



A Basis Function Generation Based Digital Predistortion Concurrent Neural Network Model for RF Power Amplifiers

SHAO Jianfeng¹, HONG Xi¹, WANG Wenjie¹,

LIN Zeyu², LI Yunhua²

(1. Xi'an Jiaotong University, Xi'an 710049, China;

2. ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTECOM.202501009

<https://kns.cnki.net/kcms/detail/34.1294.TN.20250224.1024.002.html>,
published online February 24, 2025

Manuscript received: 2023-09-23

Abstract: This paper proposes a concurrent neural network model to mitigate non-linear distortion in power amplifiers using a basis function generation approach. The model is designed using polynomial expansion and comprises a feedforward neural network (FNN) and a convolutional neural network (CNN). The proposed model takes the basic elements that form the bases as input, defined by the generalized memory polynomial (GMP) and dynamic deviation reduction (DDR) models. The FNN generates the basis function and its output represents the basis values, while the CNN generates weights for the corresponding bases. Through the concurrent training of FNN and CNN, the hidden layer coefficients are updated, and the complex multiplication of their outputs yields the trained in-phase/quadrature (I/Q) signals. The proposed model was trained and tested using 300 MHz and 400 MHz broadband data in an orthogonal frequency division multiplexing (OFDM) communication system. The results show that the model achieves an adjacent channel power ratio (ACPR) of less than -48 dB within a 100 MHz integral bandwidth for both the training and test datasets.

Keywords: basis function generation; digital predistortion; generalized memory polynomial; dynamic deviation reduction; neural network

Citation (Format 1): SHAO J F, HONG X, WANG W J, et al. A basis function generation based digital predistortion concurrent neural network model of RF power amplifier [J]. *ZTE Communications*, 2025, 23(1): 71 - 77. DOI: 10.12142/ZTECOM.202501009

Citation (Format 2): J. F. Shao, X. Hong, W. J. Wang, et al., "A basis function generation based digital predistortion concurrent neural network model of RF power amplifier," *ZTE Communications*, vol. 23, no. 1, pp. 71 - 77, Mar. 2025. doi: 10.12142/ZTECOM.202501009.

1 Introduction

With the growing demand for high-throughput wireless communications, system bandwidths continue to expand. However the use of orthogonal frequency division multiplexing (OFDM) modulation results in a high peak-to-average power ratio (PAPR)^[1]. The nonlinear behavior of power amplifiers (PAs) often leads to compression of high-dynamic-range signals, causing significant signal transmission distortion and upgraded error vector magnitude (EVM) at the receiver, even in scenarios with a high signal-to-noise ratio (SNR)^[2]. Therefore, PA behavior modeling and corresponding anti-compression techniques, such as digital predistortion (DPD), play an important role in establishing a robust wireless communication system^[3].

The wider bandwidth leads to the existing polynomial expansion models less precise for PA behavior modeling and digital predistortion techniques. Traditional DPD methods,

like the generalized memory polynomial (GMP)^[4] or dynamic deviation reduction (DDR) model^[5], rely on polynomial expansion. However, increasing bandwidth requires higher polynomial orders, which introduces a high correlation among the polynomial's high-order terms, thereby making the traditional models sensitive to noise^[6]. Additionally, conventional models require more delay taps and computational resources for high bandwidth signal transmission to radio frequency (RF) PA, complicating their integration with nonlinear bases^[7].

Recent research and data analysis indicate that neural networks (NNs) have excellent performance in data feature extraction, data fitting, and model generalization. As a result, the use of NN in DPD has received increased attention and application^[8]. For example, a feed-forward NN was proposed in Ref. [9], achieving improvements in both linearity and stability. Similarly, in Refs. [10] and [11], two-stage network models were proposed, achieving good performance metrics such as adjacent channel power ratio (ACPR) and normalized mean square error (NMSE). In Ref. [12], a novel residual NN structure connects residual learning and PA nonlinearity, providing

This work was supported by ZTE Industry-University-Institute Cooperation Funds under Grant No. HC-CN-20220722010.

better performance than conventional models.

Existing methods often achieve better performance by increasing the number of parameters, which in turn significantly raises model complexity. In real-time applications, optimizing model complexity is a crucial aspect of the integration of RF-DPD and NN. To reduce both the training data length and the number of basis functions, Ref. [13] proposed a model that combines an efficient uncorrelated equation selection mechanism with orthogonal least squares. Another model proposed in Ref. [14] is a sparse gated dynamic NN DPD model that linearizes the PA for varying transmission configurations, thereby reducing model complexity.

Several new models have been introduced to address the issues of performance and complexity that the classic NN model could not handle. These new models can effectively linearize RF PAs in broadband communications while reducing complexity. For instance, Ref. [15] introduced a novel augmented convolutional NN-based DPD that can linearize concurrent multiband PAs. Additionally, Ref. [16] proposed a novel block-oriented time-delay NN to alleviate the deterioration of linearization performance. Ref. [17] proposed a novel RNN-based behavioral model that reduces complexity and enhances linearization performance by applying the complete phase-gated Just Another Network (JANET) unit. These new models are more suitable for PA-DPD in wide bandwidths and provide better nonlinear modeling capabilities to extract PA features for DPD.

In this paper, we present a DPD concurrent NN model based on an FNN and a convolutional neural network (CNN). The basic inputs of this model are obtained through polynomial expansion of the GMP+DDR model. This concurrent NN model overcomes nonlinear distortions such as amplitude modulation to amplitude modulation (AM-AM) distortion and amplitude modulation to phase modulation (AM-PM) distortion in RF-PAs. Our experimental data comprises OFDM signals with bandwidths of 300 MHz or 400 MHz. Our model aims at reducing non-linear compression to improve the ACPR of the output signal, with a target of at least -48 dB within a 100 MHz integral bandwidth. In addition to its function of basis function generation, the proposed model seeks to have engineering feasibility and low complexity.

2 Mathematical Model of DPD

To enhance the efficiency of PAs, existing methods aim to compress the power regression range as much as possible. The PA's gain does not maintain linearity when the input signal amplitude of the amplifier output section approaches the 1 dB compression point. Typically, power amplifiers feature nonlinear effects such as AM-AM, AM-PM, and time memory. Traditional narrowband amplifiers can be modeled using polynomial expressions, with the Volterra series serving as one of the most representative mathematical models. Eq. (1) describes a P-order and M-length Volterra series.

$$\tilde{y}(n) = \sum_{p=0}^P \sum_{m=0}^M h_{p,m} \prod_{l=1}^p \tilde{x}(n - m_l) \quad (1),$$

where m denotes the length of the memory effect and p denotes the maximum order of the basis. Similar to the solution of the Wiener filter, the concatenated signal terms in Eq. (1) serve as the bases of polynomial expansion, while the corresponding coefficients $h_{p,m}$ are their respective weights.

While Volterra can effectively describe nonlinear compression with memory effects, the concatenated multiplication of signals introduces a great deal of computational effort and complexity. The memory polynomial (MP) model^[17] replaces the concatenated multiplication of signals with a modulus-valued term based on the Volterra series. The MP model can be simplified in the time domain:

$$\tilde{y}(n) = \sum_{p=0}^P \sum_{m=0}^M h_{2p,m} \tilde{x}(n - m) |\tilde{x}(n - m)|^{2p} \quad (2).$$

In the MP model, the basis function becomes the signal multiplied by the signal's ground modulus term. This modification leads to a significant reduction in the computational effort required by the network. However, as the bandwidth increases further, the MP model faces the issue of lower accuracy.

The GMP model extends the composition of the bases based on the MP model. It can describe the nonlinear compression model at larger bandwidths and can be simplified in the time domain as:

$$\tilde{y}(n) = \sum_{p=0}^P \sum_{l=0}^L \sum_{m=0}^M h_{2p,l,m} \tilde{x}(n - l) |\tilde{x}(n - m)|^{2p} \quad (3).$$

The GMP model extends the influence of time memory effects in the composition of the basis functions, which is relevant to the scenario of wideband communication.

The DDR model^[5] is also built on the MP model. However, it differs from the GMP model by placing more emphasis on the aliasing effects of wideband signals. It can be represented in the time domain as:

$$\begin{aligned} \tilde{y}(n) \approx & \sum_{p=0}^P \sum_{k=0}^K a_{2p,i,m} |x(n)|^{2p} x(n - k) + \\ & \sum_{p=0}^P \sum_{l=0}^L b_{2p-2,j,n} |x(n)|^{2p-2} x^2(n - l) x(n - l) \end{aligned} \quad (4).$$

The DDR model can be divided into two parts. As shown in Eq. (4), the first part is the MP model, while the second part describes the nonlinear compression of the signal after aliasing under the memory effect.

In this paper, since the data we use are wideband signals and the main requirement for the proposed model is better performance, we use the GMP+DDR model as the reference mathematical model for this paper. It can be written as:

$$\tilde{y}(n) \approx \sum_{p=0}^P \sum_{i=0}^I \sum_{k=0}^K a_{2p,i,m} |x(n-i)|^{2p} x(n-k) + \sum_{p=0}^P \sum_{j=1}^J \sum_{l=1}^L b_{2p-2,j,n} |x(n-j)|^{2p-2} x^2(n-l)x(n-l) \quad (5).$$

Eq. (5) combines the features of the GMP and DDR models. Both the memory effect and aliasing of broadband signals are covered to ensure that the model can achieve optimal performance. The dataset composition will also refer to the mathematical model shown in Eq. (5).

3 Designing of DPD NN Model

3.1 Basis Function Generation and Recognition

To explain the generation of basis functions, we first clarify the input dataset structure. Our basis function formulation, based on the GMP+DDR model for wideband applications, draws inspiration from the methodologies presented in Refs. [4] and [5]. The intermodulation terms in Eq. (5) are highly suitable for modeling the wideband PA. Therefore, we establish the dataset format based on the fundamental components in Eq. (6).

$$\begin{aligned} \tilde{x}(n) = & [Re(x(n-12)), \dots, Re(x(n)), \dots, Re(x(n+11)), \\ & Im(x(n-12)), \dots, Im(x(n)), \dots, Re(x^2(n)), \dots, Im(x^2(n)), \dots, \\ & |x(n)|^2, \dots, |x(n)|^4, \dots, |x(n)|^6, \dots]^T, \\ X = & [\dots \tilde{x}(n-1), \tilde{x}(n), \tilde{x}(n+1), \dots] \end{aligned} \quad (6).$$

The dataset consists of multiple vectors, as shown in Eq. (6), indicating the input terms and memory depth. The input elements include the signal, the square of the signal, and the even-square term of the signal's modulus. And they all stem from the GMP+DDR model, as detailed in Eq. (5). The model has an order of 7 and a memory depth of 24, spanning from -12 to 11. Since the neural network library we use (PyTorch) is less compatible with complex numbers, the proposed model is trained using real-valued data. For this purpose, the real and imaginary parts of the signal are split and used to construct the dataset.

To linearize the bases of individual nonlinear terms, we propose the use of fully connected (FC) layers to combine all elements. This approach enables the number of basis elements to be established by the number of neurons within the FC layer. The output of each layer is then nonlinearly activated to generate a nonlinear basis. Additionally, based on the survey re-

sults, at least three FC layers are sufficient to produce the majority of nonlinear combinations. Then the outputs would be activated by the nonlinear function to ensure their nonlinearity.

In this paper, the basis generation network (BGN) based on an FNN is illustrated in Fig. 1. The weight matrix within the FC layer adjusts the coefficients of the input terms, which are optimized through training feedback. As shown in Fig. 1, the length of the FC layers decreases in the forward direction of the arrays. Therefore, the number of neurons and the output of each FC layer are decreased. It is similar to FNN selecting bases for each hidden layer. Regarding the activation function, Rectified Linear Unit (ReLU) leads to faster loss convergence compared to other activation functions based on test results. This improvement can be attributed to ReLU's superior sparsity. Consequently, each FC layer's output in the basis function generation model is activated by ReLU. $B_{i/q-1}$ and $B_{i/q-N}$ in Fig. 1 denote the real or imaginary parts of the first and N -th substrates generated, respectively. Since the DPD model is a real-valued training NN model, the BGN has two identical structures as shown in Fig. 1. The notation " i/q " represents in-phase or quadrature components, while " -1 " or " $-N$ " serves as a label for the bases. These labels have no real physical meaning and are solely used to distinguish the bases and correspond to the weights.

3.2 Structure of Concurrent NN Model

Fig. 2 depicts the proposed concurrent neural network model, comprising an FNN and a CNN.

The left side of Fig. 2 displays the FNN model utilized to generate basis functions, as described in Section 3.1. Since the proposed model is trained using real numbers, the FNN-based function generation model has two sets of three FC layers. The basis generation function only relies on the coefficients of each hidden layer in the FNN model. On the right

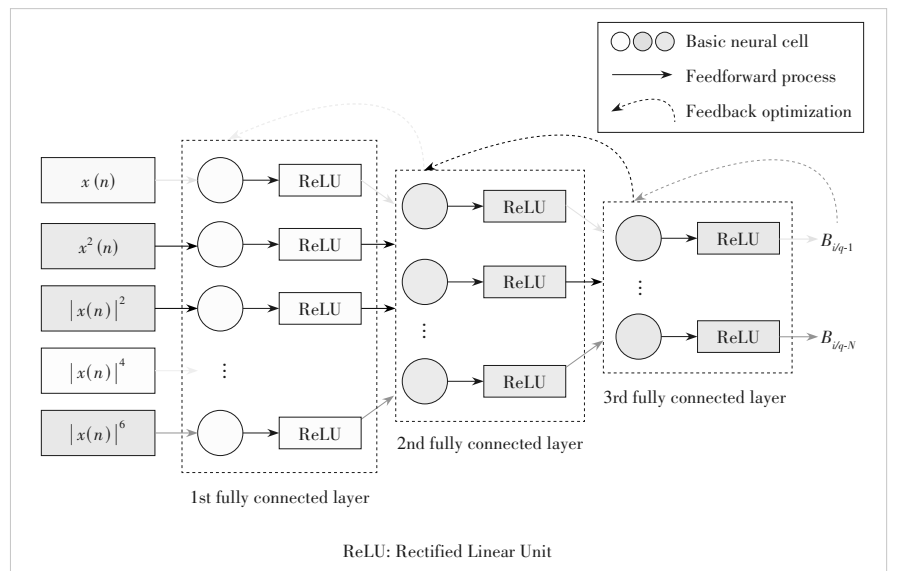


Figure 1. Proposed basis generation function

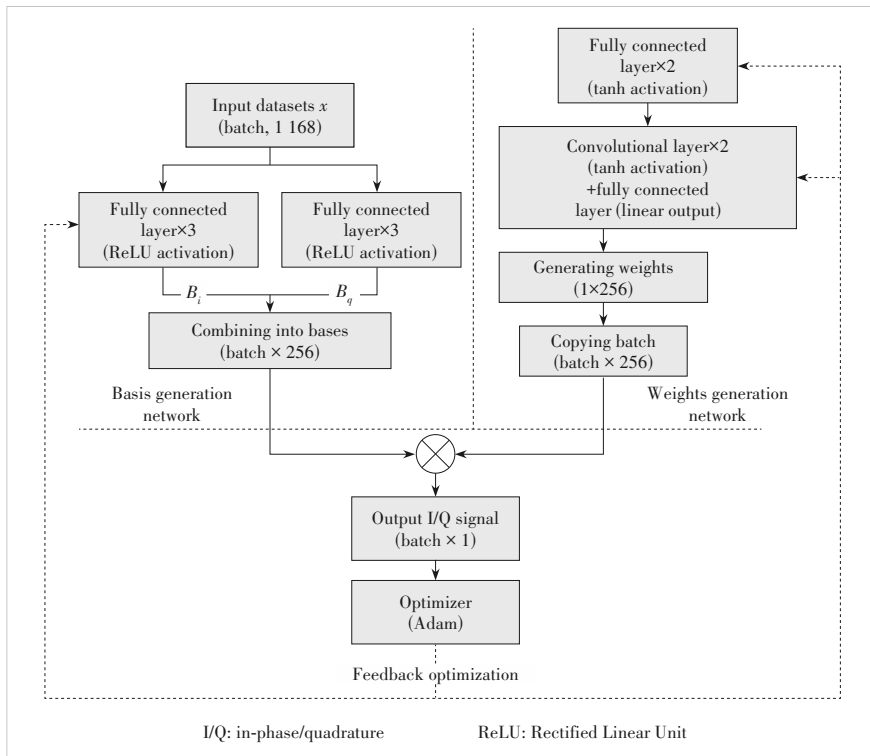


Figure 2. Proposed parallel concurrent neural network for digital predistortion

side of Fig. 2, the weights generation network (WGN) based on a CNN is illustrated. It contains three fully connected layers and two convolutional layers, as shown in Fig. 2. To ensure that the weights generation does not change by the input, the input to the WGN is fixed at a constant value (set to 1 in the following training). The input to the CNN model is derived from the output of the two FC layers positioned above it. Therefore, the input of the WGN has no real physical meaning. WGN aims to help us find a proper weight through a training process as the loss decreases. Similar to the bases, we assume that when the loss converges to a very low level, the weights will be optimized to best describe the nonlinear compression features in the trained data. Subsequently, the outputs of the FNN and CNN are trained through the projection layer at the bottom of Fig. 2 to generate the in-phase/quadrature (I/Q) data. The projection layer is tailored to perform complex multiplication accurately by incorporating appropriate dimensional changes, aligning the I/Q data, and producing the final output. The mean squared error (MSE) of the model's output is computed using the validation set (valset) as the model loss. This calculated model loss is then utilized as feedback to fine-tune the coefficients in all hidden layers of the proposed model. Furthermore, the weights generated by the CNN model are insensitive to the data fed into the model. Therefore, if the model is executed on a hardware platform such as a Field-Programmable Gate Array (FPGA), only the FNN network needs to be deployed. The trained FNN network carries out only linear operations and is easily implemented in engineer-

ing applications.

In essence, the proposed FNN-CNN concurrent model achieves DPD through time-domain fitting. The FNN model generates the basis functions based on the GMP+DDR model via model training, while the CNN model produces the weights using the coefficients from each hidden layer. Both models are jointly optimized to minimize the loss. The final output is obtained by multiplying the basis functions with the weights, after which the valset is used to compute the loss. A model with such a structure, concurrently trained by CNN and FNN, is dubbed a concurrent neural network model.

4 Training Process and Results

4.1 Dataset Use Cases

Two datasets from different RFs are available in the OFDM communication system, with bandwidths of 300 MHz and 400 MHz. Each dataset has 10×16 384 samples. We use eight of ten feedback signals as training sets (trainset), the other two of them as test sets (testset), and the corresponding transmission signal fed into the PA as the validation set.

Fig. 3 illustrates the nonlinear compression of the datasets containing both 300 MHz and 400 MHz bandwidths. The PA significantly compresses the signal amplitude. As shown in Table 1, the compressed signal produces severe out-of-band leakage and nonlinear distortion. Table 1 presents the frequency-domain performance of the datasets, and the ACPR is calculated at integral bandwidths of 100 MHz and 20 MHz. The primary goal of this paper is to minimize the ACPR (with a target of at least -48 dB) of the output generated by the proposed model by employing optimization and training techniques.

The proposed model, as outlined in Section 3, aims to eliminate the out-of-band nonlinear distortion through time-domain fitting of the trainset to the valset. Moreover, we evaluate the model's effectiveness through the ACPR at an integrated bandwidth of 100 MHz.

4.2 Training Results

The datasets shown in Section 4.1 are utilized to train the proposed model with the MSE serving as the loss function. Eight out of ten training sets are selected randomly for the training process, and the remaining two datasets are used as testsets to evaluate the trained model's performance.

Fig. 4 depicts the evolution of the model output's MSE over 20 000 training epochs. The blue curve represents the

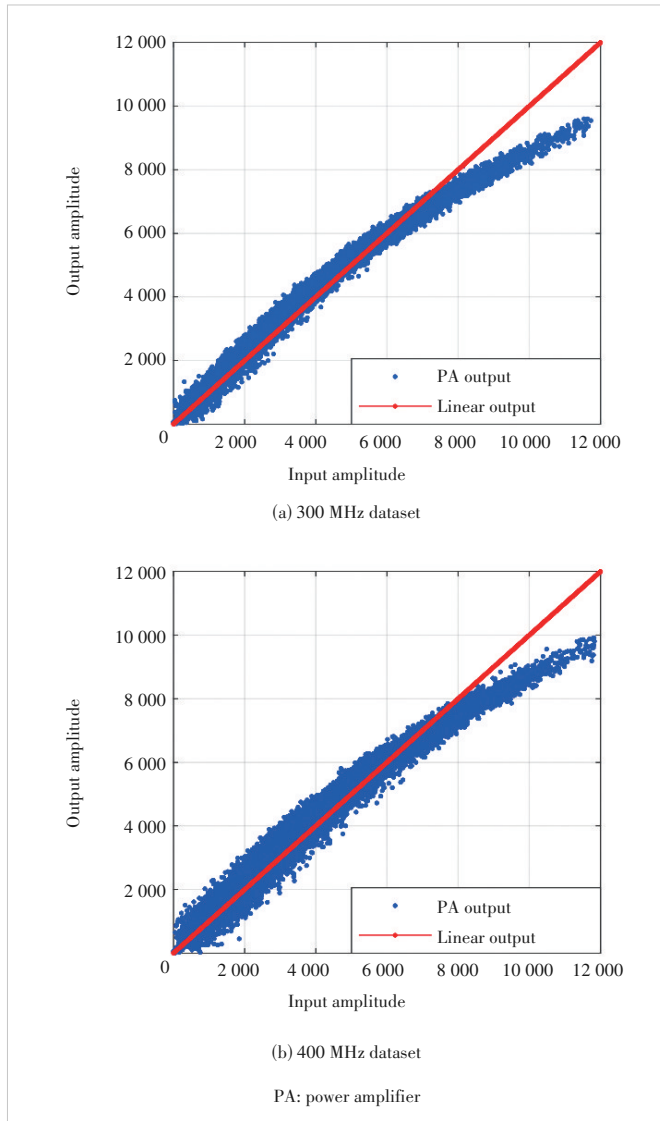


Figure 3. Demonstration of training and validation sets

300 MHz dataset, whose MSE drops sharply within the first 5 000 epochs, reaching a plateau thereafter and converging below 2×10^{-5} around 12 500 epochs. The final loss of the 300 MHz dataset after 20 000 epochs is 1.4×10^{-5} . The orange

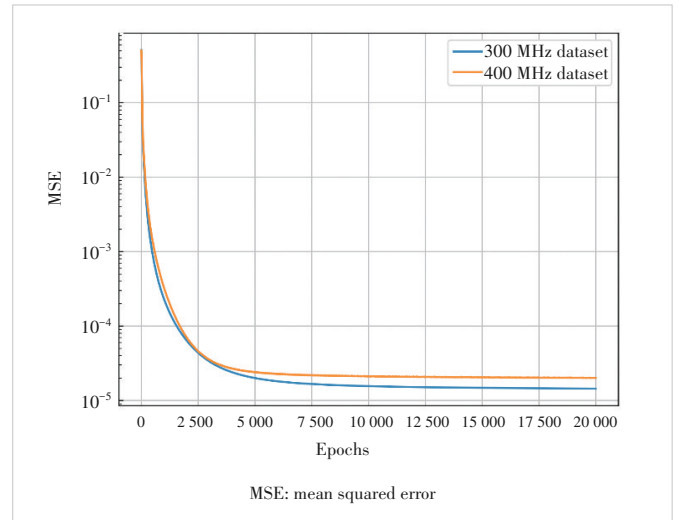


Figure 4. Model training loss

curve, representing the 400 MHz dataset, exhibits a similar downward trajectory, albeit with a poorer result than the blue curve. Its final loss after 20 000 epochs is 1.97×10^{-5} . Notably, the increase in bandwidth from 300 MHz to 400 MHz does not interfere with the convergence speed. The final convergence value is affected not only by model training but also by the differences in the datasets. Fig. 4 provides evidence that the proposed model performs well across various bandwidths, thereby highlighting its generalizability.

Fig. 5 shows the DPD results of the proposed model and a comparison with the existing Generalized Memory Polynomial (GMP) algorithm. Specifically, the out-of-band power of the proposed model (the blue curve in the figure) is approximately 40 dB lower than that of the feedback signal (the red curve) in Figs. 5a and 5b. In addition, the signal portion of the model output closely matches the signal component of the source signal (the cyan curve) for both 300 MHz and 400 MHz cases. This outcome indicates the remarkable ability of the proposed model to suppress out-of-band nonlinear distortion without compromising the fitting of the signal portion. When compared with the existing algorithms, the proposed model has obvious advantages in suppressing out-of-band leakage. There is nearly 20 dB optimization compared to the GMP model (the

Table 1. Frequency domain performance of datasets

Datasets	ACPR of Left Band/dB		ACPR of Right Band/dB		NMSE
	100 MHz Integral Bandwidth	20 MHz Integral Bandwidth	100 MHz Integral Bandwidth	20 MHz Integral Bandwidth	
300 MHz valset	-54.17	-46.18	-58.04	-48.65	
300 MHz datasets	Max=-21.45	Max=-21.00	Max=-21.94	Max=-21.01	Max=-16.40
	Min=-21.49	Min=-21.02	Min=-21.98	Min=-21.01	Min=-16.41
	Average=-21.47	Average=-21.00	Average=-21.96	Average=-21.01	Average=-16.40
400 MHz valset	-54.90	-43.29	-53.61	-46.86	
400 MHz datasets	Max=-21.36	Max=-19.39	Max=-22.61	Max=-20.13	Max=-17.02
	Min=-21.40	Min=-19.48	Min=-22.88	Min=-20.22	Min=-17.05
	Average=-21.38	Average=-19.45	Average=-22.66	Average=-20.19	Average=-17.04

ACPR: adjacent channel power ratio

NMSE: normalized mean square error

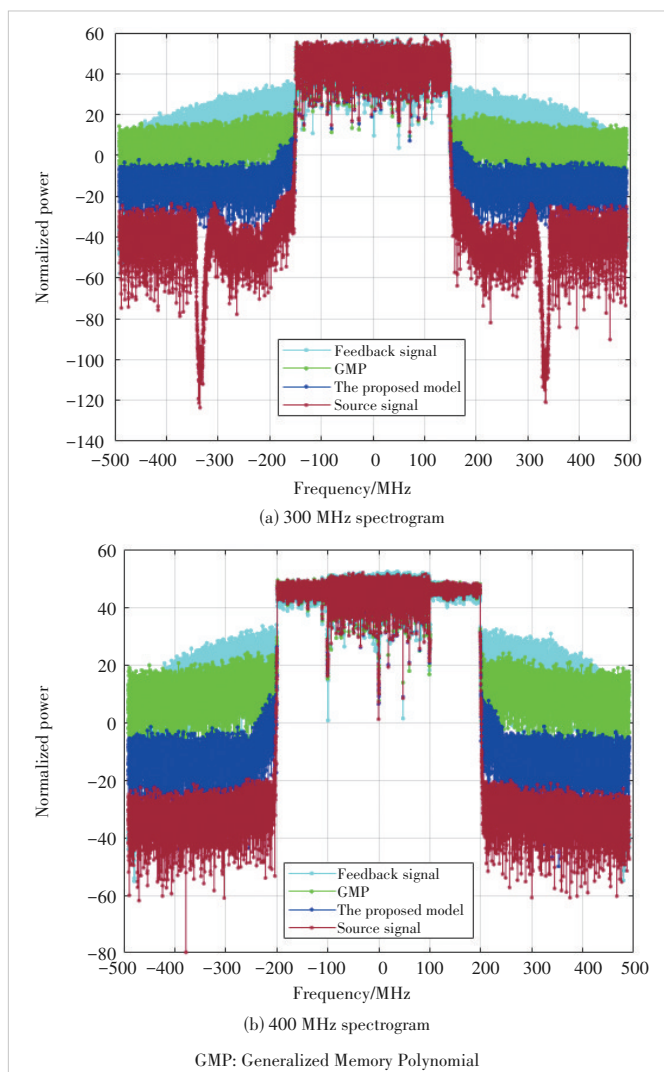


Figure 5. Digital predistortion results of the proposed model

green curve). However, the training results for the 300 MHz and 400 MHz datasets in Fig. 5 still exhibit some out-of-band non-linear distortion, approximately 20 dB higher than the out-of-band part of the validation sets.

Table 2 presents the training and testing results of the proposed model for both the 300 MHz and 400 MHz datasets. The ACPR of both the left and right frequency bands can exceed -48 dB for an integral bandwidth of 100 MHz. Additionally, the model output's NMSE indicates an improvement of nearly -30 dB compared to the initial NMSE of the trainsets, demonstrating that the model output fits the signal portion well. The ACPR of the testsets is also greater than -48 dB; however, it is approximately 1 dB worse than the ACPR of the training sets, and the NMSE training results show similar results.

In addition to evaluating the ACPR of the 100 MHz integral bandwidth, this study also calculates the ACPR of the 20 MHz integral bandwidth to identify why the ACPR produces suboptimal outcomes. By comparing Table 2 with Table 1, it is observable that the performance difference between the model output and the trainsets or testsets remains the same for both 100 MHz and 20 MHz integral bandwidths. Consequently, the expansion of 80 MHz to the periphery does not affect the ACPR results. Instead, the primary factors affecting the ACPR assessment are concentrated within the 20 MHz band boundary.

5 Conclusions

This paper presents a concurrent NN model of RF PA designed to accomplish DPD functions. The proposed model employs the enhanced DDR (GMP+DDR) model as input, which is more suited for modeling the behavior of broadband communication systems. The FNN generates the basis functions, while the CNN generates the weights, with the entire model trained to simultaneously generate their respective optimized values. This study employed eight sets of 300 MHz and 400 MHz data for 20 000 epochs and tested the model with two sets of data. After training and testing, the desired goal of achieving a -48 dB ACPR by 100 MHz integral bandwidth was met for both the trainsets and testsets. The spectrogram shows that the proposed model has a great advantage over the existing algorithms in wider bandwidth scenarios. Moreover, the ACPR was evaluated at 20 MHz integral bandwidth, revealing that the roll-off is the primary limitation of

Table 2. Frequency domain performance of model output

Datasets	ACPR of Left Band/dB		ACPR of Right Band/dB		NMSE
	100 MHz Integral Bandwidth	20 MHz Integral Bandwidth	100 MHz Integral Bandwidth	20 MHz Integral Bandwidth	
300 MHz trainsets	Max= -50.27	Max= -42.53	Max= -51.29	Max= -44.25	Max= -43.62
	Min= -50.77	Min= -43.15	Min= -51.99	Min= -45.14	Min= -46.12
	Average= -50.48	Average= -42.94	Average= -51.70	Average= -44.73	Average= -46.69
300 MHz testsets	$-49.65, -49.48$	$-42.77, -42.57$	$-50.71, -50.69$	$-44.35, -44.25$	$-44.77, -44.79$
400 MHz trainsets	Max= -48.93	Max= -40.99	Max= -48.66	Max= -42.45	Max= -43.86
	Min= -49.36	Min= -41.49	Min= -49.10	Min= -43.40	Min= -44.31
	Average= -49.17	Average= -41.24	Average= -48.82	Average= -43.10	Average= -44.15
400 MHz testsets	$-48.02, -48.17$	$-40.50, -40.64$	$-48.13, -48.12$	$-42.44, -42.54$	$-42.92, -42.85$

ACPR: adjacent channel power ratio

NMSE: normalized mean square error

ACPR. This finding can guide future efforts to optimize the proposed model.

References

- [1] LIU Z J, HU X, WANG W D, et al. A joint PAPR reduction and digital predistortion based on real-valued neural networks for OFDM systems [J]. *IEEE transactions on broadcasting*, 2022, 68(1): 223 – 231. DOI: 10.1109/TBC.2021.3132158
- [2] LIU Z J, HU X, XU L X, et al. Low computational complexity digital predistortion based on convolutional neural network for wideband power amplifiers [J]. *IEEE transactions on circuits and systems II: express briefs*, 2022, 69(3): 1702 – 1706. DOI: 10.1109/TCSII.2021.3109973
- [3] LIU X, CHEN W H, WANG D H, et al. Robust digital predistortion for LTE/5G power amplifiers utilizing negative feedback iteration [J]. *ZTE communications*, 2020, 18(3): 49 – 56. DOI: 10.12142/ZTECOM.202003008
- [4] LIU Y J, ZHOU J, CHEN W H, et al. A robust augmented complexity-reduced generalized memory polynomial for wideband RF power amplifiers [J]. *IEEE transactions on industrial electronics*, 2014, 61(5): 2389 – 2401. DOI: 10.1109/TIE.2013.2270217
- [5] ZHU A D, DRAXLER P J, YAN J J, et al. Open-loop digital predistorter for RF power amplifiers using dynamic deviation reduction-based Volterra series [J]. *IEEE transactions on microwave theory and techniques*, 2008, 56(7): 1524 – 1534. DOI: 10.1109/TMTT.2008.925211
- [6] BRAITHWAITE R N. Adaptation of a digitally predistorted RF amplifier using selective sampling [J]. *ZTE communications*, 2011, 9(3): 3 – 12
- [7] HU X, LIU Z J, WANG W D, et al. Low-feedback sampling rate digital predistortion using deep neural network for wideband wireless transmitters [J]. *IEEE transactions on communications*, 2020, 68(4): 2621 – 2633. DOI: 10.1109/TCOMM.2020.2966718
- [8] ROSOŁOWSKI D W, JĘDRZEJEWSKI K. Experimental evaluation of PA digital predistortion based on simple feedforward neural network [C]//Proc. 23rd International Microwave and Radar Conference (MIKON). IEEE, 2020: 293 – 296. DOI: 10.23919/MIKON48703.2020.9253814
- [9] FENG X, FEUVRIE B, DESCAMPS A S, et al. Digital predistortion method combining memory polynomial and feed-forward neural network [J]. *Electronics letters*, 2015, 51(12): 943 – 945. DOI: 10.1049/el.2015.0276
- [10] WU H B, CHEN W H, LIU X, et al. A uniform neural network digital predistortion model of RF power amplifiers for scalable applications [J]. *IEEE transactions on microwave theory and techniques*, 2022, 70(11): 4885 – 4899. DOI: 10.1109/TMTT.2022.3205930
- [11] JUNG S, KIM Y, WOO Y, et al. A two-step approach for DLA-based digital predistortion using an integrated neural network [J]. *Signal processing*, 2020, 177: 107736. DOI: 10.1016/j.sigpro.2020.107736
- [12] WU Y B, GUSTAVSSON U, AMAT A G I, et al. Residual neural networks for digital predistortion [C]//Proc. 2020 IEEE Global Communications Conference. IEEE, 2020: 1-6. DOI: 10.1109/globe-com42002.2020.9322327
- [13] LÓPEZ-BUENO D, MONTORO G, GILBERT P L. Training data selection and dimensionality reduction for polynomial and artificial neural network MIMO adaptive digital predistortion [J]. *IEEE transactions on microwave theory and techniques*, 2022, 70(11): 4940 – 4954. DOI: 10.1109/TMTT.2022.3209214
- [14] JIANG C Y, YANG G C, HAN R L, et al. Gated dynamic neural network model for digital predistortion of RF power amplifiers with varying transmission configurations [J]. *IEEE transactions on microwave theory and techniques*, 2023, 71(8): 3605 – 3616. DOI: 10.1109/TMTT.2023.3241612
- [15] JARAUT P, ABDELHAFIZ A, CHENINI H, et al. Augmented convolutional neural network for behavioral modeling and digital predistortion of concurrent multiband power amplifiers [J]. *IEEE transactions on microwave theory and techniques*, 2021, 69(9): 4142 – 4156. DOI: 10.1109/TMTT.2021.3075689
- [16] JIANG C Y, LI H M, QIAO W, et al. Block-oriented time-delay neural network behavioral model for digital predistortion of RF power amplifiers [J]. *IEEE transactions on microwave theory and techniques*, 2022, 70(3): 1461 – 1473. DOI: 10.1109/TMTT.2021.3124211
- [17] KOBAL T, LI Y, WANG X Y, et al. Digital predistortion of RF power amplifiers with phase-gated recurrent neural networks [J]. *IEEE transactions on microwave theory and techniques*, 2022, 70(6): 3291 – 3299. DOI: 10.1109/TMTT.2022.3161024
- [18] MONDAL R, RISTANIEMI T, DOULA M. Genetic algorithm optimized memory polynomial digital pre-distorter for RF power amplifiers[C]//Proc. International Conference on Wireless Communications and Signal Processing. IEEE, 2013: 1 – 5. DOI: 10.1109/WCSP.2013.6677117

Biographies

SHAO Jianfeng (sjf1996717@stu.xjtu.edu.cn) received his BS degree in electronic information engineering from the School of Information Engineering, North China University of Water Resources and Electric Power in 2018, MS degree in circuits and systems from the School of Physics and Electronic Engineering, Ningxia University, China in 2021. He is pursuing a PhD degree in information and communication engineering at Xi'an Jiaotong University, China. His research interests include array signal processing and digital pre-distortion.

HONG Xi received his BS degree in information engineering from the School of Electronics and Information Engineering, Xi'an Jiaotong University, China in 2012, and PhD degree in information and communication engineering from the School of Information and Communication Engineering, Xi'an Jiaotong University in 2022. He is currently an engineer at the School of Information and Communication Engineering, Xi'an Jiaotong University. His research interests include array signal processing, signal processing in communication systems, multipath mitigation in navigation, and GNSS physical layer interference detection.

WANG Wenjie received his BS, MS, and PhD degrees in information and communication engineering from Xi'an Jiaotong University, China in 1993, 1998, and 2001, respectively, where he is currently a professor. His main research interests include information theory, broadband wireless communications, signal processing in communication systems, and array signal processing.

LIN Zeyu received his BS and MS degrees at the School of Information and Communication Engineering, Xi'an Jiaotong University, China in 2018 and 2021, respectively. He is currently a senior RF algorithm architect in the RHP department of ZTE Corporation. His research interests include PA nonlinear system behavioral modeling and RF systems linearization.

LI Yunhua received his BE degree in communications engineering from Shandong University (Weihai), China in 2012 and PhD degree in military communications from the School of Telecommunications Engineering, Xidian University, China in 2017. He is currently a senior engineer of RF algorithms in ZTE Corporation. His research interests include communication signal processing, wireless communications, digital predistortion modeling of nonlinear systems, and more.