JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Digital Pre-Distortion Linearization of Power Amplifier using Neural Network

Arshdeep Singh[1], Amandeep Singh Sappal[2]

1 M.Tech Student at Department of Electronics and communication Engineering, Punjabi University, Patiala, Punjab - 147002, India

2 Associate Professor at Department of Electronics and communication Engineering, Punjabi University, Patiala, Punjab - 147002, India

Abstract: Power amplifiers lie at the front end of most radio frequency systems and are critical in ensuring the appropriate range of wireless systems. The persistent search of enhanced efficiency and linearity in power amplifiers has become a pivotal focus in modern wireless communication systems, driven by the escalating demands for higher data rates and improved signal quality. Digital pre-distortion has emerged as a well-known technique to address the inherent nonlinearities exhibited by power amplifiers, Digital pre-distortion has shown to be advantageous in systems employing higher order modulation formats. Digital pre-distortion enables power amplifiers to operate closer to their saturation point, thereby maximizing power efficiency without sacrificing signal integrity. In this paper, neural network technique has been used to model a digital pre-distorter for linearization of power amplifier used in 5G communication system. To compare the performance of proposed digital pre-distorter, cross-term memory polynomial based digital pre-distorter has also been modeled. Results show that proposed digital pre-distorter modeled using neural network has better performance. The paper is organized as follows: section I presents introduction, concept of digital pre-distortion is presented in section II, section III is concept of augmented real-valued time-delay neural network, proposed power amplifier linearization technique is presented in section IV and section V presents conclusion.

Keywords: adjacent channel power ratio, augmented real-valued time-delay neural network, digital pre-distortion, error vector magnitude, normalized mean square error, power amplifier.

I. INTRODUCTION

Power amplifiers are critical components in wireless communication systems, responsible for boosting the power of transmitted signals to ensure reliable communication over desired distances. However, power amplifiers often exhibit non-linear behavior, especially when operating at high power levels, which can introduce distortion and degrade signal quality [1]. To mitigate these issues and improve the overall performance of wireless systems, various linearization techniques have been developed and extensively studied [2]. Linearization techniques are crucial for enhancing the performance of power amplifiers in modern communication systems, where signal integrity and efficiency are paramount [3]. These techniques aim to minimize distortion and improve power efficiency, enabling the delivery of high-quality signals with minimal interference. The ongoing research and development in this field are driven by the increasing demands of wireless communication technologies, such as 5G and beyond, which require highly linear and efficient power amplifiers [4].

Feedback linearization is another powerful technique for linearizing power amplifiers, which involves using feedback control to cancel out the nonlinearities of the amplifier [5]. This approach has the advantage of linearizing non-linear system models, thereby enabling the avoidance of the complicated mathematics associated with non-linear problems [6]. However, the feedback linearization method has a singularity problem and loses the physical meaning of the obtained states [7]. The key idea behind feedback linearization is to design a controller that cancels the non-linear terms in the power amplifier's transfer function, resulting in a linear closed-loop system [8]. However, when the controller is not capable of exactly canceling the nonlinearity, such control strategy may provide unsatisfactory performance or even induce instability [9]. The effectiveness of feedback linearization depends on the accuracy of the power amplifier model and the design of the feedback controller [10]. To address the limitations of traditional feedback linearization techniques, advanced control strategies, such as adaptive control and robust control, have been proposed. Furthermore, the power amplifier's non-linearity often changes with changing operation points and conditions [11]. Feedforward linearization is yet another approach to mitigating distortion in power amplifiers, where a replica of the distorted signal is generated and subtracted from the main signal path. This technique requires careful design of the feedforward network to ensure accurate cancellation of the distortion components [12].

Digital pre-distortion is a widely used linearization technique that involves pre-distorting the input signal to compensate for the non-linear characteristics of the power amplifier. By applying an inverse distortion function to the input signal, the nonlinearities of the power amplifier can be effectively canceled out, resulting in a more linear overall response [13].

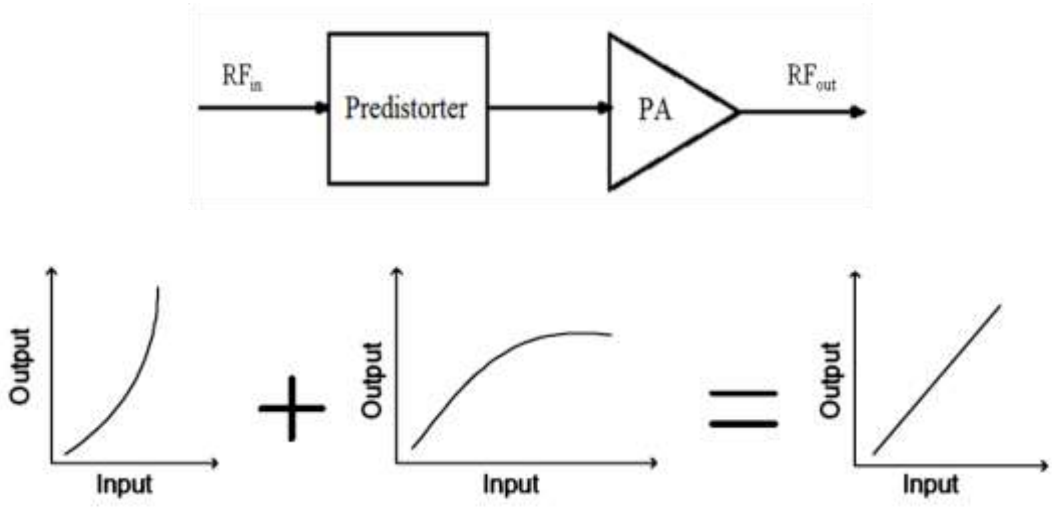

figure 1: concept of linearizer to be used with pa

Adaptive pre-distortion techniques dynamically adjust the pre-distortion function based on the changing characteristics of the power amplifier, ensuring optimal linearization performance over a wide range of operating conditions. The design of efficient and robust pre-distortion algorithms is an active area of research, with various approaches being explored, including look-up table-based methods, Volterra series-based techniques, and machine learning-based algorithms [14].

The selection of an appropriate linearization technique depends on various factors, including the specific requirements of the wireless communication system, the characteristics of the power amplifier, and the desired trade-offs between performance, complexity, and cost.

II. CONCEPT OF DIGITAL PRE-DISTORTION

The direct learning and indirect leaning architectures are used for implementation of a digital pre distorter. In direct learning architecture firstly, the behavioural model is defined and then its coefficients are estimated. Furthermore, the pre distortion function is identified from the implemented non-linear system. Whereas in indirect learning architecture, the model of digital pre distorter is assumed and then output of the PA is taken as the input of the digital pre distorter. After that the input of PA is taken as the output of the digital pre distorter. The identification of the non-linear system is not essential in the indirect learning architecture as required in direct learning architecture.

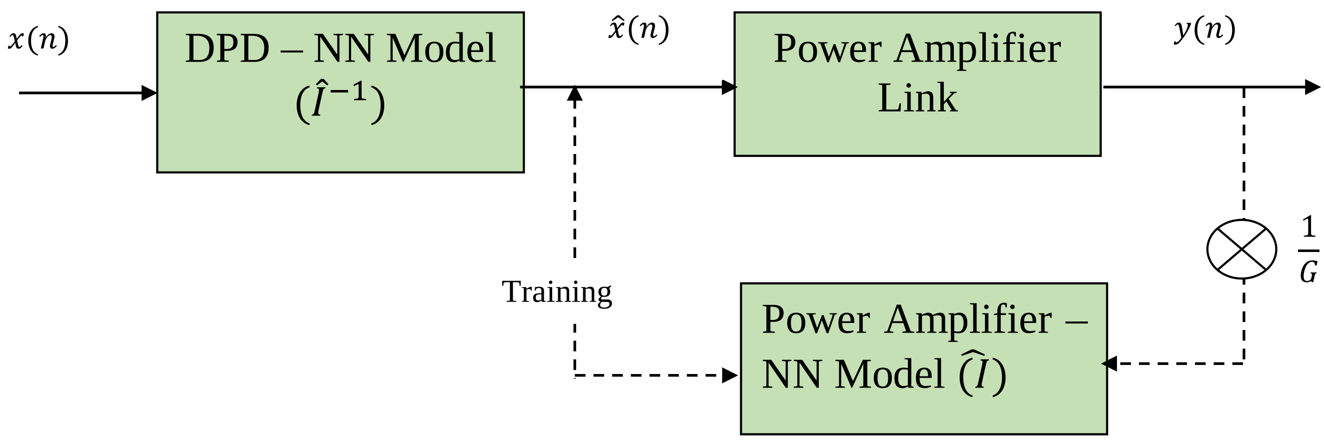

figure 2: modelling of power amplifier link

Therefore, the digital pre distorter parameters are estimated using an identification algorithm. As the use of classical adaptive learning algorithms is not possible in direct learning architecture hence in present work indirect learning architecture has been purposed to model the digital pre distorter by using neural network.

III. AUGMENTED REAL-VALUED TIME-DELAY NEURAL NETWORK (ARVTDNN)

ARVTDNN stands for Augmented Real-Valued Time-Delay Neural Network. It is a specialized neural network architecture designed primarily for applications in signal processing, particularly in the compensation of distortions and impairments in wireless communication systems. The key elements of ARVTDNN are:

1. Augmented Real-Valued: Most traditional neural networks operate on real numbers. In ARVTDNN, the network works directly with real-valued signals (instead of complex numbers or imaginary components). This is particularly useful in signal

processing tasks where real-valued inputs and outputs are more common (like in-phase and quadrature-phase components of a signal). The term augmented typically refers to adding additional features or dimensions to improve the model's performance. In ARVTDNN, augmentation may refer to adding extra parameters or modifying the network architecture to make it more robust to distortions and impairments in the communication signal.

2. Time-Delay Neural Network (TDNN): The "time-delay" aspect of the TDNN means that the network is designed to process sequences of data that have time dependencies. This is especially important in scenarios like wireless communication, where signals may undergo time shifts or delays due to factors such as fading or multi-path propagation. A TDNN is a type of neural network that is particularly effective at handling sequential data. It uses delays (memory) in its layers to process inputs over time, making it suitable for tasks where timing or sequential information is crucial. ARVTDNN consists of:

- Input: The network typically takes in real-valued components of the signal, which might include in-phase (I) and quadrature-phase (Q) components, depending on the application.
- Processing: The time-delay network uses its layers to learn the temporal relationships in the signal, compensating for time delays or phase distortions.
- Output: The output of the network is a corrected version of the input signal, where distortions and impairments have been minimized or removed.

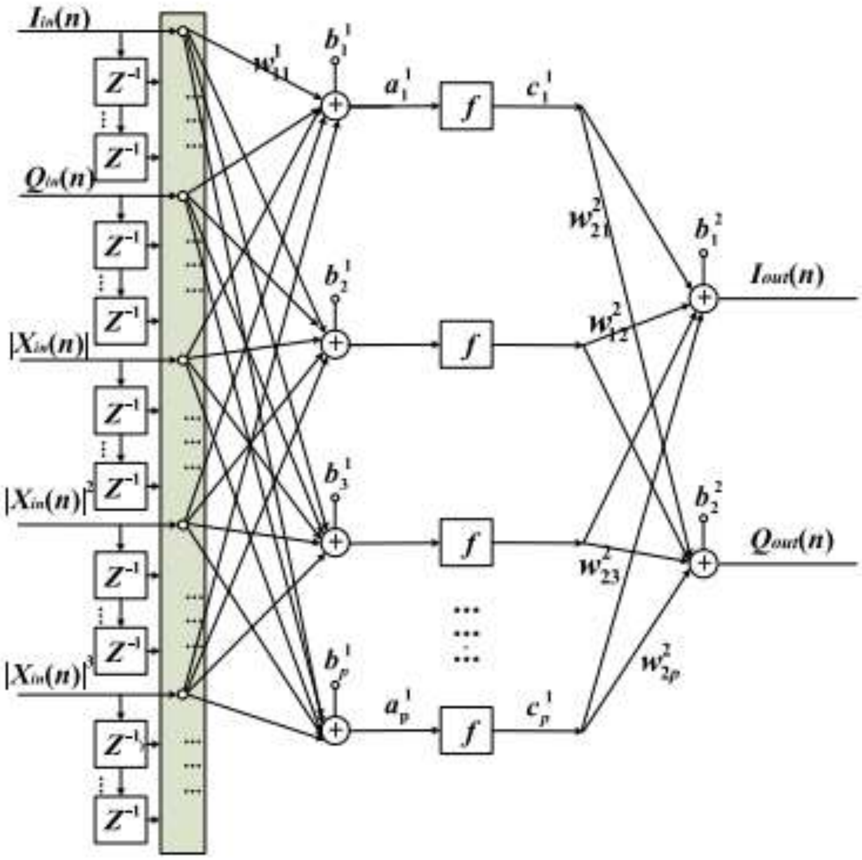

figure 3: block diagram of arvtdnn

This block diagram illustrates an Augmented Real-Valued Time-Delay Neural Network (ARVTDNN) architecture, commonly used for modelling nonlinearities and compensating impairments in wireless communication transmitters.

IV. PROPOSED POWER AMPLIFIER LINEARIZATION

In this work a power amplifier (PA) has been modeled using neural network (NN). Figure 4 shows the NN training workflow. During training, measure the input to the PA, *u*, used as the input signal and the output of the PA, *x*, used as the target signal.

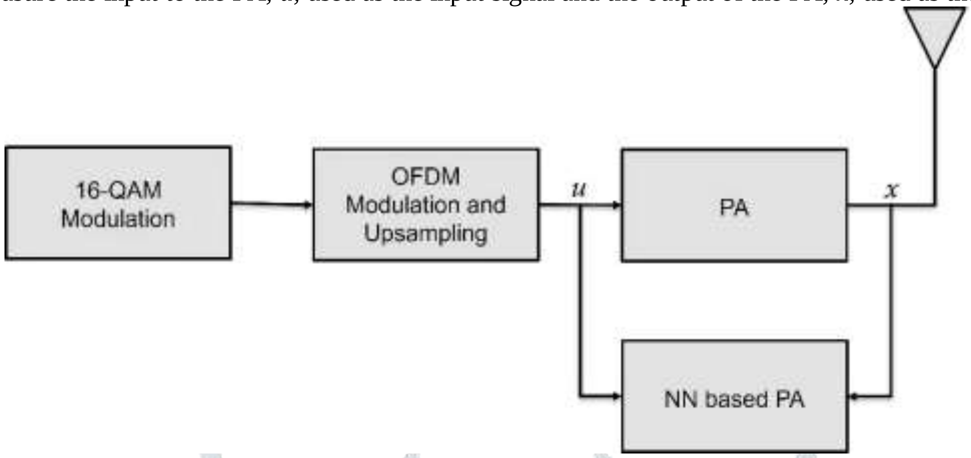

figure 4: neural network (nn) training workflow

The memory polynomial model has been commonly applied in the behavioral modeling and pre-distortion of PAs with memory effects. This equation shows the PA memory polynomial.

$$x(n) = f(u(n)) = \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} c_m u(n-m)|u(n-m)|^k \tag{1}$$

The output is a function of the delayed versions of the input signal, $u(n)$, and also powers of the amplitudes of $u(n)$ and its delayed versions. Since a neural network can approximate any function provided that it has enough layers and neurons per layer, you can input $u(n)$ to the neural network and approximate $f(u(n))$. The neural network can input $u(n-m)$ and $|u(n-m)|^k$ to decrease the required complexity. The input layer inputs the in-phase and quadrature components (I_{in} /Q_{in}) of the complex baseband samples. The I_{in} /Q_{in} samples and m delayed versions are used as part of the input to account for the memory in the PA model. Also, the amplitudes of the I_{in} /Q_{in} samples up to the k^{th} power are fed as input to account for the nonlinearity of the PA.

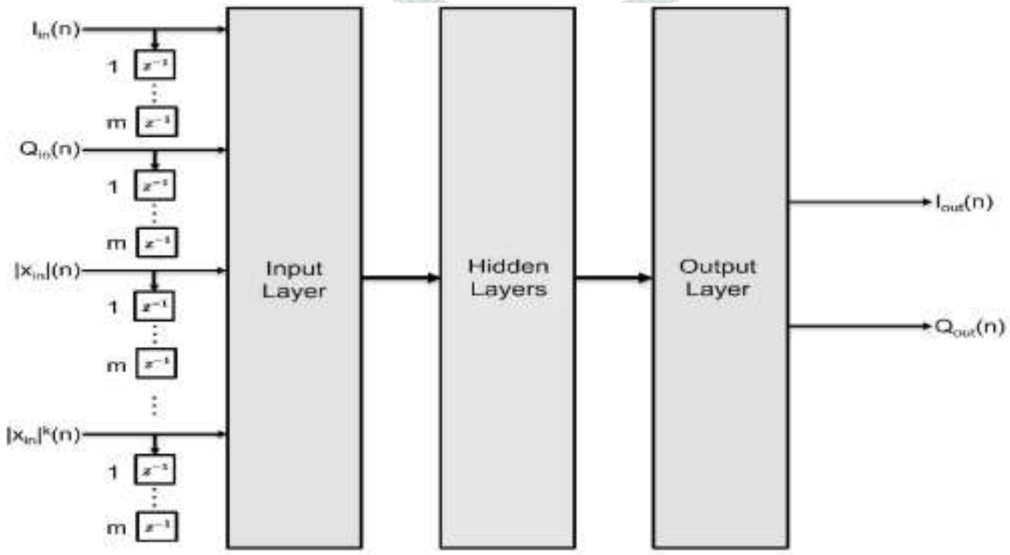

figure 5: neural network architecture

where,

$$I_{in}(n) = \Re(x(n))$$
$$Q_{in}(n) = \Im(x(n))$$
$$I_{out}(n) = \Re(u(n))$$
$$Q_{out}(n) = \Im(u(n)), \tag{2}$$

$\Re$ and $\Im$ are the real and imaginary part operators, respectively.

For comparison purpose memory polynomial and cross term memory polynomial is also used to model the DPD. Figure 6 shows the offline training workflow. First, we train an NN-DPD by using the input and output signals of the PA. Then, we use the trained NN-DPD. The upper path shows the neural network training workflow. During training, you measure the input to the PA, u, and the output of the PA, x. To train the neural network as the inverse of the PA and use it for DPD, use x as the input signal and u as the target signal. This process uses indirect learning. The lower path shows the deployed workflow with the trained NN-DPD inserted before the PA. In this configuration, the NN-DPD inputs the oversampled signal, u, and output, y, as the input to the PA. The PA output z is the linearized signal.

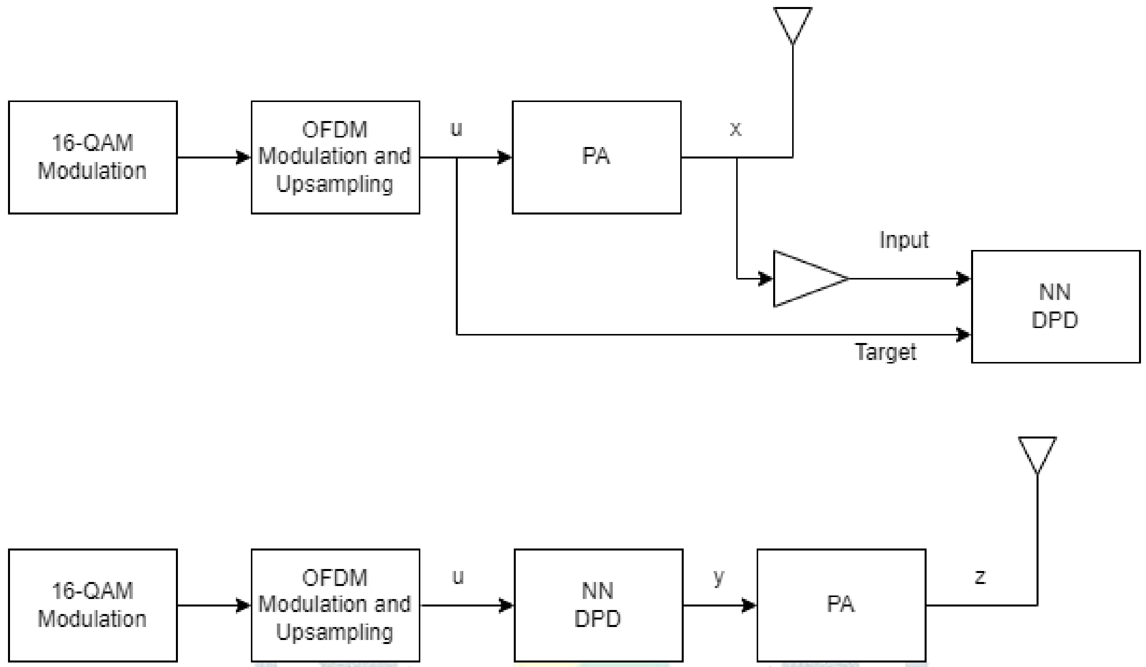

figure 6: nn-dpd offline work flow

We have designed same neural network as used for PA modeling (shown in Figure 5). The input $u(n)$ is applied to the neural network and approximate $f(u(n))$, or its inverse, which is the DPD function. To decrease the required complexity of the neural network in terms of number of layers and neurons, we have applied the memory polynomial approximation and provide the neural network extra features in the form of $u(n-m)$ and $|u(n-m)|^k$. The NN-DPD has multiple fully connected layers. The input layer inputs the in-phase and quadrature components (I_{in} /Q_{in}) of the complex baseband samples. The I_{in} /Q_{in} samples and m delayed versions are used as part of the input to account for the memory in the PA model. Also, the amplitudes of the I_{in} /Q_{in} samples up to the k^{th} power are fed as input to account for the nonlinearity of the PA.

During training, as equation (2)

$$I_{in}(n) = \Re(x(n))$$
$$Q_{in}(n) = \Im(x(n))$$
$$I_{out}(n) = \Re(u(n))$$
$$Q_{out}(n) = \Im(u(n)),$$

while during deployment

$$I_{in}(n) = \Re(u(n))$$
$$Q_{in}(n) = \Im(u(n))$$
$$I_{out}(n) = \Re(y(n))$$
$$Q_{out}(n) = \Im(y(n)), \quad (3)$$

where $\Re$ and $\Im$ are the real and imaginary part operators, respectively.

For training and validation of NN-DPD, we have used the input and output data of PA modeled using neural network. Before training the neural network DPD, we have selected memory depth of 5 and a nonlinear polynomial degree of 5. We have trained the neural network using the Adam optimizer with a mini-batch size of 1024. The initial learning rate is 4e-4 and decreases by a factor of 0.95 every five epochs. Figure 7 shows the training process. Random initialization of the weights for different layers affects the training process. To obtain the best root mean squared error (RMSE) for the final validation, we train the same network a few times.

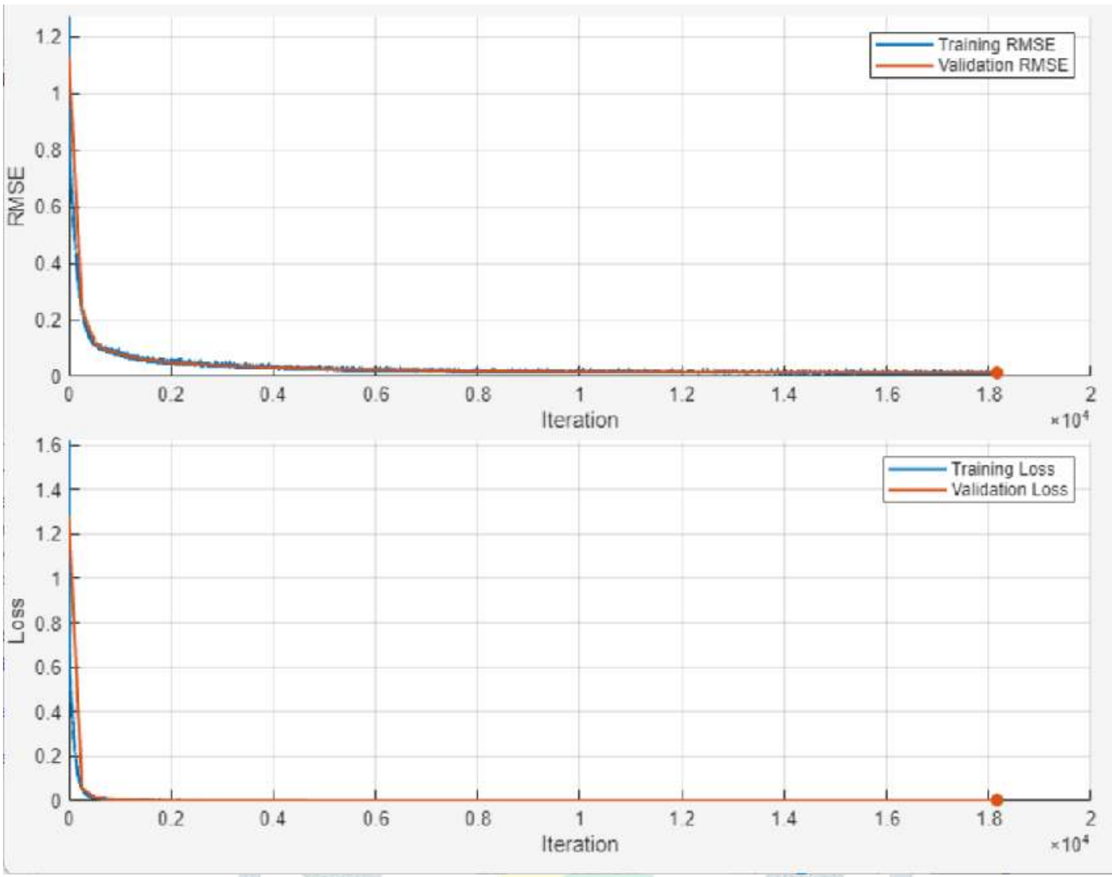

figure 7: training progress

To test the performance of NN-DPD, a simulation set-up has been implemented in Matlab as shown in Figure 8. To test the NN-DPD, we have passed the test signal through the NN-DPD and the PA and examined the normalized mean square error (NMSE) between the input to the NN-DPD and output of the PA, adjacent channel power ratio (ACPR) at the output of the PA, percent RMS error vector magnitude (EVM) which is measured by comparing the OFDM demodulation output to the 16-QAM modulated symbols. For comparison purpose, we have performed these tests for both the NN-DPD and also the memory polynomial based DPD.

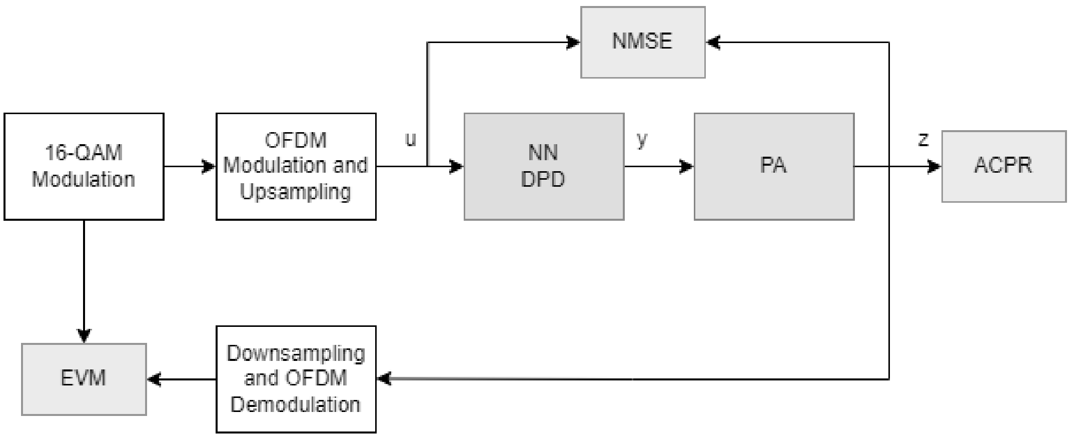

figure 8: set-up for performance evaluation of nn-dpd

table 1: comparison of NMSE, ACPR and EVM of PA with DPD

	ACPR (dB)	NMSE (dB)	EVM (%)
No DPD	-28.674	-21.287	6.5681
Cross-Term Memory Polynomial DPD	-33.889	-27.984	2.8229
Neural Network DPD	-33.889	-33.423	1.5679

Table 1 shows ACPR, NMSE and EVM values for 5G communication system with DPD and without DPD. These values are -28.674 dB, -21.287 dB and 6.5681 % respectively when no DPD is applied. When cross-term memory polynomial based DPD and neural network DPD is applied, these values show improvement. From comparison of these values, we conclude that neural network based DPD shows more improvement as compared to cross-term memory polynomial based DPD. When neural network based DPD has been used, the ACPR decreases to -33.889 dB, NMSE decrease to -33.423 dB and EVM decreases to 1.5679 %.

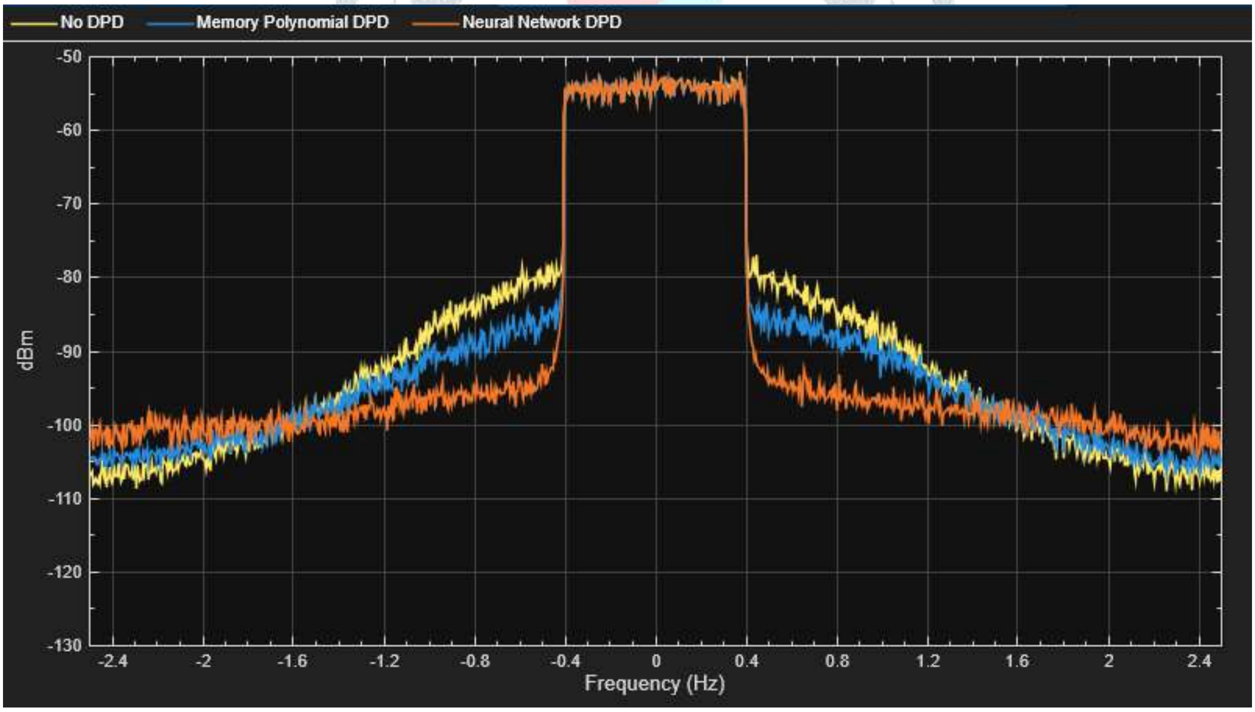

figure 9: power spectral density of pa with dpd

table 2: comparison of channel power and occupied bandwidth of pa with dpd

	Channel Power (dBm)	Occupied Bandwidth (MHz)
No DPD	-32.1404	792.2656
Cross-Term Memory Polynomial DPD	-32.1162	791.4987
Neural Network DPD	-32.1202	792.2866

Table 2 show that while using neural network based DPD, the system channel power and occupied bandwidth are almost same as no DPD is used. But these values deteriorate (decreases) when cross-term memory polynomial based DPD is used.

table 3: comparison of adjacent channel power ratio of pa with dpd

	ACPR 1		ACPR 2	
	Frequency (MHz)	**Power (dBc)**	**Frequency (MHz)**	**Power (dBc)**
No DPD	1.000	-34.6026	1.7500	-48.9291
Cross-Term Memory Polynomial DPD	1.000	-38.9680	1.7500	-48.6398
Neural Network DPD	1.000	-44.4161	1.7500	-49.8684

Table 3 show the comparison of ACPR 1 and ACPR 2 powers with DPD and without DPD. Results show that when neural network based DPD is used ACPR 1 value improves from -34.6026 dBc to -44.4161 dBc and ACPR 2 value improves from -48.9291 dBc to -49.8684 dBc.

table 4: comparison of imd products of pa with dpd

	$f1$ Power (dBc)	$f2$ Power (dBc)	$2f1-f2$ Power (dBc)	$2f2-f1$ Power (dBc)
No DPD	-52.1808	-52.7850	-27.2435	-32.1485
Cross-Term Memory Polynomial DPD	-52.4600	-52.8338	-33.9623	-34.7585
Neural Network DPD	-52.4820	-52.8465	-43.3109	-43.6242

Table 4 show that IMD products f1, f2, 2f1-f2 and 2f2-f1 occurs at 375.9766 MHz, 43.9453 MHz, 703.1250 MHz and 712.8906 MHz respectively. When neural network based DPD is applied these values shows highest improvement and these values improves to -52.4820 dBc, -52.8465 dBc, -43.3109 dBc and -43.6242 dBc respectively, which were -52.1808 dBc, -52.7850 dBc, -27.2435 dBc and -32.1485 dBc respectively when no DPD was applied.

It is worth to mention here that as the PA heats, its performance characteristics change. Thus, for parameter measurements, we have to send signal bursts through the PA repeatedly and plot system performance as a function of time. Each measurement takes about 6 s. Every 600 s, stop for 300 s to allow the PA to cool down.

V. CONCLUSION

Digital pre-distortion (DPD) allows power amplifiers to operate much closer to their saturation threshold—significantly improving their energy efficiency without compromising signal fidelity. In this work, we employ a neural-network based DPD model specifically designed to linearize a PA in a 5G communication system. To evaluate its effectiveness, we also implement a cross-term memory-polynomial predistorter as a reference model for comparison. This study reveals that a neural-network model can close to the behavior of an actual power amplifier across multiple key performance indicators—such as ACPR, EVM, NMSE, channel power, occupied bandwidth, and intermodulation distortion. In every case, the neural-network model demonstrated performance nearly identical to the real amplifier and surpassed the performance of traditional polynomial-based approaches. When integrated as a digital pre-distorter (NN-DPD), the neural-network model significantly reduced distortion in a simulated 5G system. Compared to cross-term memory polynomial DPD, the neural network achieved superior improvements in ACPR, NMSE, EVM, and intermodulation suppression. These findings support that neural-network-based modeling and pre-distortion provide a highly accurate and effective method for power amplifier linearization.

References

[1] G. Prasad, H. Johansson, and R. H. Laskar, “A General Approach to Fully Linearize the Power Amplifiers in mMIMO with Less Complexity,” arXiv (Cornell University), Jan. 2023, doi: 10.48550/arxiv.2309.04744.

[2] S. Gangadharan, R. Khanam, and V. Thangasamy, “A Study of RF Power Amplifiers for 5G and Future Generation Mobile Communication: Can FinFET Replace CMOS?,” International Journal of experimental research and review, vol. 46, p. 222, Dec. 2024, doi: 10.52756/ijerr.2024.v46.018.

[3] Z. Chang, “Ideal Advanced Amplifier Technology,” Highlights in Science Engineering and Technology, vol. 27, p. 163, Dec. 2022, doi: 10.54097/hset.v27i.3738.

[4] N. Singh et al., “Watt-Class CMOS-Compatible Power Amplifier,” p. 1, Jun. 2023, doi: 10.1109/cleo/europe-eqec57999.2023.10231735.

[5] L. Oliveira, A. Bento, V. J. S. Leite, and F. Gomide, “Comparisons of robust methods on feedback linearization through experimental tests,” IFAC-PapersOnLine, vol. 53, no. 2, p. 7983, Jan. 2020, doi: 10.1016/j.ifacol.2020.12.2218.

[6] S. Jiffri, P. Paoletti, J. E. Cooper, and J. E. Mottershead, “Feedback Linearisation for Nonlinear Vibration Problems,” Shock and Vibration, vol. 2014, p. 1, Jan. 2014, doi: 10.1155/2014/106531.

[7] Q. Quan and J. Ren, “State Compensation Linearization and Control,” arXiv (Cornell University), May 2024, doi: 10.48550/arxiv.2406.00285.

[8] H. Okajima, Y. Nishimura, and N. Matsunaga, “A Feedback Linearization Method for Non-linear Control Systems Based on Model Error Compensator,” Transactions of the Society of Instrument and Control Engineers, vol. 50, no. 12, p. 869, Jan. 2014, doi: 10.9746/sicetr.50.869.

[9] G. Innocenti and P. Paoletti, “A Novel Dissipativity-Based Control for Inexact Nonlinearity Cancellation Problems,” Mathematical Problems in Engineering, vol. 2015, p. 1, Jan. 2015, doi: 10.1155/2015/319761.

[10] G. Guardabassi and S. M. Savaresi, “Approximate linearization via feedback — an overview,” Automatica, vol. 37, no. 1, p. 1, Jan. 2001, doi: 10.1016/s0005-1098(00)00117-5.

[11] S. Akhlaghi and N. Zhou, “Adaptive Multi-Step Prediction based EKF to Power System Dynamic State Estimation,” arXiv (Cornell University), Jan. 2017, doi: 10.48550/arxiv.1702.00492.

[12] C.Y. Kai and A. Huang, “Linearization of rate-dependent nonlinearity with a compensator in feedback configuration,” Mechanical Systems and Signal Processing, vol. 39, p. 333, May 2013, doi: 10.1016/j.ymssp.2013.03.014.

[13] P. Reynaert, “Polar Modulation,” IEEE Microwave Magazine, vol. 12, no. 1, p. 46, Jan. 2011, doi: 10.1109/mmm.2010.939303.

[14] S. He and H. Chen, “Advanced Doherty Power Amplifier Architectures for 5G Handset Applications: A Comprehensive Review of Linearity, Back-Off Efficiency, Bandwidth, and Thermal Management,” Chips, vol. 4, no. 2. Multidisciplinary Digital Publishing Institute, p. 20, May 06, 2025. doi: 10.3390/chips4020020.