

Received March 20, 2020, accepted April 11, 2020, date of publication April 24, 2020, date of current version June 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2990157

Parametric Modeling of EM Behavior of Microwave Components Using Combined Neural Networks and Hybrid-Based Transfer Functions

ZHIHAO ZHAO^{1,2}, (Student Member, IEEE), FENG FENG^{1,2}, (Member, IEEE),

WEI ZHANG^{1,2}, (Student Member, IEEE),

JIANAN ZHANG^{1,2}, (Student Member, IEEE),

JING JIN^{1,2}, (Student Member, IEEE), AND QI-JUN ZHANG^{1,2}, (Fellow, IEEE)

¹School of Microelectronics, Tianjin University, Tianjin 300072, China

²Department of Electronics, Carleton University, Ottawa, ON K1S5B6, Canada

Corresponding author: Feng Feng (fengfeng@doe.carleton.ca)

ABSTRACT Neuro-transfer function (neuro-TF) approaches have become more and more popular in parametric modeling for electromagnetic (EM) behavior of microwave components. Existing pole-residue-based neuro-TF approach has better capability of dealing with high-order problem than the rational-based neuro-TF approach, but has the discontinuity issue and the associated non-smoothness issue of the poles/residues when the geometrical variations become large while the rational-based neuro-TF approach does not have. This paper addresses this situation and presents a novel hybrid-based neuro-TF technique which systematically combines both pole-residue and rational formats of the transfer functions. Starting with the pole-residue-based transfer functions, we propose a novel technique to automatically identify the poles/residues that are smooth-continuous and the poles/residues that have the discontinuity and non-smoothness issues. The proposed technique converts the poles/residues that have those issues into the coefficients of the rational-based transfer function to solve the discontinuity and non-smoothness issues in the existing pole-residue-based neuro-TF approach. The proposed technique remains the smooth-continuous poles/residues in the pole-residue format of the transfer function to maintain the capability of handling high-order problem. Compared with the existing neuro-TF modeling methods, the proposed technique can obtain better accuracy in challenging applications of large geometrical variations and high order. The proposed technique is illustrated by two examples of parametric modeling of microwave components.

INDEX TERMS Parametric modeling, microwave components, neural networks, hybrid-based transfer function, parameter extraction.

I. INTRODUCTION

Artificial neural network (ANN) has been recognized as a powerful tool for parametric modeling of electromagnetic (EM) behaviors in microwave area [1]–[5]. EM design optimization can be expensive because it requires repetitive EM simulations due to adjustments of the values of geometrical parameters. Through a systematic computer-based training process, ANN can be used to learn the relationship between EM behaviors and geometrical parameters. After training, the developed parametric model can be used to

The associate editor coordinating the review of this manuscript and approving it for publication was Chan Hwang See¹.

provide accurate and fast prediction of the EM behavior of microwave components, and can be subsequently implemented in high-level circuit and system design [6].

Another parametric modeling approach is the knowledge-based approach. The knowledge-based neural network (KBNN) is a recognized knowledge-based approach for parametric modeling. In the KBNN approach, KBNN models are developed using neural networks combined with knowledge models such as analytical formulas [7], empirical functions [8], or equivalent circuit models [9], [10]. The knowledge models can help speed up model development and enhance the capability for learning and generalization of the KBNN models [11]. Space mapping [12]–[19] is also

a recognized knowledge-based approach for parametric modeling. Research endeavors on space mapping have focused on several areas, such as implicit space mapping [13], output space mapping [14], [15], neural space mapping [16], [17], etc.

Another parametric modeling approach is the neuro-transfer function (neuro-TF) approach [20], [21], which combines neural networks and transfer functions. In this approach, EM responses of passive components versus frequency are represented by the transfer functions. This approach can be employed even if accurate knowledge models are not available. The pole-residue-based neuro-TF approach [22] is one of the commonly used neuro-TF approach. This approach can work well when the geometrical variations are small. However, as the geometrical variations become large, some poles/residues are no longer complex w.r.t. geometrical parameters, resulting in the discontinuity issue. Along with the discontinuity issue in the poles/residues, the smoothness of the neighboring poles/residues w.r.t. the geometrical parameters are also affected, resulting in the non-smoothness issue. These issues cause the difficulty in training for the pole-residue-based neuro-TF model. For this reason, the effectiveness of the approach is compromised when geometrical variations become large. Another neuro-TF approach is the rational-based neuro-TF approach [23], which does not have the discontinuity issue existed in the poles/residues. However, as the order of the transfer function increases and/or the geometrical variations become large, the transfer function is more sensitive to the coefficients, resulting in the difficulty in training for the rational-based neuro-TF model. The sensitivities of the transfer function responses w.r.t. the coefficients and the poles/residues are illustrated in [22]. For this reason, the order of the transfer function is often limited to a small number, compromising effectiveness of the rational-based neuro-TF approach to the high order applications.

In this paper, a novel parametric modeling technique is proposed to develop a combined neural network and hybrid-based transfer function (hybrid-based neuro-TF) model of EM behavior of microwave components. The proposed technique systematically combines both pole-residue and rational formats of the transfer functions. Starting with the pole-residue-based transfer functions, the proposed technique automatically identify the poles/residues that are smooth-continuous and the poles/residues that have the discontinuity and non-smoothness issues. The proposed technique converts the poles/residues that have those issues into the coefficients of the rational-based transfer function to effectively address the discontinuity and non-smoothness issues in the existing pole-residue-based neuro-TF approach. The proposed technique remains the smooth-continuous poles/residues in the pole-residue format of the transfer function to maintain the capability of handling high-order problem. Compared with the existing rational-based neuro-TF modeling method and the existing pole-residue-based neuro-TF modeling method, the proposed technique can obtain better

accuracy in challenging applications of large geometrical variations and high order, addressing the discontinuity and non-smoothness issues. After the proposed modeling process is finished, the developed model can provide accurate and fast prediction of the EM behavior of microwave components and can be subsequently used in high-level circuit and system design.

The organization of this paper is as follows. In Section II, we described the discontinuity issue and the non-smoothness issue in the poles/residues w.r.t. geometrical parameters. In Section III, we described the proposed technique for developing a hybrid-based neuro-TF model in details. In Section IV, we demonstrated the parametric modeling process by two EM examples, confirming the effectiveness of our proposed hybrid-based neuro-TF modeling technique.

II. THE ISSUES OF DISCONTINUITY AND ACCOMPANIED NON-SMOOTHNESS IN POLES AND RESIDUES W.R.T. GEOMETRICAL PARAMETERS

In the existing pole-residue-based neuro-TF modeling approach [22], a parameter extraction process is required to extract the poles/residues in the transfer function for each geometrical parameter sample. All the extracted poles/residues are required to be changed continuously in terms of the types of their values as the geometrical parameters change. The method can work well when the geometrical variations are small. It is because that all the poles/residues vary within complex values, which are continuous as the geometrical parameters change. However, as the geometrical variations become large, not all the poles/residues vary within complex values as the geometrical parameters change. In this case, the discontinuity issue occurs in the poles/residues and becomes a major issue for parametric modeling. More specifically, some poles/residues vary within complex values, while other poles/residues vary between real values and complex values, as the geometrical parameters change. The discontinuity issue is also the cause for the high nonlinearity of pole/residues from sample to sample, resulting in the difficulty of the training for the pole-residue-based neuro-TF model. For this reason, the effectiveness of the method is compromised when the discontinuity issue exists in the poles/residues as the geometrical variations become large. Along with the discontinuity issue in the poles/residues, the smoothness of the neighboring poles/residues w.r.t. the geometrical parameters are also affected. Although some poles/residues do not have the discontinuity issue as other poles/residues, some of these continuous poles/residues are not smooth w.r.t. the geometrical parameters, leading to the non-smoothness issue. This issue is also the cause for the high nonlinearity of pole/residues from sample to sample, resulting in the difficulty of the training for the pole-residue-based neuro-TF model. This is another reason why the effectiveness of the method is compromised when the discontinuity issue exists in the poles/residues as the geometrical variations become large.

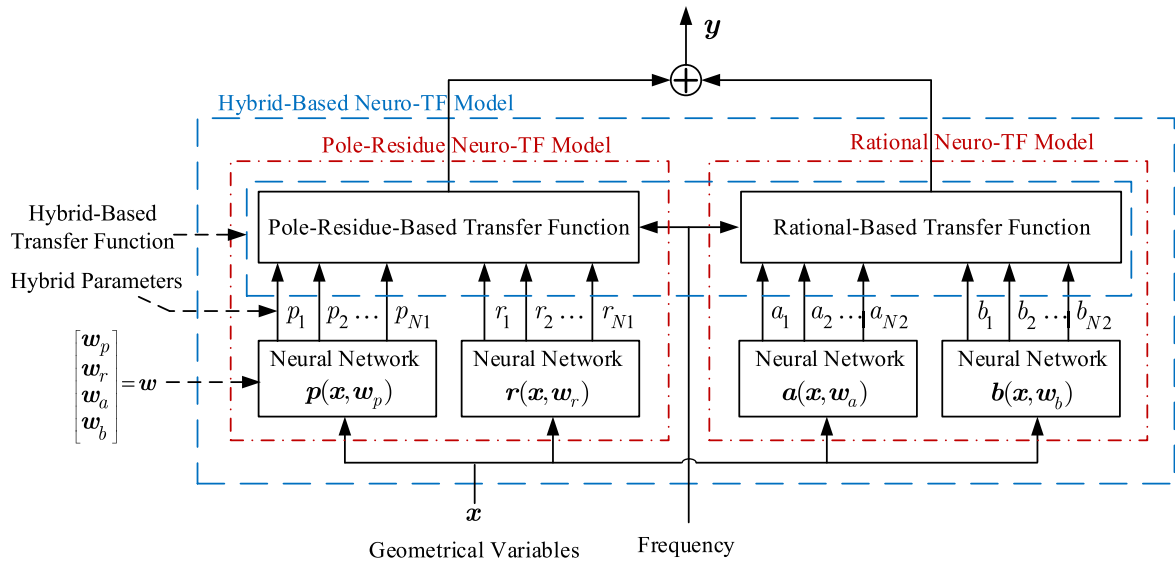


FIGURE 1. The structure of the proposed hybrid-based neuro-TF model.

Rational-based neuro-TF modeling approach is one of the methods that does not have the discontinuity issue existed in the pole-residue-based neuro-TF approach. In the existing rational-based neuro-TF modeling approach [23], the poles/residues are reformatted into the coefficients of the numerators and the denominators of the rational-based transfer function. These coefficients always vary in real values as the geometrical parameters change. In other words, the coefficients are continuous and does not have the discontinuity issue. However, as the order of the transfer function increases, the transfer function is more sensitive to the coefficients, resulting in the difficulty of the training for the rational-based neuro-TF model. For this reason, the effectiveness of the method is compromised when the order of the transfer function is high.

The focus of this paper is on developing a new and effective parametric modeling technique, addressing the discontinuity issue and the accompanied non-smoothness issue. In the subsequent section, we introduce a novel hybrid-based neuro-TF modeling technique which systematically combines both pole-residue and rational formats of the transfer functions in order to overcome those issues in the poles/residues. The main idea is that we convert the poles/residues that have the issues into the coefficients of the rational-based transfer function to effectively address the discontinuity and non-smoothness issues and we remain the smooth-continuous poles/residues in the pole-residue format of the transfer function to maintain the capability of handling high-order problem. In this way, the discontinuity and non-smoothness issues can be solved by the proposed hybrid-based neuro-TF modeling technique and high-order modeling capability can be maintained by the overall model.

III. PROPOSED HYBRID-BASED NEURO-TF METHOD FOR PARAMETRIC MODEL DEVELOPMENT

A. FORMULATION OF THE HYBRID-BASED NEURO-TF MODEL

The structure of the proposed hybrid-based neuro-TF model is illustrated in Fig. 1. Let x be the model input, representing the geometrical variables. Let y be the model output, representing the EM responses (e.g., S -parameters). The proposed model consists of the pole-residue-based transfer function, the rational-based transfer function, and four neural networks. The structure of the proposed model is a little bit more complicated than the existing pole-residue-based neuro-TF or the existing rational-based neuro-TF. The outputs of the model are the EM behaviors e.g., S -parameters, of the microwave components. The inputs of the model are the geometrical parameters of the microwave components and frequency. As the values of geometrical parameters change, some of the pole/residues may vary between real and complex values, resulting in the discontinuity issue. Along with the discontinuity issue in the poles/residues, the smoothness of the neighboring poles/residues w.r.t. the geometrical parameters are also affected, leading to the non-smoothness issue. Here, we define the continuous pole/residue to be the one whose value is always complex or real as the values of the geometrical parameters change. We define the discontinuous pole/residue to be the one whose value varies between real and complex as the value of the geometrical parameters change. Among the continuous poles/residues, there may exist one or more poles/residues whose values do not change smoothly as the values of the geometrical parameters change. The non-smoothness issue happens in these continuous poles/residues whose smoothness w.r.t. the geometrical parameters does not satisfy the smoothness requirement.

Let these poles/residues be called the non-smooth-continuous poles/residues. The rest of the continuous poles/residues satisfy the smoothness requirement. Let the remaining smooth and continuous poles/residues be called the smooth-continuous poles/residues.

We propose to maintain the smooth-continuous poles/residues to formulate the pole-residue-based transfer function, while to re-formulate the discontinuous poles/residues and the non-smooth-continuous poles/residues together into rational-based transfer function. The pole-residue-based transfer function and the rational-format-based transfer function are combined together to formulate the proposed hybrid-based transfer function. By such systematical combination of the two transfer functions, the discontinuity and non-smoothness issues can be effectively addressed by the rational-based transfer function with a relatively low order and the capability of handling high-order problem can be maintained by the pole-residue-based transfer function without the issues. In this way, the discontinuity and non-smoothness issues can be solved by the proposed hybrid-based neuro-TF modeling technique and high-order modeling capability can be maintained by the overall model. The parameters in the proposed hybrid-based transfer function are defined as the hybrid parameters, containing the poles/residues of the pole-residue-based transfer function and coefficients in the numerator/denominator of the rational-based transfer function. Since the relationship between the hybrid parameters and the geometrical parameters is nonlinear and unknown, we propose to use neural networks to learn and represent this nonlinear relationship.

Let N_1 and N_2 represent the orders of the pole-residue-based transfer function and the rational-based transfer function, respectively. Let N represent the total order of the hybrid-based transfer function, which equals the sum of the N_1 and N_2 , i.e., $N = N_1 + N_2$. In the propose model, four different neural networks are used to learn and represent the nonlinear and unknown relationship between the hybrid parameters and the geometrical parameters. More specifically, these neural networks are used to learn and represent the nonlinear and unknown relationship between the poles/residues of the pole-residue-based transfer function and the geometrical parameters, and the nonlinear and unknown relationship between the coefficients of the rational-based transfer function and the geometrical parameters. Let p_i represent the i th output of the neural network for poles corresponding to geometrical variables and neural network weights \mathbf{w}_p . Let r_i represent the i th output of the neural network for residues corresponding to geometrical variables and neural network weights \mathbf{w}_r . Let a_i represent the i th output of the neural network for the coefficients in the numerator of the rational-based transfer function corresponding to geometrical variables and neural network weights \mathbf{w}_a . Let b_i represent the i th output of the neural network for the coefficients in the denominator of the rational-based transfer function corresponding to geometrical variables and neural network weights \mathbf{w}_b . Let \mathbf{w} represent a vector containing all the weights in the proposed

hybrid-based neuro-TF model, defined as

$$\mathbf{w} = \left[\mathbf{w}_p^T \ \mathbf{w}_r^T \ \mathbf{w}_a^T \ \mathbf{w}_b^T \right]^T. \quad (1)$$

Let \mathbf{x} be a vector containing the geometrical variables, representing the inputs of the proposed hybrid-based neuro-TF model. Let y represent the output frequency response, e.g., S -parameter of the proposed model. For convenience of representation, we define \mathbf{p} and \mathbf{r} to be vectors containing the smooth-continuous poles and residues, respectively, i.e., $\mathbf{p} = [p_1 \ p_2 \ \dots \ p_{N_1}]^T$ and $\mathbf{r} = [r_1 \ r_2 \ \dots \ r_{N_1}]^T$. We define \mathbf{a} and \mathbf{b} to be vectors containing coefficients in the numerator and denominator of the rational-based transfer function, respectively, i.e., $\mathbf{a} = [a_1 \ a_2 \ \dots \ a_{N_2}]^T$ and $\mathbf{b} = [b_1 \ b_2 \ \dots \ b_{N_2}]^T$. As is known to all, the pole-residue-based transfer function and the rational-based transfer function are interconvertible to represent the frequency response in general. Instead of using these two classical transfer functions, we propose to convert some of the pole-residue terms from the original pole-residue-based transfer function into the rational-based transfer function in order to formulate the hybrid-based neuro-TF model. The output y of the hybrid-based neuro-TF model can be formulated as

$$y(\mathbf{x}, \mathbf{w}, s) = \sum_{i=1}^{N_1} \frac{r_i(\mathbf{x}, \mathbf{w}_r)}{s - p_i(\mathbf{x}, \mathbf{w}_p)} + \frac{\sum_{i=1}^{N_2} a_i(\mathbf{x}, \mathbf{w}_a) s^{i-1}}{1 + \sum_{i=1}^{N_2} b_i(\mathbf{x}, \mathbf{w}_b) s^i}, \quad (2)$$

where s represents the frequency in Laplace domain.

B. PROPOSED DERIVATIVE CALCULATION FOR HYBRID-BASED NEURO-TF MODEL DEVELOPMENT

In this subsection, we derive the derivative formulation for training of the proposed hybrid-based neuro-TF model. Let $\partial y / \partial p_i$ and $\partial y / \partial r_i$ represent the derivative of the output y w.r.t. the i th pole p_i and the i th residue r_i of the pole-residue-based transfer function, respectively. The derivatives $\partial y / \partial p_i$ and $\partial y / \partial r_i$ are formulated as

$$\frac{\partial y}{\partial p_i}(\mathbf{x}, \mathbf{w}, s) = \frac{r_i(\mathbf{x}, \mathbf{w}_r)}{(s - p_i(\mathbf{x}, \mathbf{w}_p))^2} \quad (3)$$

and

$$\frac{\partial y}{\partial r_i}(\mathbf{x}, \mathbf{w}, s) = \frac{1}{s - p_i(\mathbf{x}, \mathbf{w}_p)}. \quad (4)$$

The derivative vectors of the output y w.r.t. \mathbf{p} and \mathbf{r} are defined as $\partial y / \partial \mathbf{p}$ and $\partial y / \partial \mathbf{r}$ which contain the derivatives $\partial y / \partial p_i$ and $\partial y / \partial r_i$, respectively, defined as

$$\frac{\partial y}{\partial \mathbf{p}} = \left[\frac{\partial y}{\partial p_1} \ \frac{\partial y}{\partial p_2} \ \dots \ \frac{\partial y}{\partial p_{N_1}} \right]^T \quad (5)$$

and

$$\frac{\partial y}{\partial \mathbf{r}} = \left[\frac{\partial y}{\partial r_1} \ \frac{\partial y}{\partial r_2} \ \dots \ \frac{\partial y}{\partial r_{N_1}} \right]^T \quad (6)$$

Let α and β represent the values of the numerator and the denominator of the rational-based transfer function, respectively, expressed as

$$\alpha(\mathbf{x}, \mathbf{w}_a, s) = \sum_{i=1}^{N_2} a_i(\mathbf{x}, \mathbf{w}_a) s^{i-1} \quad (7)$$

and

$$\beta(\mathbf{x}, \mathbf{w}_b, s) = 1 + \sum_{i=1}^{N_2} b_i(\mathbf{x}, \mathbf{w}_b) s^i. \quad (8)$$

The derivatives of the output y w.r.t. the i th coefficient a_i in the numerator and the i th coefficient b_i in the denominator of the rational-based transfer function are defined as $\partial y/\partial a_i$ and $\partial y/\partial b_i$, respectively, formulated as

$$\frac{\partial y}{\partial a_i}(\mathbf{x}, \mathbf{w}, s) = \frac{1}{\beta(\mathbf{x}, \mathbf{w}_b, s)} s^{i-1} \quad (9)$$

and

$$\frac{\partial y}{\partial b_i}(\mathbf{x}, \mathbf{w}, s) = -\frac{\alpha(\mathbf{x}, \mathbf{w}_a, s)}{\beta^2(\mathbf{x}, \mathbf{w}_b, s)} s^i. \quad (10)$$

The derivative vectors of output y w.r.t. \mathbf{a} and \mathbf{b} are defined as $\partial y/\partial \mathbf{a}$ and $\partial y/\partial \mathbf{b}$ which contain the derivatives $\partial y/\partial a_i$ and $\partial y/\partial b_i$, respectively, defined as

$$\frac{\partial y}{\partial \mathbf{a}} = \left[\frac{\partial y}{\partial a_1} \quad \frac{\partial y}{\partial a_2} \quad \dots \quad \frac{\partial y}{\partial a_{N_2}} \right]^T \quad (11)$$

and

$$\frac{\partial y}{\partial \mathbf{b}} = \left[\frac{\partial y}{\partial b_1} \quad \frac{\partial y}{\partial b_2} \quad \dots \quad \frac{\partial y}{\partial b_{N_2}} \right]^T. \quad (12)$$

Let \mathbf{s} represent a vector containing different orders of frequency s , defined as,

$$\mathbf{s} = \left[s \quad s^2 \quad s^3 \quad \dots \quad s^{N_2} \right]^T \quad (13)$$

The derivative vectors $\partial y/\partial \mathbf{a}$ and $\partial y/\partial \mathbf{b}$ can be formulated as

$$\frac{\partial y}{\partial \mathbf{a}} = \frac{1}{s\beta} \mathbf{s} \quad (14)$$

and

$$\frac{\partial y}{\partial \mathbf{b}} = -\frac{\alpha}{\beta^2} \mathbf{s}. \quad (15)$$

The derivatives of neural network outputs \mathbf{p} , \mathbf{r} , \mathbf{a} , and \mathbf{b} w.r.t. neural network weighting parameters \mathbf{w} are defined as $\tilde{\mathbf{P}}$, $\tilde{\mathbf{R}}$, $\tilde{\mathbf{A}}$, and $\tilde{\mathbf{B}}$, respectively. The derivatives $\tilde{\mathbf{P}}$, $\tilde{\mathbf{R}}$, $\tilde{\mathbf{A}}$, and $\tilde{\mathbf{B}}$ can be calculated using the algorithms in [5], where $\tilde{\mathbf{P}}(\mathbf{x}, \mathbf{w}) = \partial \mathbf{p}/\partial \mathbf{w}$, $\tilde{\mathbf{R}}(\mathbf{x}, \mathbf{w}) = \partial \mathbf{r}/\partial \mathbf{w}$, $\tilde{\mathbf{A}}(\mathbf{x}, \mathbf{w}) = \partial \mathbf{a}/\partial \mathbf{w}$, and $\tilde{\mathbf{B}}(\mathbf{x}, \mathbf{w}) = \partial \mathbf{b}/\partial \mathbf{w}$. Let $\partial y/\partial \mathbf{w}$ be defined as the derivatives of model output y w.r.t. weighting parameters \mathbf{w} . Substituting (3)–(15), the derivatives $\partial y/\partial \mathbf{w}$ are formulated as

$$\frac{\partial y}{\partial \mathbf{w}} = \tilde{\mathbf{P}}^T \frac{\partial y}{\partial \mathbf{p}} + \tilde{\mathbf{R}}^T \frac{\partial y}{\partial \mathbf{r}} + \tilde{\mathbf{A}}^T \frac{\partial y}{\partial \mathbf{a}} + \tilde{\mathbf{B}}^T \frac{\partial y}{\partial \mathbf{b}}. \quad (16)$$

The derivatives formulated in this subsection are used for training of the proposed hybrid-based neuro-TF model. In order to accurately train the proposed model, the initial values for the neural network weights are very important. To obtain good initial values for the weights, preliminary training of the neural networks is necessary before training the overall model [22]. To perform the preliminary training of the neural networks, the data for the hybrid parameters need to be extracted from the EM behaviors (e.g., S-parameters). In the next subsection, we describe the proposed hybrid parameter extraction technique.

C. PROPOSED HYBRID PARAMETER EXTRACTION TECHNIQUE

The proposed hybrid parameter extraction technique begins with samples of EM data d_k , e.g., S-parameters, for different values of geometrical parameters \mathbf{x}_k , where the subscript k represents the index indicating the k th sample of geometrical parameters, i.e., $k \in T_r = \{1, 2, \dots, n_s\}$, where T_r is the index set of training samples of geometrical parameters and n_s is the total number of training samples. Training samples are generated w.r.t. the geometrical parameters while frequency is a separate variable swept by the EM simulator during data generation. In our method, frequency is an additional input of the hybrid-based transfer function.

Vector fitting process [24] is performed to obtain a group of poles and residues for each geometrical sample. In the vector fitting process, the given information is EM data d versus frequency for a certain geometrical sample. Expected solutions are poles and residues of the transfer function. Let \mathbf{c}_k represent a vector containing the poles $\mathbf{p}^{(k)}$ and residues $\mathbf{r}^{(k)}$ of the transfer function for the k th geometrical sample obtained after vector fitting, defined as

$$\mathbf{c}_k = \begin{bmatrix} \mathbf{p}^{(k)} \\ \mathbf{r}^{(k)} \end{bmatrix} = \begin{bmatrix} [p_1^{(k)} \quad p_2^{(k)} \quad \dots \quad p_N^{(k)}]^T \\ [r_1^{(k)} \quad r_2^{(k)} \quad \dots \quad r_N^{(k)}]^T \end{bmatrix}. \quad (17)$$

We propose a hybrid parameter extraction technique to identify the discontinuous poles/residues, the smooth-continuous poles/residues, and the non-smooth-continuous poles/residues. After the identification, the smooth-continuous poles/residues will be maintained into the pole-residue-based transfer function, while the discontinuous poles/residues and the non-smooth-continuous poles/residues together will be re-formulated into rational-based transfer function.

Suppose that we have already obtained the pole and residue data (i.e., $\mathbf{p}^{(k)}$ and $\mathbf{r}^{(k)}$) for all the geometrical samples after vector fitting. Since each pole $p_i^{(k)}$ usually has a related residue $r_i^{(k)}$ after vector fitting, we suppose that the residue data $\mathbf{r}^{(k)}$ are listed in the same way as their corresponding pole data $\mathbf{p}^{(k)}$. The proposed hybrid parameter extraction technique consists of three stages.

The first stage of the proposed hybrid parameter extraction technique is to identify the discontinuous poles/residues and to identify the continuous poles/residues for further

identification in the second stage. Let N be the total order of the hybrid-based transfer function. Let n_s be the total number of the geometrical samples. Usually, the discontinuous residues appear together with their corresponding discontinuous poles. Similarly, the continuous residues appear together with their corresponding continuous poles. For simplicity of representation, we use n_i^R to represent the number of the geometrical samples where the i th pole $p_i^{(k)}$ (or the i th residue $r_i^{(k)}$) is a real pole (or a real residue), i.e., $i \in I = \{1, 2, \dots, N\}$, where I is the index set of samples of pole data $\mathbf{p}^{(k)}$ (or residue data $\mathbf{r}^{(k)}$). Let η_i represent the ratio between n_i^R and n_s for $p_i^{(k)}$ (or $r_i^{(k)}$), where $n_i^R \leq n_s$, defined as

$$\eta_i = \frac{n_i^R}{n_s}. \quad (18)$$

Here we define a step function $f(\cdot)$ for η_i , expressed as

$$f(\eta_i) = \begin{cases} 1, & \text{if } 0 < \eta_i < 1 \\ 0, & \text{otherwise.} \end{cases} \quad (19)$$

If $0 < \eta_i < 1$, it indicates that $p_i^{(k)}$ (or $r_i^{(k)}$) is a real pole (or residue) for some geometrical samples and is a complex pole (or residue) for other geometrical samples. In this case, $p_i^{(k)}$ (or $r_i^{(k)}$) is a discontinuous pole (or residue) and $f(\eta_i) = 1$. Otherwise, $p_i^{(k)}$ (or $r_i^{(k)}$) is either a real pole (or residue) or a complex pole (or residue) for all the geometrical samples. In this case, $p_i^{(k)}$ (or $r_i^{(k)}$) is a continuous pole (or residue) and $f(\eta_i) = 0$. Let N_2^a be the number of discontinuous poles (or residues), formulated as

$$N_2^a = \sum_{i=1}^N f(\eta_i). \quad (20)$$

The index set I_2^a of the samples of discontinuous pole data (or discontinuous residue data) are formulated as

$$I_2^a = \{i \mid f(\eta_i) = 1, i \in I\}. \quad (21)$$

Subsequently, the index set I' of the samples of continuous pole data (or continuous residue data) can be expressed as $I' = I - I_2^a$. Both index sets I_2^a and I' are the subsets of the index set I . So far, we have identified the discontinuous pole/residue data and the continuous pole/residue data through the calculation of N_2^a , I_2^a , and I' .

In the second stage, we identify the non-smooth continuous poles/residues from the identified continuous poles/residues with the index set I' . Here we use σ_i^p and σ_i^r to represent the deviations of the i th continuous poles \mathbf{p}_i and the i th continuous residues \mathbf{r}_i , respectively, covering all the geometrical samples, where $i \in I'$. The deviations σ_i^p and σ_i^r are expressed as

$$\sigma_i^p = \max_{k \in T_r} \{|p_i^{(k)} - \mu_i^p|\} \quad (22)$$

and

$$\sigma_i^r = \max_{k \in T_r} \{|r_i^{(k)} - \mu_i^r|\} \quad (23)$$

where μ_i^p and μ_i^r are expressed as

$$\mu_i^p = \frac{1}{n_s} \sum_{k=1}^{n_s} p_i^{(k)} \quad (24)$$

and

$$\mu_i^r = \frac{1}{n_s} \sum_{k=1}^{n_s} r_i^{(k)}. \quad (25)$$

Here we define a shared deviation σ_i for both i th continuous poles and residues \mathbf{p}_i and \mathbf{r}_i in order to perform the identification, formulated as

$$\sigma_i = \max\{\sigma_i^p, \sigma_i^r\} \quad (26)$$

where $i \in I'$. Let σ_0 be the user-defined threshold of the deviation. We define a step function $g(\cdot)$ for σ_i , expressed as

$$g(\sigma_i) = \begin{cases} 1, & \text{if } \sigma_i > \sigma_0 \\ 0, & \text{otherwise.} \end{cases} \quad (27)$$

If $\sigma_i > \sigma_0$, the i th continuous poles and residues \mathbf{p}_i and \mathbf{r}_i are considered non-smooth and $g(\sigma_i) = 1$. Otherwise, the i th continuous poles and residues \mathbf{p}_i and \mathbf{r}_i are considered smooth and $g(\sigma_i) = 0$. Let N_2^b represent the number of the non-smooth-continuous poles/residues, formulated as

$$N_2^b = \sum_{i \in I'} g(\sigma_i). \quad (28)$$

It is noticed that the order N_2 of rational-based transfer function cannot be high, since the transfer function becomes more sensitive to its coefficients if the order is higher. For this reason, N_2 should be less than a certain user-defined order N_2^{\max} (e.g., N_2^{\max} equals eight). Suppose that the number of discontinuous poles (or residues) N_2^a is less than N_2^{\max} . Therefore, N_2^b should satisfy the condition

$$N_2^b \leq N_2^{\max} - N_2^a. \quad (29)$$

If this condition is not satisfied, then we increase the value of the user-defined threshold of the deviation σ_0 , recalculate N_2^b and recheck the condition with the updated N_2^b . If the condition is satisfied, we obtain the index set I_2^b of the samples of non-smooth-continuous pole data and non-smooth-continuous residue data, formulated as

$$I_2^b = \{i \mid g(\sigma_i) = 1, i \in I'\}. \quad (30)$$

Subsequently, the index set I_1 of the samples of smooth-continuous pole data and smooth-continuous residue data can be expressed as $I_1 = I' - I_2^b$. Both index sets I_2^b and I_1 are the subsets of the index set I . So far, we have identified the smooth-continuous pole/residue data and the non-smooth-continuous pole/residue data through the calculation of N_2^b , I_2^b , and I_1 .

The order N_2 is subsequently calculated as

$$N_2 = N_2^a + N_2^b. \quad (31)$$

The order N_1 is subsequently calculated as

$$N_1 = N - N_2. \quad (32)$$

The total index set I_2 of the samples of the discontinuous pole/residue data and the non-smooth-continuous pole/residue data is expressed as $I_2 = I_2^a \cup I_2^b$.

In the third stage, we gather the discontinuous poles/residues and the non-smooth-continuous poles/residues with the index set I_2 from the original pole/residue data into the new pole/residue data. These new pole/residue data are reformatted into the coefficient data for the rational-format-based transfer function. After reformatting, the coefficient data are defined as $\hat{a}_i^{(k)}$ and $\hat{b}_i^{(k)}$ for the k th geometrical sample, where $i = 1, 2, \dots, N_2$. We maintain the smooth-continuous pole/residue data for the pole-residue-based transfer function. These smooth-continuous pole/residue data are redefined as $\hat{p}_i^{(k)}$ and $\hat{r}_i^{(k)}$ for the k th geometrical sample, where $i = 1, 2, \dots, N_1$.

Let \hat{c}_k represent a vector containing the hybrid parameters of the hybrid-based transfer function for the k th geometrical sample, defined as

$$\hat{c}_k = \left[\hat{p}^{(k)T} \hat{r}^{(k)T} \hat{a}^{(k)T} \hat{b}^{(k)T} \right]^T \quad (33)$$

where $\hat{p}^{(k)}$ is a vector of smooth-continuous poles for the k th geometrical sample, defined as $\hat{p}^{(k)} = [\hat{p}_1^{(k)} \hat{p}_2^{(k)} \dots \hat{p}_{N_1}^{(k)}]^T$; $\hat{r}^{(k)}$ is a vector of smooth-continuous residues for the k th geometrical sample, defined as $\hat{r}^{(k)} = [\hat{r}_1^{(k)} \hat{r}_2^{(k)} \dots \hat{r}_{N_1}^{(k)}]^T$; $\hat{a}^{(k)}$ is a vector of coefficients in the numerator of the transfer function for the k th geometrical sample, defined as $\hat{a}^{(k)} = [\hat{a}_1^{(k)} \hat{a}_2^{(k)} \dots \hat{a}_{N_2}^{(k)}]^T$; $\hat{b}^{(k)}$ is a vector of coefficients in the denominator of the transfer function for the k th geometrical sample, defined as $\hat{b}^{(k)} = [\hat{b}_1^{(k)} \hat{b}_2^{(k)} \dots \hat{b}_{N_2}^{(k)}]^T$. The stepwise algorithm of the proposed three-stage hybrid parameter extraction technique is illustrated by a flow diagram as shown in Fig. 2.

D. PRELIMINARY TRAINING OF NEURAL NETWORKS

We propose a three-stage training process. In the first stage, we perform a preliminary training process for characterizing the relationships between the poles/residues of the transfer function and the geometrical parameters. During this process, neural networks are trained to learn these relationships. The training data for this phase is $(x_k, \hat{p}^{(k)})$ and $(x_k, \hat{r}^{(k)})$, $k \in T_r$ (i.e., geometrical parameters as model inputs and poles/residues as model outputs). The nonlinear relationship between poles and geometrical parameters are usually different from that between residues and geometrical parameters. Subsequently, poles and residues are provided by separate neural networks.

In the second stage, we perform a preliminary training process for characterizing the relationships between the coefficients of the transfer function and the geometrical parameters. During this process, different sets of neural networks are trained to learn these relationships. The training data for

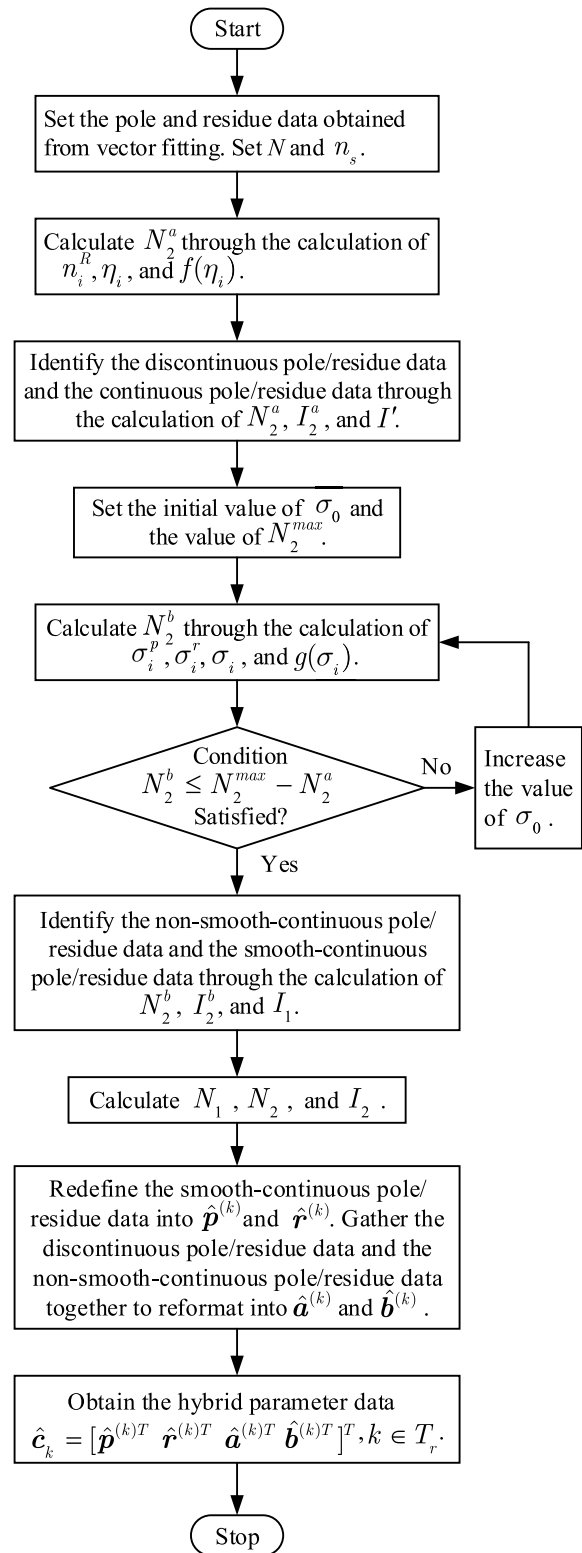


FIGURE 2. The flow diagram of the proposed three-stage hybrid parameter extraction process.

this phase is $(x_k, \hat{a}^{(k)})$ and $(x_k, \hat{b}^{(k)})$, $k \in T_r$ (i.e., geometrical parameters as model inputs and coefficients as model outputs). The nonlinear relationship between the coefficients

in the numerator and geometrical parameters are usually different from that between the coefficients in the denominator and geometrical parameters. Similarly, separate neural networks are used to train the coefficients in the numerator and denominator.

During the two stages of preliminary training, a certain amount of training error (e.g., 5%-10%) can be tolerated. With this relaxed error criteria, less hidden neurons can be used in the neural networks. The neural networks thus have lower nonlinearity which makes the hybrid-based neuro-TF more robust. After preliminary training of the neural networks, an overall model refinement process is performed to further refine the preliminary-trained model.

E. REFINEMENT TRAINING OF THE HYBRID-BASED NEURO-TF MODEL

In the third stage of the training process, a model refinement is performed to further refine the overall hybrid-based neuro-TF model. The training data for this phase is (\mathbf{x}_k, d_k) , $k \in T_r$ (i.e., geometrical parameters as model inputs and EM responses (i.e., S-parameters) as model outputs). The mechanism for the refinement training process of the overall hybrid-based neuro-TF model is shown in Fig. 4. It consists of the hybrid-based transfer function of (2) and the neural networks whose initial values are the optimal solutions from the two-stage preliminary training. This model refinement process consists of both training and testing of the model. The objective of training is to minimize the training error of the overall model. Training is performed by optimizing the weights inside the neural networks to minimize the error function

$$E_{Tr}(\mathbf{w}_p, \mathbf{w}_r, \mathbf{w}_a, \mathbf{w}_b) = \frac{1}{2n_s} \sum_{k \in T_r} \sum_{i \in \Omega} \|y(\mathbf{w}_p, \mathbf{w}_r, \mathbf{w}_a, \mathbf{w}_b, \mathbf{x}_k, s_i) - d_{k,i}\|^2 \quad (34)$$

where T_r is the index set of training samples of various geometrical parameters, and n_s is the total number of training samples. Ω is the index set of frequency samples. \mathbf{w}_p , \mathbf{w}_r , \mathbf{w}_a and \mathbf{w}_b represent the weights in the neural networks for poles \mathbf{p} , residues \mathbf{r} , coefficients in numerator \mathbf{a} , and coefficients in denominator \mathbf{b} , respectively. y represents the outputs of the overall model, which ultimately is a function of geometrical variables \mathbf{x}_k , frequency s_i , and neural network weights \mathbf{w}_p , \mathbf{w}_r , \mathbf{w}_a and \mathbf{w}_b .

The training process terminates when the training error becomes lower than a user-defined threshold E_t . After the training process, an independent set of testing data which are never used in training is used for testing the quality of the trained hybrid-based neuro-TF model. The testing error E_{Ts} is defined as the error between the model response and the testing data. If the testing error is also lower than the threshold error E_t , the model refinement process terminates and the hybrid-based neuro-TF model is ready to be used for high-level design. Otherwise, the overall model training process will be repeated with different numbers of hidden neurons.

A flow diagram illustrating the overall hybrid-based neuro-TF model development process is shown in Fig. 3.

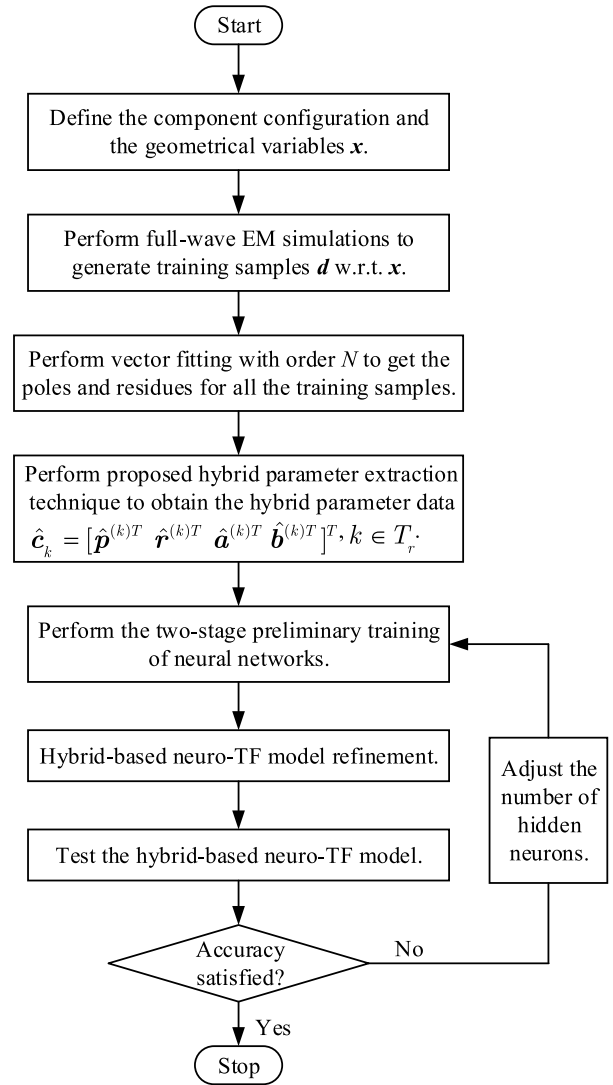


FIGURE 3. The flow diagram of the overall hybrid-based neuro-TF model development process.

IV. APPLICATION EXAMPLES

A. PARAMETRIC MODELING OF A FOUR-POLE WAVEGUIDE FILTER

In this example, the development of a parametric model for the EM behavior of a four-pole waveguide filter [25] using the proposed hybrid-based neuro-TF technique is illustrated. As shown in Fig. 5, the geometrical parameters of the four-pole waveguide filter are $\mathbf{x} = [h_1 h_2 h_3 h_{c1} h_{c2}]^T$.

In the proposed technique, a hybrid-based neuro-TF model is developed as the parametric model for the four-pole waveguide filter example. As defined in the model structure for the four-pole waveguide filter example in Fig. 6, the model has five input geometrical variables and a frequency variable (as an additional input). The model has two outputs,

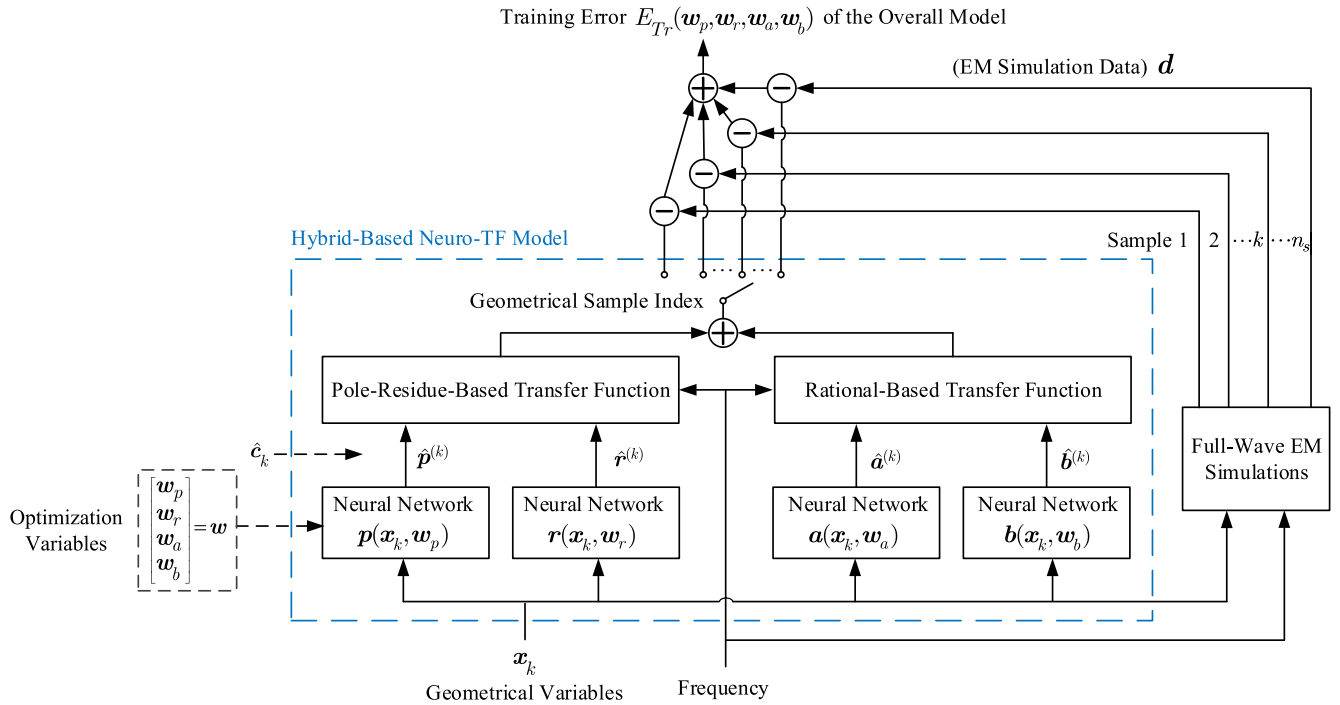


FIGURE 4. The mechanism for the refinement training process of the overall hybrid-based neuro-TF model.

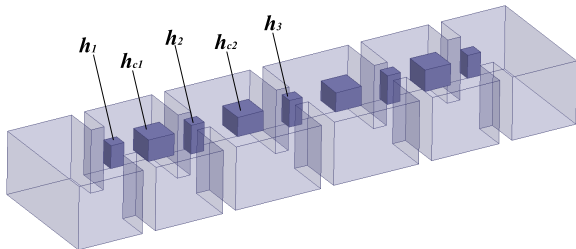


FIGURE 5. The geometrical parameters of the four-pole waveguide filter and the 3D configuration for EM simulation. The geometrical parameters are h_1, h_2, h_3, h_{c1} , and h_{c2} . A parametric model w.r.t. these five geometrical parameters is to be developed.

i.e., $y = [RS_{11} \ IS_{11}]^T$, which are the real and imaginary parts of S_{11} , respectively.

The ANSYS HFSS EM simulator is used to perform the full-wave EM simulation for generation of training and testing data for parametric modeling. A parallel computation method is used in data generation. Message passing interface (MPI) is used for parallel processing with distributed memory. Design of experiments (DOE) [26] method is used as the sampling method to sample both training data and testing data.

As can be seen in Table 1, the proposed parametric modeling technique is applied to two different cases, in terms of the parameter range. Case 1 considers a narrower parameter range and Case 2 considers an increased parameter range. In both cases, the total order N of the hybrid-based transfer functions is set to twelve for all the training and test samples. The number of frequency samples for each geometrical

sample is 101. In both cases, nine levels of DOE are used to sample the training data, resulting in a total number of 81 samples, while eight levels of DOE are used to sample the testing data, resulting in a total number of 64 samples. The detailed ranges of training data and testing data for the two cases are listed in Table 1.

TABLE 1. Definition of training and testing data for the four-pole waveguide filter example.

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

The original frequency range in this example is 10.5 GHz - 11.5 GHz. Scaling and shifting of the frequency range are performed. The total order N is split into the order N_1

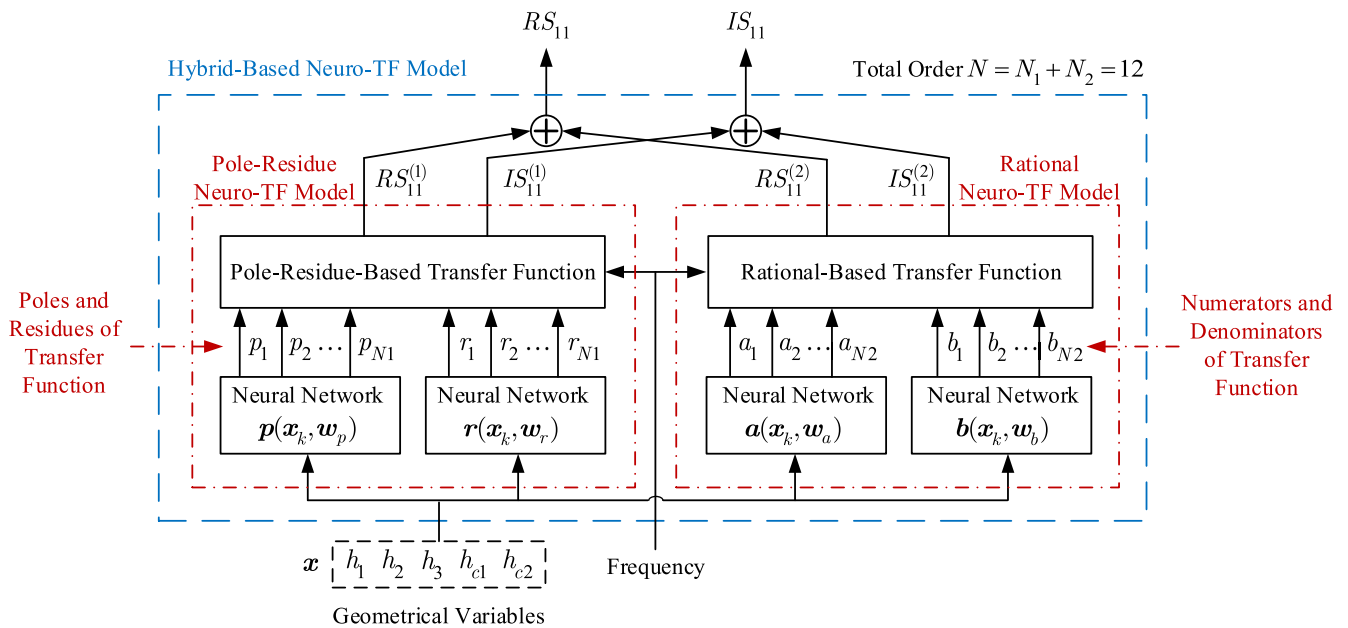


FIGURE 6. The structure of the proposed hybrid-based neuro-TF model for the four-pole waveguide filter example.

of the pole-residue-based transfer function and the order N_2 of the rational-based transfer function by using the proposed hybrid parameter extraction technique. Based on the pole/residue data (which are directly obtained from the vector fitting of training data), we identify the discontinuous poles/residues, non-smooth-continuous poles/residues, and the smooth-continuous poles/residues for both cases. The pole-residue data in Case 1 are relatively smooth as the values of the geometrical parameters change. By contrast, the pole-residue data in Case 2 have discontinuity issue in some of the imaginary parts of poles and residues, and have the accompanied non-smoothness issue in other poles and residues, as the values of the geometrical parameters change. The discontinuity issue and the non-smoothness issue in Case 2 can be observed in Fig. 7. As shown in the figure, the original poles/residue data have the discontinuity issue and the non-smoothness issue. After using the proposed extraction technique, the newly obtained rational data have better smoothness than the original poles/residue data. Based on the proposed hybrid parameter extraction technique illustrated in Section III-B, N_1 is set to eight and N_2 is set to four in Case 1. N_1 is set to six and N_2 is set to six in Case 2. The hybrid-based neuro-TF model is trained using the *NeuroModelerPlus* software for both preliminary training and refinement training. The average training error is 0.909% in Case 1 and the average testing error is 1.620% in Case 2.

For comparison purpose, the recent pole-residue-based neuro-TF modeling method and the existing rational-based neuro-TF modeling method [23] are applied to the two cases. Table 2 shows the comparisons of ANN structures, number of hidden neurons, and average training and testing errors between several parametric modeling methods and EM data.

In Case 1, since the geometrical parameters vary within a small range, discontinuity and non-smoothness issues do not exist and the pole-residue data are relatively smooth. Both the pole-residue-based neuro-TF method and the proposed method obtain comparatively small training and testing errors. The rational-based neuro-TF method obtains relatively large training and testing errors, since the order of the transfer function is relatively high.

In Case 2, the geometrical parameters vary within an increased range and the pole/residue data have the discontinuity issue and the accompanied non-smoothness issue. The pole-residue-based neuro-TF method encounters the discontinuity issue and the non-smoothness issue, resulting in large testing error. This is because that the pole/residue data (which are used in the preliminary training) have the discontinuity and non-smoothness issues. These issues are also the cause for the high nonlinearity of pole/residues from sample to sample. To achieve good training result, the number of hidden neurons is increased from 10 to 50. Such increase of the number of hidden neurons, however, leads to the overlearning of the ANN models, making the testing error even worse. The rational-based neuro-TF method obtains relatively smaller testing error than that obtained by the pole-residue-based neuro-TF method with the same number of hidden neurons. This is because that the rational data (which are converted from the pole-residue data for preliminary training) have better smoothness. However, the order of the rational-based transfer function is relatively high. The rational-based transfer function is more sensitive to its coefficients, and the coefficients are more nonlinear than the pole/residues over the increased geometrical parameter range from sample to sample. Therefore, the testing error of the rational-based

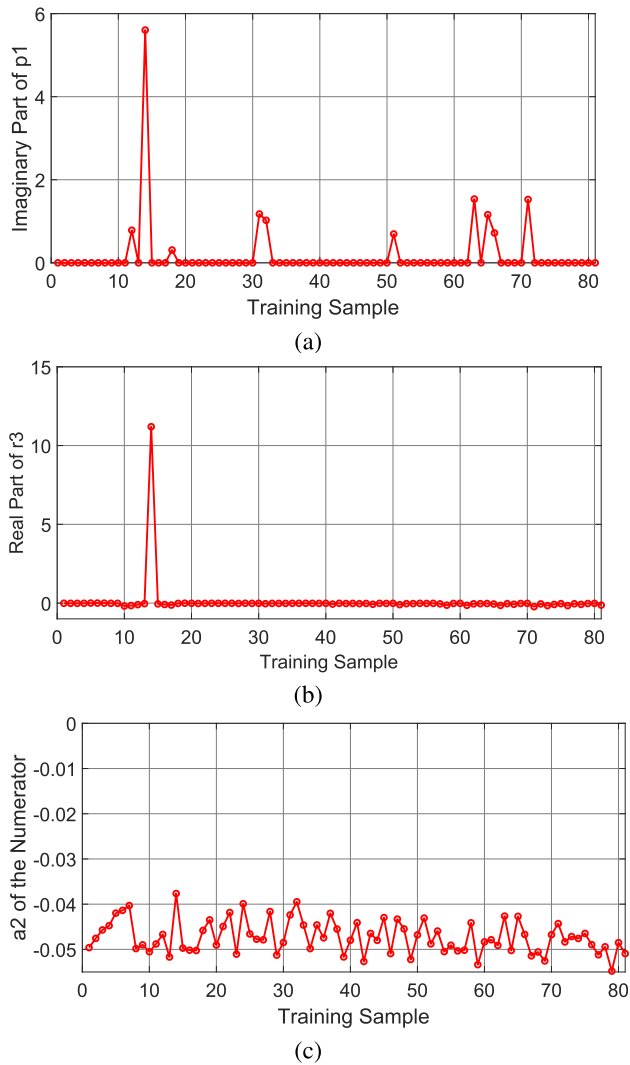


FIGURE 7. The original poles/residue data with discontinuity issue and the non-smoothness issue versus the newly obtained rational data with improved smoothness (after using the proposed hybrid parameter extraction technique) in Case 2 for the four-pole waveguide filter example. (a) the imaginary part of p_1 versus the training sample from the original poles/residue data, (b) the real part of r_3 versus the training sample from the original poles/residue data, (c) a_2 of the numerator versus the training sample from the newly obtained rational data.

neuro-TF model is still high. These are the reasons why the parametric models trained by the two existing methods cannot obtain good training and testing accuracy.

The proposed hybrid-based neuro-TF method can achieve better accuracy in training and testing than the two existing methods for both cases, as shown in the comparison in Table 2. It can be also noticed that the amount of hidden neurons used in the hybrid-based neuro-TF model is small in both cases. This indicates that the good learning of the model is achieved, leading to good training and testing results. The reason behind is that the poles/residues which have the discontinuity and non-smoothness issues in the original pole/residue data are reformatted into the new rational data (of which the smoothness is improved) with less order.

TABLE 2. Comparisons of different modeling methods for the four-pole waveguide filter example.

Training Method		Number of Hidden Neurons	Average Training Error	Average Testing Error		
Case 1 (Narrower Range)	Pole-Residue Neuro-TF Method	Pole NN	6	0.966 %	0.986 %	
		Residue NN	6			
	Rational Neuro-TF Method	Numerator NN	10	5.094 %	5.221 %	
		Denominator NN	10			
		Numerator NN	40	2.562 %		
		Denominator NN	40			
	Proposed Hybrid Neuro-TF Method	Pole NN	4	0.837 %	0.909 %	
		Residue NN	4			
Numerator NN		6				
Denominator NN		6				
Case 2 (Increased Range)	Pole-Residue Neuro-TF Method	Pole NN	10	8.823 %	9.413 %	
		Residue NN	10			
	Rational Neuro-TF Method	Pole NN	50	1.084 %	1299.1 %	
		Residue NN	50			
		Numerator NN	10	5.621 %		
		Denominator NN	10			
	Proposed Hybrid Neuro-TF Method	Numerator NN	50	1.951 %	24.263 %	
		Denominator NN	50			
Proposed Hybrid Neuro-TF Method		Pole NN	10	0.986 %		1.620 %
		Residue NN	10			
	Numerator NN	10				
	Denominator NN	10				

The poles/residues which are smooth and continuous in the original pole/residue data are gathered together to format into the new pole/residue data (which are smooth) with less order. These new pole/residue data are used to train the pole-residue-based neuro-TF model and the new rational data are used to train the rational-based neuro-TF model during the preliminary training. Since there is no discontinuity and non-smoothness issues in the new pole/residue data, the pole-residue-based neuro-TF model can achieve good learning with small amount of hidden neurons. Since the new rational data are smoother and with less order, the rational-based transfer function is less sensitive to its coefficients, leading to good learning of the rational-based neuro-TF model with small amount of hidden neurons as well. The preliminary trained pole-residue-based neuro-TF and rational-based

neuro-TF models are then assembled into the hybrid-based neuro-TF model for the refinement training. These are the reasons why the parametric models trained by the proposed method can obtain good training and testing accuracy.

Fig. 8 shows the output S_{11} of the proposed hybrid-based neuro-TF model for three different test geometrical samples in the four-pole waveguide filter example. The S_{11} response from the proposed model is compared

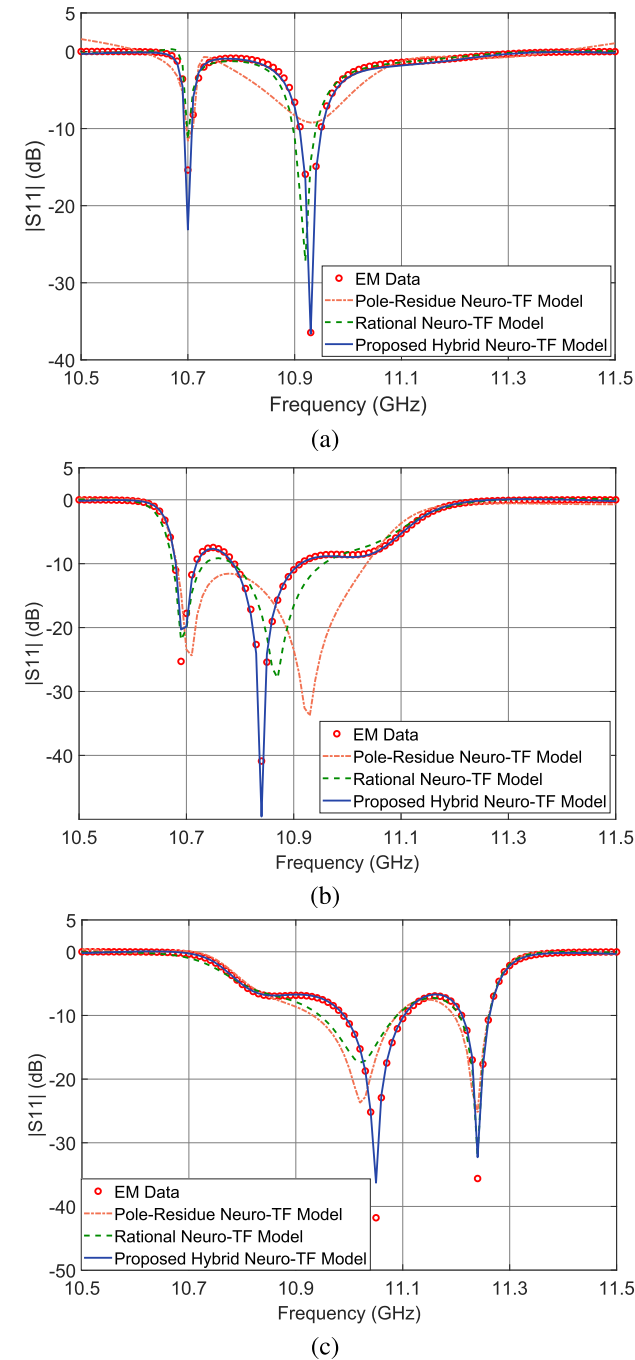


FIGURE 8. Comparison of $|S_{11}|$ (dB) of the models developed using different modeling methods and the HFSS EM data: (a) test geometrical sample #1, (b) test geometrical sample #2, and (c) test geometrical sample #3 for the four-pole waveguide filter example.

with the model responses using different methods and EM data. The geometrical variables for the three samples (#1, #2, and #3) are listed follows.

Test geometrical sample #1: $x = [3.44 \ 4.30 \ 4.12 \ 3.12 \ 3.06]^T$ (mm)

Test geometrical sample #2: $x = [3.52 \ 4.46 \ 3.96 \ 3.28 \ 3.06]^T$ (mm)

Test geometrical sample #3: $x = [3.48 \ 4.46 \ 4.04 \ 3.32 \ 2.90]^T$ (mm)

It is observed from Fig. 8 that the model using the proposed hybrid-based neuro-TF method can achieve good model accuracy for different geometrical samples even though these test samples are never used in training.

B. PARAMETRIC MODELING OF A THREE-POLE H-PLANE FILTER

In this example, the development of a parametric model for the EM behavior of a three-pole H -plane filter [27] using the proposed hybrid-based neuro-TF technique is illustrated. As shown in Fig. 9, the geometrical parameters of the three-pole H -plane filter are $x = [L_1 \ L_2 \ W_1 \ W_2]^T$ with $a = 19.05$ mm, $b = 9.525$ mm, and $t = 2.0$ mm.

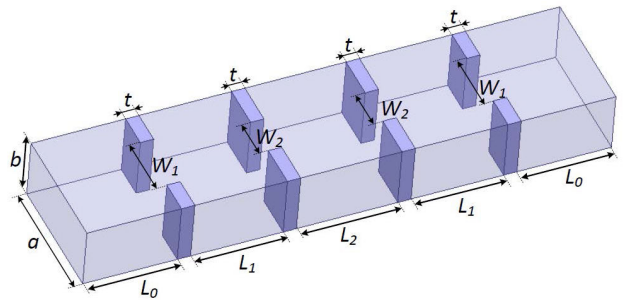


FIGURE 9. The geometrical parameters of the three-pole H -plane filter and the 3D configuration for EM simulation. The geometrical parameters are $L_1, L_2, W_1,$ and W_2 , with $a = 19.05$ mm, $b = 9.525$ mm, and $t = 2.0$ mm. A parametric model w.r.t. these four geometrical parameters is to be developed.

In the proposed technique, a hybrid-based neuro-TF model is developed as the parametric model for the three-pole H -plane filter example. As defined in the model structure for the three-pole H -plane filter example in Fig. 10, the model has four input geometrical variables and a frequency variable (as an additional input). The model has two outputs, i.e., $y = [RS_{11} \ IS_{11}]^T$, which are the real and imaginary parts of S_{11} , respectively.

The ANSYS HFSS EM simulator is used to perform the full-wave EM simulation for generation of training and testing data. A parallel computation method is used in data generation. DOE method is used as sampling method.

As can be seen in Table 3, the proposed parametric modeling technique is applied to two different cases, similar to the four-pole waveguide filter example. Case 1 considers a narrower parameter range and Case 2 considers an increased parameter range. In both cases, the total order N of the hybrid-based transfer functions is set to ten for all the training

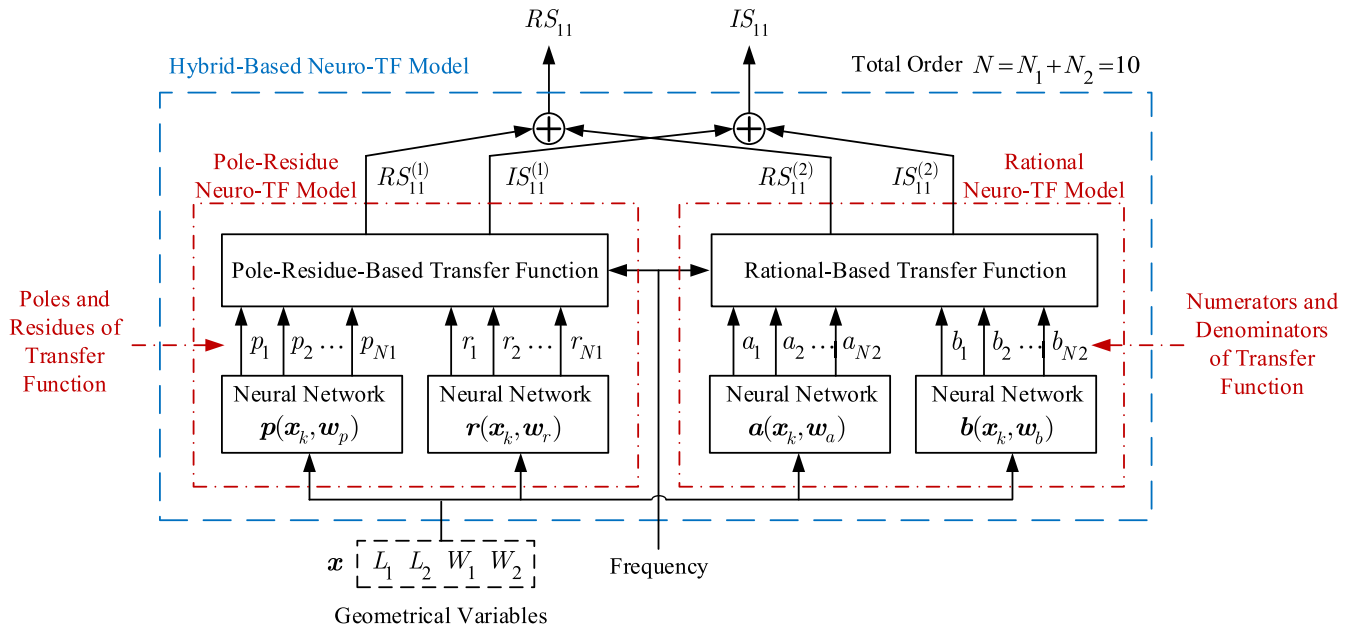


FIGURE 10. The structure of the proposed hybrid-based neuro-TF model for the three-pole H-plane filter example.

TABLE 3. Definition of training and testing data for the three-pole H-plane filter example.

Geometrical Variables		Training Data (49 samples)			Testing Data (49 samples)		
		Min	Max	Step	Min	Max	Step
Case 1 (Narrower Range)	L_1 (mm)	13.84	14.12	0.05	13.86	14.10	0.04
	L_2 (mm)	15.05	15.35	0.05	15.07	15.33	0.04
	W_1 (mm)	8.91	9.09	0.03	8.93	9.08	0.03
	W_2 (mm)	5.94	6.06	0.02	5.95	6.05	0.02
Case 2 (Increased Range)	L_1 (mm)	13.70	14.26	0.09	13.75	14.21	0.08
	L_2 (mm)	14.90	15.50	0.10	14.95	15.45	0.08
	W_1 (mm)	8.82	9.18	0.06	8.85	9.15	0.05
	W_2 (mm)	5.88	6.12	0.04	5.90	6.10	0.03

and test samples. The number of frequency samples for each geometrical sample is 71. In both cases, seven levels of DOE are used to sample the training and testing data, resulting in a total number of 49 training samples and 49 testing samples. The detailed ranges of training data and testing data for the two cases are listed in Table 3.

The original frequency range in this example is 11.6 GHz - 12.4 GHz. Scaling and shifting of the frequency range are performed. The total order N is split into the order N_1 of the pole-residue-based transfer function and the order N_2 of the rational-based transfer function by using the proposed hybrid parameter extraction technique. Based on the pole/residue data (which are directly obtained from the vec-

tor fitting of training data), we identify the discontinuous poles/residues, non-smooth-continuous poles/residues, and the smooth-continuous poles/residues for both cases. The pole-residue data in Case 1 are relatively smooth as the values of the geometrical parameters change. By contrast, the pole-residue data in Case 2 have discontinuity issue in some of the imaginary parts of poles and residues, and have the accompanied non-smoothness issue in other poles and residues, as the values of the geometrical parameters change. The discontinuity issue and the non-smoothness issue in Case 2 can be observed in Fig. 11. As shown in the figure, the original poles/residue data have the discontinuity issue and the non-smoothness issue. After using the proposed extraction technique, the newly obtained rational data have better smoothness than the original poles/residue data. Based on the proposed hybrid parameter extraction technique illustrated in Section III-B, N_1 is set to six and N_2 is set to four in both cases. Similarly, the hybrid-based neuro-TF model is trained using the *NeuroModelerPlus* software for both preliminary training and refinement training. The average training error is 0.356% in Case 1 and the average testing error is 1.729% in Case 2.

For comparison purpose, the recent pole-residue-based neuro-TF modeling method and the existing rational-based neuro-TF modeling method are also applied to the two cases of this example. Table 4 shows the comparisons of ANN structures, number of hidden neurons, and average training and testing errors between several parametric modeling methods and EM data.

In Case 1, since the geometrical parameters vary within a small range, discontinuity and non-smoothness issues do not exist and the pole-residue data are relatively smooth. The order of the transfer function is moderate (not too high).

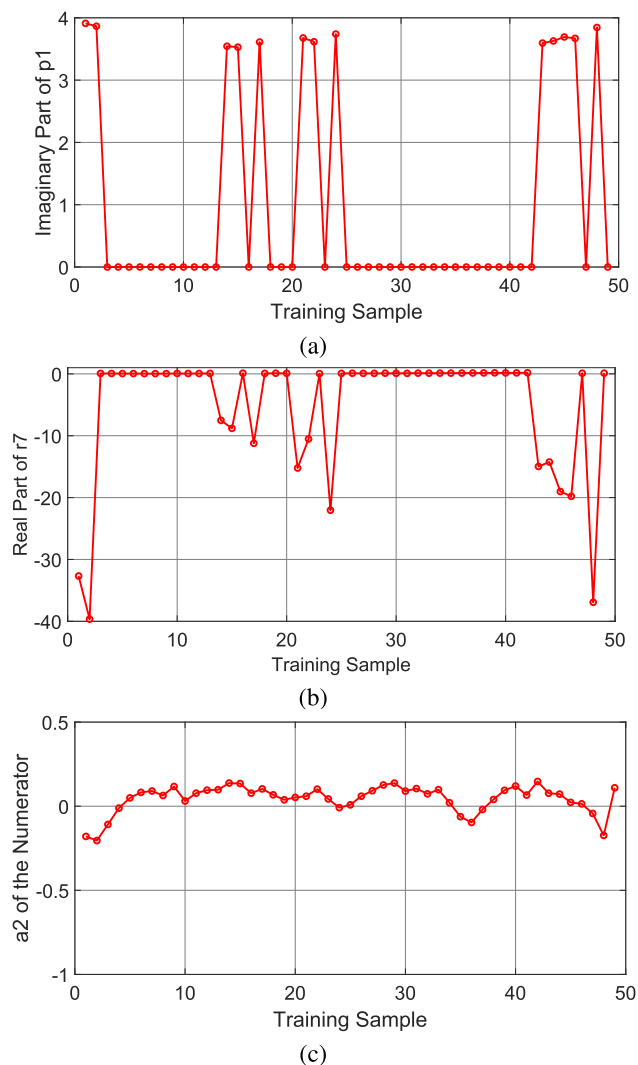


FIGURE 11. The original poles/residue data with discontinuity issue and the non-smoothness issue versus the newly obtained rational data with improved smoothness (after using the proposed hybrid parameter extraction technique) in Case 2 for the three-pole *H*-plane filter example. (a) the imaginary part of p_1 versus the training sample from the original poles/residue data, (b) the real part of r_7 versus the training sample from the original poles/residue data, (c) a_2 of the numerator versus the training sample from the newly obtained rational data.

For these reasons, all the methods in this case achieve comparatively small training and testing errors.

In Case 2, the geometrical parameters vary within an increased range and the pole/residue data have the discontinuity and non-smoothness issues. The pole-residue-based neuro-TF method, again, encounters these issues, resulting in large testing error. To achieve good training result, the number of hidden neurons is increased from 5 to 40. Such increase of the number of hidden neurons, however, leads to the overlearning of the ANN models, making the testing error even worse.

The rational-based neuro-TF method obtains relatively smaller testing error than that obtained by the pole-residue-based neuro-TF method. However, the order of the

TABLE 4. Comparisons of different modeling methods for the three-pole *H*-plane filter example.

Training Method		Number of Hidden Neurons	Average Training Error	Average Testing Error	
Case 1 (Narrower Range)	Pole-Residue Neuro-TF Method	Pole NN	10	0.128 %	0.174 %
		Residue NN	10		
	Rational Neuro-TF Method	Numerator NN	10	0.630 %	0.718 %
		Denominator NN	10		
	Proposed Hybrid Neuro-TF Method	Pole NN	10	0.100 %	0.356 %
		Residue NN	10		
Numerator NN		10			
Denominator NN		10			
Case 2 (Increased Range)	Pole-Residue Neuro-TF Method	Pole NN	5	21.751 %	29.376 %
		Residue NN	5		
		Pole NN	40	1.559 %	25.927 %
		Residue NN	40		
	Rational Neuro-TF Method	Numerator NN	10	3.680 %	3.689 %
		Denominator NN	10		
		Numerator NN	50	1.596 %	4.340 %
		Denominator NN	50		
	Proposed Hybrid Neuro-TF Method	Pole NN	10	1.708 %	1.729 %
		Residue NN	10		
Numerator NN		10			
Denominator NN		10			

rational-based transfer function is not low enough and the geometrical variations are larger than those in Case 1. The rational-based transfer function is still more sensitive to its coefficients, and the coefficients are more nonlinear than the pole/residues over the increased geometrical parameter range from sample to sample. Therefore, the testing error of the rational-based neuro-TF model is still not satisfactory.

The proposed hybrid-based neuro-TF method can achieve better accuracy in training and testing than the two existing methods for both cases, as shown in the comparison in Table 2. It can be also noticed that the amount of hidden neurons used in the hybrid-based neuro-TF model is small in both cases. This suggests that the good learning of the model is achieved, leading to good training and testing results. In the proposed technique, there is no discontinuity and non-smoothness issues in the new pole/residue data and the new rational data are smoother and with less order. The poles/residues which are non-smooth and discontinuous in the original pole/residue data are gathered

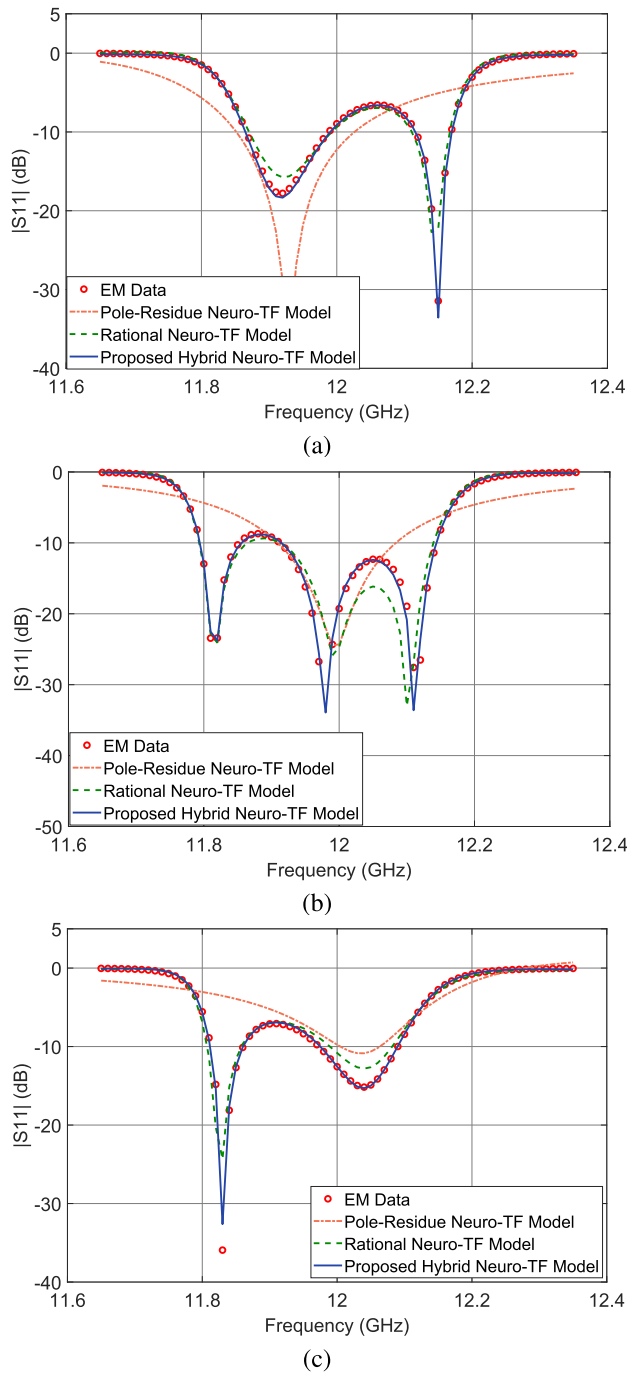


FIGURE 12. Comparison of $|S_{11}|$ (dB) of the models developed using different modeling methods and the HFSS EM data: (a) test geometrical sample #1, (b) test geometrical sample #2, and (c) test geometrical sample #3 for the three-pole H -plane filter example.

together to reformat into the new rational data (which are much smoother) with less order. Since the new rational data are smoother and with less order, the rational-based transfer function is less sensitive to its coefficients. Therefore, both the pole-residue-based neuro-TF model and the rational-based neuro-TF model can achieve good learning with small amount of hidden neurons, leading to good training and testing accuracy of the hybrid-based neuro-TF model.

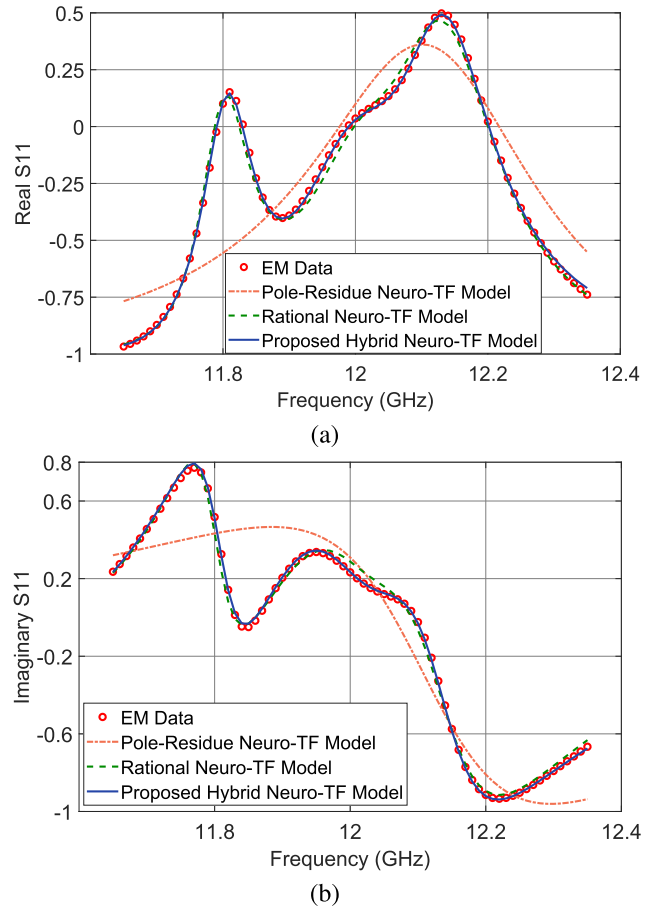


FIGURE 13. Comparison of real and imaginary parts of S_{11} of the models developed using different modeling methods and the HFSS EM data at test geometrical sample #3 for the three-pole H -plane filter example. (a) real parts of S_{11} and (b) imaginary parts of S_{11} .

Fig. 12 shows the output S_{11} of the proposed hybrid-based neuro-TF model for three different test geometrical samples in the three-pole H -plane filter example. The S_{11} response from the proposed model is compared with the model responses using different methods and EM data. The geometrical variables for the three samples (#1, #2, and #3) are listed follows.

Test geometrical sample #1: $x = [14.08 \ 15.12 \ 9.00 \ 5.93]^T$ (mm)

Test geometrical sample #2: $x = [14.06 \ 15.20 \ 8.90 \ 6.10]^T$ (mm)

Test geometrical sample #3: $x = [13.98 \ 15.37 \ 9.05 \ 5.90]^T$ (mm)

It is observed from Fig. 12 that the model using the proposed hybrid-based neuro-TF method can achieve good model accuracy for different geometrical samples even though these test samples are never used in training. As additional results, we also include the comparison of real and imaginary parts of S_{11} using different modeling methods in Fig. 13. It can also be seen from this figure that the model using the proposed hybrid-based neuro-TF method can achieve good model accuracy.

So far the focus of our work is on addressing the discontinuity issue and the accompanied non-smoothness issue in the poles/residues with the same order. Discontinuity issue coming from order changing is another situation to be considered. Order changing is needed for parametric modeling especially when different responses cannot be effectively represented with a same order from vector fitting, as the geometrical variables change. It would be an interesting future direction to consider the case of the discontinuity and on-smoothness issues coming from order changing.

V. CONCLUSION

This paper has presented a novel hybrid-based neuro-TF parametric modeling technique which systematically combines both pole-residue and rational formats of the transfer functions, addressing the discontinuity and non-smoothness issues. The proposed technique identifies the discontinuous poles/residues, the smooth-continuous poles/residues, and non-smooth-continuous poles/residues. The discontinuous and non-smooth-continuous poles/residues are gathered together to reformat into the coefficients of the rational-based transfer function to effectively address the discontinuity and non-smoothness issues. The smooth-continuous poles/residues are gathered together to be remained in the pole-residue format of the pole-residue-based transfer function to maintain the capability of handling high-order problem. In this way, the discontinuity and non-smoothness issues are solved by the proposed hybrid-based neuro-TF modeling technique and high-order modeling capability is maintained by the overall model. The developed model can provide accurate and fast prediction of the EM behavior of microwave components. As illustrated in the examples and compared with the existing modeling methods, the proposed technique can obtain better accuracy in challenging applications of large geometrical variations and high order, addressing the discontinuity and non-smoothness issues.

ACKNOWLEDGMENT

The authors thank Jianguo Ma of Guangdong University of Technology, Guangzhou, China, for his support of this research work and his technical discussions.

REFERENCES

- [1] J. E. Rayas-Sanchez, "EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 420–435, Jan. 2004.
- [2] V. Rizzoli, A. Costanzo, D. Masotti, A. Lipparini, and F. Matri, "Computer-aided optimization of nonlinear microwave circuits with the aid of electromagnetic simulation," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 362–377, Jan. 2004.
- [3] M. B. Steer, J. W. Bandler, and C. M. Snowden, "Computer-aided design of RF and microwave circuits and systems," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 3, pp. 996–1005, Mar. 2002.
- [4] P. Burrascano, S. Fiori, and M. Mongiardo, "A review of artificial neural networks applications in microwave computer-aided design (invited article)," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 158–174, May 1999.
- [5] S. A. Sadrossadat, Y. Cao, and Q.-J. Zhang, "Parametric modeling of microwave passive components using sensitivity-analysis-based adjoint neural-network technique," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 5, pp. 1733–1747, May 2013.
- [6] Q. J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*. Norwood, MA, USA: Artech House, 2000.
- [7] V. K. Devabhaktuni, B. Chattaraj, M. C. E. Yagoub, and Q.-J. Zhang, "Advanced microwave modeling framework exploiting automatic model generation, knowledge neural networks, and space mapping," *IEEE Trans. Microw. Theory Techn.*, vol. 51, no. 7, pp. 1822–1833, Jul. 2003.
- [8] J. W. Bandler, M. A. Ismail, J. E. Rayas-Sanchez, and Q.-J. Zhang, "Neuro-modeling of microwave circuits exploiting space-mapping technology," *IEEE Trans. Microw. Theory Techn.*, vol. 47, no. 12, pp. 2417–2427, Dec. 1999.
- [9] J. E. Rayas-Sanchez and V. Gutierrez-Ayala, "EM-based Monte Carlo analysis and yield prediction of microwave circuits using linear-input neural-output space mapping," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 12, pp. 4528–4537, Dec. 2006.
- [10] Y. Cao and G. Wang, "A wideband and scalable model of spiral inductors using space-mapping neural network," *IEEE Trans. Microw. Theory Techn.*, vol. 55, no. 12, pp. 2473–2480, Dec. 2007.
- [11] H. Kabir, L. Zhang, M. Yu, P. H. Aaen, J. Wood, and Q.-J. Zhang, "Smart modeling of microwave devices," *IEEE Microw. Mag.*, vol. 11, no. 3, pp. 105–118, May 2010.
- [12] J. W. Bandler, Q. S. Cheng, S. A. Dakrouy, A. S. Mohamed, M. H. Bakr, K. Madsen, and J. Sondergaard, "Space mapping: The state of the art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 337–361, Jan. 2004.
- [13] S. Koziel, J. W. Bandler, and Q. S. Cheng, "Constrained parameter extraction for microwave design optimisation using implicit space mapping," *IET Microw., Antennas Propag.*, vol. 5, no. 10, pp. 1156–1163, 2011.
- [14] S. Koziel, J. W. Bandler, and K. Madsen, "Space mapping with adaptive response correction for microwave design optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 2, pp. 478–486, Feb. 2009.
- [15] R. B. Ayed, J. Gong, S. Brisset, F. Gillon, and P. Brochet, "Three-level output space mapping strategy for electromagnetic design optimization," *IEEE Trans. Magn.*, vol. 48, no. 2, pp. 671–674, Feb. 2012.
- [16] L. Zhang, J. Xu, M. C. E. Yagoub, R. Ding, and Q.-J. Zhang, "Efficient analytical formulation and sensitivity analysis of neuro-space mapping for nonlinear microwave device modeling," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 9, pp. 2752–2767, Sep. 2005.
- [17] D. Gorissen, L. Zhang, Q.-J. Zhang, and T. Dhaene, "Evolutionary neuro-space mapping technique for modeling of nonlinear microwave devices," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 2, pp. 213–229, Feb. 2011.
- [18] S. Koziel, J. W. Bandler, and K. Madsen, "A space-mapping framework for engineering optimization-theory and implementation," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 10, pp. 3721–3730, Oct. 2006.
- [19] S. Koziel, J. W. Bandler, and Q. S. Cheng, "Tuning space mapping design framework exploiting reduced electromagnetic models," *IET Microw., Antennas Propag.*, vol. 5, no. 10, pp. 1219–1226, 2011.
- [20] X. Ding, V. K. Devabhaktuni, B. Chattaraj, M. C. E. Yagoub, M. Deo, J. Xu, and Q. J. Zhang, "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 Techn.*, vol. 52, no. 1, pp. 436–449, Jan. 2004.
- [21] V.-M.-R. Gongal-Reddy, F. Feng, and Q.-J. Zhang, "Parametric modeling of millimeter-wave passive components using combined neural networks and transfer functions," in *Proc. Global Symp. Millimeter-Waves (GSMM)*, Montreal, QC, Canada, May 2015, pp. 1–3.
- [22] F. Feng, C. Zhang, J. Ma, and Q.-J. Zhang, "Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 1, pp. 60–77, Jan. 2016.
- [23] Y. Cao, G. Wang, and Q.-J. Zhang, "A new training approach for parametric modeling of microwave passive components using combined neural networks and transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 11, pp. 2727–2742, Nov. 2009.
- [24] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Trans. Power Del.*, vol. 14, no. 3, pp. 1052–1061, Jul. 1999.
- [25] C. Zhang, F. Feng, V.-M.-R. Gongal-Reddy, Q. J. Zhang, and J. W. Bandler, "Cognition-driven formulation of space mapping for equal-ripple optimization of microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 7, pp. 2154–2165, Jul. 2015.

- [26] S. R. Schmidt and R. G. Launsby, *Understanding Industrial Designed Experiments*. Colorado Springs, CO, USA: Air Force Academy, 1992.
- [27] J. Zhang, F. Feng, W. Zhang, J. Jin, J. Ma, and Q.-J. Zhang, "A novel training approach for parametric modeling of microwave passive components using Padé via Lanczos and EM sensitivities," *IEEE Trans. Microw. Theory Techn.*, early access, Mar. 24, 2020, doi: 10.1109/TMTT.2020.2979445.



University, Ottawa, ON, Canada.

His current research interests include parametric modeling, GaN HEMT modeling, nonlinear microwave device and circuit modeling, power amplifier behavioral modeling, space mapping modeling, and neural network modeling.



University, Ottawa, ON, Canada.

His current research interests include parametric modeling, GaN HEMT modeling, nonlinear microwave device and circuit modeling, power amplifier behavioral modeling, space mapping modeling, and neural network modeling.



University, Ottawa, ON, Canada.

His current research interests include parametric modeling, GaN HEMT modeling, nonlinear microwave device and circuit modeling, power amplifier behavioral modeling, space mapping modeling, and neural network modeling.



JIANAN ZHANG (Student Member, IEEE) was born in Tieling, Liaoning, China, in 1991. He received the B.Eng. degree from Tianjin University, Tianjin, China, in 2013. He is currently pursuing the Ph.D. degree with the School of Microelectronics, Tianjin University, Tianjin, and the Cotutelle Ph.D. degree with the Department of Electronics, Carleton University, Ottawa, ON, Canada.

His current research interests include statistical modeling, electromagnetic (EM)-based yield optimization of microwave structures, uncertainty quantification using polynomial chaos, computational electromagnetics, and space mapping-based EM optimization.



JING JIN (Student Member, IEEE) was born in Suizhou, Hubei, China, in 1991. She received the B.Eng. degree from Wuhan University, Wuhan, China, in 2014. She is currently pursuing the Ph.D. degree with the School of Microelectronics, Tianjin University, Tianjin, China, and the Cotutelle Ph.D. degree with the Department of Electronics, Carleton University, Ottawa, ON, Canada.

Her current research interests include the modeling of microwave circuits, deep neural network modeling technique, and computer-aided optimization and tuning of microwave filters.



QI-JUN ZHANG (Fellow, IEEE) received the B.Eng. degree from the Nanjing University of Science and Technology, Nanjing, China, in 1982, and the Ph.D. degree in electrical engineering from Mc-Master University, Hamilton, ON, Canada, in 1987.

From 1982 to 1983, he was with the System Engineering Institute, Tianjin University, Tianjin, China. From 1988 to 1990, he was with Optimization Systems Associates (OSA) Inc., Dundas, ON, where he developed advanced microwave optimization software. In 1990, he joined the Department of Electronics, Carleton University, Ottawa, ON, where he is currently a Chancellor's Professor. He has authored or coauthored over 300 publications. He authored *Neural Networks for RF and Microwave Design* (Artech House, 2000), coedited *Modeling and Simulation of High-Speed VLSI Interconnects* (Kluwer, 1994), and contributed to the *Encyclopedia of RF and Microwave Engineering* (Wiley, 2005), *Fundamentals of Nonlinear Behavioral Modeling for RF and Microwave Design* (Artech House, 2005), and *Analog Methods for Computer-Aided Analysis and Diagnosis* (Marcel Dekker, 1988). He was a Guest Co-Editor for the Special Issue on High-Speed VLSI Interconnects for the *International Journal of Analog Integrated Circuits and Signal Processing* (Kluwer, 1994), and twice was a Guest Editor for the Special Issue on Applications of ANN to RF and Microwave Design for the *International Journal of RF and Microwave Computer-Aided Engineering* (Wiley, 1999 and 2002). His research interests are microwave design automation especially on neural-network and optimization methods for high-speed/high-frequency circuit design.

Dr. Zhang is a Fellow of the Electromagnetics Academy and a Fellow of Canadian Academy of Engineering. He is the Chair of the Technical Committee on Design Automation (MTT-2) of the IEEE Microwave Theory and Techniques Society (IEEE MTT-S). He is an Associate Editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES.

...