MOVING RECEIVER SIGNAL PREDICTION VIA WAVEFIELD–TRAJECTORY DECOUPLING

Anonymous authors
Paper under double-blind review

ABSTRACT

Predicting future electromagnetic signals at moving receivers is an important task in mobile wireless systems, supporting beam tracking, Doppler compensation, and resource allocation. The received signal varies both as the wavefield evolves over time and as receiver motion changes the sampled field amplitude and phase. In particular, motion-induced Doppler shifts can produce rapid phase variations, making accurate prediction challenging. Existing methods tend to jointly learn temporal wavefield evolution and spatial trajectory sampling to predict future signals. Encoding these two factors with different properties into a shared representation can cause different receiver velocities to induce conflicting gradient directions during optimization, thereby impairing the learned wavefield representation. To focus the network on wavefield evolution while avoiding neural approximation of trajectory sampling, we propose **Wavefield–Trajectory Decoupling (WTD)**, comprising a velocity-agnostic complex neural operator and a parameter-free geometric projector. The operator predicts Eulerian wavefields from past wavefields, and the projector samples the predicted field along known trajectories to obtain future received signals. A phase-preserving complex activation and full-field Sobolev supervision constrain high-frequency wavefield learning. WTD outperforms baselines on in-distribution and out-of-distribution velocity, with up to $13.9\times$ lower relative L_2 signal error and $11.5\times$ lower absolute Doppler error.

1 INTRODUCTION

In high-mobility scenarios of high-speed trains (Zhao et al., 2026), unmanned aerial vehicles (Zeng et al., 2019), and low Earth orbit satellites (You et al., 2020), transmitters and receivers undergo significant relative motion. Delays in signal measurement, processing, and feedback cause system parameter adjustments to lag behind actual signal variations (Xiao et al., 2024; Jin et al., 2025). Accurately predicting future electromagnetic signals at mobile receivers can therefore provide a basis for beam adjustment (Li et al., 2023), Doppler compensation (Yeh et al., 2024), and resource allocation (Minani et al., 2025). For stationary receiver, the signal varies only with the temporal evolution of the wavefield at its location. In contrast, a mobile receiver is additionally affected by relative motion, which induces Doppler shifts (Yin et al., 2020) in the signal and causes signal strength to vary with position (Zhu et al., 2024). These spatial variations and rapid phase variations combine to produce more complex signal variations, making future received signals harder to predict.

A mobile receiver samples the time-evolving electromagnetic wavefield $E(\mathbf{x}, t)$ along its trajectory $\mathbf{r}(t_n) = \mathbf{r}_0 + \mathbf{v}t_n$. The sampled field values at successive positions are $s_{\mathbf{v}}[n] = E(\mathbf{r}(t_n), t_n)$. Arranged in time order, they form the one-dimensional Lagrangian signal $\mathbf{s}_{\mathbf{v}} = [s_{\mathbf{v}}[0], s_{\mathbf{v}}[1], \ldots, s_{\mathbf{v}}[T-1]]$. Variations $\Delta s_{\mathbf{v}}[n] = s_{\mathbf{v}}[n] - s_{\mathbf{v}}[n-1]$ in this signal are jointly determined by two processes: wavefield evolution determines future field values at each location, while trajectory sampling determines where to sample field values at each time. So, accurate signal prediction requires both accurately capturing wavefield evolution and adapting to changes in sampling positions associated with different receiver velocities.

Wavefield evolution and trajectory sampling induced by receiver motion are both challenging. High-frequency wavefields contain rapid oscillations and are difficult to learn accurately (Raissi et al., 2019; Zhang et al., 2026). Meanwhile, different receiver velocities correspond to different sampling positions, thereby producing different received signals. To jointly model these two types of variation,

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.

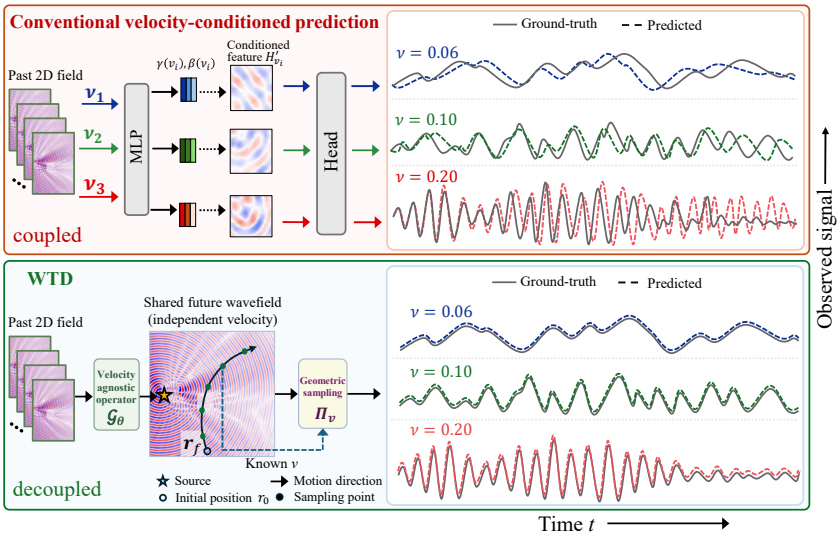


Figure 1: **Wavefield–Trajectory Decoupling. Top:** Velocity-conditioned prediction couples wavefield evolution and trajectory sampling in a learned mapping to produce 1D Lagrangian signals. **Bottom:** WTD predicts a shared velocity-independent wavefield and samples it along known receiver trajectories. The depicted trajectory is illustrative.

existing methods typically introduce velocity as a conditioning variable and combine it with past wavefields to predict future signals, using either feature concatenation (Hagnberger et al., 2024) or modulation (Perez et al., 2018; Wang et al., 2024).

However, receiver velocity changes only the sampling positions, not wavefield evolution. The above conditioned predictors use a single mapping to fit signal variations jointly determined by wavefield evolution and trajectory sampling, bringing two problems: First, within the training range, conditioned predictors jointly encode velocity-independent wavefield evolution and velocity-dependent trajectory sampling into a shared representation (Hagnberger et al., 2024; Wang et al., 2024). Although receiver velocity is known, it is encoded as a network input rather than used to determine the receiver trajectory for explicit field sampling. The network thus approximates how velocity changes the sampled field values, which can induce conflicting gradient directions during optimization and impair the learned wavefield representation, introducing in-range signal error, see Appendix H.3. Second, different frequency–velocity pairs can produce the same 1D received signal because temporal frequency and motion are coupled in the received frequency, as formalized in §3.2. This illustrates that wavefield variation and motion-induced sampling are entangled at the signal level. Conditioned predictors learn velocity-dependent mappings and must infer how changes in velocity affect the sampled field values, resulting in additional wavefield-prediction error. (Cheng et al., 2026; Meliha et al., 2025). Therefore, how to decouple wavefield evolution and trajectory sampling to achieve accurate signal prediction across receiver velocities, is the problem that needs to be addressed.

To meet this need, we propose **Wavefield–Trajectory Decoupling (WTD)** (Fig. 1), which explicitly separates wavefield evolution from trajectory sampling. WTD first predicts future wavefields without conditioning on receiver velocity and then uses the velocity only to determine the receiver trajectory for deterministic field sampling. In this way, signal differences across velocities arise solely from different sampling positions rather than being learned implicitly by a conditioned predictor. This design is therefore more robust to velocity changes beyond the training range. To further improve local field accuracy along the sampling trajectory, we employ a complex-valued neural operator with a phase-preserving activation and full-field Sobolev supervision.

We summarize our main contributions below:

Wavefield–trajectory decoupling. We first reformulate moving-receiver signal prediction from a velocity-conditioned joint mapping into a decoupled problem of shared wavefield learning and ex-

plicit sampling along known trajectories. This prevents velocity-dependent trajectory sampling from being encoded into shared representation, avoiding conflicting gradient directions during optimization and improving in-distribution(ID) and out-of-distribution(OOD) prediction accuracy.

Identifiability and mechanism. We find that, under a local plane-wave model, different frequency–velocity pairs can produce the same 1D received signal, revealing a limitation of uniquely separating wavefield variation from unknown receiver motion using 1D signal alone. We further propose that receiver motion alters trajectory sampling rather than wavefield evolution, and that motion-induced Doppler shifts arise from phase sampling along receiver trajectory. These findings motivate us to explicitly leverage known motion, learn a shared wavefield, and compute trajectory-based observations.

Empirical evaluation. Compared to baselines with matched wavefield inputs and known velocity, WTD achieves up to $13.9\times$ lower relative L_2 signal error and $11.5\times$ lower absolute Doppler error. The gains persist in a three-seed cross-solver evaluation using MEEP (Oskooi et al., 2010).

2 Related Work

Received Signal Prediction. Predicting subsequent time-domain signals from historical data can reduce the cost of electromagnetic simulation. Autoregressive models extrapolate FDTD sequences to obtain frequency responses efficiently (Jandhyala et al., 1994). Neural approaches use finite impulse response networks (Wu et al., 1999) and multilayer perceptrons (Göksu & Wunsch, 2003) as nonlinear predictors of subsequent FDTD samples. Dynamic mode decomposition enables temporal extrapolation of cavity responses (Nayak et al., 2024), while LSTM networks forecast electromagnetic compatibility test signals (Li et al., 2025). However, these methods predict signals at fixed locations or on simulation grids. WTD separates wavefield evolution from trajectory sampling to improve prediction along known receiver trajectories within and beyond the training velocity range.

Wavefield Prediction and High-Frequency Learning. Neural operators approximate PDE solution operators from data, with architectures including FNO, DeepONet, and CNO (Li et al., 2021; Lu et al., 2021; Raonic et al., 2023). For electromagnetic simulation, physics-informed methods use governing equations and boundary conditions to constrain field prediction (Raissi et al., 2019), while physics-inspired designs incorporate causality and material properties into time-domain wavefield prediction (Ma et al., 2025). Despite these advances, standard PINNs exhibit spectral bias and struggle to capture rapid high-frequency oscillations. Zhang et al. (Zhang et al., 2026) analyzed this limitation in wavefield reconstruction and alleviated it through a physics-guided envelope transformation. High-frequency wavefield prediction therefore requires accurate representation of amplitude, phase, and spatial variation. Complex-valued networks model amplitude and phase in speech enhancement and MRI reconstruction (Choi et al., 2019; Cole et al., 2021). Applying nonlinearities separately to real and imaginary parts can alter phase, whereas modReLU preserves phase for nonzero outputs (Arjovsky et al., 2016). Sobolev training supervises derivatives alongside function values to constrain local variation (Czarnecki et al., 2017). WTD combines a phase-preserving activation and full-field Sobolev supervision to support accurate high-frequency wavefield prediction.

Conditioning and Generalization. Conditioning incorporates auxiliary variables through feature concatenation (Hagnberger et al., 2024) or feature-wise modulation, including FiLM and adaptive layer normalization (Perez et al., 2018; Peebles & Xie, 2023). Symmetry-based methods exploit known transformations through data augmentation (Brandstetter et al., 2022) or equivariant neural operators (Helwig et al., 2023). Learned canonicalization selects input-dependent transformations that map inputs into a common representation, with predictions transformed back when required by the task (Kaba et al., 2023; Ma et al., 2024). These methods provide different ways to incorporate auxiliary information and known transformations into prediction. WTD uses known receiver velocity only during geometric sampling, allowing the same predicted wavefield to serve receivers at different velocities.

3 Methodology

We first formulate the problem (§3.1), discuss local unidentifiability and the decoupling principle (§3.2), then present the WTD architecture (§3.3) and training objective (§3.4).

3.1 MOVING–RECEIVER SIGNAL PREDICTION SETTING

Consider a scalar wavefield $E(\mathbf{x}, t)$, the E_z component of a 2D Maxwell TM mode. The field propagates in an inhomogeneous medium with permittivity $\varepsilon(\mathbf{x})$, where $\mathbf{x} \in \Omega \subset \mathbb{R}^2$, and is excited at carrier angular frequency ω_c. Within each local prediction window, a receiver follows the trajectory $\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}t$ with known initial position $\mathbf{r}_0$ and velocity $\mathbf{v}$. In this moving-receiver observation setting, the receiver samples the Eulerian wavefield along its trajectory, yielding the 1D Lagrangian signal: $s_{\mathbf{v}}(t) = E(\mathbf{r}(t), t)$. Compared with the stationary-receiver signal $s_{\mathbf{0}}(t) = E(\mathbf{r}_0, t)$, receiver motion changes the sampling locations and thereby affects the signal phase evolution.

Learning task. We study an offline surrogate setting. Let $\mathbf{u}_{\mathrm{past}}$ and $\mathbf{u}_{\mathrm{future}}$ denote the corresponding T_{in} past and T_{out} future field frames on a local $P \times P$ Eulerian patch. At inference, the local 2D past field and receiver velocity are available, with the velocity assumed known from onboard navigation or platform telemetry, and the prediction target is the future 1D Lagrangian signal. The ground-truth future field $\mathbf{u}_{\mathrm{future}}$ is used only to supervise field prediction during training and is not required at inference.

Evaluation. Training is restricted to receiver speeds $\|\mathbf{v}\| \in [\underline{v}, \overline{v}]$. Evaluation includes in-distribution speeds within this range and velocity-OOD extrapolation to higher speeds satisfying $\|\mathbf{v}\| > \overline{v}$. Data and task configurations are detailed in Appendix E.

Assumptions and scope. Our setting assumes that (i) the receiver's initial position is known; (ii) the receiver velocity is approximately constant within each local prediction window; and (iii) the experiments consider receiver speeds satisfying $\|\mathbf{v}\|/c \leq 0.2$, c denotes the reference wave speed.

3.2 THE UNIDENTIFIABILITY OBSTACLE AND THE DESIGN PRINCIPLE

We first examine local ambiguity in inferring receiver motion from a 1D signal, then motivate decoupling wavefield evolution from trajectory sampling when receiver motion is known.

Received-frequency relation. A receiver following $\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}t$ samples the analytic field $U(\mathbf{x}, t) = A(\mathbf{x}, t)\mathrm{e}^{\mathrm{j}\phi(\mathbf{x},t)}$, yielding $s^{\mathrm{an}}_{\mathbf{v}}(t) = U(\mathbf{r}(t), t)$. Its real part is the signal defined in §3.1. Where the amplitude is nonzero and the phase is differentiable, the instantaneous received angular frequency follows from the chain rule:

$$\omega_{\mathrm{recv}} = \frac{d\phi(\mathbf{r}(t), t)}{dt} = \underbrace{\frac{\partial \phi}{\partial t}}_{\omega_{\mathrm{loc}}\ \text{(intrinsic)}} + \underbrace{\mathbf{v} \cdot \nabla \phi}_{\text{Doppler term}} = \omega_{\mathrm{loc}} - \mathbf{v} \cdot \mathbf{k}, \tag{1}$$

where $\omega_{\mathrm{loc}} = \partial_t \phi$ and $\mathbf{k} = -\nabla \phi$ are evaluated along the trajectory. Thus, received phase evolution combines temporal phase variation at a fixed position with spatial phase variation induced by motion. The following construction examines whether these contributions can be uniquely separated from a 1D signal alone.

The unidentifiability obstacle. Consider a nondispersive plane wave over the observation window:

$$U(\mathbf{x}, t) = A \exp\left(\mathrm{j}\left[\phi_0 + \omega_0 t - \frac{\omega_0}{c_p}\hat{\mathbf{k}} \cdot (\mathbf{x} - \mathbf{r}_0)\right]\right), \tag{2}$$

where amplitude $A > 0$, initial phase ϕ_0, unit propagation direction $\hat{\mathbf{k}}$, and phase speed $c_p > 0$ are fixed. Sampling along $\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}t$ gives

$$s^{\mathrm{an}}_{\mathbf{v}}(t) = A \exp\left(\mathrm{j}\left[\phi_0 + \omega_0\left(1 - \mathbf{v} \cdot \hat{\mathbf{k}}/c_p\right) t\right]\right). \tag{3}$$

The temporal trace therefore depends on $(\omega_0, \mathbf{v})$ only through the scalar

$$g(\omega_0, \mathbf{v}) = \omega_0\left(1 - \mathbf{v} \cdot \hat{\mathbf{k}}/c_p\right). \tag{4}$$

Proposition 1 (Local unidentifiability of wave frequency and receiver velocity)**.** *For the plane-wave model above, suppose that both ω_0 and $\mathbf{v}$ are unknown. Distinct admissible parameter pairs satisfying*

$$\omega_0'\left(1 - \mathbf{v}' \cdot \hat{\mathbf{k}}/c_p\right) = \omega_0\left(1 - \mathbf{v} \cdot \hat{\mathbf{k}}/c_p\right)$$

produce identical temporal traces. Hence, on any parameter domain containing such pairs, the observation map is non-injective: wave frequency and velocity cannot be jointly recovered uniquely.

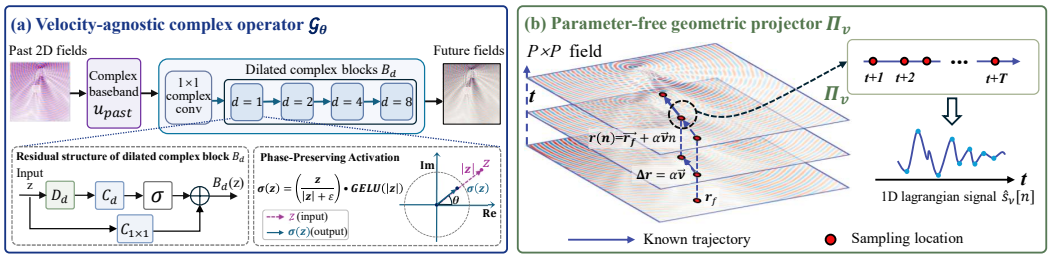


Figure 2: **Two modules of WTD.** (a) The velocity-agnostic complex operator $\mathcal{G}_\theta$ predicts future Eulerian fields from local 2D past fields with dilated blocks and phase-preserving activation. (b) The parameter-free projector $\Pi_{\mathbf{v}}$ samples them along $\mathbf{r}(n) = \mathbf{r}_{\mathrm{f}} + \alpha \mathbf{v} n$ to produce the future 1D Lagrangian signal $\widehat{s}_{\mathbf{v}}$; velocity enters only through the projector.

For example, choose $\mathbf{v}'$ with $\mathbf{v}' \cdot \hat{\mathbf{k}} \neq \mathbf{v} \cdot \hat{\mathbf{k}}$ and set

$$\omega_0' = \omega_0 \frac{1 - \mathbf{v} \cdot \hat{\mathbf{k}}/c_p}{1 - \mathbf{v}' \cdot \hat{\mathbf{k}}/c_p}.$$

This produces the same trace, provided the denominator is nonzero and ω_0' and $\mathbf{v}'$ remain within the considered range. Further details are provided in Appendix A. This proposition establishes a local ambiguity in joint frequency–velocity recovery from a 1D received signal. Our canonicalization controls examine velocity estimation from signal histories in multipath settings (Appendix I.2), providing empirical diagnostics beyond the plane-wave analysis. When receiver velocity is given, this inverse problem is avoided.

Decoupling. For a passive receiver that does not perturb the field, changing velocity changes sampling locations, not wavefield evolution, with source, medium, and boundary conditions otherwise unchanged. WTD therefore uses a velocity-agnostic operator $\mathcal{G}_\theta$ to predict future Eulerian fields from local field histories, and a parameter-free geometric projector $\Pi_{\mathbf{v}}$ to sample them along the known trajectory, producing the 1D Lagrangian signal. Using the supplied velocity avoids velocity estimation, while explicit trajectory sampling removes the need to learn a velocity-dependent sampling map.

3.3 Wavefield–Trajectory Decoupling Architecture

WTD comprises two modules (Fig. 2): a velocity-agnostic complex operator $\mathcal{G}_\theta$ predicting the Eulerian field, and a parameter-free geometric projector $\Pi_{\mathbf{v}}$ that samples the predicted wavefield along the known receiver trajectory:

$$\widehat{\mathbf{u}}_{\mathrm{future}} = \mathcal{G}_\theta(\mathbf{u}_{\mathrm{past}}), \qquad \widehat{s}_{\mathbf{v}} = \Pi_{\mathbf{v}}(\widehat{\mathbf{u}}_{\mathrm{future}}). \tag{5}$$

By construction, $\mathbf{v}$ enters only through $\Pi_{\mathbf{v}}$, following the decoupling principle in §3.2. The complete forward pass is given in Algorithm 1 in Appendix C.

Complex baseband representation. The real-valued wavefield $E(\mathbf{x}, t)$ contains rapid oscillations at the carrier angular frequency ω_c, making phase-accurate prediction challenging. We apply a temporal Hilbert transform to the past frames only, then demodulate to complex baseband:

$$U(\mathbf{x}, t) = E(\mathbf{x}, t) + \mathrm{j}\mathcal{H}_t[E](\mathbf{x}, t), \qquad u(\mathbf{x}, t) = U(\mathbf{x}, t)\mathrm{e}^{-\mathrm{j}\omega_c t}. \tag{6}$$

This removes the temporal carrier factor while retaining amplitude and phase variations required for trajectory sampling. Hereafter, $\mathbf{u}$ and $s_{\mathbf{v}}$ denote complex baseband fields and signals, respectively.

Velocity-agnostic complex neural operator $\mathcal{G}_\theta$. Following WTD's decoupling principle, $\mathcal{G}_\theta$ predicts Eulerian wavefields without velocity input. Complex-valued dilated convolutions cover the $P \times P$ patch (RF $\geq P$). Because trajectory sampling transfers wavefield phase errors to the signal, phase accuracy is essential. For $z = a + \mathrm{j}b$, componentwise activation changes its phase to

$\operatorname{atan2}(f(b), f(a))$. We adapt magnitude-based activations by applying GELU to the magnitude (Arjovsky et al., 2016):

$$\sigma(z) = \frac{z}{|z| + \epsilon}\ \mathrm{GELU}(|z|). \tag{7}$$

For $z \neq 0$, $\arg \sigma(z) = \arg z$. Implementation details appear in Appendix C.

Parameter-free geometric projector $\Pi_{\mathbf{v}}$. Given the known receiver velocity $\mathbf{v}$, the receiver position at output time step n is $\mathbf{r}(n)$. $\Pi_{\mathbf{v}}$ samples the predicted wavefield at $\mathbf{r}(n)$:

$$\mathbf{r}(n) = \mathbf{r}_{\mathrm{f}} + \alpha \mathbf{v} n, \qquad \widehat{s}_{\mathbf{v}}[n] = \mathcal{I}_{\mathrm{bilin}}\big(\widehat{u}_n, \mathbf{r}(n)\big), \quad n = 0, \ldots, T_{\mathrm{out}} - 1, \tag{8}$$

where $\mathbf{r}_{\mathrm{f}}$ is the initial forecast position in grid coordinates, $\widehat{u}_n$ is the predicted Eulerian wavefield at output time step n, and $\alpha = \Delta t_{\mathrm{snap}}/\Delta x$, so $\alpha \mathbf{v}$ is the displacement per forecast step in grid coordinates. $\mathcal{I}_{\mathrm{bilin}}$ applies bilinear interpolation separately to the real and imaginary components, using weights determined by $\mathbf{r}(n)$. Because the weights are not learned, $\Pi_{\mathbf{v}}$ has no learnable parameters. Consequently, receiver velocity enters WTD only through $\mathbf{r}(n)$.

Doppler shift from trajectory sampling. For a locally monochromatic component, let $\phi(\mathbf{x}, t) = \omega_0 t - \mathbf{k} \cdot \mathbf{x} + \phi_0$, where ω_0 is the wave angular frequency and $\mathbf{k}$ is the wave vector. Define the baseband phase $\psi = \phi - \omega_c t$ and receiver trajectory $\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}t$. The chain rule gives

$$\frac{d}{dt}\psi\big(\mathbf{r}(t), t\big) = \frac{\partial \psi}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} \psi = (\omega_0 - \omega_c) - \mathbf{v} \cdot \mathbf{k}. \tag{9}$$

$$\omega_{\mathrm{recv}} = \omega_c + \frac{d}{dt}\psi(\mathbf{r}(t), t) = \omega_0 - \mathbf{v} \cdot \mathbf{k}, \qquad \Delta\omega_{\mathrm{D}} = -\mathbf{v} \cdot \mathbf{k}. \tag{10}$$

In WTD, the Doppler shift therefore arises from trajectory sampling rather than learned velocity dependence. The discrete projector approximates this sampling at forecast times using spatial interpolation. Appendix A.6 gives the discrete formulation.

3.4 Training Objective

Ground-truth future wavefields provide training-only supervision. We optimize

$$\mathcal{L} = \mathcal{L}_{\mathrm{env}} + \beta_{\mathrm{phase}} \mathcal{L}_{\mathrm{phase}} + \lambda_{\mathrm{energy}} \mathcal{L}_{\mathrm{energy}} + \lambda_{\mathrm{field}} \left(\mathcal{L}_{\mathrm{field}} + \lambda_{\mathrm{grad}} \mathcal{L}_{\mathrm{grad}} \right). \tag{11}$$

$\mathcal{L}_{\mathrm{env}}$ and $\mathcal{L}_{\mathrm{field}}$ measure mean-squared errors in the real and imaginary parts of the 1D signal and future Eulerian wavefields, respectively. $\mathcal{L}_{\mathrm{energy}}$ matches the average $|s|^2$ of predicted and target signals. For phase supervision, define

$$\hat{q}_n = \frac{\hat{s}_n \hat{s}_{n-1}^*}{|\hat{s}_n \hat{s}_{n-1}^*| + \epsilon}, \qquad q_n = \frac{s_n s_{n-1}^*}{|s_n s_{n-1}^*| + \epsilon}, \tag{12}$$

where * denotes complex conjugation. $\mathcal{L}_{\mathrm{phase}}$ compares these increments with target-amplitude weighting (Appendix C). $\mathcal{L}_{\mathrm{grad}}$ matches spatial field differences between neighboring grid points along both axes. Signal losses supervise 1D prediction, while field losses constrain values and spatial variations across the output grid. Loss weights are given in Appendix C.

4 Experiments

We first compare WTD with baselines on signal prediction (§4.2). We then verify the velocity independence of WTD's field prediction (§4.3), analyze the separate and combined effects of decoupling and full-field supervision (§4.4), and ablate individual components (§4.5). Finally, we compare WTD with augmentation and learned canonicalization (§4.6) and perform a cross-solver evaluation of WTD and FiLM using MEEP (Appendix J).

4.1 Setup

Data. Yee FDTD generates 240 training and 60 held-out test wavefields, with reference wave speed $c = 1$. Training and ID speeds are $v = \|\mathbf{v}\| \in \{0.06, 0.08, 0.10\}$; OOD speeds are $v \in \{0.15, 0.20\}$. Data generation and spatial spectra are detailed in Appendices E and B.

Table 1: **Signal and Doppler errors under velocity shift.** Results use 60 held-out fields at ID speeds $v \in \{0.06, 0.08, 0.10\}$ and OOD speeds $v \in \{0.15, 0.20\}$. Within each run, errors are averaged over velocities per field, then over fields. Values are mean $\pm$ standard deviation over three seeds. First-block methods use known velocity; canonicalization estimates it from the past signal.

Model	Params	Signal Rel-L_2 ID (↓)	Signal Rel-L_2 OOD (↓)	Abs. Doppler Error/f_c ID (↓)	Abs. Doppler Error/f_c OOD (↓)	Doppler OOD/ID
Known-velocity methods						
WTD (ours)	0.70M	$\mathbf{0.0957 \pm 0.0016}$	$\mathbf{0.1235 \pm 0.0038}$	$\mathbf{0.0113 \pm 0.0008}$	$\mathbf{0.0113 \pm 0.0005}$	**1.00**
FiLM (plain)	0.99M	1.0718 ± 0.0431	1.5125 ± 0.1410	0.0354 ± 0.0009	0.0591 ± 0.0137	1.67
FiLM + field	∼1.0M	1.0592 ± 0.0219	1.4902 ± 0.0590	0.0361 ± 0.0007	0.0434 ± 0.0009	1.20
FiLM + LPSDA	0.99M	1.1734 ± 0.1440	1.5870 ± 0.1543	0.0321 ± 0.0033	0.0714 ± 0.0290	2.23
Cond. concat	0.45M	1.1241 ± 0.0173	1.3225 ± 0.1573	0.0317 ± 0.0027	0.1026 ± 0.0281	3.24
Cond. adaLN	0.58M	1.2457 ± 0.2465	1.7129 ± 0.0191	0.0272 ± 0.0029	0.1295 ± 0.1442	4.76
FNO (e2e)	8.45M	1.0822 ± 0.0568	1.3104 ± 0.0400	0.0351 ± 0.0016	0.0953 ± 0.0080	2.71
Estimated-velocity diagnostics						
Canon. ($+v$ sup.)†	0.74M	0.5602 ± 0.1604	1.3888 ± 0.0016	0.0200 ± 0.0003	0.0791 ± 0.0041	3.95
Canon. (e2e)†	0.74M	0.6093 ± 0.1276	1.3913 ± 0.0089	0.0192 ± 0.0007	0.0798 ± 0.0044	4.15

Signal Rel-L_2 is $\|\hat{s} - s\|_2 / \|s\|_2$.

Normalized absolute Doppler error is $\left|\widehat{\Delta f}_{\mathrm{D}} - \Delta f_{\mathrm{D}}\right| / f_c$. Predicted and target Doppler shifts are measured relative to the same ground-truth stationary-receiver frequency. Frequencies are estimated using the lag-1 correlation estimator defined in Appendix F. Here $f_c = 1$ in normalized units. Doppler OOD/ID is the ratio of mean OOD to mean ID Doppler error, computed using unrounded values.

†Both canonicalization variants share a frozen WTD field operator from one training run. Only the 0.042M-parameter canonicalizer is trained. Reported standard deviations are over three independent canonicalizer runs.

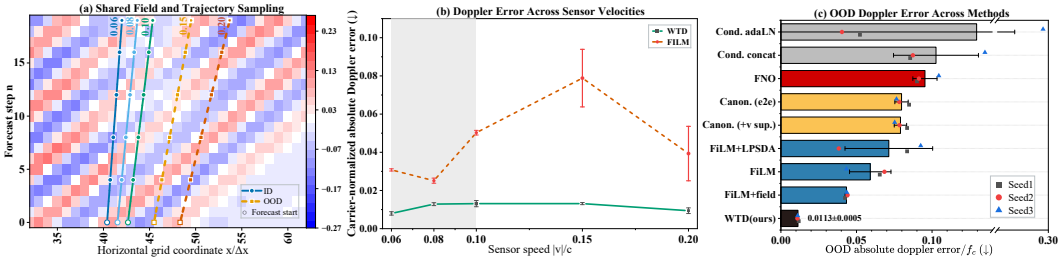


Figure 3: **Velocity-independent field prediction and OOD stability.** **(a)** A fixed-y space–time slice of WTD's predicted Eulerian field with ID (solid) and OOD (dashed) sampling trajectories; open markers indicate forecast starts. **(b)** Mean Doppler error across receiver velocities; error bars show standard deviation over three seeds. **(c)** OOD Doppler error across methods; bars show means, error bars indicate standard deviations, and markers denote individual runs.

Metrics. We report Signal Rel-L_2, carrier-normalized absolute Doppler error, and Field N-L_2 where field predictions are available. Unless stated otherwise, results are mean $\pm$ standard deviation over three seeds. Definitions and statistical tests are provided in Appendix F.

Baselines. All methods use identical data splits and past 2D fields. Baselines supplied with known velocity use FiLM, feature concatenation, adaLN, or an end-to-end FNO. FiLM+field matches WTD's full-field supervision, while FiLM+LPSDA adds trajectory augmentation within the ID velocity range. Canonicalization instead estimates velocity from past 1D signals. Detailed configurations are provided in Appendix C.

4.2 Signal Prediction under Velocity Shift

Table 1 reports signal relative L_2 error and carrier-normalized absolute Doppler error at ID and OOD velocities. WTD achieves the lowest signal relative L_2 errors: 0.0957 ± 0.0016 ID and 0.1235 ± 0.0038 OOD. These values are lower than the best known-velocity baseline in each regime by factors of 11.1 and 10.6, respectively. The OOD advantage therefore persists beyond the training velocity range. WTD's mean Doppler error remains 0.0113 in both regimes, giving an OOD/ID ratio of 1.00.

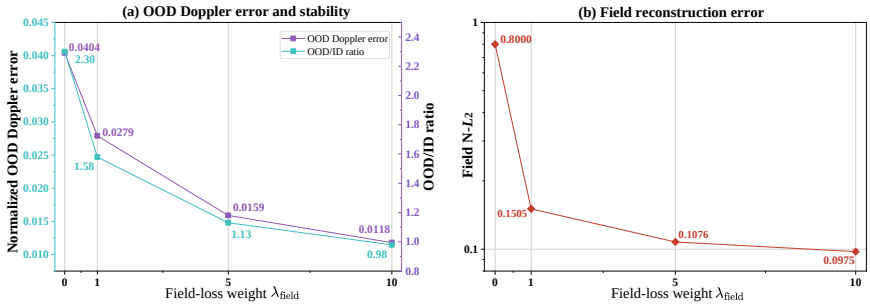


Figure 4: **Fixed-seed field-loss-weight sweep.** As λ_{field} increases, OOD Doppler error decreases and the OOD/ID ratio approaches 1 **(a)**, while Field N-L_2 decreases **(b)**; panel **(b)** is log-scaled.

Its OOD Doppler error is lower than the next-best result by a factor of 3.8. The same increase in Doppler error under velocity shift is observed for direct concatenation and adaLN, indicating that the trend is not specific to FiLM (Appendix H.1). Across seeds, even WTD's worst OOD Doppler result is better than the best result of every baseline (Fig. 3(c)). Individual results from each run for WTD and FiLM are reported in Appendix D. WTD's OOD Doppler advantage is significant under two-sided paired Wilcoxon tests against every baseline ($p_{\text{Dop}} \leq 8.7 \times 10^{-10}$). The 8.45M parameter FNO remains less accurate than the 0.70M parameter WTD. Canonicalization is analyzed separately in §4.6 because it estimates velocity instead of receiving the known value.

4.3 Velocity Independence of the Predicted Field

Receiver velocity $\mathbf{v}$ enters WTD only through the geometric projector $\Pi_{\mathbf{v}}$, not $\mathcal{G}_\theta$. Thus, a fixed input of past fields yields the same predicted Eulerian field across velocities. We verify this by fixing the past field input and changing only the velocity. We compare the spatial representation used to produce the 1D output: the predicted Eulerian field in WTD and the velocity-conditioned feature map in FiLM. The WTD field has zero relative change up to numerical precision. The FiLM feature map has a mean relative change of 1.21 over 300 comparisons involving 60 test fields and five velocities. This confirms that velocity conditioning alters FiLM's spatial representation. Appendix H.2 defines the probe and reports the results. Figure 3(a) shows how the geometric projector samples the predicted Eulerian field along different trajectories. WTD's carrier-normalized absolute Doppler error remains 0.0113 at both ID and OOD velocities (OOD/ID $= 1.00$). By contrast, FiLM has a higher Doppler error at OOD velocities than at ID velocities (Fig. 3(b)). In WTD, $\mathbf{v}$ changes only the trajectory used by $\Pi_{\mathbf{v}}$. The learned operator therefore does not need to model how the predicted field varies with velocity or extrapolate such dependence beyond the training range. This suggests that, in WTD, the remaining learned error is primarily associated with imperfect field reconstruction rather than velocity-dependent extrapolation. We test this attribution in §4.4.

4.4 Structural Decoupling and Full-Field Supervision

To disentangle the effects of structural decoupling and full-field supervision, we conduct a controlled, 2×2 ablation using a fixed seed. All variants use the same inputs of past fields and matched complex-valued backbones. The two factors are velocity handling (decoupled WTD or velocity-conditioned FiLM) and full-field supervision. Each variant is evaluated on the same 60 held-out fields at five velocities, yielding 300 matched cases across fields and velocities (Appendix H.6, Table 10). Both factors reduce OOD Doppler error, and their combination performs best. Relative to signal-only FiLM, decoupling alone reduces the OOD Doppler error by 38%, while field supervision alone reduces it by 35%. Combining both reduces the error by 82% and yields an OOD/ID ratio of 0.98. Adding field supervision reduces the OOD error by a factor of 3.4 for WTD and 1.5 for FiLM. Under matched field supervision, FiLM's OOD error is 3.6 times that of WTD, with an OOD/ID ratio of 1.18. This difference is statistically significant under a two-sided Wilcoxon signed-rank test performed over the 60 held-out fields ($p = 9.1 \times 10^{-10}$). To assess sensitivity to the weight on the field loss, we vary $\lambda_{\text{field}} \in \{0, 1, 5, 10\}$ while holding the architecture and initialization fixed. OOD Doppler error, the OOD/ID ratio, and Field N-L_2 decrease monotonically over the tested values

Table 2: **Fixed-seed component ablations.** Each row modifies WTD (full); dilation and normalization are ablated jointly.

Variant	Doppler ID ↓	Doppler OOD ↓	OOD/ID	Field N-L_2 ↓
WTD (full)	0.0121	**0.0118**	**0.98**	**0.0975**
Dilation [1]/no norm	0.0222	0.0309	1.39	0.2700
No gradient loss	**0.0119**	0.0155	1.30	0.1056
Separate GELUs	0.0407	0.0407	1.00*	12.850

Dilation [1]*/no norm*: The dilation schedule is set to [1], and max-magnitude normalization is removed. *No gradient loss*: $\lambda_{\mathrm{grad}}=0$ retains field supervision. *Separate GELUs*: Separate GELUs yield a 132-fold increase in Field N-L_2.
* ID and OOD errors are both 0.0407; the ratio does not imply low error.

(Fig. 4). The corresponding numerical values are provided in Appendix H.4. Together, these results support distinct but complementary roles: structural decoupling ensures velocity-independent field prediction, whereas full-field supervision improves the accuracy of the predicted field.

4.5 Component Ablations

Table 2 reports three component ablations using the same random seed. Table 1 instead reports averages over three trained runs. All three ablations increase both OOD Doppler error and Field N-L_2. Replacing the dilation schedule $[1, 2, 4, 8]$ with $[1]$ and removing max-magnitude normalization jointly increase OOD Doppler error by a factor of 2.6 and Field N-L_2 by a factor of 2.8. Because the dilation schedule and normalization are modified jointly, this ablation evaluates the combined effect of changing the dilation schedule and removing max-magnitude normalization. The remaining ablations examine the loss and activation design. Removing the spatial-gradient loss increases the OOD/ID ratio from 0.98 to 1.30, suggesting that gradient supervision improves stability under velocity shift. Applying GELU separately to the real and imaginary parts leads to severe field-reconstruction failure: Field N-L_2 increases to 12.850 and OOD Doppler error to 0.0407. The resulting OOD/ID ratio of 1.00 reflects similarly poor ID and OOD predictions rather than genuine OOD stability. Ablation details are provided in Appendix H.5. Appendix H.7 further discusses why the OOD/ID ratio must be interpreted jointly with the absolute error.

4.6 Alternative Strategies and Cross-Solver Evaluation

WTD passes known receiver velocity only to its parameter-free geometric projector. We compare this design with two alternatives: trajectory augmentation within the ID velocity range and learned canonicalization, which estimates velocity before trajectory sampling. A mechanistic comparison is provided in Appendix G, with quantitative results in Appendix I. For FiLM, LPSDA reduces ID Doppler error from 0.0354 to 0.0321 but increases OOD error from 0.0591 to 0.0714, yielding an OOD/ID ratio of 2.23. Thus, LPSDA does not improve FiLM's OOD Doppler error at the evaluated velocities. The supervised and end-to-end canonicalization variants attain ID errors of 0.0200 and 0.0192; their OOD errors increase to 0.0791 and 0.0798, yielding ratios of 3.95 and 4.15. For both variants, the estimated OOD velocities are biased toward the training range. Because canonicalization estimates velocity instead of using the known value, we report these variants as alternative pipelines that estimate velocity, rather than as directly matched baselines. To assess robustness to the data-generation solver, we retrain WTD and FiLM from scratch on independently generated MEEP data using three seeds. FiLM's mean OOD Doppler error is 20.1 times that of WTD, showing that WTD's advantage persists under an independent solver. Full results are in Appendix J.

5 Conclusion

We introduced Wavefield–Trajectory Decoupling, separating velocity-independent wavefield prediction from deterministic sampling along known receiver trajectories. Combined with phase-preserving complex learning and Sobolev supervision, WTD improves signal prediction under both ID and OOD velocities. Experiments across two electromagnetic solvers support decoupling known receiver motion from learned field dynamics as an effective approach to robust moving-receiver signal prediction.

AI USE STATEMENT

In this work, we used generative AI tools to aid in the writing and polishing of the manuscript. Specifically, we used generative AI tools to assist in refining the language, improving readability, ensuring clarity, rephrasing sentences, checking grammar, and enhancing the overall flow of the text. We have not used generative AI tools for research ideation, research methodology, experimental design, scientific content, or data analysis, and the rest of the required disclosure tasks are not applicable to this work. Additionally, we used generative AI tools solely to improve the linguistic quality of the manuscript. We have reviewed all AI-assisted work and ensured that the AI-generated text adheres to ethical guidelines and does not contribute to plagiarism or scientific misconduct. All research concepts, ideas, and analyses were developed and conducted by the authors. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

Our work is only aimed at learning-based electromagnetic wavefield and moving-receiver signal prediction for human good, so there is no involvement of human subjects or conflicts of interest as far as the authors are aware of.

REPRODUCIBILITY STATEMENT

We support reproducibility through detailed descriptions of the proposed method, theoretical results, data-generation process, training protocol, evaluation procedure, and run-level results. The WTD architecture, parameter-free geometric projector, training objective, and experimental protocol are specified in Sections 3.3–3.4 and Section 4.1. Appendix A states the assumptions, provides the complete proof of Proposition 1, and analyzes continuous-to-discrete trajectory sampling. Appendix C documents the architectures and forward procedures of WTD and all compared models, together with the loss definitions and weights, normalization, optimization, hyperparameters, and computational resources. Appendices E and B describe FDTD wavefield generation, patch extraction, physical and numerical parameters, data splits, and spatial spectral analysis. Appendix F defines all evaluation metrics, aggregation procedures, and statistical tests. Run-level results and additional diagnostics, including the representation probes, field-supervision sweep, component ablations, 2×2 factorial study, alternative-paradigm controls, and cross-solver MEEP evaluation, are reported in Appendices D, H, I, and J.

REFERENCES

Martin Arjovsky, Amar Shah, and Yoshua Bengio. Unitary evolution recurrent neural networks. In Maria Florina Balcan and Kilian Q. Weinberger (eds.), *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 1120–1128, New York, New York, USA, 20–22 Jun 2016. PMLR. URL https://proceedings.mlr.press/v48/arjovsky16.html.

Johannes Brandstetter, Max Welling, and Daniel E. Worrall. Lie point symmetry data augmentation for neural PDE solvers. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 2241–2256. PMLR, 2022.

Mainak Chakraborty, Masood Aryapoor, and Masoud Daneshtalab. Frequency domain complex-valued convolutional neural network. *Expert Systems with Applications*, 295:128893, 2026. ISSN 0957-4174. doi: https://doi.org/10.1016/j.eswa.2025.128893. URL https://www.sciencedirect.com/science/article/pii/S0957417425025102.

Sikai Cheng, Reza Zandehshahvar, Haoruo Zhao, Daniel A. Garcia-Ulloa, Alejandro Villena-Rodriguez, Carles Navarro Manchón, and Pascal Van Hentenryck. Csi-4cast: A hybrid deep learning model for csi prediction with comprehensive robustness and generalization testing. *IEEE Transactions on Machine Learning in Communications and Networking*, 4:882–912, 2026. doi: 10.1109/TMLCN.2026.3694830.

Hyeong-Seok Choi, Janghyun Kim, Jaesung Huh, Adrian Kim, Jung-Woo Ha, and Kyogu Lee. Phase-aware speech enhancement with deep complex u-net. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=SkeRTsAcYm.

Elizabeth Cole, Joseph Cheng, John Pauly, and Shreyas Vasanawala. Analysis of deep complex-valued convolutional neural networks for mri reconstruction and phase-focused applications. *Magnetic Resonance in Medicine*, 86(2):1093–1109, 2021. doi: https://doi.org/10.1002/mrm.28733. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.28733.

Wojciech Marian Czarnecki, Simon Osindero, Max Jaderberg, Grzegorz Świrszcz, and Razvan Pascanu. Sobolev training for neural networks, 2017. URL https://arxiv.org/abs/1706.04859.

Hüseyin Göksu and Donald Wunsch. Fdtd data extrapolation using multilayer perceptron (mlp). pp. 735 – 737 vol.2, 09 2003. ISBN 0-7803-7835-0. doi: 10.1109/ISEMC.2003.1236698.

Jan Hagnberger, Marimuthu Kalimuthu, Daniel Musekamp, and Mathias Niepert. Vectorized conditional neural fields: A framework for solving time-dependent parametric partial differential equations. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 17189–17223. PMLR, 2024. URL https://proceedings.mlr.press/v235/hagnberger24a.html.

Jacob Helwig, Xuan Zhang, Cong Fu, Jerry Kurtin, Stephan Wojtowytsch, and Shuiwang Ji. Group equivariant Fourier neural operators for partial differential equations. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 12907–12930. PMLR, 2023. URL https://proceedings.mlr.press/v202/helwig23a.html.

Vikram Jandhyala, Eric Michielssen, and Raj Mittra. Fdtd signal extrapolation using the forward-backward autoregressive (ar) model. *IEEE Microwave and Guided Wave Letters*, 4:163–165, 1994.

Yanliang Jin, Yifan Wu, Yuan Gao, Shunqing Zhang, Shugong Xu, and Cheng-Xiang Wang. Linformer: A linear-based lightweight transformer architecture for time-aware mimo channel prediction. *IEEE Transactions on Wireless Communications*, 24(9):7177–7190, 2025. doi: 10.1109/TWC.2025.3558950.

Sékou-Oumar Kaba, Arnab Kumar Mondal, Yan Zhang, Yoshua Bengio, and Siamak Ravanbakhsh. Equivariance with learned canonicalization functions. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 15546–15566. PMLR, 2023. URL https://proceedings.mlr.press/v202/kaba23a.html.

Dongyue Li, Hongxi Chen, Shuanghong Zhou, and Yuting Zhang. Electromagnetic compatibility test signal prediction method based on lstm network. *Journal of Signal Processing Systems*, 96:837–848, 03 2025. doi: 10.1007/s11265-025-01952-z.

Qiaoyu Li, Philip Sisk, Arumugam Kannan, Taesang Yoo, Tao Luo, Gaurav Shah, Badri Manjunath, Chanaka Samarathungage, Mahmoud Taherzadeh Boroujeni, Hamed Pezeshki, and Himanshu Joshi. Machine learning based time domain millimeter-wave beam prediction for 5g-advanced and beyond: Design, analysis, and over-the-air experiments. *IEEE Journal on Selected Areas in Communications*, 41(6):1787–1809, 2023. doi: 10.1109/JSAC.2023.3275613.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations (ICLR)*, 2021.

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3:218–229, 2021.

George Ma, Yifei Wang, Derek Lim, Stefanie Jegelka, and Yisen Wang. A canonicalization perspective on invariant and equivariant learning. In *Advances in Neural Information Processing Systems*, volume 37, pp. 60936–60979. Curran Associates, Inc., 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/702b67152ec4435795f681865b67999c-Abstract-Conference.html.

Pingchuan Ma, Haoyu Yang, Zhengqi Gao, Duane S. Boning, and Jiaqi Gu. PIC2O-Sim: A physics-inspired causality-aware dynamic convolutional neural operator for ultra-fast photonic device time-domain simulation. *APL Photonics*, 10(3):036104, 2025. doi: 10.1063/5.0242728.

Mehdi Meliha, Pascal Chargé, Yide Wang, Salah Eddine Bouzid, Christophe Henry, and Yejian Chen. Csi prediction with a single ai model for heterogeneous doppler frequencies and multipaths. In *2025 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, 2025. doi: 10.1109/ICMLCN64995.2025.11140519.

K. Miller and M. Rochwarger. A covariance approach to spectral moment estimation. *IEEE Transactions on Information Theory*, 18(5):588–596, 1972. doi: 10.1109/TIT.1972.1054886.

Frodouard Minani, Makoto Kobayashi, Takuya Fujihashi, Md. Abdul Alim, Shunsuke Saruwatari, Masahiro Nishi, and Takashi Watanabe. Channel prediction and fair resource allocation for ntn uplinks by lstm and deep reinforcement learning. *IEEE Transactions on Wireless Communications*, 24(10):8311–8330, 2025. doi: 10.1109/TWC.2025.3565774.

Steven C. Nardone and Vincent J. Aidala. Observability criteria for bearings-only target motion analysis. *IEEE Transactions on Aerospace and Electronic Systems*, AES-17(2):162–166, 1981.

Indranil Nayak, Fernando L. Teixeira, and Robert J. Burkholder. On-the-fly dynamic mode decomposition for rapid time-extrapolation and analysis of cavity resonances. *IEEE Transactions on Antennas and Propagation*, 72(1):131–146, 2024. doi: 10.1109/TAP.2023.3295511.

Ardavan F. Oskooi, David Roundy, Mihai Ibanescu, Peter Bermel, J.D. Joannopoulos, and Steven G. Johnson. Meep: A flexible free-software package for electromagnetic simulations by the fdtd method. *Computer Physics Communications*, 181(3):687–702, 2010. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2009.11.008. URL https://www.sciencedirect.com/science/article/pii/S001046550900383X.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4195–4205, 2023.

Ethan Perez, Florian Strub, Harm de Vries, Vincent Dumoulin, and Aaron C. Courville. FiLM: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, pp. 3942–3951, 2018. doi: 10.1609/aaai.v32i1.11671.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.

Bogdan Raonic, Roberto Molinaro, Tim De Ryck, Tobias Rohner, Francesca Bartolucci, Rima Alaifari, Siddhartha Mishra, and Emmanuel de Bézenac. Convolutional neural operators for robust and accurate learning of pdes. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (eds.), *Advances in Neural Information Processing Systems*, volume 36, pp. 77187–77200. Curran Associates, Inc., 2023. doi: 10.52202/075280-3376. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/f3c1951b34f7f55ffaecada7fde6bd5a-Paper-Conference.pdf.

Zakhar Shumaylov, Peter Zaika, James Rowbottom, Ferdia Sherry, Melanie Weber, and Carola-Bibiane Schönlieb. Lie algebra canonicalization: Equivariant neural operators under arbitrary lie groups. In *International Conference on Learning Representations (ICLR)*, 2025.

Sifan Wang, Jacob H. Seidman, Shyam Sankaran, Hanwen Wang, George J. Pappas, and Paris Perdikaris. Bridging operator learning and conditioned neural fields: A unifying perspective. *arXiv preprint arXiv:2405.13998*, 2024.

Chen Wu, Enrique A. Navarro, Joaquin Navasquillo, and John Litva. Fdtd signal extrapolation using a finite impulse response neural network model. *Microwave and Optical Technology Letters*, 21(5):325–330, 1999. doi: https://doi.org/10.1002/(SICI)1098-2760(19990605)21:5⟨325::AID-MOP6⟩3.0.CO;2-C.

Zhuoran Xiao, Zhaoyang Zhang, Zirui Chen, Zhaohui Yang, Chongwen Huang, and Xiaoming Chen. From data-driven learning to physics-inspired inferring: A novel mobile mimo channel prediction scheme based on neural ode. *IEEE Transactions on Wireless Communications*, 23(7): 7186–7199, 2024. doi: 10.1109/TWC.2023.3338419.

Bo-Heng Yeh, Jen-Ming Wu, and Ronald Y. Chang. Efficient doppler compensation for leo satellite downlink ofdma systems. *IEEE Transactions on Vehicular Technology*, 73(12):18863–18877, 2024. doi: 10.1109/TVT.2024.3437430.

Haifan Yin, Haiquan Wang, Yingzhuang Liu, and David Gesbert. Addressing the curse of mobility in massive mimo with prony-based angular-delay domain channel predictions. *IEEE Journal on Selected Areas in Communications*, 38(12):2903–2917, 2020. doi: 10.1109/JSAC.2020.3005473.

Li You, Ke-Xin Li, Jiaheng Wang, Xiqi Gao, Xiang-Gen Xia, and Björn Ottersten. Massive mimo transmission for leo satellite communications. *IEEE Journal on Selected Areas in Communications*, 38:1851–1865, 06 2020. doi: 10.1109/JSAC.2020.3000803.

Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. In *Advances in Neural Information Processing Systems*, volume 33, pp. 5824–5836, 2020.

Yongs Zeng, Qingqing Wu, and Rui Zhang. Accessing from the sky: A tutorial on uav communications for 5g and beyond. *Proceedings of the IEEE*, 107(12):2327–2375, 2019. doi: 10.1109/JPROC.2019.2952892.

Huiwen Zhang, Feng Ye, and Chu Ma. Physics-informed neural networks with architectural physics embedding for large-scale wave field reconstruction. *arXiv preprint arXiv:2603.02231*, 2026. doi: 10.48550/arXiv.2603.02231.

Haihong Zhao, Zinan Zheng, Chenyi Zi, and Jia Li. Channelmts: A multi-modal time-series framework for high-speed railway channel prediction. In *KDD 2026 - Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1*, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 2594–2605, United States, April 2026. Association for Computing Machinery. doi: 10.1145/3770854.3783957.

Lipeng Zhu, Wenyan Ma, and Rui Zhang. Modeling and performance analysis for movable antenna enabled wireless communications. *IEEE Transactions on Wireless Communications*, 23(6):6234–6250, 2024. doi: 10.1109/TWC.2023.3330887.

APPENDIX TABLE OF CONTENTS

A Local Unidentifiability: Derivation, Fisher Analysis, and Discrete Sampling 15
A.1 Received frequency along a moving-receiver trajectory 15
A.2 Constructive non-injectivity of the observation map 16
A.3 Fisher information, rank, and estimability 17
A.4 Local extension and scope beyond plane waves 18
A.5 Connection to passive-motion observability and empirical diagnostics 19
A.6 Discrete trajectory sampling and error decomposition 20
B Spatial Spectral Complexity and Directional Content 21
C Implementation and Training Details 22
C.1 Decoupled WTD Inference 24
D Results Across Independent Runs 26
E Data Generation and Task Configuration 26
F Evaluation Metrics and Statistical Protocol 28
F.1 Signal Relative L_2 Error 29
F.2 Carrier-Normalized Doppler Recovery Error 29
F.3 Field Reconstruction N-L_2Norm 29
F.4 Aggregation and Statistical Significance 30
G Qualitative Comparison with Related Work 30
H Supplementary Experimental Results 30
H.1 Robustness to Velocity-Injection Mechanism 30
H.2 Internal-Representation Velocity-Dependence Probe 30
H.3 Cross-Velocity Gradient Interference 31
H.4 Field-Supervision Strength Sweep: Full Numerical Values 31
H.5 Ablation Experimental Details 32
H.6 Full Results of the 2×2 Factorial Study 32
H.7 Interpreting the Doppler OOD/ID Ratio 32
I Competing Paradigms: Full Quantitative Results 33
I.1 Augmentation paradigm (LPSDA) 33
I.2 Learned Canonicalization 33
J Cross-Solver Robustness: MEEP Evaluation 34

A Local Unidentifiability: Derivation, Fisher Analysis, and Discrete Sampling

This appendix develops the theoretical analysis underlying Proposition 1. It proceeds in six parts. §A.1 derives the received-frequency relation using the chain rule; §A.2 constructs a non-injective local wavefield–velocity family; §A.3 characterizes the resulting Fisher-information rank deficiency and its estimable directions; §A.4 clarifies the local scope beyond the single-plane-wave construction; §A.5 relates the local ambiguity to classical passive-tracking observability and to the learned-canonicalization diagnostic; and §A.6 derives the discrete trajectory-sampling relation and separates field-prediction error from interpolation error.

Together, these results establish non-injectivity within an admissible local wavefield subfamily. They do not assert that every arbitrary multipath observation has a rank-one Fisher information matrix.

Throughout, let $U(\mathbf{x},t) = A(\mathbf{x},t)\,\mathrm{e}^{\mathrm{j}\phi(\mathbf{x},t)}$ denote the complex analytic Eulerian field in the stationary simulation frame, and let $u(\mathbf{x},t) = U(\mathbf{x},t)\mathrm{e}^{-\mathrm{j}\omega_c t}$ denote its complex baseband representation used by the neural operator. We assume that $A(\mathbf{x},t) > 0$ over the local observation window so that ϕ is well defined. The receiver follows the continuous trajectory $\mathbf{x}_s(t) = \mathbf{x}_0 + \mathbf{v}t$ with constant velocity $\mathbf{v} \in \mathbb{R}^d$ in the low-speed Galilean regime $|\mathbf{v}|/c \le 0.2$. Following the sign convention in the main text, we define the local instantaneous angular frequency and wavevector as $\omega_{\mathrm{loc}}(\mathbf{x},t) = \partial_t\phi(\mathbf{x},t)$ and $\mathbf{k}(\mathbf{x},t) = -\nabla_{\mathbf{x}}\phi(\mathbf{x},t)$, respectively. A locally monochromatic plane wave therefore takes the analytic form $\mathrm{e}^{\mathrm{j}(\omega_0 t - \mathbf{k}\cdot\mathbf{x})}$ and the corresponding baseband form $\mathrm{e}^{\mathrm{j}((\omega_0-\omega_c)t - \mathbf{k}\cdot\mathbf{x})}$.

A.1 Received frequency along a moving-receiver trajectory

The receiver samples the analytic field along its time-dependent trajectory, yielding the 1D analytic signal

$$s(t) = U(\mathbf{x}_s(t),t) = A(\mathbf{x}_s(t),t)\exp(\mathrm{j}\phi(\mathbf{x}_s(t),t))\,. \tag{13}$$

Because $A(\mathbf{x}_s(t),t) > 0$ over the local observation window, the instantaneous angular frequency is well defined as

$$\omega_{\mathrm{recv}}(t) = \frac{d}{dt}\arg s(t) = \frac{d}{dt}\phi(\mathbf{x}_s(t),t). \tag{14}$$

Since $\mathbf{x}_s(t)$ depends on time, this is a total derivative. Applying the multivariate chain rule gives

$$\omega_{\mathrm{recv}}(t) = \frac{\partial\phi}{\partial t}\big(\mathbf{x}_s(t),t\big) + \sum_{i=1}^{d}\frac{\partial\phi}{\partial x_i}\big(\mathbf{x}_s(t),t\big)\frac{dx_{s,i}(t)}{dt} \tag{15}$$

$$= \frac{\partial\phi}{\partial t}\big(\mathbf{x}_s(t),t\big) + \dot{\mathbf{x}}_s(t)\cdot\nabla_{\mathbf{x}}\phi\big(\mathbf{x}_s(t),t\big). \tag{16}$$

Using

$$\omega_{\mathrm{loc}}(\mathbf{x},t) = \partial_t\phi(\mathbf{x},t), \qquad \mathbf{k}(\mathbf{x},t) = -\nabla_{\mathbf{x}}\phi(\mathbf{x},t), \tag{17}$$

together with $\dot{\mathbf{x}}_s(t) = \mathbf{v}$, we obtain the exact trajectory-frequency relation

$$\boxed{\omega_{\mathrm{recv}}(t) = \omega_{\mathrm{loc}}\big(\mathbf{x}_s(t),t\big) - \mathbf{v}\cdot\mathbf{k}\big(\mathbf{x}_s(t),t\big)}. \tag{18}$$

Eq. 18 requires only a differentiable local phase and the receiver trajectory; no monochromatic or single-direction assumption is needed.

For the locally monochromatic plane-wave construction underlying Proposition 1, $\omega_{\mathrm{loc}}(\mathbf{x},t) = \omega_0$ and $\mathbf{k}(\mathbf{x},t) = \mathbf{k}$ are constant over the local observation window. Eq. 18 then reduces exactly to

$$\boxed{\omega_{\mathrm{recv}} = \omega_0 - \mathbf{v}\cdot\mathbf{k}}. \tag{19}$$

This is the relation stated in Eq. 1 of the main text.

The neural operator uses the complex baseband field $u(\mathbf{x},t) = U(\mathbf{x},t)\exp(-\mathrm{j}\omega_c t)$. Accordingly, the baseband signal sampled along the same trajectory is

$$s_{\mathrm{bb}}(t) = u(\mathbf{x}_s(t),t) = s(t)\exp(-\mathrm{j}\omega_c t), \tag{20}$$

with instantaneous angular frequency

$$\omega_{\mathrm{recv}}^{\mathrm{bb}}(t) = \omega_{\mathrm{recv}}(t) - \omega_c = \big[\omega_{\mathrm{loc}}\big(\mathbf{x}_s(t), t\big) - \omega_c\big] - \mathbf{v} \cdot \mathbf{k}\big(\mathbf{x}_s(t), t\big). \tag{21}$$

Carrier demodulation therefore removes only the common carrier term; the motion-induced contribution $-\mathbf{v} \cdot \mathbf{k}$ is unchanged.

Eq. 18 describes how receiver motion contributes to the received frequency. To establish non-identifiability, §A.2 constructs a plane-wave family in which distinct $(\omega_0, \mathbf{v})$ pairs yield identical local temporal traces through the same value of $g(\omega_0, \mathbf{v})$. This establishes ambiguity within the constructed family.

Narrow-band premise. The experimental source has fractional bandwidth $b_{\mathrm{frac}} = 0.05$, supporting the locally monochromatic approximation used in the constructive subfamily below. This narrow-band condition does not imply a single propagation direction after scattering. The experimental wavefields contain multiple interfering components(Appendix B), and the scope of the local argument beyond the single-direction construction is clarified in §A.4.

A.2 Constructive non-injectivity of the observation map

Consider an admissible locally monochromatic, nondispersive plane-wave component expressed relative to the receiver's reference position:

$$U(\mathbf{x}, t) = A \exp\left(\mathrm{j}\left[\phi_0 + \omega_0 t - \frac{\omega_0}{c_p}\hat{\mathbf{k}} \cdot (\mathbf{x} - \mathbf{x}_0)\right]\right), \tag{22}$$

where $A > 0$, $\phi_0 \in \mathbb{R}$, $\|\hat{\mathbf{k}}\| = 1$, and c_p is the local phase speed of this nondispersive plane-wave subfamily. Its wavevector is

$$\mathbf{k} = \frac{\omega_0}{c_p}\hat{\mathbf{k}}. \tag{23}$$

The quantities A, ϕ_0, $\hat{\mathbf{k}}$, and c_p are held fixed throughout the following construction; only ω_0 and $\mathbf{v}$ are varied.

Sampling Eq. 22 along $\mathbf{x}_s(t) = \mathbf{x}_0 + \mathbf{v}t$ gives

$$\begin{aligned} s(t; \omega_0, \mathbf{v}) &= U(\mathbf{x}_0 + \mathbf{v}t, t) && (24)\\ &= A \exp\left(\mathrm{j}\left[\phi_0 + \omega_0\left(1 - \frac{\mathbf{v} \cdot \hat{\mathbf{k}}}{c_p}\right)t\right]\right) && (25)\\ &= A \exp\left(\mathrm{j}\left[\phi_0 + g(\omega_0, \mathbf{v})t\right]\right), && (26) \end{aligned}$$

where

$$g(\omega_0, \mathbf{v}) = \omega_0\left(1 - \frac{\mathbf{v} \cdot \hat{\mathbf{k}}}{c_p}\right). \tag{27}$$

Carrier demodulation multiplies all such signals by the same factor $\exp(-\mathrm{j}\omega_c t)$ and therefore does not alter the non-injectivity argument.

To construct a distinct parameter pair producing the same signal, decompose the velocity as

$$\mathbf{v} = \mathbf{v}_\perp + q\hat{\mathbf{k}}, \qquad q = \mathbf{v} \cdot \hat{\mathbf{k}}, \qquad \mathbf{v}_\perp \cdot \hat{\mathbf{k}} = 0. \tag{28}$$

For any $\omega_0' \neq \omega_0$ sufficiently close to ω_0, define

$$q' = c_p - \frac{\omega_0}{\omega_0'}(c_p - q), \qquad \mathbf{v}' = \mathbf{v}_\perp + q'\hat{\mathbf{k}}. \tag{29}$$

Then

$$\begin{aligned} g(\omega_0', \mathbf{v}') &= \omega_0'\left(1 - \frac{q'}{c_p}\right) && (30)\\ &= \omega_0\left(1 - \frac{q}{c_p}\right) = g(\omega_0, \mathbf{v}). && (31) \end{aligned}$$

If $(\omega_0, \mathbf{v})$ lies in the interior of the admissible parameter domain, continuity ensures that $\mathbf{v}'$ also remains admissible for ω_0' sufficiently close to ω_0. Because A, ϕ_0, $\hat{\mathbf{k}}$, and c_p are unchanged, the two distinct parameter pairs produce exactly the same temporal signal:

$$s(t; \omega_0', \mathbf{v}') = s(t; \omega_0, \mathbf{v}), \qquad \forall\, t \in \mathcal{T}. \tag{32}$$

To state the result formally, let

$$\boldsymbol{\theta} = (\omega_0, \mathbf{v}^\top)^\top \tag{33}$$

and define the local observation map by

$$[\mathcal{F}(\boldsymbol{\theta})](t) = s(t; \boldsymbol{\theta}), \qquad t \in \mathcal{T}. \tag{34}$$

The construction above gives

$$\boxed{\boldsymbol{\theta}' \neq \boldsymbol{\theta}, \qquad g(\boldsymbol{\theta}') = g(\boldsymbol{\theta}) \Longrightarrow \mathcal{F}(\boldsymbol{\theta}') = \mathcal{F}(\boldsymbol{\theta})}. \tag{35}$$

Hence, the observation map restricted to this plane-wave subfamily is non-injective, even when the amplitude, initial phase, propagation direction, and phase speed are held fixed. Because this subfamily is contained in the broader admissible field class, non-injectivity of the restricted map is sufficient to establish non-injectivity of the observation operator over the full class.

The ambiguity also includes velocity components orthogonal to $\hat{\mathbf{k}}$. For any sufficiently small $\boldsymbol{\eta}$ satisfying

$$\boldsymbol{\eta} \cdot \hat{\mathbf{k}} = 0 \tag{36}$$

and such that $\mathbf{v} + \boldsymbol{\eta}$ remains admissible,

$$g(\omega_0, \mathbf{v} + \boldsymbol{\eta}) = g(\omega_0, \mathbf{v}). \tag{37}$$

Thus, within this plane-wave subfamily, the local signal contains no information about the $d - 1$ velocity components orthogonal to the propagation direction. This additional orthogonal ambiguity is not needed for the main non-injectivity construction, which already confounds ω_0 with the velocity component parallel to $\hat{\mathbf{k}}$.

The tangent directions of a level set of g satisfy

$$\left(1 - \frac{\mathbf{v} \cdot \hat{\mathbf{k}}}{c_p}\right) \delta\omega_0 - \frac{\omega_0}{c_p} \hat{\mathbf{k}} \cdot \delta\mathbf{v} = 0. \tag{38}$$

Since $\|\mathbf{v}\| < c_p$ and $\omega_0 > 0$ imply $\nabla g \neq 0$, the implicit function theorem gives a locally d-dimensional level set in the $(d+1)$-dimensional parameter space. At low speed, $\|\mathbf{v}\|/c_p \ll 1$, Eq. 38 reduces at leading order to

$$\delta\omega_0 \approx \frac{\omega_0}{c_p} \hat{\mathbf{k}} \cdot \delta\mathbf{v} = \mathbf{k} \cdot \delta\mathbf{v}. \tag{39}$$

Role of the unknown wave frequency. If ω_0, c_p, and $\hat{\mathbf{k}}$ were known, the received frequency would determine the velocity component parallel to the propagation direction:

$$\mathbf{v} \cdot \hat{\mathbf{k}} = c_p \left(1 - \frac{\omega_{\mathrm{recv}}}{\omega_0}\right). \tag{40}$$

The orthogonal velocity components would nevertheless remain unidentified unless additional geometric information were available. In the inverse setting considered here, ω_0 is also unknown, so even the parallel component is confounded with the wave frequency. Equivalently, the complete local plane-wave trace constrains $(\omega_0, \mathbf{v})$ only through the scalar $g(\omega_0, \mathbf{v})$.

A.3 Fisher information, rank, and estimability

Let

$$\boldsymbol{\theta} = (\omega_0, \mathbf{v}^\top)^\top \in \mathbb{R}^{d+1}.$$

Conditioning on the fixed A, ϕ_0, $\hat{\mathbf{k}}$, and c_p used in §A.2, consider any regular stochastic observation model for this local subfamily whose likelihood depends on $\boldsymbol{\theta}$ only through $g(\boldsymbol{\theta})$ in Eq. 27:

$$p(\mathbf{y} \mid \boldsymbol{\theta}) = \widetilde{p}\big(\mathbf{y} \mid g(\boldsymbol{\theta})\big). \tag{41}$$

This includes additive parameter-independent noise applied to the local temporal signal. Writing

$$\ell(\boldsymbol{\theta}; \mathbf{y}) = \log p(\mathbf{y} \mid \boldsymbol{\theta}),$$

the score is

$$\nabla_{\boldsymbol{\theta}} \ell = \frac{\partial \ell}{\partial g}\, \mathbf{a}, \qquad \mathbf{a} = \nabla_{\boldsymbol{\theta}} g = \begin{pmatrix} 1 - \dfrac{\mathbf{v} \cdot \hat{\mathbf{k}}}{c_p} \\ -\dfrac{\omega_0}{c_p} \hat{\mathbf{k}} \end{pmatrix}. \tag{42}$$

The Fisher information matrix is therefore

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}} \left[\nabla_{\boldsymbol{\theta}} \ell \, \nabla_{\boldsymbol{\theta}} \ell^{\top} \right] = \kappa(\boldsymbol{\theta})\, \mathbf{a}\mathbf{a}^{\top}, \tag{43}$$

where

$$\kappa(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}} \left[\left(\frac{\partial \ell}{\partial g} \right)^2 \right]. \tag{44}$$

For $0 < \kappa(\boldsymbol{\theta}) < \infty$ and $\omega_0 > 0$, $\mathbf{a} \neq 0$, so $\mathcal{I}(\boldsymbol{\theta})$ has rank one. Its null space has dimension d and consists of perturbations

$$\boldsymbol{\delta} = (\delta\omega_0, \delta\mathbf{v}^{\top})^{\top}$$

satisfying

$$\mathbf{a}^{\top}\boldsymbol{\delta} = \left(1 - \frac{\mathbf{v} \cdot \hat{\mathbf{k}}}{c_p} \right) \delta\omega_0 - \frac{\omega_0}{c_p} \hat{\mathbf{k}} \cdot \delta\mathbf{v} = 0, \tag{45}$$

which recovers Eq. 38. Thus, the data identify only $g(\boldsymbol{\theta})$ and functions that are constant on its level sets. Locally, the gradient of any identifiable scalar functional must lie in $\mathrm{range}(\mathcal{I}) = \mathrm{span}\{\mathbf{a}\}$.

Implications of the singular Fisher information. Because $\mathcal{I}(\boldsymbol{\theta})$ is singular, the standard inverse-form Cramér–Rao bound is unavailable for the full unconstrained parameter vector. The Moore–Penrose pseudoinverse may be used in generalized bounds only for identifiable functionals satisfying the appropriate range condition, or after valid parameter constraints are imposed. Its zero action on the null space must not be interpreted as zero uncertainty in non-identifiable directions.

The stronger conclusion follows directly from non-identifiability. If $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$ but $g(\boldsymbol{\theta}') = g(\boldsymbol{\theta})$, then

$$p(\mathbf{y} \mid \boldsymbol{\theta}') = p(\mathbf{y} \mid \boldsymbol{\theta}). \tag{46}$$

Any estimator must consequently have the same distribution, and hence the same expectation whenever it exists, under these two parameter values, whereas unbiased estimation of the full parameter vector would require the expectations to equal two distinct parameter values. Hence no estimator can be globally unbiased for the full $(\omega_0, \mathbf{v})$ over this parameter space.

Additional observations governed by the same g can increase the nonzero Fisher eigenvalue but cannot increase its rank. Data identifiability requires measurements with independent sensitivity directions, such as spatial field observations, multiple non-collinear propagation directions, or time-varying geometry. Priors may regularize the estimation problem, and sufficiently strong constraints can restore identifiability on a restricted parameter space, but neither changes the non-injectivity of the original unconstrained data model.

A.4 Local extension and scope beyond plane waves

The plane-wave construction in §A.2 provides a sufficient counterexample to injectivity over the admissible field class. It does not imply that every multipath observation has a rank-one Fisher information matrix.

At any point $(\mathbf{x}^*, t^*)$ where the analytic field is nonzero and its phase is differentiable, the phase admits the local spatiotemporal expansion

$$\begin{aligned} \phi(\mathbf{x}, t) = \phi(\mathbf{x}^*, t^*) + \omega_*(t - t^*) - \mathbf{k}_* \cdot (\mathbf{x} - \mathbf{x}^*) \\ + o(\|\mathbf{x} - \mathbf{x}^*\| + |t - t^*|)\,, \end{aligned} \tag{47}$$

where

$$\omega_* = \partial_t \phi(\mathbf{x}^*, t^*), \qquad \mathbf{k}_* = -\nabla_{\mathbf{x}} \phi(\mathbf{x}^*, t^*). \tag{48}$$

For a constant-velocity trajectory satisfying $\mathbf{x}_s(t^*) = \mathbf{x}^*$, substituting the trajectory into Eq. 47, equivalently evaluating Eq. 18 at $(\mathbf{x}^*, t^*)$, gives

$$\frac{d}{dt}\phi(\mathbf{x}_s(t), t)\bigg|_{t=t^*} = \omega_* - \mathbf{v} \cdot \mathbf{k}_*. \tag{49}$$

Thus, for any differentiable local phase, the received phase rate contains the same additive combination of the local temporal phase rate and the motion-induced term. This pointwise relation alone does not imply non-identifiability of a complete multipath observation. Over a short window in which $\omega_{\mathrm{loc}}(\mathbf{x}, t)$ and $\mathbf{k}(\mathbf{x}, t)$ vary slowly, the linearized phase in Eq. 47 is well approximated by a locally monochromatic plane-wave phase.

Multiple separately resolved components or time-varying wavevectors can provide additional scalar relations. Their aggregate Fisher information may have rank greater than one when the associated sensitivity vectors are linearly independent and their evolution is sufficiently constrained. We therefore do not claim that every multipath trace has a globally rank-one Fisher information matrix. Proposition 1 instead establishes that the observation map is not universally injective over the stated field class: its plane-wave subfamily already contains continuous sets of distinct $(\omega_0, \mathbf{v})$ pairs producing identical local observations.

The FDTD experiments contain multiple interfering wavevector components (Appendix B) and are therefore not restricted to the single-component subfamily used in the construction. WTD itself makes no single-direction assumption. It predicts the complete superposed Eulerian field and uses the known receiver trajectory only through the geometric projector $\Pi_{\mathbf{v}}$.

A.5 Connection to passive-motion observability and empirical diagnostics

The non-injectivity derived above parallels classical observability limits in passive tracking. For example, bearings-only target-motion analysis requires sufficient receiver motion and measurement geometry to recover all kinematic states (Nardone & Aidala, 1981). Our measurement model, observations, and unknown parameters differ from that setting, so this connection serves as a conceptual precedent rather than a direct application of the bearings-only result. In our local single-component construction, the analogous loss of identifiability is expressed explicitly by the rank-one Fisher information matrix in Eq. 43.

Empirical diagnostic. The learned-canonicalization results are consistent with this local ambiguity. The supervised and end-to-end variants use the same frozen WTD field operator, but their OOD velocity estimates remain biased toward the training range (§4.6, Fig. 6). These results show that the tested canonicalizers do not reliably recover OOD velocities. They do not, however, prove non-identifiability for arbitrary multipath observations or extend the rank-one Fisher result beyond the local single-component model.

Receiver-motion model used in the experiments. The FDTD simulator evolves an Eulerian field in a fixed simulation frame. Moving-receiver traces are then obtained by bilinearly sampling this field along

$$\mathbf{x}_s(t) = \mathbf{x}_0 + \mathbf{v}t. \tag{50}$$

For a continuous local plane-wave component, this construction gives the moving-receiver relation derived in Eq. 19,

$$\omega_{\mathrm{recv}} = \omega_0 - \mathbf{v} \cdot \mathbf{k}. \tag{51}$$

The discrete implementation approximates the corresponding continuous trajectory sampling on the spatial grid and therefore inherits ordinary spatial-discretization and bilinear-interpolation errors. The experiments therefore model a moving receiver sampling a field in a fixed frame; no equivalence with a physically moving source is assumed. The resulting discrete sampling relation and its error decomposition are derived in §A.6.

A.6 DISCRETE TRAJECTORY SAMPLING AND ERROR DECOMPOSITION

Let $t_n = n\Delta t_{\text{snap}}$ denote the snapshot times. For the following interpolation analysis, let U_h^n denote the restriction of an ideal continuous Eulerian field $U(\mathbf{x}, t_n)$ to a uniform spatial grid with spacing Δx. The neural operator predicts the corresponding complex baseband field, denoted by $\widehat{u}_h^n$. For the analytic-field derivation below, we equivalently define the remodulated prediction

$$\widehat{U}_h^n = \widehat{u}_h^n \exp(\mathrm{j}\omega_c t_n). \tag{52}$$

Let

$$\boldsymbol{\beta} = \frac{\mathbf{v}}{c} \tag{53}$$

be the normalized receiver velocity reported in the experiments. Here, c denotes the reference wave speed used to normalize the experimental velocities and is distinct from the local phase speed c_p used in the constructive analysis of §A.2. Let $\mathbf{r}_0$ denote the grid-coordinate representation of the initial receiver position $\mathbf{x}_0$. The receiver position in grid coordinates is

$$\mathbf{r}_n = \mathbf{r}_0 + \frac{\mathbf{v}\Delta t_{\text{snap}}}{\Delta x} n = \mathbf{r}_0 + \alpha_\beta \boldsymbol{\beta} n, \tag{54}$$

where

$$\alpha_\beta = \frac{c\Delta t_{\text{snap}}}{\Delta x} = c\alpha. \tag{55}$$

When the simulation is nondimensionalized with $c = 1$, this reduces to the trajectory expression used in the implementation.

Let $\mathcal{I}_h$ denote bilinear interpolation on the spatial grid. The ideal continuous signal, the signal interpolated from the reference grid samples, and the implemented prediction are respectively

$$s[n] = U(\mathbf{x}_0 + \mathbf{v}t_n, t_n)\,, \tag{56}$$

$$s_h[n] = \mathcal{I}_h[U_h^n](\mathbf{r}_n), \tag{57}$$

$$\widehat{s}[n] = \mathcal{I}_h[\widehat{U}_h^n](\mathbf{r}_n). \tag{58}$$

Their difference relative to the ideal continuous signal admits the exact decomposition

$$\begin{aligned} \widehat{s}[n] - s[n] = &\underbrace{\mathcal{I}_h\big[\widehat{U}_h^n - U_h^n\big](\mathbf{r}_n)}_{\epsilon_{\text{field}}[n]} \\ &+ \underbrace{\mathcal{I}_h[U_h^n](\mathbf{r}_n) - U(\mathbf{x}_0 + \mathbf{v}t_n, t_n)}_{\epsilon_{\text{int}}[n]}\,. \end{aligned} \tag{59}$$

Here, $\epsilon_{\text{field}}[n]$ is the field-prediction error sampled along the trajectory, whereas $\epsilon_{\text{int}}[n]$ is the deterministic interpolation error between the reference grid and the ideal continuous observation.

For the discrete learning problem used in the experiments, the reference trajectory signal is $s_h[n]$, because the target traces are obtained by applying the same bilinear sampler to the reference FDTD field. Therefore,

$$\boxed{\widehat{s}[n] - s_h[n] = \mathcal{I}_h\big[\widehat{U}_h^n - U_h^n\big](\mathbf{r}_n) = \epsilon_{\text{field}}[n].} \tag{60}$$

Thus, $\epsilon_{\text{int}}[n]$ measures the discrepancy between the discrete reference signal and an ideal continuous-field observation; it is not an additional learned error introduced by WTD relative to the dataset target.

Because the bilinear interpolation weights are nonnegative and sum to one,

$$|\epsilon_{\text{field}}[n]| \le \left\| \widehat{U}_h^n - U_h^n \right\|_{\ell_\infty(K_n)}, \tag{61}$$

where K_n is the grid cell containing $\mathbf{r}_n$. If $U(\cdot, t_n)$ has bounded second spatial derivatives on K_n, standard bilinear-interpolation estimates give

$$|\epsilon_{\text{int}}[n]| \le C\Delta x^2 \max_{|\gamma|=2} \|\partial_{\mathbf{x}}^\gamma U(\cdot, t_n)\|_{L^\infty(K_n)} \tag{62}$$

for a grid-dependent constant C. For a plane wave, this error scales as $\mathcal{O}((\|\mathbf{k}\|\Delta x)^2)$, illustrating why sufficient spatial resolution remains important for high-frequency fields.

For the ideal continuous local plane wave

$$U(\mathbf{x}, t) = A \exp(\mathrm{j}[\omega_0 t - \mathbf{k} \cdot \mathbf{x}]), \tag{63}$$

Eq. 59 becomes

$$\begin{aligned} \widehat{s}[n] = A\mathrm{e}^{-\mathrm{j}\mathbf{k}\cdot\mathbf{x}_0} \exp(\mathrm{j}(\omega_0 - \mathbf{k} \cdot \mathbf{v})n\Delta t_{\mathrm{snap}}) \\ + \epsilon_{\mathrm{field}}[n] + \epsilon_{\mathrm{int}}[n]. \end{aligned} \tag{64}$$

The ideal continuous term therefore has the discrete phase increment

$$\Delta\phi_{\mathrm{recv}} = (\omega_0 - \mathbf{k} \cdot \mathbf{v})\Delta t_{\mathrm{snap}}. \tag{65}$$

For the baseband field $u(\mathbf{x}, t) = U(\mathbf{x}, t)\mathrm{e}^{-\mathrm{j}\omega_c t}$, the corresponding increment is

$$\Delta\phi^{\mathrm{bb}}_{\mathrm{recv}} = \big[(\omega_0 - \omega_c) - \mathbf{k} \cdot \mathbf{v}\big]\Delta t_{\mathrm{snap}}. \tag{66}$$

Carrier demodulation therefore changes only the temporal frequency and leaves the motion-induced term unchanged.

For this single-component baseband plane wave, the discrete phase increment is defined modulo 2π. Unaliased recovery of its temporal frequency therefore requires

$$|(\omega_0 - \omega_c) - \mathbf{k} \cdot \mathbf{v}| < \frac{\pi}{\Delta t_{\mathrm{snap}}}. \tag{67}$$

Velocity enters the projector only through the deterministic coordinates $\mathbf{r}_n$. The projector contains no trainable parameters and does not learn a velocity-to-signal mapping. Relative to the discrete reference signal used in the experiments, its prediction error is therefore the predicted Eulerian field error sampled along the prescribed trajectory.

B Spatial Spectral Complexity and Directional Content

We characterize the spatial spectral content of the simulated wavefields using a dataset-averaged radial spectrum and the windowed two-dimensional spectrum of a representative source-excluded local patch. These diagnostics are computed from real-valued FDTD snapshots E_z, rather than from the complex analytic or baseband representation introduced in Appendix A.

Dataset-averaged radial spectrum. We uniformly select 30 independently generated field realizations and six snapshots from each realization between 25% and 90% of the stored temporal range, yielding $N_s = 180$ snapshots. Snapshots from the same realization are not treated as statistically independent samples.

For each snapshot, we subtract its spatial mean and apply a separable two-dimensional Hann window before computing the spatial Fourier transform $\widehat{E}_z^{(n)}(\mathbf{m})$, where $\mathbf{m} = (m_x, m_y)$ denotes the Cartesian FFT-bin index. We define the radial bin coordinate

$$\rho = \|\mathbf{m}\|_2 \tag{68}$$

and the unit-width annulus

$$\mathcal{A}_\rho = \left\{ \mathbf{m} : \rho - \frac{1}{2} \le \|\mathbf{m}\|_2 < \rho + \frac{1}{2} \right\}. \tag{69}$$

Figure 5(a) shows the dataset-averaged radial amplitude

$$A(\rho) = \frac{1}{N_s} \sum_{n=1}^{N_s} \frac{1}{|\mathcal{A}_\rho|} \sum_{\mathbf{m} \in \mathcal{A}_\rho} \left| \widehat{E}_z^{(n)}(\mathbf{m}) \right|. \tag{70}$$

We separately compute the pooled cumulative spectral power

$$C(\rho) = \frac{\displaystyle\sum_{n=1}^{N_s} \sum_{\|\mathbf{m}\|_2 \le \rho} \left| \widehat{E}_z^{(n)}(\mathbf{m}) \right|^2}{\displaystyle\sum_{n=1}^{N_s} \sum_{\mathbf{m}} \left| \widehat{E}_z^{(n)}(\mathbf{m}) \right|^2}. \tag{71}$$

The 90% cumulative-power radius is defined by

$$\rho_{90} = \min\{\rho : C(\rho) \geq 0.9\}. \tag{72}$$

For the 256×256 grid with eight pixels per nominal carrier wavelength, the nominal background-medium carrier wavenumber corresponds to the radial FFT-bin coordinate $\rho_c = 32$. We use $\rho_{\mathrm{Nyq}} = 128$ to denote the radius of the largest circle contained in the Cartesian Nyquist square. Thus,

$$\rho_c = 0.25\,\rho_{\mathrm{Nyq}}. \tag{73}$$

The measured cumulative-power radius is

$$\rho_{90} = 75 \approx 0.59\,\rho_{\mathrm{Nyq}} = 2.34\,\rho_c. \tag{74}$$

Thus, non-negligible spatial spectral content extends well beyond the nominal background-medium carrier radius. This broad spatial bandwidth makes accurate representation of spatial variations particularly important and motivates the spatial-gradient supervision used by WTD.

Local two-dimensional spatial spectrum. The proof of Proposition 1 uses a single plane-wave subfamily as a sufficient counterexample to injectivity; it does not assume that the experimental fields contain only one propagation direction. The FDTD data are more complex because random dielectric scatterers generate reflected and interfering components within the source-excluded local field patches.

For Figure 5(b), we subtract the spatial mean from a representative 75×75 patch and apply a separable two-dimensional Hann window to reduce boundary leakage. The normalized spatial power is displayed as

$$P_{\mathrm{dB}}(m_x, m_y) = 10\log_{10}\frac{\left|\widehat{E}_z(m_x, m_y)\right|^2}{\max\limits_{m_x, m_y}\left|\widehat{E}_z(m_x, m_y)\right|^2}. \tag{75}$$

Values below $-30\,\mathrm{dB}$ are clipped for visualization. The windowed patch is zero-padded by a factor of four to interpolate the displayed spectrum; zero-padding does not increase its physical spectral resolution. The plotted axes are expressed in FFT-bin units of the original unpadded patch. All quantitative spectral measurements are computed from unpadded transforms.

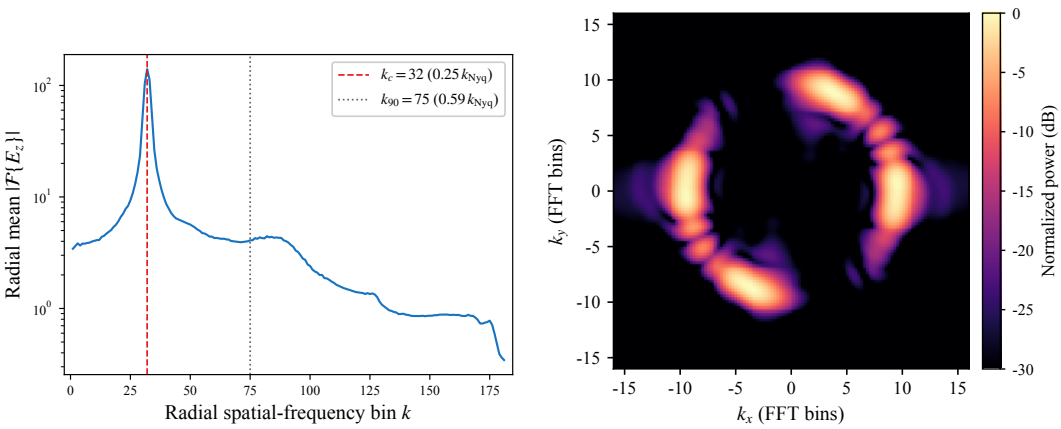
Because the analyzed FDTD snapshot is real valued, its spatial spectrum is Hermitian. A single real plane wave multiplied by the Hann window would produce one conjugate pair of broadened spectral lobes. The observed spectrum instead contains several angularly separated lobe pairs and extended angular structure at comparable radial frequencies. This pattern is consistent with the simultaneous presence of multiple spatial wavevector orientations. It should not be interpreted as a unique modal decomposition or as a continuous isotropic energy ring.

These diagnostics confirm that the simulated dataset includes multi-component spatial structure and is not restricted to the single-direction subfamily used in the non-injectivity construction. WTD itself makes no single-direction assumption: its velocity-agnostic operator predicts the superposed Eulerian field, and the geometric projector samples that field along the known receiver trajectory. The results in Section 4 show that this separation remains effective in the presence of such multi-component spatial structure.

C Implementation and Training Details

Table 3 summarizes the model architecture, loss coefficients, training settings, and evaluation protocol used in the primary WTD experiments. Unless otherwise stated, controlled ablations retain this configuration and modify only the settings specified for each ablation. Dataset-generation settings and physical constants are provided in Appendix E.

Compared models. WTD predicts a velocity-independent future complex Eulerian field and obtains the future 1D signal by sampling the field along the known trajectory using a parameter-free bilinear projector.



(a) Dataset-averaged radial spectrum **(b)** Local 2D spatial power spectrum

Figure 5: **Spatial spectrum of the simulated wavefields.** **(a)** Dataset-averaged radial amplitude over 180 Hann-windowed snapshots from 30 independently generated field realizations. The nominal background-medium carrier radius is $\rho_c = 32 = 0.25\,\rho_{\mathrm{Nyq}}$, and the 90% cumulative-power radius is $\rho_{90} = 75 \approx 0.59\,\rho_{\mathrm{Nyq}} = 2.34\rho_c$. **(b)** Normalized spatial power spectrum of a representative source-excluded 75×75 patch. Multiple angularly separated conjugate lobe pairs are consistent with simultaneous spatial wavevector orientations.

FiLM (plain) (Perez et al., 2018) instead uses the known velocity to feature-wise modulate the real and imaginary components of its latent representation and directly predicts the future 1D signal through spatial pooling and a signal head.

FiLM + field retains this direct signal-prediction pathway and adds an auxiliary, velocity-conditioned field head supervised by the same future-field value and spatial-gradient losses used for WTD; the auxiliary field is not used for trajectory sampling.

FiLM + LPSDA adapts Lie point symmetry data augmentation (Brandstetter et al., 2022) by sampling additional velocities within the training range and online resampling the stored Eulerian fields along the corresponding trajectories.

Cond. concat broadcasts the velocity components as two additional input channels, whereas Cond. adaLN uses velocity-dependent scale, shift, and residual gating following the adaLN-Zero design (Peebles & Xie, 2023).

FNO (e2e) (Li et al., 2021) stacks the real and imaginary field components as real-valued channels, applies an FNO backbone with layer-wise FiLM velocity conditioning, and directly predicts the future 1D signal through spatial pooling.

Following learned canonicalization (Kaba et al., 2023; Shumaylov et al., 2025), the two canonicalization diagnostics estimate velocity from the observed past complex 1D signal and use the estimate to construct the sampling trajectory for the frozen WTD field operator: Canon. ($+v$ sup.) additionally receives direct velocity supervision, whereas Canon. (e2e) is trained only through the final signal-prediction objective.

The phase-preserving complex activation in Table 3 adapts magnitude-based complex nonlinearities (Arjovsky et al., 2016; Chakraborty et al., 2026), using GELU as a radial magnitude nonlinearity while preserving the input phase for nonzero activations.

Phase loss. For a batch of B predicted and target signals of length T_{out}, let $\hat{q}_{b,n}$ and $q_{b,n}$ denote the normalized adjacent products in Eq. 12, with batch index b, $n = 1, \ldots, T_{\mathrm{out}} - 1$, and $\epsilon = 10^{-8}$. The target-amplitude weights are normalized across the batch and temporal pairs:

$$w_{b,n} = \frac{|s_{b,n-1}|}{\frac{1}{B(T_{\mathrm{out}}-1)} \sum_{b'=1}^{B} \sum_{m=1}^{T_{\mathrm{out}}-1} |s_{b',m-1}| + \epsilon}. \tag{76}$$

Table 3: **Experimental configuration for WTD.**

Group	Setting	Value
Architecture	backbone width d_{model}	64
	dilation schedule	$[1, 2, 4, 8]$
	prediction stages (`num_iters`)	1
	spatial patch size	75×75
	input / forecast snapshots	$T_{\text{in}} = 40$ / $T_{\text{out}} = 20$
	theoretical receptive field	$\approx 89 \times 89$
	field normalization	max-magnitude normalization
	complex activation	phase-preserving $\frac{z}{\lvert z\rvert + \epsilon}\ \text{GELU}(\lvert z\rvert)$
Loss	complex-envelope MSE λ_{env}	1.0
	phase coefficient β_{phase}	10.0
	field coefficient λ_{field}	10.0
	gradient coefficient λ_{grad}	0.5
	energy coefficient λ_{energy}	1.0
Training	optimizer	Adam
	initial learning rate	5×10^{-4}
	weight decay	1×10^{-5}
	learning-rate schedule	cosine annealing
	training epochs	500
	batch size	2
	gradient-norm clipping	1.0
Evaluation	training / test fields	240 / 60
	reported statistic	mean $\pm$ standard deviation over three independent runs
	stationary Doppler reference	corresponding ground-truth $\mathbf{v} = 0$ trace

The phase loss is

$$\mathcal{L}_{\text{phase}} = \frac{1}{B(T_{\text{out}} - 1)} \sum_{b=1}^{B} \sum_{n=1}^{T_{\text{out}}-1} w_{b,n} \left[1 - \text{Re}\left(\hat{q}_{b,n} q_{b,n}^{*}\right)\right]. \tag{77}$$

The loss compares adjacent phase increments without phase unwrapping and assigns smaller weights to pairs with smaller preceding target amplitudes.

Computational resources. All neural-network training and evaluation runs were conducted on a single NVIDIA A100 GPGPU with 40 GB of memory.

For Doppler evaluation, the shifts of both the predicted and ground-truth moving-receiver traces are computed relative to the corresponding ground-truth $\mathbf{v} = 0$ trace. This reference affects only metric computation and does not alter model inputs or predictions.

C.1 Decoupled WTD Inference

Algorithm 1 summarizes the evaluated single-stage WTD inference procedure. Its dependency structure is

$$\widehat{\mathbf{u}}_{\text{future}} = \mathcal{G}_{\theta}\left(\mathbf{u}_{\text{past}}\right), \qquad \widehat{s}_{\mathbf{v}} = \Pi_{\mathbf{v}}\left(\widehat{\mathbf{u}}_{\text{future}}\right). \tag{78}$$

Thus, the learned operator predicts the Eulerian field independently of receiver velocity, which enters only through the parameter-free geometric projector. All primary experiments use `num_iters` $= 1$, so no autoregressive feedback or cross-stage state transfer is active.

Algorithm 1: Single-Stage WTD Inference

Input: Past real-valued field frames $E_{\text{past}} \in \mathbb{R}^{T_{\text{in}} \times P \times P}$; trained field operator $\mathcal{G}_\theta$; known velocity $\mathbf{v}$; local receiver position $\mathbf{r}_{\text{f}}$ at the first forecast snapshot; coordinate scale α
Output: Predicted complex baseband signal $\widehat{s}_{\mathbf{v}} \in \mathbb{C}^{T_{\text{out}}}$

`/* 1. Past-only analytic baseband construction */`
$\mathbf{u}_{\text{past}} \leftarrow \text{Hilbert}_t(E_{\text{past}}) \odot \exp(-\text{j}\omega_c t)$
$S \leftarrow \max_{t,\mathbf{x}} |\mathbf{u}_{\text{past}}(\mathbf{x}, t)| + \epsilon$
$\widetilde{\mathbf{u}}_{\text{past}} \leftarrow \mathbf{u}_{\text{past}} / S$
`/* 2. Eulerian field prediction without velocity input */`
$\widehat{\mathbf{u}}_{\text{future}} \leftarrow S\, \mathcal{G}_\theta\left(\widetilde{\mathbf{u}}_{\text{past}}\right)$ `// v is not provided to` $\mathcal{G}_\theta$
`/* 3. Parameter-free trajectory projection */`
for $n \leftarrow 0$ **to** $T_{\text{out}} - 1$ **do**
  $\mathbf{r}(n) \leftarrow \mathbf{r}_{\text{f}} + \alpha \mathbf{v} n$
  $\widehat{s}_{\mathbf{v}}[n] \leftarrow \text{Bilinear}\left(\widehat{\mathbf{u}}_{\text{future}}[n, :, :], \mathbf{r}(n)\right)$
return $\widehat{s}_{\mathbf{v}}$

Past-only preprocessing. The analytic representation is constructed exclusively from the observed T_{in} frames; future frames are not used. In the implementation, analytic conversion and trajectory-coordinate construction precede `Model.forward`. The encoder and decoder receive only $\mathbf{u}_{\text{past}}$, while the trajectory coordinates are supplied directly to `local_trajectory_sample`.

Velocity-agnostic complex field operator. Each sample is normalized by a positive real scale

$$S_b = \max_{t,\mathbf{x}} \left|\mathbf{u}^{(b)}_{\text{past}}(\mathbf{x}, t)\right| + \epsilon, \qquad \arg\left(\frac{\mathbf{u}^{(b)}_{\text{past}}}{S_b}\right) = \arg\left(\mathbf{u}^{(b)}_{\text{past}}\right). \tag{79}$$

The normalization therefore rescales magnitude without changing phase. No separate real- and imaginary-component normalization is applied inside the complex convolution blocks.

A block with dilation d is

$$\mathcal{B}_d(z) = \rho[\mathcal{C}_d\left(\mathcal{D}_d(z)\right)] + \mathcal{C}_{1\times1}(z), \tag{80}$$

where $\mathcal{D}_d$ is the identity for $d = 1$ and a complex depthwise convolution for $d > 1$, while $\mathcal{C}_d$ denotes a dilated complex convolution. The activation

$$\rho(z) = \frac{z}{|z| + \epsilon}\, \text{GELU}(|z|) \tag{81}$$

applies a radial nonlinearity to the magnitude while preserving the phase of nonzero complex activations. For the dilation schedule $[1, 2, 4, 8]$, the receptive-field increments of the four blocks are 2, 6, 12, and 24 pixels. Applying the same schedule in the encoder and decoder gives

$$R_{\text{WTD}} = 1 + 2(2 + 6 + 12 + 24) = 89, \tag{82}$$

so the theoretical 89×89 receptive field covers the 75×75 input patch.

Parameter-free complex trajectory projection. Let $\mathcal{I}_h$ denote bilinear interpolation on the discrete field grid. The projector computes

$$\widehat{s}_{\mathbf{v}}[n] = \mathcal{I}_h\left[\widehat{\mathbf{u}}_{\text{future}}[n]\right](\mathbf{r}_{\text{f}} + \alpha \mathbf{v} n). \tag{83}$$

PyTorch `grid_sample` applies the same coordinates and interpolation weights separately to the real and imaginary components:

$$\mathcal{I}_h[a + \text{j}b] = \mathcal{I}_h[a] + \text{j}\mathcal{I}_h[b]. \tag{84}$$

By linearity, this split computation is algebraically equivalent to complex bilinear interpolation and introduces no separate decomposition error. Its numerical approximation is limited to ordinary off-grid interpolation error. All evaluated coordinates remain inside the local patch, so border padding does not affect the reported results.

Velocity sensitivity. For a fixed past field, the predicted Eulerian field satisfies

$$\frac{\partial \widehat{\mathbf{u}}_{\mathrm{future}}}{\partial \mathbf{v}} = \mathbf{0}. \tag{85}$$

Within each bilinear interpolation cell, the chain rule gives

$$\nabla_{\mathbf{v}} \widehat{s}_{\mathbf{v}}[n] = \alpha n \, \nabla_{\mathbf{x}} \mathcal{I}_h \left[\widehat{\mathbf{u}}_{\mathrm{future}}[n] \right] (\mathbf{r}_{\mathrm{f}} + \alpha \mathbf{v} n) \, . \tag{86}$$

Hence, velocity dependence arises through the deterministic trajectory and the spatial gradients of the predicted field, rather than through a learned velocity-conditioned representation. Eq. 86 also motivates spatial-gradient supervision: errors in the predicted field gradients directly perturb the signal's sensitivity to velocity.

Multi-trajectory reuse. For M admissible trajectories sharing the same past field, WTD requires one field-operator evaluation followed by M projections:

$$\mathcal{C}_{\mathrm{WTD}}(M) = \mathcal{C}_{\mathcal{G}} + M\mathcal{C}_{\Pi}, \qquad \mathcal{C}_{\Pi} = \mathcal{O}(T_{\mathrm{out}}). \tag{87}$$

Thus, for multiple admissible trajectories sharing the same past field, the predicted Eulerian field can be reused without additional evaluations of $\mathcal{G}_\theta$; only the parameter-free geometric projections are repeated.

D Results Across Independent Runs

Table 4 provides the exact run-level OOD values underlying Figure 3(c) for the primary matched comparison between WTD and FiLM; aggregate ID results are included for reference. Across these runs, WTD's standard deviation remains small from ID to OOD (0.0008 and 0.0005, respectively), whereas FiLM's increases from 0.0009 to 0.0137. Although three runs do not characterize the full distribution over random initializations, the consistent run-level results indicate that WTD's advantage is not attributable to a single favorable initialization.

Table 4: **Carrier-normalized absolute Doppler error across independent training runs.** Individual values are shown for the OOD evaluation; aggregate ID results are included for reference. Values are mean $\pm$ sample standard deviation across three runs. Lower is better.

Model	OOD				ID
	Seed 0	Seed 1	Seed 2	Mean $\pm$ std	Mean $\pm$ std
WTD (ours)	0.0118	0.0109	0.0111	$\mathbf{0.0113 \pm 0.0005}$	$\mathbf{0.0113 \pm 0.0008}$
FiLM	0.0654	0.0685	0.0434	0.0591 ± 0.0137	0.0354 ± 0.0009

E Data Generation and Task Configuration

This appendix describes the numerical solver, random-medium generation, source configuration, moving-receiver observation, and data partition used in the primary experiments. Table 5 summarizes the fixed simulation and task configuration. All methods use the same Eulerian field realizations, trajectory samples, and train–test partition.

Numerical solver and temporal sampling. Ground-truth Eulerian wavefields are generated for the TM_z components (E_z, H_x, H_y) using a two-dimensional Yee leapfrog FDTD scheme on a 256×256 Cartesian grid. The computational domain is surrounded by a 30-cell cubic-graded absorbing layer with design attenuation $R = 10^{-8}$.

We use normalized units with $c = 1$. With $\Delta x = \Delta y = 0.125$, the FDTD time step is

$$\Delta t = \frac{0.5 \, \Delta x}{c\sqrt{2}} = 0.044194. \tag{88}$$

The corresponding Courant ratio is

$$\frac{c\Delta t}{\Delta x} = \frac{0.5}{\sqrt{2}} \approx 0.353553 < \frac{1}{\sqrt{2}}, \tag{89}$$

Table 5: **Primary FDTD dataset and prediction-task configuration.**

Configuration	Value
Solver	2D Yee leapfrog FDTD, TM_z polarization
Global grid	256×256
Spatial spacing	$\Delta x = \Delta y = 0.125$
Carrier	$f_c = 1$, $\omega_c = 2\pi$, $\lambda_c = 1$
Background resolution	8 grid cells per carrier wavelength
FDTD time step	$\Delta t = 0.044194$
Courant ratio	$c\Delta t/\Delta x = 0.353553$
Absorbing boundary	30-cell cubic-graded absorbing layer; design attenuation $R = 10^{-8}$
FDTD steps / snapshots	1200 / 300
Snapshot stride	4 FDTD steps
Snapshot interval	$\Delta t_{\mathrm{snap}} = 4\Delta t = 0.176777$
Source	Point source at grid index $(45, 128)$ with 5% fractional bandwidth
receiver initial position	Grid index $(85, 128)$
Background permittivity	$\varepsilon_r = 1$
Scatterers	2–4 random rectangles per field realization
Scatterer size	Integer width and height sampled from $[8, 30]$ grid cells
Scatterer permittivity	$\varepsilon_r \sim \mathcal{U}(2, 6)$
Trajectory	Constant velocity along the positive x direction
Normalized velocity magnitudes	$\|\mathbf{v}\|/c \in \{0, 0.06, 0.08, 0.10, 0.15, 0.20\}$
ID normalized speeds	$\{0.06, 0.08, 0.10\}$
OOD normalized speeds	$\{0.15, 0.20\}$
Input / forecast length	$T_{\mathrm{in}} = 40$, $T_{\mathrm{out}} = 20$ snapshots
Local field patch	75×75 Eulerian crop centered at the initial receiver position
Field realizations	300
Train / test split	240 / 60 field realizations
Storage precision	32-bit floating point

which satisfies the two-dimensional Yee stability condition.

Each simulation contains 1200 FDTD steps. Every fourth step is stored, yielding 300 Eulerian field snapshots with temporal spacing

$$\Delta t_{\mathrm{snap}} = 4\Delta t \approx 0.176777. \tag{90}$$

With $f_c = 1$, the background carrier wavelength is $\lambda_c = c/f_c = 1$, corresponding to eight grid cells per carrier wavelength.

Random medium and narrow-band source. The background relative permittivity is $\varepsilon_r = 1$. Each field realization contains between two and four rectangular dielectric scatterers. Their integer widths and heights are sampled from $[8, 30]$ grid cells, and their relative permittivities are sampled from $\mathcal{U}(2, 6)$. A new random medium is generated for each field realization.

A point source is placed at $(45, 128)$. For each field realization, a random complex baseband envelope $s(t)$ is generated with spectral support

$$|\nu| \le \frac{b_{\mathrm{frac}} f_c}{2}, \qquad b_{\mathrm{frac}} = 0.05, \tag{91}$$

corresponding to a total fractional bandwidth of 5%. The real-valued excitation injected into the FDTD solver is

$$E_{\mathrm{src}}(t) = \mathrm{Re}\left\{ s(t) \exp(\mathrm{j}\omega_c t) \right\}. \tag{92}$$

The narrow-band setting supports the locally monochromatic construction used in Proposition 1. It does not restrict the experimental wavefields to a single propagation direction. Random dielectric scattering produces reflected and interfering components, whose spatial structure is characterized in Appendix B.

Moving-receiver observation and velocity split. For each field realization, the FDTD solver first generates a time-dependent Eulerian wavefield on the fixed spatial grid, independently of receiver velocity. Moving-receiver signals are then constructed by bilinearly sampling the stored field snapshots along prescribed constant-velocity trajectories. Consequently, the same Eulerian wavefield is shared across all evaluated velocities, while the resulting 1D Lagrangian signals differ only through trajectory sampling.

Let

$$\boldsymbol{\beta} = \frac{\mathbf{v}}{c} \tag{93}$$

denote the normalized receiver velocity. All trajectories in the primary experiments are directed along the positive x direction, so

$$\boldsymbol{\beta} = (\beta, 0). \tag{94}$$

Between consecutive stored snapshots, the receiver displacement in grid coordinates is

$$\Delta\mathbf{r} = \alpha_\beta \boldsymbol{\beta}, \qquad \alpha_\beta = \frac{c\Delta t_{\mathrm{snap}}}{\Delta x} = \sqrt{2}. \tag{95}$$

The resulting trajectory coordinates are sampled using bilinear interpolation, consistent with the parameter-free geometric projector used by WTD.

The generated normalized speed magnitudes are

$$\beta \in \{0, 0.06, 0.08, 0.10, 0.15, 0.20\}. \tag{96}$$

The case $v = 0$ serves as the stationary reference. Training and ID evaluation use

$$\mathcal{V}_{\mathrm{ID}} = \{0.06, 0.08, 0.10\}, \tag{97}$$

whereas velocity-OOD evaluation uses

$$\mathcal{V}_{\mathrm{OOD}} = \{0.15, 0.20\}. \tag{98}$$

Thus, the largest OOD speed is twice the maximum training speed, and all evaluated speeds satisfy $\|\mathbf{v}\|/c \leq 0.2$.

The velocity split varies receiver velocities and thus sampling positions at each time, without changing the underlying PDE realization. This isolates the effect of receiver velocity from changes in wavefield evolution.

Prediction task and field-level partition. Each example uses a fixed 75×75 local Eulerian crop. The first $T_{\mathrm{in}} = 40$ snapshots provide the local 2D past wavefields $\mathbf{u}_{\mathrm{past}}$, and the subsequent $T_{\mathrm{out}} = 20$ snapshots define the prediction interval. The corresponding future Eulerian wavefields $\mathbf{u}_{\mathrm{future}}$ provide dense supervision during training but are not provided to the model at inference, while the primary prediction target is the future 1D Lagrangian signal sampled along the known receiver trajectory. This matches the task formulation

$$\widehat{\mathbf{u}}_{\mathrm{future}} = \mathcal{G}_\theta(\mathbf{u}_{\mathrm{past}}), \qquad \widehat{\mathbf{s}}_{\mathbf{v}} = \Pi_{\mathbf{v}}(\widehat{\mathbf{u}}_{\mathrm{future}}). \tag{99}$$

The source lies outside the local field crop, so the model receives propagated and scattered wavefields in the local observation region rather than direct access to the excitation point.

The dataset contains 300 field realizations. The realizations are partitioned at the field level into 240 training fields and 60 held-out test fields before field–velocity examples are formed. All trajectories derived from the same Eulerian field remain in the same partition, preventing the same underlying field realization from appearing in both training and test sets.

F Evaluation Metrics and Statistical Protocol

We report three prediction metrics: the relative L_2 error of the future 1D Lagrangian signal, the absolute Doppler error, and the normalized L_2 error of the predicted Eulerian field. We additionally report the aggregation procedure and statistical tests used for the main comparisons.

F.1 Signal Relative L_2 Error

The primary signal prediction metric is the relative L_2 error over the T_{out}-step forecast trajectory:

$$E_{\text{sig}} = \frac{\|\widehat{\mathbf{s}}_{\mathbf{v}} - \mathbf{s}_{\mathbf{v}}\|_2}{\|\mathbf{s}_{\mathbf{v}}\|_2 + \epsilon}, \qquad \epsilon = 10^{-8}. \tag{100}$$

Here, $\widehat{\mathbf{s}}_{\mathbf{v}}, \mathbf{s}_{\mathbf{v}} \in \mathbb{C}^{T_{\text{out}}}$ denote the predicted and reference complex baseband Lagrangian signals, respectively. The norm is taken over the complete forecast horizon. Lower values indicate more accurate recovery of the future signal, including both its amplitude and phase.

F.2 Carrier-Normalized Doppler Recovery Error

Doppler frequency is estimated from the complex baseband trajectory signal using the classical lag-1 autocorrelation (pulse-pair) estimator (Miller & Rochwarger, 1972). For a signal $s[0], \ldots, s[T_{\text{out}} - 1]$, we compute

$$\widehat{R}_1(s) = \frac{1}{T_{\text{out}} - 1} \sum_{k=1}^{T_{\text{out}}-1} s[k]s^*[k-1], \tag{101}$$

followed by

$$\widehat{f}(s) = \frac{\angle \widehat{R}_1(s)}{2\pi \Delta t_{\text{snap}}}. \tag{102}$$

Each lag-1 product has magnitude $|s[k]||s[k-1]|$, so low-amplitude sample pairs contribute less to the accumulated correlation without requiring an additional weighting term.

Let $\mathbf{s_0}$ denote the ground-truth stationary-receiver signal for the same field realization. The target and predicted Doppler shifts are defined relative to this common stationary reference:

$$\Delta f_D = \widehat{f}(\mathbf{s}_{\mathbf{v}}) - \widehat{f}(\mathbf{s_0}), \tag{103}$$

$$\widehat{\Delta f}_D = \widehat{f}(\widehat{\mathbf{s}}_{\mathbf{v}}) - \widehat{f}(\mathbf{s_0}). \tag{104}$$

The reported Doppler metric is the carrier-normalized absolute error

$$E_D = \frac{\left|\widehat{\Delta f}_D - \Delta f_D\right|}{f_c}. \tag{105}$$

The same ground-truth stationary reference is used for all methods. This isolates Doppler recovery of the moving-receiver prediction from model-specific errors in a separately predicted stationary signal. Because $f_c = 1$ in the normalized units used in our experiments, carrier normalization does not change the numerical values.

We use absolute rather than relative Doppler error because $|\Delta f_D|$ approaches zero as the receiver speed approaches zero, making normalization by the true Doppler shift numerically unstable.

For completeness, the Doppler OOD/ID ratio reported in Table 1 is

$$R_D = \frac{\overline{E}_{D,\text{OOD}}}{\overline{E}_{D,\text{ID}}}, \tag{106}$$

where the ratio is computed from the unrounded mean errors in the two velocity regimes. A value close to one indicates that Doppler error does not increase under the evaluated velocity shift; the absolute error must be considered jointly with this ratio.

F.3 Field Reconstruction N-L_2Norm

For methods that explicitly predict the future Eulerian field, we use the averaged per-frame normalized L_2 error following the PIC2O-Sim evaluation convention (Ma et al., 2025):

$$E_{\text{field}} = \frac{1}{T_{\text{out}}} \sum_{n=0}^{T_{\text{out}}-1} \frac{\|\widehat{\mathbf{u}}_{\text{future}}[n] - \mathbf{u}_{\text{future}}[n]\|_2}{\|\mathbf{u}_{\text{future}}[n]\|_2 + \epsilon}. \tag{107}$$

The spatial L_2 norm is taken over the complete $P \times P$ complex Eulerian field at each forecast step. Equivalently, the real and imaginary components are included jointly in the Euclidean norm. The metric is reported only for methods that provide an explicit future Eulerian field aligned with the ground-truth field target.

F.4 Aggregation and Statistical Significance

Let $e_{r,i,v}$ denote a sample-level error for run r, held-out field i, and receiver speed v. For each velocity regime $\mathcal{V}$, errors within a run are first averaged over the velocities associated with each held-out field,

$$e^{\mathcal{V}}_{r,i} = \frac{1}{|\mathcal{V}|} \sum_{v \in \mathcal{V}} e_{r,i,v}, \tag{108}$$

and then averaged over the 60 held-out fields,

$$E^{\mathcal{V}}_{r} = \frac{1}{60} \sum_{i=1}^{60} e^{\mathcal{V}}_{r,i}. \tag{109}$$

Table 1 reports the mean and sample standard deviation of $E^{\mathcal{V}}_{r}$ over three independently trained runs.

For pairwise significance tests on the OOD Doppler error, the two OOD velocities are first averaged within each held-out field and the resulting field-level errors are then averaged across the three runs. This yields one OOD error per method for each of the 60 held-out fields. We apply a two-sided paired Wilcoxon signed-rank test to the 60 field-level pairs for each method comparison. Thus, the Wilcoxon tests quantify consistency of the performance difference across held-out field realizations, while variability across independent training runs is reported separately as the standard deviation over three runs.

G Qualitative Comparison with Related Work

Table 6 compares representative approaches according to how receiver kinematics enter prediction in the present velocity-OOD moving-receiver task. The comparison is mechanism-level: the relevant distinction is whether receiver motion is represented through augmented data, architectural constraints, learned conditioning or canonicalization, or an explicit trajectory-sampling operation.

Table 6: **Mechanism-level comparison of approaches to receiver kinematics.** The comparison focuses on how receiver motion enters prediction in the present velocity-OOD moving-receiver task.

Paradigm	Representative	Treatment of receiver kinematics	Limitation in the present task
Augmentation	LPSDA	Augment training data using prescribed transformations	Does not constrain behavior beyond the augmented velocity range
Equivariance	G-FNO	Encode a prescribed spatial symmetry group	Does not explicitly separate trajectory sampling from field evolution
Canonicalization	Kaba / LieLAC	Estimate an input-dependent canonical transformation	Our control requires velocity estimation from a 1D history
Conditioning	FiLM / adaLN	Inject known velocity into learned feature transformations	Learned velocity dependence must extrapolate beyond the training range
WTD	**Ours**	**Predict the Eulerian field without velocity input and sample it along the known trajectory**	**Performance depends on field prediction and deterministic trajectory sampling**

Unlike the evaluated velocity-conditioned baselines, including FiLM with trajectory augmentation, WTD does not learn how receiver velocity changes the received signal. The known velocity determines the sampling trajectory of its parameter-free geometric projector. Unlike our canonicalization controls, WTD uses the supplied receiver velocity without estimating it from the signal. Thus, trajectory sampling handles motion dependence, while the learned operator predicts Eulerian wavefields.

H Supplementary Experimental Results

H.1 Robustness to Velocity-Injection Mechanism

Table 1 compares three representative learned velocity-injection mechanisms—FiLM, direct concatenation, and adaLN—using the same dataset, local 2D inputs, and complex backbone family. Across three independently trained seeds, all three exhibit substantial OOD degradation, with OOD/ID ratios of 1.67, 3.24, and 4.76, respectively, whereas WTD remains stable at approximately 1.00. Their distinct behavior and variance indicate that the observed degradation is not specific to FiLM alone, but is shared across the tested learned-conditioning mechanisms.

H.2 Internal-Representation Velocity-Dependence Probe

We quantify how the spatial representation used for signal generation changes when the past-field input is held fixed and only the supplied receiver velocity is varied. For WTD, this representation is

the predicted future Eulerian field; for FiLM, it is the velocity-conditioned spatial feature representation immediately before signal generation.

Across 300 field–velocity comparisons derived from 60 held-out fields and five evaluated moving-receiver velocities, the FiLM representation exhibits a mean relative change of 1.21. By contrast, WTD produces identical predicted Eulerian fields across velocities up to numerical precision, as expected because velocity is excluded from $\mathcal{G}_\theta$ by construction.

This probe verifies the intended architectural distinction: changing the supplied velocity alters FiLM's learned spatial representation, whereas WTD confines velocity dependence to trajectory sampling. The probe does not imply that every velocity-conditioned architecture must fail under OOD velocities; the corresponding prediction behavior is evaluated separately in Table 1.

H.3 Cross-Velocity Gradient Interference

We further examine cross-velocity gradient alignment during optimization. For each matched training minibatch, we compute the gradient separately for each velocity condition and measure their pairwise cosine similarity, following standard gradient-conflict analyses (Yu et al., 2020),

$$C(v_i, v_j) = \frac{g_{v_i}^\top g_{v_j}}{\|g_{v_i}\|_2 \|g_{v_j}\|_2}. \tag{110}$$

Negative cosine similarity indicates conflicting local update directions. For WTD, we report both the signal-only gradient and the total gradient that additionally includes full-field supervision.

Table 7: Cross-velocity gradient alignment on matched training minibatches. Additional thresholds summarize the negative tail of the cosine-similarity distribution.

Method	Mean C	Median C	$P(C<0)$	$P(C<-0.1)$	$P(C<-0.2)$	$P(C<-0.3)$
FiLM	-0.010	-0.030	53.61%	43.06%	30.28%	20.00%
WTD (signal only)	-0.005	0.002	49.44%	27.78%	14.44%	5.83%
WTD (total)	0.004	0.011	47.50%	26.94%	14.17%	5.83%

As shown in Table 7, the negative tail differs substantially across methods. In particular, $P(C < -0.3)$ decreases from 20.00% for FiLM to 5.83% for WTD. The similar WTD signal-only and total results also indicate that dense field supervision alone does not account for this difference.

We further perform a matched continuation intervention using PCGrad (Yu et al., 2020), while keeping the remaining training conditions fixed.

Table 8: Matched continuation with and without gradient-conflict projection. OOD changes are reported relative to vanilla continuation.

Method	$P(C<0)$	$P(C<-0.3)$	Δ OOD Rel-L2	Δ OOD Doppler
Vanilla continuation	30.3%	12.6%	0	0
PCGrad continuation	0.9%	0.0%	-0.0136 (-0.99%)	-0.01096 (-13.0%)

As shown in Table 8, suppressing cross-velocity gradient conflict improves subsequent OOD performance. PCGrad reduces $P(C<0)$ from 30.3% to 0.9%, removes the $C<-0.3$ tail, and reduces OOD Doppler error by 13.0% relative to matched vanilla continuation.

Together, Tables 7 and 8 provide optimization-level evidence that WTD exhibits weaker strong cross-velocity interference, while suppressing such interference improves OOD optimization.

H.4 Field-Supervision Strength Sweep: Full Numerical Values

Table 9 reports the values underlying Fig. 4. Under the fixed architecture and initialization, increasing λ_{field} monotonically reduces OOD Doppler error, the OOD/ID ratio, and field N-L_2 over the four tested settings. The improvement from 5 to 10 is smaller than at earlier steps, indicating diminishing

returns in this controlled run. This trend links field accuracy to flatter extrapolation for the tested configuration, but does not establish a universal relationship across architectures or initializations.

Table 9: **Fixed-seed λ_{field} sweep.** OOD Doppler error, OOD/ID ratio, and field N-L_2 for the four tested supervision weights.

λ_{field}	Doppler OOD ↓	OOD/ID Ratio ↓	Field N-L_2 ↓
0	0.0404	2.30	0.8000
1	0.0279	1.58	0.1505
5	0.0159	1.13	0.1076
10	**0.0118**	**0.98**	**0.0975**

H.5 Ablation Experimental Details

Sobolev gradient ablation control. We set $\lambda_{\text{grad}} = 0$ while retaining field-value supervision with $\lambda_{\text{field}} = 10$ and keeping all other settings unchanged. This ablation isolates the contribution of spatial-gradient supervision beyond field-value supervision.

Phase-preserving activation and the ratio trap. The naive activation ablation yields an OOD/ID ratio of ≈ 1.00, which by itself could suggest flat extrapolation. However, this ratio accompanies severely degraded Eulerian field reconstruction (N-L_2 12.85) and similarly large aggregate Doppler errors in the ID and OOD regimes. In contrast, WTD-full achieves a ratio of 0.98 together with substantially lower absolute Doppler error and field N-L_2 0.0975. This comparison shows that the OOD/ID ratio must be interpreted jointly with absolute error.

H.6 Full Results of the 2×2 Factorial Study

Table 10 reports the complete 2×2 controlled comparison of velocity handling (entangled FiLM versus decoupled WTD) and supervision (signal-only versus full-field). All four cells use the same fixed training seed, simulator-generated 2D history, and matched complex-valued backbone family.

Table 10: **Architecture $\times$ field-supervision factorial study.** All four cells use the same fixed training seed. Values are absolute Doppler errors (in-band / OOD ↓), with the OOD/ID ratio in parentheses, evaluated on 60 held-out fields.

	Signal-only loss	+ Full-field supervision
FiLM (entangled)	0.0362 / 0.0654 (1.81)	0.0360 / 0.0424 (1.18)
WTD (decoupled)	0.0175 / 0.0404 (2.30)	**0.0121 / 0.0118 (0.98)**

Taking signal-only FiLM as the reference, field supervision alone reduces OOD error by 35%, while decoupling alone reduces it by 38%. Combining both reduces OOD error by 82% and yields a near-unity OOD/ID ratio of 0.98. Under matched full-field supervision, WTD achieves $\sim 3.6\times$ lower OOD error than FiLM ($p = 9.1 \times 10^{-10}$), supporting the complementary roles of structural decoupling and field supervision.

H.7 Interpreting the Doppler OOD/ID Ratio

The Doppler OOD/ID ratio is a secondary measure of how the absolute Doppler error changes under the evaluated velocity shift; it is not an absolute measure of prediction quality. In particular, a ratio near one can arise either because a model remains accurate across velocity regimes or because it performs similarly poorly in both.

The phase-activation ablation provides a concrete example. Its aggregate Doppler errors are 0.0407 in both the ID and OOD regimes, giving an OOD/ID ratio of approximately 1.00, despite a severely degraded Eulerian field (N-L_2 12.85). By contrast, WTD-full obtains a comparable ratio of 0.98 together with a much lower OOD Doppler error of 0.0118 and field N-L_2 of 0.0975. Therefore,

the OOD/ID ratio is interpreted jointly with the corresponding absolute Doppler error rather than in isolation.

The separate reason for preferring absolute over relative Doppler error is given in Appendix F.2: normalization by the true Doppler shift becomes ill-conditioned as the shift approaches zero.

I Competing Paradigms: Full Quantitative Results

This appendix provides additional implementation details and quantitative results for the augmentation and learned-canonicalization controls discussed in Section 4.6. Both controls are adapted to the present velocity-OOD moving-receiver prediction task.

I.1 Augmentation Paradigm (LPSDA)

Inspired by the symmetry-augmentation principle of Brandstetter et al. (2022), we construct a task-specific trajectory-augmentation control, labeled FiLM+LPSDA in the result tables. During training, additional receiver speeds are sampled from the ID interval $[0.06, 0.10]$. Corresponding 1D signals are obtained by resampling the stored Eulerian fields along these trajectories. This densifies trajectory coverage within the training velocity range without introducing OOD velocities or transforming the underlying PDE. The control retains FiLM's velocity-conditioned backbone and signal-level training objective; only the trajectory augmentation is added.

Table 11: **Trajectory-augmentation results.** Carrier-normalized absolute Doppler error at ID and OOD velocities. Predicted and target Doppler shifts use the same ground-truth stationary reference as defined in Appendix F.2. Values are mean $\pm$ standard deviation over three independently trained runs.

Model Configuration	ID $\downarrow$	OOD $\downarrow$	OOD/ID Ratio
FiLM (plain)	0.0354 ± 0.0009	0.0591 ± 0.0137	1.67
FiLM + LPSDA	0.0321 ± 0.0033	0.0714 ± 0.0290	2.23
WTD (Ours)	$\mathbf{0.0113 \pm 0.0008}$	$\mathbf{0.0113 \pm 0.0005}$	**1.00**

As shown in Table 11, trajectory augmentation slightly reduces the mean ID Doppler error ($0.0354 \rightarrow 0.0321$), but does not improve OOD performance: the mean OOD error increases from 0.0591 to 0.0714, with larger variation across training runs. In contrast, WTD maintains a mean Doppler error of 0.0113 in both velocity regimes. These results show that, for the tested FiLM backbone, densifying the observed trajectories within the ID velocity interval is insufficient to obtain accurate zero-shot extrapolation to the higher evaluated speeds.

Scope of augmentation. Training directly on $v \in \{0.15, 0.20\}$ would no longer constitute the zero-shot velocity-OOD protocol considered in this work. Augmentation can provide additional samples within the observed velocity interval, but it does not directly specify how the learned velocity dependence should behave beyond that interval. WTD instead uses the known receiver velocity only to determine the trajectory of the parameter-free geometric projector at inference.

I.2 Learned Canonicalization

Following the learned-canonicalization principle of Kaba et al. (2023), we construct a task-specific velocity-estimation control. A lightweight canonicalizer comprising three 1D convolutional layers operates on the real and imaginary components of the past 1D Lagrangian signal $\{s_{\mathbf{v}}[n]\}_{n=0}^{T_{\mathrm{in}}-1}$ and predicts a receiver velocity $\widehat{\mathbf{v}}$. The predicted velocity is then used to construct the trajectory along which the future Eulerian field is sampled.

For a controlled comparison, both canonicalization variants use the same pretrained WTD field operator, which remains frozen. Only the canonicalizer is trained, with gradients propagated through

trajectory sampling. The end-to-end variant uses the final signal objective, while the velocity-supervised variant additionally uses direct ID velocity supervision.

Table 12: **Velocity estimates of the learned canonicalization controls.** Values are mean ± standard deviation over three independently trained canonicalizers.

True normalized speed β	$\widehat{\beta}$ (end-to-end)	$\widehat{\beta}$ (+ velocity supervision)
0.06	0.0728 ± 0.0033	0.0712 ± 0.0050
0.08	0.0832 ± 0.0020	0.0832 ± 0.0006
0.10	0.0889 ± 0.0059	0.0917 ± 0.0055
0.15 (OOD)	$\mathbf{0.0902 \pm 0.0042}$	$\mathbf{0.0936 \pm 0.0048}$
0.20 (OOD)	$\mathbf{0.0901 \pm 0.0036}$	$\mathbf{0.0923 \pm 0.0029}$

Figure 6 provides an empirical diagnostic related to Proposition 1. Both canonicalizers respond to changes in velocity within the training interval but show little further increase once the true speed moves beyond that interval. As the normalized speed increases from $\beta = 0.10$ to $\beta = 0.20$, the end-to-end estimate remains near $\widehat{\beta} = 0.090$, while the directly supervised estimate remains near $\widehat{\beta} = 0.092$. At $\beta = 0.20$, the corresponding estimates are 0.0901 ± 0.0036 and 0.0923 ± 0.0029, respectively. The same saturation pattern is observed across the three independently trained canonicalizers, and direct ID velocity supervision does not restore accurate OOD velocity estimation.

Matched-operator control. The canonicalization controls and WTD use the same frozen velocity-agnostic field operator. Consequently, for a fixed past-field input, their predicted future Eulerian field is identical (Field N-$L_2 = 0.0975$); the difference lies in the trajectory used for sampling. WTD constructs this trajectory from the known receiver velocity, whereas the canonicalization controls construct it from $\hat{\mathbf{v}}$. The comparison therefore isolates the effect of estimating the trajectory velocity while holding Eulerian field prediction fixed.

The end-to-end and velocity-supervised canonicalizers obtain OOD Doppler errors of 0.0798 ± 0.0044 and 0.0791 ± 0.0041, respectively, compared with ID errors of 0.0192 ± 0.0007 and 0.0200 ± 0.0003. Their Doppler OOD/ID ratios are approximately 4.15 and 3.95, respectively. Thus, direct velocity supervision within the training interval provides little improvement in OOD Doppler recovery: both tested estimators remain biased toward the training velocity range as the true speed increases.

These empirical results are consistent with the local ambiguity characterized in Proposition 1: when velocity must be inferred from a limited 1D observation, accurate recovery outside the training range can be difficult. They should not be interpreted as establishing non-identifiability for arbitrary multipath observations. In the setting studied in this paper, where receiver velocity is already available, WTD avoids this additional estimation problem by supplying the known kinematics directly to the geometric projector.

J Cross-Solver Robustness: MEEP Evaluation

To test whether the observed architectural difference depends on our custom Yee FDTD implementation, we independently reproduce the prediction task using MEEP (Oskooi et al., 2010). The carrier frequency, physical geometry, snapshot interval, normalized velocity sets, and fractional bandwidth are matched to the primary FDTD setup, while the field solver and source implementation are constructed independently.

WTD and FiLM are each retrained from scratch in three independent training runs with different random seeds. We apply the same aggregation protocol as in the primary benchmark: within each run and velocity regime, errors are first averaged over the corresponding velocities for each held-out field and then over the 60 held-out fields. The resulting values for each run are reported as mean ± sample standard deviation over the three runs. The Doppler OOD/ID ratio is computed from the corresponding unrounded mean errors across the three runs, following Appendix F.2.

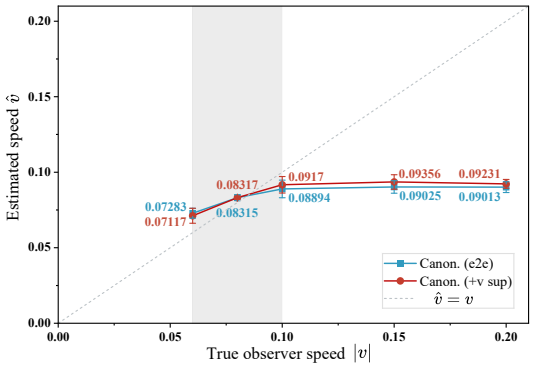


Figure 6: **OOD velocity-estimation diagnostic for learned canonicalization.** Estimated $\hat{v}$ versus true receiver speed $|v|$, averaged over three independently trained canonicalizers. Error bars denote standard deviation across runs. The shaded region marks the training velocity interval, and the dashed diagonal denotes ideal recovery $\hat{v} = v$. Both variants respond to velocity changes within the training range but saturate near its upper boundary for the evaluated OOD velocities.

Table 13: **Three-seed evaluation on the independently generated MEEP dataset.** Signal Rel-L_2 and carrier-normalized absolute Doppler errors use the same aggregation and metric definitions as Table 1. Values are mean $\pm$ sample standard deviation over three independently trained runs.

Model	Signal Rel-L_2, ID ↓	Signal Rel-L_2, OOD ↓	Doppler Error, ID ↓	Doppler Error, OOD ↓	Doppler OOD/ID
WTD	$\mathbf{0.03192 \pm 0.00040}$	$\mathbf{0.03777 \pm 0.00002}$	$\mathbf{0.00202 \pm 0.00009}$	$\mathbf{0.00197 \pm 0.00003}$	**0.98**
FiLM	1.10214 ± 0.06012	1.12782 ± 0.04451	0.01536 ± 0.00041	0.03962 ± 0.00371	2.58

Across three independent runs, WTD maintains nearly unchanged Doppler error from ID to OOD velocities on MEEP (0.00202 ± 0.00009 versus 0.00197 ± 0.00003), giving an OOD/ID ratio of 0.98. FiLM instead increases from 0.01536 ± 0.00041 to 0.03962 ± 0.00371, giving an OOD/ID ratio of 2.58. WTD therefore achieves $20.1\times$ lower mean OOD Doppler error and $29.9\times$ lower mean OOD Signal Rel-L_2 than FiLM on the independently generated MEEP dataset.

The same qualitative pattern observed with the primary FDTD solver therefore persists under an independently generated Maxwell dataset: WTD maintains a near-unity Doppler OOD/ID ratio, whereas the velocity-conditioned FiLM baseline exhibits increased Doppler error at higher OOD velocities. This cross-solver result supports the interpretation that WTD's OOD advantage is not specific to the custom Yee FDTD data generation pipeline used in the primary experiments.

WTD's Field N-L_2 on MEEP is 0.03466 ± 0.00025 across the three independent runs. This value is lower than its corresponding Field N-L_2 on the primary FDTD dataset. The lower reconstruction error is consistent with the MEEP wavefields being easier for the trained field predictor to reconstruct, but the present comparison does not isolate whether this difference arises from solver discretization, subpixel treatment, source synthesis, or other implementation details. Importantly, the MEEP source implementation is independently generated, so the two datasets should not be interpreted as differing only in their field integration engines.

For WTD, the predicted wavefields remain independent of receiver velocity by construction because velocity is not an input to $\mathcal{G}_\theta$. The MEEP experiment therefore tests whether the resulting trajectory sampling advantage persists when the underlying wavefields are produced by an independent solver, rather than altering the structural decoupling used by WTD.