Mechanical Systems and Signal Processing 180 (2022) 109437

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Physics-guided, data-refined modeling of granular material-filled particle dampers by deep transfer learning

Xin Ye [a,b], Yi-Qing Ni [a,b,*], Masoud Sajjadi [a,b], You-Wu Wang [a,b], Chih-Shiuan Lin [a,b]

[a] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[b] National Rail Transit Electrification and Automation Engineering Technology Research Center (Hong Kong Branch), Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Communicated by Xingjian Jing

Keywords:
Particle damper (PD)
Multi-fidelity modeling
Energy loss factor
Deep learning
Transfer learning

ABSTRACT

This study presents a novel transfer learning (TL)-based multi-fidelity modeling approach for a set of granular material-filled particle dampers (PDs) with varying cavity height and particle filling ratio, targeting to realize vibration/noise mitigation across a broad frequency band. The dynamic characteristics of this kind of dampers are highly nonlinear and depend on a number of features such as particle material and size, cavity configuration, filling ratio, excitation frequency and amplitude, etc. While deep neural network (DNN) has demonstrated success in a variety of fields including nonlinear dynamics, DNN is a data-hungry modeling approach and tends to yield inaccurate or inadequate models for high-dimensional nonlinear problems when data are scarce or expensive to collect. In this paper, we propose a multi-fidelity approach for characterizing the dynamics of granular material-filled PDs by combining low-fidelity data from an approximate governing/constitutive equation and high-fidelity experimental data in the context of deep TL. Making use of the low-fidelity data, a DNN is first trained to represent a mapping between input parameters (cavity height, particle filling ratio, excitation frequency and amplitude) and output parameter (damper energy loss factor). Then, in compliance with the deep TL philosophy, the weights and biases in all layers of the pre-trained DNN except a few outermost layers will be frozen, while those in the outermost layers are re-trained using the experimental data to formulate a multi-fidelity DNN. The modeling capability of this multi-fidelity DNN model developed by the deep TL strategy is compared with a DNN model with the same architecture but trained using only the experimental data. Results show that the multi-fidelity DNN model offers much better performance than the DNN model trained using only the experimental data for characterizing the PD dynamics across a broad frequency band from 100 to 2000 Hz. Since the formulated model is versatile to varying cavity height and particle filling ratio and accommodates different excitation frequencies and amplitudes, it is amenable to use in the optimal design of PDs.

1. Introduction

Particle damping promises a simple yet effective vibration/noise mitigation strategy. In particle damper (PD), a certain number of

* Corresponding author at: Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.
E-mail addresses: nixey.ye@connect.polyu.hk (X. Ye), ceyqni@polyu.edu.hk (Y.-Q. Ni), masoud.sajjadi@polyu.edu.hk (M. Sajjadi), youwu.wang@polyu.edu.hk (Y.-W. Wang), jason-cs.lin@polyu.edu.hk (C.-S. Lin).

https://doi.org/10.1016/j.ymssp.2022.109437
Received 31 January 2022; Received in revised form 7 April 2022; Accepted 9 June 2022
Available online 21 June 2022

particles are filled inside the enclosures of the PD's cavity. When it is attached to a vibrating structure, the vibration energy can be dissipated through the inelastic collision and friction effects between particles and particle-to-wall [1]. The PD takes the advantages of low cost, high effectiveness, and ability to operate in harsh working conditions. It has high potential for noise and vibration control in a broad frequency range [2], especially for granular material-filled PD. Particle damping has been investigated during the past decades [3,4], and has attained applications in machine racks [5], combustion engines [6], printed circuit boards [7], railways [8], and spacecraft [9]. Since the inception of PD, developing concise and precise phenomenological models that can reliably describe the complex behavior of this kind of dampers has been an unremitting effort [10]. The intricate motion of particles inside damper cavity during vibration is susceptible to the working condition and significantly affected by various factors such as the particle material, particle size, particle filling ratio, cavity configuration, excitation frequency and amplitude, etc. These factors make PDs highly nonlinear [11], and deservedly make it extremely difficult to develop a universal phenomenological or mechanistical model for PDs. In some endeavors, researchers simplified the motion of particles in the entire particle bed as an equivalent single impact ball inside the damper's cavity [12–14]. In other studies, the discrete element method (DEM) [15–24] was adopted for the dynamic modeling of PDs and PD-equipped structures that transforms the complex modeling task into a highly demanding computational problem.

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 [25]. 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 [26–28]. The latest scientific deep learning, such as physics-informed neural network [29], 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. [30] 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 [31], 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 [32–34]. 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 encode 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 optimal design 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.

The rest of the paper is organized as follows. Sec. 2 introduces the deep TL-based multi-fidelity modeling framework after a briefing of the common data-driven DNN method. Sec. 3 aims to select an appropriate governing/constitutive equation for granular material-filled PDs. Through comparing the attributes of two typical models, a two-phase flow equivalent viscosity model is chosen as the approximate governing/constitutive equation to generate low-fidelity data. Sec. 4 describes dynamic experiments on cylindrical-cavity PDs under various scenarios to generate high-fidelity data. Sec. 5 provides the development of the TL-based multi-fidelity DNN model and its comparison with the DNN model trained using only the experimental data, and the evaluation of the effect of various system parameters on the energy-dissipating capacity of PDs. Conclusions are summarized in Sec. 6.

2. Deep TL-based multi-fidelity modeling method for PDs

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 $\mathscr{X} = [\mathscr{X}_1, \mathscr{X}_2, \cdots, \mathscr{X}_N] : \Omega \to \mathbb{R}^N$ 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.

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

where $\eta(\mathscr{X})$ is the latent solution for the energy loss factor; $\mathscr{N}[\bullet; \lambda]$ 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 $\mathscr{D}_h = [\mathscr{X}_h; \eta_h]$, 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 Eq. (1) that reflect the unknown latent state $\eta(\mathscr{X})$ 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:

$$\mathbb{N}(\cdot; \boldsymbol{\theta}) = \left(\sigma^{[\mathrm{L}]} \circ \mathbf{W}^{[\mathrm{L}+1]}\right) \circ \cdots \circ \left(\sigma^{[0]} \circ \mathbf{W}^{[1]}\right) \tag{2}$$

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 '$\circ$' 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 $\left\{\boldsymbol{W}^{[l]}\right\}_{l=1}^{L+1}$ are concisely represented by $\boldsymbol{\theta}$. In this study, the 0th layer $\sigma^{[0]}$ represents the input vector $\mathscr{X}$ 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 $\mathscr{D}_h = [\mathscr{X}_h; \eta_h]$, the model parameters $\boldsymbol{\theta}$ can be trained by minimizing the loss function ($\mathscr{L}_d$):

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \mathscr{L}_d(\boldsymbol{\theta}) \tag{3}$$

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.

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

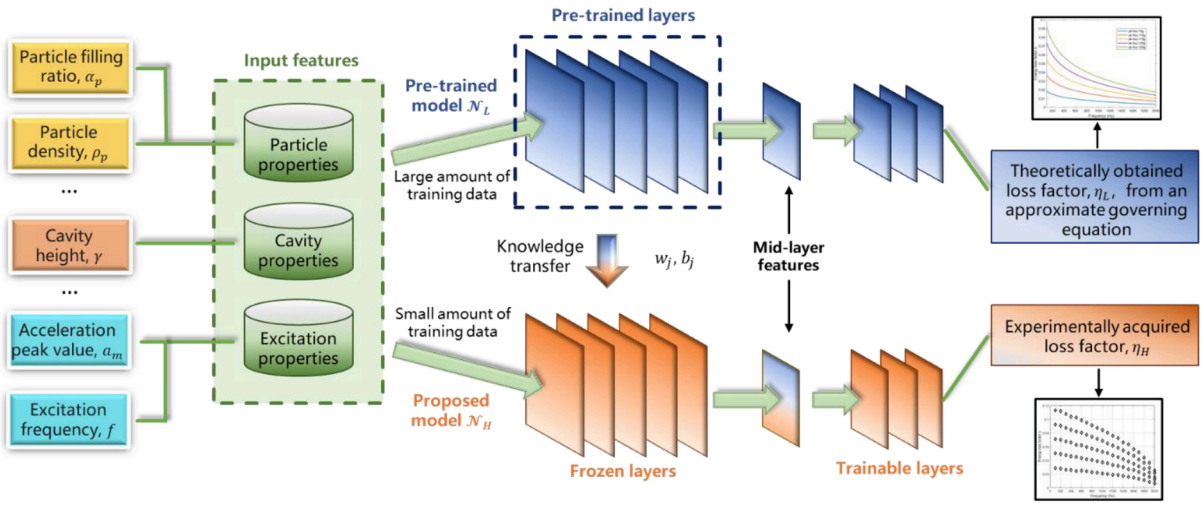

Fig. 1. Flowchart of the TL-based multi-fidelity DNN modeling method for PDs.

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

task) that has already been learned. It has been proven [33–36] 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 $\{X, Y\}$, $X = \{x_1, x_2, \cdots, x_n\} \in \mathscr{X}$ consists of n samples and $Y = \{y_1, y_2, \cdots, y_n\} \in \mathscr{Y}$ represents the corresponding label of the samples. The supervised learning is to map the samples X to their label Y. The pairs $\mathscr{X}$ and $\mathscr{Y}$ constitute the domain $\mathscr{D} = \{\mathscr{X}, \mathscr{Y}\}$ of the task which is denoted as $\mathscr{T}$. The TL in the context of DNN, i.e. deep TL, is pursued to accomplish the source task $\mathscr{T}_S$ first, which aims to learn a DNN using labeled data from the source domain $\mathscr{D}_S = \{\mathscr{X}_S, \mathscr{Y}_S\}$. 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 $\mathscr{T}_T$ targeting a multi-fidelity DNN with the use of fewer new labeled data with high fidelity from the target domain $\mathscr{D}_T = \{\mathscr{X}_T, \mathscr{Y}_T\}$. 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 $\mathscr{D}_S$ 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 $\mathscr{T}_S$. 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 $\mathscr{D}_T$, 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 Fig. 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 $\mathscr{T}_T$ differs from the source task $\mathscr{T}_S$ 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.

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 [37] 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 [38]. Recently, a novel PD modeling approach adopting two-phase or multi-phase flow theory has been proposed [9,39–41] 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 1.

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 [42–44] 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 [42,43], 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 [45] to derive the equivalent viscosity μ_c induced by particle–particle collision. The equivalent viscosity is defined as [44]:

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

where e_p denotes the coefficient of restitution which is set as 0.6 [44]; 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:

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

Substituting the particle fluctuation rate expression into Eq. (4) yields:

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

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

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

The equivalent shear viscosity μ_f generated due to the effect of interparticle friction can be expressed as [44]:

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

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:

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

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

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

where

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

With Eqs. (6) and (10), the total equivalent viscosity of particle flow, μ_p, can be obtained with considering both collision and friction

effects as:

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

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:

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

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:

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

According to Sarpkaya [46], the drag coefficient C_d in Eq. (13) can be expressed as:

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

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:

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

where

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

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

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

$$\alpha = K_1/K_3;\ \alpha_1 = K_2/K_3 \tag{20}$$

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

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

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:

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

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 [47], where the expression of average power flow is:

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

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:

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

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, Eq. (21) can be rewritten as:

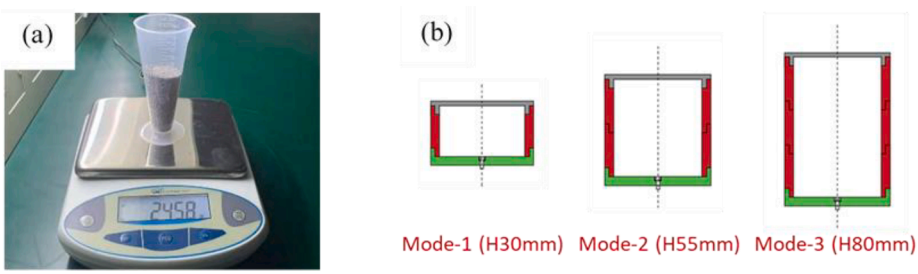

Fig. 2. (a) Weighing of filling particles; (b) schematic diagram of three cavity modes.

Table 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

Table 3
Weight of tungsten powder for different filling volume fractions (g).

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

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

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

4. Dynamic experiments 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.

4.1. Experiment 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 (Fig. 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 (Fig. 2(b)), respectively. The number of assembled modules in the three configurations is one (mode-1), two (mode-2), and three (mode-3).

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 2.

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 from 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. The weight of the accelerometer placed on top of the cavity for the measurement of acceleration in the acceleration-controlled experiments is 3.0 g.

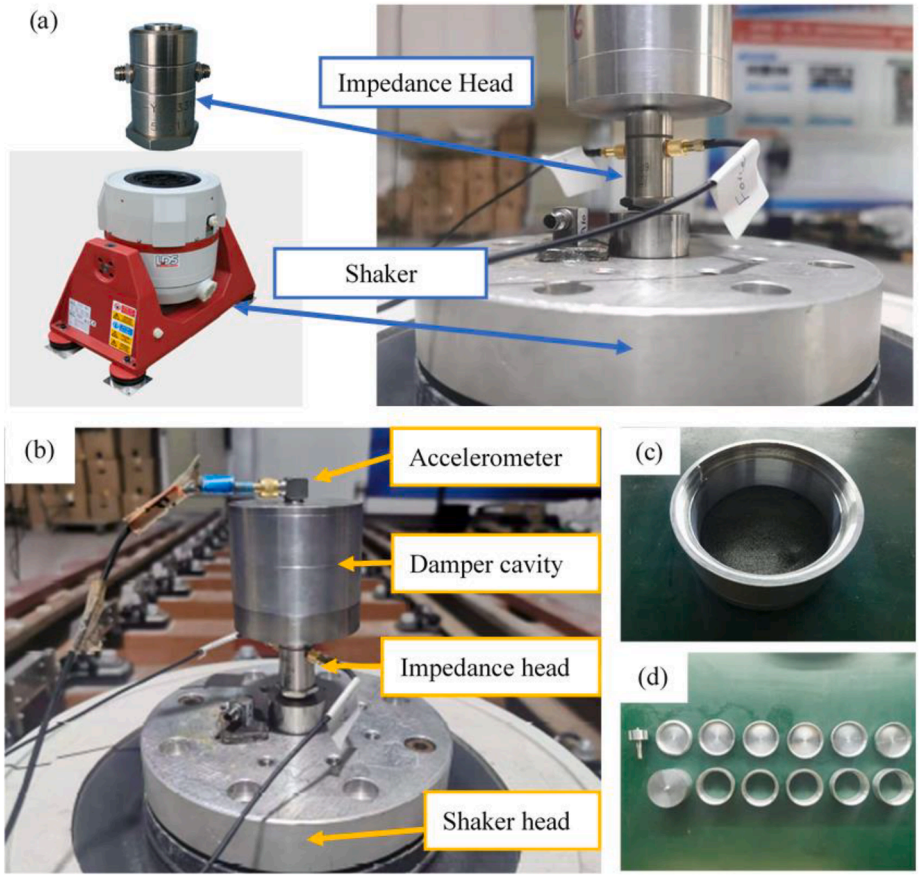

Fig. 3. Experimental apparatus: (a) assembly of impedance head and shaker; (b) setup of cylindrical PD; (c) particles filling in damper cavity; and (d) damper assembly components.

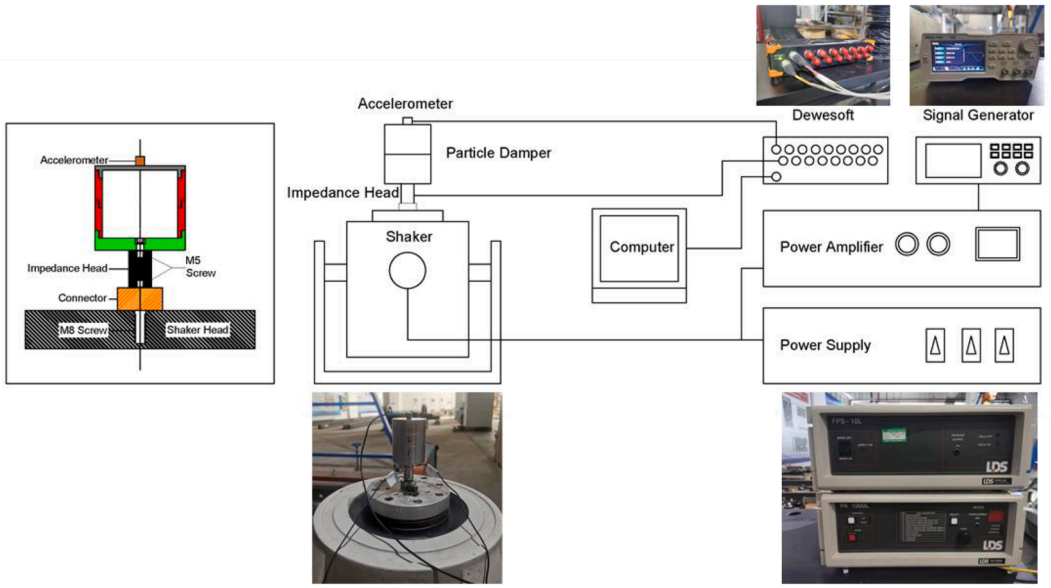

Fig. 4. Schematic diagram of experimental setup.

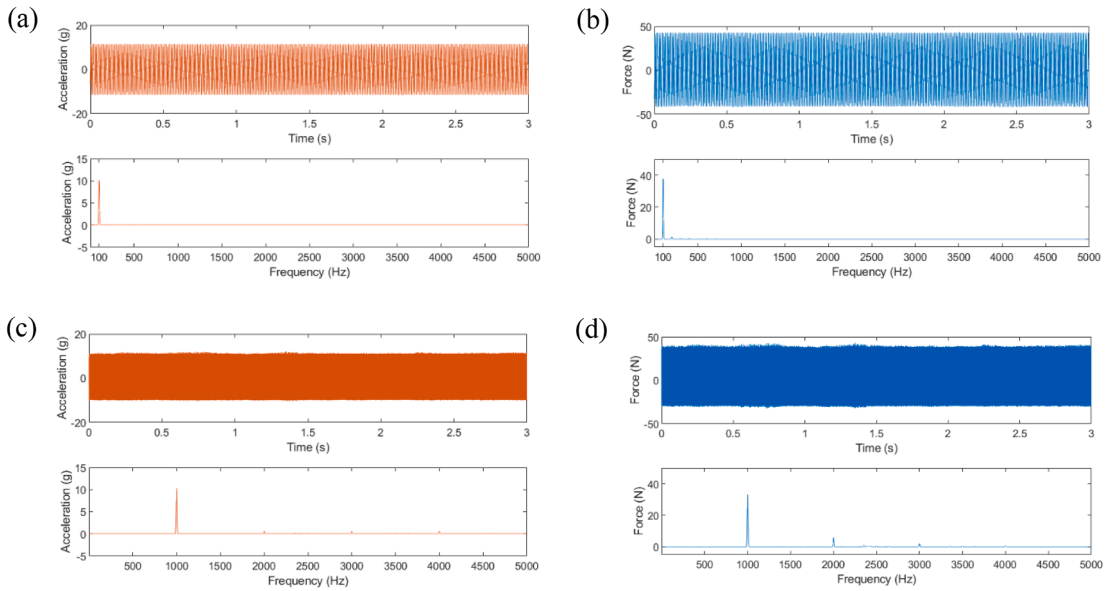

Fig. 5. (a) Acceleration in time- and frequency-domains at $f = 100$ Hz; (b) Force in time- and frequency-domains at $f = 100$ Hz; (c) Acceleration in time- and frequency-domains at $f = 1000$ Hz; and (d) Force in time- and frequency-domains at $f = 1000$ Hz.

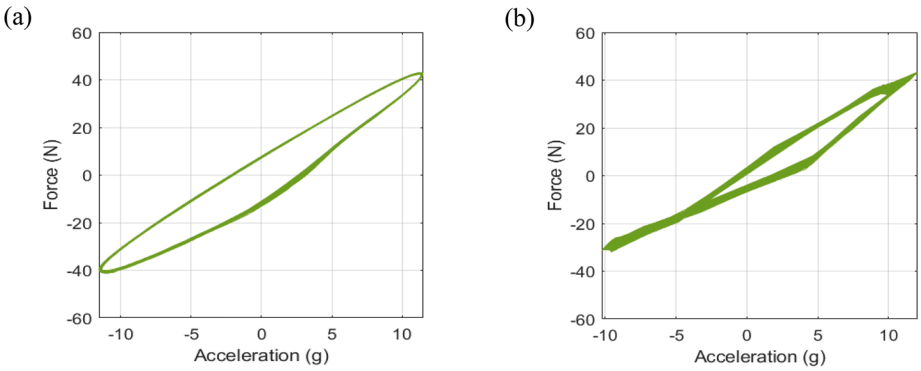

Fig. 6. Diagram of force versus acceleration: (a) $f = 100$ Hz; (b) $f = 1000$ Hz.

4.2. 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 Fig. 3.

Fig. 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

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

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

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.

Fig. 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 Fig. 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 Eqs. (23) to (25).

5. Multi-fidelity model of PDs

5.1. 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 $\mathscr{X}=\{\alpha_p,\rho_p,\gamma,a_m,f\}$, 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, $\mathscr{T}$, the size of each hidden layer, $\mathscr{C}$, 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/cm^3) 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 [48] 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 $\mathscr{T}$ is $\{3,4,5\}$, and the size of hidden layer $\mathscr{C}$ 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 4. The hyper-parameters $\{\mathscr{T}=3,\ \mathscr{C}=10\}$ give rise to the minimum MSE; therefore, the DNN architecture is set as 3 hidden layers with 10 neurons in each hidden layer.

After determining the number of hidden layers ($\mathscr{T}$) and the size of each hidden layer ($\mathscr{C}$), 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 [49] 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) [50,51], 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) [52]. 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 5. It is concluded that the activation functions of "tanh-tanh-

Table 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

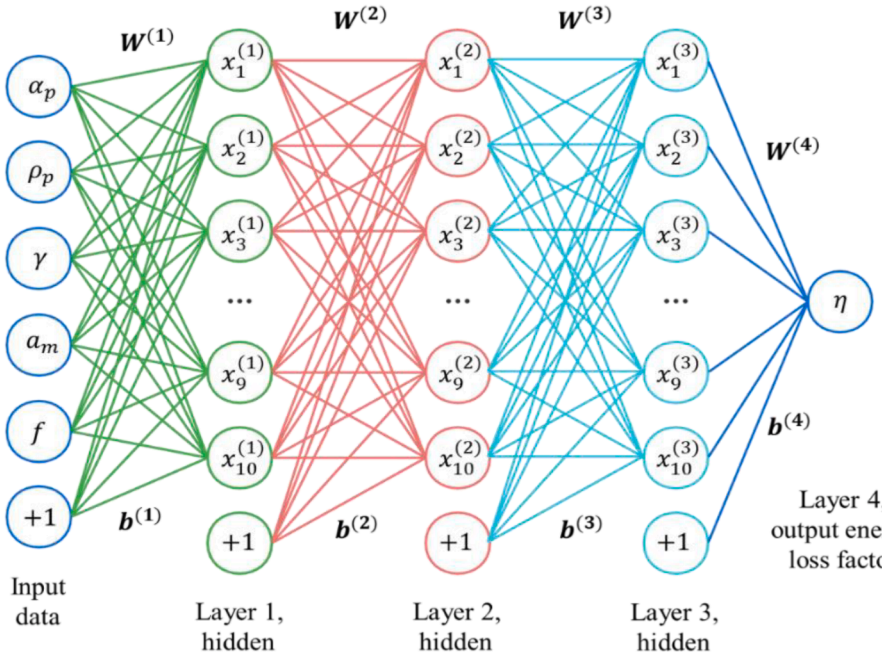

Fig. 7. Architecture of DNN model.

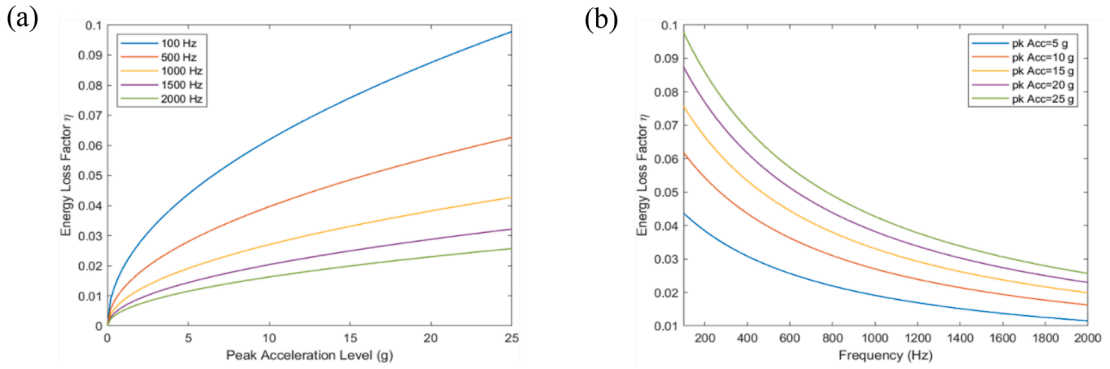

Fig. 8. Calculated values of the energy loss factor η of PD using the two-phase flow equivalent viscosity model: (a) η versus peak acceleration under various excitation frequencies; (b) η versus excitation frequency under various peak accelerations.

ReLU" for the three hidden layers together with Adam algorithm achieves the best performance. This configuration in conjunction with the specified hyper-parameters $\{\mathscr{T} = 3, \mathscr{C} = 10\}$ constitutes the architecture of the DNN (Fig. 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

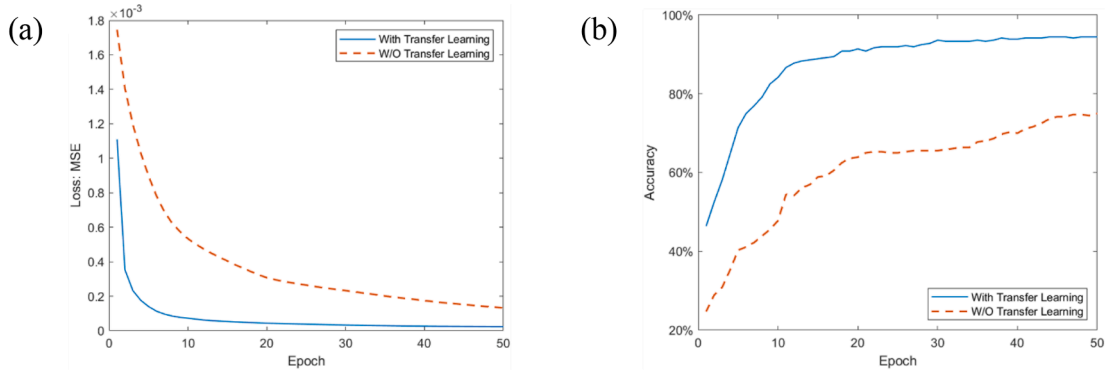

Fig. 9. Comparison of training process between the DNN models with TL and W/O TL in terms of (a) mean squared error (MSE); (b) model accuracy.

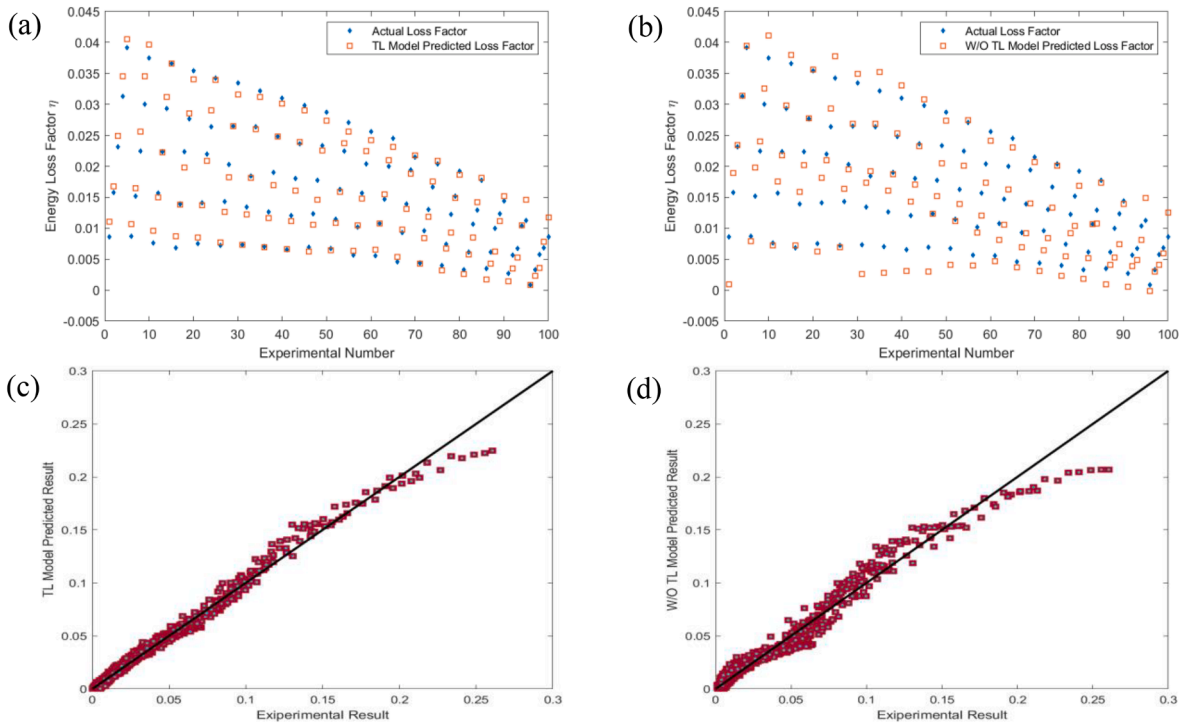

Fig. 10. Prediction of energy loss factor by (a) TL approach, and (b) W/O TL approach; R^2 score of the prediction results from (c) TL approach, and (d) W/O TL approach.

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/cm^3 (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 by using an explicit Runge-Kutta numerical method. Fig. 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 [53] 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 (Fig. 7) are specified.

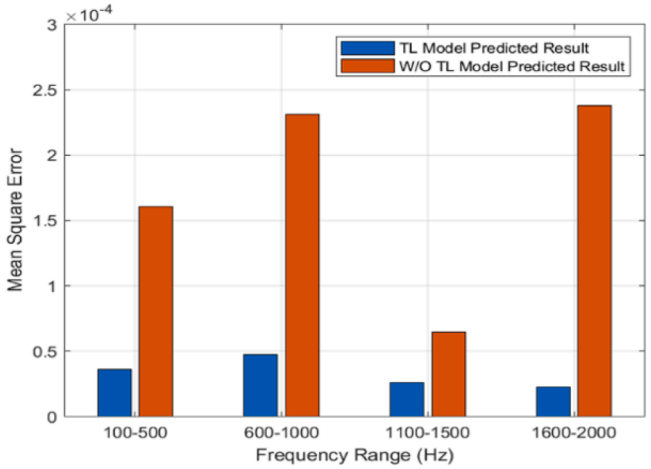

Fig. 11. Comparison of average mean squared error (MSE) generated by the two models in different frequency ranges.

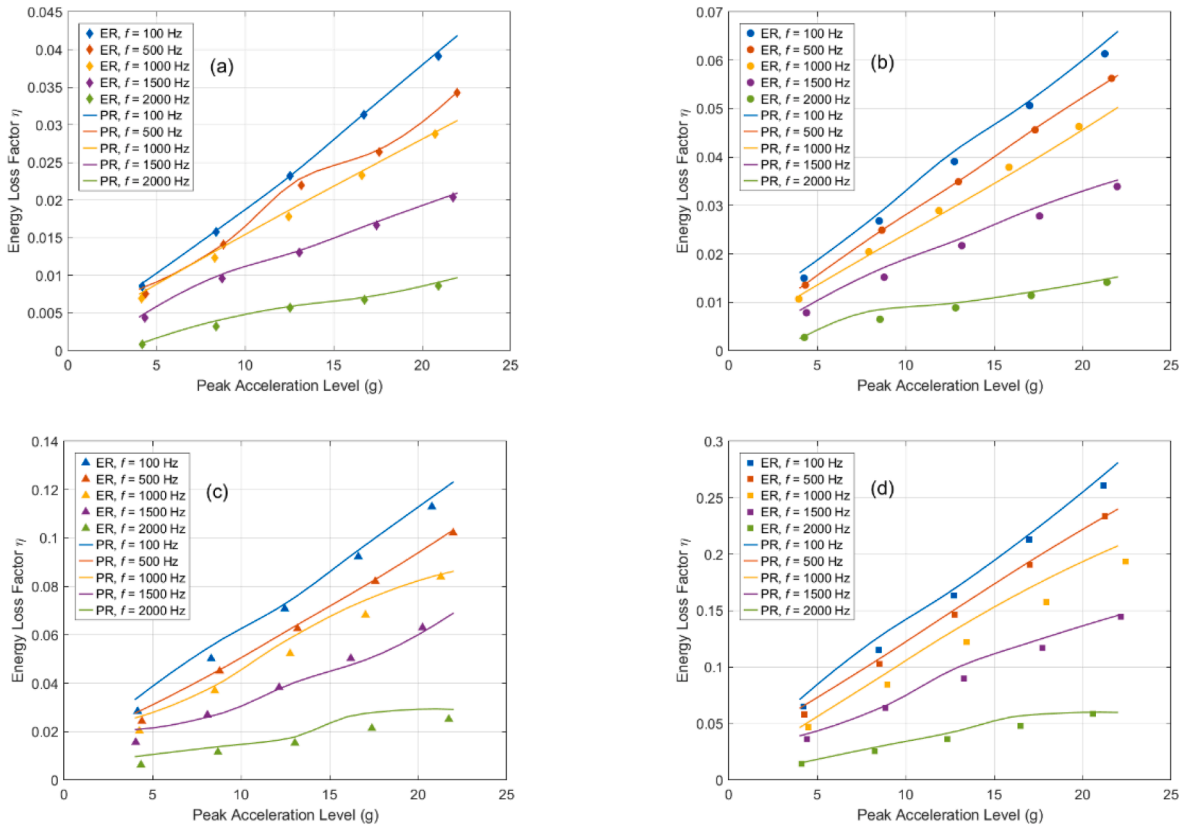

Fig. 12. Energy loss factor versus excitation level (acceleration amplitude) for mode-1 cavity under filling ratio of: (a) 10%; (b) 30%; (c) 50%; (d) 70%. Here ER stands for experimental results, and PR stands for predicted results.

5.2. Formulation of multi-fidelity model using deep TL

The pre-trained DNN, which was trained using ample numerical data (source dataset $\mathscr{D}_S$) from the two-phase flow equivalent viscosity model, is now refined using actual experimental data (target dataset $\mathscr{D}_T$) 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%

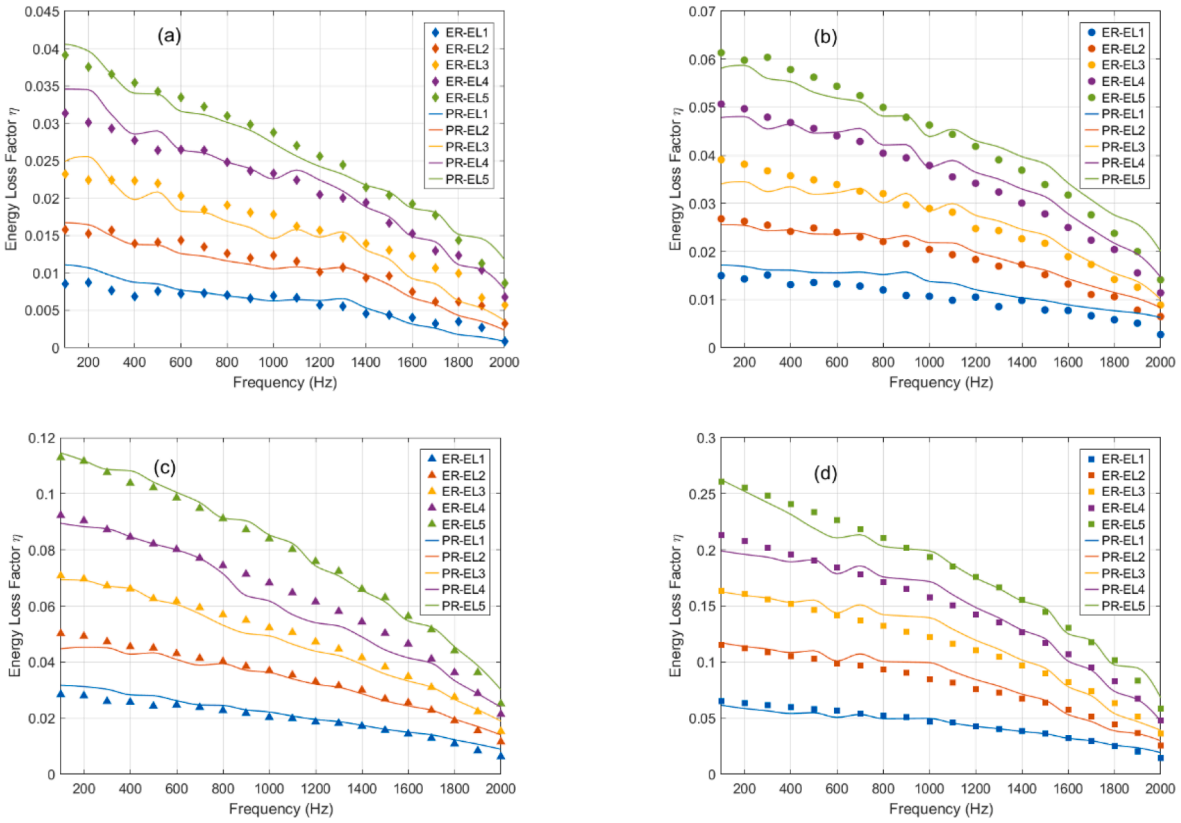

Fig. 13. Energy loss factor versus excitation frequency for mode-1 cavity under filling ratio of: (a) 10%; (b) 30%; (c) 50%; (d) 70%. Here ER stands for experimental results, PR stands for predicted results, and EL is the abbreviation of excitation level.

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. Fig. 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 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 Fig. 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%.

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. Fig. 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 R^2 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 Fig. 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.

5.3. *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. Fig. 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

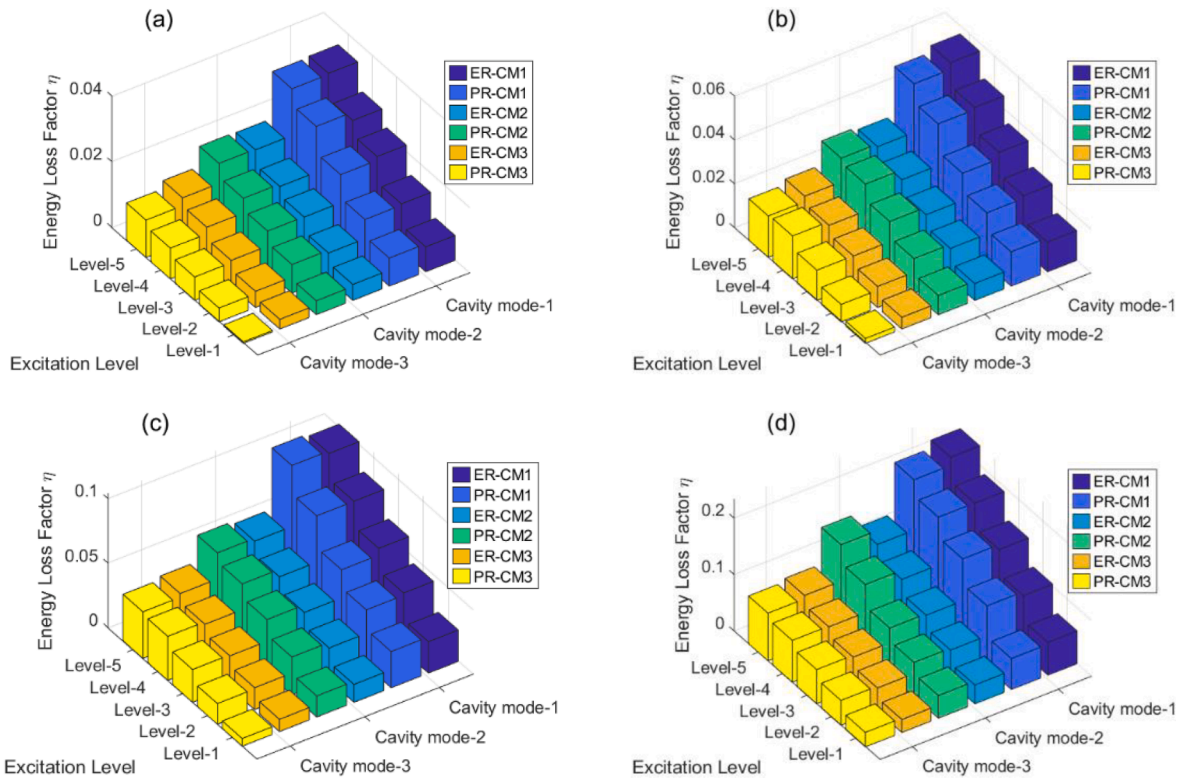

Fig. 14. Energy loss factor versus excitation level for three cavity modes under excitation frequency of 500 Hz and filling ratio of: (a) 10%; (b) 30%; (c) 50%; (d) 70%. Here ER stands for experimental results, PR stands for predicted results, and CM is the abbreviation of cavity mode.

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.

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. Fig. 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, and Figs. 14 and 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.

6. Conclusions and future work

In this paper, 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

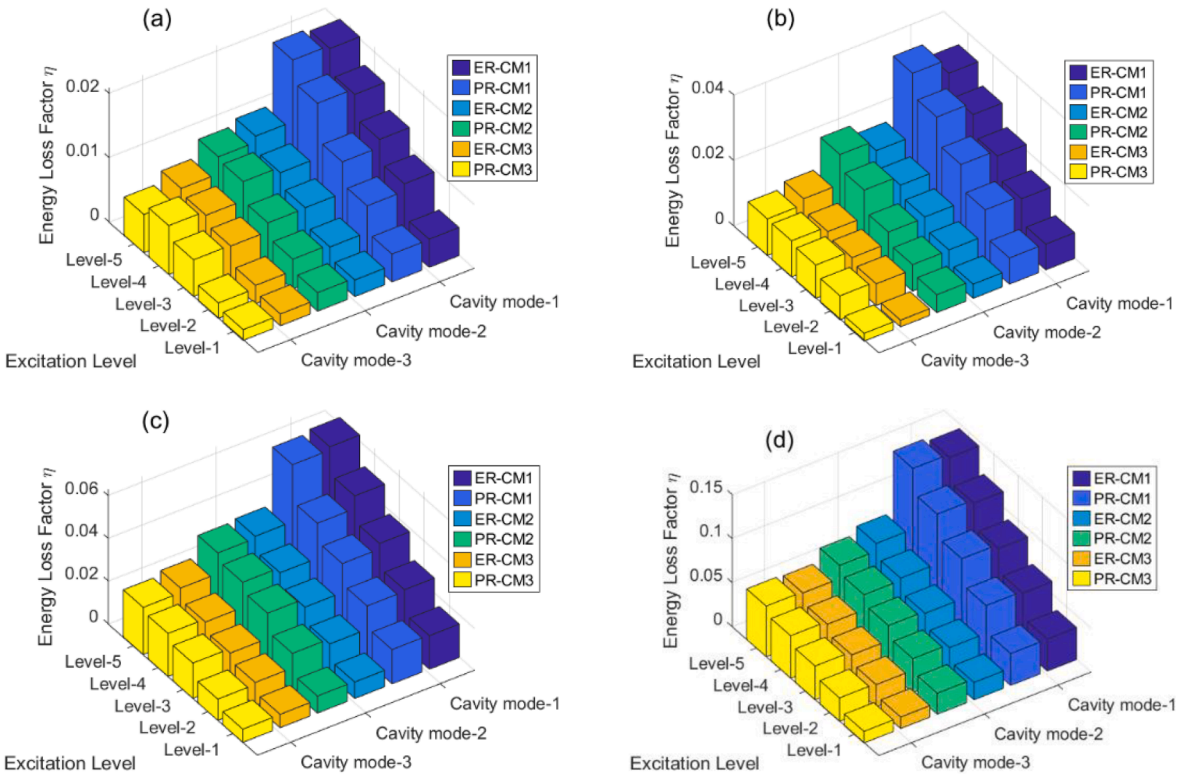

Fig. 15. Energy loss factor versus excitation level for three cavity modes under excitation frequency of 1500 Hz and filling ratio of: (a) 10%; (b) 30%; (c) 50%; (d) 70%. Here ER stands for experimental results, PR stands for predicted results, and CM is the abbreviation of cavity mode.

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-tunning 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.

CRediT authorship contribution statement

Xin Ye: Conceptualization, Methodology, Formal analysis, Software, Writing – original draft. **Yi-Qing Ni:** Conceptualization, Supervision, Resources, Project administration, Funding acquisition, Writing – review & editing. **Masoud Sajjadi:** Investigation, Visualization, Data curation. **You-Wu Wang:** Software, Validation. **Chih-Shiuan Lin:** Data curation, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research described in this paper was supported by a grant (RIF) from the Research Grants Council of the Hong Kong Special Administrative Region (SAR), China (Grant No. R-5020-18), a grant from the National Natural Science Foundation of China (Grant No. U1934209), and a grant from the Hong Kong, Macao and Taiwan Science and Technology Innovation Cooperation Key Project of Sichuan Province, China (Grant No. 2020YFH0178). The authors would also like to appreciate the funding support by the Innovation and Technology Commission (ITC) of Hong Kong SAR Government to the Hong Kong Branch of Chinese National Rail Transit Electrification and Automation Engineering Technology Research Center (Grant No. K-BBY1).

References

[1] T. Chen, K. Mao, X. Huang, M.Y. Wang, Dissipation mechanisms of nonobstructive particle damping using discrete element method, in, in: Smart Structures and Materials 2001: Damping and Isolation, SPIE, 2001, pp. 294–301, https://doi.org/10.1117/12.432713.
[2] M. Saeki, Analytical study of multi-particle damping, J. Sound Vib. 281 (2005) 1133–1144, https://doi.org/10.1016/j.jsv.2004.02.034.
[3] J.J. Hollkamp, R.W. Gordon, Experiments with particle damping, in: Smart Structures and Materials 1998: Passive Damping and Isolation, SPIE, 1998: pp. 2–12. https://doi.org/10.1117/12.310675.
[4] Z. Xu, M.Y. Wang, T. Chen, Particle damping for passive vibration suppression: Numerical modelling and experimental investigation, J. Sound Vib. 279 (2005) 1097–1120, https://doi.org/10.1016/j.jsv.2003.11.023.
[5] Z. Zheng, C. Wu, H. Wu, J. Wang, X. Lei, Vibration isolation performance of nonobstructive particle damping in a machine rack, J. Vib. Control 25 (2019) 1122–1130, https://doi.org/10.1177/1077546318814039.
[6] F. Duvigneau, S. Koch, E. Woschke, U. Gabbert, An effective vibration reduction concept for automotive applications based on granular-filled cavities, J. Vib. Control 24 (2018) 73–82, https://doi.org/10.1177/1077546316632932.
[7] P. Veeramuthuvel, K.K. Sairajan, K. Shankar, Vibration suppression of printed circuit boards using an external particle damper, J. Sound Vib. 366 (2016) 98–116, https://doi.org/10.1016/j.jsv.2015.12.034.
[8] J. Jin, W. Yang, H.I. Koh, J. Park, Development of tuned particle impact damper for reduction of transient railway vibrations, Appl. Acoust. 169 (2020), 107487, https://doi.org/10.1016/j.apacoust.2020.107487.
[9] X. Wang, X. Liu, Y. Shan, T. He, Design, simulation and experiment of particle dampers attached to a precision instrument in spacecraft, J. Vibroeng. 17 (2015) 1605–1614. https://www.jvejournals.com/article/15880.
[10] Z. Lu, Z. Wang, S.F. Masri, X. Lu, Particle impact dampers: past, present, and future, Struct. Control Health Monitor. 25 (2018) 1–25, https://doi.org/10.1002/stc.2058.
[11] R.D. Nayeri, S.F. Masri, J.P. Caffrey, Studies of the performance of multi-unit impact dampers under stochastic excitation, J. Vibr. Acoust. Trans. ASME 129 (2007) 239–251, https://doi.org/10.1115/1.2346694.
[12] R.D. Friend, V.K. Kinra, Particle impact damping, J. Sound Vib. 233 (2000) 93–118, https://doi.org/10.1006/jsvi.1999.2795.
[13] Z. Lu, X. Lu, W. Lu, S.F. Masri, Experimental studies of the effects of buffered particle dampers attached to a multi-degree-of-freedom system under dynamic loads, J. Sound Vib. 331 (2012) 2007–2022, https://doi.org/10.1016/j.jsv.2011.12.022.
[14] A. Papalou, S.F. Masri, Response of impact dampers with granular materials under random excitation, Earthq. Eng. Struct. Dyn. 25 (1996) 253–267, https://doi.org/10.1002/(SICI)1096-9845(199603)25:3<253::AID-EQE553>3.0.CO;2-4.
[15] P.A. Cundall, O.D.L. Strack, A discrete numerical model for granular assemblies, Geotechnique 29 (1979) 47–65, https://doi.org/10.1680/geot.1979.29.1.47.
[16] H. Elmqvist, A. Goteman, V. Roxling, T. Ghandriz, in: Generic modelica framework for multibody contacts and discrete element method, in, Linköping University Electronic Press, 2015, pp. 427–440, https://doi.org/10.3384/ecp15118427.
[17] D.N.J. Els, Damping of rotating beams with particle dampers: Discrete element method analysis, in: AIP Conference Proceedings, American Institute of PhysicsAIP, 2013: pp. 867–870. https://doi.org/10.1063/1.4812069.
[18] M.A.J. Holmes, R. Brown, P.A.L. Wauters, N.P. Lavery, S.G.R. Brown, Bending and twisting friction models in soft-sphere discrete element simulations for static and dynamic problems, Appl. Math. Model. 40 (2016) 3655–3670, https://doi.org/10.1016/j.apm.2015.10.026.
[19] Z. Lu, X. Lu, H. Jiang, S.F. Masri, Discrete element method simulation and experimental validation of particle damper system, Eng. Comput. 31 (2014) 810–823, https://doi.org/10.1108/EC-08-2012-0191.
[20] K. Mao, M.Y. Wang, Z. Xu, T. Chen, DEM simulation of particle damping, Powder Technol. 142 (2004) 154–165, https://doi.org/10.1016/j.powtec.2004.04.031.
[21] H.G. Matuttis, J. Chen, Understanding the discrete element method: simulation of non-spherical particles for granular and multi-body systems, Wiley Blackwell (2014), https://doi.org/10.1002/9781118567210.
[22] B. Suhr, K. Six, Friction phenomena and their impact on the shear behaviour of granular material, Comput. Particle Mech. 4 (2017) 23–34, https://doi.org/10.1007/s40571-016-0119-2.
[23] Z. Syed, M. Tekeste, D. White, A coupled sliding and rolling friction model for DEM calibration, J. Terramech. 72 (2017) 9–20, https://doi.org/10.1016/j.jterra.2017.03.003.

[24] C.M. Wensrich, A. Katterfeld, Rolling friction as a technique for modelling particle shape in DEM, Powder Technol. 217 (2012) 409–417, https://doi.org/10.1016/j.powtec.2011.10.057.
[25] L. Gagnon, M. Morandini, G.L. Ghiringhelli, A review of particle damping modeling and testing, J. Sound Vib. 459 (2019), 114865, https://doi.org/10.1016/j.jsv.2019.114865.
[26] H.B. Huang, J.H. Wu, X.R. Huang, M.L. Yang, W.P. Ding, The development of a deep neural network and its application to evaluating the interior sound quality of pure electric vehicles, Mech. Syst. Sig. Process. 120 (2019) 98–116, https://doi.org/10.1016/J.YMSSP.2018.09.035.
[27] Z. Ye, J. Yu, Deep morphological convolutional network for feature learning of vibration signals and its applications to gearbox fault diagnosis, Mech. Syst. Sig. Process. 161 (2021), 107984, https://doi.org/10.1016/J.YMSSP.2021.107984.
[28] H. Yu, J. Tao, C. Qin, M. Liu, D. Xiao, H. Sun, C. Liu, A novel constrained dense convolutional autoencoder and DNN-based semi-supervised method for shield machine tunnel geological formation recognition, Mech. Syst. Sig. Process. 165 (2022), 108353, https://doi.org/10.1016/J.YMSSP.2021.108353.
[29] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707, https://doi.org/10.1016/j.jcp.2018.10.045.
[30] P. Veeramuthuvel, K. Shankar, K.K. Sairajan, Application of RBF neural network in prediction of particle damping parameters from experimental data, J. Vib. Control 23 (2017) 909–929, https://doi.org/10.1177/1077546315587147.
[31] S.X. Chen, L. Zhou, Y.Q. Ni, X.Z. Liu, An acoustic-homologous transfer learning approach for acoustic emission-based rail condition evaluation, Struct. Health Monitor. 20 (2021) 2161–2181, https://doi.org/10.1177/1475921720976941.
[32] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, T. Darrell, DeCAF: A deep convolutional activation feature for generic visual recognition, 31st International Conference on Machine Learning, ICML 2014. 2 (2013) 988–996. http://arxiv.org/abs/1310.1531.
[33] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks? Adv. Neural Inform. Process. Syst. 4 (2014) 3320–3328. https://doi.org/10.48550/arxiv.1411.1792.
[34] M. Long, Y. Cao, J. Wang, M.I. Jordan, Learning transferable features with deep adaptation networks, 32nd International Conference on Machine Learning, ICML 2015. 1 (2015) 97–105. https://doi.org/10.48550/arxiv.1502.02791.
[35] J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, G. Zhang, Transfer learning using computational intelligence: a survey, Knowl.-Based Syst. 80 (2015) 14–23, https://doi.org/10.1016/j.knosys.2015.01.010.
[36] S. Chakraborty, Transfer learning based multi-fidelity physics informed deep neural network, J. Comput. Phys. 426 (2021) 109942, https://doi.org/10.1016/j.jcp.2020.109942.
[37] S.F. Masri, J.P. Caffrey, Response of pounding dynamic vibration neutralizer under harmonic and random excitation, J. Appl. Mech. Trans. ASME 86 (2019) 1–15, https://doi.org/10.1115/1.4041910.
[38] Z. Lu, X. Chen, D. Zhang, K. Dai, Experimental and analytical study on the performance of particle tuned mass dampers under seismic excitation, Earthq. Eng. Struct. Dyn. 46 (2017) 697–714, https://doi.org/10.1002/eqe.2826.
[39] X. Lei, C. Wu, P. Chen, Optimizing parameter of particle damping based on Leidenfrost effect of particle flows, Mech. Syst. Sig. Process. 104 (2018) 60–71, https://doi.org/10.1016/j.ymssp.2017.10.037.
[40] X. Lei, C. Wu, H. Wu, A novel composite vibration control method using double-decked floating raft isolation system and particle damper, J. Vib. Control 24 (2018) 4407–4418, https://doi.org/10.1177/1077546317724967.
[41] D. Wang, C. Wu, A novel prediction method of vibration and acoustic radiation for rectangular plate with particle dampers, J. Mech. Sci. Technol. 30 (2016) 1021–1035, https://doi.org/10.1007/s12206-016-0205-7.
[42] C.J. Wu, W.H. Liao, M.Y. Wang, Modeling of granular particle damping using multiphase flow theory of gas-particle, J. Vibrat. Acoust. Trans. ASME 126 (2004) 196–201, https://doi.org/10.1115/1.1688763.
[43] D.K. Wang, C.J. Wu, R.C. Yang, Free vibration of the damping beam using co-simulation method based on the mft, Int. J. Acoust. Vibrat. 20 (2015) 251–257. https://doi.org/10.20855/ijav.2015.20.4388.
[44] X. Lei, C. Wu, Non-obstructive particle damping using principles of gas-solid flows, J. Mech. Sci. Technol. 31 (2017) 1057–1065, https://doi.org/10.1007/s12206-017-0204-3.
[45] L.S. Fan, C. Zhu, Principles of Gas-Solid Flows, Cambridge University Press, London, 1998.
[46] T. Sarpkaya, Force on a circular cylinder in viscous oscillatory flow at low keulegan–carpenter numbers, J. Fluid Mech. 165 (1986) 61–71, https://doi.org/10.1017/S0022112086002999.
[47] M. Bustamante, S.N.Y. Gerges, E.F. Vergara, J.P. Arenas, High damping characteristics of an elastomer particle damper, Int. J. Acoust. Vibr. 21 (2016) 112–121. https://doi.org/10.20855/ijav.2016.21.1401.
[48] Y. Jung, Multiple predicting K-fold cross-validation for model selection, J. Nonparametr. Statist. 30 (1) (2018) 197–215.
[49] M. Li, T. Zhang, Y. Chen, A.J. Smola, Efficient mini-batch training for stochastic optimization, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. (2014) 661–670. https://doi.org/10.1145/2623330.2623612.
[50] L. Bottou, Stochastic gradient descent tricks, Neural Networks: Tricks of the Trade, Springer, Berlin, 2012, pp. 421–436.
[51] N. Ketkar, Stochastic gradient descent, Deep Learning with Python, in: Deep Learning with Python, Apress, Berkeley, CA, 2017, pp. 113–132.
[52] D.P. Kingma, J.L. Ba, Adam: a method for stochastic optimization, 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. (2014). https://arxiv.org/abs/1412.6980v9.
[53] S. Gopal Krishna Patro, K. Kumar Sahu, Normalization: a preprocessing stage, Int. Adv. Res. J. Sci. Eng. Technol. (2015) 20–22. https://doi.org/10.17148/iarjset.2015.2305.