©SHUTTERSTOCK.COM/NABUGU

MOR-Based Neuro-TF

Li Ma, Jingyi Feng, Jia Nan Zhang, and Qi-Jun Zhang

Neuro-transfer function (neuro-TF) approaches are useful instruments for electromagnetic (EM) parameterized modeling of microwave passive components [1], [2], [3], [4]. A vector fitting technique is commonly used to obtain the TF coefficients from the EM responses utilized in conventional neuro-TF approaches [5]. However, the use of this method can introduce an "order-changing" issue, where the TF orders vary across different regions of the design parameter space. In neuro-TF models, shifting the TF's order complicates the process of accurately training neural networks and eventually results in low modeling

Li Ma (li.ma@tju.edu.cn) and Jingyi Feng (sme_fjy@tju.edu.cn) are with the School of Microelectronics, Tianjin University, Tianjin 300072, China.
Jia Nan Zhang (jiananzhang@seu.edu.cn) is with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China.
Qi-Jun Zhang (qjz@doe.carleton.ca) is with Department of Electronics, Carleton University, Ottawa, ON K1S 5B6, Canada.

Digital Object Identifier 10.1109/MMM.2024.3486584
Date of current version: 13 January 2025

accuracy. Recently, an alternative neuro-TF method has been proposed, which employs a model-order reduction (MOR)-based TF [6], [7], [8], [9]. Instead of utilizing a vector fitting process, the MOR process acquires the poles/zeros (rational coefficients) during full-wave EM simulations. The MOR technique ensures a consistent TF order across different geometric parameter values. The compelling characteristics of MOR-based neuro-TF methods have garnered significant interest from the research community.

This article provides an extensive overview of the most recent developments in MOR-based neuro-TF techniques. In particular, the approach is examined in two formats: the rational format [9] and the pole/zero format [8]. The former technique provides a solution to the problem of "order-changing" in conventional vector fitting-based neuro-TF, but it still has problems with poles/zeros being mismatched. Thus, the latter format, an EM sensitivity-based approach, has been developed to guarantee the continuity of the poles/zeros across different geometric training samples to address the mismatch problem [8]. This MOR-based neuro-TF approach in rational format combines the MOR-based rational TF [9] with neural networks to address mismatched poles/zeros without EM sensitivity.

MOR-Based Neuro-TF in Pole/Zero Format

Model Structure

Figure 1 depicts the logic diagram of the MOR-based neuro-TF model in pole/zero format, where $\boldsymbol{x}$ is the geometric variable, y is the model output, and d is the EM training data. Neural networks and the TF in pole/zero format in the complex propagation space are the two main elements of the model. From full-wave EM simulations for the microwave component, the frequency responses (i.e., S-parameters) are used as its outputs. The inputs to the model consist of the geometric variables of the frequency and the EM structure. To minimize the discrepancies between the model outputs y and the desired outputs d at different values of $\boldsymbol{x}$, the model undergoes training by iteratively adjusting the weights of the neural network using a specific optimization algorithm.

Let us define ω as the angular frequency, $y(\omega)$ as the frequency response, and $\beta(\omega)$ as the propagation constant. Then, we can obtain $\gamma(\omega) = j\beta(\omega)$. By using the finite element method (FEM) [10] and the matrix Padé via Lanczos (MPVL) algorithm (one of the popular MOR methods) [11], the pole/zero-gain form of the TF $y(\omega)$ is as follows:

$$y(\omega) = \gamma(\omega) K \frac{\prod_{i=1}^{q-1} (\gamma(\omega) - \gamma_0 - z_i)}{\prod_{i=1}^{q} (\gamma(\omega) - \gamma_0 - p_i)} + c. \tag{1}$$

In the above formula, p_i are the poles, z_i are the zeros, c is a feed forward term, and K is the TF gain. The order q, which also corresponds to the number of poles, describes the reduced order models. γ_0 corresponds to the value of γ at the expansion point ω_0 in frequency.

Algorithm for Pole/Zero Matching Based on Sensitivity Analysis

To solve the mismatch issue, the continuation method [7] and the distance-based pole-matching method [6] can be used when geometric vary over a small/medium range. The sensitivity-analysis–based pole/zero-matching algorithm is an excellent choice for resolving the poles/zeros mismatch issue for significant geometric variations. This algorithm utilizes the EM sensitivities of zeros/poles at a specific value of a geometric parameter to estimate the updated locations of the poles/zeros at a different value of geometric parameters in multidimensional parameter spaces. The algorithm then minimizes the discrepancies between the predicted and original poles/zeros and aligns them.

To accomplish this in a given range, define $\boldsymbol{p}$ and $\boldsymbol{z}$ as all of the poles and zeros in $y(\omega)$. Let n represent the number of geometric samples. For every $\boldsymbol{x}^{(n)}$ of a set of geometric parameter values, $\boldsymbol{p}$ and $\boldsymbol{z}$ are solved by applying MPVL algorithm based on FEM. We obtain the expected locations of the poles, denoted as $\bar{\boldsymbol{p}}^{(l)}$, at $\boldsymbol{x}^{(l)}$ using the following procedure:

$$\bar{\boldsymbol{p}}^{(l)} = \boldsymbol{p}^{(k)} + \boldsymbol{J}_p^{(k)} \Delta \boldsymbol{x} \tag{2}$$

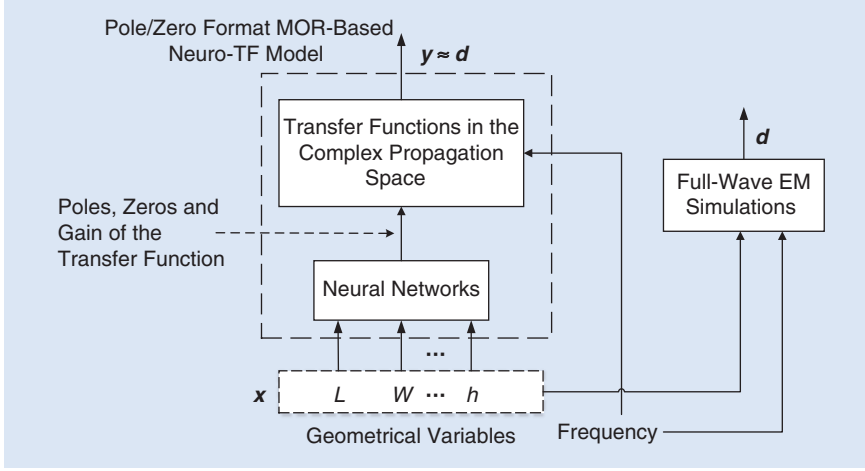

Figure 1. *Logic diagram of the MOR-based neuro-TF model in pole/zero format [8].*

where the formula $\Delta \boldsymbol{x} = \boldsymbol{x}^{(l)} - \boldsymbol{x}^{(k)}$ illustrates the shift in geometric parameters between $\boldsymbol{x}^{(k)}$ and $\boldsymbol{x}^{(l)}$. The poles' EM sensitivities at $\boldsymbol{x}^{(k)}$ are represented by $\boldsymbol{J}_p^{(k)}$. We can find suboptimal pole-matching solutions between neighboring geometric samples by using a matrix $\boldsymbol{D}$ that includes the distances at $\boldsymbol{x}^{(l)}$ between the poles' actual and predicted positions. Another algorithm can be used to generate a sequence of pairs from all of the geometric samples in the situation where two consecutive geometric samples are separated by a large distance in the geometric parameter space. This ensures that the poles are constantly inside one another's neighborhoods, addressing the problem effectively [8].

Two-Stage Training for the Model

The neuro-TF model goes through two stages of training. In the first stage of preliminary training, the neural networks are initially trained to discover the relationships between the geometric parameters and the poles, zeroes, and gains. It's crucial to remember that the training set for the zeros and poles has to be sorted before being utilized in this phase. The total model moves onto the stage of model refining after the neural networks have finished their preliminary training.

Then the MOR-based neuro-TF model, which is expressed in the pole/zero format, undergoes refinement through a dedicated training process. The objective of this refinement training is to fine-tune the previously trained neural network models, with the specific goal of minimizing the error in the S-parameters. Pairs $[\boldsymbol{x}^{(n)}, \boldsymbol{d}^{(n)}]$ where $n \in T_r$, make up the training data. Here T_r denotes the index set of all geometric parameter values. The EM responses (for example, S-parameters) are represented by $\boldsymbol{d}$ in this instance. The primary goal of the whole training procedure is to minimize the error function E_{tr} by fine-tuning the neural network's weights, $\boldsymbol{w}_p$, $\boldsymbol{w}_z$, and $\boldsymbol{w}_K$. To perform this optimization, a certain training procedure is used

$$E_{tr}(\boldsymbol{w}_p, \boldsymbol{w}_z, \boldsymbol{w}_K) = \frac{1}{2(N+1)} \times \sum_{n \in T_r} \sum_{i \in F} \left\| y(p_{NN}(\boldsymbol{x}^{(n)}, \boldsymbol{w}_p), z_{NN}(\boldsymbol{x}^{(n)}, \boldsymbol{w}_z), K_{NN}(\boldsymbol{x}^{(n)}, \boldsymbol{w}_K), \omega_i) - d_i^{(n)} \right\|^2. \quad (3)$$

In the above formula, F denotes the index set of all frequency samples, and i stands for the index value. When the training error falls below a predetermined threshold ε, as defined by the user, the training process concludes. After the refinement training process, a new group of data samples are selected to evaluate the performance of the trained model. If the measured error is likewise below the threshold ε, the refinement training process of the model terminates, and the developed model is ready for circuits design. If the testing error exceeds the predefined threshold, we change the number of hidden neurons in p_{NN}, z_{NN}, and K_{NN}, and then repeat both the preliminary and the refinement training processes. This iterative procedure continues until the target testing error threshold is attained.

MOR-Based Neuro-TF in Rational Format

The MOR-based neuro-TF method mentioned earlier effectively addresses the challenge of pole/zero mismatch. However, there are instances where obtaining or accessing EM sensitivity data is arduous or not possible. In such scenarios, the accuracy of neuro-TF modeling can be compromised. To overcome this limitation, this section explores the use of MOR-based neuro-TF in rational format, which resolves the issue of mismatch without relying on EM sensitivity data.

Model Structure

Figure 2 exhibits the MOR-based neuro-TF model composed of the rational TF and two neural networks. The inputs are $\boldsymbol{x}$ and frequency, and the output is the S-parameter of the microwave structure. Following the training phase, y represents the rational format output, which depends on γ, $\boldsymbol{x}$, and the weights of the neural network

$$y(\gamma, \boldsymbol{x}, \boldsymbol{w}_a, \boldsymbol{w}_b) = \gamma \frac{\sum_{i=1}^{q} a_i(\boldsymbol{x}, \boldsymbol{w}_a) \cdot (\gamma - \gamma_0)^{q-i}}{\sum_{i=1}^{q+1} b_i(\boldsymbol{x}, \boldsymbol{w}_b) \cdot (\gamma - \gamma_0)^{q+1-i}} - \delta \quad (4)$$

where γ_0 is the γ value at the solution frequency, and the Kronecker delta function is represented by δ. δ equals one when the input and output ports are the same port, while it equals zero when the input and output ports are different ports. In the outputs of the ith coefficient, $\boldsymbol{w}_a$ and $\boldsymbol{w}_b$ stand for the weights in the neural network related to the output values of a_i and b_i. A single neural network or separate neural networks can be used for learning the relationships of a_i and b_i with respect to

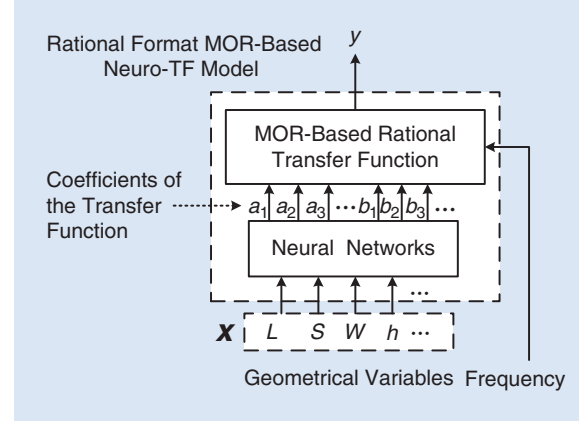

Figure 2. *Logic diagram of the MOR-based neuro-TF model in rational format [9].*

geometric variables. The detailed calculation processes of a_i and b_i are shown in the following section.

Format Transformation of the MOR-Based Neuro-TF

Consider a multiport microwave structure with the geometric parameters denoted by $\boldsymbol{x}$; the frequency response $S^{(n)}(\gamma)$ associated with a specific combination of input/output ports can be discovered [12]

$$S^{(n)}(\gamma) = \gamma\kappa\bar{\boldsymbol{\beta}}^T(\boldsymbol{K}_0^{(n)} + \gamma\boldsymbol{K}_1^{(n)} + \gamma^2\boldsymbol{K}_2^{(n)})^{-1}\boldsymbol{\beta} - \delta. \tag{5}$$

The finite element matrices in the formula above are denoted by $\boldsymbol{K}_0^{(n)}$, $\boldsymbol{K}_1^{(n)}$, and $\boldsymbol{K}_2^{(n)}$. The constant κ is based on the powers incident to the output ports with $\bar{\boldsymbol{\beta}}$ as EM excitation and to the input ports with $\boldsymbol{\beta}$ as EM excitation. With the MPVL technique, we can approximate the $S^{(n)}(\gamma)$ by a rapid frequency sweep [11]

$$S_q^{(n)}(\gamma) = \gamma\kappa\left[\boldsymbol{\eta}^{(n)}(\boldsymbol{I} - (\gamma - \gamma_0)\boldsymbol{T}_q^{(n)})^{-1}\boldsymbol{\rho}^{(n)}\right] - \delta. \tag{6}$$

In the above formula, q is the reduced-order model in the MPVL algorithm; the matrices $\boldsymbol{\rho}^{(n)} \in \mathbb{C}^{q\times 1}$, $\boldsymbol{T}_q^{(n)} \in \mathbb{C}^{q\times q}$ and $\boldsymbol{\eta}^{(n)} \in \mathbb{C}^{1\times q}$ are all acquired from MPVL. To solve the zero/pole mismatch, the S-parameter calculation in (6) is transformed into the rational format. By applying certain matrix operations to (6), we can derive a state-space representation of $S_q^{(n)}(\gamma)$ in the following manner:

$$S_q^{(n)}(\gamma) = \gamma[\mathbf{C}_q^{(n)}((\gamma - \gamma_0)\boldsymbol{I} - \boldsymbol{A}_q^{(n)})^{-1}\mathbf{B}_q^{(n)} + \mathbf{D}_q] - \delta \tag{7}$$

where $\mathbf{A}_q^{(n)} \in \mathbb{C}^{q\times q}$, $\mathbf{B}_q^{(n)} \in \mathbb{C}^{q\times 1}$, $\mathbf{C}_q^{(n)} \in \mathbb{C}^{1\times q}$, and $\mathbf{D}_q^{(n)} \in \mathbb{R}$ are referred to as the *state-space matrices* [13]. Once $\mathbf{A}_q^{(n)}$, $\mathbf{B}_q^{(n)}$, $\mathbf{C}_q^{(n)}$, and $\mathbf{D}_q$ are computed, we compute the results for the coefficients, marked as $a_i^{(n)}$ and $b_i^{(n)}$ $(n = 1, \ldots, N_s)$. As reported in [14], one can use the corresponding rational TF to represent the state-space calculated by (7).

Discussion

The MOR-based neuro-TF parameterized EM modeling techniques described in this article can provide efficient evaluations of EM responses with respect to the value changes of geometric parameters. Two MOR-based neuro-TF structures have been discussed in this article: one in pole/zero format and the other in rational format. When the order of the TF is relatively high, the sensitivity of the output of MOR-based neuro-TF in rational format with respect to rational TF coefficients is relatively high, which makes accomplishing high-accuracy model training difficult. In this situation, MOR-based neuro-TF in pole/zero format is more suitable, since the corresponding sensitivity is much lower. However, the pole/zeros for different geometric samples in the MOR-based neuro-TF do not have fixed sequences, resulting in a mismatch issue [8], especially when geometric ranges get larger. When the order of the TF is relatively low, the MOR-based neuro-TF in rational format is easier to use, since it does not need to deal with the complicated pole/zero-mismatch issue of the pole/zero format [9]. In the future, a mixed format MOR-based neuro-TF method may be developed that addresses the drawbacks of the existing formats of MOR-based neuro-TF methods.

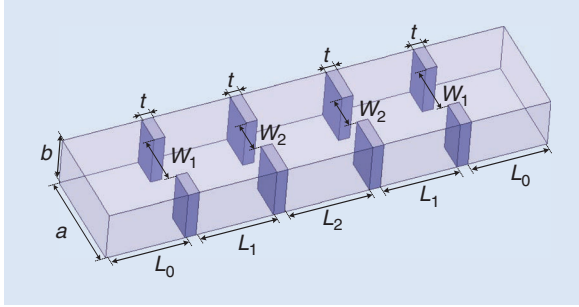

Figure 3. *EM model of the three-pole H-plane filter. $\boldsymbol{x} = [L_1 L_2 W_1 W_2]^T$ are the filter's four geometric parameters.*

Applications in EM Parameterized Modeling

Parameterized Modeling of a Three-Pole H-Plane Filter Using the MOR-Based Neuro-TF in Pole/Zero Format

Figure 3 illustrates the application of the MOR-based neuro-TF in pole/zero format using a three-pole *H*-plane filter [15], which consists of one output ($\boldsymbol{y} = |S_{11}|$) and four geometric variables [$\boldsymbol{x} = [L_1\ L_2\ W_1\ W_2]^T$ (mm)]. Among them, $a = 19.05$ mm, $b = 9.525$ mm, and $t = 2$ mm. In the MPVL method, it is suitable to set the reduced-order model order to $q = 9$. Each geometric parameter value has nine poles and eight zeros as a consequence. The sampling methodology used for the training and testing data are the design of experiments method [16].

We applied the pole/zero format MOR-based neuro-TF method to two samples (Table 1). In case 1, there are narrower geometric variations. The average training and measured errors for the proposed method are 0.16% and 0.18%, respectively, and rise to 0.62% and 1.56% in case 2 due to broader geometric variances. To capture the EM behavior of the filter in both scenarios, we additionally used three modeling techniques that are currently in use. In case 1, all methods achieved relatively low training and testing errors. This is due to the fact that the three pole-matching techniques correctly match the poles and zeros when the geometric changes are minimal. The mismatch pattern of poles and zeros, however, gets more intricate in case 2, where the geometric parameters change over a wider range. Only the pole/zero format MOR-based neuro-TF approach manages to maintain small training and testing errors. The model outputs from the four techniques at three distinct

test geometric parameter values are shown in Figure 4. The figure shows that, in comparison to the other three available approaches, the pole/zero format neuro-TF approach offers a closer fit to the EM data.

Parameterized Modeling of a Fifth-Order Waveguide Bandpass Filter Using the MOR-Based Neuro-TF in Rational Format

As seen in Figure 5(a), parameterized EM modeling of a fifth-order waveguide bandpass filter is performed using the MOR-based neuro-TF in rational format. Nine geometric variables, $x = [d_1 \ d_2 \ d_3 \ z_1 \ z_2 \ z_3 \ t_1 \ t_2 \ t_3]^T$, make up the neuro-TF model's inputs. Its output is $y = |S_{11}|$. A relatively large modeling range is considered. The rational format neuro-TF approach yields an average training error of 0.61% and a testing error of 0.78%. For comparison, parametric models of the filter are also built using the pole/zero format MOR-based neuro-TF approach and using refinement training in a distance-based approach. A comparison of the modeling accuracy attained by the three approaches is given in Table 2. Of the three approaches, the neuro-TF method with rational format has the lowest testing error. A comparison between the EM data and the outputs produced by the three neuro-TF models for a particular geometric testing sample is shown in Figure 5(b). When compared to the outputs of the other two neuro-TF models currently in use, the output of the rational format neuro-TF model clearly demonstrates a closer consensus with the EM data, as the figure makes clear. The developed MOR-based neuro-TF model is a set of analytical formulations that can obtain the output responses (accurate enough to represent the EM response) much faster than EM simulations.

TABLE 1. Comparison of modeling technique accuracy for the three-pole *H*-plane filter example.

	Modeling Method	Number of Hidden Neurons	Training Error	Testing Error
Case 1	Distance-based method	20	1.37%	1.66%
	Distance-based method with refinement training	10	0.16%	0.18%
	Continuation method with refinement training	10	0.16%	0.18%
	Pole/zero format MOR-based neuro-TF	10	0.16%	0.18%
Case 2	Distance-based method	20	35.8%	107.57%
		40	0.86%	81.91%
	Distance-based method with refinement training	10	1.94%	14.09%
		20	0.17%	26.42%
	Continuation method with refinement training	10	3.56%	12.83%
		20	1.16%	460.51%
	Pole/zero format MOR-based neuro-TF	10	0.62%	1.56%

Conclusion

MOR-based neuro-TF approaches have been extensively explored for efficient parameterized EM modeling, enabling rapid predictions of EM behaviors in microwave design scenarios involving repetitive changes in geometric parameters. These neuro-TF techniques can prevent the problem of order changes in conventional neuro-TF methods, achieving higher accuracy when applied to modeling the EM behavior of microwave passive components. The realm of neuro-TF approaches for rapid parameterized EM modeling presents vast prospects for

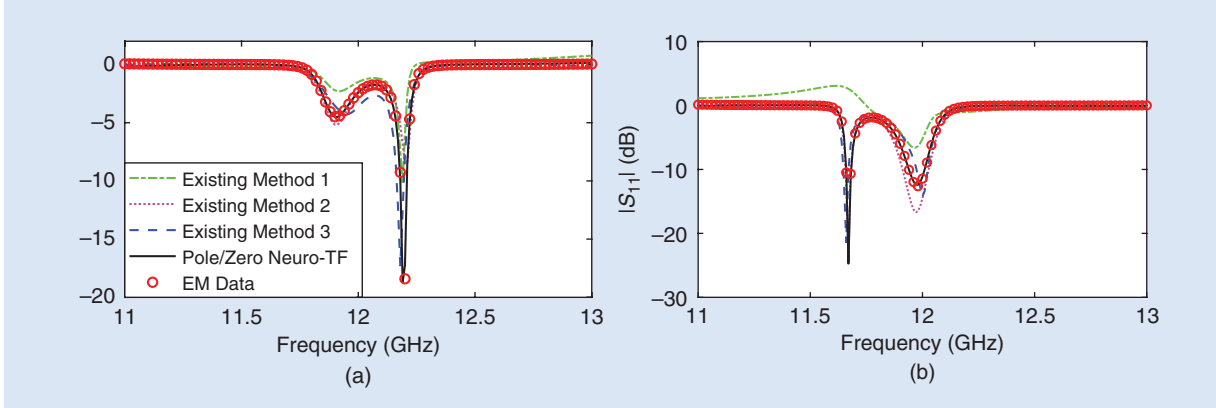

Figure 4. *Comparison between the EM data and the models' outputs for case 2 [8] (a) Test geometric parameter value #1 and (b) test geometric parameter value #2.*

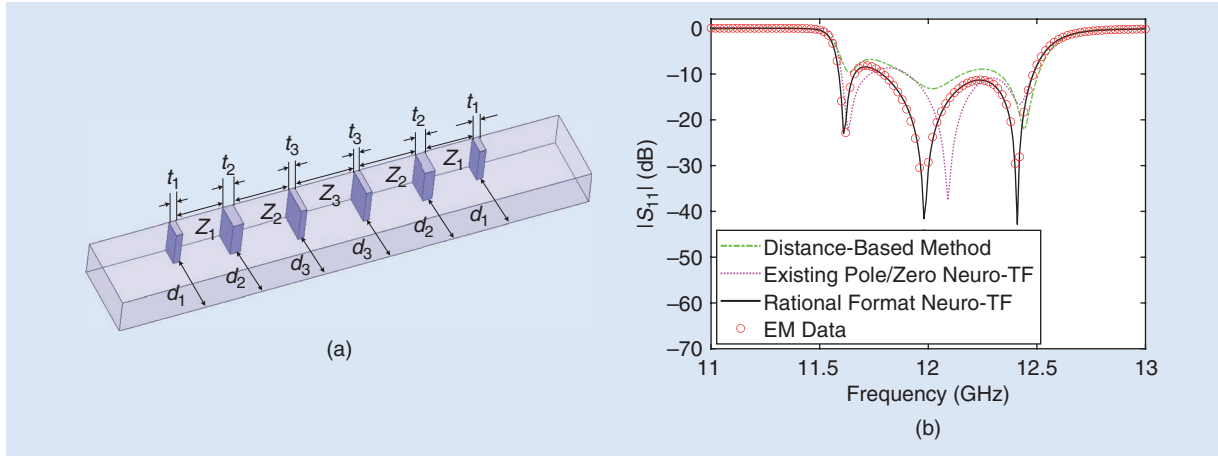

Figure 5. *(a) EM model of a fifth-order waveguide bandpass filter. (b) Comparison between the EM data and model outputs [9].*

TABLE 2. Comparison of modeling accuracy using three different approaches for the fifth-order waveguide bandpass filter example.

Modeling Method	TF Parameter Continuity?	Training Error	Testing Error
Distance-based method with refinement training	No	2.77%	13.61%
Pole-zero format neuro-TF method	No	1.97%	12.61%
Rational format neuro-TF method	Yes	0.61%	0.78%

future innovations and applications. These encompass advances in microwave-centric neuro-TF architectures, training algorithms, and novel application domains. Research into neuro-TF approaches, which mimic knowledge conceptualization and usage processes to improve the process of design, remains an ongoing and strategically significant direction.

References

[1] 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, doi: 10.1109/TMTT.2015.2504099.

[2] 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, doi: 10.1109/TMTT.2009.2032476.

[3] F. Feng, C. Zhang, W. Na, J. Zhang, W. Zhang, and Q. J. Zhang, "Adaptive feature zero assisted surrogate-based EM optimization for microwave filter design," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 1, pp. 2–4, Jan. 2019, doi: 10.1109/LMWC.2018.2884643.

[4] F. Feng, V. M. R. Gongal-Reddy, C. Zhang, J. Ma, and Q. J. Zhang, "Parametric modeling of microwave components using adjoint neural networks and pole-residue transfer functions with EM sensitivity analysis," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 6, pp. 1955–1975, Jun. 2017, doi: 10.1109/TMTT.2017.2650904.

[5] 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, doi: 10.1109/61.772353.

[6] Y. Yue, L. Feng, and P. Benner, "Interpolation of reduced-order models based on modal analysis," in *Proc. IEEE Int. Conf. Numer. Electromagn. Multiphys. Model. Optim. (NEMO)*, Reykjavik, Iceland, Aug. 2018, pp. 1–4, doi: 10.1109/NEMO.2018.8503114.

[7] Y. Yue, L. Feng, and P. Benner, "An adaptive method for interpolating reduced-order models based on matching and continuation of poles," in *Proc. IEEE Int. Conf. Numer. Electromagn. Multiphys. Model. Optim. (NEMO)*, Boston, MA, USA, May 2019, pp. 1–4, doi: 10.1109/NEMO43056.2019.9247028.

[8] J. Zhang, F. Feng, W. Zhang, J. Jin, 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.*, vol. 68, no. 6, pp. 2215–2233, Jun. 2020, doi: 10.1109/TMTT.2020.2979445.

[9] J. Zhang, J. Chen, Q. Guo, W. Liu, F. Feng, and Q. J. Zhang, "Parameterized modeling incorporating MOR-based rational transfer functions with neural networks for microwave components," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 5, pp. 379–382, May 2022, doi: 10.1109/LMWC.2022.3146376.

[10] J. M. Jin, *The Finite Element Method in Electromagnetics*. New York, NY, USA: Wiley, 2002.

[11] R. D. Slone and R. Lee, "Applying Padé via Lanczos to the finite element method for electromagnetic radiation problems," *Radio Sci.*, vol. 35, no. 2, pp. 331–340, Mar./Apr. 2000, doi: 10.1029/1999RS002258.

[12] F. Feng, J. Zhang, J. Jin, W. Zhang, Z. Zhao, and Q. J. Zhang, "Adjoint EM sensitivity analysis for fast frequency sweep using matrix Padé via Lanczos technique based on finite element method," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 5, pp. 2413–2428, Mar. 2021, doi: 10.1109/TMTT.2021.3061566.

[13] S. Grivet-Talocia and B. Gustavsen, *Passive Macromodeling: Theory and Applications*. Hoboken, NJ, USA: Wiley, 2015.

[14] D. Rowell, "State-space representation of LTI systems," MIT Department of Mechanical Engineering, Cambridge, MA, USA, Oct. 2002. [Online] Available: Http://web.mit.edu/2.14/www/Handouts/StateSpace.pdf

[15] S. Liao, H. Kabir, Y. Cao, J. Xu, Q. J. Zhang, and J. Ma, "Neural-network modeling for 3-D substructures based on spatial EM-field coupling in finite-element method," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 1, pp. 21–38, Jan. 2011, doi: 10.1109/TMTT.2010.2090405.

[16] D. C. Montgomery, *Design and Analysis of Experiments*. New York, NY, USA: Wiley, 1991.