IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 November 2022, accepted 3 December 2022, date of publication 7 December 2022, date of current version 12 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3227327

RESEARCH ARTICLE

Approach for Extreme Learning Machine-Based Microwave Power Device Modeling

QIAN LIN[1], XIAO-ZHENG WANG[1], HAI-FENG WU[2], AND LI-NING JIA[1]
[1]School of Physics and Electronic Information Engineering, Qinghai Minzu University, Xining 810007, China
[2]Chengdu Ganide Technology, Chengdu 610000, China

Corresponding author: Hai-Feng Wu (abgott@tju.edu.cn; linqian@tju.edu.cn)

This work was supported in part by the National Natural Science Foundation under Grant 62161046, and in part by the Natural Science Foundation of Qinghai Province under Grant 2021-ZJ-910.

ABSTRACT The nonlinear representation of active devices plays an important role in microwave circuit design. Whereas, it takes a long time to extract a large amount of large signal data, and the problem of memory resource and CPU occupancy becomes significant. In order to address the problems in traditional large-signal modeling methods, in this paper an X-parameter modeling method for microwave power devices based on extreme learning machine (ELM) is proposed. To demonstrate the effectiveness of this method, a double layer back propagation (BP) neural network model is established. Then, harmonic balance simulations are used to verify the accuracy of these two models. After comparisons, it is proved that the three harmonic errors of double layer BP neural network model are 9.525dBm, 1.309dBm and 14.593dBm, respectively, and the three harmonic errors of ELM model are 0.673 dBm, 0.314 dBm, 3.09 dBm, respectively. Furthermore, the three harmonic modulus errors of double layer BP neural network model are 0.031, 0.002, 7.665e-4, respectively, and the errors of ELM model are 0.005, 0.001, 8.38e-5, respectively. Finally, in order to verify the accuracy of the predicted model in circuit design, the predicted X-parameter is used in the design of power amplifier. Moreover, the errors of the double layer BP neural network prediction model at 2.5 GHz, 5 GHz and 7.5 GHz are 1.142 dBm, 1.436 dBm and 2.294 dBm, respectively. The output power error of the ELM model at 2.5 GHz, 5 GHz and 7.5 GHz are 0.089 dBm, 0.311 dBm and 0.309 dBm, respectively. These experimental results demonstrate that the established ELM model is an efficient and valid approach for modeling GaN high electron mobility transistor types of nonlinear microwave devices.

INDEX TERMS BP neural networks, ELM, large signal modeling, microwave power device, transistor, X-parameters.

I. INTRODUCTION

With the rapid development of the communication field, the demands for pretty ideal quality and efficiency of communication have been put forward. Therefore, higher requirements for communication systems output power, efficiency are proposed. Meanwhile microwave power devices are the key components in communications systems and play a decisive role in their performance. Therefore, the model accuracy of microwave power devices is highly related to the design efficiency and performance of communication systems. Consequently, improving the modeling accuracy is a goal for every microwave engineer and the key challenge for radio frequency (RF) circuit design.

Given the high efficiency and accuracy of artificial neural networks (ANNs), they are widely used in microwave power device modeling [1]. In the field of microwave radio frequency, with the continuous increase of input power, the working area of the microwave power device has gradually changed from the linear region to the nonlinear region. Therefore, it is vital to define the nonlinear characteristics of microwave power devices as accurately as possible.

Studies have shown that the common modeling methods include empirical, X-parameter, and ANN models have been

The associate editor coordinating the review of this manuscript and approving it for publication was Giambattista Gruosso.

developed [2], [3]. Furthermore, because the large-signal model established by DC data cannot fully describe the nonlinear characteristics of transistors, the Curtice empirical model was proposed in 1980 [4]. Then, the large-signal characteristics of microwave power devices were fitted using polynomials. In 1996, Angelov model combined with CAD was proposed, which is suitable for parameter extraction [5], [6], [7]. However, the empirical model is easier to combine with the CAD software, which is widely used in circuit design. While the accuracy of the model is greatly affected by the formula [8]. Consequently, considering the influence of the conjugate signal, S-parameter was extended to X-parameter with higher accuracy in 2008 [9]. In 2014, a nonlinear vector network analyzer was utilized to characterize surface acoustic wave (SAW) filters [10]. Subsequently, an X-parameters modeling technique with high accuracy and fast convergence was proposed in 2014 [11]. In 2016, X-parameter was used to accurately predict the output power of a pseudomorphic high-electron-mobility transistor (pHEMT) under different loads by Lee [12]. In 2019, X-parameter was introduced in the sound field to process large signals with high nonlinearity [13]. Therefore, X-parameter can supplement the nonlinear theory and has become an essential approach for modeling of microwave power devices. The nonlinear characteristics of microwave power devices can be accurately described by X-parameters. Consequently, an accurate X-parameter model has great significance to research the nonlinear characteristics of microwave power devices.

At the same time, with the rapid evolution of ANNs such as multi-layer perceptron (MLP), back propagation (BP) neural network, recurrent neural network (RNN), and ELM are extensively used in the modeling of microwave power devices [14], [15], [16]. Due to high-speed and high-accuracy, ELM has become a critical tool for microwave power devices modeling [17], [18]. In 2017, Xiao used ELM to model the electromagnetic behavior of a triple-mode filter [19]. Then, ELM was utilized to model the third-order intermodulation distortion (IMD3) of an RF power amplifier [20]. In 2021, an improved ELM algorithm was adopted to model ultra-wide band (UWB) antennas [21]. Notably, ELM is a powerful tool for device modeling.

In addition, the combination of neural network and X-parameters is used to establish the X-parameter model. In 2019, three-layer BP neural network was used to model the X-parameters [22]. However, this model was not verified in the actual circuit. In 2021, neural network was also used to model the X-parameters, and the circuit verification of the established model is carried out by the third harmonic error [23].

To achieve high-accuracy modeling for microwave power devices, the CGH40010F transistor produced by Cree is chosen as the modeling object here. X-parameters of this transistor are extracted by ADS, and the prediction models are constructed by ELM and double layer BP neural networks, respectively. The test results show that the mean square error

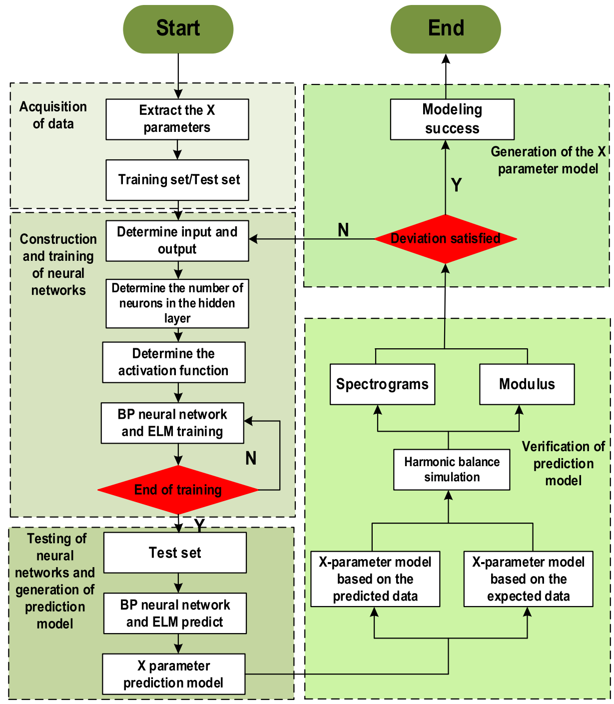


FIGURE 1. Flow chart of BP neural network and ELM for X-parameters modeling.

of ELM prediction model is 0.0027, and the fitting rate is 94%, which is higher than that of the double layer BP neural networks prediction model. Moreover, the X-parameter model is established through the predicted data preciously. Furthermore, to verify the accuracy of the model, harmonic balance simulation is carried out to measure the spectrogram and modulus for the constructed models. The errors of ELM for fundamental, second and third harmonic are 0.673 dBm, 0.314 dBm, and 3.09 dBm, respectively, which is much lower than double layer BP neural network. Therefore, the X-parameter modeling based on ELM can achieve accurate large-signal modeling. As a result, this research can characterize the non-linearity of the transistor accurately. Furthermore, this model is convenient for obtaining X-parameters of the transistor.

The rest of this paper is organized as follows. Section II introduces the X-parameter modeling. The results and discussions are described in Section III. The conclusions are presented in Section IV.

II. X-PARAMETER MODELING

Since X-parameters can characterize the nonlinear characteristics of microwave power devices, they have become a tool for large-signal modeling. Thus, BP neural network and ELM are used to model X-parameters for transistor here. The specific modeling process is shown in Fig. 1.

There are five steps involved in X-parameter modeling based on BP neural network and ELM: acquisition of data, construction and training of neural networks, testing of neural

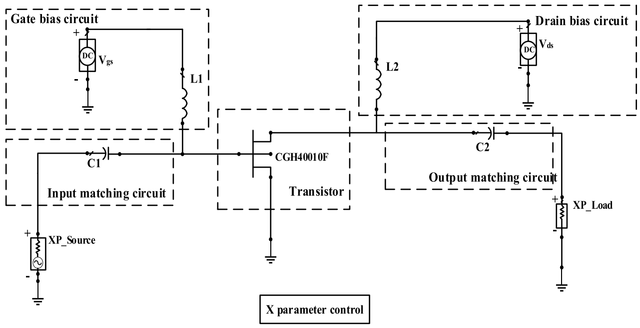


FIGURE 2. Extraction schematic of the X-parameters.

TABLE 1. The meaning of variables in the .xnp file.

Independent variable	Implication
AN_p_f	The amplitude of the incident voltage wave at the port p frequency f in the operating point of a large signal, where f is the frequency of its harmonics and refers to the frequency of a single tone
FB_p_f	$X^{(FB)}$, scattering wave of port p frequency f at large signal operating point
FI_p	$X^{(FI)}$, output DC current of port p at large signal operating point
S_pOut_fOut_pIn_fIn T_pOut_fOut_pIn_fIn	$X^{(S)}$ and $X^{(T)}$, the relationship between the input wave at the port p_{In} frequency f_{In} and the scattered wave at the port p_{Out} frequency f_{Out}
XY_pOut_pIn_fin	$X^{(Y)}$, the relationship between the input wave and the output DC current of the p_{Out} port at the port p_{In} frequency f_{In}

networks and generation of prediction model, verification of prediction model and generation of the X-parameter model.

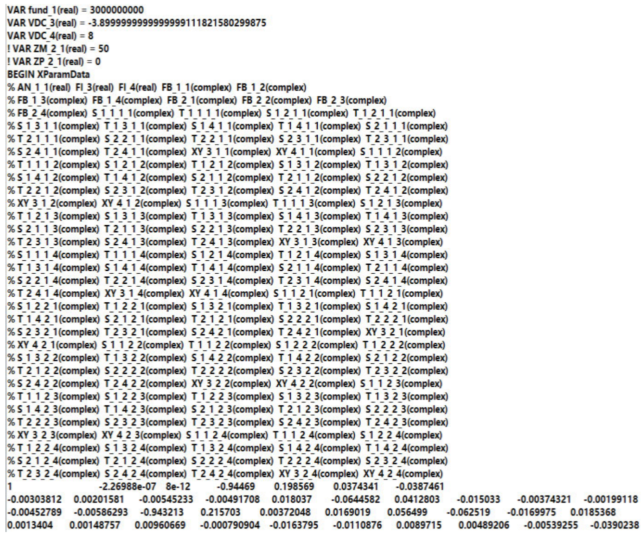
A. DEFINITION OF X-PARAMETERS

An X-parameter model that fully characterizes the nonlinear behavior of the device under test (DUT) can be expressed by equation (1) [24]:

$$B_{p,k} = X_{p,k}^{(F,B)} \left(\left\{ DCS_q \right\}, \left| A_{1,1} \right|, \left\{ A_{\substack{q,k \\ (q,k)\neq(1,1)}} P^{-k} \right\} \right) P^{-k} \tag{1}$$

where $B_{p,k}$ represents the X-parameter model of the DUT, $X_{p,k}^{(F,B)}$ is X-parameters, q is the number of ports, $F_{p,k}$ is the scattered wave behavior of the amplifier at port p, k is the k^{th} harmonic of the amplifier port, A is the large signal excitation, P is the phase factor, and DCS_q is DC bias excitation.

Equation (1) is original nonlinear model, and according to (1), X-parameters can be extracted. Typically, the nonlinear vector network analyzer (NVNA) and ADS are the major methods to extract X-parameters. In this paper, ADS software is used to do this. Then, the extracted data is divided into the training set and the test set.



```
VAR fund_1(real) = 3000000000
VAR VDC_3(real) = -3.89999999999999911821580299875
VAR VDC_4(real) = 8
! VAR ZM_2_1(real) = 50
! VAR ZP_2_1(real) = 0
BEGIN XParamData
% AN_1_1(real) FI_3(real) FI_4(real) FB_1_1(complex) FB_1_2(complex)
% FB_1_3(complex) FB_1_4(complex) FB_2_1(complex) FB_2_2(complex) FB_2_3(complex)
% FB_2_4(complex) S_1_1_1_1(complex) T_1_1_1_1(complex) S_1_2_1_1(complex) T_1_2_1_1(complex)
% S_1_3_1_1(complex) T_1_3_1_1(complex) S_1_4_1_1(complex) T_1_4_1_1(complex) S_2_1_1_1(complex)
% T_2_1_1_1(complex) S_2_2_1_1(complex) T_2_2_1_1(complex) S_2_3_1_1(complex) T_2_3_1_1(complex)
% S_2_4_1_1(complex) T_2_4_1_1(complex) XY_3_1_1(complex) XY_4_1_1(complex) S_1_1_1_2(complex)
% T_1_1_1_2(complex) S_1_2_1_2(complex) T_1_2_1_2(complex) S_1_3_1_2(complex) T_1_3_1_2(complex)
% S_1_4_1_2(complex) T_1_4_1_2(complex) S_2_1_1_2(complex) T_2_1_1_2(complex) S_2_2_1_2(complex)
% T_2_2_1_2(complex) S_2_3_1_2(complex) T_2_3_1_2(complex) S_2_4_1_2(complex) T_2_4_1_2(complex)
% XY_3_1_2(complex) XY_4_1_2(complex) S_1_1_1_3(complex) T_1_1_1_3(complex) S_1_2_1_3(complex)
% T_1_2_1_3(complex) S_1_3_1_3(complex) T_1_3_1_3(complex) S_1_4_1_3(complex) T_1_4_1_3(complex)
% S_2_1_1_3(complex) T_2_1_1_3(complex) S_2_2_1_3(complex) T_2_2_1_3(complex) S_2_3_1_3(complex)
% T_2_3_1_3(complex) S_2_4_1_3(complex) T_2_4_1_3(complex) XY_3_1_3(complex) XY_4_1_3(complex)
% S_1_1_1_4(complex) T_1_1_1_4(complex) S_1_2_1_4(complex) T_1_2_1_4(complex) S_1_3_1_4(complex)
% T_1_3_1_4(complex) S_1_4_1_4(complex) T_1_4_1_4(complex) S_2_1_1_4(complex) T_2_1_1_4(complex)
% S_2_2_1_4(complex) T_2_2_1_4(complex) S_2_3_1_4(complex) T_2_3_1_4(complex) S_2_4_1_4(complex)
% T_2_4_1_4(complex) XY_3_1_4(complex) XY_4_1_4(complex) S_1_1_2_1(complex) T_1_1_2_1(complex)
% S_1_2_2_1(complex) T_1_2_2_1(complex) S_1_3_2_1(complex) T_1_3_2_1(complex) S_1_4_2_1(complex)
% T_1_4_2_1(complex) S_2_1_2_1(complex) T_2_1_2_1(complex) S_2_2_2_1(complex) T_2_2_2_1(complex)
% S_2_3_2_1(complex) T_2_3_2_1(complex) S_2_4_2_1(complex) T_2_4_2_1(complex) XY_3_2_1(complex)
% XY_4_2_1(complex) S_1_1_2_2(complex) T_1_1_2_2(complex) S_1_2_2_2(complex) T_1_2_2_2(complex)
% S_1_3_2_2(complex) T_1_3_2_2(complex) S_1_4_2_2(complex) T_1_4_2_2(complex) S_2_1_2_2(complex)
% T_2_1_2_2(complex) S_2_2_2_2(complex) T_2_2_2_2(complex) S_2_3_2_2(complex) T_2_3_2_2(complex)
% S_2_4_2_2(complex) T_2_4_2_2(complex) XY_3_2_2(complex) XY_4_2_2(complex) S_1_1_2_3(complex)
% T_1_1_2_3(complex) S_1_2_2_3(complex) T_1_2_2_3(complex) S_1_3_2_3(complex) T_1_3_2_3(complex)
% S_1_4_2_3(complex) T_1_4_2_3(complex) S_2_1_2_3(complex) T_2_1_2_3(complex) S_2_2_2_3(complex)
% T_2_2_2_3(complex) S_2_3_2_3(complex) T_2_3_2_3(complex) S_2_4_2_3(complex) T_2_4_2_3(complex)
% XY_3_2_3(complex) XY_4_2_3(complex) S_1_1_2_4(complex) T_1_1_2_4(complex) S_1_2_2_4(complex)
% T_1_2_2_4(complex) S_1_3_2_4(complex) T_1_3_2_4(complex) S_1_4_2_4(complex) T_1_4_2_4(complex)
% S_2_1_2_4(complex) T_2_1_2_4(complex) S_2_2_2_4(complex) T_2_2_2_4(complex) S_2_3_2_4(complex)
% T_2_3_2_4(complex) S_2_4_2_4(complex) T_2_4_2_4(complex) XY_3_2_4(complex) XY_4_2_4(complex)
1             -2.26988e-07  8e-12     -0.94469   0.198569   0.0374341   -0.0387461
-0.00303812  0.00201581  -0.00545233  -0.00491708  0.018037  -0.0644582  0.0412803  -0.015033  -0.00374321  -0.00199118
-0.00452789  -0.00586293  -0.943213  0.215703  0.00372048  0.0169019  0.056499  -0.062519  -0.0169975  0.0185368
0.0013404  0.00148757  0.00960669  -0.000790904  -0.0163795  -0.0110876  0.0089715  0.00489206  -0.00539255  -0.0390238
```

FIGURE 3. X-parameter data block.

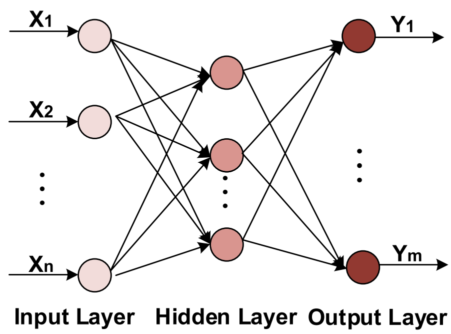


FIGURE 4. Basic structure of BP neural network.

The schematic of X-parameters extraction is established in ADS. In this case, [frequency, P_{in}, V_{gs}, V_{ds}] are chosen as the input vector, a group of X-parameter is the output vector. Specifically, the frequency is 1GHz to 5GHz with a step of 1GHz, P_{in} is 30 dBm to 39 dBm with 3 dBm, V_{gs} is -3.9 V to -0.9 V with step of -1 V, and V_{ds} is 8 V to 20 V with step of 4 V.

As shown in Fig. 2, the schematic is comprised by a CGH40010F transistor, input matching circuit, output matching circuit, gate bias circuit, and drain bias circuit. By setting the corresponding bias voltage and frequency, X-parameter of this transistor can be extracted by the X-parameter control. Finally, a total of 320 groups of X-parameters are extracted, among which a group of X-parameters is composed of 307 data. In building the neural network model, 307 data are trained and tested together as outputs. Moreover, plus setting the parameters and creating the.xnp file, it took a total of 2 minutes to extract this data. And uniform sampling method is used to divide the data into training sets and test sets. Among them, the ratio of training set and test set is 2:1.

The names and specific meanings of the variables in the .xnp file are displayed in Table 1. The nonlinear characteristics of the device is characterized by these parameters.

Taking frequency is 3 GHz, V_{gs} is −3.9 V, V_{ds} is 8 V and P_{in} is 30 dBm as an example, the X- parameter data block is shown in Fig. 3. It is can be seen that the data block header is the name of the variable, and the variable name followed by "complex" is a negative type, "real" is the real type. Combining with Table 1, the extracted X-parameters not only include $X^{(FB)}$ term, but also include $X^{(T)}$ term, AN term, $X^{(FI)}$ term, $X^{(S)}$ term and XY term. It is worth noting that Table 1 and Fig. 1 show the X-parameter terms when the frequency is 3 GHz, V_{gs} is −3.9 V, V_{ds} is 8 V, and P_{in} is 30 dBm. If different conditions and order are used, which contain different terms.

Furthermore, among the 307 data, there are 1 AN term, 2 $X^{(FI)}$ terms, 16 $X^{(FB)}$ terms, 32 XY terms, 128 $X^{(S)}$ terms and 128 $X^{(T)}$ terms. It is worth noting that the values after the data block correspond one-to-one with the variables in the header.

And because X-parameters are the black box model, all the extracted terms form X-parameters, so each term is non-independent.

B. BACK PROPAGATION NEURAL NETWORKS PREDICTION MODEL

BP neural networks are widely used to solve nonlinear problems, such as speech recognition [25], image classification [26], fault detection [27] and risk early warning [28].

The basic structure of the BP neural network is illustrated in Fig. 4. It is mainly composed of the input layer, hidden layer and output layer. The number of neurons in the input and output layers is determined by the input data [X_1, $X_2, \ldots, X_n$] and output data [Y_1, $Y_2, \ldots, Y_m$], and the number of hidden layers is set according to the complexity of nonlinear relationship.

According to the train set, BP neural network starts training. The training of the BP neural network includes forward propagation and back propagation. Forward propagation can obtain the predicted value of the output layer, this value can be calculated from the equation (2):

$$O_k = \sum_{j=1}^{l} H_j \omega_{jk} - b_k \quad k = 1, 2, \cdots, m \tag{2}$$

where O_k is the output layer result, j is the number of neurons in the hidden layer, k is the number of neurons in the output layer, H_j is the output value from the input layer to the hidden layer, ω_{jk} is the weight between the hidden layer neurons and the output layer neurons and b is the output layer threshold.

Then, the error between the predicted output and the expected output of the output layer is calculated by formula (3) as:

$$e_k = Y_k - O_k \tag{3}$$

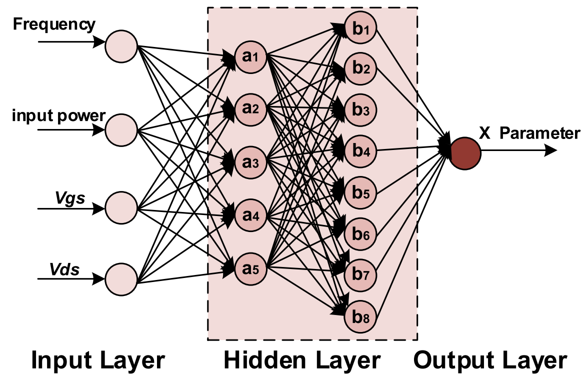


FIGURE 5. X-parameters prediction model based on double layer BP neural network.

where Y_k is the expected output, the weights and thresholds are updated according to e_k.

The purpose of back propagation is to repeatedly modify the weights and thresholds to minimize the error. The weights are updated as the equation (4) and (5):

$$\omega_{ij} = \omega_{ij} + \eta H_j \left(1 - H_j\right) x\,(i) \sum_{k=1}^{m} \omega_{jk} e_k \tag{4}$$

$$\omega_{jk} = \omega_{jk} + \eta H_j e_k \tag{5}$$

where ω_{ij} is the connection weight between the input layer neurons and hidden layer neurons, x_i is the input variable and η is the learning rate. After the weights are updated and the thresholds can be calculated by the equation (6) and (7):

$$a_j = a_j + \eta H_j \left(1 - H_j\right) \sum_{k=1}^{m} \omega_{jk} e_k \tag{6}$$

$$b_k = b_k + e_k \tag{7}$$

where a is the hidden layer threshold. BP neural network repeats these two processes until it converges. By several training and testing, an X-parameter prediction model is generated.

C. EXTREME LEARNING MACHINE PREDICTION MODEL

Next, the ELM prediction model is built, and its structure is the same as that of the BP neural network. There are three steps in ELM modeling. First, the number of hidden layer neurons is set, and the weight ω of the neurons and the threshold b of hidden layer neurons are randomly generated. Second, the activation function f is selected so that the hidden layer output matrix H can be calculated by equation (8):

$$H = f\left(\sum_{i=1}^{n} \omega x_i + \mathrm{b}_j\right) \quad j = 1, 2, \cdots, l \tag{8}$$

where x is input data, j is the number of neurons in the hidden layer, n is the number of neurons in the input layer and f is the excitation function. Third, the weight matrix of neurons between the hidden layer and the output layer can be

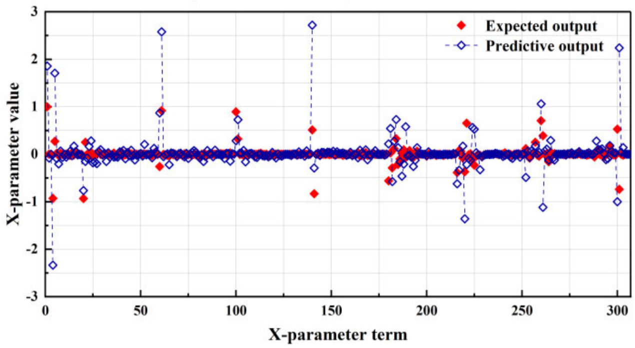


FIGURE 6. Fitting results of double layer BP neural network at frequency=3 GHz, $P_{in} = 30$ dBm, $V_{gs} = -1.9$ V and $V_{ds} = 20$ V.

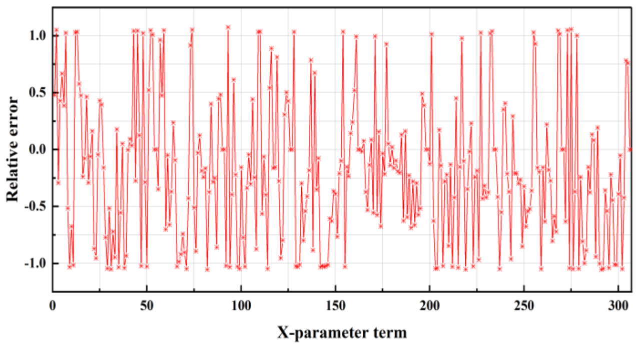


FIGURE 7. Relative error of double layer BP neural network at frequency=3 GHz, $P_{in} = 30$ dBm, $V_{gs} = -1.9$ V and $V_{ds} = 20$ V.

calculated by equation (9):

$$\beta = \min_{\beta} \|H\beta - L\| \tag{9}$$

where L is the expected output of ELM and β is the weight matrix between the hidden layer and the output layer.

In short, the learning process of ELM is simpler than the BP neural network, and it can model for the nonlinear problems within a shorter time. Furthermore, the complex problem of locally optimal solutions will be largely avoided by ELM [29]. Thus, in order to verify the prediction model, an X-parameter model based on the predicted data and expected data is built. Thus, the harmonic balance simulation is carried out to obtain the corresponding spectrograms and modulus. Finally, the deviation of model can be obtained by spectrograms and modulus comparison. If the deviation of spectrograms is in [0, 4], and the error of modulus is ideal, the modeling is successful. Otherwise, it needs to be re-modeled.

III. RESULTS AND DISCUSSIONS

A. PREDICTIVE MODELING BASED ON DOUBLE LAYER BP NEURAL NETWORK AND ELM

In order to achieve X-parameter modeling for transistors, a double layer BP neural network is constructed firstly.

The X-parameter prediction model based on the double layer BP neural network is illustrated in Fig. 5. It can be seen that the input layer is [frequency, P_{in}, V_{gs}, V_{ds}], and the output layer is a set of X-parameter composed of

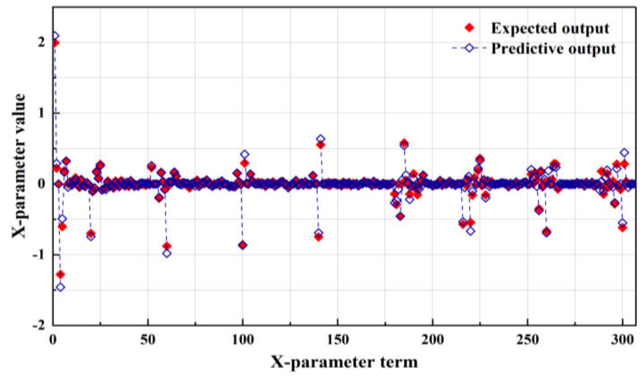


FIGURE 8. Fitting results of double layer BP neural network at frequency=1 GHz, $P_{in} = 36$ dBm, $V_{gs} = -3.9$ V and $V_{ds} = 8$ V.

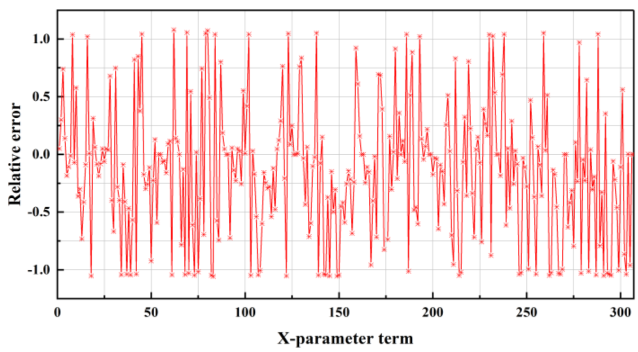


FIGURE 9. Relative error of double layer BP neural network at frequency=1 GHz, $P_{in} = 36$ dBm, $V_{gs} = -3.9$ V and $V_{ds} = 8$ V.

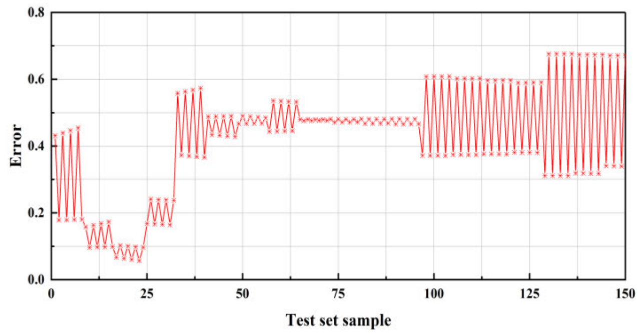


FIGURE 10. Maximum prediction error in 307 terms of X-parameter with double layer BP neural network.

307 data. LM (Levenberg-Marquardt) algorithm is utilized as the training algorithm. By persistent training and testing, the number of neurons in the two hidden layers is five and eight, respectively.

Let the frequency of transistor is 3 GHz, input power is 30 dBm, V_{gs} is -1.9 V and V_{ds} is 20 V. The fit rate between the predicted and expected data is shown in Fig. 6. The horizontal axis of Fig. 6 represents the 307 X-parameter terms extracted, and the vertical axis is the corresponding values of X- parameter. It can be shown that the fitting rate of some predicted data is good when it is near 0, but the fitting rate of data deviating from 0 is not optimal.

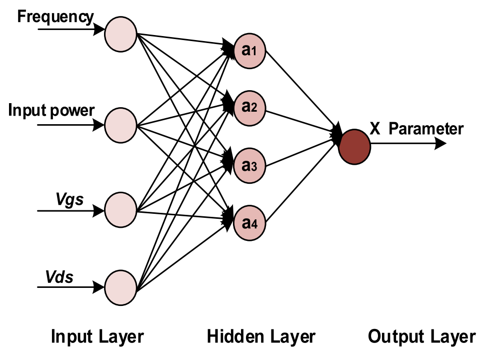


FIGURE 11. X-parameters prediction model based on ELM.

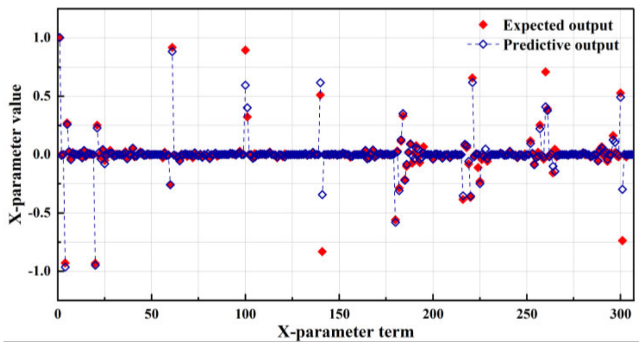


FIGURE 12. Fitting results of ELM prediction model at frequency=3 GHz, $P_{in} = 30$ dBm, $V_{gs} = -1.9$ V and $V_{ds} = 20$ V.

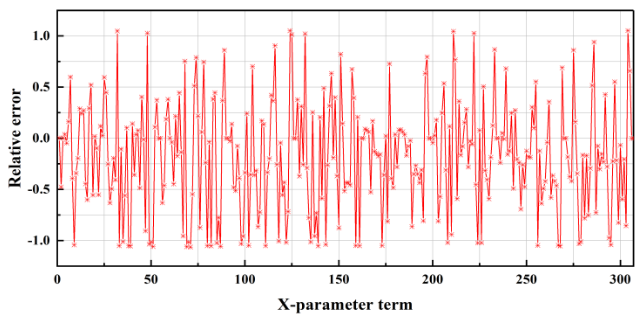


FIGURE 13. Relative error of ELM prediction model at frequency=3 GHz, $P_{in} = 30$ dBm, $V_{gs} = -1.9$ V and $V_{ds} = 20$ V.

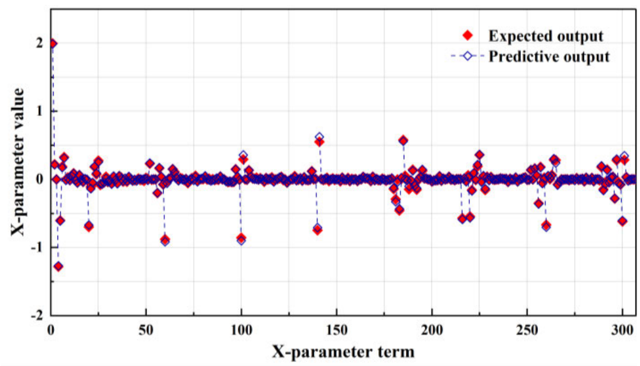


FIGURE 14. Fitting results of ELM prediction model at frequency=1 GHz, $P_{in} = 36$ dBm, $V_{gs} = -3.9$ V and $V_{ds} = 8$ V.

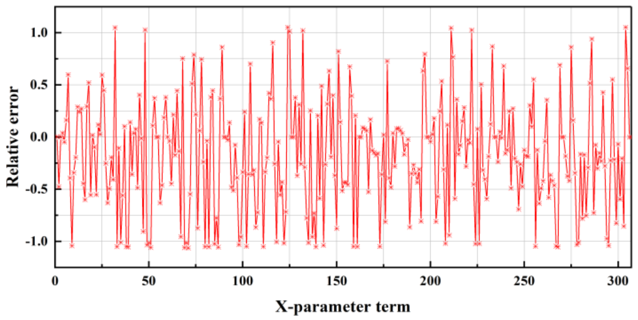


FIGURE 15. Relative error of ELM prediction model at frequency=1 GHz, $P_{in} = 36$ dBm, $V_{gs} = -3.9$ V and $V_{ds} = 8$ V.

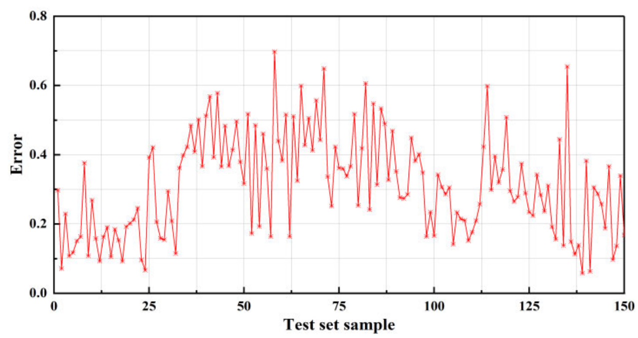


FIGURE 16. Maximum prediction error in 307 terms of X-parameter with double ELM.

In addition, the relative error is calculated, as shown in equation (10):

$$E_R = \frac{C_i - T_i}{T_i} \tag{10}$$

where E_R is relative error, i is number of samples, C represents the predictive output, and T is expected output.

And the relative error of double layer BP neural network is displayed in Fig.7. Typically, the maximum relative error is allowed within [0.95, 1.05]. It can be seen that the relative error of double layer BP neural network at this case is basically within $[-1, 1]$, and more points are concentrated around -1 or 1. Therefore, the relative errors of most prediction data are distributed around the boundary of maximum error. On the contrary, the relative error centered near 0 is less.

In addition, when the frequency is 1GHz, P_{in} is 36 dBm, V_{gs} is -3.9V, and V_{ds} is 8V, the fitting situation of the double layer BP neural network is shown in Fig.8. The overall fit is better than Fig.6. But some point errors much larger than 0 are still significant, which reflects the limitation of double layer BP neural network in X-parameter modeling. And the relative error of double layer BP neural network is displayed in Fig.9.It is seen that in this case the relative error of double layer BP neural network is also basically within $[-1,1]$. Although there are fewer points near -1 and 1 than in Fig.7, most of the points are concentrated between $[-0.75, 0.75]$. And a small part of the relative errors is concentrated near 0. Hence the relative error in this case is better than Fig. 7.

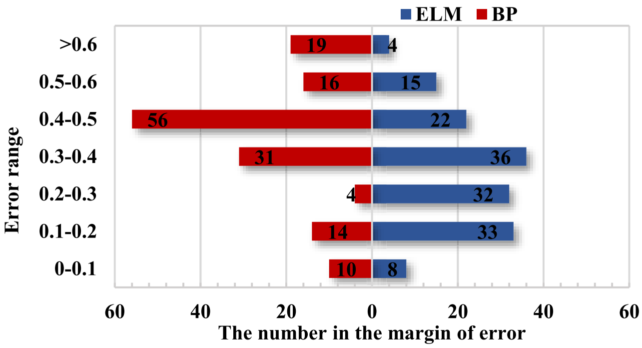


FIGURE 17. Maximum prediction error distribution of X-parameter.

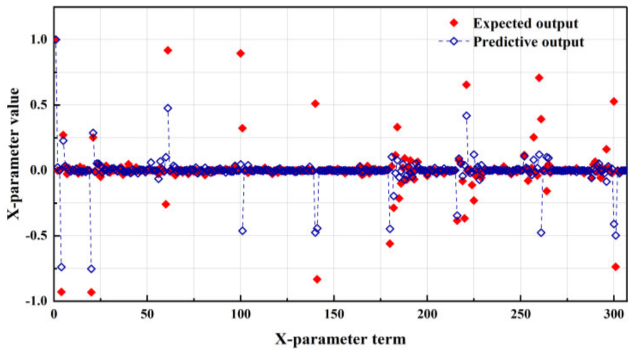


FIGURE 18. Fitting results of interpolation at frequency=3 GHz, P_{in} = 30 dBm, $V_{gs} = -1.9$ V and $V_{ds} = 20$ V.

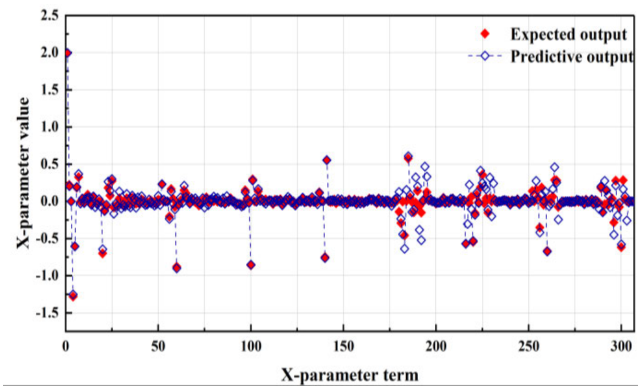


FIGURE 19. Fitting results of interpolation at frequency=1 GHz, P_{in} = 36 dBm, $V_{gs} = -3.9$ V and $V_{ds} = 8$ V.

When the test set is used to verify the BP neural network prediction model, the maximum prediction error in each group of 307 predicted values is obtained, and the results are shown in Fig.10. The maximum error range is within [0, 0.8], but most of the errors are concentrated around [0.5,0.6], while the errors around [0,0.3] are few. Moreover, the mean square error of the prediction model is 0.1455. And the double layer BP model takes 65 seconds in training and testing.

The X-parameter prediction model based on ELM is illustrated in Fig. 11. With the same input and output, the number of neurons in the hidden layer is four finally.

When working at the frequency of 3 GHz, with P_{in} of 30 dBm, V_{gs} of −1.9 V and V_{ds} of 20 V, the fitting curve

TABLE 2. Resource occupancy of double layer BP neural network, ELM and ADS.

Model	Spend time	Memory usage	CPU usage
Double layer BP neural network	65 seconds	10%	17%
ELM	2 seconds	7%	15%
ADS	120 seconds	10%	30%

between the predicted data and the expected data for ELM is illustrated in Fig. 12. It can be seen that the fitting rate of the model near 0 is ideal, and some data deviating from 0 can be roughly fitted, but few data have errors. Even so, the error of ELM is smaller than the double layer BP neural network in the same case. And the relative error of ELM prediction model is shown in Fig.13. It is seen that some points are around 1 or −1, but most of them are around [−0.5, 0.5]. Relative errors concentrated around 1 or −1 are reduced compared with double layer BP neural network.

In case of frequency is 1 GHz, P_{in} is 36 dBm, V_{gs} is −3.9 V and V_{ds} is 8 V, the fitting results of ELM is shown in Fig.14. It can be seen that the fitting results are very ideal, basically all points can be well matched. However, only very few points have errors, and the errors are within the acceptable range. The relative error of ELM prediction model in this case is shown in Fig.15. It can be seen that most points are around [−0.5, 0.5], the overall situation is better than double layer BP neural network. So according to the result of relative error, the prediction credibility of ELM is better than double layer BP neural network.

Similarly, when ELM prediction is adopted, the maximum error of each test set is obtained, which the result is shown in Fig.16. Although the maximum error range of each group is [0,0.8], it can be roughly seen that the maximum error of ELM greater than 0.6 is relatively small. Finally, the mean square error of ELM is calculated as 0.0027, and training and testing of the ELM model takes 10 seconds.

In order to clearly show the maximum error distribution of BP and ELM, Fig.17 is depicted. The error of double layer BP neural network is mostly concentrated in the range of [0.4, 0.6], while the error range of ELM is mostly in the range of [0, 0.4]. It is demonstrated that the error of double layer BP neural network is larger than ELM.

Besides building the neural network model, an interpolation experiment of X-parameter is performed. It is used to verify whether interpolation method can achieve better prediction of X-parameter. The interpolation method is utilized to calculate X-parameter of when $f = 3$ GHz, Pin=30 dBm, $V_{gs} = -1.9$ V and $V_{ds} = 20$ V, and the fitting result is shown in Fig.18. The error of X-parameter calculated by the interpolation method is relatively large. And the fitting effect of points far from 0 is not good, even the fitting of some points closes to 0 is not ideal.

The interpolation experiment with $f = 1$ GHz, P_{in} = 36 dBm, $V_{gs} = -3.9$ V and $V_{ds} = 8$ V is carried out here,

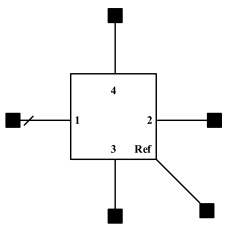


FIGURE 20. X-parameter model.

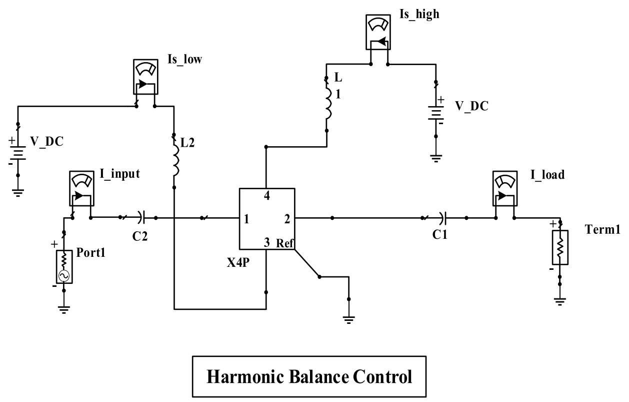


FIGURE 21. Schematic of the harmonic balance simulation.

and the fitting result is shown in Fig.19. It can be seen that although the fitting situation of the data in the former part is generally ideal, the fitting situation of the data in the latter part is not optimistic, and the error is relatively large. It is proved from the interpolation experiments that the neural network model has a greater advantage than the interpolation method in predicting X-parameters.

In the above experiments, although the operation of the interpolation method is not restricted to a single software, its prediction ability for X-parameter is limited, and it is far less accurate than ELM and double layer BP neural network. And through the experiments on double layer BP neural network and ELM, it can be concluded that ELM is better than double layer BP neural network in terms of fitting situation, error or training time.

Moreover, the resource occupancy of X-parameters obtained by using double-layer BP neural network, ELM and ADS is shown in Table 2. In general, it can be seen that ADS spends more time and occupies more computer resources when extracting data. And using neural network to obtain X-parameter occupies the least resources. Among them, BP neural network takes more computer resources than ELM because it takes longer time to train and test.

In conclusion, when the two methods are used to establish a prediction model, ELM modeling is relatively ideal and has small resource occupation, which means that ELM is more effective than the double layer BP neural network in X-parameter modeling.

B. VALIDATE PREDICTIVE MODEL

Harmonic balance simulation focuses on signal frequency domain characteristics and is good at handling the analysis of nonlinear circuits. Therefore, harmonic balance simulation has become the most common method for nonlinear system analysis. This method can be used to simulate noise, gain compression, harmonic distortion, etc. in nonlinear circuits, which is faster than traditional SPICE simulation. Based on these advantages of harmonic simulation, the X-parameter model is verified by this method.

TABLE 3. Spectrogram errors of double layer BP neural network and ELM.

Model	Fundamental Harmonic Error	Second Harmonic Error	Third Harmonic Error
Double layer BP neural network	9.525 dBm	1.309 dBm	14.593 dBm
ELM	0.673 dBm	0.314 dBm	3.09 dBm

TABLE 4. Modulus errors of double layer BP neural network and ELM.

Model	Fundamental Harmonic Modulus Error	Second Harmonic Modulus Error	Third Harmonic Modulus Error
Double layer BP neural network	0.031	0.002	7.665e-4
ELM	0.005	0.001	8.38e-5

In order to verify the error of the model in the actual circuit, harmonic balance simulation is carried out. Finally, spectrogram of the third harmonic and the modulus of the output voltage for the X-parameter models are obtained.

Accordingly, a verification and prediction model of X-parameters is constructed. The prediction data and expected data of the double layer BP neural network and ELM are generated into the corresponding the X-parameter model. When the X-parameter model is generated, the X4P module in ADS is adopted, and then the.xnp parameter file is imported into the module. Because the.xnp file is a GMDIF file, the file type must be selected as GMDIF when imported into the X4P module. Otherwise, the generation of the X-parameter model will fail. It is worth noting that these X-parameter models are generated by X-parameter with a frequency of 3 GHz, P_{in} of 30 dBm, V_{gs} of -1.9 V and V_{ds} of 20 V.

The generated X-parameter model is shown in Fig.20. Where terminal 1 of the X-parameter model is the gate input, terminal 2 is the drain output, terminal 3 is the gate voltage, terminal 4 is the drain voltage, and the Ref terminal is grounded. The X-parameter model can be used to simulate CGH40010F transistors in ADS, such as load pull, harmonic balance simulation and so on. Thus, the relationship between X-parameter file and X-parameter model can be seen. X-parameter file is obtained by ADS to represent the large signal characteristics of the transistor under certain conditions, while X-parameter model is generated by importing X-parameter file into X4P module.

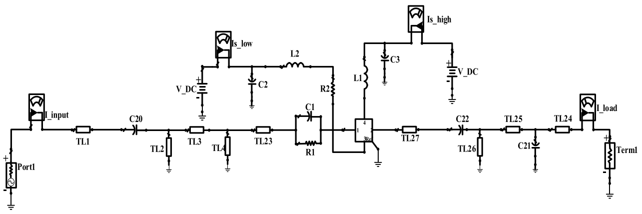


FIGURE 22. Schematic of power amplifier by using X-parameter model.

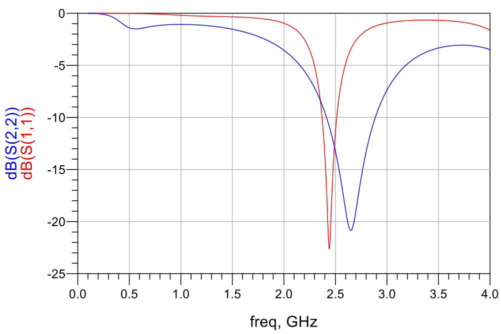


FIGURE 23. Small Signal Simulation Results of S_{11} and S_{22}.

After that, the CGH40010F transistor in Fig. 2 is replaced by the X-parameter models. Then the balance simulation circuit is designed in ADS to obtain the harmonic balance simulation schematic as shown in Fig. 21. After simulation, the corresponding spectrograms for the double layer BP neural network and ELM are obtained.

For the sake of comparison, the standard spectrum has been acquired. Where the fundamental wave is −21.670 dBm, the second harmonic is −21.566 dBm, and the third harmonic is −41.919 dBm.

Then X-parameter spectrogram predicted by the double layer BP neural network model is obtained. And the fundamental, second, and third harmonics of X-parameter spectrogram predicted by the double layer BP neural network model are −12.145 dBm, −20.257 dBm, and −27.326 dBm, respectively. The spectrogram error can be calculated by (11):

$$D_i = E_i - P_i \tag{11}$$

where D_{i} is the error of the i^{th} harmonic, E_i is the i^{th} harmonic of the standard spectrum, and P_i is the i^{th} harmonic of the predicted X-parameter model.

Compared with the standard spectrogram, the error can be obtained as 9.525 dBm, 1.309 dBm, and 14.593 dBm, respectively, as shown in Table 1. That is to say the error of the second harmonic is small, but the fundamental error and the third harmonic error are significant.

Finally, the spectrogram obtained by X-parameters modeled by ELM can be obtained. The fundamental, second, and third harmonics are −20.997 dBm, −21.880 dBm, and −38.829 dBm, respectively. Similarly, the errors for fundamental, second and third harmonic are 0.673 dBm, 0.314 dBm, and 3.09 dBm, respectively, as shown in Table 3. It is concluded that the three harmonic errors of X-parameters are controlled within the range of [0, 4], the modeling accuracy is higher than double layer BP neural network.

Therefore, the double layer BP neural network model only has the ideal error of the second harmonic. In contrast, the fundamental and the third harmonic errors are relatively large, so the model cannot accurately characterize the nonlinear characteristics of the transistor. While, the fundamental error, the second harmonic error and the third harmonic error of the ELM prediction model are all in the range of [0, 4].

For the purpose of further verifying the accuracy of the prediction model, the three-harmonic modulus of the prediction data are compared with the expected data, and the result is displayed in the Table 4. It can be shown that the three harmonic modulus error of double layer BP is much higher than ELM. Through the above verification experiments, it can be concluded that the error of ELM prediction model is lower than the double layer BP neural network, so the model can be used to model the X-parameters of transistors. The BP neural network model is easy to fall into local optimum and cannot reach global optimum, which leads to poor accuracy. However, except for the number of neurons in the hidden layer, other parameters of ELM are randomly generated, so the optimal solution can be obtained quickly and the model has a good accuracy.

C. DESIGN OF RF POWER AMPLIFIER BY USING X-PARAMETER MODEL

X-parameter models can be used in nonlinear circuit designs [30]. To verify the accuracy of the X-parameter prediction model in circuit design, it is used in power amplifier design.

Firstly, the predicted X-parameter is manually collated into a.xnp format. Then, the .xnp file is imported into the X4P module for the design of power amplifier. Because the expected X-parameter model is used first when designing amplifiers. Therefore, the expected X-parameter should be imported into the X4P module.

After the bias circuit by using the expected X-parameter is designed, the source pull and load pull simulation are implemented. Then the optimal source impedance and load impedance are obtained. Next, source impedance and load impedance matching are performed. Finally, the schematic of the power amplifier designed with the X-parameter model is displayed in Fig. 22. It can be seen that the circuit structure is similar to Fig.2, which is also composed of four parts. In addition, to enhance the stability of the circuit, resistor R_1 is added to the gate, and a capacitor C_1 is connected in parallel at R_1 to reduce the gain and power reduction.

When the design of power amplifier by using the expected X-parameter is completed, the small signal simulation is carried out. The small signal simulation results of S_{11} and S_{22} is shown in Fig. 23. It can be seen that when the power amplifier is running at 2.5 GHz, both S_{11} and S_{22} are less than −10 dB, which indicates that most of the signal in the amplifier can be output.

TABLE 5. Output power errors of double layer BP neural network and ELM.

Model	2.5 GHz Output Power Error	5 GHz Output Power Error	7.5 GHz Output Power Error
Double layer BP neural network	1.142 dBm	1.436 dBm	2.294 dBm
ELM	0.089 dBm	0.311 dBm	0.309 dBm

In addition, harmonic balance simulation is carried out. The output power simulated by the expected X-parameter model at 2.5 GHz, 5.0 GHz and 7.5 GHz is 8.917 dBm, 5.206 dBm and −2.899 dBm, respectively.

In order to verify whether the predicted X-parameters can be used in circuit design, the expected X-parameter model is replaced by the predicted X-parameter model for harmonic balance simulation. Finally, under the same conditions, the output power of the double layer BP neural network X-parameter prediction model at 2.5 GHz, 5.0 GHz and 7.5 GHz is 7.775 dBm, 3.770 dBm and 0.605 dBm, respectively. And the output power of the ELM X-parameter prediction model at 2.5 GHz, 5.0 GHz and 7.5 GHz is 8.828 dBm, 5.517 dBm and −3.208 dBm, respectively.

The output power error at different frequencies used by double layer BP neural network and ELM predictive models for amplifier design are shown in Table 5 . The error of the ELM model is below 0.35 dBm, while the error of the double layer BP neural network is below 2.5 dBm. It is concluded that the accuracy of the ELM predictive model in the actual circuit design is higher than the double layer BP neural network. Therefore, the precision of ELM model is sufficient for the power amplifier design.

IV. CONCLUSION

To characterize the nonlinear characteristics of microwave power devices, double layer BP neural network and ELM models are built here. In addition, the harmonic balance simulation is carried out to verify the result. By comparison, ELM has higher accuracy. Therefore, ELM can accurately characterize the nonlinear characteristics of microwave power devices than double layer BP neural networks. It can serve as an essential method for further improving design efficiency and performance for RF microwave circuit. In future, the memory effect can be combined into the proposed ELM model. As a future potential direction, the proposed modeling technique can be further studied for other transistor modeling.

REFERENCES

[1] C. Sarris and Q. J. Zhang, "Machine learning in microwave engineering," *IEEE Microw. Mag.*, vol. 22, no. 10, pp. 20–22, Oct. 2003.

[2] L. Zhu, Q. Zhang, K. Liu, Y. Ma, B. Peng, and S. Yan, "A novel dynamic neuro-space mapping approach for nonlinear microwave device modeling," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 2, pp. 131–133, Feb. 2016.

[3] F. Feng, J. Zhang, W. Zhang, Z. Zhao, J. Jin, and Q. Zhang, "Recent advances in parametric modeling of microwave components using combined neural network and transfer function," *Int. J. Numer. Model., Electron. Netw., Devices Fields*, vol. 33, no. 6, pp. 1–17, Feb. 2020.

[4] W. R. Curtice, "A MESFET model for use in the design of GaAs integrated circuits," *IEEE Trans. Microw. Theory Techn.*, vol. MTT-28, no. 5, pp. 448–456, May 1980.

[5] I. Angelov, L. Bengtsson, and M. Garcia, "Extensions of the Chalmers nonlinear HEMT and MESFET model," *IEEE Trans. Microw. Theory Techn.*, vol. 44, no. 10, pp. 1664–1674, Oct. 1996.

[6] Q. H. Le, D. K. Huynh, S. Lehmann, Z. Zhao, T. Kampfe, and M. Rudolph, "Empirical large-signal modeling of mm-wave FDSOI CMOS based on Angelov model," *IEEE Trans. Electron Devices*, vol. 68, no. 4, pp. 1446–1453, Apr. 2021.

[7] H. Statz, P. Newman, I. W. Smith, R. A. Pucel, and H. A. Haus, "GaAs FET device and circuit simulation in SPICE," *IEEE Trans. Electron Devices*, vol. ED-34, no. 2, pp. 160–169, Feb. 1987.

[8] P. Yang, W. Ming, J. Liang, I. Ludtke, S. Berry, and K. Floros, "Hybrid data-driven modeling methodology for fast and accurate transient simulation of SiC MOSFETs," *IEEE Trans. Power Electron.*, vol. 37, no. 1, pp. 440–451, Jan. 2022.

[9] J. Verspecht and D. E. Root, "Polyharmonic distortion modeling," *IEEE Microw. Mag.*, vol. 7, no. 3, pp. 44–57, Jun. 2006.

[10] W. Akstaller, C. Musolff, R. Weigel, and A. Hagelauer, "X-parameter characterization of TC SAW filters with enhanced dynamic range," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 11, pp. 4541–4549, Nov. 2017.

[11] P. Barmuta, F. Ferranti, A. Lewandowski, L. Knockaert, and D. Schreurs, "Efficient generation of X-parameters transistor models by sequential sampling," *IEEE Microw. Wireless Compon. Lett.*, vol. 24, no. 8, pp. 530–532, Aug. 2014.

[12] C. I. Lee, Y. T. Lin, and W. C. Lin, "Large-signal characterization of pHEMT under different load conditions by using X-parameters," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 2, pp. 125–127, Feb. 2016.

[13] R. Moezzi, J. Hlava, and T. Vu, "Implementation of X-parameters principle for non-linear vibroacoustic membrane using two-port measurement," *Traitement Signal*, vol. 36, no. 4, pp. 297–301, Oct. 2019.

[14] J. Jin, F. Feng, W. Na, S. Yan, W. Liu, L. Zhu, and Q. Zhang, "Recent advances in neural network-based inverse modeling techniques for microwave applications," *Int. J. Numer. Model., Electron. Netw., Devices Fields*, vol. 33, no. 6, pp. 1–18, Nov. 2020.

[15] W. Liu, W. Na, L. Zhu, J. Ma, and Q.-J. Zhang, "A Wiener-type dynamic neural network approach to the modeling of nonlinear microwave devices," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 6, pp. 2043–2062, Feb. 2017.

[16] W. Y. Liu, L. Zhu, F. Feng, W. Zhang, Q. J. Zhang, Q. Lin, and G. Liu, "A time delay neural network based technique for nonlinear microwave device modeling," *Micromachines*, vol. 11, no. 9, pp. 1–14, Aug. 2020.

[17] S. Ding, X. Xu, and R. Nie, "Extreme learning machine and its applications," *Neural Comput. Appl.*, vol. 25, nos. 3–4, pp. 549–556, Sep. 2014.

[18] G. Huang, G.-B. Huang, S. Song, and K. You, "Trends in extreme learning machines: A review," *Neural Netw.*, vol. 61, pp. 32–48, Jun. 2015.

[19] L.-Y. Xiao, W. Shao, T.-L. Liang, and B.-Z. Wang, "Efficient extreme learning machine with transfer functions for filter design," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 551–553.

[20] C.-Y. Zhang, Y.-Y. Zhu, Q.-F. Cheng, H.-P. Fu, J.-G. Ma, and Q.-J. Zhang, "Extreme learning machine for the behavioral modeling of RF power amplifiers," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 554–557.

[21] J. Nan, H. Xie, M. Gao, Y. Song, and W. Yang, "Design of UWB antenna based on improved deep belief network and extreme learning machine surrogate models," *IEEE Access*, vol. 9, pp. 126541–126549, 2021.

[22] N. Lei, F. Jiang, and L. Sun, "X-parameter modelling of GaN HEMT based on neural network," *J. Eng.*, vol. 2019, no. 23, pp. 8955–8958, Dec. 2019.

[23] B. Wang, L. Sun, P. Liang, and F. Nian, "Nonlinear model method of microwave power device based on extreme learning machine," in *Proc. Int. Conf. Microw. Millim. Wave Technol. (ICMMT)*, Sep. 2020, pp. 1–3.

[24] D. E. Root, J. Verspecht, J. Horn, and M. Marcu, "X-parameters-fundamental concepts," in *X-Parameters: Characterization, Modeling, and Design of Nonlinear RF and Microwave Components*, 11st ed. Cambridge, U.K.: CUP, 2013, pp. 20–44.

[25] A. Argha, B. G. Celler, and N. H. Lovell, "A novel automated blood pressure estimation algorithm using sequences of Korotkoff sounds," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 4, pp. 1257–1264, Apr. 2021.

[26] J. Zhu, A. Wu, X. Wang, and H. Zhang, "Identification of grape diseases using image analysis and BP neural networks," *Multimedia Tools Appl.*, vol. 79, nos. 21–22, pp. 14539–14551, Jan. 2019.

[27] H. Yu and T. Wang, "A method for real-time fault detection of liquid rocket engine based on adaptive genetic algorithm optimizing back propagation neural network," *Sensors*, vol. 21, no. 15, pp. 1–15, Jul. 2021.

[28] G. Du, Z. Liu, and H. Lu, "Application of innovative risk early warning mode under big data technology in internet credit financial risk assessment," *J. Comput. Appl. Math.*, vol. 386, pp. 1–11, Apr. 2021.

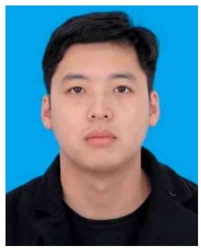
[29] H. Bai, F. Feng, J. Wang, and T. Wu, "Modeling M(3000)F2 based on extreme learning machine," *Adv. Space Res.*, vol. 65, no. 1, pp. 107–114, Jan. 2020.

[30] M. Wu, G. Crupi, C. Yu, and J. Cai, "Application of X-parameter model of gallium nitride device for a continuous broadband Doherty power amplifier design," *Int. J. Numer. Model., Electron. Netw., Devices Fields*, vol. 2022, pp. 1–15, Oct. 2022.

QIAN LIN received the B.S. degree in electronic information science and technology from Qinghai Nationalities University, Xining, China, in 2004, the M.S. degree in electronic information technology from Qinghai Normal University, Xining, in 2010, and the Ph.D. degree in circuits and systems from Tianjin University, Tianjin, China, in 2017. She is currently a Professor at Qinghai Minzu University. Her research interests include RF circuit design, RF device modeling, testing, simulation of circuit reliability, and electronic technology application.

XIAO-ZHENG WANG was born in Shandong, China, in 1997. She is currently pursuing the master's degree with Qinghai Minzu University. Her main research interest includes neural network modeling of microwave power devices.

HAI-FENG WU received the B.E. and M.E. degrees in electronics from the University of Electronic Science and Technology of China, Chengdu, China, in 2007 and 2010, respectively, and the Ph.D. degree in microelectronics and solid state electronics from Tianjin University, Tianjin, China, in 2016. From 2010 to 2011, he was an Engineer at Accelicon Technologies, Inc., Beijing, China, where he was involved with the transistor model building. Since 2016, he has been an Engineer at the Chengdu Ganide Technology, Chengdu, where he is involved with the RFIC design. His research interests include RF transistors modeling and the power amplifier design in MMIC.



LI-NING JIA was born in Hebei, China, in 1998. He is currently pursuing the master's degree with Qinghai Minzu University. His main research interests include temperature reliability of RF circuit, ANN modeling application, and RF circuit design.