

Article

# A Novel High-Speed and High-Accuracy Mathematical Modeling Method of Complex MEMS Resonator Structures Based on the Multilayer Perceptron Neural Network

Qingsong Li <sup>1,2,†</sup> , Kuo Lu <sup>1,†</sup>, Kai Wu <sup>1</sup>, Hao Zhang <sup>1</sup>, Xiaopeng Sun <sup>1</sup>, Xuezhong Wu <sup>1,2,3</sup> and Dingbang Xiao <sup>1,2,3,\*</sup>

- <sup>1</sup> College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China; liqingsong12@nudt.edu.cn (Q.L.); lukuo13@nudt.edu.cn (K.L.); wukai@nudt.edu.cn (K.W.); zhanghao19@nudt.edu.cn (H.Z.); sxp@nudt.edu.cn (X.S.); xzww@nudt.edu.cn (X.W.)
- <sup>2</sup> Hunan MEMS Research Center, Changsha 410073, China
- <sup>3</sup> Laboratory of Science and Technology on Integrated Logistics Support, National University of Defense Technology, Changsha 410073, China
- \* Correspondence: dingbangxiao@nudt.edu.cn; Tel.: +86-0731-8457-4958
- † Qingsong Li and Kuo Lu contributes equally to this work.



**Citation:** Li, Q.; Lu, K.; Wu, K.; Zhang, H.; Sun, X.; Wu, X.; Xiao, D. A Novel High-Speed and High-Accuracy Mathematical Modeling Method of Complex MEMS Resonator Structures Based on the Multilayer Perceptron Neural Network. *Micromachines* **2021**, *12*, 1313. <https://doi.org/10.3390/mi12111313>

Academic Editors: Marc Desmulliez and David Flynn

Received: 13 September 2021  
Accepted: 24 October 2021  
Published: 26 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** MEMS resonators have become core devices in a large number of fields; however, due to their complex structures, the finite element analysis (FEA) method is still the main method for their theoretical analysis. The traditional finite element analysis method faces the disadvantages of large calculation amount and long simulation time, which limits the development of high-performance MEMS resonators. This paper demonstrates a high-speed and high-accuracy simulation tool based on the artificial neural network, where a multilayer perceptron (MLP) neural network model is constructed. The typical structural parameters of MEMS resonator are used as the input layer, and its performance indicators produced by the finite element analysis method are the output layer. After iteratively trained with 4000 samples, the cumulative error of the neural network decreases to 0.0017 and a prediction network model is obtained. Compared with the finite element analysis results, the structural accuracy error predicted by the neural network model can be controlled within 6%, but its runtime is shortened by 15,000 times. This high-speed and high-accuracy mathematical modeling method can effectively improve the analyzing efficiency and provide a promising tool for the design and optimization of different complex MEMS resonators, which exhibit remarkable accuracy and speed.

**Keywords:** mathematical modeling; finite element analysis; artificial neural network; multilayer perceptron neural network; structural optimization design

## 1. Introduction

With the development of the modern technology, micro-electro-mechanical system (MEMS) devices have become an important branch of the core sensor field [1–9]. Due to its advantages of small size, low cost and mass production, MEMS devices have attracted the attention of many research teams and commercial companies, dramatically increasing their demands [10,11]. Among these devices, the structure of MEMS resonators is the core, which determines their performance. Therefore, analyzing and evaluating the performance of MEMS resonator structures is a key step to promote the application and development of MEMS devices, which is also the direction that many research teams are committed to study and has been reported in many literatures [12]. However, the complex topological structure and numerous parameters make this research particularly difficult.

According to previous research, the performance of the MEMS resonators is usually obtained by finite element analysis method with related simulation software. The finite

element analysis is a numerical solution method that uses mathematical approximations to simulate real physical systems [13,14]. However, as for a complex MEMS resonator structure, under the combined influence of topological structure and mesh density, the finite element analysis method will take a long time to do the simulation, which severely limits the progress rate of structural design. As a result, the development of a novel high-speed and high-accuracy mathematical modeling method for complex MEMS resonator structures is urgent.

The artificial neural network (ANN) is an algorithmic mathematical model that imitates the behavioral characteristics of animal neural networks and performs distributed and parallel information processing [15–17]. It has preliminary self-adaptation and self-organization capabilities. In the learning or training process, by changing synaptic weight values, this model can be adapted to the requirements of the environment. As a result, the artificial neural network and the finite element analysis method can be combined flexibly. The artificial neural network model is trained on the basis of an appropriate amount of finite element analysis result dataset, so that it can accurately predict the finite element analysis process and provide an efficient and convenient simulation tool. In this case, it can effectively replace the long-term finite element simulation process and improve the efficiency of the analysis.

In this paper, a high-speed and high-accuracy method based on the multilayer perceptron neural network is put forward to mathematically simulate complex MEMS resonator structures. The disk MEMS resonator is taken as the research object. It is a complex topological structure determined by multiple structural parameters, with typical representativeness and research value. In the following sections, the basic structural parameters and performance indicators of the disk MEMS resonator are analyzed. At the same time, the finite element analysis method is introduced and used to simulate the disk MEMS resonators' performance, to provide enough training and testing data. Then, the theoretical analysis of the multilayer perceptron neural network and its model structure are studied. Furthermore, the structural performance indicators learned by the neural network according to the prepared dataset are compared with the finite element analysis results. Finally, the basic principles for the application of the multilayer perceptron neural network in MEMS resonators are concluded.

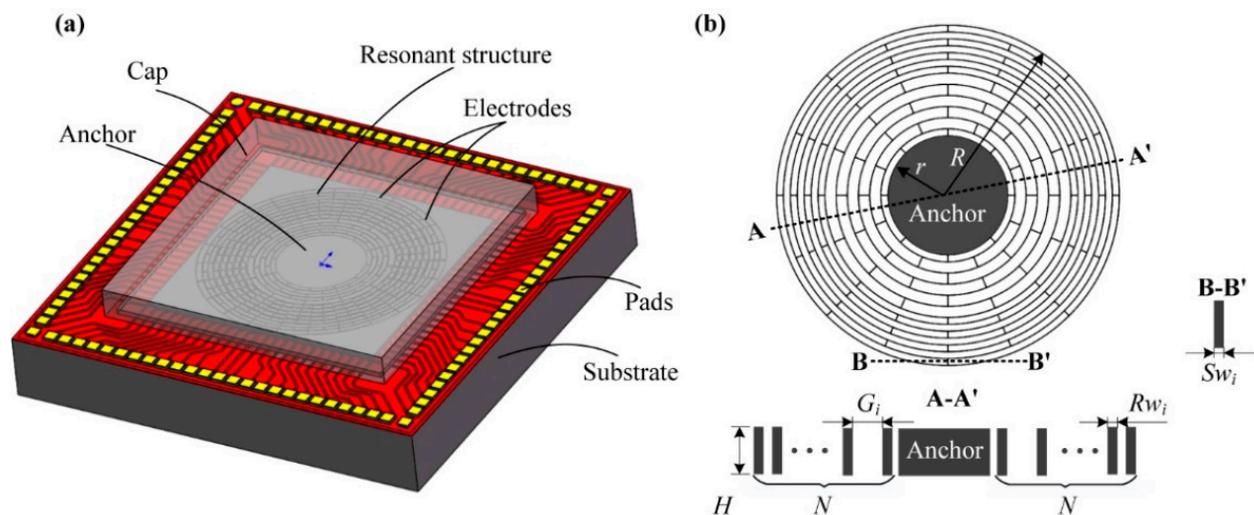
## 2. Disk MEMS Resonator

### 2.1. Structure Description

The disk MEMS resonator is a typical distributed mass structure, which is composed of multiple connected beams and rings [18,19]. Its key structural parameters include height, spoke width, ring width, ring gap, anchor radius and ring number, which contribute to its final performance indicators. Numerous structural parameters make the structure of the disk MEMS resonator relatively complex and hard to be modeled. The structure of a typical wafer-level disk MEMS resonator is shown in Figure 1a, which is mainly composed of a substrate, an oxidizing isolation layer, a resonant structure, and multiple electrodes [20]. The silicon-based substrate is used to carry the resonant structure and electrodes. Its oxidation isolation layer is used to be insulated. The resonant structure is the core component of the resonator, as shown in Figure 1b, which is mainly connected by a plurality of concentric rings and spokes. The gaps between those rings are used to place electrodes to drive, detect and tune the disk MEMS resonator.

Obviously, the structure of the disk MEMS resonator is relatively complicated, with numerous structural parameters. In order to facilitate the subsequent analysis, the main structural parameters are defined as follows: the thickness of the structure layer  $H$ , the number of resonant rings  $N$ , the ring thickness  $Rw$ , the gap between rings (spoke length)  $G$ , the radius of the anchor  $r$ , and the width of the spoke  $Sw$ . It is worth noting that because the structure of disk MEMS resonator is composed of multiple nested rings, a typical structure contains multiple  $Rw$ ,  $G$ , and  $Sw$ . These values at the position of the  $i$ th ring can be expressed as  $Rw_i$ ,  $G_i$ , and  $Sw_i$ , where  $i = 1, 2, \dots, N$ ; these values can be

different from each other. Through these parameters, a typical disk MEMS resonator can be completely designed.



**Figure 1.** The schematic diagram of the disk MEMS resonator. (a) The wafer-level disk MEMS resonator; (b) the resonant structure of the disk MEMS resonator.

In addition, it is reasonable that the structural parameters of disk MEMS resonators are not unlimited. According to prior knowledge and previous experience, these structural parameters can be constrained in reasonable value ranges. Using values in the interval for structural design can be purposeful and directional, and can effectively reduce the number of trial and error. The value ranges of different structural parameters of the disk MEMS resonator used in the design process of this paper are shown in Table 1.

**Table 1.** The structural parameter range table of the disk MEMS resonator.

Structural Parameter	Symbol	Value Range	Unit
Height	$H$	100:100:500	$\mu\text{m}$
Anchor radius	$r$	1.5:0.5:5	mm
Ring number	$N$	5:1:20	1
Ring width	$Rw$	3:2:30	$\mu\text{m}$
Ring gap	$G$	100:20:600	$\mu\text{m}$
Spoke width	$Sw$	3:2:30	$\mu\text{m}$

The complex structure and numerous parameters pose a huge challenge to the disk MEMS resonator’s design and analysis process. To achieve a reasonable design and improve the resonator’s performance, an accurate performance evaluation is required.

### 2.2. Core Performance Indicators

The characteristics of the resonant structure determine the ultimate performance of the MEMS resonator. As a typical representative, the performance of a disk MEMS resonator can be mainly characterized by its mechanical structure performance indicators, thermodynamic performance indicators, and electrical performance indicators [12].

#### 2.2.1. Fundamental Frequency

The disk MEMS resonator is a typical representative of MEMS sensors that works in a specific operating mode [21]. The fundamental frequency of the operating mode is one of the core performance indicators. It is mainly determined by the resonator’s material properties and structural characteristics. When the disk MEMS resonator is excited by the resonance frequency of its operating mode, its dynamic response can be amplified and its

sensitivity can be improved. For a common traditional mass-stiffness system, its resonant frequency can be calculated as:

$$f_i = \sqrt{\frac{k_i}{m_i}} \quad (1)$$

### 2.2.2. Quality Factor

In addition, the quality factor ( $Q$ ) is the key indicator to evaluate the dynamic characteristics of the resonator, and a resonator with a higher quality factor will have better performance characteristics. It is an evaluation of the energy dissipation and damping characteristics in the micromechanical resonator system, which is mainly affected by the damping loss. The energy of the resonator will be dissipated in the form of heat during a period of movement. There are different energy loss mechanisms in MEMS resonators, including thermoelastic damping, air damping, anchor loss, surface loss, etc. For the disk MEMS resonator, in order to improve its vibration stability and reduce the power consumption of the system, it is generally expected that it has a higher quality factor characteristic. It can be obtained from Equation (2) that the total quality factor is limited to the smallest one of its many quality factors:

$$\frac{1}{Q} = \frac{1}{Q_{Thermoelastic}} + \frac{1}{Q_{air}} + \frac{1}{Q_{anchor}} + \frac{1}{Q_{surface}} + \frac{1}{Q_{other}} \quad (2)$$

For the disk MEMS resonator, due to its symmetrical planar structure, its support loss can be neglected; at the same time, since the MEMS resonator usually adopts a mature high-vacuum packaging process, its air loss can also be controlled in a very small range. Therefore, the thermoelastic damping is the key factor, which could be explained by the Zener's formalism [22]. It is mainly determined by the resonant structure parameters and material properties [23,24]. Therefore, in order to improve the theoretical quality factor of the disk MEMS resonator, optimizing its geometry and size is a common method to reduce the thermoelastic damping.

### 2.2.3. Mechanical Sensitivity

The output of the disk MEMS resonator contains effective signals and error signals, so that one of the designing goals is to improve the resonator's signal-to-noise ratio, which could be described by the mechanical sensitivity. The mechanical sensitivity is an important parameter that characterizes the resolution of the detection axis displacement of the resonator in the force feedback control mode. It determines the accuracy of the resonator system. The expression of the mechanical sensitivity of the classic disk MEMS resonator can be expressed in Equation (3) [25]:

$$S_{mech} = \frac{2k|\bar{x}|}{\sqrt{\left(\frac{\omega_0^2}{\omega_d} - \omega_d\right)^2 + \frac{4}{\tau_0^2}}} \quad (3)$$

where  $k$  is the Coriolis coupling coefficient,  $\bar{x}$  is the amplitude of the resonator when it is working,  $\omega_0$  is the natural frequency of the working mode,  $\omega_d$  is the frequency of the driving signal, and  $\tau_0$  is the decay time constant of the working mode, which is related to the quality factor of the resonator.

### 2.2.4. Mechanical Thermal Noise

The resolution of the disk MEMS resonator is mainly determined by the mechanical thermal noise and the circuit noise. The mechanical thermal noise determines the limit resolution of the disk MEMS resonator and is closely related to its structural parameters and material properties. The mechanical thermal noise is random motion caused by the

Brownian thermal motion of the MEMS resonator structure itself, the expression of which is shown in Equation (4) [26]:

$$\Omega_{Brown} = \frac{1}{k|\bar{x}|} \sqrt{\frac{2k_B T}{k_{eff} \tau_0}} \quad (4)$$

where  $k_B$  is the Boltzmann's constant,  $T$  is the absolute temperature, and  $k_{eff}$  is the equivalent stiffness coefficient of the working mode of the resonator.

In summary, the core performance indicators of the disk MEMS resonant gyroscope studied in this paper are shown in Table 2.

**Table 2.** The core performance indicators table of the disk MEMS resonator.

Core Performance Indicators	Symbol	Unit
Fundamental frequency	$f_0$	Hz
Quality factor	$Q$	1
Mechanical sensitivity	$S_{mech}$	m/(°/s)
Mechanical thermal noise	$\Omega_{brown}$	°/√h

### 2.3. Finite Element Analysis Method

In the designing process, it is necessary to obtain theoretical performance indicators of the disk MEMS resonator. These performance indicators can be used to obtain feedback and optimize the structural design parameters. Therefore, the finite element analysis method is applied to provide related simulation research.

The basic idea of the finite element analysis method is to discretize the structure and use a finite number of easy-to-analyze elements to represent complex objects. The elements are connected to each other through a finite number of nodes, and then they are comprehensively solved according to the deformation coordination conditions. However, the meshing density has a great influence on the accuracy of simulating results. The higher accuracy of finite element analysis, the more computing time and computing resources are required.

In this paper, the finite element analysis simulation is used to provide enough simulation results for different disk MEMS resonators, which will be used as the training dataset and testing dataset. To make samples more representative and disorderly, the structural parameters of the disk resonator studied in the paper are randomly selected from the corresponding value range, and then simulated by simulation software to extract its performance indicators. In order to visually show the changes in disk MEMS resonators' structures, some typical disk resonator structures with only one single structural parameter changed are selected and displayed in Figure 2. In the actual simulation process, the structural parameters of the disk resonator are changed together by multiple parameters, and are not limited to changing only one single parameter. At the same time, their changing order and trend are also random.

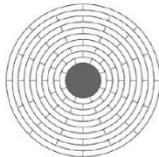
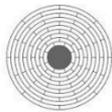
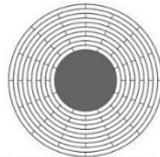
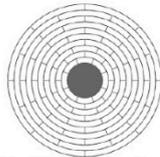
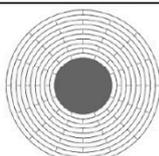
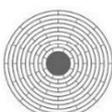
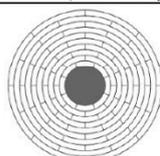
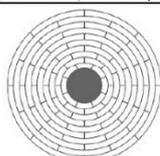
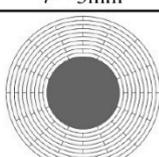
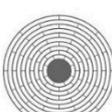
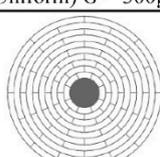
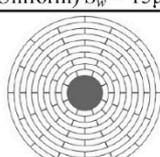
Changing height	Changing anchor radius	Changing ring number	Changing ring width	Changing ring gap	Changing spoke width
 $H = 100\mu\text{m}$	 $r = 1.5\text{mm}$	 $N = 5$	 (Uniform) $R_w = 10\mu\text{m}$	 (Uniform) $G = 100\mu\text{m}$	 (Uniform) $S_w = 5\mu\text{m}$
 $H = 300\mu\text{m}$	 $r = 3\text{mm}$	 $N = 10$	 (Uniform) $R_w = 20\mu\text{m}$	 (Uniform) $G = 300\mu\text{m}$	 (Uniform) $S_w = 15\mu\text{m}$
 $H = 500\mu\text{m}$	 $r = 4.5\text{mm}$	 $N = 15$	 (Uniform) $R_w = 30\mu\text{m}$	 (Uniform) $G = 500\mu\text{m}$	 (Uniform) $S_w = 25\mu\text{m}$

Figure 2. Typical structure examples of the disk MEMS resonator with only one single structural parameter changed.

### 3. Multilayer Perceptron Neural Network Model

A neural network is a widely parallel interconnected network composed of adaptable simple units [27]. According to the input of the system, by constructing a neural network structure and adjusting the distribution weight of neurons and the number of propagation layers, the corresponding system response output can be obtained. Based on the characteristics of the neural network, it can be applied to other fields of engineering and scientific research.

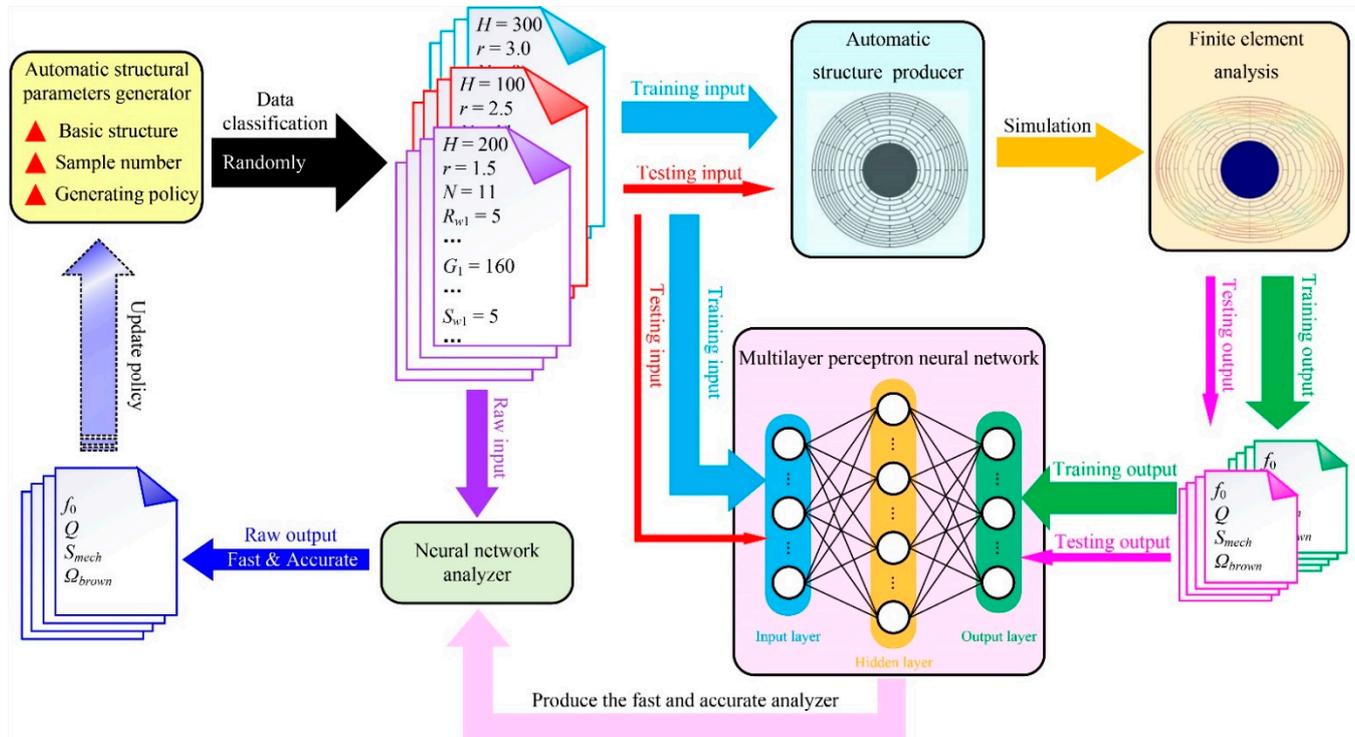
Through extensive training datasets, the neural network structure is optimized, and the ideal output value can be reasonably predicted based on the input value. Therefore, proper neural network structure parameters are the key to the high speed and high accuracy of prediction. Due to the complex structure, it is extremely difficult to manually modify and optimize its structure parameters, and machine learning methods can be used to optimize it [28]. The machine learning method can effectively improve the optimization efficiency, greatly reduce the optimization time, and realize the rapid convergence of the neural network structure. It is a typical regression task using the neural network to learn the simulation process of the finite element analysis method. Therefore, based on the neural network and machine learning method, this paper proposes a high-speed and high-accuracy simulation analyzer, and its top-level architecture is shown in Figure 3.

#### 3.1. Dataset Definition

Machine learning is an algorithm that uses a large amount of historical data to dig out the hidden laws, which could be used for prediction or classification [27]. The generalization error of the learning algorithm could be evaluated through an experimental test. The overall dataset can be divided into two mutually exclusive sets. One set is used as the training set  $S$  to guide the update iteration of the learning algorithm; the other set is used as the test set  $T$  to test the learning algorithm's ability to analyze new samples. In this case, the basic idea is to first train the model within the training set  $S$ , and then verify it on the test set  $T$ , and evaluate the test error as an estimate of the generalization error.

In order to ensure the validity of the dataset, it is necessary to maintain the consistency of the data distribution when dividing the training set and the test set. This paper adopts a dual random classification method: firstly, in the process of generating the parameters of the dataset, the method of randomly picking values in the parameter interval is used to generate the dataset; then, when dividing the training set and the test set, the random

sampling method is used to divide. These methods can effectively reduce the subjective error introduced when the dataset is classified, and ensure the uniformity and consistency of the dataset distribution. In this case, the additional bias introduced by the input division process can be avoided.



**Figure 3.** The top-level architecture for the high-speed and high-accuracy disk MEMS resonator simulation analyzer enabled by the multilayer perceptron neural network.

In this paper, we randomly generated 4000 sets of data, and randomly selected 3200 sets of data as the training set for the machine learning on the neural network structure; the remaining 800 sets of data were used as the test set to detect neural network models' accuracy and reliability.

### 3.2. Multilayer Perceptron Neural Network

The multilayer perceptron neural network is also called the artificial neural network, which introduces one or more hidden layers on the basis of a single-layer neural network [27,29]. In addition to the input and output layers, there are multiple hidden layers in the middle. The simplest MLP model only contains one hidden layer, which is a three-layer structure. It is a composite architecture composed of multiple neurons in an orderly distribution. Its core mainly includes two parts: a feedforward neural network and an error back propagation algorithm. The following subsections will systematically describe the multilayer perceptron model constructed in this paper.

#### 3.2.1. M-P Neuron Model

A neural network is a composite structure composed of multiple neurons, and its basic unit is a neuron [27]. The used neuron model in this paper is the M-P neuron model proposed by McCulloch and Pitts [30], which could be described in Figure 4. In this model, the neuron receives input signals from  $n$  other neurons, and these input signals are transmitted through weighted connections. The total input value received by the neuron will be compared with the neuron's threshold  $\theta$ , and then processed through the activation function  $f(\cdot)$  to produce the neuron's output.

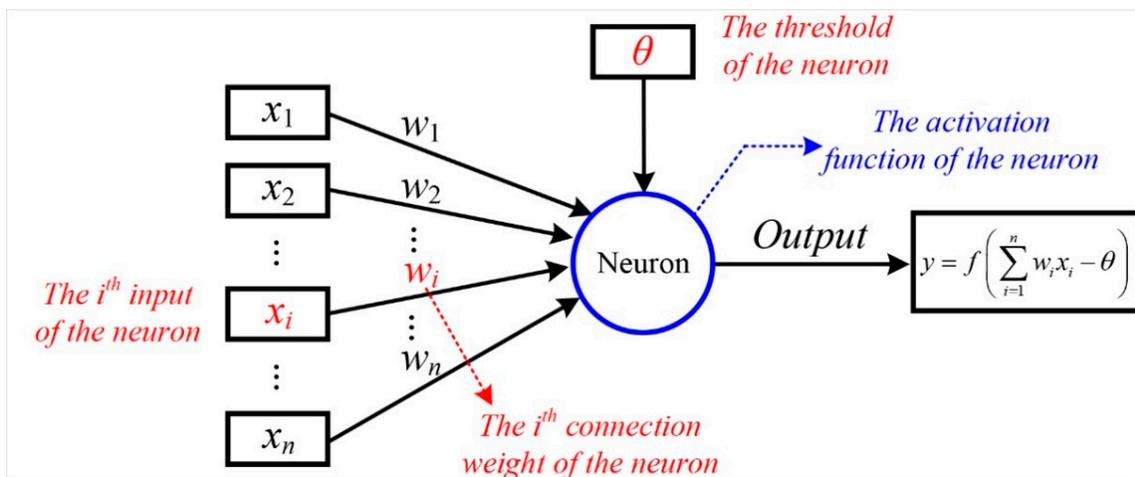


Figure 4. The M-P neuron model.

In our model, the output of a single neuron can be expressed as:

$$y = f\left(\sum_{i=1}^n w_i x_i - \theta\right) \quad (5)$$

where  $x_i$  and  $w_i$  ( $i = 1, 2, 3, \dots, n$ ) are the inputs of the neuron and the weight coefficients, respectively.  $\theta$  is the threshold of the neuron and  $f(\cdot)$  is its activation function, which is used to process the input signal of the neuron to generate its output signals.

However, since the structural analysis of the MEMS resonator is a complex nonlinear problem, this simple linear model is obviously limited. Therefore, in order to make the model better reflect the actual condition, the Sigmoid function is taken as the activation function in this paper, which can make the neural network arbitrarily approximate any nonlinear function [28]. The mathematical expression of Sigmoid activation function is as follows:

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

### 3.2.2. Multi-Layer Feedforward Neural Networks

Generally, it is believed that the typical structural parameters of disk MEMS resonators are independent of each other and do not affect each other; at the same time, all structural parameters have an impact on the final performance indicators of the resonator. As a result, the multilayer perceptron architecture is applied in this paper. The multilayer perceptron model is a typical multilayer feedforward neural network. In this neural network structure, neurons in each layer are fully interconnected with neurons in the next layer, indicating that the neurons in the bottom layer all contribute to the state of the neurons in the next layer; there is no connection between neurons in the same layer, indicating that neurons in the same layer are independent of each other and do not affect each other; at the same time, there is no cross-layer connection.

The first step of the multilayer perceptron architecture is to transfer the signal from the input layer to the output layer, which is the process of forward propagation. Forward propagation refers to the process of gradually transferring information from the first layer to the high-level layer. The forward propagation schematic diagram of the three-layer multilayer perceptron model is shown in Figure 5.

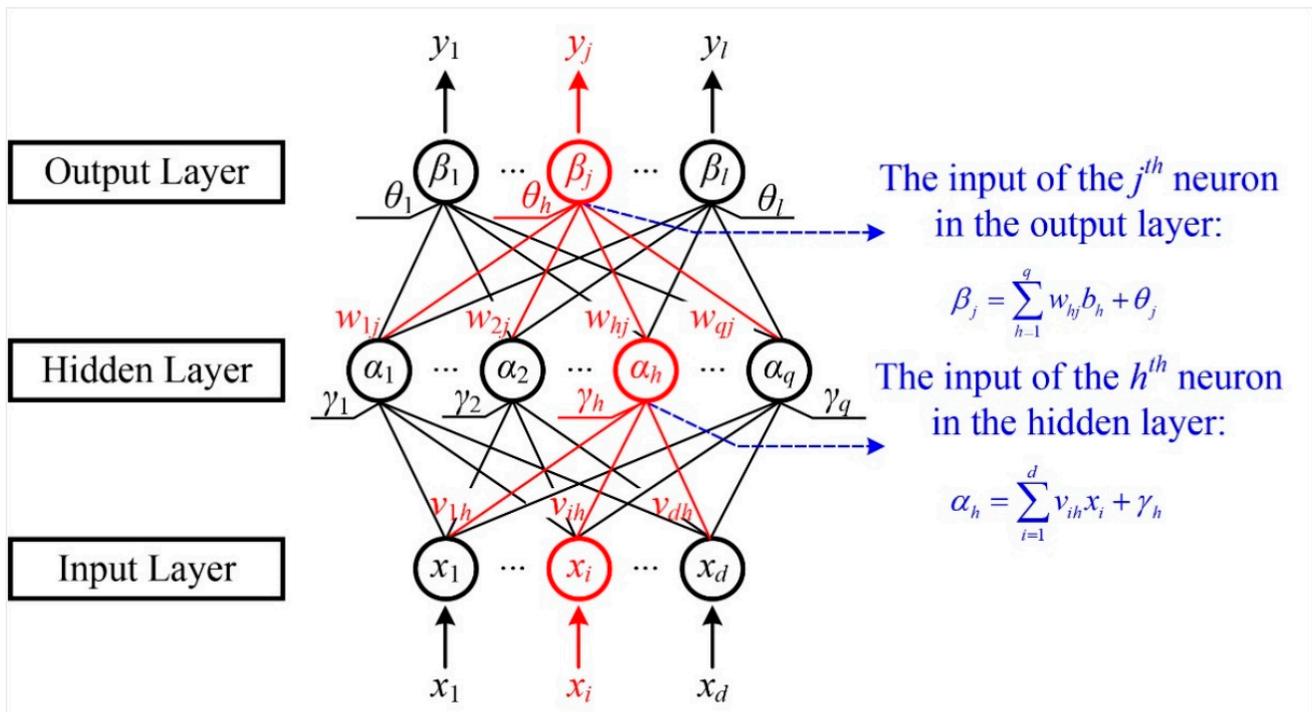


Figure 5. The forward propagation schematic diagram of the three-layer multilayer perceptron model.

Figure 5 exhibits the three-layer feedforward propagation network structure used in this paper, which contains  $d$  input neurons,  $q$  hidden neurons, and  $l$  output neurons. The initial input information of the model is  $[x_1, x_2, \dots, x_d]$ , and the final output information is  $[y_1, y_2, \dots, y_l]$ . In this model, the threshold of the  $h$ th neuron in the hidden layer is represented by  $\gamma_h$ , and the threshold of the  $j$ th neuron in the output layer is represented by  $\theta_j$ . In addition, the connection weight between the  $i$ th neuron in the input layer and the  $h$ th neuron in the hidden layer is  $v_{ih}$ , and the connection weight between the  $h$ th neuron in the hidden layer and the  $j$ th neuron in the output layer is  $w_{hj}$ . The activation function of the hidden layer is  $f(\cdot)$  and the activation function of the output layer is  $g(\cdot)$ . Therefore, the input signal received by the  $h$ th neuron in the hidden layer is shown in Equation (7):

$$\alpha_h = \sum_{i=1}^d v_{ih} x_i \tag{7}$$

The input signal received by the  $j$ th neuron in the output layer is as follows:

$$\beta_j = \sum_{h=1}^q w_{hj} b_h \tag{8}$$

Among them,  $b_h$  is the output of the  $h$ th neuron in the hidden layer, which can be expressed as:

$$b_h = f(\alpha_h - \gamma_h) = f\left(\sum_{i=1}^d v_{ih} x_i - \gamma_h\right) \tag{9}$$

As a result, the final output signal of the  $j$ th neuron in the output layer of the perceptron can be expressed as:

$$y_j = g(\beta_j - \theta_j) = g\left(\sum_{h=1}^q w_{hj} b_h - \theta_j\right) = g\left[\sum_{h=1}^q w_{hj} \cdot f\left(\sum_{i=1}^d v_{ih} x_i - \gamma_h\right) - \theta_j\right] \tag{10}$$

Finally, the matrix expression of the forward propagation model of the entire three-layer perceptron can be obtained:

$$\begin{aligned}
 \mathbf{y} &= \begin{bmatrix} y_1 \\ \vdots \\ y_j \\ \vdots \\ y_l \end{bmatrix} = g \left( \begin{bmatrix} w_{11} & \cdots & w_{1h} & \cdots & w_{1q} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{j1} & \cdots & w_{jh} & \cdots & w_{jq} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{l1} & \cdots & w_{lh} & \cdots & w_{lq} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_h \\ \vdots \\ b_q \end{bmatrix} - \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_j \\ \vdots \\ \theta_l \end{bmatrix} \right) = g \left( \mathbf{w} \begin{bmatrix} b_1 \\ \vdots \\ b_j \\ \vdots \\ b_q \end{bmatrix} - \boldsymbol{\theta} \right) \\
 &= g \left( \mathbf{w} \cdot f \left( \begin{bmatrix} v_{11} & \cdots & v_{1i} & \cdots & v_{1d} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ v_{h1} & \cdots & v_{hi} & \cdots & v_{hd} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ v_{q1} & \cdots & v_{qi} & \cdots & v_{qd} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_d \end{bmatrix} - \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_h \\ \vdots \\ \gamma_q \end{bmatrix} \right) - \boldsymbol{\theta} \right) = g(\mathbf{w} \cdot f(\mathbf{v}\mathbf{x} - \boldsymbol{\gamma}) - \boldsymbol{\theta})
 \end{aligned} \tag{11}$$

### 3.2.3. Error Back Propagation Algorithm

Another core task is updating the multilayer perceptron model’s structural parameters, where the error back propagation algorithm is applied in this paper. The back propagation of the loss function from the top layer to the bottom layer was used to realize the updating optimization of these parameters [27,31]. In this model, the mean squared error characterizes the degree of difference between the predicted value of the neural network and the theoretical value, which is used to express the learning accuracy of the multilayer perceptron neural network. According to Equation (5), for the training example  $(x_k, y_k)$ , the output of the neural network can be expressed as:

$$\hat{\mathbf{y}}_k = (\hat{y}_1^k, \hat{y}_2^k, \dots, \hat{y}_l^k) \tag{12}$$

where:

$$\hat{y}_j^k = f(\beta_j - \theta_j), j = 1, 2, \dots, l \tag{13}$$

Therefore, the mean squared error of this neural network on the training set  $(x_k, y_k)$  is as follows:

$$E_k = \frac{1}{2} \sum_{j=1}^l (\hat{y}_j^k - y_j^k)^2 \tag{14}$$

In the three layers perceptron neural network designed in this paper, as shown in Figure 5, a total of  $(d + l + 1) \times q + l$  network parameters need to be optimized:  $d \times q$  weight values from the input layer to the hidden layer,  $q \times l$  weight values from the hidden layer to the output layer,  $q$  thresholds of hidden layer neurons, and  $l$  thresholds of output layer neurons.

The error back propagation algorithm used in this paper is based on a gradient descent strategy, which optimizes the parameters in the direction of the negative gradient of the target. The learning rate  $\eta \in (0, 1)$  controls the update step size in each iteration of the algorithm. For the connection weight  $w_{hj}$  from the hidden layer to the output layer, when the learning rate is  $\eta$ , its updated value can be expressed as:

$$\Delta w_{hj} = -\eta \frac{\partial E_k}{\partial w_{hj}} \tag{15}$$

According to the forward propagation network model, it is obvious that  $w_{hj}$  first affects the input value  $\beta_j$  of the  $j$ th output layer neuron, then affects its output value  $\hat{y}_j^k$ , and finally affects  $E_k$ , so:

$$\frac{\partial E_k}{\partial w_{hj}} = \frac{\partial E_k}{\partial \hat{y}_j^k} \cdot \frac{\partial \hat{y}_j^k}{\partial \beta_j} \cdot \frac{\partial \beta_j}{\partial w_{hj}} \tag{16}$$

Combined with Equation (8), it can be seen that:

$$\frac{\partial \beta_j}{\partial w_{hj}} = b_h \quad (17)$$

In this paper, the activation functions of both the hidden layer and the output layer adopt the Sigmoid function, which has a good property to satisfy the following derivative:

$$f'(x) = f(x)(1 - f(x)) \quad (18)$$

On this basis, combining Equations (13) and (14), we obtain:

$$g_j = -\frac{\partial E_k}{\partial \hat{y}_j^k} \cdot \frac{\partial \hat{y}_j^k}{\partial \beta_j} = -(\hat{y}_j^k - y_j^k) f'(\beta_j - \theta_j) = \hat{y}_j^k (1 - \hat{y}_j^k) (y_j^k - \hat{y}_j^k) \quad (19)$$

Substituting Equations (17) and (19) into Equation (16), and then substituting the result into Equation (15), the updated equation for  $w_{hj}$  in the error back propagation algorithm can be expressed as:

$$\Delta w_{hj} = \eta g_j b_h \quad (20)$$

In the same way, other updated equations can be obtained:

$$\Delta \theta_j = -\eta g_j \quad (21)$$

$$\Delta v_{ih} = \eta e_h x_i \quad (22)$$

$$\Delta \gamma_h = -\eta e_h \quad (23)$$

where:

$$e_h = -\frac{\partial E_k}{\partial b_h} \cdot \frac{\partial b_h}{\partial \alpha_h} = -\sum_{j=1}^l \frac{\partial E_k}{\partial \beta_j} \cdot \frac{\partial \beta_j}{\partial b_h} f'(\alpha_h - \gamma_h) = \sum_{j=1}^l w_{hj} g_j f'(\alpha_h - \gamma_h) = b_h (1 - b_h) \sum_{j=1}^l w_{hj} g_j \quad (24)$$

As for the multilayer perceptron neural network, the goal of the error back propagation algorithm is to minimize the cumulative error on the training set:

$$E = \frac{1}{m} \sum_{k=1}^m E_k \quad (25)$$

Therefore, the cumulative error back propagation algorithm gets the errors of the entire training set before updating and optimizing the parameters. The working flow chart of the error back propagation algorithm used in this paper is displayed in Table 3. For each training example, the error back propagation algorithm performs the following operations: (1) first provide the input example to the input layer neuron, and then forward the signal layer by layer until the output layer result is produced; (2) calculate the output layer's error (corresponding to Process 4 to Process 5 in the Table 3), and then propagate the error back to the hidden layer neuron (corresponding to Process 6 in the Table 3); (3) finally, adjust the connection weights and thresholds based on the error of the hidden layer neuron (corresponding to Process 7 in the Table 3). This iterative process is repeated until the stopping condition is satisfied. The stop condition used in this paper is that the cumulative error is less than  $10^{-3}$  or the number of iterations exceeds  $10^5$ .

**Table 3.** The working flow chart of the error back propagation algorithm used in this paper.

<b>Input:</b>	The training dataset $D = \{(x_k, y_k)\}_{k=1}^m$ and the learning rate $\eta$ .
1	Randomly initialize all connection weights and thresholds in the network within the range of (0, 1).
2	<b>repeat</b>
3	<b>for all</b> $(x_k, y_k) \in D$ <b>do</b>
4	Calculate the output value $\hat{y}_k$ of the current sample according to the current parameters and Equation (13).
<b>Process:</b>	5 Calculate the gradient index $g_j$ of the neurons in the output layer according to Equation (19).
6	Calculate the gradient index $e_h$ of the hidden layer neuron according to Equation (24).
7	Update the connection weights $w_{hj}$ , $v_{ih}$ and thresholds $\theta_j$ , $\gamma_h$ in the neural network according to Equations (20)–(23).
8	<b>end for</b>
9	<b>until</b> satisfies the stop condition (the cumulative error is less than $10^{-3}$ or the number of iterations exceeds $10^5$ ).
<b>Output:</b>	Multi-layer perceptron neural network after optimizing connection weights and thresholds.

## 4. Discussion

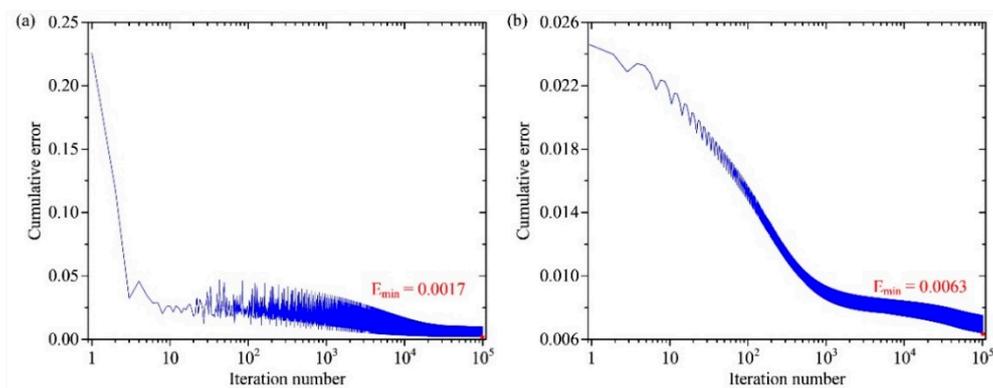
### 4.1. Neural Network Structure Parameters

In this paper, a traditional three-layer perceptron neural network as the deep learning tool to build a high-speed and high-accuracy mathematical modeling analysis model is explored. In this neural network structure, the number of the input layer's neurons  $d$  is equal to the number of the typical disk MEMS resonator's basic structural parameters, while the number of the output layer's neurons  $l$  is equal to the number of the resonator's core performance indicators. Considering the complexity and accuracy of the model, the number of neurons in the hidden layer of the neural network built in this paper is set as  $(d + l) \times 30$  [27,32]. The learning rate is 0.01 and the activation functions are both Sigmoid functions. The stop condition of the training process is that its cumulative error is less than  $10^{-3}$  or the number of iterations exceeds  $10^5$ .

### 4.2. Neural Network Learning Performance

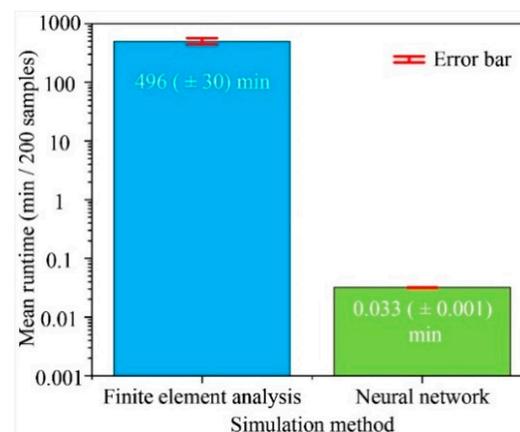
The multilayer perceptron neural network needs to be trained before performing the prediction process. The automatic structural parameters generator produces 4000 different typical disk MEMS resonator samples, which are divided as the training dataset and the testing dataset. With these sample data, the multilayer perceptron neural network steps into the training process. To obtain an effective mathematical modeling tool with high speed and high accuracy, the entire training process takes about 516 min by using an NVIDIA GeForce GTX 1660Ti graphics cards. With more advanced equipment, both the training time and the runtime can be potentially further reduced. The cumulative error curves of the multilayer perceptron neural network are displayed in Figure 6. It is obvious that with the increasing of the iterations, the cumulative errors of the training data and the testing data both decrease gradually, which are 0.0017 and 0.0063 in the end, respectively. It proves that the cumulative error of the multilayer perceptron neural network reaches a satisfying level.

Generally, the evaluation indicators of the neural network learning performance are the prediction speed and the prediction accuracy. In the evaluating process, the performance of our high-speed and high-accuracy mathematical modeling tool generated by the multilayer perceptron neural network is compared with the corresponding performance produced by the traditional finite element analysis, which is set as the reference.



**Figure 6.** The cumulative error curves of the multilayer perceptron neural network. (a) The cumulative error curve of the training dataset; (b) the cumulative error curve of the test dataset.

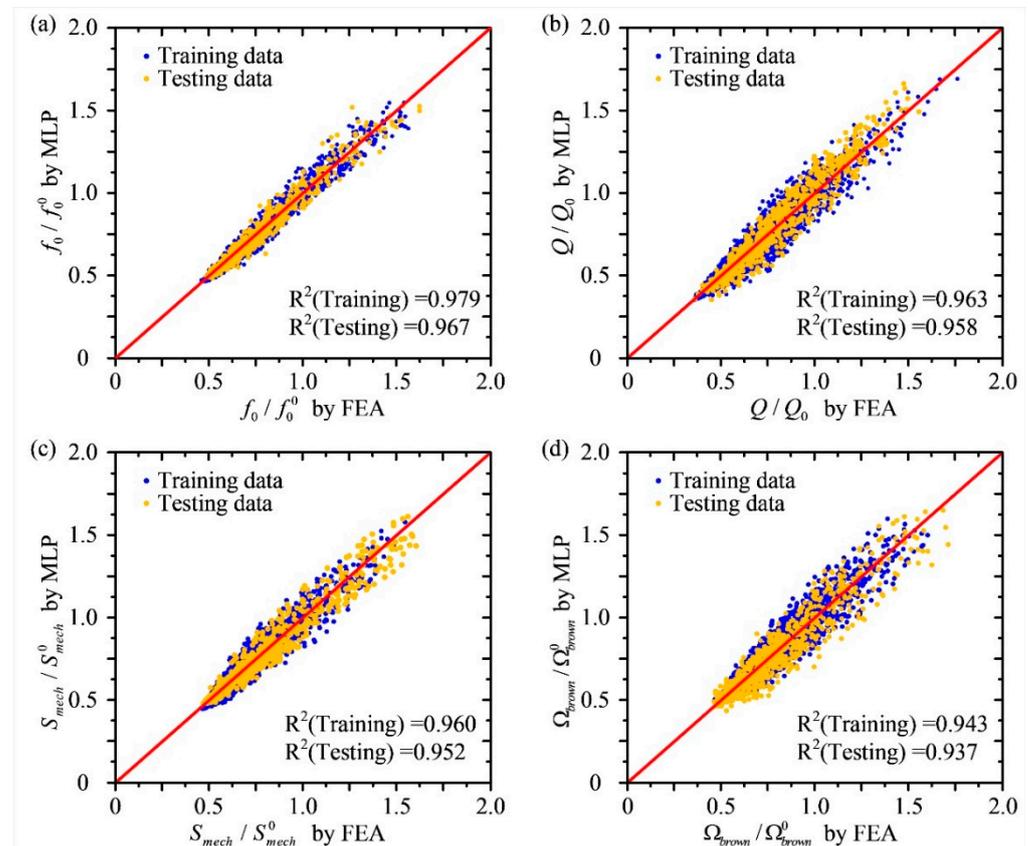
In order to fairly compare the runtime of different methods, these two methods dealt with the same structural models under the same computer configuration. Firstly, the automatic structural parameters generator produces 800 new typical disk MEMS resonator's structural parameters. Then, these two methods both simulate or predict these samples' performance indicators. The runtime for 200 samples comparison between the traditional finite element analysis method and the neural network analyzer based on the multilayer perceptron neural network is shown in Figure 7. The average runtime for the traditional finite element analysis method is 496 ( $\pm 30$ ) minutes per 200 samples, while the time consumption of the neural network for the same work is only 0.033 ( $\pm 0.001$ ) minutes. It is obvious that the comparison speed of the neural network is much higher than the traditional finite element analysis speed. The calculation speed of the neural network is nearly about 15,000 times that of the finite element analysis speed, exhibiting excellent calculation characteristics and effectively shortening the simulation time.



**Figure 7.** Runtime comparison between the traditional finite element analysis method and the neural network analyzer based on the multilayer perceptron neural network.

In the process of evaluating the accuracy of the predictions produced by the multilayer perceptron neural network, the corresponding simulation results produced by the traditional finite element analysis method are chosen as the reference. The comparison between performance indicators calculated by the finite element analysis method and predicted by the multilayer perceptron neural network is depicted in Figure 8. In order to better characterize the cumulative error of the neural network, all the performance index parameters of the disk MEMS resonator are normalized. The accuracy of the multilayer perceptron neural network is defined by the mean squared error  $E_k$ , as shown in Equation (17), which could be represented by the matching degree of the predictions from the neural network and the simulation results produced by the finite element analysis method. It

is obvious that the training dataset and test dataset for different performance indicators both scatter around the straight line  $y = x$ . Moreover, the mean squared errors for different performance indicators of this neural network are all less than 0.06. It has been proved that the multilayer perceptron neural network has a high prediction accuracy and there is no overfitting problem.



**Figure 8.** Predicted performances of automatically produced disk MEMS resonators given by the neural network analyzer based on the multilayer perceptron neural network compared with the traditional finite element analysis method. (a) The comparison with the fundamental frequency, where  $f_0^0$  is normalized fundamental frequency reference; (b) the comparison with the quality factor, where  $Q_0$  is normalized quality factor reference; (c) the comparison with the mechanical sensitivity, where  $S_{mech}^0$  is normalized mechanical sensitivity reference; (d) the comparison with the mechanical thermal noise, where  $\Omega_{brown}^0$  is normalized mechanical thermal noise reference.

Obviously, the neural network model after training has reached a very good accuracy and speed, which is of great significance for reducing repetitive work. Since the neural network is a learning model based on prior knowledge, it needs to provide certain data for its learning and training, so that the time cost of pre-training needs to be taken into consideration. Although it takes a certain amount of time for the neural network to grasp the internal characteristics of the resonator, this time is very short in a long time span. It can realize the fast simulation of the resonator's performance under fixed constraints, greatly reducing the time-consuming and long repetitive work. In addition, the multilayer perceptron network also has great application potential in other fields. For the multilayer perceptron network, it does not care about the internal physical connection, but focuses on the connection between inputs and outputs. Therefore, for the problem of clarifying the input and output of the system, the multilayer perceptron neural network can obtain an ideal system architecture through learning and training with sufficient sample data [27–32]. It has important significance and value for the optimization and simulation of structural parameters in specific scenarios, and greatly improves design efficiency.

## 5. Conclusions

The complete procedures of the multilayer perceptron neural network have been discussed in detail in this paper. To provide enough samples, an automatic structural parameter generator is applied to produce different data within reasonable regions. The error back propagation algorithm is used to update the connection weights and thresholds in the neural network. The time consumption for optimizing the neural network structure is about 516 min after iterating  $10^5$  dealing with 4000 samples. It is worth noting that the cumulative error of the multilayer perceptron neural network decreases to 0.0017 in the end, proving that the training performance of the neural network reaches a satisfying level.

This study has proven that when dealing with the same samples, the time consumption of the neural network is less than 1/15,000 of the time required for the finite element analysis, exhibiting excellent calculation characteristics. Moreover, the prediction accuracy of the neural network remains a remarkable level after an adequate training process. Compared with the traditional finite element analysis method, the mean squared errors for different performance indicators of this neural network are all less than 0.06. As a result, it can be concluded that the prediction tool based on the multilayer perceptron neural network achieves a high speed and accuracy level.

The high-speed and high-accuracy mathematical modeling technology is a product of the artificial neural network in the traditional MEMS resonator design process, and it has great potential in enhancing the designing performance. It has been proved that it is a good method to accelerate the prediction speed while considering the accuracy. Moreover, it is demonstrated that the artificial neural network model can provide a promising tool for the design and optimization of structures. This mathematical modeling method presented in this work is suitable for other MEMS devices, which can be applied in the optimization design after suitable training process.

**Author Contributions:** Conceptualization, Q.L. and K.L.; methodology, Q.L.; software, K.L.; validation, K.W., H.Z. and X.S.; formal analysis, Q.L.; investigation, X.W.; resources, K.L.; data curation, D.X.; writing—original draft preparation, Q.L. and K.L.; writing—review and editing, Q.L., K.L. and K.W.; visualization, K.L.; supervision, X.W. and D.X.; project administration, X.W.; funding acquisition, X.W. and D.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key R&D Program of China, grant number 2018YFB2002304; the Excellent Youth Foundation of Hu'nan Scientific Committee, grant number 2020JJ2033.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data are contained within the article.

**Acknowledgments:** The authors would like to thank the Laboratory of Microsystem, National University of Defense Technology, China, for equipment access and technical support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Middlemiss, R.P.; Samarelli, A.; Paul, D.J.; Hough, J.; Rowan, S.; Hammond, G.D. Measurement of the earth tides with a MEMS gravimeter. *Nature* **2016**, *531*, 614–618. [[CrossRef](#)] [[PubMed](#)]
2. Wei, X.; Zhang, T.; Jiang, Z.; Ren, J.; Huan, R. Frequency latching in nonlinear micromechanical resonators. *Appl. Phys. Lett.* **2017**, *110*, 143506. [[CrossRef](#)]
3. Lu, K.; Li, Q.; Zhou, X.; Song, G.; Wu, K.; Zhuo, M.; Wu, X.; Xiao, D. Modal coupling effect in a novel nonlinear micromechanical resonator. *Micromachines* **2020**, *11*, 472. [[CrossRef](#)] [[PubMed](#)]
4. Hanay, M.S.; Kelber, S.; Naik, A.K.; Chi, D.; Hentz, S.; Bullard, E.C.; Colinet, E.; Duraffourg, L.; Roukes, M.L. Single-protein nanomechanical mass spectrometry in real time. *Nat. Nanotechnol.* **2012**, *7*, 602–608. [[CrossRef](#)] [[PubMed](#)]
5. Wang, Z.; Jia, H.; Zheng, X.; Yang, R.; Ye, G.J.; Chen, X.H.; Feng, P.X. Resolving and tuning mechanical anisotropy in black phosphorus crystal via nanomechanical multimode resonance spectromicroscopy. *Nano Lett.* **2016**, *16*, 5394–5400. [[CrossRef](#)]
6. Zhang, T.; Zhou, B.; Song, M.; Lin, Z.; Jin, S.; Zhang, R. Structural parameter identification of the center support quadruple mass gyro. *IEEE Sens. J.* **2017**, *17*, 3765–3775. [[CrossRef](#)]

7. Zhou, X.; Zhao, C.; Xiao, D.; Sun, J.; Sobreviela, G.; Gerrard, D.D.; Chen, Y.; Flader, I.; Kenny, T.W.; Wu, X.; et al. Dynamic modulation of modal coupling in microelectromechanical gyroscopic ring resonators. *Nat. Commun.* **2019**, *10*, 4980. [[CrossRef](#)]
8. Zhang, H.; Huang, J.; Yuan, W.; Chang, H. A high-sensitivity micromechanical electrometer based on mode localization of two degree-of-freedom weakly coupled resonators. *J. Microelectromech. Syst.* **2016**, *25*, 937–946. [[CrossRef](#)]
9. Nagourney, T.; Cho, J.Y.; Shiari, B.; Darvishian, A.; Najafi, K. 259 Second ring-down time and 4.45 million quality factor in 5.5 kHz fused silica birdbath shell resonator. In Proceedings of the 19th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS), Kaohsiung, Taiwan, 18–22 June 2017; pp. 790–793.
10. Lu, K.; Zhou, X.; Li, Q.; Xiao, D. Coherent phonon manipulation in a disk resonator gyroscope with inertial resonance. In Proceedings of the 7th IEEE International Symposium on Inertial Sensors and Systems (INERTIAL), Hiroshima, Japan, 23–26 March 2020; p. 9090063.
11. Li, Q.; Xiao, D.; Zhou, X.; Xu, Y.; Zhuo, M.; Hou, Z.; He, K.; Zhang, Y.; Wu, X. 0.04 degree-per-hour MEMS disk resonator gyroscope with high-quality factor(510k) and long decaying time constant (74.9s). *Microsyst. Nanoeng.* **2018**, *4*, 32. [[CrossRef](#)]
12. Li, Q.; Xiao, D.; Zhou, X.; Hou, Z.; Zhuo, M.; Xu, Y.; Wu, X. Dynamic modeling of the multiring disk resonator gyroscope. *Micromachines* **2019**, *10*, 181. [[CrossRef](#)]
13. Dennis, J.O.; Rabih, A.A.S.; Khir, M.H.M.; Ahmed, M.G.A.; Ahmed, A.Y. Modeling and finite element analysis simulation of MEMS based acetone vapor sensor for noninvasive screening of diabetes. *J. Sens.* **2016**, *2016*, 9563938. [[CrossRef](#)]
14. Yoon, G.H.; Jensen, J.S.; Sigmund, O. Topology optimization of acoustic-structure interaction problems using a mixing finite element formulation. *Int. J. Numer. Methods Eng.* **2007**, *70*, 1049–1075. [[CrossRef](#)]
15. Mao, Y.; He, Q.; Zhao, X. Designing complex architecture materials with generative adversarial networks. *Sci. Adv.* **2020**, *6*, eaaz4169. [[CrossRef](#)]
16. Hanakata, P.Z.; Cubuk, E.D.; Campbell, D.K.; Park, H.S. Accelerated search and design of stretchable graphene kirigami using machine learning. *Phys. Rev. Lett.* **2018**, *121*, 255304. [[CrossRef](#)]
17. Wilt, J.K.; Yang, C.; Gu, G.X. Accelerating auxetic metamaterial design with deep learning. *Adv. Energy Mater.* **2020**, *22*, 1901266.
18. Challoner, A.D.; Ge, H.H.; Liu, J.Y. Boeing disc resonator gyroscope. In Proceedings of the 2014 IEEE/ION Position, Location and Navigation Symposium-PLANS 2014, Monterey, CA, USA, 5–8 May 2014; pp. 504–514.
19. Nitzan, S.; Ahn, C.H.; Su, T.H.; Li, M.; Ng, E.J.; Wang, S.; Yang, Z.M.; O'Brien, G.; Boser, B.E.; Kenny, T.W.; et al. Epitaxially-encapsulated polysilicon disk resonator gyroscope. In Proceedings of the 2013 IEEE 26th International Conference on Micro Electro Mechanical Systems (MEMS), Taipei, Taiwan, 20–24 January 2013; pp. 625–628.
20. Li, Q.; Xiao, D.; Zhou, X.; Hou, Z.; Xu, Y.; Wu, X. Quality factor improvement in the disk resonator gyroscope by optimizing the spoke length distribution. *J. Microelectromech. Syst.* **2018**, *27*, 414–423. [[CrossRef](#)]
21. Ahn, C.H.; Shin, D.D.; Hong, V.A.; Yang, Y.; Ng, E.J.; Chen, Y.; Flader, I.B.; Kenny, T.W. Encapsulated disk resonator gyroscope with differential internal electrodes. In Proceedings of the 2016 IEEE 29th International Conference on Micro Electro Mechanical Systems (MEMS 2016), Shanghai, China, 24–28 January 2016; pp. 962–965.
22. Zhou, X.; Xiao, D.; Wu, X.; Wu, Y.; Hou, Z.; He, K.; Li, Q. Stiffness-mass decoupled silicon disk resonator for high resolution gyroscopic application with long decay time constant (8.695 s). *Appl. Phys. Lett.* **2016**, *109*, 263501. [[CrossRef](#)]
23. Zhou, X.; Xiao, D.; Hou, Z.; Li, Q.; Wu, Y.; Yu, D.; Li, W.; Wu, X. Thermoelastic quality-factor enhanced disk resonator gyroscope. Proceedings of 2017 IEEE 30th International Conference on Micro Electro Mechanical Systems (MEMS), Las Vegas, NV, USA, 22–26 January 2017; pp. 1009–1012.
24. Zhou, X.; Xiao, D.; Li, Q.; Hou, Z.; He, K.; Chen, Z.; Wu, Y.; Wu, X. Decaying time constant enhanced MEMS disk resonator for high precision gyroscopic application. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 452–458. [[CrossRef](#)]
25. Zhou, X.; Wu, Y.; Xiao, D.; Hou, Z.; Li, Q.; Yu, D.; Wu, X. An investigation on the ring thickness distribution of disk resonator gyroscope with high mechanical sensitivity. *Int. J. Mech. Sci.* **2016**, *117*, 174–181. [[CrossRef](#)]
26. Wang, Z.; Lee, J.; Feng, P.X.L. Spatial mapping of multimode Brownian motions in high-frequency silicon carbide microdisk resonators. *Nat. Commun.* **2014**, *5*, 5158. [[CrossRef](#)] [[PubMed](#)]
27. Zhou, Z. *Machine Learning*, 1st ed.; Tsinghua University Press: Beijing, China, 2016; pp. 97–106.
28. Guo, R.; Xu, R.; Wang, Z.; Sui, F.; Lin, L. Accelerating MEMS design process through machine learning from pixelated binary images. In Proceedings of the 2021 IEEE 34th International Conference on Micro Electro Mechanical Systems (MEMS), Gainesville, FL, USA, 25–29 January 2021; pp. 153–156.
29. Liu, F.; Wang, X.; Wang, X.; Wang, L. Machine learning-based design and optimization of curved beams for multistable structures and metamaterials. *Extreme Mech. Lett.* **2020**, *41*, 101002. [[CrossRef](#)]
30. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [[CrossRef](#)]
31. Chen, C.T.; Gu, G.X. Generative deep neural networks for inverse materials design using backpropagation and active learning. *Adv. Sci.* **2020**, *7*, 1902607. [[CrossRef](#)]
32. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [[CrossRef](#)]