Parametric Modeling Incorporating Joint Polynomial-Transfer Function With Neural Networks for Microwave Filters

Yan Zhuo, Feng Feng, *Member, IEEE*, Jianan Zhang, *Member, IEEE*, and Qi-Jun Zhang, *Fellow, IEEE*

***Abstract*— This article proposes a novel parametric modeling technique incorporating a joint polynomial-transfer function with neural networks (short for neuro-PTF) for electromagnetic (EM) behaviors of microwave filters. In the proposed technique, the polynomial function is introduced together with the pole-residue-based transfer function to represent the EM responses. The pole-residue-based transfer function is used to represent the whole EM response at the beginning and is subsequently divided into multiple subtransfer functions where each subtransfer function contains one pair of pole/residue. A novel smoothness-discriminating algorithm is proposed to judge the smoothness of each subtransfer function response and separate the pole/residue pairs whose subtransfer function response is determined to be smooth. The proposed method introduces low nonlinear polynomial functions to re-fit the smooth parts of subresponse and remains the nonsmooth parts of subresponse for the highly nonlinear transfer functions to represent. By this way, the proposed method avoids the discontinuity problems of nonunique parameter extraction caused by fitting smooth curves using the highly nonlinear transfer function. Neural networks are proposed to learn the relationship between the polynomial coefficients/poles/residues and the geometrical parameters. Utilizing both advantages of the polynomial functions and transfer functions, the proposed method can produce more accurate models than the existing neuro-TF methods, especially with large geometrical variations. The accuracy and robustness of the proposed technique are demonstrated using three EM application examples of microwave filters.**

***Index Terms*— Electromagnetic (EM) modeling, microwave filters, neural networks, neuro-PTF, polynomial function, transfer function.**

I. Introduction

ARTIFICIAL neural network (ANN) is recognized as a powerful tool in electromagnetic (EM)-based parametric modeling [1], [2], [3], [4]. The ANN is trained to learn the relationship between geometrical parameters and EM response, which is beneficial for providing accurate and fast prediction of the EM behavior of microwave filters with geometrical parameters as variables. Using the ANN parametric models, the designer can subsequently perform high-level circuit and system design efficiently.

Various ANN-based parametric modeling methods have been introduced for developing accurate parametric models. To increase the accuracy of ANN modeling interpolation, a sensitivity-analysis-based adjoint neural-network technique is discussed [5]. To increase the smoothness of ANN modeling extrapolation, a multidimensional extrapolation technique for neural-based microwave modeling and design is described [6]. In [7], an ANN-based inverse modeling is presented for efficient design purposes. As a further approach, the knowledge-based neural network has been introduced to combine the ANN with prior knowledge to improve the modeling accuracy and robustness [8], [9], [10], [11], [12]. Space mapping [13], [14], [15] using a neural network as the mapping function is recognized as a special kind of knowledge-based neural network modeling method which is very effective to be used for good modeling accuracy [16].

Recently, an advanced knowledge-based parametric modeling method combining neural networks with transfer function (neuro-TF) has been developed to establish parametric models of EM responses [17], [18], [19], [20]. Recent research studies focus on neuro-TF parametric modeling methods for their efficiency and robustness. In [21], a neuro-TF training method based on rational transfer function formulation is discussed, providing the solution to the discontinuity of coefficients in transfer functions over the geometrical variables. However, when the order of transfer function becomes high, the coefficients in the rational format of transfer function become very sensitive and may lead to inaccurate models. In [22], the transfer function based on the pole–residue format is applied to the establishment of the neuro-TF model. The sensitivities of the EM response w.r.t. the pole/residues are much lower than that of rational coefficients, which makes the pole–residue-based transfer function training more robust. The vector fitting technique [23], [24], [25] is used to extract the pole/residues from S-parameter curves of EM simulations. The order of the transfer function obtained by vector fitting may vary as the values of geometrical parameters change, resulting in the discontinuous problem of pole/residues. In [26], a pole/residue tracking method is introduced to solve this discontinuous problem to obtain accurate models. To properly determine the

Manuscript received 4 April 2022; revised 11 July 2022; accepted 3 September 2022. Date of publication 3 October 2022; date of current version 4 November 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62101382, in part by the Key Research and Development and Transformation Project of Qinghai Province under Grant 2022-QY-212, and in part by the Fundamental Research Funds for the Central Universities under Grant 2242022R10023. *(Corresponding author: Feng Feng.)*

Yan Zhuo and Feng Feng are with the School of Microelectronics, Tianjin University, Tianjin 300072, China (e-mail: emoyu@tju.edu.cn; ff@tju.edu.cn).

Jianan Zhang is with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China (e-mail: jiananzhang@seu.edu.cn).

Qi-Jun Zhang is with the Department of Electronics, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: qjz@doe.carleton.ca).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TMTT.2022.3207761.

Digital Object Identifier 10.1109/TMTT.2022.3207761

minimum order of the transfer function for the pole/residue tracking technique, a systematic order fitting algorithm is discussed [27].

The neuro-TF method utilizes the advantage of transfer functions to accurately fit the highly nonlinear EM responses with sharp and nonsmooth curves at reflection/transmission zeros (e.g., inside filter band). However, when fitting flat and smooth parts of EM responses (e.g., outside filter band), the usage of transfer function representation is not effective. To force the transfer function to fit the flat and smooth curve, nonunique extraction of pole/residues may occur, resulting in the discontinuous pole/residues as the values of geometrical parameters change. To address this problem, a neuro-TF algorithm using a hybrid-based transfer function is presented by transforming part of the pole/residues which are discontinuous into rational coefficients [28]. The transformed rational coefficients behave more smoothly than poles and residues do w.r.t. the change of the values of geometrical parameters. However, since the transfer function is still used to represent the flat and smooth part of EM response, the method in [28] has not completely solved the discontinuous problem. How to solve the discontinuity problem of nonunique parameter extraction using the transfer function representation for EM responses to perform accurate parametric modeling is still an open topic.

To address the discontinuity problem of nonunique parameter extraction using transfer function representation, a novel parametric modeling technique incorporating joint polynomial-transfer function with a neural network (short for neuro-PTF) is proposed. In the proposed technique, the polynomial function is introduced together with the pole–residue-based transfer function to represent the EM responses. The pole–residue-based transfer function is used to represent the whole EM response at the beginning and is subsequently divided into multiple subtransfer functions where each subtransfer function contains one pair of pole/residue. A novel smoothness-discriminating algorithm is proposed to judge the smoothness of each subtransfer function response and separate the pole/residue pairs whose subtransfer function response is determined to be smooth. The proposed method introduces low nonlinear polynomial functions to re-fit the smooth parts of subresponse and remains the nonsmooth parts of subresponse for the highly nonlinear transfer functions to represent. By this way, the proposed method avoids the discontinuity problems of nonunique parameter extraction caused by fitting smooth curves using the highly nonlinear transfer function. Neural networks are proposed to learn the relationship between the polynomial coefficients/poles/residues and the geometrical parameters. Utilizing both advantages of the polynomial functions and transfer functions, the proposed method can produce more accurate models than the existing neuro-TF methods [26], [28], especially with large geometrical variations.

This article is organized as follows. Section II illustrates the discontinuity issue of nonunique parameter extraction using the transfer function. In Section III, the details of the proposed neuro-PTF technique by combining the joint polynomial-transfer function with neural networks are introduced. In Section IV, the accuracy and robustness of the proposed technique are demonstrated using three EM application examples of microwave filters. Section V concludes this article.

II. Discontinuity Issue of Nonunique Parameter Extraction Using Transfer Function

In this section, we illustrate the nonunique parameter extraction using pure transfer function representation, which results in the discontinuity issue of pole/residues w.r.t. geometrical variations. Typically, the S-parameter curve consists of not only the sharp and nonsmooth part of EM responses at reflection/transmission zeros (e.g., inside filter band), but also the flat and smooth part of EM responses (e.g., outside filter band) to well represent the characteristics of a microwave component. To fit the sharp and nonsmooth part of EM responses, the transfer functions have proved their advantage using pole/residue pairs. However, the usage of transfer function representation is not effective for fitting flat and smooth parts of EM responses. To force the transfer function to fit the flat and smooth curves, nonunique extraction of pole/residues may occur. For different geometrical samples, the vector fitting process [23] is carried out individually to extract the pole/residues of the transfer function to represent the EM response curves (e.g., S-parameters) without considering the geometrical parameter variations among different samples. Therefore, the nonunique extraction of pole/residues will result in the discontinuity issue of pole/residues as the values of geometrical parameters change. Since neural networks are used to learn the relationship between pole/residues and geometrical parameters in the existing neuro-TF modeling method, such discontinuity issue will lead to a low accuracy for the trained neural networks, consequently resulting in a low accuracy for the overall neuro-TF model.

A fifth-order microstrip filter example is used to illustrate the discontinuity issue, as shown in Fig. 1. For simplicity of the illustration, one geometrical parameter (i.e., the width w) is considered to perform multiple geometrical samples of EM simulations. Vector fitting with the 14th-order transfer function is used to extract pole/residues for all the geometrical samples. The S-parameter curve represented by a transfer function can be divided into a summation of multiple subtransfer functions formed by each pole/residue pair. Fig. 2 shows the subtransfer function responses formed using four pairs of pole/residues at one geometrical sample. From Fig. 2(a) and (d), we see that the transfer function response formed by the first (or the 13th) pole/residue pair is smooth. As shown in Fig. 2(b) and (c), the transfer function formed using the seventh (or the ninth) pole/residue pairs is nonsmooth with steep peaks/valleys. As illustrated in Fig. 2, subtransfer function responses with smooth and nonsmooth curves both exist after the vector fitting process.

Fig. 3(a) and (b) illustrates the values of imaginary parts of the first pole p_1 and residue r_1 w.r.t. the change in the value of geometrical parameter w, respectively. Since the subtransfer function corresponding to the first pole/residue pair is used to fit the flat and smooth curve as shown in Fig. 2(a), nonunique extraction of the first pole/residues occurs. The nonunique

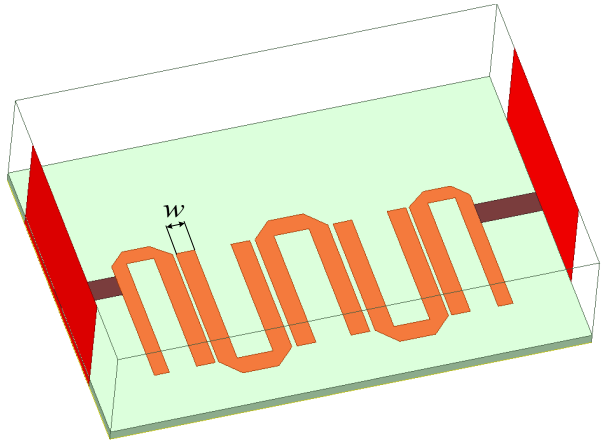


Fig. 1. 3-D configuration of a fifth-order microstrip filter example for illustration. One geometrical parameter, that is, the width w, is used in this example.

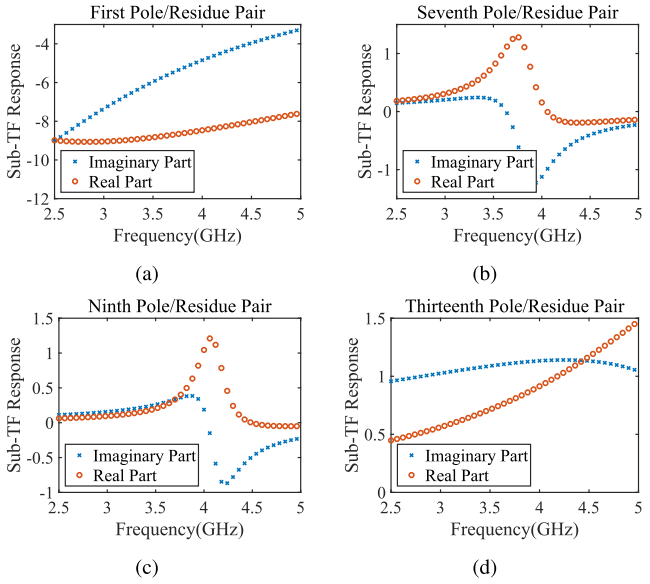


Fig. 2. Real part and imaginary part of the sub-TF response formed using different pole/residue pairs at one geometrical sample. (a) First pole/residue pair. (b) Seventh pole/residue pair. (c) Ninth pole/residue pair. (d) Thirteenth pole/residue pair.

extraction results in the pole/residues discontinuously changing as the value of the geometrical parameter w changes as shown in Fig. 3(a) and (b). This discontinuity of pole/residues leads to the low accuracy of neural network modeling. To address this discontinuity issue, this article proposes to introduce polynomial functions to fit the flat and smooth parts of the EM response. Since the EM response is a complex value, a complex-valued polynomial function is used to fit the EM response. Fig. 3(c) and (d) illustrates the real and imaginary parts of the first coefficients c_0 of the polynomial functions fitting the real and imaginary parts of the EM response w.r.t. the change in values of geometrical parameter w, respectively. As shown in Fig. 3(c) and (d), fitting a smooth curve with a polynomial function obtains continuous coefficients, which avoids the discontinuity issue using the transfer function.

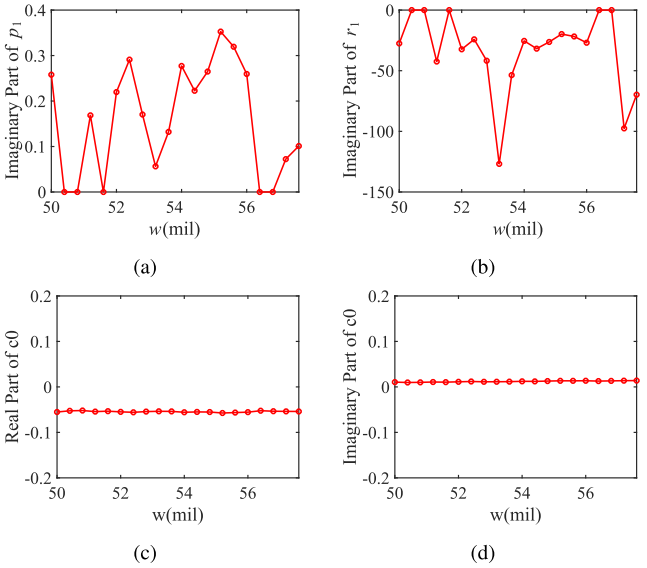


Fig. 3. Comparison of the coefficient of two forms. (a) Imaginary part of the first pole p_1, (b) imaginary part of the first residue r_1, (c) real part of the first polynomial coefficients c_0, and (d) imaginary part of the first polynomial coefficients c_0.

In this article, we proposed a novel formulation to combine the polynomial functions with the transfer function to represent the EM response. The transfer function is used to represent the sharp and nonsmooth parts of EM responses, while the polynomial function is used to represent the flat and smooth parts of EM responses. The detailed modeling structure and the training algorithm are described in Section III.

III. Proposed Neuro-PTF Modeling Technique

The structure of the proposed neuro-PTF modeling technique that incorporates the joint polynomial-transfer function with neural networks is shown in Fig. 4. The model consists of the pole–residue-based transfer functions, polynomial functions, and neural networks. Let $\boldsymbol{x}$ be a vector containing input geometrical parameters. Let y represent the output of the neuro-PTF model which is a sum of the polynomial function response and the pole–residue-based transfer function response. Let $\boldsymbol{w}$ be a vector containing the weighting parameters of the neural networks.

A. Vector Fitting for Pole–Residue Extraction

The proposed algorithm begins with the data generation of EM responses (i.e., S-parameters) of microwave filters by EM simulation for all the geometrical samples. Let k represent the index of geometrical samples, that is, $k = 1, 2, \ldots, n_s$, where n_s represents the total number of geometrical samples used for training the neuro-PTF model. Vector fitting [23] is used to obtain the pole/residues from the EM response for each geometrical sample. We use individual vector fitting with the minimum order of transfer function to generate pole/residues for each sample. With a minimum order of transfer function, we can avoid nonunique and arbitrary numerical solutions of pole/residues from the vector fitting process [26]. Let $\tilde{\boldsymbol{p}}_k$ and

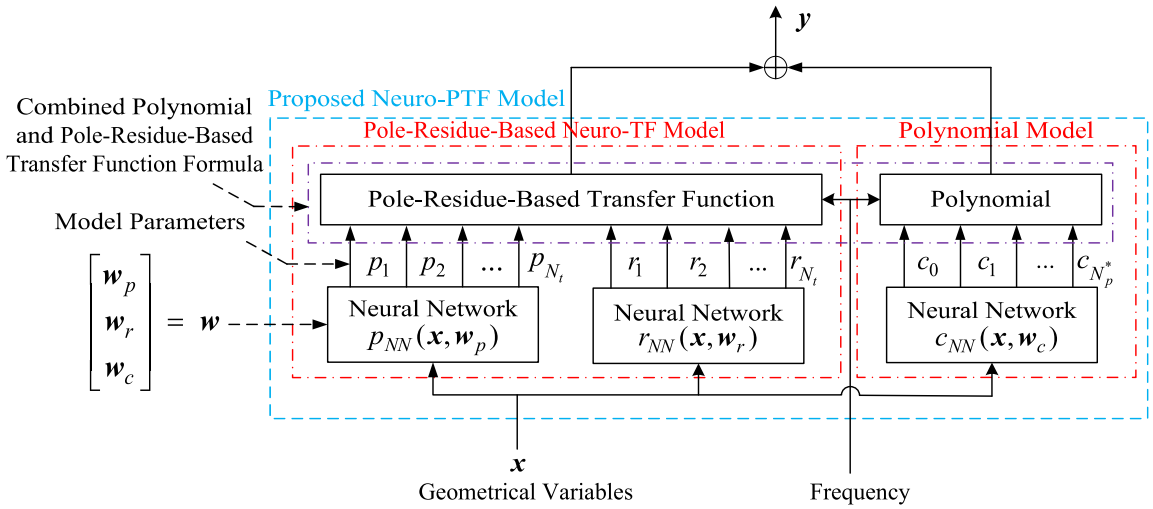


Fig. 4. Structure of the proposed neuro-PTF model. The model consists of the pole–residue-based transfer functions, polynomial functions, and neural networks. $\boldsymbol{x}$ represents the geometrical variables. $\boldsymbol{y}$ represents the real and imaginary parts of the outputs of the neuro-PTF model. $\boldsymbol{w}$ represents neural network weighting parameters.

$\tilde{\boldsymbol{r}}_k$ represent the vectors of original poles and residues for the kth geometrical sample, respectively, formulated as

$$\tilde{\boldsymbol{p}}_k = \left[\tilde{p}_1^{(k)} \quad \tilde{p}_2^{(k)} \quad \cdots \quad \tilde{p}_i^{(k)} \quad \cdots \quad \tilde{p}_N^{(k)}\right]^T \tag{1}$$

$$\tilde{\boldsymbol{r}}_k = \left[\tilde{r}_1^{(k)} \quad \tilde{r}_2^{(k)} \quad \cdots \quad \tilde{r}_i^{(k)} \quad \cdots \quad \tilde{r}_N^{(k)}\right]^T \tag{2}$$

where $\tilde{p}_i^{(k)}$ and $\tilde{r}_i^{(k)}$ represent the ith pole and residue for the kth geometrical sample, respectively, and N is the order of transfer function with the same order after pole–residue tracking technique [26] for all the geometrical samples. Let T represent the index set for the pole/residue pair, that is, $T = \{1, 2, \ldots, N\}$. The pole/residues are sorted in pairs according to the absolute value of the imaginary part of the poles, since the imaginary parts of poles have high correspondence with the resonance frequencies for microwave filters [26].

Let $H^{(k)}$ represent the response of the transfer function for the kth geometrical sample, formulated as

$$H^{(k)}(f) = \sum_{i=1}^{N} \frac{\tilde{r}_i^{(k)}}{j2\pi f - \tilde{p}_i^{(k)}} \tag{3}$$

where j is the imaginary symbol and f is the frequency.

B. Proposed Smoothness-Discriminating Algorithm

After obtaining the original pole/residues, we propose a novel smoothness-discriminating algorithm to identify the pole/residue pairs whose transfer function response curves are flat and smooth that need to be re-fit by polynomial functions.

The proposed smoothness-discriminating algorithm begins with the derivation of the formulation to calculate the smoothness index, defined as s_i, of the transfer function response for the ith pole/residue pair. We propose to use the maximum absolute value of the second-order derivatives for the real part and imaginary part of the transfer function w.r.t. frequency in the frequency band of interest for calculating the smoothness index s_i. Let m represent the index of frequency samples inside the frequency band of interest, $m = 1, 2, \ldots, n_f$, where n_f is the total number of frequency samples. Let $\alpha_{i,m}^{(k)}$ and $\beta_{i,m}^{(k)}$ be defined as the second-order derivatives of the real and imaginary parts of the transfer function response w.r.t. frequency of the ith pole/residue pair at the mth frequency sample for the kth geometrical sample. According to (3), $\alpha_{i,m}^{(k)}$ and $\beta_{i,m}^{(k)}$ are obtained as

$$\alpha_{i,m}^{(k)} = \text{real}\left[\frac{-8\pi^2 \tilde{r}_i^{(k)}}{\left(j2\pi f_m - \tilde{p}_i^{(k)}\right)^3}\right] \tag{4}$$

$$\beta_{i,m}^{(k)} = \text{imag}\left[\frac{-8\pi^2 \tilde{r}_i^{(k)}}{\left(j2\pi f_m - \tilde{p}_i^{(k)}\right)^3}\right]. \tag{5}$$

Let $D_i^{(k)}$ be defined as the maximum absolute value of the second-order derivatives for the real part and imaginary part of the transfer function response w.r.t. frequency of the ith pole/residue pair for the kth geometrical sample among all the frequency samples, formulated as

$$D_i^{(k)} = \max_{m=1,2,\ldots,n_f} \left\{\left|\alpha_{i,m}^{(k)}\right|, \left|\beta_{i,m}^{(k)}\right|\right\}. \tag{6}$$

After calculating $D_i^{(k)}$, s_i is formulated to be the average square values of $D_i^{(k)}$ considering all the geometrical samples for the ith pole/residue pair as

$$s_i = \frac{1}{n_s} \sum_{k=1}^{n_s} \left[D_i^{(k)}\right]^2. \tag{7}$$

After s_i is calculated, the proposed smoothness-discriminating algorithm then exploits s_i to separate the pole/residue pairs into two sets representing smooth and nonsmooth function responses. A user-defined threshold ξ is used to judge whether s_i is small or large, that is, whether the transfer function curve is smooth or nonsmooth. If s_i is larger than the user-defined threshold ξ, which indicates

that the transfer function curve is nonsmooth, then the ith pole/residue pair remains in its transfer function format. We name this kind of pole/residue pair the remaining pole/residue pair. If s_i is smaller than ξ, which indicates that the transfer function curve is smooth, we re-fit the subtransfer function formulated using the ith pole/residue pair by a polynomial function. We name this kind of pole/residue pair the re-fitting pole/residue pair. Let T_p represent the index set containing the index of the re-fitting pole/residue pairs, formulated as

$$T_p = \{i \in T \mid s_i < \xi\}. \tag{8}$$

Let T_t represent the index set of the remaining pole/residue pairs to formulate transfer functions, calculated as

$$T_t = T \setminus T_p \tag{9}$$

where "\" denotes the difference between two sets. Let N_t represent the number of elements in T_t, that is, $N_t = |T_t|$.

After the smoothness discrimination and the pole/residue set separation using (4)–(9), the proposed smoothness-discriminating algorithm systematically re-formulates the vectors of the remaining poles and residues. Let $\boldsymbol{p}_k$ and $\boldsymbol{r}_k$ represent the vectors of the remaining poles and residues for the kth geometrical sample, respectively, formulated as

$$\boldsymbol{p}_k = \begin{bmatrix} p_1^{(k)} & p_2^{(k)} & \cdots & p_i^{(k)} & \cdots & p_{N_t}^{(k)} \end{bmatrix}^T \tag{10}$$

$$\boldsymbol{r}_k = \begin{bmatrix} r_1^{(k)} & r_2^{(k)} & \cdots & r_i^{(k)} & \cdots & r_{N_t}^{(k)} \end{bmatrix}^T \tag{11}$$

where $p_i^{(k)}$ and $r_i^{(k)}$ represent the ith remaining pole and residue for the kth geometrical sample, respectively, formulated as

$$p_i^{(k)} = \tilde{p}_{L_i}^{(k)} \tag{12}$$

$$r_i^{(k)} = \tilde{r}_{L_i}^{(k)} \tag{13}$$

where

$$L_i = \begin{cases} \min\limits_{l \in T_t} \{l\}, & i = 1 \\ \min\limits_{\substack{l \in T_t \\ l \notin \{L_1, \ldots, L_{i-1}\}}} \{l\}, & i \geq 2. \end{cases} \tag{14}$$

Based on the above calculation, the proposed smoothness-discriminating algorithm finally reformulates the response $H^{(k)}$ of the transfer function for the kth geometrical sample as

$$H^{(k)}(f) = H_p^{(k)}(f) + H_t^{(k)}(f) \tag{15}$$

where $H_p^{(k)}$ and $H_t^{(k)}$ represent the subtransfer functions formulated by the re-fitting pole/residue pairs and remaining pole/residue pairs, respectively, formulated as

$$H_p^{(k)}(f) = \sum_{i \in T_p} \frac{\tilde{r}_i^{(k)}}{j2\pi f - \tilde{p}_i^{(k)}} \tag{16}$$

$$H_t^{(k)}(f) = \sum_{i=1}^{N_t} \frac{r_i^{(k)}}{j2\pi f - p_i^{(k)}}. \tag{17}$$

Using the proposed smoothness-discriminating algorithm, the original transfer function is divided and formulated as a summation of two subtransfer functions, that is, a smooth subtransfer function $H_p^{(k)}$ and a nonsmooth subtransfer function $H_t^{(k)}$. The smooth subtransfer function $H_p^{(k)}$ will be further re-fit using a polynomial function, which is described in detail in Section III-C.

C. Proposed Polynomial Fitting for Representing Smooth Subtransfer Function Response

In this section, a novel polynomial fitting module is carried out to replace the smooth subtransfer function $H_p^{(k)}$. Let $G^{(k)}$ be defined as the polynomial function output versus frequency f to fit $H_p^{(k)}$, formulated as

$$G^{(k)}(f) = \sum_{i=0}^{N_p} c_i^{(k)} f^i \tag{18}$$

where N_p is the order of polynomial function, f^i represent the ith power of frequency f, and $c_i^{(k)}$ represents the ith complex polynomial coefficient for the kth geometrical sample. To obtain continuous polynomial coefficients corresponding to the change in geometrical variables, the same order of polynomial function is used for all the geometrical samples. To determine the best order for the polynomial function, we start with a relatively lower order of the polynomial function and perform the polynomial fitting. If the fitting accuracy does not satisfy the requirement, we increase the order of the polynomial function and re-perform the polynomial fitting iteratively until the required accuracy is satisfied. We initialize the order of the polynomial function to be $N_p = 2$. Let $\boldsymbol{c}_k$ represent the vectors containing all the polynomial coefficients $c_i^{(k)}$, that is,

$$\boldsymbol{c}_k = \begin{bmatrix} c_0^{(k)} & c_1^{(k)} & \cdots & c_i^{(k)} & \cdots & c_{N_p}^{(k)} \end{bmatrix}^T. \tag{19}$$

To perform the polynomial fitting to obtain the polynomial coefficients $\boldsymbol{c}_k$, we need to formulate a polynomial characteristic matrix w.r.t. frequency f, containing the ith power of the frequency, $i = 0, 1, 2, \ldots, N_p$, for all the frequency samples. Let $\boldsymbol{A}$ be defined as the characteristic matrix, formulated as

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix} \boldsymbol{a}_0 & \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_i & \cdots & \boldsymbol{a}_{N_p} \end{bmatrix} \\ quad &= \begin{bmatrix} 1 & f_1^1 & f_1^2 & \cdots & f_1^{N_p} \\ 1 & f_2^1 & f_2^2 & \cdots & f_2^{N_p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & f_{n_f}^1 & f_{n_f}^2 & \cdots & f_{n_f}^{N_p} \end{bmatrix} \end{aligned} \tag{20}$$

where $\boldsymbol{a}_i$ represents the vector containing the ith power of the frequency f for all the geometrical samples.

Let $\boldsymbol{b}_k$ represent a vector containing the smooth subtransfer function response $H_p^{(k)}(f)$ at all the frequency samples for the kth geometrical sample, formulated as

$$\boldsymbol{b}_k = \begin{bmatrix} H_p^{(k)}(f_1) & H_p^{(k)}(f_2) & \cdots & H_p^{(k)}(f_{n_f}) \end{bmatrix}^T. \tag{21}$$

The polynomial fitting for the kth geometrical sample is to calculate the polynomial coefficients $\boldsymbol{c}_k$ using the data of

the smooth subtransfer function response $\boldsymbol{b}_k$ by solving the following:

$$\boldsymbol{A}\boldsymbol{c}_k = \boldsymbol{b}_k. \tag{22}$$

Since $\boldsymbol{A}$ is not a square matrix, we decompose the matrix $\boldsymbol{A}$ into an orthogonal matrix denoted as $\boldsymbol{Q}$ and a nonsingular upper triangular matrix denoted as $\boldsymbol{R}$, that is, $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$, based on Gram–Schmidt orthogonalization method [30], to solve (22) to obtain $\boldsymbol{c}_k$.

To obtain the orthogonal matrix $\boldsymbol{Q}$, we process $\boldsymbol{Q}$ column by column until the orthogonal matrix is generated. The formulation to calculate $\boldsymbol{Q}$ is formulated as

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_i & \cdots & \boldsymbol{q}_{N_p} \end{bmatrix} \tag{23}$$

where

$$\boldsymbol{q}_i = \begin{cases} \dfrac{\boldsymbol{a}_i}{\|\boldsymbol{a}_i\|}, & i = 1 \\ \dfrac{\boldsymbol{a}_i - \sum_{l=1}^{i-1} \boldsymbol{q}_l^T \boldsymbol{a}_i \boldsymbol{q}_l}{\left\| \boldsymbol{a}_i - \sum_{l=1}^{i-1} \boldsymbol{q}_l^T \boldsymbol{a}_i \boldsymbol{q}_l \right\|}, & i \geq 2. \end{cases} \tag{24}$$

After the orthogonal matrix $\boldsymbol{Q}$ is obtained, the nonsingular upper triangular matrix $\boldsymbol{R}$ can be consequently calculated. Since $\boldsymbol{Q}$ is an orthogonal matrix, the inversion of $\boldsymbol{Q}$ equals the transpose of $\boldsymbol{Q}$, that is, $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^T$

$$\begin{aligned} \boldsymbol{R} &= \boldsymbol{Q}^T \boldsymbol{A} \\ &= \begin{pmatrix} \boldsymbol{q}_1^T \boldsymbol{a}_1 & \boldsymbol{q}_1^T \boldsymbol{a}_2 & \cdots & \cdots & \boldsymbol{q}_1^T \boldsymbol{a}_{N_p} \\ 0 & \boldsymbol{q}_2^T \boldsymbol{a}_2 & \cdots & \cdots & \boldsymbol{q}_2^T \boldsymbol{a}_{N_p} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \boldsymbol{q}_{N_p}^T \boldsymbol{a}_{N_p} \end{pmatrix}. \end{aligned} \tag{25}$$

By decomposing $\boldsymbol{A}$ matrix into the orthogonal matrix $\boldsymbol{Q}$ and the nonsingular upper triangular matrix $\boldsymbol{R}$. The polynomial coefficients $\boldsymbol{c}_k$ can be calculated as

$$\boldsymbol{c}_k = \boldsymbol{R}^{-1}\left(\boldsymbol{Q}^T \boldsymbol{b}_k\right). \tag{26}$$

Since $\boldsymbol{R}$ is a nonsingular upper triangular matrix, $\boldsymbol{c}_k$ calculated using (26) is more accurate than that directly calculated using (22). After obtaining the coefficients $\boldsymbol{c}_k$, we need to calculate the error between the fit polynomial function response and the original smooth subtransfer function response, denoted as e_{N_p}, to judge if the fitting is accurate enough. The calculation for the error e_{N_p} is formulated as

$$e_{N_p} = \max_{k \in \{1,2,\ldots,n_s\}} \left\{ \left\| \sum_{i=0}^{N_p} c_i^{(k)} f^i - \sum_{i \in T_p} \frac{\tilde{r}_i^{(k)}}{j2\pi f - \tilde{p}_i^{(k)}} \right\|^2 \right\}. \tag{27}$$

If the error is small enough, that is, $e_{N_p} \leq \eta$, where η is a user-defined threshold for the polynomial fitting, the order of polynomial function for the proposed model is set as $N_p^* = N_p$. The coefficients $\boldsymbol{c}_k = [c_0^{(k)}, c_1^{(k)}, .., c_{N_p^*}^{(k)}]$ of polynomial function for each geometrical sample can be used for the training of the proposed model. If $e_{N_p} > \eta$, we increase the order $N_p = N_p + 1$ and re-perform the process above (i.e., (18)–(27)). The polynomial fitting process of (18)–(27) is performed iteratively until $e_{N_p} \leq \eta$ is satisfied. After the iterative polynomial fitting process is finished, the overall model function response for the kth geometrical sample is formulated as

$$\begin{aligned} H^{(k)}(f) &= G^{(k)}(f) + H_t^{(k)}(f) \\ &= \sum_{i=0}^{N_p^*} c_i^{(k)} f^i + \sum_{i=1}^{N_t} \frac{r_i^{(k)}}{j2\pi f - p_i^{(k)}}. \end{aligned} \tag{28}$$

The order of polynomial-transfer function is calculated as $N_A = N_T + N_p^*$. Fig. 5 shows the flowchart of the proposed algorithm of smoothness-discriminating and polynomial fitting for neuro-PTF model development. The fit polynomial coefficients $\boldsymbol{c}_k$ together with the remaining pole/residues $\tilde{\boldsymbol{p}}_k$ and $\tilde{\boldsymbol{r}}_k$ are used for training the proposed neuro-PTF in the next section.

D. Two-Stage Training Process for Developing the Proposed Neuro-PTF Model

After the proposed polynomial fitting process, we convert the original smooth subtransfer function with discontinuous pole/residues into a polynomial function with continuous coefficients. The two-stage training process is proposed to develop the neuro-PTF model.

In the first-stage training, that is, the preliminary training [26], neural networks are trained to learn the relationships between polynomial-transfer-function parameters (i.e., polynomial coefficients, poles, and residues) and the geometrical parameters. Let the training data of neural network as $\{\boldsymbol{x}_k, \boldsymbol{c}_k\}$, $\{\boldsymbol{x}_k, \boldsymbol{p}_k\}$, and $\{\boldsymbol{x}_k, \boldsymbol{r}_k\}$, where $k \in \{1, 2, \ldots, n_s\}$. Let $\boldsymbol{c}_{\mathrm{NN}}$, $\boldsymbol{p}_{\mathrm{NN}}$, and $\boldsymbol{r}_{\mathrm{NN}}$ represent the outputs of the neural network for polynomial coefficients, poles, and residues, with the ith output as c_i, p_i, and r_i, respectively. Let $\boldsymbol{w}$ represent the neural network internal weight parameters, that is,

$$\boldsymbol{w} = [\boldsymbol{w}_c \quad \boldsymbol{w}_p \quad \boldsymbol{w}_r]^T \tag{29}$$

where $\boldsymbol{w}_c$, $\boldsymbol{w}_p$, and $\boldsymbol{w}_r$ represent the neural network internal weight parameters corresponding to $\boldsymbol{c}_{\mathrm{NN}}$, $\boldsymbol{p}_{\mathrm{NN}}$, and $\boldsymbol{r}_{\mathrm{NN}}$, respectively.

In the second stage of training, that is, the refinement training [26], the overall model consists of the polynomial-transfer function and the neural networks whose initial values are the optimal solutions from the preliminary training. Let y represent the output of the neuro-PTF model, which is formulated as

$$y(\boldsymbol{x}, \boldsymbol{w}, f) = \sum_{i=0}^{N_p^*} c_i(\boldsymbol{w}_c, \boldsymbol{x}) f^i + \sum_{i=1}^{N_t} \frac{r_i(\boldsymbol{w}_r, \boldsymbol{x})}{j2\pi f - p_i(\boldsymbol{w}_p, \boldsymbol{x})}. \tag{30}$$

The second stage of training is used to refine the accuracy of the overall neuro-PTF model. The training data for the second stage training is $\{\boldsymbol{x}_k, d_k\}$, where $k \in \{1, 2, \ldots, n_s\}$ and d_k represents the output of the EM data, that is, S-parameters from the EM simulation, at the kth training geometrical sample. The process of the second-stage training is to minimize the sum of errors between $y(\boldsymbol{x}_k)$ and d_k for all the training geometrical samples, by adjusting the neural network internal

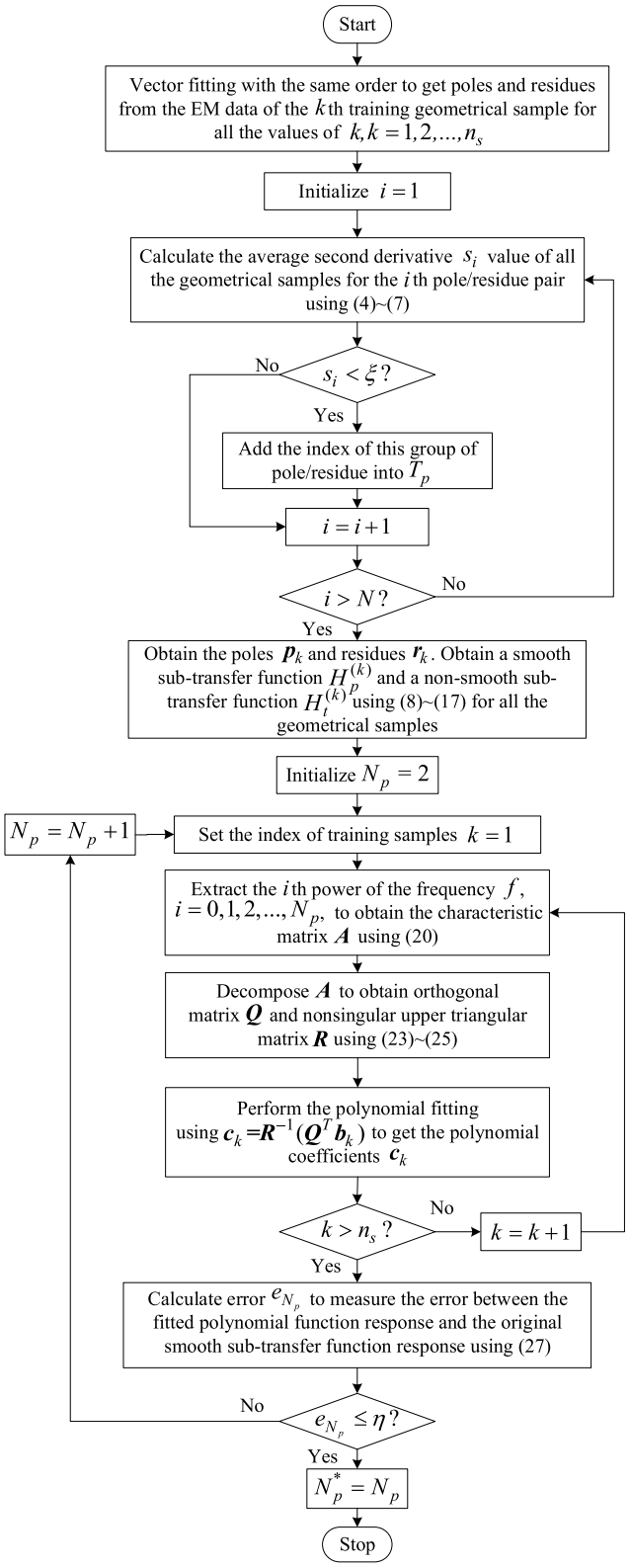


Fig. 5. Flowchart of the proposed algorithm of smoothness-discriminating and polynomial fitting for neuro-PTF model development.

weights w. The second-stage training process terminates when the training error between the neuro-PTF and the EM data is lower than a user-defined threshold ε. A new group of data samples that have never been used in the training process is used to test the trained model. If the test error is also lower than the user-defined threshold ε, the two-stage training process terminates. The trained neuro-PTF model can be used for high-level circuit and system designs. Otherwise, we adjust the number of hidden neurons and repeat the two-stage training process.

TABLE I

DEFINITION OF TRAINING AND TESTING DATA FOR THE FIFTH-ORDER HAIRPIN BANDPASS FILTER EXAMPLE

	Geometrical Variables	Training Data (81 Samples)			Testing Data (64 Samples)		
		Min	Max	Step	Min	Max	Step
Case 1 (Small)	w_1(mil)	50	51.6	0.2	50.1	51.5	0.2
	w_2(mil)	51	51.8	0.1	51.05	51.75	0.1
	s_1(mil)	8.8	9.2	0.05	8.825	9.175	0.05
	s_2(mil)	13.2	14	0.1	13.25	13.95	0.1
	l_1(mil)	77	77.4	0.05	77.025	77.375	0.05
	l_2(mil)	382.8	384.4	0.2	382.9	384.3	0.2
	l_3(mil)	99	99.8	0.1	99.05	99.75	0.1
	l_4(mil)	370.8	374.8	0.5	371.05	374.55	0.5
	l_5(mil)	48.6	49	0.05	48.625	48.975	0.05
Case 2 (Large)	w_1(mil)	50	56.4	0.8	50.4	56	0.8
	w_2(mil)	51	55.8	0.6	51.3	55.5	0.6
	s_1(mil)	8.8	10.4	0.2	8.9	10.3	0.2
	s_2(mil)	13.2	16.4	0.4	13.4	16.2	0.4
	l_1(mil)	77	78.6	0.2	77.1	78.5	0.25
	l_2(mil)	382.8	395.6	1.6	383.6	394.8	1.6
	l_3(mil)	99	105.4	0.8	99.4	105	0.8
	l_4(mil)	370.8	394.8	3	372.3	393.3	3
	l_5(mil)	48.6	50.2	0.2	48.7	50.1	0.2

TABLE II

COMPARISONS OF TRAINING AND TESTING ERRORS OF DIFFERENT MODELING METHODS FOR THE FIFTH-ORDER HAIRPIN BANDPASS FILTER EXAMPLE

	Training Methods	Coefficient (or Pole-Residue) Continuity?	Number of Hidden Neurons	Training Time	Mean-Square Training Error	Mean-Square Testing Error
Case 1 (Small)	Pole-Residue Neuro-TF	Good	8	2.16min	1.49%	1.87%
	Hybrid Neuro-TF	Good	8	2.51min	1.41%	1.74%
	Proposed Neuro-PTF	Good	8	2.35min	0.82%	1.41%
Case 2 (Large)	Pole-Residue Neuro-TF	Bad	8	3.08min	3.83%	3.69%
	Hybrid Neuro-TF	Bad	8	2.66min	3.17%	3.07%
	Proposed Neuro-PTF	Good	8	2.73min	1.16%	1.60%

IV. APPLICATION EXAMPLES

A. Parametric Modeling of Fifth-Order Hairpin Bandpass Filter

In the first example, we develop the proposed model for EM parametric modeling of a fifth-order hairpin bandpass

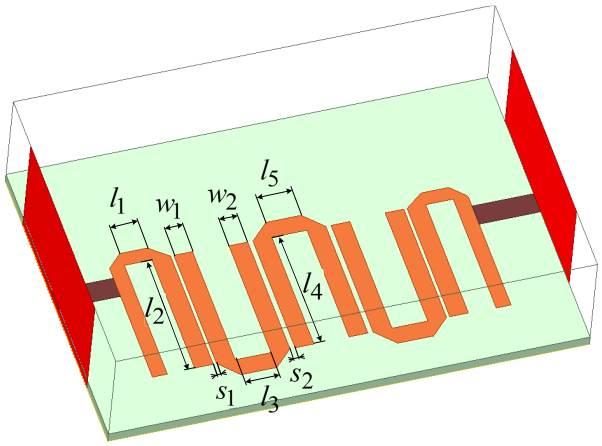


Fig. 6. 3-D configuration with nine geometrical parameters of the fifth-order hairpin bandpass filter example for EM simulation.

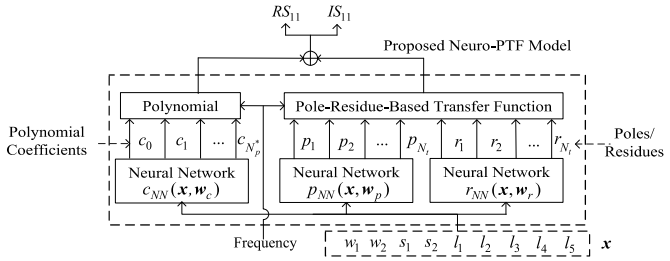


Fig. 7. Structure of the proposed neuro-PTF model for the fifth-order hairpin bandpass filter example.

filter [31] with the 3-D configuration shown in Fig. 6. S_1 and S_2 are the gaps between two groups of coupling lines. w_1 and w_2 are the widths of the coupling lines. l_1, l_2, l_3, l_4, and l_5 are the lengths of the coupling lines. This hairpin bandpass filter is a symmetrical structure. The bandwidth of the hairpin filter is about 1 GHz and the model frequency range is 2.5–5 GHz.

Fig. 7 shows the structure of the proposed neuro-PTF model, where geometrical parameters $\boldsymbol{x}$ and frequency f are the inputs. The input of the ANN module only contains geometrical parameters. The overall model contains both geometrical parameters and frequency as input variables. This proposed neuro-PTF model has nine input geometrical variables, that is, $\boldsymbol{x} = [w_1, w_2, s_1, s_2, l_1, l_2, l_3, l_4, l_5]^T$ (mil). The model has two outputs, which are the real and imaginary parts of the S_{11}, denoted as RS_{11} and IS_{11} respectively.

The full-wave EM simulation for generating training and testing data is performed using the *HFSS* software. Since this example has up to nine geometrical parameters to be solved, the design of experiments (DOE) [29] method with orthogonal distribution is used as the sampling method for both training and testing data. We use nine-level of DOE on defining training data, that is, a total of 81 samples of training data. Eight-level DOE is used to define 64 samples of testing data. The detailed geometrical range is given in Table I. Two cases of different geometrical ranges are used to demonstrate the advantage of the proposed method. The total CPU time of EM data generation for both cases is around 4.5 h. After generating data from EM simulation, the proposed technique is applied to develop the neuro-PTF model. The threshold ξ can be set to satisfy that s_i of a smooth curve is far smaller than ξ and the s_i of a nonsmooth curve is far larger than ξ. The threshold ξ is set to be 10 in this example. When frequency sampling is used to calculate the second-order derivatives, more sampling will improve accuracy. However, the accuracy will converge until the number of frequency samples becomes large enough. The number of frequency sampling to calculate the second-order derivatives is 251 in this example. *NeuroModelerPlus* software is used to train and test the neuro-PTF model. The overall model building time is around 5 min, including parameter extraction, algorithm operation, preliminary training, and refinement training.

To demonstrate the advantage of the proposed method, we compared the proposed method with the other two recent modeling methods, which are the pole–residue-based neuro-TF modeling method [26], the hybrid neuro-TF modeling method [28] using transfer functions in rational and pole/residue formats. Table II shows the comparison of the three methods. In Case 1, since the geometrical parameters vary within a smaller range, the discontinuity of the coefficients is not obvious. All three methods obtain better training accuracy. In Case 2, since the geometrical parameters vary within a large range, the pole–residue-based neuro-TF modeling method [26] has obvious discontinuity problems of nonunique parameter extraction which result in a high error. Compared with the pole–residue-based neuro-TF method, the hybrid neuro-TF modeling method [28] improves the continuity of the transfer function parameters to a certain extent by the transfer part of pole/residues to rational coefficients. However, the hybrid neuro-TF modeling method still does not fully solve the discontinuity problem of the transfer function parameters. The proposed method re-fit the discontinuous transfer function parameters with continuous polynomial coefficients to solve the discontinuity problem of transfer function parameters. Consequently, the proposed method obtains the best modeling accuracy among all the modeling methods in the comparison as shown in Table II.

To more intuitively compare the differences among the three different modeling methods, we show three different test geometrical samples #1, #2, and #3, which are shown in Fig. 8. The values of input geometrical variables for three different test geometrical samples are shown as follows.

Test Geometrical Sample #1:

$$\boldsymbol{x} = [50.4\ 51.3\ 8.9\ 13.4\ 77.1\ 383.6\ 99.4\ 372.3\ 48.7]^T \text{ (mil)}.$$

Test Geometrical Sample #2:

$$\boldsymbol{x} = [53.6\ 53.7\ 8.9\ 16.2\ 77.7\ 391.6\ 101.0\ 375.3\ 49.9]^T \text{ (mil)}.$$

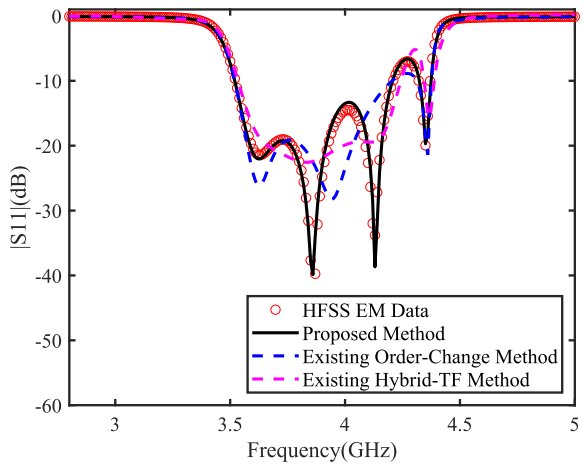
Test Geometrical Sample #3:

$$\boldsymbol{x} = [50.4\ 53.1\ 9.5\ 15.0\ 78.1\ 393.2\ 105.0\ 375.3\ 49.1]^T \text{ (mil)}.$$

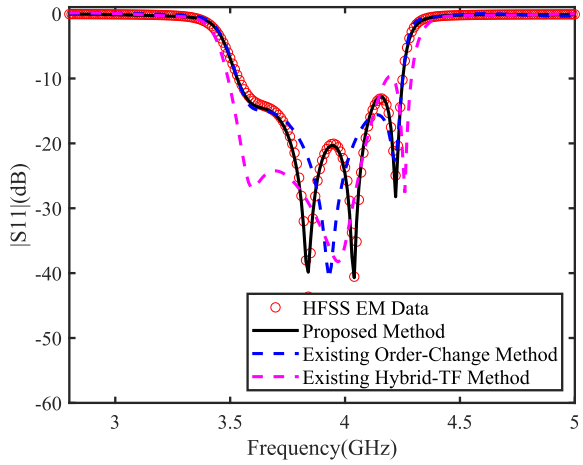
The above test geometrical samples are never used during the training process. From Fig. 8, it is observed that the proposed method can achieve better modeling accuracy than the existing methods.

B. Parametric Modeling of an Interdigital Bandpass Filter

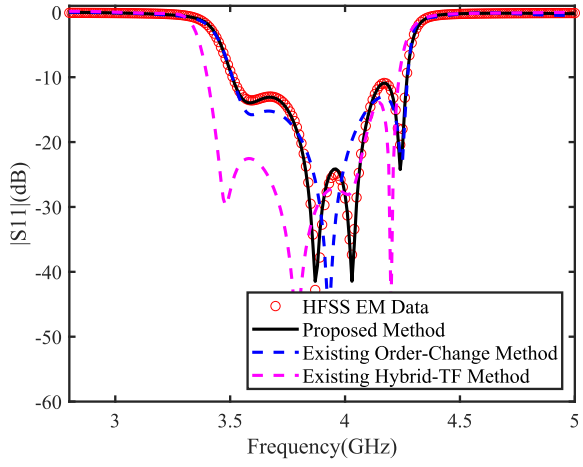
In the second example, we consider the parametric modeling of an interdigital bandpass filter example [32]. Fig. 9 shows the



(a)



(b)



(c)

Fig. 8. Comparison of the magnitude of S_{11} in decibels of the models developed using different modeling methods and HFSS EM data. (a) Test geometrical sample #1, (b) test geometrical sample #2, and (c) test geometrical sample #3 for the hairpin bandpass filter example.

3-D configuration of the interdigital bandpass filter example, where s_1, s_2, and s_3 are the distances between two adjacent resonant plates; w_1, w_2, and w_3 are the widths of the adjacent

TABLE III

DEFINITION OF TRAINING AND TESTING DATA FOR THE INTERDIGITAL BANDPASS FILTER EXAMPLE

	Geometrical Variables	Training Data (81 Samples)			Testing Data (64 Samples)		
		Min	Max	Step	Min	Max	Step
Case 1 (Small)	s_1(mm)	1.2	1.216	0.002	1.201	1.215	0.002
	s_2(mm)	2.4	2.432	0.004	2.402	2.43	0.004
	s_3(mm)	2.25	2.282	0.004	2.252	2.28	0.004
	l_1(mm)	42.5	42.82	0.04	42.52	42.8	0.04
	l_2(mm)	46.5	46.82	0.04	46.52	46.8	0.04
	l_3(mm)	46.5	46.82	0.04	46.52	46.8	0.04
	w_1(mm)	4.8	4.88	0.01	4.805	4.875	0.01
	w_2(mm)	4.9	4.98	0.01	4.905	4.975	0.01
	w_3(mm)	4.7	4.78	0.01	4.705	4.775	0.01
Case 2 (Large)	s_1(mm)	1.2	1.28	0.01	1.205	1.275	0.01
	s_2(mm)	2.4	2.56	0.02	2.41	2.55	0.02
	s_3(mm)	2.25	2.41	0.02	2.26	2.4	0.02
	l_1(mm)	42.5	44.1	0.2	42.6	44	0.2
	l_2(mm)	46.5	48.1	0.2	46.6	48	0.2
	l_3(mm)	46.5	48.1	0.2	46.6	48	0.2
	w_1(mm)	4.8	5.12	0.04	4.83	5.1	0.04
	w_2(mm)	4.9	5.22	0.04	4.92	5.2	0.04
	w_3(mm)	4.7	5.02	0.04	4.72	5	0.04

TABLE IV

COMPARISONS OF TRAINING AND TESTING ERRORS OF DIFFERENT MODELING METHODS FOR THE INTERDIGITAL BANDPASS FILTER EXAMPLE

	Training Methods	Coefficient (or Pole-Residue) Continuity?	Number of Hidden Neurons	Training Time	Mean-Square Training Error	Mean-Square Testing Error
Case 1 (Small)	Pole-Residue Neuro-TF	Good	8	1.86min	0.69%	0.76%
	Hybrid Neuro-TF	Good	10	1.58min	0.56%	0.62%
	Proposed Neuro-PTF	Good	8	1.78min	0.32%	0.33%
Case 2 (Large)	Pole-Residue Neuro-TF	Bad	10	3.33min	11.58%	11.6%
	Hybrid Neuro-TF	Bad	8	2.01min	3.05%	3.36%
	Proposed Neuro-PTF	Good	8	2.08min	1.35%	1.45%

resonant plates; l_1, l_2, and l_3 are the lengths of the adjacent resonant plates. The bandwidth of the interdigital filter is about 0.8 GHz and the model frequency range is 1–2.2 GHz.

The proposed neuro-PTF model for this example has nine geometrical variables, that is, $\boldsymbol{x} = [s_1\ s_2\ s_3\ w_1\ w_2\ w_3\ l_1\ l_2\ l_3]^T$ (mm), and frequency f as inputs and two outputs, which are the real part and imaginary part of S_{11}, denoted as RS_{11} and IS_{11}, respectively. The structure of the proposed neuro-PTF model is shown in Fig. 10.

The *HFSS* software is used to perform the full-wave EM simulation. The DOE method is also used to obtain the training data and testing data. Nine-level DOE is used to generate 81 training data and eight-level DOE is used to generate 64 testing data. The threshold ξ is set to be 10 in

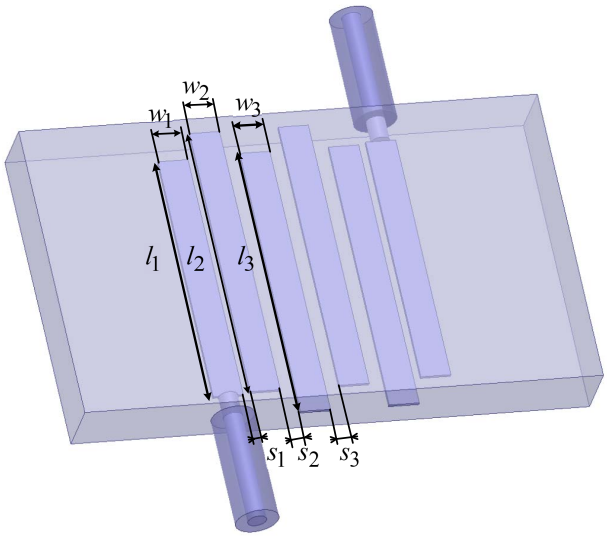


Fig. 9. 3-D configuration with nine geometrical parameters of the interdigital bandpass filter example for EM simulation.

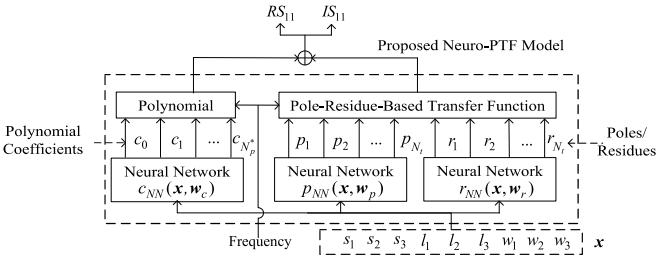


Fig. 10. Structure of the proposed neuro-PTF model for the interdigital bandpass filter example.

this example. The number of frequency sampling to calculate the second-order derivatives is 121. The proposed modeling method is applied to two different cases, that is, Case 1 with a narrower parameter range and Case 2 with a wider parameter range. The range and sweep steps of all the nine geometrical parameters for the two cases are defined in Table III. The total CPU time for generating EM data is around 5.5 h for both cases.

We carry out the proposed method to develop the parametric model. For comparison purposes, we also apply the pole–residue-based neuro-TF modeling method [26] and the hybrid neuro-TF modeling method [28] using transfer functions in rational and pole/residue formats. Table IV shows the comparison of the three methods. In Case 1, since the geometrical parameters vary within a smaller range, the discontinuity of the coefficients is not obvious. All three methods obtain better training accuracy. In Case 2, since the geometrical parameters vary within a large range, the pole–residue-based neuro-TF modeling method [26] has obvious discontinuity problems of nonunique parameter extraction which result in a high error. Compared with the pole–residue-based neuro-TF method, the hybrid neuro-TF modeling method [28] improves the continuity of the transfer function parameters to a certain extent by the transfer part of pole/residues to rational coefficients. However, the hybrid neuro-TF modeling method still does not fully solve the discontinuity problem of the transfer function parameters. The proposed method re-fit the discontinuous transfer function

TABLE V
DEFINITION OF TRAINING AND TESTING DATA FOR THE MINIATURIZED UWB FILTER EXAMPLE

	Geometrical Variables	Training Data (81 Samples)			Testing Data (64 Samples)		
		Min	Max	Step	Min	Max	Step
Case 1 (Small)	l_1(mm)	2.68	2.72	0.005	2.6825	2.7175	0.005
	l_2(mm)	4.1	4.14	0.005	4.1025	4.1375	0.005
	l_3(mm)	3.5	3.54	0.005	3.5025	3.5375	0.005
	w_g(mm)	0.28	0.288	0.001	0.2805	0.2875	0.001
	w_1(mm)	0.42	0.428	0.001	0.4205	0.4275	0.001
	w_2(mm)	0.18	0.188	0.001	0.1805	0.1875	0.001
	l_c(mm)	0.18	0.188	0.001	0.1805	0.1875	0.001
	w_c(mm)	0.35	0.358	0.001	0.3505	0.3575	0.001
Case 2 (Large)	l_1(mm)	2.68	2.84	0.02	2.69	2.83	0.02
	l_2(mm)	4.1	4.26	0.02	4.11	4.25	0.02
	l_3(mm)	3.5	3.66	0.02	3.51	3.65	0.02
	w_g(mm)	0.28	0.312	0.004	0.282	0.31	0.004
	w_1(mm)	0.42	0.452	0.004	0.422	0.45	0.004
	w_2(mm)	0.18	0.212	0.004	0.182	0.21	0.004
	l_c(mm)	0.18	0.212	0.004	0.182	0.21	0.004
	w_c(mm)	0.35	0.382	0.004	0.352	0.38	0.004

parameters with continuous polynomial coefficients to solve the discontinuity problem of transfer function parameters. Consequently, the proposed method obtains the best modeling accuracy among all the modeling methods in the comparison as shown in Table IV.

Fig. 11 shows the comparison of the three different modeling methods with the EM data at three test geometrical samples within the training range. The geometrical values of the three test geometrical samples are as follows.

Test Geometrical Sample #1:

$$\boldsymbol{x} = [1.20\ 2.40\ 2.25\ 42.5\ 46.5\ 46.5\ 4.80\ 4.90\ 4.70]^T \text{ (mm)}.$$

Test Geometrical Sample #2:

$$\boldsymbol{x} = [1.26\ 2.56\ 2.35\ 43.3\ 46.5\ 46.7\ 4.92\ 5.18\ 4.78]^T \text{ (mm)}.$$

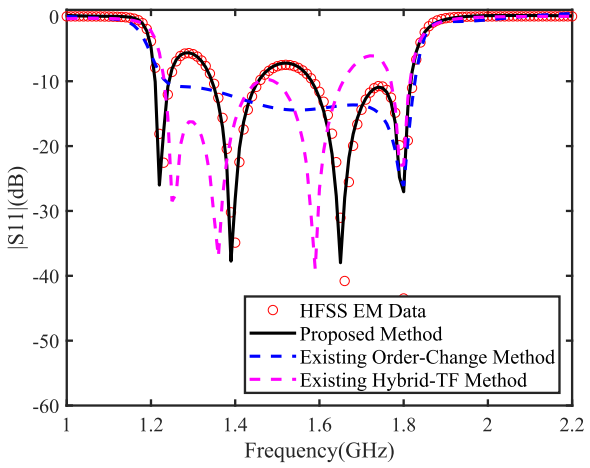
Test Geometrical Sample #3:

$$\boldsymbol{x} = [1.23\ 2.56\ 2.37\ 44.1\ 47.3\ 47.9\ 5.00\ 4.98\ 4.74]^T \text{ (mm)}.$$

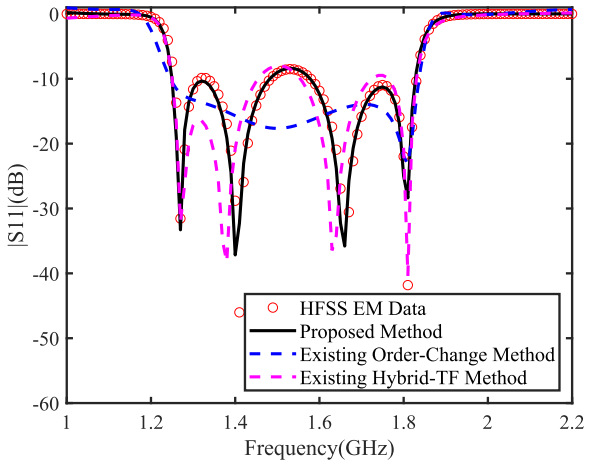
From Fig. 11, we can see that the proposed method can fit the EM data better than the other two methods in the comparison. This further proves the accuracy of the proposed method.

C. Parametric Modeling of a Miniaturized UWB Filter Using E-Shape Structures

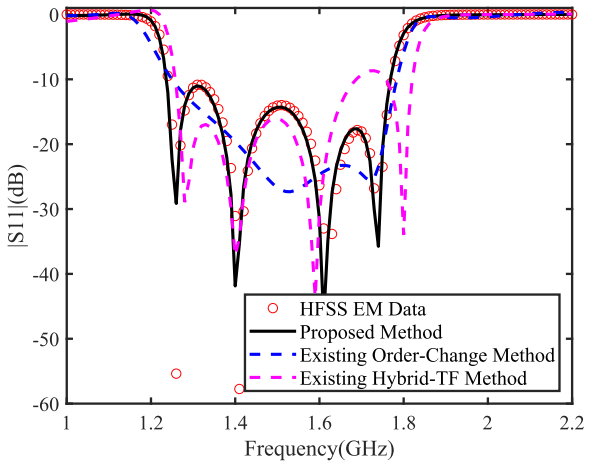
In the third example, we illustrate the proposed method using a miniaturized ultrawide-band (UWB) filter with E-shape structures [33], as shown in Fig. 12. l_1, l_2, and l_3 are the lengths of adjacent resonant plates. w_1 and w_2 represent the width of the adjacent resonant plates. l_c and w_c represent the width and length of the connected microchip line, respectively. w_g represents the distance from the centerline to both sides of each E-shape structure. The bandwidth of the UWB filter is about 8 GHz and the model frequency range is 1–13 GHz.



(a)



(b)



(c)

Fig. 11. Comparison of the magnitude of S_{11} in decibels of the models developed using different modeling methods and HFSS EM data. (a) Test geometrical sample #1, (b) test geometrical sample #2, and (c) test geometrical sample #3 for the interdigital bandpass filter example.

The proposed neuro-PTF model for this example has eight geometrical variables, that is, $\boldsymbol{x} = [l_1\ l_2\ l_2\ w_g\ w_1\ w_2\ l_c\ w_c]^T$ (mm), and frequency as inputs, and two outputs, which are the real part and imaginary part of S_{11}, denoted as RS_{11} and IS_{11}, respectively. The structure of the proposed model is shown in Fig. 13. Nine-level DOE is used to generate 81 training data and eight-level DOE is used to generate 64 testing data. The threshold ξ is set to be 10 in this example. The number of frequency sampling to calculate the second-order derivatives

TABLE VI

COMPARISONS OF TRAINING AND TESTING ERRORS OF DIFFERENT MODELING METHODS FOR THE MINIATURIZED UWB FILTER EXAMPLE

	Training Methods	Coefficient (or Pole-Residue) Continuity?	Number of Hidden Neurons	Training Time	Mean-Square Training Error	Mean-Square Testing Error
Case 1 (Small)	Pole-Residue Neuro-TF	Good	8	4.08min	0.41%	0.74%
	Hybrid Neuro-TF	Good	8	5.25min	0.5%	1.02%
	Proposed Neuro-PTF	Good	6	4.63min	0.37%	0.75%
Case 2 (Large)	Pole-Residue Neuro-TF	Bad	40	13.38min	9.11%	16.08%
	Hybrid Neuro-TF	Bad	10	6.42min	4.62%	4.38%
	Proposed Neuro-PTF	Good	10	7.50min	0.24%	0.26%

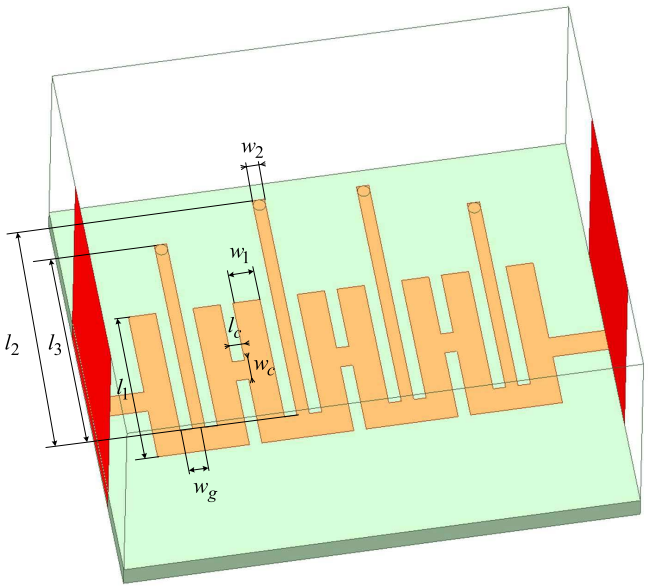


Fig. 12. 3-D configuration with eight geometrical parameters of the miniaturized UWB filter example for EM simulation.

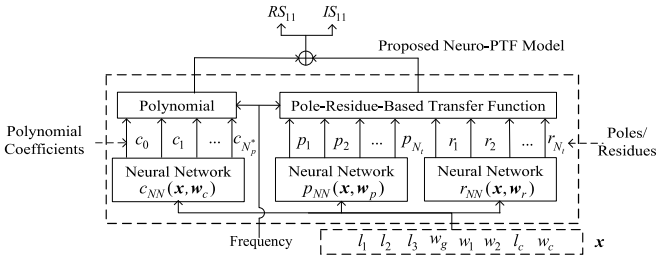
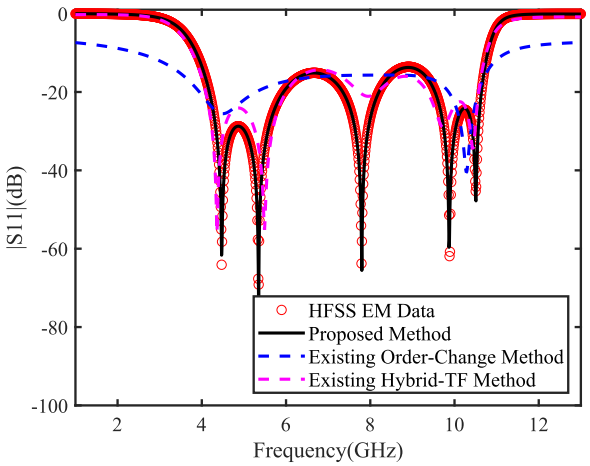
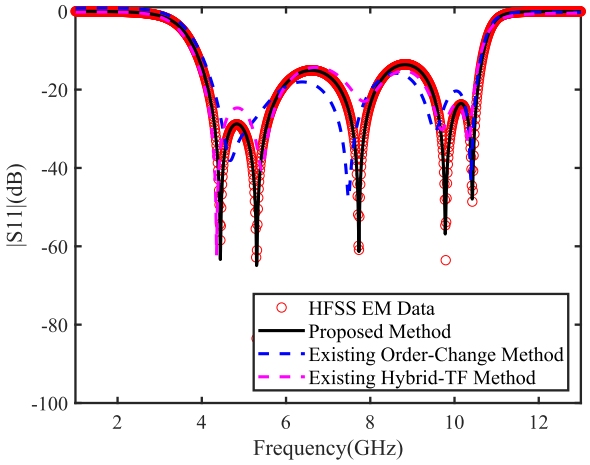


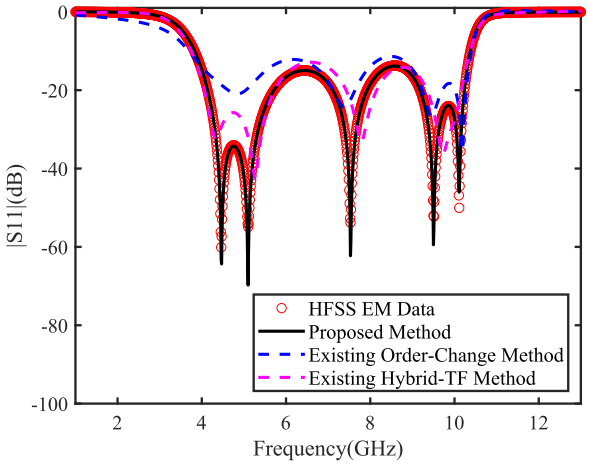
Fig. 13. Structure of the proposed neuro-PTF model for the miniaturized UWB filter example.



(a)



(b)



(c)

Fig. 14. Comparison of the magnitude of S_{11} in decibels of the models developed using different modeling methods and HFSS EM data. (a) Test geometrical sample #1, (b) test geometrical sample #2, and (c) test geometrical sample #3 for the UWB filter example.

is 1201. Table V shows the ranges of the eight geometrical parameters for two cases, that is, Case 1 with a narrower parameter range and Case 2 with a wider parameter range. The total CPU time for EM data generation is around 4 h for both cases.

The proposed modeling method is applied for both cases. We also apply the pole–residue-based neuro-TF modeling method [26] and the hybrid neuro-TF modeling method [28] using transfer functions in rational and pole/residue formats for comparison purposes. Table VI shows the comparison of the three methods. In Case 1, since the geometrical parameters vary within a smaller range, the discontinuity of the coefficients is not obvious. All three methods obtain better training accuracy. In Case 2, since the geometrical parameters vary within a large range, the pole–residue-based neuro-TF modeling method [26] has obvious discontinuity problems of nonunique parameter extraction which result in a high error. Compared with the pole–residue-based neuro-TF method, the hybrid neuro-TF modeling method [28] improves the continuity of the transfer function parameters to a certain extent by the transfer part of pole/residues to rational coefficients. However, the hybrid neuro-TF modeling method still does not fully solve the discontinuity problem of the transfer function parameters. The proposed method re-fit the discontinuous transfer function parameters with continuous polynomial coefficients to solve the discontinuity problem of transfer function parameters. Consequently, the proposed method obtains the best modeling accuracy among all the modeling methods in the comparison as shown in Table VI.

In Fig. 14, we compare the magnitude of S_{11} in decibels using the three different methods with EM data. Three test geometrical samples, which are never used in the training process, are shown in Fig. 14. The values of geometrical variables for the three test samples are as follows.

Test Geometrical Sample #1:

$$\boldsymbol{x} = [2.69\ 4.11\ 3.51\ 0.282\ 0.422\ 0.182\ 0.182\ 0.352]^T\ \text{(mm)}.$$

Test Geometrical Sample #2:

$$\boldsymbol{x} = [2.71\ 4.11\ 3.53\ 0.286\ 0.426\ 0.186\ 0.186\ 0.356]^T\ \text{(mm)}.$$

Test Geometrical Sample #3:

$$\boldsymbol{x} = [2.77\ 4.15\ 3.53\ 0.306\ 0.422\ 0.210\ 0.194\ 0.372]^T\ \text{(mm)}.$$

From Fig. 14, we can see that the proposed method has the best modeling accuracy among the three modeling methods to fit the EM data.

V. Conclusion

In this article, a neuro-PTF technique incorporating joint polynomial-transfer function with neural network for EM parametric modeling has been presented. In the proposed technique, the polynomial function has been introduced together with the pole–residue-based transfer function to represent the EM responses. A novel smoothness-discriminating algorithm has been proposed to judge the smoothness of each subtransfer function response and separate the pole/residue pairs whose subtransfer function response is determined to be smooth. The proposed method has introduced low nonlinear polynomial functions to re-fit the smooth parts of subresponse and left the nonsmooth parts of subresponse for the highly nonlinear

transfer functions to represent. By this way, the proposed method has avoided the discontinuity problems of nonunique parameter extraction caused by fitting smooth curves using the highly nonlinear transfer function. Utilizing both advantages of the polynomial functions and transfer functions, the proposed method has been demonstrated to produce more accurate models than the existing neuro-TF methods, especially with large geometrical variations. Once trained, the proposed model can provide a more accurate prediction of the EM response w.r.t. the geometrical variables. The developed models can be further used for high-level optimization with geometrical parameters.

REFERENCES

[1] Q. J. Zhang, K. C. Gupta, and V. K. Devabhaktuni, "Artificial neural networks for RF and microwave design-from theory to practice," *IEEE Trans. Microw. Theory Techn.*, vol. 51, no. 4, pp. 1339–1350, Apr. 2003.

[2] 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.

[3] V. Rizzoli, A. Costanzo, D. Masotti, A. Lipparini, and F. Mastri, "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.

[4] P. Burrascano, S. Fiori, and M. Mongiardo, "A review of artificial neural networks applications in microwave computer-aided design," *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] W. Na, W. Liu, L. Zhu, F. Feng, J. Ma, and Q.-J. Zhang, "Advanced extrapolation technique for neural-based microwave modeling and design," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 10, pp. 4397–4418, Oct. 2018.

[7] C. Zhang, J. Jin, W. Na, Q. J. Zhang, and M. Yu, "Multivalued neural network inverse modeling and applications to microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 8, pp. 3781–3797, Aug. 2018.

[8] 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.

[9] J. W. Bandler, M. A. Ismail, J. E. Rayas-Sánchez, and Q.-J. Zhang, "Neuromodeling of microwave circuits exploiting space-mapping technology," *IEEE Trans. Microw. Theory Techn.*, vol. 47, no. 12, pp. 2417–2427, Dec. 1999.

[10] J. E. Rayas-Sánchez and V. Gutiérrez-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.

[11] 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.

[12] H. Kabir, L. Zhang, M. Yu, P. H. Aaen, J. Wood, and Q. J. Zhang, "Smart modeling of microwave device," *IEEE Microw. Mag.*, vol. 11, no. 3, pp. 105–118, May 2010.

[13] J. W. Bandler *et al.*, "Space mapping: The state of the art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 337–361, Jan. 2004.

[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] 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.

[16] 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.

[17] F. Feng, W. Na, J. Jin, W. Zhang, and Q. J. Zhang, "ANNs for fast parameterized EM modeling," *IEEE Microw. Mag.*, vol. 22, no. 10, pp. 37–50, Oct. 2021.

[18] X. Ding *et al.*, "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.

[19] F. Feng *et al.*, "Multifeature-assisted neuro-transfer function surrogate-based EM optimization exploiting trust-region algorithms for microwave filter design," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 2, pp. 531–542, Feb. 2020.

[20] F. Feng, W. Na, W. Liu, S. Yan, L. Zhu, and Q.-J. Zhang, "Parallel gradient-based EM optimization for microwave components using adjoint-sensitivity-based neuro-transfer function surrogate," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 9, pp. 3606–3620, Sep. 2020.

[21] 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.

[22] Z. Guo, J. Gao, Y. Cao, and Q. J. Zhang, "Passivity enforcement for passive component modeling subject to variations of geometrical parameters using neural networks," in *IEEE MTT-S Int. Microw. Symp. Dig.* Montreal, QC, Canada, Jun. 2012, pp. 1–3.

[23] 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.

[24] P. Zhao and K.-L. Wu, "Model-based vector-fitting method for circuit model extraction of coupled-resonator diplexers," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 6, pp. 1787–1797, Jun. 2016.

[25] P. Zhao and K.-L. Wu, "Circuit model extraction of parallel-connected dual-passband coupled-resonator filters," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 2, pp. 822–830, Feb. 2018.

[26] 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.

[27] Y. Zhuo, F. Feng, J. Zhang, and Q.-J. Zhang, "Systematic order fitting algorithm in neuro-TF for parametric modeling of microwave components," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 6, pp. 487–490, Jun. 2022.

[28] Z. Zhao, F. Feng, W. Zhang, J. Zhang, J. Jin, and Q.-J. Zhang, "Parametric modeling of EM behavior of microwave components using combined neural networks and hybrid-based transfer functions," *IEEE Access*, vol. 8, pp. 93922–93938, 2020.

[29] S. R. Schmidt and R. G. Launsby, *Understanding Industrial Designed Experiments*. Colorado Springs, CO, USA: Air Force Academy, 1992.

[30] S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra*. Cambridge, U.K.: Cambridge Univ. Press, 2018.

[31] M. Yousefi, H. Aliakbarian, and R. Sadeghzadeh, "Design and integration of a high-order hairpin bandpass filter with a spurious suppression circuit," in *Proc. Loughborough Antennas Propag. Conf. (LAPC)*, Nov. 2015, pp. 1–4.

[32] F. Feng, W. Na, W. Liu, S. Yan, L. Zhu, and Q.-J. Zhang, "Parallel gradient-based EM optimization for microwave components using adjoint-sensitivity-based neuro-transfer function surrogate," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 9, pp. 3606–3620, Sep. 2020.

[33] R. T. Hammed and D. Mirshekar-Syahkal, "Multiple band rejection notches in miniaturized UWB fifth-order filter using E-shape microstrip structures," in *Proc. IEEE Radio Wireless Symp.*, Jan. 2013, pp. 94–96.

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

Yan Zhuo was born in Ningde, Fujian, China, in 2001. He is currently pursuing the B.Eng. degree at the School of Microelectronics, Tianjin University, Tianjin, China.

His current research interests include neural-network-based methods for microwave modeling and electromagnetic (EM)-based optimization of microwave structures.

Feng Feng (Member, IEEE) received the B.Eng. degree from Tianjin University, Tianjin, China, in 2012, and the Ph.D. degree from the School of Microelectronics, Tianjin University, and the Department of Electronics, Carleton University, Ottawa, ON, Canada, in 2017.

From 2017 to 2020, he was a Post-Doctoral Fellow with the Department of Electronics, Carleton University. He is currently an Associate Professor with the School of Microelectronics, Tianjin University. His research interests include electromagnetic parametric modeling and design optimization algorithms, deep neural network modeling method, space mapping algorithm and surrogate model optimization, the finite-element method in electromagnetic simulation and optimization, and multiphysics modeling and optimization.

Jianan Zhang (Member, IEEE) received the B.Eng. degree from Tianjin University, Tianjin, China, in 2013, and the Ph.D. degree from the School of Microelectronics, Tianjin University, and the Department of Electronics, Carleton University, Ottawa, ON, Canada, in 2020.

From 2020 to 2022, he was a Research Assistant with the Department of Electronics, Carleton University. He is currently an Associate Researcher with the School of Information Science and Engineering, Southeast University, Nanjing, China. He has authored or coauthored two book chapters and over 30 refereed journal articles and conference papers. His research interests include neural network-based electromagnetic (EM) parametric modeling and optimization, surrogate modeling and surrogate-assisted optimization, finite-element analysis in EM, and quantum computing with applications to EM problems.

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 McMaster 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 *Simulation-Driven Design Optimization and Modeling for Microwave Engineering* (Imperial College Press, 2013), 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 Coeditor of the Special Issue on High-Speed VLSI Interconnect for the *International Journal of Analog Integrated Circuits and Signal Processing* (Kluwer, 1994), and twice was a Guest Editor of 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 neural networks and optimization methods for high-speed/high-frequency circuit design.

Dr. Zhang is a Fellow of the 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.