

Dynamic neural network switching for active control of nonlinear systems

Xander Pike and Jordan Cheer

*Institute of Sound and Vibration Research, University of Southampton,
Southampton SO17 1BJ, United Kingdom*

(Dated: 3 May 2025)

1 Feedforward active noise and vibration control systems have been developed for many
2 applications, but are generally designed using linear digital filters, most typically im-
3 plementing the Filtered reference Least Mean Squares (FxLMS) algorithm. When the
4 system under control exhibits nonlinearities, linear controllers cannot fully capture
5 the system dynamics to maximize performance. Previous work has shown that Neu-
6 ral Network (NN) based controllers can improve control performance in the presence
7 of nonlinearities. However, inferring the outputs of NN controllers can be computa-
8 tionally expensive, limiting their practicality, particularly when control is required
9 across a range of nonlinear behaviors. In this paper, a control strategy is proposed
10 where performance is maintained across a nonlinear range of operation by dynam-
11 ically switching between a set of smaller, and therefore more efficient, NNs that
12 are individually trained over specific ranges of the nonlinear system behavior. It is
13 demonstrated via both simulations of a system with a simple nonlinear stiffness in
14 the primary path and offline simulations using a physical nonlinear dynamical system
15 in the primary path, that the performance of the proposed switching approach offers
16 a control performance advantage compared to both a larger generalized individual
17 NN controller and a Functional Link Artificial NN (FLANN) based controller.

I. INTRODUCTION

The requirements for the control of unwanted noise and vibration are becoming increasingly stringent in modern engineering systems, with higher levels of control performance required to address user expectations and to manage the increasing requirement for lightweight engineering designs that exacerbate noise and vibration. High frequency noise and vibration can typically be attenuated effectively using passive control solutions. However, the control of low frequency disturbances via passive methods can often require the implementation of large or heavy systems. By comparison, active control solutions are typically capable of achieving effective control at low frequencies, and can benefit from being relatively lightweight and compact.

Historically, feedforward active noise and vibration control systems have been implemented using linear control filters and system models. However, it is well understood that nonlinearities present in either the plant or primary path of the control system can have a significant impact on control performance¹⁻⁴. A wide range of approaches have been proposed to overcome this limitation, including polynomial, cross-term or trigonometric expansion of the reference signal^{5,6}, genetic algorithms⁷ and fuzzy logic-based methods⁸. Another methodology that has shown promising results is the application of machine learning methods to these nonlinear control problems. In particular, Neural Networks (NNs) are well-motivated for both modeling and control of nonlinear systems⁹ due to the fact that they are known to possess the property of being ‘universal approximators’¹⁰.

Many different uses of NNs have been studied in the literature, including system modeling^{4,11–15}, feedforward controller design^{4,11,16–18}, inverse modeling¹⁹, signal prediction and feedback control^{20–25}, linear filter selection²⁶, adaptive parameter estimation for linear controllers^{24,27}, frequency-domain control²⁸, multichannel controller design²⁹, and signal classification³⁰. However, the ability of such NN control systems to generalize well across a range of system behaviors has not been extensively explored in the associated literature. Nonlinear systems can exhibit rich and varied behavior as the input excitation changes, so training individual NNs to achieve acceptable control performance under such conditions is not straightforward. A training or design approach that provides effective control performance is clearly desirable in practical implementations where the properties of the excitation, and therefore the behavior of the system nonlinearity, may change over time.

It has previously been demonstrated that it is possible to train individual NN controllers to produce control performance over a range of nonlinear system behaviors that approaches the performance of identical NN controllers trained at a single level of nonlinear behavior³¹. However, the performance of such generalized controllers is dependent on the range of nonlinear behavior over which the controller is trained to perform well, and the size (and therefore computational cost of inference) of the network. In this paper, a possible solution to this problem is proposed where a set of relatively small NN controllers are trained over distinct operating ranges and a simple method of switching between these controllers to achieve control performance across a much wider range of operational conditions is proposed¹. Section II describes the simulated nonlinear control problem that is utilised to initially explore

the proposed control strategy. Section III describes the proposed controller switching approach, individual controller architecture and the adopted training methodology. Section IV presents simulation results, which explore how the switched-controller training ranges influence performance and then demonstrate how the performance of the proposed control architecture compares to a larger generalized NN and to an adaptive nonlinear controller. Section V presents a study where the proposed switched-controller approach is applied to a physical nonlinear system, which demonstrates the practicability of the approach for realistic nonlinearities with more complicated characteristics. Finally, Section VI discusses conclusions from the work.

II. NONLINEAR CONTROL PROBLEM DEFINITION

A diagram of the simulated nonlinear system to be controlled is presented in Figure 1. The system consists of a primary source, generating the unwanted disturbance, and a secondary source, which is used to cancel the disturbance. To introduce a nonlinearity into the primary path of the system, the primary source is modeled as a Duffing oscillator. While the type of nonlinearity studied is clearly important, this type of nonlinearity has been utilized in the first instance due to its widespread use within the literature to represent systems with dynamically varying stiffness, as well as its simplicity, with the aim of ensuring that the results of this work are not constrained to an overly-specific problem. Additionally, this system represents a simple use case for the proposed controller-switching approach, since the degree of nonlinear behavior of the system is governed by the magnitude of the system excitation or floor motion, $x(t)$, which is equivalent to the magnitude of the reference signal

81 used by the feedforward control in this case. The secondary source is modeled as a simple
 82 harmonic oscillator such that the plant response is linear.

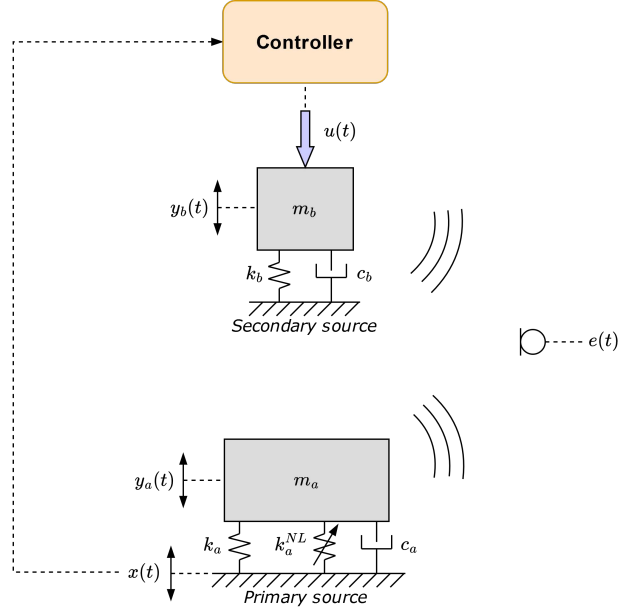


FIG. 1. Diagram of the simulated nonlinear system, consisting of a nonlinear primary acoustic source, and a linear secondary source. System parameters are defined in Table I.

83

84

85 The displacement of the primary source, $y_a(t)$, is induced by motion of the floor, $x(t)$,
 86 to which it is attached. While the motion of the secondary source is induced by the control
 87 force, $u(t)$, which acts upon it. The equations of motion for the total system can be expressed
 88 as

$$m_a \ddot{y}_a(t) + k_a \theta(t) + k_a^{NL} \theta^3(t) + c_a \dot{\theta}(t) = 0 \quad (1)$$

89

$$m_b \ddot{y}_b(t) + k_b y_b(t) + c_b \dot{y}_b(t) + F_c(t) = 0 \quad (2)$$

90 where $\theta(t) = y_a(t) - x(t)$ and the remaining variables are defined in Figure 1 and Table I. It
 91 can be seen from equation 1 that the degree of nonlinearity is dependent on the term that

is proportional to the nonlinear stiffness, $k(a)^{NL}$, which depends on both the motion of the floor, $x(t)$, and the motion of the primary mass, $y_a(t)$. However, since the motion of the primary mass is directly dependent on the motion of the floor, the floor motion, which is used as the reference signal in the considered feedforward controller, is an effective measure for the degree of nonlinearity. The parameters that define the dynamics of the simulated system were selected such that the two oscillators have unity mass, but distinct resonance frequencies of 60 Hz and 80 Hz. The damping coefficients c_a and c_b were selected such that each oscillator is subject to 20% of critical damping, such that the oscillators are neither significantly underdamped or overdamped.

The two oscillators are assumed to behave as monopole acoustic sources, with the simplifying assumption that any pressure measurements are made in the far field. The complex far-field pressure field generated by an acoustic monopole oscillating at an angular frequency ω at radius r and time t can be expressed as

$$\tilde{p}(r, \omega, t) = i \frac{Q \rho c k}{4\pi r} e^{i(\omega t - kr)}, \quad (3)$$

where Q is the scalar volume velocity of the oscillator, ρ and c are the density and speed of sound of the acoustic medium respectively, and $k = \frac{\omega}{c}$. For a sphere of radius a oscillating radially with a surface velocity of magnitude U_0 , the volume velocity can be expressed as

$$Q = 4\pi a^2 U_0 \quad (4)$$

and so equation 3 can be rewritten as

$$\tilde{p}(r, \omega, t) = i \frac{a^2 U_0 \rho c k}{r} e^{i(\omega t - kr)} = i \frac{a^2 \rho c k}{r} e^{-ikr} \tilde{u}(a, \omega, t) \quad (5)$$

where $\tilde{u}(a, t) = U_0 e^{i\omega t}$ is the complex velocity of the surface of the oscillating sphere. Under the far-field assumption, the complex pressure and complex source velocity are therefore related by a magnitude scaling factor of $\frac{a^2 \rho c k}{r}$ and a phase change of $\pi/2 - kr$.

TABLE I. Simulated system parameters

Parameter	Symbol	Value
Primary source mass	m_a	1 kg
Secondary source mass	m_b	1 kg
Primary source linear stiffness	k_a	$1.42 \times 10^5 \text{ Nm}^{-1}$
Primary source cubic stiffness	k_a^{NL}	$1.42 \times 10^{14} \text{ Nm}^{-3}$
Secondary source stiffness	k_b	$2.53 \times 10^5 \text{ Nm}^{-1}$
Primary source damping	c_a	151 Nsm^{-1}
Secondary source damping	c_b	201 Nsm^{-1}

112

113

In the simulations presented within this paper, $x(t)$ is assumed to be a band-limited Gaussian white noise with a frequency range of $[0, 250]$ Hz. To obtain the pressure produced by each of the two sources, the source velocity, $\dot{y}_a(t)$ or $\dot{y}_b(t)$, is transformed into the frequency domain using a Discrete Fourier Transform (DFT), multiplied by $i \frac{a^2 \rho c k}{r} e^{-ikr}$ as a function of frequency, and inverse Fourier transformed to recover the time-domain pressure. It is assumed that $a = 1$ for both sources, and the distances from the primary and secondary sources to the error sensor are $r_{\text{primary}} = 2$ m and $r_{\text{secondary}} = 1$ m, respectively. The system

120

dynamics are simulated in the time-domain using a 4th order Runge-Kutta method at a sample rate of $f_s = 2$ kHz.

III. PROPOSED CONTROLLER DESIGN

As noted in the Section I, it is challenging to train a single NN controller with performance that generalizes across a range of nonlinear behaviors. To overcome this challenge, a dynamic controller switching approach is proposed here where a simple switching process is utilized to select the most suitable controller from a bank of relatively small NN controllers that have been trained to perform over distinct operating ranges, as shown by the proposed controller architecture in Figure 2. Although this general approach could be used to maintain control performance when the dynamics of the system under control change due various factors, the focus here is on the case where the degree of nonlinear behavior in the system is determined by the magnitude of the signal exciting the primary system, as discussed in Section II, which is given by the motion of the floor for the considered system as shown in Figure 1. From Figure 2 it can be seen that this signal provides the sampled reference signal, $x[n]$, which is used to both drive the feedforward controller and determine the selection of the most appropriate controller from the controller bank. Specifically, the selection of the controller to be used at a given time instant is determined by comparing an estimate of the RMS of the reference signal, x_{RMS} , to the range of reference signal magnitudes over which each controller in the bank of controllers has been trained. x_{RMS} is estimated using a simple moving average process in this paper, the operation of which is determined by two parameters – t_{update} and t_{RMS} . t_{update} is the period between successive calculations of x_{RMS} , and t_{RMS} is the length

of the window used for each estimation. This moving average can be expressed as

$$x_{RMS}[n] = \sqrt{\frac{\sum_{i=n-n_{RMS}+1}^n x^2[i]}{n_{RMS}}} \quad (6)$$

where n_{RMS} is the number of digital samples in the time period t_{RMS} . The controller trained over the magnitude range containing the current estimate of the reference signal RMS is then utilized to generate the control signal, $u[n]$, for the following time period, t_{update} . Effectively, this results in the weights and biases of the NN being updated dynamically, depending on changes in the magnitude of the reference signal. Although more advanced approaches could be used for the estimation of the RMS, or indeed for selecting the highest performing controller adaptively or using a neural network as in³³ to maximize the control performance, the simple approach has been utilized here to demonstrate that the performance of the proposed strategy is not strongly reliant on a complicated estimation approach.

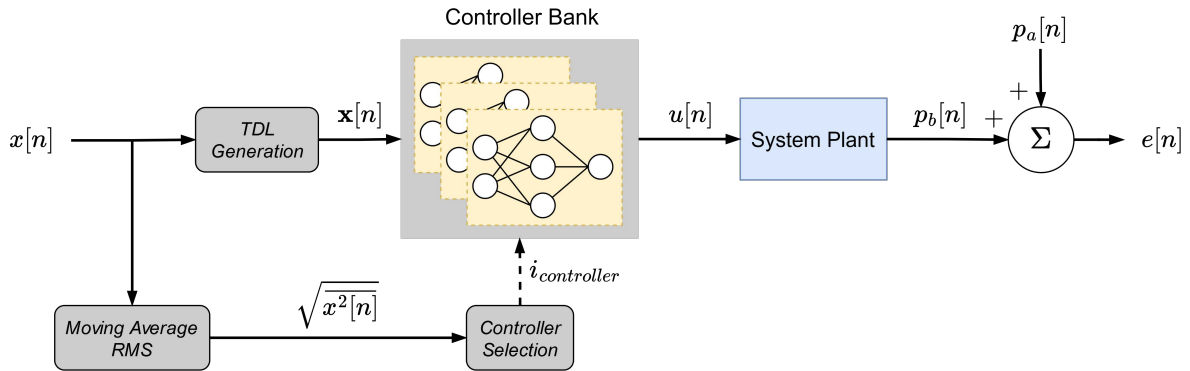


FIG. 2. Block diagram of the controller switching architecture.

In order to realize the full dynamically switched controller described above, and shown in Figure 2, it is necessary to design the individual controllers and to specify the ranges over which they are trained. The methodology utilized to train and test the individual

controllers is described in Section III A, whilst consideration of the controller training ranges is dependent on the specific system and so is discussed separately in Section IV A for the simulated nonlinear system and in Section V for the physical nonlinear system.

A. Individual Controller Architecture and Training

The focus of this work is not on the specific network architecture of the individual NN controllers, but on the dynamic switching between controllers. Therefore, all controllers have been implemented as Multi-Layer Perceptron (MLP) networks. Although alternative controller architectures could be utilized, even with the potential to mix different controllers for use over different operational ranges, this is left for future work. Each MLP controller has an input layer of size 160, with further increases having a negligible impact on controller attenuation. Each MLP uses a single hidden layer with the number of hidden nodes being variously explored, but no significant increase in performance being achieved for the considered system with more than 100 nodes in the hidden layer. A simplified diagram of the individual MLP controller network is provided in Figure 3 for example.

Similarly to a Finite Impulse Response (FIR) filter, as shown in Figure 3, the MLP controller takes a tapped delay line of the digitally sampled reference signal, $\mathbf{x}[n]$, as its input, which is given by

$$\mathbf{x}[n] = \begin{bmatrix} x[n], x[n-1], \dots, x[n-L+1] \end{bmatrix}^T, \quad (7)$$

and the output is the control signal $u[n]$. However, the MLP differs from an FIR filter in that it also contains a ‘hidden’ layer of values, $\mathbf{h}[n]$, which are calculated from the input

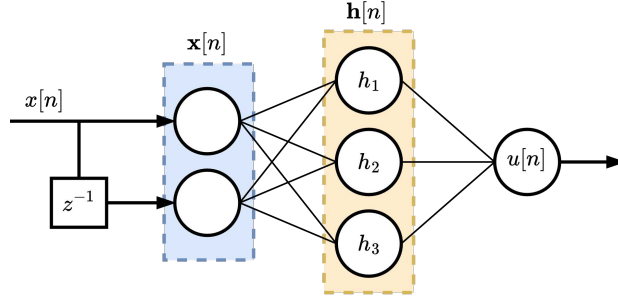


FIG. 3. Block diagram of an example MLP controller network, with an input tapped delay line size of 2, and 3 nodes in the hidden layer.

177 tapped delay line as

$$\mathbf{h}[n] = \sigma(\mathbf{W}\mathbf{x}[n] + \mathbf{b}_h) \quad (8)$$

178 where $\mathbf{W}$ is a matrix of hidden layer weights, $\mathbf{b}_h$ is a vector of biases associated with the
 179 hidden layer, and $\sigma(\bullet)$ is a nonlinear activation function that allows for the network to
 180 generate nonlinear mappings and has been defined as a hyperbolic tangent in this case. The
 181 network output, $u[n]$, is then given by

$$u[n] = \mathbf{w}_o^T \mathbf{h}[n] + b_o \quad (9)$$

182 where $\mathbf{w}_o$ is a vector of output weights, and b_o is an output bias. Combining equations 8
 183 and 9 then gives the network output as

$$u[n] = \mathbf{w}_o^T \sigma(\mathbf{W}\mathbf{x}[n] + \mathbf{b}_h) + b_o. \quad (10)$$

184 1. Controller training

185 A diagram of the architecture used to train the NN controller is presented in Figure 4. The
 186 training is undertaken using the generated reference signal, $x[n]$, and the simulated primary

187 source pressure at the error microphone, $p_a[n]$. To calculate an estimate of the error signal,
 188 a Hankel matrix, $\mathbf{X}[n]$ of size $N \times L$ is generated, where N is the tapped delay line length,
 189 or input layer size, of the MLP controller, and L is the order of the fixed plant model, $\hat{\mathbf{g}}$, in
 190 the discrete time domain. $\mathbf{X}[n]$ can be written as

$$\mathbf{X}[n] = \begin{bmatrix} x[n] & x[n-1] & \dots & x[n-N+1] \\ x[n-1] & x[n-2] & \dots & x[n-N] \\ \vdots & \vdots & \ddots & \vdots \\ x[n-L+1] & x[n-L] & \dots & x[n-N-L+2] \end{bmatrix}. \quad (11)$$

191 This matrix is passed to the MLP controller, generating a vector $\mathbf{u}[n]$ of length L which is
 192 a tapped delay line of the control signal generated by the current iteration of the controller.
 193 The vector $\mathbf{u}[n]$ is subsequently passed to the plant model, generating an estimate, $\hat{p}_b[n]$, of
 194 the pressure generated by the secondary source. An estimate of the error signal at the error
 195 microphone can then be calculated via the linear summation of the primary and estimated
 196 secondary source pressures as, $\hat{e}[n] = p_a[n] + \hat{p}_b[n]$. Each controller is trained to minimize
 197 the mean squared error (MSE) signal, which is defined as

$$J = \overline{\hat{e}^2[n]} \quad (12)$$

198 where the mean is calculated over 128 instances of the estimation of the error signal, col-
 199 lectively referred to as a mini-batch. The backpropagation algorithm used to update the
 200 controller weights and biases was the Adam algorithm³⁴ with parameters $\alpha = 1 \times 10^{-4}$,
 201 $\beta_1 = 0.9$, $\beta_2 = 0.99$, and $\epsilon = 10^{-7}$. These parameters were selected through trial and error
 202 with a view to reaching an effective trade-off between controller performance and training

speed. In all cases, the plant model used for controller training was an FIR filter with 160 taps, which was capable of achieving high levels of modeling accuracy due to the linear nature of the simulated plant response.

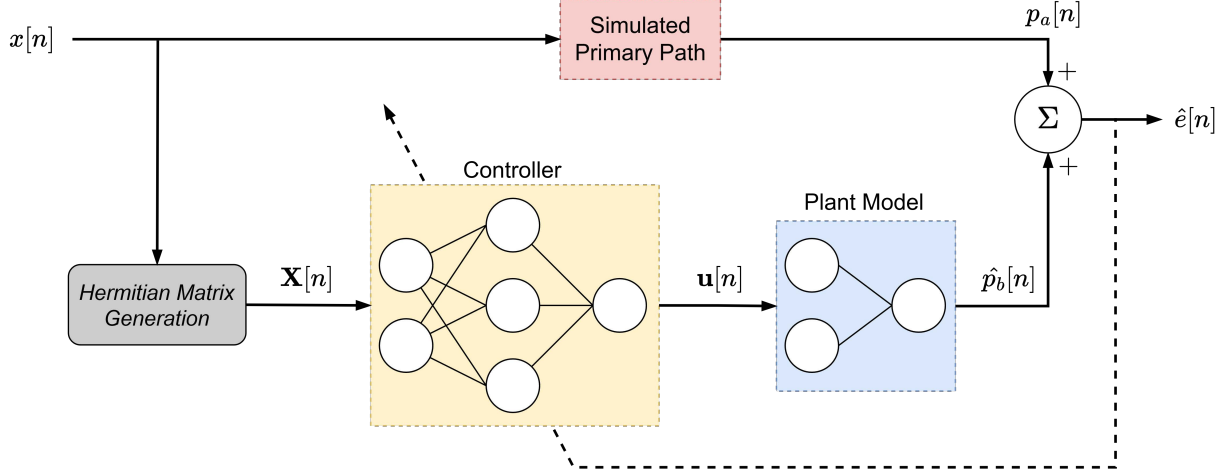


FIG. 4. Block diagram of the controller training method.

To train each network, two sets of 900 s of simulated data were generated. The first of the two datasets was used for network training, and the second for validation to assess overfitting. As a result of using this relatively large amount of training data considering the size of the networks, there was no apparent overfitting during the network training, as assessed via the training and validation losses, and, therefore, network regularization techniques were not applied. Each dataset consists of the reference signal, $x[n]$, which as noted above is a band-limited Gaussian white noise with a frequency range of $[0, 250]$ Hz, and the disturbance signal, $p_a[n]$. In each simulation, the magnitude of the reference signal, or system excitation, increases linearly over time so that it covers the targeted training range in each case. Each update to the weights and biases of the controller networks was

undertaken using an average over a mini-batch of a tapped delay line of the reference signal, and a sample of the disturbance signal. These mini-batches were selected randomly from the generated data set. In each training step, 1000 such mini-batches were used, and the full network training was undertaken over 500 steps. The term ‘step’ here is used in place of the typical term ‘epoch’ to clarify that the full dataset is not used in each training step, which is explained further below.

When training the MLP controllers it was found that, for a given controller and training range, if the random selection of the training data used to update the network weights and biases had a uniform distribution, then the control attenuation achieved was approximately equal over the training range. However, when training a controller at a single excitation level, it was found that the maximum control attenuation achievable is not uniform over excitation level and in fact decreases as the magnitude of the reference signal increases. This means that when using a uniformly distributed selection of training data to train the generalized controllers, their performance approaches the maximum at the upper end of the training range, but falls below the maximum at the bottom. For a set of N training examples with reference signal magnitudes x_{mag} in the range $a < x_{mag} < b$, the probability of training example q being included in a training batch (up to a normalizing factor) can be defined as

$$P(q) \propto 10^{-\gamma(x_{mag}-a)} \quad (13)$$

where γ is a factor controlling the shape of the probability distribution. Modifying the selection of the training data in this way affects the resultant control attenuation achieved by the MLP controllers across the training range, and an appropriate selection of γ for a given

training range results in generalized control performance that approaches the maximum MLP controller performance across the training range. This approach has previously been explored in³¹.

IV. SWITCHED CONTROLLER TUNING AND PERFORMANCE

This section presents simulation results demonstrating the performance of the proposed switch-controller approach. In the first instance the effect of the controller training range on potential performance is explored and then the performance of the switched-controller is evaluated. In all cases, simulated data not utilized in the training phase has been utilized.

A. Switched controller training range analysis

A key parameter in designing the switched-controller system is the selection of the reference signal magnitude ranges over which each controller is trained. This selection should consider both the number of training ranges utilized to cover the overall range of operational conditions and the relative widths of the training ranges. In terms of the relative widths of the training ranges, it has been found that if the overall range is subdivided equally, as shown in the upper image in Figure 5, this results in under-performance at lower magnitude system excitation levels compared to those trained at higher magnitude levels. This can be related to the fact that the range of behavior exhibited by the nonlinear system is broader for a fixed training width at lower reference signal magnitudes, primarily due to the nature of the assumed nonlinear stiffness. Motivated by this, it has been found that subdividing the overall magnitude range such that the ratio between the widths of neighboring ranges is

constant as the reference signal magnitude increases provides more consistent performance across the overall range. An example of this proposed subdivision of the full training range is shown in the lower image in Figure 5.

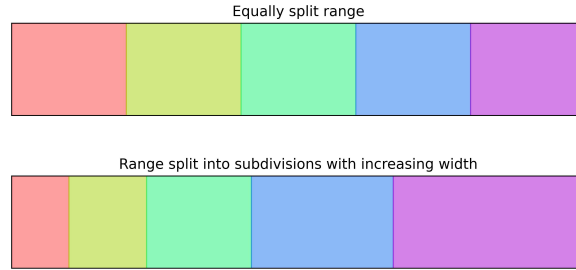


FIG. 5. Examples showing how the full training range of interest may be split into equal subdivisions, or subdivisions with increasing width. In the lower example, the ratio between the widths of neighboring ranges is constant.

To explore the effect of varying the number of training ranges, Figure 6 shows the generalized performance of MLP controllers with between 12 and 100 hidden nodes, trained over subdivisions of three, six and nine magnitude ranges. In each case, the black dashed lines represent the maximum potential control attenuation achievable, which is determined by using controllers trained and tested at individual reference signal magnitudes, and the colored lines represent the generalized performance of the individual controllers trained over the magnitude ranges denoted by the corresponding colored regions. In both cases, the controllers are tested using 300 s of newly simulated data with a constant excitation magnitude. From the results presented in Figure 6(a) [Three training ranges](#) it can be seen that with three subdivisions, even increasing the number of hidden nodes substantially from 12 to 100 does not

achieve a generalized performance that reaches the estimated maximum performance across
 the full range. Furthermore, the increase in control performance produced by increasing the
 number of hidden nodes from 50 to 100 in this case is minimal, suggesting that a further in-
 crease in the number of hidden nodes is unlikely to significantly improve performance. From
 the results presented in Figures 6(b) Six training ranges and 6(c) Nine training ranges for
 a subdivision of the overall magnitude range into six and nine training ranges respectively,
 it can be seen that increasing the number of hidden nodes in the controllers improves their
 generalized control performance and this approaches the estimated maximum performance
 as the number of hidden nodes approaches 100. This is perhaps a predictable result, as in-
 creasing the number of subdivisions of the full training range approaches the case where the
 controllers are each trained at a single level. However, comparison of the results presented in
 Figures 6(b) Six training ranges and 6(c) Nine training ranges illustrates that increasing the
 number of subdivisions does not necessarily increase the generalized control attenuation for
 a given number of hidden nodes. Moreover, as there is a computational cost and additional
 time associated with training each of the individual controllers, it is clear that for any partic-
 ular application there will be some optimal number of subdivisions of the full training range
 that maximizes generalized control attenuation whilst minimizing the number of networks
 required to be trained. Furthermore, although increasing the number of subdivisions may
 mean that the individual networks require fewer hidden nodes, narrow training regions may
 result in rapid switching between networks during control, compromising performance.

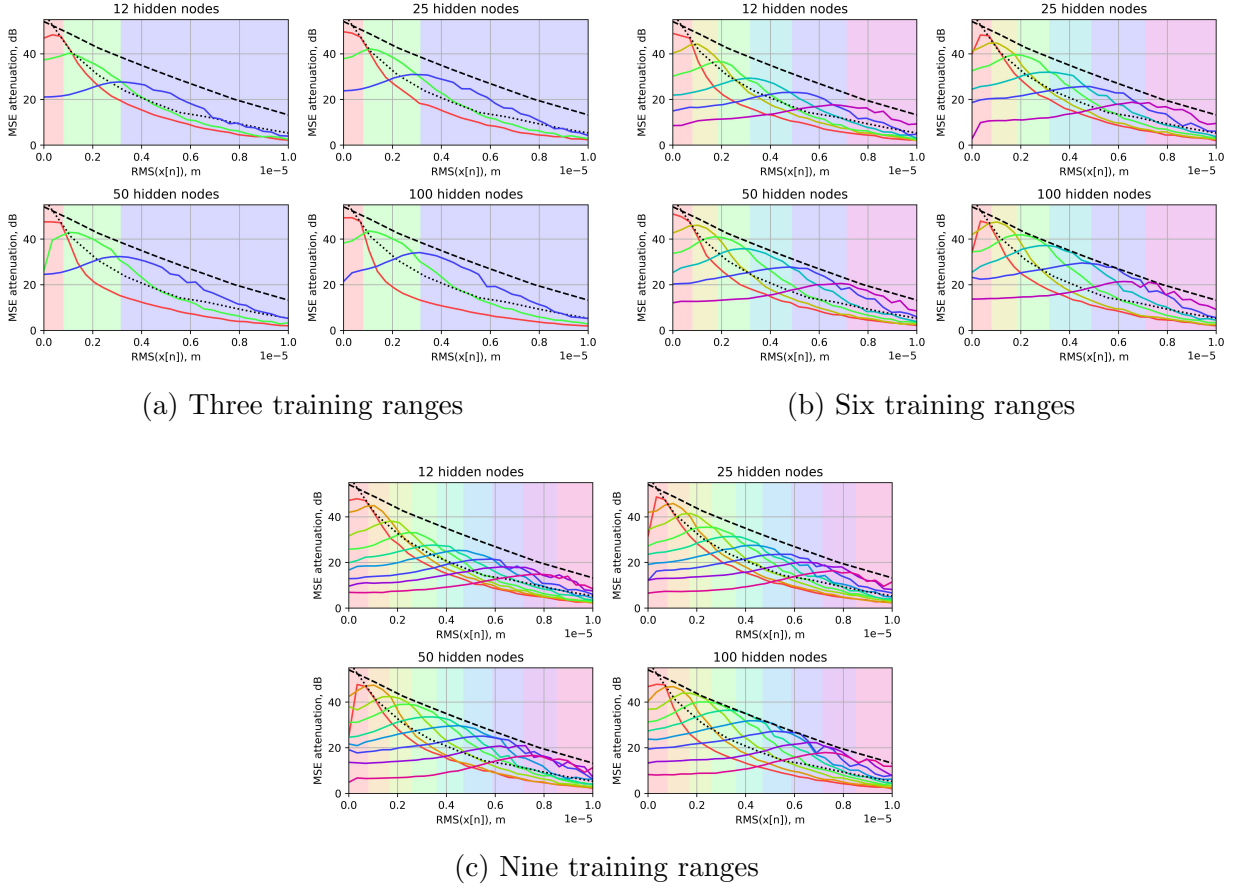


FIG. 6. Generalized performance of controllers trained across three, six and nine training ranges.

The black dashed lines indicate the maximum performance achievable by the MLP trained at discrete levels with 100 hidden nodes. The black dotted lines indicate the maximum performance achievable by an FIR controller trained at discrete levels. The colored lines indicate the generalized performance achieved by the individual controllers trained over the excitation signal magnitude ranges defined by the corresponding colored region.

B. Switched controller performance analysis

To test the performance of the switched controller, the set of six controllers with 50 hidden nodes described in Section IV A has been utilized, since it offers close to maximum

performance over the full magnitude range without unnecessarily high training costs. In this instance, the moving average process utilized to dynamically estimate the magnitude of the reference signal as described in Section III has been implemented with t_{RMS} and t_{update} set to 0.2 s. To test the performance of the switched controllers, a dynamically varying excitation condition was simulated over 60 s. Specifically, the magnitude of excitation was first slowly increased from a low to a high value, and then reduced again slowly before rapidly increasing and decreasing. The control attenuation achieved by the switched-controller is presented in Figure 7, along with the performance achieved by an adaptive Normalized-Step-Size FLANN-based (NSS-FLANN) controller³⁵ and a generalized NN trained across the full range of excitation levels. The FLANN-based controller has been included to provide a benchmark against a commonly utilized adaptive nonlinear control algorithm, whilst the larger single MLP controller has been included to demonstrate the performance of the switching approach compared to simply utilizing a larger network trained across the full range of operating conditions. From these results it can be seen that the switched MLP controller is able to achieve effective control attenuation across the range of excitation levels and as such closely tracks the maximum steady-state control attenuation for the MLP denoted by the dashed black line. In comparison to the larger generalized MLP controller and the FLANN, the proposed switched MLP controller achieves a consistently higher level of control attenuation. The performance is consistently around 10 dB above the generalized MLP controller, whilst utilizing half the number of hidden nodes and, therefore, a significantly reduce computational load. A similar performance advantage is also achieved compared to

the implemented FLANN, which can also be observed to be relatively slow to regain effective control attenuation following periods of high excitation level.

Despite the effective performance of the switched MLP controller, it is clear from the colored regions in Figure 7 that there is occasionally quite rapid switching between controllers. It should be noted that this can be addressed by adjusting the t_{update} and t_{rms} parameters used in the RMS estimation procedure. Specifically, with a fixed update rate (t_{update}), a trade-off between accuracy and speed of estimation can be realized by adjusting t_{rms} . With a larger t_{rms} the rapid switching can be reduced by obtaining a more accurate estimate of the reference signal RMS, but this will introduce a delay in the estimate and, therefore, the selection of the most appropriate controller. In the case of t_{update} it is also possible to reduce controller switching by reducing the update rate via an increase in t_{update} , but this would also reduce the ability of the switched-controller to respond to rapid changes in the excitation level. Therefore, both estimation parameters must be tuned for the application considering the required speed of controller switching to maximize control performance. Alternatively, as noted in Section III, it is possible to use a more advanced method of selecting the appropriate controller at any time instant, either via an adaptive approach or in more complicated scenarios via an intelligent approach based on machine learning, as utilized in³³ to select the most suitable controller for different types of noise. That said, it is clear from the results presented in Figure 7 that even with the controller switching based on the simple RMS estimation procedure performance close to the steady-state maximum can be achieved.

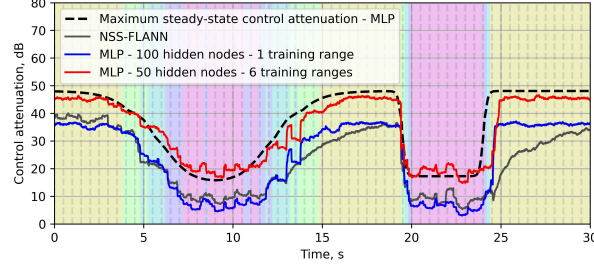


FIG. 7. The control performance achieved when the excitation level, or reference signal magnitude, varies over time for the proposed switched MLP controller, a large MLP trained over the full excitation range, and an NSS-FLANN controller. The MSE control attenuation achieved has been calculated using a 0.5 s moving average. The colored regions represent the selected controller at the corresponding time instance for the switched-controller implementation.

V. APPLICATION TO THE CONTROL OF A PHYSICAL NONLINEAR SYSTEM

To further validate the proposed control strategy, its performance when applied to a physical system with a nonlinear response has been explored. The considered experimental system consisted of a thin aluminum plate clamped along its edges by a thick aluminum frame, as shown in Figure 8. The primary excitation was provided by a small electrodynamic shaker (Tectonic Elements TEAX19C01-8) attached to the surface of the plate, which was overdriven to introduce the physical nonlinearity. The secondary source was provided by a larger electrodynamic shaker (Tectonic Elements TEAX32C30-4/B), which was also attached to the surface of the plate. The error sensor was provided by an accelerometer, located with the secondary source on the underside of the plate and therefore not visible in Figure 8. The primary source was driven by low-pass filtered random Gaussian noise, with a

350 cutoff at 250 Hz, and this signal was taken as the reference signal. The system was measured
 351 at a sample rate of 2 kHz, and the MLP controllers utilized a hyperbolic tangent activation
 352 function and a reference signal tapped delay line length of 0.15 seconds, corresponding to
 353 300 samples. The estimation of the reference signal RMS was undertaken using the moving
 354 average approach described in Section III, with $t_{update} = 0.1$ s, and $t_{RMS} = 0.5$ s in this
 355 case. The data used for the training of the generalized MLP controllers was provided by a
 356 measurement of the system excited by a signal that linearly increased in magnitude from a
 357 low level to a high level over 150 seconds, then decreased in magnitude over a further 150 s
 358 back to the lower magnitude level. The lower level limit was defined by the noise floor in the
 359 system and the upper limit was chosen to avoid destroying the electrodynamic shaker. The
 360 NN training was undertaken following the same methodology as outlined in Section III A 1.

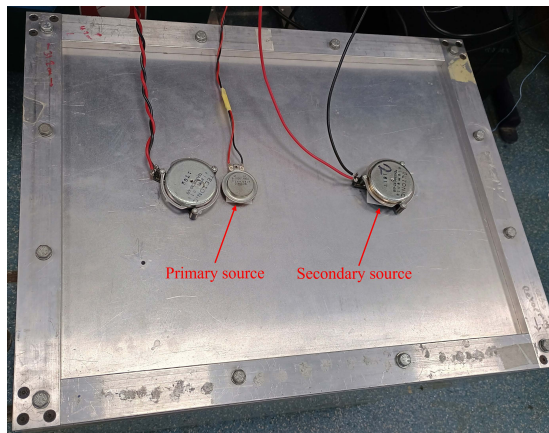


FIG. 8. A photograph of the physical system consisting of a thin aluminum plate with clamped boundaries and electrodynamic shakers providing the nonlinear primary source (left) and secondary source (right). (nb. the third shaker on the left with the same dimensions as the secondary source was not used in this study).

Figure 9 presents the generalized control attenuation calculated via offline simulations for the MLP controllers trained over 4 ranges with 40 hidden nodes, and 9 ranges with 80 hidden nodes, respectively. As was similarly noted in the simulation study presented in Section IV A, it can be seen that by increasing the number of training ranges and hidden nodes beyond 4 ranges and 40 hidden nodes, only a small increase in generalized control performance is achieved. Therefore, the controllers trained over 4 ranges with 40 hidden nodes were implemented in the following evaluation of the switched MLP controller.

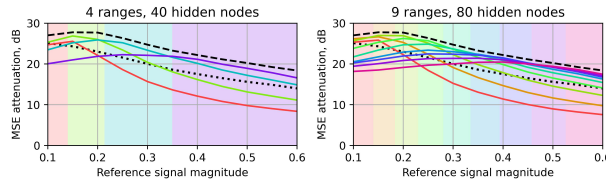


FIG. 9. Generalized performance of controllers trained across 4 ranges with 40 hidden nodes, and 9 ranges with 80 hidden nodes. The black dashed lines indicate the maximum performance achievable by the MLP trained at discrete reference signal magnitudes with 80 hidden nodes. The black dotted lines indicate the maximum performance similarly achievable with an FIR controller.

As in Section IV B, to test the performance of the switched MLP controller the excitation signal level was increased from a low level to a high level and back down again, first slowly, then quickly. The upper plot in Figure 10 presents the results of offline simulations using the measured time-history of the primary disturbance signal and the measured system responses, and shows the attenuation in the MSE achieved by the switched MLP controller, along with the attenuation achieved by the NSS-FLANN controller and a larger single MLP controller trained over the full range of excitation signal levels. These results show that the switched

MLP controller achieves a control performance that is comparable to that of the NSS-FLANN controller during the first 5 s where the excitation level is low and remains relatively constant; however, after both the slow and fast decreases in the excitation signal level (14–19 s and 24.5–30 s), the switched MLP controller shows a lower level of control performance. It is possible that this is due to some form of hysteretic behavior induced by the preceding periods of high-level excitation temporarily changing the dynamics of the system. This is suggested by the fact that during these time periods, the control attenuation shows a slow upward trend as the original system response returns. This change in dynamics was not represented in the training data, where the change in excitation signal level was extremely slow, and therefore it is perhaps unsurprising that the MLP controllers underperform after these changes. To overcome this limitation, rather than retraining the MLP controllers with different excitation signal time-histories, a small augmentation was made to the controller switching approach. Specifically, an adaptive output gain μ was implemented for each MLP controller, with the gain updated at each sample according to

$$\Delta\mu \propto -e[n]\mathbf{g}^T\mathbf{u}[n] \quad (14)$$

where $\mathbf{g}$ is an FIR filter modeling the system plant, and $\mathbf{u}[n]$ is a tapped delay line of the previous MLP outputs. It may be noted that, as this update term depends only on the current error signal sample, an FIR plant model and a TDL of the MLP output, this adaptation will typically come at only a small increase in computational cost. Moreover, the output gain associated with each controller is only updated when that controller is implemented to avoid unnecessary computation. The MSE attenuation levels achieved when the switched MLP controller and the larger single range MLP controller are implemented

with an adaptive output gain are presented in the lower plot of Figure 10. From these results it can be seen that this simple modification produces a considerable increase in the control performance in both cases. The switched MLP controller with an adaptive output gain now closely follows the MLP maximum attenuation curve, whilst the larger MLP controller trained over the full range now achieves a comparable performance to the NSS-FLANN.

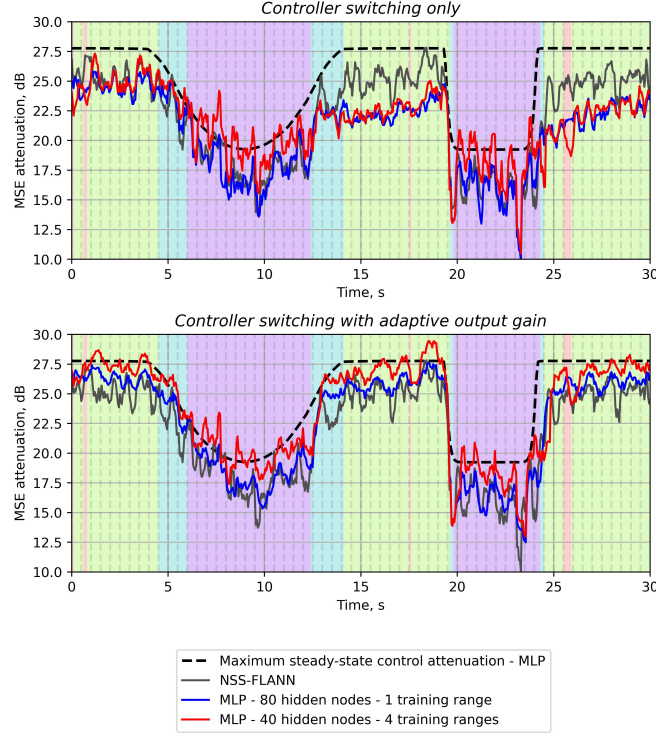


FIG. 10. The control performance achieved for the physical nonlinear system when the excitation level, or reference signal magnitude, varies over time for the proposed switched MLP controller, a large MLP trained over the full excitation range, and an NSS-FLANN controller. The MSE control attenuation achieved has been calculated using a 0.5 s moving average. The colored regions represent the selected controller at the corresponding time instance for the switched-controller implementation.

401 VI. CONCLUSIONS

402 In this paper, a method of switching between a set of pre-trained MLP controllers to
 403 improve the control performance achievable for a nonlinear system excited across a range
 404 of levels has been presented. Its potential performance has been demonstrated via a set
 405 of discrete-time simulations representing the active control of noise produced by a simple
 406 nonlinear system and also via a set of offline simulations using data measured for a more
 407 complicated physical nonlinear system. For the presented switched MLP controller, a sim-
 408 ple RMS estimator of the magnitude of the reference signal has been used to select from
 409 a set of MLP controllers trained to have near-maximal control performance over a set of
 410 system excitation magnitude ranges. The effect of increasing the number of training ranges
 411 on controller performance has been explored, and the trade-off between computational cost
 412 and performance has been discussed. In the numerical simulations of the simple nonlinear
 413 system, the switching approach has been shown to achieve control performance that approx-
 414 imates the maximum achievable performance as the reference signal magnitude varies over
 415 time. To further validate the proposed control strategy, an offline control simulation has
 416 been undertaken using measurements of a physical system with a nonlinear response, which
 417 includes both saturation and hysteretic like nonlinear behaviors. In this more realistic case,
 418 the proposed switched MLP controller achieved control performance that was significantly
 419 lower than the maximum potential attenuation and was outperformed by a well-known adap-
 420 tive nonlinear control strategy. This was noted to be possibly due to changes in the system
 421 dynamics caused by high excitation levels which were not included in the MLP training data.

However, by introducing a simple and computationally efficient adaptive output gain into the MLP controllers, it has been shown that the performance of the switched MLP controller could be significantly increased, outperforming both the well-known adaptive nonlinear controller and a single fixed larger MLP trained across the full range of excitation levels.

ACKNOWLEDGMENTS

This work was partially supported by the Intelligent Structures for Low Noise Environments EPSRC Prosperity Partnership (EP/S03661X/1) and by the Department of Science, Innovation and Technology (DSIT) Royal Academy of Engineering under the Research Chairs and Senior Research Fellowships programme. The authors also acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work.

CONFLICT OF INTEREST

JC and XP have patent WO 2025/012603/A1 pending.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- ¹X. Pike and J. Cheer, “Limitations of FxLMS in feedforward active vibration control of a nonlinear two-degree-of-freedom system,” INTER-NOISE and NOISE-CON Congress and Conference Proceedings **265**(2), 5219–5229 (2023) doi: [10.3397/in_2022_0761](https://doi.org/10.3397/in_2022_0761).
- ²M. Costa, J. Bermudez, and N. Bershad, “Stochastic analysis of the filtered-x LMS algorithm in systems with nonlinear secondary paths,” IEEE Transactions on Signal Processing **50**(6), 1327–1342 (2002) doi: [10.1109/tsp.2002.1003058](https://doi.org/10.1109/tsp.2002.1003058).
- ³O. Tobias and R. Seara, “On the LMS algorithm with constant and variable leakage factor in a nonlinear environment,” IEEE Transactions on Signal Processing **54**(9), 3448–3458 (2006) doi: [10.1109/tsp.2006.879274](https://doi.org/10.1109/tsp.2006.879274).
- ⁴S. Snyder and N. Tanaka, “Active control of vibration using a neural network,” IEEE Transactions on Neural Networks **6**(4), 819–828 (1995) doi: [10.1109/72.392246](https://doi.org/10.1109/72.392246).
- ⁵L. Tan and J. Jiang, “Adaptive volterra filters for active control of nonlinear noise processes,” IEEE Transactions on Signal Processing **49**(8), 1667–1676 (2001) doi: [10.1109/78.934136](https://doi.org/10.1109/78.934136).
- ⁶E. Reddy, D. Das, and K. Prabhu, “Fast exact multichannel FSLMS algorithm for active noise control,” Signal Processing **89**(5), 952–956 (2009) doi: [10.1016/j.sigpro.2008.11.013](https://doi.org/10.1016/j.sigpro.2008.11.013).
- ⁷C. Wangler and C. Hansen, “Genetic algorithm adaptation of non-linear filter structures for active sound and vibration control,” in *Proceedings of ICASSP '94. IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE (1994), doi:

459 [10.1109/icassp.1994.389979](https://doi.org/10.1109/icassp.1994.389979).

460 ⁸Q.-Z. Zhang and W.-S. Gan, “Active noise control using a simplified fuzzy neural network,”

461 Journal of Sound and Vibration **272**(1-2), 437–449 (2004) doi: [10.1016/s0022-460x\(03\)](https://doi.org/10.1016/s0022-460x(03)00742-9)

462 [00742-9](https://doi.org/10.1016/s0022-460x(03)00742-9).

463 ⁹S. Chen and S. A. Billings, “Neural networks for nonlinear dynamic system modelling

464 and identification,” International journal of control **56**(2), 319–346 (1992) doi: [10.1080/](https://doi.org/10.1080/00207179208934317)

465 [00207179208934317](https://doi.org/10.1080/00207179208934317).

466 ¹⁰K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are univer-

467 sal approximators,” Neural Networks **2**(5), 359–366 (1989) doi: [10.1016/0893-6080\(89\)](https://doi.org/10.1016/0893-6080(89)90020-8)

468 [90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).

469 ¹¹M. Bouchard, B. Paillard, and C. L. Dinh, “Improved training of neural networks for the

470 nonlinear active control of sound and vibration,” IEEE Transactions on Neural Networks

471 **10**(2), 391–401 doi: [10.1109/72.750568](https://doi.org/10.1109/72.750568).

472 ¹²K. Narendra and K. Parthasarathy, “Identification and control of dynamical systems using

473 neural networks,” IEEE Transactions on Neural Networks **1**(1), 4–27 doi: [10.1109/72.](https://doi.org/10.1109/72.80202)

474 [80202](https://doi.org/10.1109/72.80202).

475 ¹³D. Hong, H. Lee, Y. Han, and B. Kim, “Numerical feasibility study for transverse vibration

476 control of rotating shaft with a neural network-based tracking algorithm,” INTER-NOISE

477 and NOISE-CON Congress and Conference Proceedings **263**(5), 1293–1298 (2021) doi:

478 [10.3397/in-2021-1803](https://doi.org/10.3397/in-2021-1803).

¹⁴H.-S. Kim, “Development of seismic response simulation model for building structures with semi-active control devices using recurrent neural network,” *Applied Sciences* **10**(11), 3915 doi: [10.3390/app10113915](https://doi.org/10.3390/app10113915).

¹⁵J. Y. Oh, H. W. Jung, M. H. Lee, K. H. Lee, and Y. J. Kang, “Enhancing active noise control of road noise using deep neural network to update secondary path estimate in real time,” *Mechanical Systems and Signal Processing* **206**, 110940 (2024) doi: [10.1016/j.ymssp.2023.110940](https://doi.org/10.1016/j.ymssp.2023.110940).

¹⁶C.-Y. Chang and F.-B. Luoh, “Enhancement of active noise control using neural-based filtered-x algorithm,” *Journal of Sound and Vibration* **305**(1), 348–356 doi: [10.1016/j.jsv.2007.04.007](https://doi.org/10.1016/j.jsv.2007.04.007).

¹⁷D. Chen, L. Cheng, D. Yao, J. Li, and Y. Yan, “A secondary path-decoupled active noise control algorithm based on deep learning,” *IEEE Signal Processing Letters* **29**, 234–238 (2021) doi: [10.1109/LSP.2021.3130023](https://doi.org/10.1109/LSP.2021.3130023).

¹⁸Y. Guo, D. Shi, X. Shen, J. Ji, and W.-S. Gan, “A survey on adaptive active noise control algorithms overcoming the output saturation effect,” *Signal Processing* 109525 (2024) doi: [10.1016/j.sigpro.2024.109525](https://doi.org/10.1016/j.sigpro.2024.109525).

¹⁹Y. Yan, L. Dong, Y. Han, and W. Li, “A general inverse control model of a magnetorheological damper based on neural network,” *Journal of Vibration and Control* **28**(7), 952–963 doi: [10.1177/1077546320986380](https://doi.org/10.1177/1077546320986380).

²⁰K. Na and S.-I. Chae, “Single-sensor active noise cancellation using recurrent neural network predictors,” in *Proceedings of International Conference on Neural Networks*

(*ICNN'97*), IEEE, Vol. 4, pp. 2153–2156, doi: [10.1109/ICNN.1997.614239](https://doi.org/10.1109/ICNN.1997.614239).

²¹C. Chen and T.-D. Chiueh, “Multilayer perceptron neural networks for active noise cancellation,” in *1996 IEEE International Symposium on Circuits and Systems. Circuits and Systems Connecting the World. ISCAS 96*, IEEE, Vol. 3, pp. 523–526, doi: [10.1109/ISCAS.1996.541648](https://doi.org/10.1109/ISCAS.1996.541648).

²²M. Salmasi, H. Mahdavi-Nasab, and H. Pourghassem, “Comparison of feed-forward and recurrent neural networks in active cancellation of sound noise,” in *2011 International Conference on Multimedia and Signal Processing*, IEEE, pp. 25–29, doi: [10.1109/CMSP.2011.96](https://doi.org/10.1109/CMSP.2011.96).

²³T. Matsuura, T. Hiei, H. Itoh, and K. Torikoshi, “Active noise control by using prediction of time series data with a neural network,” in *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*, IEEE, Vol. 3, pp. 2070–2075, doi: [10.1109/ICSMC.1995.538084](https://doi.org/10.1109/ICSMC.1995.538084).

²⁴K. Hiramoto and T. Matsuoka, “Active vibration control of structural systems with a preview of a future seismic waveform generated by remote waveform observation data and an artificial intelligence-based waveform estimation system,” *Journal of Vibration and Control* **26**(17), 1602–1613 (2020) doi: [10.1177/1077546319901024](https://doi.org/10.1177/1077546319901024).

²⁵Y.-J. Cha, A. Mostafavi, and S. S. Benipal, “Dnoisenet: Deep learning-based feedback active noise control in various noisy environments,” *Engineering Applications of Artificial Intelligence* **121**, 105971 (2023) doi: [10.1016/j.engappai.2023.105971](https://doi.org/10.1016/j.engappai.2023.105971).

- ²⁶Q. Wang, J. Wang, X. Huang, and L. Zhang, “Semiaactive nonsmooth control for building structure with deep learning,” *Complexity* 1–8 (2017) doi: [10.1155/2017/6406179](https://doi.org/10.1155/2017/6406179).
- ²⁷J. Liu, X. Li, X. Zhang, and X. Chen, “Modeling and simulation of energy-regenerative active suspension based on BP neural network PID control,” *Shock and Vibration* 1–8 (2019) doi: [10.1155/2019/4609754](https://doi.org/10.1155/2019/4609754).
- ²⁸H. Zhang and D. Wang, “Deep ANC: A deep learning approach to active noise control,” *Neural Networks* **141**, 1–10 (2021) doi: [10.1016/j.neunet.2021.03.037](https://doi.org/10.1016/j.neunet.2021.03.037).
- ²⁹H. Zhang and D. Wang, “Deep MCANC: A deep learning approach to multi-channel active noise control,” *Neural Networks* **158**, 318–327 (2023) doi: [10.1016/j.neunet.2022.11.029](https://doi.org/10.1016/j.neunet.2022.11.029).
- ³⁰R. Ranjan, J. He, T. Murao, L. Bhan, and W.-S. Gan, “Selective active noise control system for open windows using sound classification,” in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (2016), Vol. 253, pp. 1921–1931.
- ³¹X. Pike and J. Cheer, “Generalised performance of neural network controllers for feed-forward active noise control of nonlinear systems,” *Acta Acustica* **8**, 51 (2024) doi: [10.1051/aacus/2024024](https://doi.org/10.1051/aacus/2024024).
- ³²An earlier version of this work has been published in the POMA article³⁶.
- ³³Z. Luo, D. Shi, and W.-S. Gan, “A hybrid sfanc-fxlms algorithm for active noise control based on deep learning,” *IEEE Signal Processing Letters* **29**, 1102–1106 (2022) doi: [10.1109/LSP.2022.3169428](https://doi.org/10.1109/LSP.2022.3169428).

- ³⁴D. Kingma and J. Ba, “Adam: A method for stochastic optimization” (2014), doi: [10.48550/ARXIV.1412.6980](https://doi.org/10.48550/ARXIV.1412.6980).
- ³⁵D. Das and G. Panda, “Active mitigation of nonlinear noise processes using a novel filtered-s lms algorithm,” *IEEE Transactions on Speech and Audio Processing* **12**(3), 313–322 (2004) doi: [10.1109/TSA.2003.822741](https://doi.org/10.1109/TSA.2003.822741).
- ³⁶X. Pike and J. Cheer, “Dynamic neural network switching for active noise control of nonlinear systems,” *Proceedings of Meetings on Acoustics* **52**(1), 055007 (2024) doi: [10.1121/2.0001855](https://doi.org/10.1121/2.0001855).

548 **REFERENCES**

549 ¹An earlier version of this work has been published in the POMA article^{[36](#)}.