finite-sample optimization variation and does not indicate that removing height information lowers conditional uncertainty.

VII. CONCLUSION

This paper investigated deterministic radiomap blind prediction under incomplete observation. Under squared loss, we identified its population-optimal target as the conditional-mean radiomap. Based on this target, we established an exact risk decomposition into reducible predictor approximation and irreducible uncertainty induced by hidden propagation factors. Building on this decomposition, we characterized the train-test risk gap through domain-dependent changes in both terms and proved that informative observation enrichment reduces the uncertainty term.

With the observation fixed, predictor design can further reduce approximation error. We therefore examined propagation priors as physically grounded guidance for this reduction, while showing that their implementations may bias the attainable predictor. To retain this guidance while correcting such deviations, we proposed RadioDecomp, which augments a prior-guided base with residual refinement. Its feature-guided and LoS-Shadow structured instantiations, RadioFR and RadioLSR, improved upon their corresponding bases under random, distribution-shifted, and limited-data settings. Controlled studies of model capacity, training-support coverage, and observation coarsening further supported the analysis. Overall, observation enrichment and correctable propagation-prior learning provide complementary routes to reducing irreducible uncertainty and predictor-approximation error, respectively.

REFERENCES

- [1] Y. Zeng, J. Chen, J. Xu, D. Wu, X. Xu, S. Jin, X. Gao, D. Gesbert, S. Cui, and R. Zhang, "A Tutorial on Environment-Aware Communications via Channel Knowledge Map for 6G," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 3, pp. 1478–1519, 2024.

- [2] H. Quan, W. Ni, T. Zhang, X. Ye, Z. Xie, S. Wang, Y. Liu, and H. Song, "Large Language Model Agents for Radio Map Generation and Wireless Network Planning," *IEEE Networking Letters*, vol. 7, no. 3, pp. 166–170, 2025.
- [3] S. Zhang, X. Ning, X. Zheng, Q. Shi, T.-H. Chang, and Z.-Q. Luo, "A Physics-Based and Data-Driven Approach for Localized Statistical Channel Modeling," *IEEE Transactions on Wireless Communications*, vol. 23, no. 6, pp. 5409–5424, 2024.
- [4] B. Li and J. Chen, "Radio Map-Assisted Routing and Predictive Resource Allocation Over Dynamic Low-Altitude Networks," *IEEE Transactions on Wireless Communications*, vol. 25, pp. 9955–9970, 2026.
- [5] R. Ding, F. Zhou, M. Giordani, Q. Wu, and M. Zorzi, "Toward 6G: Spectrum Management in Low-Altitude Aerial Information Networks," *IEEE Network*, pp. 1–9, 2026.
- [6] H. Jiang, J. Wang, J. Mu, C. Xing, H. Wymeersch, N. Zhao, and D. Niyato, "Covert Communications: A Generative Spectrum Map Paradigm," *IEEE Wireless Communications*, pp. 1–8, 2026.
- [7] K. Wang, T.-X. Zheng, J. Shi, Z. Yang, J. An, and Z. Li, "Movable Antenna Assisted Covert Communication With Wavelength-Scale Beamforming Resolution," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 12, pp. 19864–19869, 2025.
- [8] Z. Xing, G. Zhang, L. Xie, X. Ma, W. Zhao, Y. Liu, Y. Chen, Y. Zhao, and L. Cui, "Constructing Radio Map via Spatio-Temporal Learning of Crowdsourced Trajectory data," *IEEE Transactions on Mobile Computing*, pp. 1–18, 2026.
- [9] T. Hu, Y. Huang, J. Chen, Q. Wu, and Z. Gong, "3D Radio Map Reconstruction Based on Generative Adversarial Networks Under Constrained Aircraft Trajectories," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 6, pp. 8250–8255, 2023.
- [10] D. Schäufele, R. L. G. Cavalcante, and S. Stanczak, "Tensor Completion for Radio Map Reconstruction using Low Rank and Smoothness," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2019, pp. 1–5.
- [11] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering Radio Maps for Wireless Resource Management," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 133–141, 2019.
- [12] X. Li, Y. Han, Z. Lu, S. Jin, and C.-K. Wen, "U6G XL-MIMO Radiomap Prediction: Multi-Config Dataset and Beam Map Approach," 2026. [Online]. Available: <https://arxiv.org/abs/2603.06401>
- [13] V. Abhayawardhana, I. Wassell, D. Crosby, M. Sellars, and M. Brown, "Comparison of empirical propagation path loss models for fixed wireless access systems," in *2005 IEEE 61st Vehicular Technology Conference*, vol. 1, 2005, pp. 73–77 Vol. 1.
- [14] J. Bian, C.-X. Wang, X. Gao, X. You, and M. Zhang, "A General 3D Non-Stationary Wireless Channel Model for 5G and Beyond," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3211–3224, 2021.
- [15] J. Zhang, J. Lin, P. Tang, W. Fan, Z. Yuan, X. Liu, H. Xu, Y. Lyu, L. Tian, and P. Zhang, "Deterministic Ray Tracing: A Promising Approach to THz Channel Modeling in 6G Deployment Scenarios," *IEEE Communications Magazine*, vol. 62, no. 2, pp. 48–54, 2024.
- [16] R. Levie, c. Yapar, G. Kutyniok, and G. Caire, "RadioUNet: Fast Radio Map Estimation With Convolutional Neural Networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 4001–4015, 2021.
- [17] S. Zhang, A. Wijesinghe, and Z. Ding, "RME-GAN: A Learning Framework for Radio Map Estimation Based on Conditional Generative Adversarial Network," *IEEE Internet of Things Journal*, vol. 10, no. 20, pp. 18 016–18 027, 2023.
- [18] X. Wang, K. Tao, N. Cheng, Z. Yin, Z. Li, Y. Zhang, and X. Shen, "RadioDiff: An Effective Generative Diffusion Model for Sampling-Free Dynamic Radio Map Construction," *IEEE Transactions on Cognitive Communications and Networking*, vol. 11, no. 2, pp. 738–750, 2025.
- [19] F. Jaensch, G. Caire, and B. Demir, "Radio Map Prediction From Aerial Images and Application to Coverage Optimization," *IEEE Transactions on Wireless Communications*, vol. 25, pp. 308–320, 2026.
- [20] X. Jiang, T. Li, Z. Xiao, K. Chen, S. Ma, Z. Wang, and K. Li, "UniRM: A Universal Large Model for Multiband 3D Radio Map Construction," *IEEE Journal on Selected Areas in Communications*, vol. 44, pp. 2244–2258, 2026.
- [21] H. Wang, X. Shi, H. Zhang, Y. Cao, S. Yang, J. Wang, and K. Huang, "BeamCKM: A Framework of Channel Knowledge Map Construction for Multi-Antenna Systems," *IEEE Transactions on Wireless Communications*, vol. 25, pp. 21 261–21 275, 2026.
- [22] L. Zhao, Y. Wang, X. Wang, Z. Fei, and Y. Zeng, "BeamCKMDiff: Beam-Aware Channel Knowledge Map Construction via Diffusion Transformer," in *IEEE INFOCOM 2026 - IEEE Conference on Computer Communications*, 2026, pp. 1–6.
- [23] X. Li, Z. Cai, N. Qi, C. Dong, G. Zhu, H. Ma, Q. Wu, and S. Jin, "A Disentangled Representation Learning Framework for Low-Altitude Network Coverage Prediction," *IEEE Transactions on Mobile Computing*, vol. 25, no. 5, pp. 6261–6276, 2026.
- [24] W. Chen and J. Chen, "Diffraction and Scattering Aware Radio Map and Environment Reconstruction Using Geometry Model-Assisted Deep Learning," *IEEE Transactions on Wireless Communications*, vol. 23, no. 12, pp. 19 804–19 819, 2024.
- [25] W. Lu, Z. Lu, J. Yan, and S. Gao, "SIP2Net: Situational-Aware Indoor Pathloss-Map Prediction Network for Radio Map Generation," in *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025, pp. 1–2.
- [26] X. Li, S. Zhang, H. Li, X. Li, L. Xu, H. Xu, H. Mei, G. Zhu, N. Qi, and M. Xiao, "RadioGAT: A Joint Model-Based and Data-Driven Framework for Multi-Band Radiomap Reconstruction via Graph Attention Networks," *IEEE Transactions on Wireless Communications*, vol. 23, no. 11, pp. 17 777–17 792, 2024.
- [27] Y. Zhou, A. Wijesinghe, Y. Ma, S. Zhang, and Z. Ding, "TiRE-GAN: Task-Incentivized Generative Learning for Radiomap Estimation," *IEEE Wireless Communications Letters*, vol. 14, no. 5, pp. 1401–1405, 2025.
- [28] X. Wang, Q. Zhang, N. Cheng, R. Sun, Z. Li, S. Cui, and X. Shen, "RadioDiff-k2: Helmholtz Equation Informed Generative Diffusion Model for Multi-Path Aware Radio Map Construction," *IEEE Journal on Selected Areas in Communications*, vol. 44, pp. 2318–2333, 2026.
- [29] M. Li, T. Wu, Z. Dong, X. Liu, Y. Lu, S. Zhang, Z. Wu, Y. Zhang, L. Yu, and J. Zhang, "DeepRT: A Hybrid Framework Combining Large Model Architectures and Ray Tracing Principles for 6G Digital Twin Channels," *Electronics*, vol. 14, no. 9, 2025. [Online]. Available: <https://www.mdpi.com/2079-9292/14/9/1849>
- [30] W. Chen and J. Chen, "Learning Blockage and Reflection Geometry for MIMO Beam Map Construction," in *ICC 2025 - IEEE International Conference on Communications*, 2025, pp. 2321–2326.
- [31] J. Zhou, C. Hu, G. Wu, Z. Ren, H. Hu, J. Zhang, R. Zhang, and J. Xu, "6D Channel Knowledge Map Construction via Bidirectional Wireless Gaussian Splatting," *arXiv preprint arXiv:2510.26166*, 2025.
- [32] X. Wang, T. Yuan, Y. Cao, N. Cheng, R. Sun, and W. Zhuang, "iRadioDiff: Physics-Informed Diffusion Model for Indoor Radio Map Construction and Localization," in *ICC 2026 - IEEE International Conference on Communications*, 2026, pp. 1–6.
- [33] F. Jiang, T. Li, X. Lv, H. Rui, and D. Jin, "Physics-Informed Neural Networks for Path Loss Estimation by Solving Electromagnetic Integral Equations," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 15 380–15 393, 2024.
- [34] L. Liu, X. Chen, Z. Tang, M. Ma, and W. Zhang, "PINN and GNN-based RF Map Construction for Wireless Communication Systems," in *2025 International Conference on Future Communications and Networks (FCN)*, 2025, pp. 1–6.
- [35] H. Jia, W. Chen, Z. Huang, L. Wang, H. Xiao, N. Jia, K. Wu, S. Lai, B. Tian, and Y. Yue, "Physics-Informed Representation Alignment for Sparse Radio-Map Reconstruction," in *Proceedings of the 33rd ACM International Conference on Multimedia*, ser. MM '25. New York, NY, USA: Association for Computing Machinery, 2025, p. 12352–12360. [Online]. Available: <https://doi.org/10.1145/3746027.3758161>
- [36] W. Li, B. Wang, H. Fan, Y. Zhang, G. Zhu, and Y. Zhang, "Cross-Domain Radiomap Prediction for Multi-Scatterer Environment: A Spherical-Wave-Based Approach," 2026. [Online]. Available: <https://arxiv.org/abs/2605.02401>
- [37] L. Zhao, Z. Fei, X. Wang, J. Luo, and Z. Zheng, "3D-RadioDiff: An Altitude-Conditioned Diffusion Model for 3D Radio Map Construction," *IEEE Wireless Communications Letters*, vol. 14, no. 7, pp. 1969–1973, 2025.
- [38] Z. Liu, Q. Liu, S. Zhang, H. Zhang, and L. Song, "A Fine-Grained 3D Radio Map Construction Paradigm With Ultra-Low Sampling Rates by Large Generative Models," *IEEE Journal on Selected Areas in Communications*, vol. 44, pp. 4397–4413, 2026.
- [39] J. Jeong, L. Tsang, X. Xu, S. H. Yueh, and S. A. Margulis, "Full-Wave Electromagnetic Model Simulations of P-Band Radio Signal Propagation Through Forest Using the Fast Hybrid Method," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 6711–6722, 2023.