Novel Decomposition Technique on Rational-Based Neuro-Transfer Function for Modeling of Microwave Components

Zhihao Zhao [1,2], **Feng Feng** [2,*], **Jianan Zhang** [2], **Wei Zhang** [2], **Jing Jin** [1,2], **Jianguo Ma** [3] **and Qi-Jun Zhang** [1,2]

[1] School of Microelectronics, Tianjin University, Tianjin 300072, China; zhihaozhao@doe.carleton.ca (Z.Z.); jingjin5@cmail.carleton.ca (J.J.); qjz@doe.carleton.ca (Q.-J.Z.)

[2] Department of Electronics, Carleton University, Ottawa, ON K1S5B6, Canada; jiananzhang@doe.carleton.ca (J.Z.); weizhang13@doe.carleton.ca (W.Z.)

[3] School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China; mjg@gdut.edu.cn

[*] Correspondence: fengfeng@doe.carleton.ca

check for updates

Abstract: The rational-based neuro-transfer function (neuro-TF) method is a popular method for parametric modeling of electromagnetic (EM) behavior of microwave components. However, when the order in the neuro-TF becomes high, the sensitivities of the model response with respect to the coefficients of the transfer function become high. Due to this high-sensitivity issue, small training errors in the coefficients of the transfer function will result in large errors in the model output, leading to the difficulty in training of the neuro-TF model. This paper proposes a new decomposition technique to address this high-sensitivity issue. In the proposed technique, we decompose the original neuro-TF model with high order of transfer function into multiple sub-neuro-TF models with much lower order of transfer function. We then reformulate the overall model as the combination of the sub-neuro-TF models. New formulations are derived to determine the number of sub-models and the order of transfer function for each sub-model. Using the proposed decomposition technique, we can decrease the sensitivities of the overall model response with respect to the coefficients of the transfer function in each sub-model. Therefore, the modeling approach using the proposed decomposition technique can increase the modeling accuracy. Two EM parametric modeling examples are used to demonstrate the proposed decomposition technique.

Keywords: decomposition; microwave components; neural networks; parameter extraction; parametric modeling; rational-based transfer function

1. Introduction

Parametric modeling of electromagnetic (EM) behaviors for microwave components has become important for EM-based design in the microwave area. EM-based designs can be time-consuming, since repetitive simulations of EM are usually required due to value adjustments of geometrical parameters. Parametric models are developed to characterize the relationship between geometrical variables and EM responses. The developed parametric models allow faster simulation and optimization with varying values of geometrical parameters and can be subsequently implemented in high-level design of circuit and system.

An artificial neural network (ANN) has been an important vehicle for parametric modeling of EM behavior in radio frequency and microwave area [1–6]. ANN benefits from its strong learning and generalization capabilities and it has been used for a wide variety of microwave applications [7–16].

The universal approximation theorem [17] of ANNs provides the theoretical foundation that if sufficient data are used in training, the good accuracy of ANN models can be achieved within the training region. To develop an accurate parametric model, we train ANNs to learn the nonlinear relationship between geometrical variables and EM behavior. Apart from modeling microwave passive components, ANN has also been used for modeling microwave active components, such as power amplifiers [18], field-effect transistor [19], high electron-mobility transistors [20], etc.

Another popular parametric modeling method is the neuro-transfer function (neuro-TF) method [21–25], which combines neural networks and transfer functions. In the neuro-TF method, EM responses of passive components versus frequency are represented by the transfer functions. The transfer functions in the neuro-TF model are used as prior knowledge, which allows less hidden neurons and less training data to be used in the model development than the ANN method. The prior knowledge can help speed up model development and enhance the capability for learning and generalization of the overall model. One of the popular neuro-TF methods is the pole-residue-based neuro-TF method [22,23]. This method can deal with high-order problem. However, when the geometrical variations become large, some poles and residues obtained from vector fitting may vary discontinuously with respect to geometrical variables. This discontinuity issue leads to the difficulty of training for the overall model. Hybrid-based neuro-TF method [24] is used to further reduce the discontinuity issue of poles and residues by converting the most discontinuous poles and residues into the rational format. Rational-based neuro-TF method [25] is an alternative neuro-TF method, which does not have the discontinuity issue of poles and residues. The rational-based neuro-TF works well when the order in the neuro-TF model is low. However, as the order increases, the transfer function response is very sensitive to the coefficients of the transfer function, leading to the high-sensitivity issue. The high-sensitivity issue can result in the difficulty of training for the overall model when the order in the neuro-TF model is high. How to solve the high-sensitivity issue in the rational-based neuro-TF modeling method is a challenging topic.

This paper proposes a new decomposition technique for developing a rational-based neuro-TF model for EM microwave components, addressing the challenge of high-sensitivity issue. The proposed technique decomposes the single neuro-TF model with high order into multiple sub-neuro-TF models with much lower order and reformulates the overall neuro-TF model as the combination of the sub-neuro-TF models. New formulations have been derived for determination of the number of sub-neuro-TF models and the order of each sub-neuro-TF model. The proposed decomposition technique can decrease the sensitivities of the overall model response with respect to the coefficients of the transfer function, improving the overall model accuracy over the existing rational-based neuro-TF method. Compared with the existing pole-residue-based neuro-TF method, the proposed neuro-TF method has fewer sub-neuro-TF models and less discontinuity issue in each sub-neuro-TF model, achieving better overall model accuracy.

We organize this paper as follows. Section 2 describes the high-sensitivity issue in the existing rational-based neuro-TF method. Section 3 provides a detailed description of the proposed decomposition technique and the proposed decomposition technique for developing the proposed rational-based neuro-TF model. Section 4 demonstrates the proposed decomposition technique by two EM examples. Section 5 concludes the paper.

2. The High-Sensitivity Issue in the Existing Rational-Based Neuro-TF Method

The existing rational-based neuro-TF modeling method [25] requires a parameter extraction process for extraction of the coefficients of the transfer function numerator and denominator in the neuro-TF model for each geometrical parameter sample. The existing method can be well-functioning if the geometrical variations are small and/or the order of the transfer function in the neuro-TF model is low. However, the existing neuro-TF method may not be suitable to address the situations where geometrical variations become large and/or where the order of the transfer function becomes high.

When the geometrical variations are small, the relationship between the geometrical variables and the coefficients of the transfer function is of low nonlinearity. Since neural networks can easily learn this relationship, the training accuracy of the neural networks is high. This results in the high training accuracy of the overall neuro-TF model. When the geometrical variations become larger, the relationship between the geometrical variables and the coefficients of the transfer function becomes more nonlinear. The accuracy of the neural networks becomes lower. Under this situation, the order of the neuro-TF model becomes an important factor in the accuracy of the overall model. When the transfer function order becomes high, the sensitivities of the neuro-TF model response with respect to the coefficients of the transfer function become high. With the high sensitivity, a relatively small error of the neural networks will result in a large error of the overall neuro-TF model. How to solve this high-sensitivity issue to increase the accuracy of the rational-based neuro-TF is a challenging topic.

To address the high-sensitivity issue, this paper proposes a novel decomposition technique for the rational-based neuro-TF modeling method. The main idea is that we decompose the neuro-TF model into sub-neuro-TF models and reformulate the overall neuro-TF model as the combination of the sub-neuro-TF models. In this way, the order of the overall neuro-TF model is decomposed into lower orders of the sub-neuro-TF models, decreasing the sensitivities of the overall model response with respect to the coefficients of the transfer function. By decreasing sensitivities of the overall model response with respect to coefficients of the transfer function using the proposed decomposition technique, the accuracy of the overall neuro-TF model can be improved.

3. Proposed Decomposition Technique for Development of Rational-Based Neuro-TF Model

3.1. Concept of the Decomposition Technique for Rational-Based Neuro-TF Model

Let x be a vector which contains the geometrical variables. Let y represent the frequency response, e.g., *S*-parameter. A general rational-based neuro-TF model [25] in frequency domain can be expressed as

$$y(x, w_a, w_b, s) = \frac{\sum_{j=1}^{N} a_j(x, w_a) s^{j-1}}{1 + \sum_{j=1}^{N} b_j(x, w_b) s^j} \tag{1}$$

where N represents the order of the transfer function; a_j and b_j represent the jth coefficients of the transfer function numerator and denominator, respectively; w_a and w_b represent the neural network weights; and s represents the Laplace domain frequency.

As can be seen from (1), the direct decomposition of the rational-based neuro-TF is not easy. In this case, we propose to reformat the rational-based transfer function and decompose the neuro-TF model indirectly. To reformat the rational-based transfer function, we use the pole-residue-based transfer function [23], which is another format of the transfer function. Here we define the effective poles as the poles whose imaginary parts are positive [23]. We define the effective residues as the residues corresponding to the effective poles [23]. We define the effective order as the order of the pole-residue transfer function which consists only of the effective poles and residues. In order to derive the decomposition technique for the rational-based neuro-TF model, we need to first convert the rational-based transfer function into the pole-residue format, expressed as

$$y = \frac{\sum_{j=1}^{N} a_j s^{j-1}}{1 + \sum_{j=1}^{N} b_j s^j} = \sum_{j=1}^{N_{eff}} \frac{r_j}{s - p_j} + \sum_{j=1}^{N_{eff}} \frac{r_j^*}{s - p_j^*} \tag{2}$$

where N_{eff} is the effective order of the pole-residue-based transfer function; p_j and r_j are the jth effective pole and residue of the pole-residue-based transfer function, respectively; and p_j^* and r_j^* are the complex conjugate parts of jth effective pole and residue, respectively. To ensure that all the poles and residues in (2) are in complex values, order N should be an even number and thus the effective order $N_{eff} = N/2$. It is noticed that the pole-residue-based transfer function itself is expressed as a summation of sub-transfer functions. We can easily decompose the pole-residue-based transfer function into multiple sub-transfer functions in pole-residue format. Let M represent the number of the sub-transfer functions. The decomposition of the pole-residue-based transfer function can be expressed as

$$y = \sum_{i=1}^{M} y_i = \sum_{i=1}^{M} \left(\sum_{j=1}^{N_i} \frac{r_{n(i,j)}}{s - p_{n(i,j)}} + \sum_{j=1}^{N_i} \frac{r_{n(i,j)}^*}{s - p_{n(i,j)}^*} \right) \tag{3}$$

where

$$n(i,j) = \begin{cases} j, & \text{if } i = 1 \\ j + \sum_{t=1}^{i-1} N_t, & \text{if } i > 1 \end{cases}, \tag{4}$$

y_i represents the frequency response of the ith sub-transfer function; and N_i represents the effective order of the ith sub-transfer function. The relationship between N_{eff}, N_i, and M can be expressed as

$$N_{eff} = \sum_{i=1}^{M} N_i. \tag{5}$$

After the decomposition, we convert these sub-transfer functions back into the rational format, expressed as

$$y_i = \sum_{j=1}^{N_i} \frac{r_{n(i,j)}}{s - p_{n(i,j)}} + \sum_{j=1}^{N_i} \frac{r_{n(i,j)}^*}{s - p_{n(i,j)}^*} = \frac{\sum_{j=1}^{2N_i} \hat{a}_{ij} s^{j-1}}{1 + \sum_{j=1}^{2N_i} \hat{b}_{ij} s^{j}} \tag{6}$$

where $\hat{a}_{ij}$ and $\hat{b}_{ij}$ represent the jth coefficients of the numerator and denominator for the ith sub-transfer function, respectively. Now the rational-based transfer function can be formulated as the summation of M sub-rational-based transfer functions. Substituting (6) into (3), we derive the rational-based neuro-TF model incorporating decomposition technique, formulated as

$$y(\boldsymbol{x}, \boldsymbol{w}_a, \boldsymbol{w}_b, s) = \sum_{i=1}^{M} y_i(\boldsymbol{x}, \boldsymbol{w}_a, \boldsymbol{w}_b, s) = \sum_{i=1}^{M} \frac{\sum_{j=1}^{2N_i} \hat{a}_{ij}(\boldsymbol{x}, \boldsymbol{w}_a) s^{j-1}}{1 + \sum_{j=1}^{2N_i} \hat{b}_{ij}(\boldsymbol{x}, \boldsymbol{w}_b) s^{j}}. \tag{7}$$

Table 1 illustrates the sensitivities of the proposed and the existing neuro-TF model responses with respect to coefficients of the transfer function. In Table 1, for the existing neuro-TF model formulated in (2), when the frequency range is small (e.g., varying around 1 to ignore the influence of s^j or s^{j-1}), all the effective poles (i.e., $p_j, \forall j \in N$) are close to each other. When the value of frequency s is in the middle of all the effective poles, the value of $\prod_{j=1}^{N_{eff}} (s - p_j)$ becomes very small as N_{eff} is high, which makes the y too sensitive with respect to the coefficients a_j and b_j. This high sensitivity-issue cannot be easily solved by simply reducing the order N in vector fitting, since the minimum value of N has been used in vector fitting. The high-sensitivity issue can be effectively overcome by the neuro-TF model incorporating the proposed decomposition technique, formulated in (7). For the proposed neuro-TF model, the value of N_i is much smaller than that of N_{eff}. In that case, the value of $\prod_{j=1}^{N_i} (s - p_{n(i,j)})$ is

much larger than that of $\prod_{j=1}^{N_{eff}} (s - p_j)$. Therefore, the sensitivities of y with respect to the coefficients $\hat{a}_{ij}$ and $\hat{b}_{ij}$ of the proposed neuro-TF model are much lower than those with respect to the coefficients a_j and b_j of the existing neuro-TF model. With lower sensitivities of the model response with respect to its coefficients of the transfer function, the proposed neuro-TF model can achieve higher accuracy than the existing neuro-TF model.

Table 1. Comparison of the sensitivity of rational-based neuro-transfer function (neuro-TF) model response with respect to the coefficients with/without decomposition.

	Coeff.	Transfer Function	Sensitivity of Transfer Function Response w.r.t the Coeff.
Existing Rational-Based Neuro-TF model (without Decomposition)	a_j	$y = \dfrac{\sum\limits_{j=1}^{N} a_j s^{j-1}}{1 + \sum\limits_{j=1}^{N} b_j s^j}$	$\dfrac{\partial y}{\partial a_j} = \dfrac{s^{j-1}}{1 + \sum\limits_{j=1}^{N} b_j s^j} = \dfrac{s^{j-1}}{\prod\limits_{j=1}^{N_{eff}} \left(s - p_j\right) \cdot \prod\limits_{j=1}^{N_{eff}} \left(s - p_j^*\right)}$
	b_j		$\dfrac{\partial y}{\partial b_j} = \dfrac{-s^j \sum\limits_{j=1}^{N} a_j s^{j-1}}{\left(1 + \sum\limits_{j=1}^{N} b_j s^j\right)^2} = \dfrac{-y \cdot s^j}{\prod\limits_{j=1}^{N_{eff}} \left(s - p_j\right) \cdot \prod\limits_{j=1}^{N_{eff}} \left(s - p_j^*\right)}$
Proposed Rational-Based Neuro-TF model (with Decomposition)	$\hat{a}_{ij}$	$y = \sum\limits_{i=1}^{M} y_i = \sum\limits_{i=1}^{M} \dfrac{\sum\limits_{j=1}^{2N_i} \hat{a}_{ij} s^{j-1}}{1 + \sum\limits_{j=1}^{2N_i} \hat{b}_{ij} s^j}$	$\dfrac{\partial y}{\partial \hat{a}_{ij}} = \dfrac{s^{j-1}}{1 + \sum\limits_{j=1}^{2N_i} \hat{b}_{ij} s^j} = \dfrac{s^{j-1}}{\prod\limits_{j=1}^{N_i} \left(s - p_{n(i,j)}\right) \cdot \prod\limits_{j=1}^{N_i} \left(s - p_{n(i,j)}^*\right)}$
	$\hat{b}_{ij}$		$\dfrac{\partial y}{\partial \hat{b}_{ij}} = \dfrac{-s^j \sum\limits_{j=1}^{2N_i} \hat{a}_{ij} s^{j-1}}{\left(1 + \sum\limits_{j=1}^{2N_i} \hat{b}_{ij} s^j\right)^2} = \dfrac{-y_i \cdot s^j}{\prod\limits_{j=1}^{N_i} \left(s - p_{n(i,j)}\right) \cdot \prod\limits_{j=1}^{N_i} \left(s - p_{n(i,j)}^*\right)}$

We use $\hat{\boldsymbol{a}}_i$ and $\hat{\boldsymbol{b}}_i$ to represent vectors which contain all the coefficients of the transfer function numerator and denominator in the ith sub-neuro-TF model, respectively, defined as

$$\hat{\boldsymbol{a}}_i = \left[\hat{a}_{i1}\ \hat{a}_{i2}\ \cdots\ \hat{a}_{ij}\ \cdots\ \hat{a}_{i(2N_i)}\right]^T \tag{8}$$

and

$$\hat{\boldsymbol{b}}_i = \left[\hat{b}_{i1}\ \hat{b}_{i2}\ \cdots\ \hat{b}_{ij}\ \cdots\ \hat{b}_{i(2N_i)}\right]^T \tag{9}$$

where $i = 1, 2, ..., M$ and $j = 1, 2, ..., 2N_i$. Based on the definition of coefficients $\hat{\boldsymbol{a}}_i$ and $\hat{\boldsymbol{b}}_i$, we use $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ to represent vectors which contain all the coefficients of the transfer function numerators and denominators in all the sub-neuro-TF models, respectively, defined as

$$\hat{\boldsymbol{a}} = \left[\ \hat{\boldsymbol{a}}_1^T\ \hat{\boldsymbol{a}}_2^T\ \cdots\ \hat{\boldsymbol{a}}_i^T\ \cdots\ \hat{\boldsymbol{a}}_M^T\ \right]^T \tag{10}$$

and

$$\hat{\boldsymbol{b}} = \left[\ \hat{\boldsymbol{b}}_1^T\ \hat{\boldsymbol{b}}_2^T\ \cdots\ \hat{\boldsymbol{b}}_i^T\ \cdots\ \hat{\boldsymbol{b}}_M^T\ \right]^T \tag{11}$$

where $i = 1, 2, ..., M$.

Figure 1 shows the structure of the neuro-TF model incorporating the proposed decomposition technique.
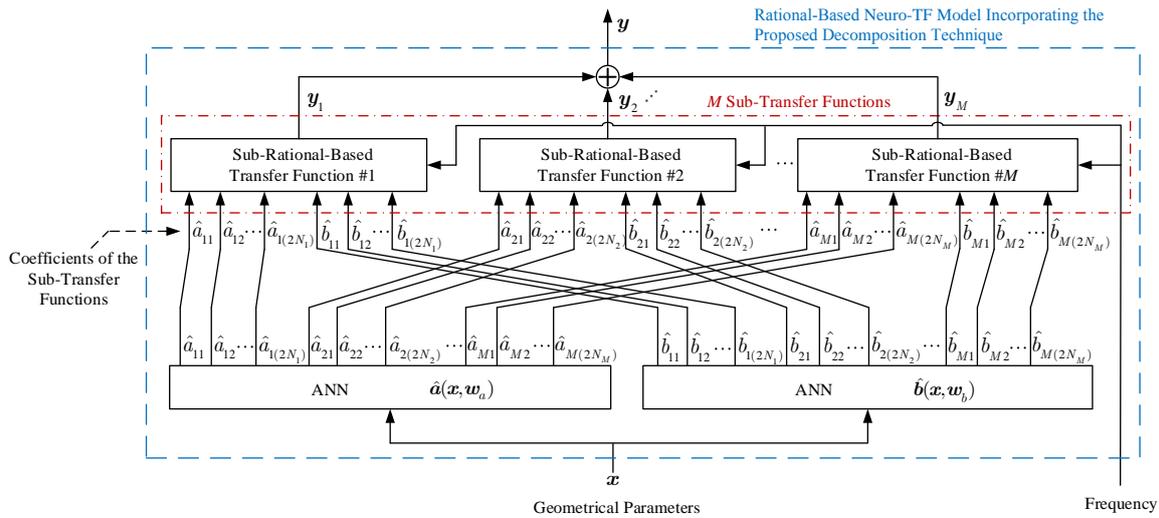
Figure 1. Structure for the rational-based neuro-TF model incorporating the proposed decomposition technique.

As can be seen in Figure 1, the proposed model consists of two separate neural networks. Neural network $\hat{a}(x, w_a)$ is used to learn and represent the nonlinear relationship between the geometrical variables x and the coefficients $\hat{a}$ of the transfer function numerators in all the sub-neuro-TF models. Neural network $\hat{b}(x, w_b)$ is used to learn and represent the nonlinear relationship between the geometrical variables x and the coefficients $\hat{b}$ of the transfer function denominators in all the sub-neuro-TF models.

In the next subsection, we propose a novel decomposition technique for the parameter extraction and development of the proposed rational-based neuro-TF model. New formulations are derived to determine the number M of sub-neuro-TF models and the effective order N_i of the transfer function for each sub-neuro-TF model. Using the proposed decomposition technique, we can decrease the sensitivities of the overall model response with respect to the coefficients of the transfer function in each sub-model, achieving better overall model accuracy.

3.2. Proposed Decomposition Technique for Parameter Extraction and Model Development

In this subsection, we propose a decomposition technique for parameter extraction and model development. The proposed decomposition technique starts with the EM data samples, e.g., *S*-parameter data for different training samples of the geometrical parameter. Let n_s represent the total number of training samples. Let T_r denote the index set of the training samples of the geometrical variables, i.e., $T_r = \{1, 2, ..., n_s\}$. Let x_k and d_k represent the *k*th sample of the geometrical variables and EM data, respectively, where $k \in T_r$. During data generation, frequency is swept by the EM simulator as a separate variable.

We perform the vector fitting process [26] to obtain poles and residues for each geometrical parameter sample. A scaling-and-shifting process for the frequency range with an even number of order N is set up during the vector fitting process to get all the poles and residues in complex values. Let $\tilde{p}^{(k)}$ and $\tilde{r}^{(k)}$ represent vectors which contain the poles and the residues obtained after vector fitting, respectively, at the *k*th geometrical sample, defined as

$$\tilde{p}^{(k)} = [\tilde{p}_1^{(k)} \; \tilde{p}_2^{(k)} \; \cdots \; \tilde{p}_{\tilde{\ell}}^{(k)} \; \cdots \; \tilde{p}_N^{(k)}]^T \tag{12}$$

and

$$\tilde{r}^{(k)} = [\tilde{r}_1^{(k)} \; \tilde{r}_2^{(k)} \; \cdots \; \tilde{r}_{\tilde{\ell}}^{(k)} \; \cdots \; \tilde{r}_N^{(k)}]^T \tag{13}$$

where $\tilde{p}_{\tilde{\ell}}^{(k)}$ and $\tilde{r}_{\tilde{\ell}}^{(k)}$ denote the $\tilde{\ell}$th related pole and residue, respectively; and $\tilde{\ell} \in \tilde{I} = \{1, 2, ..., N\}$. The obtained poles/residues contain both the effective poles/residues (i.e., the poles whose imaginary

parts are positive/the residues that related to the effective poles) and the complex conjugate parts of the effective poles/residues.

In the first stage of the proposed decomposition technique, we sort the effective poles and residues in ascending sequence according to the values of the imaginary parts of the effective poles. The effective poles obtained from vector fitting may not be in the same sequence from sample to sample, which will result in the discontinuity of the effective poles over the geometrical parameters from sample to sample. The sorting process can make the effective poles located in the same sequence among different geometrical samples, minimizing the discontinuity of the effective poles. With minimum discontinuity issue in the effective poles, the proposed decomposition technique can be more robust, and the obtained sub-neuro-TF models can achieve better accuracy. Let the effective poles and residues after sorting be denoted as $p^{(k)}$ and $r^{(k)}$, respectively. The nth elements of $p^{(k)}$ and $r^{(k)}$ are denoted as $p_n^{(k)}$ and $r_n^{(k)}$, respectively, calculated as

$$p_n^{(k)} = \tilde{p}_{\ell_n}^{(k)} \tag{14}$$

and

$$r_n^{(k)} = \tilde{r}_{\ell_n}^{(k)} \tag{15}$$

where

$$\ell_n = \arg \min_{\tilde{\ell} \in D \setminus D_n} \left\{ \mathrm{Im}(\tilde{p}_{\tilde{\ell}}^{(k)}) \right\} \tag{16}$$

$$D = \left\{ \tilde{\ell} \,\middle|\, \mathrm{Im}(\tilde{p}_{\tilde{\ell}}^{(k)}) > 0, \, \tilde{\ell} \in \tilde{I} \right\} \tag{17}$$

$$D_n = \begin{cases} \varnothing, & \text{if } n = 1 \\ D_{n-1} \cup \{\ell_{n-1}\} & \text{if } n > 1 \end{cases} \tag{18}$$

and $n \in I = \left\{ 1, 2, ..., N_{eff} \right\}$, where $N_{eff} = N/2$.

The second stage of the proposed decomposition technique is to quantify the degree of smoothness of the sorted effective pole/residue data with respect to the changing of the geometrical variables. The sorted effective poles and residues are correlated with the values of geometrical variables when the values of geometrical variables vary continuously. The quantified degree of smoothness can be used as an indicator for decomposition. Let σ_n^p represent the deviation of the nth effective pole as the geometrical parameter changes. Similarly, let σ_n^r represent the deviation of the nth effective residue as the geometrical parameter changes. The deviations σ_n^p and σ_n^r are formulated as

$$\sigma_n^p = \max_{k \in T_r} \left\{ \left(\left| \frac{\mathrm{Re}(p_n^{(k)}) - \mathrm{Re}(\mu_n^p)}{\mathrm{Re}(\mu_n^p)} \right|^2 + \left| \frac{\mathrm{Im}(p_n^{(k)}) - \mathrm{Im}(\mu_n^p)}{\mathrm{Im}(\mu_n^p)} \right|^2 \right)^{1/2} \right\} \tag{19}$$

and

$$\sigma_n^r = \max_{k \in T_r} \left\{ \left(\left| \frac{\mathrm{Re}(r_n^{(k)}) - \mathrm{Re}(\mu_n^r)}{\mathrm{Re}(\mu_n^r)} \right|^2 + \left| \frac{\mathrm{Im}(r_n^{(k)}) - \mathrm{Im}(\mu_n^r)}{\mathrm{Im}(\mu_n^r)} \right|^2 \right)^{1/2} \right\} \tag{20}$$

where μ_n^p and μ_n^r are expressed as

$$\mu_n^p = \frac{1}{n_s} \sum_{k=1}^{n_s} p_n^{(k)} \tag{21}$$

and

$$\mu_n^r = \frac{1}{n_s} \sum_{k=1}^{n_s} r_n^{(k)}. \tag{22}$$

The deviations σ_n^p and σ_n^r are computed using all the geometrical samples to represent the variations of each sorted effective pole and residue. Here we define the deviation vector σ as

$$\sigma = [\sigma_1, \sigma_2, ..., \sigma_n, ..., \sigma_{N_{eff}}]^T \tag{23}$$

where

$$\sigma_n = \sqrt{(\sigma_n^p)^2 + (\sigma_n^r)^2} \tag{24}$$

and $n \in I$. The deviation vector σ reflects the degree of smoothness of the effective pole/residue data ($p^{(k)}$ and $r^{(k)}$). It can be seen that the size of σ is consistent with the size of the index set $I = \{1, 2, ..., N_{eff}\}$. Based on the different values of σ_n, the index set I will be decomposed into multiple subsets for parameter extraction.

In the third stage, we decompose the index set $I = \{1, 2, ..., N_{eff}\}$ into M subsets based on σ and M. The initial value of M is set to two. The ultimate value of M will be determined through an iterative decomposition process for the proposed rational-based neuro-TF model, which will be explained later on. Since the size of the ith subset depends on the effective order N_i of the ith sub-neuro-TF model, we determine the effective order N_i, formulated as

$$N_i = \begin{cases} \left\lceil \dfrac{N_{eff}}{M} \right\rceil, & \text{if } i = 1 \\[4mm] \left\lceil \dfrac{N_{eff} - \sum\limits_{j=1}^{i-1} N_j}{M - i + 1} \right\rceil, & \text{if } i > 1 \end{cases} \tag{25}$$

where $i = 1, 2, ..., M$. By using (25), N_i is distributed as evenly as possible for each sub-neuro-TF model.

After the determination of N_i, we sort the elements of the deviation vector σ in descending sequence. Let $\hat{u}$ denote a vector which contains the indices of the elements of σ after sorting. Let $\hat{u}_m$ denote the mth element of $\hat{u}$, formulated as

$$\hat{u}_m = \arg \max_{u \in I \backslash J_m} \{\sigma_u\} \tag{26}$$

where

$$J_m = \begin{cases} \varnothing, & \text{if } m = 1 \\ J_{m-1} \cup \{\hat{u}_{m-1}\}, & \text{if } m > 1 \end{cases} \tag{27}$$

and $m \in I$. We define the ith subset $I_i \subseteq I$, where $i = 1, 2, ..., M$; $|I_i| = N_i$; $\bigcup_{i=1}^{M} I_i = I$; and $\left| \bigcup_{i=1}^{M} I_i \right| = |I| = N_{eff}$. Using the information of the index vector u and the effective order N_i, the ith subset I_i can be expressed as

$$I_i = \{\hat{u}_{q+1}, \hat{u}_{q+2}, ..., \hat{u}_{q+N_i}\} \tag{28}$$

where q can be formulated as

$$q = \begin{cases} 0, & \text{if } i = 1 \\ \sum\limits_{t=1}^{i-1} N_t, & \text{if } i > 1 \end{cases}. \tag{29}$$

In the fourth stage, we group the effective poles $p^{(k)}$ and their complex conjugate parts $(p^*)^{(k)}$ into pole subsets and group the effective residues $r^{(k)}$ and their complex conjugate parts $(r^*)^{(k)}$ into

residue subsets, based on the information of subsets $I_i, \forall i = 1, 2, ..., M$. Let $\widehat{p}_i^{(k)}$ and $\widehat{r}_i^{(k)}$ denote the pole and residue data for the ith sub-neuro-TF model, respectively, expressed as

$$\widehat{p}_i^{(k)} = \left[p_{\hat{u}_{q+1}}^{(k)} \ (p_{\hat{u}_{q+1}}^*)^{(k)} \ p_{\hat{u}_{q+2}}^{(k)} \ (p_{\hat{u}_{q+2}}^*)^{(k)} \ \cdots \ p_{\hat{u}_{q+N_i}}^{(k)} \ (p_{\hat{u}_{q+N_i}}^*)^{(k)} \right]_{1 \times 2N_i}^T \tag{30}$$

and

$$\widehat{r}_i^{(k)} = \left[r_{\hat{u}_{q+1}}^{(k)} \ (r_{\hat{u}_{q+1}}^*)^{(k)} \ r_{\hat{u}_{q+2}}^{(k)} \ (r_{\hat{u}_{q+2}}^*)^{(k)} \ \cdots \ r_{\hat{u}_{q+N_i}}^{(k)} \ (r_{\hat{u}_{q+N_i}}^*)^{(k)} \right]_{1 \times 2N_i}^T \tag{31}$$

where $i = 1, 2, ..., M$; and q is calculated by (29).

We convert these grouped pole and residue data (i.e., $\widehat{p}_i^{(k)}$ and $\widehat{r}_i^{(k)}$) into the coefficient data of the transfer function numerator $\hat{a}_i^{(k)}$ and the transfer function denominator $\hat{b}_i^{(k)}$ for the ith sub-neuro-TF model. Since the nonlinear relationships between the geometrical variables x and the coefficients of the numerators $\hat{a}_i$ are usually similar among different sub-neuro-TF models, we use one single neural network $\hat{a}(x, w_a)$ to learn the relationships between the geometrical variables x and the coefficients of the numerators $\hat{a}$ of all the sub-neuro-TF models [25]. We use another single neural network $\hat{b}(x, w_b)$ to learn the relationships between the geometrical variables x and the coefficients of the denominators $\hat{b}$ of all the sub-neuro-TF models. Let $(x_k, \hat{a}^{(k)})$ and $(x_k, \hat{b}^{(k)})$ denote the training data which are used to train the two neural networks $\hat{a}(x, w_a)$ and $\hat{b}(x, w_b)$, respectively, expressed as

$$\hat{a}^{(k)} = \left[(\hat{a}_1^{(k)})^T \ (\hat{a}_2^{(k)})^T \ \cdots \ (\hat{a}_i^{(k)})^T \ \cdots \ (\hat{a}_M^{(k)})^T \right]^T \tag{32}$$

and

$$\hat{b}^{(k)} = \left[(\hat{b}_1^{(k)})^T \ (\hat{b}_2^{(k)})^T \ \cdots \ (\hat{b}_i^{(k)})^T \ \cdots \ (\hat{b}_M^{(k)})^T \right]^T \tag{33}$$

where $i = 1, 2, ..., M$.

In the fifth stage, we train the proposed neuro-TF model. First, a preliminary training process [22] is performed for the two neural networks $\hat{a}(x, w_a)$ and $\hat{b}(x, w_b)$ using the obtained training data $(x_k, \hat{a}^{(k)})$ and $(x_k, \hat{b}^{(k)})$, respectively. The two neural networks are trained separately to learn the relationships between the geometrical variables x and the coefficients (i.e., $\hat{a}$ and $\hat{b}$). After the preliminary training, we combine the trained neural networks with the transfer functions to obtain the overall proposed neuro-TF model. Even though the neural networks are trained well, the training and testing errors of the overall proposed model may still not satisfy the accuracy criteria. To make the overall proposed model satisfy the accuracy criteria, we perform a model refinement training process [22] to refine the overall proposed model. The training data for the proposed overall model are (x_k, d_k), where $k \in T_r$. The objective of the training is to minimize the training error E_{Tr} of the overall model by optimizing the neural network weights w_a and w_b. The training error function E_{Tr} is formulated as

$$E_{Tr}(w_a, w_b) = \frac{1}{2n_s} \sum_{k \in T_r} \sum_{\lambda \in \Omega} \left\| \sum_{i=1}^{M} y_i(w_a, w_b, x_k, s_\lambda) - d_{k,\lambda} \right\|^2 \tag{34}$$

where n_s denotes the total number of training samples; Ω denotes the index set of frequency samples; M represents the number of the sub-neuro-TF models; and y_i represents the output of the ith sub-neuro-TF model, which is also a function of x_k, w_a and w_b, and s_λ.

We use the training data (x_k, d_k) to verify the quality of trained overall model. If the training error E_{Tr} is lower than a user-defined threshold E_t, the whole training process is terminated and the model is ready for testing. Otherwise, the two neural networks are under-learned. In that case, we should increase the number of hidden neurons for the two neural networks and repeat the two training processes.

After the model is trained, we test the quality of the model using independent testing data which are never used in training. We define E_{Ts} to be the testing error. If the testing error E_{Ts} is also lower

than the threshold E_t, the model development is finished and the current value of M is regarded as the final M. Otherwise, we increase the number M of sub-neuro-TF models by one (i.e., $M = M + 1$) and repeat the parameter extraction and training processes from the third stage to the fifth stage (25)–(34) of the proposed decomposition technique.

Here we summarize the proposed decomposition technique in a stepwise algorithm as follows.

Step 1: Select the effective poles and residues ($\boldsymbol{p}^{(k)}$ and $\boldsymbol{r}^{(k)}$) by (12)–(18).

Step 2: Obtain the deviation vector $\boldsymbol{\sigma}$ by (19)–(24).

Step 3: Initialize the number M of the sub-neuro-TF models to two.

Step 4: Obtain the effective order N_i by (25) and subset I_i by (26)–(29) with the current value of M for the ith sub-neuro-TF model.

Step 5: Obtain the pole and residue data $\hat{\boldsymbol{p}}_i^{(k)}$ and $\hat{\boldsymbol{r}}_i^{(k)}$ by (30)–(31) for the ith sub-neuro-TF model. Convert data $\hat{\boldsymbol{p}}_i^{(k)}$ and $\hat{\boldsymbol{r}}_i^{(k)}$ into data $\hat{\boldsymbol{a}}_i^{(k)}$ and $\hat{\boldsymbol{b}}_i^{(k)}$. Obtain the training data $(\boldsymbol{x}_k, \hat{\boldsymbol{a}}^{(k)})$ and $(\boldsymbol{x}_k, \hat{\boldsymbol{b}}^{(k)})$ by (32)–(33) for the two neural networks $\hat{\boldsymbol{a}}(\boldsymbol{x}, \boldsymbol{w}_a)$ and $\hat{\boldsymbol{b}}(\boldsymbol{x}, \boldsymbol{w}_b)$.

Step 6: Perform the preliminary training of the two neural networks and refinement training of the overall model by (34).

Step 7: Use the training data $(\boldsymbol{x}_k, \boldsymbol{d}_k)$ to verify the trained overall model. If the training error E_{Tr} is lower than a user-defined threshold E_t, go to *Step 8*. Otherwise, increase the number of hidden neurons and go to *Step 6*.

Step 8: Use the testing data to verify the overall model. If the testing error E_{Ts} is lower than the user-defined threshold E_t, go to *Step 9*. Otherwise, increase the number M of sub-neuro-TF models by one (i.e., $M = M + 1$) and go to *Step 4*.

Step 9: Stop the modeling process.

Figure 2 shows a flow diagram of the development process for the overall neuro-TF model using the proposed decomposition technique.
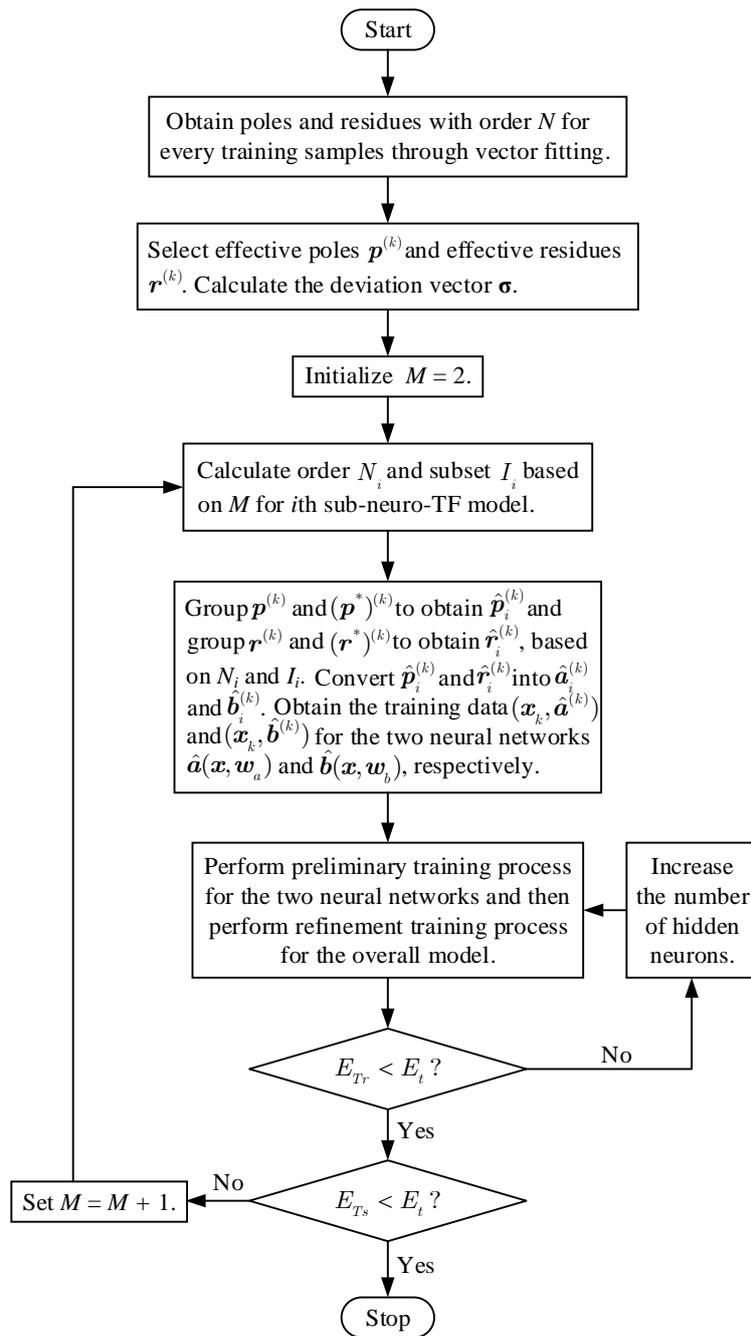
Figure 2. The flow diagram of the overall model development process using the proposed decomposition technique.

4. Application Examples

4.1. Three-Order Waveguide Filter Modeling

A three-order waveguide filter [27] is used to illustrate the rational-based neuro-TF method using the proposed decomposition technique. As can be seen in Figure 3, a = 19.05 mm, b = 9.525 mm, and t = 2.0 mm and the geometrical variables are $x = [L_1\ L_2\ W_1\ W_2]^T$ for the three-order waveguide filter. Four geometrical variables, i.e., $x = [L_1\ L_2\ W_1\ W_2]^T$, are used as model inputs and the real and imaginary parts of S_{11}, i.e., $y = [RS_{11}\ IS_{11}]^T$, are used as model outputs.
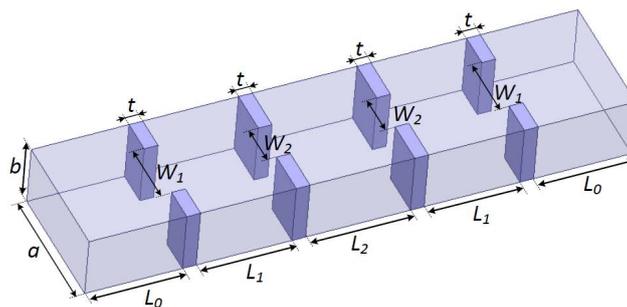
Figure 3. Geometrical variables and 3D configuration of the three-order waveguide filter example.

Full-wave EM simulations are performed using the ANSYS HFSS EM simulator to generate training and test samples for parametric modeling. Both training samples and test samples are generated using the design of experiments (DOE) [28] sampling method. As shown in Table 2, the proposed parametric modeling technique is performed for three cases according to different geometrical parameter ranges. Case 1 is with a narrower range, Case 2 is with an increased range, and Case 3 is with a wider range. In all cases, we set the total order N of the overall transfer function to 10 for every sample. Subsequently, N_{eff} equals five for the three cases. In Case 1 and Case 2, the frequency range is from 11.65 GHz to 12.35 GHz and the number of frequency samples per geometrical sample equals 71. In Case 3, the frequency range is from 11.0 GHz to 13.0 GHz and the number of frequency samples per geometrical sample equals 201. We use seven levels of DOE for all cases to generate the training and test samples. The numbers for training samples and testing samples are both 49. Table 2 lists the specific ranges of training samples and test samples for the three cases.

Table 2. Definition of training and test samples for the three-order waveguide filter example.

Geometrical Parameters (mm)		Training Samples (49 Samples)			Test Samples (49 Samples)		
		Min	Max	Steps	Min	Max	Steps
Case 1 (Narrower Range)	L_1	13.84	14.12	0.05	13.86	14.10	0.04
	L_2	15.05	15.35	0.05	15.07	15.33	0.04
	W_1	8.91	9.09	0.03	8.93	9.08	0.03
	W_2	5.94	6.06	0.02	5.95	6.05	0.02
Case 2 (Increased Range)	L_1	13.70	14.26	0.09	13.75	14.21	0.08
	L_2	14.90	15.50	0.10	14.95	15.45	0.08
	W_1	8.82	9.18	0.06	8.85	9.15	0.05
	W_2	5.88	6.12	0.04	5.90	6.10	0.03
Case 3 (Wider Range)	L_1	13.28	14.68	0.23	13.40	14.56	0.19
	L_2	14.44	15.96	0.25	14.57	15.83	0.21
	W_1	8.55	9.45	0.15	8.63	9.38	0.13
	W_2	5.70	6.30	0.10	5.75	6.25	0.08

We scale and shift the frequency range and set the total order N to ten (an even number) during the vector fitting process, ensuring that the obtained poles and residues are in complex values.

Since the initial value of the number M of sub-neuro-TF models is set to two, the proposed model is initially trained with $M = 2$. Using the proposed decomposition technique with $M = 2$, the effective orders are determined as $N_1 = 3$ and $N_2 = 2$ for all the three cases and the subsets are determined

as $I_1 = \{2,4,5\}$ and $I_2 = \{1,3\}$ for Case 1 and Case 2 and $I_1 = \{2,3,4\}$ and $I_2 = \{1,5\}$ for Case 3. The proposed neuro-TF model with $M=2$ has a frequency variable for each sub-neuro-TF model, four input geometrical variables, and two outputs. The overall transfer function consists of two sub-transfer functions. The high effective order N_{eff} of 5 is decomposed into lower effective orders N_1 of 3 and N_2 of 2.

The NeuroModelerPlus software is used for training of the proposed model. Suitable numbers of hidden neurons are used to achieve a good learning of the proposed neuro-TF model. The user-defined threshold E_t for the three-order waveguide filter example is set to 1%. Three parametric models are developed for the three cases using the proposed technique. The average training errors of the trained models with $M = 2$ are 0.239%, 0.467%, and 1.490% for Case 1, Case 2, and Case 3, respectively. The average testing errors of the trained models with $M = 2$ are 0.267%, 0.496%, and 1.840% for Case 1, Case 2, and Case 3, respectively. On the one hand, since the training and testing errors for both Case 1 and Case 2 are lower than E_t, the two trained models with $M = 2$ are the final models for both Case 1 and Case 2. On the other hand, since the testing error for Case 3 is higher than E_t, we increased M to three and redeveloped the model with $M = 3$ for Case 3. Figure 4 illustrates the model structure with $M = 3$ for the three-order waveguide filter example. As illustrated in Figure 4, the overall transfer function consists of three sub-transfer functions. The high effective order N_{eff} of 5 is decomposed into lower effective orders N_1 of 2, N_2 of 2, and N_3 of 1. The subsets are determined as $I_1 = \{2,4\}$, $I_2 = \{3,5\}$, and $I_3 = \{1\}$ for Case 3. The average training and testing errors of the trained models with $M = 3$ are 0.746% and 0.962%, which are lower than E_t. Therefore, the trained model with $M = 3$ is the final model for Case 3. The number of different parameter extractions in the proposed technique depends on both M and the number n_s of total training samples. Specifically, the number of different parameter extractions equals n_s multiplied by (M-1). In this example, we performed 49 different parameter extractions for Cases 1 and 2, and 98 different parameter extractions for Case 3.
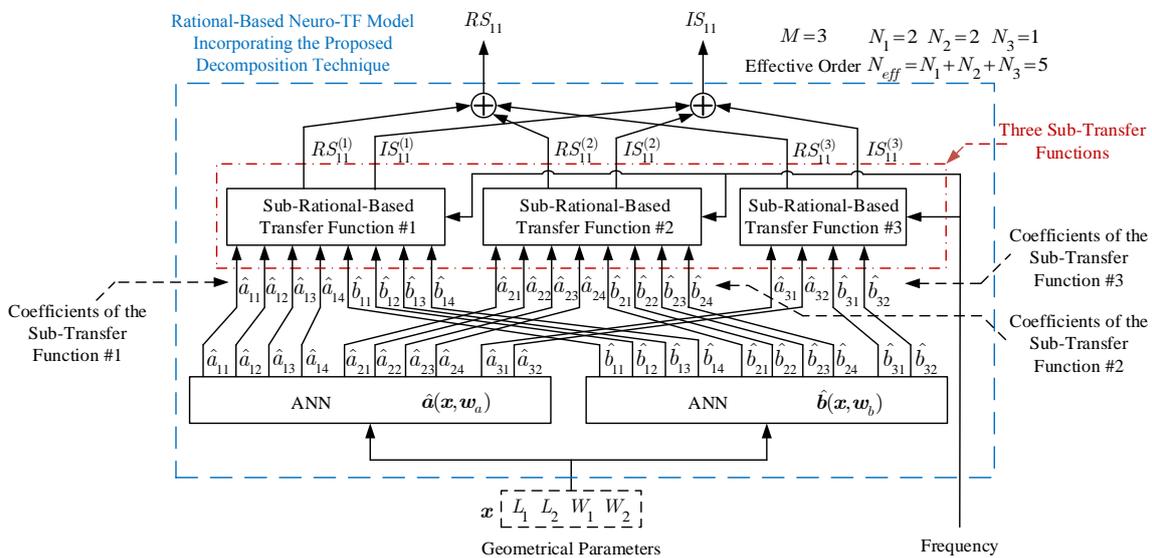
Figure 4. Structure for the proposed rational-based neuro-TF model with $M = 3$ of the three-order waveguide filter example.

For comparison purpose, we apply the existing modeling approach using rational-based neuro-TF [25] to the three separate cases of this example.

Comparisons of number M of sub-neuro-TF models, effective order N_i for ith sub-neuro-TF model, hidden neuron numbers, and average errors for training and testing between different rational-based neuro-TF modeling methods and EM data are shown in Table 3.

Table 3. Comparisons of different rational-based neuro-TF modeling approaches for the three-order waveguide filter example.

Modeling Methods		No. of Sub-Models	N_{eff} or N_i	No. of Hidden Neurons		Average Training Error	Average Testing Error
Case 1 (Narrower Range)	Existing Rational Neuro-TF Method	1	5	NN for Numerator	10	0.630 %	0.718 %
				NN for Denominator	10		
	Proposed Rational Neuro-TF Method	2	3 2	NN for Numerator	10	0.239 %	0.267 %
				NN for Denominator	10		
Case 2 (Increased Range)	Existing Rational Neuro-TF Method	1	5	NN for Numerator	10	1.953 %	2.017 %
				NN for Denominator	10		
	Proposed Rational Neuro-TF Method	2	3 2	NN for Numerator	10	0.467 %	0.496 %
				NN for Denominator	10		
Case 3 (Wider Range)	Existing Rational Neuro-TF Method	1	5	NN for Numerator	10	5.073 %	6.604 %
				NN for Denominator	10		
		1	5	NN for Numerator	40	3.222 %	50.69%
				NN for Denominator	40		
	Proposed Rational Neuro-TF Method	2	3 2	NN for Numerator	10	1.490 %	1.840 %
				NN for Denominator	10		
		3	2 2 1	NN for Numerator	10	0.746 %	0.962 %
				NN for Denominator	10		

In Case 1, since the geometrical variations are small, the relatively low nonlinear relationship between the geometrical variables and the coefficients of the transfer function can be easily represented by the neural networks with less hidden neurons of both the existing model and the proposed model in the preliminary training. With the high training accuracy of the neural networks of both models, both models can have high training accuracy in the refinement training. For these reasons, both methods in this case can obtain good training and testing accuracy.

In Case 2, since the geometrical variables vary within an increased range, the relationship between the geometrical variables and the coefficients of the transfer function becomes more nonlinear than that in Case 1. In this case, the sensitivities of the overall model response with respect to the coefficients of the transfer function have higher effects on the accuracy of the overall model. Since the orders and the sensitivities for the proposed model are much lower than those for the existing model, the proposed model can obtain more accurate training and testing results than the existing method.

In Case 3, the geometrical variables vary within a wider range, the relationship between the geometrical variables and the coefficients of the transfer function becomes highly nonlinear. In this case, the sensitivities of the overall model response with respect to the coefficients of the transfer function have large effects on the accuracy of the overall model. The proposed model with $M = 2$ have smaller training and testing errors than the existing model but still can not satisfy the model accuracy criteria. Therefore, we increased the M to be 3. With $M = 3$, the proposed model can satisfy the model accuracy criteria. It can be concluded from the above three cases that the neuro-TF model using the proposed decomposition technique has much lower sensitivities of the overall model response with respect to the coefficients of the transfer function, improving the modeling accuracy over the existing method.

Figure 5 shows the output $|S_{11}|$ in dB of the proposed neuro-TF model for three different test geometrical samples in Case 3 of the three-order waveguide filter example. We compare the model response using the proposed technique with $M = 3$, using the proposed technique with $M = 2$, using the existing neuro-TF method, and EM data. We also provide comparisons of the real and imaginary parts of S_{11}, as shown in Figure 6. The geometrical variables for the three test samples are $x = [14.37\ 14.57\ 9.25\ 6.17]^T$ (mm), $x = [14.17\ 15.20\ 8.75\ 6.25]^T$ (mm), and $x = [14.56\ 15.20\ 8.63\ 6.17]^T$ (mm). These three test samples belong to Case 3.
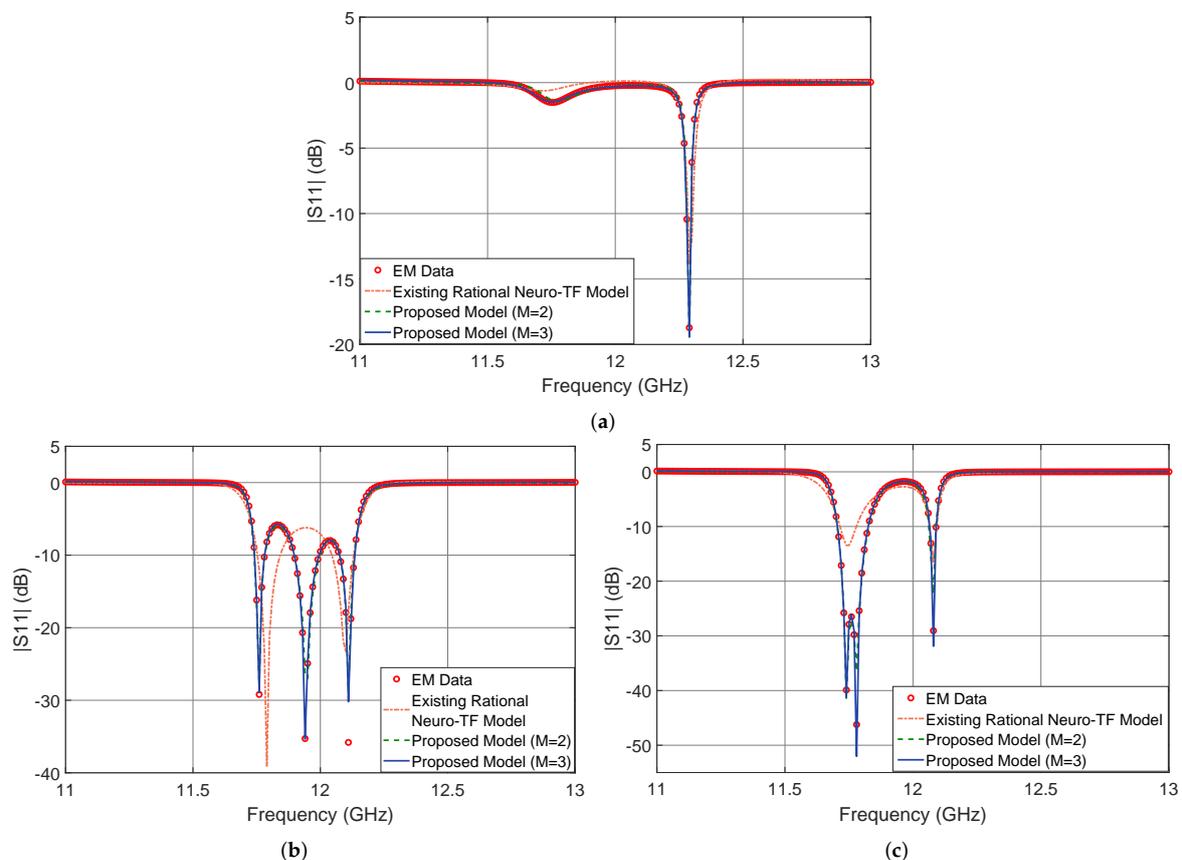
Figure 5. Comparison of magnitude in decibels of S_{11} different modeling approaches and EM data: (**a**) test sample $x = [14.37\ 14.57\ 9.25\ 6.17]^T$ (mm), (**b**) test sample $x = [14.17\ 15.20\ 8.75\ 6.25]^T$ (mm), and (**c**) test sample $x = [14.56\ 15.20\ 8.63\ 6.17]^T$ (mm) for the three-order waveguide filter example.

It is observed from Figures 5 and 6 that the neuro-TF model using the proposed decomposition method can achieve better model accuracy for different geometrical samples which are never used in training.
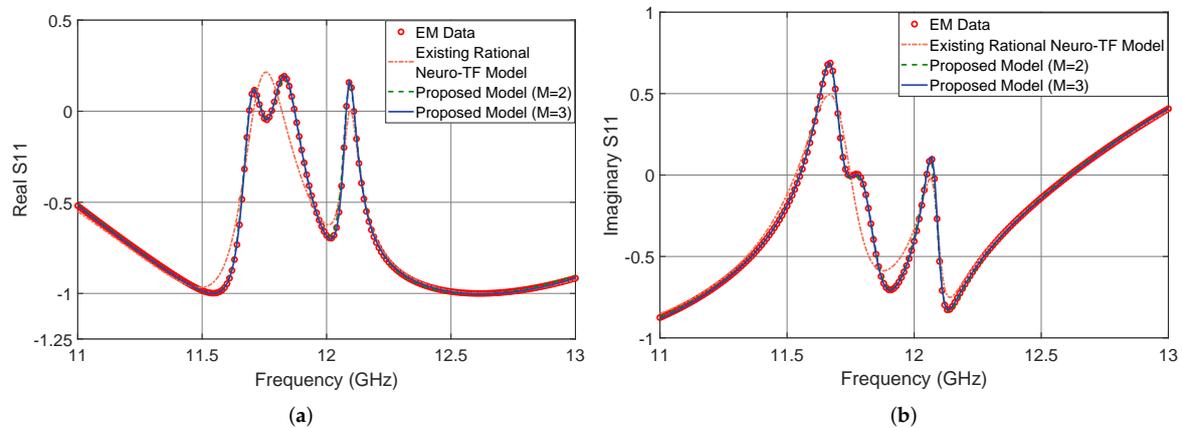
Figure 6. Comparison of $\text{Re}(S_{11})$ and $\text{Im}(S_{11})$ for different modeling approaches and EM data at test sample $x = [14.56\ 15.20\ 8.63\ 6.17]^T$ (mm) for the three-order waveguide filter example. (**a**) $\text{Re}(S_{11})$ and (**b**) $\text{Im}(S_{11})$.

4.2. Four-Order Bandpass Filter Modeling

A four-order bandpass filter [29] is used to illustrate the rational-based neuro-TF method using the proposed decomposition technique. As can be seen in Figure 7, the geometrical variables are $x = [h_1\ h_2\ h_3\ h_{c1}\ h_{c2}]^T$ for the four-order bandpass filter, where h_1, h_2, and h_3 denote the heights of the posts between the coupling windows, and h_{c1} and h_{c2} denote the heights of the posts in the resonant cavities. These five geometrical variables are used as model inputs and the real and imaginary parts of S_{11} are used as model outputs.
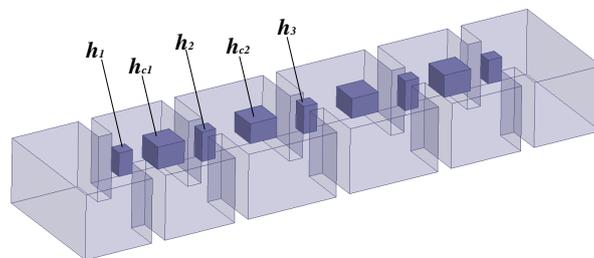
Figure 7. Geometrical variables and 3D configuration of the four-order bandpass filter example.

Similar to the first example, ANSYS HFSS EM simulator and DOE sampling method are used for generation of training and test samples. As shown in Table 4, the proposed parametric modeling technique is performed for two different cases according to different geometrical parameter ranges. Case 1 is with a narrower parameter range and Case 2 is with a wider parameter range. In both cases, we set the total order N of the overall transfer function to 12 for every sample. Subsequently, N_{eff} equals six for the two cases. In both cases, the frequency range is from 10.50 GHz to 11.50 GHz and the number of frequency samples per geometrical sample equals 101. We use nine levels of DOE to generate the training samples and eight levels of DOE to generate the test samples for both cases. The numbers for the training samples and the testing samples are 81 and 64, respectively. Table 4 lists the specific ranges of training samples and test samples for the two cases.

Table 4. Definition of training and test samples for the four-order bandpass filter example.

Geometrical Parameters (mm)		Training Samples (81 Samples)			Test Samples (64 Samples)		
		Min	Max	Steps	Min	Max	Steps
Case 1 (Narrower Range)	h_1	3.4	3.56	0.02	3.41	3.55	0.02
	h_2	4.3	4.46	0.02	4.31	4.45	0.02
	h_3	4.0	4.16	0.02	4.01	4.15	0.02
	h_{c1}	3.2	3.36	0.02	3.21	3.35	0.02
	h_{c2}	2.9	3.06	0.02	2.91	3.05	0.02
Case 2 (Wider Range)	h_1	3.3	3.62	0.04	3.32	3.6	0.04
	h_2	4.2	4.52	0.04	4.22	4.5	0.04
	h_3	3.9	4.22	0.04	3.92	4.2	0.04
	h_{c1}	3.1	3.42	0.04	3.12	3.4	0.04
	h_{c2}	2.8	3.12	0.04	2.82	3.1	0.04

Similar to the first example, the proposed model is initially trained with $M = 2$. With $M = 2$, the effective orders are determined as $N_1 = 3$ and $N_2 = 3$ for both cases and the subsets are determined as $I_1 = \{2,4,5\}$ and $I_2 = \{1,3,6\}$ for both cases. The proposed neuro-TF model with $M = 2$ has a frequency variable for each sub-neuro-TF model, five input geometrical variables, and two outputs. The overall transfer function consists of two sub-transfer functions. The high effective order N_{eff} of 6 is decomposed into lower effective orders N_1 of 3 and N_2 of 3.

The NeuroModelerPlus software is used for training of the proposed model. The user-defined threshold E_t for the four-order bandpass filter example is set to 2%. Two parametric models are developed for the two cases using the proposed technique. The average training errors of the trained models with $M = 2$ are 1.487% and 2.252% for Case 1 and Case 2, respectively. The average testing errors of the trained models with $M = 2$ are 1.672% and 4.397% for Case 1 and Case 2, respectively. On the one hand, since the training and testing errors for Case 1 are lower than E_t, the trained model with $M = 2$ is the final model for Case 1. On the other hand, since the testing error for Case 2 is higher than E_t, we increased M to three and redeveloped the model with $M = 3$ for Case 2. Figure 8 illustrates the model structure with $M = 3$ for the four-order bandpass filter example. As illustrated in Figure 8, the overall transfer function consists of three sub-transfer functions. The high effective order N_{eff} of 6 is decomposed into lower effective orders N_1 of 2, N_2 of 2, and N_3 of 2. The subsets are determined as $I_1 = \{2,5\}$, $I_2 = \{3,4\}$, and $I_3 = \{1,6\}$ for Case 2. The average training and testing errors of the trained model with $M = 3$ are 1.624% and 1.982%, respectively, which are lower than E_t. Therefore, the trained model with $M = 3$ is the final model for Case 2. In this example, we performed 162 different parameter extractions for each case.
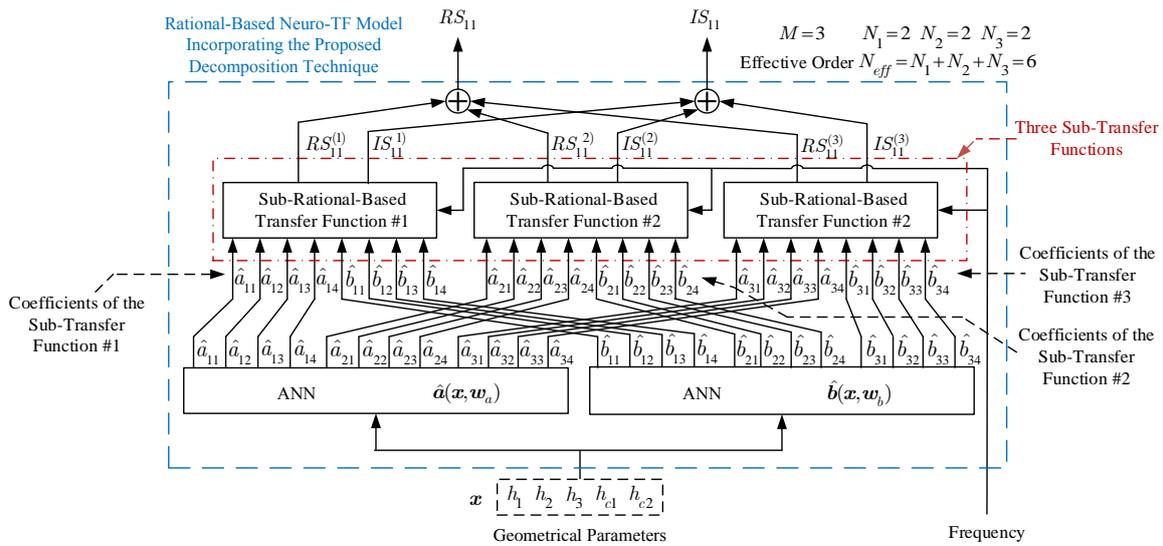
Figure 8. Structure for the proposed rational-based neuro-TF model with $M = 3$ of the four-order bandpass filter example.

For comparison purpose, we apply the existing modeling approach using rational-based neuro-TF to the two separate cases of this example. Comparisons of number M of sub-neuro-TF models, effective order N_i for ith sub-neuro-TF model, hidden neuron numbers, and average errors for training and testing between different rational-based neuro-TF modeling methods and EM data are shown in Table 5.

In Case 1, since the geometrical variations are relatively small, the relationship between the geometrical variables and the coefficients of the transfer function is moderately nonlinear. In this case, the sensitivities of the overall model response with respect to the coefficients of the transfer function have moderate effects on the accuracy of the overall model. Since the orders and the sensitivities for the proposed model are much lower than those for the existing model, the proposed model can obtain more accurate training and testing results than the existing method.

In Case 2, the geometrical variables vary within a wider range, the relationship between the geometrical variables and the coefficients of the transfer function becomes highly nonlinear. In this case, the sensitivities of the overall model response with respect to the coefficients of the transfer function have large effects on the accuracy of the overall model. The proposed model with $M = 2$ have smaller training and testing errors than the existing model but still can not satisfy the model accuracy criteria. Therefore, we increased the M to be 3. With $M = 3$, the proposed model can satisfy the model accuracy criteria. It can be again concluded from the above two cases that the neuro-TF model using the proposed decomposition technique has much lower sensitivities of the overall model response with respect to the coefficients of the transfer function, improving the modeling accuracy over the existing method.

Figure 9 shows the output $|S_{11}|$ in dB of the proposed neuro-TF model for three different test geometrical samples in Case 2 of the four-order bandpass filter example. We compare the model response using the proposed technique with $M = 3$, using the proposed technique with $M = 2$, using the existing neuro-TF method, and EM data. We also provide comparisons of the real and imaginary parts of S_{11}, as shown in Figure 10. The geometrical variables for the three test samples are $x = [3.52\ 4.46\ 3.96\ 3.28\ 3.07]^T$ (mm), $x = [3.60\ 4.22\ 4.20\ 3.40\ 3.10]^T$ (mm), and $x = [3.36\ 4.38\ 4.00\ 3.36\ 3.02]^T$ (mm).

(**a**)

(**b**)   (**c**)
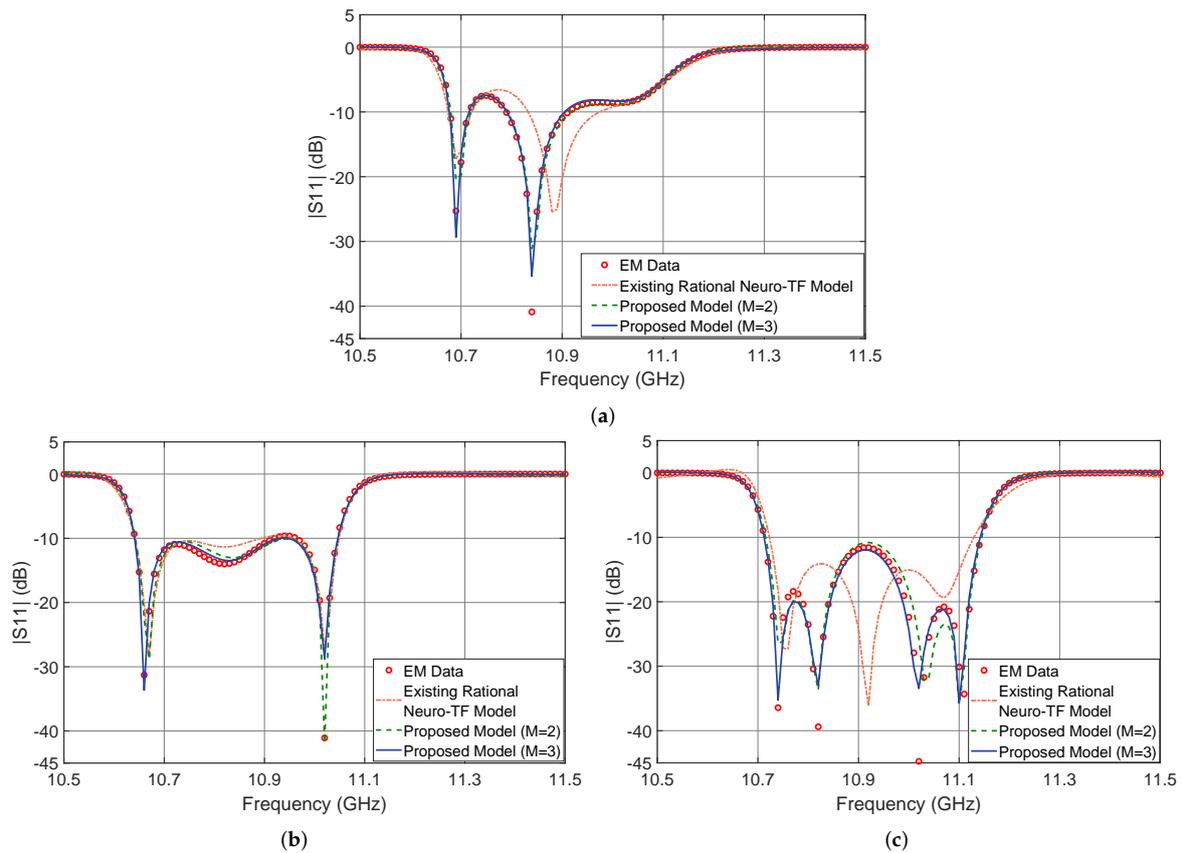
Figure 9. Comparison of magnitude in decibels of S_{11} for different modeling approaches and EM data:
(**a**) test sample $x = [3.52\ 4.46\ 3.96\ 3.28\ 3.07]^T$ (mm), (**b**) test sample , $x = [3.60\ 4.22\ 4.20\ 3.40\ 3.10]^T$ (mm),
and (**c**) test sample $x = [3.36\ 4.38\ 4.00\ 3.36\ 3.02]^T$ (mm) for the four-order bandpass filter example.

(**a**)   (**b**)
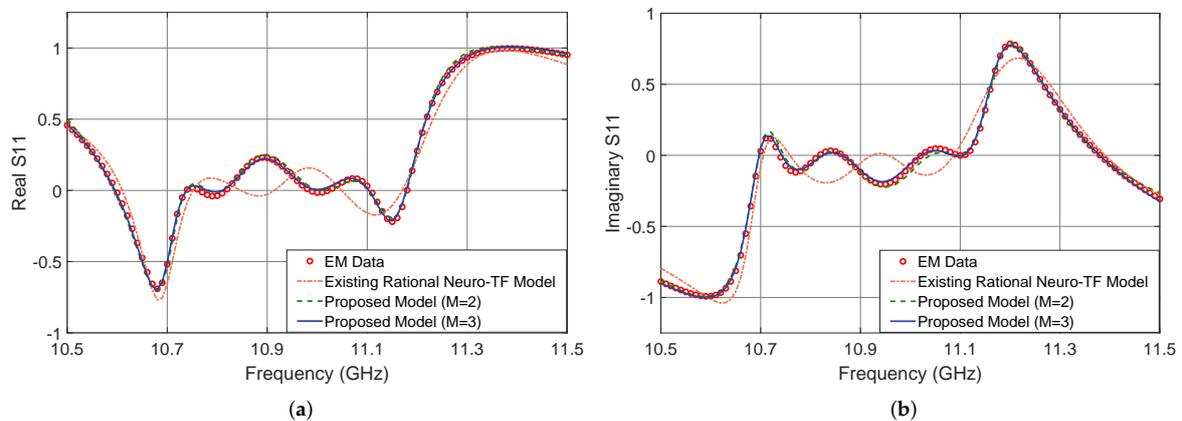
Figure 10. Comparison of $\mathrm{Re}(S_{11})$ and $\mathrm{Im}(S_{11})$ for different modeling approaches and EM data at test
sample $x = [3.36\ 4.38\ 4.00\ 3.36\ 3.02]^T$ (mm) for the four-order bandpass filter example. (**a**) $\mathrm{Re}(S_{11})$ and
(**b**) $\mathrm{Im}(S_{11})$.

It is observed from Figures 9 and 10 that the neuro-TF model using the proposed decomposition
method can achieve better model accuracy for different geometrical samples which are never used
in training.

Table 5. Comparisons of different rational-based neuro-TF modeling approaches for the four-order bandpass filter example.

Modeling Methods		No. of Sub-Models	N_{eff} or N_i	No. of Hidden Neurons		Average Training Error	Average Testing Error
Case 1 (Narrower Range)	Existing Rational Neuro-TF Method	1	6	NN for Numerator	10	4.448 %	4.674 %
				NN for Denominator	10		
		1	6	NN for Numerator	40	2.562 %	9.121 %
				NN for Denominator	40		
	Proposed Rational Neuro-TF Method	2	3 3	NN for Numerator	10	1.487 %	1.672 %
				NN for Denominator	10		
		3	2 2 2	NN for Numerator	10	0.876 %	1.015 %
				NN for Denominator	10		
Case 2 (Wider Range)	Existing Rational Neuro-TF Method	1	6	NN for Numerator	10	6.809 %	8.382 %
				NN for Denominator	10		
		1	6	NN for Numerator	40	4.791 %	32.23%
				NN for Denominator	40		
	Proposed Rational Neuro-TF Method	2	3 3	NN for Numerator	10	2.252 %	4.397 %
				NN for Denominator	10		
		2	3 3	NN for Numerator	40	1.877 %	16.21 %
				NN for Denominator	40		
		3	2 2 2	NN for Numerator	10	1.624 %	1.982 %
				NN for Denominator	10		

5. Conclusions

This paper has proposed a novel decomposition technique for rational-based neuro-TF modeling method. The proposed technique has decomposed the single neuro-TF model with high order into multiple sub-neuro-TF models with the much lower order and reformulated the overall neuro-TF model as the combination of the sub-neuro-TF models. New formulations have been derived for the determination of the number of sub-neuro-TF models and the order of each sub-neuro-TF model. Since the high order of the overall neuro-TF model has been decomposed into multiple lower orders of the sub-neuro-TF models, the proposed technique has decreased the sensitivities of the overall model response with respect to the coefficients of the transfer function. By decreasing the sensitivities using the proposed decomposition technique, the accuracy of the overall neuro-TF model has been improved. In this paper, the proposed decomposition technique has been applied to two microwave filter applications. The potential use of the proposed technique for other EM applications, such as diplexers and antennas, could be an interesting future extension of this work.

References

1. Rayas-Sanchez, J.E. EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art. *IEEE Trans. Microw. Theory Tech.* **2004**, *52*, 420–435. [CrossRef]

2. Rizzoli, V.; Costanzo, A.; Masotti, D.; Lipparini, A.; Mastri, F. Computer-aided optimization of nonlinear microwave circuits with the aid of electromagnetic simulation. *IEEE Trans. Microw. Theory Tech.* **2004**, *52*, 362–377. [CrossRef]

3. Steer, M.B.; Bandler, J.W.; Snowden, C.M. Computer-aided design of RF and microwave circuits and systems. *IEEE Trans. Microw. Theory Tech.* **2002**, *52*, 996–1005. [CrossRef]

4. Burrascano, P.; Fiori, S.; Mongiardo, M. A review of artificial neural networks applications in microwave computer-aided design. *Int. J. RF Microw. Comput. Aided Eng.* **1999**, *9*, 158–174. [CrossRef]

5. Ding, X.; Devabhaktuni, V.K.; Chattaraj, B.; Yagoub, M.C.E.; Deo, M.; Xu, J.; Zhang, Q.J. Neural-network approaches to electromagnetic based modeling of passive components and their applications to high frequency and high-speed nonlinear circuit optimization. *IEEE Trans. Microw. Theory Tech.* **2004**, *52*, 436–449. [CrossRef]

6. Sadrossadat, S.A.; Cao, Y.; Zhang, Q.J. Parametric Modeling of Microwave Passive Components Using Sensitivity-Analysis-Based Adjoint Neural-Network Technique. *IEEE Trans. Microw. Theory Tech.* **2013**, *61*, 1733–1747. [CrossRef]

7. Zhang, Q.J.; Gupta, K.C.; Devabhaktuni, V.K. Artificial neural networks for RF and microwave design-From theory to practice. *IEEE Trans. Microw. Theory Tech.* **2003**, *51*, 1339–1350. [CrossRef]

8. Bandler, J.W.; Ismail, M.A.; Rayas-Sanchez, J.E.; Zhang, Q.J. Neuromodeling of microwave circuits exploiting space-mapping technology. *IEEE Trans. Microw. Theory Tech.* **1999**, *47*, 2417–2427. [CrossRef]

9. Zhang, L.; Xu, J.; Yagoub, M.C.E.; Ding, R.; Zhang, Q.J. Efficient analytical formulation and sensitivity analysis of neuro-space mapping for nonlinear microwave device modeling. *IEEE Trans. Microw. Theory Tech.* **2005**, *53*, 2752–2767. [CrossRef]

10. Gorissen, D.; Zhang, L.; Zhang, Q.J.; Dhaene, T. Evolutionary neuro-space mapping technique for modeling of nonlinear microwave devices. *IEEE Trans. Microw. Theory Tech.* **2011**, *59*, 213–229. [CrossRef]

11. Kabir, H.; Zhang, L.; Yu, M.; Aaen, P.H.; Wood, J.; Zhang, Q.J. Smart modeling of microwave device. *IEEE Microw. Mag.* **2010**, *11*, 105–118. [CrossRef]

12. Liu, W.; Na, W.; Zhu, L.; Ma, J.; Zhang, Q.J. A Wiener-type dynamic neural network approach to the modeling of nonlinear microwave devices. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 2043–2062. [CrossRef]

13. Na, W.; Feng, F.; Zhang, C.; Zhang, Q.J. A unified automated parametric modeling algorithm using knowledge-based neural network and l1 optimization. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 729–745. [CrossRef]

14. Zhang, C.; Jin, J.; Na, W.; Zhang, Q.-J.; Yu, M. Multivalued neural network inverse modeling and applications to microwave filters. *IEEE Trans. Microw. Theory Tech.* **2018**, *66*, 3781–3797. [CrossRef]

15. Jin, J.; Zhang, C.; Feng, F.; Na, W.; Ma, J.; Zhang, Q.-J. Deep neural network technique for high-dimensional microwave modeling and applications to parameter extraction of microwave filters. *IEEE Trans. Microw. Theory Tech.* **2019**, *67*, 4140–4155. [CrossRef]

16. Huang, A.-D.; Zhong, Z.; Wu, W.; Guo, Y.-X. An artificial neural network-based electrothermal model for GaN HEMTs with dynamic trapping effects consideration. *IEEE Trans. Microw. Theory Tech.* **2016**, *64*, 2519–2528. [CrossRef]

17. Zhang, Q.J.; Gupta, K.C. *Neural Networks for RF and Microwave Design*; Artech House: Norwood, MA, USA, 2000.

18. Mkadem, F.; Boumaiza, S. Physically inspired neural network model for RF power amplifier behavioral modeling and digital predistortion. *IEEE Trans. Microw. Theory Tech.* **2011**, *59*, 913–923. [CrossRef]

19. Xu, J.; Horn, J.; Iwamoto, M.; Root, D.E. Large-signal FET model with multiple time scale dynamics from nonlinear vector network analyzer data. In Proceedings of the 2010 IEEE MTT-S International Microwave Symposium, Anaheim, CA, USA, 23–28 May 2010; pp. 417–420.

20. Zhu, L.; Zhang, Q.J.; Liu, K.; Ma, Y.; Peng, B.; Yan, S. A novel dynamic neuro-space mapping approach for nonlinear microwave device modeling. *IEEE Microw. Wirel. Compon. Lett.* **2016**, *26*, 131–133. [CrossRef]

21. Gongal-Reddy, V.-M.-R.; Feng, F.; Zhang, Q.J. Parametric modeling of millimeter-wave passive components using combined neural networks and transfer functions. In Proceedings of the Global Symposium on Millimeter Waves (GSMM), Montreal, QC, Canada, 25–27 May 2015; pp. 1–3.

22. Feng, F.; Zhang, C.; Ma, J.; Zhang, Q.J. Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions. *IEEE Trans. Microw. Theory Tech.* **2016**, *64*, 60–77. [CrossRef]

23. Feng, F.; Gongal-Reddy, V.-M.-R.; Zhang, C.; Ma, J.; Zhang, Q.-J. Parametric modeling of microwave components using adjoint neural networks and pole-residue transfer functions with EM sensitivity analysis. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 1955–1975. [CrossRef]

24. Zhao, Z.; Feng, F.; Zhang, W.; Zhang, J.; Jin, J.; Zhang, Q.J. Parametric Modeling of EM Behavior of Microwave Components Using Combined Neural Networks and Hybrid-Based Transfer Functions. *IEEE Access* **2020**, *8*, 93922–93938. [CrossRef]

25. Cao, Y.; Wang, G.; Zhang, Q.J. A new training approach for parametric modeling of microwave passive components using combined neural networks and transfer functions. *IEEE Trans. Microw. Theory Tech.* **2009**, *57*, 2727–2742.

26. Gustavsen, B.; Semlyen, A. Rational approximation of frequency domain responses by vector fitting. *IEEE Trans. Power Deliv.* **1999**, *14*, 1052–1061. [CrossRef]

27. Zhang, J.; Feng, F.; Zhang, W.; Jin, J.; Ma, J.; Zhang, Q.J. A novel training approach for parametric modeling of microwave passive components using Pade via Lanczos and EM sensitivities. *IEEE Trans. Microw. Theory Tech.* **2020**, *68*, 2215–2233. [CrossRef]

28. Schmidt, S.R.; Launsby, R.G. *Understanding Industrial Designed Experiments*; Air Force Academy: Colorado Springs, CO, USA, 1992.

29. Zhang, C.; Feng, F.; Gongal-Reddy, V.-M.-R.; Zhang, Q.J.; Bandler, J.W. Cognition-driven formulation of space mapping for equal-ripple optimization of microwave filters. *IEEE Trans. Microw. Theory Tech.* **2015**, *63*, 2154–2165. [CrossRef]