TEACHING DIFFUSION TO UNDERSTAND WIRELESS PHYSICS: HELMHOLTZ-GUIDED SIGNAL GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Generative models have recently demonstrated strong capabilities in learning complex data distributions and synthesizing signals. However, for wireless signals with real electromagnetic propagation characteristics, purely data-driven approaches can learn statistical patterns from observations but often struggle to capture the underlying propagation structures behind signal variations. Meanwhile, physics-based methods provide propagation-consistent signal modeling but typically require accurate scene information, limiting their applicability in practical wireless environments. Therefore, how to incorporate propagation principles into generative models without scene priors and enable signal generation with both statistical and propagation characteristics remains a key challenge. In this work, we propose Helm-Diffusion, a novel diffusion process guided by the Helmholtz equation that incorporates signal propagation principles into the generation trajectory. Helm-Diffusion exploits the propagation consistency implied by the Helmholtz equation to construct approximate representations of propagation structures from finite time frequency observations, and further derives static consistency for characterizing shared cross-frequency propagation structures and dynamic consistency for describing the evolution patterns of different frequency responses under changing propagation environments. Based on these propagation representations, Helm-Diffusion reformulates the stochastic generation process of conventional DDPMs by embedding propagation-structure degradation and recovery mechanisms into the progressive diffusion trajectory. This enables the generation process to incorporate propagation structure information while learning the statistical characteristics of wireless signals, achieving a synergistic integration of physics guidance and data-driven generation. Experiments on multiple wireless sensing and communication datasets demonstrate that Helm-Diffusion effectively improves generation quality over existing generative approaches while enhancing the consistency between generated signals and propagation characteristics, highlighting the effectiveness of physical laws as structured priors for generative models.

1 INTRODUCTION

Wireless signal generation is a fundamental capability for research in wireless communication and sensing systems, providing essential support for a wide range of key tasks, including channel modeling, system simulation, algorithm design, and performance evaluation (Poddar et al., 2023; Chi et al., 2024). As emerging wireless systems continue to evolve toward higher frequency bands, wider bandwidths, larger antenna arrays, and greater intelligence (Tataria et al., 2021), wireless signals exhibit increasingly complex propagation characteristics and time-frequency structures, placing higher demands on the realism and fidelity of signal generation. Therefore, how to efficiently generate high-fidelity wireless signals that accurately reflect real-world propagation characteristics has become an important research problem in wireless communication and sensing (Chi et al., 2024).

Existing wireless signal generation approaches mainly follow two paradigms: physics-based models and data-driven models.

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

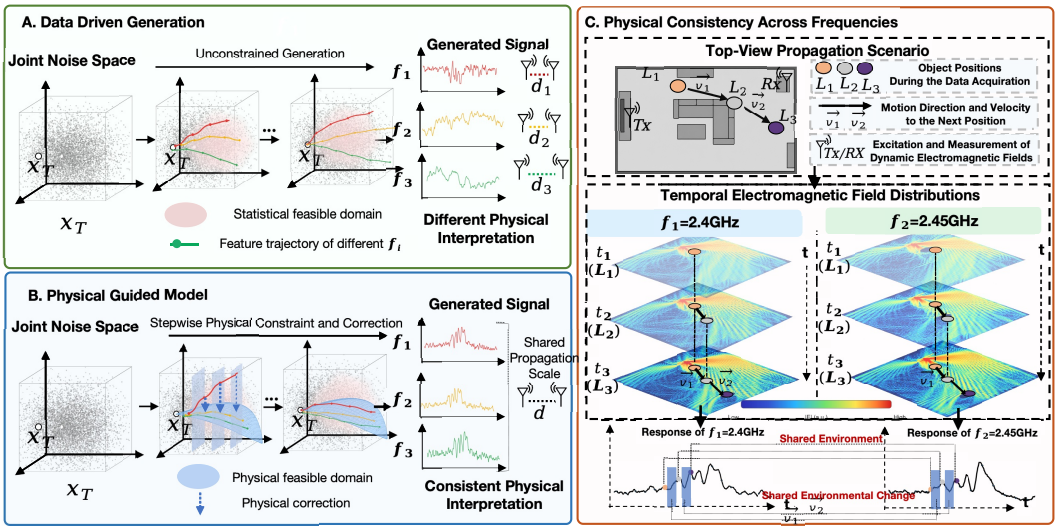


Figure 1: Motivation of physics-consistent wireless signal generation. Data-driven models learn statistical patterns but may violate physical propagation consistency, while physics-based simulation requires costly environmental modeling. Our approach learns shared propagation structures and incorporates physical consistency during generation.

Physics-based Models. Physics-based wireless signal generation methods typically employ deterministic propagation modeling approaches, such as Ray Tracing (Seidel & Rappaport, 1994; Hoydis et al., 2023), to explicitly simulate signal propagation. By modeling propagation paths, scene structures, material properties, and antenna configurations, these methods can generate signals with realistic propagation characteristics. However, high-fidelity physical modeling relies on accurate environmental priors, including 3D scene geometry, electromagnetic material properties, and antenna characteristics, which are often difficult to obtain in practical scenarios. Moreover, dynamic environments require continuous updates of propagation paths and channel states, further increasing the complexity of physical modeling.

Data-driven Models. Recent wireless signal generation methods leverage generative models, including Diffusion models (Chi et al., 2024), GANs (Wang et al., 2023; Yang et al., 2019), and VAEs (Xia et al., 2022), to learn wireless signal distributions from large-scale observations without explicit environmental modeling. For example, in indoor human activity scenarios, wireless propagation is jointly affected by complex structures and dynamic objects, such as walls, furniture, devices, and moving targets, making fine-grained physical modeling difficult to apply directly. Data-driven methods can learn statistical signal patterns from measurements, but primarily focus on distribution characteristics and dynamic fluctuations, while overlooking the physical consistency imposed by the shared propagation environment. Specifically, different subcarriers at the same time correspond to frequency-dependent responses of the same propagation environment and should preserve consistent propagation relationships. Consequently, generated signals may match the statistical characteristics of real measurements while lacking physically consistent propagation structures, limiting their applicability to physical-scale-sensitive tasks such as localization and navigation.

We observe that different frequency responses are jointly governed by the same propagation environment and should therefore satisfy shared physical consistency. And such consistency should not be imposed only on the final output, but continuously constrain the evolution of the generated signal throughout the generation process as shown in Fig. 1.

However, realizing this idea involves two key challenges: 1) The theoretical challenge is how to extract compact physical consistency structures from limited wireless observations. Although electromagnetic propagation can be described by physical equations, accurately solving these equations usually requires extensive prior knowledge of the propagation environment, such as spatial geometry, material properties, and boundary conditions, which are difficult to obtain in practical scenarios. 2) The modeling challenge is how to embed these structures into the signal generation process and continuously guide the evolution of generated signals toward physically consistent solutions. Since the generation process involves a gradual transition from stochastic states to realistic signals, con-

straints imposed only on the final output cannot prevent intermediate states from deviating from the physically valid space.

In this paper, we propose Helm-Diffusion, a physics-informed diffusion framework for high-fidelity wireless signal generation. To address the theoretical challenge, we exploit the cross-frequency consistency of electromagnetic propagation and leverage the Helmholtz equation to extract compact propagation structures from limited wireless observations, providing an implicit approximation of propagation environments without requiring explicit environmental parameters. To address the modeling challenge, we observe that the multi-step denoising trajectory of diffusion naturally provides opportunities for continuous physical intervention. Therefore, we design a physics–statistics diffusion process that incorporates the extracted propagation structures into each generation step, progressively correcting intermediate signal states and steering the generation trajectory toward physically consistent solutions.

Our contributions are summarized as follows:

Physical consistency modeling. We reveal and exploit the intrinsic physical consistency of wireless signals across frequency and time dimensions, and develop compact propagation representations from limited observations without requiring complete environmental information.

Physics-informed diffusion generation. We propose **Helm-Diffusion**, a physics-informed diffusion framework that injects propagation consistency into the generation trajectory, enabling wireless signals with both high statistical fidelity and physical consistency.

Comprehensive evaluation. We conduct extensive experiments across multiple wireless signal generation tasks and datasets. The results show that Helm-Diffusion consistently improves generation quality over existing state-of-the-art methods, while better preserving physical-scale consistency and enhancing robustness in downstream localization tasks. Ablation studies further validate the effectiveness of our diffusion process and denoising design.

2 Related Work

Wireless Signal Generation. GANs, VAEs, and diffusion models have been used for wireless signal and channel generation (Yang et al., 2019; Wang et al., 2023; Xie et al., 2023; Xia et al., 2022; Chi et al., 2024). RF-Diffusion (Chi et al., 2024) introduced time-frequency diffusion for WiFi and FMCW data, while physics-informed latent models further incorporate structured channel priors (Böck et al., 2025). These methods mainly learn signal distributions or predefined representations, without explicitly maintaining propagation consistency across frequency and time throughout generation.

Physics-guided Learning for Wireless Systems. Physical priors have been introduced through structured channel representations, Helmholtz-based radio maps, and environment-aware information (Böck et al., 2025; Wang et al., 2026b;a). Such methods typically rely on predefined physical representations or explicit environmental information. Helm-Diffusion instead derives propagation constraints from measured signals and applies them throughout denoising.

Diffusion Models with Physical Constraints. Physics-informed diffusion models incorporate governing equations through physical residuals, PDE priors, or constrained sampling (Bastek et al., 2025; Huang et al., 2024; Cheng et al., 2025). These methods generally assume that the governing equation or its residual can be directly evaluated. Helm-Diffusion addresses wireless settings where propagation structure must instead be inferred from measurements.

3 Methods

Helm-Diffusion incorporates propagation structure directly into the wireless signal generation trajectory. As illustrated in Fig. 2, we first derive static cross-frequency consistency and local dynamic consistency from the Helmholtz equation. These physical structures are then used to construct a propagation degradation before conventional statistical corruption. During reverse generation, the Denoise Module first recovers the signal statistical characteristics of the wireless signal, after which the Physics Guided Module restores the corresponding propagation structure. Throughout this section, τ denotes physical time and t denotes the diffusion step.

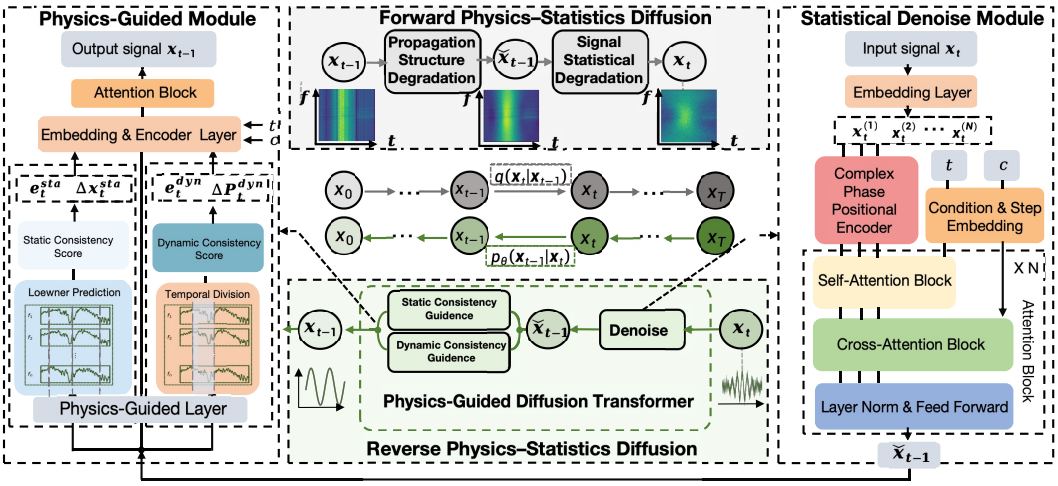


Figure 2: Overview of Helm-Diffusion.

3.1 Helmholtz-Based Physical Consistency

Wireless propagation in a time-varying environment follows

$$\nabla \times \left[\mu^{-1}(\mathbf{r},\tau)\nabla \times \mathbf{E}(\mathbf{r},\omega,\tau)\right] - \omega^2 \epsilon(\mathbf{r},\tau)\mathbf{E}(\mathbf{r},\omega,\tau) = \mathbf{q}(\mathbf{r},\omega), \qquad s = \omega^2. \tag{1}$$

Applying Galerkin discretization (Hiptmair, 2002; Nédélec, 1980) yields

$$[\mathbf{K}(\tau) - s\mathbf{M}(\tau)]\mathbf{e}(s,\tau) = \mathbf{b}, \qquad H(s,\tau) = \mathbf{c}^H[\mathbf{K}(\tau) - s\mathbf{M}(\tau)]^{-1}\mathbf{b}. \tag{2}$$

Here, $\mathbf{K}(\tau)$ and $\mathbf{M}(\tau)$ characterize the propagation environment, while $\mathbf{b}$ and $\mathbf{c}$ denote transmitter excitation and receiver observation. Thus, responses at different frequencies are frequency-dependent realizations of the same underlying propagation system. The complete discretization is provided in Appendix A.1.

Static propagation consistency. At a fixed physical time τ, different frequency responses share the same $\mathbf{K}(\tau)$, $\mathbf{M}(\tau)$, $\mathbf{b}$, and $\mathbf{c}$. Because these responses arise from the same propagation environment, their shared cross-frequency dependence can be approximated by a common finite-order Loewner representation. For two frequency parameters μ_i and λ_j, the Loewner and shifted-Loewner matrices are

$$L_{ij} = \frac{H(\mu_i,\tau) - H(\lambda_j,\tau)}{\mu_i - \lambda_j}, \qquad (L_\sigma)_{ij} = \frac{\mu_i H(\mu_i,\tau) - \lambda_j H(\lambda_j,\tau)}{\mu_i - \lambda_j}. \tag{3}$$

Substituting Eq. (2) gives

$$\mathbf{L} = \mathbf{OMR}, \qquad \mathbf{L}_\sigma = \mathbf{OKR}, \qquad \mathbf{L}_\sigma - s\mathbf{L} = \mathbf{O}(\mathbf{K} - s\mathbf{M})\mathbf{R}, \tag{4}$$

For a square Loewner pencil formed from equal numbers of left and right observations, let $\mathbf{V}$ collect the responses at $\{\mu_i\}$ and $\mathbf{W}$ collect those at $\{\lambda_j\}$. Where the pencil is nonsingular, we define the finite-order approximation

$$\widehat{H}(s,\tau) = \mathbf{W}(\mathbf{L}_\sigma - s\mathbf{L})^{-1}\mathbf{V}. \tag{5}$$

The complete Helmholtz–Loewner derivation is given in Appendix A.1.

We assess static propagation consistency by comparing each observed response with its approximation obtained from the other frequencies. Let

$$\mathcal{F} = \{s_1 < s_2 < \cdots < s_F\}. \tag{6}$$

For target frequency s_i, we remove s_i and use the remaining observations to approximate the held-out response:

$$\widehat{H}^{\text{sta}}_{m,i} = \mathcal{P}^{(-i)}_{\text{adj-Loe}}\left(s_i; \{H(s_j,\tau_m)\}_{j\neq i}\right). \tag{7}$$

Repeating the procedure for all frequencies gives

$$\Delta H^{\mathrm{sta}}_{m,i} = \widehat{H}^{\mathrm{sta}}_{m,i} - H(s_i, \tau_m), \qquad e^{\mathrm{sta}}_{m,i} = \frac{|\Delta H^{\mathrm{sta}}_{m,i}|^2}{|H(s_i, \tau_m)|^2 + \varepsilon}. \tag{8}$$

The adjacent-frequency construction and leave-one-frequency-out prediction are detailed in Appendix A.2.

Dynamic propagation consistency. When the propagation environment evolves over time, $\mathbf{K}(\tau)$ and $\mathbf{M}(\tau)$ vary accordingly. Differentiating Eq. (2) with respect to physical time gives

$$\dot{\mathbf{e}}(s, \tau) = \mathbf{D}(s, \tau)\mathbf{e}(s, \tau), \qquad \mathbf{D}(s, \tau) = -[\mathbf{K}(\tau) - s\mathbf{M}(\tau)]^{-1}[\dot{\mathbf{K}}(\tau) - s\dot{\mathbf{M}}(\tau)]. \tag{9}$$

The derivation of the temporal propagation operator is provided in Appendix A.1.

Within a local temporal window centered at τ_m, the received response can be locally expressed as

$$H(f, \tau_m + u) \approx \sum_{\ell} \beta_\ell(f, \tau_m) e^{-j2\pi f \dot{\tau}_\ell(\tau_m) u}, \qquad \nu_\ell(f, \tau_m) = -f\dot{\tau}_\ell(\tau_m). \tag{10}$$

Hence,

$$\frac{\nu_\ell(f_i, \tau_m)}{f_i} = -\dot{\tau}_\ell(\tau_m). \tag{11}$$

We define

$$\xi = \frac{\nu}{f_i}, \qquad \widetilde{S}_i(\tau_m, \xi) = S_i(\tau_m, f_i\xi), \tag{12}$$

so that dynamic modes corresponding to the same propagation variation are aligned across frequencies. The derivation is provided in Appendix A.1.

To capture propagation variations at different temporal resolutions, we use multiple local windows

$$\mathcal{W} = \{W_1, \ldots, W_R\}. \tag{13}$$

For each location (τ_m, f_i) and window scale W_r, we compute a normalized dynamic distribution $P^{(r)}_{m,i}$ and compare it with the cross-frequency shared pattern $\bar{P}^{(r)}_m$. The resulting multi-scale dynamic consistency is

$$e^{\mathrm{dyn}}_{m,i} = \frac{1}{R}\sum_{r=1}^{R} \left\| P^{(r)}_{m,i} - \bar{P}^{(r)}_m \right\|_2^2, \qquad \Delta P^{\mathrm{dyn}}_{m,i} = \frac{1}{R}\sum_{r=1}^{R} \left(P^{(r)}_{m,i} - \bar{P}^{(r)}_m \right). \tag{14}$$

The multi-scale window construction and dynamic normalization are detailed in Appendix A.2.

3.2 Forward Physics-Statistics Diffusion

As shown in the upper branch of Fig. 2, Helm-Diffusion first degrades the propagation structure and subsequently applies conventional signal statistical corruption. Different from conventional statistical corruption, we construct a frequency-domain propagation degradation that progressively attenuates propagation structures along the frequency dimension.

Propagation structure degradation. The static and dynamic consistency derived above show that finite time–frequency observations can approximately recover the shared cross-frequency propagation structure and the common local propagation-evolution pattern. We progressively degrade these structures using a frequency-domain Gaussian blur operator

$$[\mathcal{M}_t(H)]_{m,i} = \sum_j G_t(i, j) H(\tau_m, f_j), \tag{15}$$

where $G_t(i, j)$ denotes the frequency-domain Gaussian blur kernel at diffusion step t.

The frequency-domain propagation degradation progressively removes the cross-frequency propagation structure captured by the static consistency. By attenuating the frequency-dependent response variation, the degraded signal gradually loses the finite-frequency representation of the shared propagation environment. Details of $\mathcal{M}_t$ are provided in Appendix A.4.

Signal statistical degradation. After propagation structure degradation, Gaussian corruption is applied to progressively remove the signal statistical characteristics of the wireless response:

$$\widetilde{x}_{t-1} = \mathcal{M}_t(x_{t-1}), \qquad x_t = \sqrt{\alpha_t}\,\widetilde{x}_{t-1} + \sqrt{1-\alpha_t}\,\epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{16}$$

The two degradation mechanisms therefore operate on complementary signal properties: $\mathcal{M}_t$ progressively removes the propagation structure, whereas Gaussian corruption progressively destroys the statistical characteristics of the wireless response. The frequency-domain degradation construction and degradation schedule are detailed in Appendix A.4. We further show in Appendix A.3 that the resulting forward process approaches a Gaussian noise state and admits a well-defined reverse recovery process.

3.3 Reverse Physics-Statistics Diffusion

As shown in the lower branch of Fig. 2, reverse generation first restores signal statistical characteristics and then recovers the propagation structure. At each diffusion step, the Denoise Module first estimates an intermediate reverse state, after which the Physics Guided Module projects it toward a propagation-consistent state.

Signal statistical recovery. Given the current diffusion state x_t, the Denoise Module first predicts the clean wireless response $\widetilde{x}_0$. It encodes the complex-valued wireless response to preserve its amplitude–phase structure. Self-attention captures dependencies within the wireless signal, while cross-attention incorporates the task condition and diffusion step to guide statistical recovery at different denoising stages.

Propagation structure recovery. For the statistically recovered candidate $\widetilde{x}_0$, the Physics Guided Module extracts the same static and dynamic propagation quantities defined in Sec. 3.1:

$$\mathcal{C}_t^{\mathrm{sta}} = [e_t^{\mathrm{sta}}, \Delta x_t^{\mathrm{sta}}], \qquad \mathcal{C}_t^{\mathrm{dyn}} = [e_t^{\mathrm{dyn}}, \Delta P_t^{\mathrm{dyn}}]. \tag{17}$$

The static cue is obtained from adjacent-frequency Loewner prediction, while the dynamic cue is obtained from the multi-scale normalized local dynamics.

The recovered signal and both propagation cues are jointly embedded:

$$z_t^{\mathrm{phy}} = \mathcal{E}_{\mathrm{phy}}\left(\widetilde{x}_0, e_t^{\mathrm{sta}}, \Delta x_t^{\mathrm{sta}}, e_t^{\mathrm{dyn}}, \Delta P_t^{\mathrm{dyn}}, t, c\right). \tag{18}$$

An attention-based projection block maps the prediction toward a propagation-consistent response:

$$z_t^{\mathrm{proj}} = \mathcal{A}_\phi(z_t^{\mathrm{phy}}), \qquad \widehat{x}_0 = \mathcal{P}_\phi(z_t^{\mathrm{proj}}). \tag{19}$$

The projected clean response is then used in the reverse transition defined in Appendix A.5.

The complete Physics-Guided Module is described in Appendix A.5.

3.4 Training Objective

Helm-Diffusion is trained on the final physics-corrected prediction. Given a degraded input x_t, the Denoise Module first estimates a clean response $\widetilde{x}_0$, which is further refined by the Physics Guided Module to obtain $\widehat{x}_0$.

The training objective jointly considers signal reconstruction and propagation consistency:

$$\mathcal{L}_{\mathrm{train}} = \mathcal{L}_{\mathrm{rec}} + \lambda_{\mathrm{sta}}\mathcal{L}_{\mathrm{sta}} + \lambda_{\mathrm{dyn}}\mathcal{L}_{\mathrm{dyn}}, \tag{20}$$

where $\mathcal{L}_{\mathrm{rec}}$ measures the signal reconstruction error between $\widehat{x}_0$ and the ground-truth response, while $\mathcal{L}_{\mathrm{sta}}$ and $\mathcal{L}_{\mathrm{dyn}}$ directly impose static and dynamic propagation consistency on the generated response. The detailed formulation of these objectives is provided in Appendix A.6.

4 EXPERIMENT

4.1 EXPERIMENT DESIGN

We evaluate Helm-Diffusion across multiple wireless signal generation tasks, including wireless sensing tasks (i.e., human activity recognition and localization) and communication signal generation tasks (i.e., FDD channel generation). We conduct extensive experiments on multiple datasets with diverse propagation characteristics and compare Helm-Diffusion with representative signal generation models, including diffusion-, GAN-, and VAE-based approaches. We evaluate the generated signals from both sample-level structural fidelity and distribution-level generation quality, validating the effectiveness of Helm-Diffusion in generating physically consistent wireless signals across diverse wireless applications.

4.1.1 COMPARATIVE METHODS

Baselines. We compare Helm-Diffusion with five representative generative models spanning diffusion-, GAN-, and VAE-based approaches. RF-Diffusion (Chi et al., 2024) is an RF-specific diffusion model that performs signal generation through time-frequency diffusion, and serves as our primary baseline since Helm-Diffusion is built upon its framework. CSVAE (Böck et al., 2025) is a physics-informed VAE that incorporates channel priors by learning a compressible channel representation under a predefined physical dictionary. DDPM (Ho et al., 2020) is included as a generic diffusion baseline without RF-specific designs. DCGAN (Radford et al., 2015) represents adversarial generative models, while CVAE (Sohn et al., 2015) represents conventional conditional latent-variable generation.

4.1.2 EVALUATION METRICS

To evaluate the quality of generated wireless signals, we adopt two widely used distribution-level metrics in signal generation, including Wasserstein Distance (WD) and Maximum Mean Discrepancy (MMD).

We evaluate Helm-Diffusion across three categories of wireless tasks. For HAR tasks, we conduct experiments on the Widar (Zhang et al., 2021), Falldar (Yang et al., 2022), and GaitSense (Zhang et al., 2022) datasets. For localization tasks, we use the ESPARGOS (Euchner et al., 2025) and DICHASUS (Euchner et al., 2021) datasets. For communication-oriented signal generation, we evaluate on the Argos (Shepard et al., 2016) and COST2100 (Liu et al., 2012) FDD channel datasets.

4.2 OVERALL PERFORMANCE

Table 1 summarizes the overall generation performance across sensing and communication tasks. Helm-Diffusion consistently achieves the lowest WD and MMD on all HAR and localization datasets, while attaining the highest SNR on both FDD datasets. Compared with RF-Diffusion, the average WD and MMD over the five sensing datasets are reduced by approximately 34.6% and 28.0%, respectively, showing that the improvement generalizes across different sensing scenarios rather than being tied to a particular dataset.

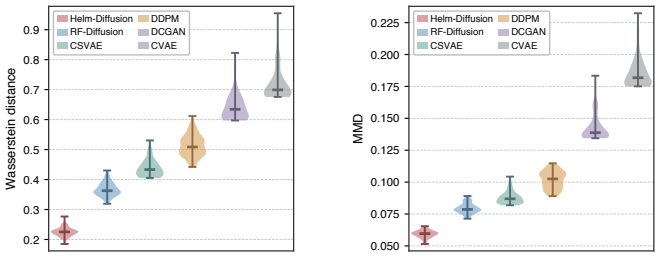


Figure 3: WD and MMD distributions in Table 1.

Fig. 3 further examines this advantage from the distribution perspective. The WD and MMD distributions of Helm-Diffusion are consistently shifted toward lower values and remain concentrated in the low-error region, whereas the competing methods exhibit progressively larger distribution discrepancies. This indicates that Helm-Diffusion improves not only the aggregate metrics in Table 1, but also the consistency of generated signals in matching the measured data distribution.

Model	HAR						Localization				FDD	
	Widar		Falldar		GaitSense		ESPARGOS		DICHASUS		Argos	COST2100
	WD↓	MMD↓	WD↓	MMD↓	WD↓	MMD↓	WD↓	MMD↓	WD↓	MMD↓	SNR↑	SNR↑
Helm-Diffusion	**0.2257**	**0.0594**	**0.2316**	**0.0548**	**0.2474**	**0.0517**	**0.2174**	**0.0506**	**0.2295**	**0.0529**	**28.81**	**27.94**
RF-Diffusion	0.3670	0.0794	0.3518	0.0756	0.3421	0.0728	0.3427	0.0719	0.3564	0.0747	26.97	26.21
CSVAE	0.4387	0.0879	0.4215	0.0834	0.4098	0.0807	0.3976	0.0788	0.4142	0.0815	20.35	19.82
DDPM	0.5126	0.1018	0.4964	0.0967	0.4819	0.0925	0.4938	0.0951	0.5072	0.0984	22.86	22.31
DCGAN	0.6438	0.1426	0.6215	0.1378	0.6087	0.1342	0.6194	0.1385	0.6358	0.1417	18.74	18.25
CVAE	0.7215	0.1863	0.7048	0.1812	0.6895	0.1764	0.7126	0.1835	0.7251	0.1872	15.62	15.18

Table 1: Overall performance comparison across HAR, localization, and FDD channel generation datasets. Lower WD and MMD and higher SNR indicate better performance.

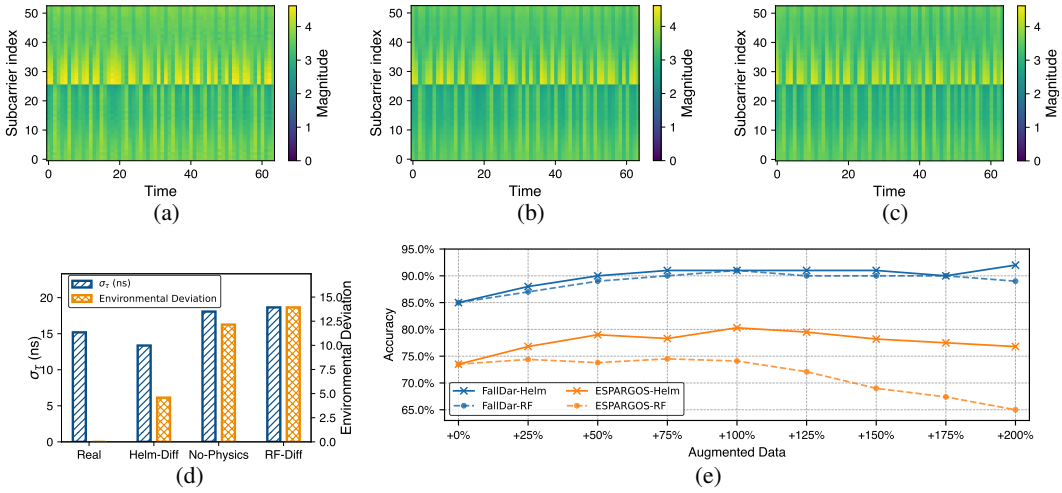


Figure 4: Analysis of statistical similarity, cross-subcarrier physical consistency, and downstream localization and HAR performance of generated wireless signals. (a) is real data, (b) and (c) are generated by Helm-Diffusion and RF-Diffusion respectively.

The same advantage extends to FDD channel generation, where Helm-Diffusion achieves 28.81 dB and 27.94 dB on Argos and COST2100, exceeding RF-Diffusion by 1.84 dB and 1.73 dB, respectively. Together, these results show that incorporating propagation consistency improves both distribution matching for sensing signals and signal fidelity for communication channels.

4.3 Beyond Statistical Similarity: The Importance of Physical Consistency

To examine this distinction, we analyze generated signals from three perspectives: signal-level statistics, cross-frequency physical structure, and downstream task performance. As shown in Fig. 4, both RF-Diffusion and Helm-Diffusion reproduce amplitude distributions and time-frequency patterns similar to real measurements, indicating that existing generative models can already capture the statistical characteristics of wireless signals.

To further examine whether generated signals preserve the underlying physical propagation structure, we specifically select data collected under a fixed propagation environment. Under this setting, different frequency components may exhibit different responses, but they are generated by the same environmental geometry, boundary conditions, and propagation structure, and therefore should reflect a common physical propagation scale. We divide each sample into multiple wide frequency bands and estimate the propagation delay (ToF) from the phase–frequency relationship within each band. As shown in Fig. 4, σ_τ measures the dispersion of ToF estimates across frequency bands within the same sample, and thus characterizes whether the cross-frequency propagation relationship is consistent with that of real signals. Real multipath signals also exhibit nonzero cross-band ToF dispersion; the proximity of Helm-Diffusion's dispersion to this real-data level indicates that it preserves realistic frequency-dependent variation rather than forcing identical delays across bands.

Environmental Deviation further uses ToF as a physical-scale indicator and measures the deviation between the mean ToF of a generated signal and the reference ToF obtained from real measurements in the same fixed environment.

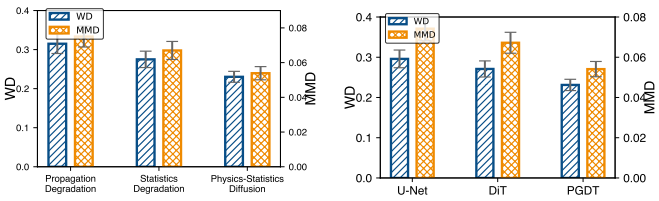
For comparison, No-Physics denotes a variant of Helm-Diffusion that removes only the Physics-Guided Module from the reverse process, while retaining the Denoise Module and the same forward diffusion process, including Propagation Structure Degradation. This comparison evaluates the contribution of explicit physics-guided refinement during reverse generation. Compared with No-Physics and RF-Diffusion, Helm-Diffusion exhibits cross-frequency delay behavior closer to real data while achieving a substantially smaller Environmental Deviation. In particular, the comparison with No-Physics supports the contribution of the Physics-Guided Module to recovering the propagation structure of the target environment. The results for RF-Diffusion further illustrate that statistical similarity alone does not necessarily imply a propagation scale consistent with the real environment.

This difference is further reflected in downstream tasks. As shown in Fig. 4, HAR mainly relies on motion-induced temporal variations and statistical signal patterns, and therefore both RF-Diffusion and Helm-Diffusion benefit from additional synthetic samples. Localization, however, depends more strongly on spatial propagation information jointly encoded across frequency components. As the proportion of synthetic data increases, the degradation in localization performance with RF-Diffusion is consistent with inaccurate cross-frequency propagation structures interfering with the spatial representations learned by the localization model. Helm-Diffusion better preserves the propagation structure of the target environment and therefore provides more reliable synthetic data.

4.4 Ablation Study

Beyond the No-Physics comparison in Fig. 4, we conduct ablation studies to examine the contributions of the diffusion method and network design in Helm-Diffusion.

Additional ablations of the static and dynamic physics designs are reported in Appendix A.7.



(a) Diffusion method. (b) Network design.

Figure 5: Ablation study of Helm-Diffusion.

Impact of diffusion methods. To validate the effectiveness of the proposed physics-guided diffusion process, we keep the network architecture unchanged and only replace the forward degradation strategy with three variants: statistical degradation, propagation degradation, and their combination. As shown in Fig. 5a, the complete design achieves lower WD and MMD than either degradation alone. Statistical degradation mainly destroys the signal distribution, while propagation degradation further perturbs the propagation-dependent structures across time and frequency.

Impact of network design. To evaluate the contribution of PGDT, we fix the diffusion process and replace the denoising backbone with PGDT, vanilla DiT, and U-Net, respectively. As shown in Fig. 5b, PGDT achieves lower WD and MMD than both alternatives. Compared with U-Net, which mainly models local patterns through convolution, and vanilla DiT, which performs generic token-level modeling without explicit physical structure awareness, PGDT better captures long-range dependencies across the temporal and frequency dimensions and aligns with the propagation structures perturbed by the proposed diffusion process, leading to more accurate signal restoration.

5 Conclusion

We proposed Helm-Diffusion, which incorporates cross-frequency and temporal propagation consistency into wireless signal generation. Experiments across sensing and communication datasets show improvements in generation quality and propagation realism. Our results offer a promising direction for generating wireless signals by guiding diffusion with physical propagation structure.

AI USE STATEMENT

Generative AI tools were used during the preparation of this manuscript only for language refinement, grammar checking, and improving the clarity of presentation. The authors independently developed all research ideas, theoretical formulations, algorithmic designs, experimental settings, and analyses. All AI-assisted revisions were carefully reviewed by the authors, who take full responsibility for the accuracy, originality, and integrity of the manuscript.

ETHICAL CONSIDERATIONS

This work studies physics-consistent wireless signal generation and does not involve the collection of new personal data or human-subject experiments. The proposed method is designed to improve the reliability of synthetic wireless data for research purposes, including wireless sensing, channel modeling, and communication system evaluation.

Although generated signals may facilitate simulation and algorithm development, they should be properly validated before deployment in real-world or safety-critical systems. We encourage responsible use of synthetic wireless data and transparent reporting of generated data in downstream applications.

REFERENCES

Jan-Hendrik Bastek, WaiChing Sun, and Dennis M. Kochmann. Physics-informed diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025.

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

Chaoran Cheng, Boran Han, Danielle C. Maddix, Abdul Fatir Ansari, Andrew Stuart, Michael W. Mahoney, and Bernie Wang. Gradient-free generation for hard-constrained systems. In *The Thirteenth International Conference on Learning Representations*, 2025.

Guoxuan Chi, Zheng Yang, Chenshu Wu, Jingao Xu, Yuchong Gao, Yunhao Liu, and Tony Xiao Han. Rf-diffusion: Radio signal generation via time-frequency diffusion. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, pp. 77–92, 2024.

Florian Euchner, Marc Gauger, Sebastian Dörner, and Stephan ten Brink. A Distributed Massive MIMO Channel Sounder for "Big CSI Data"-Driven Machine Learning. In *WSA 2021; 25th International ITG Workshop on Smart Antennas*, 2021.

Florian Euchner, David Kellner, Phillip Stephan, and Stephan ten Brink. CSI Dataset espargos-0007: Passive Target with Four Synchronized ESPARGOS Antenna Arrays in a Lab Room, 2025. URL `https://doi.org/10.18419/DARUS-4973`.

Boris Gralak. Analytic properties of the electromagnetic green's function. *Journal of Mathematical Physics*, 58(7):071501, 2017. doi: 10.1063/1.4993199.

Ralf Hiptmair. Finite elements in computational electromagnetism. *Acta Numerica*, 11:237–339, 2002.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

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

Jiahe Huang, Guandao Yang, Zichen Wang, and Jeong Joon Park. Diffusionpde: Generative pde-solving under partial observation. *Advances in Neural Information Processing Systems*, 37: 130291–130323, 2024.

Lingfeng Liu, Claude Oestges, Juho Poutanen, Katsuyuki Haneda, Pertti Vainikainen, François Quitin, Fredrik Tufvesson, and Philippe De Doncker. The COST 2100 MIMO Channel Model. *IEEE Wireless Communications*, 19(6):92–99, 2012. doi: 10.1109/MWC.2012.6393523.

Jean-Claude Nédélec. Mixed finite elements in 3. *Numerische Mathematik*, 35(3):315–341, 1980.

Hitesh Poddar, Shihao Ju, Dipankar Shakya, and Theodore S. Rappaport. A tutorial on nyusim: Subterahertz and millimeter-wave channel simulator for 5g, 6g, and beyond. *IEEE Communications Surveys & Tutorials*, 26(2):824–857, 2023.

Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

Scott Y. Seidel and Theodore S. Rappaport. Site-specific propagation prediction for wireless in-building personal communication system design. *IEEE Transactions on Vehicular Technology*, 43(4):879–891, 1994.

Clayton Shepard, Jian Ding, Ryan E. Guerra, and Lin Zhong. Understanding Real Many-Antenna MU-MIMO Channels. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pp. 461–467, 2016. doi: 10.1109/ACSSC.2016.7869082.

Kihyuk Sohn, Honglak Lee, and Xinchen Yan. Learning structured output representation using deep conditional generative models. *Advances in Neural Information Processing Systems*, 28, 2015.

Harsh Tataria, Mansoor Shafi, Andreas F. Molisch, Mischa Dohler, Henrik Sjöland, and Fredrik Tufvesson. 6g wireless systems: Vision, requirements, challenges, insights, and opportunities. *Proceedings of the IEEE*, 109(7):1166–1199, 2021.

Kiran Venugopal, Nuria Gonzalez-Prelcic, and Robert W Heath. Optimal frequency-flat precoding for frequency-selective millimeter wave channels. *IEEE Transactions on Wireless Communications*, 18(11):5098–5112, 2019.

Satyavrat Wagle, Akshay Malhotra, Shahab Hamidi-Rad, Aditya Sant, David J. Love, and Christopher G. Brinton. Physics-based generative models for geometrically consistent and interpretable wireless channel synthesis. In James Kwok (ed.), *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI-25*, pp. 9384–9392. International Joint Conferences on Artificial Intelligence Organization, 8 2025. doi: 10.24963/ijcai.2025/1043. URL https://doi.org/10.24963/ijcai.2025/1043. AI4Tech: AI Enabling Technologies.

Weidong Wang, Jiancheng An, Hongshu Liao, Lu Gan, and Chau Yuen. Radio generation using generative adversarial networks with an unrolled design. *arXiv preprint arXiv:2306.13893*, 2023.

Xiucheng Wang, Tingwei Yuan, Yang Cao, Nan Cheng, Ruijin Sun, and Weihua Zhuang. iradiodiff: Physics-informed diffusion model for indoor radio map construction and localization. In *IEEE International Conference on Communications (ICC)*, pp. 1–6. IEEE, 2026a. doi: 10.1109/ICC59461.2026.11587950.

Xiucheng Wang, Qiming Zhang, Nan Cheng, Ruijin Sun, Zan Li, Shuguang Cui, and Xuemin Shen. Radiodiff-k^2: Helmholtz equation informed generative diffusion model for multi-path aware radio map construction. *IEEE Journal on Selected Areas in Communications*, 44:2318–2333, 2026b. doi: 10.1109/JSAC.2025.3641105.

William Xia, Sundeep Rangan, Marco Mezzavilla, Angel Lozano, Giovanni Geraci, Vasilii Semkin, and Giuseppe Loianno. Generative neural network channel modeling for millimeter-wave uav communication. *IEEE Transactions on Wireless Communications*, 21(11):9417–9431, 2022.

Wenwu Xie, Ming Xiong, Zhihe Yang, Wei Liu, Lvrong Fan, and Jian Zou. Real and fake channel: Gan-based wireless channel modeling and generating. *Physical Communication*, 61:102214, 2023. doi: 10.1016/j.phycom.2023.102214.

Yang Yang, Yang Li, Wuxiong Zhang, Fei Qin, Pengcheng Zhu, and Cheng-Xiang Wang. Generative-adversarial-network-based wireless channel modeling: Challenges and opportunities. *IEEE Communications Magazine*, 57(3):22–27, 2019.

Zheng Yang, Yi Zhang, and Qian Zhang. Rethinking fall detection with wi-fi. *IEEE Transactions on Mobile Computing*, 22(10):6126–6143, 2022.

Yi Zhang, Yue Zheng, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. Widar3.0: Zero-effort cross-domain gesture recognition with wi-fi. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11):8671–8688, 2021.

Yi Zhang, Yue Zheng, Guidong Zhang, Kun Qian, Chen Qian, and Zheng Yang. Gaitsense: Towards ubiquitous gait-based human identification with wi-fi. *ACM Transactions on Sensor Networks*, 18 (1):1:1–1:24, February 2022. doi: 10.1145/3466638.

A Appendix

A.1 Derivation of Helmholtz-Based Propagation Consistency

This section provides the derivation underlying the static and dynamic propagation consistency used in Helm-Diffusion.

From the Helmholtz equation to the propagation pencil. Wireless propagation in a time-varying environment satisfies

$$\nabla \times \left[\mu^{-1}(\mathbf{r},\tau)\nabla \times \mathbf{E}(\mathbf{r},\omega,\tau)\right] - \omega^2 \epsilon(\mathbf{r},\tau)\mathbf{E}(\mathbf{r},\omega,\tau) = \mathbf{q}(\mathbf{r},\omega), \tag{21}$$

where $\mathbf{r}$ denotes the spatial coordinate, $\omega = 2\pi f$ is the angular frequency, $\mathbf{E}$ denotes the electromagnetic field, and ϵ and μ characterize the propagation environment.

We expand the field using a finite set of spatial basis functions

$$\mathbf{E}_N(\mathbf{r},\omega,\tau) = \sum_{j=1}^{N} e_j(\omega,\tau)\mathbf{N}_j(\mathbf{r}), \qquad \mathbf{e}(\omega,\tau) = [e_1,\ldots,e_N]^{\mathsf{T}}. \tag{22}$$

Applying Galerkin projection gives

$$\sum_{j=1}^{N} K_{ij}(\tau)e_j - \omega^2 \sum_{j=1}^{N} M_{ij}(\tau)e_j = b_i, \tag{23}$$

where

$$K_{ij}(\tau) = \int_{\Omega} (\nabla \times \mathbf{N}_i)^H \mu^{-1}(\mathbf{r},\tau)(\nabla \times \mathbf{N}_j)\, d\Omega, \qquad M_{ij}(\tau) = \int_{\Omega} \mathbf{N}_i^H \epsilon(\mathbf{r},\tau)\mathbf{N}_j\, d\Omega. \tag{24}$$

Defining $s = \omega^2$, the finite-dimensional propagation system becomes

$$[\mathbf{K}(\tau) - s\mathbf{M}(\tau)]\mathbf{e}(s,\tau) = \mathbf{b}. \tag{25}$$

Under the standard narrowband approximation and a fixed transmitter–receiver configuration, the excitation and observation responses vary negligibly across the considered subcarriers and can be approximated as frequency-flat (Venugopal et al., 2019). Therefore, $\mathbf{b}(\omega) \approx \mathbf{b}$ and $\mathbf{c}(\omega) \approx \mathbf{c}$ within the observed frequency band. The corresponding wireless response is

$$H(s,\tau) = \mathbf{c}^H[\mathbf{K}(\tau) - s\mathbf{M}(\tau)]^{-1}\mathbf{b}. \tag{26}$$

Thus, at a fixed physical time, responses across the considered frequencies approximately share the same $\mathbf{K}(\tau)$, $\mathbf{M}(\tau)$, $\mathbf{b}$, and $\mathbf{c}$, while their frequency dependence is represented by s. Therefore, within this approximation, responses at different frequencies share the same propagation matrices $\mathbf{K}(\tau)$ and $\mathbf{M}(\tau)$, as well as the same excitation and observation operators $\mathbf{b}$ and $\mathbf{c}$, while their frequency dependence enters through $s = \omega^2$.

Helmholtz–Loewner static consistency. At a fixed physical time τ, consider two disjoint sets of frequency parameters $\{\mu_i\}_{i=1}^{p}$ and $\{\lambda_j\}_{j=1}^{q}$. In the following algebra, $\mathbf{K}$ and $\mathbf{M}$ abbreviate $\mathbf{K}(\tau)$ and $\mathbf{M}(\tau)$. The Loewner and shifted-Loewner matrices are

$$L_{ij} = \frac{H(\mu_i,\tau) - H(\lambda_j,\tau)}{\mu_i - \lambda_j}, \qquad (L_\sigma)_{ij} = \frac{\mu_i H(\mu_i,\tau) - \lambda_j H(\lambda_j,\tau)}{\mu_i - \lambda_j}. \tag{27}$$

Let

$$\mathbf{A}_{\mu_i} = \mathbf{K} - \mu_i \mathbf{M}, \qquad \mathbf{A}_{\lambda_j} = \mathbf{K} - \lambda_j \mathbf{M}. \tag{28}$$

Using

$$\mathbf{A}_{\mu_i}^{-1} - \mathbf{A}_{\lambda_j}^{-1} = \mathbf{A}_{\mu_i}^{-1}(\mathbf{A}_{\lambda_j} - \mathbf{A}_{\mu_i})\mathbf{A}_{\lambda_j}^{-1}, \tag{29}$$

and

$$\mathbf{A}_{\lambda_j} - \mathbf{A}_{\mu_i} = (\mu_i - \lambda_j)\mathbf{M}, \tag{30}$$

we obtain

$$L_{ij} = \mathbf{c}^H(\mathbf{K} - \mu_i \mathbf{M})^{-1}\mathbf{M}(\mathbf{K} - \lambda_j \mathbf{M})^{-1}\mathbf{b}. \tag{31}$$

Similarly,

$$(L_\sigma)_{ij} = \mathbf{c}^H(\mathbf{K} - \mu_i \mathbf{M})^{-1}\mathbf{K}(\mathbf{K} - \lambda_j \mathbf{M})^{-1}\mathbf{b}. \tag{32}$$

Defining

$$\mathbf{O} = \begin{bmatrix} \mathbf{c}^H(\mathbf{K} - \mu_1 \mathbf{M})^{-1} \\ \vdots \\ \mathbf{c}^H(\mathbf{K} - \mu_p \mathbf{M})^{-1} \end{bmatrix}, \qquad \mathbf{R} = \begin{bmatrix} (\mathbf{K} - \lambda_1 \mathbf{M})^{-1}\mathbf{b} & \cdots & (\mathbf{K} - \lambda_q \mathbf{M})^{-1}\mathbf{b} \end{bmatrix}, \tag{33}$$

gives

$$\mathbf{L} = \mathbf{OMR}, \qquad \mathbf{L}_\sigma = \mathbf{OKR}, \qquad \mathbf{L}_\sigma - s\mathbf{L} = \mathbf{O}(\mathbf{K} - s\mathbf{M})\mathbf{R}. \tag{34}$$

Let

$$\mathbf{V} = [H(\mu_1,\tau), \ldots, H(\mu_p,\tau)]^{\mathsf{T}}, \qquad \mathbf{W} = [H(\lambda_1,\tau), \ldots, H(\lambda_q,\tau)]. \tag{35}$$

For prediction, we select the same number d of left and right observations, so that $p = q = d$ and the Loewner pencil is square. The divided-difference definitions imply that the j-th column of $\mathbf{L}_\sigma - \lambda_j \mathbf{L}$ equals $\mathbf{V}$, and the i-th row of $\mathbf{L}_\sigma - \mu_i \mathbf{L}$ equals $\mathbf{W}$. Thus, wherever the selected pencil is nonsingular, the associated finite-order rational response is

$$\widehat{H}(s,\tau) = \mathbf{W}(\mathbf{L}_\sigma - s\mathbf{L})^{-1}\mathbf{V}. \tag{36}$$

It interpolates the selected frequency responses when the pencil is nonsingular at those frequencies. The factorization in Eq. (34) connects these data-driven differences to the common propagation matrices. If the finite-dimensional transfer response has a minimal order resolved by sufficiently many nondegenerate observations, a Loewner realization of sufficient order can represent that response. With the local order used below, $\widehat{H}(s,\tau)$ provides an approximation at unobserved frequencies; its accuracy is evaluated by holding out measured responses.

Temporal propagation operator. We next characterize how the propagation response evolves when the environment varies with physical time τ. We assume that the transmitter excitation and receiver observation operators remain fixed within the considered local temporal interval, such that the temporal variation is attributed to the propagation environment represented by $\mathbf{K}(\tau)$ and $\mathbf{M}(\tau)$. Define

$$\mathbf{A}(s,\tau) = \mathbf{K}(\tau) - s\mathbf{M}(\tau). \tag{37}$$

The discretized propagation system can then be written as

$$\mathbf{A}(s,\tau)\mathbf{e}(s,\tau) = \mathbf{b}. \tag{38}$$

Differentiating this relation with respect to physical time gives

$$\dot{\mathbf{A}}(s,\tau)\mathbf{e}(s,\tau) + \mathbf{A}(s,\tau)\dot{\mathbf{e}}(s,\tau) = 0. \tag{39}$$

Since

$$\dot{\mathbf{A}}(s,\tau) = \dot{\mathbf{K}}(\tau) - s\dot{\mathbf{M}}(\tau), \tag{40}$$

we obtain

$$\dot{\mathbf{e}}(s,\tau) = \mathbf{D}(s,\tau)\mathbf{e}(s,\tau), \qquad \mathbf{D}(s,\tau) = -[\mathbf{K}(\tau) - s\mathbf{M}(\tau)]^{-1}[\dot{\mathbf{K}}(\tau) - s\dot{\mathbf{M}}(\tau)]. \tag{41}$$

Therefore, within the finite-dimensional Helmholtz model, $\mathbf{D}(s,\tau)$ characterizes the local temporal evolution induced by changes in the propagation environment. In particular, its frequency dependence is inherited from the same propagation pencil $\mathbf{K}(\tau) - s\mathbf{M}(\tau)$ that governs the cross-frequency response.

From the Helmholtz propagation operator to normalized local dynamics. Eq. 41 describes propagation evolution at the discretized operator level. We next connect this representation to the path-wise temporal modes used to construct the dynamic consistency in Helm-Diffusion.

Consider the continuous Helmholtz operator

$$\mathcal{A}(\omega,\tau) = \nabla \times \mu^{-1}(\mathbf{r},\tau)\nabla \times -\omega^2\epsilon(\mathbf{r},\tau). \tag{42}$$

Subject to the corresponding boundary or radiation conditions, and at frequencies for which the inverse exists, the associated Green operator is formally the inverse of the Helmholtz operator (Gralak, 2017),

$$\mathcal{G}(\omega,\tau) = \mathcal{A}^{-1}(\omega,\tau). \tag{43}$$

More precisely, the electromagnetic dyadic Green's function $\mathbf{G}(\mathbf{r},\mathbf{r}';\omega,\tau)$ is the integral kernel of this inverse operator, such that

$$[\mathcal{G}(\omega,\tau)\mathbf{q}](\mathbf{r}) = \int_\Omega \mathbf{G}(\mathbf{r},\mathbf{r}';\omega,\tau)\mathbf{q}(\mathbf{r}',\omega)\, d\mathbf{r}'. \tag{44}$$

This formulation also provides a direct operator-level connection to temporal propagation evolution. Since $\mathcal{A}(\omega,\tau)\mathcal{G}(\omega,\tau) = \mathcal{I}$, differentiating with respect to physical time yields

$$\dot{\mathcal{G}}(\omega,\tau) = -\mathcal{G}(\omega,\tau)\dot{\mathcal{A}}(\omega,\tau)\mathcal{G}(\omega,\tau). \tag{45}$$

This inverse-operator identity is the continuous counterpart of the finite-dimensional evolution in Eq. 41. Under Galerkin discretization, $\mathbf{A}(s,\tau) = \mathbf{K}(\tau) - s\mathbf{M}(\tau)$ represents the Helmholtz operator, while $[\mathbf{K}(\tau) - s\mathbf{M}(\tau)]^{-1}$ is the corresponding finite-dimensional resolvent. Consequently, for fixed excitation and observation operators, the measured channel response

$$H(s,\tau) = \mathbf{c}^H[\mathbf{K}(\tau) - s\mathbf{M}(\tau)]^{-1}\mathbf{b} \tag{46}$$

is a transmitter–receiver projection of the same Green-operator response.

To obtain a physically interpretable local representation of this operator evolution, we consider a standard effective multipath representation of the projected Green-function response. Over a sufficiently narrow frequency band and a short temporal interval, the response can be approximated by a finite collection of effective propagation components,

$$H(\omega,\tau) \approx \sum_{\ell=1}^{L} A_\ell(\omega,\tau)\exp[-jk(\omega)L_\ell(\tau)]\,, \tag{47}$$

where $A_\ell(\omega,\tau)$ contains the complex gain and transmitter–receiver coupling associated with the ℓ-th effective propagation component, and $L_\ell(\tau)$ denotes its effective propagation length.

Within a locally non-dispersive frequency band, the wavenumber can be approximated as

$$k(\omega) \approx \frac{\omega}{v}, \tag{48}$$

where v denotes the corresponding effective propagation velocity. Defining the propagation delay of the ℓ-th component as

$$\tau_\ell(\tau) = \frac{L_\ell(\tau)}{v}, \tag{49}$$

we have

$$e^{-jk(\omega)L_\ell(\tau)} = e^{-j\omega\tau_\ell(\tau)} = e^{-j2\pi f\tau_\ell(\tau)}. \tag{50}$$

Eq. 47 can therefore be written in the familiar path-wise channel form

$$H(f,\tau) \approx \sum_{\ell=1}^{L} a_\ell(f,\tau) e^{-j2\pi f \tau_\ell(\tau)}. \tag{51}$$

Importantly, Eq. 51 is not introduced as an independent propagation law. Rather, it serves as a local reduced representation of the same Green-operator response whose finite-dimensional evolution is described by Eq. 41. To make this connection explicit, under the fixed excitation and observation operators used above,

$$\dot{H}(s,\tau) = \mathbf{c}^H \dot{\mathbf{e}}(s,\tau) = \mathbf{c}^H \mathbf{D}(s,\tau)\mathbf{e}(s,\tau). \tag{52}$$

Thus, at the operator level, temporal channel variations are induced by the environmental perturbations $\dot{\mathbf{K}}(\tau)$ and $\dot{\mathbf{M}}(\tau)$.

On the other hand, differentiating the local path-wise representation in Eq. 51 gives

$$\dot{H}(f,\tau) \approx \sum_{\ell=1}^{L} \left[\dot{a}_\ell(f,\tau) - j2\pi f \dot{\tau}_\ell(\tau) a_\ell(f,\tau)\right] e^{-j2\pi f \tau_\ell(\tau)}. \tag{53}$$

Eq. 52 and 53 therefore describe the same local channel evolution at two different levels: the former characterizes the variation of the underlying Helmholtz propagation operator, whereas the latter expresses its locally observed effect through changes in effective propagation gains and delays. Within this local multipath reduction, variations represented by $\dot{\mathbf{K}}(\tau)$ and $\dot{\mathbf{M}}(\tau)$ manifest themselves through $\dot{a}_\ell(f,\tau)$ and $\dot{\tau}_\ell(\tau)$.

We now consider a short temporal window centered at τ_m. Within this window, we assume that the complex path gain varies slowly and that the path delay admits a first-order local approximation,

$$\tau_\ell(\tau_m + u) = \tau_\ell(\tau_m) + \dot{\tau}_\ell(\tau_m)u + \mathcal{O}(u^2). \tag{54}$$

Approximating $a_\ell(f,\tau_m+u) \approx a_\ell(f,\tau_m)$ and defining

$$\beta_\ell(f,\tau_m) = a_\ell(f,\tau_m) e^{-j2\pi f \tau_\ell(\tau_m)}, \tag{55}$$

the local response becomes

$$H(f,\tau_m+u) \approx \sum_{\ell=1}^{L} \beta_\ell(f,\tau_m) e^{-j2\pi f \dot{\tau}_\ell(\tau_m)u}. \tag{56}$$

Hence, the local temporal frequency associated with the ℓ-th propagation component is

$$\nu_\ell(f,\tau_m) = -f\dot{\tau}_\ell(\tau_m). \tag{57}$$

For two subcarriers f_i and f_j observing the same local propagation variation, this gives

$$\frac{\nu_\ell(f_i,\tau_m)}{f_i} = \frac{\nu_\ell(f_j,\tau_m)}{f_j} = -\dot{\tau}_\ell(\tau_m). \tag{58}$$

Therefore, although the absolute temporal frequency of a propagation component scales with the carrier frequency, the scaling is removed by normalization with respect to the subcarrier frequency. Components associated with the same local propagation variation are consequently aligned in the normalized temporal-frequency domain.

For subcarrier f_i, define its local temporal spectrum as

$$S_i(\tau_m,\nu) = \int H(f_i,\tau_m+u) w(u) e^{-j2\pi\nu u}\, du, \tag{59}$$

where $w(u)$ denotes the local analysis window. We further define the normalized temporal-frequency coordinate and the corresponding resampled spectrum as

$$\xi = \frac{\nu}{f_i}, \qquad \widetilde{S}_i(\tau_m,\xi) = S_i(\tau_m, f_i\xi). \tag{60}$$

According to Eq. 58, the ℓ-th propagation component is aligned at

$$\xi_\ell(\tau_m) = -\dot{\tau}_\ell(\tau_m) \tag{61}$$

for all subcarriers. The normalized local spectrum therefore provides a finite-observation representation of the temporal propagation variation induced by the underlying time-varying Helmholtz operator.

We emphasize that this cross-frequency alignment is a local first-order propagation relation: it follows from the common Green-operator evolution under the locally non-dispersive, slowly-varying multipath approximation, rather than assuming an exact global invariance for arbitrary time-varying wireless channels.

A.2 Static and Dynamic Consistency Computation

Adjacent-frequency Loewner prediction. At physical time τ_m, let

$$\mathcal{F} = \{s_1 < s_2 < \cdots < s_F\} \tag{62}$$

denote the ordered frequency parameters, where $s_i = (2\pi f_i)^2$. To evaluate the response at s_i, we remove it from the observed set:

$$\mathcal{F}_{-i} = \mathcal{F} \setminus \{s_i\} = \{s_{j_1} < \cdots < s_{j_{F-1}}\}. \tag{63}$$

The Loewner data are constructed from consecutive observations in $\mathcal{F}_{-i}$. For every adjacent pair $(s_{j_k}, s_{j_{k+1}})$, we compute

$$\ell_k^{(-i)} = \frac{H(s_{j_{k+1}}, \tau_m) - H(s_{j_k}, \tau_m)}{s_{j_{k+1}} - s_{j_k}}, \qquad \ell_{\sigma,k}^{(-i)} = \frac{s_{j_{k+1}} H(s_{j_{k+1}}, \tau_m) - s_{j_k} H(s_{j_k}, \tau_m)}{s_{j_{k+1}} - s_{j_k}}. \tag{64}$$

The adjacent differences retain the divided-difference information of all observed frequency pairs. For any $1 \le a < b \le F-1$, the Loewner difference between s_{j_a} and s_{j_b} satisfies

$$\frac{H(s_{j_b}, \tau_m) - H(s_{j_a}, \tau_m)}{s_{j_b} - s_{j_a}} = \sum_{k=a}^{b-1} \frac{s_{j_{k+1}} - s_{j_k}}{s_{j_b} - s_{j_a}} \ell_k^{(-i)}. \tag{65}$$

Likewise, the shifted Loewner difference satisfies

$$\frac{s_{j_b} H(s_{j_b}, \tau_m) - s_{j_a} H(s_{j_a}, \tau_m)}{s_{j_b} - s_{j_a}} = \sum_{k=a}^{b-1} \frac{s_{j_{k+1}} - s_{j_k}}{s_{j_b} - s_{j_a}} \ell_{\sigma,k}^{(-i)}. \tag{66}$$

Both identities follow by summing the differences over consecutive intervals, which cancels the intermediate response terms. Consequently, every nonadjacent divided difference among the observed frequencies can be recovered algebraically from the adjacent differences, without enumerating random frequency pairs. For an interior held-out frequency s_i, the two observed frequencies that bracket it are consecutive in $\mathcal{F}_{-i}$. Their divided differences are therefore included explicitly.

For each interior target, we select the two closest observed frequencies on each side of s_i and use Eqs. (65) and (66) to form the required 2×2 Loewner matrices. We evaluate the finite-order response in Eq. (36) at s_i using a numerically stabilized solve. If only one observation is available on one side, the closest bracketing pair gives an order-one predictor. At a band endpoint, the two closest observed frequencies give a one-sided order-one Loewner extrapolation. Denoting this procedure by $\mathcal{P}^{(-i)}_{\text{adj-Loe}}$, we obtain

$$\widehat{H}^{\text{sta}}_{m,i} = \mathcal{P}^{(-i)}_{\text{adj-Loe}}\left(s_i; \{H(s_j, \tau_m)\}_{j\neq i}\right). \tag{67}$$

Repeating the procedure at every physical time produces

$$\Delta H^{\text{sta}}_{m,i} = \widehat{H}^{\text{sta}}_{m,i} - H(s_i, \tau_m), \qquad e^{\text{sta}}_{m,i} = \frac{|\Delta H^{\text{sta}}_{m,i}|^2}{|H(s_i, \tau_m)|^2 + \varepsilon}. \tag{68}$$

These values form the static-consistency and discrepancy maps used by the Physics Guided Module.

Table 2: Average NMSE of held-out CSI prediction across tasks (lower is better).

	Average NMSE ↓	
Task	Loewner	Linear interpolation
HAR	**0.0323**	0.1588
Localization	**0.0174**	0.1325
FDD	**0.0087**	0.0134

Held-out frequency prediction. We evaluate the empirical predictability of CSI by holding out each interior frequency in turn. Table 2 compares a local order-two Loewner predictor with linear interpolation using the same available observations. The comparison uses all adjacent pairs among the remaining observations and reports

$$\mathrm{NMSE} = \frac{\sum_{\kappa,i} \left| \widehat{H}_\kappa(s_i) - H_\kappa(s_i) \right|^2}{\sum_{\kappa,i} |H_\kappa(s_i)|^2}, \tag{69}$$

where κ indexes evaluated CSI traces and i indexes held-out interior frequencies. The metric is computed separately for each method using the same targets and observed frequencies.

Multi-scale dynamic consistency. A single temporal window provides only one temporal resolution. We therefore use a set of local window lengths

$$\mathcal{W} = \{W_1, \ldots, W_R\}. \tag{70}$$

For the r-th scale, let $w_r(u)$ denote the corresponding local window. Writing the STFT S_i with an explicit scale index, we define

$$S_i^{(r)}(\tau_m, \nu) = \int H(f_i, \tau_m + u) w_r(u) e^{-j2\pi\nu u}\, du, \qquad \widetilde{S}_i^{(r)}(\tau_m, \xi) = S_i^{(r)}(\tau_m, f_i \xi). \tag{71}$$

The normalized dynamic distribution is

$$P_{m,i}^{(r)}(\xi) = \frac{|\widetilde{S}_i^{(r)}(\tau_m, \xi)|^2}{\sum_{\xi'} |\widetilde{S}_i^{(r)}(\tau_m, \xi')|^2 + \varepsilon}. \tag{72}$$

At each scale, the shared local dynamic pattern is

$$\bar{P}_m^{(r)}(\xi) = \frac{1}{F} \sum_{j=1}^{F} P_{m,j}^{(r)}(\xi). \tag{73}$$

We evaluate the discrepancy of each time–frequency element as

$$e_{m,i}^{(r)} = \left\| P_{m,i}^{(r)} - \bar{P}_m^{(r)} \right\|_2^2, \qquad \Delta P_{m,i}^{(r)} = P_{m,i}^{(r)} - \bar{P}_m^{(r)}. \tag{74}$$

The multi-scale dynamic cue is obtained by aggregation:

$$e_{m,i}^{\mathrm{dyn}} = \frac{1}{R} \sum_{r=1}^{R} e_{m,i}^{(r)}, \qquad \Delta P_{m,i}^{\mathrm{dyn}} = \frac{1}{R} \sum_{r=1}^{R} \Delta P_{m,i}^{(r)}. \tag{75}$$

All scales are evaluated on the same normalized ξ grid before aggregation. This produces a dynamic-consistency score and discrepancy feature for every time–frequency element.

A.3 Diffusion Process with Propagation Degradation

In this section, we analyze the statistical properties of the proposed propagation-aware diffusion process. Unlike conventional diffusion models that introduce only isotropic Gaussian perturbations, Helm-Diffusion incorporates a deterministic propagation degradation operator to gradually remove propagation structures during the forward trajectory. We show that this modification preserves the Gaussian transition property and results in a Gaussian terminal distribution with a structured covariance.

A.3.1 LINEAR GAUSSIAN TRANSITION

For simplicity, we vectorize the real and imaginary components of a wireless response into $\mathbf{x}_t \in \mathbb{R}^d$. The proposed propagation-aware degradation applies frequency-domain Gaussian blur kernels over the time–frequency grid. Since these kernels are determined only by the frequency index and diffusion step, they are independent of signal samples and random Gaussian perturbations.

Therefore, the propagation degradation can be represented as a deterministic linear operator:

$$\mathcal{M}_t(\mathbf{x}) = \mathbf{M}_t\mathbf{x}, \tag{76}$$

where $\mathbf{M}_t$ denotes the propagation-aware mixing operator at diffusion step t.

Accordingly, the forward transition of Helm-Diffusion is formulated as:

$$\mathbf{x}_t = \sqrt{\alpha_t}\mathbf{M}_t\mathbf{x}_{t-1} + \sqrt{1-\alpha_t}\boldsymbol{\epsilon}_t, \qquad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{77}$$

Since the propagation degradation is linear and the injected perturbation is Gaussian, the transition distribution remains Gaussian:

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}\left(\sqrt{\alpha_t}\mathbf{M}_t\mathbf{x}_{t-1}, (1-\alpha_t)\mathbf{I}\right). \tag{78}$$

Therefore, the propagation-aware degradation modifies the evolution path of diffusion states while preserving the fundamental Gaussian transition property required by diffusion models.

A.3.2 FORWARD DISTRIBUTION AND GAUSSIAN CONVERGENCE

We next write the forward process in a compact closed form. Define the linear transition operator at diffusion step t as

$$\mathbf{A}_t = \sqrt{\alpha_t}\mathbf{M}_t, \tag{79}$$

and the accumulated transition operator as

$$\boldsymbol{\Phi}_t = \mathbf{A}_t\mathbf{A}_{t-1}\cdots\mathbf{A}_1 = \prod_{k=1}^{t}\sqrt{\alpha_k}\mathbf{M}_k. \tag{80}$$

The one-step transition in Eq. (77) can then be written as

$$\mathbf{x}_t = \mathbf{A}_t\mathbf{x}_{t-1} + \sqrt{1-\alpha_t}\boldsymbol{\epsilon}_t, \qquad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{81}$$

Recursively applying the transition gives

$$\mathbf{x}_t = \boldsymbol{\Phi}_t\mathbf{x}_0 + \sum_{k=1}^{t}\sqrt{1-\alpha_k}\,\boldsymbol{\Phi}_{t:k+1}\boldsymbol{\epsilon}_k, \tag{82}$$

where

$$\boldsymbol{\Phi}_{t:k+1} = \mathbf{A}_t\mathbf{A}_{t-1}\cdots\mathbf{A}_{k+1}, \qquad \boldsymbol{\Phi}_{t:t+1} = \mathbf{I}. \tag{83}$$

Since the accumulated perturbation remains Gaussian, we use the compact reparameterization

$$\mathbf{x}_t = \boldsymbol{\Phi}_t\mathbf{x}_0 + \bar{\boldsymbol{\sigma}}_t\boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{84}$$

where $\bar{\boldsymbol{\sigma}}_t$ denotes the effective accumulated Gaussian noise weight induced by the predefined Gaussian-noise schedule and frequency-domain degradation schedule. Its induced covariance is denoted by

$$\boldsymbol{\Sigma}_t = \bar{\boldsymbol{\sigma}}_t \bar{\boldsymbol{\sigma}}_t^{\mathsf{T}}. \tag{85}$$

Accordingly,

$$q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}\left(\boldsymbol{\Phi}_t \mathbf{x}_0, \boldsymbol{\Sigma}_t\right). \tag{86}$$

We next examine the contribution of the original signal. Because the Gaussian convolution kernel is normalized, each row of $\mathbf{M}_t$ forms a weighted aggregation over neighboring frequency responses:

$$[\mathbf{M}_t]_{ij} \geq 0, \qquad \sum_j [\mathbf{M}_t]_{ij} = 1. \tag{87}$$

Thus, for any signal vector $\mathbf{z}$,

$$|[\mathbf{M}_t \mathbf{z}]_i| \leq \sum_j [\mathbf{M}_t]_{ij} |z_j| \leq \|\mathbf{z}\|_\infty, \tag{88}$$

which gives

$$\|\mathbf{M}_t \mathbf{z}\|_\infty \leq \|\mathbf{z}\|_\infty. \tag{89}$$

Therefore,

$$\|\boldsymbol{\Phi}_t \mathbf{x}_0\|_\infty \leq \left(\prod_{k=1}^{t} \sqrt{\alpha_k}\right) \|\mathbf{x}_0\|_\infty. \tag{90}$$

For a diffusion schedule satisfying

$$\prod_{k=1}^{t} \alpha_k \to 0, \tag{91}$$

the contribution of the original wireless response progressively vanishes:

$$\boldsymbol{\Phi}_t \mathbf{x}_0 \to \mathbf{0}. \tag{92}$$

Consequently, the terminal state is parameterized by the accumulated Gaussian noise weight:

$$\mathbf{x}_T \approx \bar{\boldsymbol{\sigma}}_T \boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{93}$$

Equivalently,

$$q(\mathbf{x}_T|\mathbf{x}_0) \approx \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_T\right), \qquad \boldsymbol{\Sigma}_T = \bar{\boldsymbol{\sigma}}_T \bar{\boldsymbol{\sigma}}_T^{\mathsf{T}}. \tag{94}$$

The terminal distribution is therefore determined by the same predefined forward degradation and Gaussian-noise schedules used throughout training.

A.3.3 GAUSSIAN INITIALIZATION FOR REVERSE SAMPLING

The reverse process starts from the same terminal Gaussian parameterization derived above. We first sample

$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{95}$$

and construct

$$\mathbf{x}_T = \bar{\boldsymbol{\sigma}}_T \boldsymbol{\epsilon}. \tag{96}$$

Hence,

$$\mathbf{x}_T \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_T\right), \tag{97}$$

which is the same terminal distribution as Eq. (94). The reverse trajectory then progressively removes the Gaussian corruption while recovering the propagation structures degraded during the forward process.

A.4 FORWARD PROPAGATION-STATISTICS DIFFUSION

Frequency-domain Gaussian blur construction. The propagation degradation in Helm-Diffusion is implemented as a frequency-domain Gaussian blur operator. Different from the stochastic Gaussian corruption, this degradation explicitly removes fine-grained propagation structures along the frequency dimension before the reverse recovery process.

At diffusion step t, the Gaussian blur kernel is defined as

$$G_t(v) = \frac{\exp\left(-\frac{v^2}{2\sigma_{f,t}^2}\right)}{Z_t}, \tag{98}$$

where v denotes the frequency offset and Z_t is the normalization factor. The blur strength increases along the diffusion trajectory

$$\sigma_{f,t} = \rho_t \sigma_f, \qquad \rho_t = \sqrt{\frac{t}{T-1}}, \qquad t = 0, \ldots, T-1. \tag{99}$$

The propagation degradation operator is therefore defined as

$$[M_t(x)]_{m,i} = \sum_v G_t(v) x_{m,i+v}, \tag{100}$$

where m and i denote the temporal and frequency indices, respectively. Since the Gaussian kernel is normalized and independent of the measured response, M_t is a deterministic linear operator shared by all wireless responses.

The normalized Gaussian blur progressively suppresses local frequency variations and weakens the recoverability of propagation structures. Since the degradation is applied along the frequency dimension, it directly reduces the cross-frequency propagation consistency while maintaining the Gaussian transition formulation of diffusion models.

A.5 REVERSE PHYSICS-STATISTICS DIFFUSION

The forward process gradually removes propagation structures and statistical characteristics through the propagation-aware degradation operator and Gaussian noise injection. In the reverse process, Helm-Diffusion recovers the clean wireless response through iterative denoising and propagation-consistent projection.

Clean response prediction. At reverse diffusion step t, given the noisy state x_t, the denoising network predicts the clean response

$$\widetilde{x}_0 = D_\theta(x_t, t, c), \tag{101}$$

where c denotes the generation condition. The predicted response restores the statistical characteristics learned from the data distribution.

Propagation-consistent projection. The Physics Guided Module extracts propagation-related information from the predicted clean response

$$C_{\mathrm{sta}} = C_{\mathrm{sta}}(\widetilde{x}_0), \qquad C_{\mathrm{dyn}} = C_{\mathrm{dyn}}(\widetilde{x}_0). \tag{102}$$

The propagation information is incorporated through the physics-guided projection module

$$\widehat{x}_0 = P_\phi(\widetilde{x}_0, C_{\mathrm{sta}}, C_{\mathrm{dyn}}, t, c). \tag{103}$$

The projected clean response is then used for the reverse diffusion transition.

Reverse posterior sampling. The propagation-aware forward process in Appendix A.3 forms a linear Gaussian transition

$$x_t = A_t x_{t-1} + \eta_t, \qquad A_t = \sqrt{\alpha_t} M_t, \qquad \eta_t \sim \mathcal{N}(0, Q_t), \tag{104}$$

where

$$Q_t = (1 - \alpha_t) I. \tag{105}$$

Since the forward process remains Gaussian, the reverse transition also follows a Gaussian posterior. Given the projected clean response $\widehat{x}_0$, the reverse posterior is formulated as

$$q(x_{t-1} | x_t, \widehat{x}_0) = \mathcal{N}(\mu_{t-1}, \Sigma_{t-1}^{\mathrm{post}}). \tag{106}$$

The posterior mean is computed as:

$$\mu_{t-1} = \Phi_{t-1} \widehat{x}_0 + K_t \left(x_t - A_t \Phi_{t-1} \widehat{x}_0 \right), \tag{107}$$

where

$$K_t = \Sigma_{t-1} A_t^T \left(A_t \Sigma_{t-1} A_t^T + Q_t \right)^{-1}. \tag{108}$$

The posterior covariance is given by

$$\Sigma_{t-1}^{\mathrm{post}} = \Sigma_{t-1} - K_t A_t \Sigma_{t-1}, \tag{109}$$

where Φ_{t-1} and Σ_{t-1} denote the accumulated transition operator and Gaussian covariance defined in Appendix A.3.2.

The reverse state is sampled from this posterior

$$x_{t-1} \sim q(x_{t-1} | x_t, \widehat{x}_0). \tag{110}$$

Different from conventional diffusion models, the transition matrix A_t contains the frequency-domain Gaussian blur operator M_t. Therefore, the reverse posterior explicitly considers the structured degradation introduced during the forward trajectory while recovering the statistical characteristics of wireless responses.

Table 3: Dataset and training configurations of Helm-Diffusion.

Dataset	Total	Train	Val.	Test	Updates	Blocks
WiDAR	18,400	14,720	1,840	1,840	2,400	32
FallDar	12,600	10,080	1,260	1,260	2,600	32
GaitSense	15,700	12,560	1,570	1,570	2,800	32
ESPARGOS	19,200	15,360	1,920	1,920	2,500	16
DICHASUS	17,300	13,840	1,730	1,730	2,700	32
Argos	11,800	9,440	1,180	1,180	2,200	32
COST2100	16,500	13,200	1,650	1,650	2,300	32

Conditioning strategy. For human activity recognition and localization tasks, the generation process is conditioned on the corresponding activity category or spatial condition. For FDD channel generation, the model is conditioned on the uplink channel response, while the downlink channel response is used only as the target during training and evaluation.

A.6 Training and Implementation Details

Datasets and data splits. We evaluate Helm-Diffusion on three HAR datasets (WiDAR, FallDar, and GaitSense), two localization datasets (ESPARGOS and DICHASUS), and two FDD channel datasets (Argos and COST2100). A separate model is trained for each dataset.

For each dataset, the original recordings are first partitioned into fixed training, validation, and test subsets with an 8:1:1 ratio. Window extraction and dataset-specific preprocessing are then performed independently within each subset. The same partitions are used for all compared methods and all random seeds.

Input representation and preprocessing. Wireless responses are represented in complex form and stored as

$$x \in \mathbb{R}^{B\times N\times F\times 2}, \tag{111}$$

where the last dimension contains the real and imaginary components. The two components are normalized using statistics computed from the training partition and the same normalization is subsequently applied to the validation and test partitions. The model preserves the complex-valued representation throughout the network using complex linear layers, complex nonlinearities, complex normalization, and complex attention.

Task conditions. The conditional variable c is defined according to the downstream task. For HAR, c is a six-dimensional categorical condition describing the activity-related state and is directly embedded by the condition network. For localization,

$$c = (x, y, z), \tag{112}$$

where (x, y, z) is the mean three-dimensional receiver position associated with the current measurement window. The coordinates are normalized using training-set statistics before being provided to the model.

For FDD channel generation, the uplink channel is used as the condition,

$$c = H_{\mathrm{UL}}, \qquad x_0 = H_{\mathrm{DL}}, \tag{113}$$

such that the model generates the corresponding downlink response conditioned on the observed uplink channel. The signal-to-noise ratio (SNR) for FDD channel generation is calculated as

$$\mathrm{SNR} = -10\log_{10}\left(\frac{\|x_{\mathrm{down}} - x_{\mathrm{est}}\|^2}{\|x_{\mathrm{down}}\|^2}\right), \tag{114}$$

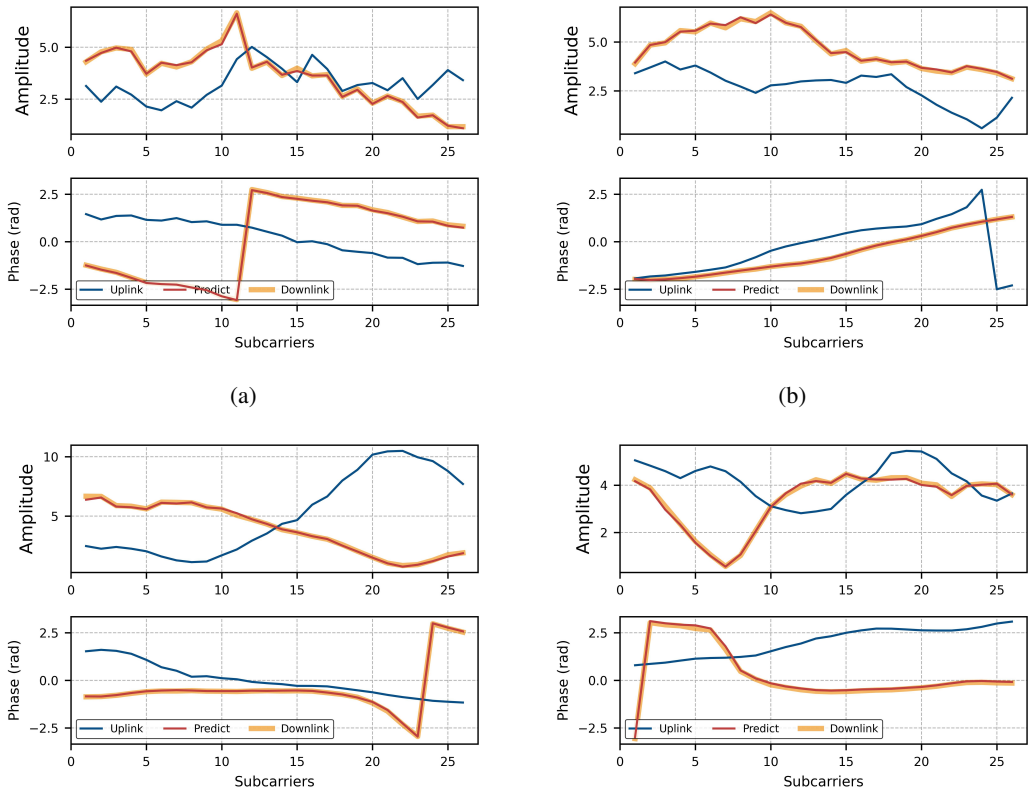


Figure 6: Visualization of generated FDD channel responses. The results demonstrate the fidelity of generated downlink channels under different propagation conditions.

and the results are shown in Fig. 6.

All conditions are mapped through a complex-valued condition embedding network with the structure

$$d_c \to d_{\text{hid}} \to 4d_{\text{hid}} \to d_{\text{hid}}, \tag{115}$$

and are combined with the diffusion-step embedding for adaptive normalization.

Network architecture. The input response is first mapped to the hidden representation by a ComplexLinear layer. Temporal positions are encoded using a complex exponential positional representation and multiplied element-wise with the latent response.

The diffusion step is represented by a 256-dimensional embedding and mapped to the hidden dimension through

$$256 \to d_{\text{hid}} \to d_{\text{hid}}. \tag{116}$$

Each DiA block contains complex LayerNorm, adaptive normalization, eight-head complex self-attention, and a complex MLP with an expansion ratio of 4. The hidden dimension is 128 for HAR and FDD models and 256 for localization models. ESPARGOS uses 16 DiA blocks, while the remaining datasets use 32 blocks, as summarized in Table 3.

The diffusion process contains

$$T = 300 \tag{117}$$

steps. We use a linear Gaussian noise schedule

$$\beta_t = \beta_{\min} + \frac{t}{T-1}\left(\beta_{\max} - \beta_{\min}\right), \qquad t = 0, \ldots, T-1, \tag{118}$$

where

$$\beta_{\min} = 10^{-4}, \qquad \beta_{\max} = 6.5 \times 10^{-2}. \tag{119}$$

The noise schedule is selected such that the residual contribution of the original signal at the terminal diffusion step is negligible. Specifically, defining

$$\bar{\alpha}_T = \prod_{t=0}^{T-1} (1 - \beta_t), \tag{120}$$

the above schedule with $T = 300$ gives

$$\bar{\alpha}_T \approx 4.6 \times 10^{-5}, \qquad \sqrt{\bar{\alpha}_T} \approx 6.8 \times 10^{-3}. \tag{121}$$

Therefore, the residual contribution of the original signal in the terminal state is bounded to below 0.7% in amplitude, which makes the terminal distribution well approximated by the Gaussian initialization used in the reverse process.

We further define

$$\alpha_t = 1 - \beta_t, \qquad \bar{\alpha}_t = \prod_{k=0}^{t} \alpha_k. \tag{122}$$

The Gaussian schedule controls the stochastic corruption strength during the forward process, while the propagation degradation operator $\mathcal{M}_t(\cdot)$ progressively removes propagation consistency. Together, stochastic corruption and propagation degradation suppress the original signal contribution along the diffusion trajectory. The reverse generation process uses the same 300 diffusion steps and starts from a Gaussian noise initialization.

Optimization and training schedule. All models are trained using the AdamW optimizer with an initial learning rate of

$$\eta = 1 \times 10^{-4},$$

Adam coefficients $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and a weight decay of

$$\lambda_{\text{wd}} = 1 \times 10^{-4}.$$

We use a mini-batch size of

$$B = 24$$

for all datasets. The learning rate is linearly warmed up during the first 5% of the training updates and is subsequently decayed using a cosine schedule.

Equivalent Gaussian noise scale. The instantaneous Gaussian noise scale at diffusion step t is determined by the predefined noise schedule

$$\sigma_t = \sqrt{1 - \alpha_t} = \sqrt{\beta_t}, \qquad t = 0, \ldots, T-1. \tag{123}$$

Using the same notation as Appendix A.3.2, let

$$\mathbf{A}_t = \sqrt{\alpha_t}\mathbf{M}_t, \qquad \mathbf{\Phi}_{t:k+1} = \mathbf{A}_t \mathbf{A}_{t-1} \cdots \mathbf{A}_{k+1}, \tag{124}$$

with $\mathbf{\Phi}_{t:t+1} = \mathbf{I}$. The effective accumulated Gaussian noise weight $\bar{\boldsymbol{\sigma}}_t$ is the same schedule-derived quantity used in Eq. (84). In implementation, it is precomputed from the fixed Gaussian-noise schedule and the cumulative frequency-domain degradation, and is not treated as an additional tunable hyperparameter.

Equivalently, its induced covariance follows the forward recursion

$$\mathbf{\Sigma}_t = \mathbf{A}_t \mathbf{\Sigma}_{t-1} \mathbf{A}_t^{\mathsf{T}} + (1-\alpha_t)\mathbf{I}, \qquad \mathbf{\Sigma}_0 = (1-\alpha_0)\mathbf{I}, \tag{125}$$

with

$$\mathbf{\Sigma}_t = \bar{\boldsymbol{\sigma}}_t \bar{\boldsymbol{\sigma}}_t^{\mathsf{T}}. \tag{126}$$

For consistency with the terminal-state notation used throughout the paper, the final implemented diffusion step is denoted by

$$\bar{\boldsymbol{\sigma}}_T \equiv \bar{\boldsymbol{\sigma}}_{T-1}. \tag{127}$$

The terminal state used for reverse sampling is therefore

$$\mathbf{x}_T = \bar{\boldsymbol{\sigma}}_T \boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{128}$$

Thus, the terminal Gaussian scale is fully determined by the same forward diffusion configuration used during training, and no additional terminal variance is manually specified.

Propagation-degradation settings. The frequency-domain Gaussian blur kernel is predefined according to the frequency sampling interval and is shared across all signal samples.

The propagation-degradation schedule is defined as

$$\rho_t = \sqrt{\frac{t}{T-1}}, \qquad t = 0, \ldots, T-1, \tag{129}$$

such that $\rho_0 = 0$ and $\rho_{T-1} = 1$.

Propagation-degradation settings. The base standard deviation of the frequency-domain Gaussian blur is set to

$$\sigma_f = 2.0$$

subcarrier bins for all experiments. At diffusion step t, the blur strength follows

$$\sigma_{f,t} = \rho_t \sigma_f, \qquad \rho_t = \sqrt{\frac{t}{T-1}},$$

where $T = 300$. Therefore, the frequency-domain degradation gradually increases from $\sigma_{f,0} = 0$ to $\sigma_{f,T-1} = 2.0$ subcarrier bins.

For numerical implementation, the Gaussian kernel is truncated to a radius of $3\sigma_{f,t}$ and normalized to unit sum before convolution along the frequency dimension. The same degradation schedule and base blur scale are used for all datasets.

Sampling protocol. All reported generation results are obtained through random generation rather than reconstruction from test responses. The reverse process starts from the same terminal Gaussian parameterization used in Appendix A.3.3

$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \qquad \mathbf{x}_T = \bar{\boldsymbol{\sigma}}_T \boldsymbol{\epsilon}. \tag{130}$$

The model then performs $T = 300$ reverse diffusion steps. At reverse step s, the denoising network predicts the clean wireless response

$$\widetilde{x}_0 = D_\theta(x_s, s, c). \tag{131}$$

The prediction is further refined by the Physics Guided Module

$$\widehat{x}_0 = P_\phi(\widetilde{x}_0, C_{\mathrm{sta}}, C_{\mathrm{dyn}}, s, c). \tag{132}$$

The previous diffusion state is sampled according to the reverse transition

$$x_{s-1} \sim q(x_{s-1}|x_s, \widehat{x}_0). \tag{133}$$

Therefore, the implemented sampler follows a predict–project–sample procedure. The initialization contains only Gaussian noise and does not use any measured response. For FDD generation, the uplink channel is provided only as the condition, while the downlink response is never used for initialization or reconstruction.

WD and MMD evaluation. For distribution-level evaluation, we adopt the 2-Wasserstein distance (WD), following prior wireless channel generation studies (Wagle et al., 2025), together with Maximum Mean Discrepancy (MMD). Both metrics are computed between the complete generated sample set and the corresponding real test set using the same evaluation implementation for all compared methods.

For MMD, we use the Gaussian RBF kernel

$$k(x, y) = \exp\left(-\frac{\|x - y\|_2^2}{2\sigma_k^2}\right), \tag{134}$$

where the kernel bandwidth is determined using the median heuristic on the pooled real and generated samples. Let $X = \{x_i\}_{i=1}^N$ and $Y = \{y_j\}_{j=1}^M$ denote the two sample sets, and let $Z = (z_1, \ldots, z_{N+M})$ denote their concatenation. We compute all pairwise squared Euclidean distances and set

$$\sigma_k^2 = \text{median}\left\{\|z_a - z_b\|_2^2 \;\middle|\; 1 \le a < b \le N + M\right\} + 10^{-12}, \tag{135}$$

where the additive constant prevents a zero bandwidth.

The squared empirical MMD is computed as

$$\begin{aligned}\text{MMD}^2(X, Y) = &\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} k(x_i, x_j) + \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} k(y_i, y_j) \\ &- \frac{2}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} k(x_i, y_j).\end{aligned} \tag{136}$$

Both within-set sums include the diagonal terms, corresponding to averaging all entries of the respective kernel matrices.

For brevity, we refer to the squared MMD statistic in Eq. 136 as MMD throughout the main text, tables, and figures. All reported MMD values are computed without taking the square root.

Independent runs and statistical reporting. Each experiment is independently trained and evaluated using five random seeds,

$$\{0, 1, 2, 3, 4\}. \tag{137}$$

Each seed determines an independent model initialization, training run, and set of sampling-noise realizations. Metrics are first calculated separately for each complete run. For metric m_r from the r-th run, we report

$$\bar{m} = \frac{1}{5}\sum_{r=1}^{5} m_r, \tag{138}$$

together with the sample standard deviation

$$s_m = \sqrt{\frac{1}{4}\sum_{r=1}^{5}(m_r - \bar{m})^2}. \tag{139}$$

Accordingly, quantitative results are reported as $\text{mean} \pm \text{std}$ over the five independent runs. The detailed dataset-wise results for HAR, localization, and FDD channel generation are summarized in Tables 4, 5, and 6, respectively.

Table 4: Generation performance on HAR datasets over five independent runs. Results are reported as mean ± standard deviation. Lower WD and MMD indicate better performance.

Model	Widar		Falldar		GaitSense	
	WD↓	MMD↓	WD↓	MMD↓	WD↓	MMD↓
Helm-Diffusion	$\mathbf{0.2257 \pm 0.0108}$	$\mathbf{0.0594 \pm 0.0031}$	$\mathbf{0.2316 \pm 0.0114}$	$\mathbf{0.0548 \pm 0.0030}$	$\mathbf{0.2474 \pm 0.0122}$	$\mathbf{0.0517 \pm 0.0028}$
RF-Diffusion	0.3670 ± 0.0147	0.0794 ± 0.0045	0.3518 ± 0.0139	0.0756 ± 0.0042	0.3421 ± 0.0151	0.0728 ± 0.0040
CSVAE	0.4387 ± 0.0179	0.0879 ± 0.0051	0.4215 ± 0.0164	0.0834 ± 0.0048	0.4098 ± 0.0182	0.0807 ± 0.0046
DDPM	0.5126 ± 0.0213	0.1018 ± 0.0062	0.4964 ± 0.0207	0.0967 ± 0.0058	0.4819 ± 0.0198	0.0925 ± 0.0055
DCGAN	0.6438 ± 0.0286	0.1426 ± 0.0081	0.6215 ± 0.0269	0.1378 ± 0.0078	0.6087 ± 0.0277	0.1342 ± 0.0075
CVAE	0.7215 ± 0.0324	0.1863 ± 0.0096	0.7048 ± 0.0311	0.1812 ± 0.0092	0.6895 ± 0.0306	0.1764 ± 0.0088

Table 5: Generation performance on localization datasets over five independent runs. Results are reported as mean ± standard deviation. Lower WD and MMD indicate better performance.

Model	ESPARGOS		DICHASUS	
	WD↓	MMD↓	WD↓	MMD↓
Helm-Diffusion	$\mathbf{0.2174 \pm 0.0097}$	$\mathbf{0.0506 \pm 0.0026}$	$\mathbf{0.2295 \pm 0.0105}$	$\mathbf{0.0529 \pm 0.0032}$
RF-Diffusion	0.3427 ± 0.0128	0.0719 ± 0.0038	0.3564 ± 0.0142	0.0747 ± 0.0043
CSVAE	0.3976 ± 0.0158	0.0788 ± 0.0043	0.4142 ± 0.0171	0.0815 ± 0.0049
DDPM	0.4938 ± 0.0204	0.0951 ± 0.0057	0.5072 ± 0.0210	0.0984 ± 0.0060
DCGAN	0.6194 ± 0.0258	0.1385 ± 0.0079	0.6358 ± 0.0281	0.1417 ± 0.0082
CVAE	0.7126 ± 0.0320	0.1835 ± 0.0094	0.7251 ± 0.0332	0.1872 ± 0.0097

Training objective. Given a clean response x_0, diffusion step t, task condition c, and the corresponding degraded state x_t, the denoising backbone first predicts a clean response

$$\widetilde{x}_0 = D_\theta(x_t, t, c). \tag{140}$$

The static and dynamic propagation cues described in Appendix A.2 are then extracted from $\widetilde{x}_0$. The Physics Guided Module projects the denoised prediction toward a propagation-consistent response:

$$\widehat{x}_0 = P_\phi\left(\widetilde{x}_0, C_{\mathrm{sta}}(\widetilde{x}_0), C_{\mathrm{dyn}}(\widetilde{x}_0), t, c\right), \tag{141}$$

where $P_\phi(\cdot)$ denotes the physics guided module. The response $\widehat{x}_0$ is used for subsequent reconstruction and propagation-consistency supervision.

Reconstruction loss. The reconstruction objective directly supervises the final corrected response

$$\mathcal{L}_{\mathrm{rec}} = \mathbb{E}_{x_0,c,t,\epsilon}\left[\|\widehat{x}_0 - x_0\|_2^2\right]. \tag{142}$$

Static consistency loss. For a response x, let $e^{\mathrm{sta}}_{m,i}(x)$ denote the normalized static consistency score obtained using the adjacent-frequency leave-one-out Loewner procedure in Appendix A.2. We compute the score for both the generated response and the corresponding clean response

$$\mathcal{L}_{\mathrm{sta}} = \mathbb{E}_{x_0,c,t,\epsilon}\left[\frac{1}{MF}\sum_{m=1}^{M}\sum_{i=1}^{F}\left|e^{\mathrm{sta}}_{m,i}(\widehat{x}_0) - e^{\mathrm{sta}}_{m,i}(x_0)\right|^2\right]. \tag{143}$$

This objective matches the finite-observation cross-frequency consistency of the generated response to that of the measured response rather than forcing the Loewner approximation residual to vanish.

Dynamic consistency loss. For response x, the multi-scale dynamic consistency score is

$$e^{\mathrm{dyn}}_{m,i}(x) = \frac{1}{R}\sum_{r=1}^{R}\left\|P^{(r)}_{m,i}(x) - \bar{P}^{(r)}_{m}(x)\right\|_2^2. \tag{144}$$

The corresponding dynamic consistency objective is

$$\mathcal{L}_{\mathrm{dyn}} = \mathbb{E}_{x_0,c,t,\epsilon}\left[\frac{1}{MF}\sum_{m=1}^{M}\sum_{i=1}^{F}\left|e^{\mathrm{dyn}}_{m,i}(\widehat{x}_0) - e^{\mathrm{dyn}}_{m,i}(x_0)\right|^2\right]. \tag{145}$$

Table 6: Generation performance on FDD channel datasets over five independent runs. Results are reported as mean ± standard deviation. Higher SNR indicates better performance.

Model	Argos SNR↑	COST2100 SNR↑
Helm-Diffusion	$\mathbf{28.81 \pm 0.42}$	$\mathbf{27.94 \pm 0.47}$
RF-Diffusion	26.97 ± 0.51	26.21 ± 0.55
CSVAE	20.35 ± 0.63	19.82 ± 0.68
DDPM	22.86 ± 0.71	22.31 ± 0.76
DCGAN	18.74 ± 0.84	18.25 ± 0.88
CVAE	15.62 ± 0.91	15.18 ± 0.95

Combined objective. The complete training objective is

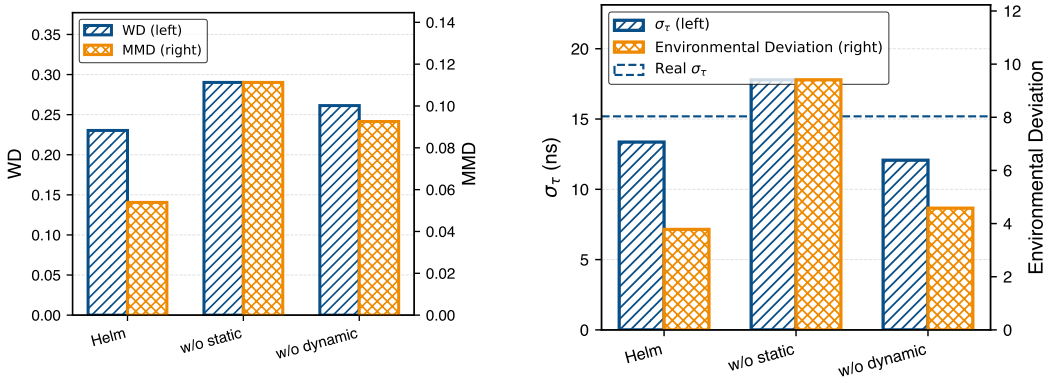
$$\boxed{\mathcal{L}_{\text{train}} = \mathcal{L}_{\text{rec}} + 0.5\,\mathcal{L}_{\text{sta}} + 0.5\,\mathcal{L}_{\text{dyn}}} \tag{146}$$

or equivalently,

$$\lambda_{\text{sta}} = \lambda_{\text{dyn}} = 0.5. \tag{147}$$

The two consistency objectives therefore provide equally weighted supervision for the static cross-frequency propagation structure and the local dynamic propagation-evolution structure, while reconstruction remains the primary signal-domain objective.

A.7 Ablation of Static and Dynamic Propagation Designs



(a) Generation quality.

(b) Propagation-scale consistency.

Figure 7: Ablation of the static and dynamic physics designs.

We examine the contributions of the static and dynamic consistency branches by removing each branch from the generation process. As shown in Fig. 7a, Helm-Diffusion achieves a WD of approximately 0.23 and an MMD of 0.05. Removing the static branch increases these values to approximately 0.29 and 0.12, while removing the dynamic branch increases them to approximately 0.26 and 0.09. Both branches improve distribution matching, with the static branch having a larger effect on these metrics.

Fig. 7b compares cross-band ToF dispersion and Environmental Deviation. The real-data ToF dispersion is approximately 15 ns. Helm-Diffusion yields approximately 13 ns, compared with 17.5 ns without the static branch and 12 ns without the dynamic branch. Its Environmental Deviation is also lower: approximately 3.8, versus 9.4 and 4.5, respectively. Although removing the dynamic branch produces a smaller ToF dispersion, that value is farther from the real-data reference. Together, these results show that both branches improve generation quality and help preserve realistic cross-band propagation characteristics.