



Physics-Guided Residual DCT Spectral Neural Operator for Air-Gap Response Prediction in EMS Maglev Trains

Rang Zhang¹ Jiwei Liu¹ Chenxu Lu¹ Dilai Chen¹ Qin Li²

(1.Faculty of Intelligent Technology, Shanghai Institute of Technology, Shanghai 201418, China; 2.College of Intelligent Robotics and Advanced Manufacturing, Fudan University, Shanghai, 200433 China)

*Corresponding author: Dilai Chen. E-mail: chendilai@163.com.

Abstract

Accurate and efficient prediction of levitation air-gap responses is important for the dynamic assessment of electromagnetic suspension (EMS) maglev systems. However, repeated high-fidelity vehicle–guideway–electromagnetic simulations are computationally expensive, while conventional sequence models have limited ability to preserve vibration spectral characteristics. This study proposes a physics-guided Residual Discrete Cosine Transform Spectral Neural Operator (DCT-SNO) to learn the function-to-function mapping from track irregularity to time-domain air-gap response. The model employs a real-valued DCT spectral operator with learnable weights restricted to retained low-order modes, thereby embedding the low-frequency-dominant characteristic of maglev dynamics, while a residual time-domain branch captures nonlinear response features. The model is trained and evaluated using high-fidelity UM–Simulink co-simulation data covering multiple track-irregularity spectra and operating speeds of 300–500 km/h. Results show that DCT-SNO accurately reconstructs held-out air-gap responses, preserves characteristic narrowband peaks and frequency-domain energy distributions, and achieves better spectral consistency than FNO and recurrent baselines. It also reduces inference latency by approximately one order of magnitude relative to FNO, demonstrating its potential as an efficient surrogate for repeated maglev dynamic-response evaluations.

Keywords

Maglev train, air-gap dynamics, spectral neural operator, track irregularity, discrete cosine transform

1. Introduction

The electromagnetic-suspension (EMS) maglev train-guideway system couples guideway geometry, vehicle motion, and active suspension control through the levitation air gap. Excessive gap reduction risks magnet-guideway contact, whereas excessive enlargement weakens levitation capability; the gap is therefore a key indicator of suspension safety and running quality (Huang et al., 2018; Xin et al., 2013; Zhou et al., 2010). Track irregularity is a major external excitation, with medium- and long-wavelength components strongly affecting vehicle-guideway vibration and gap fluctuation (Sheng et al., 2026; Qin et al., 2025; Jia et al., 2023; Huang et al., 2018). Efficient prediction of the air-gap response from track irregularity is consequently of direct engineering interest.

High-fidelity multibody, flexible-guideway, and electromagnetic co-simulation models capture these coupled dynamics (Jia et al., 2023; Sun et al., 2023a; Kong et al., 2011), and



stochastic methods further address irregularity and parameter uncertainty (Zhang and Yu, 2025; Lu et al., 2022; Yu et al., 2021). Their computational cost, however, limits repeated evaluation over many conditions. Data-driven models have therefore been introduced for state prediction and levitation regulation (Ren et al., 2024; Sun et al., 2023b, c), but conventional sequence models remain tied to finite-dimensional sampled inputs and outputs rather than the complete track-irregularity-to-air-gap mapping.

Neural operators provide a function-to-function alternative (Kovachki et al., 2023). DeepONet and the Fourier neural operator (FNO) are representative formulations (Li et al., 2021; Lu et al., 2021), supported by operator-approximation theory (Boullé and Townsend, 2024; Kovachki et al., 2021) and applications in structural mechanics (You et al., 2022). For the present one-dimensional vibration problem, however, classical FNO uses complex-valued Fourier operations and a generic spectral representation that does not explicitly exploit the dominant spectral characteristics of EMS air-gap dynamics.

Real-valued frequency-domain representations have been used in temporal-learning architectures (Yu et al., 2026; Jiang et al., 2023; Pan et al., 2022), while physics-informed operator variants show the value of embedding problem structure (Vo and Lee, 2025). Measurements and dynamic analyses of EMS maglev systems indicate dominant low-frequency response energy and attenuation of high-frequency disturbances (Sheng et al., 2026; Zhang et al., 2025; Xin et al., 2013). These observations motivate a real-valued operator that emphasizes retained low-order spectral content.

Accordingly, this study proposes a physics-embedded Residual Discrete Cosine Transform Spectral Neural Operator (DCT-SNO). High-fidelity UM-Simulink track-irregularity and air-gap histories are treated as input and output functions on a common grid. Each block uses a real-valued DCT kernel restricted to retained low-order modes together with a residual time-domain branch. The contributions are: (1) a function-to-function surrogate for the controlled track-irregularity-to-air-gap mapping; and (2) a residual DCT spectral operator that embeds the low-frequency prior while avoiding complex-valued spectral operations.

2. Dynamics Modeling and High-Fidelity Data Generation for Maglev

Systems

Two modeling levels are used. Section 2.1 gives a reduced-order representation to clarify the physical coupling among track irregularity, guideway/vehicle motion, electromagnetic force, and air gap. Section 2.2 describes the three-dimensional UM-Simulink closed-loop model that actually generates the learning data; the reduced equations are therefore interpretive rather than the numerical data generator.

2.1 Mechanism-Oriented Reduced-Order Representation of the Maglev Vehicle-Guideway-Electromagnetic System

The surrogate targets the vertical levitation air gap because it enters the suspension feedback loop and directly reflects levitation safety and ride quality; reconstruction of the complete vehicle state is outside the present scope.

Figure 1 gives a reduced vertical representation of the maglev train-guideway system. The carbody is treated as a rigid body with vertical and pitch motion and is supported by several levitation frames through secondary suspension elements. Levitation electromagnets are

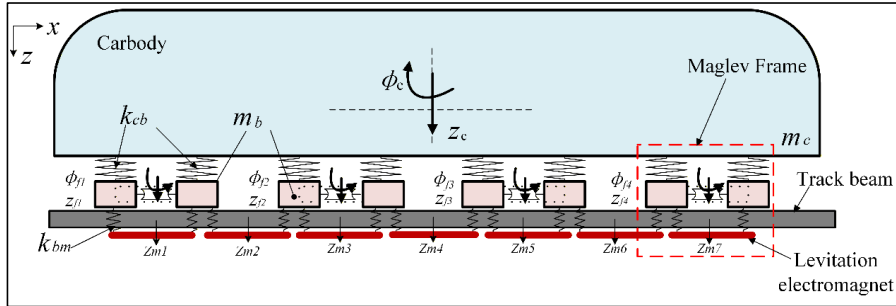


mounted beneath the frames and interact with an elastic guideway. Track irregularity is imposed
 85 as a prescribed vertical geometric excitation at the guideway, so the local air gap is determined
 by the relative displacement between the electromagnet and the guideway. The model is
 intentionally compact: it exposes the quantities that govern the vertical suspension channel
 without reproducing every spatial degree of freedom included in the three-dimensional co-
 simulation.

90 This reduced model isolates the vertical levitation channel. Lateral irregularity and the
 associated lateral, roll, and yaw responses are not DCT-SNO inputs or outputs; the full three-
 dimensional reference model is described in Section 2.2.

$$g_k(t) = z_{r,k}(t) - z_{m,k}(t) \quad (1)$$

where the electromagnet-core displacement enters the relative gap. Gap variation changes
 95 electromagnetic force and, through the suspension, the vehicle response. The resulting
 mechanism can be summarized as track irregularity → guideway motion → air-gap fluctuation
 → vehicle vibration. In the high-fidelity model, coil current is generated by the closed-loop
 controller rather than prescribed independently.



100 **Figure 1. Mechanism-oriented reduced-order vertical model of the maglev train-guideway system**

For mechanism-level interpretation, the vertical equations of the carbody, levitation frames,
 electromagnets, and guideway can be assembled in matrix form as

$$\mathbf{M}\ddot{\mathbf{z}} + \mathbf{C}\dot{\mathbf{z}} + \mathbf{K}\mathbf{z} = \mathbf{F}_{\text{em}}(g, i) + \mathbf{F}_{\text{ext}}(z_r) \quad (2)$$

$$\mathbf{F}_{\text{em}}(g, i) = \frac{\mu_0 N_e^2 A_e}{4} \left[\frac{\mathbf{I}(t)}{z_m - z_b} \right]^2 \quad (3)$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are the mass, damping, and stiffness matrices; z_m and z_b denote the
 vertical displacements of the electromagnet and the guideway beam at the corresponding
 levitation point, respectively; $\mathbf{F}_{\text{em}}$ is the levitation-force vector determined jointly by air gap
 and coil current; and $\mathbf{F}_{\text{ext}}$ represents the equivalent excitation introduced through guideway
 110 motion. Equation (3) highlights the principal nonlinear dependence of electromagnetic force on
 current and gap. This reduced relation is used to explain the feedback physics and is not the
 numerical model from which the learning samples are generated.

$$I(t) = I_{\text{bal}} + K_{sp}(z_m - z_b - s_0) + K_{sd} \frac{d(z_m - z_b)}{dt} + K_{sd} \frac{d^2(z_m - z_b)}{dt^2} \quad (4)$$



where K_{sp} and K_{sd} are the proportional and derivative gap-feedback gains, I_{bal} is the
 115 equilibrium current, and s_0 is the nominal gap. Equation (4) closes the mechanism-level link
 between the measured gap state and the current entering Eq. (3). The complete nonlinear
 electrical subsystem and suspension controller, including the voltage-current evolution used in
 data generation, remain implemented in Simulink rather than being replaced by this reduced
 expression.

120 At the co-simulation level, the instantaneous gap is passed to the controller, the coil current
 and electromagnetic force are updated, and the resulting multibody motion produces the next
 air-gap state. This bidirectional exchange closes the structural-electromagnetic-control loop.
 The main parameters of the vehicle are shown in Table 1

Table 1. Main mechanical and electromagnetic parameters of the maglev vehicle model

Name	Symbol	Value	Unit
Train tare mass (leading/end car)	m_{c1}	52900	kg
Train tare mass (middle car)	m_{c2}	50300	kg
Mass of a single levitation electromagnet	m_m	653	kg
Mass of a single levitation frame	m_b	1427	kg
Carbody pitch moment of inertia	I_c	1711025	kg·m ²
Levitation frame pitch moment of inertia	I_b	1200	kg·m ²
Carbody length (leading/end car)	L_{c1}	27.5	m
Carbody length (middle car)	L_{c2}	24.5	m
Half the center distance between adjacent levitation frames	L_b	3.096	m
Secondary suspension vertical stiffness	k_{cb}	6×10^4	N/m
Secondary suspension vertical damping	c_{cb}	3×10^4	Ns/m
Primary suspension vertical stiffness	k_{bm}	2×10^7	N/m
Primary suspension vertical damping	c_{bm}	5×10^3	Ns/m
Magnetic leakage ratio	η	0	-
Electromagnetic coil turns	N_e	356	-
Magnetic permeability of free space	μ_0	1.26×10^{-6}	H/m
Coil resistance	R_e	1	Ω
Effective pole area	A_e	0.021	m ²

125 2.2 High-Fidelity Three-Dimensional UM-Simulink Co-Simulation for Data



Generation

Reference air-gap data were generated via the 3D UM–Simulink co-simulation (Fig. 2): UM models the flexible guideway, carbody, and levitation frames under vertical track irregularities, Simulink implements the electromagnetic subsystem and controller, and two-way state exchange at each step closes the structure–electromagnet–control loop. The dataset captures full vehicle–guideway–control coupling and serves as the reference for DCT-SNO training and validation.

Accordingly, Sections 2.1 and 2.2 have different roles: the former defines the dominant mechanism and variables, while the latter provides the numerical reference data for operator learning. DCT-SNO is trained on the closed-loop UM-Simulink input-output responses and not on direct integration of Eqs. (2)–(4). The present dataset is restricted to vertical irregularity excitation and vertical air-gap response; multi-directional irregularities and multi-output vehicle-response prediction are left for future extensions.

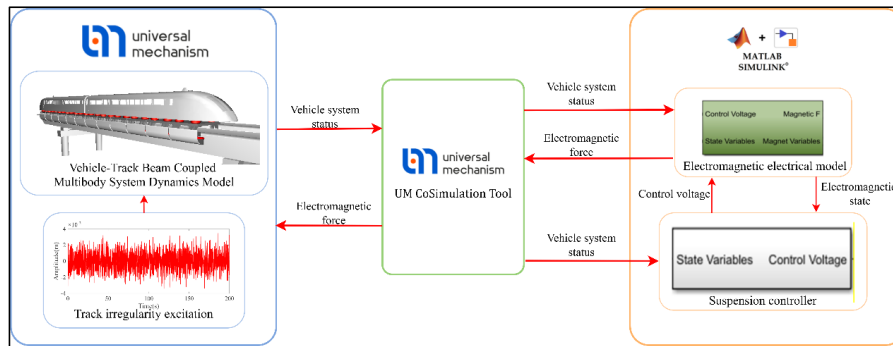
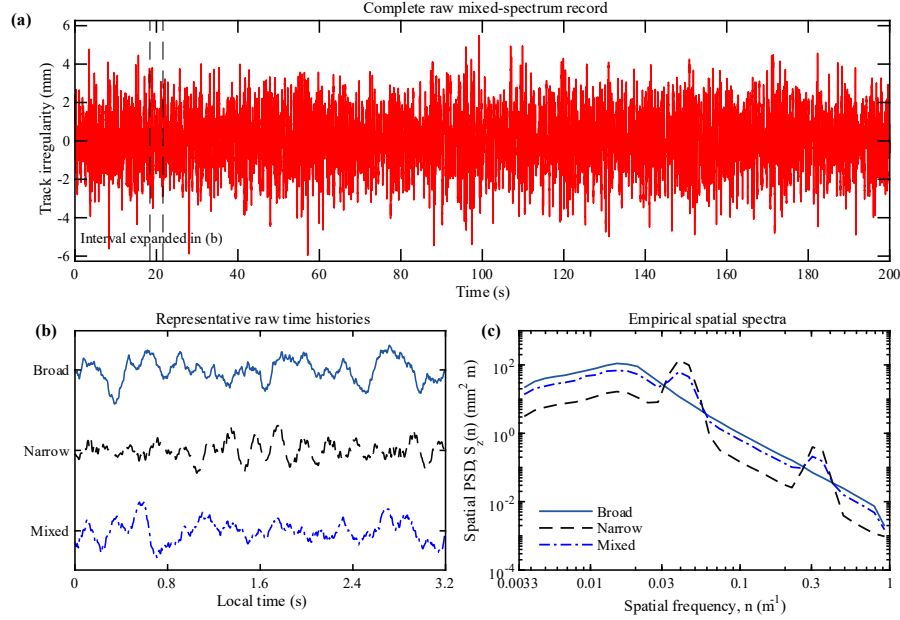


Figure 2. High-fidelity UM-Simulink closed-loop co-simulation framework used for data generation

2.3 Track-Irregularity Synthesis and Standardized Dataset Construction

Track-irregularity histories were synthesized from statistical characteristics of measured Shanghai high-speed maglev data using inverse Fourier reconstruction. Three spectral forms were retained to represent distinct excitation structures: a broadband component describing distributed background roughness, a narrowband component emphasizing periodic spectral peaks, and a mixed spectrum combining both. Figure 3 illustrates representative time histories and empirical spatial spectra.



150 **Figure 3. Construction and spectral characteristics of synthetic track-irregularity**
 155 **excitations: (a) complete mixed-spectrum record; (b) representative broadband,**
narrowband, and mixed time histories; (c) empirical spatial power spectral densities.

Figure 3 characterizes the synthesized broadband, narrowband, and mixed irregularities with spatial PSDs. A full factorial design (3 speeds $\times$ 4 amplitudes $\times$ 3 spectra) yields 36 paired
 155 1024-point sequences. Following an 85/15 split, 30 cases train the model, while the test set strictly reserves 500 km/h and $A=1.5/2.0$ across all spectral types to evaluate unseen-condition generalization. The result is a standardized dataset of co-grid track-irregularity and air-gap pairs.

3. Residual DCT-SNO

DCT-SNO learns the mapping from track irregularity to time-domain air-gap response on
 160 a standardized grid, with the output strictly in the time domain. The DCT serves solely as an internal representation within residual blocks, where features are updated via real-valued spectral kernels and returned by inverse transform, while reported spectra are post-processing diagnostics only. “Physics-embedded” refers to an architectural inductive bias rather than equation-based constraints. By restricting learnable spectral operations to the first K DCT
 165 modes via a prescribed hyperparameter K , the low-frequency dominance of maglev vibration is structurally encoded, and its effect is evaluated in Section 3.5.

3.1 Function-to-Function Mapping on a Standardized Time Grid

Let the sampled track-irregularity and air-gap histories represent functions on the common
 170 grid. The nonlinear operator G maps the irregularity function space A to the air-gap response space U .

The parameterized neural operator is trained to approximate G by minimizing

$$\min_{\theta} E_{z_r \sim \mu} \left[\|G_{\theta}(z_r) - G(z_r)\|^2 \right] \quad (5)$$



At the discrete level, the training objective becomes

$$\hat{g} = G_{\theta}(z_{r,n}) \approx g_n, \quad n = 1, 2, \dots, N \quad (6)$$

175 where N is the number of training samples. The formulation maps an entire irregularity history to an entire response history. All results use the fixed grid defined in Section 2.3; generalization to unseen temporal discretizations is not claimed.

3.2 Adaptive Normalization of Air-Gap Response Sequences

180 Air-gap sequences vary in mean level and fluctuation amplitude across operating conditions. Direct training on these unequal scales can bias optimization toward large-amplitude cases. A time-step-wise Z-score normalization is therefore applied using training-set statistics only.

For N training responses stacked in $\mathbf{Y}_{\text{train}}$, the mean and standard deviation at each grid point are computed as

$$185 \quad \mathbf{Y}_{\text{train}} = \begin{bmatrix} \mathbf{g}_r^{(1)T} \\ \vdots \\ \mathbf{g}_r^{(N)T} \end{bmatrix} \quad (7)$$

where $\mathbf{g}^{(i)}$ is the i -th air-gap sequence. For each time point, the training-set mean and standard deviation are

$$\begin{cases} \mu_{y,j} = \frac{1}{N} \sum_{i=1}^N g_r^{(i)}(t_j) \\ \sigma_{y,j} = \sqrt{\frac{1}{N} \sum_{i=1}^N (g_r^{(i)}(t_j) - \mu_{y,j})^2} \end{cases} \quad (8)$$

A small lower bound is imposed on the standard deviation to avoid numerical instability.

190 The normalized targets are

$$\mathbf{g}_r^{(i)\text{norm}} = \frac{\mathbf{g}_r^{(i)} - \boldsymbol{\mu}_y}{\boldsymbol{\sigma}_y}, \quad i = 1, \dots, N \quad (9)$$

All training losses are evaluated in this normalized response space.

3.3 Residual DCT-SNO Kernel: Physics-Guided Low-Frequency Spectral Operator

195 Measured maglev track irregularities and air-gap responses exhibit pronounced low-frequency dominance. Power-spectral analysis indicates that irregularity energy is concentrated mainly in long-wavelength components (Sheng et al., 2026), while dynamic studies show that the suspension system attenuates high-frequency disturbances and that medium- and long-wavelength excitation is more influential on the levitation gap (Zhang et al., 2025; Xin et al., 200 2013). These observations motivate an operator with an explicit low-frequency inductive bias rather than a generic full-band spectral parameterization.

DCT-II provides a real-valued orthogonal basis for this one-dimensional vibration signal. Learnable weights act only on a retained low-order mode set, whereas the classical FNO uses a complex-valued Fourier representation. The retained bandwidth is tied explicitly to the low-frequency prior and is subsequently checked by the K-sensitivity study. The comparison with 205



FNO therefore examines not only prediction error but also whether a real-valued retained-mode kernel can preserve the response-energy structure with lower computational overhead.

The architecture also differs from using DCT as a fixed preprocessing or feature-extraction step. DCT and inverse DCT are embedded inside every residual operator block, the retained coefficients are transformed by trainable kernels, and the spectral branch is optimized jointly with the nonlinear time-domain branch. Consequently, DCT coordinates form part of the learned input-output operator rather than static descriptors supplied to a conventional predictor.

For the b -th block in Figure 4, DCT-II maps the time-domain feature to spectral coefficients,

$$\hat{x}_k^{(b)} = D(\mathbf{x}^{(b)}) \in \mathbf{R}^L \quad (10)$$

with the k -th coefficient defined by

$$\hat{x}_k^{(b)} = \sum_{n=0}^{L-1} x_n^{(b)} \cos \frac{\pi k(2n+1)}{2L}, \quad k=0, \dots, L-1 \quad (11)$$

Learnable spectral action is restricted to the first K coefficients, while higher modes are set to zero:

$$\bar{x}_k^{(b)} = \begin{cases} W_k^{(b)} \hat{x}_k^{(b)}, & 0 \leq k < K \\ 0, & k \geq K \end{cases} \quad (12)$$

K fixes the retained spectral bandwidth, while only the weights inside this band are learned from data. Coefficients at and above the cutoff are suppressed in the spectral branch, which makes the low-frequency preference explicit. The modified coefficients are then returned to the time domain by IDCT,

$$h^{(b)} = D^{-1}(\bar{x}^{(b)}) \quad (13)$$

yielding the spectral-branch output on the standardized time grid.

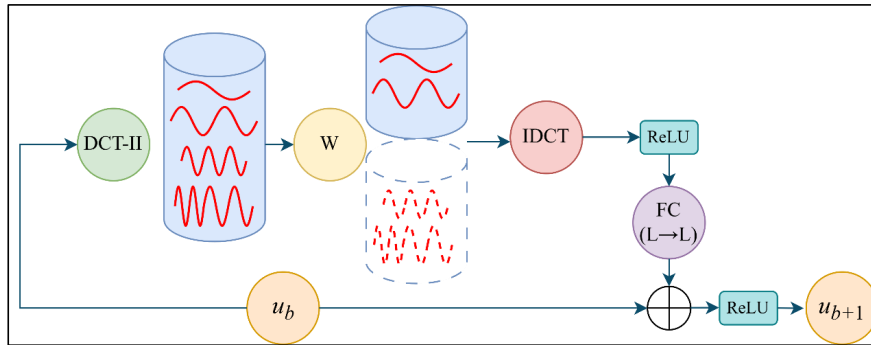


Figure 4. Single DCT-SNO block structure

A time-domain fully connected residual branch is added outside the spectral branch to represent nonlinear coupling effects that cannot be expressed by a purely linear spectral transform and to support stable stacking of multiple operator blocks:

$$u^{(b)} = W_{fc}^{(b)} h^{(b)} + b_{fc}^{(b)} \quad (14)$$



$$x^{(b+1)} = \text{ReLU}(u^{(b)} + x^{(b)}) \quad (15)$$

where the fully connected mapping remixes information across the time-domain feature
 235 vector. The residual update adds an incremental nonlinear correction without discarding the
 spectral structure already represented by the preceding block. This combination of mode-
 restricted spectral modulation and time-domain residual correction allows the network to model
 global low-frequency evolution together with local response variations.

After B residual blocks, a final linear head maps the features to the normalized output,

$$\hat{g}_{\text{norm}} = W_{\text{head}} x^{(B)} + b_{\text{head}} \quad (16)$$

producing the normalized time-domain air-gap prediction after the stacked spectral and residual
 updates.

3.4 Loss Construction and Time-Domain Training–Inference Pipeline

Figure 5 summarizes the training and inference pipeline. Both the supervised target and
 245 the final model output are time-domain air-gap sequences; DCT operations remain internal to
 the residual blocks.

During training, the predicted and reference responses are compared in the normalized
 space using mean squared error (MSE):

$$\mathcal{L}_{\text{MSE}} = \frac{1}{NL} \sum_{n=1}^N \sum_{j=1}^L \left(\hat{g}_n^{\text{norm}}(t_j) - g_n^{\text{norm}}(t_j) \right)^2 \quad (17)$$

250 where n and j index samples and grid points. MSE is the only optimization objective; no
 frequency-domain target or spectral loss is imposed during training. Through backpropagation,
 this time-domain loss jointly updates the retained-mode spectral kernels and the time-domain
 residual mappings. The frequency-domain measures in Section 4 are therefore independent
 checks of whether accurate time-domain reconstruction also preserves physically relevant
 255 spectral structure.

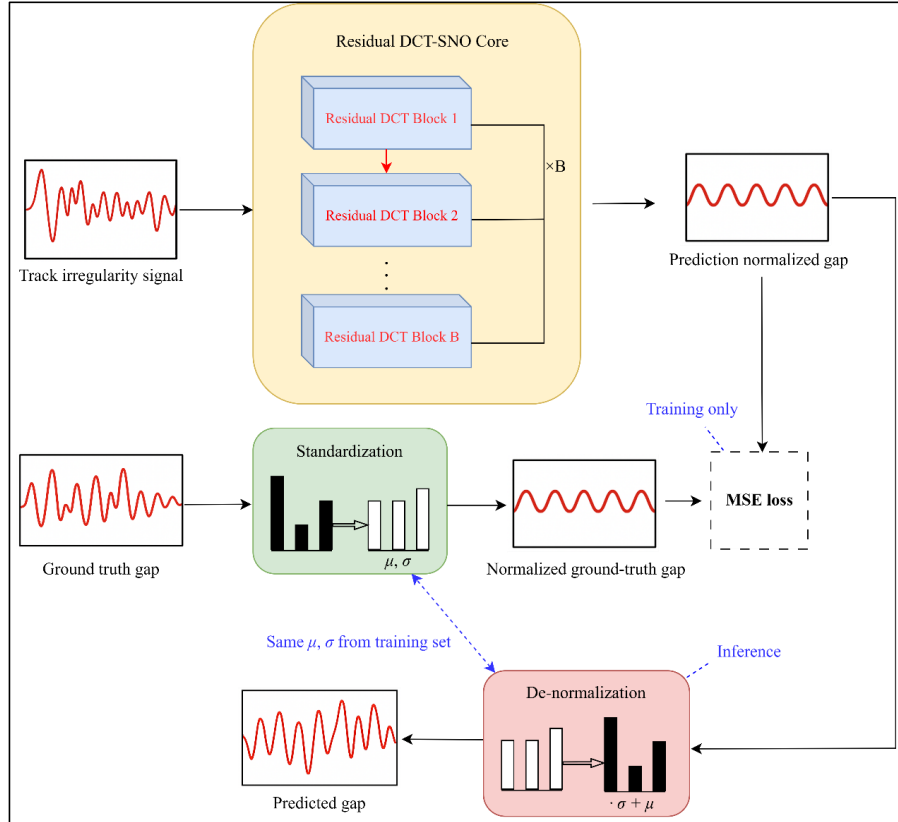


Figure 5. Multi-layer residual DCT-SNO training and inference process

At inference, the normalized output is transformed back to physical air-gap units using training-set statistics:

$$\hat{\mathbf{g}} = \hat{\mathbf{g}}^{norm} \odot \sigma_y + \mu_y \quad (18)$$

The same training-set statistics are used for every test sample, preventing test-data leakage.

3.5 Sensitivity Analysis and Training Configuration

Table 2 summarizes the algorithm. Model configuration was examined for retained modes K, learning rate, optimizer, and mini-batch size using RMSE and PSD cosine consistency as complementary configuration-selection criteria. Figure 6 shows the median and interquartile range (IQR), with filled markers indicating the adopted baseline.

Table 2. Multilayer residual DCT-SNO algorithm for air-gap prediction

Multilayer residual DCT-SNO algorithm for air-gap prediction

Input: Track irregularity sequence $z_r(t) \in R^L$, where L is the unified sequence length

Initialization: Set $x^{(0)} = z_r(t)$ as the input feature of layer 0

for $b=1$ **to** B **do** (construct the b -th neural-operator layer):

- Function-to-spectrum transform: $\hat{x}^{(b)} = DCT(x^{(b-1)})$
- Spectral operator kernel: $\tilde{z}^{(b)} = IDCT(W^{(b)} \odot \hat{x}^{(b)})$,



where $\tilde{h}$ is the learnable spectral kernel acting on the first K DCT coefficients

- Nonlinear activation: $h^{(b)} = \text{ReLU}(\tilde{z}^{(b)})$
- Fully connected mapping in function space: $u^{(b)} = \text{FC}(h^{(b)})$
- Residual update: $x^{(b)} = \text{ReLU}(u^{(b)} + x^{(b-1)})$

end for

Output layer: $\hat{y} = \text{FC}_{\text{head}}(x^{(B)})$

De-normalization: $\hat{g} = \hat{y} \odot \sigma_Y + \mu_Y$

Output: Predicted air-gap sequence $\hat{g} \in \mathbb{R}^L$

Figure 6(a) shows that increasing K from 32 to 64 improves PSD consistency while maintaining a near-minimum RMSE, whereas retaining still more modes gives no consistent joint benefit. The selected $K = 64$ therefore represents a compromise between reconstruction accuracy and the intended low-frequency bias. Figures 6(b)–(d) show analogous trade-offs for the training parameters: the adopted configuration uses an initial learning rate of 1×10^{-3} , Adam optimization, and a mini-batch size of 16. The selection is based on joint time-domain and spectral behavior rather than on minimizing a single error metric.

The baseline used in subsequent comparisons is $K = 64$, Adam, an initial learning rate of 1×10^{-3} , and mini-batch size 16.

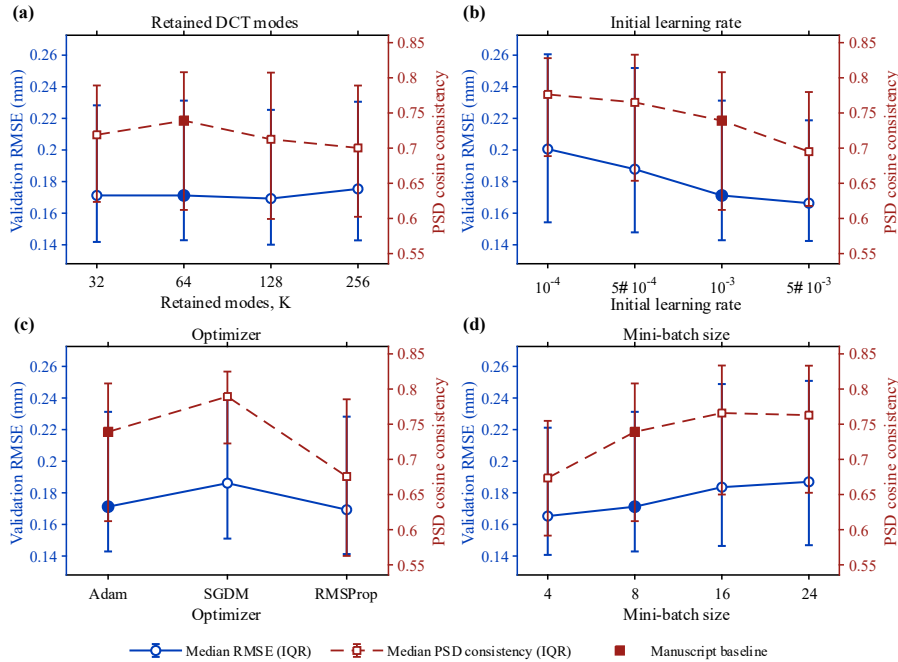


Figure 6. Sensitivity analysis of the Residual DCT-SNO configuration using validation RMSE and PSD cosine consistency: (a) retained DCT modes K ; (b) initial learning rate; (c) optimizer; (d) mini-batch size.

4 Results and Generalization Analysis

The final test set contains six held-out samples, TC1–TC6, corresponding to broadband, narrowband, and mixed irregularities at 500 km/h for $A = 1.5$ and $A = 2.0$. These six



combinations are excluded from network training and are used consistently in Figures 7-10 and
Appendix Figures A1-A2. Table 3 lists their operating conditions so that the same physical case
can be followed across the time-domain, spectral, and energy-based diagnostics.

Section 4.1 evaluates time histories, Sections 4.2-4.3 examine spectral and energy fidelity,
Section 4.4 compares baselines, and Section 4.5 defines the supported generalization scope.

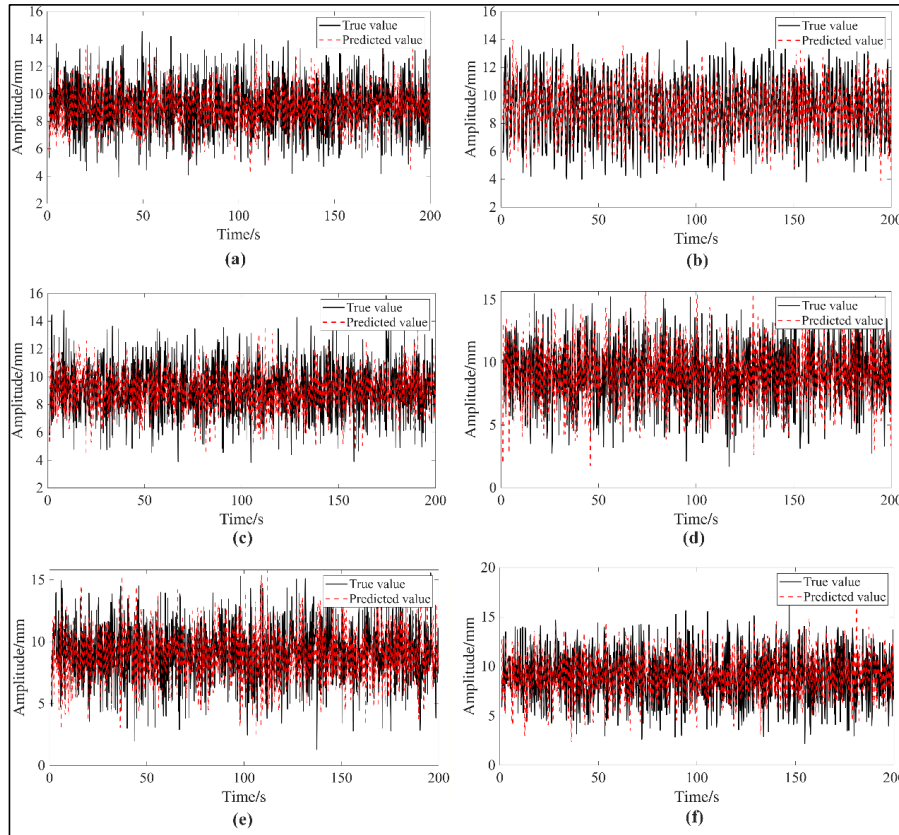
**Table 3. Operating conditions of the six held-out test samples used in Figures 7-10 and
Appendix Figures A1-A2**

Test sample	Operating condition	Spectrum type
TC1	500 km/h, $A = 1.5$	Broadband
TC2	500 km/h, $A = 1.5$	Narrowband
TC3	500 km/h, $A = 1.5$	Mixed
TC4	500 km/h, $A = 2.0$	Broadband
TC5	500 km/h, $A = 2.0$	Narrowband
TC6	500 km/h, $A = 2.0$	Mixed

4.1 Time-Domain Reconstruction of Held-Out Air-Gap Responses

Figure 7 compares DCT-SNO predictions with the high-fidelity reference responses for all
six held-out samples. The cases include the three spectral types at two high-amplitude settings,
so their waveform characteristics differ substantially. Across the six records, the predicted
sequences remain aligned with the dominant phase and overall amplitude envelope, indicating
that one trained operator can reproduce qualitatively different air-gap trajectories without case-
specific adjustment.

The main visible limitation occurs at sharp local extrema, where the prediction exhibits
mild peak-amplitude compression. This behavior is consistent with the retained-mode bias:
dominant low-frequency and band-limited content is reconstructed accurately, whereas highly
localized higher-frequency components can be attenuated. The effect is therefore treated as a
modeling trade-off and is assessed together with the spectral and energy-distribution
diagnostics rather than from the time trace alone.



305 **Figure 7. Time-domain air-gap predictions and high-fidelity reference responses for the six held-out test samples: (a) TC1; (b) TC2; (c) TC3; (d) TC4; (e) TC5; (f) TC6.**

4.2 Frequency-Domain Fidelity of the Predicted Time-Domain Responses

DCT-SNO outputs time-domain air-gap responses, while the corresponding amplitude spectra and PSDs are obtained by post-processing. As shown in Figures 8 and 9, the predictions reproduce the overall broadband spectral envelope and retain the dominant narrowband features around 10, 23, and 32 Hz with little frequency shift. Despite minor discrepancies in local amplitudes, the principal peak locations and surrounding energy levels are well preserved. These results confirm that DCT-SNO maintains the characteristic spectral structure of the air-gap response in addition to accurate time-domain reconstruction, supporting the retained-mode DCT design.

310

315

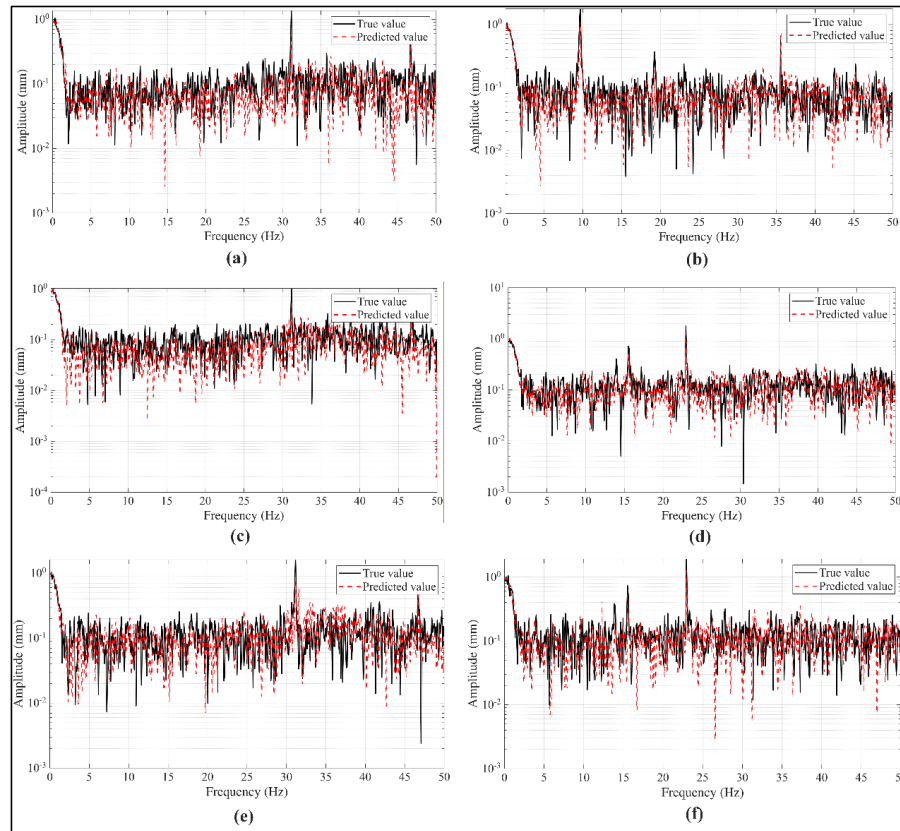
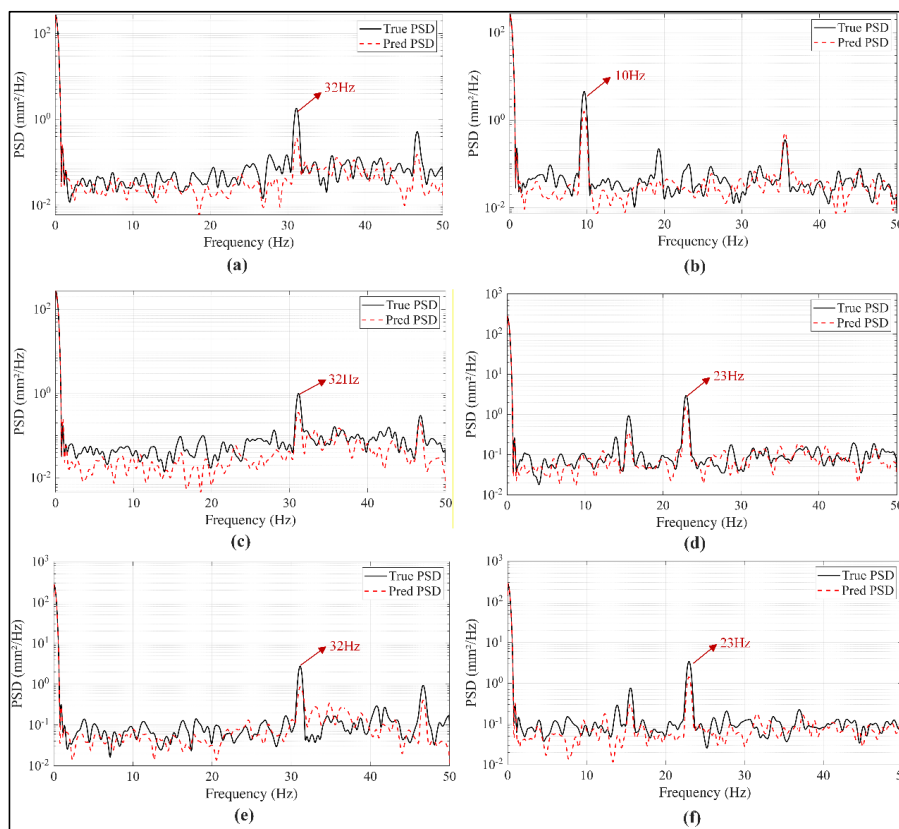


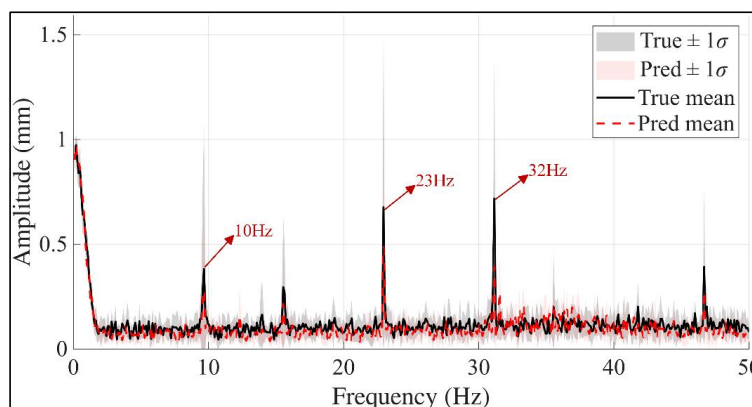
Figure 8. Amplitude spectra derived from the time-domain air-gap predictions and reference responses: (a) TC1; (b) TC2; (c) TC3; (d) TC4; (e) TC5; (f) TC6.

320 A closer examination of the PSDs in Figure 9 further supports this observation. The principal marked peak locations are reproduced without evident frequency shifts, and the surrounding spectral levels remain in reasonable agreement with the reference responses. The remaining local amplitude differences are consistent with the mild smoothing observed in the time-domain predictions.



325 **Figure 9. Power spectral densities derived from the time-domain air-gap predictions and reference responses: (a) TC1; (b) TC2; (c) TC3; (d) TC4; (e) TC5; (f) TC6.**

Figure 10 aggregates the six samples. Predicted and reference mean spectra agree over most of the frequency range, and the $\pm 1\sigma$ across-record bands overlap substantially. These bands describe deterministic variation across test samples, not probabilistic uncertainty.



330

Figure 10. Mean amplitude spectra and $\pm 1\sigma$ across-record variability bands for the



predicted and reference air-gap responses.

4.3 Energy-Distribution Fidelity of the Predicted Responses

Two PSD-based measures assess energy-distribution fidelity. Normalized cumulative power $C(f)$ describes the accumulation of spectral energy with frequency:

$$C(f) = \frac{\int_0^f S(v) dv}{\int_0^{f_{\max}} S(v) dv} \in [0, 1] \quad (19)$$

Band-energy ratio quantifies the fraction of normalized energy in specified frequency intervals:

$$P_{[f_1, f_2]} = \frac{\int_{f_1}^{f_2} S(f) df}{\int_0^{50} S(f) df}, \quad \sum_{B \in \{[0, 5], [5, 20], [20, 50]\}} P_B = 1 \quad (20)$$

Detailed sample-wise results are provided in Appendix Figures A1 and A2. The predictions follow the reference cumulative-power trends and preserve the dominant 0-5 Hz energy share; remaining differences are consistent with the mild smoothing seen in the time histories and PSDs.

4.4 Distributional Consistency and Model Comparison

Three metrics summarize spectral consistency across the test samples: band-energy-share consistency, cumulative-spectrum Kolmogorov-Smirnov distance, and normalized earth-mover distance. They measure agreement in band allocation, cumulative spectral shape, and frequency-wise energy displacement.

The band-energy-share consistency index is

$$CI_{share} = 1 - \frac{1}{2} \sum_{b \in B} |p_b^{pred} - p_b^{true}| \quad (21)$$

where B denotes the frequency bands and p_b the corresponding energy fraction; values closer to 1 indicate better agreement.

The cumulative-spectrum distance D_{KS}^p measures the maximum discrepancy between normalized cumulative spectra; smaller values are better.

The normalized earth-mover distance measures the frequency displacement required to transform the predicted spectrum into the reference distribution:

$$EMD_{norm} = \frac{1}{f_{\max}} \int_0^{f_{\max}} |C^{pred}(f) - C^{true}(f)| df \quad (22)$$

where larger values indicate greater spectral mismatch.

Table 4 compares DCT-SNO with MLP, GRU, LSTM, and FNO. DCT-SNO gives the best median result for all three spectral-consistency metrics: 0.977 for band-energy consistency, 0.078 for cumulative-spectrum distance, and 0.034 for normalized earth-mover distance. FNO reaches 0.953, 0.177, and 0.083, respectively, while the recurrent baselines show larger cumulative and transport distances. The results therefore support improved preservation of the response-energy distribution by the retained-mode DCT operator within the evaluated test set.

Table 4. Comparison of different models on the three frequency-domain consistency metrics



Model	CI_{share}		D_{KS}^p		EMD_{norm}	
	Value	IQR	Value	IQR	Value	IQR
DCT-SNO	0.977	[0.975, 0.979]	0.078	[0.072, 0.086]	0.034	[0.031, 0.045]
FNO	0.953	[0.951, 0.964]	0.177	[0.164, 0.196]	0.083	[0.072, 0.094]
MLP	0.958	[0.952, 0.965]	0.164	[0.150, 0.176]	0.083	[0.076, 0.093]
GRU	0.939	[0.930, 0.954]	0.262	[0.207, 0.284]	0.137	[0.118, 0.145]
LSTM	0.936	[0.925, 0.952]	0.275	[0.216, 0.312]	0.146	[0.124, 0.158]

Table 5 compares DCT-SNO with FNO under identical hardware. Per-sample latency decreases from 0.261 s to 0.025 s and inference throughput increases from 22.95 to 236.66 samples/s, corresponding to an approximately tenfold improvement in the present implementation. The lower overhead is consistent with the real-valued retained-mode spectral formulation and is important when repeated evaluations are required over many excitation cases.

Table 5. Comparison of computational performance and inference efficiency between DCT-SNO and FNO

Metric category	DCT-SNO	FNO
Per-sample latency (s)	0.025	0.261
Compute utilization (GFLOPs/s)	18.38	1.78
Inference throughput (samples/s)	236.66	22.95
Total wall-clock time	2min50s	15min7s

4.5 Generalization Scope and Engineering Interpretation

The six test samples are unseen combinations within the predefined factorial design: all three spectral types at 500 km/h with $A = 1.5$ and 2.0 . The training set still contains these amplitude levels at 300 and 400 km/h and contains 500 km/h cases for $A = 0.5$ and 1.0 . Consequently, the experiment does not represent extrapolation to an entirely new speed or spectrum family. Instead, it tests whether the learned operator can interpolate across known factor ranges and predict held-out speed-amplitude-spectrum combinations without retraining. The consistent time-domain and spectral results support this within-design generalization claim.

Performance outside 300–500 km/h, beyond the four amplitude levels and three spectral families, or after changes in vehicle loading, suspension-control law, electromagnetic parameters, guideway stiffness or geometry, and lateral/combined irregularity remains unverified. Such changes may alter the underlying input-output operator itself. Stronger claims of out-of-distribution generalization therefore require dedicated experiments such as leave-one-speed-out evaluation, parameter-randomized training, and validation against additional experimental or in-service data.

5. Conclusions

This study developed a physics-embedded Residual DCT-SNO surrogate for mapping track irregularity to EMS maglev air-gap response using high-fidelity UM-Simulink data. The main findings are:

(1) Operator-learning framework for air-gap response modeling

The function-to-function formulation provides a structured surrogate for the controlled track-irregularity-to-air-gap mapping. Restricting the learnable spectral kernel to low-order DCT modes embeds the observed low-frequency prior while the residual branch represents nonlinear time-domain corrections.



(2) Long-horizon accuracy and frequency-domain fidelity

Across the six held-out test samples, the model reproduces the dominant time-domain trajectories and characteristic spectral content. The energy-distribution metrics show better consistency than FNO and recurrent baselines, while the main limitation is mild smoothing at sharp local extrema.

(3) Computational efficiency and engineering applicability

DCT-SNO reduces per-sample inference latency from 0.261 s to 0.025 s relative to FNO in the reported implementation. The demonstrated generalization is limited to held-out combinations within the 300-500 km/h, four-amplitude, three-spectrum design; broader operating conditions and experimental validation remain necessary before deployment claims can be made.

Appendix A. Detailed Energy-Distribution Diagnostics

Appendix Figure A1 compares normalized cumulative power for the six test samples. The predicted curves reproduce the main accumulation trends; convergence to 1 at the upper limit follows from normalization and is not itself evidence of energy conservation.

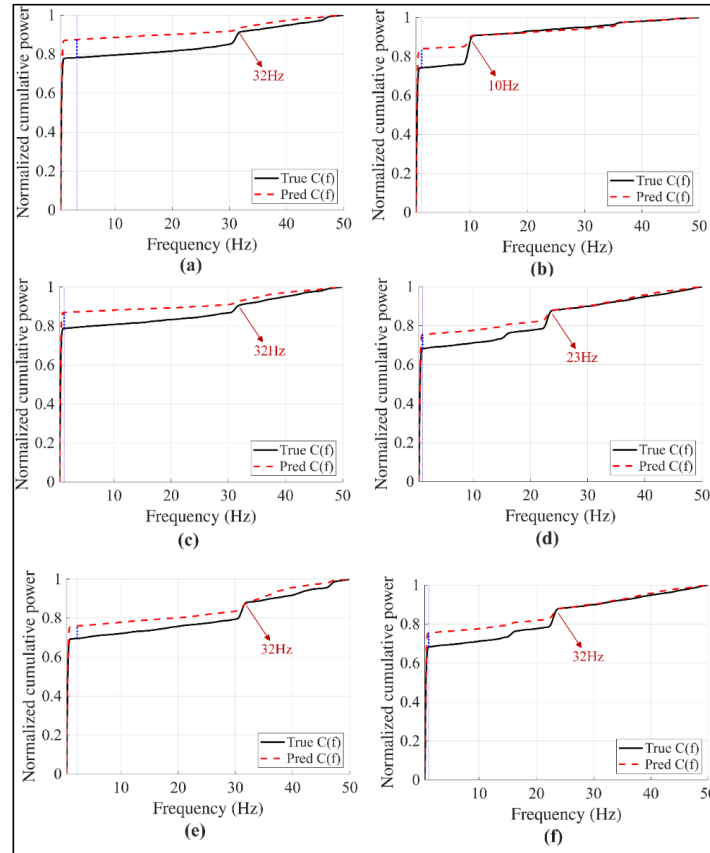
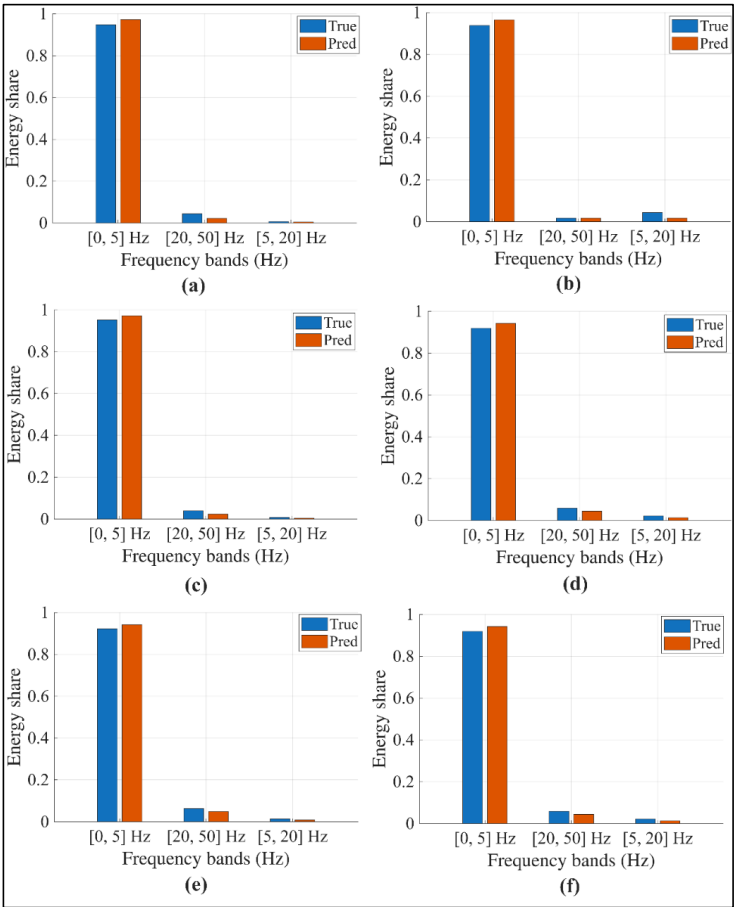


Figure A1. Normalized cumulative power curves of the predicted and reference air-gap responses: (a) TC1; (b) TC2; (c) TC3; (d) TC4; (e) TC5; (f) TC6.



415 Appendix Figure A2 compares band-energy shares. Approximately 90% of response
energy lies in 0-5 Hz for these cases, and the predictions preserve this dominant low-frequency
contribution with small redistribution toward the lowest band.



420 **Figure A2. Frequency-band energy shares of the predicted and reference air-gap
responses: (a) TC1; (b) TC2; (c) TC3; (d) TC4; (e) TC5; (f) TC6.**

Author contributions:

R. Zhang: Conceptualization, Methodology, Writing – original draft.

J. Liu: Validation, Formal analysis, Investigation.

C. Lu: Resources, Data curation, Software.

425 D. Chen: Writing – review & editing, Visualization., Funding acquisition.

Q. Li: Supervision, Project administration.

All authors have read and agreed to the published version of the manuscript.

Competing interests: The authors declare that they have no conflict of interest.

Financial support:

430 This research has been supported by the National Natural Science Foundation of China (grant
no. 52502482) under the project "Instability mechanism and flexible cooperative control of



multi-field coupled vibration in high-speed maglev vehicle-guideway systems".

Data availability

435 The data supporting the findings of this study are available from the corresponding author upon reasonable request. The data are not made fully publicly available because some of them contain or are derived from sensitive train operational information that is subject to confidentiality and data-sharing restrictions. Data that can be shared within the applicable restrictions will be provided by the corresponding author upon reasonable request.

440 Code availability

The code used in this study is available from the corresponding author upon reasonable request. Some parts of the code are closely associated with restricted train operational data and related data-processing procedures and therefore cannot be made fully publicly available. The corresponding author will provide the shareable code upon reasonable request, subject to the applicable restrictions.

References

- Boullé, N. and Townsend, A.: A mathematical guide to operator learning, *Handb. Numer. Anal.*, 25, 83–125, <https://doi.org/10.1016/bs.hna.2024.05.003>, 2024.
- 450 Huang, J. Y., Wu, Z. W., Shi, J., Gao, Y., and Wang, D.-Z.: Influence of track irregularities in high-speed Maglev transportation systems, *Smart Struct. Syst.*, 21, 571–582, <https://doi.org/10.12989/sss.2018.21.5.571>, 2018.
- Jia, X. Y., Qiang, H. Y., Xiao, C., Zhuang, C. L., Yang, P. Y., Gao, X. Y., and Wang, S. M.: Simulation analysis of track irregularity in high-speed Maglev systems based on Universal Mechanism software, *Mechatron. Intell. Transp. Syst.*, 2, 236–249, <https://doi.org/10.56578/mits020405>, 2023.
- 455 Jiang, M., Zeng, P., Wang, K., Liu, H., Chen, W., and Liu, H.: FECAM: Frequency enhanced channel attention mechanism for time series forecasting, *Adv. Eng. Inform.*, 58, 102158, <https://doi.org/10.1016/j.aei.2023.102158>, 2023.
- 460 Kong, E., Song, J.-S., Kang, B.-B., and Na, S.: Dynamic response and robust control of coupled maglev vehicle and guideway system, *J. Sound Vib.*, 330, 6237–6253, <https://doi.org/10.1016/j.jsv.2011.05.031>, 2011.
- Kovachki, N., Lanthaler, S., and Mishra, S.: On universal approximation and error bounds for Fourier neural operators, *J. Mach. Learn. Res.*, 22, 1 – 76, 2021.
- 465 Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., and Anandkumar, A.: Neural operator: Learning maps between function spaces with applications to PDEs, *J. Mach. Learn. Res.*, 24, 1–97, 2023.
- Li, Z., Kovachki, N. B., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A. M., and Anandkumar, A.: Fourier neural operator for parametric partial differential equations, 9th International Conference on Learning Representations (ICLR 2021), virtual conference, 3–7 May 2021, <https://openreview.net/forum?id=c8P9NQVtmnO>, 2021.
- 470 Lu, L., Jin, P., Pang, G., Zhang, Z., and Karniadakis, G. E.: Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators, *Nat. Mach. Intell.*, 3, 218–229, <https://doi.org/10.1038/s42256-021-00302-5>, 2021.



- Lu, Z.-R., Chen, R., and Su, C.: Random vibration analysis of maglev vehicle-bridge coupled
 475 systems with the explicit time-domain method, *Engineering Mechanics*, 39, 19–30,
<https://doi.org/10.6052/j.issn.1000-4750.2021.04.0321>, 2022.
- Pan, H., Zhu, X., Atici, S., and Cetin, A. E.: DCT perceptron layer: A transform domain
 approach for convolution layer, *arXiv [preprint]*, <https://doi.org/10.48550/arXiv.2211.08577>,
 15 November 2022.
- 480 Qin, T., Kong, D., Song, Y., Pan, L., and Zhang, C.: Effect of stochastic guideway irregularity
 on dynamic performance of Maglev train, *Infrastructures*, 10, 285,
<https://doi.org/10.3390/infrastructures10110285>, 2025.
- Ren, Q., Zhang, J., Zhou, H., and Mo, F.: Robust adaptive levitation control for medium and
 low-speed Maglev with magnetic saturation and eddy current effect, *J. Vib. Eng. Technol.*, 12,
 485 2835–2849, <https://doi.org/10.1007/s42417-023-01017-0>, 2024.
- Sheng, X., Feng, Y., Zhao, C., Liang, X., Shi, J., and Xu, L.: Extraction of characteristic and
 analysis of distribution characteristic of high-speed maglev track irregularity, *Journal of
 Railway Science and Engineering*, 23, 491–502, <https://doi.org/10.19713/j.cnki.43-1423/u.T20250668>, 2026.
- 490 Sun, Y., He, Z., Xu, J., Sun, W., and Lin, G.: Dynamic analysis and vibration control for a
 maglev vehicle-guideway coupling system with experimental verification, *Mech. Syst. Signal
 Process.*, 188, 109954, <https://doi.org/10.1016/j.ymssp.2022.109954>, 2023a.
- Sun, Y., Qiang, H., Wang, L., Ji, W., and Mardani, A.: A fuzzy-logic-system-based cooperative
 control for the multielectromagnets suspension system of Maglev trains with experimental
 495 verification, *IEEE Trans. Fuzzy Syst.*, 31, 3411–3422,
<https://doi.org/10.1109/TFUZZ.2023.3257036>, 2023b.
- Sun, Y., Xu, J., Lin, G., and Sun, N.: Adaptive neural network control for maglev vehicle
 systems with time-varying mass and external disturbance, *Neural Comput. Appl.*, 35, 12361–
 12372, <https://doi.org/10.1007/s00521-021-05874-2>, 2023c.
- 500 Vo, D. T. and Lee, J.: Physics-informed Fourier neural operator: A neural operator framework
 for non-prismatic beam bending analysis, *Eng. Struct.*, 343, 121084,
<https://doi.org/10.1016/j.engstruct.2025.121084>, 2025.
- Xin, L., Luo, S.-H., Wu, Q., and Ma, W.-H.: Dynamic response of maglev vehicles excited by
 irregularities, *International Journal of Advanced Railway*, 1, 69–73, 2013.
- 505 You, H., Zhang, Q., Ross, C. J., Lee, C.-H., and Yu, Y.: Learning deep implicit Fourier neural
 operators (IFNOs) with applications to heterogeneous material modeling, *Comput. Methods
 Appl. Mech. Eng.*, 398, 115296, <https://doi.org/10.1016/j.cma.2022.115296>, 2022.
- Yu, W., Fu, B., Li, Z., Zhou, J., Wang, J., and Zhang, J.: DTFPNet: Temporal and frequency
 dynamic graph neural network for time series classification, *Pattern Recognit.*, 172, 112588,
 510 <https://doi.org/10.1016/j.patcog.2025.112588>, 2026.
- Yu, Z., Zhang, P., Mao, J., Spanos, P. D., and Chen, Y. F.: An efficient approach for stochastic
 vibration analysis of high-speed Maglev vehicle-guideway system, *Int. J. Struct. Stab. Dyn.*, 21,
 2150080, <https://doi.org/10.1142/S0219455421500802>, 2021.
- Zhang, P. and Yu, Z.: Stochastic dynamic analysis for a three-dimensional maglev vehicle-
 515 guideway interaction system based on probability density evolution method considering
 uncertain parameters, *Structures*, 73, 108371, <https://doi.org/10.1016/j.istruc.2025.108371>,
 2025.



- 520 Zhang, Z., Huang, J., and Wang, D.: Stability analysis and field test of high-speed maglev
vehicle-guideway system, in: Structural Health Monitoring 2025: Ensuring Mobility and
Autonomy with Sustainability, Proceedings of the 15th International Workshop on Structural
Health Monitoring (IWSHM 2025), Stanford, California, USA, 9–11 September 2025, 1534–
1541, <https://doi.org/10.12783/shm2025/37452>, 2025.
- 525 Zhou, D., Hansen, C. H., Li, J., and Chang, W.: Review of coupled vibration problems in EMS
maglev vehicles, *Int. J. Acoust. Vib.*, 15, 10–23, <https://doi.org/10.20855/ijav.2010.15.1255>,
2010.