DEEP LEARNING-BASED INVESTIGATION OF AN INNOVATIVE RAIL DAMPER USING PARTICLE DAMPING TECHNOLOGY FOR NOISE AND VIBRATION CONTROL

YE XIN

PhD

The Hong Kong Polytechnic University

2024

The Hong Kong Polytechnic University

Department of Civil and Environmental Engineering

Deep Learning-based Investigation of an Innovative Rail Damper Using Particle Damping Technology for Noise and Vibration Control

YE Xin

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

October 2023

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

____________________ (Signed)

YE Xin (Name of student)

Dedicated to my beloved family,

for their love and support

ABSTRACT

Particle damping is a passive vibration control strategy that possesses various pronouncing merits in field application due to its straightforward mechanism. This technology is achieved by a granular material-filled enclosure that dissipates vibration energy through the collision and friction effects between the particles inside the cavity. To address the noise and vibration issue induced by the urban railway system, a novel rail particle damper (RPD) was developed to alleviate the vibration of the rail track. However, the intricate motion of the particle bed renders the modeling of particle dampers (PDs) a daunting challenge. This study aims to establish the surrogate particle damping models under the deep learning (DL) framework. The data-driven approach circumvents the time-consuming attempt that models the movement of all particles inside a PD, and directly gets to the desired target. The investigation of the PD is carried out progressively in this thesis, i.e., from a simple case to the field application of the rail particle dampers, and a few issues of exploiting DL will be settled during the procedure. Finally, the developed RPDs will be systematically evaluated.

The first thing to learn about a PD is its energy dissipation ability. Various parameters influence the performance of a PD. A series of tests on simple cylindrical-shaped PDs were conducted to obtain their energy loss factor corresponding to different particle properties, damper cavity properties, and external excitation properties. The collected data was utilized for the training of the neural network (NN) surrogate model. Inductive transfer learning (TL) is

applied here to remedy the lack of expensive high-fidelity experimental data, as the training of NNs is data hungry. The TL can leverage the low-fidelity knowledge from an approximate governing/constitutive equation to facilitate the learning in the target task, which forms a multi-fidelity approach for the modeling of PDs.

Moving one step further, exploring the response of a PD helps to understand its mechanism. For example, the response force of a PD exhibits hysteretic behavior under dynamic excitation. However, there is a long-standing pathology called spectra bias that hampers NNs from reproducing the complex hysteresis loops of a PD. Analyzing the NN through the neural tangent kernel's (NTK) perspective reveals why NNs are perplexed in recognizing high-frequency features. The Fourier features are thus embedded to extricate NNs from the shackle of spectra bias. The Fourier features-embedded NN (*ff*NN) is then combined with the TL-incorporated PINN (TLPINN) to enhance its performance. With these treatments, it is proved that the proposed *ff*-TLPINN can reconstruct the hysteresis loops of a PD under various excitation levels in a broad frequency band.

However, the limitation of *ff*-TLPINN is that it can only model the response force of one PD. Introducing more parameters related to PD configurations into the model will make the task too bulky for the NN. The developed RPDs are equipped with cavities of different sizes, so the models for PDs with different configurations are required. Here, a divide-and-conquer strategy is applied. The basic model is firstly established on one PD with *ff*-TLPINN, the models for other PDs are then extended from the basic model. The mc-ESDN is proposed for this model extension. The recurrent neural network (RNN) based ESDN aims to decode the sequential information from the experimental data, and the imposition of multi-dimensional and multi-scale convolutions ameliorates the feature extraction performance of the proposed method. With the mc-ESDN, the surrogated models for the developed RPDs of different cavity sizes can be established.

The effectiveness of the novel RPDs was thoroughly evaluated through experiments from laboratory to field. Numerous types of filling materials were investigated through the single-bay rail test, and the modal analysis on the RPD studied the movement of the damper cavity relative to the rail. The selected filling materials were further tested on a 6 m rail testbed to identify the optimum type of particle and its filling ratio. A noise-sensitive urban metro line was selected to evaluate the PRDs' performance on noise and vibration control of the railway system. Through the in-situ test, the vibration of the track and surrounding noise were compared before and after the installation of RPDs. The vibration wave propagation is a critical factor in evaluating the noise emission of the track. This factor was measured and predicted by a periodic track-damper coupled model with the aid of the DL surrogate model of the RPD. Results show that RPDs are functional in suppressing the vibration of rail tracks.

LIST OF PUBLICATIONS

Journal Papers:

Ye, X., Ni, Y.Q., Sajjadi, M., Wang, Y.W., & Lin, C.S. (2022). Physics-guided, data-refined modeling of granular material-filled particle dampers by deep transfer learning. *Mechanical Systems and Signal Processing*, *180*, 109437. https://doi.org/10.1016/J.YMSSP.2022.109437 (SCI, Q1)

Ye, X., Ni, Y. Q., Ao, W. K., & Yuan, L. (2024). Modeling of the hysteretic behavior of nonlinear particle damping by Fourier neural network with transfer learning. *Mechanical Systems and Signal Processing*, 208, 111006. https://doi.org/10.1016/j.ymssp.2023.111006 (SCI, Q1)

Li, G. Z., **Ye, X.**, Deng, E., Yang, W. C., Ni, Y. Q., He, H., & Ao, W. K. (2023). Aerodynamic mechanism of a combined buffer hood for mitigating micro-pressure waves at the 400 km/h high-speed railway tunnel portal. *Physics of Fluids*, 35(12). https://doi.org/10.1063/5.0177766 (SCI, Q1)

Conference Paper:

Ye, X., Ni, Y.Q. (2021), Experimental study on a particle damper for high-frequency vibration mitigation. *Proceedings of the 2nd ZHITU Symposium on Advances in Civil Engineering*, 28-29 September 2021, UNIST, Ulsan, South Korea.

ACKNOWLEDGMENTS

First and foremost, I want to express my sincere gratitude to my supervisor, Prof. Yi-Qing Ni, for offering me the precious opportunity to pursue the Ph.D. study at The Hong Kong Polytechnic University. His support and instructions have led me to overcome all kinds of obstacles during the past years. Prof. Ni is the most hard-working and open-minded person I have ever met in my life. The first time I joined this group, I learned from Prof. Ni that we should never limit our minds in research, and it is our obligation to delve into the philosophy of matters. The prudent attitude and self-motivated spirit of Prof. Ni are the most prominent things I learned in the Ph.D. study period, and they will keep illuminating my life journey in the future. It is my great honor to have the chance to study from Prof. Ni.

I would like to give my thanks to the financial support from the National Rail Transit and Electrification and Automation Engineering Technology Center (Hong Kong Branch). Colleges like Mr. Tai-Tung Wai and Mr. Wing-Hong Kwan greatly supported me. My sincere thanks also go to Dr. Wai Kei Ao, Dr. You-Wu Wang and Dr. Masoud for their guidance.

I am grateful to Dr. Si-Xin Chen, Dr. Bei-Yang Zhang, and Dr. Yun-Ke Luo. They were like my elder brothers when I was new to Hong Kong. I also want to pass my gratitude to Mr. Yu-Ling Wang, Mr. Lei Yuan, Mr. En-Ze Rui, Mr. Wei-Jia Zhang, Mr. Shuo Hao, Mr. Yan-Ke Tan, Mr. Da-Zhi Dang, Mr. Yu-Xuan Liang and Mr. Xiang-Xiong Li for their friendship and mutual support during my daily life.

I want to thank the colleagues and friends, Dr. Su-Mei Wang, Dr. Zheng-Wei Chen, Dr. Hong-Wei Li, Dr. E Deng, Dr. Duo Zhang, Dr. Zi-Yu Tao, Dr. Yuan-Peng He, Dr Wen-Qiang Liu, Dr. Masoud Sajjadi, Dr. Jason Lin, Mr. Qi-Fan Zhou, Mr. Guang-Zhi Zeng, Mr. Sheng-Yuan Liu, Mr. Yang Lu, Mr. You-Liang Zheng, Mr. Zhen Lin, Mr. Gao-Feng Jiang, Ms. Si-Yi Chen, Ms. Qi Zhu, Mr. Zhen-Bin Zhou, Mr. Guang Zhou, Mr. Chao Zhang and the other current and previous members of Prof. Ni's group. Working with you has been a pleasure to me.

Special thanks should be given to Dr. Masoud, who gave the original design of rail particle dampers, and led and managed many of the related experimental projects. Also, Dr. Wai Kei Ao instructed and helped on the modal analysis, Dr. Si-Xin Chen has inspired and guided me on transfer learning, Mr. Yan-Ke Tan has provided precious suggestions on the methodology presented in the CHAPTER 5, and Mr. Lei Yuan has inspired us on the physics-informed neural networks.

I am more than lucky to have my fiancée, Ms. Zhu's love and appreciation. Without her understanding and company, I would not be able to complete my Ph.D. study. I wish her all the best on the path to her Ph.D. degree.

I would like to dedicate my most sincere gratitude to my beloved parents. Their everlasting love and care are the most important thing to me. There is nothing I could give them in return, and I hope nothing more but to be their proud.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

The main symbols used in this thesis are listed below:

CHAPTER 3	Type	Description
η	Constant	The energy loss factor
$\mathcal{D}_h$	Constant	The high-fidelity domain
μ_c	Constant	The equivalent viscosity relates to collision
e_p	Constant	The coefficient of restitution
α_p	Constant	The volume fraction of filling particles
ρ_p	Constant	The density of particles
d_p	Constant	The diameter of particles
Θ	Constant	The particle fluctuation rate
g_p	Constant	The radial distribution function
μ_f	Constant	The equivalent viscosity relates to friction
ϕ	Constant	The inner friction angle
I_{2D}	Constant	The quadratic invariants of deviatoric stresses
p_p	Constant	The solid phase pressure
μ_p	Constant	The total equivalent viscosity

c_{eq}	Constant	The equivalent viscous damping coefficient
ρ_m	Constant	The overall density of the mixture flow
C_d	Constant	The drag coefficient
$E_{dissipated}$	Constant	The energy dissipated
$E_{kinetic}$	Constant	The total maximum kinetic energy
F_{rms}	Constant	The RMS of the force
V_{rms}^{*}	Constant	The RMS of the complex conjugate velocity
P	Constant	The value of power
γ	Constant	The cavity height
f	Constant	The excitation frequency
$\mathcal{F}$	Constant	The number of hidden layers
$\mathcal{C}$	Constant	The size of hidden layer
$\mathcal{N}[\cdot;\lambda]$	Function	The nonlinear operator parameterized by λ
$\mathbb{N}(\cdot;\boldsymbol{\theta})$	Function	The DNN parameterized by $\boldsymbol{\theta}$
$\sigma^{[l]}$	Function	The activation function output of the lth layer
$\mathcal{L}_d(\boldsymbol{\theta})$	Function	The loss function of DNN
$\boldsymbol{\mathcal{X}}$	Matrix	The input of the DNN
$\boldsymbol{W}^{[l]}$	Matrix	The weight matrix of the lth layer
$\mathcal{T}_s$	Task	The source task in transfer learning
$\mathcal{T}_T$	Task	The target task in transfer learning

$\{X, Y\}$	Vector	The dataset including input and output
CHAPTER 4	**Type**	**Description**
η_t	Constant	The learning rate
F_{tpf}	Constant	The two-phase flow model damper force
c_{eq}	Constant	The two-phase flow equivalent damping
c_c	Constant	The equivalent viscosity relates to the collision
c_f	Constant	The equivalent viscosity relates to the friction
$\mathcal{F}$	Constant	The number of hidden layers
$\mathcal{C}$	Constant	The size of hidden layer
$\epsilon_{train/test}$	Constant	The index of training/testing MSE
$f(\cdot, \boldsymbol{\theta})$	Function	The output of NN
$\mathcal{L}(\boldsymbol{\theta})$	Function	The loss function of NN
$\gamma_f(\cdot)$	Function	The Fourier mapping
$\boldsymbol{\theta}$	Matrix	The NN parameters
$\boldsymbol{K}$	Matrix	The NTK operator
$\boldsymbol{X_{train}}$	Matrix	The input of training data
$\boldsymbol{Y_{train}}$	Matrix	The output of training data
$\boldsymbol{Q}$	Matrix	The eigenmatrix of NTK operator
$\widehat{\boldsymbol{y}}$	Vector	The labeled data
CHAPTER 5	**Type**	**Description**

$\mathcal{N}_t$	Constant	The strength of extension
F_t	Constant	The yield extension force of screw
λ	Constant	The relaxation coefficient of ESN
$\rho(\cdot)$	Constant	The spectra radius of ESN
m_1	Constant	Temporal delay step for input
m_2	Constant	Temporal delay step for states
m_3	Constant	Temporal delay step for output
$\delta_{i,j}$	Constant	The Kronecker delta
k	Constant	The number of all channels
k_v	Constant	The number of selected channels
$\Psi_a(\cdot)$	Function	The activation function
$\boldsymbol{\mathcal{E}}(\cdot,\cdot)$	Function	The error of current state
$Dive(\cdot,\cdot)$	Function	The diversity value of two slices
$\boldsymbol{s_t}$	Matrix	The latent state of ESN
$\boldsymbol{W_{res}}$	Matrix	The reservoir weight matrix
$\boldsymbol{W_{in}}$	Matrix	The input weight matrix
$\boldsymbol{W_{back}}$	Matrix	The back-propagate weight matrix
$\boldsymbol{f_t}$	Matrix	The output of *t* state
$\boldsymbol{W_{out}}$	Matrix	The output weight matrix
$\boldsymbol{R_i}$	Matrix	The parameters for matrix updating
$\boldsymbol{O_t}$	Matrix	The correspond time index
$\boldsymbol{F_{d1}}$	Matrix	The basic damper force

$\boldsymbol{C_k}$	Matrix	The channels in MDSC
$\boldsymbol{S_{ik}}$	Matrix	The slices matrix
CHAPTER 6	**Type**	**Description**
$\overline{W}$	Constant	The acoustic radiation power
A_r	Constant	The area of acoustic radiation
σ_r	Constant	The acoustic efficiency constant
ρ_0	Constant	The density of air
μ	Constant	The wave propagation constant
S_t	Constant	The translational stiffness of track
S_r	Constant	The rotational stiffness of track
η	Constant	The energy loss factor
ζ	Constant	The damping ratio
$u(\cdot)$	Constant	The vibration amplitude
k_b	Constant	The wavenumber of track
$\theta(\cdot)$	Constant	The slope of track
$M(\cdot)$	Constant	The moment of track
$V(\cdot)$	Constant	The shear force of track
m_d	Constant	The effective mass of the damper
S_{ta}	Constant	The translational stiffness of the damper
Δl	Constant	The length of a track bay
Ψ	Constant	The track decay rate

ρ_r	Constant	The density of rail
$A(\cdot)$	Constant	The acceleration amplitude
$\boldsymbol{T_{ij}}$	Matrix	The transfer matrix
$\boldsymbol{r_i}$	Vector	The response vector

LIST OF ABBREVIATIONS

BBD	Bean Bag Damper
BGD	Batch Gradient Descent
CM	Cavity Mode
DEM	Discrete Element Method
DL	Deep Learning
DNN	Deep Neural Network
DoG	Difference of Gaussian
EL	Excitation Level
ER	Experimental Results
ESDN	Echo State Deep Network
ESN	Echo State Network
FBS	Frequency-based Substructuring
FF	Fourier Feature
FRF	Frequency Response Function
MDOF	Multiple Degree of Freedoms
MDSC	Multiple Dimensions and Scales Convolution
ML	Machine Learning
MSE	Mean Squared Error
NN	Neural Network
NTK	Neural Tangent Kernel

PCB	Printed Circuit Board
PD	Particle Damper
PINN	Physics-informed Neural Network
PR	Prediction Results
RMS	Root Mean Square
RNN	Recurrent Neural Network
RNN	Recurrent Neural Network
RPD	Rail Particle Damper
SDOF	Single Degree of Freedom
SGD	Stochastic Gradient Descent
SPL	Sound Pressure Level
TDR	Track Decay Rate
TL	Transfer Learning

CHAPTER 1. Introduction

1.1 Research Background

Rail transportation is widely regarded as an environmentally friendly mode of transport, with far less pollution, energy consumption, and carbon dioxide emissions per passenger kilometer than by road or air. It has been found that high-speed rail can effectively compete with air transport on routes that less than three hours. Railway can achieve the advantages of economy and environmental protection while ensuring timeliness. Public railway transportation systems are the key to realize urban mobility. The vigorous development of urban rail transit not only solves the problem of traffic congestion but also improves the utilization of land resources, which has incomparable advantages over other transportation modes. However, the environmental problems caused by urban railway transportation are becoming more and more severe because most of the rail lines are located in the area with a dense number of residents. The most crucial problem comes from the environmental problems caused by the noise of the railway system. The control of noise and vibration caused by urban rail transit has been the focus of worldwide attention.

The noise and vibration problem of railway systems have already been well documented

(Thompson, 2013). The wheel-rail contact is one of the most critical sources of railway noise. Wide-band vibration is generated when the track and wheel are in contact. The vibration energy is transferred to the environment through sound waves. In response to this mean of noise, many noise and vibration control related technologies have been developed. For example, sound barriers have been developed to block the transmission of sound waves. These products are effective but often expensive and are ineffective against the noise transmitted by the structure. Many works are focused on controlling the noise from the source, which means reducing the noise generation at the source by controlling the vibration of the structure, including utilizing composite materials wheel and rail parts, improving the smoothness of the contact surface, and other methods. In these vibration control technology methods, implementing damper is a straightforward approach. Dampers can reduce vibration and noise by dissipating vibration energy. Now dampers are widely used in many engineering fields, including rail transportation.

In numerous damping technology, particle damping is a concept that originated long time ago, which can be traced back to 1937. The damper was first designed to contain a single mass inside the cavity. The impact of the moving mass dissipates the vibration energy. However, the impact can also produce high-level noise and significant impact force. This kind of damper is sensitive to the change of excitation amplitude and coefficient of restitution (Paget, 1937). Then in 1945, Lieber and Jensen extended this concept by proposing multiple masses that moves between the walls of a container to dissipate the energy of vibrations (Lieber & Jensen, 1945). Following this idea, a multi-particle designed damper replaced the traditional single-mass damper and thus resulting in the particle dampers (PDs). Its principle is to fill the hollow cavity with many solid particles. The friction and collision between particles will dissipate the energy of the vibration system. The concept of PD takes the advantages of good durability, high reliability, not being sensitive to temperature changes, can tolerate extreme environmental conditions, and the particle damping effect is not restricted in any directions. These advantages make it possible for PDs to be widely used in engineering practice.

PDs are particularly suitable for complicated and harsh working environments. The simple working mechanism makes the PD relatively insensitive to the working conditions. It can work in different scenes like places with very high or low ambient temperatures, underwater with extremely high pressure, or in outer space. It can also control a wide frequency band of vibration. Theoretically, PDs are effective in any form of external excitation, whether steady-state excitation, shock excitation, or random chaos excitation. The vibration direction is also not limited for the PDs to be functional. Unlike some dampers, when a PD is mounted on the structure, it can mitigate the vibration in arbitrary directions. PDs also have irreplaceable advantages in terms of economy and durability. Due to their low cost, PDs have the potential for mass manufacture. At the same time, PDs also hold the advantage of durability, as they can withstand long-term wearing and tearing without affecting their effectiveness. Therefore, the lifecycle cost of this kind of damper remains low. More importantly, they would only add a little extra weight to the main structure while simultaneously ensuring the efficiency of vibration control. This is a significant advantage in applying many engineering fields sensitive to mass.

Due to their various advantages, PDs have been widely used in aviation, automation, energy, electronics, medicine, and other industries. Applications include but are not limited to structural seismic protection, vibration reduction with aviation instruments, turbofan engine blade vibration reduction, helicopter blade vibration reduction, and vibration control with aerospace instruments. However, the PDs also have their shortcomings. For example, the particles in the traditional PD are stacked together, limiting the particles' movement ability. The collision between the particles and the cavity is predominantly elastic, so the energy dissipation ability of the collision is limited.

Although PDs are convincing in enhancing the engineering application, there still lacks an impeccable analytical approach that can describe the PDs' usage performance and facilitate the PD's design. The essential fact for this problem is the highly nonlinear characteristics of the PD. As a result, in past studies, the analysis and numerical models often come to contradictory conclusions with experimental results. Trial and errors are still the most commonly used method in previous designs on PDs. At the same time, the response of the PD is very sensitive to external excitation, which leads to the end that no theoretical model can universally describe the response of the PD. Most of the past research only focus on the PD response to steady-state and low-frequency excitation, making the rest of the work conditions a blank area for PD investigation.

The application of PDs in rail transit has just commenced. The developed rail PDs are mounted on the rail track. Track is one of the primary noise sources in railway operation, especially for rolling noise. However, the vibration of rail tracks is different from normal structures. The vibration of the rail track mainly concentrates in the vertical direction on a straight line since the wheels apply the excitation vertically. The vertical stiffness of a rail track section is quite stiff, making the vibration always has a high-frequency character. In such circumstances, even when the acceleration of the vibration is significant, the displacement amplitude of the vibration is still minimal. The micro-displacement vibration is not preferred for PDs to dissipate energy, since the energy dissipation is counted on the moving of the particle bed. It is proved that the friction effect mainly contributes to the damping mechanism of PD in high-frequency vibration.

Although Masri has established the optimal steady-state solution of the multi-particle impact damper model decades ago (Masri, 1969). This identified mode is related to the particle bed's periodic impacts with the cavity's two ends. Relevant experiments have revealed that in the working environment of rail transit, this kind of movement is not achievable for the particle bed inside the PD as it requires a considerable displacement amplitude. However, even though it is recognized that the vibration energy is mainly dissipated through the friction effect, more research on the mechanism of PD energy dissipation in high-frequency vibration is needed to guide its application in rail transit.

In this research, acceleration rather than displacement will be chosen to assess track vibration based on two primary reasons. Firstly, acceleration can be more easily measured using on-site accelerometers, which offer higher precision and lower cost compared to devices used for measuring displacement. Displacement can be derived from the integration of accelerometer data, making acceleration the most common method for vibration measurement in the railway industry. Secondly, and more importantly, the high-frequency nature of rail track vibration results in minimal displacements that are often imperceptible to passengers. Typically, displacement amplitudes normally do not exceed 2 mm, whereas acceleration amplitudes can easily reach dozens of g. As a result, acceleration is the critical factor influencing passengers' experience. Therefore, the research of this thesis will be carried out in this background.

1.2 Research Objectives

This study aims to establish the surrogate model of the granular material-filled PDs for a better understanding of particle damping. This type of damper will eventually be applied to the railway industry for noise and vibration control. Since the noise is generated from the vibration of the structure, the idea of applying dampers is to suppress the structural vibration and thus control the noise from the source. The reason to select the granular material filled PD is that the granular material has been proven to be the most suitable filling material in the scope of high-frequency vibration. However, understanding this kind of damper and the damper application in this scenario is currently inadequate. Based on this situation, the detailed objectives of this study are:

1. To identify the energy dissipation ability of the PDs concerning various parameters,

including internal parameters related to dampers' cavity and particles, and external parameters related to excitation. Since the energy-dissipating ability is the primary factor for an engineer to consider a damper, the understanding of the granular material-filled PD must be built in this aspect starting from a simple shape PD. The parameters to be investigated include particle properties, damper cavity properties, and external excitation properties.

2. To propose a method for modeling the PDs with limited experimental data. After the experimental investigation of the PDs, the surrogate model based on the obtained experimental dataset is desired. However, the expensive high-fidelity data may not be sufficient to establish the deep learning-based surrogate model. A multi-fidelity modeling approach that utilizes two fidelity levels of learning resources is adopted to integrate low-fidelity learning resources into the model.

3. To further test the PDs to explore their damping mechanism. The energy consumption factor is an indirect character of PDs for application use. The response of a PD under dynamic excitation reveals more about its essence of the damping mechanism, e.g., the hysteretic behavior of the particle damping.

4. To propose another method for reconstructing the response of the PDs. Unlike the situation in the energy consumption experiment. The dynamic response test is time dependent. The high-frequency sampling rate will make the dataset relatively enormous and make the learning task too bulky for a neural network. A divide-and-

conquer strategy is applied here to solve the problem.

5. To thoroughly evaluate the proposed rail PD. Rail particle dampers (RPDs) are developed for the noise and vibration control of the railway system. Its efficacy needs a systematic investigation from laboratory to field. Since the damper cavity is fixed, different types of filling materials and their filling ratios should be tested to identify the optimum set. It is also required to know its real effectiveness in field applications. A simple model coupling the track and the dampers with the aid of the surrogate PD model is also needed to predict the actual performance of RPDs on the field.

1.3 Thesis Outline

The thesis comprises seven chapters. A brief description of the content of each chapter is given as follows.

CHAPTER 1 gives an overall introduction to this thesis, which includes the background and motivation. The introduction of particle damping and the issues of applying particle damping in the railway industry are given. The challenges for the modeling are also briefly discussed.

CHAPTER 2 reviews the development of the PDs and the investigation conducted on particle damping. The PD has a long history and has been applied to various industries. Various improvements are invented on the PDs during their application, including the configurations and materials. The attempt to model PDs is mainly carried out through the discrete element method (DEM). The DEM simulates the contact of particles by considering all aspects, resulting in a significant computational load of this approach. The various experimental methods to investigate PDs are also reviewed in this chapter.

In CHAPTER 3, a novel approach for modeling granular material-filled PDs with varying cavity height and particle filling ratio is presented. This approach is based on transfer learning (TL) and utilizes multi-fidelity techniques. The dynamic test conducted in this study encompassed frequencies ranging from 100 Hz to 2000 Hz. Simple cylindrical PDs were used in the experiment to establish fundamental knowledge about particle damping. The deep neural network (DNN) is proposed for the modeling task. Nonlinear dynamics have shown promising results with DNNs. However, a DNN relies heavily on the sufficiency of learning resources, i.e., the size of the training dataset. To address this, a multi-fidelity approach is proposed, which combines low-fidelity data from an approximate governing/constitutive equation with high-fidelity experimental data within the framework of deep TL. Initially, a DNN is trained using low-fidelity data. Subsequently, the weights and biases of the pre-trained DNN are frozen except for a few outermost layers. The outermost layers are then re-trained using the experimental data to construct a multi-fidelity DNN.

CHAPTER 4 studies the hysteretic behavior of PDs under dynamic excitation. The hysteresis means that the damper's response depends not only on the current excitation but also on its excitation history. The hysteresis loops of a PD vary with the excitation frequency due to its nonlinear nature. To model the particle damping hysteresis, this study proposes using neural networks (NN) to reconstruct the hysteresis loops of a cylindrical PD. However, NNs suffer from a long-standing pathology called spectra bias, which means they struggle to recognize high-frequency components in the target function. To address this issue, the recently developed theory of neural tangent kernel (NTK) revealed why NNs are perplexed by the spectra bias. Based on this theory, Fourier features-embedded neural network (*ff*NN) is proposed to expedite the learning of NNs on high-frequency features. After implementing the Fourier features embedding, an investigation on the use of transfer learning (TL), incorporated with the physics-informed neural network (PINN), is conducted to improve the proposed model's performance. These approaches formulate the proposed *ff*-TLPINN.

The limitation of *ff*-TLPINN is that it can only model the response force of one PD. CHAPTER 5 discusses a model extension from the basic model to achieve the surrogate for multiple dampers. In this chapter, the testing target switched from the simple cylinder-shaped PD to the cavities of RPDs. The experimental data collected from the dynamic experiments is a group of time series. However, the conventional structure of NN ignores the sequential information of time series. The echo state network (ESN) in the recurrent neural network (RNN) framework is proposed to decode the sequential information in the collected dataset. To strengthen the learning ability of ESN, a deep fully connected layer replaces the linear regression output layer and thus forms an echo state deep network (ESDN). A multi-dimensional and scale convolution (MDSC) layer is introduced in front of the reservoir of the ESDN to facilitate the feature extraction. These approaches formulate the proposed mc-ESDN.

CHAPTER 6 studies the effectiveness of the developed RPDs. Numerous types of filling materials were investigated through the single-bay rail test, and the modal analysis on the RPD studied the movement of the damper cavity relative to the rail track. The selected filling materials were further tested on a 6 m rail testbed to identify the optimum type of particle and its filling ratio. A noise-sensitive urban metro line was selected to evaluate the PRDs' performance on noise and vibration control of the railway system. Through the in-situ test, the vibration of the track and surrounding noise were compared before and after the installation of RPDs. The vibration wave propagation is a critical factor in evaluating the noise emission of the track. This factor was measured on the field test, and later compared with the model prediction by a periodic track-damper coupled model.

CHAPTER 7 gives the conclusion of this study. The contributions and findings are summarized, followed by recommendations for future research.

CHAPTER 2.
Literature Review

2.1 Development of Particle Dampers

Lu (2020) has given a detailed summary on the development of particle dampers (Figure 2-1), The traditional particle dampers can be roughly categorized into four types regarding the number of particles in the damper's cavity according to Lu's definition (2020): the impact damper (Lieber, 1945), which has only one single moving mass inside the damper and the vibration energy will be attenuated by the inelastic impact between the moving mass and cavity wall; the multi-unit impact damper (Nayeri et al., 2007), in which the working strategy is still inelastic impact between mass and damper cavity. However, the cavity is divided into several sub-cavity, and each of the sub-cavity is filled with one moving mass, so the damper can be regarded as an assembly of multiple impact dampers; the particle damper (Lu, 2020), which is filled with numerous small particles instead of one single moving mass. The advantage of this kind of dampers is evident, since the amount of inelastic collision and the effect of friction are significantly increased compared to the impact damper, without adding more extra mass to the damper. However, the advanced performance of the particle damper comes with a side effect: it is tough to model the motion of particles inside the damper while under excitation. The fourth type is the multi-unit particle damper (Saeki, 2005). The design is the same concept of particle damper. It assembles multiple particle dampers and thus using the superimposed effect.

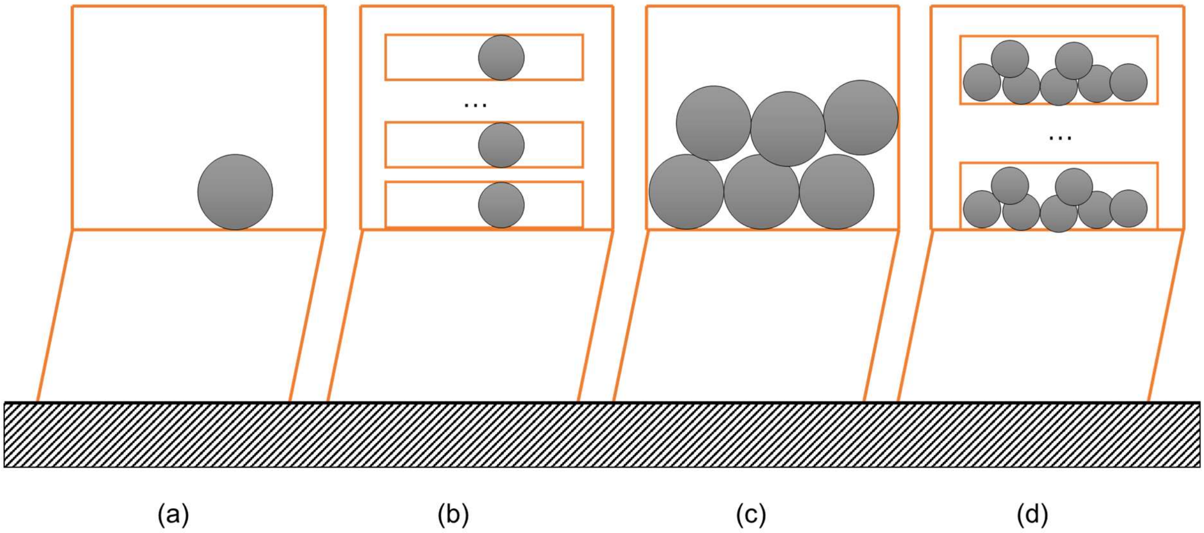


Figure 2-1 Diagrammatic drawing of four kinds of commonly used particle dampers: (a) Impact damper; (b) Multi-unit impact damper; (c) Particle damper; (d) Multi-unit particle damper (Lu, 2020).

The following literature review on particle dampers' development has referred the content in Lu's book (2020). Multi-unit impact dampers are developed based on impact dampers. In 1969, Masri (1969) successfully derived the explicit solution of the motion of a multi-unit impact damper under the harmonic steady-state excitation. In his calculation, he assumes the steady-state vibration when two collisions between particles and the cavity wall occurs every period. This is the condition when the multi-unit impact damper can dissipate the most energy and obtain its best performance, and the experimental results agreed well with his simulation in computer. Moreover, this kind of design is able to achieve a superior performance compared to the original impact damper. Based on this pioneering research, the concept of particle damper has attracted more attention.

Particle dampers with multiple particles have been widely proven to perform better than impact dampers with only one particle. The factors that could influence the performance of a particle damper were studied (Saeki, 2005). In this study, the discrete element method was used to analyze the horizontal vibration of multi-unit particle dampers. The analytical results were compared with the experimental results to prove its reliability. In the experiment, particles were filled inside several cavities during vibration, and a series of tests revealed the effectiveness of various parameters. The results turned out that cavity radius was preferred to be lower when there were more cavities inside the multi-unit particle damper. Besides, the mass ratio of particles and the cavities numbers are key factor to the performance of particle dampers (Saeki, 2005).

2.1.1 Development on the shape of particle dampers

In the scope of traditional particle dampers, the configuration only means the particle damper cavity's size and shape. However, innovating the particle dampers' structure configuration can improve the vibration control performance under some specified scenarios. The bean bag impact damper (BBD) is a flexible restraint particle impact damper (Popplewell & Semercigil, 1989). Instead of directly filling particles inside a cavity, BBD utilities a soft bag to wrap the particles (Figure 2-2). The soft bag holds a certain elastic resilience. Therefore, BBD generally works as a traditional impact damper but replaces the rigid mass block with a bag of particles, so that friction effect is introduced during the collision. Besides, this change also reduces impact force and noise generated by the impact thanks to the soft surface of the bag. In the study, Popplewell and Semercigil found that the vibration of the main structure under the action of BBD could be well analyzed through simple equivalent substitution. However, the deformation of the bean bag itself under the impact is beyond their ability to solve.

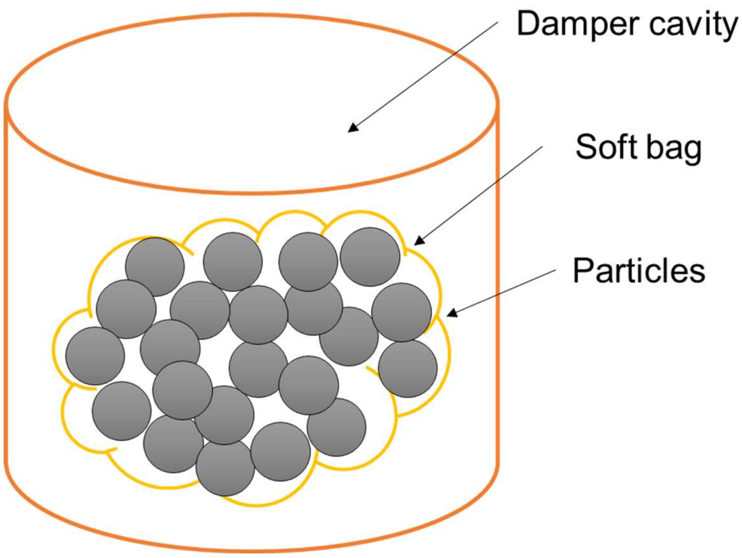


Figure 2-2 Diagrammatic drawing of BBD (Popplewell & Semercigil, 1989).

In another modification, Shah et al added a piston that inserts into the particle bed (Shah et al., 2009). This piston would not move during the vibration, as a result, when the damper vibrates with the structure, the particle bed will interact with the rigid piston and dissipates more energy. It is equivalent to extend the outer cavity wall that contacts with the granular bed. The expanded contact area improves the damping effect. This design's significant change is that the damper still has damping effect even if the external vibration is very small. Furthermore, they also achieved the semi-active controlling damper by applying a magnetic field (Shah et al., 2011). The whole system is magnetized in the magnetic field that generated by the electromagnetic coil (Figure 2-3). The pistons and particles are made of ferromagnetic materials without hysteresis. They found that the effectiveness of the damper increases with the magnetization of the piston and the particle, so controlling the energy of the magnetic field by switching the magnitude of the current in the electromagnetic coil can achieve semi-active control of magnetization.

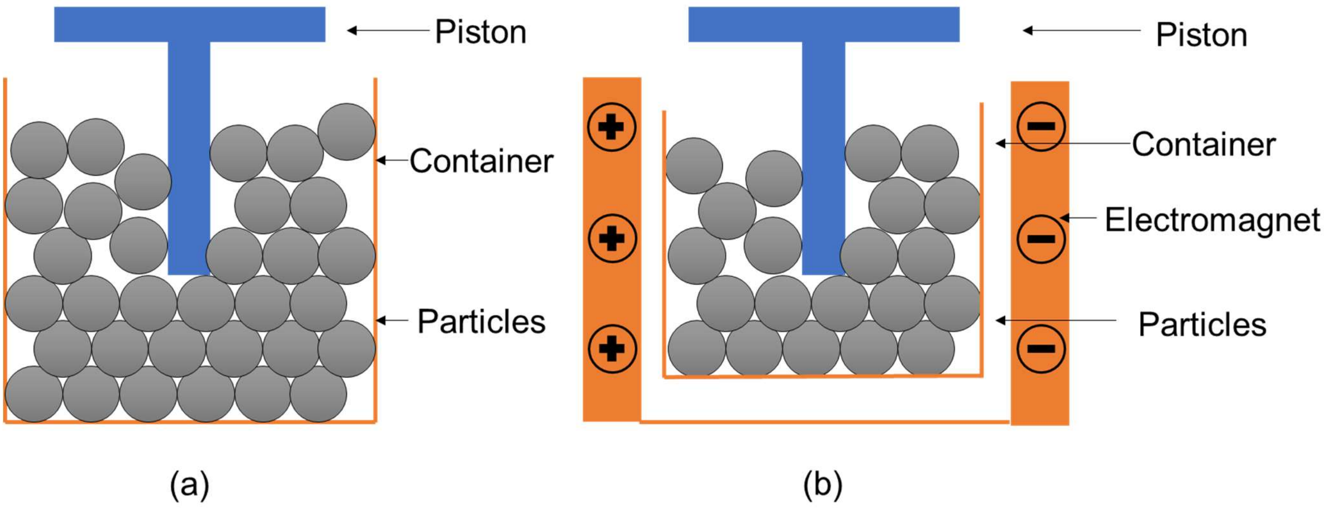


Figure 2-3 (a) Diagrammatic drawing of piston-based particle damper; (b) Semi-active control of magnetic field damper (Shah et al., 2011).

Particles of different sizes are also used in particle dampers. Theoretically, particles of various diameters can significantly reduce the porosity of granular beds, but such an approach will also significantly reduce the number of effective collisions. Therefore, Gharib and Ghani (2013) chose to place particles with different diameter sizes on the same axis and restrict the external excitation in the same direction, thus reducing the equivalent clearance in this direction (Figure 2-4). This treatment increases the number of impacts, and thus improves the damper effectiveness.

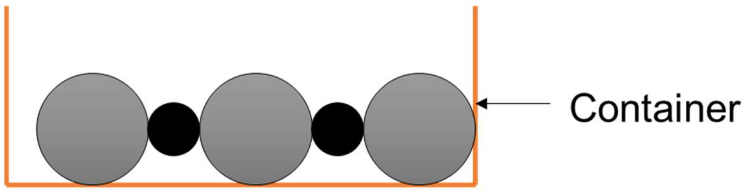


Figure 2-4 Diagrammatic drawing of particle damper designed by Gharib and Ghani (2013).

In addition to the new particle dampers described above, researchers have also developed ball vibration absorber (Zuo et al., 2019), the tuned rolling ball damper (Chen & Georgakis, 2013) etc. These dampers have achieved better vibration reduction performance through their novel design.

2.1.2 Development on the material of particle dampers

A lot of scholars contributed to the investigation on the material development of the particle dampers. Li and Darby (2008) adopted a kind of innovatively designed damper by laying a thin layer of rubber material on the inner wall of the particle damper cavity. This kind of design is based on the idea to increase the energy dissipation of collision between particles and damper cavity. In their research background, which is the vibration control of multi-degree-of-freedom structure control, the vibration is usually low frequency with relatively large amplitude. Therefore, the impact effect plays the main role in energy dissipation. But adopting such a novel design, Li and Dardy (2008) managed to enhance the damping effect in a broad frequency band. Du and Wang (2010) made improvement to a single-particle impact damper. Instead of putting one steel ball inside the damper cavity, they also filled in a certain number of fine particles. Consequently, these fine particles serve as additional energy absorber during vibration and dissipates more energy compared to a single-particle impact damper. Darabi and Rongong (2012) applied spherical elastomeric particles in their research. The physical properties of elastomeric particles were tested in a special designed test rig, parameters that could influence the performance of the particle damper including Young's modulus, loss factor and dynamic stiffness, etc. were determined before the testing of the particle damper. After the shaker test, they found out that energy consumption is related to the amplitude and frequency, but not sensitive to the friction coefficient (Darabi and Rongong, 2012).

Bustamante et al. (2013) also adopted elastomer particles in their research. However, they didn't limit to the sphere particles, but also tested some irregular shaped elastomer particles with the dynamic shaker. They reported a fluidization phenomenon in their experiments. This phenomenon indicates the condition when the movement of particles can suppress the vibration of the primary structure, a beam tested in their experiments, to achieve the lowest response.

Michon et al. (2013) used soft hollow particles to replace traditional particles in aluminum honeycomb cantilever beams. They identified the viscos-elastic behavior of the particle bed through the hysteretic response. Abbas et al. (2014) compared the effectiveness of rubber material-made particle with metal particles, including aluminum, mild steel and stainless-steel swarf for the vibration control of a hollow aluminum beam.

2.2 Theoretical Analysis on Particle Damping

Various literatures have repeatedly emphasized that the physical modeling of particle dampers is a highly nonlinear problem, which brings great difficulties to the establishment of the damper model. One of the reasons for this situation is that the particle damper is very sensitive to multiple parameters, and its performance is divided into several different stages, in which the major damping effect is generated from different physics mechanisms. Therefore, a reasonable model can only accurately describe the performance of a particular stage.

As a matter of fact, many of the current modeling approaches are based on trial and error. Because the results of theoretical models are often quite different from the results of practical experiments. In this section, the modeling method of particle damper will be illustrated in analytical and numerical aspects.

2.2.1 Analytical modeling approaches

A rheological-based model was proposed in the research reported by Olędzki et al. (1999). This model is established based on the assumption that particles will deform in a collision. In conclusion, the motion of the overall damper is described by the following rheology theory. The parameters in the equation are calibrated by experimental results. The model, which should be used for impact dampers, was then introduced into the simulation system, and obtained a good agreement with the experimental results.

Ibrahim (2009) has concluded a series of modeling and analytical approaches regarding the vibration an impact. He suggested that the existence of impact and collision can make a linear or weakly nonlinear system becomes strongly nonlinear. In order to analyze such a motion, one can try with power-law phenomenological modeling, which attempts to comprehend the chaotic motion in the energy aspect. Besides, he also introduced other analytical tools including Zhuravlev non-smooth coordinate transformation, Ivanov transformation and Hertz contact. The last one is the foundation of the discrete element method. Ibrahim (2009) suggested that the challenge of impact damping could be addressed through the application of multi-body dynamics, which analyze the movement of every single particle in a stepwise manner. However, this kind of analyzation is not easy when involves the friction effect, as it frequently involves viscous and sliding problems related to friction effects.

Chung et al. (2010) added a sand-sawdust damping layer to the lightweight floor/ceiling system. They measured the structure with the laser to obtain its velocity response. This response revealed the features of the low-frequency vibration, like the natural frequency and modal shapes. In this analytical work on the structure equipped with particle damping, the root mean square velocity was selected as the indicator to evaluate the average sound power level. Other than the theoretical analysis of the acoustic performance of the lightweight structure, Chung et al. also invited recipients to give a subjective evaluation of the structure's sound-blocking ability. Finally, the experiment results showed that with a proper design, the proposed lightweight floor/ceiling systems can reach or even surpass the performance of a concrete system.

Kollmer et al. (2013) attempted to analyze the vibration decay of a spring system by applying particle dampers. They derived two decay modes, including linear decay and linear relaxation. These decay modes are the consequence of the particle dynamics in different phases. Kollmer et al. (2013) adopted a high-speed camera to catch the movement of the cavity attached to the spring and the particle inside. Their report identified two motion modes of the particle bed, i.e., the collect-and-collide and gas phases. These two motion modes correspond to the two different decay modes of the structure on the spring. This research has provided a hint to the community that the movement of particle beds could go through different phases during vibration.

Wang and Wu (2016) introduced a new modeling approach for particle dampers based on the multi-phase flow theory. Inside the damper cavity are particles of the solid phase and air of the gas phase. Therefore, Wang and Wu regarded the particle bed as the solid-gas flow during the vibration. Following the two-phase flow theory, they separately defined and calculated the equivalent viscous damping of the particle bed related to collision and friction. This modeling approach was verified with the numerical simulation and experiments on a cantilever plate with a particle damper. The equivalent damping parameters concerning different factors were also analyzed in this study. Following this work, Lei and Wu (2017) further enhanced the theory of calculating equivalent damping parameters. Compared to the previous work, they added the equivalent damping factor based on the particle beds kinetic. The experiment validation was conducted on a non-obstructive particle damping in this study.

Furthermore, these researchers explored the optimization design of particle dampers based on the two-phase flow equivalent model (Lei, Wu, & Chen, 2018). They estimate the vibration response of the structure concerning different damper parameters. Lei et al. reported that the particle dampers reached the optimum performance when the Leidenfrost effect occurred. The Leidenfrost effect describes that at the interaction surface between the particle bed and the damper cavity, the contacting layer of particles went into a fluid phase of motion at certain conditions. With this theory, Lei et al. (2018) identified the Leidenfrost state of particle dampers with the modified two-phase flow model.

Zhang et al. (2020) followed the two-phase flow theory in their modeling attempt. They also focused on the optimization of the particle dampers. However, other than the Leidenfrost effect, they searched for the optimum condition of particle dampers through the particle swarm algorithm, an optimization algorithm based on swarm intelligence. This method works by nominating a bunch of candidates, i.e., damper parameters in this case, and then simulating the evolution of these candidates through their current state to find the optimum solution. Zhang et al. reported a 10 dB acceleration mitigation in the target frequency range by this method.

Djemal et al. (2019) also utilized multi-phase flow theory to establish a numerical model. This work can be regarded as an application of the model proposed by Wang and Wu (2016).

Although, Djemal et al. used metal balls to fill inside the damper rather than the tungsten powder. This modification made a change in the analysis of the material aspect.

Zhang et al. (2018) established an equivalent model to describe the vibration of a cantilever beam. In this study, the novel tuned particle damper was combined with the theory of conventional dynamic absorber. They applied a series of sweep sinusoidal excitations to the cantilever beam to study the performance of the tuned particle damper, the excitation level being defined by amplitude and frequency. They expected the tuned particle damper to behave like a particle damper and a dynamic vibration absorber in different scenarios. In general, particle dampers are effective in a broad frequency band, while conventional vibration absorbers perform better when the vibration amplitude is small. By conducting this analytical work, Zhang et al. (2018) summarized the changing rules of the dynamic equivalent mass and the tuned particle damper damping.

Sack et al. (2020) presented an interesting study exploring particle dampers' performance in microgravity. The particle damper was excited with sinusoidal excitation at different levels. They identified two phases of the particle's motion. The first phase was the gas phase, in which particles were evenly distributed inside the damper cavity with low vibration amplitude. Another was the collective phase that occurred with large amplitude vibration. In the second phase, particles moved following the external excitation and collided periodically with the damper cavity. The energy dissipation was undoubtedly larger in the second phase since a larger number of particles impacted the cavity in each period.

Wang et al. (2020) adopted particle damping to formulate the phononic crystal for vibration control. The phononic crystal consists of periodic distributed particle dampers filled with steel balls. This design utilized both bandgap features of a periodic structure and particle damping. For the analytical process, Wang et al. focused on equivalent modulus and Rayleigh damping ratio. The polynomial coefficients were fitted according to the simulation of particle dampers. Various parameters, including the damper's shape, size, and filling ratio, were studied in this literature. They concluded that the damping ratio coefficient of this structure was a nonlinear function concerning the vibration frequency.

Huang et al. (2021) proposed an equivalent model by considering the rolling effect of a muti-particle damper. The mechanism of particle rolling was usually ignored in the previous literature due to its complexity. In this study, the rolling of particles was investigated during the non-collision phase of the particle motion. The inertia of particles was also discussed for its influence on the dynamic amplification factor, which describes the response amplitude amplification of a structure in harmonic excitation compared to the static load. Huang et al. concluded that the inertia of particles would increase the equivalent frequency of the structure and decrease the relative displacement.

2.2.2 Numerical modeling approaches

Besides analytical models of particle dampers, scholars also endeavored to explore the numerical simulation methods for modeling particle damping. As a matter of fact, this approach might be the first choice for a researcher who first dealt with particle damping. Numerical simulation is straightforward compared to the analytical approach: it transfers the intricate theories and phenomena into a computational load of modern computers. The primary numerical method is called the discrete element method (DEM). It defines the contact between particles and calculates the motion of particles step by step based on simulation software.

The study of particle damping with DEM can at least be traced back to decades ago. Olson (2003) simulated the motion of particles and the interaction between them based on the Hertz contact, this approach considered the normal force and tangential force between particles. He adopted Maxwell model to describe the viscos-elastic behavior of material and substituted it into the Hertz model. Similarly, Liu et al. (2005) defined the effective damping coefficient of a particle damper through the two-dimensional particle dynamics. In their simulation, the particle motion was calculated step by step: the moving state of each particle at each time step was determined by the initial condition. By this approach, the total motion of the particle bed was separated to the motion of individual particles. Liu et al. reported that the simulation results agreed well with the experimental results. However, the above mentioned two studies focused on the single-degree-of-freedom case.

Lu et al. (2012) considered both collision and friction in their simulation. They adopted Hertz-Mindlin model to describe the normal contact between particles. As for the friction between particles, Coulomb friction model was selected. They established the finite-element model (FEM) of the primary structure, which is a three-story steel frame structure. The particle damper was installed on the top of this frame. The dynamic loads applied to this structure was on one direction, but since the structure contains three layers, it was still a multi-degree-of-freedom motion. The movement of the structure was calculated by the Newmark-β approach. The mass ratio, volume fraction and material of particles were investigated in their simulation.

Sánchez et al.(2012) simulated a particle damper in the single-degree-of-freedom vibration. In their DEM simulation, particles were sealed inside the damper cavity enclosure. The average and tangential elastic coefficient, friction coefficient, and coefficient of restitution were adjusted to observe the amplitude of the structure vibration and motion of particles. However, Sánchez et al. found that the performance of the particle damper was irrelevant to the interaction between particles. Once the number of particles is effectively large, or the density of particles is large enough, the particle damper can reach a considerable performance. They suppose the reason for this phenomenon was related to the inelastic collapse. According to this theory, the particle bed could dissipate the vibration energy quickly even though the energy dissipation was small in each collision between particles. Sánchez et al. concluded this effect with a single mass equivalent model and presented a suitable situation for adopting this model.

These scholars further examined the nonlinear effect of particle damping (Sánchez & Manuel Carlevaro, 2013). Again, they adopted DEM simulation on a single-degree-of-freedom particle damper that was filled with multiple particles in the enclosure. After the simulation of the complex dynamic behavior of the particle bed, they adopted the Poincare map and Lyapunov exponent to investigate this chaotic motion. From the state-space reconstruction, they pictured the vibration characteristics of particle beds in different excitation frequencies. In conclusion, they found that the particles transition from periodic to chaotic movement according to the excitation frequency. The optimum damping effect was achieved in the particle bed's periodic phase. Therefore, they proposed an optimization strategy regarding the space of the cavity. In a proper damper design, the vibration of the particle bed resides in the solid phase with a periodic pattern.

Wang et al. (2015) applied a FEM-DEM coupled method in their model. The FEM for a spacecraft was first developed. DEM then modeled the particles inside the spacecraft. In this application research, they mainly explored the effect of system parameters, including mass ratio, particle material, depth of cavity, etc. They reported that DEM could successfully simulate the behavior of particle dampers that were attached to a multi-degree-of-freedom structure. Their work provided a reference to the design and optimization of particle dampers.

Wang et al. (2016) simulated the vibration of a particle damper's motion in a two-direction, i.e., vertical and horizontal, vibration with DEM. Most of the research focused on the vibration in one direction before. They defined a term named specific damping capacity to evaluate the damping performance of the particle damper, this term was described as the energy consumption of the particle damper between two successive period. The maximum velocity of the structure was selected to estimate the kinetic energy. The DEM was adopted to simulate the energy consumption of the particle and aligned with the experimental results. Seven factors including acceleration amplitude, mass ratio of particles, clearance in vertical and horizontal directions, coefficient of restitution, friction coefficient and excitation amplitude in vertical and horizontal directions were studied. This literature introduced the idea to investigate the particle damping in the energy aspect.

Masmoudi et al. (2016) also simulated the particle damping through the interaction between each particle. Particularly, they propose a simplification model that equips the particles inside the cavity to a stack of sphere balls. With this method, the interaction between particles was significantly reduced by Masmoudi et al. Based on the DEM simulation of this study, they identified the energy loss factor as a nonlinear function of the driving frequency and mass of the particle bed.

Lu et al. (2018) studied a 5-story frame structure with the free vibration and shaker experiments. The simulation was conducted on the designed particle dampers attached to the frame structure. Unlike a conventional DEM model, they equated the particle bed inside the damper to a single mass sphere. This simplification ignored the interaction between particles, and the energy was only dissipated by the collision between the single mass and the damper cavity. Such a method is only applied to structures with a large displacement amplitude when the collision effect significantly influences energy consumption. This kind of simplification would bring significant error to the simulation for vibration in a high-frequency regime with micro-displacement. However, this equivalent model reduced the computation load of the simulation so that there was no need to involve simulation software to obtain the result. In their report, this model has an acceptable accuracy and thus can be utilized in engineering applications. These scholars studied the particle filling ratio, mass ratio, and density of particles in this literature.

Eberhard et al. (2019) proposed a novel design of particle damper. They used a stereolithography 3D printer to manufacture the tetrapod particles for damper filling. Moreover, the liquid was also filled inside the damper to enhance the damping performance. In order to simulate this configuration of particle damper, a combination of smoothed particle hydrodynamics and discrete element method was proposed. The first one was based on the Reynolds-averaged Navier-Stokes equations to describe the motion of liquids. This method could simulate the particle-liquid coupled movement and successfully predict the energy consumption of the particle damper. They reported that the addition of liquid effectively improved the relative displacement of particles and thus enhanced the collision and friction effect.

Wang et al.(2019) set up an experiment to test the performance of particle dampers in one direction. The particle dampers were installed in different locations with various response levels. These authors applied DEM simulation to establish the model for the particle dampers. A cyclic iterations approach was adopted to investigate the interaction between particles in different phases of the multi-unit particle dampers. In general, this approach calculates the damping effect at each step, and then the corresponding acceleration condition on this step is updated until it reaches the steady state. The model validation proved the accuracy of this DEM model.

Moreover, this study has presented a semi-empirical prediction method based on the damping factor of different normalized accelerations. This prediction model was able to depict the nonlinearity character of particle damping. The damping increased with the normalized acceleration until it reached a peak. These authors claim that the advantage of this semi-empirical method was its calculation efficiency. However, adequate experimental data is required to calibrate the parameters in this method.

Ferreyra et al.(2020) evaluated particle dampers' performance with two factors: apparent mass and loss factor. The modeling of particle damping was conducted with DEM, and a variety of boundary conditions of particle damper were explored in this simulation work. These boundary conditions include 3D, quasi-2D, and quasi-1D, and these conditions were compared with the theoretical prediction based on the inelastic bouncing ball model. This study showed that when the driving frequency is higher or lower than the structure's natural frequency, a strong nonlinearity of particle damping appears. They observed that the damping effect in the high-frequency regime could be well defined by a universal curve. However, the damping in the low-frequency regime was more complex but still obeyed to an upper limit.

Yan et al.(2020) studied the particle mass, coefficient of restitution, rolling friction coefficient, and radius of particles. The fiction effect was considered in their simulation. For the normal contact force, five models were compared for the numerical modeling: linear spring model, Kelvin model, Hertz model, Hertz-damper model, and Jan-Hertz-damper model. Based on the energy method, this study calculated the vibration energy of each component of the single-degree-of-freedom system. They found that particle dampers were not as good as tuned mass dampers at the natural frequency, while in other frequency ranges, particle dampers were better. It was also concluded that the particle damping performed better when the coefficient of restitution was low.

Gorla and Nicoletti (2021) developed a numerical model based on the combination of FEM and DEM. In their experiment, a hollow beam was filled with particles of relatively large size. The reason for adopting large particles was to reduce the space needed in dampers. Still, this selection came with a side effect: increased collision noise because of the enormous inertia and small clearance. The model of the beam was established by FEM simulated beam and DEM simulated particles. The hollow cavity inside the beam was 5.1 mm and 5 mm for particles. Therefore, the clearance between particles and beam was only 0.1 mm. Since the collision effect was the central aspect of energy dissipation, this kind of damper design performed better in the low-frequency range where the amplitude was more significant, and damping was low in the high-frequency range.

Liao et al. (2023) proposed a coupled model by adopting DEM and muti-body dynamics. A slider that formed the particle damper was installed on the top of a spring supported structure. The simulation results were compared with benchmark vibration tests, including single-degree-of-freedom vibration with friction, normal collision between two identical balls with different coefficient of restitution, and oblique collision of a ball with a rigid plane with stable resultant velocity but at different angles. These benchmark tests defined the interaction occurred inside the particle damper. After the modeling of shear and normal contact forces between particles, the energy definition of particles and the cavity was also given.

2.2.3 Neural network-based modeling approaches

Wang and Wu (2015) reported a new method to find the optimization design of particle dampers. Regarding various parameters that affect the performance of a particle damper, they adopted a genetic algorithm, a classical optimization method, to explore the optimum damper design. They considered the influence of internal parameters (filling ratio, material, and size) and external factors (shaker location, amplitude, and excitation frequency) in their study. In particular, they introduced back-propagate neural networks to explore the damping effect. To the authors' knowledge, it was the first time that neural networks were utilized to model particle dampers.

Since the nonlinearity of the particle damper design parameters to the damper performance, the particle damper is extremely difficult to model with a simplified method. Therefore, Veeramuthuvel et al. (2016; 2017) employed the neural network method to investigate the relationship between input parameters (such as particle size, particle density, packing ratio, etc.) and the resulting damping ratio in particle dampers. They proposed a novel application of the radial basis function neural network, comparing its performance with the traditional backpropagation neural network. The results demonstrated that this approach offers superior accuracy while reducing the computational effort required.

2.3 Experimental Investigation on Particle Dampers

2.3.1 Experiments for particle dampers on SDOF

The shaker test was widely adopted in the investigation of particle dampers. Many scholars chose to excite the particle damper directly without mounting it on a primary structure or a simple support. This approach enabled the investigation of the damper itself and could get a better understanding of the damper parameters. For example, Liu et al. (2005) applied horizontal force to a particle damper with a dynamic shaker. The particle damper was horizontally connected to the wall with a transverse spring. Different damper geometries were tested at different excitation levels to obtain their frequency response function (FRF). It was found that when the excitation level was high enough, the particle bed inside the damper would be completely excited and thus reduce the apparent mass of the damper. In this case, the natural frequency of the system increased. This phenomenon revealed the nonlinearity of the particle damping.

Xiao et al. (2015) tested particle dampers with a shaker that applied horizontal excitation force. The energy dissipation of particle dampers was calculated under this experimental setup. This study tested three different particle materials, including aluminum, stainless-steel, and tungsten. They reported that the particle density, particle diameter, and particle filling ratios were all influential parameters to energy dissipation. It was concluded that the energy consumption increased with the particle density, and there was an optimum particle diameter.

For the stainless-steel balls, this optimum number was 4 mm. For the filling ratio, the maximum energy dissipation of particle dampers occurred at the 90% filling of particles.

Bustamante et al. (2016) also used a shaker in their experiments. Several options were presented in this study for evaluation of the effectiveness of particle damping. The half-power bandwidth method was a conventional approach that estimated the equivalent damping ratio introduced by particle damping. The power input method could give the amount of energy the particle damper dissipates, which could generate the energy loss factor. This method required the damper to be directly attached to the shaker head, and the energy loss was evaluated through the phase difference between the excitation force and velocity. The spatial-average responses were also estimated for the vibration decay of the structure in this study. The white noise, sweep sinusoidal, and steady-state excitation were selected to test the nonlinearity of the particle damping.

Compared with the innovation in the shape and application scenarios of particle dampers, Snoun and Trigui (2018), improved the design method. Because the performance of a particle damper is affected by many different parameters, and the relationship between the parameters and the target is often nonlinear, it is difficult to optimize the design of a particle damper. However, the test results of particle dampers, whether through experiments or DEM simulation, can only be tested for limited parameter combinations, which means that other methods must be used to assist in optimizing the design parameters of particle dampers. To that end, Snoun and Trigu (2018) proposed the method of genetic algorithm to evaluate each design parameter.

The genetic algorithm is a parameter optimization algorithm in computational mathematics. It randomly evolves the parameters in each iteration, evaluates the best evolution direction, defines it as the next evolutionary state, and searches for the optimal state through such continuous iteration. The engineering background they designed is relatively simple, namely a clamped beam coupled with a particle impact damper. Finally, their test results were compared and verified by numerical simulation and experimental results, and it was found that using this method to find the optimal damper design can be utilized.

Genetic algorithms are further extended in the design of particle dampers. Oltmann et al. (2020) also used this method to find the optimal design parameters of particle dampers. However, they believe using simulation methods such as DEM to model a particle damper is too cumbersome and cannot guarantee accuracy. Therefore, they chose a design method that combines numerical simulation and experiment, namely frequency-based substructuring (FBS), in the application of the scenario that they were faced with, i.e., small mass main structure and particle damper.

In the theory of FBS, the main structure and the particle damper are regarded as two independent substructures, which are defined as several significant freedom degrees, respectively. The frequency response function (FRF) of the two substructures is established according to the degrees of freedom of these substructures, respectively. Then the FRF of the two substructures is combined by using the Boolean matrix and Laplace operator so that the coupled structure of FRF is formed. Since the main structure is relatively simple, traditional finite element modeling can be used to obtain the FRF function according to its corresponding degree-of-freedom. Moreover, the modeling of particle dampers, as mentioned above, is relatively complex. Therefore, Oltmann et al. chose the experimental method to obtain the FRF of the damper itself by applying a sine sweep excitation by shaker; this method regards the particle damper as a black box, ignores its internal motion model and its response can be measured directly as a result, reduces the difficulty of the research.

With the help of FBS, they can omit the complex modeling of particle damper and directly obtain the response of structural after coupling the damper and structure. They choose the peak value of FRF as the index to evaluate particle damper's performance. When the peak value of the FRF function is low, it shows that the particle damper behaves a more significant effect on the vibration control of the structure. Based on this result, they applied genetic algorithms to iteratively search the optimum damper design parameters that had the best effect on vibration control. The results show that this design method has an excellent guiding effect on the problems they set. However, the limitation of this method is that it can only facilitate the lightweight structure and particle damper design. Once the weight of the particle damper is considerable, testing itself to obtain FRF function is not feasible due to the operative limitation.

Another problem comes from the definition of structural freedom. In the application scenario introduced by Oltmann et al., the main structure is a cantilever thin plate, and the particle damper is a small box lightweight. Both are simple structures, so they are defined as three-degree-of-freedom and single-degree-of-freedom, respectively. Moreover, the particle damper and the main structure are point-connected, and only one degree-of-freedom is needed to define the connection position. However, in practical application, the main structure and particle damper design may be more complicated, such as the rail damper designed by Jin et al.. The surface contact between the damper and the rail cannot be defined with only one degree-of-freedom, which also makes applying FBS difficult.

Meyer and Seifried (2022) conducted a shaker test in the horizontal direction. The advantage of testing particle dampers in this direction was that the gravity of the particle bed could be ignored. The cubic-shaped damper cavity was connected to a linear driver with a load cell. The particle damper was filled with different numbers of steel balls. In this study, a variety of displacement amplitudes were tested; the amplitude varied from 5 mm to 50 mm with a frequency of 2 Hz. Each test applied excitation for 20 periods, and the first two periods were excluded for the steady-state analysis.

Meyer and Seifried identified two motion patterns of the particle bed. In the scattered mode, the particles' motion was irregular, and the energy dissipation was low. In the rolling collection-and-collide mode, the particle bed rolled on the damper bottom and collided with the damper cavity as an entity. In this mode, the energy dissipation was considerably high. The transmission point between these two motion patterns was identified based on the complex power method. The filling ratio determined the transmission point, which was irrelevant to the excitation frequency.

Recently, Hu et al. (2023) designed a simple test setup to investigate the energy dissipation ability of particle damper. The particle damper was vertically connected to the dynamic shaker. The absence of a primary structure made it a straightforward investigation of the damper itself, and various related characteristics were studied. An adjustment to the damper cavity was made in this research. They added a grid into the cavity, similar to the idea of a piston-inserted particle damper; this amendment improved the performance of the particle damper.

2.3.2 Experiments for particle dampers on MDOF

Michon et al. (2013) created a non-obstructive particle damper based on a beam structure. The damper cavity was designed as a honeycomb shape that spread inside the aluminum beam. A small piece of the particle damper was first tested on a shaker. The cavity was directly attached to the shaker in the vertical direction, and the viscous equivalent damping was evaluated through this test. A beam structure that consisted of honeycomb cells was then tested. In the modeling of the beam, a combination of single-degree-of-freedom equivalent models generated from the shaker test was adopted, i.e., the entire beam was divided into several sections, and the single-degree-of-freedom model represented each section in this study.

Xiao et al.(2016) embedded the non-obstructive particle dampers inside gears. As a pioneer study of particle damping in the centrifugal field, Xiao et al. drilled holes in the exact radius locations of the gear and filled them with damping particles. Experiments were conducted to investigate particle damping in mitigating vibrations in gear transmission. The different rotation speeds of the gears represented the excitation level. Special characteristics of particle damping were identified in the centrifugal field. For instance, the particles were extruded at the end from the center when the gear rotated. Xiao et al. reported that the filling ratio was a critical factor for the damping. An improper filling ratio could significantly reduce the damping effect. Based on a series of tests, they provided a guideline for optimizing the damper design in the centrifugal field.

Wang et al. (2016) designed an L-shaped cantilever beam for their experimental investigation. The long arm of the L-shaped beam was rigidly connected to the wall, and the particle damper was installed on the short arm, which pointed up. During the test, Wang et al. hang a certain mass at the joint of the L-shaped beam. By releasing the mass, the structure undergoes a two-directional (vertical and horizontal) vibration at the location of the particle damper. By changing the weight of the mass, this setup could achieve different excitation levels, which was defined as the normalized acceleration in this study. A laser vibrometer was adopted to measure the vibration response of the cantilever beam.

Duvigneau et al. (2016) introduced a granular material-filled particle damper to an oil pan of an automotive. The granular materials possess the advantage of being cheap and robust, and they are available everywhere. Most importantly, the collision effect inside a granular particle bed was significantly reduced. Thus, no extra noise would be generated from the interaction between particles. The experiment results showed the broadband vibration control ability of this kind of particle dampers.

Other than the analytical model, Zhang et al. (2020) also provided a good reference for the experimental setup. They tested the particle dampers on an aluminum plate which had an irregular shape. Particle dampers that were filled with tungsten powder were placed in different locations on the plate, which bore excitation from the shaker. By this means, this study investigated the optimum damper configuration on the primary structure.

Meyer and Seifried (2020) attached a particle damper to a beam structure. The beam was hung on its two ends with rope to achieve a free boundary condition. Shaker excitation was applied to one end of the beam while the particle damper was installed on the midspan. A laser scanning vibrometer was selected to measure the velocity profile of the beam.

As mentioned in the modeling section, Lu et al. (2017) installed the particle damper on a 5-story steel frame structure on the shaking table. The shaking table applied excitation in one direction and caused a multi-degree-of-freedom vibration on the frame. The particle damper was set on top of the frame and was divided into several sub-cavities to increase the collision between the particle bed and the cavity. They input four seismic wave signals to the frame, including EL Centro, Wenchuan, Japan 311, and SHW2. Particle dampers were capable of mitigating the structure response under seismic waves.

Lu et al. further studied the vibration mitigation of particle dampers on the frame structure (Lu et al., 2021). They set up a 76-story high-rise structure model that simulated a 306 m tall building in the wind tunnel. The frame was made of steel with a shell. This frame structure model weighed 19.2 kg with a first modal mass of 6.71 kg. The particle damper, a rectangular cavity, was again attached to the top of the frame. The diameter of the particle balls was 6 mm. Lu et al. analyzed the vibration energy of the structure under wind excitation; the energy components include kinetic energy, elastic potential energy, and energy dissipation by structural damping and particle damping. The damping induced by particle damping was divided into the kinetic energy of particles and energy dissipation through collisions. Liu et al. (2023) also tested the particle damper on top of a frame structure. The frame structure consisted of 40 stories. The main difference between this study and the previous studies was that the interaction between the frame structure and the solid ground was considered in the analysis. A soil-structure coupled model was proposed to analyze the structure response under excitation. As reported in this study, a 30% vibration mitigation could be achieved by particle dampers.

Jin et al. (2021) designed a particle damper for pipeline vibration control. The damper wraps the pipeline with filling particles. The application scenarios include hydraulic power source pipelines and aircraft hydraulic pipelines, which vibrate under the excitation of motors. During the experiment, the excitation was applied to the pipeline equipped with a particle damper in three directions, and the response in vertical and horizontal directions was compared before and after the installation of the particle damper. Since the amplitude of pipeline vibrations was relatively low, the friction effect mainly dissipated the vibration energy. Therefore, a large filling ratio was preferred in this kind of application. Jin et al. reported an optimum filling ratio between 94.9% and 97.9%.

Wang et al. (2023) also designed a vibration controller for pipelines based on particle damping. However, this kind of particle damper differed from the design of Jin et al. (2021). Wang et al. installed the particle damper on the end of an L-shaped cantilever beam. The vibration at this point contained vertical and horizontal components. The new design particle damper consisted of several cylinder-shaped cavities. Each cavity was filled with one particle. The clearance of each cavity was adjustable through a piston at the end of the cylinder cavity. Apparently, Wang et al. expected to utilize the collision between the particle and cavity to consume energy. This method was feasible since the vibration amplitude at the L-shaped pipeline was effectively large. Due to the existence of the gravity effect, two natural frequencies were identified in this structure. Finally, a 97.54% vibration mitigation was achieved by this setup based on particle damping.

Harduf et al. (2023) fabricated a structure with an embedded particle damper inside through selective laser sintering. The cross-section of the primary structure was rectangular; Harduf et al. drilled three pipes in this structure to formulate the embedded particle damper with powders as particles. During the test, the beam was hung on an elastic support, and the shaker applied excitation to the midspan of the beam. Various excitations with different frequencies and amplitudes were applied. Harduf et al. also chose a laser vibrometer to measure the response since this approach won't affect the structure. The Hilbert-transform analysis was adopted by Harduf et al. to reveal that this target structure held a similar behavior to a two-mass model. A threshold of excitation level divided the linear and nonlinear regimes of the structure dynamics. This threshold existed when the inertial force on the particle bed exceeded the friction force between the particle bed and the damper cavity.

Papalou (2023) conducted an experimental investigation on a multi-degree-of-freedom structure. Papalou suspended a particle damper on a frame. The particle damper was divided into two cavities with different sizes. In order to express the effectiveness of particle dampers, Papalou replaced the particle bed with a fixed weight. The root mean square acceleration response under random excitation for both configurations evaluated the structure’s vibration. It turned out that the particle bed could suppress the vibration amplitude around the structure’s natural frequency.

Boussollaa et al. (2023) tested a particle damper on a cantilever beam. The particle damper was installed on the beam's end, where the vibration amplitude was most significant. The damper cavity was slightly larger than the particles in their damper design. Therefore, particles were stacked and piled up inside the damper. A small clearance was left for the particles. This design achieved major energy dissipation through the impacts between the particle chain. The parameters of particles were determined by a falling test on a particle. Generally, testing particle dampers on a cantilever beam is a popular approach to evaluate particle dampers.

2.3.3 Experimental investigation in railway applications

Jin et al. (2020) first developed a particle damper for rail vibration and noise reduction. The shape of the damper is custom-designed to be more easily mounted on the track surface.

The damper is perforated with spherical particles of a slightly smaller diameter than the particle cavity. In this way, when the train passes by, and the rail vibrates, the particles in the damper will collide and rub with the cavity, and the vibration energy will be dissipated in the process. The damper design is shown in Figure 2-5. Through the modal analysis of the rail, they studied the effect of the damper under different mass blocks, analyzed the damping excitation of the damper, and also analyzed the interaction between the transverse vibration of the rail and the damper.

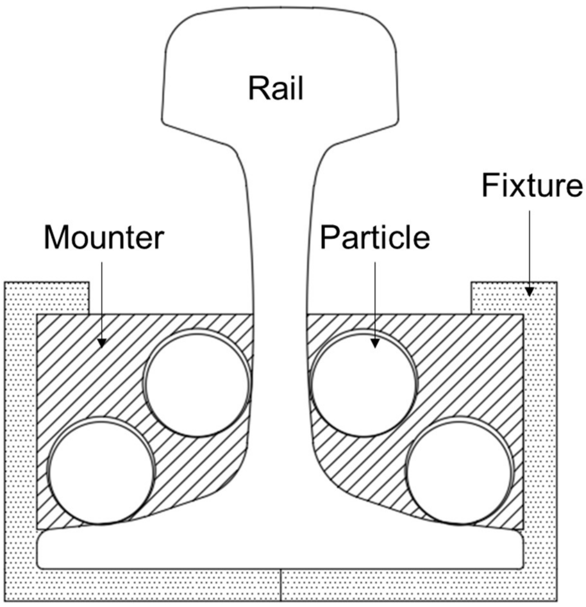


Figure 2-5 Schematic digram of the rail particle damper design (Jin et al., 2020).

They evaluated the performance of the damper by hammering the structure into free vibration under damping. Since the amplitude of free vibration displacement of the damped system exhibits natural exponential attenuation, they fitted the vibration displacement attenuation envelope obtained by experiment with the least square method and identified the damping coefficient. In addition, they established a multi-degree-of-freedom numerical model to analyze the structure's response under the damper. They defined the contact and coefficient of restitution between the particle and the damper cavity through the DEM method. Their findings reveal that the particle damper effectively mitigates structural vibrations across a wide frequency range. The inelastic collisions occurring between the particles and the cavity play a crucial role in dissipating energy. Additionally, they analyzed the influence of the clamping force exerted by the damper clamp on the vibration spectrum of the entire structure.

Following the above work, these scholars moved one more step in the application of rail particle dampers (Jin et al., 2022). The particle dampers were first mounted on a 3-m length rail track in the laboratory. This rail track was not fastened on sleepers. Through the modal analysis on the damper cavity, the scholars identified two main vibration modes of the damper cavities that were installed on the rail track: the first one was the bouncing mode at around 620 Hz; the second one was the torsion mode at about 1082 Hz. These two modes also correspond to the first and second modes of the standard UIC60 rail track. Based on this modal analysis, the equivalent connection model of the rail particle dampers was defined by a complex spring and a complex rotational spring. The complex stiffness implemented the loss factor to represent the modal damping.

Structural parameters like bending stiffness and loss factor of the coupled structure were studied at different damper parameters, including mass, diameter, coefficient of restitution, and geometry. A wave propagation theory was adopted to analyze the effectiveness of rail particle dampers in mitigating rail vibration and noise emission. The rail track was considered an infinite periodically supported Euler beam; the fasteners were also modeled as transverse complex springs. The wave propagation constants were calculated based on the periodic segments. It was found that installing rail particle dampers can effectively increase the vibration wave decay on the rail. The developed rail particle dampers were installed on an operating line for testing; the results showed that this damper suppressed the rail track vibration and reduced 5.7 dBA noise in the residential area.

Lu et al. (2023) also attempted to explore the utilization of particle dampers in railway applications. However, rather than a damper cavity that fits the rail track profile, Lu et al. only designed a rectangular damper attached to the rail web. Lu et al. are experienced experts in the field of particle damping. However, this study seems to be a preliminary investigation of rail particle dampers since a damper fixture design was not presented in this study, and the installation of particle dampers hasn't considered the safety zone of the railway operation. Numerical simulation-based modeling was adopted to explore the motion of the particle bed. Individual tests on particles determined the coefficient of restitution used in the DEM simulation.

Lu et al. defined three states of particle bed's motion, including active state, partially active state, and solid-like state. The state motion pattern was determined by the excitation levels which involves frequency and amplitude. In different states, the energy consumption was calculated through the power flow theory. The results showed that the energy dissipation ability of rail particle dampers increased with excitation levels.

Due to the high computational demand of the DEM-based numerical simulation, Lu et al. (2023) further proposed a simple analytical equivalent model to analyze the rail track coupled with particle dampers. This model was a two-degree-of-freedom vibrator in the vertical direction, which is the main direction for rail track vibration in straight line. The model parameters were fitted to align with the experimental results in the energy aspect.

2.4 Research Gaps

Based on the previous review of the particle dampers development history and the theoretical progress of simulation analysis, and in combination with the problem this study focuses on, i.e., the application of particle dampers in rail vibration reduction, the following research gaps can be summarized:

There still lacks a theoretical model that can accurately describe the behavior of the particle damper. Because the particle damper is affected by multiple internal and external parameters and most of these function relations are highly nonlinear. It is not easy to model the particle bed motion in a theoretical way. Currently, the modeling methods can be roughly divided into two ideas. One is the DEM simulation method, which attempts to reproduce the motion of paticles inside the damper cavity. This method has a higher accuracy theoretically, but many problems will occur in actual practice. First is that the DEM itself needs a standard modeling method, there are many alternative models and parameter choices for particle-particle and particle-wall contact. The existing DEM modeling mainly tunes the parameters to make the simulation results close to the experimental results as far as possible, making the simulation less useful in practice. Another defect of DEM is that it requires a tremendous amount of computation capacity, especially when the cavity is filled with granular material. The contact times in the simulation increase by several orders of magnitude, making the DEM simulation requires even more computation time.

Another approach, represented by the research of Masri and Lu, is to simplify the filling particle bed. That is to equivalent the whole particle bed to a single mass particle and then use an impact damper model to analyze the particle damper. The benefits of this approach are apparent, it reduces the complexity of the model and the number of parameters that need to be adjusted. However, the influence brought by this simplification must be addressed. For example, it completely ignores the friction effect between particles. Therefore, for the particle damper whose friction is the primary energy dissipation mode, such an analysis method is completely unsuitable.

Most of the existing models for particle dampers focus on structural vibration. Such vibration tends to be low frequency and have large displacement amplitude. In this circumstance, particle dampers work mainly in the mode of impact damping. However, the rail track is a structure with very high stiffness, the vibration frequency of its pin-pin mode is usually around 1000 Hz, and a small displacement always accompanies its vibration. These conditions mean that the particle damper will mainly work in friction damping. There is not much literature focused on this aspect.

In conclusion, a numerical method that can accurately describe the particle damper's behavior, especially the friction effect, is pending to be developed. A surrogate model that avoids the modeling of particle damper's complex internal motion may offer the potential to issue this gap.

CHAPTER 3.
Modeling of PDs with Limited Experimental Data: TL-based Multi-fidelity Deep Learning

3.1 Introduction

Based on the review of the modeling approaches of PDs, none of the current modeling approaches has demonstrated the effectiveness to work under all operating conditions. Instead, the modeling of PDs can be well conducted in a preset and limited regime of interest (Gagnon, Morandini, and Ghiringhelli 2019). It is feasible to build one or a set of continuous multi-parameter equations to link various input features of a PD with its target output as a representation of physical mechanism. Data-driven modeling approaches have advanced rapidly in recent years, among which deep learning models including deep/convolutional/recurrent neural networks have shown strong capability to accurately formulate a nonlinear mapping between observed input and output pairs, providing an appealing alternative to skip the challenge in modeling from physical mechanism (Ye & Yu 2021)The latest scientific deep learning, such as physics-informed neural network (Raissi, Perdikaris, and Karniadakis 2019), further leverages the formidable regression ability of the state-of-the-art deep learning techniques to tackle the forward and inverse problems of nonlinear systems when their physical laws are explicitly available in terms of ordinary or partial differential equations. This paradigm sheds light on system modeling in the presence of an understanding about the underlying physical laws.

In the modeling of PDs, the relevant input parameters that influence damper performance include three categories: particle properties, damper cavity properties, and external excitation properties. The output index that describes the performance of the damper can be defined freely as needed. Veeramuthuvel et al. (2017) were the first to use the neural network approach to model a PD by using experimental data and predict the vibration suppression efficiency of the damper on print circuit boards. To identify an accurate phenomenological model, high fidelity of the observed datasets must be guaranteed. However, acquisition of high-fidelity experimental data is often expensive and time-consuming, and thus only a limited amount of high-fidelity data is obtained. The commonly used deep neural network (DNN) methods for modeling are very data-hungry during training, and the resulting models from the limited amount of data may not be versatile enough to portray the physics mechanism under different working conditions, especially for highly nonlinear dynamic systems. In the model-based optimal design of PDs, a model which can accommodate varying damper cavity properties (e.g., different cavity height) and varying particle properties (e.g., different particle filling ratio) is preferred, but experiment data available for formulating such a model are always scarce since it is highly expensive to fabricate and test a series of PDs with various cavity and particle properties.

One possible solution to insufficient learning resources is that, once a computationally efficient approximate governing/constitutive equation is available, it can provide supplemental information by generating extra training data from the governing/constitutive equation. However, the data generated by an approximate governing/constitutive equation can only reflect the damper behavior in an approximate manner, that is to say, they are low-fidelity data compared with high-fidelity experimental data. Simply merging the low-fidelity data and high-fidelity data is naïve and not substantially helpful to enhance the model accuracy. Instead, the deep transfer learning (TL) technique, which enables hierarchical neural network training using multi-fidelity data (Chakraborty, 2021), is more promising for leveraging low-fidelity and high-fidelity data in deep learning. The deep TL can be built on DNN. For a DNN model consisting of several hidden layers, each layer receives the output of the preceding layer and processes it with an activation function. This execution abstracts the information and outputs it to the next layer afterward. It is known that the lower the level of a layer is, the lesser task-specific the layer is (Donahue et al., 2013). This property inspires the application of TL in deep learning to enhance learning efficiency and lower the risk of over-fitting.

In this study, a TL-based multi-fidelity modeling approach in the context of DNN will be developed for characterizing the dynamic performance of a set of granular material-filled PDs with varying cavity height and particle filling ratio. With the aid of an approximate governing/constitutive equation depicting the dynamics of granular material-filled PDs, a physics-guided DNN is first trained using the numerical train data generated from the governing/constitutive equation. This pre-trained DNN model captures the basis of complex physics mechanisms of the granular material-filled PDs yet without promising a high accuracy. Afterwards, deep TL is pursued to re-train the DNN with the use of high-fidelity experimental data, where the network parameters (weights and biases) in all layers of the pre-trained DNN except a few outermost layers are frozen, while those in the outermost layers are released for re-training using the experimental data in compliance with the deep TL philosophy, thus yielding a data-refined multi-fidelity DNN model. The proposed TL-based multi-fidelity approach enables to decode useful information about the physics of the problem from the approximate governing/constitutive equation and hence, enhance the ability to reflect the underlying physical mechanisms of the experimental results. To facilitate the understanding of damper, the input variables considered in the DNN model include cavity height, particle filling ratio, excitation amplitude and frequency, while the output variable is the energy loss factor of damper, which is defined as the ratio between the energy dissipated by particle damping and the maximum kinetic energy of the vibrating system. In the case study, a two-phase flow equivalent viscosity physical model is considered as the approximate governing/constitutive equation to generate low-fidelity numerical data, while acceleration-controlled experiments on granular material-filled PDs with different cavity heights and particle filling ratios are carried out to generate high-fidelity experimental data. The efficiency and accuracy of the formulated multi-fidelity DNN model are compared with those of a DNN model which is trained using only the experimental data, under a wide range of particle filling ratio (from 10% to 70%) and excitation frequency (from 100 to 2000 Hz). In particular, the modeling capability of the formulated multi-fidelity model in high-frequency and micro-displacement vibrations is evaluated.

3.2 Establishment of Deep TL-based Multi-fidelity Modeling

3.2.1 Data-driven DNN

The data-driven DNN explored in this study aims to train a transfer learning paradigm by utilizing the hierarchy structure of DNN to settle the conflict between the scarceness of experimental data and the data-hungry nature of deep learning methods. In our study, the energy loss factor η will be defined as a performance indicator of PDs, i.e., the output variable of the DNN. Let $\boldsymbol{\mathcal{X}} = [\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_N]$: $\Omega \rightarrow \mathbb{R}^N$ be N input parameters of the DNN which represent damper cavity properties, particle properties, and excitation properties. Yet the expression of η about the input parameters is unknown, the nonlinear function of the energy loss factor η in terms of the input parameters can be expressed in the following general form

$$\eta(\boldsymbol{\mathcal{X}}) + \mathcal{N}[\eta; \lambda] = 0, \boldsymbol{\mathcal{X}} \in \Omega \tag{3-1}$$

where $\eta(\boldsymbol{\mathcal{X}})$ is the latent solution for the energy loss factor; $\mathcal{N}[\cdot\,; \lambda]$ is a nonlinear operator parameterized by λ. In general, such a nonlinear ordinary/partial differential equation is built on the spatio-temporal domain; for the energy loss factor function, however, a high dimensional domain is considered to encapsulate the entire parameter space in which all parameters that influence the dynamic performance of PD are included. Suppose a series of observations in the parameter space are obtained as $\mathcal{D}_h = [\boldsymbol{\mathcal{X}}_h; \eta_h]$, where the subscript h denotes the observed input and output pairs. In this setting, the goal of modeling becomes to discover the parameters λ in Equation (3-1) that reflect the unknown latent state $\eta(\boldsymbol{\mathcal{X}})$ via smoothing, filtering, or data-driven solution.

Consider a DNN architecture with L hidden layers that is represented by a sequence of activation functions and linear transformations as follows:

$$\mathbb{N}(\cdot\,;\boldsymbol{\theta}) = (\sigma^{[L]} \circ \boldsymbol{W}^{[L+1]}) \circ \cdots \circ (\sigma^{[0]} \circ \boldsymbol{W}^{[1]}) \tag{3-2}$$

where $\sigma^{[l]}: \mathbb{R} \to \mathbb{R}$ and $\boldsymbol{W}^{[l]}$ denote the activation function output of the lth layer and the weight matrix connecting the lth and (l+1)th layers; the symbol '∘' denotes the operator composition. For brevity, the bias of each layer has been absorbed into the weight matrix, and all the weight and bias parameters $\{\boldsymbol{W}^{[l]}\}_{l=1}^{L+1}$ are concisely represented by $\boldsymbol{\theta}$. In this study, the

$$\boldsymbol{\theta}^* = arg\,\min_{\boldsymbol{\theta}} \mathcal{L}_d(\boldsymbol{\theta}) \tag{3-3}$$

0th layer $\sigma^{[0]}$ represents the input vector $\boldsymbol{\mathcal{X}}$ which includes various parameters (particle properties, cavity properties, and excitation properties) affecting the dynamic performance of PDs, while the (L+1)th layer is the DNN output η which is the energy loss factor of PDs. With a set of labeled experimental data $\mathcal{D}_h = [\boldsymbol{\mathcal{X}}_h; \eta_h]$, the model parameters $\boldsymbol{\theta}$ can be trained by minimizing the loss function ($\mathcal{L}_d$):

As aforementioned, the DNN-based modeling approach is data-hungry. When a data-driven DNN model for PDs involves a lot of input parameters (particle, cavity, and excitation features), it is highly expensive and time-consuming to acquire complete experimental data by exploring different particle, cavity, and excitation features associated with the whole parameter space. A DNN model formulated using insufficient data is usually incompetent to accurately describe the dynamic performance of PDs under various working conditions. The solution to this challenge is described in the next sub-section.

3.2.2 Multi-fidelity modeling by deep TL

A variety of models in terms of governing or constitutive equations have been proposed in the past decade to characterize the physics mechanism of PDs. While being capable of capturing the basic physics law of PDs, all these governing/constitutive equations can only depict the intrinsic physics phenomenon in an approximate manner. For example, when a PD vibrates in a very high frequency and with micro-displacement, the friction effect between particles is the main mechanism of energy dissipation and this damping mechanism is hard to be portrayed credibly by the existing models. In this connection, we propose a deep TL-based multi-fidelity modeling approach which leverages the low-fidelity knowledge from the approximate governing/constitutive equations and high-fidelity experimental data of PDs.

TL aims to improve learning in a new task (target task) through the transfer and leverage of knowledge from a related task (source task) that has already been learned. It has been proven (Yosinski et al., 2014; Long et al., 2015; Chakraborty, 2021) that DNNs trained with sufficient data tend to learn feature representations at the low-level layers of a DNN with generality, which are common to similar (related) tasks, and learn feature representations at the high-level layers with specificity, which depend heavily on the underlying task which the training data stem from. As such, one can transfer the knowledge across domains by freezing the weights and biases in the low-level layers of a pre-trained DNN using plentiful data from the source domain and fine-tuning the high-level layers using possibly scarce new data from the target domain. In the TL framework, given a dataset $\{X, Y\}$, $X = \{x_1, x_2, \cdots, x_n\} \in \mathcal{X}$ consists of n samples and $Y = \{y_1, y_2, \cdots, y_n\} \in \mathcal{Y}$ represents the corresponding label of the samples. The supervised learning is to map the samples X to their label Y. The pairs $\mathcal{X}$ and $\mathcal{Y}$ constitute the domain $\mathcal{D} = \{\mathcal{X}, \mathcal{Y}\}$ of the task which is denoted as $\mathcal{T}$. The TL in the context of DNN, i.e. deep TL, is pursued to accomplish the source task $\mathcal{T}_s$ first, which aims to learn a DNN using labeled data from the source domain $\mathcal{D}_s = \{\mathcal{X}_s, \mathcal{Y}_s\}$. Since the labeled data in the source domain are generated from approximate governing/constitutive equations in the present study, they are abundant yet low-fidelity, and hence the resulting DNN (referred to as pre-trained DNN) is only an approximate model without promising a high accuracy. The deep TL is then to leverage this knowledge to facilitate the target task $\mathcal{T}_T$ targeting a multi-fidelity DNN with the use of fewer new labeled data with high fidelity from the target domain $\mathcal{D}_T = \{\mathcal{X}_T, \mathcal{Y}_T\}$. In this study, the high-fidelity data in the target domain are directly obtained by experiments on actual PDs. The pre-trained DNN that was trained in the source domain $\mathcal{D}_s$ with a large amount of labeled data from an approximate governing/constitutive equation of PDs extracts and retains some underlying physics features in the low-level layers of the DNN; these learned features are general and are not specific to $\mathcal{T}_s$. Thus, in compliance with the deep TL philosophy, the multi-fidelity DNN can be elicited by re-training the DNN using high-fidelity yet possibly scarce experimental data in the target domain $\mathcal{D}_T$, where the parameters (weights and biases) in the low-level layers of the network are frozen while those in the high-level layers (outermost layers) are fine-tuned with the experimental data.

The implementation flowchart of the proposed TL-based multi-fidelity modeling method for PDs in the context of DNN is shown in Figure 3-1, where the input features include various parameters that would influence the energy loss factor of PDs and the output is the predicted energy loss factor. Although the target task $\mathcal{T}_T$ differs from the source task $\mathcal{T}_S$ in multi-fidelity learning, they share the same input feature space. As a result, only inductive TL is needed. Since the model includes the damper cavity properties and particle properties as input variables, it is suitable for optimal parameter design of PDs.

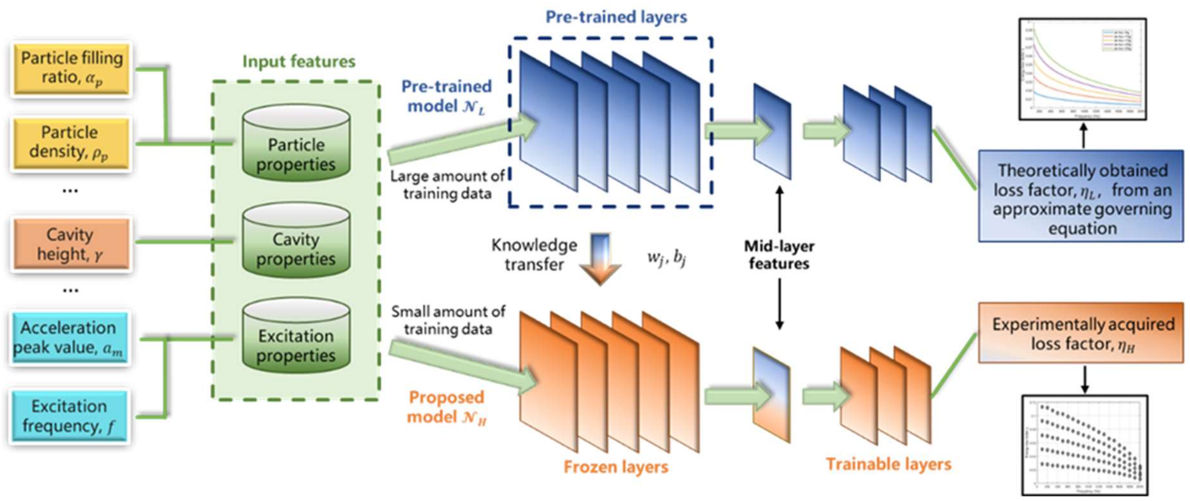


Figure 3-1 Flowchart of the TL-based multi-fidelity DNN modeling method for PDs.

3.2.3 Approximate governing/constitutive equation of PDs

Most of the current analytical models for PDs in terms of governing/constitutive equations are incapable of accurately describing the particle motion in damper cavity. The most popular modeling approach is approximating the motion of particles in the entire particle bed as an equivalent single impact ball inside the damper cavity. With this simplified model, Masri and Caffrey (2019) studied the PD's motion pattern and optimum operation condition. This model has also facilitated a great deal of studies exploring the structural vibration attenuation effectiveness of PDs (Lu et al., 2017). Recently, a novel PD modeling approach adopting two-phase or multi-phase flow theory has been proposed (Wang & Wu 2016; Lei et al., 2018a; Lei et al., 2018b) where particle motion inside the damper cavity is interpreted from the perspective of effective viscosity, and models developed under this philosophy have been adopted to evaluate the vibration characteristics of target structures mounted with grain (tungsten powder)-filled PDs. A comparison of the attributes of the single-ball impact model and the two-phase flow equivalent viscosity model is given in Table 3-1.

Table 3-1 Attributes of single-ball impact model and two-phase flow equivalent viscosity model

Single-ball impact model	**Two-phase flow equivalent viscosity model**

• Contact force occurs during particle-cavity interaction;	• Continuous drag force generated and applied on target structure;
• Intergranular interaction ignored and only particle-wall collision considered;	• Intergranular interaction considered in both collision and friction effects;
• Applicative to small-quantity, large-size particles and low-level excitations when particle bed moves as an entity.	• Applicative to granular material particles and high-level excitations.

The research addressed in this study focuses on granular material-filled PDs. Since this kind of PDs involves constant interaction between particles, the single-ball impact model is apparently inapplicable. We are also interested in noise reduction using PDs, which inevitably involves high frequency and vibration amplitude in micro-displacement level. In such scenarios, the damping of PDs is mainly contributed by the friction effect inside the cavity, and the multi-phase flow theory is more appropriate to interpret the damping mechanism. In view of this, a two-phase flow equivalent viscosity model proposed in (Wu et al., 2004; Wang et al., 2015; Wang & Wu 2016; Lei et al., 2018a; Lei et al., 2018b) will be used as the approximate governing/constitutive equation to generate numerical low-fidelity data for the pre-training of a DNN.

The two-phase flow equivalent viscosity model was inspired by the solid-gas flow theory that describes particle-particle and particle-gas interactions (Wu et al., 2004; Wang et al., 2015), which can predict the damping effect of granular material-filled non-obstructive PDs. In general, the mixture of gas and granular particle inside the vibrating cavity can be regarded as low Reynolds solid-gas flow where there are highly concentrated particles, namely a dense multi-phase flow. Under this assumption, pseudo-shear stress and equivalent viscosity can be introduced to describe the momentum exchange between particles. In this formation, all components inside a cavity can be treated as a united entity, rather than treating every particle discretely, thus avoiding the computational obstacle of the latter approach where the computational complexity is exponentially increasing when there is a large number of particles, such as in the case of granular material-filled particles in PDs.

In the two-phase flow equivalent viscosity model, instead of addressing interparticle collision traversal, it exploits the kinetic theory of dense multi-phase flow (Fan & Zhu, 1998) to derive the equivalent viscosity μ_c induced by particle-particle collision. The equivalent viscosity is defined as (Fan & Zhu, 1998):

$$\mu_c = \frac{6}{5}(1 + e_p)\sqrt{\frac{\Theta}{\pi}}{\alpha_p}^2 g_p \rho_p d_p \tag{3-4}$$

where e_p denotes the coefficient of restitution which is set as 0.6 (Lei & Wu, 2017) α_p is the volume fraction of filling particles; ρ_p and d_p are respectively the density and diameter of particles; Θ represents the particle fluctuation rate, and in harmonic motions, it is defined as $\Theta = \frac{|\dot{x}|^2}{6}$. The parameter g_p represents the radial distribution function which is expressed as

$$g_p = \frac{1}{1 - \alpha_p} + \frac{3\alpha_p}{2(1 - \alpha_p)^2} + \frac{{\alpha_p}^2}{2(1 - \alpha_p)^3} \tag{3-5}$$

Substituting the particle fluctuation rate expression into Equation (4) yields

$$\mu_c = K_1|\dot{x}| \tag{3-6}$$

in which $|\dot{x}|$ denotes the absolute velocity of PD, and the parameter K_1 is

$$K_1 = \frac{1}{5}(1 + e_p)\sqrt{\frac{6}{\pi}}\alpha_p{}^2 g_p \rho_p d_p \tag{3-7}$$

The equivalent shear viscosity μ_f generated due to the effect of interparticle friction can be expressed as (Lei & Wu, 2017):

$$\mu_f = \frac{p_p \sin\phi}{2\sqrt{I_{2D}}} \tag{3-8}$$

where ϕ is the inner friction angle of particle determined by its material; I_{2D} represents quadratic invariants of deviatoric stresses; p_p is the solid phase pressure which is a sum of kinetic term and collision term and is written as

$$p_p = \alpha_p \rho_p \Theta + 2\rho_p(1 + e_p) g_p \alpha_p{}^2 \Theta \tag{3-9}$$

Similar to the equivalent collision viscosity μ_c, the equivalent shear viscosity μ_f can be re-expressed as

$$\mu_f = K_2|\dot{x}|^2 \tag{3-10}$$

where,

$$K_2 = \frac{(\alpha_p \rho_p + 2\rho_p(1 + e_p) g_p \alpha_p{}^2)\sin\phi}{12\sqrt{I_{2D}}} \tag{3-11}$$

With Equation (3-6) and Equation (3-10), the total equivalent viscosity of particle flow, μ_p, can be obtained with considering both collision and friction effects as

$$\mu_p = \mu_c + \mu_f = K_1|\dot{x}| + K_2|\dot{x}|^2 \tag{3-12}$$

It is apparent that the viscosity of gas is ignorable in regard to the particle laminar flow. Thus, the mixture solid-gas flow viscosity μ_m is approximately equal to the particle laminar flow μ_p. Furthermore, the effective viscous damping force is formulated as

$$F_d = -\frac{1}{2}\rho_m S C_d |\dot{x}|\dot{x} = -c_{eq}\dot{x} \tag{3-13}$$

where c_{eq} is the equivalent viscous damping coefficient; S represents the cross-section area of damper cavity; ρ_m is the overall density of the mixture flow, which is expressed as

$$\rho_m = \left(1-\alpha_p\right)\rho_g + \alpha_p\rho_p \tag{3-14}$$

According to Sarpkaya (1986), the drag coefficient C_d in Equation (3-13) can be expressed as

$$C_d = \frac{fd\pi^3}{|\dot{x}|}\left(\frac{3}{2}\beta^{-1/2} + \frac{3}{2}\beta^{-1} - \frac{3}{8}\beta^{-3/2}\right) \tag{3-15}$$

where $\beta = \pi d^2 f \rho_m/\mu_m$; d represents the diameter of the cavity; and f is the fundamental frequency of the target structure. Thus, the equivalent viscous damping coefficient c_{eq} can be derived as

$$c_{eq} = c_1|\dot{x}|^{1/2} + c_2|\dot{x}| - c_3|\dot{x}|^{3/2} + c_{11}|\dot{x}| + c_{21}|\dot{x}|^2 - c_{31}|\dot{x}|^3 \tag{3-16}$$

where,

$$c_1 = 4\bar{c}\alpha^{1/2}f^{1/2} \qquad c_2 = 4\bar{c}\alpha \qquad c_3 = \bar{c}\alpha^{3/2}f^{-1/2} \tag{3-17}$$

$$c_{11} = 4\bar{c}{\alpha_1}^{1/2}f^{1/2} \qquad c_{21} = 4\bar{c}\alpha_1 \qquad c_{31} = \bar{c}{\alpha_1}^{3/2}f^{-1/2} \tag{3-18}$$

$$\bar{c} = (3/16)\pi^3 d^2 h\rho_m \qquad K_3 = \pi d^2\rho_m \tag{3-19}$$

$$\alpha = K_1/K_3 \qquad \alpha_1 = K_2/K_3 \tag{3-20}$$

In this study, the damping performance of granular material-filled PDs is characterized by the energy loss factor η, which is expressed as

$$\eta = \frac{E_{dissipated}}{E_{kinetic}} \tag{3-21}$$

where $E_{dissipated}$ is the energy dissipated by the drag force generated from particle damping, $E_{kinetic}$ indicates the total maximum kinetic energy for a steady-state harmonic vibration. The condition for the simulation case to obtain the low-fidelity energy loss factor η_l (subscript l denotes low-fidelity) is kept the same as in the experiment, i.e., the PD is connected rigidly to a dynamic exciter without any relative displacement, thus the system force of PD can be calculated by

$$F = m\ddot{x} + c_{eq}\dot{x} \tag{3-22}$$

Here, the dynamic exciter is vibrating in given frequency f and acceleration amplitude a_m. The PD's force would contain multi-frequency components due to its nonlinearity.

The energy loss factor can be determined by the power input method (Bustamante et al. 2016), where the expression of average power flow is

$$P = \frac{1}{2} F_{pk} V_{pk}^{*} \tag{3-23}$$

This quantity is calculated by the inner product of system force and system velocity in the frequency domain. For a single-harmonic signal, the root-mean-square (RMS) value is $1/\sqrt{2}$ of its peak value. In the acceleration-controlled harmonic tests, since the velocity has value only at frequency f in the frequency domain, the average power flow can be concerned only with the frequency f, and the average power flow can alternatively be written as

$$P = F_{rms}V_{rms}^{*} \tag{3-24}$$

where F_{rms} is the RMS of force (F) which is a complex number, and V_{rms}^{*} is the RMS of V*, which is the complex conjugate of the velocity. The real part of P is the value of active power, or “dissipated power”, which is dissipated during vibration; the imaginary part of P is the value of reactive power or “trapped power”, which is stored in the vibrating system by means of potential energy.

Since power is the time derivative of energy, Equation (3-21) can be rewritten as

$$\eta = \frac{Real(P)/\omega}{E_{kinetic}} \tag{3-25}$$

where $Real(P)$ is the real part of the average power flow P, and ω is the angular frequency.

3.3 Power Measurements of PDs

This section describes dynamic experiments on a number of granular material-filled PDs. The energy loss factor of the PDs is obtained based on the experimental results under various damper configurations and excitation conditions.

3.3.1 Experimental preparation

For the PDs used in this study, tungsten powder with a diameter of 0.2 mm was chosen as the grain to be filled in the damper cavity. The apparent density, instead of material density, is used to calculate the filling weight of the powder in the damper cavity according to the volume fraction. The apparent density of tungsten powder is 8.9 g/cm^3, which is obtained by measuring the mass of the sample with the pile by a precise digital scale (Figure 3-2(a)).

A cylindrically shaped cavity is selected as the damper body to reduce the side-effect from lateral contact between the particle bed and cavity wall since the vibration generated in the test is limited in vertical direction. To realize a customized PD cavity, the cavity is designed to be an assembly of several modules which are aluminum pipes with identical diameter and height. This assemblage design facilitates the fabrication of PDs with different cavity heights. In the experiments, three PDs were tested with the cavity height equal to 30 mm, 55 mm, and 80 mm (Figure 3-2(b)), respectively. The number of assembled modules in the three configurations is one (mode-1), two (mode-2), and three (mode-3).

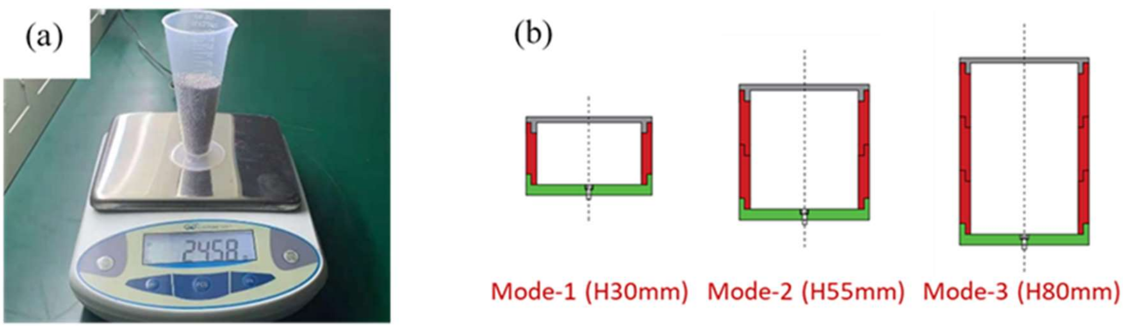


Figure 3-2 (a) Weighing of filling particles; (b) Schematic diagram of three cavity modes.

Before the experiments, all the damper cavity pieces under three cavity assembly modes were weighed. The weight of the damper cavity was later used to calculate the inertial force under acceleration excitation during the acceleration-controlled experiments. The weight and cavity dimension are listed in Table 3-2.

Table 3-2 Weight and cavity dimension in three modes

	Mode-1	Mode-2	Mode-3
Weight (g)	275.4	321.3	367.2
Cavity height (cm)	3.0	5.0	8.0
Cavity volume (cm^3)	48.11	88.20	128.29

The volume of particles to be filled in each cavity mode is determined according to the given filling volume fraction, and then the weight of required particles for each filling volume fraction is obtained for particle apparent density. Except for the empty cavity case that was tested for calibration, four particle volume fractions, namely 10%, 30%, 50%, and 70%, were considered in the test for each cavity mode. The weight of particles for different filling volume fractions is given in Table 3-3. The weight of the accelerometer placed on top of the cavity for the measurement of acceleration in the acceleration-controlled experiments is 3.0 gr.

Table 3-3 Weight of tungsten powder for different filling volume fractions (gr)

Volume fraction	Mode-1	Mode-2	Mode-3

10%	42.8	78.5	114.1
30%	128.5	235.6	342.6
50%	214.1	392.5	570.9
70%	299.7	549.4	799.2

3.3.2 Experimental apparatus and setup

The B&K LDS V-650 vibration shaker was used in the dynamic experiments, which can apply sinusoidal force up to 2.2 kN with the operation frequency ranging between 5 Hz and 4 kHz. The shaker is equipped with a built-in pneumatic support system that can accommodate payloads of up to 50 kg with full relative displacement. The damper was mounted on the shaker with an impedance head KD3001B, which acquires the force signal. The force limitation of the impedance head is 2.5 kN in association with the acceleration limitation of 100 g and working frequency ranging from 0.5 Hz to 6 kHz. The damper cavity was connected to the impedance head with an M5 screw. The accelerometer Dytran 3273 with its measurement range up to 100 g is mounted on the damper cavity top to monitor the vibration of the damper. The experimental apparatus is shown in Figure 3-3.

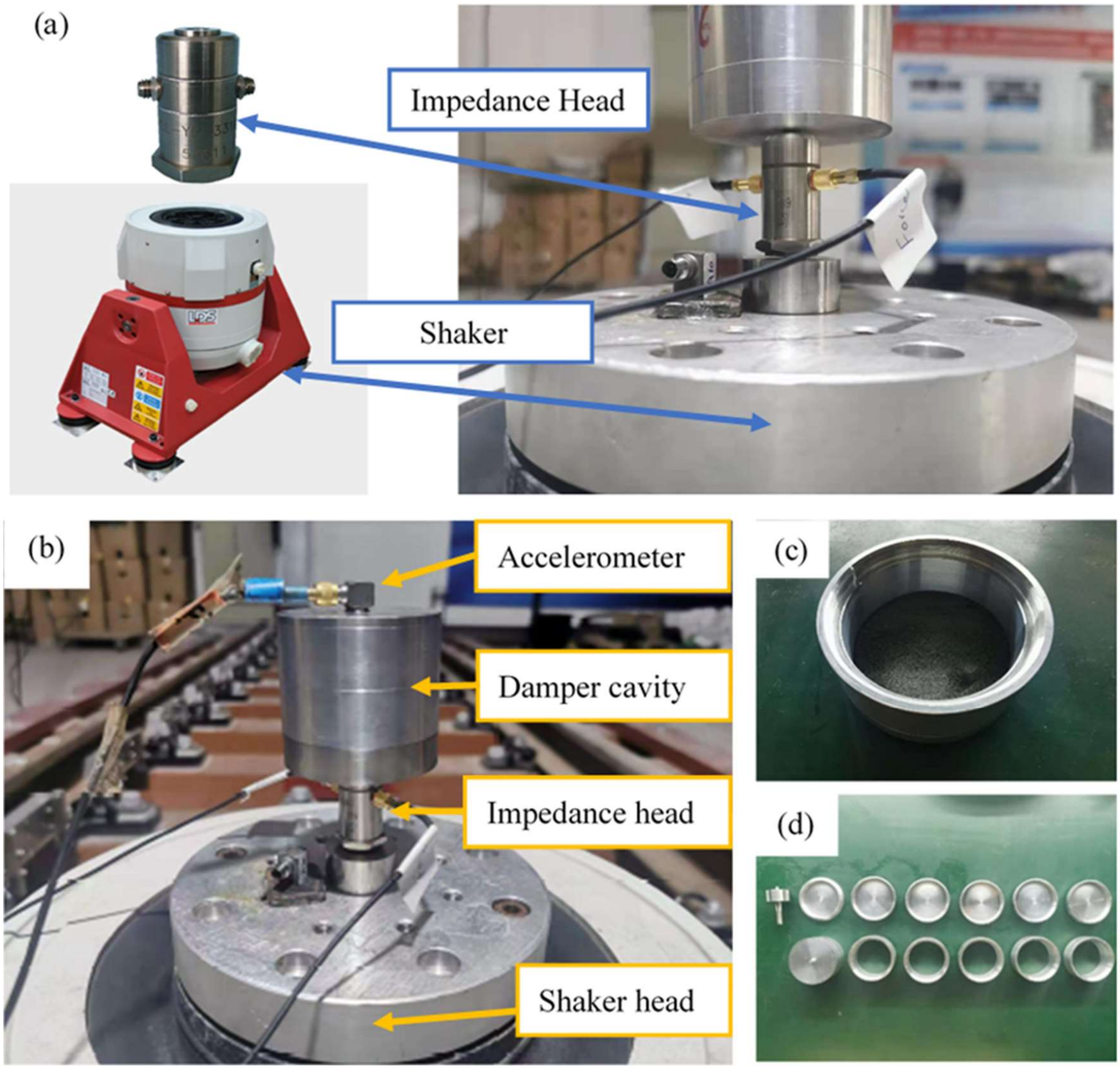


Figure 3-3 Experimental apparatus: (a) Assembly of impedance head and shaker; (b) Setup of cylindrical PD; (c) Particles filling in damper cavity; and (d) Damper assembly components.

Figure 3-4 illustrates the experimental setup. In the experiment, the PD is bolted onto the impedance head through an M5 screw. As the KD3001B impedance head has M5 thread at both ends, a stainless-steel connector is used to bolt the impedance head onto the shaker head as well. The electrical signal output from the signal generator is amplified by a power amplifier and then exerted to the vibration shaker. The shaker head motivated by an electromagnet coil thereby vibrates in accordance with the input signal. The amplitude of the vibration signal can be adjusted by switching the electrical power in the power amplifier. The acceleration signal from the accelerometer and force signal from the impedance head were collected by the DEWEsoft DAQ system. Sinusoidal signals were selected as input exerted to the shaker, with the excitation frequency varying from 100 to 2000 Hz with an interval of 100 Hz. The excitation level is controlled by acceleration amplitude to facilitate analysis and comparison. The peak acceleration level is set as 5 g to 25 g with an interval of 5 g. It is worth noting that since the peak acceleration value is manually controlled with the shaker's gain, it is impossible to control the peak acceleration value to a precise value. In view of this, excitation levels 1 to 5 will be used later to denote the five levels of acceleration amplitude around 5 g, 10 g, 15 g, 20 g, and 25 g, respectively. In each case, five seconds of the force and acceleration signals were recorded, but only the middle three seconds of the signals after removing the transient portions are used in the subsequent analysis.

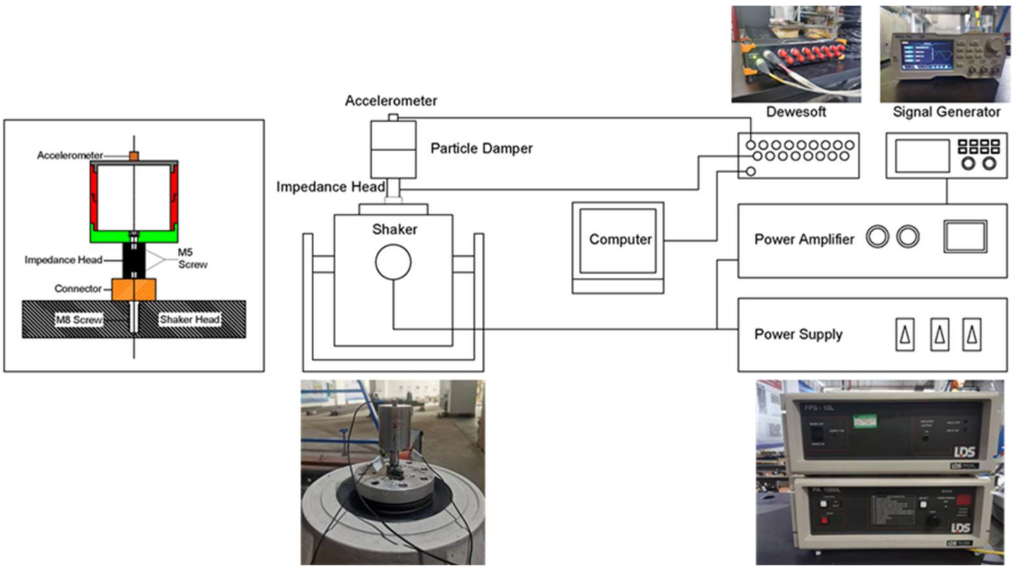


Figure 3-4 Schematic diagram of experimental setup.

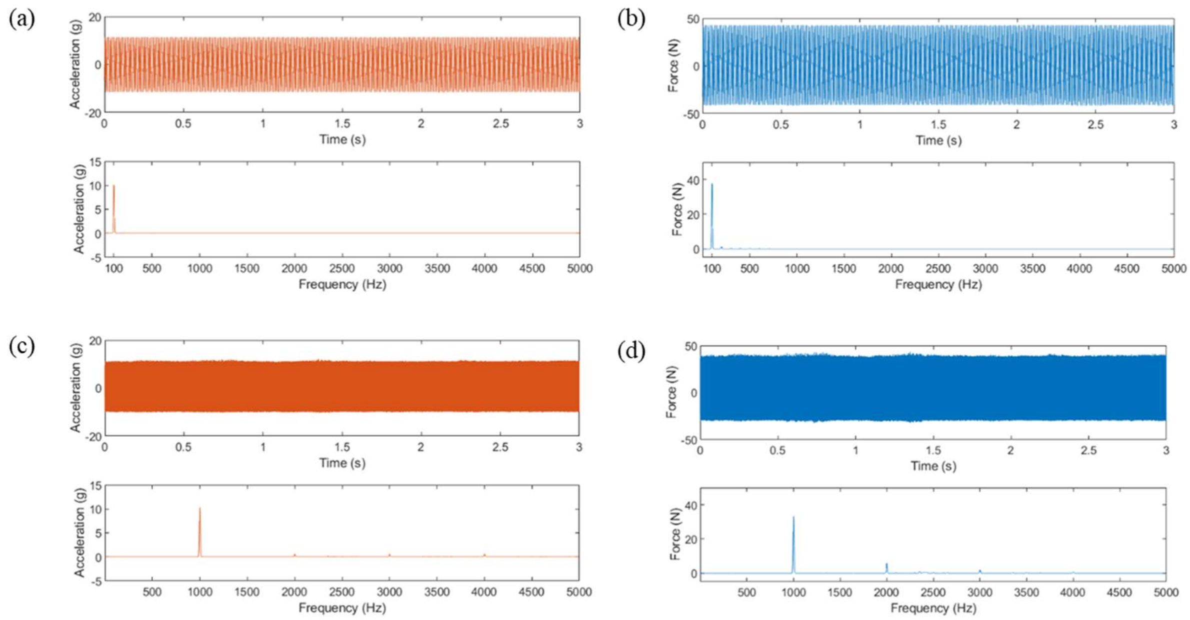


Figure 3-5 (a) Acceleration in time- and frequency-domains at f = 100 Hz; (b) Force in time- and frequency-domains at f = 100 Hz; (c) Acceleration in time- and frequency-domains at f = 1000 Hz; and (d) Force in time- and frequency-domains at f = 1000 Hz.

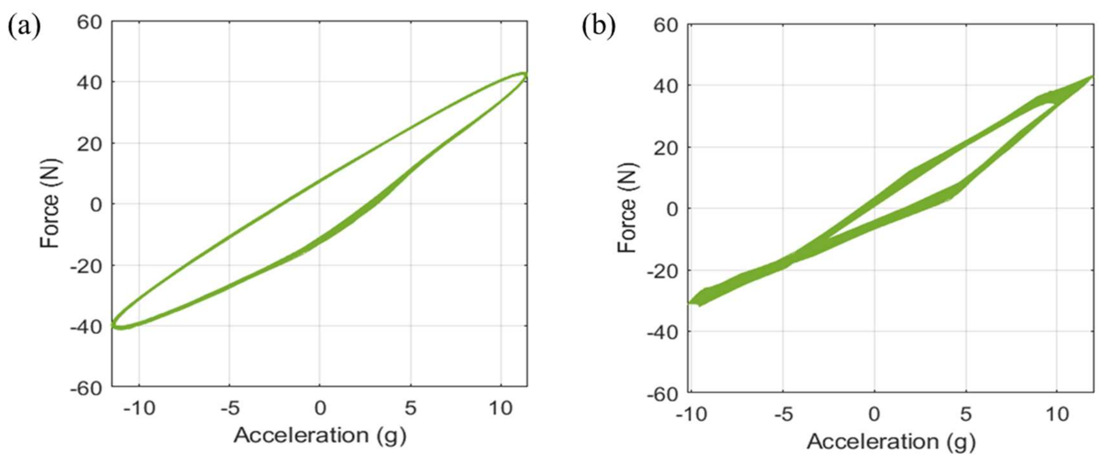


Figure 3-6 Diagram of force versus acceleration: (a) f = 100 Hz; (b) f = 1000 Hz.

Figure 3-5 illustrates the acceleration and force signals in both time and frequency domains when the acceleration amplitude is controlled around 10 g and the excitation frequency is 100 Hz and 1000 Hz, respectively, and **Figure 3-6** shows the diagrams of force versus acceleration at the same excitation level and the same excitation frequencies. It is seen that while a sinusoidal signal was generated by the signal generator as input to the shaker, the actual force signals contain multi-frequency components, especially in the case of high excitation frequency, due to the strong nonlinearity of the PD. However, the acceleration signals keep almost single-harmonic in the acceleration-controlled experiments. In the experiments, the PD was bolted directly to the shaker head without it mounting on any target structure. This test arrangement helps to understand the effects of PD's design parameters straightforwardly, but the conventional damping indicator, e.g., damping ratio, is inapplicable. In this regard, the energy loss factor η, which is also easy to be derived from the two-phase flow model, is used to represent the damper performance of the tested PDs from the measured acceleration and force signals by using Equations. (3-23) to (3-25).

3.4 Proof of Concept

3.4.1 Pre-training of DNN using two-phase flow model

The first step in implementing the proposed TL-based multi-fidelity modeling method is to pre-train a DNN using low-fidelity data generated from an approximate governing/constitutive equation. In this study, the two-phase flow equivalent viscosity model is used to generate such training data. To comply with the experimental data, the input vector of the DNN is set as $\mathcal{X} = \{\alpha_p, \rho_p, \gamma, a_m, f\}$, where α_p is the particle filling ratio, ρ_p is the particle apparent density, γ is the cavity height, a_m is the excitation (acceleration) amplitude, and f is the excitation frequency (Hz). The output parameter of the DNN is the energy loss factor η of PD.

Before training the DNN, its architecture (the number of hidden layers, $\mathcal{F}$, the size of each hidden layer, $\mathcal{C}$, and the activation functions) and optimization algorithm for DNN training should be selected. A DNN with deep layers and large neuron numbers would gain vigorous regression ability but tends to be over-fitting, especially when the size of the training dataset is small. To be fair in comparison with the DNN model that is trained using only the experimental data, we determine these hyper-parameters using the experimental data. In the experiments, α_p was taken as 10%, 30%, 50% and 70%, respectively (4 settings in total); ρ_p was constant (8.9 g/cm3) since only tungsten powder was used as the filling particles (1 setting in total); γ was equal to 30 mm, 55 mm and 80 mm, respectively (3 settings in total); a_m was ranging from 5 g to 25 g with an interval of 5 g (5 settings in total); and f was ranging from to 100 to 2000 Hz with an interval of 100 Hz (20 settings in total). As a result, the size of the experimental dataset is 1200. We apply the K-fold cross-validation (Jung, 2017) to determine the hyper-parameters. It first divides the training dataset into K parts with equal size. One part is used as the validation dataset, and the remaining $K - 1$ parts are used for training. By taking each part to be the validation dataset in turn, an average prediction error could be obtained to evaluate the model. In this study, the K-fold cross-validation is applied to the experimental dataset with K = 5. With the total size of experimental dataset being 1200, 80% of the experiment dataset with a size of

960 is used for training and validation while 20% of the data with a size of 240 is used for testing. The candidate number of hidden layers $\mathcal{F}$ is $\{3,4,5\}$, and the size of hidden layer $\mathcal{C}$ is kept the same for all hidden layers and is selected from $\{5,10,15\}$. By temporarily setting the activation functions for all layers as "ReLU" function, the average predicted mean squared error (MSE) after 50 training epochs for each selection is given in Table 3-4. The hyper-parameters $\{\mathcal{F}=3, \mathcal{C}=10\}$ give rise to the minimum MSE; therefore, the DNN architecture is set as 3 hidden layers with 10 neurons in each hidden layer.

Table 3-4 Average MSE under different DNN architectures after 50 training epochs

Number of hidden layers	$\mathcal{C}=5$	$\mathcal{C}=10$	$\mathcal{C}=15$
$\mathcal{F}=3$	3.7×10^{-5}	$\mathbf{2.4\times10^{-5}}$	2.9×10^{-5}
$\mathcal{F}=4$	4.5×10^{-5}	4.3×10^{-5}	3.4×10^{-5}
$\mathcal{F}=5$	4.3×10^{-5}	3.7×10^{-5}	3.6×10^{-5}

After determining the number of hidden layers ($\mathcal{F}$) and the size of each hidden layer ($\mathcal{C}$), we proceed to the selection of activation functions and optimization algorithm. We try the commonly used activation functions "Sigmoid", "tanh", and "ReLU" in different combinations (the activation function for the output layer is fixed as "ReLU" since the energy loss factor as output is always non-negative). In this study, the mini-batch technique (Li et al., 2014) will be used in the training of DNN. It divides the entire dataset into small batches, ensuring the stability of gradient descent and enhancing the calculation efficiency. One of the most commonly used optimization algorithms is the stochastic gradient descent (SGD) (Bottou, 2012; Ketkar, 2017), which achieves enhanced training efficiency compared to the batch gradient descent (BGD). In combination with the mini-batch technique, the SGD can ensure high accuracy in gradient computation. Another well-known optimization algorithm is the adaptive moment estimation (Adam) (Kingma & Ba, 2014). This algorithm combines the ability to deal with sparse gradients and non-stationary objectives and has been proven to work well in practice. We compare SGD and Adam algorithms along with different combinations of the three activation functions. The comparison is made in terms of two indices: MSE loss and model accuracy. The latter is defined as the ratio between the amount of successful prediction and the total size of the testing dataset. Here, successful prediction implies that the predicted energy loss factor error is less than 10% of the true value. The performance of each setting after 50 training epochs is shown in Table 3-5. It is concluded that the activation functions of "tanh-tanh-ReLU" for the three hidden layers together with Adam algorithm achieves the best performance. This configuration in conjunction with the specified hyper-parameters $\{\mathcal{F} = 3, \mathcal{C} = 10\}$ constitutes the architecture of the DNN (Figure 3-7).

Again, to be fair in comparison with the DNN model that is trained using only the experimental data, the numerical low-fidelity data used to train the physics-guided DNN will be generated from the two-phase flow equivalent viscosity model with the values of all input parameters not exceeding their corresponding ranges explored in the experiments, but sampled with a higher resolution because these training data can be cheaply generated by numerical computation.

Specifically, in generating the numerical training data, α_p is taken from 10% to 70% at intervals of 5% (13 settings in total); ρ_p is still taken as a single value equal to 8.9 g/cm3 (1 setting in total); γ is taken as 30 mm, 55 mm, and 80 mm same as in the experiments (3 settings in total); a_m is taken from 5 g to 25 g at intervals of 1 g (21 settings in total); and f is taken from 100 to 2000 Hz at intervals of 50 Hz (39 settings in total). As a result, the size of the numerical dataset is 31941, of which 3000 is used for training and the rest is used for validation. For each set of given input parameters, the energy loss factor is calculated from the two-phase flow equivalent viscosity model. Figure 3-8 shows the calculated values of energy loss factor under different acceleration levels and excitation frequencies. To eliminate the scale difference among different parameter features, the input dataset is normalized by the Z-score method (Gopal et al., 2015) before being presented to train the physics-guided DNN. The dataset after normalization will have a mean value of 0 and a standard deviation of 1. After completing the training of the DNN, all the weights and biases in this pre-trained DNN model (Figure 3-7) are specified.

Table 3-5 Comparison of different activation functions and optimization algorithms

Hidden layer 1 (10 neurons)	Hidden layer 2 (10 neurons)	Hidden layer 3 (10 neurons)	Optimization algorithm	MSE (10^{-4})	Model accuracy (%)

Sigmoid	Sigmoid	ReLU	Adam	57	87.2
tanh	tanh	ReLU	Adam	**24**	**94.4**
ReLU	ReLU	ReLU	Adam	27	93.9
tanh	tanh	tanh	Adam	30	92.6
Sigmoid	Sigmoid	ReLU	SGD	50	90.5
tanh	tanh	ReLU	SGD	59	86.4
ReLU	ReLU	ReLU	SGD	81	82.8
tanh	tanh	tanh	SGD	50	89.6

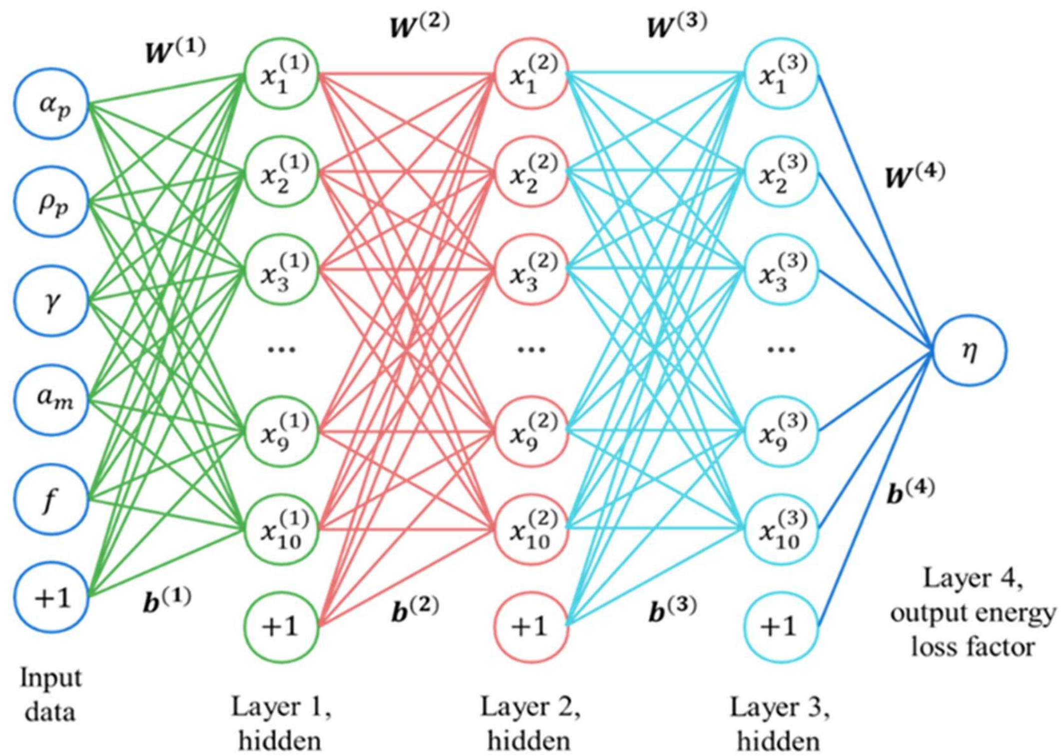


Figure 3-7 Architecture of DNN model.

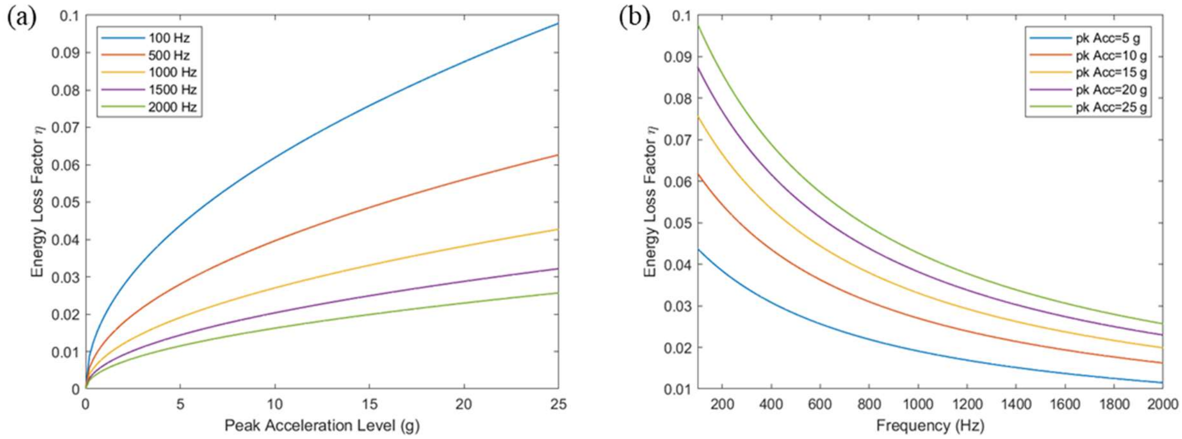


Figure 3-8 Calculated values of the energy loss factor η of PD using the two-phase flow equivalent viscosity model: (a) η versus peak acceleration under various excitation frequencies; (b) η versus excitation frequency under various peak accelerations.

3.4.2 Formulation of multi-fidelity model using deep TL

The pre-trained DNN, which was trained using ample numerical data (source dataset $\mathcal{D}_S$) from the two-phase flow equivalent viscosity model, is now refined using actual experimental data (target dataset $\mathcal{D}_T$) to formulate a multi-fidelity DNN model with the same architecture as the pre-trained DNN. In compliance with the deep TL philosophy, all low-level layers of the DNN are frozen with fixed weights and biases same as in the pre-trained DNN, while the last hidden layer and the output layer are released with their weights and biases to be re-trained using the experimental data. Recall that the size of the experimental dataset is 1200 (of which 80% is used for training and validation and 20% for testing). For verification of the formulated multi-fidelity DNN model, a new single-fidelity DNN model with the same architecture is also trained using only the experimental dataset. The former capitalizes on TL, while the latter does not. Figure 3-9 provides a comparison of training process between the two DNN models in terms of MSE and model accuracy. It is observed that the initial MSE loss of the DNN with TL (multi-fidelity DNN model) is already lower than that of the DNN without (marked as W/O) TL (single-fidelity DNN model) at the beginning of the network training. Moreover, the DNN with TL converges much faster than the DNN without TL. For any number of training epochs, the model accuracy achieved by the DNN with TL is much better than that achieved by the DNN without TL. What is more, the DNN with TL can reach stable model accuracy after a small number of training epochs, and the model accuracy finally achieved by the DNN with TL after convergence is much higher than that achieved by the DNN without TL. As shown in Figure 3-9, the model accuracy of the DNN with TL after 50 training epochs reaches 94.4%, while the model accuracy of the DNN without TL is only about 75%.

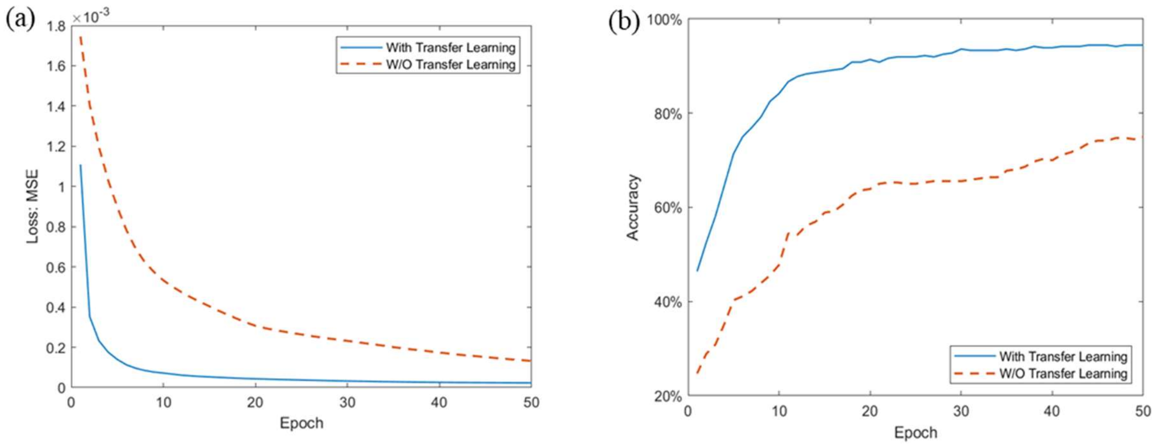


Figure 3-9 (a) Mean squared error (MSE) of DNN models with TL and W/O TL; (b) Model accuracy with TL and W/O TL.

After formulating the two models, all 1200 input features explored in the experiments are fed into the models to obtain the predicted values of the energy loss factor. Figure 3-10 shows a comparison between the experimentally obtained energy loss factors and the prediction results obtained by the DNN model with TL and the DNN model without TL, respectively. The R2 score, which indicates the total variance explained by a model, reaches 0.9885 for the DNN model with TL and 0.9675 for the DNN model without TL. A comparison of average MSE on the predicted energy loss factor generated by the two models within different frequency ranges is given in Figure 3-11. It is seen that in all frequency ranges of interest, the average MSE generated by the DNN model with TL is largely less than that generated by the DNN model without TL. In particular, the former achieves only about one tenth of the MSE generated by the latter in the high frequency region of 1600 to 2000 Hz (the average MSE decrement ratio for the entire frequency range from 100 to 2000 Hz is 19.4%). In summary, the DNN model with TL is much superior to the DNN model without TL.

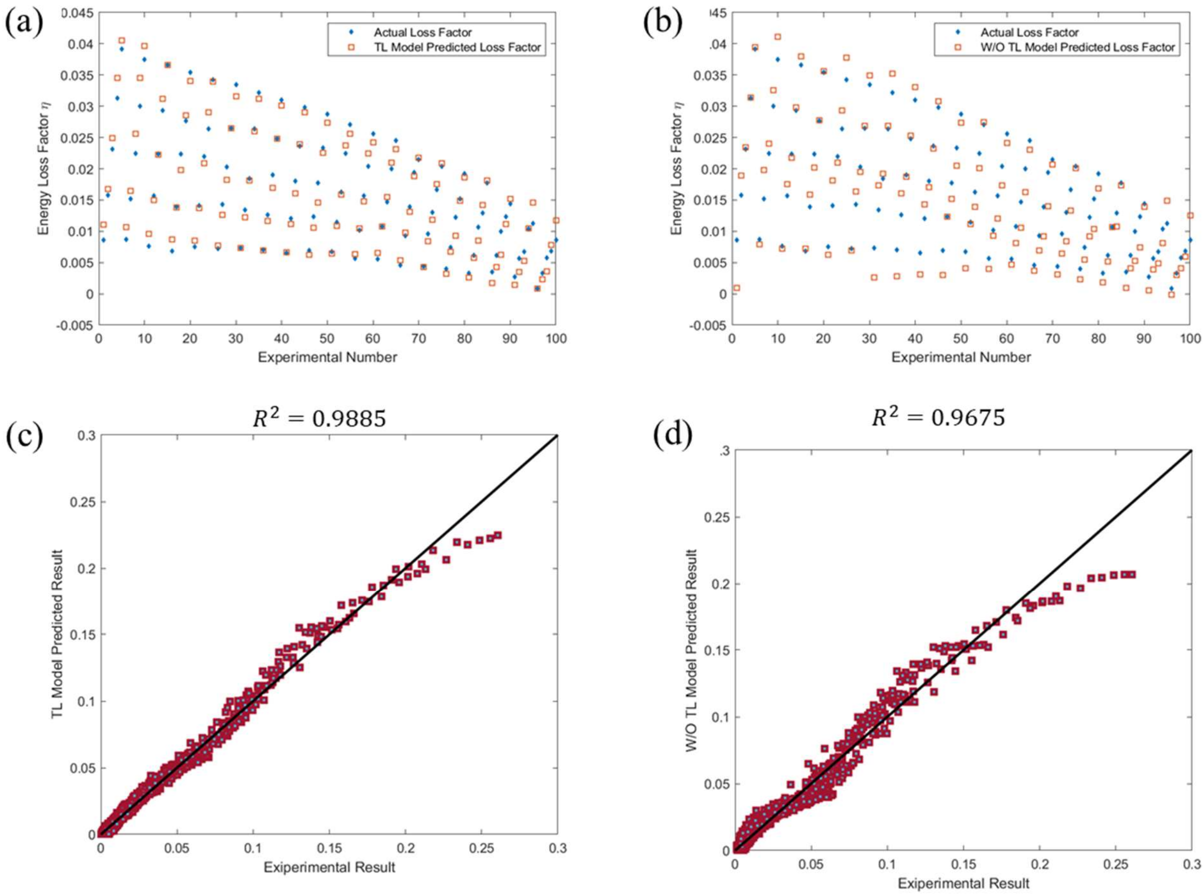


Figure 3-10 Prediction of energy loss factor by (a) TL approach, and (b) W/O TL approach; R^2 score of the prediction results from (c) TL approach, and (d) W/O TL approach.

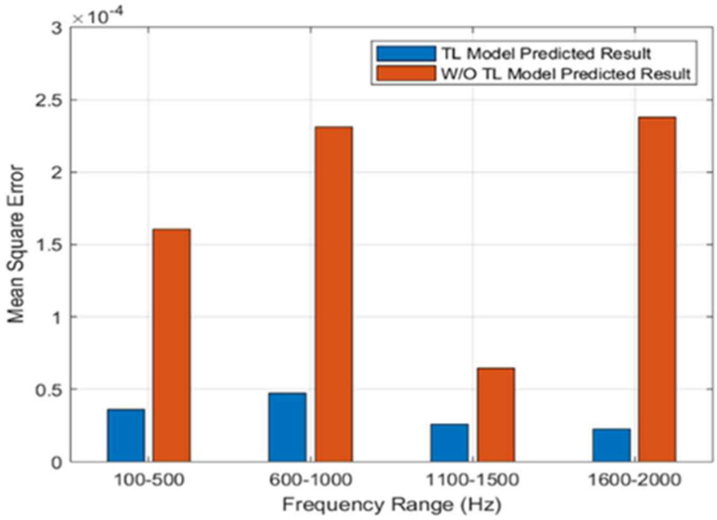


Figure 3-11 Comparison of average mean squared error (MSE) generated by the two models in different frequency ranges.

3.4.3 Effect of damper parameters on damping performance

With the formulated multi-fidelity DNN model of high accuracy, we investigate the effect of damper parameters on the damping performance in this section, where the prediction results are also compared with the experimentally obtained values. Figure 3-12 illustrates the energy loss factor versus excitation level (acceleration amplitude) for mode-1 cavity under different excitation frequencies and particle filling ratios. From the results under different filling ratios, it is seen that the PD with more filling particles in a cavity with fixed volume leads to a larger energy loss factor. This is quite reasonable since, with the increase of filling particles, the friction effect of the particle bed would increase. Furthermore, it is discovered that the energy loss factor decreases with the excitation frequency. This can be explained by the fact that in high-frequency vibration, the velocity and displacement amplitudes are lower than those in low-frequency vibration for the same acceleration level in acceleration-controlled tests. Hence, even though the acceleration level is identical, the motion of the cavity that induces the friction effect is lower in a higher frequency scenario.

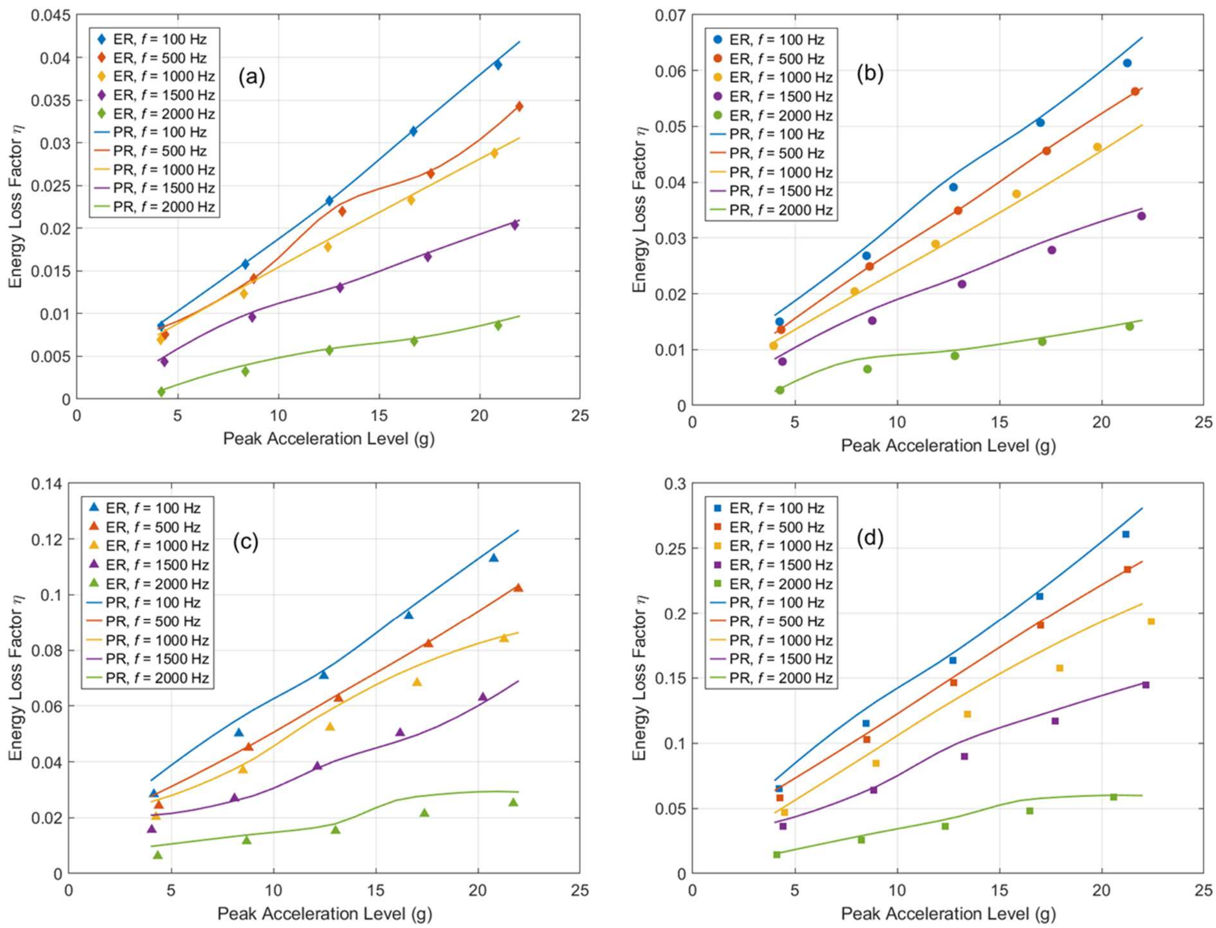


Figure 3-12 Energy loss factor versus excitation level (acceleration amplitude) for mode-1 cavity under filling ratio of: (a) 10%; (b) 30%; (c) 50%; (d) 70%. Here ER stands for experimental results, and PR stands for predicted results.

Excitation frequency is a factor significantly affecting the damping performance of PDs. The energy loss factor obtained under a broad frequency band from 100 to 2000 Hz is compared under the same acceleration level. Since it is difficult to maintain the acceleration amplitude to be exactly the same under different excitation frequencies in the experiments (the acceleration level was controlled by manually adjusting the shaker's gain to the nearest level), we adopt the term "excitation level" instead of providing the exact acceleration amplitudes. Figure 3-13 shows the energy loss factor versus excitation frequency for mode-1 cavity under different excitation levels and particle filling ratios. With the filling ratio from 10% to 70%, different excitation levels lead to distinct values of the energy loss factor. The relationship between the energy loss factor and the excitation frequency is nonlinear. It can be found that the energy dissipated by high-frequency vibration is less than that by low-frequency vibration at the same excitation level (acceleration amplitude). At higher frequencies, the difference of the energy loss factor induced by different excitation levels is reduced.

A comparison of the damping performance of the three cavity modes, i.e., the cavity mode-1 with a height of 30 mm, cavity mode-2 with a height of 55 mm, and cavity mode-3 with a height of 80 mm, is explored. We have obtained the energy loss factor versus excitation level for the three cavity modes under excitation frequency from 100 to 2000 Hz, Figure 3-14 and Figure 3-15 show the results when the excitation frequency is 500 and 1500 Hz, respectively. Overall, the energy loss factor decreases with the cavity height. However, since the volume of cavity is larger for the taller cavity, more particles are filled in the PD under the same volume fraction so that the particle bed has a larger mass, and in this sense, the net energy dissipated in the taller cavity is larger. The reason why damper with a higher cavity height produces lesser energy loss factor can be explained by the observation that the particle bed with large mass is not easy to be excited under the same excitation level, thus resulting in a reduced friction effect.

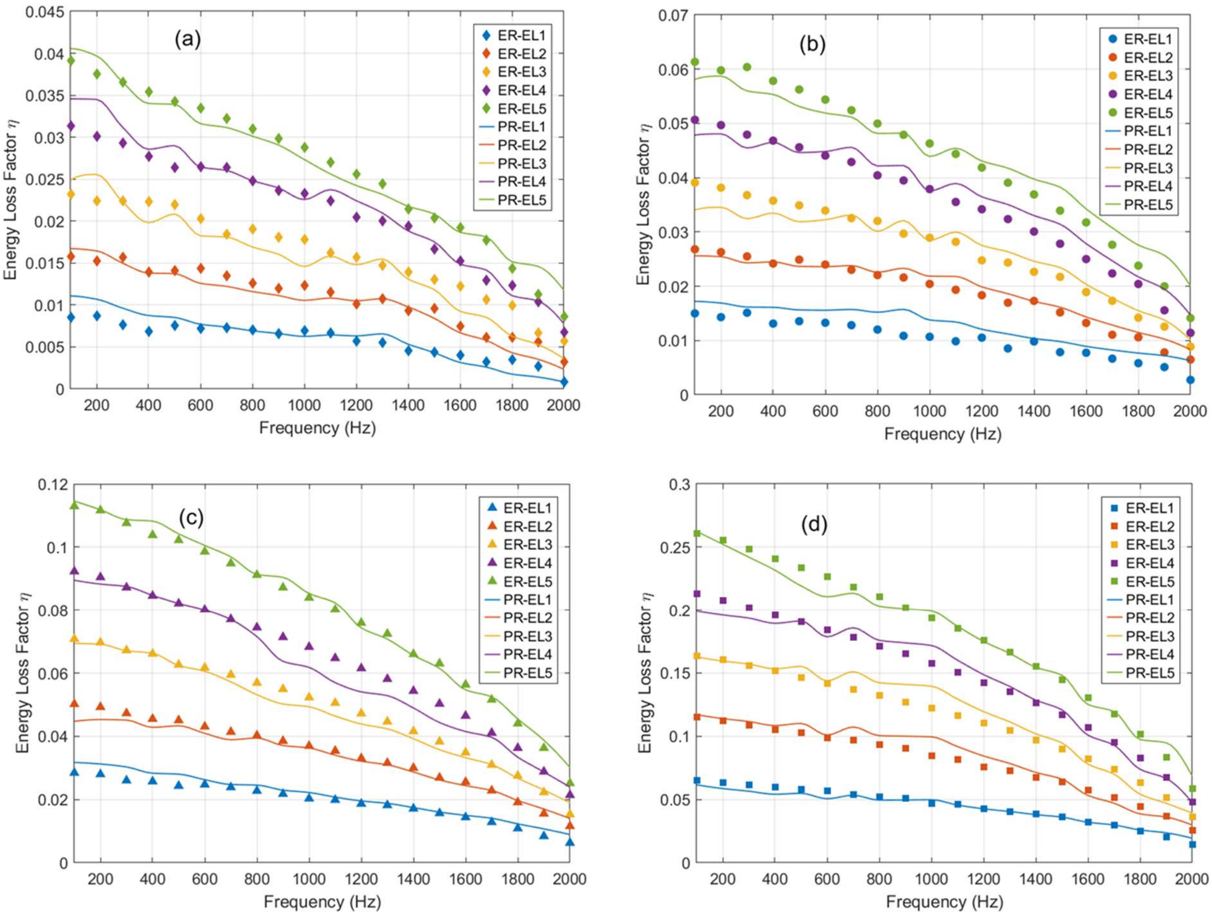


Figure 3-13 Energy loss factor versus excitation frequency for mode-1 cavity under filling ratio of: (a) 10%; (b) 30%; (c) 50%; (d) 70%. Here ER stands for experimental results, PR stands for predicted results, and EL is the abbreviation of excitation level.

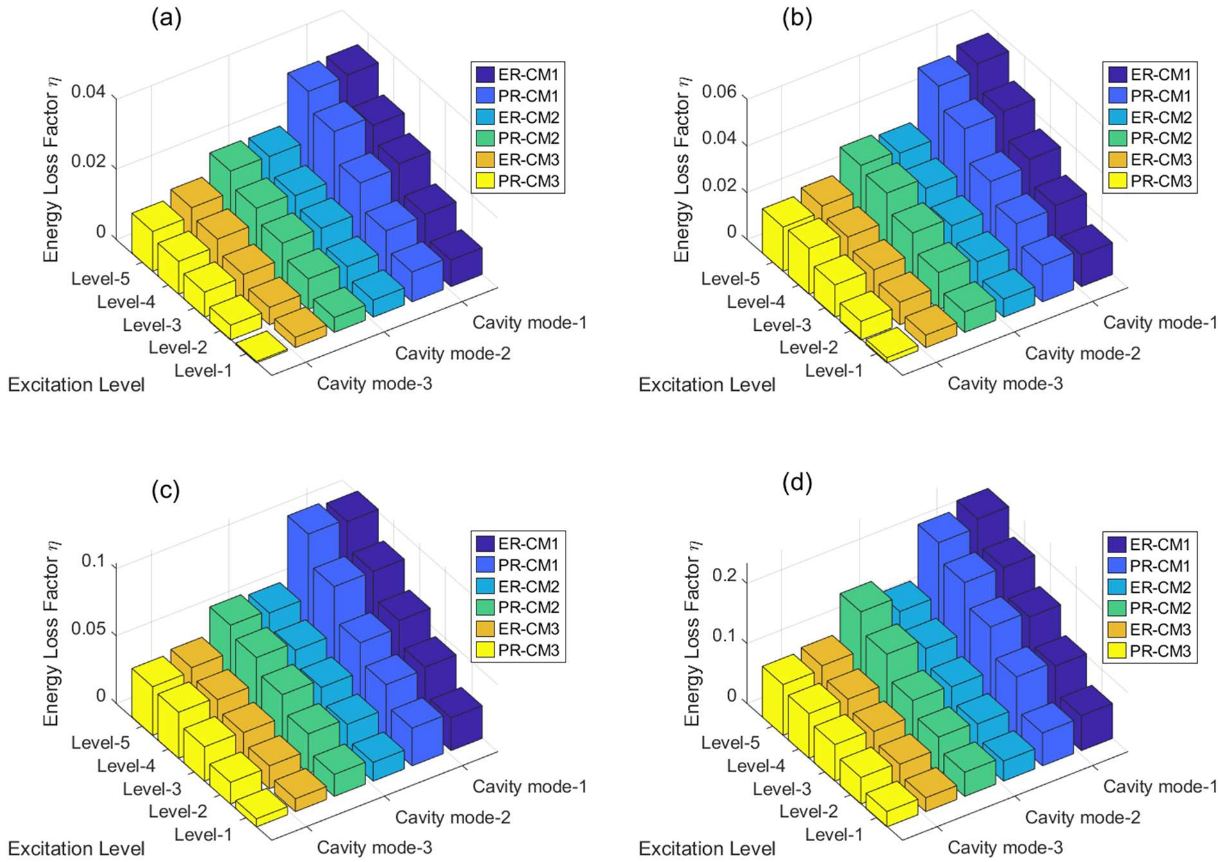


Figure 3-14 Energy loss factor versus excitation level for three cavity modes under excitation frequency of 500 Hz and filling ratio of: (a) 10%; (b) 30%; (c) 50%; (d) 70%. Here ER stands for experimental results, PR stands for predicted results, and CM is the abbreviation of cavity mode.

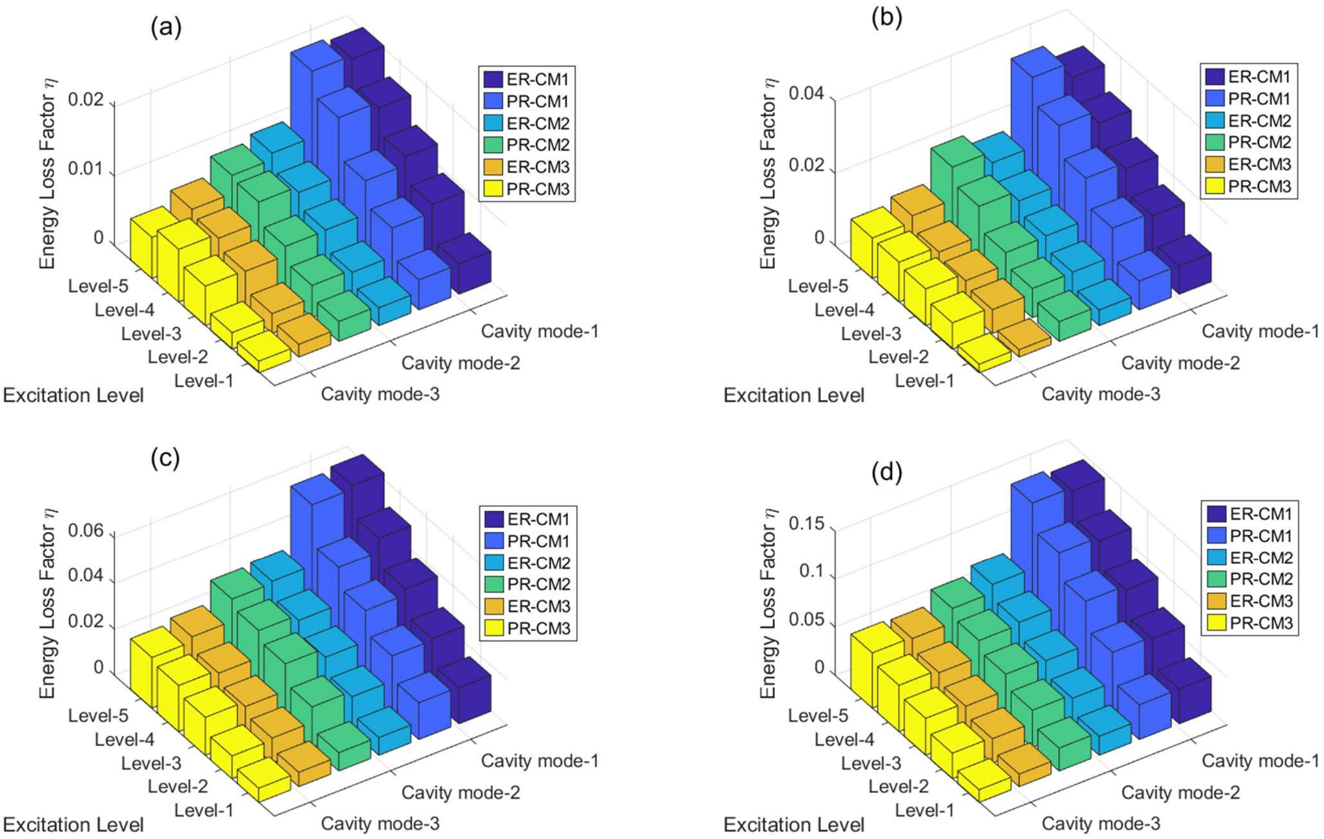


Figure 3-15 Energy loss factor versus excitation level for three cavity modes under excitation frequency of 1500 Hz and filling ratio of: (a) 10%; (b) 30%; (c) 50%; (d) 70%. Here ER stands for experimental results, PR stands for predicted results, and CM is the abbreviation of cavity mode.

3.5 Summary

In this section, a novel transfer learning (TL)-based multi-fidelity modeling method in the framework of deep neural network (DNN) was proposed for characterizing the dynamic performance of a set of granular material-filled particle dampers (PDs) with strong nonlinearity. The proposed method leverages the knowledge from an approximate governing/constitutive equation characterizing the PDs and that from experiments on the PDs. On account of the data-hungry nature of DNNs, the approximate governing/constitutive equation generates low-fidelity numerical data which supplement high-fidelity experimental data in compliance with the deep TL philosophy. A physics-guided DNN is first trained by using the low-fidelity yet plentiful data generated from the approximate governing/ constitutive equation, which inherits common underlying physics features in its low-level layers. Then the DNN is refined using the high-fidelity yet possibly scarce experimental data by freezing the low-level layers and fine tuning the hyper-parameters in the high-level layers (outermost layers), resulting in a multi-fidelity DNN model.

The proposed multi-fidelity modeling method was applied to develop a model for a set of granular material-filled PDs with the use of a two-phase flow equivalent viscosity model and experimental data. The model establishes a nonlinear mapping between the characteristic features (including particle, cavity, and excitation properties) and the damper performance in terms of energy loss factor. The performance and capability of the formulated multi-fidelity DNN model were verified by comparison with the experimental data and with a single-fidelity DNN model that was trained using only the experimental data. The investigation comes to the following conclusions:

1. The multi-fidelity model can accurately characterize the damping performance of the PDs across a broad frequency band from 100 to 2000 Hz and for a range of vibration amplitudes. The multi-fidelity model predicts the PD performance with 94.4% accuracy

and 2.3×10^{-5} MSE, while the single-fidelity DNN model trained using the same experimental data can only achieve 75.0% accuracy and 1.3×10^{-4} MSE. In particular, in the high frequency range of 1600 to 2000 Hz (it corresponds to micro-displacements in the acceleration-controlled case), the MSE arising from the multi-fidelity model is only one tenth of that generated by the single-fidelity model. The deep TL strategy benefits a lot to the enhancement of performance.

2. The two-phase flow equivalent viscosity model is appropriate to generate low-fidelity data for multi-fidelity modeling of the granular material-filled particle dampers (PDs) under various particle, cavity, and excitation properties. Incorporating the low-fidelity data generated by this approximate model can significantly improve the accuracy of modeling and prediction.
3. The formulated multi-fidelity model can accommodate continuously varying particle, cavity, and excitation features in the input parameter space; in particular, it enables both damper properties (particle and cavity features) and excitation properties to be cast into the input vector. As a result, the developed model is amenable to the optimal design of PD's parameters.

In this study, the behavior of PDs subjected to single-frequency harmonic excitations under a broad frequency band was investigated. Whilst in other situations, especially in practical problems, such ideal excitation mode seldom appears. Dynamic loads encountered in actual applications often contain multiple frequency components. The principle of linear combination on the response spectra of single-frequency modes is not applicable to PDs because of their highly nonlinear nature. However, to progressively evaluate and interpret the intricate highly nonlinear effects, the findings documented in this paper can serve as an intermediate step towards addressing the PD design problem under arbitrary excitation conditions. The particle damping characteristics under harmonic excitations can be considered as low-fidelity knowledge to other complex problems, with which TL or other multi-fidelity approaches can be executed to elicit high-fidelity models for the complex situations without needing much extra experiment data, e.g., adding a compensating DNN to explore the linear/nonlinear relationship between the multi-fidelity results.

CHAPTER 4.
Modeling Hysteresis of a PD: *ff*-TLPINN for High-frequency Feature Recognition

4.1 Introduction

The previous chapter has preliminarily explored the feasibility of adopting the NN to model the PD with limited experimental data. However, in the view of PD's response, the energy loss is only an indirect factor of a PD. Following the previous work, Veeramuthuvel et al. (2016) introduced a framework for modeling PDs using NNs. In their research, they attached a particle damper capsule to a printed circuit board (PCB). They developed a backpropagation NN and a radial basis function NN based on the experimental database, and the established NN models were able to accurately predict the acceleration response of the PCB under simple harmonic excitation. Similarly, Wang and Wu (2015) also used back-propagation NN and incorporated genetic algorithms to determine the optimal damper design parameters for attenuating the vibration of a cantilever beam.

The challenge of adopting NNs to model PDs lies in the conflict that arises when there is a shortage of high-fidelity experimental data, making it difficult to provide the necessary resources to train the data hungry NNs. To address this issue, deep transfer learning (TL) was introduced in the previous chapter, which aimed to establish a multi-fidelity modeling approach. The TL utilizes a trained model's knowledge to nurture the learning of a new dataset (Wang et al., 2023). Despite the previously mentioned limitations of NNs, there is growing evidence of another problematic behavior exhibited by deep fully connected networks, which is known as spectral bias (Rahaman et al., 2018; Basri et al., 2019; Cao et al., 2019; Xu et al., 2019). The NNs are hindered in recognizing solution that contains high-frequency components, which has profoundly restricted the performance of NNs. Since PD modeling involves a broad frequency band. Thus, NNs require a collaborative effort to address this limitation for PD modeling. Recent studies (Li et al., 2020; Liu et al., 2020; Moseley et al., 2021) have attempted to alleviate spectra bias by scaling input features to approximate high-frequency components. Wang et al. (2021) introduced an alternative approach by analyzing the fundamental weakness of NNs using the theory of neural tangent kernel (NTK) (Jacot et al., 2018). The observation indicates that NNs have a bias toward learning the dominant frequencies of the most significant eigenvectors of the NTK. As a result, researchers have proposed the use of Fourier features embedding to customize the NTK of the fully connected networks (Tancik et al., 2020; Wang et al., 2021). The Fourier NN successfully elevated the performance of NNs.

Based on the findings of these recent studies, an attempt is made in this work to model the hysteretic response force of a PD under a shaker's excitation in a wide frequency bandwidth (100-2000 Hz). Two forms of excitation, i.e., steady-state simple harmonic excitation and sweep-sinusoidal excitation, are applied to a PD that filled with granular material (tungsten powder). The response force is investigated to reveal the hysteretic behavior of PD. To model PD, the Fourier NN is examined and discussed its potential to improve training accuracy. Additionally, the effectiveness of wavelet basis features embedding was evaluated in this study. These approaches formulate the Fourier/Wavelet neural network (*ff/wf*NN) method. Afterwards, this study presents an investigation on the *ff*NN coping with the former proposed deep TL.

4.2 Establishment of *ff*-TLPINN

4.2.1 Principles of NTK

Spectral bias is a ubiquitous pathology observed in neural network training. The gradient flow of NN becomes increasingly stiff when the objective function exhibits a high-frequency character. Such a bias is also associated with a broad input feature domain. As the input feature domain expands, the normalization of the input variables effectively turns low-frequency features into high-frequency features. This bias mentioned has an impact on the output of NNs right from their initialization stage. The derivation given in (Wong et al., 2021) proved when the parameters of a fully connected network $\boldsymbol{\theta}$ are i.i.d. sampled from Gaussian distribution $\mathcal{N}(0,1)$, the output of the initialized network is surely a flat function. This implies that an initialized NN's output function already exhibits characteristics of low frequency.

A novel insight is built on this limitation of NNs through the perspective by analyzing their NTK (Wang et al., 2021). According to Jacot et al. (2018), the continuous gradient descent flow is demonstrated with an infinitesimal learning rate η (gradient flow). Then, the gradient flow is defined by the neural tangent kernel operator $\boldsymbol{K}$ as follows:

$$\boldsymbol{K}_{ij} = \boldsymbol{K}(x_i, x_j) = \langle \frac{\partial f(x_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \frac{\partial f(x_j, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \rangle \tag{4-1}$$

For illustrating the effect of the neural tangent kernel operator $\boldsymbol{K}$, a one-hidden-layer NN (Figure 4-1) with m hidden neurons are considered, biases are included in the input vector a and output vector b. The single output of the NN is $y = f(\boldsymbol{x}, \boldsymbol{\theta})$ parameterized by $\boldsymbol{\theta}$.

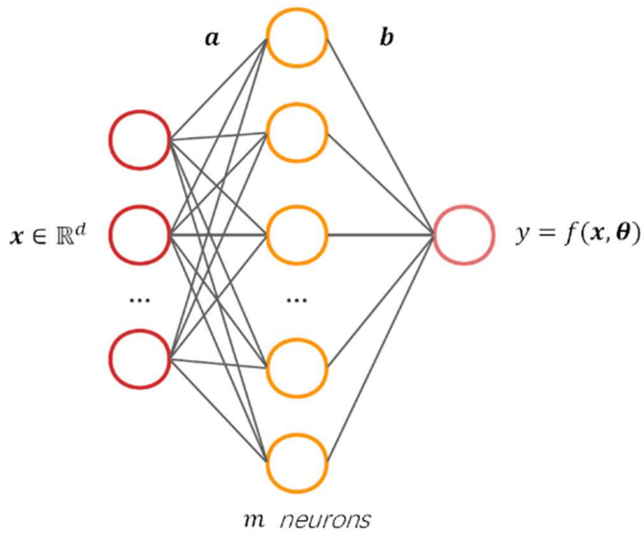


Figure 4-1 One-hidden-layer NN with m hidden neurons.

From this structure, the output of the network can be expressed as:

$$y = \sum_{i=1}^{m} b_i \delta(a_i^T x) \tag{4-2}$$

where $\delta(\cdot)$ is the activation function. Training the network is to adjust the output to meet the labeled data $\widehat{\boldsymbol{y}} \in \mathbb{R}^n$. Normally, the training is controlled by the L2 norm loss function, which can be flexibly written in (auxiliary coefficient $\frac{1}{2}$ is implemented to comfort the following derivation):

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{2}\sum_{i=1}^{n}(f(x_i, \boldsymbol{\theta}) - \hat{y}_i)^2 \tag{4-3}$$

Consider the gradient descend procedure of the network parameters, which in noted as:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta_t \nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}_t) \tag{4-4}$$

$$\Rightarrow \boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta_t (f(\boldsymbol{x}, \boldsymbol{\theta}) - \hat{\boldsymbol{y}}) \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta}_t) \tag{4-5}$$

During training of the network, the input $\boldsymbol{x}$ remains stable, and only the parameter space $\boldsymbol{\theta}_t$ is updated every iteration. If a gradient flow of the $f(\boldsymbol{x}, \boldsymbol{\theta})$ with an infinitely small learning rate η_t is defined, it is in fact a flow of the $\boldsymbol{\theta}_t$. The rate of the $\boldsymbol{\theta}_t$ is given by:

$$\frac{d\boldsymbol{\theta}}{dt} = \frac{\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_t}{\eta_t} = -(f(\boldsymbol{x}, \boldsymbol{\theta}) - \hat{\boldsymbol{y}}) \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta}_t) \tag{4-6}$$

Also, the first order Taylor expansion of $f(\boldsymbol{x}, \boldsymbol{\theta})$ around the initial parameter value $\boldsymbol{\theta}_0$ is considered:

$$f(\boldsymbol{x}, \boldsymbol{\theta}) \approx f(\boldsymbol{x}, \boldsymbol{\theta}_0) + \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta}_0) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \tag{4-7}$$

Therefore, the dynamic of this network output is:

$$\frac{df(\boldsymbol{x}, \boldsymbol{\theta})}{dt} \approx \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta}_0) \cdot \frac{d\boldsymbol{\theta}}{dt} \tag{4-8}$$

$$\Rightarrow \frac{df(\boldsymbol{x}, \boldsymbol{\theta})}{dt} \approx -\nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta}) \cdot \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta}_0)^T (f(\boldsymbol{x}, \boldsymbol{\theta}) - \hat{\boldsymbol{y}}) \tag{4-9}$$

As derived by Jacot et al. (2018), the term $\nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta})$ becomes deterministic and remains static when the network is infinitely wide. Even not like this, the parameters also hardly vary for an over-parameterized network during training, which is the so-called "lazy training" (Chizat, Oyallon, and Bach 2018). Under such circumstances, $\nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta}_0) \approx \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}, \boldsymbol{\theta})$ can be assumed. Since the neural tangent kernel (NTK) operator $\boldsymbol{K}$ is defined as:

$$\boldsymbol{K}_{ij} = \boldsymbol{K}(x_i, x_j) = \langle \frac{\partial f(x_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \frac{\partial f(x_j, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \rangle \tag{4-10}$$

The Equation (4-9) can be reformed as:

$$\frac{df(\boldsymbol{x}, \boldsymbol{\theta})}{dt} \approx -\boldsymbol{K} \cdot (f(\boldsymbol{x}, \boldsymbol{\theta}) - \hat{\boldsymbol{y}}) \tag{4-11}$$

where $f(\boldsymbol{x}, \boldsymbol{\theta})$ is the output of a fully connected NN that parameterized by $\boldsymbol{\theta}$, and $\boldsymbol{x} \in \mathbb{R}^d$ is the input to the NN. The training of the network is focused on the parameters $\boldsymbol{\theta}$ since $\boldsymbol{X_{train}}$ is stable. In that case, with the asymptotic conditions (Lee et al., 2017), the dynamic of the $f(\boldsymbol{x}, \boldsymbol{\theta})$ can be derived by:

$$\frac{df(\boldsymbol{X_{train}}, \boldsymbol{\theta}(t))}{dt} \approx -\boldsymbol{K} \cdot (f(\boldsymbol{X_{train}}, \boldsymbol{\theta}(t)) - \boldsymbol{Y_{train}}) \tag{4-12}$$

It directly follows that:

$$f(\boldsymbol{X_{train}}, \boldsymbol{\theta(t)}) \approx (I - e^{-\boldsymbol{K}t}) \cdot \boldsymbol{Y_{train}} \tag{4-13}$$

Where $\boldsymbol{X_{train}} = (x_i)_{i=1}^N$ are inputs and $\boldsymbol{Y_{train}} = (y_i)_{i=1}^N$ are the corresponding outputs, N is the dataset size. $\boldsymbol{K}$ is a semi-definite positive kernel, so $e^{-\boldsymbol{K}t}$ can be written into $\boldsymbol{Q^T} e^{-\boldsymbol{\Lambda}t} \boldsymbol{Q}$, where the matrix $\boldsymbol{Q}$ is orthogonal. The ith column $\boldsymbol{q_i}$ of the matrix $\boldsymbol{Q}$ is related to the eigenvector of $\boldsymbol{K}$. The matrix $\boldsymbol{\Lambda}$ is diagonal, and its diagonal entries are eigenvalues λ_i of the NTK operator $\boldsymbol{K}$. Herein, Equation (4-13) can be reformed as:

$$\begin{bmatrix} \boldsymbol{q}_1^T \\ \boldsymbol{q}_2^T \\ \vdots \\ \boldsymbol{q}_N^T \end{bmatrix} \left(f\left(\boldsymbol{X}_{train}, \boldsymbol{\theta}(\boldsymbol{t})\right) - \boldsymbol{Y}_{train}\right) = \begin{bmatrix} e^{-\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-\lambda_N t} \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_1^T \\ \boldsymbol{q}_2^T \\ \vdots \\ \boldsymbol{q}_N^T \end{bmatrix} \boldsymbol{Y}_{train} \tag{4-14}$$

$$\Rightarrow f\left(\boldsymbol{X}_{train}, \boldsymbol{\theta}(\boldsymbol{t})\right) - \boldsymbol{Y}_{train} = \sum_{i=1}^{N} (e^{-\lambda_i t} \boldsymbol{q}_i^T \boldsymbol{Y}_{train})\, \boldsymbol{q}_i \tag{4-15}$$

The left-hand side of this equation represents the error of the NN model. Obviously, along the NTK's eigendirections (eigenvectors) or eigenfunction of the network with a larger eigenvalue will earn an eminent convergence speed than others. However, in the case of a conventional fully connected network, the primary eigenfunctions tend to possess low frequency attributes. Meanwhile, as the frequency of these eigenfunctions increases, the corresponding eigenvalues will gradually decrease in a monotonic manner (Wang, et al., 2021). This theoretical analysis reveals the essence of the spectral bias and provides potential solutions to this underlying drawback of NNs. By manipulating the eigenspace of the NTK, it is possible to overcome the spectral bias that affects NNs.

The NTK theory suggests that when the NN width is infinitely large, the kernel $\boldsymbol{K}$ becomes deterministic (Jacot et al., 2018). Unfortunately, this condition typically does not align with the given content. Still, the analysis on the initialization of the NN to glimpse its ability to recognize high-frequency components can still be exploited. The corresponding proof can be found in in Appendix A through an illustrative example utilizing a physics-informed neural network (PINN) (Raissi et al., 2019).

4.2.2 Fourier/Wavelet neural network

In this study, a Fourier NN (*ff*NN) is utilized, which incorporates a random Fourier mapping on the input features. This technique is shown to enhance the eigenspace of the NTK, resulting in superior performance when compared to a conventional NN. Further details regarding this can be found in Appendix A. Additionally, in this work, the wavelet mapping (Wang et al., 2013) will also be evaluated.

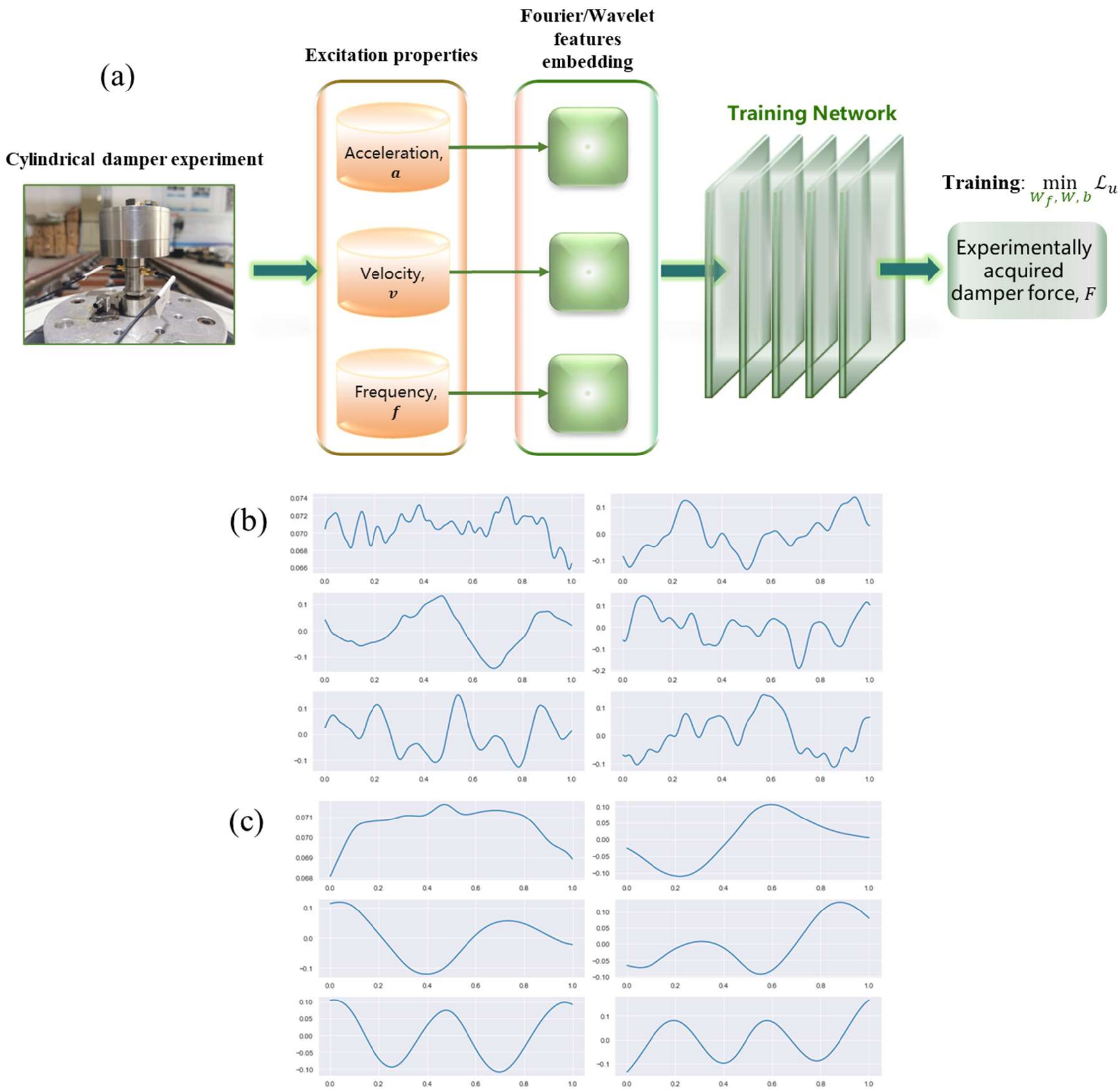


Figure 4-2 (a) Flowchart of a Fourier/Wavelet network; The six leading NTK eigenvectors of

a fully connected NN (5 layers, 100 units per hidden layer, *tanh* activations): (b) With embedded Fourier features that initialized by $\sigma = 10$; (c) Without Fourier features embedding.

1. Fourier features embedding

Following the work given in the literature (Tancik et al., 2020; Wang et al., 2021). To consider an *ff*NN with 100 hidden units per hidden layer, the random Fourier mapping is defined by $\gamma_f(x) = \begin{bmatrix} \cos(\boldsymbol{B}x) \\ \sin(\boldsymbol{B}x) \end{bmatrix}$, where $\boldsymbol{B} \in \mathbb{R}^{m \times d}$ is initialized by complying with the Gaussian distribution $\mathcal{N}(0, \sigma^2)$. Here, $m = 50$ is half of the layer width. The mapping γ_f in the first layer transfers the input data into Fourier feature space, which is subsequently fed into the fully connected NN (Figure 4-2(a)). The introduction of the hyperparameter σ results in an increased likelihood of initializing a value $b_i \in \boldsymbol{B}$ with a higher value. Figure 4-2(b) and 1(c) show the efficacy of this simple method in adjusting the eigenspace of the NTK. The eigenvectors of NTK exhibit more high-frequency features when σ is larger. The frequency characteristics of the leading NTK eigenvectors (eigenvectors with the largest eigenvalues) are the key indicator of the spectra bias of NN. This trick can enhance the NN's ability to recognize the high-frequency features, as it successfully modified the shape of the leading eigenvectors of a NN (Figure 4-2(b) and Figure 4-2(c)). This character is determined at the initialization stage.

In order to explore the effect of Fourier features on different NN architectures, the leading eigenvectors of different NN with the number of hidden layers $\mathcal{F}$ chosen from $\{3,4,5,6,7\}$, and the width of the hidden layer C chosen from $\{50,100,150\}$ are depicted in Figure 4-3. It is observed that Fourier features can modify the eigenspace of the NN with different architectures. For the effect of the σ value, it can be seen that with a higher σ, the frequency components of the leading eigenvector also extend to higher frequency range (Figure 4-4). A detailed discussion on the effect of σ value is given in Figure 4-7

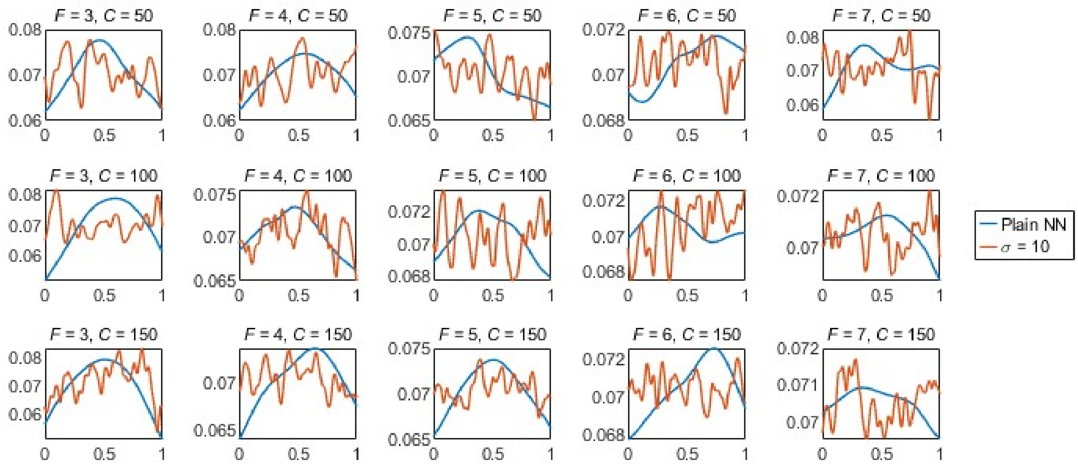


Figure 4-3 The leading NTK eigenvector of each NN architecture with and without Fourier features embedding.

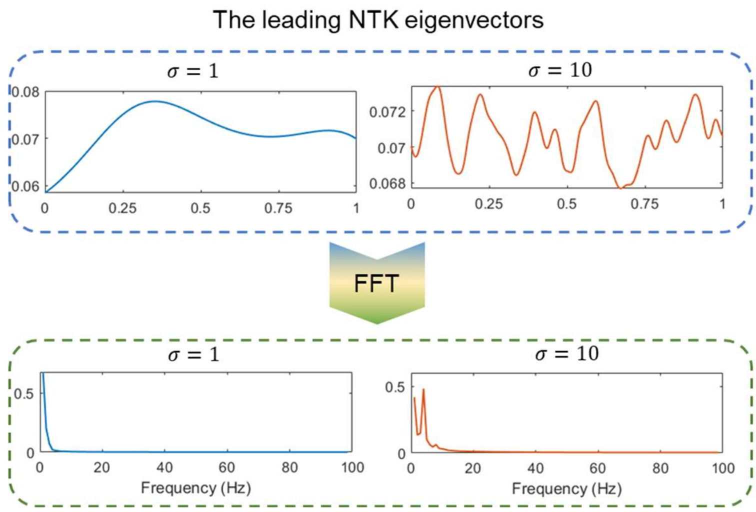


Figure 4-4 Effect of different σ values on the leading eigenvector.

2. Gaussian wavelet features embeddings

$$\gamma_g(x) = \frac{\boldsymbol{B}x - \boldsymbol{T}}{\sqrt{2\pi}} \cdot e^{-\frac{(\boldsymbol{B}x-\boldsymbol{T})^2}{2}}, \qquad \boldsymbol{B} \sim \mathcal{N}(0, \sigma^2), \boldsymbol{T} \sim \mathcal{N}(0,1) \tag{4-16}$$

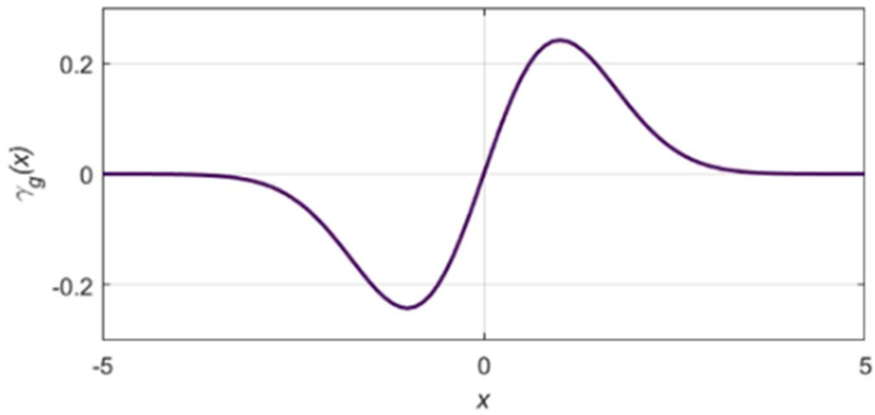


Figure 4-5 Gaussian wavelet basis function ($b = 1, \tau = 0$)

3. Morlet wavelet features embeddings

$$\gamma_m(x) = cos[1.75 \cdot (\boldsymbol{B}x - \boldsymbol{T})] \cdot e^{-\frac{(\boldsymbol{B}x-\boldsymbol{T})^2}{2}}, \qquad \boldsymbol{B} \sim \mathcal{N}(0, \sigma^2), \boldsymbol{T} \sim \mathcal{N}(0,1) \tag{4-17}$$

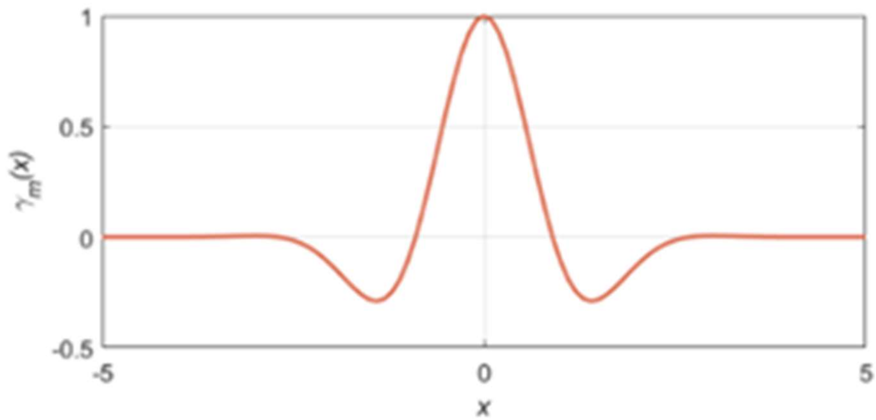


Figure 4-6 Morlet wavelet basis function ($b = 1, \tau = 0$)

Similar to the mapping of Fourier features, the utilization of wavelet mapping produces a comparable effect on the eigenspace of the NTK. For instance, a *wf*NN composed of 100 hidden units per hidden layer is considered. The continuous wavelet transform involves the Gaussian (Figure 4-5) and Morlet wavelet (Figure 4-6) functions (Jin et al.,2007) which serve as basic functions. These functions are scaled by a same vector $\boldsymbol{B} \in \mathbb{R}^{m \times d}$, which is sampled from a normal distribution with mean 0 and variance σ^2, $\mathcal{N}(0, \sigma^2)$. Additionally, the wavelet basis function can be shifted by a translation factor $\boldsymbol{T} \in \mathbb{R}^{m \times 1}$, where $\tau_i \in \boldsymbol{T}$ is initially sampled from $\mathcal{N}(0,1)$ and set as trainable in the *wf*NN. Here, $m = 100$ equals to the network width.

The frequency analysis of the eigenvector possessing the largest eigenvalue is depicted in Figure 4-7. According to Equation (4-15), this eigenvector is the most representative vector for the spectra character of NN. The frequency analysis result can roughly reflect the NN's ability to recognize high-frequency features according to different σ values. However, it should be noted that this frequency analysis does not represent the frequency range that NN can recognize. For both *ff*NN and *wf*NN, the spectrum is clearly extended with a larger σ, implying that *ff*NN and *wf*NN are excelling in accommodating high-frequency functions with the larger σ. It is noteworthy to mention that in our approach, the scaling vector $\boldsymbol{B}$ is set to be trainable, which makes the feature embedding layer a specific hidden layer of the NN. Moreover, when $\sigma = 1$, the resulting network is equivalent to a plain NN, with Fourier/Wavelet activation functions employed in the first hidden layer. The value of these hyper-parameters will significantly dictate the model's performance with different datasets. The selection of hyper-parameters will be discussed in section 4.4, after obtaining the experiment dataset.

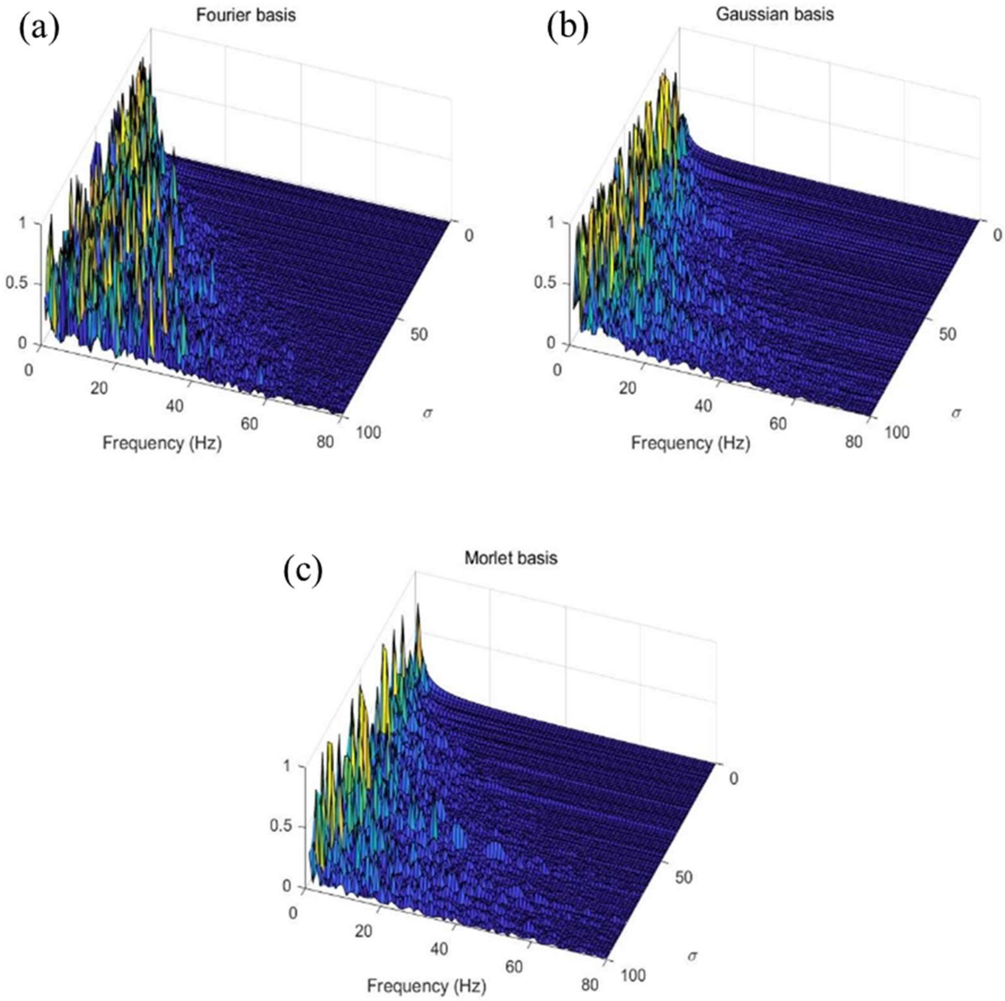


Figure 4-7 Frequency analysis of the leading NTK eigenvector with the most significant eigenvalue for a NN.

4.2.3 Incorporating *ff*NN with TL

The knowledge acquired from the source task may be leveraged to enhance performance on the target task through TL. Previous works (Tan et al., 2018; Zhuang et al., 2021) have revealed that the layers at the low level of a pre-trained NN model learn features with generality. The learned knowledge resides in the weights/biases of the hidden layers is transferrable across tasks. After transferring the layers from the pre-trained NN, the layers at the low levels are frozen. In the context of the target task, only fine-tuning the high-level layers are needed. Analogous to the framework introduced by the authors in a prior study (Ye et al., 2022), a low fidelity *ff*NN model is pre-trained with the approximate governing equation of PD. Subsequently, this model is refined using the high-fidelity dataset obtained from the experiment observations. Specifically, the training of the source task utilizing the approximate governing equation is carried out through the physics-informed neural networks (PINNs) (Raissi et al., 2019), which obviate the need for generating a database through the approximate governing/constitutive equations. In contrast to conventional numerical methods, PINNs produce a solution across an entire domain without requiring a pre-specified grid of points. The methodology of utilizing PINNs to address differential equations is presented upon in Appendix A.

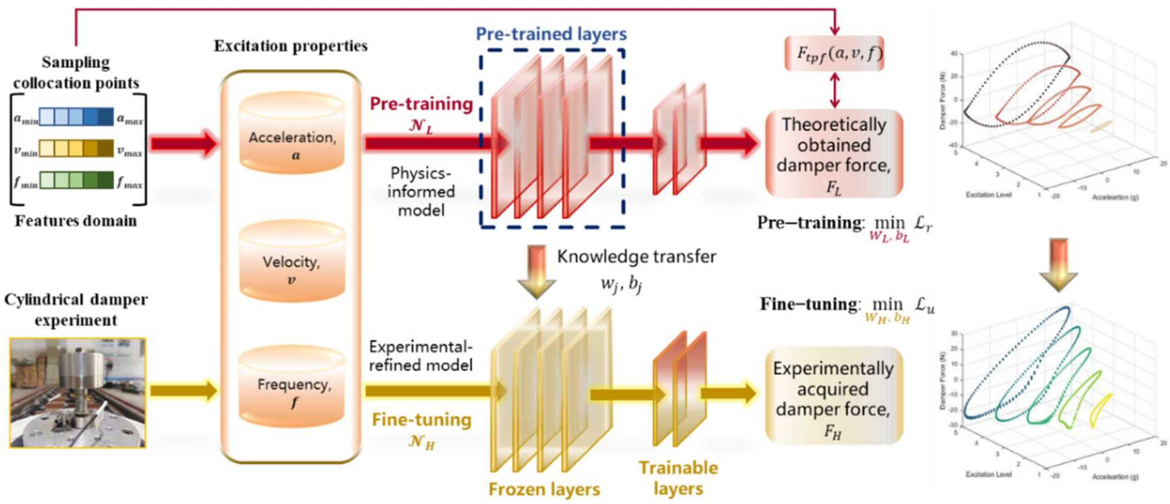


Figure 4-8 Flowchart of the TL-incorporated PINN modeling approach for the PD.

Following the discussion presented in the prior study, the approximate constitutive equation for two-phase flow equivalent viscosity model is proposed by Wu et al. (Wu et al., 2004; Lei & Wu, 2017). This equivalent model provides low-fidelity knowledge of the PD nature within the scope discussed in this study (high-frequency vibration). According to the experimental setup, the rigid connection between the PD and the shaker eliminates the relative displacement term in the dynamic force equation. Therefore, the force response of the PD is computed based on the two-phase flow model, which is given by:

$$F_{tpf} = ma + c_{eq}v \tag{4-18}$$

where m is the total mass of the damper, a and v are vibration acceleration and velocity, respectively, c_{eq} is the total equivalent viscous damping coefficient which relates to the damper response (velocity and frequency), it is written as:

$$c_{eq} = c_c + c_f \tag{4-19}$$

where c_c represents the equivalent viscosity relates to the inter-granular collision, and c_f is related to the inter-granular friction. These terms are calculated according to the two-phase flow theory (Wu et al., 2004; Lei & Wu, 2017). In this PINN, the features domain is maintained identical to that of the experimental setup, i.e., acceleration ranges from 5 to 25 g, and frequency ranges from 100 to 2000 Hz. Consequently, the features domain is kept consistently across both the source and target tasks, making this study a prototypical example of inductive TL (Pan & Yang, 2010). The domain adaption across tasks is no requirement. The PINN outputs the theoretical damper force F_{tpf} based on the two-phase flow model. Additionally, model testing is unnecessary for the source training task.

4.3 Experiment on Hysteretic Response Force of a PD

4.3.1 Experimental setup

At the current stage, the response-induced damping mechanism of the granular material bed remains an unsolved puzzle, especially in the high-frequency range. Therefore, a comprehensive investigation on the response of granular particle beds is required before applying PD. In this regard, a simple experiment was designed to obtain the interaction forces that the particle bed applies to the cavity during simple harmonic vibration. The recorded force signal consists of the interaction force between the particle bed and cavity, and the inertia force of the damper cavity. Since the inertia force of the damper cavity and the accelerometer is not of interest, it will be deduced, and the response force mentioned afterward all refer to the interaction force from the particle bed.

In this experiment, the PD used was directly attached to the acceleration-controlled shaker (B&K LDS V-650) without being mounted onto any primary structure. The utilization of the shaker on a PD allows a straightforward examination of the damper mechanism (Hu et al., 2023). The weight of the damper was effectively small compared to the mass of the shaker moving element. Hence, the PD was able to vibrate synchronously with the shaker head as an entity. It made the vibration of PD follow the demanded pattern. The PD was installed vertically on the impedance head (KD3001B) with the M5 screw. This sensor records the force and acceleration signal in the vertical direction. The impedance head was also bolted to a stainless-steel connector that connects to the shaker (see Figure 4-9). An accelerometer (i.e., Dytran 3273) is attached to the top of the damper. This sensor records the vertical acceleration signal. This sensor's purpose was to ensure the PD vibrated in sync with the shaker head. As seen in Figure 4-11, there was no relative displacement between the PD and the shaker during vibration. The sampling rate of the force and acceleration signal was 10 kHz. The Dewesoft SIRIUS data acquisition system was applied in the test to collect data, the signal-to-noise ratio of this system is higher than 130 dB.

The vibration was constrained in the vertical direction. A cylinder shapes the damper cavity to mitigate the influence of contact between the particle bed and cavity in the lateral direction. Aluminum alloy was chosen for the material of the damper cavity wall to reduce the inertia force generated by the cavity. However, the support base (marked in green in Figure 4-10(c)) was made of stainless steel to bear the high-frequency load. The filling particle is 0.2 mm tungsten powder, and the filling volume was 30% of the cavity volume, making the weight of the filled tungsten powder 128.5 gr.

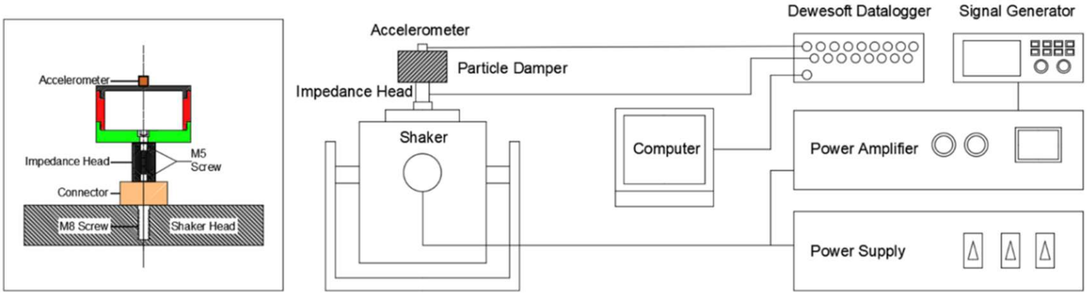


Figure 4-9 Schematic diagram of experimental setup.

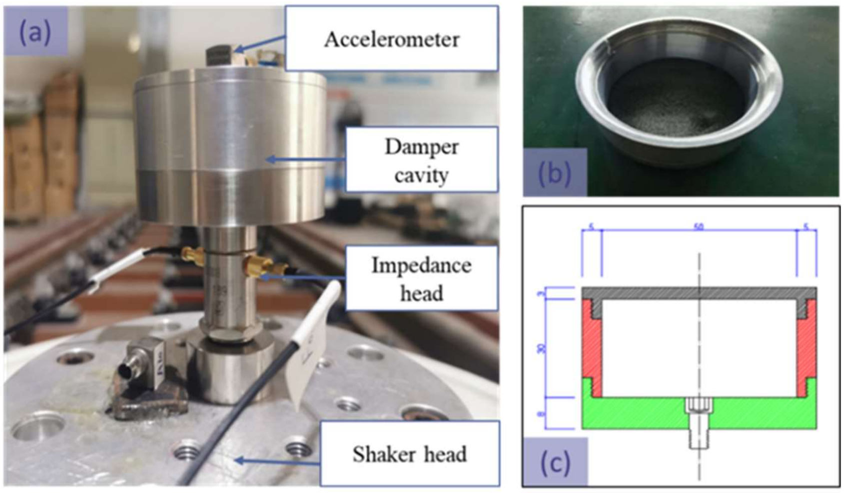


Figure 4-10 (a) Setup of the cylindrical PD; (b) Tungsten powder filling in the damper cavity; (c) Schematic diagram of the damper cavity.

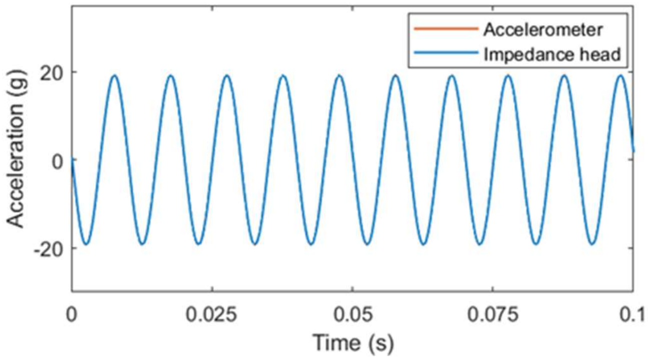


Figure 4-11 Acceleration signal recorded by the accelerometer and the impedance head.

The detailed parameters of the experimental apparatus are given in Table 4-1. In this test, two types of excitations were adopted to excite the PD: steady-state simple harmonic excitation and sweep sinusoidal excitation.

Table 4-1 Specifications of experimental apparatus

Experimental apparatus	Parameter name	Parameter value
B&K LDS V-650 shaker	Sine force peak	2.2 kN
	Max acceleration sine peak	50 g
	Usable frequency range	5 Hz – 4 kHz
	Mass of moving element	4,460 g
KD3001B impedance head	Max force peak	2.5 kN
	Max acceleration peak	100 g
	Working frequency range	0.5 Hz – 6 kHz
	Weight	10.8 gr
	Sensitivity	2 mV/N
Dytran 3273 accelerometer	Max acceleration peak	100 g
	Weight	3.0 g
Cylindrical damper cavity	Inner cavity diameter	5.0 cm
	Inner cavity height	3.0 cm
	Weight	275.4 g

4.3.2 Steady-state simple harmonic excitation

The steady-state simple harmonic excitation frequency ranges between 100 and 2000 Hz with an interval of 100 Hz. The excitation amplitude is controlled manually by switching the gain on the amplifier; five levels of amplitude are applied to the PD: The peak acceleration value exhibits a range of approximately 5 g to 25 g, with intervals of 5 g. The excitation period is of 5 s for each condition, but only the middle 0.1 s of the signal starting at the π phase is selected for the investigation, with the aim of avoiding the transient effect. Each acceleration-based vibration signal contains 1,000 temporal stamps as the signal sampling rate is 10 kHz.

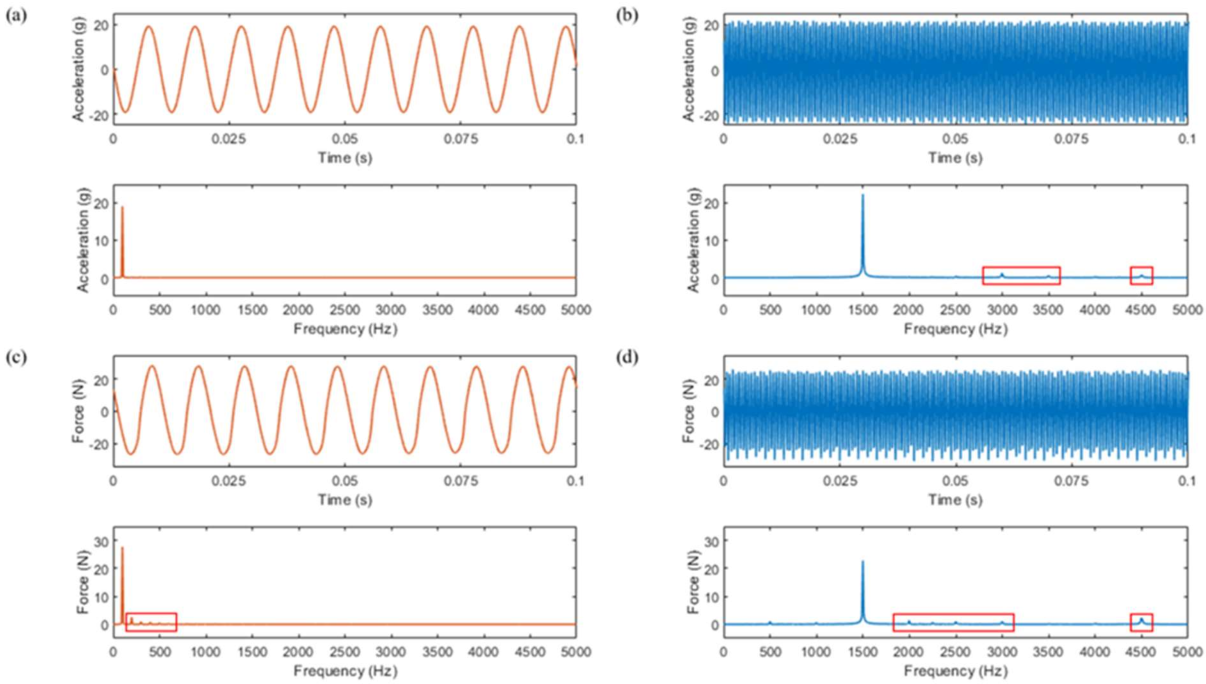


Figure 4-12 (a) Recorded acceleration in time and frequency domain at 100 Hz; (b) Acceleration in time and frequency domain at 1500 Hz; (c) Force in time and frequency domain at 100 Hz; (d) Force in time and frequency domain at 1500 Hz.

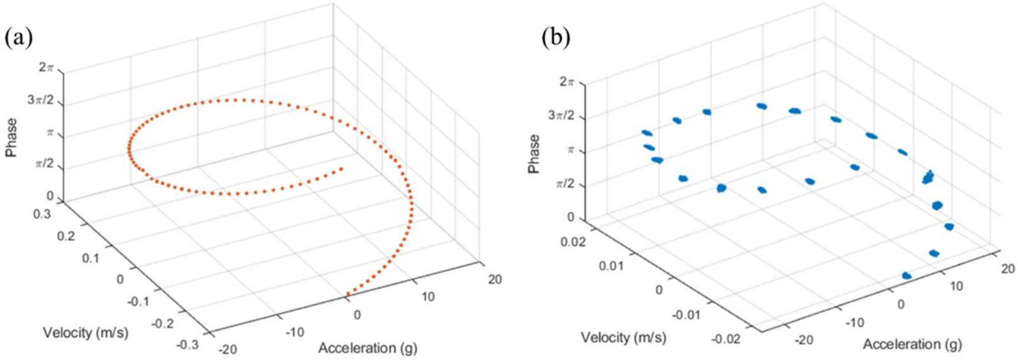


Figure 4-13 Phase of the acceleration signal with respect to the value of acceleration and velocity at (a) 100 Hz; (b) 1500 Hz.

Figure 4-12 shows the acceleration and force signal that recorded at the frequency of 100 Hz and 1500 Hz, with an acceleration amplitude of approximately 20 g. It can be seen that there contains a certain degree of multi-frequency components (ultra-harmonics, e.g., 3000 Hz and 4500 Hz) in a high frequency regime (see Figure 4-12). The appearance of ultra-harmonics is due to the nonlinear nature of the particle damping, which was generated from the complex interaction force of the particle bed. The reactive force inevitably influenced the vibration of the shaker head. However, since the mass of PD is small, the influence is limited, and the vibration can still be roughly regarded as simple harmonic vibration according to the acceleration spectrum. The acceleration signal extracted for analyzation starts from the π phase, and its phase goes forward on the timeline. Because the signal is periodic, its phase is equivalent to the range of $(0,2\pi]$. Fig. 11 illustrates the acceleration phase with respect to the value of acceleration and velocity at the frequency of 100 Hz and 1500 Hz. Similar as observed in Fig. 10. At the frequency of 1500 Hz, nonlinear effects are more significant than that at the frequency of 100 Hz. It is presented by a larger noise in the phase of the acceleration signal. This nonlinear effect-induced noise would influence the model reconstruction. For this uncertainty phase reconstruction, a Bayesian-based approach can be a solution, which will be discussed in future work. This work mainly focuses on the improvement brought by Fourier features embedding.

Figure 4-14 gives the hysteresis loop of the PD under different excitation frequencies and levels, where the response force is dependent on the current excitation and the history of excitation. Relative displacement doesn't exist in this experimental set. The velocity amplitude varies 20 times in the frequency range from 100 Hz to 2000 Hz. However, since the velocity amplitude is negligible compared to the acceleration in such high-frequency vibration. The amplitude of the response force is dependent on the equivalent inertia force of PD, which doesn't change with frequency. Since there was no relative displacement between PD and the shaker, the amplitude of the response force was mainly determined by the inertia force of the damper cavity. It is noted that the acceleration amplitude is manually controlled and does not precisely meet the desired value in the test. Therefore, the term "excitation levels" (ELs) is chosen to represent the different acceleration amplitudes. The nonlinear damping effect of the PD varies in the frequency range: it exhibits a viscous behavior at low frequencies, while at high frequencies, the damping effect is restrained by the hysteretic phenomenon. The underlying rationale for the reduced damping effect observed in the acceleration-controlled test scheme can be attributed to the scaling down of velocity and displacement with respect to frequency. Therefore, it can be conceptualized that the particle bed emulates a lump mass during micro-displacement vibration, thereby contributing to the observed phenomenon of low damping effect. The shape of hysteresis loops in different frequencies varies a lot, which reflects the intricate motion of the particle bed inside the cavity. The particle bed's movement is more intense in lower frequencies with the higher displacement amplitude. Thereupon, A larger area of the hysteresis loop can be observed due to the elevated energy dissipation level by the particle bed. However, it is rather hard to dissect the shape of every hysteresis loop. Luckily, the data-driven deep learning approach offers an opportunity to avoid expressing the intricate particle damping explicitly. The application of the proposed deep learning method will be discussed in the next section.

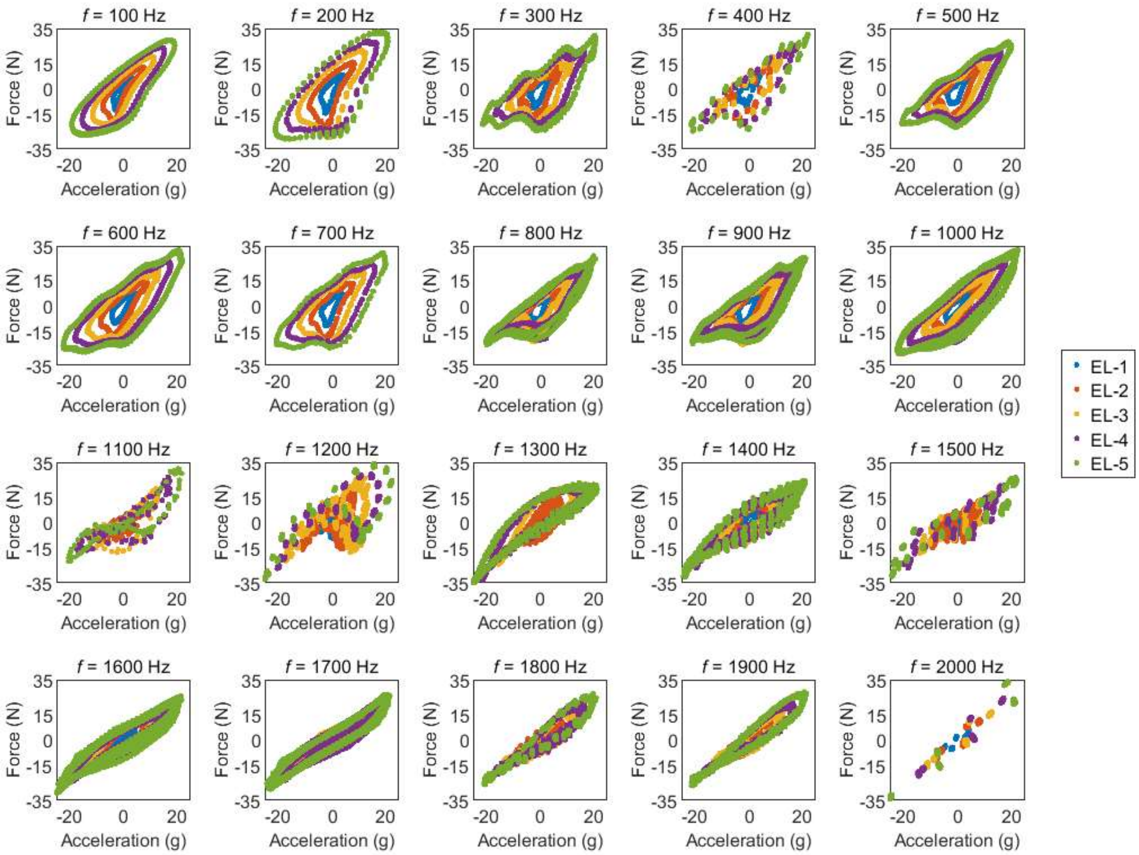


Figure 4-14 Hysteresis loop (output force versus excitation acceleration) under different excitation frequencies and levels.

4.3.3 Sweep sinusoidal excitation

A sweep-sinusoidal signal is also selected as an input signal of the shaker. It keeps the same frequency range as the steady-state excitation: the signal frequency linearly sweeps from 100 to 2000 Hz, the total sweeping time is 20 s, and the sweep-frequency rate is 95 Hz/s. The acceleration amplitude is not controllable during the sweeping. In the present situation, the voltage applied to the shaker, which governs the electrical signal, was maintained at a consistent level equivalent to the first level of excitation during the steady-state condition.

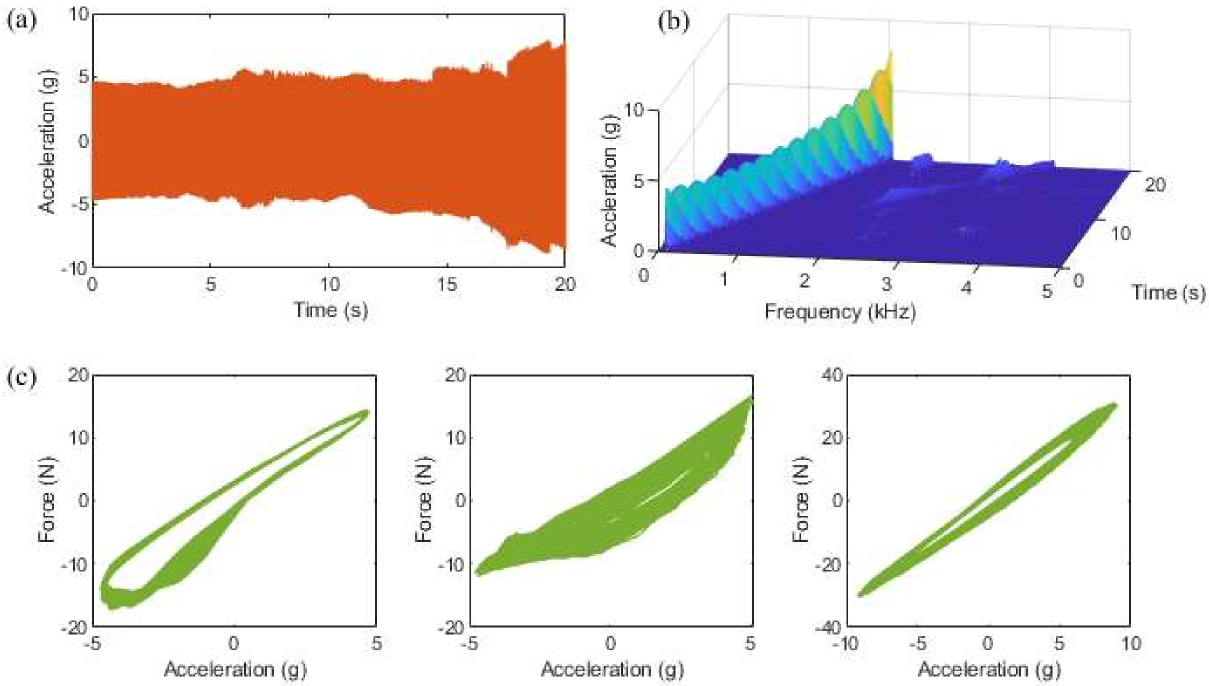


Figure 4-15 (a) Sweep sine acceleration signal in time; (b) Time-frequency spectrum of the acceleration signal; (c) Diagram of force versus acceleration at frequency range of 100 – 150 Hz (left), 1000 – 1050 Hz (middle), and 1950 – 2000 Hz (right).

Nonlinear characteristics of PD can be observed through the detection of ultra-harmonic signals, which occur when the fundamental frequency surpasses 1000 Hz. However, the sweeping frequency rate is not significantly high according to the setting of 95 Hz/s. The resultant hysteresis loops in different frequency ranges are similar to those in the steady-state excitation case (see Figure 4-15(c)), which are not the smooth incline ellipse shape.

4.4 Proof of Concept

4.4.1 Selection of the NN architecture

For selecting a network architecture, large networks (in regard to width and depth) exhibit greater robustness in terms of their regression capability. However, training such networks is inevitably more laborious due to the larger number of parameters that must be updated during each iteration. Meanwhile, expanding the network may result in only marginal improvements in performance, particularly when a smaller NN is already capable of accomplishing the given task. Moreover, a solid strategy to guide the network selection is currently absent to the best of the authors' knowledge. Therefore, the plain NNs on the steady-state dataset (total size of 1×10^5) are tested. This dataset consists of 20 sets of frequencies (100, 200, …, 2000 Hz); 5 sets of acceleration levels (approximately 5, 10, …, 25 g); each piece of the acceleration-controlled vibration data is 0.1 s long starting at the π phase with 1,000 temporal stamps. In consideration of representing the hysteretic effect, the excitation acceleration a, and its integration (i.e., velocity) v are chosen to be the input features. The purpose of inputting velocity is to provide the phase information of the signal. Calculating the phase directly and inputting it into the NN is unsuitable. Since the vibration in the test is not purely simple harmonic according to Figure 4-12. The additional input, excitation frequency f, is intended to reflect the nonlinearity of PD in the frequency domain (see Figure 4-2(a)). The network outputs the force response of the damper. The data corresponding to the third-level excitation (EL-3, 20% of the steady-state dataset) is extracted for testing the model. Training is conducted on the rest excitation levels (EL-1, EL-2, EL-4, EL-5, 80% of the steady-state dataset). The entire dataset is preprocessed by the min-max normalization (Henderi et al., 2021).

Various architectures with different widths and depths are empirically selected from the candidate set. The number of hidden layers $\mathcal{F}$ is chosen from $\{3,4,5,6,7\}$, and the width of the hidden layer $\mathcal{C}$ is chosen from $\{50,100,150\}$. In this study, all hidden layers in one NN architecture are set to be the identical width. The tanh activation function is applied for these NNs. The training and testing mean squared error (MSE) after 15,000 iterations for each NN architecture is given in Figure 4-16. The training of NN intuitively receives a reduced loss with a larger NN size. However, this merit also comes along with over-fitting, as demonstrated by the increasing testing loss after meeting the inflection point. The selection of the NN architecture should consider these two aspects simultaneously. By comparing the average MSE of different NN architectures on the entire dataset (see Table 4-2), it can be identified that NN with $\{\mathcal{F} = 5, \mathcal{C} = 100\}$ obtains the best model accuracy. Therefore, the network architecture of 5 hidden layers with 100 neurons in this study is chosen. This selection is based on the plain NNs. In addition, for a fair comparison, all methodologies will adopt the selected architecture afterwards, and the training and testing loss are presented in MSE.

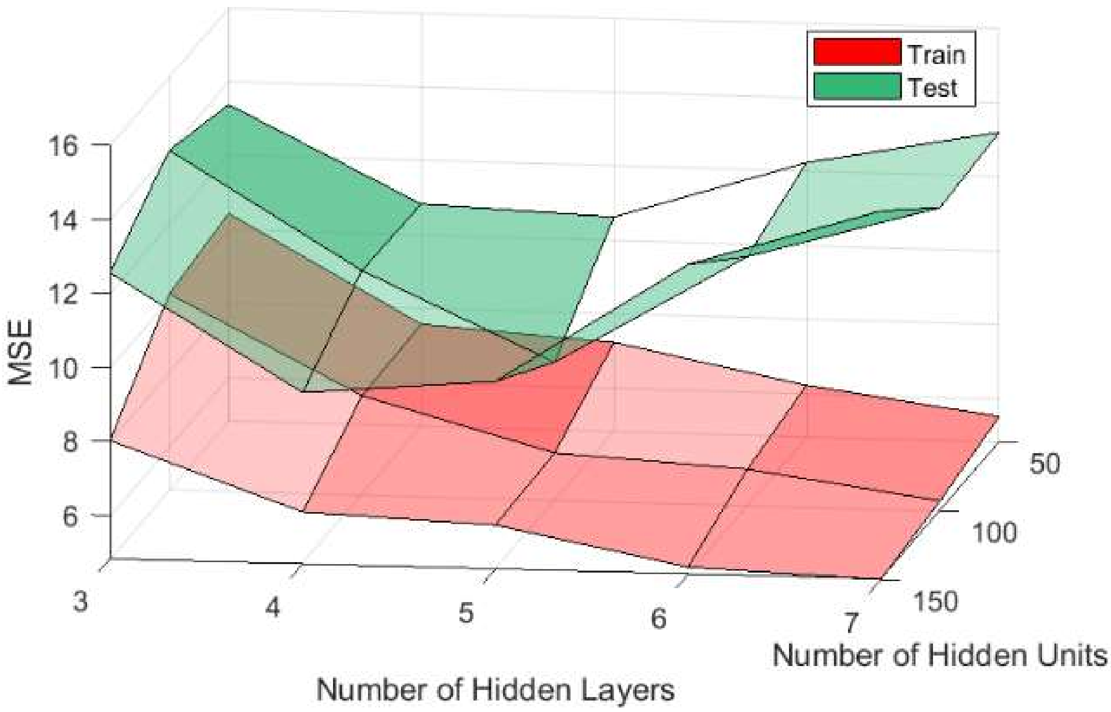


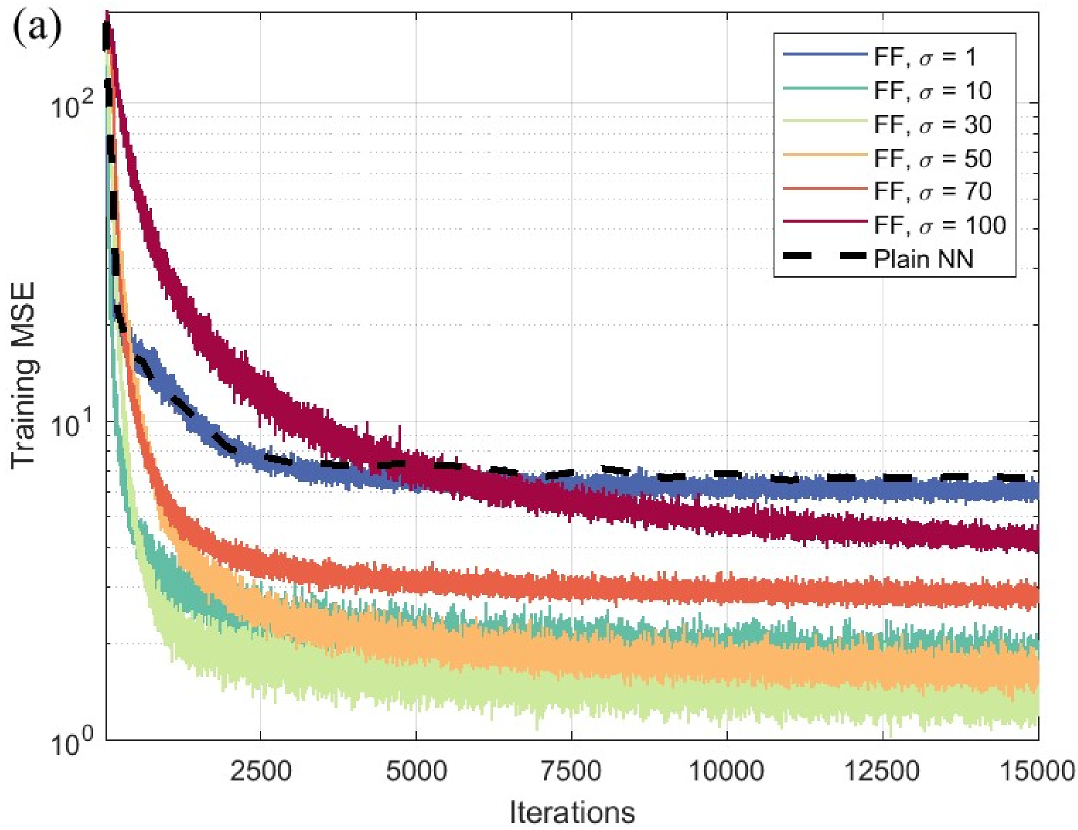
Figure 4-16 Training and testing MSE of each NN architecture after 15,000 iterations.

Table 4-2 Average MSE on the entire dataset of each NN architecture after 15,000 iterations

	$\mathcal{C} = 50$	$\mathcal{C} = 100$	$\mathcal{C} = 150$
$\mathcal{F} = 3$	11.02	10.90	8.93
$\mathcal{F} = 4$	8.23	8.19	6.88
$\mathcal{F} = 5$	7.91	**6.60**	6.80
$\mathcal{F} = 6$	7.42	6.96	6.67
$\mathcal{F} = 7$	7.05	6.69	6.83

4.4.2 Validation of the optimum features embedding strategy

As mentioned in Table 4-2, the NN architecture ($\mathcal{F} = 5, \mathcal{C} = 100$) is adopted in this section to explore the optimum features embedding strategy. The search is conducted on the dataset of steady-state excitation test (5 excitation levels, 20 excitation frequencies, and 1,000 temporal stamps for each set), in which the EL-3 dataset is assigned as the testing dataset while the rest is for training. Ten independent trial tests are conducted on the NNs embedded with Fourier (FF), Gaussian (Ga), and Morlet (Mo) features that are initialized by different $\sigma \in [1,100]$.



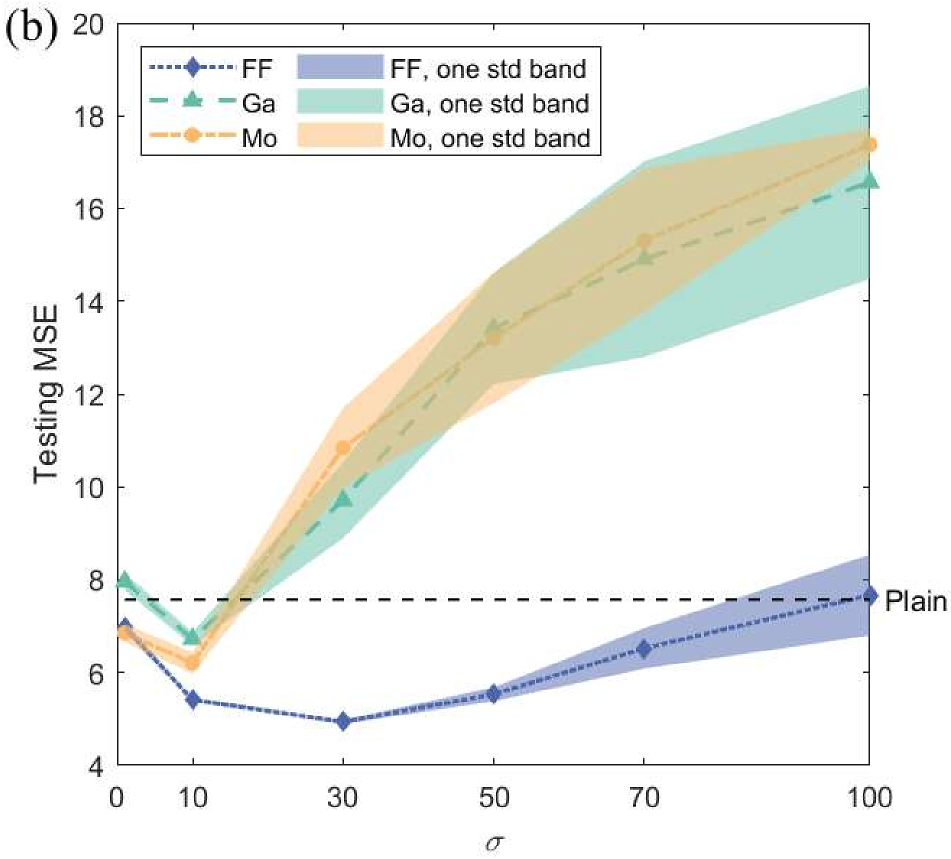


Figure 4-17 (a) Evolution of the training MSE from one trial test of a NN embedded with Fourier features initialized by the different $\sigma \in [1,100]$ and a plain network; (b) Testing MSE with the three feature functions by the different $\sigma \in [1,100]$ after 15,000 iterations of gradient descent over 10 independent trials. The black dash line represents the final testing loss of the plain network after 15,000 iterations.

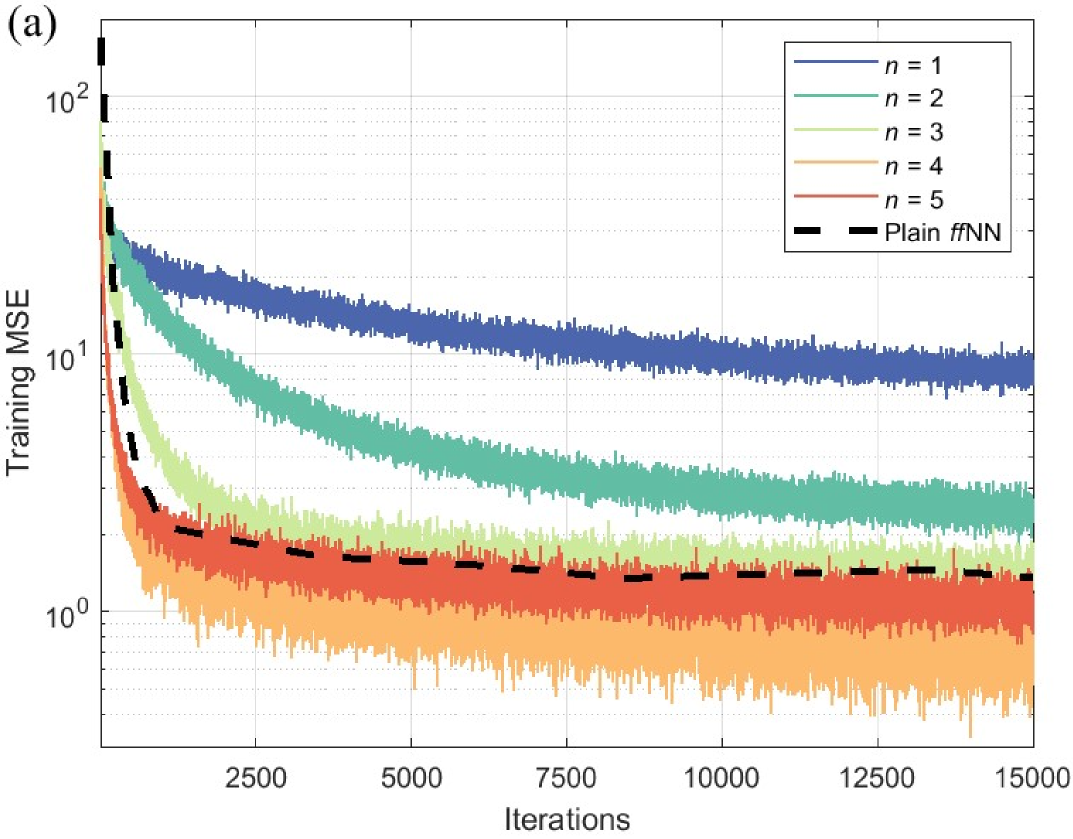
The evolution of the training MSE from one trial test with various Fourier features embedding shown in Figure 4-17(a) reveals that with a proper σ, the *ff*NN is faster in convergence and consequently achieves a better accuracy after a certain number of iterations. Testing results in Figure 4-17(b) show that Fourier features embedding is more robust (thinner one standard band) than wavelet features embedding (i.e., Ga and Mo). The latter one even shows worse performance than the plain network in most cases. A possible explanation for this phenomenon is that when the wavelet basis function is scaled with a large-value factor $b_i \in \boldsymbol{B}$, the function becomes a "thin" basis containing a wide range of zero values in its domain. As a result, the function in question exhibits limited capacity to accommodate the objective function effectively. Another flaw of the *wf*NN is that, unlike the matrix operation of the scaling factor $\boldsymbol{B}$, the translation factor $\boldsymbol{T}$ can only be applied to each neuron in the embedding layer. It means that different dimensions of input $\boldsymbol{x}$ must share the same translation factor $\tau_i \in \boldsymbol{T}$ on one neuron, which also restricts the performance of the *wf*NN.

In alignment with the preceding discussion, it can be observed that for the case where $\sigma = 1$, the *ff/wf*NN is virtually a variant of plain networks, with the activation functions of the first hidden layer being the Fourier/Wavelet features mapping. Demonstrating the validity of this assertion, the training and testing results of the *ff/wf*NNs are in the vicinity of the plain NN. Among the various experimental trials conducted, the Fourier features initialized with $\sigma = 30$ yield the best performance.

4.4.3 Validation of the optimum free layers number

The selected *ff*NN architecture (5 layers, 100 units per hidden layer, tanh activations, $\sigma = 30$) is adopted in this subsection. Remarkably, the number of frozen layers transferred from the source task is a crucial factor for the performance of the target task. A reduced number of free layers in the target model can use the low-fidelity knowledge to a greater extent and expedite the training. Conversely, this approach impedes the learning capability of the high-fidelity dataset.

This section aims to explore the influence of the free-layer number and confirm the optimum number of free layers. The search is conducted on the same dataset from the steady-state excitation test (EL-3 for testing and the rest for training). The performance of the different free-layer number $n \in [1,5]$ over 10 independent trials is tested. For the purpose of comparison, the mean squared error (MSE) evaluation of the plain *ff*NN is also provided.



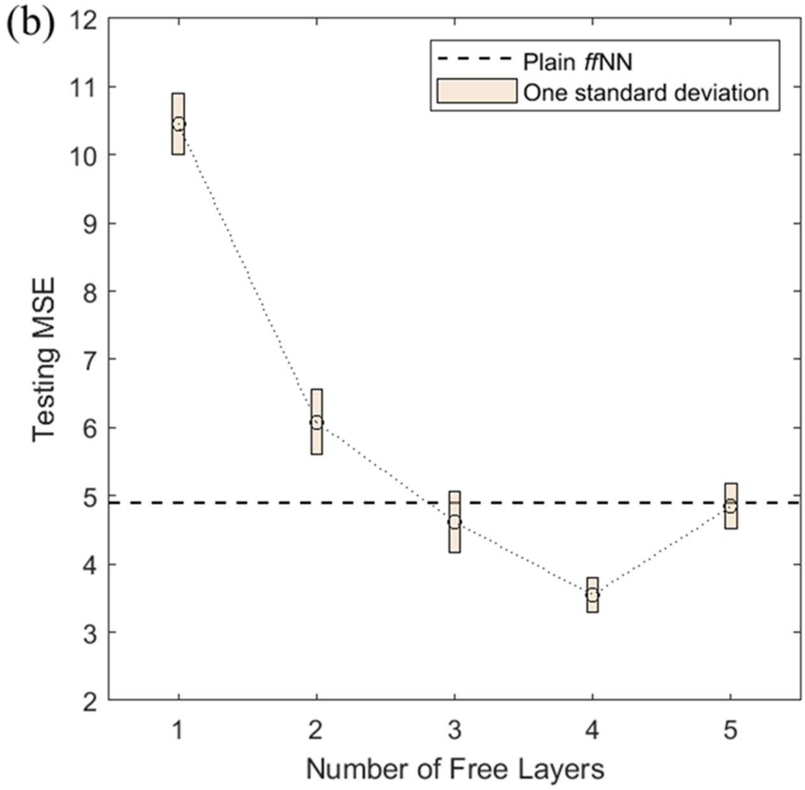


Figure 4-18 (a) Evaluation of the training MSE from one trial test with the different free-layer number $n \in [1,5]$ and the plain *ff*NN; (b) Testing MSE with the different free-layer number $n \in [1,5]$ after 15,000 iterations of gradient descent over 10 independent trials. The black dash line represents the final testing loss of the plain *ff*NN after 15,000 iterations.

The evolution history of the training MSE from one trial test in Figure 4-18(a) shows that the transferred layers from the source task facilitate learning in the target task with a decreased initial error. Nevertheless, in the case of a complex and highly nonlinear issue such as PD, a reduction in the number of trainable layers leads to a decline in the accuracy of both training and testing using the high-fidelity dataset (see Figure 4-18), even though the increased efficiency of training as fewer parameters are needed. The attractive advantage of TL is its potential in training data hungry NNs when high-fidelity data is limited (Ye et al., 2022). It is noted that the present size of the steady-state dataset amounts to 1×10^5, which is abundant for NN training purpose. An intuitive explanation is that the plain *ff*NN reaches almost the identical MSE where $n = 5$, indicating that the plain *ff*NN is well-driven, eventually converging to the performance level of the TL approach, even though its higher initial MSE. It is reasonable to anticipate that in the case of a smaller training dataset, the TL approach would be more appealing.

In addition, the TL provides a slightly enhanced performance where $n = 4$. Therefore, despite the fact that the improvement is not as significant as that achieved through proper Fourier features embedding, the *ff*NN with TL-PINN (*ff*-TLPINN) is chosen to incorporate as it ultimately yields superior results.

Following the discussion on the dataset size, the size of the training dataset is 8×10^4 (EL-3 extracted for testing). The considerable dataset size goes against the TL approach showing its merits. Herein, the training dataset is randomly cut down to lessen the dataset size, namely 80%, 60%, 40%, and 20%. The efficacy of the TL under different number of free layers is dissected on these reduced datasets based on *ff*NN with $\sigma = 30$.

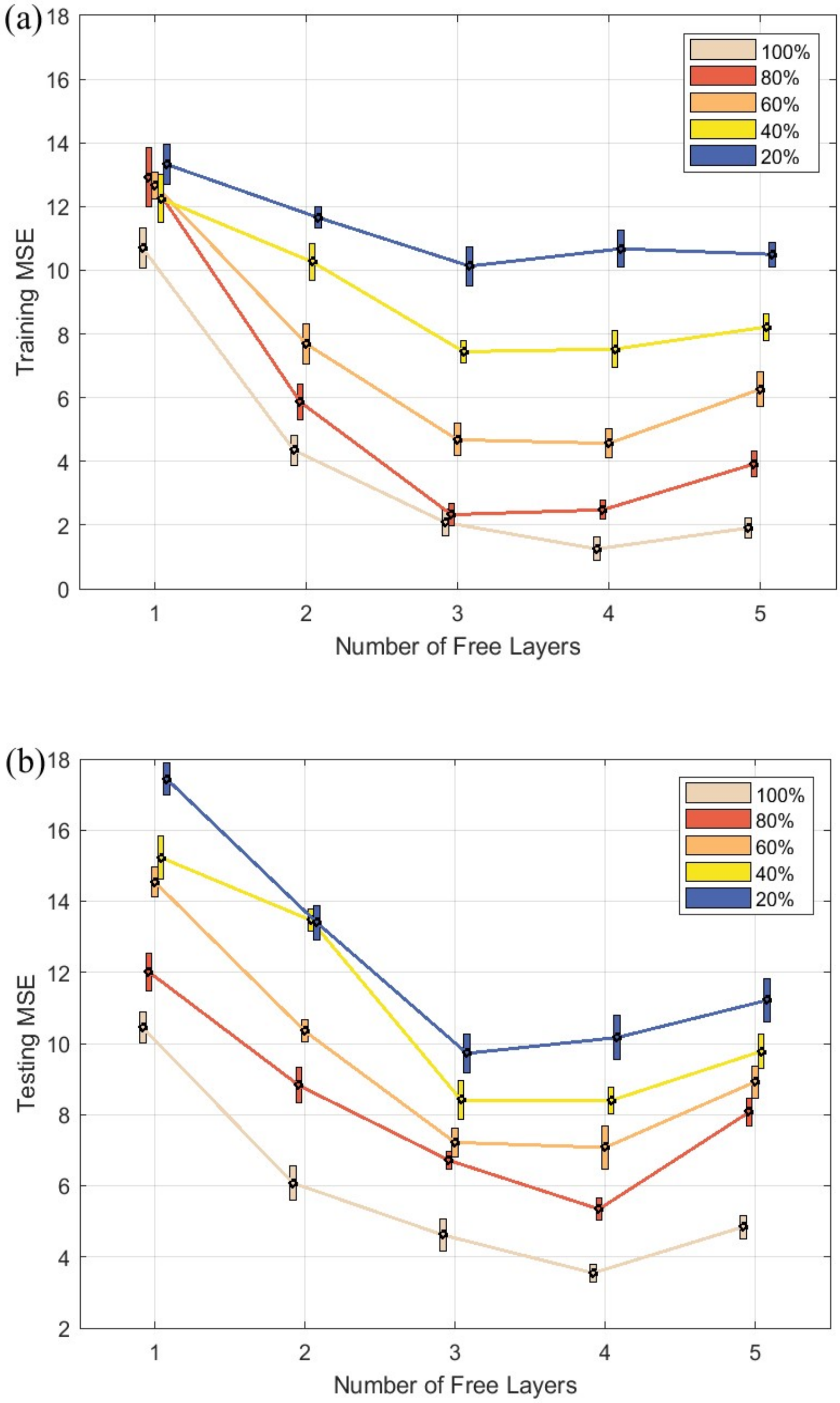


Figure 4-19 (a) Efficacy of free-layer numbers on various dataset sizes on model training; (b) Efficacy of free-layer numbers on various dataset sizes on model testing. Evaluated with MSE and one standard deviation.

Firstly, the most intuitive conclusion is that reducing the dataset will raise the MSE in training and testing. However, such influence is decreased when training the model with just one free layer, indicating that only limited data is required for training in this condition. But even though the model is well-trained to capture the features in the training dataset, the lessened training dataset still fails to provide a complete view of the task. So that the rising testing MSE with one free layer can be found when reducing the training dataset.

It can also be observed that the effect of TL becomes more apparent with a lessened dataset. Compared with the intact training dataset, the optimum free-layer number is decreased on the lessened dataset. It means the source task knowledge plays a more vital role when the training dataset of the target task is reduced. To better present this conclusion, for each dataset size, the result on the plain model without the TL approach is compared with the TL approach. An index $\epsilon_{train/test}$ is defined, which is calculated by the training/testing MSE of the non-TL model deducting the training/testing MSE of the TL models. When $\epsilon > 0$, the TL outperforms the non-TL approach and vice versa.

The comparison in Figure 4-20 provides compelling proof that $\epsilon_{train/test}$ is increasing with the decreasing dataset size. In short, the TL approach can facilitate establishing a better surrogate when the dataset size is small, no matter the number of free layers.

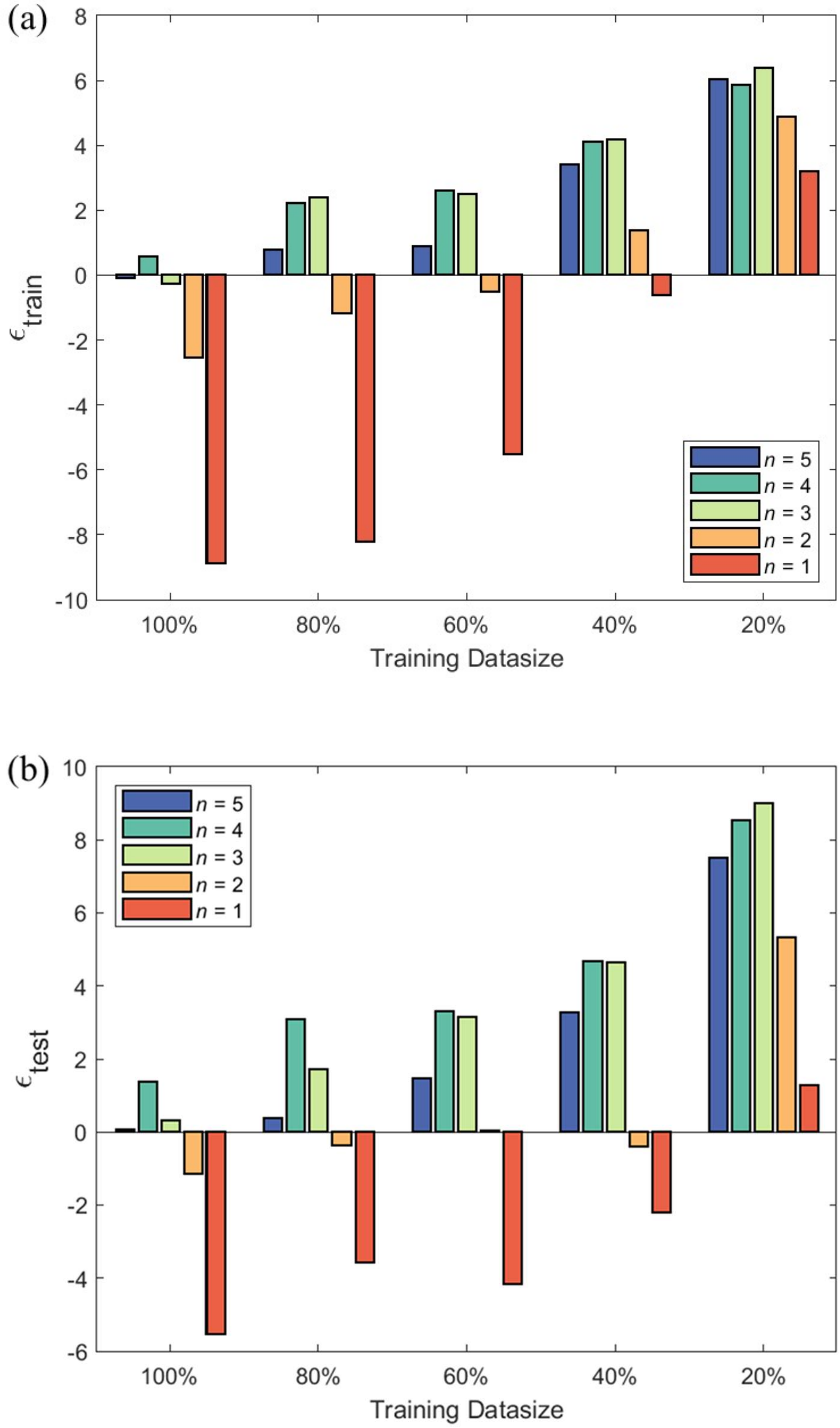


Figure 4-20 (a) Performance of TL and non-TL approaches on model training; (b) Performance of TL and non-TL approaches on model testing.

4.4.4 Comparison between the surrogate models

After the discussion on *ff*NN and TL, the optimum strategy is achieved by the fusion of *ff*NN and TL-PINN (*ff*-TLPINN), in which the first layer is the Fourier-features mapping and followed by the TL-PINN approach. To be acquainted with these methods, all results from the different methods are compared in this subsection, including the NN, *ff*NN, and the *ff*-TLPINN (Fourier features embedding with $\sigma = 30$, transferring layers with the free-layer number $n = 4$). Furthermore, to reveal the effect of the TL-PINN approach on the plain NN, a similar discussion as in section 4.4.3 is conducted. Finally, the optimum result of pure TL-PINN is also obtained with $n = 4$. Consequently, the results are presented for the sake of comparison. All surrogate models are trained using the steady-state dataset for 15,000 iterations, employing the same training/testing dataset and network architecture. The model evaluation is then conducted on the entire dataset.

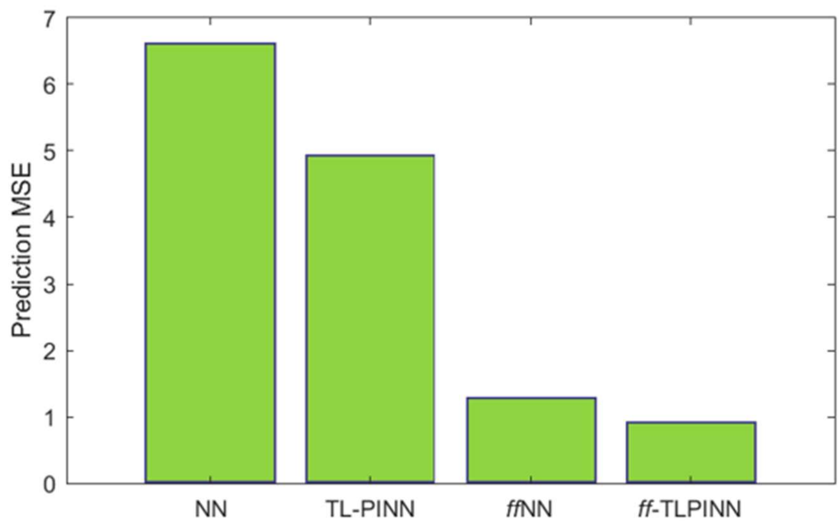


Figure 4-21 Prediction MSE for different approaches with the same NN architecture.

Figure 4-21 shows the model prediction of the MSE on the steady-state dataset of various methods. It can be seen that the TL-PINN (MSE 4.92) enhanced the model accuracy regarding a plain NN (MSE 6.60), but a more remarkable result is presented by the properly initialized *ff*NN (MSE 1.27). Furthermore, blending the TL approach with the *ff*NN brings a slight improvement in the model accuracy (MSE 0.90) regarding the plain *ff*NN, indicating the most critical challenge for this task is the spectral bias of the NN. As a result, the best accuracy is given by *ff*-TLPINN. This method is selected to reconstruct the hysteretic effect of PD.

4.5 Model Reconstruction for Response Force of PD

The damper force response is reconstructed by implementing the proposed Fourier features-embedded, transfer learning-incorporated PINN (*ff*-TLPINN). An *ff*NN (5 layers, 100 units per hidden layer, *tanh* activations) with the Fourier feature embedding $\gamma_f(\boldsymbol{X})$ initialized by $\sigma = 30$ was firstly trained according to the low-fidelity physics model, i.e., the two-phase flow equivalent viscosity model. Then, the pre-trained layers were leveraged to the second stage of training, where the low-level layers were frozen. The number of free layers to be updated by the high-fidelity experimental data is 4. The input vector $\boldsymbol{X}$ of the steady-state excitation case consists of excitation frequency $\boldsymbol{f}$ (20 settings in total), excitation acceleration $\boldsymbol{a}$, and the corresponding velocity $\boldsymbol{v}$ (5 levels in total). Each set has 1,000 temporal stamps. After training with the steady-state dataset for 15,000 iterations on the designated training dataset (excluding EL-3), the sweep-sinusoidal dataset is used to validate the model.

4.5.1 Reconstruction of the steady-state excitation case

The proposed model manages to reconstruct the hysteresis loop of the PD. The hysteresis loop at 100 Hz shows a typical viscous effect (see Figure 4-22), while the shape of the hysteresis loop at 1200 Hz (see Figure 4-23) is significantly different from other frequencies due to the nonlinearity nature of the PD. The most considerable prediction error is found in this frequency case. As observed in Figure 4-24, the hysteretic phenomenon is significantly restrained at 2000 Hz. Since the excitation is acceleration-controlled, the PD actually undergoes a micro-displacement vibration at this frequency. Most of the force response is contributed by the inertia force of the damper. In this case, since the sampling frequency (10 kHz) happens to be five times the primary frequency (2 kHz), there can only be a sample of five points in one loop. Although the dataset contains an enormous number of temporal stamps, there are a lot of duplicated data since the input features are periodic signals. The error resides in these reduplicative cases and results in the prediction error. For instance, the force response of PD under the fifth level excitation at 2000 Hz (phase $-\pi$) is shown in Figure 4-25. The experimental error is the systematic error in terms of the model training.

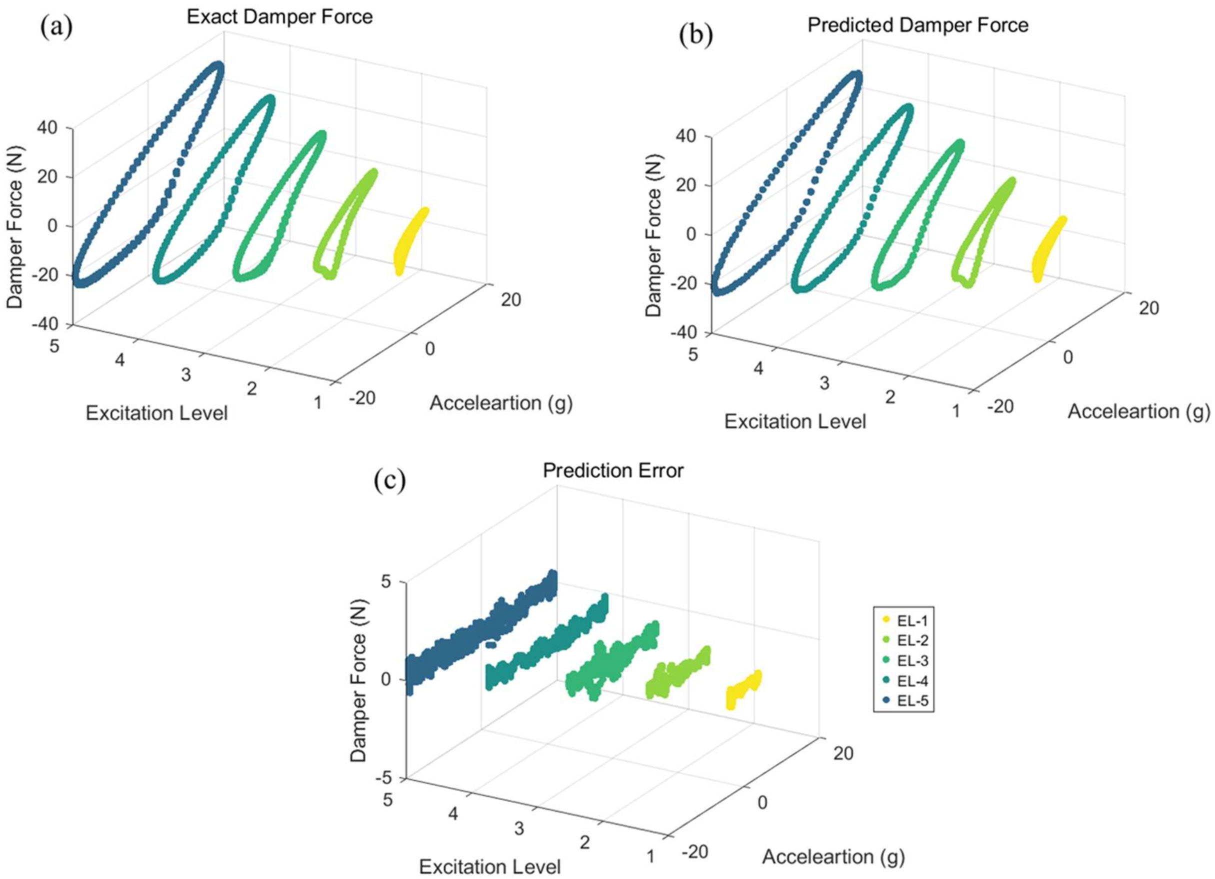


Figure 4-22 Hysteresis loop of 100 Hz: (a) Exact damper response force recorded in the experiment; (b) The predicted damper force response from the proposed model; (c) The prediction error of the surrogate model.

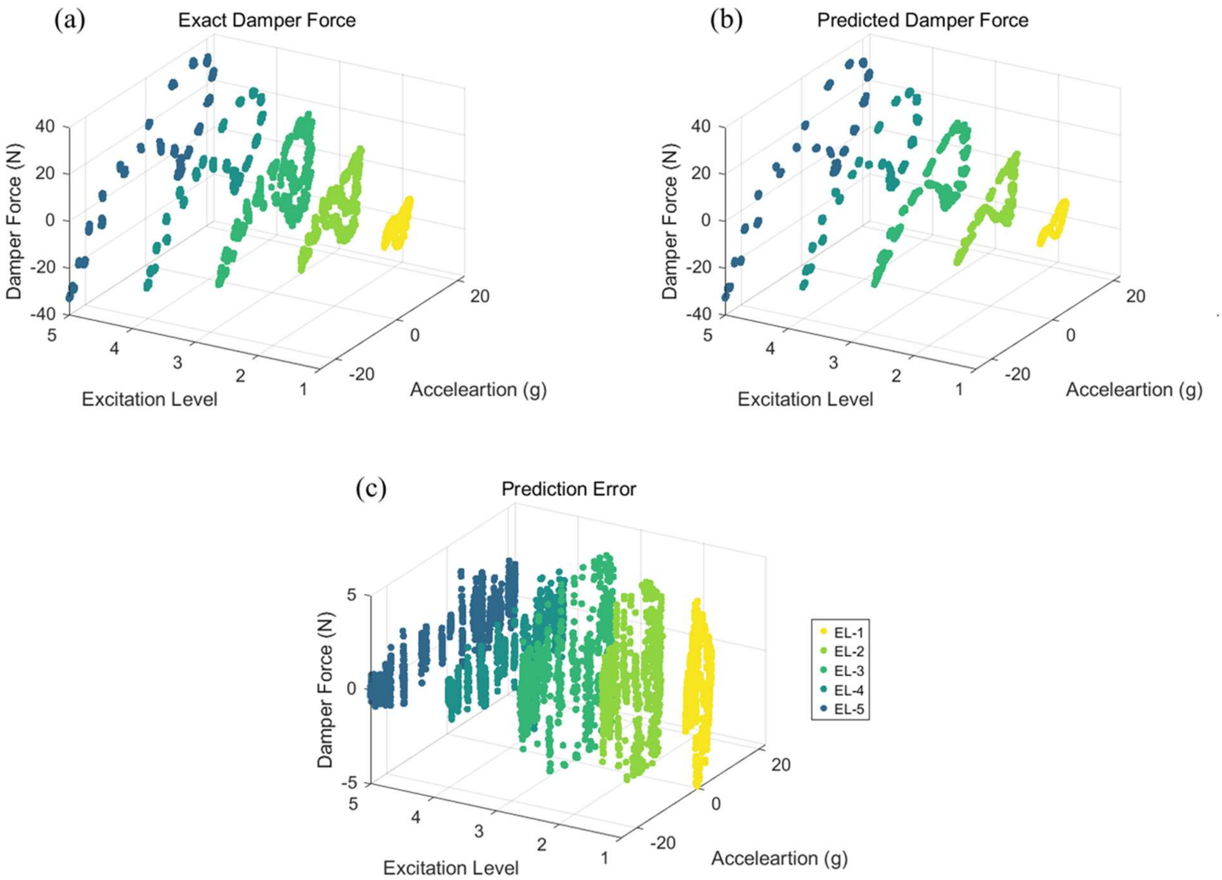


Figure 4-23 Hysteresis loop of 1200 Hz: (a) Exact damper response force recorded in the experiment; (b) The predicted damper force response from the proposed model; (c) The prediction error of the surrogate model.

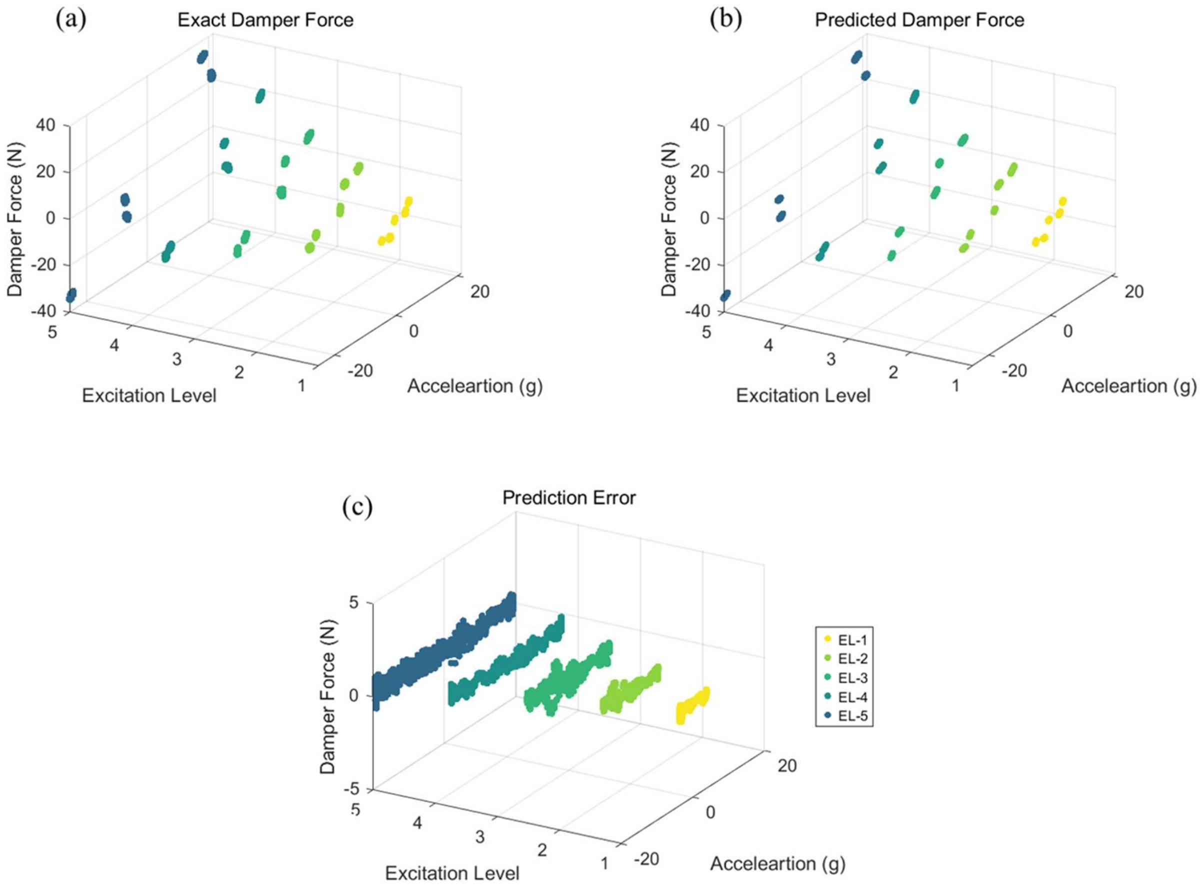


Figure 4-24 Hysteresis loop of 2000 Hz: (a) Exact damper response force recorded in the experiment; (b) The predicted damper force response from the proposed model; (c) The prediction error of the surrogate model.

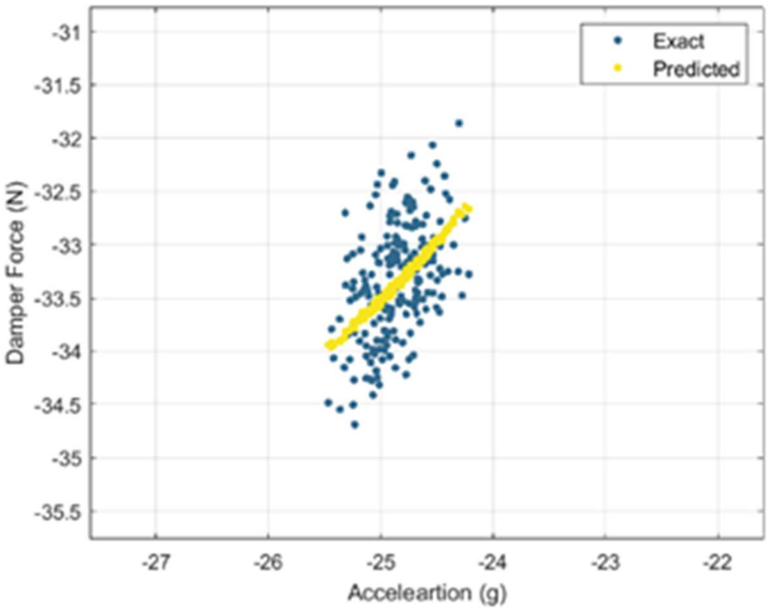


Figure 4-25 Exact and predicted damper force under fifth excitation level at 2000 Hz (phase $-\pi$).

The above prediction error is related to the shape of the hysteresis loop and the experimental error. Figure 4-26 presents the MSE prediction on the entire dataset of the proposed surrogate model and a plain NN model for each frequency. The error is high for the proposed model in the 800 to 1900 Hz range. The experimental result (see Figure 4-14) shows that the experimental error is relatively high in this range (thick trajectory). However, the signal is sampled five times per period at the frequency of 2000 Hz, reducing the difficulty in regression. Both models find the most considerable error at 1200 Hz, reflecting the identical influence of the unequable hysteresis loop shape.

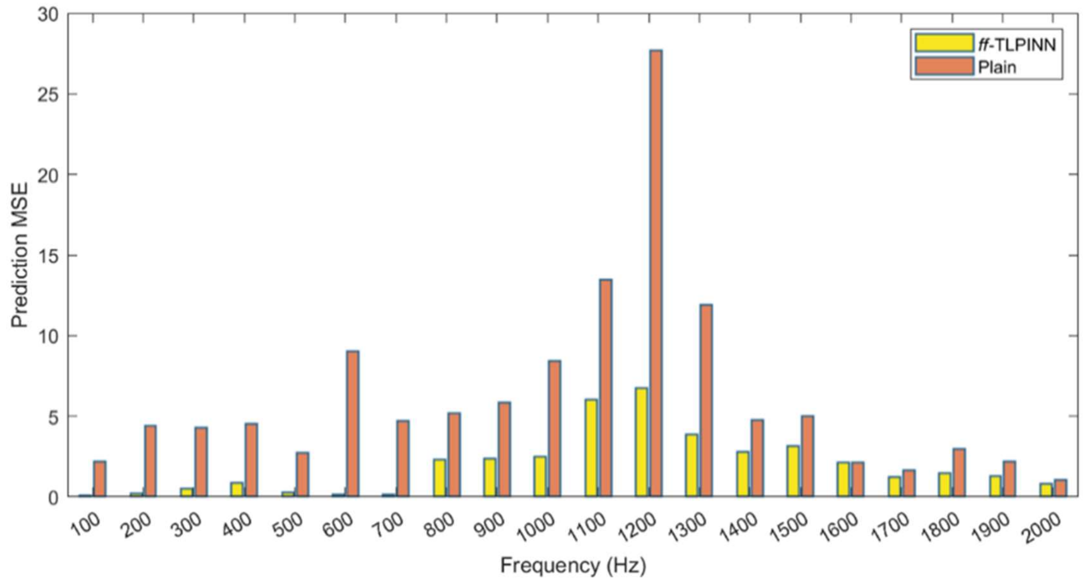


Figure 4-26 Prediction MSE of the proposed model and the plain NN model with respect to each frequency.

4.5.2 Model validation using sweep-sinusoidal dataset

Section 4.3.3 mentioned that the hysteretic effect of PD under a slow sweep-sinusoidal excitation is similar to the steady-state case. Therefore, the proposed surrogate model can be validated by the sweep-sinusoidal dataset. After establishing the surrogate model based on the steady-state dataset, the excitation signal from the sweep-sinusoidal test is input to the surrogate model to predict the corresponding force response. The result presented in Figure 4-27 is expressed by force response in terms of the excitation velocity in place of acceleration in order to have better visualization. However, the test is still acceleration controlled, as same as the steady-state test. In general, the model output captures the feature of the damper force. However, an enlarged prediction error can be observed compared to the steady-state case in which the model is trained. Specifically, the MSE prediction on the two datasets is 4.83 for the sweep-sinusoidal dataset and 0.90 for the steady-state dataset, respectively.

In this case, the prediction error in the high-frequency range is more significant with respect to a lower frequency. The error could come from the transient effect at this range when the damper response tends to be more complex in high-frequency vibration. The hysteresis loop recorded in Figure 4-15 supports this assumption, where the shape of the hysteresis loop shows a thick trajectory, indicating the instability from the transient response of the PD.

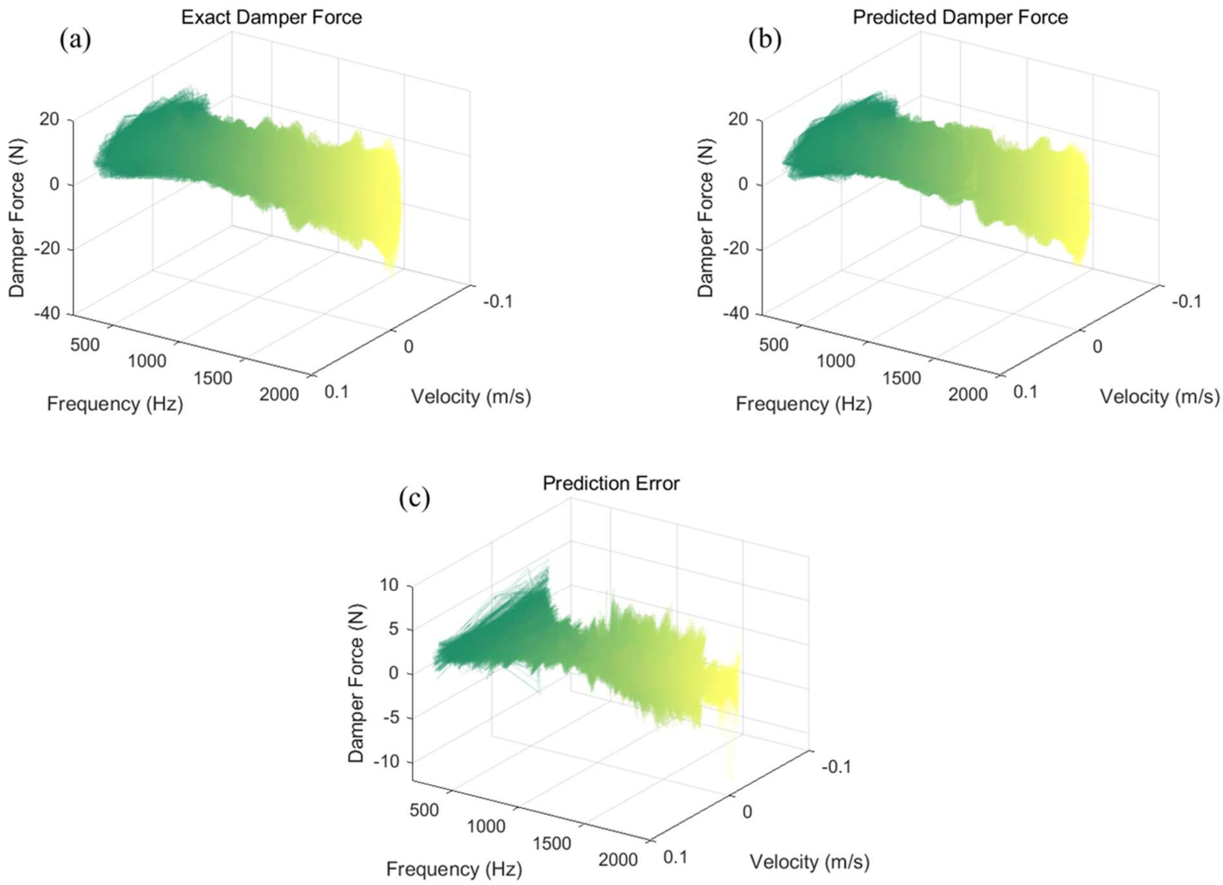


Figure 4-27 (a) Exact damper force response recorded in the sweep-sinusoidal experiment; (b) The predicted damper force response from the proposed model; (c) The prediction error of the surrogate model.

4.6 Summary

The framework exploiting NN to establish the surrogate model for characterizing a particle damper (PD) was investigated in this paper. After in-depth exploration of the long-standing issue of spectra bias in neural networks (NNs), the proposed Fourier neural network (*ff*NN) is equipped with the capability to acquire high-frequency features. This simple technique is concatenated with the TL approach, which integrated PINN. The findings indicate that the efficacy of TL-PINN is limited when dealing with a task that has an abundant amount of high-fidelity experimental data. The conclusions can be summarized as follows:

1. Compared to the wavelet features, the Fourier features embedding performs better in this case. Besides, the hyperparameter σ should keep in line with the frequency feature in the objective function. The proper imposition of this method can address the pathology of spectra bias of NNs.
2. The number of free layers released in the second training step in TL-PINN influences the result. For a case with abundant high-fidelity learning resources, fewer free layers can speed up the training but hinder the accommodating ability, which worsens the performance of NNs. Therefore, more free layers are preferred in such situations.
3. The concatenated method *ff*-TLPINN is reliable for reconstructing the hysteretic effect of PDs in a broad frequency band. The MSE arising from the *ff*-TLPINN model (0.90) is less than one-seventh of the plain network (6.60) with the same architecture. Correspondingly, the MSE is 4.92 for TL-PINN and 1.27 for *ff*NN.

This study investigates a PD with 30% of tungsten powder under various excitations. However, the performance of a PD is related to plentiful parameters, e.g., particle material, particle diameter, filling ratio, size, and shape of the damper cavity, etc. Therefore, a surrogate model considering all these parameters should be presented to design an optimum PD in noise/vibration control. However, merging all these factors into the input feature space will exponentially increase the size of the input dataset. Such a task will be too bulky for NNs as the input size on the temporal space is enormous (1,000 stamps). Confronting the issue of dimension exploration, a re-design of the input-output scheme, which excludes the temporal space, is required. Moreover, the experimental error observed in Figure 4-25 is another burden for the identification by the proposed approach. The experimental error should also be considered in NN training.

The future improvement of this study can be pursued by utilizing the phase of the periodic signal in place of the temporal stamps as the input-related factor. This approach enables the response obtained at the same phase to be viewed as a distribution (see Figure 4-25). The experimental error can thus be integrated into model learning. This idea is naturally oriented to the Bayesian theory-supported neural networks, where probability is imported to regulate the parameters in a network that differs from a point-estimate network. In another line of work, a large class of hysteresis loops are described by fractional derivatives. PINN can pursue the equation discovery of particle damping in high-frequency range based on fractional derivation equations. The Fourier features embedding can also be investigated in facilitating that attempt.

CHAPTER 5.
Model Extension from Basic Model: mc-ESDN for Sequential Information Decoding

5.1 Introduction

The rail particle damper (RPD) based on particle damping technology was developed for controlling the noise and vibration issues that occurred in the railway transit industry. Two patents have been issued on this kind of rail dampers, i.e., “modular steel rail particle damper for vibration and noise reduction of rail transit” (ZL202120908352.2) and “rail transit vibration and noise reduction method based on modular steel rail particle damper” (ZL202110466428.5). The inventors are Dr. Masoud Sajjadi, Prof. Yi-Qing Ni, Dr. Jason Lin and Mr. Chao Zhang. The damper case follows the CHN60 rail profile, which is widely used in China’s railway systems. The main body of the RPD is a hollow box comprising twelve cavities arranged in two rows with varied sizes, as shown in Figure 5-1. The rail web of the CHN60 has a curved-shaped profile, so a tailor-made fixture with a specific clamping mechanism needs to be designed. The purpose is to ensure that the RPDs will be mounted tightly and safely on the rail track in long-term operation. To that end, a special fixture is designed and fabricated. The designed fixture is robust, reliable, and safe. Furthermore, it is easy to install dampers with this kind of fixture on all kinds of rail track systems just by using a simple screwdriver.

Figure 5-1 Main body of the RPD (Designed by Seyed Masoud Sajjadi Alehashem).

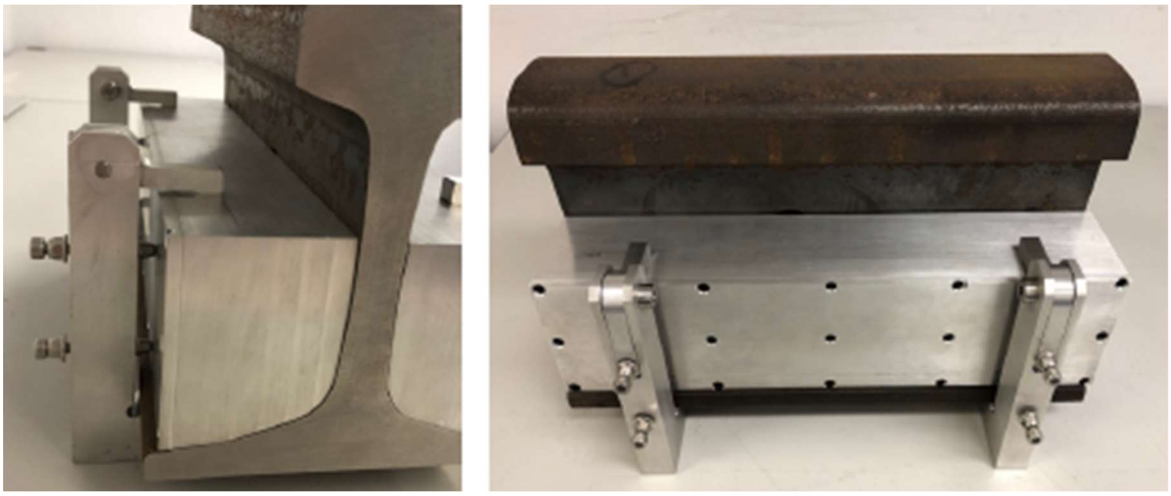


Figure 5-2 Fixture of the RPD (Designed by Seyed Masoud Sajjadi Alehashem).

As shown in Figure 5-2, the fixture consists of two main components, the U-shape body underneath the rail track serves as the frame of the fixture, and the L-shape clamps are screwed on the U-shape body. By applying force to the screw, the dampers are tightly attached to the rail track and vibrate along with the railway during operation.

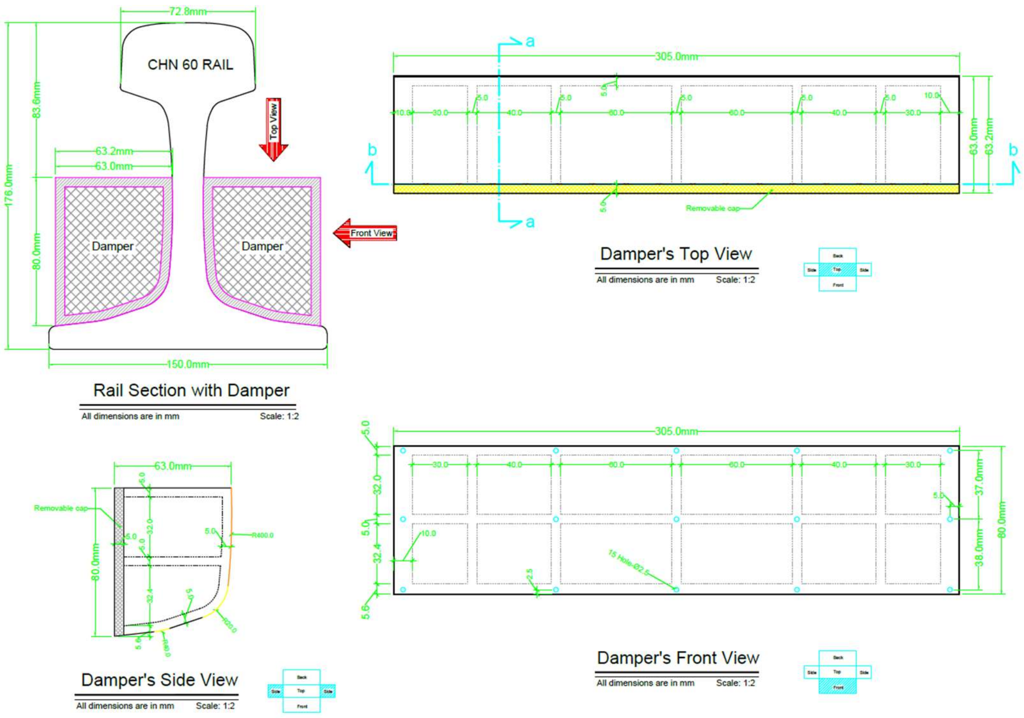


Figure 5-3 Dimensions of the RPD cavity (Designed by Seyed Masoud Sajjadi Alehashem).

The total length of the damper cavity takes half the length of the rail track bay (600 mm) so that it is able to apply sufficient effective area on the track while not interfering with the track clips (Figure 5-3). The damper cavity is divided into three different width units (30 mm, 40 mm, and 60 mm). This design is to reflect the concept that different width units can be more effective in a specific frequency band. However, this concept still requires further verification.

In the previous chapters, the framework of adopting a data-driven deep-learning method based on the experimental data collected through the shaker test is proposed to establish the surrogate model of the particle damper. Here, a similar methodology can be applied to the modeling of this RPD. Still and all, several issues need to be clarified for the feasibility of conducting the shaker test to obtain the dataset for training. First, the RPD is not centrosymmetric. Such a configuration will bring a considerable side-effect during the vertical excitation. Especially when the particle bed inside the damper is excited, the barycenter is not statical, which may bring shifting or torsion motion to the damper cavity. However, in this interpretation research, it is better to eliminate the multiple motion modes and constrain the vibration of the RPD in the vertical direction, the same as what was carried out in previous chapters. The solution to this issue is to connect two RPDs symmetrically to form the testing piece. A specialized modification to the damper configuration is thus demanded.

However, this modification will bring another issue, which is the weight of the testing piece. Taking an RPD filled with 50% 1.5 mm steel particles for example, the total weight of such an RPD is around 4.5 kg. The KD3001B impedance head that mounted on the shaker is connected with an M5 screw. For a standard M5×8 screw that made of 304 stainless steel, the strength of extension $\mathcal{N}_t$ is 400 MPa with a yield ratio of 0.8. The yield extension force F_t of this screw is calculated as follows:

$$F_t = 0.8\mathcal{N}_t \times \frac{\pi}{4}\left(d - \frac{13p}{24\sqrt{3}}\right)^2 \approx 4.5\ KN \tag{5-1}$$

The weight of two RPDs, including the modified fixture, would exceed 10 kg. In this condition, the vibration with a peak acceleration value of 50 g would bring the inertia force higher than 5 KN, which is beyond the endurance limit of the standard M5×8 stainless steel screw. Recall the need to explore the efficacy of the cavity size. These factors motivate us to select the section with the same width of the RPD as a unit. With this, the RPD is divided into three units with different width in this way. The modified testing piece is designed as two asymmetrically combined units of the same size, as shown in Figure 5-4.

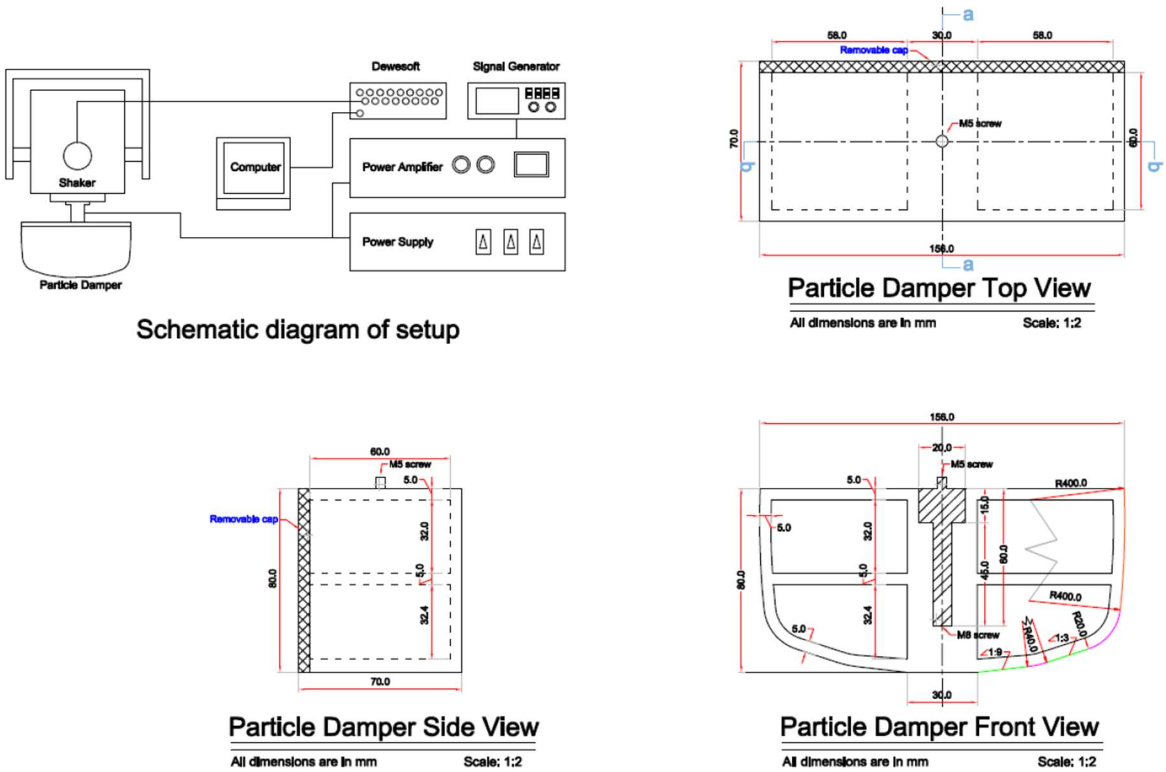


Figure 5-4 Drawing of the modified testing piece.

The testing piece is hung under the shaker to avoid the bending moment applied on the impedance head. The cavity is made of aluminum alloy, the same as the RPD. An M5 screw is set in the middle of the upper face of the testing piece for connection with the impedance head. This connector is made of stainless steel to bear the high-frequency load. The connector is embedded in the testing piece with the M8×45 screw. In the following description, each testing piece, according to its cavity width, will be briefed as RPD-30, RPD-40, and RPD-60, respectively.

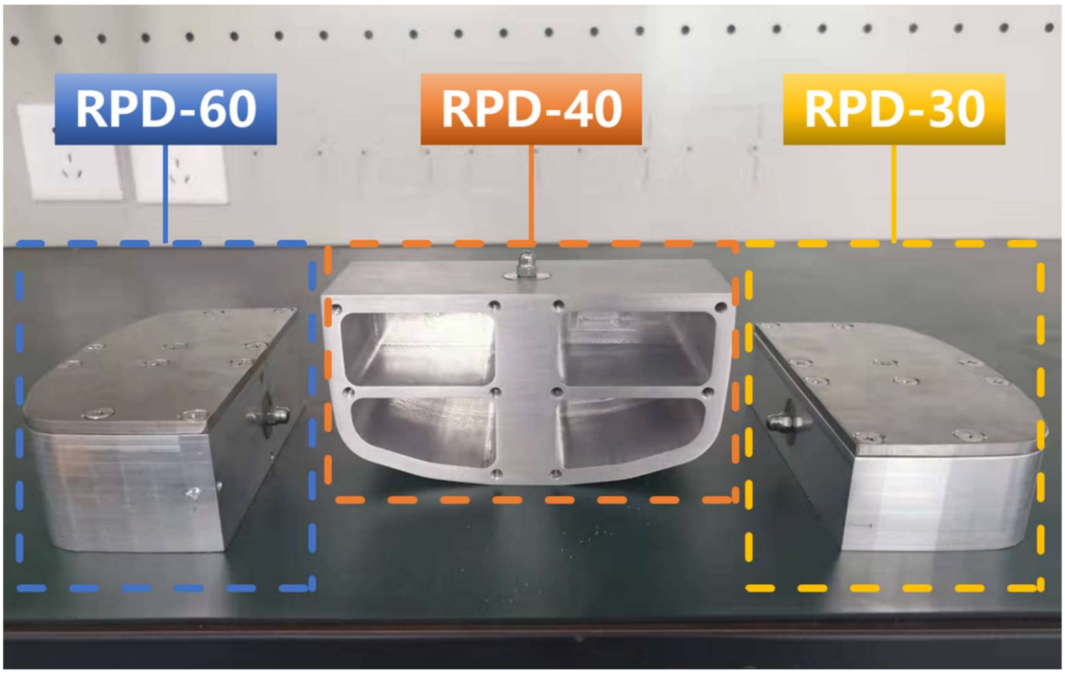


Figure 5-5 The modified testing pieces.

Figure 5-6 The cavity of the testing piece fitting the CHN-60 rail profile.

In the previous chapter, the hysteretic response force of the particle damper was reconstructed with the data-driven DNN, i.e., *ff*-TLPINN. However, such a model only works for a specific configuration of particle damper since no parameter related to the damper itself is fed to the training of the DNN. Nevertheless, in the confronting scenario, three particle dampers should be modeled. A trivial attempt in this case is to repeat the work three times and obtain three models for the task. However, in this chapter, an integral model that extends the model from a single particle damper to another particle damper is explored.

In *ff*-TLPINN, the underlying relationship between the input and the output is analyzed point-wisely: every input-output pair is regarded as independent of other pairs. Yet the dataset obtained from the shaker test is, in fact, a group of time series. The sequential information is completely neglected in this way of training. The recurrent neural network (RNN) is a potent tool that integrates large dynamical memory and adaptable computational abilities to recognize sequential information (Salehinejad et al., 2018). Based on the topological structure of RNN, the delay loop embedded in the neurons enables RNN to form a stacked flow to process the series data. In the development of RNN, gating is applied in the gated recursive neuron to improve the long-term memory of the structure (Cho et al., 2014). The gates can link the current state with another state from a long temporal distance before. In the bidirectional RNN structure, even the future state can be linked to the current stage. Therefore, the long short-term relationship along the series can be reconciled by the RNN. The appearance of gates can enhance the states updating when facing noise-polluted data and prevent undesired updates.

However, conventional RNN is inherently tedious to train, especially when processing the long-term interrelation series. Since the parameter space updating with an extensive size are prone to fall into the pathology of gradients vanishing or exploding easily (Bengio et al., 1994). In this regard, a variation of the RNN that introduces reservoir computing is formed and proposed (Yeo, 2019), which is the echo-state neural network (ESN). In ESN, the reservoir is defined by a large-scale latent state space, and the neurons inside this latent state space are sparsely connected to ensure a limited spectral radius. In practice, the connection weights inside the reservoir usually have only 1~2% non-zero values. Such a feature of the reservoir is required since the sparsity amends the decoupling of the sub-network, and in this way, the state space dynamics can be under control (Jaeger, 2002).

A typical ESN can be regarded as a three-layer network: the input signal links to the latent state space with input weights, and the reservoir serves as a feature extractor. Finally, the readout layer outputs the target signal. The input weights and the reservoir are randomly generated and remain unchanged during training, and only the readout layer will be updated. Furthermore, the readout layer usually is a linear regression from the reservoir to the target signal. Thus, the training load of the ESN is significantly reduced compared to standard RNN. The ESN is conceptually similar to the Gaussian process. They all project the input signal into a high-dimensional space, where the nonlinear features are decoupled and more easily captured.

Based on the ESN, improvements such as introducing multi-scale convolutional kernels are proposed (He et al., 2022). Using multiple convolutional kernels with various dimensions and scales (MDSC) can extract features in different receptive fields, and the different scale spatial-temporal correlations can be focused separately.

In this section, the principles of ESN and MDSC will be first discussed. Then follows the structure of the proposed mc-ESDN. The dynamic experiments on the RPDs are conducted to acquire datasets for the model establishment. Two baseline methods utilizing the plain-NN and the TL-NN are compared with the proposed method in this chapter. These models are trained and tested on the steady-state simple harmonic dataset. The dataset obtained from the sweep-sinusoidal experiments will be used to validate the established surrogate model.

5.2 Establishment of mc-ESDN

5.2.1 Principles of ESN

The echo state network has been extensively studied since it has pronounced the amendable ability in learning complex nonlinear dynamics in series signal. The nonlinearity of the ESN is embedded in its activation function for the state updating. As aforementioned, the latent state space is randomly generated, along the time axis of the series signal. Following the work of Yeo (2019), the evolution of the latent state is defined as:

$$\boldsymbol{s}_{t+1} = \lambda \boldsymbol{s}_t + (1-\lambda)\Psi_a(\boldsymbol{s}_t, \boldsymbol{x}_t, \boldsymbol{y}_t) \tag{5-2}$$

Here, $\boldsymbol{s}_t \in \mathbb{R}^{N_s}$ represents the latent state, $\boldsymbol{x}_t \in \mathbb{R}^{N_x}$ is the input series signal to the ESN, the relaxation coefficient $\lambda \in [0,1)$ aims to call back the memory from the last timestamp to pursue the continuity of the state evolution. $\boldsymbol{y}_t \in \mathbb{R}^{N_y}$ is the vector or signal value of the target signal at the current state. The activation function $\Psi_a(\cdot)$, such as the common hyperbolic tangent function, provides nonlinear property to the ESN. The subscript t presents the general time stamp.

$$\Psi_a(\boldsymbol{s}_t, \boldsymbol{x}_t, \boldsymbol{y}_t) = \tanh(\boldsymbol{W}_{\boldsymbol{res}} \cdot \boldsymbol{s}_t + \boldsymbol{W}_{\boldsymbol{in}} \cdot \boldsymbol{x}_t + \boldsymbol{W}_{\boldsymbol{back}} \cdot \boldsymbol{y}_t) \tag{5-3}$$

Aside from calling back the last state, the activation of the state space contains the information of all three layers. Here, $\boldsymbol{W}_{\boldsymbol{res}} \in \mathbb{R}^{N_s \times N_s}$, $\boldsymbol{W}_{\boldsymbol{in}} \in \mathbb{R}^{N_s \times N_x}$, and $\boldsymbol{W}_{\boldsymbol{back}} \in \mathbb{R}^{N_s \times N_y}$ are independent weight matrices that are generated by the uniform distribution. The structure of ESN presented by Equations ((5-2) and ((5-3) are very similar to a standard RNN. Although in RNN, the weight matrices are calculated through the maximum likelihood method, the weights mentioned in Equation (5-2) will not be updated since they were generated. The reservoir weights, $\boldsymbol{W}_{\boldsymbol{res}}$, define the connectivity inside the latent space. As aforementioned, the $\boldsymbol{W}_{\boldsymbol{res}}$ should be a sparse matrix to ensure a richer internal dynamic of the reservoir (Jaeger & Haas, 2004).

In summary, the structure of a standard ESN is depicted in Figure 5-7.

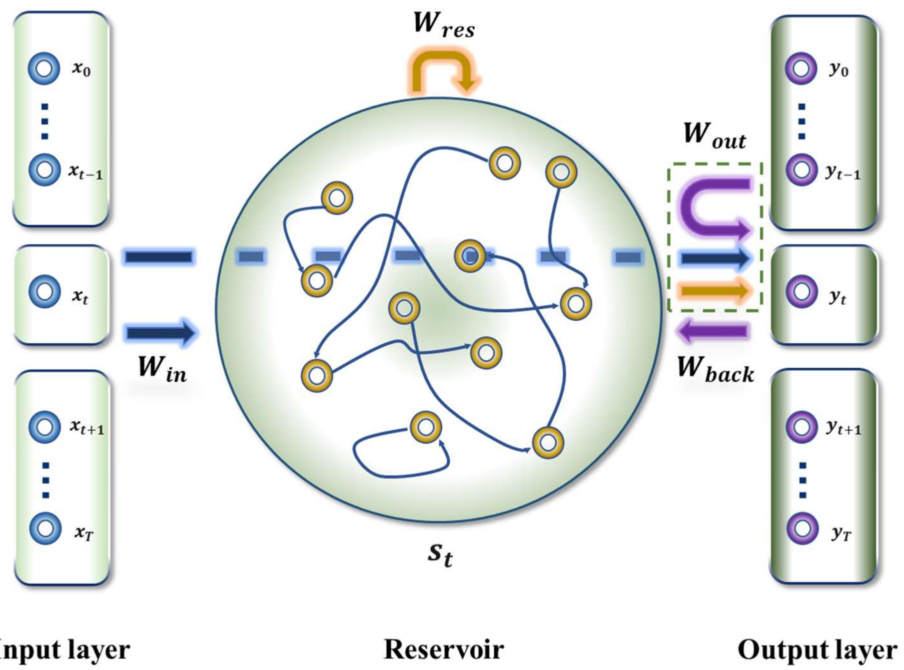


Figure 5-7 Structure of a standard ESN.

Once the latent space is generated, it should go through a certain free-running period before it is ready to generate the output signal.

$$\widehat{\boldsymbol{y}}_t = \boldsymbol{W_{out}} \cdot \left[\boldsymbol{x}_{t-m_1}, \dots, \boldsymbol{x}_{t+m_1}, \boldsymbol{s}_{t-m_2}, \dots, \boldsymbol{s}_{t+m_2}, \widehat{\boldsymbol{y}}_{t-m_3}, \dots, \widehat{\boldsymbol{y}}_{t-1}\right] + \boldsymbol{b_Y} \tag{5-4}$$

Following the observation given by Lukoševičius and Jaeger (2009), taking the input, current state, the previous output, and their delayed form together can improve the model accuracy. Here, m_1, m_2, and m_3 are the delayed time steps of input, current state, and the previous output time series, respectively. Counting the output series stops at the last time step $\widehat{\boldsymbol{y}}_{t-1}$, since it is impossible to forecast the future output in the model output stage.

The model training step is only focused on the $\boldsymbol{W_{out}} \in \mathbb{R}^{N_y \times (2m_1N_x + 2m_2N_s + m_3N_y)}$ in a standard ESN. The task is to find a linear map of the target signal. The training is relatively straightforward even though the reservoir has extracted plenty of information. Consider a training series $\boldsymbol{Y} = (\boldsymbol{y_0}, \dots, \boldsymbol{y_N})$ with N+1 segments. Here N+1 is the number of training data. This linear map can be determined by following the regularized optimization problem, which can be presented by (Yeo, 2019):

$$\min_{\boldsymbol{W_{out}}} \sum_{i=0}^{N} \frac{1}{2} \|\boldsymbol{y_i} - \widehat{\boldsymbol{y}}_{\boldsymbol{i}}\|_2^2 + \frac{\beta}{2} \|\boldsymbol{W_{out}}\|_2^2 \tag{5-5}$$

In which $\|\cdot\|_F$ is the Frobenius norm, the regularization parameter β ensures the sparsity of the calculated $\boldsymbol{W_{out}}$, so that the generated model has a better ability of generalization. The analytical solution of Equation ((5-5) is as follows:

$$\boldsymbol{W_{out}} = \left(\sum_{i=1}^{N} \boldsymbol{R_i}\,\boldsymbol{R_i^T} + \beta\boldsymbol{I}\right)^{-1} \left(\sum_{i=1}^{N} \boldsymbol{R_i}\,\boldsymbol{y_i^T}\right) \tag{5-6}$$

where,

$$\boldsymbol{R_i} = \left(\boldsymbol{x}_{i-m_1}, \dots, \boldsymbol{x}_{i+m_1}, \boldsymbol{s}_{i-m_2}, \dots, \boldsymbol{s}_{i+m_2}, \widehat{\boldsymbol{y}}_{i-m_3}, \dots, \widehat{\boldsymbol{y}}_{i-1}\right) \tag{5-7}$$

The free parameters of an ESN are listed in Table 5-1.

Table 5-1 Free parameters of an ESN

λ	Temporal relaxation parameter
β	L_2-norm regularization parameter
ρ_{max}	Spectral radius of $\boldsymbol{W_{res}}$
ξ_{in}	$\boldsymbol{W_{in}} \sim \mathcal{U}(-\xi_{in}, \xi_{in})$
ξ_{res}	$\boldsymbol{W_{res}} \sim \mathcal{U}(-\xi_{res}, \xi_{res})$
ξ_{back}	$\boldsymbol{W_{back}} \sim \mathcal{U}(-\xi_{back}, \xi_{back})$
m_1	Temporal delay step for input
m_2	Temporal delay step for states
m_3	Temporal delay step for output

The ESN holds hailed good performance on lost data reconstruction from sparse observation data. It elucidates that ESN has a pronounced ability to recognize a nonlinear map. Usually, the fixed-point reconstruction with ESN starts with linear interpolating of the observed data as the initial output $\boldsymbol{Y^0}$. After each iteration of weight updating with Equation (5-5), an updated output $\boldsymbol{Y}^U$ would be obtained, which can be replaced by $\boldsymbol{\mathcal{E}}(\boldsymbol{Y}^{U,k}, \boldsymbol{R})$. The fixed-point iteration reconstruction is then formulated as (Yeo, 2019):

$$\boldsymbol{Y}^{U,k+1} = \boldsymbol{\mathcal{G}}(\boldsymbol{Y}^{U,k}, \boldsymbol{R}) = \alpha \boldsymbol{Y}^{U,k} + (1-\alpha)\boldsymbol{\mathcal{E}}(\boldsymbol{Y}^{U,k}, \boldsymbol{R}) \tag{5-8}$$

The newly introduced $0 < \alpha < 1$ is to make sure the Jacobian of $\boldsymbol{\mathcal{G}}$ exists. Here, the convergence of this proposed method is derived according to the Banach fixed-point theorem, which gives the sufficient condition for convergence. This theorem requires a contraction mapping:

$$\|\boldsymbol{\mathcal{G}}(\boldsymbol{Y}^{a}, \boldsymbol{R}) - \boldsymbol{\mathcal{G}}(\boldsymbol{Y}^{b}, \boldsymbol{R})\| \leq L\|\boldsymbol{Y}^{a} - \boldsymbol{Y}^{b}\| \tag{5-9}$$

In which $0 \leq L < 1$ is a Lipschitz constant. In this part, we are trying to give the convergence condition in a limited scope around the fixed-point ground truth $\boldsymbol{Y}^{*}$, as the assumption is too strong that a globally uniform contraction generally meets the desired nonlinear map. The Jacobian of $\boldsymbol{\mathcal{G}}(\boldsymbol{Y}, \boldsymbol{R})$ is $\mathcal{J}_{\mathcal{G}}(\boldsymbol{Y})$, and according to the contraction mapping theorem, the fixed-point iteration would converge to the observed time series, $\boldsymbol{Y}^{*}$, if $\left\|\mathcal{J}_{\mathcal{G}}(\boldsymbol{Y}^{*})\right\| < 1$ with the subordinate matrix norm $\|\cdot\|$. The convergence of ESN is based on a proper selection of the free parameters. The above derivations are given by Kyongmin Yeo (2019).

5.2.2 Multi-dimensional and multi-scale convolution

Convolutional neural networks have emerged as the leading algorithm in recent years. Its merit resides in the vigorous feature extraction in high-dimensional data. Therefore, there went a recent surge in developing this algorithm in computer vision and other application scenarios.

Therefore, an extensive study has been conducted by implementing the convolutional kernel with a flexible receptive field (Lu et al., 2018; Elboushaki et al., 2020).

In the ESN, each channel is separately inputted into the reservoir. It is a limited approach, as the correlation across channels has been ignored. Besides, some hidden features also exist along a particular channel of series data. A pre-processing of the single channel can also help the ESN recognize the underlying pattern. Regarding these factors, the convolutional kernels can be applied to the input features to extract the high-dimensional information. Convolutional kernels of different shapes and dimensions distill unique features concerning their receptive field and kernel values.

In the input feature space, there are channels of different information. In our case, it includes the excitation signal and the response force of the basic damper. The combination of them forms the input space, $\boldsymbol{X} \in \mathbb{R}^{N_{in} \times T}$. Consider a convolutional kernel $\boldsymbol{W_i} \in \mathbb{R}^{U \times V}$. The convolution of the signal is calculated by:

$$X_{C,i}(j,k) = \sum_{u=1}^{U_i} \sum_{v=1}^{V_i} W_i(u,v) \times X(j-u+1, k-v+1) \tag{5-10}$$

After this operation, the original series would be shortened to $T - V + 1$. Therefore, the start and the end of the convolution should be compensated by the original signal to ensure the length of the convolution is equal to the original signal. With this, the convolution $\boldsymbol{X_C}$ and original input $\boldsymbol{X}$ can be concatenated for the ESN.

There are kernels with one dimension in time, which learns the time variation laws of a specific channel. Kernels with one dimension in space aim to extract the correlation across channels (Elboushaki et al., 2020). The 2-D kernels combined the information from both dimensions. Moreover, the various conceptions of different scales enable the kernels to obtain local and general feature laws. In detail, short-scale kernels focus on short-term features, and more high-frequency information would be learned. The long-scale kernels serve oppositely. Blending the kernels formulates the muti-dimensional and multi-scale convolutional kernels (MDSC).

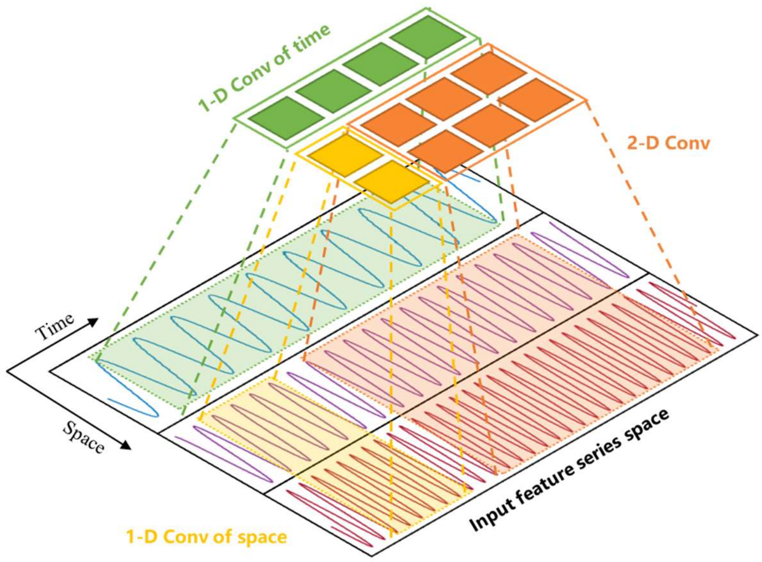


Figure 5-8 Applying MDSC to the input series.

Up to now, a great many of useful kernel operators have been developed. The Prewitt (Song et al., 2019) is one of the most widely adopted convolutional detectors to approximate the gradient of a 2-D matrix. Besides that, there also are Robert cross operator (Albdour & Zanoon, 2020), Sobel (Wang et al., 2018), Laplacian (Qi & Liu, 2020) and difference of Gaussian (DoG)

(Assirati et al., 2014) operators. In this study, various MDSCs are employed as the features' extractors. The entries of the useful candidate kernels are shown in Figure 5-9.

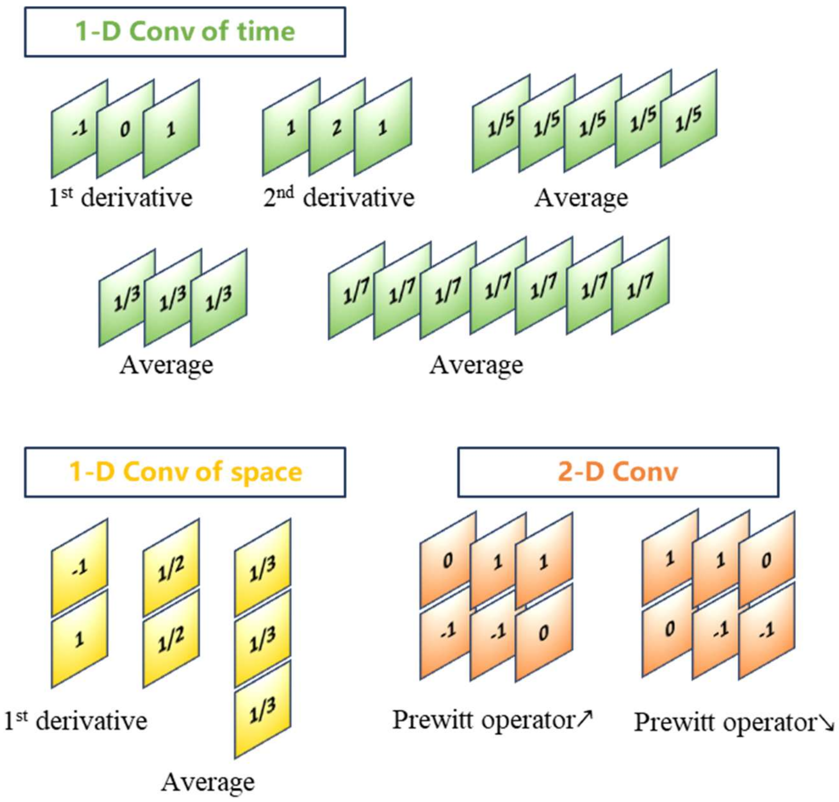


Figure 5-9 MDSC candidate kernels for feature extraction.

5.2.3 Architecture of the mc-ESDN

In this section, a model extension based on the *ff*-TLPINN is explored. First, the *ff*-TLPINN proposed in the last chapter would be applied to model the response force of one of the dampers based on the data collected from a similar shaker test. Next, the model will be extended to two other dampers. In this task, the input of the extended model contains excitation properties and the basic damper's response force under such excitation, which is generated through the *ff*-TLPINN. It should be noted that the excitation signals are not identical for different dampers, even though the frequency or amplitude is the same, because of the phase difference. The difference from the previous works is that the inputs are no longer regarded as individual "points" but as a set of time series according to the different excitation settings. That is, 100 different series (20 sets of frequencies and 5 sets of excitation levels) with a length of 1000 would be evaluated.

However, the ESN, as a sequence processing method, cannot reflect the correlation of so many sequences with different correlations. To that end, synergizing is added to the ESN by replacing the linear regression readout layer with deep fully connected layers. It lets the ESN evolve to the echo state deep neural network (ESDN) method. Besides, MDSC was also adopted ahead of the reservoir to enhance the features' distilling ability. As such, it formulates the proposed multi-dimensional and multi-scale convolutional echo state deep neural network (mc-ESDN).

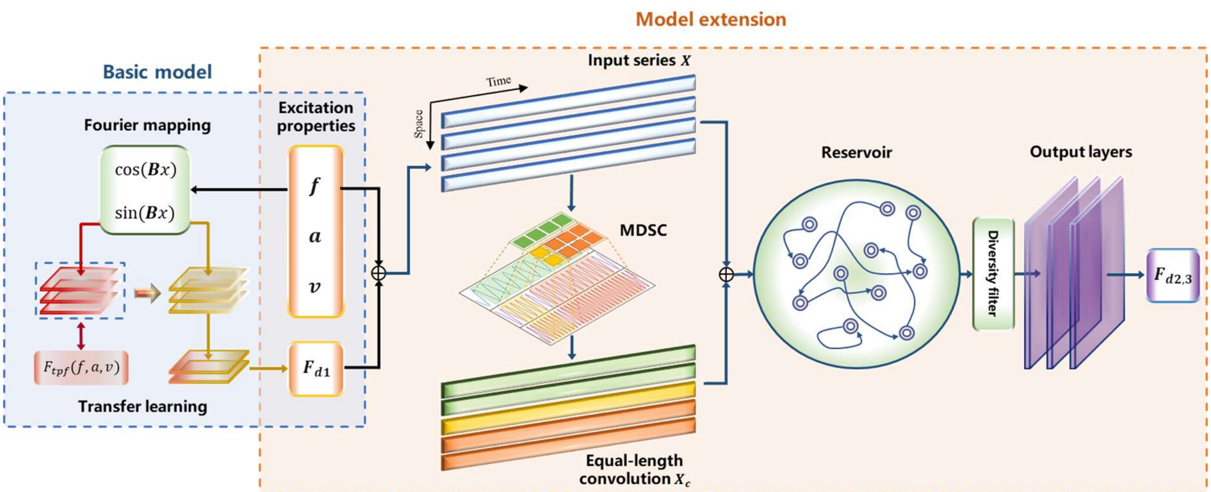


Figure 5-10 Framework of the mc-ESDN.

Figure 5-10 shows that the basic model is established through the *ff*-TLPINN with the input of the excitation properties. The basic model's output, the basic damper's response force, is combined with the corresponding excitation properties and becomes the input of the extended model. The following model aims to identify the relationship between the basic damper and the target dampers. After the procedure of MDSC, a bunch of equal-length convolutions would be generated and ready to insert into the reservoir together with the input series.

However, since the reservoir is a large latent state space, the width of the output features would be very extensive (1046 in this case). The feature space's large size would hamper the readout layer's training. Moreover, due to the sparse nature of the reservoir, a considerable number of feature channels contain duplicated information, further reducing the readout layer's performance. For a standard ESN, these sparse outputs can be easily filtered in linear regression by introducing norm regularization. However, the concept of norm regularization conflicts with the deep readout layers. Therefore, a filter after the reservoir should be manually adopted to reduce the width of features.

In this case, the original input features contain four channels: the frequency $\boldsymbol{f}$, acceleration $\boldsymbol{a}$, velocity $\boldsymbol{v}$, and response force of the basic damper $\boldsymbol{F_{d1}}$. All channels comprise 100 time series slices (20 sets of frequencies and 5 sets of excitation levels). These slices are cut out from the middle of the steady-state simple harmonic excitations with a length of 0.1 s. Since the sampling frequency is 10 kHz, every slice has 1000 time steps. After the original input features are treated by the MDSC and the reservoir, the number of channels expands to $k \gg 4$, according to the hyper-parameters. For the *k*th channel $\boldsymbol{C_k} \in \mathbb{R}^{100\times1000}$, we expect it provides more information when its 100 slices are diverse. However, if the slices in one channel are similar to each other, we decide it is a vain channel and should be filtered.

All slices $\boldsymbol{S_{ik}} \in \mathbb{R}^{1\times1000}, i \in [1,100]$ are vectors with the same dimension, so the diversity can be evaluated by calculating the L2 norm of the vector difference between two slices. Here we define the diversity value, $Dive$, which is given by:

$$Dive(\boldsymbol{S_{ik}}, \boldsymbol{S_{jk}}) = \left\|\boldsymbol{S_{ik}} - \boldsymbol{S_{jk}}\right\|_2 \quad (5\text{-}11)$$

The *Dive* value for a channel would be a matrix, recording the diversity of every slice pair. Slice pairs with higher Dive values mean the pair of slices are different from each other. The average of this matrix represents the diversity value of one channel.

$$V_k = E\left(Dive\left(\boldsymbol{S_{ik}}, \boldsymbol{S_{jk}}\right), i \in [1,100], j \in [1,100]\right) \quad (5\text{-}12)$$

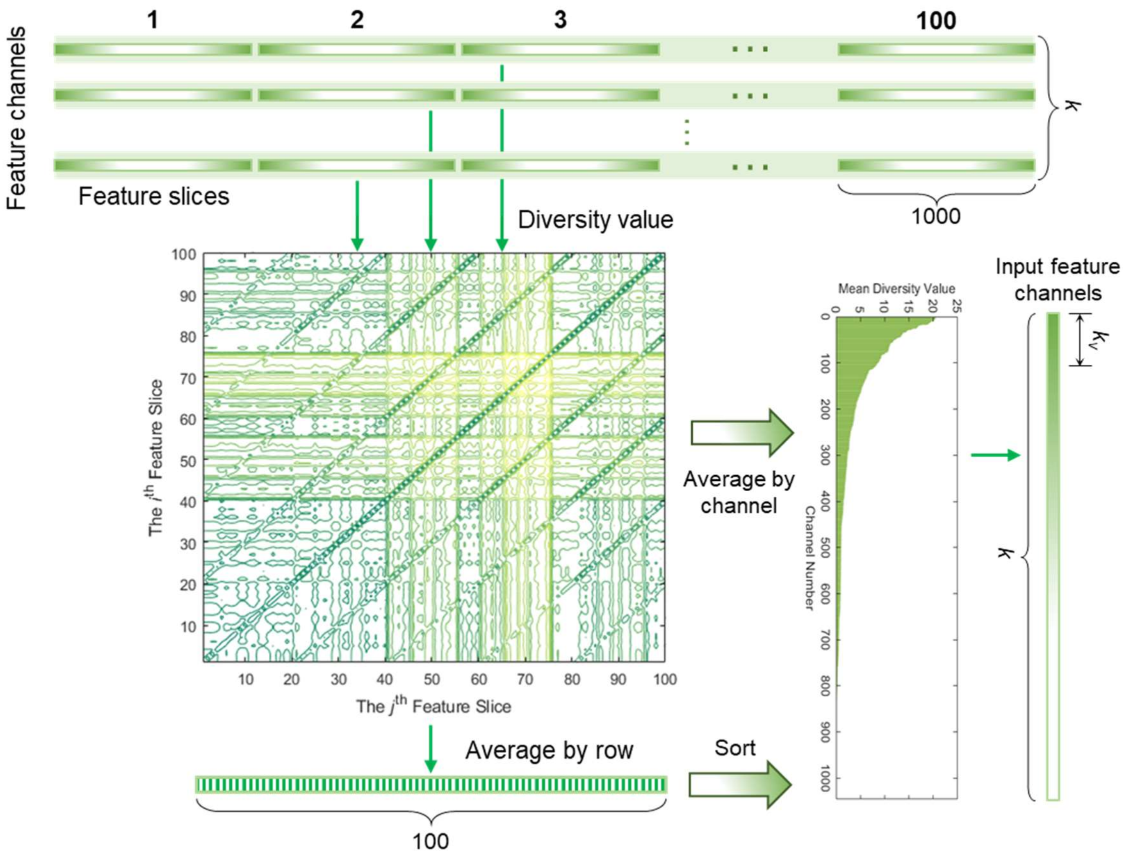


Figure 5-11 Framework of the diversity filter.

In order to make sure the diversity value is comparable across channels. All channels should be normalized to the range from 0 to 1 before calculating V_k. The diversity filter is a high-pass filter. Among all k channels, only the first k_v channels with the largest diversity value would be selected as the input feature channels into the deep readout layer. Notably, the original feature channels should be selected mandatorily.

5.3 Experiment on RPDs

5.3.1 Experimental setup

For the PDs used in this study, stainless steel balls with a diameter of 1.5 mm were chosen as the particles to be filled in the damper cavity, and the filling ratio was set as 50%. Similarly, the amount of filling particles is determined by its apparent density (4.9 g/cm^3). Thus, the weight of filling particles can be calculated according to the volume of the damper cavity. As mentioned above, there are three sizes of rail damper cavities, RPD-30, RPD-40, and RPD-60, respectively. According to the damper design, the damper is divided into upper and lower cavities inside every damper mode. The volume of the different upper and lower cavities is shown in Table 5-2. This experiment was conducted by Mr. Xiang-Xiong Li and me.

Table 5-2 Volume of damper cavity in three modes

Volume (cm^3)	RPD-30	RPD-40	RPD-60
Upper cavity	51	68	102
Lower cavity	42	56	83

Before the experiments, all the damper cavity modes were weighted in empty and filled conditions. The cavity's weight can reflect the dampers' inertia force during vibration. However, the effective mass should be determined by the response force of the damper since the particle bed inside the cavity are not steady. It means the damper is not a lumped mass to the primary structure but exhibits nonlinearity and the ability to consume vibration energy.

Table 5-3 Weight of damper in three modes

Weight (g)	RPD-30	RPD-40	RPD-60
Empty	1037.6	1071.4	1167.2
Filled	1493.3	1679.0	2073.7

Figure 5-12 Weighting of the dampers.

The HEV-2000N-H high-energy constant-resistance shaker was used in the dynamic experiments. This kind of shaker has a wide range of working frequency bands, and its response phase shift lag is small. The parameters of the HEV-2000N-H shaker are listed below.

Table 5-4 Parameters of the HEV-2000N-H shaker

Parameter name	Parameter value
Peak excitation force	2.5 kN
Force constant	40.9 N/A

Peak displacement	±8 mm
Usable frequency range	0 – 3 kHz
Movable mass	9.45 kg
Total mass	147 kg

The shaker is hanging on the gantry. The RPD is connected vertically to the shaker head with an impedance head (KD3001B). This impedance head will record the force and acceleration signal during the dynamic test. For the parameters of this sensor: the maximum force peak is 2.5 kN, the maximum acceleration peak is 100 g, and the frequency range is from 0.5 Hz to 6 kHz. A Dytran 3273 accelerometer is attached to the bottom of the RPD to monitor the acceleration signal. The maximum acceleration peak of this sensor is also 100 g. This arrangement ensures the correctness of the recorded signal from the impedance head. The sampling frequency of this dynamic test is 10 kHz. The experimental setup is shown in Figure 5-13.

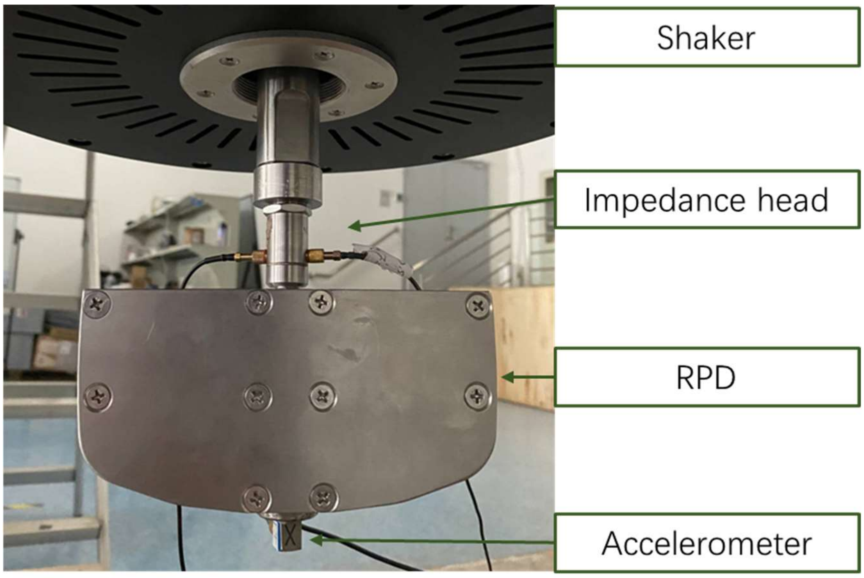


Figure 5-13 Setup of the dynamic test of the RPD.

The experiment aims to explore the response force of RPDs under a specific condition of steady-state simple harmonic excitation. Before the dynamic test, a transparent cap was installed on the RPD to observe the particles' moving pattern inside the damper cavity. The observation was conducted with an excitation frequency of 100 Hz, and three excitation levels were applied to excite the RPD. Namely, 5 g, 20 g, and 40 g, respectively. From the recorded video, it can be concluded that during a steady-state simple harmonic excitation, the particle bed undergoes a certain vibration pattern with respect to different excitation levels. This kind of pattern can be regarded as the specific mode of the particle bed. This finding implies the need of future investigation of this complex mechanism. Also, it is found that the particle bed barely moves when the acceleration amplitude is 5 g. Therefore, an elevated excitation level is preferred in the dynamic test of the RPD.

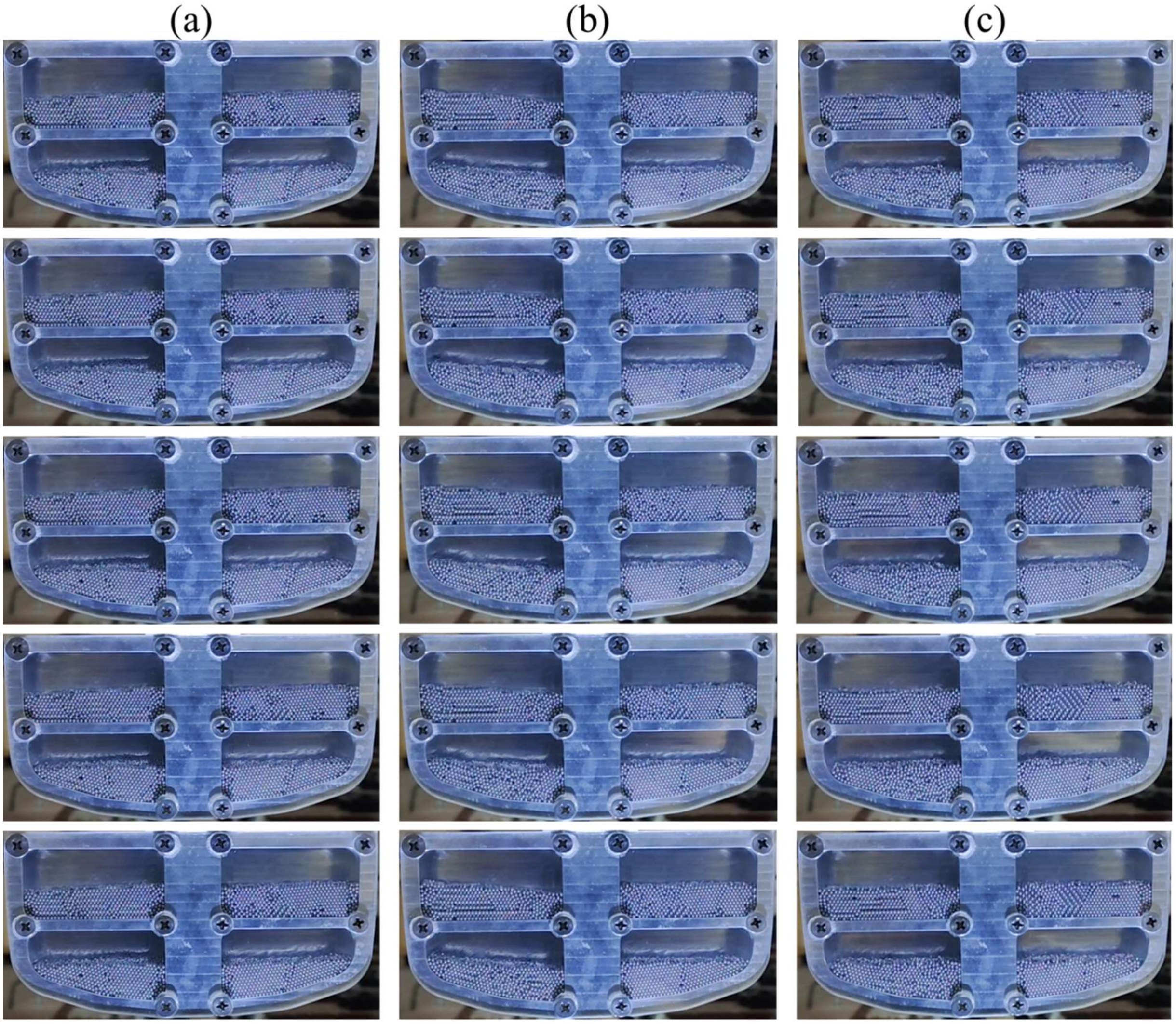


Figure 5-14 Frames of RPD vibration video under 100 Hz excitation with amplitude of: (a) 5 g, (b) 20 g, and (c) 40 g.

A sweep-sinusoidal test will also be applied to the three damper cases. The generated dataset will be used to validate the generalization ability of the established model on the frequency domain. Similar to the setting applied on the cylindrical damper experiment, the signal frequency sweeps linearly from 100 to 2000 Hz in 20 s. The sweep-frequency rate is 95 Hz/s.

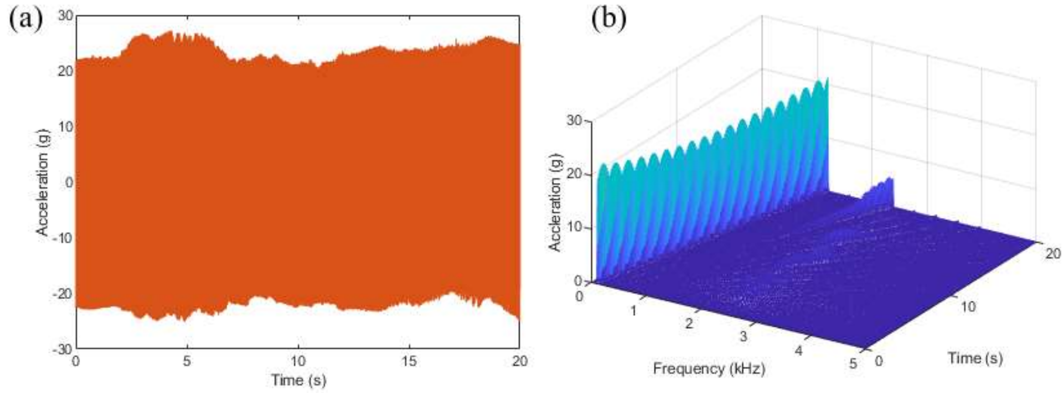


Figure 5-15 (a) Sweep sine acceleration signal in time; (b) Time-frequency spectrum of the acceleration signal.

5.3.2 Hysteresis loops of RPDs

The steady-state simple harmonic excitation frequency ranges between 100 and 2000 Hz with an interval of 100 Hz. The excitation amplitude is controlled manually by switching the gain on the amplifier. Five levels of amplitude are applied to the PD in this test: The peak acceleration value approximately ranges from around 10 g to around 50 g with a 10 g interval. Similarly, the excitation level is abbreviated as "EL". Excitation lasts 5 s on each condition, and 0.1 s of the signal in the middle starting at the π phase is selected for the investigation to prevent the transient effect. Each acceleration-controlled vibration signal piece contains 1,000 temporal stamps as the signal sampling rate is 10 kHz.

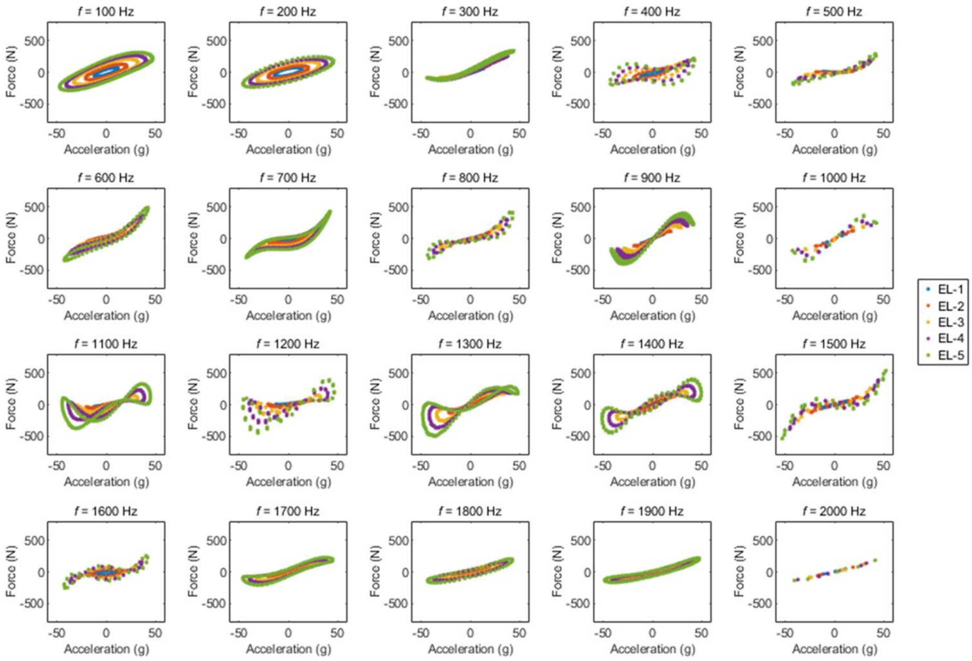


Figure 5-16 Diagram of force versus excitation acceleration under different frequencies and excitation levels of RPD-30.

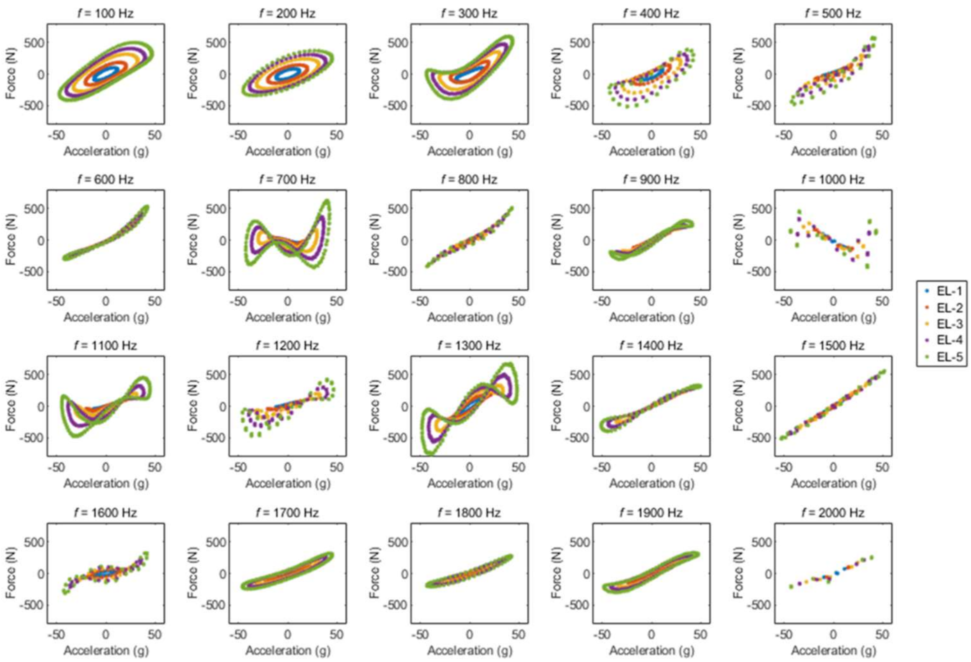


Figure 5-17 Diagram of force versus excitation acceleration under different frequencies and excitation levels of RPD-40.

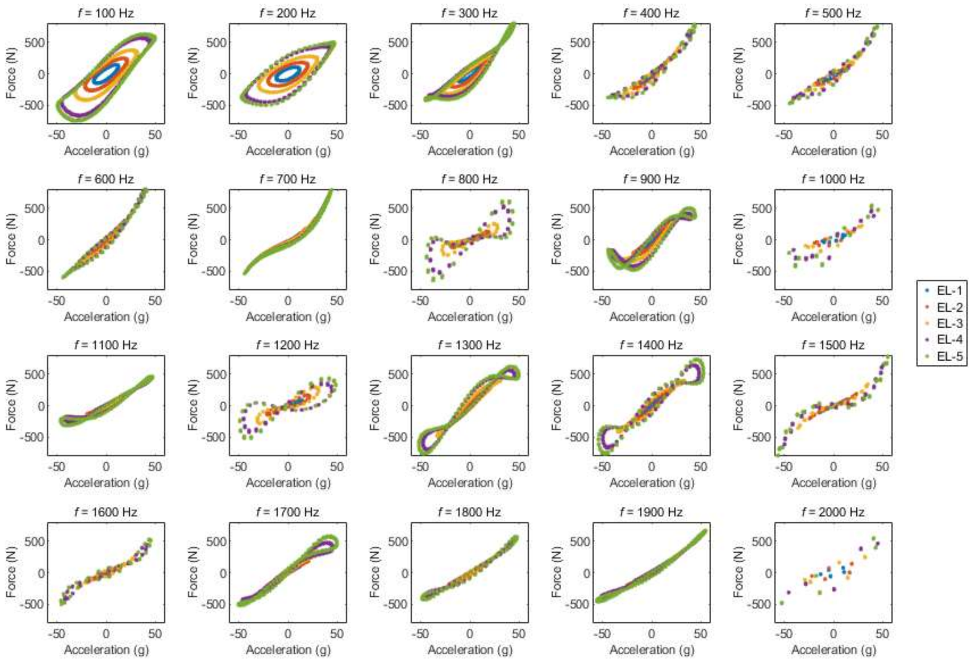


Figure 5-18 Diagram of force versus excitation acceleration under different frequencies and excitation levels of RPD-60.

Unlike the hysteresis loops presented by the tungsten powder, the stainless-steel balls are lightweight particles compared to the tungsten powder and more easily to be excited. Therefore, fickle hysteresis loop shapes were observed in this dynamic experiment. This finding indicates that the particle bed of stainless-steel balls undergoes various modes at specific excitation frequencies. Specifically, two “ears” appears at both ends of some hysteresis loops, e.g., RPD-30 in the frequency of 1400 Hz, which shows an apparent non-viscous behavior.

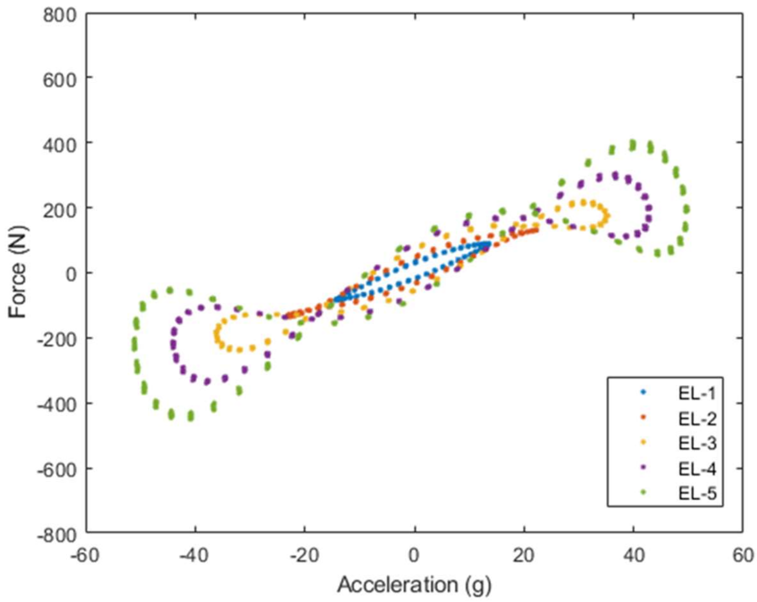


Figure 5-19 Hysteresis loops of RPD-30 under various excitation levels at 1400 Hz.

Take the hysteresis loops of RPD-30 at 1400 Hz as an example. The appearance of two "ears" is dependent on the excitation level. When the excitation level is low, the particle bed behaves like a viscous friction-combined entity. While the acceleration rises, the ears show at the ends of the loops where the peaks of the acceleration values are. To understand this phenomenon, one must first realize that the hysteresis loop does not present the acceleration-force relationship in one period but combines multiple periods at different phases.

In Figure 5-20 Time series of acceleration and damper force of RPD-30 at 1400 Hz. we can find that except for the primary frequency component, the damper force was superimposed with a lower frequency component in the time domain. This component was due to the nonlinearity of the damper, which can be interpreted as the response force from the particle bed in a specific moving mode. Influenced by the damper force, the acceleration signal also contains this component and is no longer a strict simple harmonic signal.

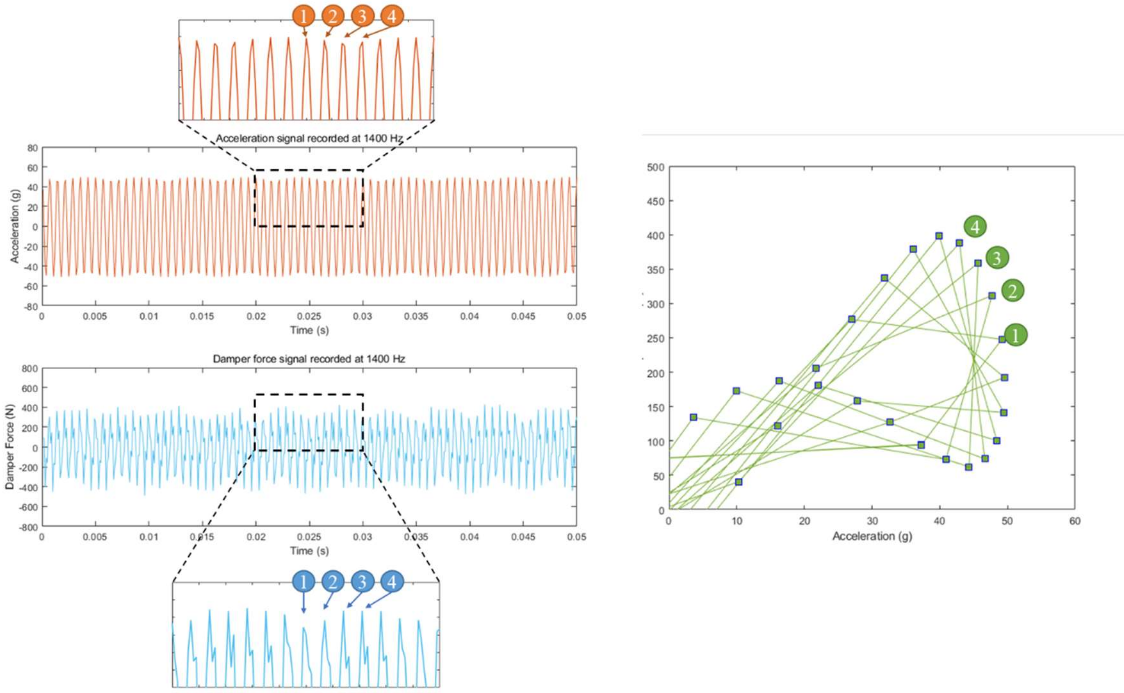


Figure 5-20 Time series of acceleration and damper force of RPD-30 at 1400 Hz.

The existence of this component results in the different peak values of acceleration and force signal at the primary frequency. Take the case of RPD-30 in 1400 Hz as an example. The force signal reaches a peak, and then after one period (consider 1400 Hz as the primary frequency), the next peak value would be lower or higher. On the hysteresis loop, this effect is presented as the separate points that caused the appearance of ears. Connecting the adjacent points on the hysteresis loop with lines can make the forming of the acceleration-force loop more straightforward.

Worthy of mention that the vibration is very high frequency. Moreover, the movement of the particle bed is challenging to describe. This theory only explains the general pattern, which reconditely arises in the complex high-frequency vibration. A deepen and more detailed investigation of this pattern is expected in future research.

5.3.3 Baseline methods

The target of this section is to extend the basic model to the target model, after the establishment of the basic model by *ff*-TLPINN. In this study, the RPD-30 is selected as the basic damper, and the response force models of the RPD-40 and RPD-60 are target model-1 and target model-2. Since the dataset of these three dampers is obtained from three individual experiments, their input features, i.e., excitation signals, must be consistent. In this regard, training the two target models together is impossible because their outputs do not correspond to the same input. For the same reason, the basic damper force F_{d1} input should be generated from the basic model with the target excitation signals rather than the original basic damper force.

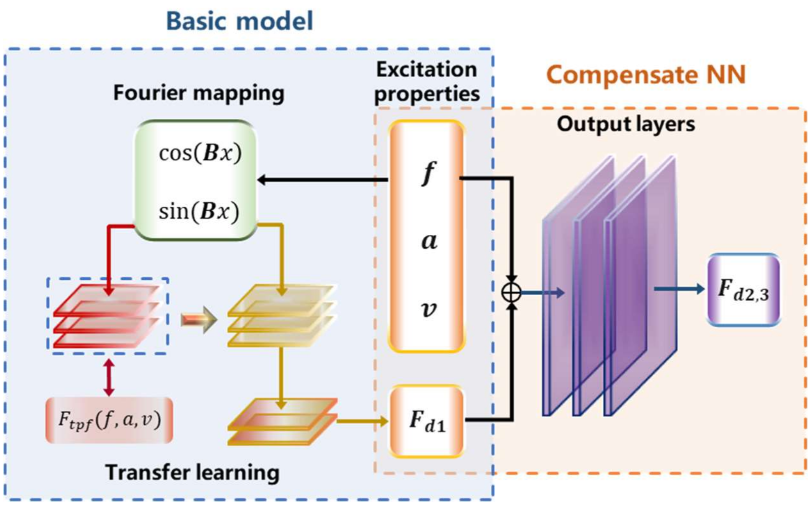


Figure 5-21 Framework of baseline method 1: compensate NN.

The first baseline method is based on the most fundamental NN. A compensated NN directly connects after the basic model, taking the excitation properties and the generated F_{d1} as input. A discussion of the architecture of the compensated NN should be discussed. Since all methods involve deep learning layers, the identified optimum structure will be adopted for all methods for easy comparison.

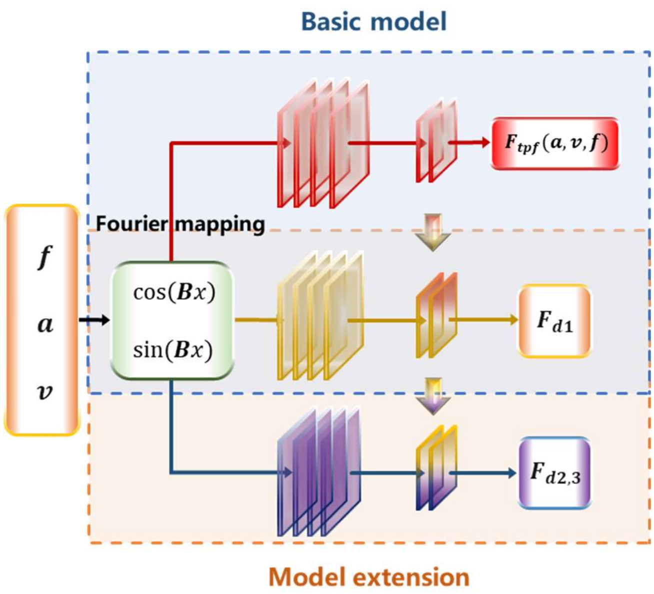


Figure 5-22 Framework of baseline method 2: TL-NN.

Following the discussion in previous chapters, it is also straightforward to adopt transfer learning here. The basic model can be a source model in this two-step transfer learning. The Fourier mapping is applied to every step. However, the number of frozen layers in the second transfer step should differ from the *ff*-TLPINN. Otherwise, the first step of transfer learning would lose its meaning since it equals directly transferring knowledge from the approximate governing equations to the target models. Therefore, the number of free layers in the second step should be less than in the first step transfer.

5.4 Proof of Concept

5.4.1 Selection of the NN architecture

For the choice of network architecture, large networks (in terms of width or depth) hold more sturdy regression ability but also are inevitably more tedious in training. An extensive network has more parameters to update in each iteration. Meanwhile, the enlarged network could bring only limited enhanced performance when a smaller NN is already qualified for the task. Moreover, a solid strategy to guide the network selection is currently absent to the best of the authors' knowledge. Therefore, we test the plain NNs on the dataset. This dataset consists of 20 sets of frequencies (100, 200, …, 2000 Hz); 5 sets of acceleration levels (approximately 10, 20, …, 50 g), a total of 100 slices, each slice of the acceleration-controlled vibration data is a 0.1 s long time series starting at the π phase with 1,000 temporal stamps.

As mentioned above, the input features to this NN include the excitation properties ($\boldsymbol{f}$, $\boldsymbol{a}$, and $\boldsymbol{v}$) and the generated response force of the basic damper $\boldsymbol{F_{d1}}$ under the corresponding excitation. Again, the signal slices of the third excitation level (EL-3) will be extracted as the testing dataset. It makes the testing dataset contains 20 slices. Before the training of the model, all input feature channels are normalized by the min-max method (Henderi et al., 2021).

In the choice of the NN architecture, the number of hidden layers is selected from $\mathcal{F} = \{3,4,5,6,7\}$, and the width of the hidden layer is selected from $C = \{50,100,150\}$. Again, the activation function is tanh. The training and testing MSE after 25,000 iterations for each architecture is given in Figure 5-23.

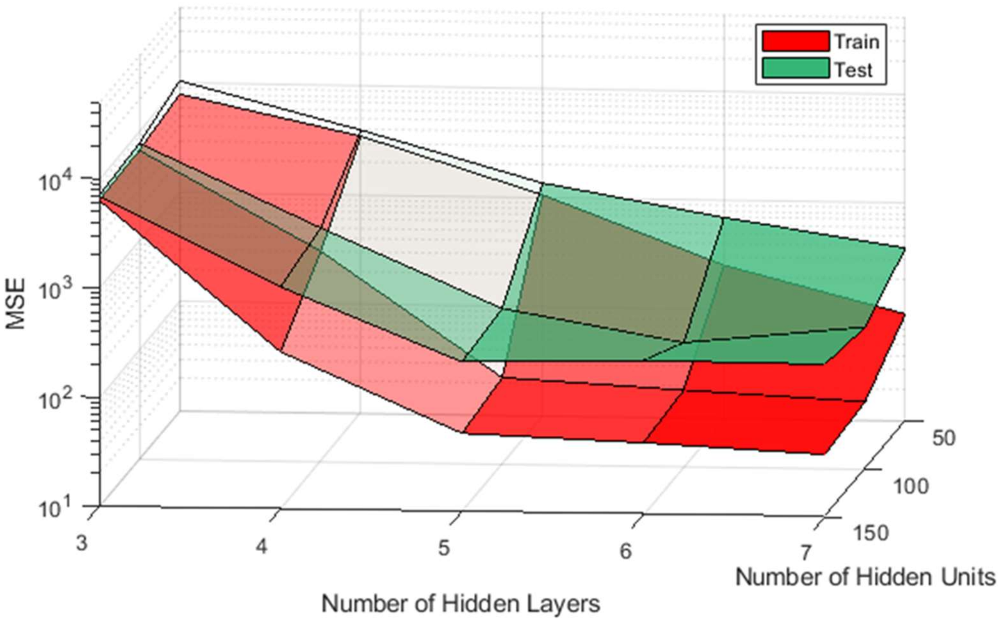


Figure 5-23 Training and testing MSE of each NN architecture after 25,000 iterations.

Table 5-5 Average MSE on the entire dataset of each NN architecture after 25,000 iterations

	$\mathcal{C} = 50$	$\mathcal{C} = 100$	$\mathcal{C} = 150$
$\mathcal{F} = 3$	8460	6848	6468
$\mathcal{F} = 4$	3554	943.2	436.3
$\mathcal{F} = 5$	1123	102.1	88.70
$\mathcal{F} = 6$	342.8	**67.29**	87.36
$\mathcal{F} = 7$	154.7	73.18	79.70

With the increase of the NN size, the training loss does receive a decreasing trend on the training dataset. However, such a trend only sometimes stands in testing, representing the model's generalization ability. It can be found that when the NN is deeper, the rise of NN size will deteriorate the performance. By this means, the structure with $\{\mathcal{F} = 6, \mathcal{C} = 100\}$ will be selected to be applied to all methodologies for comparison, including the basic model. In the basic *ff*-TLPINN, the number of free layers is set as 4, and the initialization standard deviation $\sigma = 10$.

5.4.2 Performance of models

The second baseline method that employs transfer learning was based on the basic model that already conducted one stage of transfer learning from the approximate governing equations to the F_{d1}. Therefore, if in the second stage transferring the free layer $n > 4$, the first stage transferring would be pointless. Hereby in the baseline method 2 (TL-NN), the number of free layers will be discussed in the range of $n \leq 4$.

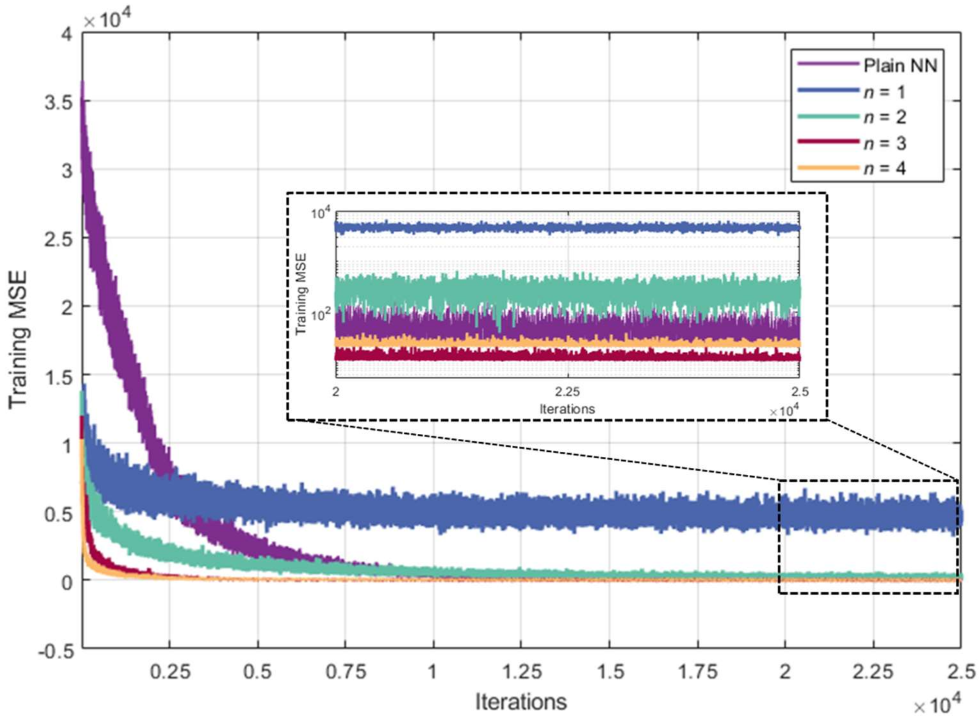


Figure 5-24 Evolution of the training MSE with the different free-layer number $n \in [1,4]$ and the plain NN.

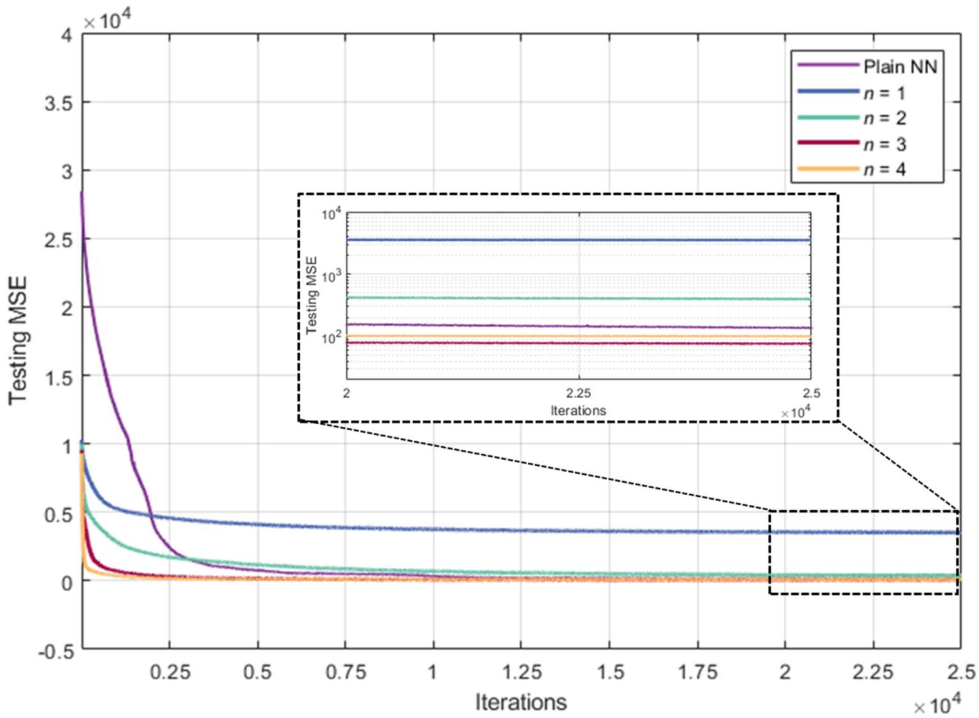


Figure 5-25 Evolution of the testing MSE with the different free-layer number $n \in [1,4]$ and the plain NN.

Similar conclusions as the last chapter can be confirmed here. Firstly, transfer learning elevated the convergence rate compared to plain NN. Partially related to the fact that the MSE started at a lower initial point. Identical to the last chapter, since the size of the dataset is quite large. The shallow free layers showed a deteriorating performance, both in terms of training and testing. As the mini-batch method with a batch size of 1000 is applied here for training, an enlarged deviation can be observed in the training evolution than the testing evolution. Note that when $n = 4$, it is essentially a transfer from the approximate governing equations to the target model, where the first stage transferring is totally passed over. Consequently, we can observe that $n = 3$ is better than $n = 4$, presenting the contribution of the first stage transferring.

After discussing the two baseline models, we now turn to the mc-ESDN based on the ESN structure. Here, the size of the reservoir N_s is set as 1000, with the spectral radius equal to 0.8. No time delay is applied to the input features in this case. As for the MDSC kernels, we selected average kernels in time with lengths of 3, 5, 7, and 9, together with the 1st derivative and 2nd derivative kernels for 1-D kernels of time; Average kernels in space with lengths of 2 and 3, the 1st derivative and 2nd derivative kernels for 1-D kernels of space; For the 2-D kernels, Prewitt operator with up and down directions, and the modified DoG operator are selected.

Afterall, the feature channel numbers expanded from 4 ($\boldsymbol{f}, \boldsymbol{a}, \boldsymbol{v},$ and $\boldsymbol{F_{d1}}$) to 1046 by the above operations. The diversity filter is applied here to choose the channels to be input to the deep readout layer.

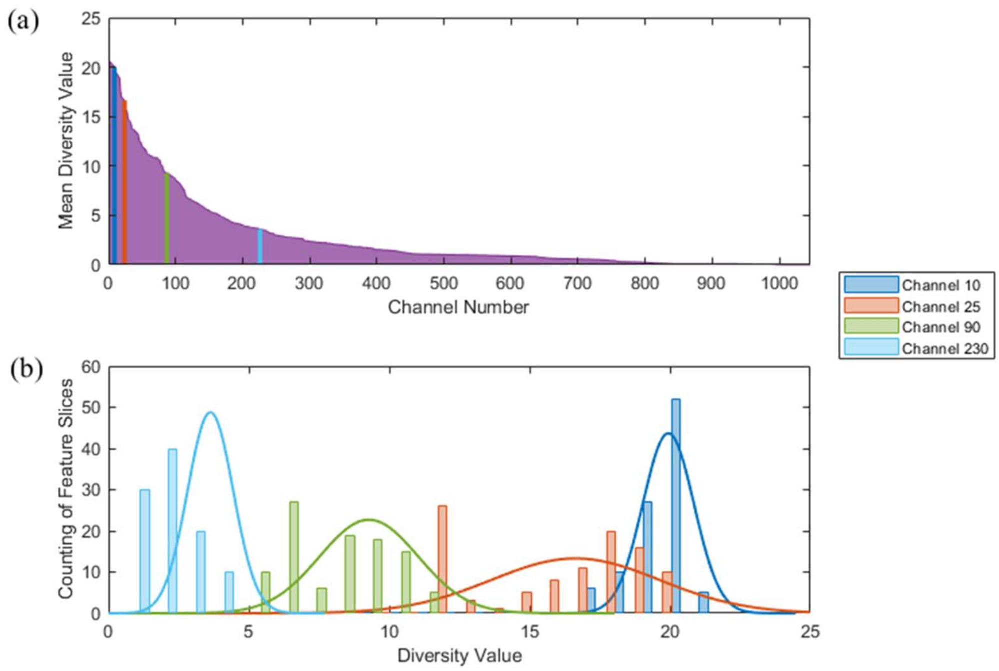


Figure 5-26 (a) Mean diversity values of all 1046 channels; (b) Distributions of the diversity value on some representative channels.

As shown in Figure 5-26, many channels hold a limited diversity value, some even close to zero. Every channel contains 100 slices corresponding to different excitation conditions. The low diversity value means that in these channels, the 100 slices are quite similar to each other. For the deep readout layer, such characterless information harms identifying the features. Figure 5-26(b) shows the distribution of the 100 slices on some representative channels. Worthy of mentioning here that the distribution does not strictly follow the Gaussian distribution. It is only a straightforward presentation of each channel's mean value and standard deviation. Based on the filter result here, the k_v is set as 20 here to select the channels with high diversity value, in which the 4 original channels are mandatorily included.

Based on these settings, the evolution of the training and testing MSE of the mc-ESDN is shown in Figure 5-27. The deviation of the training MSE is also larger than the testing MSE due to the mini-batch method. Finally, at the last iteration, the training MSE equals 3.24, and the testing MSE equals 36.68. Therefore, for the entire dataset, the average MSE is 9.93 for the mc-ESDN.

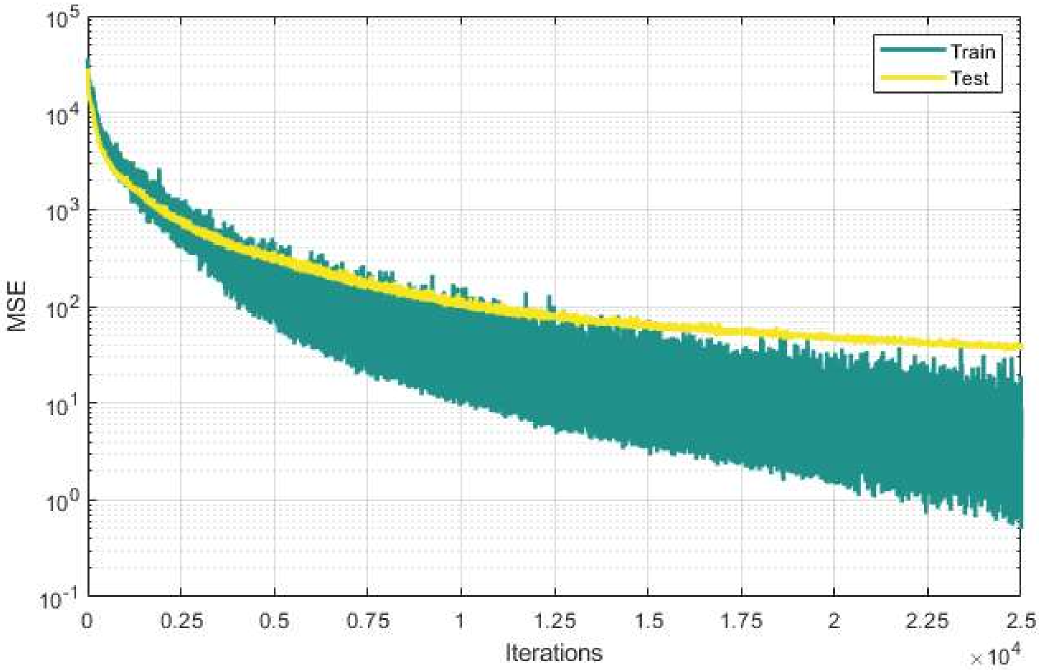


Figure 5-27 Evolution of the training and testing MSE of the mc-ESDN.

For comparison, the training/testing MSE for baseline method 1 and baseline method 2 are 34.74/137.26 and 14.41/62.51, respectively. Their average MSE are 55.20 and 24.03 for these two baseline models. The performance of mc-ESDN is evidently better than the baseline methods on the designated dataset.

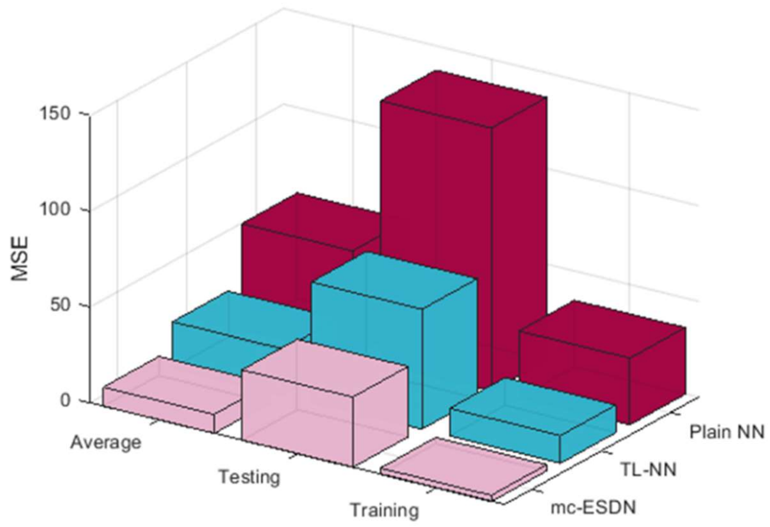


Figure 5-28 Comparison of performance for all three methods.

5.4.3 Model reconstruction for response force of the RPD

The damper response force is reconstructed by implementing the proposed mc-ESDN. Since the datasets of the two target models (RPD-40 and RPD-60) are collected in separate tests, the obtained response force of these two dampers is not corresponding to the same excitation features. Therefore, the deep readout layers must be trained separately for these two models. However, they share the parts in the mc-ESDN structure before the readout layers. The models are established based on the steady-state datasets, with the third excitation level as the testing dataset. After establishing the surrogate model, the sweep-sinusoidal dataset is applied to validate the model's generalization ability on the frequency domain.

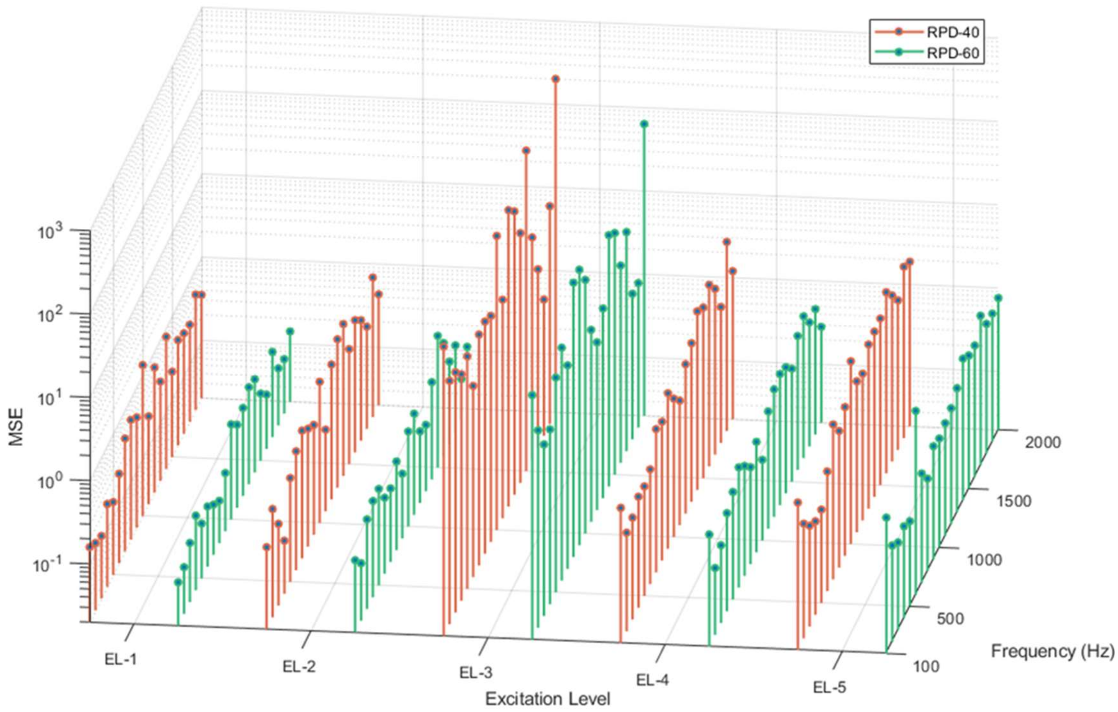


Figure 5-29 Reconstruction MSE of slices for the two target models.

There are 100 slices in the steady-state dataset for both target models. Figure 5-29 presents the reconstruction MSE for all slices of the two target models. Since the slices related to the EL-3 are the testing dataset, the MSE is higher than the other excitation levels. The proposed method manages to reconstruct the hysteresis loops of the target RPDs. Even considering the testing dataset, the recorded reconstruction error is relatively low compared to the absolute value. Therefore, we can find that the predicted hysteresis loops match well with the exact loops obtained from the experiment. It elucidates the vigorous ability of the mc-ESDN to reconstruct the series data.

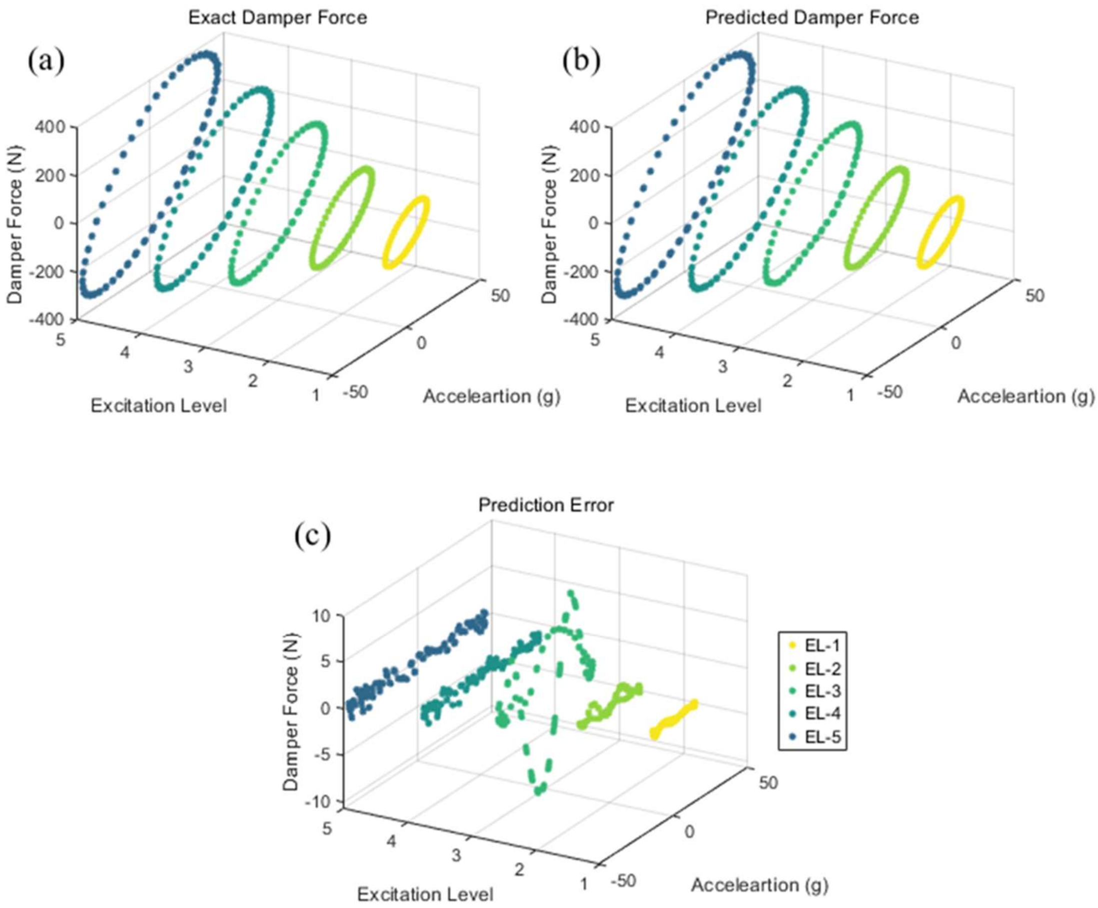


Figure 5-30 Hysteresis loop of RPD-40 on 100 Hz: (a) Exact damper response force recorded in the experiment; (b) The predicted damper response force from the proposed model; (c) The prediction error of the surrogate model.

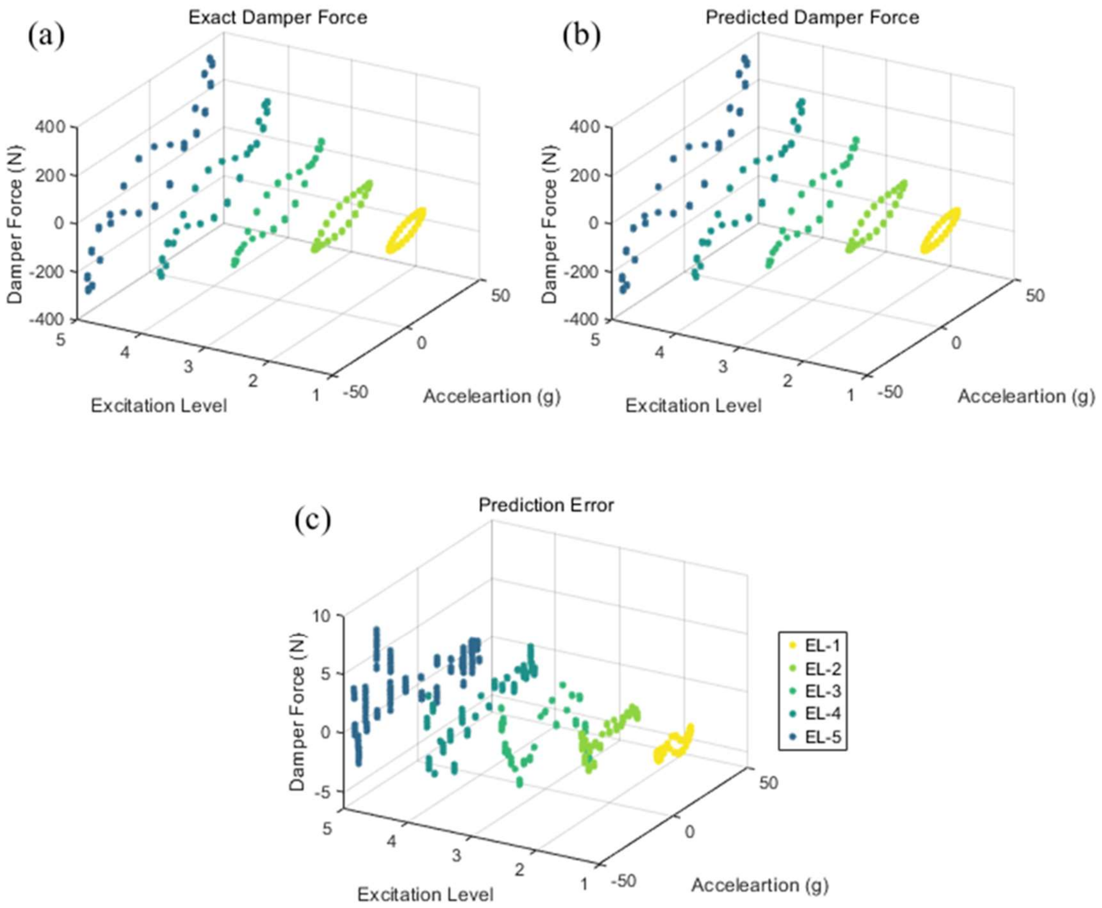


Figure 5-31 Hysteresis loop of RPD-40 on 1600 Hz: (a) Exact damper response force recorded in the experiment; (b) The predicted damper response force from the proposed model; (c) The prediction error of the surrogate model.

From the results, we find the reconstructed hysteresis loops for RPD-40. The MSE for the slice on 1600 Hz, EL-5 is quite large. The prediction error mainly concentrates on the lower part of the hysteresis loop. The reason may be the relatively large force recorded here. As deep learning is a typical interpolation method, the performance may be reduced outer the main range of the dataset. Still, the shape of the reconstructed loop is acceptable.

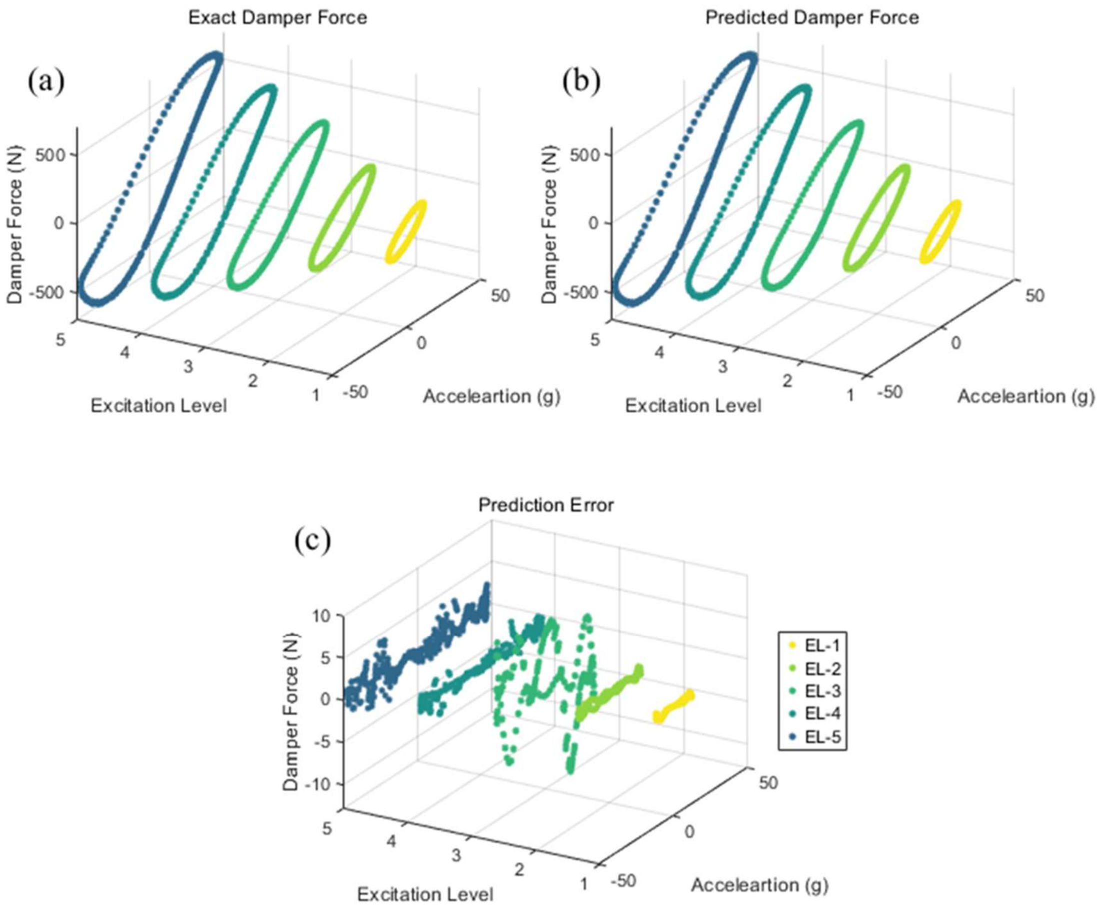


Figure 5-32 Hysteresis loop of RPD-60 on 100 Hz: (a) Exact damper response force recorded in the experiment; (b) The predicted damper response force from the proposed model; (c) The prediction error of the surrogate model.

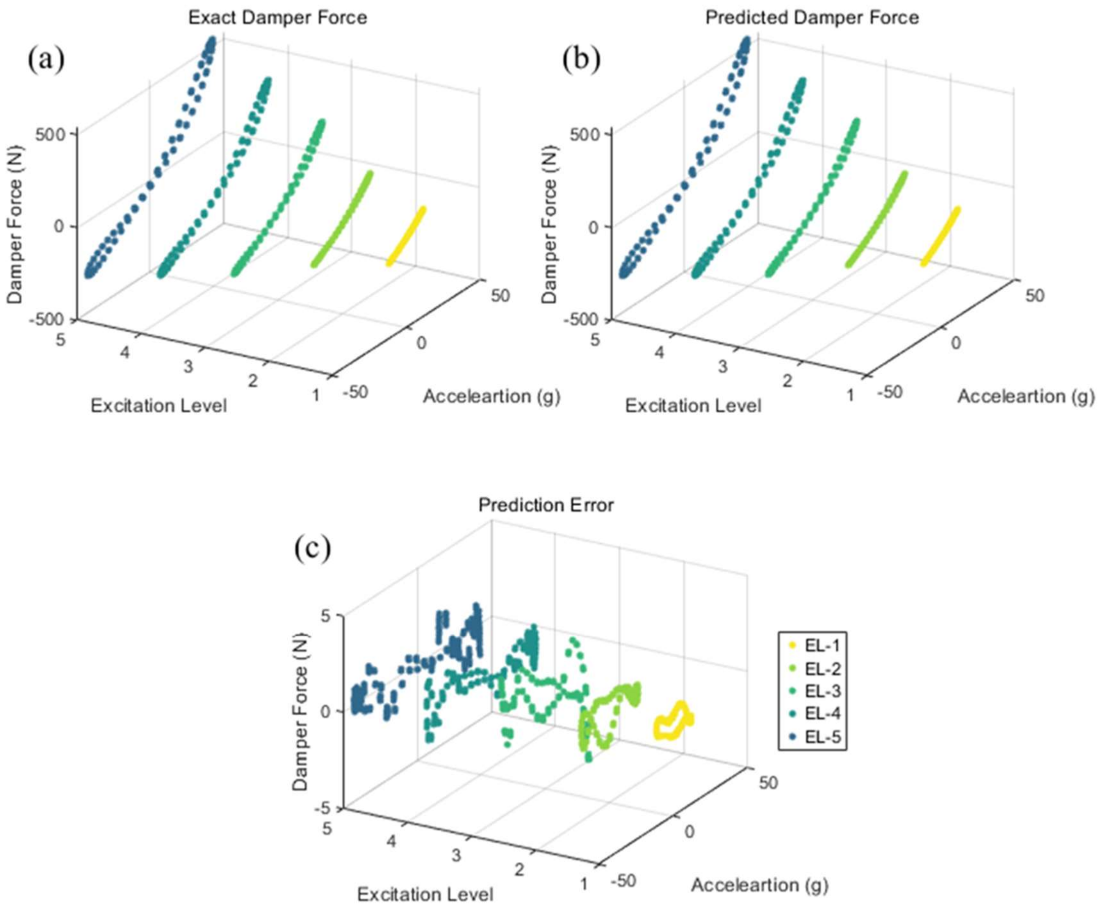


Figure 5-33 Hysteresis loop of RPD-60 on 600 Hz: (a) Exact damper response force recorded in the experiment; (b) The predicted damper response force from the proposed model; (c) The prediction error of the surrogate model.

As for the reconstruction result of RPD-60, the MSE is considerable at 600 Hz. However, no particular high error section appears on the loop shape. The loop shape difference between this frequency and the other frequencies may be the reason for the high MSE.

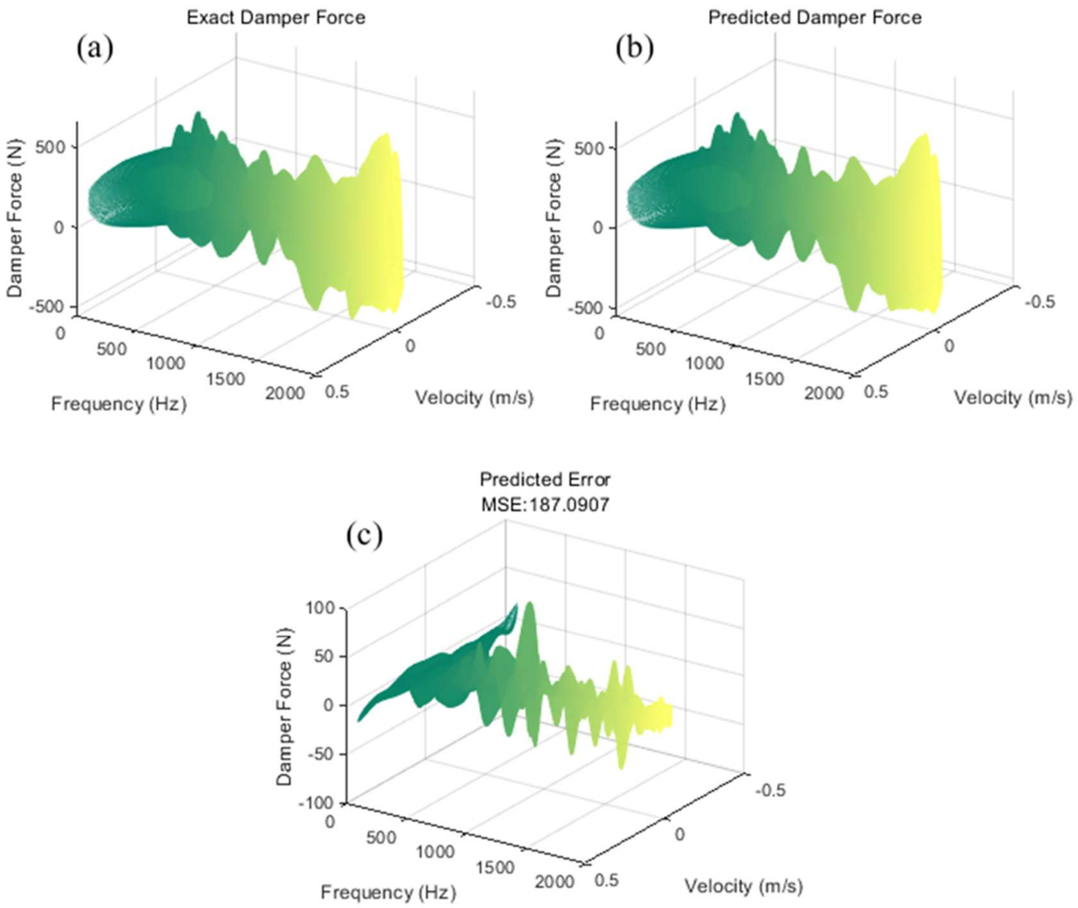


Figure 5-34 Sweep-sinusoidal dataset validation of RPD-40: (a) Exact damper response force; (b) The predicted damper response force from the proposed model; (c) The prediction error of the surrogate model.

Next, the sweep-sinusoidal dataset will be input into the established model for validation. The excitation signals are firstly input to the basic model to generate the corresponding damper force of RPD-30, which is required for the extended model. The error of this reconstruction is two-fold: the error in the reconstruction of F_{d1}, and the error in the reconstruction of $F_{d2,3}$. The observed MSE is significantly larger than the training datasets.

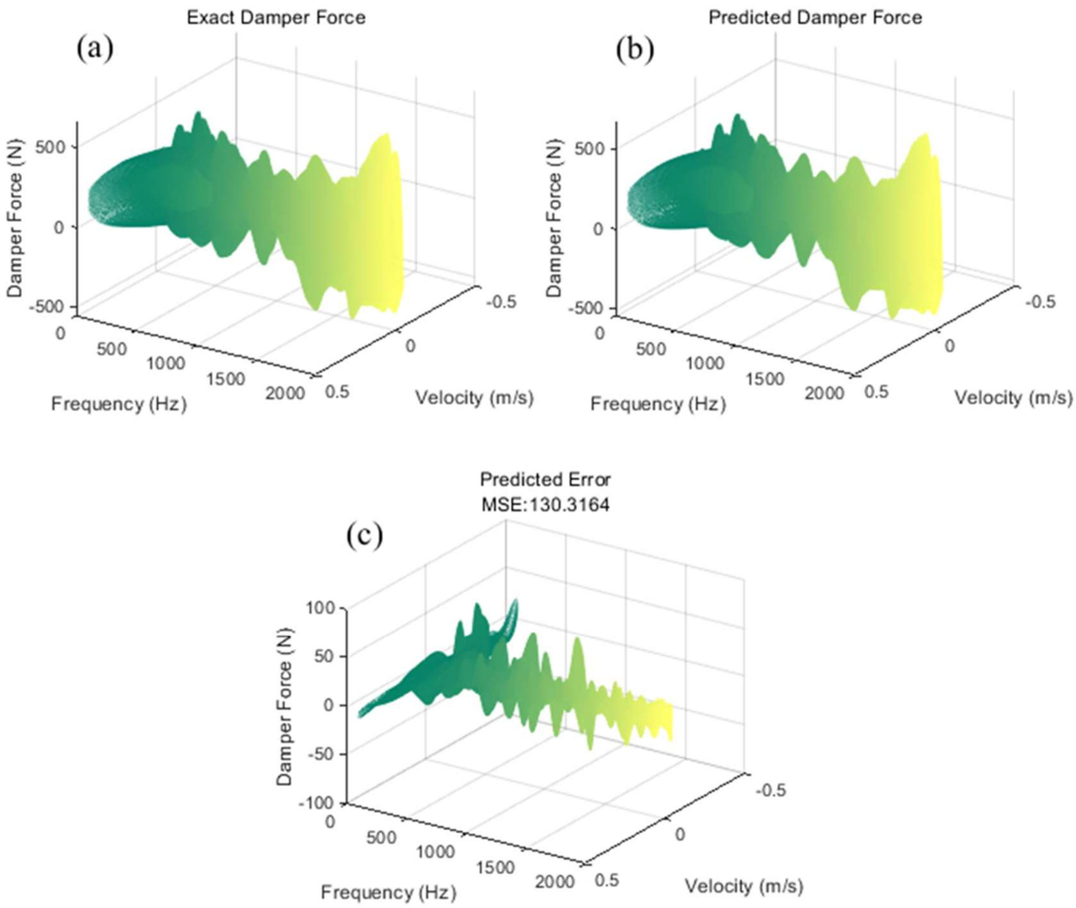


Figure 5-35 Sweep-sinusoidal dataset validation of RPD-40: (a) Exact damper response force; (b) The predicted damper response force from the baseline model 2; (c) The prediction error of the baseline model 2.

However, compared to the proposed model, baseline model 2 with transfer learning achieved a lower prediction error on the validation dataset. It elucidates that the proposed method can strongly reconstruct series data thanks to its RNN nature. However, for a series with different forms, such as sweep-sinusoidal to steady-state harmonic, the generalization ability of the proposed method is limited. Similar findings can be found from the results on RPD-60.

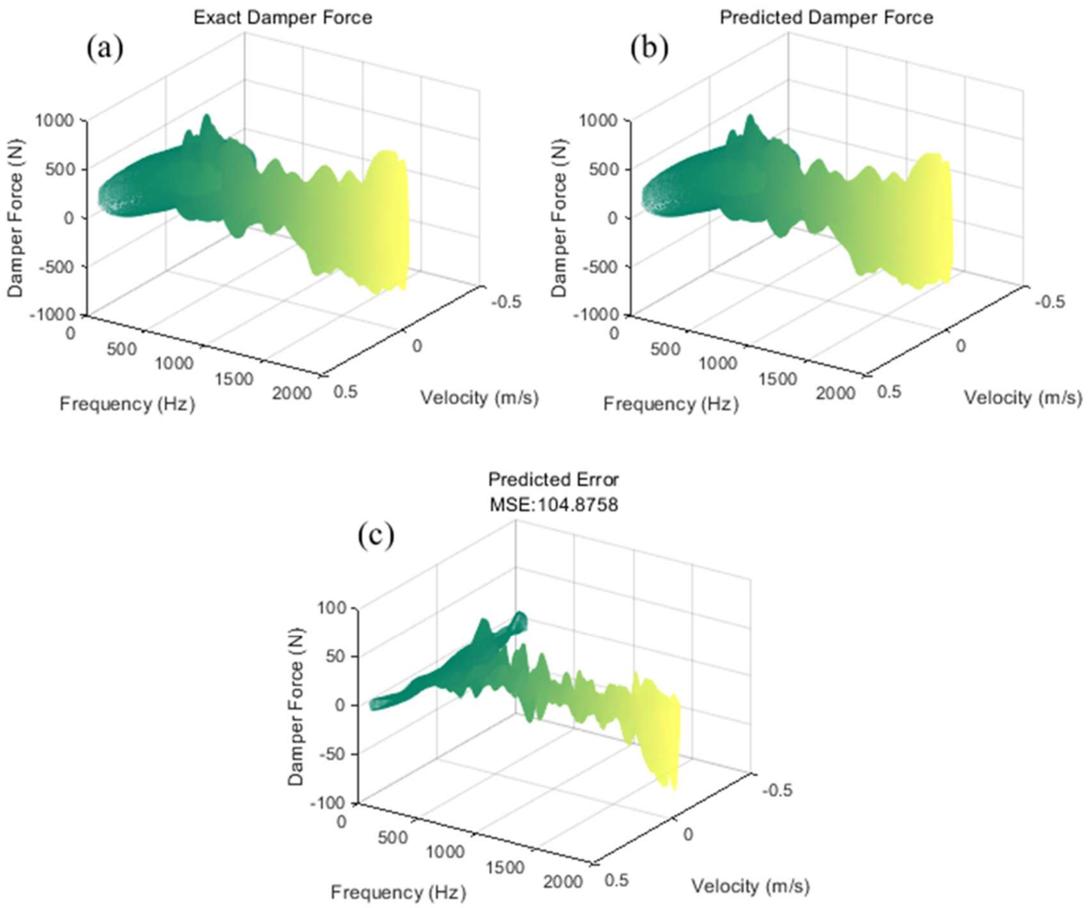


Figure 5-36 Sweep-sinusoidal dataset validation of RPD-60: (a) Exact damper response force; (b) The predicted damper response force from the proposed model; (c) The prediction error of the surrogate model.

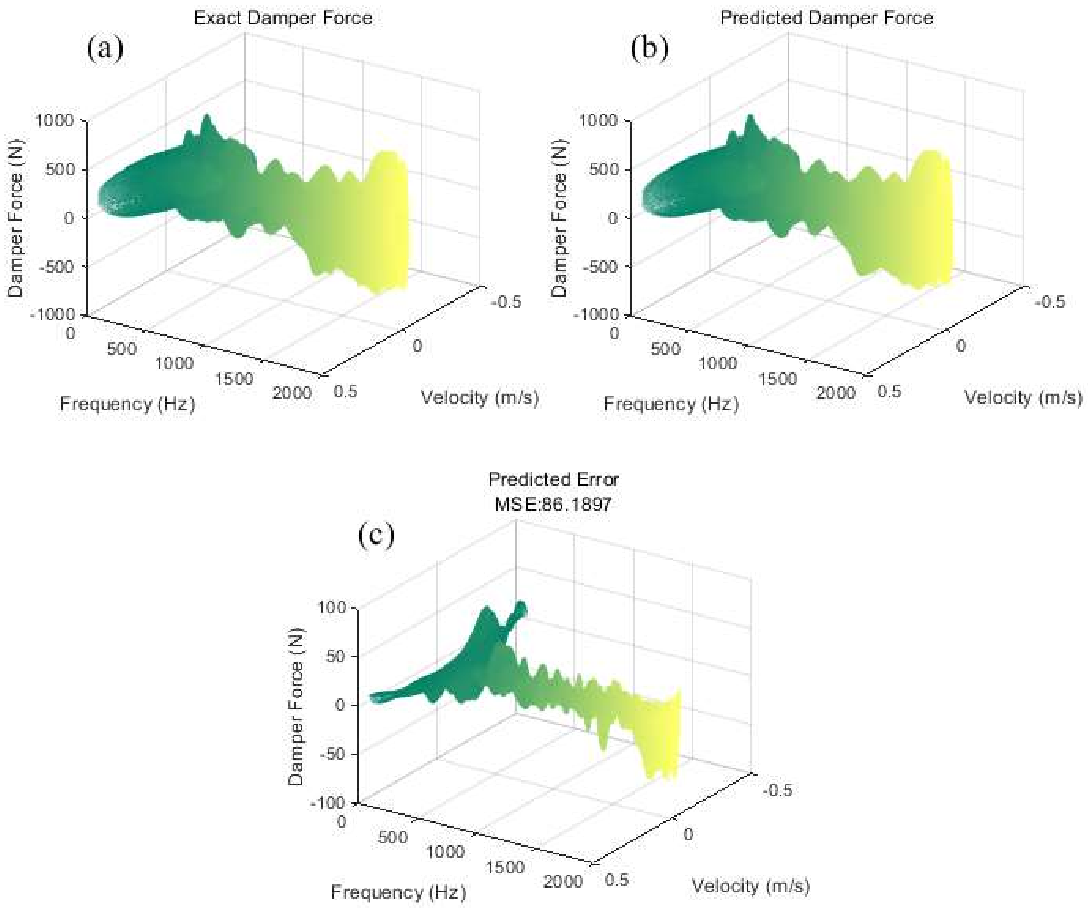


Figure 5-37 Sweep-sinusoidal dataset validation of RPD-60: (a) Exact damper response force; (b) The predicted damper response force from the baseline model 2; (c) The prediction error of the baseline model 2.

5.5 Summary

This chapter exploits the using of ESN to establish the extension model based on the basic model, which is proposed in the last chapter, i.e., *ff*-TLPINN. The standard ESN is synergized by a multi-dimensional and multi-scale convolution feature extractor and a deep readout layer.

Results show that the proposed mc-ESDN can effectively establish the model for multiple damper cases. The conclusions can be summarized as follows:

1. The proposed mc-ESDN is reliable in reconstructing the hysteresis loops of multiple RPDs based on the series data acquired from the dynamic experiments. The training and testing MSE on the designated dataset are 3.24 and 36.68, respectively.

2. The two baseline models that apply the plain NN and TL-NN also accomplished the task of establishing the target model with the same dataset, but their performance is not as good as the proposed method. The training/testing MSE for the two baseline methods are 34.74/137.26 and 14.41/62.51, respectively.

3. The RNN nature of mc-ESDN offers vigorous ability to reconstruct series data. But when confronting the dataset with a different form, the generalization ability of the proposed method is limited compared to the baseline method 2. The mc-ESDN achieves 187.09 MSE on the sweep-sinusoidal dataset, while the later method achieves only 130.32 MSE.

This study investigates the sections of the RPD with different cavity sizes under various excitations. Moreover, the surrogate model of the RPD is established accordingly. From this work, we can have the basic information of the RPD, which will facilitate its application on rail tracks. This work concentrates the RPD filled with 50% 1.5 mm stainless steel balls. For other filling ratios or particle types, the concept presented in this chapter can easily be repeated to model the damper force.

CHAPTER 6.
Evaluation of RPD from Laboratory Test to Field Application

6.1 Introduction

In the CHAPTERS 3, 4 and 5, the understanding of particle damping was built up progressively. The surrogate models of the particle dampers were established under the deep learning framework. The modeling methods were explored from the cylindrical cavity damper to the rail particle damper (RPD) and from the energy loss factor to the damper response aspect. The deep learning methodology provides a vigorous mean to identify the mechanism of particle dampers based on experiments. The established surrogate model circumvents describing the specific motion of the particle bed, which is too intricate for traditional methods to achieve. However, deep learning methods cannot foresee cases that have not been tested. Therefore, it is still needed to conduct experiments to determine the optimum configuration for the rail particle damper.

In this chapter, the effectiveness of RPD was first tested on a single-bay rail track. The vibration suppression ability of RPD was evaluated under various filling materials and ratios. After this step, the RPDs would be further tested on a 6 m rail testbed with a few selected filling particles. The purpose of these experiments was to compare the different filling particles. Then, the modal analysis of the RPD cavity was conducted to reveal its displacement and deformation relative to the rail track. This modal test was conducted on the single-bay rail track. The full range of particle filling ratio would also be explored in this test.

Finally, a field test on a noise-sensitive viaduct section of an operating urban metro line was conducted for a systematic noise and vibration mitigation performance evaluation of the RPDs. The rail track structure forms a periodic structure by implementing the RPDs. The wave propagation on this periodic structure can be analyzed to model the track decay rate of the rail, which is a crucial factor in evaluating the noise generated by the rail track.

6.2 Laboratory Vibration Test of RPD on Rail Track

6.2.1 Single-bay rail impact test

A series of impact hammer tests were meticulously designed and conducted to assess the vibration reduction capabilities of the rail system equipped with RPD. The performance of the RPD is influenced by various properties of the particles within a specific damper cavity, including particle density, friction coefficient, restitution coefficient, and particle size. These properties are determined by the particle material, making them crucial factors in the RPD's performance. To comprehensively investigate the influence of these parameters, a range of materials with different sizes were utilized in the impact hammer experiment. The tested filling substances encompassed powders (nickel and tungsten), liquids (silicone oil), and ball-shaped particles with varying sizes (ceramic and stainless steel). Furthermore, for each material case, the damper cavities were filled at different filling ratios, allowing for a thorough examination of their impact on the RPD's performance. The single-bay rail impact test was conducted by Dr. Seyed Masoud Sajjadi Alehashem and Dr. Jason Lin. The following results and analysis are referring to the internal technical report of CNERC-Rail (Masoud, 2019).

During this test, the RPDs were installed on a single-bay rail system. The rail test specimen consisted of a rail section of CHN60 with a unit weight of 60 kg/m and a length of 0.8 m. The cut section was placed on two support pads with a distance of 0.6 m, mirroring the standard distance of the rail bay in the railway industry. The rail fastening system was based on a spike that is inserted into the concrete sleeper with a screw, which secures the clamp in place.

The accelerometers are mounted at three different points on the rail track to measure the response under the impact excitation. The measuring points include the rail head, rail web, and rail foot. In this test, the impact direction and response measurement focus on the vertical (Z) direction, the primary vibration direction in railway operations. The layout of accelerometers and the defined testing directions are shown in Figure 6-1. The impact point locates 25 mm away from the central point of the midspan of the rail head, as illustrated in Figure 6-2.

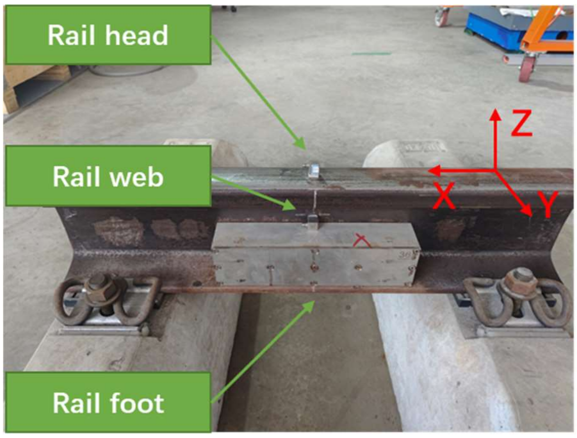


Figure 6-1 The layout of accelerometers.

The hammer used in the test is equipped with a stainless steel impact tip. The choice of material for the impact tip determines the effective frequency band of the impact. Specifically, the amplitude decay of the force spectrum in this frequency band should not exceed 10 dB. Typically, the harder the tip, the wider the effective frequency band will be. In this particular test, the stainless steel impact tip has an effective frequency band of up to 2 kHz. It’s worth noting that the rolling noise emitted by the rail track is concentrated below 2 kHz in the straight rail line.

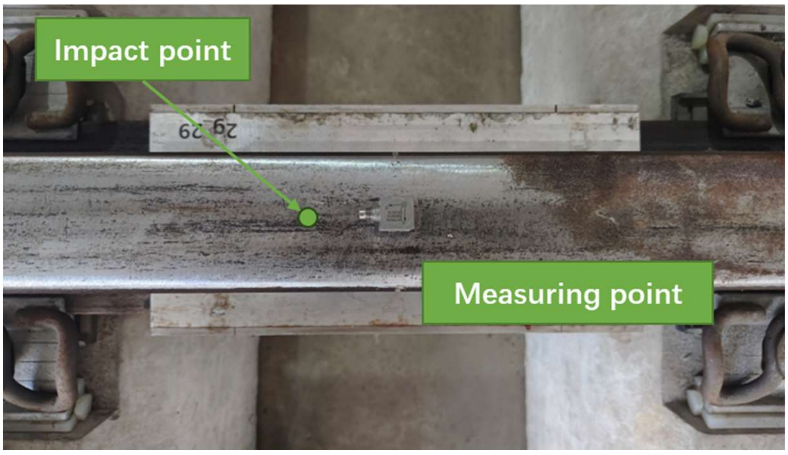


Figure 6-2 Impact and measuring points.

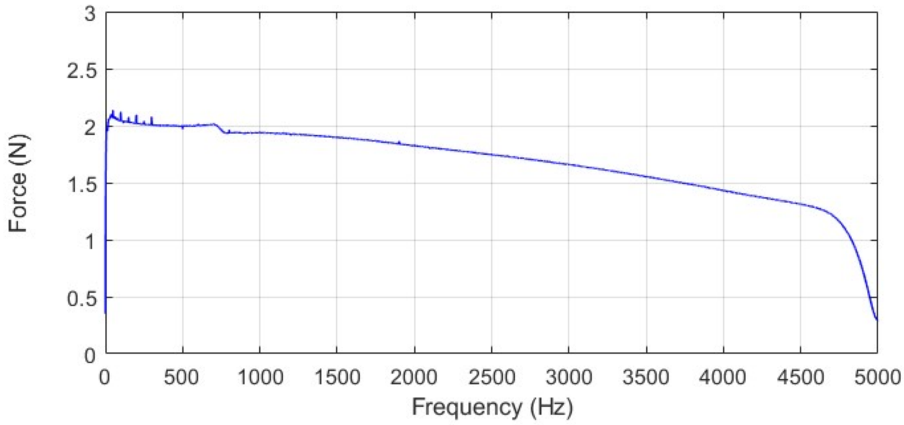


Figure 6-3 Frequency spectrum of impact force with steel tip.

In the research on railway vibration, the rail track system is always regarded as a linear system even though there are various nonlinear components, including rail pads and clips (Wu & Thompson, 2002; Jin et al., 2022). The reason is that the stiffness of the rail track is considerably high, so the vibration amplitude of the rail track is generally within 2 mm. Many of the nonlinear effects will not appear in such a micro-displacement vibration. For example, the particle bed inside the damper cannot impact the cavity ceiling in this scenario. Therefore, the nonlinearity of the particle damper is greatly reduced. In practice, it was also found that the response function of the rail track-damper coupled structure barely remained stable under different amplitudes of impact. Consequently, the system will be regarded as a linear structure in this section, and the frequency response function (FRF) will be evaluated under the same amplitude of impact force level.

Various types of filling particles were tested in this experiment. For each material case, the damper was filled with a certain amount of particles according to the volume of each cavity. The filling ratio was determined according to the apparent density of the particles, i.e., calculated by the weight of a pile of particles divided by the volume of this pile. Then the number of filling particles can be confirmed by measuring the weight of particles, which is more straightforward than measuring the volume, especially for solid particles.

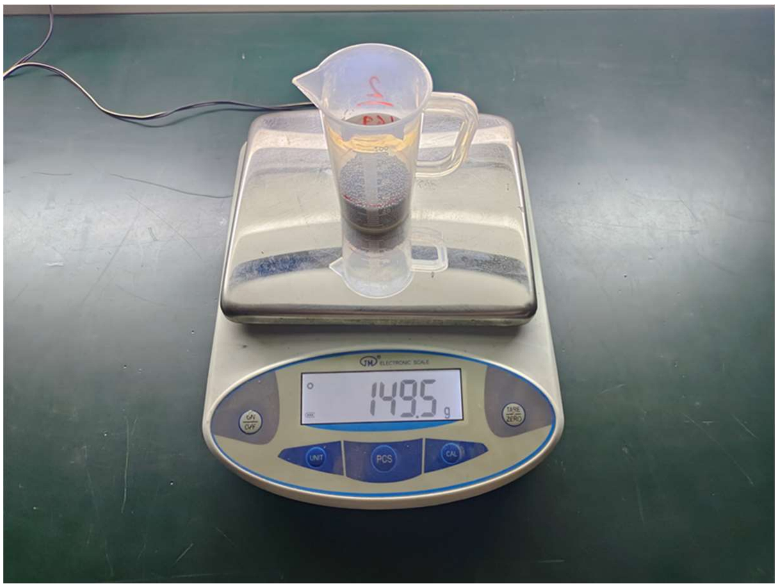


Figure 6-4 Measuring of filling particles.

A high-precision digital scale acquired the weight of filling particles. After filling the dampers with a specific amount of particles, the RPDs were mounted on the two sides of the rail track with the specially designed damper fixture mentioned in CHAPTER 5, shown in Figure 6-5.

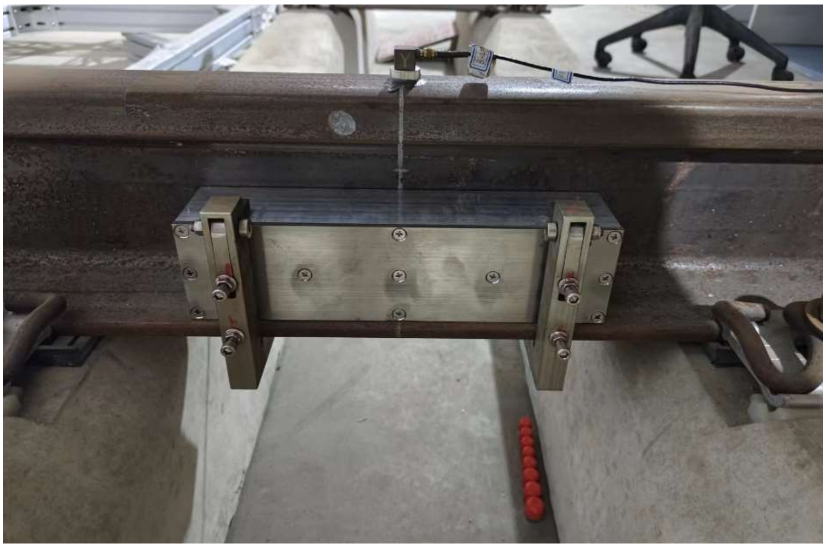


Figure 6-5 The mounted RPDs on the rail track.

As discussed in the CHAPTER 3, an increased filling rate brings larger weight to the damper. However, the energy consumption efficiency is reducing with the increasing filling ratio, and the cost of dampers will also be higher. Futhermore, a large weight particle bed is not likely to be excited, which brings a deteriorated energy consumption performance. Therefore, it is not preferred to set the filling ratio be too large when applying RPDs. The filling ratio of each material will be controlled under 50%. For the liquid case, four levels of filling ratios are set, i.e., 20%, 40%, 60%, and 80%, respectively. The reason is that the density of silicone oil is relatively small compared to the powders and balls, so weighting the 10% filling ratio increment is challenging. This experiment aims to compare the effectiveness of different materials, these results will be compared with the empty damper cases mounted on the rail track. The studied cases are:

- Rail system with empty damper (indicated as **Empty** in comparative graphs).

- Rail system with RPDs filled with **0.2 mm tungsten** powder (10%, 20%, 30%, 40%, and 50% filling ratio with powder).
- Rail system with RPDs filled with **0.2 mm nickel** powder (10%, 20%, 30%, 40%, and 50% filling ratio with powder).
- Rail system with RPDs filled with **silicone oil** (20%, 40%, 60%, and 80% filling ratio with silicone oil).
- Rail system with RPDs filled with **ceramic balls** (10%, 20%, 30%, 40%, and 50% filling ratio with balls). The size of the ceramic balls includes **0.2 mm, 0.4 mm, 0.8 mm, 1.5 mm, and 3.0 mm.**
- Rail system with RPDs filled with **stainless steel balls** (10%, 20%, 30%, 40%, and 50% filling ratio with balls). The size of the stainless steel balls includes **0.5 mm, 0.8 mm, 1.5 mm, 3.0 mm, and 5.0 mm.**

6.2.2 Test results and analysis

In the impact hammer testing, the response measured near the impact location serves as the reference point, which is the rail head in this case. The measurement of other locations can provide the vibration level of the structure. In this section, the response on the rail web and rail foot, where close to the RPDs, is worthy of analysis to obtain the effectiveness of the RPDs.

The FRF responses of the rail system without damper were collected and presented in this section. The vertical FRF responses of the rail system on the rail web and rail foot are shown in Figure 6-6. Two distinct peaks around 500 Hz and 1500 Hz can be identified. These peaks are referring to the first bouncing mode and first bending mode (pin-pin mode) of the rail track (Thompson, 2002). The most significant response occurred at the second peak around 1500 Hz. This peak will be regarded as the target response of this structure system to evaluate the RPD's performance.

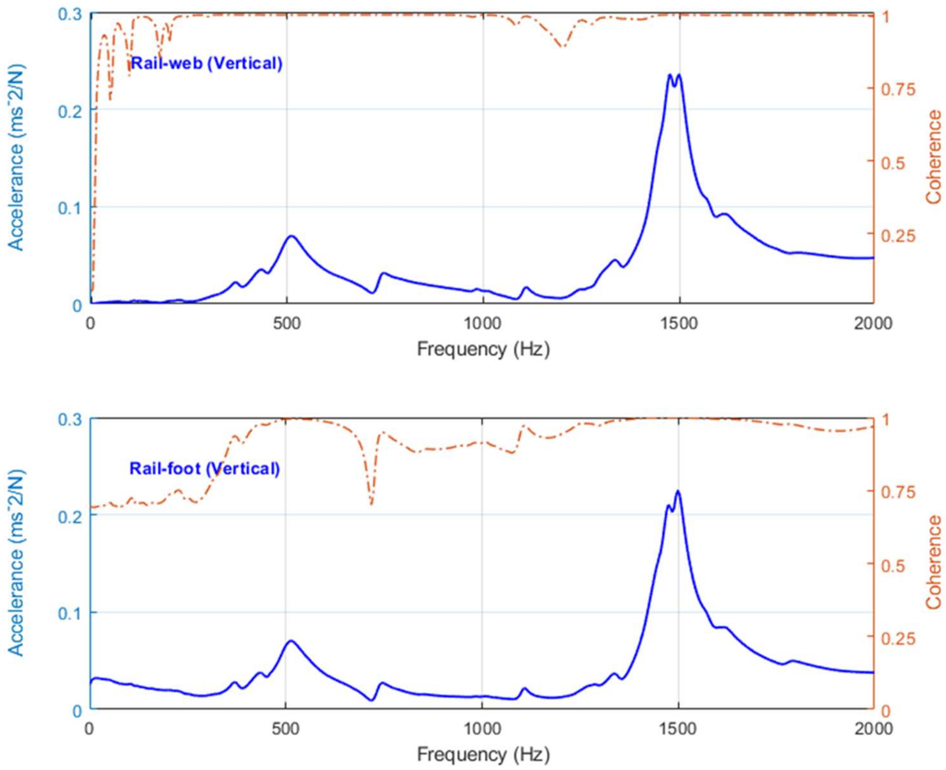


Figure 6-6 The vertical (Z) FRF of the rail system without damper.

The stiffness of the rail section makes the vibration in different locations consistent with each other. As shown in the FRF of rail web and rail foot, the values are almost the same at these two locations, indicating a similar evaluation result on these measuring points. The primary purpose of this test is to compare the vibration suppression ability of RPDs filled with different types of materials. Therefore, the FRF with all kinds of RPDs on the rail web will be presented in the following content for comparison.

The obtained FRF result was averaged from five impacts. During the test, the impact force was controlled within the range from 500 N to 800 N to ensure the input energy to the rail system was stable. Before the test, three different force levels were applied to the rail system equipped with RPD to verify the linear assumption. The RPD filled with 50% 1.5 mm stainless steel ball was mounted on the rail in this trial test. The three excitation levels (ELs) are set as EL-1: 500-800 N, EL-2: 1000-1300 N, and EL-3: 1700-2000 N.

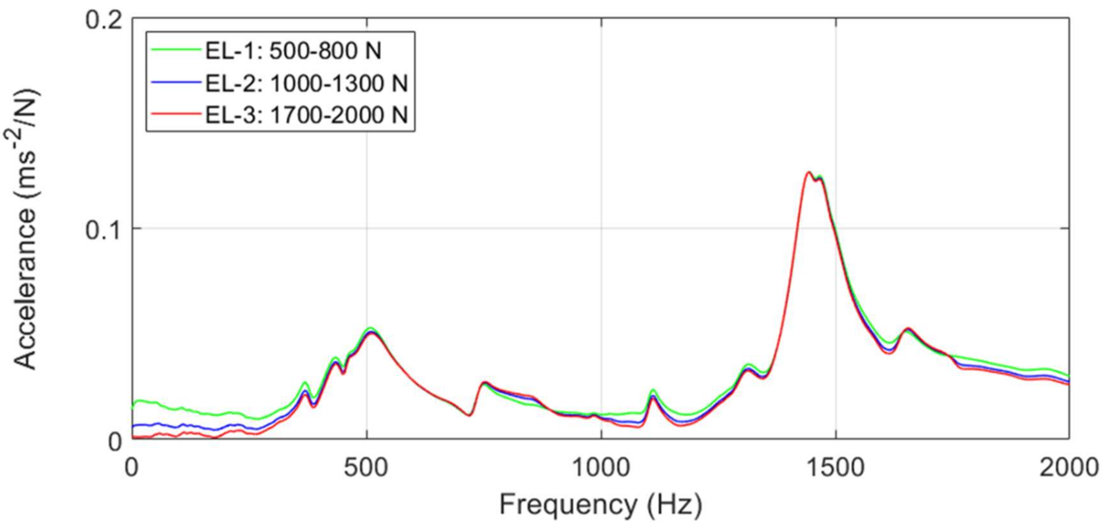


Figure 6-7 The FRF of rail track under various excitation levels.

It can be observed that at both peaks (around 500 Hz and 1500 Hz), the different impact force level brought almost the same response on the structure. Therefore, the rail track system can be regarded as a linear structure. The reason is that the mass of particles (2.2 kg) is not comparable to the mass of rail track system. Therefore, the nonlinear behaviour of particles has limited effect on the system.

Next, the FRF of the rail system equipped with various RPDs that filled with different types of particles and filling ratios will be discussed to evaluate the effectiveness of particles.

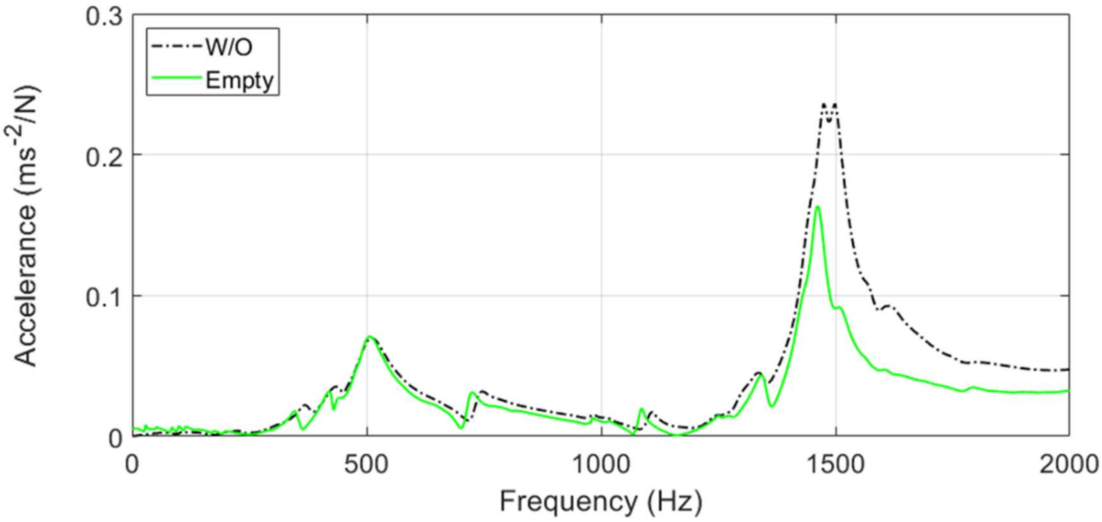


Figure 6-8 The vertical accelerance of the rail system with empty RPD case.

First, the empty RPD cases were attached to the single-bay track. The adding of the empty case changed the rail structure system. However, it can be seen in Figure 6-8 that the empty RPDs barely had any influence on the first peak of the rail track. It did lower the track's pin-pin mode at around 1500 Hz, but the effect is relatively small, which reduced the peak by approximately 25%.

The tungsten powder used in this test is an ultra-fine powder with a particle size of 0.2 mm. A photograph of the RPD filled with tungsten powder is shown in Figure 6-9.

Figure 6-9 Tungsten powder.

The vertical accelerance of the rail system with RPD filled with different ratios of tungsten powder is shown in Figure 6-10. It can be seen that the accelerence at two peaks effectively decreased proportionally with the filling ratio. Also, a slight frequency shift has occurred because of the increment of the rail system's mass after the filling of particles.

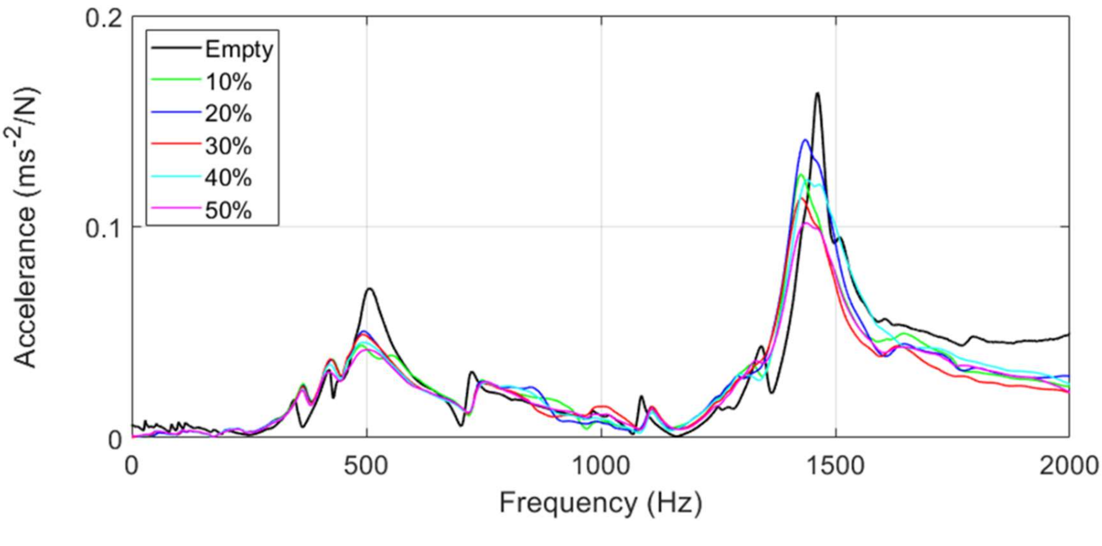


Figure 6-10 The vertical accelerance of the rail system with RPD filled with tungsten powder.

The nickel powder used in this test is also an ultra-fine powder with the same diameter of 0.2 mm. A photograph of the RPD filled with nickel powder is shown in Figure 6-11.

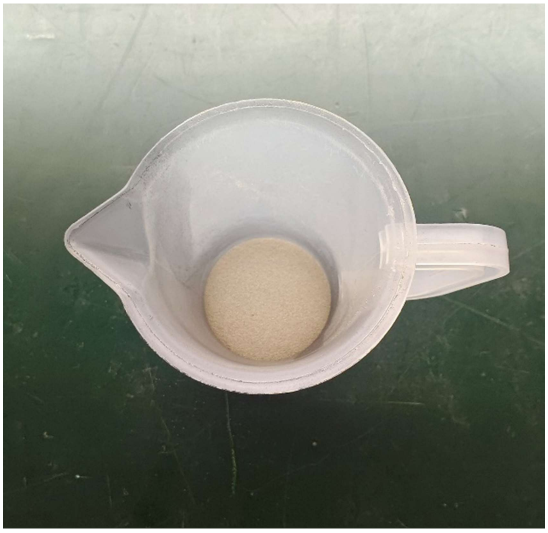


Figure 6-11 Nickel powder.

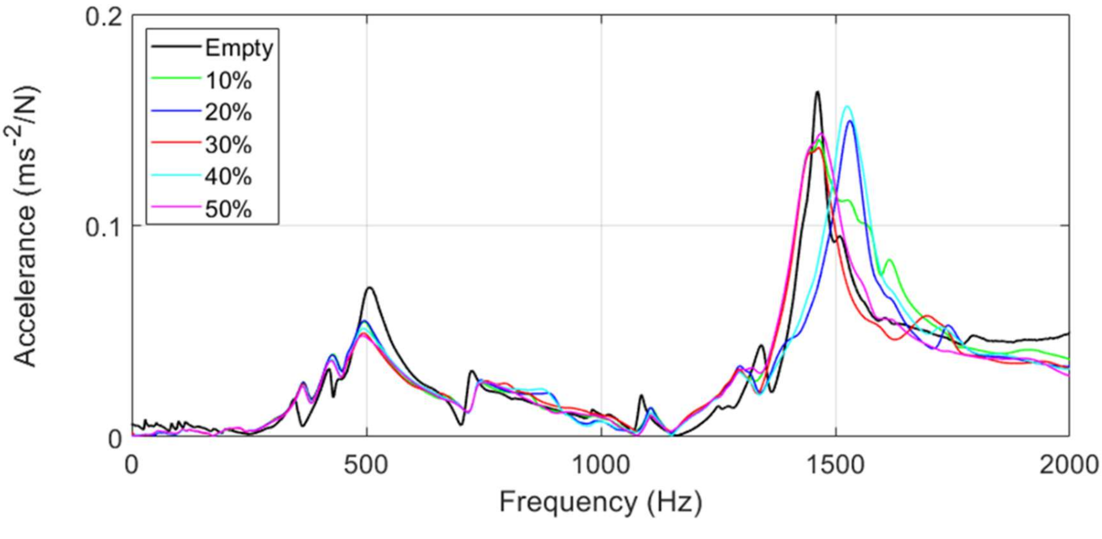


Figure 6-12 The vertical accelerance of the rail system with RPD filled with nickel powder.

The vertical accelerance of the rail system with RPD filled with different ratios of nickel powder is shown in Figure 6-12. The response at the second peak was reduced but not significant. Furthermore, the reduction relationship between the filling ratio is not clear. It can already be concluded based on these results that nickel powder is not preferred compared to tungsten powder to fill the RPDs.

Silicone oil was selected as a filling substance and was poured into the damper cavity to investigate the effects of the liquid substance on the performance of the RPD. The silicone oil is chemically stable and has no effect or reaction with the damper body. Also, the silicone oil is cost-effective. The used silicone oil is XIAMETER, model PMX-200. This kind of oil has a kinematic viscosity of 10 cSt. A photograph of the RPD filled with silicone oil is shown in Figure 6-13.

Figure 6-13 Silicone oil.

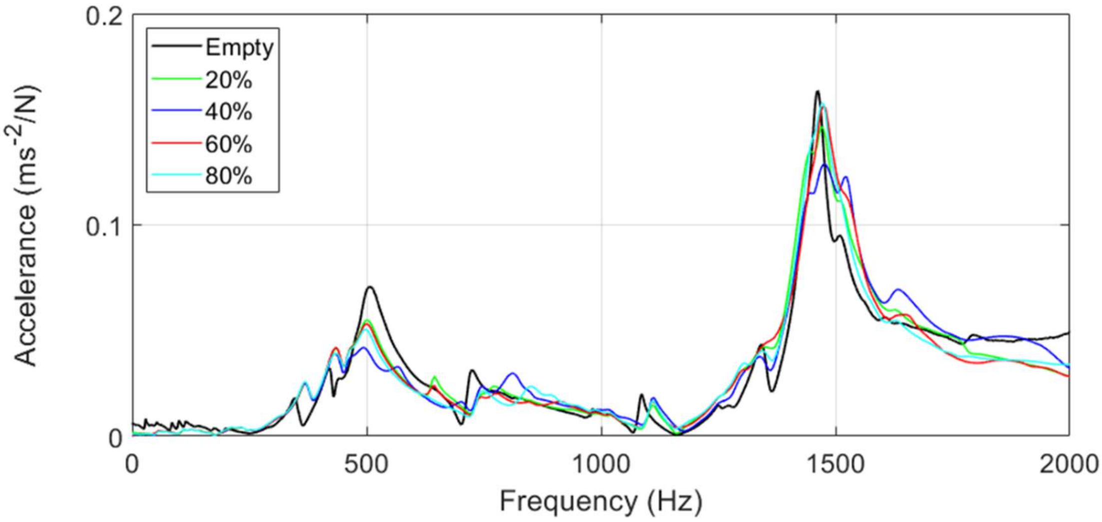


Figure 6-14 The vertical accelerance of the rail system with RPD filled with silicone oil.

To prevent any leakage between the cavities. A gasket with a thickness of 2 mm was applied to seal the cap of the damper. The vertical accelerance of the rail system with RPD filled with different ratios of silicone oil is shown in Figure 6-14. The response decreased at the target frequency. However, the vibration mitigation degree is low and not clearly related to the filling ratios. In practice, if the liquid substance does not significantly enhance the performance, it is usually not preferred due to the chance of leakage and the accompanying increase in the maintenance cost.

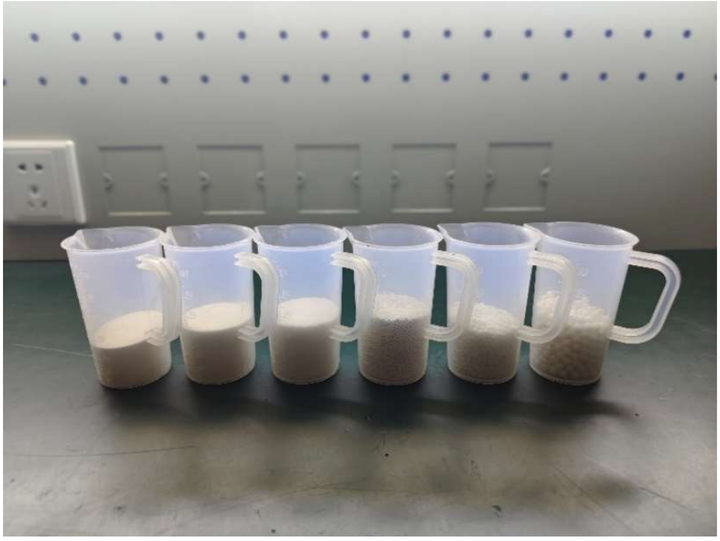


Figure 6-15 Ceramic balls.

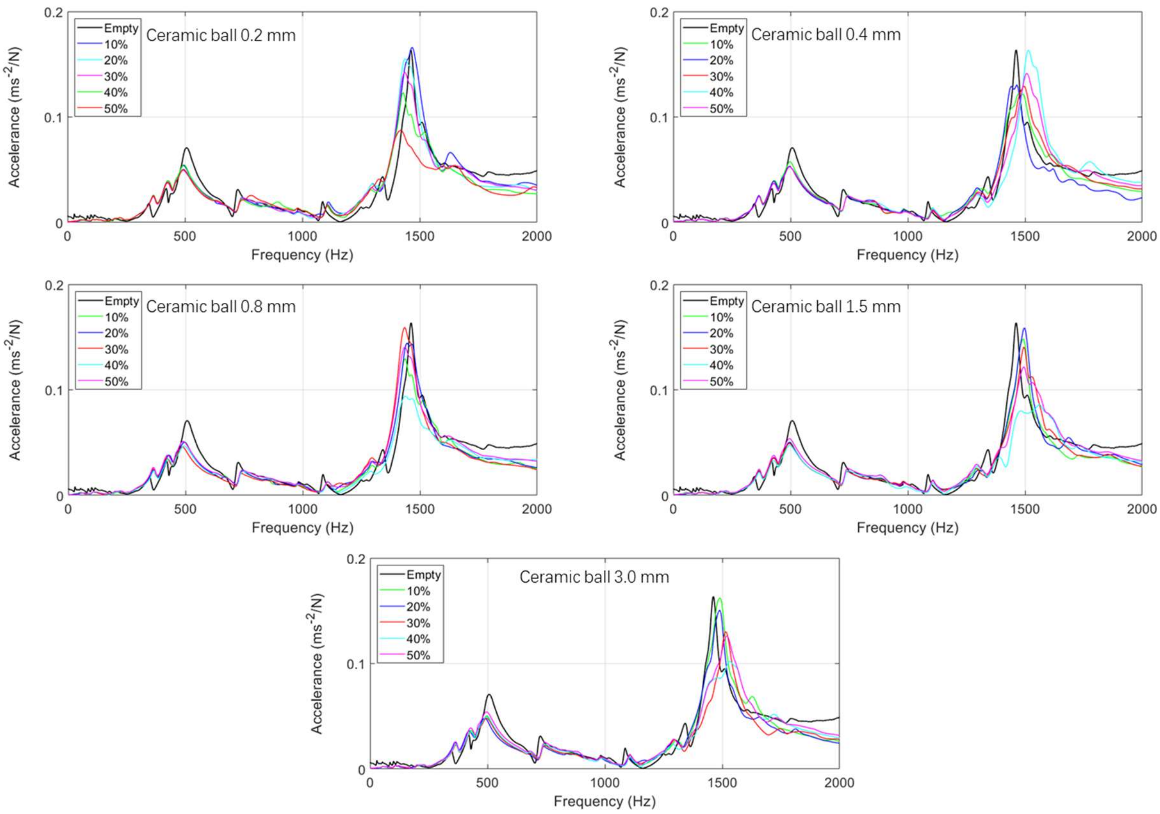


Figure 6-16 The vertical accelerance of the rail system with RPD filled with different sizes of ceramic balls.

Five sizes of ceramic balls (0.2, 0.4, 0.8, 1.5, and 3.0 mm) were investigated. Ceramic is a lightweight material. Its apparent density is 5.68 g/cm^3. It is found that some typical size of ceramic balls exhibited good performance in reducing the vibration of rail tracks. For instance, 0.2 mm and 1.5 mm ceramic balls. The influence of the ball size is two-fold. Thanks to the decreased clearance inside the particle bed, more particles can be filled under the same filling ratio with a smaller ball. On the other hand, the reduced clearance increases the contact area between particles, making the friction effect more significant. The influence of the smaller ball size on the impact effect is hard to determine. It weaks the effect of every single impact between particles since the weight of the ball is reduced. However, the amount of the impact occurred increased. It is difficult to conclude how the impact effect would affect the damper performance.

However, the experimental results do not present a straightforward trend regarding the ball size. In general, the 0.2 mm and 1.5 mm ceramic balls performed well, and this two kinds of ceramic balls would be selected for further exploration.

Figure 6-17 Stainless steel balls.

The FRF responses of the rail system with RPD filled with stainless steel balls are presented here. Since the stainless steel balls smaller than 0.5 mm are uncommon in the market. So that a larger ball size with diameter equal to 5.0 mm is added in this case. Again, five sizes of stainless steel balls (0.5, 0.8, 1.5, 3.0, and 5.0 mm) were investigated.

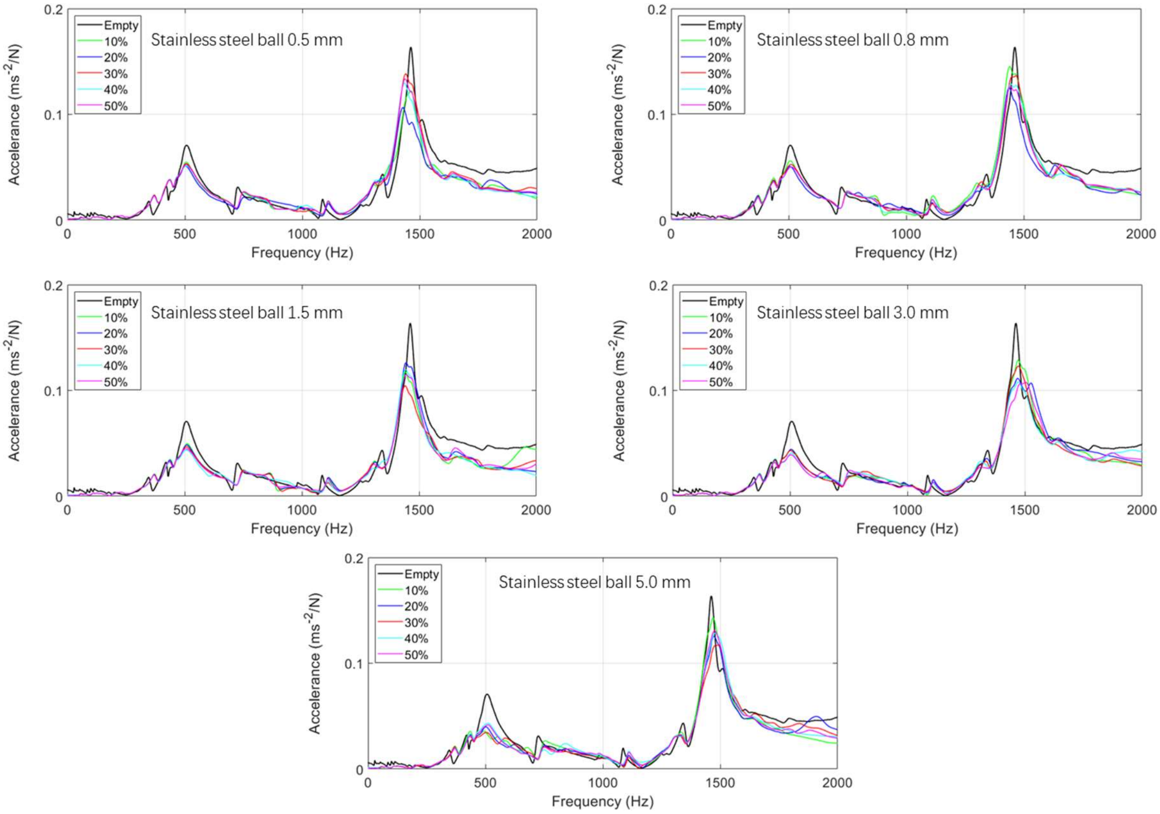


Figure 6-18 The vertical accelerance of the rail system with RPD filled with different sizes of stainless steel balls.

The stainless steel balls also show a good performance in vibration reduction. Its density (7.5-8.0 g/cm^3) is larger than the ceramic, meaning that the weight of this kind of RPD is larger. The larger mass makes the stainless steel balls less likely to be excited and impact each other. Consequently, the friction effect plays a major role in this kind of damping.

In conclusion, the energy dissipates for the liquid is based on the laminar shear of the liquid's molecules. Therefore, if the excitation amplitude is not big enough in the vertical direction, the liquid cannot effectively contribute to energy dissipation because of the fluid's weight and molecular bonding. Another energy dissipation mechanism in the liquid is fluid splashing, which only happens on the surface of the fluid which is insignificant in this case.

Energy dissipates through the friction and impact effects for powders and solid balls (ceramic and stainless steel). In this scenario, the amplitude of vibration in the vertical direction is small due to the stiffness of the rail system. The impact effect is significantly reduced in such micro-displacement vibration. Therefore, the friction effect plays a major role in dissipating energy.

Table 6-1 Summary of the best performance of each experiment case in single bay rail track

Substance	Material	Size (mm)	Apparent Density (g/cm3)	Optimum Filling Ratio (%)	Average Peak Accelerance Reduction around 1500 Hz (%)
Powder	Tungsten	0.1	8.90	50	32
	Nickel	0.1	5.15	50	17
Liquid	Silicone oil	$\nu = 10cSt$	0.90	40	21

Solid ball	Ceramic	**0.2**	**3.82**	**30**	**46**
		0.4	3.82	30	23
		0.8	3.80	40	42
		1.5	**3.80**	**40**	**43**
		3.0	3.78	40	32
	Stainless steel	0.5	4.99	30	36
		0.8	4.96	40	33
		1.5	**4.96**	**40**	**40**
		3.0	4.88	40	34
		5.0	4.74	40	28

By referring to Table 6-1, the performance of balls is generally better than powders and liquids. As mentioned above, liquid is not suitable for field applications compared to solid materials. Two types of the ceramic balls (0.2 mm and 1.5 mm) and one type of the stainless steel balls (1.5 mm) will be further tested in the next subsection to confirm the optimum filling substance for the RPDs.

6.2.3 Vibration test on 6 m full-scale rail testbed

After the impact hammer test on the single bay rail, three types of particles have been identified to hold the pronouncing ability in mitigating the rail track vibration, namely, 0.2 mm ceramic balls, 1.5 mm ceramic balls, and 1.5 mm stainless steel balls. These types of particles would be further evaluated on a full-scale rail testbed. The testbed contains a 6 m CHN60 rail track section, placed on 10 sleepers equipped with rail fastening systems. The distance between two sleepers is the same as in the railway industry, i.e., 0.6 m. 9 pairs of RPDs will be mounted on the rail track at every rail bay. This experiment was conducted by Dr. Masoud and me.

The B&K LDS V-650 shaker was hung on a gantry to apply dynamic force to the rail track. The shaker was the same type used in CHAPTER 3, its basic parameters are listed in Table 6-2. The shaker head was connected to the track with a connecting rod to transmit the force. The KD3001B impedance head was placed on top of the rod to record the excitation force on the rail.

Table 6-2 Specifications of the dynamic shaker

Parameter name	Parameter value
Sine force peak	2.2 kN
Max acceleration sine peak	50 g
Usable frequency range	5 Hz – 4 kHz

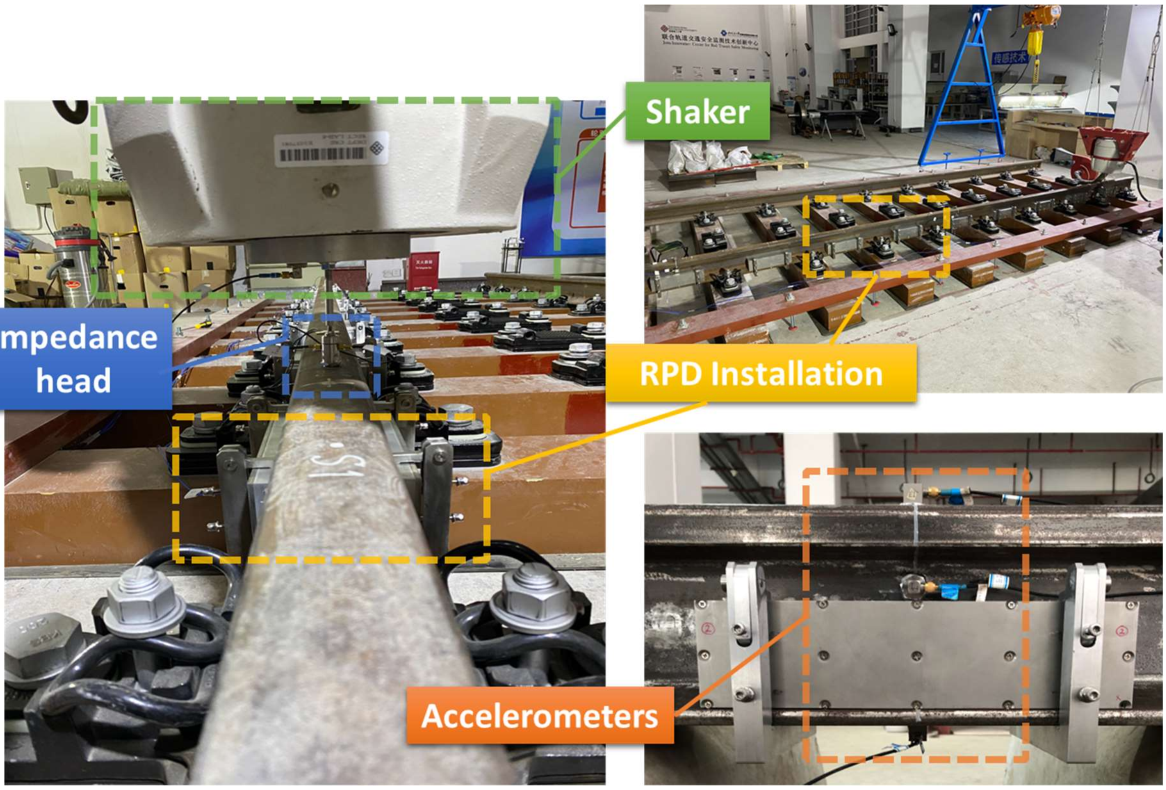


Figure 6-19 Setup of the rail testbed.

The 10 sleepers were named S1 to S10, and the rail bays between sleepers were named B1 to B10, respectively. The dynamic shaker was attached to the rail head on B1. 9 accelerometers were mounted on the rail head, rail web, and rail foot on B1, B5, and B9. The layout of the sensors and the defined testing directions are shown in Figure 6-20.

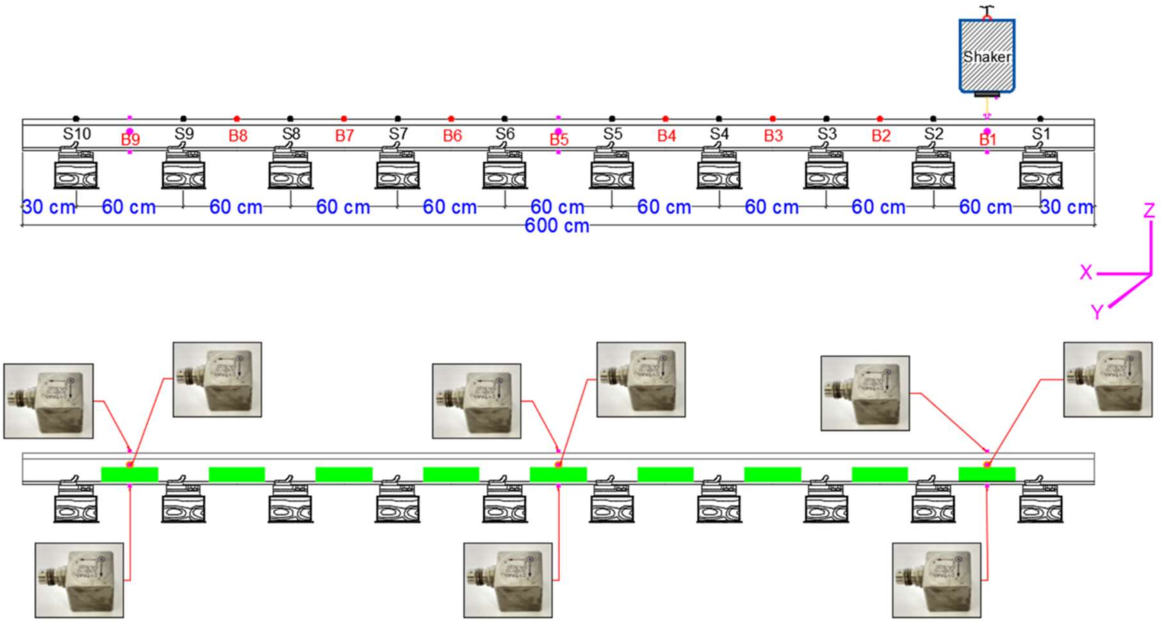


Figure 6-20 Layout of the sensors and the rail track (Masoud).

Two forms of excitations will be applied to the testbed in this test. The dynamic shaker is able to apply sinusoidal sweep excitation to the structure and obtain the structural response on the entire frequency band in the vertical (Z) direction. Similarly, the impact hammer could simulate impulse excitation to the structure. For comparison, the impact location was also on the rail head at B1.

This test aims to find the most effective particle material for mitigating rail track vibration. The particle types mentioned above would be filled in the RPD and evaluated on this testbed. Again, the investigated range of particle filling ratio was set within 50% under the same reason in the section 6.2.1. Furthermore, to investigate the influence of the empty RPD. The empty RPDs without any particle filled in the cavity were also mounted on the testbed to evaluate the vibration of the rail track. Therefore, the studied cases in this test include:

- Rail testbed with empty RPDs (indicated as **Empty** in comparative graphs).
- Rail testbed with RPDs filled with **ceramic balls** (10%, 20%, 30%, 40%, and 50% filling ratio with balls). The size of the ceramic balls includes **0.2 mm, and 1.5 mm.**
- Rail testbed with RPDs filled with **stainless steel balls** (10%, 20%, 30%, 40%, and 50% filling ratio with balls). The size of the stainless steel balls is **1.5 mm.**

6.2.4 Test results and analysis

This test recorded and compared the responses of the 6 m rail track equipped with different RPDs. The sensors were mounted on three bays of the rail track, namely, B1, B5, and B9. The B1 is also the location where excitation was applied to the track. The FRF on the rail head and rail foot of each case will be presented in this subsection.

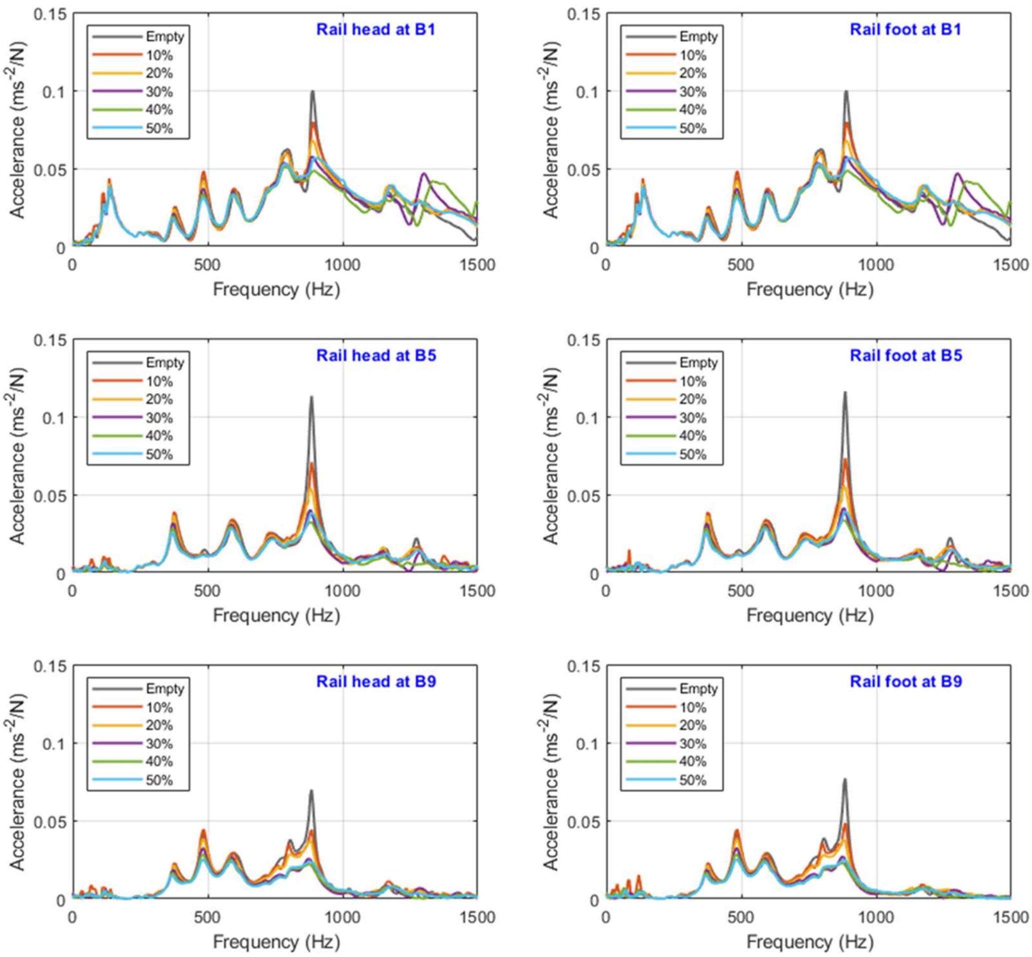


Figure 6-21 The vertical accelerance of the 6 m rail track on different locations with RPD filled with 0.2 mm ceramic balls under sweep sinusoidal excitation.

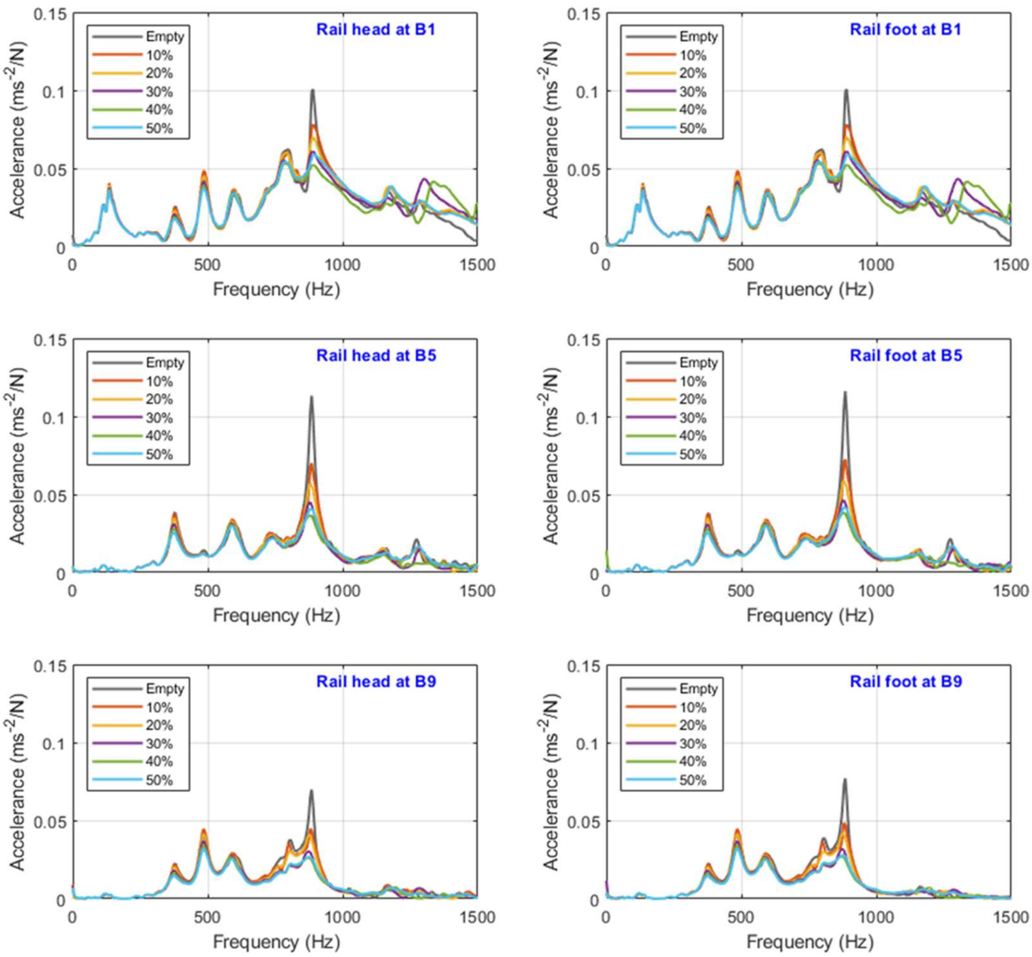


Figure 6-22 The vertical accelerance of the 6 m rail track on different locations with RPD filled with 0.2 mm ceramic balls under impact hammer excitation.

For the results with 0.2 mm ceramic balls, the rail head and rail foot responses are consistent, especially on the main peaks of the FRF, which are in the range of 0 to 1500 Hz. It indicates the results are reliable as the rail section presents uniformity during the vibration.

Considering the response of bare rail track without mounting any dampers, more peaks on FRF were excited at B1 compared to the signal-bay rail test since the 6 m rail is a more complex structure that closer to the actual application. However, many of these peaks did not appear at B5; only the major modes related to the entire structure were observed on the middle bay. When it came to the end of the track, i.e., B9, the vibration energy decayed due to the damping of the rail system, including the clips and pads and the rail track itself.

The responses of the rail system under shaker and impacte hammer excitations are barely equal to each other. Since sweep sinusoidal excitation is quite a different form of input compared to impact impulse excitation in many aspects, this finding is another proof of the linear assumption in this circumstance. It is equivalent to presenting the FRF of the rail system under any form of excitation. In the following contents, the FRFs of the rail system with the impact hammer test will be presented since the impact hammer is the most common choice in the field tests of railway system.

More dampers were installed on the testbed compared to the single bay, increasing the total weight of the auxiliary weight on the rail system. However, adding particles into the damper did not move the peak on the frequency spectrum. This finding indicates that the particle bed did not perform as a lumped mass but as an energy dissipator in the dampers. Generally, a larger filling ratio leads to enhanced performance in mitigating vibration on the peak around 800 Hz.

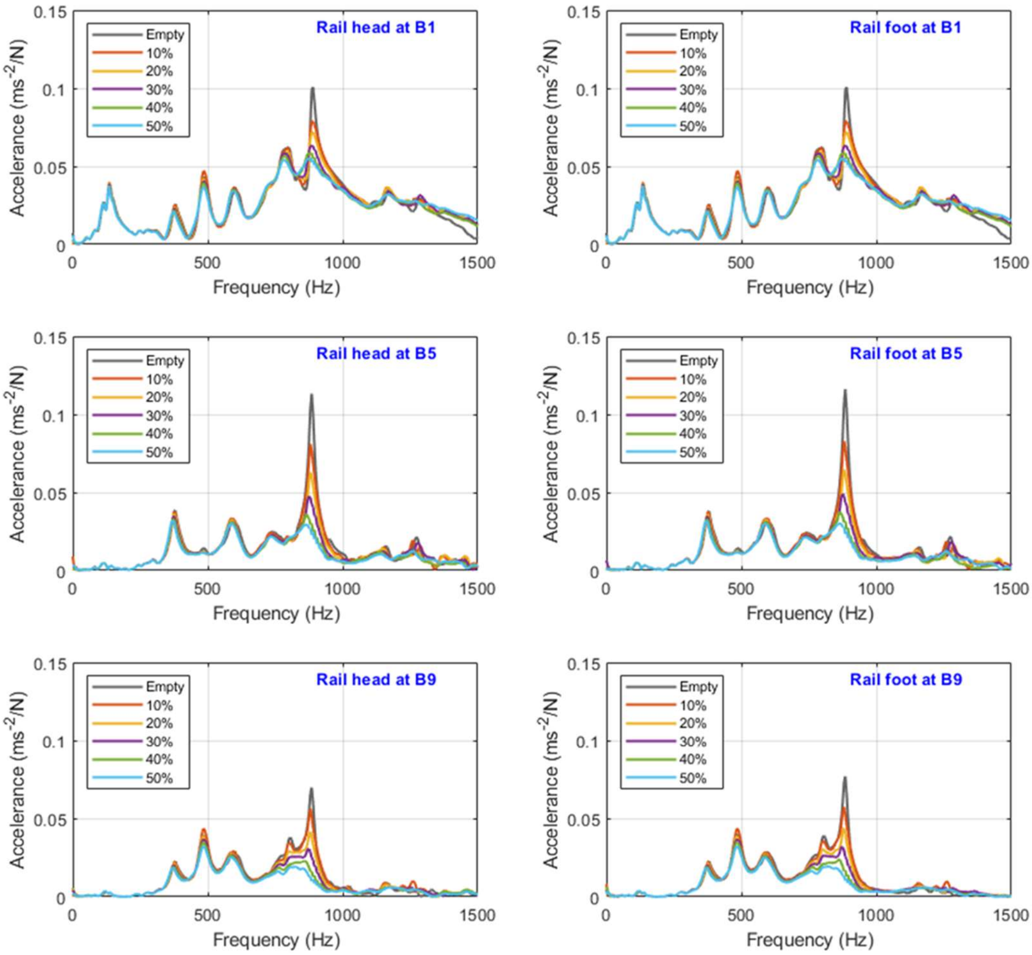


Figure 6-23 The vertical accelerance of the 6 m rail track on different locations with RPD filled with 1.5 mm ceramic balls under impact hammer excitation.

It is observed that there was not much difference between the results of 0.2 mm and 1.5 mm ceramic balls. Indicating the size of the particles did not play an important role in the performance of the RPDs in this case.

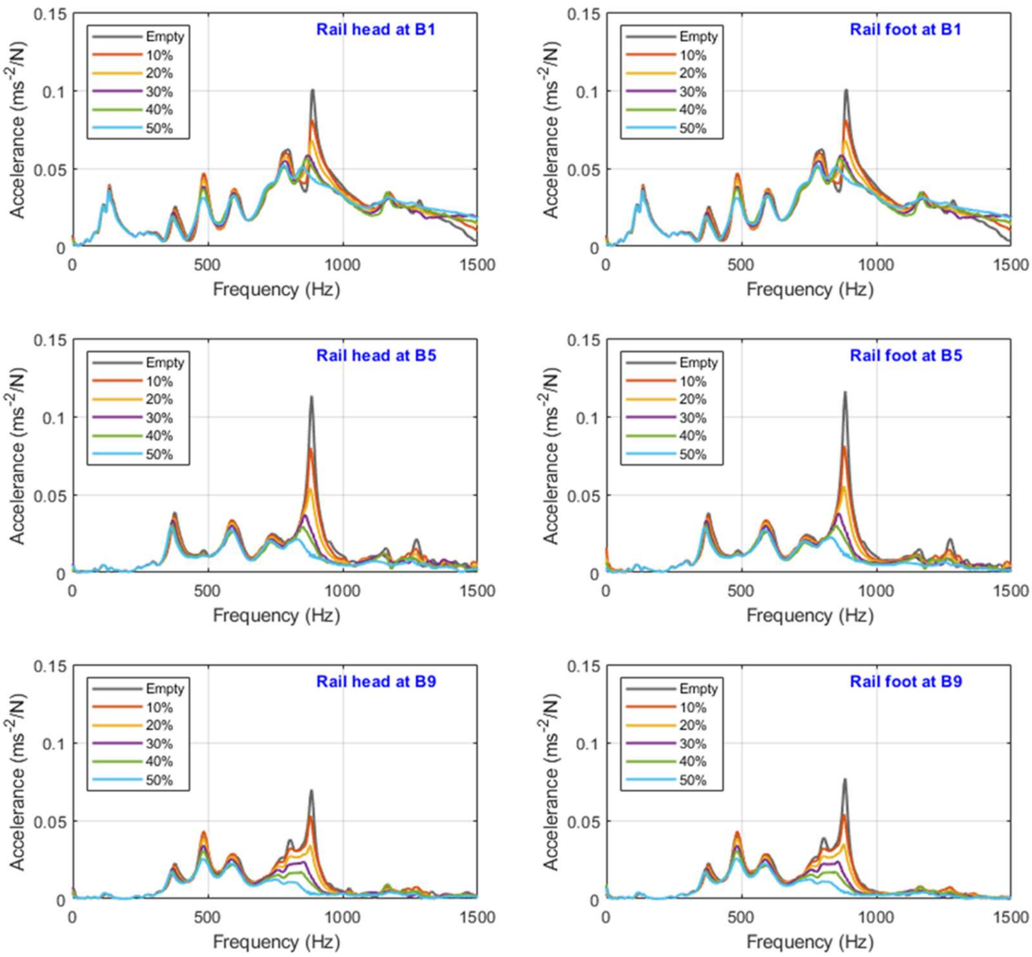


Figure 6-24 The vertical accelerance of the 6 m rail track on different locations with RPD filled with 1.5 mm stainless steel balls under impact hammer excitation.

Clear trend of the filling ratios can be found with the cases of 1.5 mm stainless steel balls, which were the heaviest particles in this test. The RPDs filled with 50% of the 1.5 mm stainless steel balls reduced this peak to a considerable low level. It indicates that this kind of particle is a suitable choice for mitigating rail vibration.

In conclusion, the weight of particles is heavier for the 1.5 mm stainless steel balls.

Compared to ceramic balls, the mass of the stainless-steel particle bed played a more significant role in affecting the vibration of the rail system. The influence of mass for the ceramic balls was limited, and it was found that a 40% filling ratio achieves the best performance with this type of particle. The best performance of each type of particle is presented in Table 6-3.

Table 6-3 Summary of the best performance of each experiment case on 6 m rail testbed

Material	Size (mm)	Apparent Density (g/cm^3)	Pin-pin Accelerance Reduction at B1 (%)	Pin-pin Accelerance Reduction at B5 (%)	Pin-pin Accelerance Reduction at B9 (%)
Ceramic	0.2	3.82	48	65	64
	1.5	3.80	45	72	76
Stainless steel	1.5	4.96	50	80	85

After a series of laboratory experiments, it can be concluded that the 1.5 mm stainless steel balls are the most suitable type of particles to be applied to mitigate the rail track's vibration among all tested filling substances. In the field application, a large amount of RPDs will be fabricated and transported. As discussed above, To demonstratively elucidate the effectiveness of RPDs in the railway industry, a 50% filling ratio of 1.5 mm stainless steel balls will be applied to control the cost. Since this configuration already presented a good performance in controlling the pin-pin mode of the rail track. A futher study on this filling particles considering the filling ratio from 0 to 90% is presented in the next subsection.

6.2.5 Modal analysis of the RPD on rail track

The 1.5 mm stainless steel ball was identified to be a suitable material as the filling particles in the RPD. To evaluate the effect of installing RPDs on the track, the RPD as an accessory component to the rail track has to be investigated. To that end, the modal analysis of the RPD, when it is attached to the track, is conducted here. The motivation is to understand the vibration mode of the damper on the track during vibration, and the modal damping ratio is obtained to reflect the energy dissipation by the damper, including the fixture system.

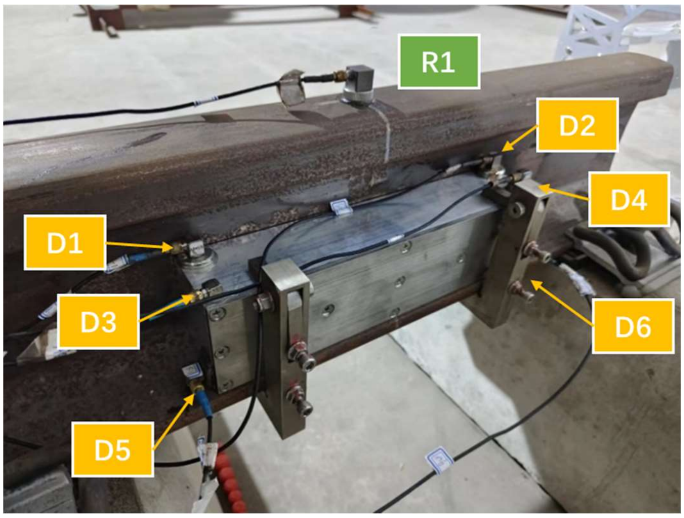


Figure 6-25 Sensor arrangement for the modal analysis on the RPD.

The analysis result of this modal test will serve for the modeling periodical rail track equipped with RPDs. The infinite track consists of multiple single-bay track units is applied in that periodic model. Therefore, the modal analysis of the RPD was conducted on the single-bay rail track. As shown in Figure 6-26, six accelerometers were mounted on the damper cavity at its six corner points. The acceleration corrected at these points can reflect the mode shape of the RPD during vibration.

The impact was applied on the top of the rail head. This excitation aims to simulate the wheel-rail impact in operation. One accelerometer was placed at the impact location as the reference response. The 1.5 mm stainless steel ball was filled in the damper cavity in this test. A thorough investigation on the influence of the filling ratio was conducted, which covered the filling ratio from 0 (empty cavity) to 90%. Worth noting here that with the 90% filling ratio, the cavity was almost full of particles considering the clearance between particles. The particle bed could barely move at this filling ratio.

Upon the impact excitation, the damper cavity would exhibit various vibration modes on the frequency band under different filling ratios. However, according to Figure 6-18, the difference brought by various filling ratios can only be observed at around 1500 Hz. Therefore, we only focus on the mode around this frequency.

Table 6-4 Modal analysis of the RPD on single-bay track around the pin-pin mode

Particle filling ratio (%)	Damper weight (g)	Modal frequency (Hz)	Modal damping ratio (%)
0	2355.9	1422	1.19
10	2730.6	1465	1.69

20	3115.8	1441	1.63
30	3578.4	1488	3.28
40	4007.2	1454	3.42
50	**4412.3**	**1465**	**3.62**
60	4809.1	1434	2.53
70	5211.9	1485	2.03
80	5606.1	1482	1.79
90	5998.5	1478	1.70

The vibration of the track induced the vibration of the RPD, and the particle bed inside the cavity. However, the modal damping ratio of the RPD varies a lot with different filling ratios, indicating the influence of the particle bed inside the cavity. According to the results presented in Table 6-4, the 50% filling ratio achieved the highest modal damping ratio. Recall the findings observed in CHAPTER 3. The large mass particle bed is less likely to be excited in vibration. This phenomenon is not beneficial for the particle bed to dissipate energy. Consequently, the filling ratio higher than 50% did not dampen the track system more than the 50% filling ratio.

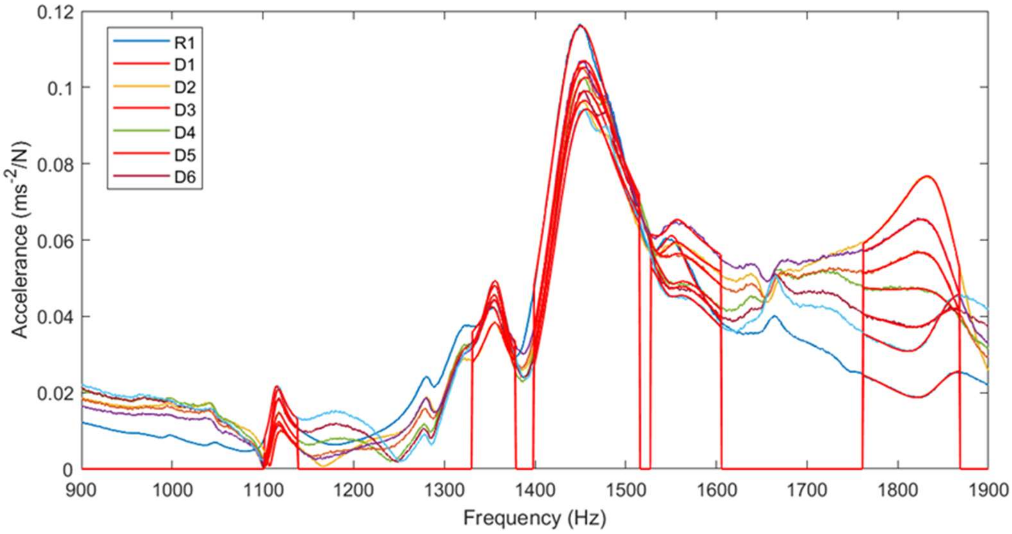


Figure 6-26 Curve fitting results of modal analysis on the damper cavity (Ao).

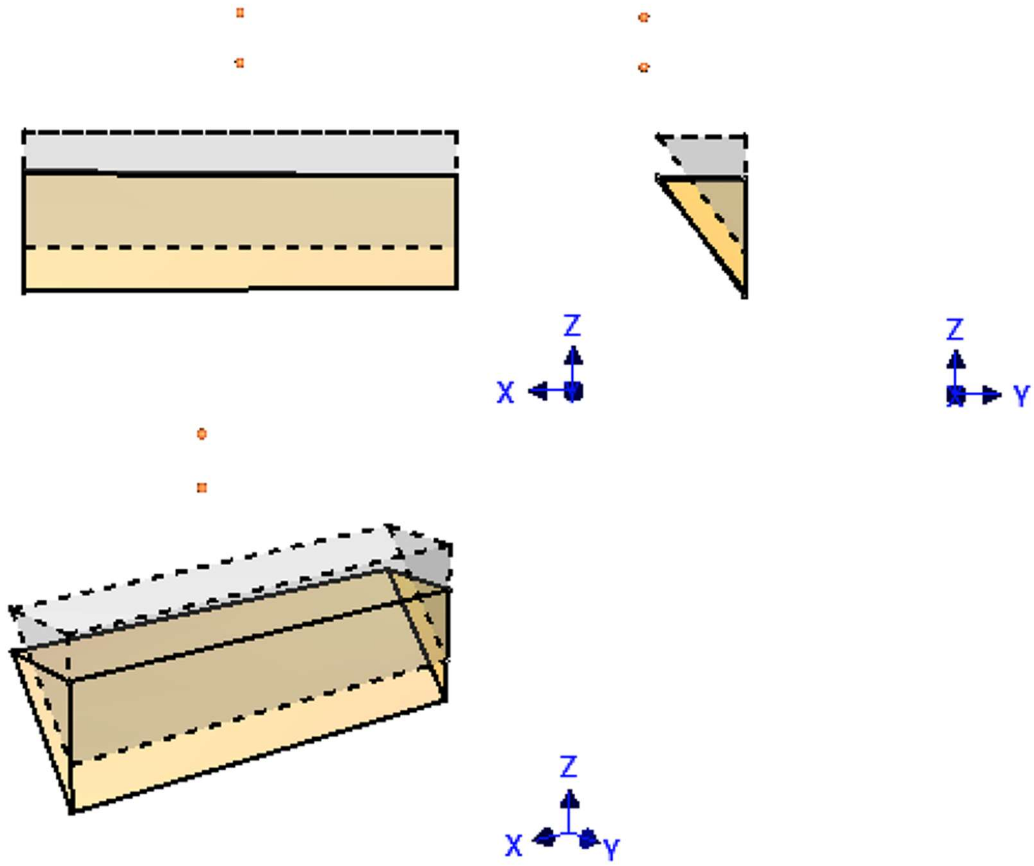


Figure 6-27 Modal shape of the RPD cavity with 50% filling ratio of 1.5 mm stainless steel balls at 1465 Hz (Ao).

Figure 6-26 shows the curve fitting results of the modal analysis. It is found that the damper cavity exhibited multiple modes. However, only the mode around the target mode of the track significantly affected the rail track as discuss in section 6.2.2.

The model shape of the RPD (Figure 6-27) reveals that the RPD mainly performs the bouncing movement in the vertical direction relative to the rail track. The fixtures behave like vertical springs that connect the dampers to the rail track. This relative displacement transmits the vibration energy to the RPDs, and then the energy is consumed by the vibration of the particle bed.

6.3 Wave Propagation Model on Periodical Structure

The vibration mitigation ability of the RPD and its dynamic property has been well studied in the laboratory. In the following content, two remaining issues should be settled for developing this kind of rail damper to control the noise and vibration of the railway. First, the previous investigation focused on the vibration aspect. Since the vibration is the origin and essence of the noise and vibration problem, the rail track acoustic radiation analysis should be conducted based on the rail track vibration. Secondly, the RPD should be evaluated on an operating line to prove its effectiveness in real applications.

For the first issue, the acoustic radiation of the rail track is related to a factor named track decay rate (TDR). The TDR describes the vibration decay along the rail track, which can be obtained by experimental and theoretical means. The TDR can be measured through a series of impact hammer tests along the rail track according to the standard BS EN 15461. Theoretically, the rail track system, with or without dampers, can be considered a periodical structure. The wave propagation along this periodical structure can be calculated through Floquet's theorem (Barone et al., 1977). Eventually, the TDR can be obtained from the wave propagation constant.

The TDR's practical and theoretical analysis can be achieved with a field experiment of the RPD. The following section selects a noise-sensitive operating urban metro line to test the effect of RPDs. Details will be presented afterward, and this section will mainly discuss the theoretical model.

6.3.1 Relationship between TDR and rail track acoustic radiation

The TDR measures the vibration decay rate along the rail track. The higher TDR means the vibration decays rapidly on the track due to damping. The acoustic radiation of the track is thus lower in this condition. The acoustic radiation ability of a unit-length rail track can be expressed as:

$$\overline{W} = A_r \sigma_r \rho_0 c_0 v^2 \tag{6-1}$$

Here, $\overline{W}$ is the acoustic radiation power of unit length rail track; A_r is the area of acoustic radiation; σ_r is a constant related to the acoustic efficiency of the rail track; ρ_0 represents the density of air, and c_0 is the sound velocity; v^2 is the average value of the vibration velocity on the track surface.

Thereby, the total acoustic radiation along a track is the integration of the unit length.

$$W = \bar{W} \int |v(x)|^2 \, dx \tag{6-2}$$

Consider the exponential form of the damping attenuation, and the vibration is assumed to decay exponentially along the track. By defining an attenuation constant or wave propagation constant μ. We will have,

$$|v(x)|^2 = v(0)e^{-2\mu|x|} \tag{6-3}$$

Combining the Equation (6-2) and Equation (6-3), the total acoustic radiation is expressed as:

$$W = \bar{W} \int |v(x)|^2 \, dx = \bar{W} v^2(0) \frac{1}{2\mu} \tag{6-4}$$

According to Equation (6-4), it is clear that with a larger μ, the acoustic radiation of the rail track will be lower. However, it is common to use another factor to replace the μ in practice, which is the TDR with the unit of dB/m. The transition from wave propagation constant μ to TDR is expressed as:

$$TDR = 20log_{10}(e^{\mu}) \tag{6-5}$$

From the above derivation, it can be concluded that the TDR is an important physical factor in predicting the noise behavior of the railway system. The wave propagation constant mentioned here is the real part of the complex constant. A detailed discussion will be presented afterward.

6.3.2 Free wave propagation model

The rail track on site is a periodically supported infinite beam structure. Also, the RPDs are evenly installed on every bay of the rail track. The periodic rail structure consists of infinite units with a length of one bay, similar to the single-bay rail track in the laboratory experiment. Therefore, the single-bay rail test results are a good reference for the modeling. As seen in Figure 6-18, the installed RPDs mainly affect the rail track around its pin-pin mode, but have limited influence on the FRF at other frequency ranges, besides, according to the modal analysis of the installed RPDs. The mode of the RPDs around the pin-pin mode is related to the bouncing movement of the damper cavity. Therefore, in modeling the vertical vibration of the rail track mounted with RPDs, the RPDs are vertically connected to the rail track with a spring to reflect its influence on the track.

Here, a simplified model to study the wave propagation is proposed here based on a Euler-Bernoulli beam for practical use. The ground vibration and damping are neglected in this model, since the ground vibration frequency is too low to significantly affect the vibration of rail tracks.

The wave propagation constant μ is obtained by analyzing the periodic segment of the infinite track, as shown in Figure 6-28. In this case, the rail is securely attached to a concrete slab. The fastening system, which consists of discretely positioned rail clips, applies two constraints to the rail track: the translational stiffness S_t and the rotational stiffness S_r.

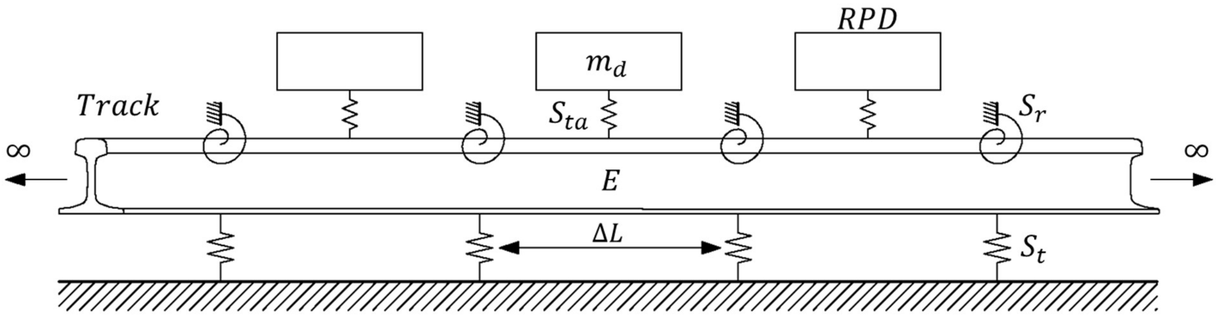


Figure 6-28 Periodic model of the track placed on concrete slab with discrete supports and RPDs.

The vibration along the rail track is analyzed by the Euler-Bernoulli beam. The complex stiffness is applied in this model for convenient derivation, which contains the energy loss factor η as the imaginary part. The relationship between the loss factor η and damping ratio ζ is (Petrone et al., 2015):

$$\eta(\omega) = 2\zeta(\omega)\sqrt{1-\zeta(\omega)^2} \tag{6-6}$$

Figure 6-29 illustrates the force analysis of the track segment with the length of one bay. The track segment is divided into five sections to analyze the vibration and force interaction with RPD.

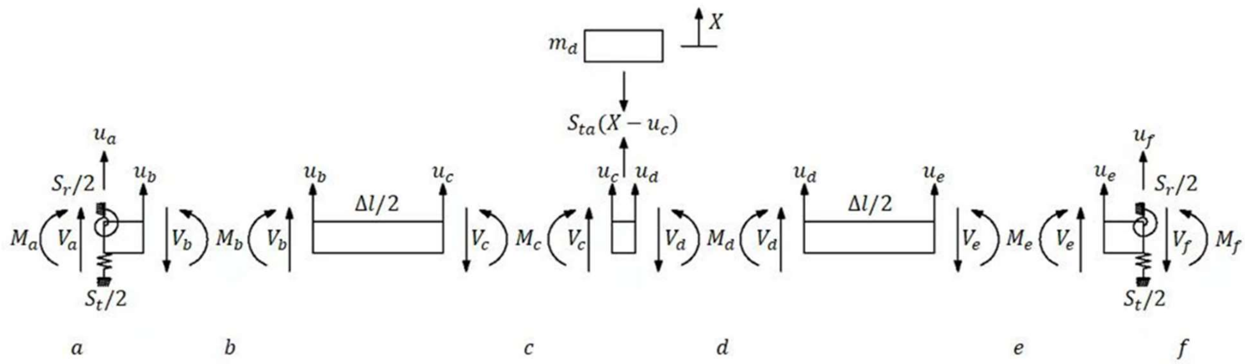


Figure 6-29 Displacements and forces of the periodic segment on different sections.

Following the derivation given in the literature (Jin et al., 2022), consider the section bc, it's simply a beam section. The transverse displacement is derived as follows:

$$u(x) = A_1 \cos k_b x + A_2 \sin k_b x + A_3 \cosh k_b x + A_4 \sinh k_b x \tag{6-7}$$

Where $k_b = (\omega^2 M_r / EI)^{1/4}$ is the wavenumber of the track. Here E and I are the elastic modulus and area moment of inertia of the rail section, M_r is the weight of unit length track. According to the definition, we also can obtain the slope $\theta(x) = \partial u(x)/\partial x$, the moment $M(x) = EI\partial^2 u(x)/\partial x^2$, and the shear force $V(x) = -EI\partial^3 u(x)/\partial x^3$. At the left end of the beam, the amplitudes are calculated as follows:

$$A_1 = \frac{u_b}{2} - \frac{M_b}{2EIk_b^2}, A_2 = \frac{\theta_b}{2k_b} - \frac{V_b}{2EIk_b^3},$$
$$A_3 = \frac{u_b}{2} + \frac{M_b}{2EIk_b^3}, A_4 = \frac{\theta_b}{2k_b} + \frac{V_b}{2EIk_b^3} \tag{6-8, a-d}$$

Similarly, applying Equation (6-7) and Equation (6-8, a-d) to the right end of the beam, the responses at the other end of the rail segment are obtained. The relationship between the responses at the two ends can be provided as $\boldsymbol{r_c} = \boldsymbol{T_{cb}} \boldsymbol{r_b}$. The response vector $\boldsymbol{r_b} = \{u_b, \theta_b, M_b, V_b\}^T$ contains both displacement and force information. The wave transfer matrix $\boldsymbol{T_{cb}}$ at this segment is calculated as follows (Wang et al., 2017):

$$\boldsymbol{T_{cb}} = \frac{1}{2}\begin{bmatrix} \cosh\Lambda + \cos\Lambda & k_b^{-1}\sinh\Lambda + k_b^{-1}\sin\Lambda & P^{-1}\cosh\Lambda - P^{-1}\cos\Lambda & Q^{-1}\sinh\Lambda + Q^{-1}\sin\Lambda \\ k_b\sinh\Lambda - k_b\sin\Lambda & \cosh\Lambda + \cos\Lambda & U^{-1}\sinh\Lambda + U^{-1}\sin\Lambda & P^{-1}\cosh\Lambda - P^{-1}\cos\Lambda \\ P\cosh\Lambda - P\cos\Lambda & U\sinh\Lambda - U\sin\Lambda & \cosh\Lambda + \cos\Lambda & k_b^{-1}\sinh\Lambda + k_b^{-1}\sin\Lambda \\ Q\sinh\Lambda + Q\sin\Lambda & P\cosh\Lambda - P\cos\Lambda & k_b\sinh\Lambda - k_b\sin\Lambda & \cosh\Lambda + \cos\Lambda \end{bmatrix} \tag{6-9}$$

Here $\Lambda = k_b \Delta l/2$, $U = EIk_b$, $P = EIk_b^2$, and $Q = EIk_b^3$. The wave transfer matrix at the third section is obtained as $\boldsymbol{T_{ed}} = \boldsymbol{T_{cb}}$.

Now consider the first segment, which is actually a cross-section without length. The displacement is continuous at two ends of this segment. The force balance is established via $\left[M_b = M_a - \left(\frac{S_r}{2}\right)\theta_a, V_b = V_a - \left(\frac{S_t}{2}\right)u_a\right]$. Therefore, the transfer matrix $\boldsymbol{T_{ba}}$ is established as:

$$\boldsymbol{T_{ba}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -S_r/2 & 1 & 0 \\ -S_t/2 & 0 & 0 & 1 \end{bmatrix} \tag{6-10}$$

It is also straightforward to conclude that $\boldsymbol{T_{fe}} = \boldsymbol{T_{ba}}$. These two transfer matrixes represent the effect of the fastening system on the track. As for the interaction between the RPD and the rail track response, the equation of motion of the RPD are first applied:

$$m_d \ddot{X} = S_{ta}(u_c - X) \tag{6-11}$$

Where m_d is the effective mass of the damper, which can be obtained from the surrogate model established for the RPD in the previous chapters. S_{ta} is the damper's dynamic translational stiffness based on the damper cavity's modal analysis together with the fixture. X is the actual displacement of the RPD in vertical direction.

With the translational reaction force that RPD applies to the track, the transfer matrix at the damper-attached segment is calculated as:

$$\boldsymbol{T_{dc}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ K_d & 0 & 0 & 1 \end{bmatrix} \tag{6-12}$$

Here $K_d = S_{ta} m_d \omega^2 / (S_{ta} - m_d \omega^2)$. The boundary conditions at the rail track unit are applied, Assume the vibration starts at point a, the vibration wave transfer matrix between the right end to the left end is expressed as:

$$\boldsymbol{r_f} = \boldsymbol{T_{fe} T_{ed} T_{dc} T_{cb} T_{ba} r_a} \tag{6-13}$$

Based on the periodic boundary conditions of the rail track, the wave propagation constant of the rail track with RPD is derived from the transfer matrix on one track unit. When the fastening system is implemented with intervals of Δl, it is the length of a track unit. According to the Bloch theorem (Ho et al., 2003):

$$\boldsymbol{r}(x + \Delta l) = e^{-\mu \Delta l} \boldsymbol{r}(x) \tag{6-14}$$

The wave propagation constant $\mu = \mu_r + i\mu_i$ is a complex coefficient between two adjacent periodic units. The real part μ_r is the attenuation constant that represents the amplitude decay rate, and the imaginary part μ_i is the phase constant that represents the phase difference. Based on Equation (6-13) and Equation (6-14), it is derived that:

$$[\boldsymbol{T} - e^{-\mu\Delta l}\boldsymbol{I}]\boldsymbol{r}(x) = 0 \tag{6-15}$$

Where $\boldsymbol{T} = \boldsymbol{T}_{\boldsymbol{fe}}\boldsymbol{T}_{\boldsymbol{ed}}\boldsymbol{T}_{\boldsymbol{dc}}\boldsymbol{T}_{\boldsymbol{cb}}\boldsymbol{T}_{\boldsymbol{ba}}$, non-trivial solutions only exist when:

$$\left|\boldsymbol{T} - e^{-\mu\Delta l}\boldsymbol{I}\right| = 0 \tag{6-16}$$

Therefore, the propagation constant μ is calculated from the eigenvalue analyzation of the $\boldsymbol{T}$. The results give two pairs of the Bloch waves $(\pm\mu_1, \pm\mu_2)$ for the periodically supported rail track. The two pairs of propagation constants represent exactly the same characteristic wave, only the travel direction along the rail track is opposite. Physically, according to the form of the propagation constant, three regions in the frequency domain are distinguished for each pair of the propagation constant (Xiao et al., 2013):

1. Attenuation region: when $\mu_r > 0$ and μ_i in the form of $n\pi$, where n is an integer. In this region, the wave attenuated along the structure with the adjacent unit vibration in phase or out of phase.

2. Propagation region: when the wave propagation is purely imaginary, $\mu_r = 0$ and $2n\pi < |\mu_i| < (2n+1)\pi$. In this region, the wave is propagating without attenuation, only have a phase chase on each adjacent unit.

3. Complex region: when $\mu_r > 0$ and $2n\pi < |\mu_i| < (2n+1)\pi$. The wave is propagating and attenuating simultaneously along the track.

Xiao et al. (2013) presented an alternative approach to calculate wave propagation constants, which is discussed in Appendix B. It should be noted that the above discussion assumes that the damping is missing in the structure. Once the damping is introduced, the wave propagation condition will also be changed. For instance, the wave will not be able to propagate to infinite distances as in the propagation region. As mentioned above, the TDR is related to the wave propagation constant. If use Ψ to represent the TDR, then we have:

$$\Psi = 20log_{10}\left[e^{Re\{\mu\}}\right] \tag{6-17}$$

6.3.3 Parametric study of the model

The parametric study of the model helps to understand the wave propagation on the rail track. In this case study, the rail is assumed to be installed on the concrete slab. The clip stiffness also varies according to different clip types, influencing wave propagation. Other parameters also have an impact on the wave constant. The baseline parameters are selected from the literature for a standard urban metro line in China (Zhai et al., 2009).

Table 6-5 Main parameters of the concrete slab track

Notation	Parameter	Value
E	Elastic modulus of rail	$2.059 \times 10^{11}\ N/m^2$

ρ_r	Density of rail	$7.860 \times 10^3\ kg/m^3$
m_r	Rail mass per unit length	$60.64\ kg$
Δl	Sleeper spacing	$0.625\ m$
I_0	Torsional inertia of rail	$3.741 \times 10^{-5}\ m^4$
S_t	Fastener stiffness in vertical direction	$2.5 \times 10^7\ N/m$
S_r	Fastener rotational stiffness	$7.0 \times 10^4\ N/rad$
η_r	Loss factor of rail track	0.02
η_f	Loss factor of fastener	0.1

The two pairs of wave constant $\pm\mu_1$ and $\pm\mu_2$ of the bare track without RPD are calculated based on these parameters. These two pairs of wave constants represent vertical wave propagation for near and far fields. The curve that depicts the wave propagation constant in the frequency range is called the dispersion curve, which can be seen in Figure 6-30 and Figure 6-31.

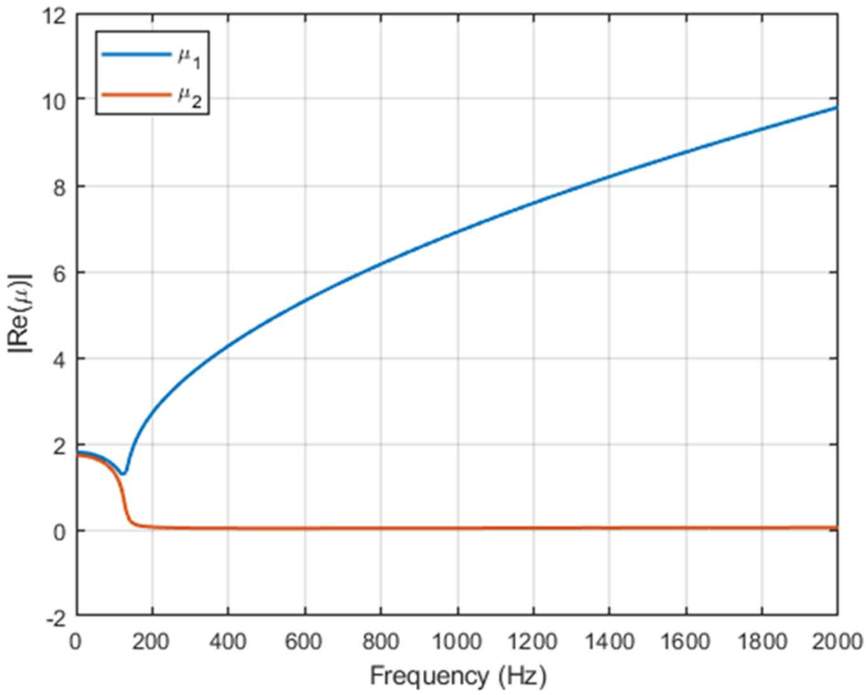


Figure 6-30 Dispersion curve of the real part of the wave propagation constant.

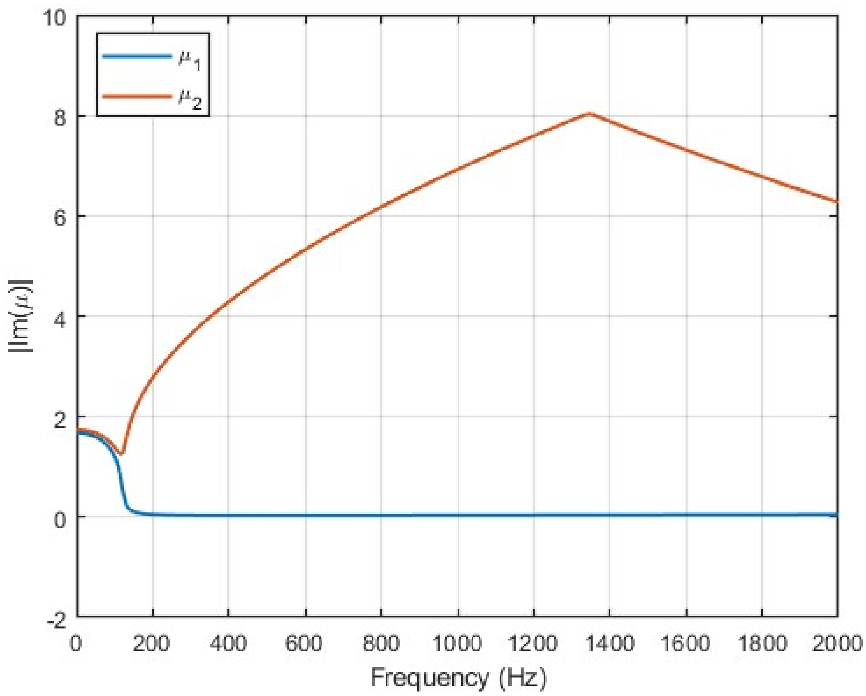


Figure 6-31 Dispersion curve of the imaginary part of the wave propagation constant.

The μ_1 shows the vertical wave propagation in the near field. The wave attenuates quickly along the track. The μ_2 shows the wave propagation in the far field. Due to structural damping, the value of the real constant is higher than zero on the entire frequency band, which means the wave propagates with attenuation along the rail track. A band gap is observed around 1340 Hz on the imaginary part of the wave constant, indicating the appearance of the pin-pin mode of the track at this frequency. The TDR (Ψ) of the track can be calculated according to the real part of the μ_2.

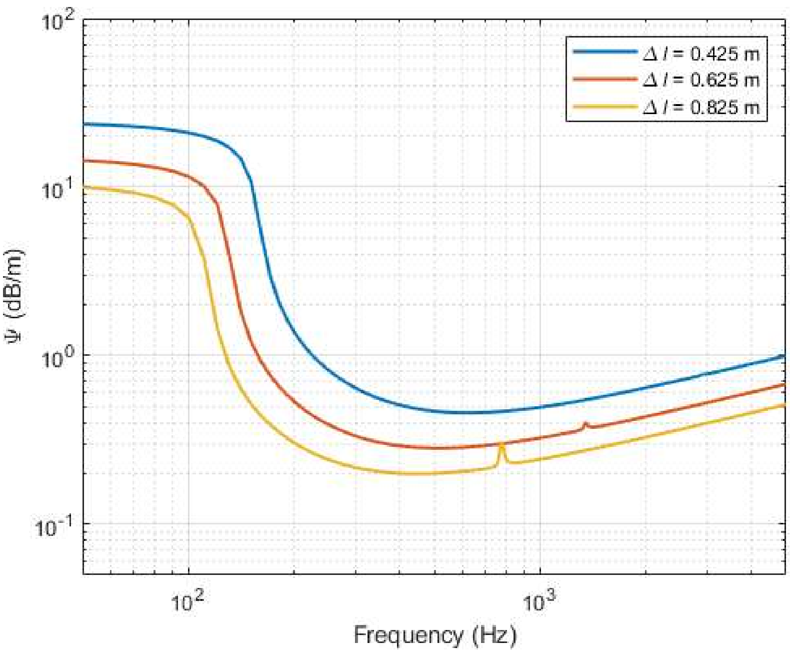


Figure 6-32 Effect of the sleeper spacing on the track TDR.

First, for the effect of the sleeper spacing, a small value indicates more fasteners are installed on the track. The fasteners are equipped with a pad with damping. More fasteners will dissipate more vibration energy, so it can be seen that the TDR is more prominent with a smaller spacing of the fasteners. On the rail track, there is a cut-on frequency related to wave propagation (Thompson, 2013). When the frequency is higher than the cut-on frequency, the damping in the rail has a negligible effect. According to the result, adding the fasteners will push the cut-on frequency to a higher value.

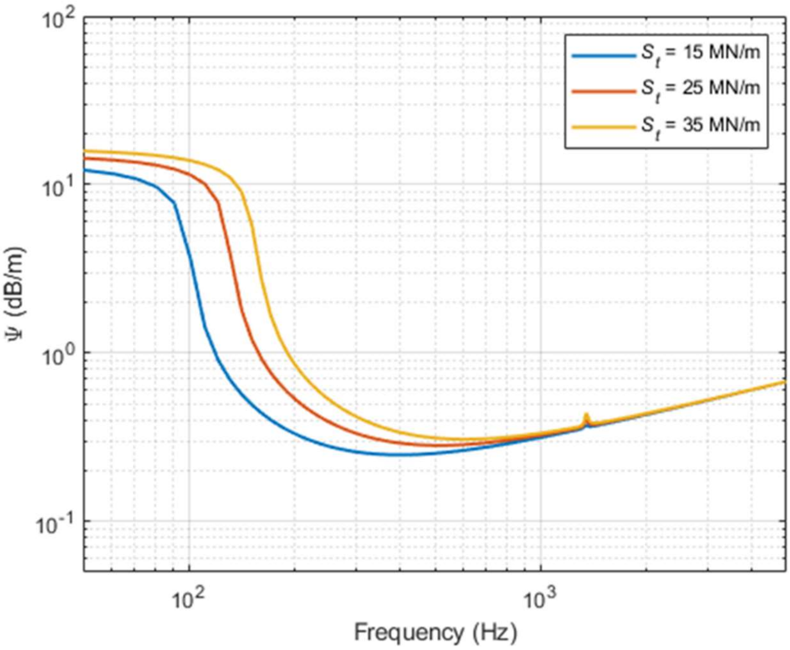


Figure 6-33 Effect of the fastener stiffness on the track TDR.

Figure 6-33 shows the influence of fastener stiffness on the track TDR. The value of S_t mainly affects the cut-on frequency of the track. With a smaller value of clip stiffness, the cut-on frequency will also be smaller. When it goes to a high-frequency range, the effect of clip stiffness gradually wears off.

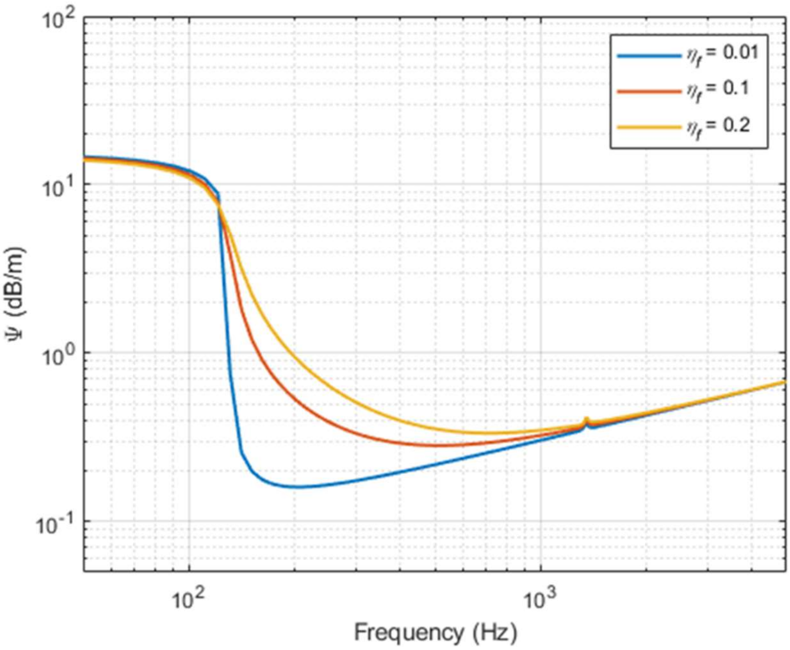


Figure 6-34 Effect of the fastener loss factor on the track TDR.

The damping pad dissipates the vibration energy whenever the wave passes a fastener. The results from Figure 6-34 again indicate that the influence of fasteners mainly concentrates on the cut-off frequency. When the frequency is larger, specifically higher than the pin-pin mode frequency, the properties of the track determine the wave propagation on the track.

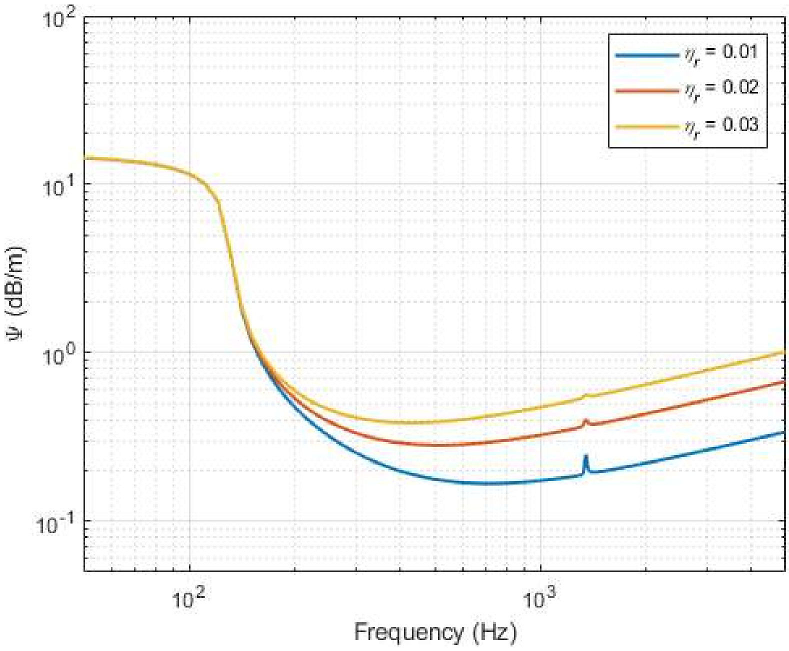


Figure 6-35 Effect of the rail track loss factor on the track TDR.

The analysis of the effect of the rail track loss factor agrees with the statement. The track damping affects the wave propagation in the high-frequency range, where above the cut-on frequency. Of course, the TDR increases with a large track loss factor. In the frequency range lower than the cut-on frequency, the damping of the track barely does not affect the wave propagation. It can be explained by the fact that the near field vibration wave in this frequency range attenuates quickly on the track. The vibration is dissipated before the track damping can show a significant influence.

The results indicate a noticeable spike on the TDR around 800 Hz, signifying the presence of a bandgap in the periodic system. Within this bandgap range, vibrations exhibit a greater decay. It is clear that the positioning of supports determines the location of the bandgap, as alterations in spacing directly impact the structure of the periodic system.

6.4 Experimental Validation of RPD Performance on Field

6.4.1 In-situ investigation

The selected noise-sensitive section is located in Shenzhen Metro Line 5, a viaduct with a slab track system between Tanglang Station and University town station. There is a residential area in the vicinity of this metro line. The test location is at the departing line, where about 500 m away from the Tanglang Station. The passing train reaches the test location at approximately 70 km/h. This experiment was conducted by Dr. Masoud, Mr. Owen Wang, Mr. Xiang-Xiong Li, Mr. Guang Zhou, Mr. Chao Zhang and me.

Figure 6-36 Test location and the surrounding situation (Masoud).

The length of the main span of each bridge is 30 m. To cover the entire length of one bridge span, the installation length of RPD is 40 m on the rail track, which contains 67 bays of rail track. The RPDs were mounted on every bay on both sides of the rail track. In this test, the RPDs were filled with 50 % of the 1.5 mm stainless steel balls.

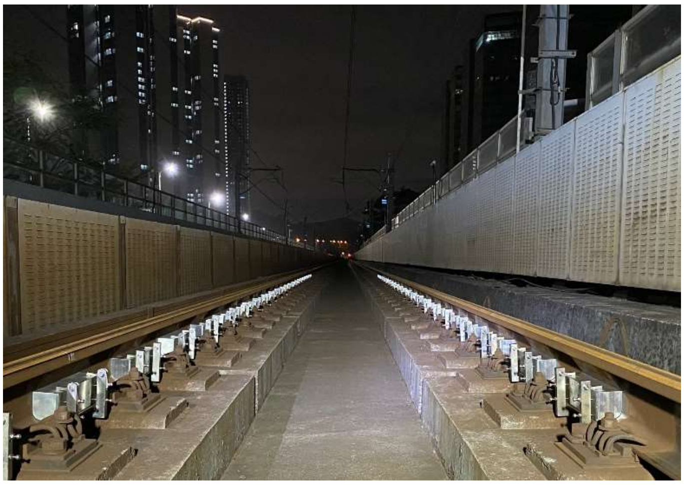


Figure 6-37 Installation of RPDs on rail track.

A series of acceleration sensors and microphones were deployed at the 34th bay of the test section (the mid-bay of the 40 m test section) to evaluate the performance of RPD in reducing the broadband rolling noise and controlling the rail vibration under dynamic excitation during a train passing by. When the train passes the test section, the installed accelerometers and microphones measure the rail vibration responses and the emitted noise level. The sensors and microphones' deployment positions are shown in Figure 6-38. Four acceleration sensors (A1, A2, A3, A4) are installed on the north track; two on each rail, one on the rail web, and another on the rail foot. The acceleration sensors are triaxial DYTRAN accelerometers with a 1000 g range mounted on the rail using epoxy glue, as shown in Figure 6-39. Two microphones (M1, M2) were set on the rail side at the rail head level, 60 cm away from the rail, as shown in Figure 6-40.

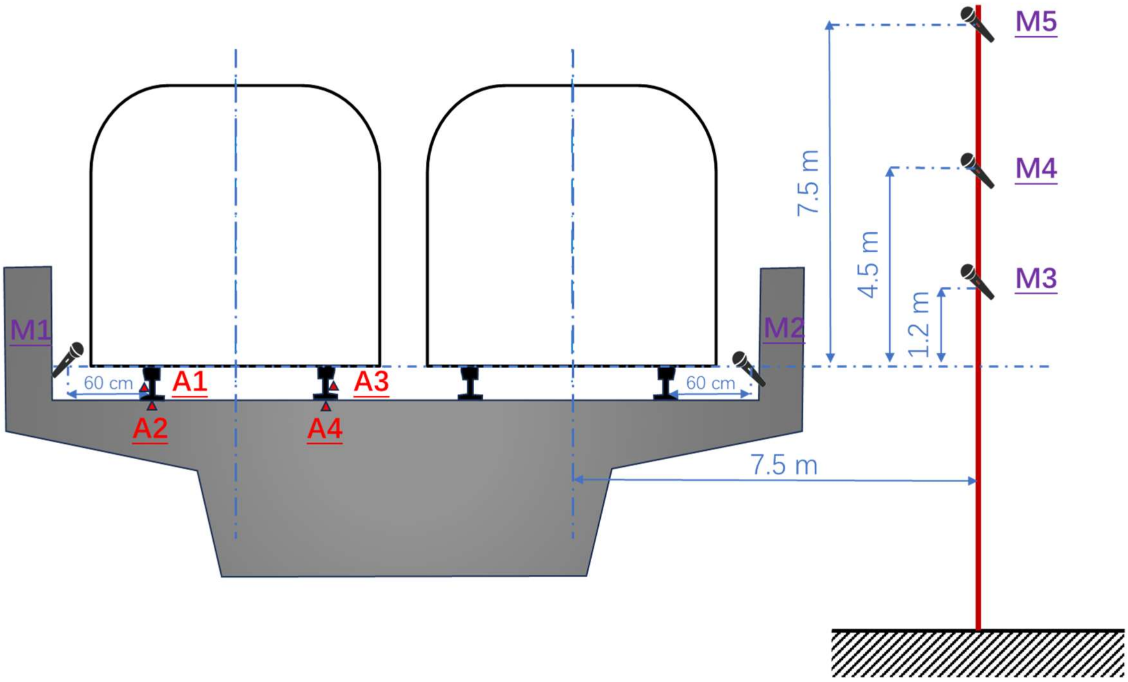


Figure 6-38 Dynamic test layout and microphones deployment position.

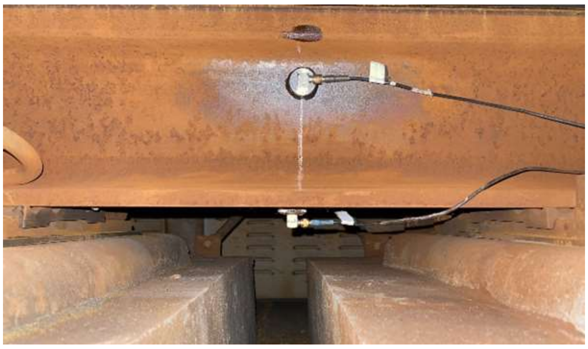


Figure 6-39 Positions of rail web and rail foot accelerometers.

Figure 6-40 Position of the rail side microphone.

Three microphones (M3, M4, M5) were mounted on a pole, set 7.5 m away from the central line of the south track on the viaduct side, and fixed in place using a temporary scaffold (see Figure 6-41). The microphones' height is set at 1.2 m, 3.5 m, and 7.5 m, respectively, from the top surface of the rail head, as shown in Figure 6-38. All microphones are the free-field type made by Brüel & Kjær.

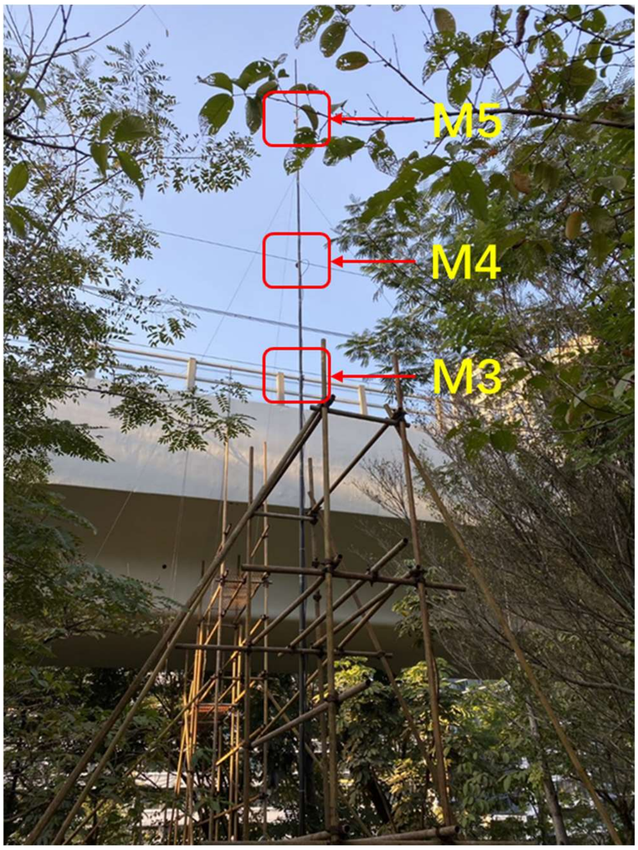


Figure 6-41 Positions of viaduct side microphones.

6.4.2 Track vibration evaluation

Evaluating the track vibration before and after the RPDs installation is essential to evaluating the RPD. According to the laboratory investigation of the RPDs, this king of dampers can suppress the vibration in relatively high frequency but have limited effect in the frequency range lower than 1000 Hz.

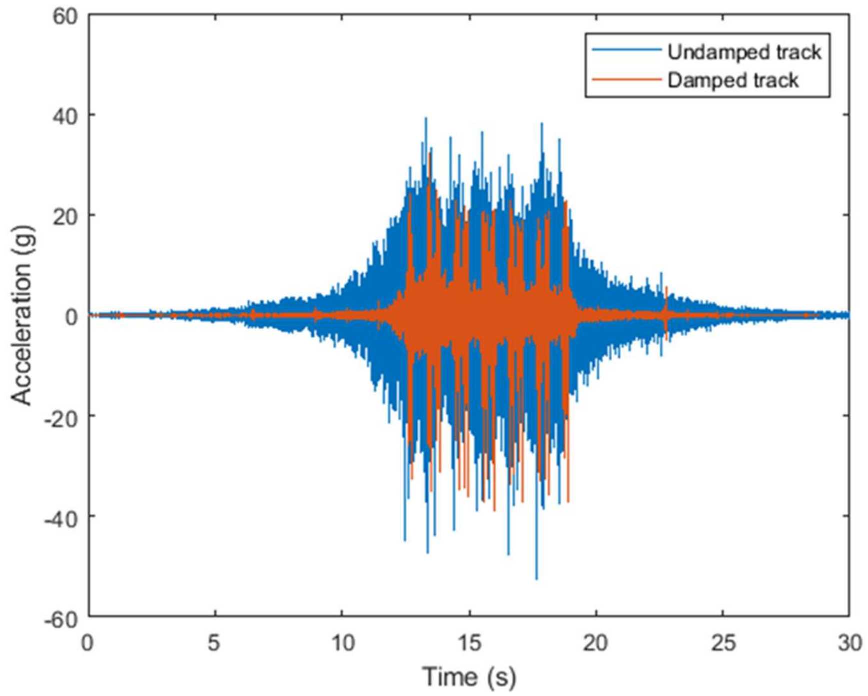


Figure 6-42 Track vibration in time domain before and after the installation of RPDs.

Figure 6-42 shows the variation of the track vibration before and after the installation of RPDs. It can be seen that the RPDs have a profound vibration mitigation ability on the rail track. Clear peaks from the passing wheel pairs impact excitation can be observed on the damped track. Other than this type of vibration, the other vibrations are significantly reduced by the RPDs.

The frequency analysis is presented in Figure 6-43. The frequency components of the vibration signal reveal the main effective range of the RPDs.

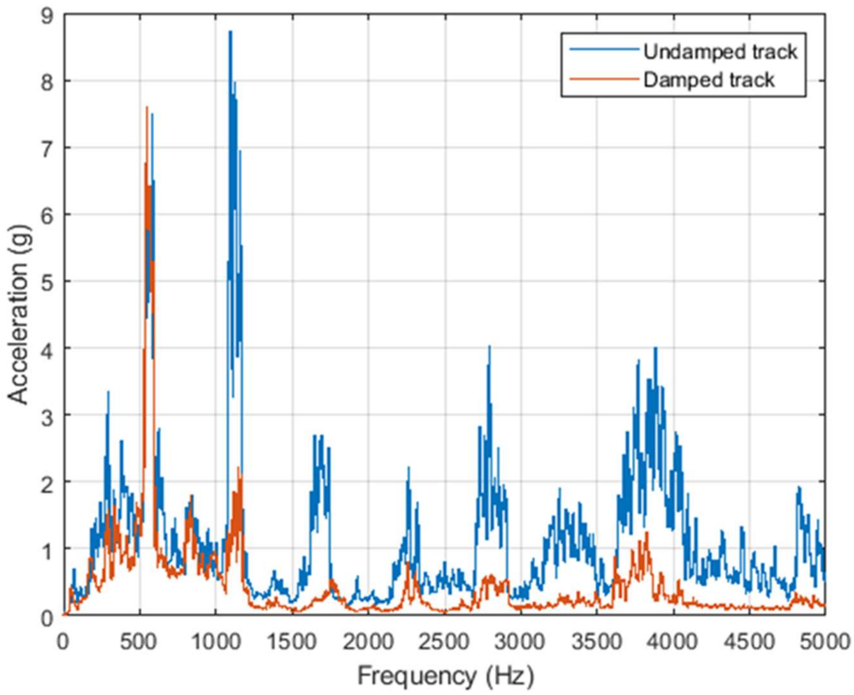


Figure 6-43 Track vibration in frequency domain before and after the installation of RPDs.

Similar to the results observed in the laboratory, the RPDs have a limited effect on frequency under 500 Hz. The mode around 500 Hz is empirically identified as the bouncing mode of the track, which is relevant to the passing wheel pairs. However, the RPDs suppressed most of the vibration energy in the high-frequency range.

Spikes can be observed in the vibration spectrum of the track. These spikes are super-harmonic responses of the rail track caused by nonlinear effects. This common phenomenon that can be identified during field tests of rail tracks.

6.4.3 Track decay rate evaluation

The TDR reflects the vibration attenuation on the rail track, which is a significant factor in evaluating the noise emission of the track. Following the BS EN 15461:2008+A1:2010 standard, the TDR can be measured through a series of impact hammer tests on the track. One accelerometer was attached to the rail head. The vibration amplitudes were recorded under the impacts at gradually increased distances from the accelerometer. The result represents the vibration attenuation with respect to the distance.

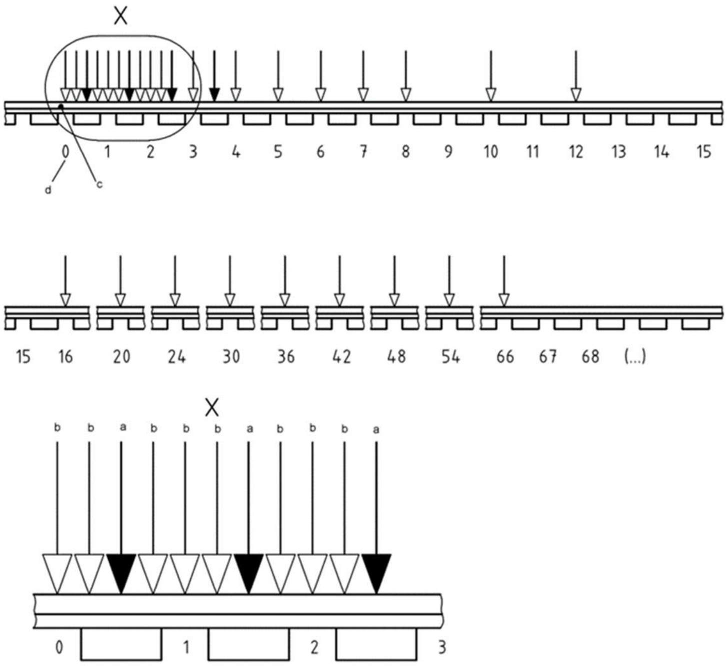


Figure 6-44 Impact locations with respect to the location of the reference accelerometer for the measurement of TDR (BS EN 15461 standard).

The measured TDR is calculated with the following equation:

$$\Psi = \frac{4343}{\sum_{x=0}^{x_{max}} \frac{|A(x_n)|^2}{|A(x_0)|^2} \cdot x_n} \tag{6-18}$$

Where $A(x_n)$ is the acceleration amplitude, and x_n is the distance to the accelerometer. In practical, the TDR is presented in 1/3 octave frequency band.

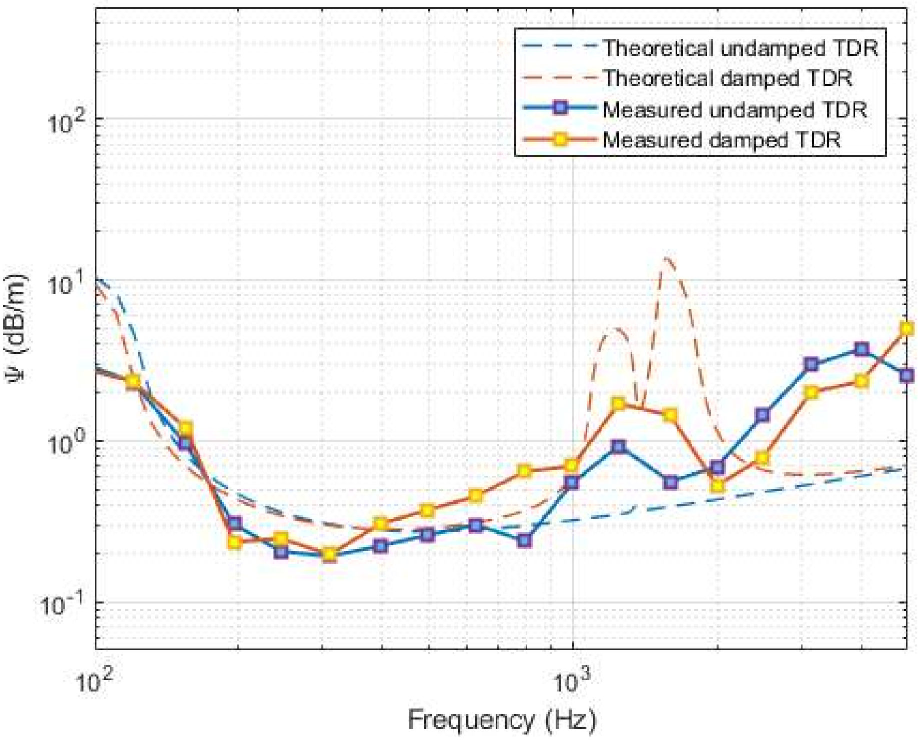


Figure 6-45 Variations in the theoretical and measured TDR with and without the installation of RPDs.

Figure 6-45 presents the variations in the theoretical and measured TDR before and after the installation of the RPDs. Both theoretical and measured results show that the RPDs increased the TDR in the frequency range higher than the cut-on frequency, indicating that the dampers elevated the damping effect on the rail tracks rather than the fasteners. The measured TDR of the damped track was increased in a wide frequency range from 1000 Hz to 5000 Hz. It is known that the vibration energy of the track concentrates on this range. The high-level vibration amplitude induces a more significant damping effect of the particle damping.

However, the theoretical TDR of the damped track only showed the effect of RPDs around 1600 Hz, which illustrates the limitation of the simplified model. The connection and interaction between the RPD and the rail track is much more complex. Many factors will introduce energy dissipation in this system, like the fixture's deformation, the friction between components, etc. Moreover, the nonlinear effect of the RPD and the railway system is totally ignored in this simple model. In the future, it is desired to develop a more sophisticated model to study the effect of installing RPDs on the rail track.

6.4.4 Near and far field noise evaluation

The imposition of M4 and M5 was meant to explore the noise emission in a higher direction. It may be tedious to show the results collected by all microphones. In this case, the sound pressure recorded by M1 and M3 was selected to represent the near-field and far-field noise. The evaluation of the RPDs on noise mitigation is based on these experimental results.

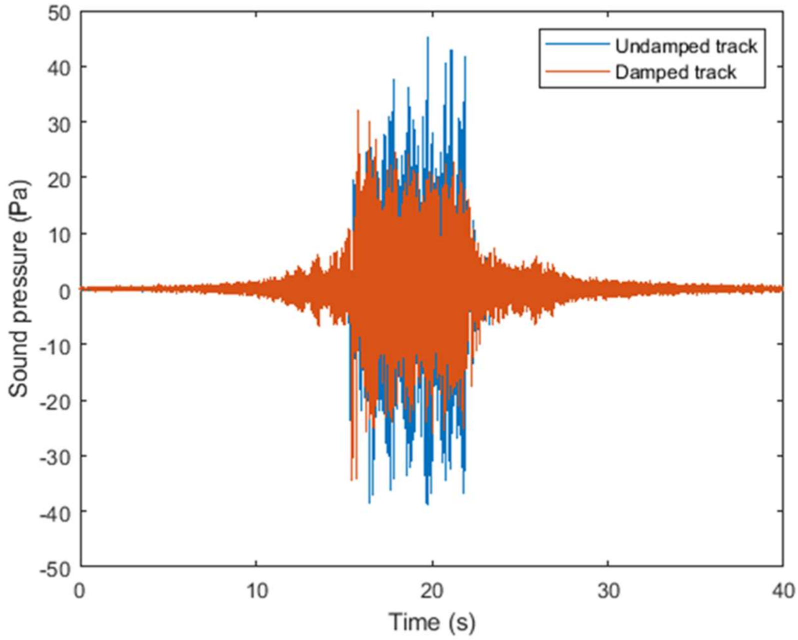


Figure 6-46 Variation of the near field sound pressure in time domain.

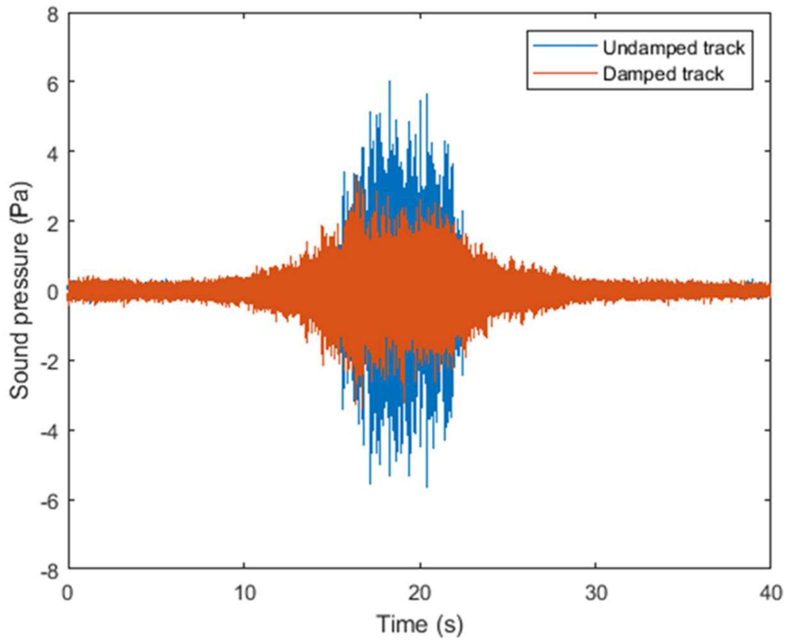


Figure 6-47 Variation of the far field sound pressure in time domain.

According to the results, the near-field sound pressure is deservedly much higher than the far-field sound pressure. The noise emission includes all directions in the surroundings, while the far-field microphone can only capture a part of it. The effect of passing wheel pairs can also be identified at the near field. The peak sound pressure can reach 45 Pa at the trackside. The peak sound pressure level at the far field is around 6 Pa. The installed RPDs effectively mitigate the sound pressure at the near and far fields. Since the installation of RPDs successfully reduces the vibration of the track.

The wavelet analysis of the sound pressure is applied to evaluate the frequency and time domain noise. The sound pressure level (SPL) is calculated from the sound pressure.

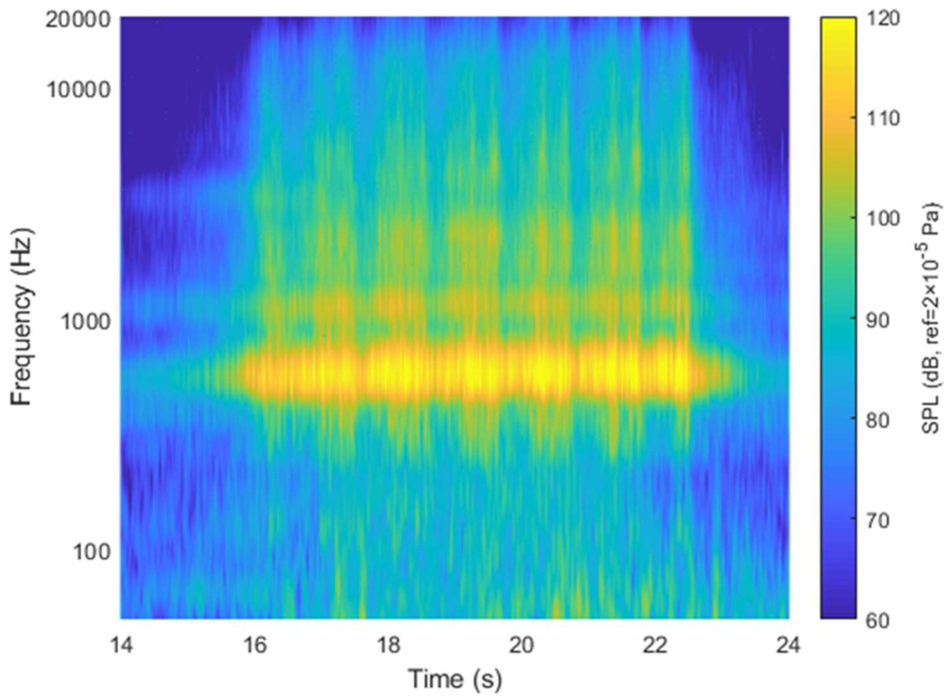


Figure 6-48 Wavelet analysis of the SPL measured at near field without installation of RPDs.

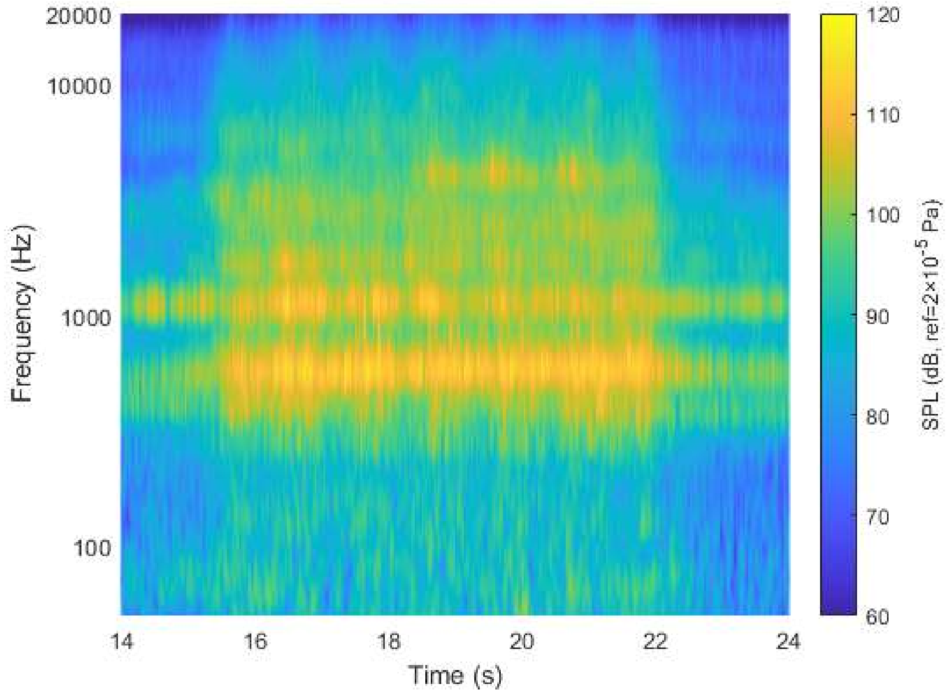


Figure 6-49 Wavelet analysis of the SPL measured at near field with installation of RPDs.

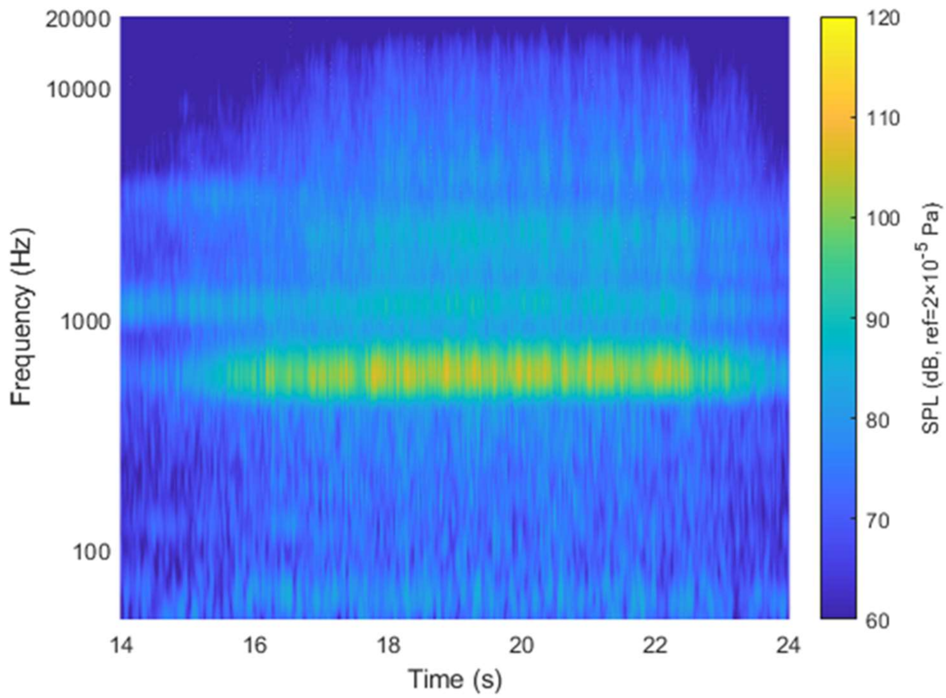


Figure 6-50 Wavelet analysis of the SPL measured at far field without installation of RPDs.

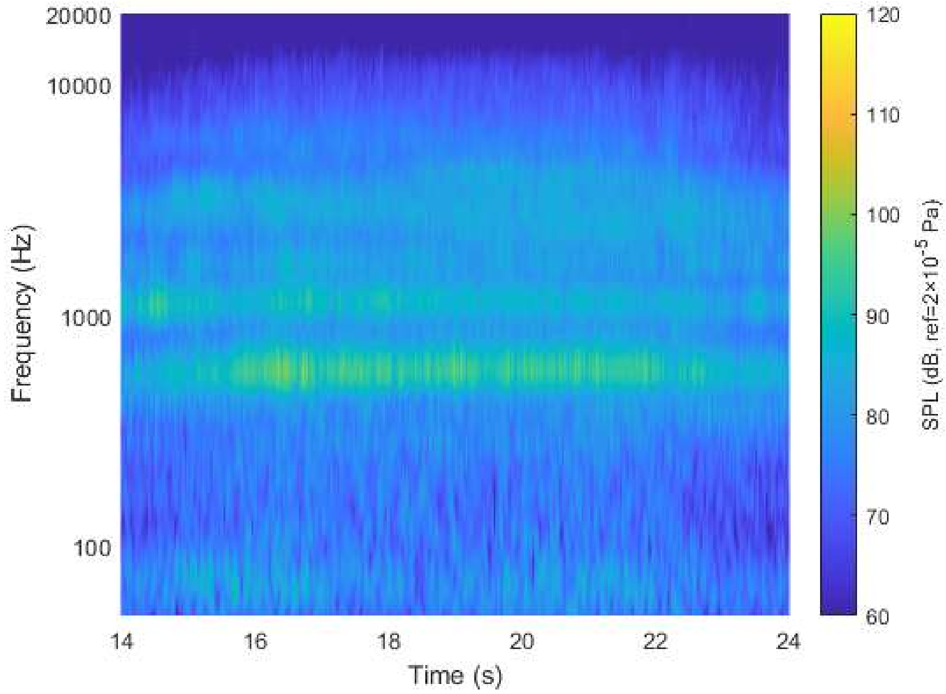


Figure 6-51 Wavelet analysis of the SPL measured at far field with installation of RPDs.

As shown in the figures, the noise from the wheel track interaction mainly concentrates around 600 Hz and 1000 Hz. Recall the vibration of the track. It is related to the bouncing mode and the first bending mode of the track concerning the fasteners. When there are no RPDs on the track, the noise at this range can reach 120 dB, which is quite a considerable level considering the train speed. The installation of RPDs affects mitigating the noise at the sensitive frequency. Around 8 dB reduction is achieved with RPDs at 600 Hz. However, it also can be observed that the noise energy was spread on the frequency band after the installation of RPDs. This kind of reduction is outstanding compared to conventional tuned mass dampers, which tend to require much more mass than the proposed RPDs.

The far-field SPL shows the same characteristic as the near-field SPL, with a reduced noise level. The presented results indicate that the rail vibration reduction effectively makes a quiescent passenger environment.

6.5 Summary

A novel rail particle damper was developed to control the noise and vibration of the railway system. A series of experiments were conducted to evaluate the ability of RPDs to mitigate the vibration of the rail track and, thus, control noise emission. The RPDs were tested on a single-bay rail track first, and various types of particles were compared through impact hammer tests. Then the selected particle types were further tested on a 6 m rail testbed. Finally, the RPDs were mounted on an operating urban metro line to control the noise and vibration. A periodic structure model was produced with RPDs to predict the TDR of the rail track. The conclusions through these tests can be summarized as follows:

1. The impact hammer tests on the single-bay track revealed that the granular material outperforms the other type of materials, including powders, solid balls, and liquid. Installing RPDs can effectively suppress the vibration of the track by decreasing the peak of the track's FRF.
2. The three selected material types were tested on the 6 m rail testbed. RPDs with these filling materials can mitigate the pin-pin mode of the rail track but have a limited effect on the vibration in the lower frequency band. The 1.5 mm stainless steel ball was identified as the best material, which can mitigate 50% of the pin-pin mode near and

85% at the far end.

3. The in-situ test has proved the effectiveness of RPDs in noise and vibration control. The installation of RPDs elevated the TDR of the rail track, which reduced the noise emission of the track. To be specific, around 8 dB of noise reduction is achieved with RPDs at 600 Hz.

CHAPTER 7.
Conclusions and Recommendations

7.1 Conclusions

This thesis presents a study on PDs for noise and vibration control in the railway system. The investigation is carried out in two aspects, i.e., the experimental investigation and the surrogate model establishment via DL methods. Both contents are pioneering attempts in the development of PDs. To meet the frequency range of the rail track vibration, the experiments conducted on the PDs reached a high-frequency range of up to 2000 Hz. This range is seldom explored in previous studies since the micro-displacement vibration makes friction the main effect in energy dissipation. The movement of the particle bed in this high-frequency vibration is intricate to model, especially for the granular materials. Despite applying the DEM that requires a significant computational load, this study chooses the DL methods to circumvent describing the movement of every single particle inside the damper cavity but to surrogate the desired target through the experimental results. This study has presented a sophisticated DL modeling framework with the aid of a series of state-of-art methodologies. The primary contributions and findings a summarized as follow.

The findings from the experiments are first concluded. The granular material-filled dampers dissipate energy mainly by the friction between the particles during vibration. Various parameters influence the energy loss factor. In general, more energy is dissipated with more filling particles. However, the energy loss factor decreases with the filling ratio because a large-weight particle bed is less likely to be excited. Moreover, a high excitation level also results in a better performance of PDs.

The response force of a PD under dynamic excitation holds hysteretic behavior. The hysteresis loops vary with excitation frequency due to the nonlinear nature of particle damping. By comparing two different types of particles, namely, 0.2 mm tungsten powder and 1.5 mm stainless steel balls, it is found that the lightweight particle bed is more easily to be excited. Therefore, the hysteresis loops of stainless-steel balls exhibit more nonlinear effects.

Various types of particles were evaluated on impact hammers. It is found that the granular material outperforms the powders, solid balls, and liquids. The primary effect of the developed RPD is to mitigate the peak of the track's FRF. However, for the frequency range that is away from the vibration mode, the effect of the RPD is limited. Therefore, it can be concluded that only when the primary structure has a considerable vibration, the RPDs can show its efficacy since the performance of RPDs is excitation-induced. Naturally, mounting RPDs on the track adds mass to the entire system, so the vibration mode moves toward a lower frequency. The modal analysis on the damper cavity reveals that the cavity performs a bouncing vibration relative to the rail track at the effective mode. This relative vibration between RPD and the track consumes the vibration energy of the system.

For the field application of RPDs, the installation of RPDs can effectively reduce the vibration level of the rail track, except for the vibration caused by the impact excitation from the passing wheels. The vibration wave propagating on the rail track is attenuated with the help of the periodically installed RPDs. The TDR of the track is elevated after the installation of RPDs. The main effective frequency range is the range higher than 1000 Hz. Since the vibration propagation is suppressed on the rail, the noise emission of the track is consequently reduced.

Then, other conclusions can be made through the modeling attempts with the DL method. First of all, introducing TL can solve the dilemma of a lack of training resources. With a similar source task, the general knowledge learned in the low-level layers in a DNN can be leveraged to facilitate learning in another target task. Compared with the plain NN that trains from scratch, the NN with TL receives a higher accuracy and convergence speed.

When modeling the response force of a PD, the inserted Fourier features layer can enhance the ability of a NN to recognize the high-frequency features. As the NNs are perplexed by the long-standing pathology called spectra bias, their performance is significantly restricted. Based on the analysis of NNs from the perspective of NTK, it is revealed that NNs can identify more high-frequency features with the imposition of the Fourier layer. Combining this approach with TL-incorporated PINN, the proposed method can reconstruct the hysteresis loops of a PD. This surrogate model is validated with a dataset obtained from an experiment that applied different forms of excitation. It proves that the established model has a good generalization ability.

From the experiments, the obtained dataset is a group of time series data. However, the aforementioned NNs ignored the sequential information. Here, the RNN-based ESN is utilized to decode the sequential information resides in the experimental results. The proposed ESN is enhanced with deep fully connected layers to formulate the ESDN. The MDSC is also placed in front of ESDN to facilitate feature extracting. It is found that on the training dataset, the mc-ESDN outperforms a conventional NN. However, for another dataset with the different excitation form, the mc-ESDN has poorly generalization ability compared to the baseline methods.

7.2 Recommendations

This study has laid a good foundation for investigating granular material-filled PD and its application in railway systems. However, this research still comes with blemishes. Therefore, further studies could be carried out regarding the following aspects.

1. The most obvious limitation of this work is that the investigation of the PD focused on the vertical direction. In a straight rail line, the vibration is mainly concentrated on the vertical direction. In the curvature section of the rail line however, lateral vibration would play a more significant role in the rail track vibration. The dominant noise type will no longer be the rolling noise in this case, but the squeal noise that generated from the friction between wheel and track instead. It is doubtful whether PDs are effective in mitigating such noise. Introducing the lateral vibration to the PD will significantly

change the system. In this case, the vibration mode of the particle bed will be inevitably influenced the vibration in another direction, and more complex nonlinear phenomenon will be generated. How to combine the vibration in two directions could also be a problem. It is difficult to decided what forms and amplitudes of excitations should be applied in each direction to simulate the railway vibration.

2. The simulation of the track vibration brings another problem. Currently, the modeling of PDs is carried out based on the datasets obtained from simple harmonic excitation experiments. Although the established model showed certain ability of generalization, it is still uncertain if these models can work in the real vibration of rail track. It should be noted that the vibration of tracks varies a lot with different rail lines, and always comes with a complex frequency component. The current models can serve as low-fidelity knowledge to this problem by following the idea of TL. Feeding the actual track vibration into the NN model is still a challenge.
3. The combination of surrogate and traditional models for the primary structure remains at an infant level. This study indirectly used the output of the surrogate model of PD as a reference to facilitate the analysis of the rail track vibration. There is currently no feasible way to unite the DL models into traditional theoretical calculations. The appearance of PINN may have provided the potential, but more endeavor is still expected to settle this issue.
4. The data-driven DL reconstructs the target function through point-wise estimation, which means the experimental data was directly input into the model for training

without an analysis of the confidence level of experimental results. However, the experimental dataset always comes with errors. These errors will become systematic errors in the training of surrogate models, consequently deteriorating performance. To this end, introducing Bayesian methods could alleviate the cause of random error in experiments. This approach embeds the distribution information of the dataset and reconstructs the target function in a probability-based manner. As discussed in CHAPTER 4, the input data should be modified to accommodate the Bayesian neural network structure. Taking the phase of the input signal to identify data of identical states could be considered as a solution.

5. In our area of interest, the vibration frequency far exceeds that of typical structures, reaching into the thousands of hertz. In these high-frequency vibrations, the displacement is minimal compared to other scenarios. In traditional application of particle dampers, energy consumption is primarily due to impact and collision effects. However, in high-frequency vibrations, friction between particles becomes the dominant factor in energy dissipation. This presents a challenge for particle dampers to achieve optimal performance. It indicates the need for future improvements in this technology. One potential improvement could involve developing new materials with enhanced friction properties to enable more efficient energy dissipation at higher frequencies. Additionally, optimizing the design of the particle damper itself, such as adjusting the size, shape, and arrangement of the particles, could better suit high-frequency vibrations. Another potential enhancement could involve integrating

advanced control systems to actively adjust the damping properties of the particle damper in real-time, ensuring optimal performance across a wide range of frequencies. Overall, while high-frequency vibrations pose a challenge for traditional particle dampers, there are opportunities for innovation and improvement to enhance their performance in these conditions.

Appendix A

In this illustrative example, a damped SDOF vibrator under the forced vibration is considered. This simple case is driven by the following ordinary differential equation (ODE):

$$m\ddot{x} + c\dot{x} + kx = F\sin(wt) \qquad \text{(A-1)}$$

With $m = 1\ kg,\ k = 200\ {}^{N}/_{m}$, damping ratio $\xi = 0.1$, $F = 100\ N, w = 120\ rad/s$. The initial condition is $x_0 = 0\ m, \dot{x}_0 = 0\ m/s$. According to these settings, the nature frequency of this vibrator is approximately 2.3 Hz, and the excitation frequency is roughly 19.1 Hz. Here, the physics-informed neural network (PINN) to pursue the solution to this ODE is adopted. As an emerging paradigm that promises profound influence on computational physics, PINNs have shown seductive efficacy in various disciplines. However, PINN also suffers from spectra bias in the framework of NNs.

To present a brief overview of the PINN, Equation (A-1) can be reformed by the general form:

$$\mathcal{N}[\boldsymbol{x}](\boldsymbol{t}) = \boldsymbol{f}(\boldsymbol{t}), \qquad \boldsymbol{t} \in \Omega \qquad \text{(A-2)}$$

$$\mathcal{B}[\boldsymbol{x}](\boldsymbol{t}) = \boldsymbol{g}(\boldsymbol{t}), \qquad \boldsymbol{t} \in \partial\Omega \qquad \text{(A-3)}$$

where $\mathcal{N}[\cdot]$ is the differential operator and $\mathcal{B}[\cdot]$ denotes the boundary condition, or initial condition. To seek the latent solution $x(t)$, a network with the input of t and output of x (see Figure A-1) is built. The derivative of x with respect to t can be obtained through auto differentiation (Güne¸ et al. 2018). Then, the PINN can be trained by minimizing the physics-induced residual loss and boundary/initial condition loss.

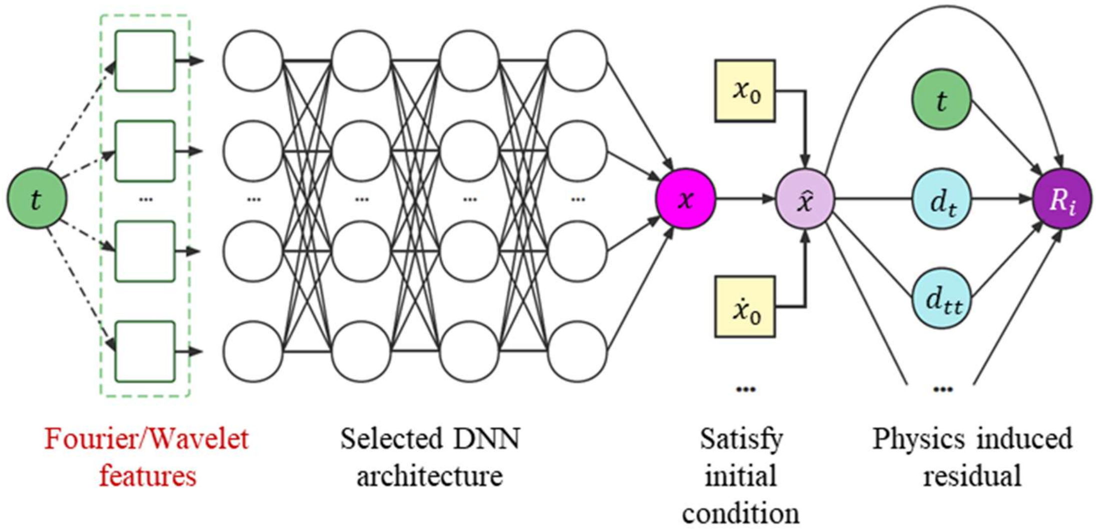


Figure A-1 Structure of Fourier/Wavelet features embedded PINN.

The calculation of this ODE is in the range of $t \in [0,3]$. The *ff/wf*PINN to solve the ODE is adopted, where Fourier/Wavelet features are embedded in the PINN. A vanilla PINN with the same architecture is also employed for comparison. The Gaussian basis function is selected for the *wf*PINN, both *ff*PINN and *wf*PINN are initialized by $\sigma = 10$. The fully connected network contains 3 hidden layers, 100 units per hidden layer, and *tanh* activation function is chosen in the network.

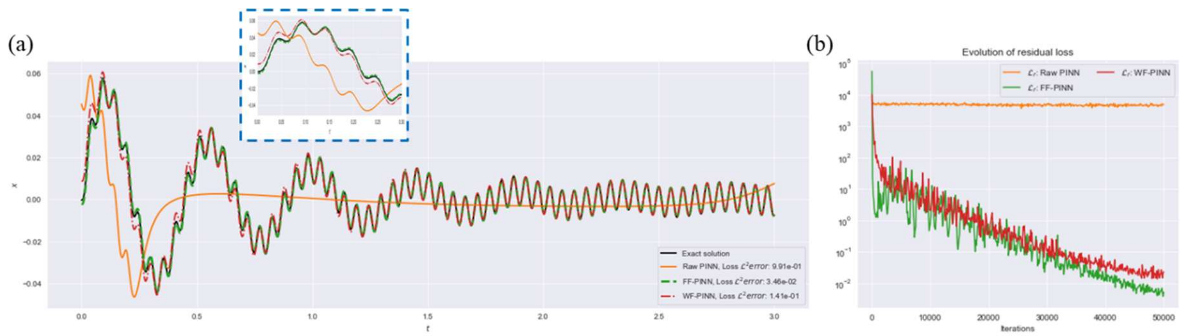


Figure A-2 (a) Calculation results of different PINNs after 50,000 iterations; (b) Evolution of the residual loss of different PINNs.

The raw PINN fails in capturing the solution to this ODE, as its residual loss barely has no descending in the training. On the contrary, *ff*PINN and *wf*PINN are capable to reconstruct the relatively high-frequency components in the objective function. Besides, the *ff*PINN is also outperforms the *wf*PINN in this case.

Appendix B

Following the derivation given by Xiao et al. (2013). For a Euler- Bernoulli beam, consider its dynamic stiffness matrix as:

$$\mathbf{S}_{track} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} \tag{B-1}$$

and its entrence is written as:

$$\boldsymbol{S}_{11} = \frac{\boldsymbol{EI}}{\boldsymbol{L}^3}\begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\gamma L} \\ \boldsymbol{\gamma L} & \boldsymbol{\beta L}^2 \end{bmatrix}, \boldsymbol{S}_{12} = \boldsymbol{S}_{21}^T = \frac{\boldsymbol{EI}}{\boldsymbol{L}^3}\begin{bmatrix} -\overline{\boldsymbol{\alpha}} & \overline{\boldsymbol{\gamma}}\boldsymbol{L} \\ -\overline{\boldsymbol{\gamma}}\boldsymbol{L} & \overline{\boldsymbol{\beta}}\boldsymbol{L}^2 \end{bmatrix}, \boldsymbol{S}_{22} = \frac{\boldsymbol{EI}}{\boldsymbol{L}^3}\begin{bmatrix} \boldsymbol{\alpha} & -\boldsymbol{\gamma L} \\ -\boldsymbol{\gamma L} & \boldsymbol{\beta L}^2 \end{bmatrix} \tag{B-2}$$

with

$$\begin{gathered}
\boldsymbol{\alpha} = \frac{(\cos \boldsymbol{\Lambda} \sinh \boldsymbol{\Lambda} + \sin \boldsymbol{\Lambda} \cosh \boldsymbol{\Lambda})\boldsymbol{\Lambda}^3}{\boldsymbol{\Delta}}; \\
\overline{\boldsymbol{\alpha}} = \frac{(\sin \boldsymbol{\Lambda} + \sinh \boldsymbol{\Lambda})\boldsymbol{\Lambda}^3}{\boldsymbol{\Delta}}; \\
\boldsymbol{\beta} = \frac{(-\cos \boldsymbol{\Lambda} \sinh \boldsymbol{\Lambda} + \sin \boldsymbol{\Lambda} \cosh \boldsymbol{\Lambda})\boldsymbol{\Lambda}}{\boldsymbol{\Delta}}; \\
\overline{\boldsymbol{\beta}} = \frac{(-\sin \boldsymbol{\Lambda} + \sinh \boldsymbol{\Lambda})\boldsymbol{\Lambda}}{\boldsymbol{\Delta}}; \\
\boldsymbol{\gamma} = \frac{(\sin \boldsymbol{\Lambda} \sinh \boldsymbol{\Lambda})\boldsymbol{\Lambda}^2}{\boldsymbol{\Delta}}; \\
\overline{\boldsymbol{\gamma}} = \frac{(-\cos \boldsymbol{\Lambda} + \cosh \boldsymbol{\Lambda})\boldsymbol{\Lambda}^2}{\boldsymbol{\Delta}}; \\
\boldsymbol{\Lambda} = \boldsymbol{k_b L}; \boldsymbol{\Delta} = \mathbf{1} - \cos \boldsymbol{\Lambda} \cosh \boldsymbol{\Lambda}
\end{gathered} \tag{B-3}$$

In this definition, $k_b = \left(\frac{m_r \omega^2}{EI}\right)^{\frac{1}{4}}$ is the wave number solution of a rail track, and m_r is the weight of a unit length rail track. If we substitute the left and right sides of a rail track unit on the periodic system into the above equations, we can have the governing equation of the track unit:

$$\begin{bmatrix} \boldsymbol{S_{LL}} & \boldsymbol{S_{LR}} \\ \boldsymbol{S_{RL}} & \boldsymbol{S_{RR}} \end{bmatrix} \begin{Bmatrix} \boldsymbol{u_L} \\ \boldsymbol{u_R} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{f_L} \\ \boldsymbol{f_R} \end{Bmatrix} \tag{B-4}$$

where $\mathbf{u}_{\mathrm{L}} = \{w_L \quad \theta_L\}^{\mathrm{T}}, \mathbf{f}_{\mathrm{L}} = \{V_L \quad M_L\}^{\mathrm{T}}$ denote the displacement and external force of this track unit. Consider the wave decay on this track unit:

$$\mathbf{u}_{\mathrm{R}} = e^{\mu}\mathbf{u}_{\mathrm{L}}, \mathbf{f}_{\mathrm{R}} = -e^{\mu}\mathbf{f}_{\mathrm{L}} \tag{B-5}$$

this gives the relationship between force and displacement on the left side:

$$\mathbf{f}_{\mathrm{L}} = (\boldsymbol{S}_{\mathrm{LL}} + e^{\mu}\boldsymbol{S}_{\mathrm{LR}})\mathbf{u}_{\mathrm{L}} \tag{B-6}$$

Substituting this relationship into above equations, we can have the similar form as Equation (6-14). The difference between this calculation and the derivation given in CHAPTER 6 is that, rather than the dynamic stiffness matrix, the wave propagation on the rail track structure is defined by the transfer matrix $\boldsymbol{T}$. However, the calculation of wave propagation constant is all based on the non-trivial solution of the equilibrium equation.

REFERENCES

Abbas, H., Hai, H., Rongong, J., & Xing, Y. (2014). Damping performance of metal swarfs in a horizontal hollow structure. *Journal of Mechanical Science and Technology*, *28*(1), 9–13. https://doi.org/10.1007/s12206-013-0980-3

Albdour, N., & Zanoon, N. (2020). A Steganographic Method Based on Roberts Operator. *Jordan Journal of Electrical Engineering*, *6*(3). https://doi.org/10.5455/jjee.204-1583873433

Assirati, L., Silva, N. R., Berton, L., Lopes, A. A., & Bruno, O. M. (2014). Performing edge detection by Difference of Gaussians using q-Gaussian kernels. *Journal of Physics: Conference Series*, *490*(1). https://doi.org/10.1088/1742-6596/490/1/012020

Barone, S. R., Narcowich, M. A., & Narcowich, F. J. (1977). Floquet theory and applications. *Physical Review A*, *15*(3), 1109. https://doi.org/10.1103/PhysRevA.15.1109

Basri, R., Jacobs, D., Kasten, Y., & Kritchman, S. (2019). The Convergence Rate of Neural Networks for Learned Functions of Different Frequencies. *Advances in Neural Information Processing Systems*, *32*. https://doi.org/10.48550/arxiv.1906.00425

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning Long-Term Dependencies with Gradient Descent is Difficult. *IEEE Transactions on Neural Networks*, *5*(2), 157–166. https://doi.org/10.1109/72.279181

Bottou, L. (2012). Stochastic gradient descent tricks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in*

Bioinformatics), *7700 LECTURE NO*, 421–436. https://doi.org/10.1007/978-3-642-35289-8_25/COVER/

Boussollaa, C., Debut, V., Antunes, J., Fakhfakh, T., & Haddar, M. (2023). Dynamics of structures coupled to particle-chain impact dampers: A parametric numerical analysis. *Https://Doi.Org/10.1177/09544062231207340*. https://doi.org/10.1177/09544062231207340

Bustamante, M., Gerges, S. N. Y., Cordioli, J., Martin, O. P., Weisbeck, J., & Ott, M. (2013). Experimental study on some parameters that affect the performance of an elastomer particle damper. *Proceedings of Meetings on Acoustics*, *19*(1), 65020. https://doi.org/10.1121/1.4801066

Bustamante, M., Gerges, S. N. Y., Vergara, E. F., & Arenas, J. P. (2016). High damping characteristics of an elastomer particle damper. *International Journal of Acoustics and Vibrations*, *21*(1), 112–121. https://doi.org/10.20855/ijav.2016.21.1401

Cao, Y., Fang, Z., Wu, Y., Zhou, D. X. X., & Gu, Q. (2019). Towards Understanding the Spectral Bias of Deep Learning. *IJCAI International Joint Conference on Artificial Intelligence*, 2205–2211. https://doi.org/10.48550/arxiv.1912.01198

Chakraborty, S. (2021a). Transfer learning based multi-fidelity physics informed deep neural network. *Journal of Computational Physics*, *426*. https://doi.org/10.1016/j.jcp.2020.109942

Chakraborty, S. (2021b). Transfer learning based multi-fidelity physics informed deep neural network. *Journal of Computational Physics*, *426*, 109942.

https://doi.org/10.1016/j.jcp.2020.109942

Chen, J., & Georgakis, C. T. (2013). Tuned rolling-ball dampers for vibration control in wind turbines. *Journal of Sound and Vibration*, *332*(21), 5271–5282. https://doi.org/10.1016/j.jsv.2013.05.019

Chizat, L., Oyallon, E., & Bach, F. (2018). On Lazy Training in Differentiable Programming. *Advances in Neural Information Processing Systems*, *32*. https://doi.org/10.48550/arxiv.1812.07956

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*. https://doi.org/10.3115/v1/d14-1179

Chung, H., Fox, C., Dodd, G., & Emms, G. (2010). Lightweight Floor/Ceiling Systems with Improved Impact Sound Insulation. *Journal of Vibroengineering*, *17*(2), 129–141. https://doi.org/https://doi.org/10.3813/AAA.918742

Darabi, B., & Rongong, J. A. (2012). Polymeric particle dampers under steady-state vertical vibrations. *Journal of Sound and Vibration*, *331*(14), 3304–3316. https://doi.org/10.1016/j.jsv.2012.03.005

Djemal, F., Chaari, R., Gafsi, W., Chaari, F., & Haddar, M. (2019). Passive vibration suppression using ball impact damper absorber. *Applied Acoustics*, *147*, 72–76. https://doi.org/10.1016/J.APACOUST.2017.09.011

Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., & Darrell, T. (2013). DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. *31st International Conference on Machine Learning, ICML 2014*, *2*, 988–996.

Du, Y., & Wang, S. (2010). Modeling the fine particle impact damper. *International Journal of Mechanical Sciences*, *52*(7), 1015–1022. https://doi.org/10.1016/j.ijmecsci.2010.04.004

Duvigneau, F., Koch, S., Woschke, E., & Gabbert, U. (2016). An effective vibration reduction concept for automotive applications based on granular-filled cavities. *Https://Doi.Org/10.1177/1077546316632932*, *24*(1), 73–82. https://doi.org/10.1177/1077546316632932

Eberhard, P., Gnanasambandham, C., Gnanasambandham, C., & Eberhard, P. (2019). Modelling a partially liquid-filled particle damper using coupled lagrangian methods. *PARTICLES VI : Proceedings of the VI International Conference on Particle-Based Methods : Fundamentals and Applications*, 744–753. https://upcommons.upc.edu/handle/2117/186881

Elboushaki, A., Hannane, R., Afdel, K., & Koutti, L. (2020). MultiD-CNN: A multi-dimensional feature learning approach based on deep convolutional networks for gesture recognition in RGB-D image sequences. *Expert Systems with Applications*, *139*. https://doi.org/10.1016/j.eswa.2019.112829

Fan, L.-S., & Zhu, C. (1998). Principles of Gas-Solid Flows. In *Principles of Gas-Solid Flows*. Cambridge University Press. https://doi.org/10.1017/cbo9780511530142

Ferreyra, M. V., Baldini, M., Pugnaloni, L. A., & Job, S. (2020). Effect of lateral confinement on the apparent mass of particle dampers. *Granular Matter*, *23*(2). https://doi.org/10.1007/s10035-021-01090-w

Gagnon, L., Morandini, M., & Ghiringhelli, G. L. (2019). A review of particle damping modeling and testing. *Journal of Sound and Vibration*, *459*, 114865. https://doi.org/10.1016/j.jsv.2019.114865

Gharib, M., & Ghani, S. (2013). Free vibration analysis of linear particle chain impact damper. *Journal of Sound and Vibration*, *332*(24), 6254–6264. https://doi.org/10.1016/j.jsv.2013.07.013

Gopal, S., Patro, K., & Kumar Sahu, K. (2015). Normalization: A Preprocessing Stage. *IARJSET*, 20–22. https://doi.org/10.17148/iarjset.2015.2305

Gorla, F. D., & Nicoletti, R. (2021). Dampening of a cantilever beam with large particles in a small cavity: model and experiment. *Archive of Applied Mechanics*, *91*(7), 2933–2942. https://doi.org/10.1007/S00419-021-01946-W/FIGURES/6

Güne¸, A., Baydin, G., Pearlmutter, B. A., & Siskind, J. M. (2018). Automatic Differentiation in Machine Learning: a Survey. In *Journal of Machine Learning Research* (Vol. 18).

Harduf, Y., Setter, E., Feldman, M., & Bucher, I. (2023). Modeling additively-manufactured particle dampers as a 2DOF frictional system. *Mechanical Systems and Signal Processing*, *187*, 109928. https://doi.org/10.1016/J.YMSSP.2022.109928

He, Y., Zhang, L., Chen, Z., & Li, C. Y. (2022). A framework of structural damage detection for civil structures using a combined multi-scale convolutional neural network and echo state network. *Engineering with Computers*. https://doi.org/10.1007/s00366-021-01584-4

Henderi, H., Wahyuningsih, T., & Rahwanto, E. (2021). Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer. *International Journal of Informatics and Information Systems*, *4*(1), 13–20. https://doi.org/10.47738/IJIIS.V4I1.73

Ho, K. M., Cheng, C. K., Yang, Z., Zhang, X. X., & Sheng, P. (2003). Broadband locally resonant sonic shields. *Applied Physics Letters*, *83*(26), 5566–5568. https://doi.org/10.1063/1.1637152

Hu, Y., Zan, H., Guo, Y., Jiang, J., Xia, Z., Wen, H., & Peng, Z. (2023). Energy dissipation characteristics of particle dampers with obstacle grids. *Mechanical Systems and Signal Processing*, *193*. https://doi.org/10.1016/j.ymssp.2023.110231

Huang, X. hong, Xu, W. bing, Wang, J., Yan, W. ming, & Chen, Y. jiang. (2021). Equivalent model of a multi-particle damper considering particle rolling and its analytical solution. *Structural Control and Health Monitoring*, *28*(6), e2718. https://doi.org/10.1002/STC.2718

Ibrahim, R. A. (2009). *Vibro-Impact Dynamics* (Vol. 43). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-00275-5

Jacot, A., Gabriel, F., & Hongler, C. (2018). Neural Tangent Kernel: Convergence and Generalization in Neural Networks. *Advances in Neural Information Processing Systems*, 8571–8580. https://doi.org/10.48550/arxiv.1806.07572

Jaeger, H. (2002). Short term memory in echo state networks. In *GMD Report 152*.

Jaeger, H., & Haas, H. (2004). Harnessing Nonlinearity: Predicting Chaotic Systems and

Saving Energy in Wireless Communication. *Science*, *304*(5667), 78–80. https://doi.org/10.1126/SCIENCE.1091277/SUPPL_FILE/JAEGER_SOM.PDF

Jin, G., Zhao, Z., Liu, B., Cun, W., Zhao, Z., Hou, M., & Chen, G. (2021). Design of a particle damper and experimental study on vibration damping of the pipeline. *Advances in Mechanical Engineering*, *13*(9), 1–14. https://doi.org/10.1177/16878140211044923/ASSET/IMAGES/LARGE/10.1177_16878140211044923-FIG2.JPEG

Jin, J., Kim, H., Koh, H. I., & Park, J. (2022). Railway noise reduction by periodic tuned particle impact damper with bounce and pitch-coupled vibration modes. *Composite Structures*, *284*, 115230. https://doi.org/10.1016/J.COMPSTRUCT.2022.115230

Jin, J., Yang, W., Koh, H. I., & Park, J. (2020). Development of tuned particle impact damper for reduction of transient railway vibrations. *Applied Acoustics*, *169*, 107487. https://doi.org/10.1016/j.apacoust.2020.107487

Jin, N., Liu, D., Pang, Z., & Huang, T. (2007). Wavelet basis function neural networks. *IEEE International Conference on Neural Networks - Conference Proceedings*, 500–505. https://doi.org/10.1109/IJCNN.2007.4371007

Jung, Y. (2017). Multiple predicting K-fold cross-validation for model selection. *Https://Doi.Org/10.1080/10485252.2017.1404598*, *30*(1), 197–215. https://doi.org/10.1080/10485252.2017.1404598

Ketkar, N. (2017). Stochastic Gradient Descent. *Deep Learning with Python*, 113–132. https://doi.org/10.1007/978-1-4842-2766-4_8

Kingma, D. P., & Ba, J. L. (2014). Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://arxiv.org/abs/1412.6980v9

Kollmer, J. E., Sack, A., Heckel, M., & Pöschel, T. (2013). Relaxation of a spring with an attached granular damper. *New Journal of Physics*, *15*, 93023–93032. https://doi.org/10.1088/1367-2630/15/9/093023

Lee, J., Xiao, L., Schoenholz, S. S., Bahri, Y., Novak, R., Sohl-Dickstein, J., Pennington, J., & Brain, G. (n.d.). *Wide Neural Networks of Any Depth Evolve as Linear Models Under Gradient Descent*. https://g.co/airesidency

Lei, X., & Wu, C. (2017). Non-obstructive particle damping using principles of gas-solid flows. *Journal of Mechanical Science and Technology*, *31*(3), 1057–1065. https://doi.org/10.1007/s12206-017-0204-3

Lei, X., Wu, C., & Chen, P. (2018). Optimizing parameter of particle damping based on Leidenfrost effect of particle flows. *Mechanical Systems and Signal Processing*, *104*, 60–71. https://doi.org/10.1016/j.ymssp.2017.10.037

Lei, X., Wu, C., & Wu, H. (2018). A novel composite vibration control method using double-decked floating raft isolation system and particle damper. *JVC/Journal of Vibration and Control*, *24*(19), 4407–4418. https://doi.org/10.1177/1077546317724967

Li, K., & Darby, A. P. (2008). A buffered impact damper for multi-degree-of-freedom structural control. *Earthquake Engineering and Structural Dynamics*, *37*(13), 1491–1510. https://doi.org/10.1002/eqe.823

Li, M., Zhang, T., Chen, Y., & Smola, A. J. (2014). Efficient mini-batch training for stochastic optimization. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 661–670. https://doi.org/10.1145/2623330.2623612

Li, X.-A., Xu, Z.-Q. J. J., & Zhang, L. (2020). A multi-scale DNN algorithm for nonlinear elliptic equations with multiple scales. *Communications in Computational Physics*, *28*(5), 1886–1906. https://doi.org/10.4208/cicp.OA-2020-0187

Liao, C. C., Chung, Y. C., & Weng, C. H. (2023). A study on the energy dissipation mechanism of dynamic mechanical systems with particle dampers by using the novel energy method. *Nonlinear Dynamics*, *111*(17), 15955–15980. https://doi.org/10.1007/S11071-023-08698-Y/FIGURES/20

Lieber, P., and D. P. J. (1945). An acceleration damper: Development, design and some applications. *Transactions of the ASME*, *67*, 523–530.

Liu, S., Li, H., Zhang, F., Lu, Z., Yang, S., & Li, P. (2023). The optimum design of particle tuned mass damper for structural seismic control considering soil-structure interaction. *JVC/Journal of Vibration and Control*. https://doi.org/10.1177/10775463231164198/ASSET/IMAGES/LARGE/10.1177_10775463231164198-FIG9.JPEG

Liu, W., Tomlinson, G. R., & Rongong, J. A. (2005). The dynamic characterisation of disk geometry particle dampers. *Journal of Sound and Vibration*, *280*(3–5), 849–861. https://doi.org/10.1016/J.JSV.2003.12.047

Liu, Z., Cai, W., & Xu, Z.-Q. J. J. (2020). Multi-scale Deep Neural Network (MscaleDNN) for

Solving Poisson-Boltzmann Equation in Complex Domains. *Communications in Computational Physics*, *28*(5), 1970–2001. https://doi.org/10.4208/cicp.OA-2020-0179

Long, M., Cao, Y., Wang, J., & Jordan, M. I. (2015). Learning Transferable Features with Deep Adaptation Networks. *32nd International Conference on Machine Learning, ICML 2015*, *1*, 97–105. https://doi.org/10.48550/arxiv.1502.02791

Lu, Z. (2020). *Particle Damping Technology Based Structural Control*.

Lu, Z., Chen, X., Zhang, D., & Dai, K. (2017a). Experimental and analytical study on the performance of particle tuned mass dampers under seismic excitation. *Earthquake Engineering and Structural Dynamics*, *46*(5), 697–714. https://doi.org/10.1002/eqe.2826

Lu, Z., Chen, X., Zhang, D., & Dai, K. (2017b). Experimental and analytical study on the performance of particle tuned mass dampers under seismic excitation. *Earthquake Engineering and Structural Dynamics*, *46*(5), 697–714. https://doi.org/10.1002/eqe.2826

Lu, Z., Chen, X., & Zhou, Y. (2018). An equivalent method for optimization of particle tuned mass damper based on experimental parametric study. *Journal of Sound and Vibration*, *419*, 571–584. https://doi.org/10.1016/J.JSV.2017.05.048

Lu, Z., Lu, X., Lu, W., & Masri, S. F. (2012). Experimental studies of the effects of buffered particle dampers attached to a multi-degree-of-freedom system under dynamic loads. *Journal of Sound and Vibration*, *331*(9), 2007–2022. https://doi.org/10.1016/j.jsv.2011.12.022

Lu, Z., Zhang, J., & Wang, D. (2021). Energy analysis of particle tuned mass damper systems with applications to MDOF structures under wind-induced excitation. *Journal of Wind

Engineering and Industrial Aerodynamics, *218*, 104766. https://doi.org/10.1016/J.JWEIA.2021.104766

Lu, Z., Zhang, Q., Fan, Q., & Li, Q. (2023). Studies on dissipative characteristics and equivalent model of particle damper in railway application. *Journal of Sound and Vibration*, *560*. https://doi.org/10.1016/j.jsv.2023.117788

Lukoševičius, M., & Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, *3*(3), 127–149. https://doi.org/10.1016/J.COSREV.2009.03.005

Masmoudi, M., Job, S., Abbes, M. S., Tawfiq, I., & Haddar, M. (2016). Experimental and numerical investigations of dissipation mechanisms in particle dampers. *Granular Matter*, *18*(3), 1–11. https://doi.org/10.1007/s10035-016-0667-4

Masri, S. F. (1969). Analytical and Experimental Studies of Multiple-Unit Impact Dampers. *The Journal of the Acoustical Society of America*, *45*(5), 1111–1117. https://doi.org/10.1121/1.1911581

Masri, S. F., & Caffrey, J. P. (2019). Response of pounding dynamic vibration neutralizer under harmonic and random excitation. *Journal of Applied Mechanics, Transactions ASME*, *86*(2), 1–15. https://doi.org/10.1115/1.4041910

Meyer, N., & Seifried, R. (2020). Numerical and experimental investigations in the damping behavior of particle dampers attached to a vibrating structure. *Computers & Structures*, *238*, 106281. https://doi.org/10.1016/J.COMPSTRUC.2020.106281

Meyer, N., & Seifried, R. (2022). Energy dissipation in horizontally driven particle dampers of low acceleration intensities. *Nonlinear Dynamics*. https://doi.org/10.1007/s11071-022-07348-z

Michon, G., Almajid, A., & Aridon, G. (2013). Soft hollow particle damping identification in honeycomb structures. *Journal of Sound and Vibration*, *332*(3), 536–544. https://doi.org/10.1016/j.jsv.2012.09.024

Moseley, B., Markham, A., & Nissen-Meyer, T. (2021). *Finite Basis Physics-Informed Neural Networks (FBPINNs): a scalable domain decomposition approach for solving differential equations*. https://doi.org/10.48550/arxiv.2107.07871

Nayeri, R. D., Masri, S. F., & Caffrey, J. P. (2007). Studies of the performance of multi-unit impact dampers under stochastic excitation. *Journal of Vibration and Acoustics, Transactions of the ASME*, *129*(2), 239–251. https://doi.org/10.1115/1.2346694

Olędzki, A. A., Siwicki, I., & Wiśniewski, J. (1999). Impact dampers in application for tube, rod and rope structures. *Mechanism and Machine Theory*, *34*(2), 243–253. https://doi.org/10.1016/s0094-114x(98)00014-7

Olson, S. E. (2003). An analytical particle damping model. *Journal of Sound and Vibration*, *264*(5), 1155–1166. https://doi.org/10.1016/S0022-460X(02)01388-3

Oltmann, J., Hartwich, T., & Krause, D. (2020). Optimizing lightweight structures with particle damping using frequency based substructuring. *Design Science*, 1–24. https://doi.org/10.1017/dsj.2020.13

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Papalou, A. (2023). An Experimental Study on the Effectiveness of Particle Tuned Mass Dampers. *KSCE Journal of Civil Engineering*, *27*(10), 4437–4443. https://doi.org/10.1007/S12205-023-0261-3/METRICS

Petrone, G., D'Alessandro, V., Franco, F., & De Rosa, S. (2015). Damping evaluation on eco-friendly sandwich panels through reverberation time (RT 60) measurements. *JVC/Journal of Vibration and Control*, *21*(16), 3328–3338. https://doi.org/10.1177/1077546314522507

Popplewell, N., & Semercigil, S. E. (1989). Performance of the bean bag impact damper for a sinusoidal external force. *Journal of Sound and Vibration*, *133*(2), 193–223. https://doi.org/10.1016/0022-460X(89)90922-X

Qi, X., & Liu, X. (2020). Evolution of the first eigenvalue of the Laplace operator and the p-Laplace operator under a forced mean curvature flow. *Open Mathematics*, *18*(1). https://doi.org/10.1515/math-2020-0090

Rahaman, N., Baratin, A., Arpit, D., Draxlcr, F., Lin, M., Hamprecht, F. A., Bengio, Y., & Courville, A. (2018). On the Spectral Bias of Neural Networks. *36th International Conference on Machine Learning, ICML 2019*, *2019-June*, 9230–9239. https://doi.org/10.48550/arxiv.1806.08734

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019a). 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. https://doi.org/10.1016/j.jcp.2018.10.045

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019b). 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. https://doi.org/10.1016/j.jcp.2018.10.045

Sack, A., Windows-Yule, K., Heckel, M., Werner, D., & Pöschel, T. (2020). Granular dampers in microgravity: sharp transition between modes of operation. *Granular Matter*, *22*(2), 1–6. https://doi.org/10.1007/S10035-020-01017-X/FIGURES/4

Saeki, M. (2005). Analytical study of multi-particle damping. *Journal of Sound and Vibration*, *281*(3–5), 1133–1144. https://doi.org/10.1016/j.jsv.2004.02.034

Salehinejad, H., Sankar, S., Barfett, J., Colak, E., & Valaee, S. (2018). *Recent Advances in Recurrent Neural Networks*.

Sánchez, M., & Manuel Carlevaro, C. (2013). Nonlinear dynamic analysis of an optimal particle damper. *Journal of Sound and Vibration*, *332*(8), 2070–2080. https://doi.org/10.1016/j.jsv.2012.09.042

Sánchez, M., Rosenthal, G., & Pugnaloni, L. A. (2012). Universal response of optimal granular damping devices. *Journal of Sound and Vibration*, *331*(20), 4389–4394. https://doi.org/10.1016/j.jsv.2012.05.001

Sarpkaya, T. (1986). Force on a circular cylinder in viscous oscillatory flow at low keulegan—carpenter numbers. *Journal of Fluid Mechanics*, *165*, 61–71. https://doi.org/10.1017/S0022112086002999

Seyed, M.S.A., (2019). Investigation Report on The Performance of Rail Particle Impact Damper (Rail-PID) Through Impact Hammer Test. *CNERC-Rail Internal Technical*

Report.

Shah, B. M., Nudell, J. J., Kao, K. R., Keer, L. M., Jane Wang, Q., & Zhou, K. (2011). Semi-active particle-based damping systems controlled by magnetic fields. *Journal of Sound and Vibration*, *330*(2), 182–193. https://doi.org/10.1016/j.jsv.2010.08.009

Shah, B. M., Pillet, D., Bai, X. M., Keer, L. M., Jane Wang, Q., & Snurr, R. Q. (2009). Construction and characterization of a particle-based thrust damping system. *Journal of Sound and Vibration*, *326*(3–5), 489–502. https://doi.org/10.1016/j.jsv.2009.06.007

Snoun, C., & Trigui, M. (2018). Design parameters optimization of a particles impact damper. *International Journal on Interactive Design and Manufacturing*, *12*(4), 1283–1297. https://doi.org/10.1007/s12008-018-0463-y

Song, Y., Ma, B., Gao, W., & Fan, S. (2019). Medical image edge detection based on improved differential evolution algorithm and prewitt operator. *Acta Microscopica*, *28*(1).

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11141 LNCS*, 270–279. https://doi.org/10.1007/978-3-030-01424-7_27/FIGURES/5

Tancik, M., Srinivasan, P. P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T. T., & Ng, R. (2020). Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains. *Advances in Neural Information Processing Systems*. https://doi.org/10.48550/arxiv.2006.10739

Thompson, D. J. (2013). Railway noise and vibration mechanisms, modelling and means of control. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9).

Veeramuthuvel, P., Sairajan, K. K., & Shankar, K. (2016). Vibration suppression of printed circuit boards using an external particle damper. *Journal of Sound and Vibration*, *366*, 98–116. https://doi.org/10.1016/j.jsv.2015.12.034

Veeramuthuvel, P., Shankar, K., & Sairajan, K. K. (2017). Application of RBF neural network in prediction of particle damping parameters from experimental data. *JVC/Journal of Vibration and Control*, *23*(6), 909–929. https://doi.org/10.1177/1077546315587147

Wang, B., Wang, H., He, Z., Liu, D., & Pan, P. (2023). Test and analysis of multi-cavity particle damper for vertical vibration control of pipeline structures. *Engineering Structures*, *281*, 115744. https://doi.org/10.1016/J.ENGSTRUCT.2023.115744

Wang, C., Xiao, W., Wu, D., Wang, S., Lin, C., Luo, Y., Xiao, J., Yao, K., & Xu, Z. (2020). Study on bandgap characteristics of particle damping phononic crystal. *Applied Acoustics*, *166*, 107352. https://doi.org/10.1016/J.APACOUST.2020.107352

Wang, D. K., Wu, C. J., & Yang, R. C. (2015). Free Vibration of the Damping Beam Using Co-simulation Method Based on the MFT. *The International Journal of Acoustics and Vibration*, *20*(4), 251–257. https://doi.org/10.20855/ijav.2015.20.4388

Wang, D., & Wu, C. (2015). Parameter estimation and arrangement optimization of particle dampers on the cantilever rectangular plate. *Journal of Vibroengineering*, *17*(5), 2503–2520.

Wang, D., & Wu, C. (2016a). A novel prediction method of vibration and acoustic radiation for rectangular plate with particle dampers. *Journal of Mechanical Science and Technology*, *30*(3), 1021–1035. https://doi.org/10.1007/s12206-016-0205-7

Wang, D., & Wu, C. (2016b). A novel prediction method of vibration and acoustic radiation for rectangular plate with particle dampers. *Journal of Mechanical Science and Technology*, *30*(3), 1021–1035. https://doi.org/10.1007/s12206-016-0205-7

Wang, G., Guo, L., & Duan, H. (2013). Wavelet neural network using multiple wavelet functions in target threat assessment. *The Scientific World Journal*, *2013*. https://doi.org/10.1155/2013/632437

Wang, P., Yi, Q., Zhao, C., Xing, M., & Tang, J. (2017). Wave propagation in periodic track structures: band-gap behaviours and formation mechanisms. *Archive of Applied Mechanics*, *87*(3), 503–519. https://doi.org/10.1007/S00419-016-1207-8

Wang, R., Yan, F., Yu, L., Shen, C., Hu, X., & Chen, J. (2023). A federated transfer learning method with low-quality knowledge filtering and dynamic model aggregation for rolling bearing fault diagnosis. *Mechanical Systems and Signal Processing*, *198*, 110413. https://doi.org/10.1016/J.YMSSP.2023.110413

Wang, S., Teng, Y., & Perdikaris, P. (2021). Understanding and Mitigating Gradient Flow Pathologies in Physics-Informed Neural Networks. *Https://Doi.Org/10.1137/20M1318043*, *43*(5), 3055–3081. https://doi.org/10.1137/20M1318043

Wang, S., Wang, H., & Perdikaris, P. (2021). On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, *384*, 113938.

https://doi.org/10.1016/J.CMA.2021.113938

Wang, S., Yu, X., & Perdikaris, P. (n.d.). When and why PINNs fail to train: A neural tangent kernel perspective. *Journal of Computational Physics*, *449*.

Wang, X., Liu, X., Shan, Y., & He, T. (2015). *Design, simulation and experiment of particle dampers attached to a precision instrument in spacecraft*. *17*(4).

Wang, Y., Liu, B., Tian, A., & Tang, W. (2016). Experimental and numerical investigations on the performance of particle dampers attached to a primary structure undergoing free vibration in the horizontal and vertical directions. *Journal of Sound and Vibration*, *371*, 35–55. https://doi.org/10.1016/J.JSV.2016.01.056

Wang, Y., Liu, B., Tian, A., Wei, D., & Jiang, X. (2019). Prediction methods for the damping effect of multi-unit particle dampers based on the cyclic iterations of a single-unit particle damper. *Journal of Sound and Vibration*, *443*, 341–361. https://doi.org/10.1016/J.JSV.2018.10.035

Wang, Z., Wang, K., Yang, F., Pan, S., & Han, Y. (2018). Image segmentation of overlapping leaves based on Chan–Vese model and Sobel operator. *Information Processing in Agriculture*, *5*(1). https://doi.org/10.1016/j.inpa.2017.09.005

Wong, J. C., Ooi, C., Gupta, A., & Ong, Y.-S. (2021). *Learning in Sinusoidal Spaces with Physics-Informed Neural Networks*. https://doi.org/10.48550/arxiv.2109.09338

Wu, C. J., Liao, W. H., & Wang, M. Y. (2004). Modeling of granular particle damping using multiphase flow theory of gas-particle. *Journal of Vibration and Acoustics, Transactions of the ASME*, *126*(2), 196–201. https://doi.org/10.1115/1.1688763

Wu, T. X., & Thompson, D. J. (2002). A hybrid model for the noise generation due to railway wheel flats. *Journal of Sound and Vibration*, *251*(1). https://doi.org/10.1006/jsvi.2001.3980

Xiao, W., Huang, Y., Jiang, H., Lin, H., & Li, J. (2016). Energy dissipation mechanism and experiment of particle dampers for gear transmission under centrifugal loads. *Particuology*, *27*, 40–50. https://doi.org/10.1016/j.partic.2015.10.007

Xiao, W., Jin, L., & Chen, B. (2015). Theoretical Analysis and Experimental Verification of Particle Damper-Based Energy Dissipation with Applications to Reduce Structural Vibration. *Shock and Vibration*, *2015*. https://doi.org/10.1155/2015/186356

Xiao, Y., Wen, J., Yu, D., & Wen, X. (2013). Flexural wave propagation in beams with periodically attached vibration absorbers: Band-gap behavior and band formation mechanisms. *Journal of Sound and Vibration*, *332*(4). https://doi.org/10.1016/j.jsv.2012.09.035

Xu, Z.-Q. J., Zhang, Y., Luo, T., Xiao, Y., & Ma, Z. (2019). Frequency Principle: Fourier Analysis Sheds Light on Deep Neural Networks. *Communications in Computational Physics*, *28*(5), 1746–1767. https://doi.org/10.4208/cicp.OA-2020-0085

Yan, W., Wang, B., & He, H. (2020). Research on Damping Mechanism and Parameter Analysis of Particle Damper Based on Energy Theory. *Journal of Engineering Mechanics*, *146*(6), 04020054. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001772

Ye, X., Ni, Y.-Q., Sajjadi, M., Wang, Y.-W., & Lin, C.-S. (2022). Physics-guided, data-refined modeling of granular material-filled particle dampers by deep transfer learning.

Mechanical Systems and Signal Processing, *180*, 109437. https://doi.org/10.1016/J.YMSSP.2022.109437

Ye, Z., & Yu, J. (2021). Deep morphological convolutional network for feature learning of vibration signals and its applications to gearbox fault diagnosis. *Mechanical Systems and Signal Processing*, *161*, 107984. https://doi.org/10.1016/J.YMSSP.2021.107984

Yeo, K. (2019). Data-driven reconstruction of nonlinear dynamics from sparse observation. *Journal of Computational Physics*, *395*, 671–689. https://doi.org/10.1016/J.JCP.2019.06.039

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems*, *4*(January), 3320–3328. https://doi.org/10.48550/arxiv.1411.1792

Zhai, W., Wang, K., & Cai, C. (2009). Fundamentals of vehicle-track coupled dynamics. *Vehicle System Dynamics*, *47*(11). https://doi.org/10.1080/00423110802621561

Zhang, K., Xi, Y., Chen, T., & Ma, Z. (2018). Experimental studies of tuned particle damper: Design and characterization. *Mechanical Systems and Signal Processing*, *99*, 219–228. https://doi.org/10.1016/J.YMSSP.2017.06.007

Zhang, R., Zhang, Y., Zheng, Z., Mo, L., & Wu, C. (2020). Parametrical optimization of particle dampers based on particle swarm algorithm. *Applied Acoustics*, *160*, 107083. https://doi.org/10.1016/J.APACOUST.2019.107083

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, *109*(1), 43–76.

https://doi.org/10.1109/JPROC.2020.3004555

Zuo, H., Bi, K., & Hao, H. (2019). Mitigation of tower and out-of-plane blade vibrations of offshore monopile wind turbines by using multiple tuned mass dampers. *Structure and Infrastructure Engineering*, *15*(2), 269–284. https://doi.org/10.1080/15732479.2018.1550096