IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 18, 2020, accepted August 28, 2020, date of publication September 1, 2020, date of current version September 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3020880

Complex-Valued Pipelined Chebyshev Functional Link Recurrent Neural Network for Joint Compensation of Wideband Transmitter Distortions and Impairments

MINGYU LI[1], (Member, IEEE), ZHENDONG CAI[1], YAO YAO[2], CHANGZHI XU[3,4], (Member, IEEE), YI JIN[4], AND XUGUANG WANG[1]

[1]School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China
[2]School of Intelligent Technology and Engineering, Chongqing University of Science and Technology, Chongqing 401331, China
[3]State Key Laboratory of Millimeter Waves, School of Information Science and Engineering, Southeast University, Nanjing 210096, China
[4]Xi'an Branch of China Academy of Space Technology, Xi'an 710000, China

Corresponding author: Mingyu Li (myli@cqu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61801377.

ABSTRACT In this article, a novel digital predistortion (DPD) model based on complex-valued pipelined Chebyshev functional link recurrent neural network (CPCFLRNN) for joint compensation of wideband transmitter distortions and impairments is proposed. The functional link neural network (FLNN) model has attracted much attention from scholars, and many improved models using this structure, such as Chebyshev FLNN, have been applied in the DPD of power amplifiers (PAs). However, these existing neural network models cannot deal with complex-valued input signals simultaneously, and the real-valued model structure will introduce cumbersome training algorithm and result in a long training time. The pipelined recurrent neural network (PRNN) has been successfully applied to nonlinear signal prediction because of its excellent ability for dealing with nonlinear nonstationary signals. Therefore, the PRNN model containing Chebyshev structure is extended to complex domain for the first time to obtain the CPCFLRNN model for DPD application. Considering the strong correlation of in-phase and quadrature phase (I/Q) components of the transmitter signal, the real time recurrent learning (RTRL) algorithm based on fully complex activation function is selected and extended to complex domain to obtain the complex-valued RTRL (CRTRL) algorithm for CPCFLRNN model training. A GaN PA was employed to verify the effectiveness of the proposed models. And the input signal is a 30MHz LTE signals which consists of I/Q imbalance and dc offsets. The experimental results show that the proposed CPCFLRNN model have more accurate modeling effect and better linearization performance compared with the conventional DPD models.

INDEX TERMS Digital predistortion (DPD), complex-valued pipelined Chebyshev functional link recurrent neural network (CPCFLRNN), I/Q imbalance, dc offset, power amplifier (PA).

I. INTRODUCTION

In 4G and 5G mobile communication systems, the demand for spectrum resources is increasing rapidly. Quadrature Amplitude Modulation (QAM) and Orthogonal frequency division multiplexing (OFDM) modulation are widely used in modern communication systems for improving spectrum efficiency. Such modulation techniques will make it more difficult to design power amplifier (PA), which play a crucial role in RF front-end components, and will also result in spectrum regeneration. At the same time, this kind of modulation signals also have the higher peak-to-average power ratio (PAPR) characteristics, which is sensitive to the nonlinear distortion of the RF PA. In addition, due to the difference between the I/Q channels, for example, the phase shifter is not the ideal 90 degrees, and the amplitude and phase response of the filter are not exactly the same, which will lead to I/Q imbalance of transmitter. This imbalance will result in the

The associate editor coordinating the review of this manuscript and approving it for publication was Young Jin Chun.

serious interactive distortion of the main signals, and will reduce the dynamic range of the communication system [1].

Digital predistortion (DPD) has become the mainstream technology of PA linearization because of its advantages of low cost, good linearization performance and high flexibility [2]. The basic principle of DPD is to insert a predistorter before the PA, and the predistorter is inverse to the nonlinear characteristic of the PA, so as to obtain the linearly amplified output at the end of the transmitter [3]. Early digital predistortion usually adopted a look-up table (LUT) [4], [5], but the LUT requires a large storage space and the convergence speed is slow. In order to reduce the requirement of memory space and to jointly compensation for PA and I/Q impairments, the conjugate memory polynomial (CMP) model is proposed in [6].

On the other hand, artificial neural network (ANN) is a nonlinear system with powerful intelligent information processing function, strong robustness, memory ability and self-learning ability, and can map complex nonlinear relations. The combination of neural network and PA predistortion benefits from the powerful approximation effect of neural network in nonlinear system modeling [7], [8]. The neural network models used for PA behavioral modeling mainly include bidirectional long short-term memory (BiLSTM) neural network [9], radial basis function (RBF) neural network [10] and time delay neural network (TDNN) [8]. However, there are few DPD models to jointly compensate I/Q imbalance and PA nonlinearities [11]–[13].

The functional link neural network (FLNN) models using orthonormal functions for functional expansion has been used to identify nonlinear system. The Chebyshev FLNN using orthogonal Chebyshev polynomial expansion has been applied in [14], and shows better fitting ability and faster convergence speed compared to traditional neural network models. The combination of Chebyshev polynomial and recurrent neural network (RNN) is also proved to be effective for nonlinear adaptive filter [15]. The complex-Chebyshev FLNN has been successfully applied to power amplifier behavioral modeling [16].

Meanwhile, the complex-valued domain signal processing have been introduced into current data analysis, and many learning algorithms have been extended for parameter extraction. Therefore, the real-valued the real time recurrent learning (RTRL) should be extended to complex domain for complex-valued parameter extraction. In complex domain, the nonlinear activation functions have the two necessary properties: bounded and differentiable everywhere. According to Liouville's theorem, the only such function in complex domain is the constant function [17]. Whereas a bounded function or an analytic function must be selected, but if the activation function is analytic, then it must be unbounded, and if it is bounded, then it should be nonanalytic. In order to overcome this conflict, two concepts are proposed: split complex activation function and the fully complex activation function. Due to the strong correlation between I/Q signals, the performance of the split complex activation function is poor for transmitter modeling [18].

Based on the characteristics of the traditional PA model, a new model structure aiming at joint compensation for nonlinear memory effects of PA and I/Q impairments is proposed. The proposed model can directly process the complex-valued signals, which can provide a new research idea for the transmitter behavioral modeling. For easy reading, Table 1 gives a list of important abbreviations. In order to evaluate the model performance, the CMP model, the CPRNN model, the Chebyshev functional link fully connected recurrent neural network (CFL-FCRNN), and the proposed CPCFLRNN model were used to compare the performance of transmitter modeling and nonlinear compensation. Various graphical and numerical results show that the proposed CPCFLRNN model based on the CRTRL algorithm can give improved performance compared to traditional models.

TABLE 1. A list of important abbreviations.

ACPR	adjacent channel power ratio
CFL-FCRNN	chebyshev functional link fully connected recurrent neural network
CPCFLRNN	complex-valued pipelined Chebyshev functional link recurrent neural network
CPRNN	complex-valued pipelined recurrent neural network
CRTRL	complex-valued real time recurrent learning
DPD	digital predistortion
FCRNN	fully connected recurrent neural network
FLNN	functional link neural network
CMP	conjugate memory polynomial
NMSE	normalized mean square error
PAPR	peak-to-average power ratio
PA	power amplifier
PRNN	pipelined recurrent neural network
PSD	power spectrum density
RVFTDNN	real-valued focused time-delay neural network

This article is organized as follows: In Section II, the CFL-FCRNN structure is firstly proposed, and then the CPCFLRNN structure is proposed based on this structure. In Section III, the CRTRL learning algorithm of the two structures is derived, and the weight update equation is obtained. Section IV introduces the optimal model parameters and simulation results of the four models, and then the experimental results are given. The full paper is summarized in section V.

II. BEHAVIORAL MODEL FOR PAS

A. CFL-FCRNN MODEL FOR PAS

The precondition that the PA model can be constructed is that the amplitude and phase information can be extracted from the complex-valued waveform. The most common solution is to introduce dual-input and dual-output neural network

structure [19]. During the training process, the coefficients of two networks are determined independently by the amplitude and phase of the input and output signals. The basic network structure shown in Fig. 1, which can be used for PA modeling with high nonlinear degree by using the I/Q components. This kind of topological structure can be used to model the PA with high degree of nonlinearity. For PA with a strong memory effect, the modeling effect will be poor. In addition, as the two neural networks are trained separately, asynchronous convergence may occur [8].

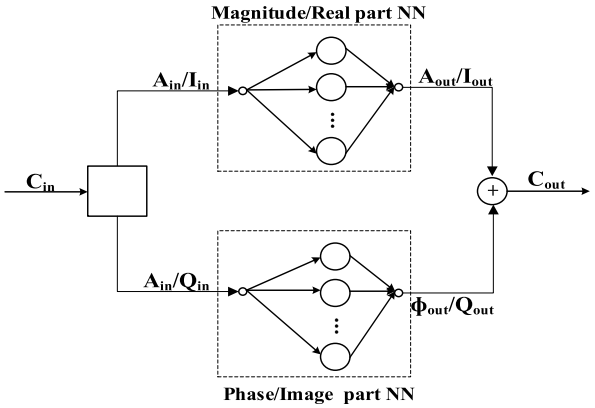


FIGURE 1. Basic neural network structure for PA.

Considering the effect of PA nonlinear memory, TDNN model shown in Fig. 2 was used for PA modeling. Due to the memory effect of the system, the output of the amplifier depends on the input values at the current and previous moments. Therefore, the delay lines should be added to extract information from the current and past inputs, and the topological structure of input with in-phase and quadrature information is adopted [8]. Although this network topology takes memory effect into account through input delay, it does not actually take into account the feedback delay of the network output, which will cause the delay effect of this incomplete structure, and also increase the complexity and instability of the system. Based on the TDNN model, a new PA model structure, the real-valued focused time-delay neural network (RVFTDNN), was proposed in [7]. This model considers the output feedback delay by introducing tap delay line (TDL). However, when the input signal is complex, this structure will inevitably encounter the problem of overtraining or undertraining.

The application of CFLNN in the recognition of nonlinear systems has been proved to have greater advantages than traditional neural network models such as multilayer perceptron (MLP) and RNN [20], [21]. Since the baseband input and output data of PA are complex signals, the most suitable method is to use complex neural network structure and the complex training algorithm [16]. In complex-valued neural networks, the inputs, the weights, and the outputs are all complex-valued, and the training algorithm directly extends to the complex domain. Fig. 3 shows CFL-FCRNN model, consisting of N neurons with P external inputs and N feedback connections. The "FE" is the function extension to increase the dimension of the input pattern, so it is easier to identify complex nonlinear dynamic systems. The neuron input and its expansion using Chebyshev polynomial are given

$$\begin{aligned} S(k) &= [s(k-1), s(k-2), \ldots, s(k-p)]^T \\ &= S^r(k) + jS^i(k) \end{aligned} \tag{1}$$

$$\begin{aligned} S_{FE}(k) &= FE\,(s(k-1), s(k-2), \ldots, s(k-p)) \\ &= [S_{FE,1}(k), S_{FE,2}(k), \ldots, S_{FE,p}(k)]^T \\ &= S^r_{FE}(k) + jS^i_{FE}(k) \end{aligned} \tag{2}$$

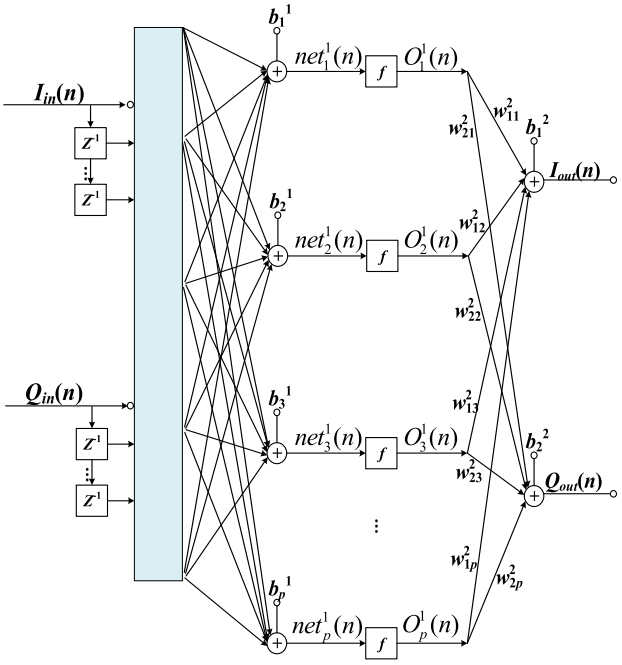


FIGURE 2. RVTDNN model for PA.

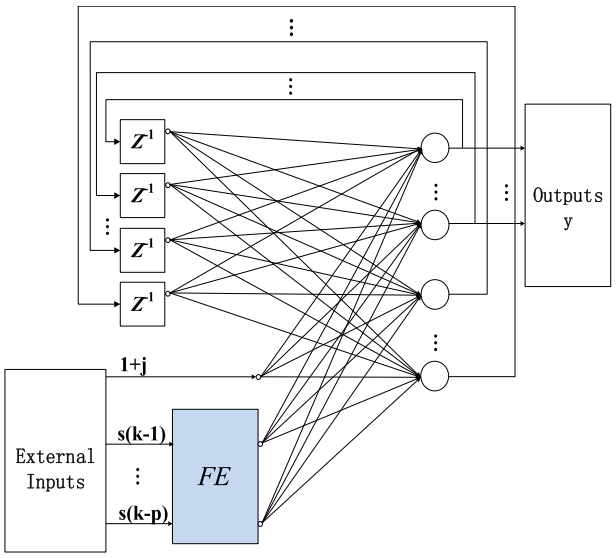


FIGURE 3. CFL-FCRNN model for PA.

The approximate formula of Chebyshev polynomial of any order is [16]

$$C_b(\omega) = \sum_{m=0}^{[b/2]} (-1)^m \binom{b}{2m} \omega^{b-2m}(1-\omega^2)^m \tag{3}$$

where $[b/2]$ is the integer part of $(b/2)$, and b is the order of Chebyshev polynomial expansion. The Chebyshev expansion is given as

$$\left.\begin{array}{l} C_0(\omega) = 1 \\ C_1(\omega) = \omega \\ C_2(\omega) = 2\omega^2 - 1 \\ C_3(\omega) = 4\omega^3 - 3\omega \\ \cdots \end{array}\right\} \tag{4}$$

Chebyshev polynomials are used to expand the input signal into any order, and the enhanced input is [14]

$$\begin{aligned} &S_{FE,b}(k) \\ &\quad = C_b(S(k)) \\ &\quad = [1\,|C_1(s(k-1))\,, C_1(s(k-2)), \ldots, C_1(s(k-p))\,| \\ &\qquad \times |C_1(s(k-1))\; C_1(s(k-2)), \ldots, \\ &\qquad C_1(s(k-(p-1)))C_1(s(k-p))\,|, \\ &\qquad |C_2(s(k-1))\,, C_2(s(k-2)), \ldots, C_2(s(k-p))\,|, \\ &\qquad \cdots \\ &\qquad |C_b(s(k-1))\,, C_b(s(k-2)), \ldots, C_b(s(k-p))|]^T \end{aligned} \tag{5}$$

To simplify the calculation, take $b = 3$ to get the third-order function expansion input

$$\begin{aligned} S_{FE,3}(k) &= [1, C_1(S(k)), C_1(S(k)), C_1(S(k-1)), \\ &\quad C_2(S(k)), C_3(S(k))]^T \\ &= [1, s(k-1), s(k-2), \ldots, s(k-p), \\ &\quad s(k-1)s(k-2), \ldots, s(k-(p-1))s(k-p), \\ &\quad 2s^2(k-1)-1, \ldots, 2s^2(k-p)-1, \\ &\quad 4s^3(k-1)-3s(k-1), \ldots, \\ &\quad 4s^3(k-p)-3s(k-p)]^T \end{aligned} \tag{6}$$

The application of neural network for PA modeling also needs to take into account the nonlinearity of PA, which is reflected in the conversion of AM/AM and AM/PM. Therefore, it is necessary to apply appropriate conjugate transformation to the above equation [16], so that the nonlinearity of passband can be expressed under the baseband as follows

$$\begin{aligned} S_{FE}(k) &= [1, s(k-1), s(k-2), \ldots, s(k-p), \\ &\quad s(k-1)|s(k-2)|, \ldots, s(k-(p-1))|s(k-p)|, \\ &\quad 2s(k-1)|s(k-1)|-1, \ldots, \\ &\quad 2s(k-p)|s(k-p)|-1, \\ &\quad 4s(k-1)|s^2(k-1)|-3s(k-1), \ldots, \\ &\quad 4s(k-p)|s^2(k-p)|-3s(k-p)]^T \end{aligned} \tag{7}$$

The entire network is a two-layer structure, which includes the external delay input layer and output feedback layer. The external complex-valued input is delayed and then extended by Chebyshev function, the bias input is the $1+j$, and the complex output of each neuron is represented by $y_l(k)$. The total input of the whole network is composed of Chebyshev functional expansion input, bias and feedback, expressed as follows

$$\begin{aligned} X(k) &= [S_{FE}(k), 1+j, y_1(k-1), y_2(k-1), \ldots, y_N(k-1)]^T \\ &= X_n^r(k) + jX_n^i(k), \quad n = 1, \ldots, p+N+1 \end{aligned} \tag{8}$$

The output of the *lth* neuron can be written as:

$$\begin{aligned} y_l(k) &= \psi^r(u_l^r(k)) + j\psi^i(u_l^i(k)) \\ &= y_l^r(k) + jy_l^i(k), \quad l = 1, \ldots, N \end{aligned} \tag{9}$$

$$u_l(k) = \sum_{n=1}^{p+N+1} w_{l,n}(k) X_n(k) \tag{10}$$

ψ represents the complex-valued nonlinear activation function of the neuron, and (10) is the input of the activation function at time k, that is, the linear sum of all the inputs of node after the weights are applied. The weight matrix of the whole neural network is

$$W = [w_1, \ldots, w_N] \tag{11}$$

where the weight vector of the *lth* neuron is

$$w_l = [w_{l,1}, \ldots, w_{l,p+N+1}]^T \tag{12}$$

The length of the whole weight matrix is $(p+N+1)*N$.

B. CPCFLRNN MODEL FOR PAS

A nonlinear adaptive prediction model called pipelined recurrent neural network (PRNN), which deals with real-valued non-stationary signals, has been successfully applied to nonlinear systems and has achieved remarkable effects [22]. Due to its spatial representation of time and feedback connection within the structure, the PRNN structure can better deal with the gradient vanishing problem, and has strong robust neural network prediction ability. The advantage of PRNN is that it is composed of M neural networks with the same structure, and can reduce the computational complexity.

It has been proved through analysis that CFL-FCRNN has better modeling performance than RVFTDNN. In this article, the complex-valued PRNN is introduced as an extension of real-valued PRNN, and each module adopts CFL-FCRNN structure shown in Fig. 4. Each module is designed as a CFL-FCRNN with N neuron, in which the previous $M-1$ module is a non-fully connected CFL-FCRNN. The $N-1$ outputs of its output neurons are used to feed back to the input, and the output of the remaining neurons (i.e. the output of the first neuron) is passed directly to the next module. The last module is a fully connected CFL-FCRNN, where the output of all neurons is fed back to the input. In the CPCFLRNN structure, all modules use the same complex-valued weight matrix

$$W(k) = [w_1(k), \ldots, w_l(k), \ldots, w_N(k)] \tag{13}$$

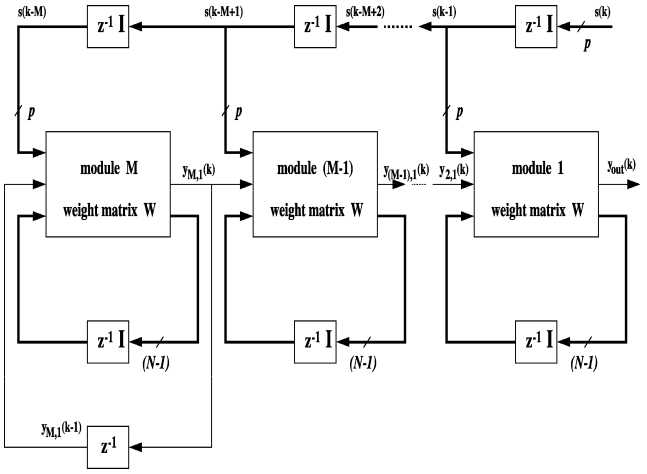


FIGURE 4. CPCFLRNN model for PA.

Refer to (9), the mathematical expression of CPCFLRNN is

$$\begin{aligned} y_{t,l}(k) &= \psi^r(u_l^r(k)) + j\psi^i(u_l^i(k)) \\ &= y_{t,l}^r(k) + jy_{t,l}^i(k), \quad t = 1, 2, \ldots, M \end{aligned} \tag{14}$$

$y_{t,l}(k)$ represents the output of the *lth* neuron of module t at time k. The input expression of each module is given as follows

$$\begin{aligned} X_t^T(k) = [&S_{FE,t}(k), 1 + j, y_{t+1,1}(k-1), y_{t,2}(k-1), \ldots, \\ &y_{t,N}(k-1)] \end{aligned} \tag{15}$$

$$\begin{aligned} X_M^T(k) = [&S_{FE,M}(k), 1 + j, y_{M,1}(k-1), y_{M,2}(k-1), \ldots, \\ &y_{M,N}(k-1)] \end{aligned} \tag{16}$$

where $X_t^T(k)$ represents the input vector by the module t at time k. $X_M^T(k)$ is input vector of the *Mth* module. It can be easily seen that the input of the ahead $M-1$ modules contains the output $y_{t+1,l}(k-1)$ of the latter module and replaces the feedback delay of the first output of this module. Since the last module M has no input from the latter module, the input retains the feedback delay of all the output of this module. Therefore, the final output signal of CPCFLRNN is represented by the output of the first neuron of the first module

$$y_{out}(k) = y_{1,1}(k) \tag{17}$$

III. COMPLEX-VALUED RTRL ALGORITHM

After deriving the output expression of the neuron, the optimal weight parameters should be obtained according to the learning process. Since the CPCFLRNN structure contains CFL-FCRNN structure, the CRTRL algorithm applicable to CFL-FCRNN is firstly deduced in this section, and then the CRTRL algorithm for CPCFLRNN is derived in the similar way.

A. CRTRL ALGOTITHM FOR CFL-FCRNN

Define the output of each neuron in the CFL-FCRNN is $y_l(k)$. The error signal can be obtained by subtracting the output of the model from the reference signal provided by the external information source [23]. The error signal at time k is given as follows

$$\begin{aligned} \varepsilon_l(k) &= d(k) - y_l(k) \\ &= d^r(k) - y_l^r(k) + d^i(k) - y_l^i(k) \end{aligned} \tag{18}$$

$d^r(k)$ and $d^i(k)$ are the real and imaginary parts of the actual output, respectively, and the cost function can be defined as

$$\begin{aligned} J(k) &= \frac{1}{2}\sum_{l=1}^{N} |\varepsilon_l(k)|^2 \\ &= \frac{1}{2}\sum_{l=1}^{N} \varepsilon_l(k)\varepsilon_l^*(k) \\ &= \frac{1}{2}\sum_{l=1}^{N} [(\varepsilon_l^r)^2 + (\varepsilon_l^i)^2] \end{aligned} \tag{19}$$

the CRTRL algorithm extracts the weight coefficients based on the gradient descent algorithm recursively. It can be seen that $J(k)$ is a non-constant function, and it is need to calculate its complex-valued reciprocal, so its gradient to weight w can be calculated as

$$\nabla_{w_{l,n}} J(k) = \frac{\partial J(k)}{\partial w_{l,n}^r} + j\frac{\partial J(k)}{\partial w_{l,n}^i} \tag{20}$$

Then the real and imaginary parts can be separated as

$$\frac{\partial J(k)}{\partial w_{l,n}^r(k)} = \frac{\partial J}{\partial y_l^r}\left(\frac{\partial y_l^r(k)}{\partial w_{l,n}^r(k)}\right) + \frac{\partial J}{\partial y_l^i}\left(\frac{\partial y_l^i(k)}{\partial w_{l,n}^r(k)}\right) \tag{21}$$

$$\frac{\partial J(k)}{\partial w_{l,n}^i(k)} = \frac{\partial J}{\partial y_l^r}\left(\frac{\partial y_l^r(k)}{\partial w_{l,n}^i(k)}\right) + \frac{\partial J}{\partial y_l^i}\left(\frac{\partial y_l^i(k)}{\partial w_{l,n}^i(k)}\right) \tag{22}$$

The above expression with sensitivity can be defined as

$$\Lambda_{l,n}^{rr}(k) = \frac{\partial y_l^r(k)}{\partial w_{l,n}^r(k)} = \psi'^r(u_l(k))\frac{\partial u_l^r(k)}{\partial w_l^r(k)} \tag{23}$$

$$\Lambda_{l,n}^{ir}(k) = \frac{\partial y_l^i(k)}{\partial w_{l,n}^r(k)} = \psi'^i(u_l(k))\frac{\partial u_l^i(k)}{\partial w_l^r(k)} \tag{24}$$

$$\Lambda_{l,n}^{ri}(k) = \frac{\partial y_l^r(k)}{\partial w_{l,n}^i(k)} = \psi'^r(u_l(k))\frac{\partial u_l^r(k)}{\partial w_l^i(k)} \tag{25}$$

$$\Lambda_{l,n}^{ii}(k) = \frac{\partial y_l^i(k)}{\partial w_{l,n}^i(k)} = \psi'^i(u_l(k))\frac{\partial u_l^i(k)}{\partial w_l^i(k)} \tag{26}$$

To calculate the gradient in the complex domain, a complex activation function is required to be resolved in the complex domain to satisfy the Cauchy-Riemann equation. Therefore, the partial derivative along the real and imaginary axes, that is the sensitivity, must be the same

$$\Lambda_{l,n}^{t}(k) = \Lambda_{l,n}^{t,rr}(k) + j\Lambda_{l,n}^{t,ir}(k) = \Lambda_{l,n}^{t,ii}(k) - j\Lambda_{l,n}^{t,ri}(k) \tag{27}$$

$$\begin{aligned} \Lambda_{l,n}^{t,rr}(k) &= \Lambda_{l,n}^{t,ii}(k) \\ \Lambda_{l,n}^{t,ir}(k) &= -\Lambda_{l,n}^{t,ri}(k) \end{aligned} \tag{28}$$

Bring formula (21)-(28) into (20), and (20) becomes

$$\begin{aligned}\nabla w_{l,n}J(k) &= \Lambda^{rr}_{l,n}(k)\frac{\partial J}{\partial y^r_l} + \Lambda^{ir}_{l,n}(k)\frac{\partial J}{\partial y^i_l} \\ &\quad + j\Lambda^{ri}_{l,n}(k)\frac{\partial J}{\partial y^r_l} + j\Lambda^{ii}_{l,n}(k)\frac{\partial J}{\partial y^i_l} \\ &= \left(\frac{\partial J}{\partial y^r_l} + j\frac{\partial J}{\partial y^i_l}\right)\left(\Lambda^{rr}_{l,n}(k) + j\Lambda^{ri}_{l,n}(k)\right) \\ &= -\sum_{l=1}^{N}\varepsilon_l(k)(\Lambda^{rr}_{l,n}(k) - j\Lambda^{ir}_{l,n}(k)) \\ &= -\sum_{l=1}^{N}\varepsilon_l(k)\left(\Lambda_{l,n}(k)\right)^* \end{aligned} \tag{29}$$

$$\begin{aligned}\left(\Lambda_{l,n}(k)\right)^* &= \begin{bmatrix}\Lambda^{rr}_{l,n}(k) & \Lambda^{ri}_{l,n}(k) \\ \Lambda^{ir}_{l,n}(k) & \Lambda^{ii}_{l,n}(k)\end{bmatrix} \\ &= \begin{bmatrix}\dfrac{\partial y^r_l(k)}{\partial w^r_{l,n}(k)} & \dfrac{\partial y^r_l(k)}{\partial w^i_{l,n}(k)} \\ \dfrac{\partial y^i_l(k)}{\partial w^r_{l,n}(k)} & \dfrac{\partial y^i_l(k)}{\partial w^i_{l,n}(k)}\end{bmatrix}\end{aligned} \tag{30}$$

where (30) is called sensitive function. Expand it to get

$$\begin{aligned}&\left(\Lambda_{l,n}(k)\right)^* \\ &= \begin{bmatrix}\psi'^r(k) & 0 \\ 0 & \psi'^i(k)\end{bmatrix} \times \begin{bmatrix}\dfrac{\partial u^r_l(k)}{\partial w^r_{l,n}(k)} & \dfrac{\partial u^r_l(k)}{\partial w^i_{l,n}(k)} \\ \dfrac{\partial u^i_l(k)}{\partial w^r_{l,n}(k)} & \dfrac{\partial u^i_l(k)}{\partial w^i_{l,n}(k)}\end{bmatrix}\end{aligned} \tag{31}$$

$$\begin{aligned}&\begin{bmatrix}\Lambda^{rr}_{l,n}(k) & \Lambda^{ri}_{l,n}(k) \\ \Lambda^{ir}_{l,n}(k) & \Lambda^{ii}_{l,n}(k)\end{bmatrix} \\ &= \begin{bmatrix}\psi'^r(k-1) & 0 \\ 0 & \psi'^i(k-1)\end{bmatrix} \\ &\quad \times \left\{\sum_{\alpha=1}^{N}\left(\begin{bmatrix}w^r_{l,\alpha+p+1}(k-1) & -w^i_{l,\alpha+p+1}(k-1) \\ w^r_{l,\alpha+p+1}(k-1) & w^i_{l,\alpha+p+1}(k-1)\end{bmatrix}\right.\right. \\ &\quad \left.\left.\times \begin{bmatrix}\Lambda^{\alpha,rr}_{l,n}(k-1) & \Lambda^{\alpha,ri}_{l,n}(k-1) \\ \Lambda^{\alpha,ir}_{l,n}(k-1) & \Lambda^{\alpha,ii}_{l,n}(k-1)\end{bmatrix}\right) \\ &\quad + \begin{bmatrix}\delta_{\ln}X^r_n(k-1) & -\delta_{\ln}X^i_n(k-1) \\ \delta_{\ln}X^i_n(k-1) & \delta_{\ln}X^r_n(k-1)\end{bmatrix}\Bigg\}\end{aligned} \tag{32}$$

By extending the method in [24] from real-value to complex-value, the updating formula of complex sensitive function can be given as (32). Its simple expression form can be written as

$$\begin{aligned}(\Lambda_{l,n}(k))^* &= \left\{\psi^*(k)\right\}' \\ &\quad \times \left[\sum_{\alpha=1}^{N} w^*_{l,\alpha+p+1}(k)\left(\Lambda^{\alpha}_{l,n}(k-1)\right)^* + \delta_{\ln}X^*_n(k)\right]\end{aligned} \tag{33}$$

where

$$\delta_{\ln} = \begin{cases}1, & l = n \\ 0, & l \neq n\end{cases} \tag{34}$$

is the Kronecker delta [15]. The weight update equation is

$$w_{l,n}(k+1) = w_{l,n}(k) + \Delta w_{l,n}(k) \tag{35}$$

$$\Delta w_{l,n}(k) = \eta\sum_{l=1}^{N}\varepsilon_l(k)\left(\Lambda_{l,n}(k)\right)^* \tag{36}$$

The weight update of CFL-FCRNN is finally shown in the following formula

$$\begin{aligned}w_{l,n}(k+1) &= w_{l,n}(k) + \eta\sum_{l=1}^{N}\Big\{\varepsilon_l(k) \times \left\{\psi^*(k)\right\}' \\ &\quad \times \sum_{\alpha=1}^{N} w^*_{l,\alpha+p+1}(k)\left(\Lambda^{\alpha}_{l,n}(k-1)\right)^* \\ &\quad + \delta_{\ln}X^*_n(k)\Big\}\end{aligned} \tag{37}$$

B. CRTRL ALGORITHM FOR CPCFLRNN

According to the above method, the CRTRL learning algorithm for CPCFLRNN can be derived. Let $y_{t,1}(k)$ be the output of module t, and the error of this module at time k can be obtained by subtracting the model output signals from the PA output signals

$$\varepsilon_t(k) = d(k-t+1) - y_{t,1}(k) = \varepsilon^r_t(k) + j\varepsilon^i_t(k) \tag{38}$$

$$\varepsilon^r_t(k) = d^r(k-t+1) - y^r_{t,1}(k)$$

$$\varepsilon^i_t(k) = d^i(k-t+1) - y^i_{t,1}(k) \tag{39}$$

Since the baseband output signal is complex, the cost function should be extended to complex domain, which is given as follows

$$\begin{aligned}J(k) &= \sum_{t=1}^{M}\gamma^{t-1}(k)|\varepsilon_t(k)|^2 \\ &= \sum_{t=1}^{M}\gamma^{t-1}(k)[\varepsilon_t(k)\varepsilon^*_t(k)] \\ &= \sum_{t=1}^{M}\gamma^{t-1}(k)[(\varepsilon^r_t)^2 + (\varepsilon^i_t)^2]\end{aligned} \tag{40}$$

where $\gamma(k)(0 < \gamma \leq 1)$ is a forgetting factor for determine the weight of the individual modules. Updating the weight in the steepest descent direction

$$\Delta w_{l,n}(k) = -\eta\frac{\partial}{\partial w_{l,n}(k)}\left(\sum_{t=1}^{M}\gamma^{t-1}(k)|\varepsilon_t(k)|^2\right) \tag{41}$$

With reference to (23)-(26), the expression of sensitive function at the time k of each module in CPCFLRNN can be given as

$$\begin{bmatrix}\Lambda^{rr,j}_{l,n,t}(k) & \Lambda^{ri,j}_{l,n,t}(k) \\ \Lambda^{ir,j}_{l,n,t}(k) & \Lambda^{ii,j}_{l,n,t}(k)\end{bmatrix} = \begin{bmatrix}\dfrac{\partial y^r_{t,j}(k)}{\partial w^r_{l,n}(k)} & \dfrac{\partial y^r_{t,j}(k)}{\partial w^i_{l,n}(k)} \\ \dfrac{\partial y^i_{t,j}(k)}{\partial w^r_{l,n}(k)} & \dfrac{\partial y^i_{t,j}(k)}{\partial w^i_{l,n}(k)}\end{bmatrix} \tag{42}$$

The element in the sensitive function matrix represents the degree of change of the output of *lth* neuron relative to the

weight at time k. Referring to (32), the update equation of sensitive functions of each module is

$$\begin{aligned}
&\begin{bmatrix} \Lambda^{rr}_{l.n,t}(k) & \Lambda^{ri}_{l.n,t}(k) \\ \Lambda^{ir}_{l.n}(k) & \Lambda^{ii}_{l.n}(k) \end{bmatrix} \\
&= \begin{bmatrix} \psi'^{r}(k-1) & 0 \\ 0 & \psi'^{i}(k-1) \end{bmatrix} \\
&\quad \times \left\{ \sum_{\alpha=1}^{N} \left(\begin{bmatrix} w^{r}_{l,\alpha+p+1}(k-1) & -w^{i}_{l,\alpha+p+1}(k-1) \\ w^{r}_{l,\alpha+p+1}(k-1) & w^{i}_{l,\alpha+p+1}(k-1) \end{bmatrix} \right.\right. \\
&\quad \left.\left. \times \begin{bmatrix} \Lambda^{rr,\alpha}_{l.n,t}(k-1) & \Lambda^{ri,\alpha}_{l.n,t}(k-1) \\ \Lambda^{ir,\alpha}_{l.n,t}(k-1) & \Lambda^{ii,\alpha}_{l.n,t}(k-1) \end{bmatrix} \right) \right. \\
&\quad \left. + \begin{bmatrix} \delta_{\ln} X^{r}_{n}(k-1) & -\delta_{\ln} X^{i}_{n}(k-1) \\ \delta_{\ln} X^{i}_{n}(k-1) & \delta_{\ln} X^{r}_{n}(k-1) \end{bmatrix} \right\}
\end{aligned} \tag{43}$$

Simplify it to:

$$\begin{aligned}
(\Lambda^{t}_{l,n}(k))^{*} = \left\{\psi^{*}(k)\right\}' \times \left[\sum_{\alpha=1}^{N} w^{*}_{l,\alpha+p+1}(k) \left(\Lambda^{t,\alpha}_{l,n}(k-1)\right)^{*} \right. \\
\left. + \ \delta_{\ln} X^{*}_{t,n}(k) \right]
\end{aligned} \tag{44}$$

Finally, the weight update equation of CPCFLRNN can be given as

$$\begin{aligned}
&w_{l,n}(k+1) \\
&= w_{l,n}(k) \\
&+ \eta \left(\begin{array}{l} \sum_{t=1}^{M} \lambda^{t-1}(k) e_t(k) \{\psi^{*}(u_{t,l}(k))\}' \\ \times \left[\sum_{\alpha=1}^{N} w^{*}_{1,\alpha+p+1}(k) \left(\Lambda^{t,\alpha}_{l,n}(k-1)\right)^{*} + \delta_{\ln} X^{*}_{t,n}(k) \right] \end{array} \right)
\end{aligned} \tag{45}$$

C. COMPLEXITY ANALYSIS

On the other hand, the CRTRL algorithm also has some certain limitations. When the number of total neurons is N, its computational complexity increases as $O(N^4)$. As the total number of neurons increases, so does the complexity increases too. Therefore, the structure of the selected neural network needs to reduce the complexity of the learning algorithm. For the total number of MN neurons in complex PRNN (CPRNN) model using the CRTRL, it needs $O(MN^4)$ arithmetic operations. By contrast, the computational requirement of conventional fully connected recurrent neural network (FCRNN) with the CRTRL algorithm is $O(M^4N^4)$ [23]. Therefore, the modular and recursive structure of CPRNN keeps the memory size from increasing with the length of the training sequence, which is suitable for real time processing. According to (44) and (45), the calculation requirement of the proposed model can be calculated as $O(M^4N^4 + 3MP + 7M + 3)$. Although the computational complexity is slightly increased, the model can provide the better compensation performance.

IV. SIMULATION AND EXPERIMENTAL RESULTS

In order to verify the modeling performance of the above neural network and the appropriate model parameters, the first step is to carry out forward modeling for PA. Forward modeling also called the behavioral model of the PA, which refers to using the input signal and the measured PA output signal to conduct the model under the same sampling rate of the system [7]. Under the condition that the model is sufficiently accurate, the calculated model output signal will

$$NMSE_{dB} = 10 \times \log_{10} \frac{\frac{1}{10} \sum_{j=1}^{N} |y_{est}(n) - y_{rea}(n)|^2}{\frac{1}{10} \sum_{j=1}^{N} |y_{rea}(n)|^2} \tag{46}$$

approach the actual output of the PA with a minor error. As the popular metrics, the normalized mean square error (NMSE) between the expected signal and the estimated signal is adopted to evaluate the model accuracy [16]. The NMSE is defined as (46). The CMP model, the CPRNN model, the CFL-FCRNN model and CPCFLRNN model are selected for PA model performance comparison. The sample signal comes from the actual measured GaN Class-F PA signals, with a total of 20000 data. The signal is a dual carrier LTE signal with 30MHz bandwidth, where the peak-to-average power radio (PAPR) of the signal is 9.91 dB. Both the I/Q imbalance, dc-offset and PA nonlinear distortions are presented in the transmitter. The amplitude imbalance is 2 dB, as well as the I/Q phase imbalance is 3°. And the dc-offset values of 3% and 5% are set for the I and Q channel, respectively. In the model process, 10000 data from the sample signals are used for transmitter modeling, another 4000 data are used to verify the model performance of different models.

A. SETTING AND SIMULATION

The neural network model has good approximation ability for nonlinear system by introducing activation function with nonlinear factors. Therefore, how to determine the activation function is the first step. In order to process the complex PA signals, the complex form of log sig, which is commonly used in neural networks, together with the elementary transcendental function (ETF) tanh, which is provided by [18], are compared as activation functions.

$$\log sig(z) = \frac{1}{1 + e^{-z}} \tag{47}$$

$$\tanh(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \tag{48}$$

The model was trained in MATLAB environment, and the performance of the model was compared and observed. The number of input neurons and output neurons in the CFL-FCRNN is 4, the forgetting factor $\gamma = 0.5$, and the learning rate $\eta = 0.05$. The CPCFLRNN contains M CFL-FCRNN modules, and the structure and parameter settings of each module are the same. In order to achieve the optimal modeling effect, the appropriate values of M need to be selected. Here, the range of M value is set to 2-6, and the simulation results are shown in Table 2. According to

TABLE 2. Performance of different parameters.

Activation Function	Number of Modules	NMSE(dB)
tanh(z)	M=2	-34.35
	M=3	-38.02
	M=4	-38.76
	M=5	-38.81
	M=6	-38.83
logsig(z)	M=2	-32.03
	M=3	-37.69
	M=4	-38.55
	M=5	-38.58
	M=6	-38.65

the simulation results, as the value of M increases, the modeling effect of CPCFLRNN will be moderately better, but the increasing number of modules will increase the computational complexity of the model, thus increasing the modeling run time. But the modeling performance will actually only be slightly improved. After comprehensive consideration, the value of M is set to 4. In addition, the simulation results also prove that it has better performance to use the elementary transcendental function, which can satisfy the requirements for dealing with the signal with I/Q imbalance using the fully complex activation function. Therefore, the complex value form of tanh function is selected as the activation function.

The M, N parameters is (3, 7) for CMP, which is determined by the optimization process. The CPRNN model parameter settings are the same as CPCFLRNN model. The NMSE value of the four models is shown in Table 3. Compared with the CMP model, the CPRNN and CFL-FCRNN model can give 2 dB improvement of NMSE. And the CPCFLRNN model has more accurate and better modeling effect, can give 4 dB improvement of NMSE.

TABLE 3. Performance of different models.

Model	Parameter	NMSE(dB)
CMP	N=7,M=3	-34.65
CPRNN	p=4,N=4,M=4	-36.83
CFL-FCRNN	p=4,N=4	-36.55
CPCFLRNN	p=4,N=4,M=4	-38.76

The notable characteristic of PA in the transmitter is the nonlinear with memory effects. Under the influence of the nonlinear characteristic, the amplitude and phase of the transmitter output signals do not change linearly with the amplitude of the input signal, which will take on nonlinear distortions. Therefore, the the dynamic AM/AM and AM/PM curves are intuitively used to describe the characteristics of the transmitter, which are given in Fig. 5 and Fig. 6. As can be seen from the figures, the nonlinear characteristics of the transmitter are reflected in the following features. That is to say, when the amplitude of input signal is small, the amplitude of output signal increases linearly. And when the amplitude of the input signal increases to a certain extent, the amplitude of the output signal no longer grows linearly, and the slope of the curve is reduced. At the same time, the signal enters the nonlinear area of the PA. Finally, the amplitude of the output signal stops growing, and it enters the saturation area of the PA. The memory effect of PAs is reflected in the divergence of the AM/AM and AM/PM curves. As the signal amplitude increases, the divergence degree decreases, which appears a compression trend. This phenomenon is called "gain compression" [25]. It can be seen that the CPCFLRNN model can not only well describe the nonlinear memory effects of the PA, but also prove that the I/Q imbalance and dc offset have the influence

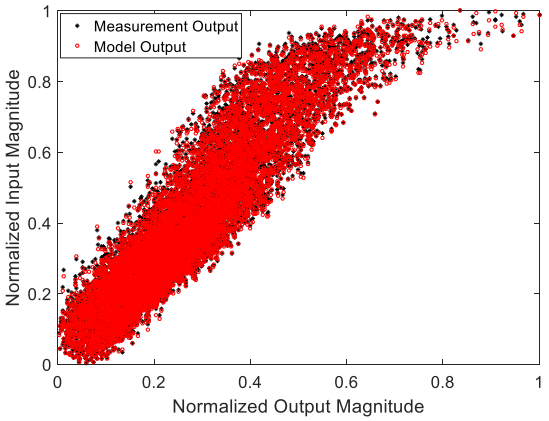


FIGURE 5. The dynamic AM/AM characteristics of the transmitter system from the measurement output and model output.

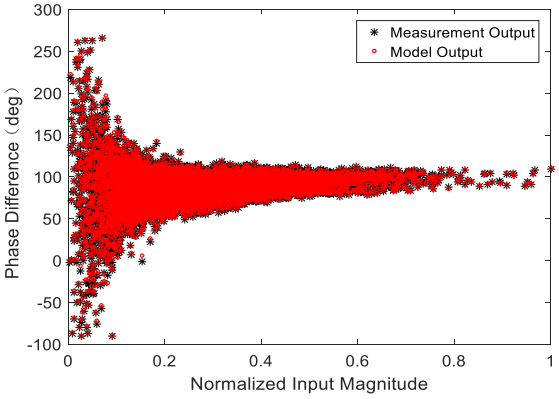


FIGURE 6. The dynamic AM/PM characteristics of the transmitter system from the measurement output and model output.

on the gain and phase of transmitter. In order to observe the accuracy of the model more directly, the model results can also be described by the power spectrum density (PSD) of the signals. The CPCFLRNN model prediction outputs, the measurement outputs and error signals between the model and the measurement output for the two-carrier LTE signals are given in Fig. 7. It can be seen that the PSD can be well matched in both in-band channels and alternate channels, and the PSD of the error signals are below −50dBm /Hz.

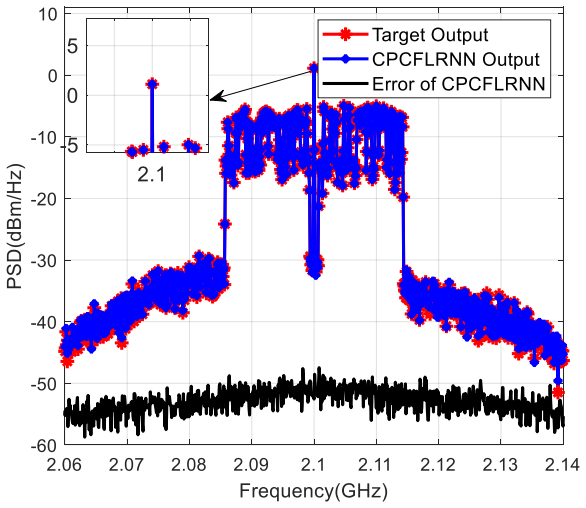


FIGURE 7. The measured and modeled PSD of the transmitter output using CPCFLRNN model.

The modeling performance of CMP degrades to a certain extent. The CPRNN model can deal with complex-valued signals, which do not have Chebyshev polynomial structure and cannot map the nonlinear relation of input signal, so the performance is slightly better. The CFL-FCRNN and CPCFLRNN models, which has the ability to process complex-valued signals, and can better represent the nonlinear memory effects of the transmitter. Moreover, the CRTRL algorithm using the fully complex activation function can process the I/Q signal concurrently, which can give optimal performance.

B. EXPERIMENTAL MEASUREMENT RESULTS

The feasibility of the CPCFLRNN model in transmitter behavioral modeling has been verified by the simulation results. Theoretically, the higher the accuracy of the model, the better the DPD linearization effect. However, the behavioral model is only an abstract model that is close to the transmitter, and it does not prove that the DPD can work well. In order to verify the performance of the model in the DPD system, the proposed model is applied to a complete DPD experimental system. The experimental system includes RF PAs, computer, vector signal generator (VSG), vector signal analyzer (VSA). The single-device GaN Class-F PA worked at 2.1GHz was used in the experimental validation. Two-carrier LTE signal with 30MHz bandwidth can be generated through ADS2017, which is further downloaded to VSG

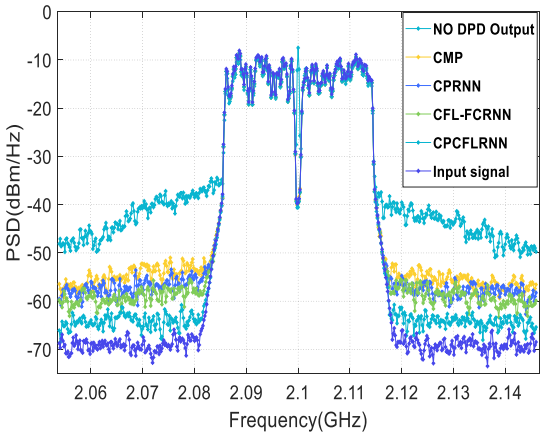


FIGURE 8. DPD comparison of different models for LTE signal.

and modulated to the RF frequency. Meanwhile, the different I/Q imbalance value can be set in the VSG to mimic the real modulator. Then, the RF signals pass through the target PA to obtain the output signals, which are attenuated by the attenuator to enable the VSA (N9010A) to collect the signal for conducting DPD model. The Matlab can perform the inverse model according to the input and output signals of the transmitter. The signal is processed through the inverse model to get the DPD signal. Then, the DPD output signals pass through the PA according to the above process, so as to get the output data of the PA after the DPD.

In order to compare the linearization performance of the different models, the same LTE signals are used for the predistortion verification using the CMP model, the CPRNN model, the CFL-FCRNN model and the CPCFLRNN model, respectively. Fig. 8 shows the compensation effect of four models for the transmitter with I/Q imbalance and dc offset. It can be seen from that all four models have the ability to compensate the transmitter distortions and impairments. The output signals of transmitter without linear compensation have the high out-of-band distortion compared with the input signal. The CMP model presents the worst compensation effect. The compensation results of the CPRNN and the CFL-FCRNN model almost coincide, and the correction ability is similar, while the compensation curve of the proposed CPCFLRNN model is the closest to the input signal, and the compensation effect is the best. These conclusions are also consistent with those obtained by simulation. As can be seen from Table 4, the CMP model gives the worst effect, which can only give the adjacent channel power ratio (ACPR) improvement about 15 dB. And the CPRNN and the CFL-FCRNN models can give the similar ACPR improvement of 16 dB. The proposed CPCFLRNN model can give the best compensation effect, where the ACPR can be improvement about 19 dB. Both the graphical and quantitative results prove that the CPCFLRNN model has satisfactory compensation effects for transmitter with nonlinear distortions and impairments.

TABLE 4. ACPR performance of different models.

DPD Models	ACPR(dB) -15MHz	ACPR(dB) +15MHz
Without DPD	-30.44	-29.84
CMP	-45.78	-45.73
CPRNN	-46.07	-46.14
CFL-FCRNN	-46.19	-46.20
CPCFLRNN	-49.79	-49.62

V. CONCLUSION

The CPCFLRNN model including the M-order CFL-FCRNN module is proposed in this article, and the CRTRL learning algorithm for the model parameter extraction is derived in detaill. The Chebyshev structure is used to fit the nonlinearity of the transmitter input, and the PRNN structure is used to process the complex-valued signals. The simulation and experiment results prove that the proposed CPCFLRNN structure can achieve the obvious advantages compared with the other models. The optimal model order M and the best fully complex activation function of the CPCFLRNN model can be obtained through simulations. To verify the linearization ability of the proposed model, the class-F PA driven by the 30MHz LTE signals including I/Q imbalance and dc offset is used for the experimental verication. Experimental results show that the CPCFLRNN DPD model can not only process complex-valued signals including I/Q imbalance and dc-offset, but also achieve better linearization compensation for the transmitter compared with the other models.

REFERENCES

[1] Z. Zhu, X. Huang, and H. Leung, "Joint I/Q mismatch and distortion compensation in direct conversion transmitters," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2941–2951, Jun. 2013.

[2] S. Wang, M. Roger, J. Sarrazin, and C. Lelandais-Perrault, "An efficient method to study the tradeoff between power amplifier efficiency and digital predistortion complexity," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 11, pp. 741–744, Nov. 2019.

[3] X. Yu and H. Jiang, "Digital predistortion using adaptive basis functions," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 12, pp. 3317–3327, Dec. 2013.

[4] P. Jardin and G. Baudoin, "Filter lookup table method for power amplifier linearization," *IEEE Trans. Veh. Technol.*, vol. 56, no. 3, pp. 1076–1087, May 2007.

[5] J. Ren, "Digital predistorter for short-wave power amplifier with improving index accuracy of lookup table based on FPGA," *IEEE Access*, vol. 7, pp. 182881–182885, 2019.

[6] L. Anttila, P. Handel, and M. Valkama, "Joint mitigation of power amplifier and I/Q modulator impairments in broadband direct-conversion transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 4, pp. 730–739, Apr. 2010.

[7] M. Rawat, K. Rawat, and F. M. Ghannouchi, "Adaptive digital predistortion of wireless power Amplifiers/Transmitters using dynamic real-valued focused time-delay line neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 1, pp. 95–104, Jan. 2010.

[8] Y. Zhang, Y. Li, F. Liu, and A. Zhu, "Vector decomposition based time-delay neural network behavioral model for digital predistortion of RF power amplifiers," *IEEE Access*, vol. 7, pp. 91559–91568, 2019.

[9] J. Sun, W. Shi, Z. Yang, J. Yang, and G. Gui, "Behavioral modeling and linearization of wideband RF power amplifiers using BiLSTM networks for 5G wireless systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 10348–10356, Nov. 2019.

[10] M. Isaksson, D. Wisell, and D. Ronnow, "Wide-band dynamic modeling of power amplifiers using radial-basis function neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 11, pp. 3422–3428, Nov. 2005.

[11] D. Wang, M. Aziz, M. Helaoui, and F. M. Ghannouchi, "Augmented real-valued time-delay neural network for compensation of distortions and impairments in wireless transmitters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 1, pp. 242–254, Jan. 2019.

[12] P. Jaraut, M. Rawat, and F. M. Ghannouchi, "Composite neural network digital predistortion model for joint mitigation of crosstalk, I/Q imbalance, nonlinearity in MIMO transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 11, pp. 5011–5020, Nov. 2018.

[13] S. Lajnef, N. Boulejfen, A. Abdelhafiz, and F. M. Ghannouchi, "Two-dimensional Cartesian memory polynomial model for nonlinearity and I/Q imperfection compensation in concurrent dual-band transmitters," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 1, pp. 14–18, Jan. 2016.

[14] H. Zhao and J. Zhang, "Functional link neural network cascaded with chebyshev orthogonal polynomial for nonlinear channel equalization," *Signal Process.*, vol. 88, no. 8, pp. 1946–1957, Aug. 2008.

[15] H. Zhao and J. Zhang, "Pipelined chebyshev functional link artificial recurrent neural network for nonlinear adaptive filter," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 40, no. 1, pp. 162–172, Feb. 2010.

[16] M. Li, J. Liu, Y. Jiang, and W. Feng, "Complex-chebyshev functional link neural network behavioral model for broadband wireless power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 6, pp. 1979–1989, Jun. 2012.

[17] T. Needham, *Visual Complex Analysis*. Oxford, U.K.: Oxford Univ. Press, 1997.

[18] T. Kim and T. Adalá, "Approximation by fully complex multilayer perceptrons," *Neural Comput.*, vol. 15, no. 7, pp. 1641–1666, Jul. 2003.

[19] E. G. Lima, T. R. Cunha, and J. C. Pedro, "A physically meaningful neural network behavioral model for wireless transmitters exhibiting PM–AM/PM–PM distortions," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 12, pp. 3512–3521, Dec. 2011.

[20] J. Zhang, P. Lei, S. Hu, M. Zhu, Z. Yu, B. Xu, and K. Qiu, "Functional-link neural network for nonlinear equalizer in coherent optical fiber communications," *IEEE Access*, vol. 7, pp. 149900–149907, 2019.

[21] S. Zhang and W. X. Zheng, "Recursive adaptive sparse exponential functional link neural network for nonlinear AEC in impulsive noise environment," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 4314–4323, Sep. 2018.

[22] D. G. Stavrakoudis and J. B. Theocharis, "Pipelined recurrent fuzzy neural networks for nonlinear adaptive speech prediction," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 37, no. 5, pp. 1305–1320, Oct. 2007.

[23] S. Lee Goh and D. P. Mandic, "Nonlinear adaptive prediction of complex-valued signals by complex-valued PRNN," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1827–1836, May 2005.

[24] S. Haykin, *Kalman Filtering and Neural Networks*. Hoboken, NJ, USA: Wiley, 2001.

[25] R. W. Santucci, M. K. Banavar, C. Tepedelenlioalu, and A. Spanias, "Energy-Efficient Distributed Estimation by Utilizing a Nonlinear Amplifier," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 1, pp. 302–311, Jan. 2014.

MINGYU LI (Member, IEEE) received the Ph.D. degree in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2009.

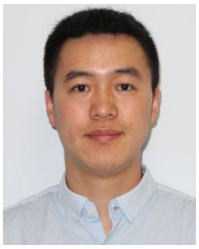
From 2012 to 2013, he was a Research Fellow with The University of Kitakyushu. He is currently an Associate Professor with the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China. His current research interests include RF/microwave transceiver design, statistical and adaptive signal processing for wireless communications, and behavioral modeling and linearization for RF power amplifiers.

ZHENDONG CAI received the B.Sc. degree in opto-electronics information science and engineering from the Hunan University of Science and Technology, Xiangtan, China, in 2018. He is currently pursuing the master's degree with the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China.

His research interests include communication signal processing, artificial intelligence, and digital predistortion implementation in FPGA.

YAO YAO received the M.S. degree from the School of Electronic Engineering, Chengdu University of Technology, Sichuan, China, in 2011, and the Ph.D. degree in circuits and systems from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2019.

She is currently a Lecturer with the School of Intelligent Technology and Engineering, Chongqing University of Science and Technology, Chongqing, China. Her current research interests include digital predistortion techniques and software defined radio.



CHANGZHI XU (Member, IEEE) received the B.E. degree in electronic information engineering and the M.S. degree in biomedical engineering from Xidian University, Xi'an, China, in 2006 and 2009, respectively. He is currently pursuing the Ph.D. degree with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing, China.

He is with Xi'an Branch of China Academy of Space Technology, Xi'an. His research interests include satellite communications, laser communications, and software defined radio systems.

YI JIN received the B.E. degree in communication and information engineering from the Nanjing University of Information Science and Technology, Nanjing, China, in 2005, and the Ph.D. degree in communication and information engineering from Southeast University, Nanjing, in 2013.

He is currently with Xi'an Branch of China Academy of Space Technology, Xi'an. His research interests include communication signal processing, satellite communications, and networking.

XUGUANG WANG received the B.Sc. degree in electronic engineering from the Hunan University of Science and Technology, Xiangtan, China, in 2019. He is currently pursuing the master's degree with the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China.

His research interests include communication signal processing and digital predistortion implementation in FPGA.