Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

A Hybrid Modeling Method Based on Neural Networks and Its Application to Microwave Filter Tuning ⋆

Leyu Bi [*,**], **Weihua Cao** [*,**,†], **Wenkai Hu** [*,**], **Yan Yuan** [*,**], **Min Wu** [*,**]

[*] *School of Automation, China University of Geosciences, Wuhan, Hubei 430074, China (e-mail: wumin@ cug.edu.cn).*
[**] *Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan, Hubei 430074, China (e-mail: weihuacao@ cug.edu.cn).*

Abstract: In performance tuning of many electromechanical devices, well-trained operators are in great demand. However, manual tuning is costly and time-consuming, and thus do not conform to the trend of smart manufacturing. Microwave filters are typical electromechanical devices. Their tuning performance is limited by low extraction accuracy and high dimensionality of circuit features. In this paper, a hybrid modeling method based on neural networks is proposed to get better tuning performance. First, a curve-shape-based modeling method using Convolutional Neural Networks is presented to bypass the cumbersome extraction of circuit features. Second, a multi-model optimized fusion model based on Elman Neural Networks is constructed to cope with the high-dimensional property of circuit features, and improve modeling accuracy. The effectiveness of the hybrid modeling method is demonstrated through experiments. It achieves better tuning performance with fewer samples compared with two single modeling methods.

Keywords: Intelligent Manufacturing, Electromechanical Devices Tuning, Hybrid Modeling, Convolutional Neural Networks, Optimized fusion modeling.

1. INTRODUCTION

Performance tuning is critical to many electromechanical devices (Shahriari et al. 2018), to cope with design errors and manufacturing tolerances. For convenience of tuning, many tunable mechanical components are designed in the device structure to adjust device electromagnetic characteristics. In practical tuning processes, the tunable components are constantly adjusted by well-trained operators until all required performance indicators are met. However, manual tuning is costly and time-consuming, and thus does not conform to the trend of smart manufacturing in the era of Industry 4.0 (Ruiz-Sarmiento et al. 2020).

Microwave filters are typical electromechanical devices and widely deployed in filtering units of base stations. In the era of 5G, the number of base stations increases significantly, and the market share of microwave filters is boosting exponentially. Hence, automatic tuning is necessary to reduce costs. There exist extensive studies on tuning methods, and the model-based tuning method is more efficient(Jin et al. 2020). The outputs of the model are geometrical parameters of tunable components, and the inputs are the features reflecting device performance. Thus, based on the developed model, tuning information is directly obtained by inputting the features corresponding to the required performance. Therefore, features extraction and modeling method are critical to the model-based tuning method.

The analysis of the microwave filter is based on its equivalent circuit, and thus the feature extraction is mainly focused on the scattering matrix (S-matrix) and the admittance matrix (Y-matrix) in circuit theory (Bostani et al. 2017). The S-matrix is directly measured by the Vector Network Analyzer (VNA), describing the transmission response of the signal (Bachiller et al. 2007). The Y-matrix is transformed from S-matrix, reflecting the relationship between voltage and current of the equivalent circuit of the microwave filter (Cao et al. 2018a). The Y-matrix contains more internal mechanism characteristics, and thus is more suitable for the tuning of microwave filters. However, the Y-matrix is a high-dimensional complex matrix. Its features, namely Y-parameters, need to be further extracted. The relationship among S-matrix, Y-matrix, and Y-parameters is shown in Fig. 1. There are a series of approaches to extract Y-parameters from the Y-matrix,

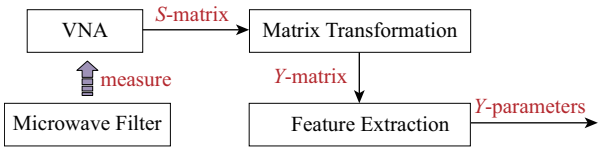

Fig. 1. Relationship among S-matrix, Y-matrix, and Y-parameters.

⋆ This work was supported by the National Natural Science Foundation of China under Grant 61773354 and the 111 project under Grant B17040.
† Corresponding author: Weihua Cao (weihuacao@ cug.edu.cn).

and show promising results when microwave filters are slightly detuned (Cao et al. 2018b). Unfortunately, when microwave filters are seriously detuned, Y-parameters are seriously attenuated, resulting in low extraction accuracy.

The tuning model is a multi-input-multi-output model with strong nonlinearity. On one hand, the essence of adjust tunable components is to change the electromagnetic field distribution in microwave filters, and thus the relationship between tunable components and Y-parameters is strongly nonlinear. On the other hand, the dimension of Y-parameters is $5N_{\mathrm{o}}$, where N_{o} is the order of filters and $N_{\mathrm{o}} > 5$ in general (Cao et al. 2018b). The tuning model is difficult to express by the analytical formula or function, and thus data-driven approaches are effective tools to develop the tuning model. Since tunable components are continuously adjusted, the modeling task is a regression problem. Artificial Neural Networks (ANNs) are extensively used in the modeling for their strong nonlinear mapping capabilities (Li et al. 2019). However, Y-parameters are high-dimensional. The single shallow ANNs cannot represent the high-dimensional relationship effectively. Thus, a deep neural network is introduced for the high-dimensional modeling (Jin et al. 2019). But more training samples are needed, limiting the tuning efficiency.

There are two problems for the model-based tuning method, namely, low extraction accuracy and high dimensionality of Y-parameters. When microwave filters are seriously detuned, the extracting accuracy of Y-parameters is low. Considering that the shape features of output response curves of microwave filters are utilized by operators in the tuning, a modeling method based on curve shape features is proposed. Specifically, shape features are automatically extracted by Convolutional Neural Networks (CNNs) (Dong et al. 2016), and thus get rid of the low-precision Y-parameters. When filters are slightly detuned, shape features cannot meet high-precision tuning requirements, because they do not characterize enough internal electromagnetic information. Nonetheless, the extraction precision of Y-parameter is sufficient, and thus the high-dimensional property of Y-parameters becomes the focus. In this paper, the Y-parameters are divided into multiple parts, and a multi-submodel optimized fusion modeling method is presented to further improve modeling accuracy. The main contributions of this paper are as follows:

- A hybrid modeling framework based on curve shape features and circuit features is created.
- A shape-based modeling method is proposed to cope with low extraction accuracy of Y-parameters.
- A multi-submodel optimized fusion modeling method is presented to deal with the high-dimensional property of Y-parameters.

The hybrid modeling method relates the current modeling difficulties to the filter resonance state, and balances tuning accuracy and efficiency.

This paper is structured as follows: In Section 2, the tuning process is abstracted, and the characteristics of output response curves are analyzed. In Section 3, the hybrid modeling method is proposed. In Section 4, the effectiveness of the method is demonstrated based on a simulation platform. Finally, in Section 5, concluding remarks and discussions are provided.

2. PROBLEM DESCRIPTION

This section abstractly defines a general tuning process. Then, the tuning process of microwave filters is presented and the characteristics of output response are analyzed.

A general tuning process can be described as

$$\boldsymbol{y}(k+1) = h\big(\boldsymbol{u}(k), \boldsymbol{y}(k)\big), \tag{1}$$

where k is the kth tuning step, $\boldsymbol{u}=[u_1,\ u_2,...,\ u_{N_v}]^{\mathrm{T}}$ is a vector representing the tuning action of tunable components, and $\boldsymbol{y}$ is the measurement output of tuned devices. The tuning method is utilized to decide the next tuning action $\boldsymbol{u}(k+1)$, making the device output move closer to the required performance indicators vector $\boldsymbol{y}_{\mathrm{d}}=[y_{\mathrm{d}1},\ y_{\mathrm{d}2},...,\ y_{\mathrm{d}N_\varsigma}]^{\mathrm{T}}$. N_v and N_ς are the number of tuning actions and performance indicators, respectively.

The tuning process of a microwave filter is shown in Fig. 2. The filter performance is adjusted by changing the lengths of tunable components in the metal shell, reflected by S-matrix measured by VNA. Thus, the tuning action $\boldsymbol{u}$ denotes the lengths of tunable components in the metal shell, and the output $\boldsymbol{y}$ in (1) is S-matrix given by

$$\begin{bmatrix} \boldsymbol{V}_1^- \\ \boldsymbol{V}_2^- \end{bmatrix} = \boldsymbol{S} \begin{bmatrix} \boldsymbol{V}_1^+ \\ \boldsymbol{V}_2^+ \end{bmatrix}, \tag{2}$$

$$\text{and } \boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{V}_1^-/\boldsymbol{V}_1^+ & \boldsymbol{V}_1^-/\boldsymbol{V}_2^+ \\ \boldsymbol{V}_2^-/\boldsymbol{V}_1^+ & \boldsymbol{V}_2^-/\boldsymbol{V}_2^+ \end{bmatrix}, \tag{3}$$

where $\boldsymbol{V}_1^+$ and $\boldsymbol{V}_1^-$ are the incident and reflected voltage singles of the input port in Fig. 2 ; $\boldsymbol{V}_2^+$ and $\boldsymbol{V}_2^-$ are output ports. The signal $\boldsymbol{V}_2^+$ indicates the incident voltage of the output port that is unloaded in the tuning process, $\boldsymbol{V}_2^+ == \boldsymbol{0}$. Therefore, vectors $\boldsymbol{S}_{12}$ and $\boldsymbol{S}_{22}$ are theoretically infinite, and thus S-matrix is simplified as $[\boldsymbol{S}_{11},\ \boldsymbol{S}_{21}]$; $\boldsymbol{S}_{11} = \boldsymbol{V}_1^-/\boldsymbol{V}_1^+$ and $\boldsymbol{S}_{21} = \boldsymbol{V}_2^-/\boldsymbol{V}_1^+$ represent the reflectivity and transmission characteristic of the input signal, respectively. The vectors $\boldsymbol{S}_{11}(\omega)$ and $\boldsymbol{S}_{21}(\omega)$ are complex vectors, where ω is the angular frequency. The amplitude-frequency response (AFR) of $\boldsymbol{S}_{11}(\omega)$ and $\boldsymbol{S}_{21}(\omega)$ visually show the frequency selection characteristics of the filter.

The AFR of a band-pass filter is shown in Fig. 3. The angular frequency is normalized, and the pass-band is transformed to [-1, 1] (Zhao and Wu 2018). Based on the AFR, the indicators $\boldsymbol{y}_{\mathrm{d}}=[y_{\mathrm{d}1},\ y_{\mathrm{d}2},\ y_{\mathrm{d}3}]^{\mathrm{T}}$ are defined.

(1) *Pass-band indicator*: The cutoff frequencies after normalization, denoted as f_{c1} and f_{c2}, should meet

$$y_{\mathrm{d}1} = |f_{c1} + 1| < \xi_1, \tag{4}$$

$$y_{\mathrm{d}2} = |f_{c2} - 1| < \xi_2, \tag{5}$$

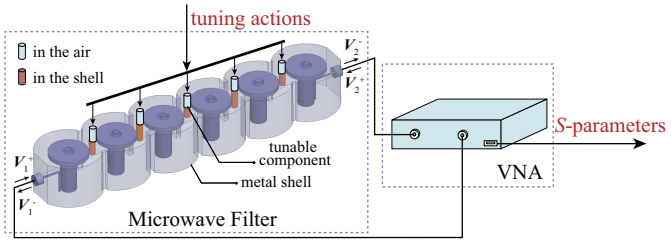

Fig. 2. Tuning process of a microwave filter. VNA is Vector Network Analyzer.

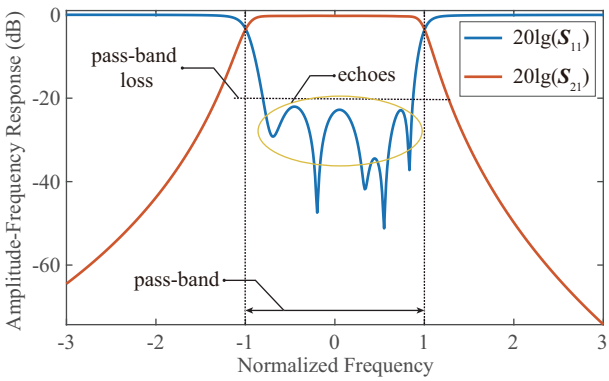

Fig. 3. AFR versus normalized frequency.

where $f_{c1} < f_{c2}$; ξ_1 and ξ_2 are selected by application requirements.

(2) *Pass-band loss indicator*: For the AFR of $\boldsymbol{S}_{11}$, larger echo amplitude comes greater signal loss in the pass-band. Thus, the maximum amplitude of echoes y_{d3} should meet

$$y_{\mathrm{d}3} < \xi_3, \tag{6}$$

which guarantees efficient signal transmission.

The satisfaction of above indicators depends on the model-based tuning method in this paper. The output of the tuning model is the tuning action $\boldsymbol{u}$. The inputs, Y-parameters or shape features of AFR, are determined by the resonance state of the microwave filter. Thus, the tuning model is a hybrid model, and the modeling method is presented in Section 3.

3. HYBRID MODEL CONSTRUCTION

Considering the different resonance states of microwave filters, a hybrid modeling method for microwave tuning is proposed. First, the hybrid modeling framework is presented. Second, the tuning model-1 based on shape features is established using CNNs. Third, the Y-parameters are extracted. Last, the tuning model-2 based on Y-parameters is built by optimized fusion modeling method.

3.1 Hybrid Modeling Framework

The hybrid modeling framework is shown in Fig. 4. $\mathcal{D}_1$ and $\mathcal{D}_2$ are datasets for modeling in two phases; $\mathcal{D}_1$ has N_ϱ samples including N_ϱ^1 training samples and N_ϱ^2 test samples; $\mathcal{D}_2$ has N_τ samples including N_τ^1 training samples and N_τ^2 test samples. Filters are tuned to a slight detuned state by $\boldsymbol{u}^1$, the output of the tuning model-1. $\mathcal{D}_2$ is composed of data obtained around $\boldsymbol{u}^1$. This ensures the extraction precision of Y-parameters. Filters are successfully tuned by $\boldsymbol{u}^2$, the output of the tuning model-2. Next, the modeling methods are shown in details.

3.2 Modeling Based on Response Curve Shapes

The AFR of S-matrix has rich shape features, such as undulations, peaks, and troughs. CNNs are end-to-end networks and can automatically extract features from raw data, without manual feature extraction. Based on CNNs, the shape features of the AFR of S-matrix are extracted,

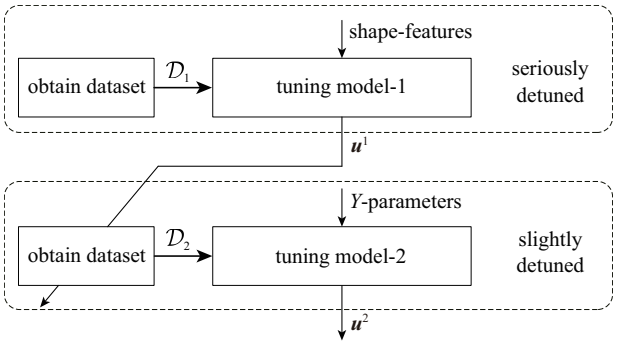

Fig. 4. Framework of the hybrid modeling method.

and the model between $\boldsymbol{u}$ and shape features is established. Thus, the problem that Y-parameters are difficult to extract accurately when filters are seriously detuned can be bypassed. CNNs are hierarchical networks that extract features by layer-by-layer abstraction, consisting of convolutional layers, activation layers, pooling layers, and fully-connected layers (Zou and Zhou 2019).

The structure of our model composed of three convolutional layers, three activation layers, two pooling layers, and three fully-connected layers. The Rectified Linear Unit (ReLU) is adopted as the activation function. Each convolutional layer is followed by the ReLU layer as depicted in Fig. 5. Considering the sparsity of the response curves, the kernel with larger size and fewer numbers is utilized in the first layer of convolution. Specifically, the first convolutional layer (Conv1) filters a 128×128 input image with 16 kernels of size 5×5 with a stride of 2 pixels. As the hierarchy deepens, the convolutional layers (Conv2 and Conv3) decrease kernel size and increase kernel numbers gradually. The first and second pooling layers (Pooling1 and Pooling2) down-sample the outputs of Conv1 and Conv2 by 2×2 kernel with a stride of 2 pixels. The output image size of Conv3 is already small, and thus does not require pooling. The last two fully-connected layers have 128 and 64 neurons, respectively. The output layer predicts the tuning action $\boldsymbol{u}$, resulting in N_v neurons.

After the model structure is completed, the model is trained by the random gradient descent method. The training objective function is defined by the L_2 norm as

$$\mathfrak{L}_{loss}^1 = \frac{1}{N_\varrho^1} \sum_{j=1}^{N_\varrho^1} \sum_{i=1}^{N_v} (l_i^j)^2, \tag{7}$$

where N_ϱ^1 is the sample number of the training set in $\mathcal{D}_1$; l_i^j is the residual error of the u_i of the jth sample as

$$l_i^j = u_i^j - \hat{u}_i^j, \tag{8}$$

where $\hat{u}_i^j$ is the predictive output of the model. The model is trained until

$$\mathfrak{L}_{loss}^1 \leq \varepsilon_1, \tag{9}$$

where ε_1 is the expected modeling error, or until the maximum number of training is reached.

After training, the model-1 predicts the tuning action $\boldsymbol{u}^1$ by inputting the response curve $\boldsymbol{I}_{\mathrm{d}}$ satisfying $\boldsymbol{y}_{\mathrm{d}}$. The model-1 has strong adaptability but low tuning accuracy. Thus, the model-1 works when filters are severely detuned, and filters are tuned to slightly detuned. In the slightly detuned state, Y-parameters have great impacts

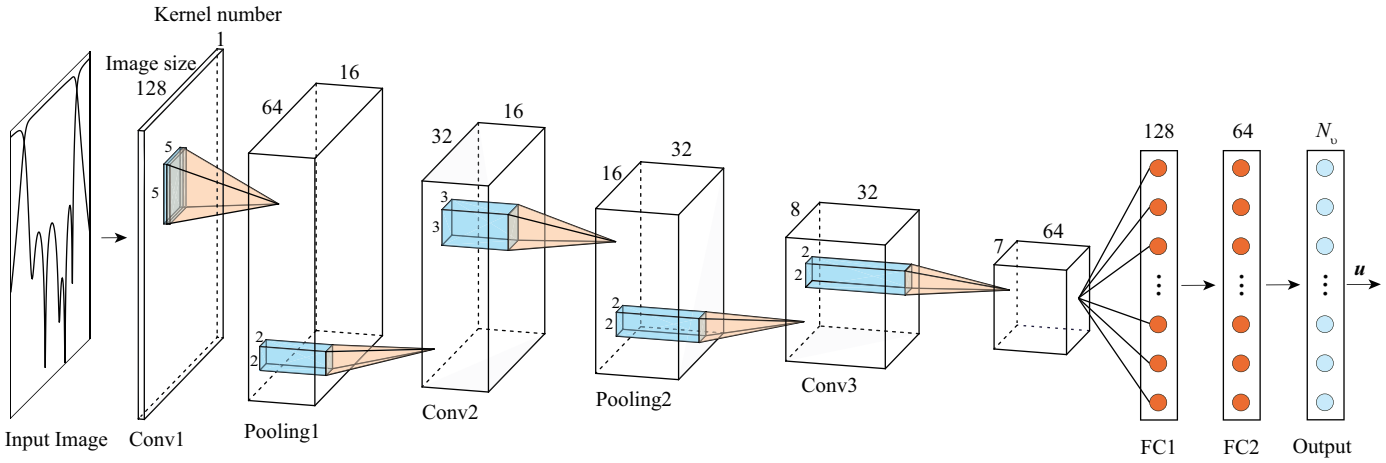

Fig. 5. Structure of the tuning model-1 using CNNs.

in high precision tuning. On one hand, they are extracted accurately; on the other hand, they reflect more tuning mechanism information. Hence, the tuning model based on Y-parameters is devised in the next subsection.

3.3 Extraction of Y-Parameters

The extraction of Y-parameters from S-matrix involves two steps as shown in Fig. 1. The Y-matrix is converted from the S-matrix (De Oliveira et al. 2013) as

$$\boldsymbol{Y}=\frac{\boldsymbol{S}_0-\boldsymbol{S}}{\boldsymbol{S}_0+\boldsymbol{S}}\frac{1}{z_0}, \tag{10}$$

where $\boldsymbol{S}_0$ has the same dimension with $\boldsymbol{S}$, and the elements in $\boldsymbol{S}_0$ are all 1's; $z_0 = 50\Omega$. The Y-matrix of the N_o-order filter is expressed as

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_{11} & \boldsymbol{Y}_{12} \\ \boldsymbol{Y}_{21} & \boldsymbol{Y}_{22} \end{bmatrix} = \sum_{k=1}^{N_o} \frac{1}{s-j\lambda_k} \begin{bmatrix} r_{11}^k & r_{12}^k \\ r_{21}^k & r_{22}^k \end{bmatrix}, \tag{11}$$

where $\boldsymbol{Y}_{11}, \boldsymbol{Y}_{12}, \boldsymbol{Y}_{21}$, and $\boldsymbol{Y}_{22} \in \mathbb{C}^{N_\sigma}$, and N_σ is the sample number; $\boldsymbol{s} = j\boldsymbol{\omega}$ where j is an imaginary unit and $\boldsymbol{\omega} \in \mathbb{R}^{N_\sigma}$ is a vector representing sample angle frequency; $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_{N_o}]^{\mathrm{T}}$ contains the poles of $\boldsymbol{Y}$; $\boldsymbol{r}_{11} = [r_{11}^1, r_{11}^2, ..., r_{11}^{N_o}]^{\mathrm{T}}$ is the residues of $\boldsymbol{Y}_{11}$. Therefore, the poles and residues represent the Y-matrix, and refer to as Y-parameters. Obviously, Y-parameters are high-dimensional with $5N_o$, where $N_o > 5$ in general.

The equivalent two-port network of microwave filters has reciprocity and symmetry. This leads to

$$\boldsymbol{Y}_{22}=\boldsymbol{Y}_{11},\ \boldsymbol{Y}_{12} = \boldsymbol{Y}_{21}. \tag{12}$$

Therefore, the Y-matrix can be simplified as $\boldsymbol{Y}=[\boldsymbol{Y}_{11}, \boldsymbol{Y}_{21}]$ and represented by $[\boldsymbol{\lambda}, \boldsymbol{r}_{11}, \boldsymbol{r}_{21}]$, and $\boldsymbol{\lambda}$, $\boldsymbol{r}_{11}$ and $\boldsymbol{r}_{21}$ are all complex numbers. Vector Fitting (VF) is a popular tool to extract the complex system poles and residues with the distinct impacts of speediness and accuracy. The Y-parameters are extracted by VF (Zhao and Wu 2018). The poles' real part and the residues' imaginary part are zero, and thus Y-parameters are further simplified as

$$\tilde{\boldsymbol{Y}} = [\mathrm{imag}(\boldsymbol{\lambda}),\ \mathrm{real}(\boldsymbol{r}_{11}),\ \mathrm{real}(\boldsymbol{r}_{21})], \tag{13}$$

where $\mathrm{real}(\cdot)$ and $\mathrm{imag}(\cdot)$ get the real and imaginary parts of a complex, respectively. The Y-parameters are eventually reduced to $3N_o$ dimensions. Next, the model between $\boldsymbol{u}$ and $\tilde{\boldsymbol{Y}}$ is constructed.

3.4 Fusion of Models Based on Y-Parameters

In the slightly detuned state, Y-parameters vary within a small neighborhood, but have strong nonlinear relation with $\boldsymbol{u}$. Elman Neural Networks (ENNs) can cope with strong nonlinear mapping, and utilize more information in the small neighborhood (Seker and Trkcan 2003). Thus, ENNs are suitable for the modeling.

Because the numerical distribution of $\mathrm{imag}(\boldsymbol{\lambda})$, $\mathrm{real}(\boldsymbol{r}_{11})$, and $\mathrm{real}(\boldsymbol{r}_{21})$ are different, it is difficult to build the model between $\boldsymbol{u}$ and $\tilde{\boldsymbol{Y}}$ by one network. Therefore, the models between $\boldsymbol{u}$ and each part of $\tilde{\boldsymbol{Y}}$ are created separately and then fused. The framework is shown in Fig. 6. The three sub-models are trained by the gradient descent method, and the training objective is

$$\mathfrak{L}_{loss}^2 = \frac{1}{N_\tau^1}\sum_{j=1}^{N_\tau^1}\sum_{i=1}^{N_\upsilon}(l_i^j)^2, \tag{14}$$

where N_τ^1 is the sample number of the training set in $\mathcal{D}_2$. Each sub-model is trained until

$$\mathfrak{L}_{loss}^2 \le \varepsilon_2, \tag{15}$$

where ε_2 is the expected modeling error, or until the maximum number of training is reached.

Three sub-models after training are referred to as ENN1, ENN2, and ENN3. The outputs of three sub-models are $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{u}_3$. The final tuning action $\boldsymbol{u}$ is obtained as

$$\boldsymbol{u} = \alpha_1 \cdot \boldsymbol{u}_1 + \alpha_2 \cdot \boldsymbol{u}_2 + \alpha_3 \cdot \boldsymbol{u}_3, \tag{16}$$

where α_1, α_2, and α_3 are the model weights, and $\alpha_1+\alpha_2+\alpha_3 = 1$. Optimal weights are solved by Particle Swarm Optimization (PSO) that is a heuristic optimization algorithm for such continuous optimization problems (Shi and Eberhart 1998). The weights optimization is defined as

$$\begin{aligned} &\widehat{\boldsymbol{\alpha}} = arg\ \underset{\boldsymbol{\alpha}}{min}\ \frac{1}{N_\tau^2}\sum_{i=1}^{N_\tau^2}(\mathbf{g}_i\mathbf{g}_i^{\mathrm{T}}), \\ &\mathbf{g}_i = \alpha_1\boldsymbol{u}_{1i} + \alpha_2\boldsymbol{u}_{2i} + \alpha_3\boldsymbol{u}_{3i} - \boldsymbol{u}_i, \\ &\alpha_1 + \alpha_2 + \alpha_3 = 1, \end{aligned} \tag{17}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \alpha_3]$ and $\widehat{\boldsymbol{\alpha}}$ is optimal model weights; N_τ^2 is the number of test sample in $\mathcal{D}_2$. For the ith test sample, $\mathbf{g}_i$ is the test squared error, and $\boldsymbol{u}_{1i}$, $\boldsymbol{u}_{2i}$, and $\boldsymbol{u}_{3i}$ are predictive outputs; $\boldsymbol{u}_i$ is the expected output.

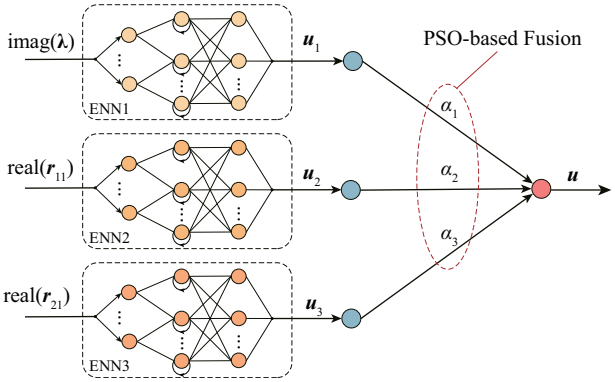

Fig. 6. Framework of the tuning model-2. Three sub-models are merged using PSO.

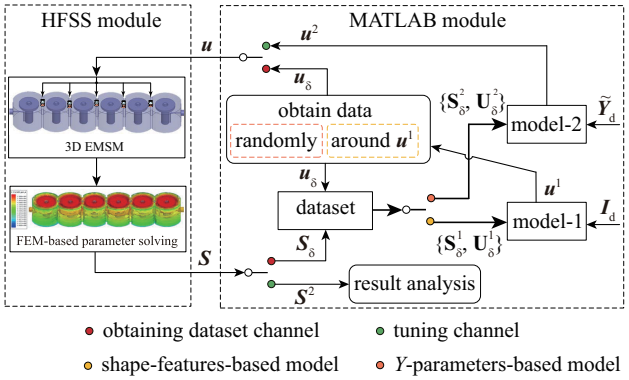

Fig. 7. Tuning process co-simulation platform. The hybrid modeling method is presented at the right hand side.

After solving $\widehat{\boldsymbol{\alpha}}$, the model is recorded as model-2, works when filters are slightly detuned, and predicts the tuning action $\boldsymbol{u}^2$ by inputting $\tilde{\boldsymbol{Y}}_{\mathrm{d}}$ satisfied $\boldsymbol{y}_{\mathrm{d}}$.

4. EXPERIMENTAL RESULTS AND ANALYSIS

This section builds a co-simulation tuning platform, and conducts experiments to demonstrate the effectiveness of the hybrid modeling method for tuning.

4.1 Experimental Set-up

The structure of the co-simulation platform composed of MATLAB and High-Frequency Structure Simulator (HFSS), as shown in Fig. 7. The three-dimensional Electromagnetic Simulation Model (EMSM) is designed in HFSS, and its S-matrix is accurately calculated by Finite Element Method (FEM). MATLAB changes the structure of EMSM, receives S-matrix, and solves tuning action $\boldsymbol{u}$.

In the MATLAB module, the dataset is composed of $(\boldsymbol{S}_\delta, \boldsymbol{u}_\delta)$. For the modeling of model-1, $\boldsymbol{u}_\delta$ is randomly changed for N_ϱ times within the allowable range, and the dataset $\mathcal{D}_1 = \{\mathbf{S}_\delta^1, \mathbf{U}_\delta^1\}$ is obtained. The filter is tuned from severely to slightly detuned based on $\boldsymbol{u}^1$. For the modeling of model-2, $\boldsymbol{u}_\delta$ is randomly changed for N_τ times in the neighborhood around $\boldsymbol{u}^1$, and the dataset $\mathcal{D}_2 = \{\mathbf{S}_\delta^2, \mathbf{U}_\delta^2\}$ is obtained. The filter is successfully tuned based on $\boldsymbol{u}^2$.

4.2 Experimental Results

With reference to indicators in practical application, pass-band indicators ξ_1 and ξ_2 are set to 0.01, and loss indicator ξ_3 is set to -20dB. The S-matrix $\boldsymbol{S}_{\mathrm{d}}$ satisfied $\boldsymbol{y}_{\mathrm{d}}$ is solved based on Chebyshev synthesis method (Cameron 1999). Then, $\boldsymbol{I}_{\mathrm{d}}$ and $\tilde{\boldsymbol{Y}}_{\mathrm{d}}$ are obtained based on the $\boldsymbol{S}_{\mathrm{d}}$.

For the model-1, the samples number N_ϱ is 200. After 100 training iterations, the loss function $\mathfrak{L}_{loss}^1$ converges to 0.2. The $\boldsymbol{I}_{\mathrm{d}}$ is the tuning input of the model-1, and the tuning result is shown in Fig. 8. The tuning response is close to $\boldsymbol{y}_{\mathrm{d}}$, but the loss of first echo is -18.3dB which is unsatisfied. The tuning response has the same variation trends with the expected response satisfied $\boldsymbol{y}_{\mathrm{d}}$. This indicates that shape features of response curve are useful for the tuning process. The effectiveness and flexibility of the shape-features-based modeling method is proven.

For the model-2, the dataset $\mathcal{D}_2$ varies randomly within 0.5 around $\boldsymbol{u}^1$ and N_τ=30. After 3000 training epoches, the loss functions $\mathfrak{L}_{loss}^2$ of three sub-models converge to 0.003, 0.002 and 0.005, respectively. The $\tilde{\boldsymbol{Y}}_{\mathrm{d}}$ is the tuning input of three sub-models and the model-2, and the tuning results are shown in the Fig. 9. The tuning results based on any of the three sub-models do not satisfy $\boldsymbol{y}_{\mathrm{d}}$. Based on the fused model-2, all indicators are satisfied. Moreover, the fusion weights, α_1=0.25, α_2=0.30, and α_3=0.45, are positively correlated with tuning results of three sub-models. The rationality and accuracy of the optimized fusion modeling method is proven.

4.3 Performance Comparison

The hybrid modeling method is compared with other two single modeling methods with only shape-features and only Y-parameters. For the two compared methods, the sample number increases to 400 to enhance contrast. The tuning results are shown in Fig. 10. On one hand, the echo loss of the only Y-parameters is worse than the model-1, indicating that Y-parameters with insufficient precision weaken modeling accuracy. On the other hand, the echo loss of the only shape-features is better than the model-1, indicating that increasing sample number improves the modeling accuracy for CNNs. But it still unsatisfied for the high-precision tuning. The hybrid method achieves excellent tuning results with fewer samples, and its effectiveness is demonstrated.

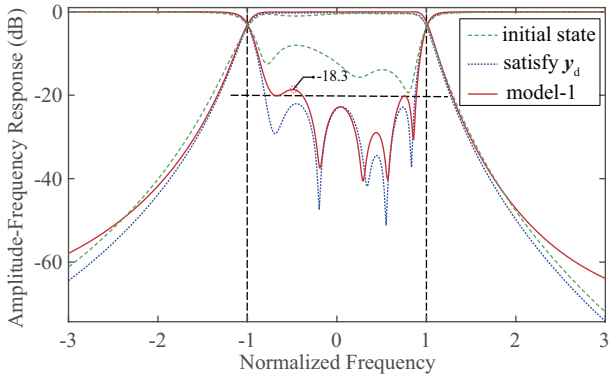

Fig. 8. Tuning result based on the model-1. The filter is tuned from severely to slightly detuned.

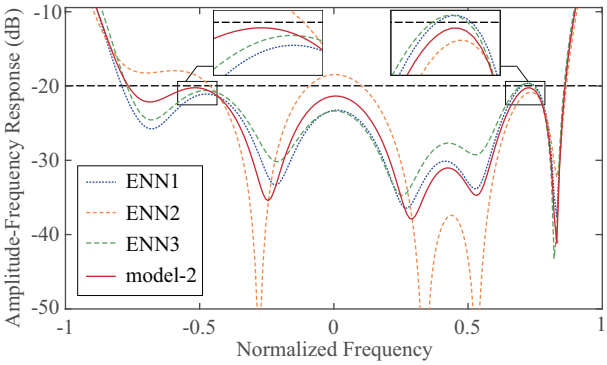

Fig. 9. Tuning result based on model-2. The filter is successfully tuned.

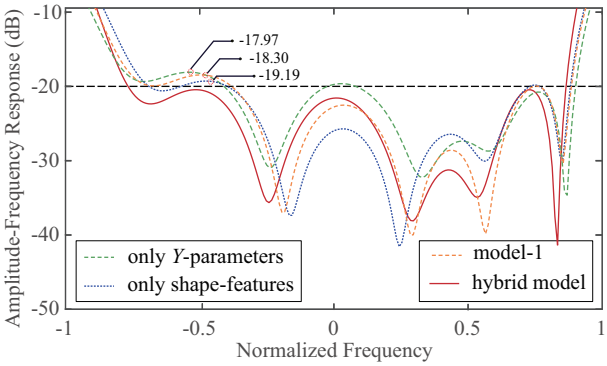

Fig. 10. Tuning results comparison. The hybrid model gets the best tuning performance.

5. CONCLUSION

The proposed hybrid modeling method reduces the precision requirement of the circuit features extraction in the tuning process, and thus has greater flexibility. The optimized fusion modeling method effectively solves the high-dimensional problem of circuit features and further improves the modeling accuracy. The proposed method achieves efficient tuning with less data in the simulation platform, and thus shows great potential in actual tuning processes. Further, there still remain many research opportunities for practical applications. Some possible problems deserving efforts include: how to adjust the tunable components to obtain the high-quality training dataset, how to mine and utilize more response curve information to facilitate modeling accuracy, how to implement a quick model update for similar tuning devices to reduce the training data and advance tuning efficiency; and how to evaluate and promote the tuning quality under tuning indicators are mutually constrained.

REFERENCES

Bachiller, C., Gonzalez, H., and Esbert, V. (2007). Efficient technique for the cascade connection of multiple two-port scattering matrices. *IEEE Transactions on Microwave Theory and Techniques*, 55(9), 1880–1886.

Bostani, R., Ardeshir, G., and Miar-Naimi, H. (2017). Analysis of millimeter-wave lc oscillators based on two-port network theory. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 64(3), 239–243.

Cameron, R.J. (1999). General coupling matrix synthesis methods for chebyshev filtering functions. *IEEE Transactions on Microwave Theory and Techniques*, 47(4), 433–442.

Cao, W.H., Liu, C., Yuan, Y., and Wu, M. (2018a). Extracting coupling matrix from lossy filters with unevenqs using differential evolution optimization technique. *International Journal of RF and Microwave Computer-Aided Engineering*, 28(6), e21269.

Cao, W.H., Liu, C., Yuan, Y., Wu, M., and Wu, S.B. (2018b). Parametric modeling of microwave filter using combined mls-svr and pole-residue-based transfer functions. *International Journal of RF and Microwave Computer-Aided Engineering*, 28(5), e21246.

De Oliveira, C.H.L., Artuzi, W.A., and Carvalho, C.A.T. (2013). An efficient technique to calculate the admittance and scattering parameters. In *2013 Brazilian Power Electronics Conference*, 1330–1333.

Dong, C., Loy, C.C., He, K., and Tang, X. (2016). Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2), 295–307.

Jin, J., Feng, F., Na, W., and Yan, S. (2020). Recent advances in neural network-based inverse modeling techniques for microwave applications. *International Journal of Numerical Modelling Electronic Networks Devices and Fields*. doi:10.1002/jnm.2732.

Jin, J., Zhang, C., Feng, F., and Zhang, Q. (2019). Deep neural network technique for high-dimensional microwave modeling and applications to parameter extraction of microwave filters. *IEEE Transactions on Microwave Theory and Techniques*, 67(10), 4140–4155.

Li, S., Wang, Y., Yu, M., and Panariello, A. (2019). Efficient modeling of ku-band high power dielectric resonator filter with applications of neural networks. *IEEE Transactions on Microwave Theory and Techniques*, 67(8), 3427–3435.

Ruiz-Sarmiento, J.R., Monroy, J., Moreno, F.A., Galindo, C., Bonelo, J.M., and Gonzalez-Jimenez, J. (2020). A predictive model for the maintenance of industrial machinery in the context of industry 4.0. *Engineering Applications of Artificial Intelligence*, 87, 103289.

Seker, S. and Trkcan, E. (2003). Elman's recurrent neural network applications to condition monitoring in nuclear power plant and rotating machinery. *Engineering Applications of Artificial Intelligence*, 16(7), 647 – 656.

Shahriari, Z., Leewe, R., Moallem, M., and Fong, K. (2018). Automated tuning of resonance frequency in an rf cavity resonator. *IEEE/ASME Transactions on Mechatronics*, 23(1), 311–320.

Shi, Y. and Eberhart, R. (1998). A modified particle swarm optimizer. In *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360)*, 69–73.

Zhao, P. and Wu, K. (2018). Circuit model extraction of parallel-connected dual-passband coupled-resonator filters. *IEEE Transactions on Microwave Theory and Techniques*, 66(2), 822–830.

Zou, Y. and Zhou, W. (2019). Automatic seam detection and tracking system for robots based on laser vision. *Mechatronics*, 63, 102261.