electronics

Communication

Efficient Neural Network DPD Architecture for Hybrid Beamforming mMIMO

Tamara Muškatirović-Zekić [1,2,*], Nataša Nešković [1] and Djuradj Budimir [1,3]

[1] School of Electrical Engineering, University of Belgrade, 11120 Belgrade, Serbia
[2] Regulatory Agency for Electronic Communications and Postal Services, 11103 Belgrade, Serbia
[3] Wireless Communications Research Group, University of Westminster, London W1B 2HW, UK
* Correspondence: muskatirovic@gmail.com

Abstract: This paper presents several different Neural Network based DPD architectures for hybrid beamforming (HBF) mMIMO applications. They are formulated, tested and compared based on their ability to compensate nonlinear distortion of power amplifiers in a single user (SU) and multiuser (MU) Fully-Connected (FC) HBF mMIMO transmitters. The proof-of-concept is provided with a 64×64 FC HBF mMIMO system, with 2 RF chains. The complexity of DPD solution is reduced by using a single Real-Valued Time-Delay Neural Network with two hidden layers (RVTDNN2L) instead of using as many different DPD blocks as there are RF chains in the HBF mMIMO transmitter and it is shown that the proposed architecture better compensates nonlinear distortion compared to the traditional memory polynomial DPD. Two RVTDNN2L DPD architectures are developed and tested for linearization of MU FC HBF mMIMO systems, and it is also shown that the proposed RVTDNN2L DPD architecture efficiently linearizes MU FC HBF mMIMO transmitters in terms of Normalized Mean-Squared Error (NMSE) and Error Vector Magnitude (EVM).

Keywords: neural networks; digital predistortion (DPD); massive MIMO; hybrid beamforming (HBF); power amplifiers (PAs)

Citation: Muškatirović-Zekić, T.; Nešković, N.; Budimir, D. Efficient Neural Network DPD Architecture for Hybrid Beamforming mMIMO. *Electronics* **2023**, *12*, 597. https://doi.org/10.3390/electronics12030597

Academic Editor: Adão Silva

Received: 12 December 2022
Revised: 18 January 2023
Accepted: 19 January 2023
Published: 25 January 2023

1. Introduction

Massive multiple-input multiple-output (mMIMO) combined with hybrid beamforming (HBF) has been widely accepted as one of the key technologies in fifth generation (5G) and beyond networks to improve the throughput and increase spectral efficiency [1–7]. Hybrid beamforming is a combination of analog beamforming in RF domain and digital beamforming in baseband. There are two types of HBF MIMO transmitter architectures: (a) fully connected (FC) architecture, when each antenna is connected to each of the P RF chains, and (b) partially connected or subarray architecture, when a set of L antennas is connected to one RF chain. On the other hand, in order to reduce power consumption, the power amplifiers (PAs), that are the main power consumers in the transmitters, must operate in a region close to saturation, causing increased nonlinear distortion and reduced spectral efficiency. Digital predistortion (DPD), as one of the most promising solutions and the most effective PA linearization technique, has been extensively studied in literature [8]. Conventional DPD requires dedicated predistorter for observing and linearizing the nonlinear distortion of each PA, which, in the case of HBF mMIMO systems, is completely unsuitable. Therefore, a lot of research work has been done to resolve DPD implementation problems in HBF mMIMO systems [9–15].

Most of these DPD techniques are proposed for subarray or reduced complexity HBF architectures, meaning that for the whole HBF mMIMO systems, as many DPD blocks as there are RF chains will be needed. In [9] DPD learning is based on the feedback from a single PA in the array. The decorrelation DPD model introduced in [10] uses a combined output signal of individual PAs in conjunction with a decorrelation-based learning rule.

The same group of authors proposed a closed-loop (CL) DPD technique in [11] as a solution for multi-user hybrid MIMO that uses a single DPD block per RF transmission subarray. The Beam-Oriented DPD (BO-DPD) technique presented in [12] linearizes the signal of the main beam, not each PA individually, using the outputs of all PAs to obtain the feedback signal. Authors in [13] proposed a Power Scalable Beam-Oriented DPD (PSBO-DPD) that linearizes the signal of the main beam of a given subarray using feedback from only one PA. As a fully-connected (FC) HBF architecture allows higher spectral efficiency compared to a subarray HBF architecture, the authors in [15] proposed an approach for linearization of FC HBF mMIMO transmitters. However, the approach still uses as many DPD blocks as there are RF chains. Based on the type of feedback signal, over-the-air (OTA) DPD and direct transmit-end DPD are distinguished. Authors in [16] proposed a best linear approximation (BLA) based equalization strategy for OTA DPD in 5G Beamformer Array at 28 GHz and showed that using this approach enhances the linearization performance when a single and fixed set of DPD coefficients is used across different beam angles and receiver positions. Nevertheless, this paper considers the linearization of the main beam using direct transmit-end DPD and does not consider such equalization.

Neural networks (NNs) have high modelling accuracy as well as good distortion compensation performance in PA linearization [17–24]. Current investigations are based on the applications of different NN in DPD of a single PA. Authors in [20] proposed real-valued time-delay feedforward backpropagation-based NN for DPD in MIMO transmitters. However, as to the best of the authors knowledge, there have been no studies on the implementation of NNs DPD for HBF mMIMO yet.

In this paper, we propose NN based DPD for FC HBF mMIMO. Firstly, we present and analyze different NN configurations, and compare their predistortion capabilities in a single user (SU) FC HBF mMIMO transmitter, in terms of Error Vector Magnitude (EVM) and Normalized Mean-Squared Error (NMSE). It is shown that the proposed NN based DPD solution efficiently linearizes FC HBF mMIMO (with P RF chains) using a single NN instead of P different DPD blocks, thus reducing the complexity compared to the state-of-the-art DPD solutions for HBF mMIMO, which uses P different DPD blocks. Therefore, we proposed an advantageous Real-Valued Time-Delay Neural Network with two hidden layers (RVTDNN2L) and showed that this NN DPD solution has smaller NMSE and EVM compared to the state-of-the-art DPD solutions for SU HBF mMIMO. Secondly, for the proposed and selected RVTDNN2L DPD approach we investigated the influence of the number of RF branches on the linearization of the FC HBF mMIMO. We increased the number of RF branches to 4 and 8 and showed that, with the implementation of this approach, NMSE decreases by increasing the number of RF branches. At the end, we considered and examined a multiuser (MU) FC HBF mMIMO using two different DPD architectures with the RVTDNN2L approach. It is shown that the proposed approach better compensates for the nonlinear distortion of PAs in MU FC HBF mMIMO. One of the advantages of the RVTDNN2L DPD approach is also its adaptability, which means that with the change of environmental conditions, it can be relatively easily retrained and adapted to new conditions.

The rest of this paper is structured as follows. In Section 2, different NN based DPD models for PA linearization are presented and their performance and complexity are analyzed. Simulation results and comparisons between different NN DPD models for SU FC HBF mMIMO are given in Section 3 and RVTDNN2L DPD is proposed as the most efficient one. In Section 4, simulation results for the proposed RVTDNN2L DPD approach for MU FC HBF mMIMO are shown. Finally, Section 5 gives the conclusions.

2. NN DPD Models

A commonly used NN model has an input layer, one or more hidden layers and an output layer. Based on the number of hidden layers, NNs can be divided into two main categories: shallow and deep neural networks. While shallow NNs have one or two hidden layers, deep neural networks (DNNs) have at least three hidden layers with a high

number of neurons in each layer. With the increase in the number of hidden layers, the network behavior is better modeled, while on the other hand, the requirements for signal processing increase. So, in this paper we analyze several different NN models such as, two shallow networks: real-valued time-delay neural network (RVTDNN) and augmented real-valued time-delay neural network (ARVTDNN), and one deep neural network (DNN), and propose the use of the RVTDNN network with two hidden layers (RVTDNN2L).

Input signals for the DNN, as well as for the RVTDNN and RVTDNN2L, are I and Q components of the signal. Beside I and Q components of the signal, ARVTDNN considers the envelope-dependent terms of the input signal, i.e., $|x(n)|$, $|x(n)|^2$ and $|x(n)|^3$, where $|x(n)|$ is amplitude of the input signal. The activation function used in these models is hyperbolic tangent (*tanh*), while for DNN model, sigmoid (*logsig*) activation function is used.

Complexity is one of the more important parameters that needs to be considered when choosing an appropriate NN DPD model. It can be expressed by the number of coefficients, as well as the number of floating-point operations (FLOPs). The total number of coefficients is equal to the sum of the used coefficients in every layer, and can be computed as follows:

$$N_{coef} = (N_i + 1)N_1 + \sum_{f=2}^{F}\left(N_{f-1} + 1\right)N_f + (N_F + 1)N_O \quad (1)$$

where Ni is the number of neurons in the input layer, F is the number of hidden layers, N_f is the number of neurons in f-th hidden layer and No is the number of neurons in the output layer.

The number of FLOPs is used for estimation of the computational complexity of the model and is different for every particular operation performed [24,25]. For the NNs with hyperbolic tangent activation function, number of FLOPs can be computed as follows:

$$\text{FLOPs} = 2N_iN_1 + 2N_FN_O + 2\sum_{f=2}^{F} N_{f-1}N_f + 15\sum_{f=1}^{F} N_f \quad (2)$$

For the NNs with sigmoid activation function, number of FLOPs can be computed as follows:

$$\text{FLOPs} = 2N_iN_1 + 2N_FN_O + 2\sum_{f=2}^{F} N_{f-1}N_f + 13\sum_{f=1}^{F} N_f \quad (3)$$

Due to its simple implementation, the indirect learning architecture (ILA) was used to form the DPD architecture based on single NN for the whole HBF mMIMO system, as shown in Figure 1. First, the appropriate NN model for DPD was formed and was trained using combined output data signals from all antennas and input data to the HBF MIMO transmitter. Then the trained NN model was used to update the input signal to HBF MIMO transmitter.

The algorithm used for training NN models is Levenberg-Marquardt optimization algorithm, a well-known and widely used type of back-propagation algorithm [18,21,22,24]. The Levenberg–Marquardt algorithm has been shown to perform exceptionally well because of its fastest convergence with lower mean square errors for training feedforward neural networks [26]. It also has an efficient implementation in Matlab ANN Toolbox. To update the values of the synaptic weights and biases in neural network, the Levenberg-Marquardt algorithm uses the following approximation:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left[\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}\right]^{-1}\mathbf{J}^T\mathbf{e} \quad (4)$$

where $\mathbf{J}$ is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and $\mathbf{e}$ is a vector of network errors. Initial μ is set equal to 0.001 and increase factor for μ is 10.

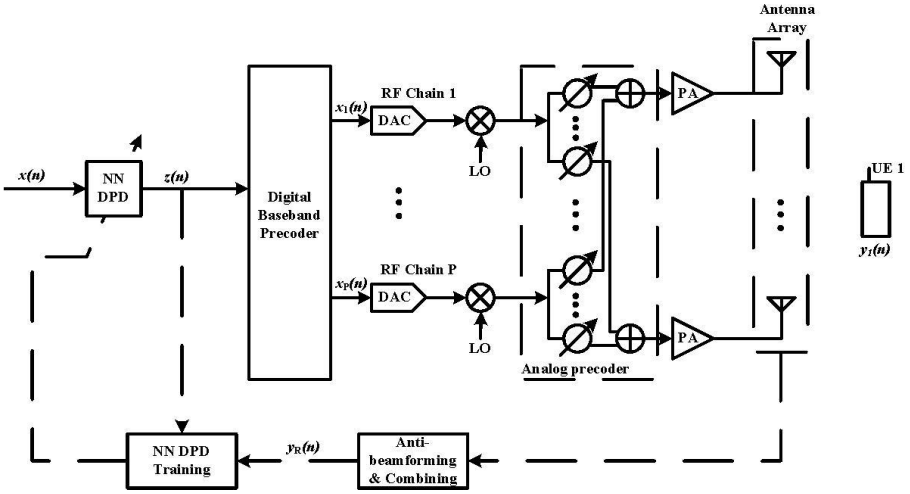

Figure 1. The proposed NN DPD architecture for SU FC HBF MIMO transmitter.

The optimal number of neurons in hidden layers are determined during training of the network and are set to give the best performance. As a result of extensive simulations, we determine following characteristics of NN models: DNN with 3 hidden layers, *logsig* activation function and 12 neurons in each hidden layer; RVTDNN with one hidden layer, *tansig* activation function and 35 neurons in hidden layer; ARVTDNN with one hidden layer, *tansig* activation function and 28 neurons in hidden layer; and RVTDNN2L with two hidden layers, *tansig* activation function, 18 neurons in the first hidden layer and 9 neurons in the second hidden layer. Block diagrams of these NN models are shown in Figure 2. As well, we propose the use of RVTDNN2L network for digital predistortion of FC HBF mMIMO transmitter. The performance and complexity of different analyzed NN models are listed in Table 1. A large number of simulations were performed, varying the number of neurons per layer from 1 to 30 for the first hidden layer and from 1 to 15 for the second hidden layer, for RVTDNN2L network. For different combinations of neurons per layer, the number of coefficients according to formula (1) was calculated, as well as the number of FLOPs according to formula (2) and values of NMSE and EVM. Observing the obtained results, the optimal numbers of neurons were chosen so that NMSE and EVM values are comparable to the values existing in literature in the literature for HBF mMIMO transmitters, but also that the number of coefficients and FLOPs are as small as possible, in order to reduce the complexity of the system. The numbers of neurons per layer for other types of networks were chosen in a similar way, ensuring that their complexity is similar to that of the RVTDNN2L network.

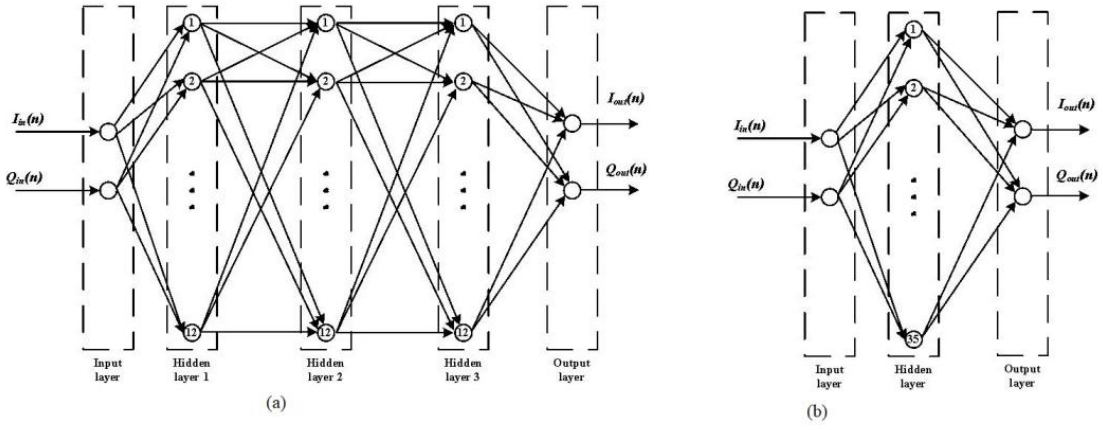

Figure 2. *Cont.*

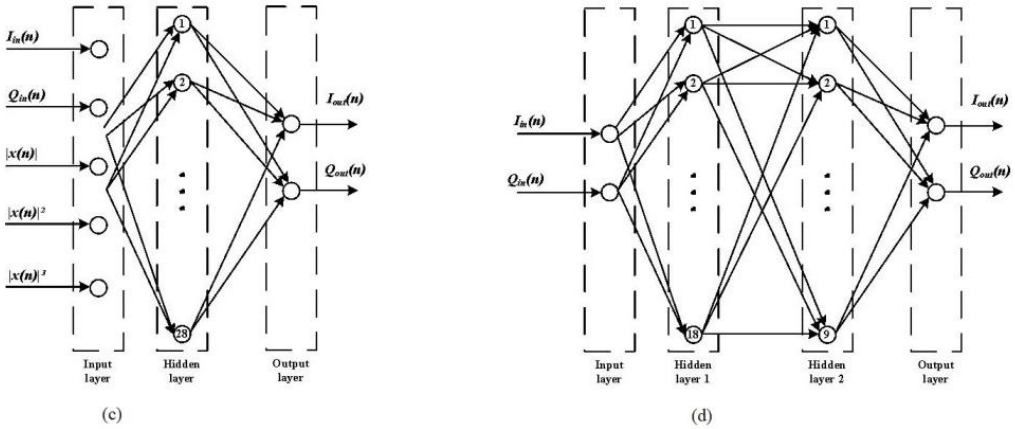

Figure 2. Block diagrams of the proposed NN models: (**a**) DNN, (**b**) RVTDNN, (**c**) ARVTDNN, (**d**) RVTDNN2L.

Table 1. Performance and complexity of different NNs.

Type of NN DPD	Input Data	Number of Neurons in Input Layer	Number of Neurons in Hidden Layers	Activation Function	Number of Neurons in Output Layer	N_{coef}	FLOPs
DNN	*I*/*Q*	2	12 12 12	*logsig*	2	374	852
RVTDNN	*I*/*Q*	2	35	*tansig*	2	177	805
ARVTDNN	I/Q, $\|x(n)\|$, $\|x(n)\|^2$, $\|x(n)\|^3$	5	28	*tansig*	2	226	812
RVTDNN2L	*I*/*Q*	2	18 9	*tansig*	2	245	837

3. Results

In this section, some DPD simulation results of the SU FC HBF mMIMO transmitter are presented. Individual PA units in transmitter are modeled using NN, based on measurement data from an actual PA. A two-stage PA (CFH 2162-P3 PA) with a 14 dB gain, and P1dB of 37 dBm driven in a hard compression region was used. The modulated waveform with channel bandwidth of 20 MHz, was fed into this power amplifier model at 2140 MHz. The process of PA modeling is shown in Algorithm 1 in Appendix A. Single PA has been modeled using one RVTDNN2L with two hidden layers, *tansig* activation function, 32 neurons in the first hidden layer and 15 neurons in the second hidden layer. To consider nonlinear crosstalk that occurs before PA, the output of each PA is modeled as:

$$y_i = f_i\left(x_i + \sum_{\substack{n=1 \\ n \neq i}}^{Nt} \alpha_{i,n} \cdot x_n\right) \tag{5}$$

where x_i is the input to i-th PA, f_i is the PA response function, $\alpha_{i,n}$ is the crosstalk factor reflecting the impact of n-th PA to i-th PA [27]. Then NN based DPD model was created, using different NN architectures, as an integral solution to jointly compensate all PAs nonlinearity, crosstalk, I/Q imbalance and dc offset imperfections, similarly to the authors work presented in [22]. Considered DPD architecture is shown in Figure 1. To obtain the feedback signal, used for the DPD learning, the outputs of the individual PAs are firstly extracted, then co-phased and combined in the analog domain before being applied, meaning that direct transmit-end feedback is considered.

The basic concept of hybrid beamforming is to split both the precoding and combining weights to baseband digital weights and RF band analog weights. The problem of determining these weights is quite a large topic and is not considered here. In this paper, the orthogonal matching pursuit algorithm which is implemented in Matlab and

proposed in [28] is used to obtain hybrid precoding and combining weights. Next, a clustered Saleh-Valenzuela channel model is adopted, with impulse response $h(t)$ defined as follows [29]:

$$h(t) = \sum_{l=0}^{N_{cl}-1} \sum_{k=0}^{N_{ray}-1} \beta_{kl} \cdot e^{j\theta_{kl}} \cdot \delta(t - T_l - \tau_{kl}) \quad (6)$$

where: N_{cl} and N_{ray} are maximum number of clusters and multipath components inside clusters, respectively, β_{kl}—the amplitude of k-th ray in l-th cluster, θ_{kl}—the phase of the kl-th ray, T_l—time arrival of the l-th cluster, τ_{kl}—arrival delay of k-th ray with respect to first ray of the l-th cluster. The model contains N_{cl} = 6 clusters, with N_{ray} = 5 rays in each cluster. It should be noted that a similar channel modeling approach is utilized in 3GPP standardization [30].

A 64-QAM modulated OFDM signal with a bandwidth of 20 MHz, generated in Matlab is used in DPD simulations. Subcarrier spacing is 15 kHz, and the number of used subcarriers is 1200. Matlab code for generating OFDM signal is given in Appendix B. Simulations of a 64 × 64 fully-connected hybrid beamforming MIMO system, with 2 RF chains were performed. Four different NN models for DPD are considered: DNN, RVTDNN, ARVTDNN and RVTDNN2L, and compared with each other and with the memory polynomial (MP) DPD approach based on following formula [31]:

$$x_{out}(n) = \sum_{m=0}^{M} \sum_{k=1}^{K} a_{mk} \cdot x_{in}(n-m) \cdot |x_{in}(n-m)|^{k-1} \quad (7)$$

where: a_{mk}—model coefficients, M—memory depth, K—nonlinearity order, x_{in} and x_{out}—input and output signal, respectively. Simulations were performed for MP model used in [32] with memory depth M = 3 and order of nonlinearity K = 11. Block diagram of used MP model is shown in Figure 3.

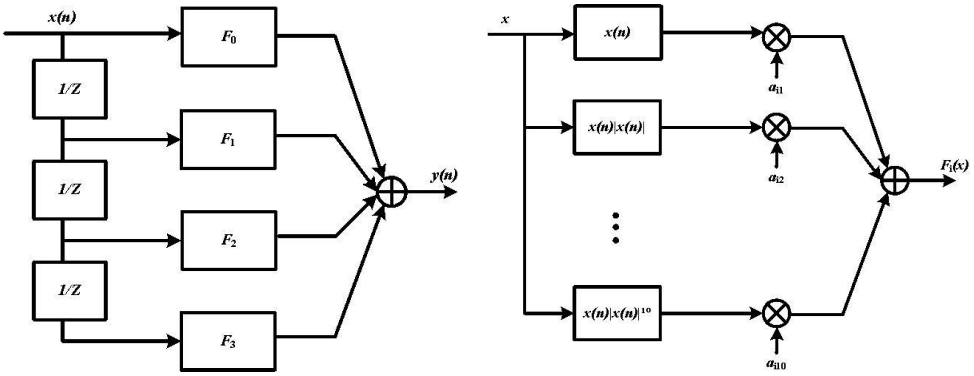

Figure 3. Block diagram of MP model.

Figure 4 shows obtained power spectrum densities (PSDs) without DPD, with MP DPD model and with analyzed NN DPD models. Table 2 summarizes the performance of MP DPD model, different NN DPD models and the proposed RVTDNN2L DPD model, in terms of EVMs and NMSEs. As it can been seen, PA linearization is significantly improved with NN DPD models compared to MP DPD model, especially with the proposed RVTDNN2L DPD model. It should be noted that observed NN DPD models are similar in complexity, i.e., in number of FLOPs. As it can be seen from the obtained results, the proposed RVTDNN2L DPD outperforms other analyzed DPD models.

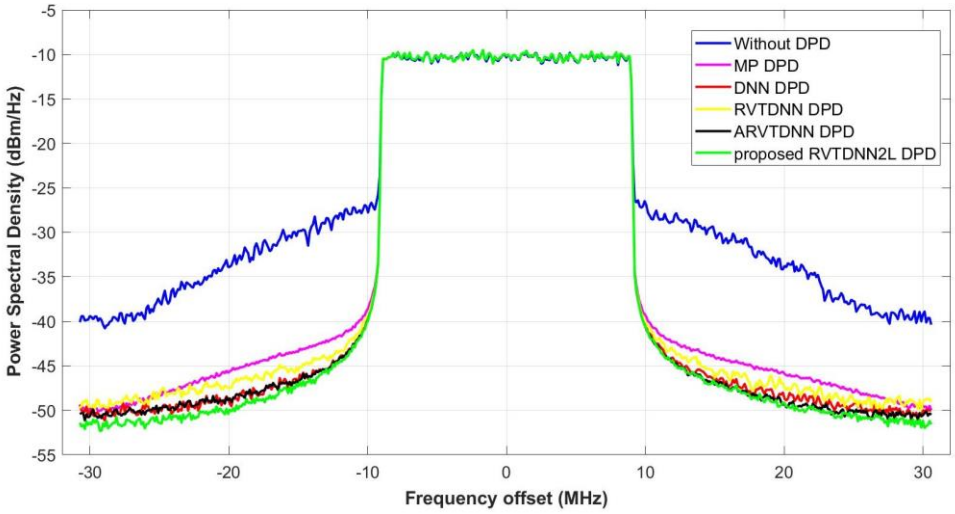

Figure 4. Power Spectral Density of the FC HBF mMIMO using different DPDs.

Table 2. Performance of different NN DPD architectures for 64 × 64 FC HBF MIMO system.

Type of DPD	EVM [%]	NMSE [dB]
Without DPD	22.09	-
MP	2.04	−33.79
DNN	1.37	−37.25
RVTDNN	1.99	−33.98
ARVTDNN	1.70	−35.39
RVTDNN2L	1.23	−38.19

Further, the proposed RVTDNN2L DPD technique has been applied on FC HBF mMIMO system with 4 and 8 RF branches. Simulations were performed in Matlab, with the same input signal and same parameters of RVTDNN2L DPD as in the previous analysis. The same orthogonal matching pursuit algorithm and Saleh-Valenzuela channel model were used to generate the hybrid beamforming coefficients for system with 4 and 8 RF branches. Figure 5 shows the power spectrum densities (PSDs) for FC HBF mMIMO system with 2, 4 and 8 RF branches with the proposed RVTDNN2L DPD model. Table 3 summarizes the performance of the proposed RVTDNN2L DPD model for system with 2, 4 and 8 RF branches, in terms of EVMs and NMSEs. It can be seen that EVM without DPD increases with increasing number of RF branches, meaning that nonlinear distortion increases with increasing number of RF branches, which was expected. The obtained simulation results showed that EVM with the proposed RVTDNN2L DPD model decreases with increasing number of RF branches, which means that the proposed model better compensates nonlinear distortion if the number of RF branches increases, since the RF signals are more correlated with each other.

Although the use of more RF branches achieves better results in EVM and NMSE with the application of digital predistortion, due to the cost of implementing a physical system with more RF branches, in practice HBF MIMO transmitters with 2 RF branches are used. Consequently, FC HBF mMIMO with 2 RF branches will be analyzed in further research of RVTDNN2L DPD for MU case.

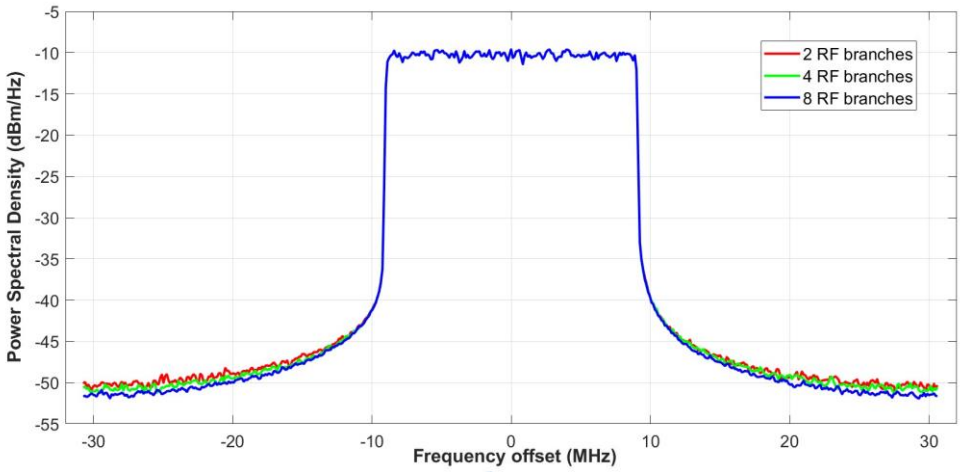

Figure 5. Power Spectral Density of the FC HBF mMIMO using 2,4 and 8 RF branches.

Table 3. Performance of RVTDNN2L DPD architecture for 64 × 64 FC HBF MIMO system with 2, 4 and 8 RF branches.

	2 RF Branches	4 RF Branches	8 RF Branches
EVM without DPD [%]	22.09	22.39	23.39
EVM with RVTDNN2L DPD [%]	1.23	1.04	0.74
NMSE [dB]	−38.19	−39.63	−42.60

4. RVTDNN2L DPD Approach for MU Case

In MU FC HBF MIMO transmitter it is necessary to take into account the signals from all users during the construction of the DPD module, due to the fact that during HBF the data to be sent are formed on the basis of data from all users. In this paper we considered a FC HBF MIMO system with only two users, but with certain minor changes it can be adapted to the case of more than two users.

We proposed and analyzed two architectures for a multi-user system:

- DPD system with one RVTDNNL network with inputs from both users and outputs for both users, shown in Figure 6 (arch1),
- DPD system with two RVTDNNL networks with inputs from both users and output for one user, shown in Figure 7 (arch2).

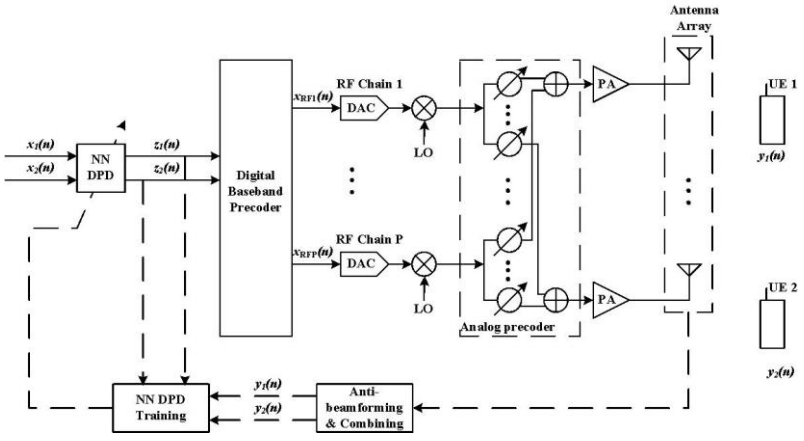

Figure 6. Architecture 1 of RVTDNN2L DPD for MU FC HBF MIMO transmitter with one NN.

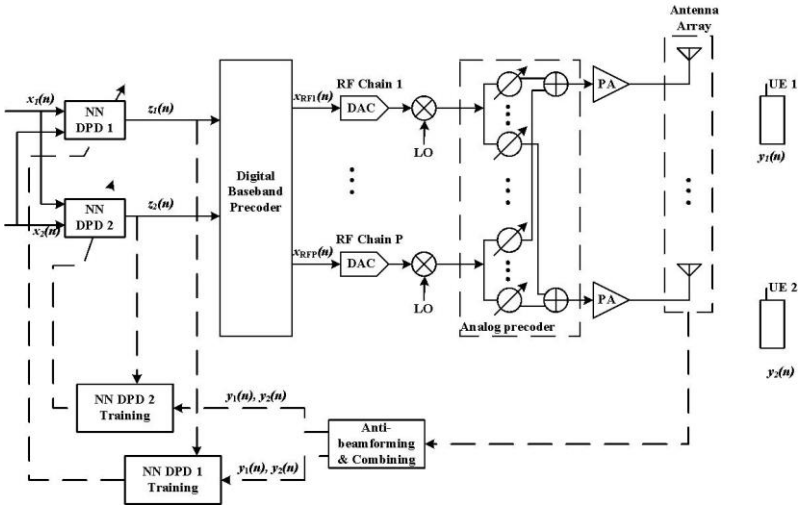

Figure 7. Architecture 2 of RVTDNN2L DPD for MU FC HBF MIMO transmitter with two NNs.

Levenberg-Marquardt optimization algorithm was applied for training RVTDNN2L. *Tansig* activation function was used in both hidden layers. We tested model with different number of neurons in hidden layers of network and proposed the optimal number to obtain the best performance. Graphics and tables are given for three combinations of numbers of neurons in hidden layers, i.e., for I = 18 neurons in the first hidden layer and II = 9 neurons in the second hidden layer; I = 32 neurons in the first hidden layer and II = 15 neurons in the second hidden layer; and I = 48 neurons in the first hidden layer and II = 21 neurons in the second hidden layer. As for SU FC HBF MIMO, a clustered Saleh-Valenzuela channel model is adopted. For MU FC HBF MIMO, the model contains N_{cl} = 6 clusters, with N_{ray} = 10 rays in each cluster. For the scheme in Figure 6, the proposed RVTDNN2L DPD model with 4 inputs and 4 outputs was used, i.e., we considered I and Q components of signals from both users, both inputs and outputs. While for the scheme in Figure 7, the proposed RVTDNN2L DPD model has two NNs, one for each user, both with 4 inputs and 2 outputs, i.e., we considered I and Q components of output signals from both users, and I and Q components of input signals for specific user.

The same PA model as for the SU case was used. Two 64 QAM modulate OFDM signals with a bandwidth of 20 MHz were generated in Matlab. Simulations of a 64 × 64 FC HBF MIMO system, with 2 users and 2 RF chains were performed, for both architectures shown in Figures 6 and 7.

Figures 8 and 9 show obtained power spectrum densities (PSDs) without DPD and with proposed RVTDNN2L DPD model for both types of architecture. While Tables 4 and 5 summarize the performance of the proposed RVTDNN2L DPD models for each user separately, in terms of EVMs and NMSEs. As can been seen, PA linearization is slightly better with RVTDNN2L DPD model with two NNs compared to RVTDNN2L DPD model with one NN. On the other hand, it should be noted that the architecture with two NNs is more complex, in terms of number of coefficients and FLOPs, compared to architecture with one NN, which is shown in Table 6.

From the above results, we conclude that compared to SU case it is necessary to further increase the complexity of the RVTDNN2L DPD system for MU case in order to efficiently linearize PAs, either by increasing the number of neurons in each layer, or by introducing another RVTDNN2L DPD for another user. Based on the results of extensive simulations, taking into account the complexity of the system, the best results were shown by the architecture with 2 RVTDNN2L networks, with 32 and 15 neurons in the first and the second hidden layer, respectively.

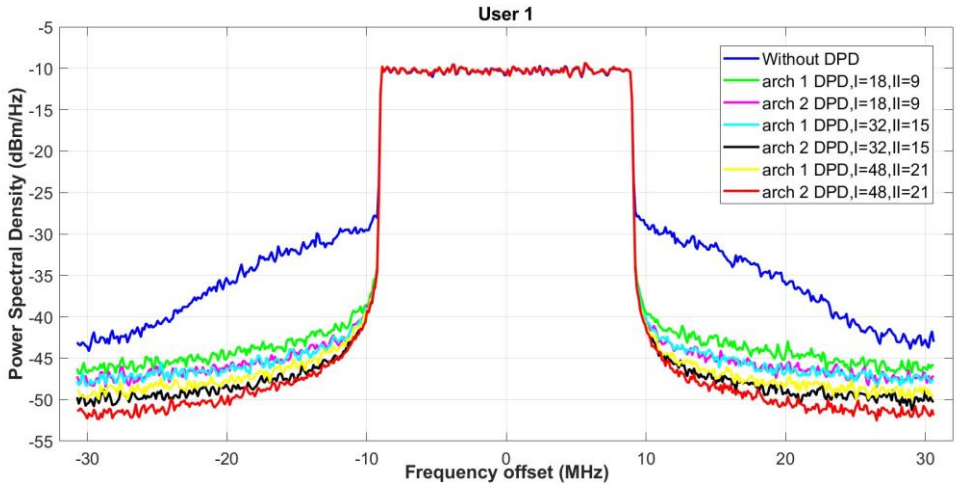

Figure 8. Power Spectral Density of the MU FC HBF mMIMO—user 1.

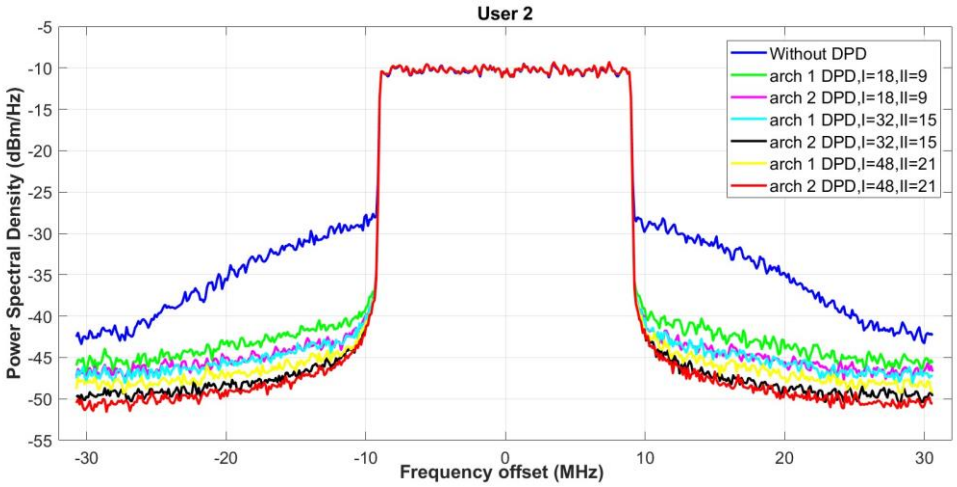

Figure 9. Power Spectral Density of the MU FC HBF mMIMO—user 2.

Table 4. Performance of RVTDNN2L DPD for 64 × 64 MU FC HBF MIMO system—user 1.

RVTDNN2L	Arch1		Arch2	
	EVM [%]	NMSE [dB]	EVM [%]	NMSE [dB]
I = 18, II = 9	4.07	−27.81	3.16	−30.02
I = 32, II = 15	3.08	−30.22	2.13	−33.45
I = 48, II = 21	2.48	32.11	1.78	−34.97

EVM without DPD is 17.43 %.

Table 5. Performance of RVTDNN2L DPD for 64 × 64 MU FC HBF MIMO system—user 2.

RVTDNN2L	Arch1		Arch2	
	EVM [%]	NMSE [dB]	EVM [%]	NMSE [dB]
I = 18, II = 9	4.56	−26.81	3.47	−29.19
I = 32, II = 15	3.34	−29.52	2.24	−32.98
I = 48, II = 21	2.69	−31.40	1.97	−34.11

EVM without DPD is 18.58 %.

Table 6. Performance and complexity of different RVTDNN2L DPD models.

RVTDNN2L	Arch1		Arch2	
	Ncoef	**FLOPs**	**Ncoef**	**FLOPs**
I = 18, II = 9	301	945	562	1818
I = 32, II = 15	719	2041	1374	3962
I = 48, II = 21	1357	3603	2626	7038

5. Conclusions

In this paper, the different NN based DPD architectures are formulated and tested. The RVTDNN2L DPD architecture in particular, which has been developed and implemented for the first time, best compensates nonlinear distortion of PAs in FC HBF mMIMO transmitters.

The developed RVTDNN2L DPD architecture was tested for linearization of FC HBF mMIMO system with more than 2 RF branches, and it was shown that, even though EVM without DPD increases with an increasing number of RF branches, the application of the proposed RVTDNN2L DPD model decreases total NMSE. Both SU and MU FC HBF mMIMO were analyzed and comprehensive simulations have been performed. It was shown that the proposed RVTDNN2L DPD model efficiently linearizes SU, as well as MU FC HBF mMIMO transmitters, where NMSE was of the same order of magnitude as in the existing DPD solutions for HBF mMIMO. As a result of all the above contributions, this study has shown that the developed RVTDNN2L DPD solution has very good PA modeling and linearization performance, in terms of NMSE and EVM, for FC HBF mMIMO transmitters.

Author Contributions: Conceptualization, T.M.-Z., N.N. and D.B.; methodology, T.M.-Z., N.N. and D.B.; software, T.M.-Z.; validation, T.M.-Z., N.N. and D.B.; formal analysis, T.M.-Z., N.N. and D.B.; investigation, T.M.-Z., N.N. and D.B.; resources, T.M.-Z., N.N. and D.B.; data curation, D.B.; writing—original draft preparation, T.M.-Z., N.N. and D.B.; writing—review and editing, T.M.-Z., N.N. and D.B.; visualization, T.M.-Z., N.N. and D.B.; supervision, N.N. and D.B.; project administration, T.M.-Z., N.N. and D.B.; funding acquisition, T.M.-Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Algorithm 1: NN learning algorithm for PA simulation

1. Determine the structure of NN as RVTDNN2L with 32 neurons in the first hidden layer and 15 neurons in the second hidden layer.
2. Get input data X_n and output data Y_n.
3. Define the cost function as MSE (Mean Square Error) function.
4. Define the optimization algorithm as LM (Levenberg-Marquardt) optimization algorithm.
5. Training NN:

- calculate NN output
- calculate cost function
- judgment: if performance requirements are met, exit the loop
- update network coefficients.

Appendix B

Matlab code for generating OFDM signal:

```
function [OFDMsignal]= createOFDM(Ndata, Nused_sc, FFT, M, m)
% Ndata=7200; % # of all transmitted data
```

```
% Nused_sc=1200; % #of used subcarrier 5MHz/15kHz LTE
Nsim=Ndata/Nused_sc;
% FFT=1024;
% M=64;
% m=2;
G=1/2^m;
CPx=G*FFT;
P1=zeros(Ndata,1);
P1 = randi(M-1, Ndata, 1);
% digital symbol mapped as analog symbol
Map = qammod(P1,M);
I1=imag(Map);
Q1=real(Map);
C1=I1.^2+Q1.^2;
U=mean(C1);
V=sqrt(U);
I=I1/V;
Q=Q1/V;
X1=[I Q];
X2=zeros(Ndata,1);
X2=X1(:,1)+i*X1(:,2);
X3=zeros(FFT,1);
x4=zeros(FFT,1);
m=zeros((FFT-Nused_sc),1);
for ii=0:(Nsim-1)
    X3=[X2(ii*Nused_sc/2+1 : Nused_sc/2+ii*Nused_sc/2,1); m; X2(Nused_sc/2+ii*Nused_sc+1:
Nused_sc+ii*Nused_sc,1)];
    x4=ifft(X3,FFT);
    x5(ii*(FFT+CPx)+1 : CPx+ii*(FFT+CPx),1)=x4(FFT-CPx+1 : FFT,1);
    x5(CPx+ii*(FFT+CPx)+1 : FFT+CPx+ii*(FFT+CPx),1)=x4;
end;
OFDMsignal = x5(:)';
```

References

1. Larsson, E.G.; Edfors, O.; Tufvesson, F.; Marzetta, T.L. Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **2014**, *52*, 186–195. [CrossRef]
2. Molisch, A.F.; Ratnam, V.V.; Han, S.; Li, Z.; Le Hong Nguyen, S.; Li, L.; Haneda, K. Hybrid Beamforming for Massive MIMO: A Survey. *IEEE Commun. Mag.* **2017**, *55*, 134–141. [CrossRef]
3. 3GPP Report TR 21.915 V15.0.0; *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects*; Release 15 Description; Summary of Rel-15 Work Items (Release 15); ETSI: Sophia Antipolis, France, 2019.
4. Zhang, T.; Dong, A.; Zhang, C.; Yu, J.; Qiu, J.; Li, S.; Zhou, Y. Hybrid Beamforming for MISO System via Convolutional Neural Network. *Electronics* **2022**, *11*, 2213. [CrossRef]
5. Hefnawi, M. Hybrid Beamforming for Millimeter-Wave Heterogeneous Networks. *Electronics* **2019**, *8*, 133. [CrossRef]
6. Masoudi, M.; Khafagy, M.G.; Conte, A.; El-Amine, A.; Francoise, B.; Nadjahi, C.; Salem, F.E.; Labidi, W.; Sural, A.; Gati, A.; et al. Green Mobile Networks for 5G and Beyond. *IEEE Access* **2019**, *7*, 107270–107299. [CrossRef]
7. Salh, A.; Audah, L.; Shah, N.S.M.; Alhammadi, A.; Abdullah, Q.; Kim, Y.H.; Al-Gailani, S.A.; Hamzah, S.A.; Esmail, B.A.F.; Almohammedi, A.A. A Survey on Deep Learning for Ultra-Reliable and Low-Latency Communications Challenges on 6G Wireless Systems. *IEEE Access* **2021**, *9*, 55098–55131. [CrossRef]
8. Borel, A.; Barzdėnas, V.; Vasjanov, A. Linearization as a Solution for Power Amplifier Imperfections: A Review of Methods. *Electronics* **2021**, *10*, 1073. [CrossRef]
9. Liu, L.; Chen, W.; Ma, L.; Sun, H. Single-PA-feedback digital predistortion for beamforming MIMO transmitter. In Proceedings of the 2016 IEEE International Conference on Microwave and Millimeter Wave Technology (ICMMT), Beijing, China, 5–8 June 2016; pp. 573–575.
10. Abdelaziz, M.; Anttila, L.; Brihuega, A.; Tufvesson, F.; Valkama, M. Digital Predistortion for Hybrid MIMO Transmitters. *IEEE J. Sel. Top. Signal Process.* **2018**, *12*, 445–454. [CrossRef]

11. Brihuega, A.; Anttila, L.; Abdelaziz, M.; Eriksson, T.; Tufvesson, F.; Valkama, M. Digital Predistortion for Multiuser Hybrid MIMO at mmWaves. *IEEE Trans. Signal Process.* **2020**, *68*, 3603–3618. [CrossRef]
12. Liu, X.; Zhang, Q.; Chen, W.; Feng, H.; Chen, L.; Ghannouchi, F.M.; Feng, Z. Beam-Oriented Digital Predistortion for 5G Massive MIMO Hybrid Beamforming Transmitters. *IEEE Trans. Microw. Theory Tech.* **2018**, *66*, 3419–3432. [CrossRef]
13. Liu, X.; Chen, W.; Chen, L.; Ghannouchi, F.; Feng, Z. Power Scalable Beam-Oriented Digital Predistortion for Compact Hybrid Massive MIMO Transmitters. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2020**, *67*, 4994–5006. [CrossRef]
14. Qiao, W.; Li, G.; Zhang, Y.; Li, H.; Liu, F. A Band-limited Digital Predistortion Method for Hybrid MIMO Transmitters. In Proceedings of the 2019 International Conference on Microwave and Millimeter Wave Technology (ICMMT), Guangzhou, China, 19–22 May 2019.
15. Liu, X.; Chen, W.; Chu, J.; Ghannouchi, F.; Feng, Z. Multi-Stream Spatial Digital Predistortion for Fully-Connected Hybrid Beamforming Massive MIMO Transmitters. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2021**, *68*, 2998–3011. [CrossRef]
16. Mengozzi, M.; Gibiino, G.; Angelotti, A.; Florian, C.; Santarelli, A. Over-the-Air Digital Predistortion of 5G FR2 Beamformer Array by exploiting Linear Response Compensation. In Proceedings of the 2022 IEEE/MTT-S International Microwave Symposium—IMS 2022, Denver, CO, USA, 19–24 June 2022; pp. 394–397.
17. Hongyo, R.; Egashira, Y.; Hone, T.; Yamaguchi, K. Deep Neural Network-Based Digital Predistorter for Doherty Power Amplifiers. *IEEE Microw. Wirel. Compon. Lett.* **2019**, *29*, 146–148. [CrossRef]
18. Wang, D.; Aziz, M.; Helaoui, M.; Ghannouchi, F. Augmented Real-Valued Time-Delay Neural Network for Compensation of Distortions and Impairments in Wireless Transmitters. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 242–254. [CrossRef]
19. Liu, Z.; Hu, X.; Liu, T.; Li, X.; Wang, W.; Ghannouchi, F. Attention-Based Deep Neural Network Behavioral Model for Wideband Wireless Power Amplifiers. *IEEE Microw. Wirel. Compon. Lett.* **2020**, *30*, 82–85. [CrossRef]
20. Hu, X.; Liu, Z.; Wang, W.; Helaoui, M.; Ghannouchi, F. Low-Feedback Sampling Rate Digital Predistortion Using Deep Neural Network for Wideband Wireless Transmitters. *IEEE Trans. Commun.* **2020**, *68*, 2621–2633. [CrossRef]
21. Jaraut, P.; Abdelhafiz, A.; Chenini, H.; Hu, X.; Helaoui, M.; Rawat, M.; Chen, W.; Boulejfen, N.; Ghannouchi, F.M. Augmented Convolutional Neural Network for Behavioral Modeling and Digital Predistortion of Concurrent Multiband Power Amplifiers. in IEEE Transactions on Microwave Theory and Techniques. *IEEE Trans. Microw. Theory Tech.* **2021**, *69*, 4142–4156. [CrossRef]
22. Jaraut, P.; Rawat, M.; Ghannouchi, F.M. Composite Neural Network Digital Predistortion Model for Joint Mitigation of Crosstalk, I/Q Imbalance, Nonlinearity in MIMO Transmitters. *IEEE Trans. Microw. Theory Tech.* **2018**, *66*, 5011–5020. [CrossRef]
23. Brihuega, A.; Anttila, L.; Valkama, M. Neural-Network-Based Digital Predistortion for Active Antenna Arrays Under Load Modulation. *IEEE Microw. Wirel. Compon. Lett.* **2020**, *30*, 843–846. [CrossRef]
24. Hu, X.; Liu, Z.; Yu, X.; Zhao, Y.; Chen, W.; Hu, B.; Du, X.; Li, X.; Helaoui, M.; Wang, W.; et al. Convolutional Neural Network for Behavioral Modeling and Predistortion of Wideband Power Amplifiers. *IEEE Trans. Neural Networks Learn. Syst.* **2021**, *33*, 3923–3937. [CrossRef]
25. Tehrani, A.S.; Cao, H.; Afsardoost, S.; Eriksson, T.; Isaksson, M.; Fager, C. A Comparative Analysis of the Complexity/Accuracy Tradeoff in Power Amplifier Behavioral Models. *IEEE Trans. Microw. Theory Tech.* **2010**, *58*, 1510–1520. [CrossRef]
26. Ghannouchi, F.M.; Hammi, O.; Helaoui, M. *Behavioral Modeling and Predistortion of Wideband Wireless Transmitters*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2015; ISBN 9781118406274.
27. Amin, S.; Landin, P.N.; Händel, P.; Rönnow, D. Behavioral Modeling and Linearization of Crosstalk and Memory Effects in RF MIMO Transmitters. *IEEE Trans. Microw. Theory Tech.* **2014**, *62*, 810–823. [CrossRef]
28. Ayach, O.E.; Rajagopal, S.; Abu-Surra, S.; Pi, Z.; Heath, R. Spatially Sparse Precoding in Millimeter Wave MIMO Systems. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 1499–1513. [CrossRef]
29. Cho, Y.S.; Kim, J.; Yang, W.Y.; Kang, C.G. *MIMO-OFDM Wireless Communications with MATLAB*; John Wiley & Sons (Asia) Pte Ltd.: Singapore, 2010; ISBN 9780470825617.
30. 3GPP Technical Report 38.901; *Study on Channel Model for Frequencies from 0.5 to 100 GHz. v16.1.0 (Release 16)*; ETSI: Sophia Antipolis, France, 2019.
31. Ghannouchi, F.M.; Hammi, O. Behavioral modeling and predistortion. *IEEE Microw. Mag.* **2009**, *10*, 52–64. [CrossRef]
32. Muškatirović-Zekić, T.; Čabarkapa, M.; Nešković, N.; Budimir, Đ. Compensation of nonlinear distortion in hybrid beamforming MIMO transmitters. *Int. J. Electr. Eng. Comput.* **2022**, *6*, 26–30. [CrossRef]