Behavioral Modeling of Direct-Conversion Transmitter by Incorporating Augmented concept with MoE

Manoj Bhatt and Sanjay Mathur*

Department of Electronics and Communication Engineering, College of Technology, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, India

mb914491@gmail.com, sanjaymathur.ece@gbpuat-tech.ac.in

* *Corresponding author.*

Abstract

In this research, an Augmented Mixture of Experts (A-MoE), a modified form of a committee machine i.e., (MoE), is used to provide a one-step solution for predicting dynamic non-linear behavior of wideband RF power amplifier (PA) along with other impairments of the direct-conversion transmitter (DCT) such as I/Q imbalance and DC offset. For data acquisition for modeling, we have utilized a class AB-PA-driven transmitter with a wideband three-carriers Long-Term Evolution (LTE) signal. The modeling performance for behavioral modeling of DCT obtained for the A-MoE model show better performance than multilayer perceptron (MLP) neural networks (NNs) and other machine learning methods for actual device data in term of normalized mean squared error (NMSE). Hence, the purposed method provides a noval solution for efficient behavioral modeling of PA or DCT's, reporting the performance level much better than comparative methods.

Keywords. PA nonlinearity, Mixture of Expert, Direct-Conversion Transmitter, transmitter impairments, Long-Term Evolution (LTE) signal.

1. INTRODUCTION

In modern communication systems, for enhancing spectrum efficiency with restricted spectrum resources, researchers have utilized higher-order modulation techniques like quadrature amplitude modulation (QAM), wideband code division multiplexing (WCDM), and orthogonal frequency division (OFDM). Due to wideband envelope varying signals with high peak-to-average power ratio (PAPR), these higher-order modulation techniques are more sensitive to the non-linear behavior of PA and produce more non-linear distortion than other similar average power leveled low PAPR modulation-based communication systems. Hence to compensate for nonlinearity in PA, different linearization techniques have been used in the digital domain of the transmitter section, in which digital predistortion is the most popular and widely used linearization technique [1].

In a wideband communication system, the high PA starts to show the memory effect because of the input signal's larger bandwidth. In such cases, memoryless models [2-4] attain limited linearization performance, and hence memory capable models comprising structures capable of modeling them are used. The Volterra series is taken as a reference model for the accurate modeling of dynamic (memory-based) non-linear systems [5]. A large number of coefficients (Volterra kernels) is one major drawback of the Volterra model and the number of coefficients increases exponentially in the Volterra model with an increase in the nonlinearity degree and memory depth of the series, which in turn surges the computational complexity of the model with slow convergence. A comparative study between various Volterra-based methods, such as Memory polynomial (MP), generalized memory polynomial (GMP), and Dynamic Deviation Reduction (DDR), etc. was done based on the complexity and performance of the models by Ghannouchi et al. [6] and concluded that these polynomial models present higher modeling accuracy of the narrow band signal or less complex PA or moderately non-linear PA. But these methods still have large hardware complexity and numerical instability due to dispersion coefficient and data matrix ill-conditioning.

A DCT chain/system of communication system has several imperfections, i.e., DC offset and I/Q imbalance, etc., and these imperfections are generated by the gain and phase mismatch behavior of the non-ideal local oscillator (LO) carrier leakage and the modulator [7], respectively. In various instances, PA nonlinearity-based distortions also additionally distort the performance of the communication system degrading the output signal. Most of the aforementioned techniques only emphasize PA nonlinearity-based distortions, which essentially generate the need for developing a convenient solution for compensating all the mentioned impairments associated with the DCT chain in parallel.

Recently, due to the capability of universal approximation and excellent adaptive nature, NNs are proficiently used for the behavioral modeling of static and dynamic PAs having time-delay tap [8]. Unlike existing linearization methods, such as Volterra and its variants, NN-based linearization methods [9,10] provide a one-step process solution for estimating and linearizing the DCT without any extra branch. NN models have also been integrated with the augmentation concept [10] etc. to provide enhanced performance.

This article uses the MoE method with an augmented concept for digital-domain behavioral modeling of PA or DCT's imperfections. In this line, the remaining paper is organized as follows. Section 2 gives the details of transmitter distortion parameters values and PA device/signal details utilized for implementing the behavioral modeling of the amplifier. Section 3 demonstrates the modeling methodology of MoE with the aforementioned concept. Section 4 presents the results and discussion related to the A-MoE method and compares its results with different neural networks and machine learning methods like Multilayer perceptron (MLP), Adaboost, K-Neighbors, Linear Regression, Linear Support Vector Regression (SVR), and Decision tree and lastly, a brief conclusion is presented in Section 5.

2. TRANSMITTER DISTORTION PARAMETERS AND PA DEVICE/SIGNAL INFORMATION

The complete equation of the impaired signal at the output of the modulator ($\mathrm{V}_{\mathrm{t-impair}}(\mathrm{t})$) of the DCT can be given as (1) [9].

$$\mathrm{V}_{\mathrm{t-impair}}(\mathrm{t}) = I_n(\mathrm{t})\cos(\omega_{\mathrm{c}}\mathrm{t}) - Q_d(\mathrm{t})\sin(\omega_{\mathrm{c}}\mathrm{t}) + \zeta(\mathrm{t}) \tag{1}$$

where $I_n(\mathrm{t})$ and $Q_d(\mathrm{t})$ are the I/Q components of the baseband input signal respectively, ω_{c} is the carrier's angular frequency, and $\zeta(\mathrm{t})$ is the overall impairment error demonstrated by the DC offset and I/Q imbalance [9].

Table 1 Different DCT distortion considerations

	Distortion Parameters	**Level of Distortion**
Condition 1	PA nonlinearity	PA nonlinearity: 4 dB compression
Condition 2	PA nonlinearity and I/Q imbalance	PA nonlinearity: Same as condition 1 I/Q imbalance: 1 dB gain compression and 3-degree phase compression
Condition 3	PA nonlinearity and DC offsets	PA nonlinearity: Same as condition 1 DC offsets: 3 and 5 % for I and Q respectively
Condition 4	PA nonlinearity, I/Q imbalance, and DC offsets	PA nonlinearity: Same as condition 1 I/Q imbalance: same as condition 2 DC offsets: same as condition 3

Table1 shows the several conditions of distortion utilized for behavioral modeling of DCT in the present work. The effect of these distortion parameters can be observed as AM/AM and AM/PM characteristics of the DCT in the form of distortion [11]. Since DPD linearization techniques are based on the capability of performing behavioral modeling using the considered model, and if the considered model is inept in precisely recording the effect of these modifications, the system performance might severely degrade [7].

The present work utilizes class AB-PA for developing behavioral modeling of a DCT system. The whole chain of the transmitter considered a device under test (DUT), is operated through a three-carriers wideband LTE input signal. This input signal carries two carriers in the ON state with the middle carrier in the OFF state (LTE101) at two bandwidths, 11 MHz (10.68 dB PAPR) and 16 MHz (11.39 dB PAPR). The other system conditions are 2 GHz center frequency with 92.16 MHz sampling frequency.

3. MODELING METHODOLOGIES OF A-MOE

Many methods have been introduced to enhance the performance of the solo DPD method. Out of them, the augmentation concept is one that has been successfully implemented [12] to show better behavioral modeling and mitigating abilities with respect to individual behavioral/DPD methods. The improved performance results inspire us to inculcate an augmentation scheme in the MoE method for enhancing the performance of MoE behavioral modeling. The architecture associated with A-MoE is presented in Figure1 (a). MoE [13] is an example of supervised learning and has a modular structure [14]. MoE has two core parts in its structure i.e., a set of experts network and a single gate network having a single-layer NN, comprising of K neurons where every neuron is allocated a specific expert. Both the experts and gate networks act synchronously to solve a non-linear supervised problem statement through Divide-and-Conquer operating principle [14]. Figure1(b) and (c) are representing the expert's signal flow graph and the gating network's signal flow graph respectively. If, the input vector to the kth expert is given by ($\mathbf{x}$), and (w_k) is the synaptic weight vector of this expert network along with the bias term (b), then the output (y_k) generated by expert k is given by equation (2)

$$y_k = w_k^T \boldsymbol{x} + b \tag{2}$$

where the input vector *($\boldsymbol{x}(n)$)* form utilized in the proposed work is given as,

$$\boldsymbol{x}(n) = \begin{bmatrix} I_{in}(n), I_{in}(n-1), I_{in}(n-2), I_{in}(n-3), \dots, I_{in}(n-M) \\ Q_{in}(n), Q_{in}(n-1), Q_{in}(n-2), Q_{in}(n-3), .., Q_{in}(n-M) \\ I_{in}(n)^3, I_{in}(n)^5, \dots., I_{in}(n)^N \\ Q_{in}(n)^3, Q_{in}(n)^5, \dots, Q_{in}(n)^N \\ |\boldsymbol{x}_{in}(n)|, |\boldsymbol{x}_{in}(n)|^3, \dots, |\boldsymbol{x}_{in}(n)|^L \end{bmatrix} \tag{4}$$

where $I_{in}(n)$ and $Q_{in}(n)$ are in and quadrature-phase elements of the current samples, $I_{in}(n-M)$ and $Q_{in}(n-M)$ are in and quadrature-phase elements of the past samples, *M*, *N*, and *L* represent the input signal's memory depth, phase component's odd orders, and input signal's absolute value respectively.

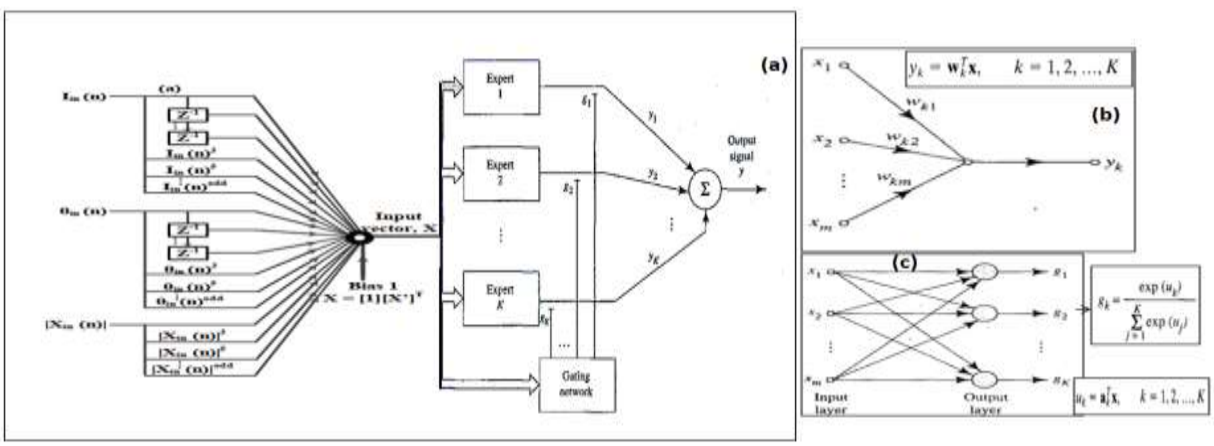

Figure 1. (a) Architecture of A-MoE, (b) Expert network's signal flow graph,
(c) Gating network's signal flow graph

The probability of determining the target vector corresponding to a specific input through the MoE model is calculated through the following equation (5) [13],

$$P\left(y|\boldsymbol{x}, \emptyset_g, \emptyset_e\right) = \sum_{l=1}^{L} g_l\left(\boldsymbol{x}, \emptyset_{\boldsymbol{g}}\right) P\left(y|\boldsymbol{x}, l, \emptyset_e\right) \tag{5}$$

where $g_l\left(\boldsymbol{x}, \emptyset_{\boldsymbol{g}}\right)$ represents gating function providing selection probability for l^{th} expert, $P\left(y|\boldsymbol{x}, l, \emptyset_e\right)$ is the occurrence probability of output y for a given input $\boldsymbol{x}$ using l^{th} expert, $\emptyset_g$ and $\emptyset_e$ are gate and expert's set of parameters. The modified gating form for the modeling has been adapted from Xu *et al*. [15]. Further, detailed information on the MoE method has been given in [11, 13].

Here, we are using two A-MoE networks for behavioral modeling of the DCT, where one network has been utilized for the input signal's in-phase components while the other is for the input signal's quadrature-phase component. The final output signal is obtained by combining these two outputs from individual A-MoE networks, i.e., in-phase and quadrature-phase components of the output signal, and the final form of the output signal is given by the following equation (8)

$$y_{out} = I_{out} + j * Q_{out} \tag{8}$$

4. SIMULATION RESULTS AND DISCUSSION RELATED TO DYNAMIC BEHAVIORAL (CLASS AB-PA) MODELING CAPABILITY OF A-MOE FOR VARIOUS SIGNALS

For modeling purpose, 13 number of experts have been optimized in the A-MoE model by analyzing NMSE performance, the most vital system performance element. The values of *M*, *N*, and *L* are 2, 5, and 3 respectively. The number of samples which are used for the training and validation purpose is 38 K and 35 K, respectively. NMSE is determined as,

$$\boldsymbol{NMSE = 10\,log_{10}\left(\frac{\sum_{i=1}^{Z}(I-I_{desired})^2+(Q-Q_{desired})^2}{\sum_{i=1}^{N}(I_{desired})^2+(Q_{desired})^2}\right)} \tag{9}$$

where *Z* is the sample length.

The performance related to behavioral modeling of the A-MoE model in terms of NMSE has been mentioned in Table2 for three carriers' LTE signals. The values of NMSE in Table2 reflect that the A-MoE model is efficiently capable of performing behavioral modeling corresponding to every signal with different impairment conditions of DCT which have already been mentioned in section 2.1.

Table 2. A-MoE behavioral modeling performance for DCT/ PA characteristics in terms of NMSE

A-MoE (LTE (4G) Signal)		**NMSE (dB)**	
		Training	**Validation**
LTE101 (3-5-3)	**Condition 1**	-37.15	-37.01
	Condition 2	-37.57	-38.10
	Condition 3	-36.92	-36.78
	Condition 4	-38.53	-38.04
LTE101 (3-3-10)	**Condition 1**	-38.63	-37.16
	Condition 2	-36.50	-36.01
	Condition 3	-35.92	-35.56
	Condition 4	-36.88	-35.62

Now, we are comparing the behavioral modeling performance of the proposed method with NNs and machine learning-based methods in terms of NMSE for both the LTE101 signals with different bandwidths for various impairment conditions**.** Table3 shows the NMSE of different methods with the A-MoE method, which shows that the A-MoE method, overshadows all the considered methods.

A-MoE methods performance capability was further analyzed in the time domain through AM/AM and AM/PM characteristics (Figure2); and frequency domain using a power spectral density (PSD) plot (Figure3). From Figure2, it can be observed that the proposed model precisely captures PA/DCT's non-linear characteristics in presence of I/Q imbalance and DC offset for the validation dataset and expresses its efficient capability for modeling.

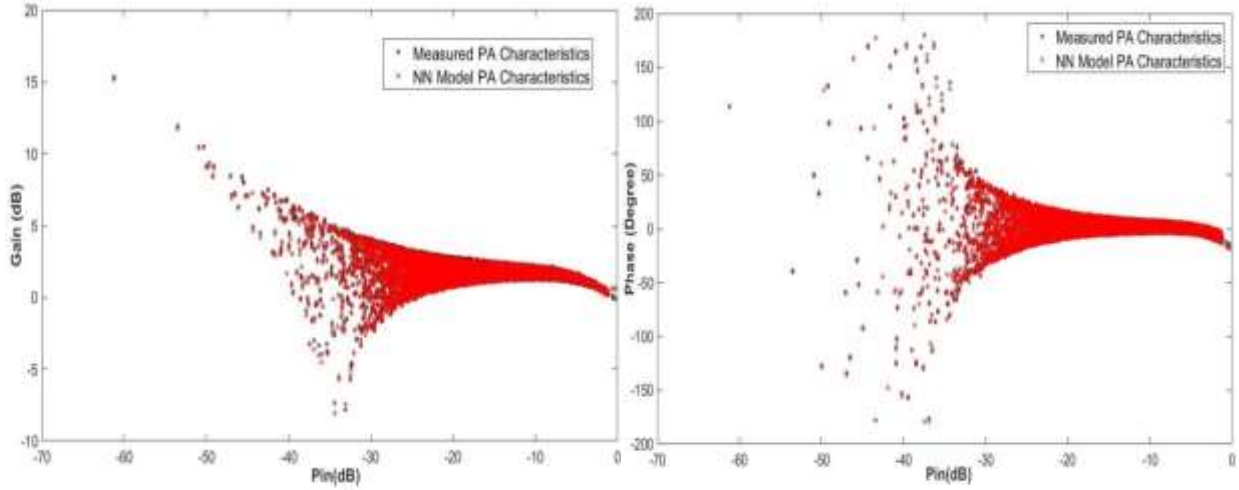

Figure 2. (a) AM/AM, (b) AM/PM characteristics of DCT's non-linear characteristics using proposed A-MoE model corresponding to LTE (101) 3-5-3 signal

Table 3. NMSE Performance comparison of A-MoE with different NNs and Machine learning methods for modeling of PA/DCT characteristics

Methods	Signal Bandwidth (LTE (4G) Signal) & Distortion Components							
	LTE101 (3-5-3) (TWO Memory elements in Input Data Vector)				LTE101 (3-3-10) (TWO Memory elements in Input Data Vector)			
	Condition 1	Condition 2	Condition 3	Condition 4	Condition 1	Condition 2	Condition 3	Condition 4
	NMSE	NMSE	NMSE	NMSE	NMSE	NMSE	NMSE	NMSE
A-MoE	**-37.01**	**-38.10**	**-36.78**	**-38.04**	**-37.16**	**-36.01**	**-35.56**	**-35.62**
MLP	-34.78	-35.85	-35.83	-35.80	-35.97	-35.91	-35.68	-34.61
Adaboost	-18.96	-18.38	-18.45	-18.33	-17.30	-17.74	-17.49	-17.58
K Neighbors	-34.22	-35.21	-35.21	-35.22	-35.44	-35.46	-35.48	-34.26
Linear Regression	-23.82	-23.93	-23.8	-23.85	-23.98	-23.99	-23.85	-23..80
Linear SVR	-22.77	-22.88	-22.77	-22.76	-22.97	-23.02	-22.82	-22.72
Decision Tree	-33.8	34.1	-34.20	-34.27	-34.37	-34.39	-34.39	-33.26

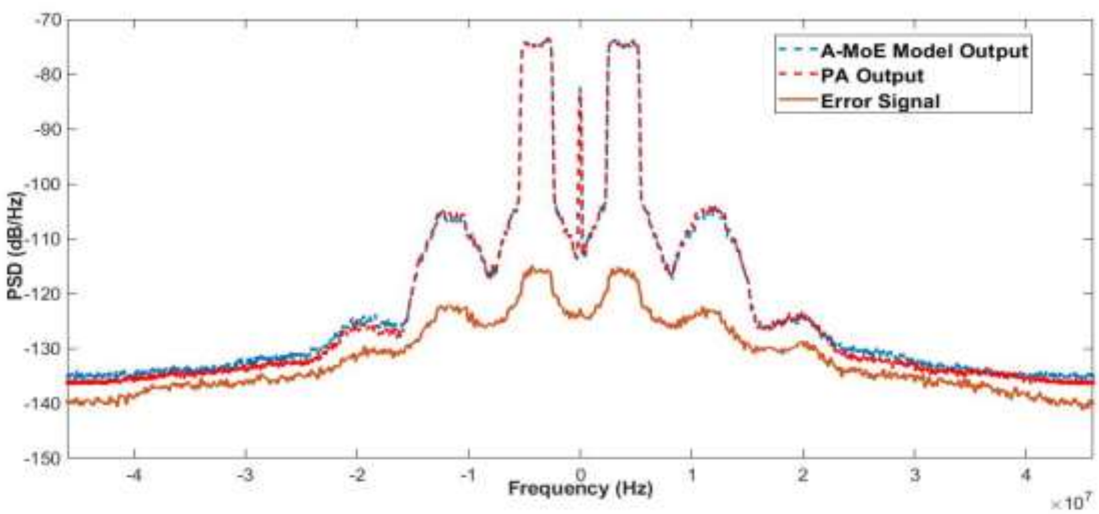

Figure 3. PSD plot with DCT's output corresponding to proposed A-MoE for LTE (101) 3-5-3 (Distortion consideration: Condition 4)

Figure3 depicts the frequency domain performance characteristics of the proposed A-MoE model. Both, the PSD plot of model output and measured output at DCT output accurately follow each other for in-band frequency as well as out-band frequency. This determines that the proposed A-MoE method is very much capable of performing behavioral modeling for entire-band data and the results significantly outperform as compared to the conventional NN methods.

5. CONCLUSION

In this paper, behavioral modeling analysis of DCT is achieved through the A-MoE framework. The modeling results of A-MoE has an excellent limit of resemblances with AM/AM and AM/PM characteristic of PA/DCT's imperfections. The method provides a single-step modeling solution to model any kind of distortion/imperfection as tabulated in the DCT chain without modifying the basic architecture of the system, which is an additional advantage as compared to other methods, where additional RF circuitry is always required to model I/Q imbalance and DC offsets. Based on performance metioned in Table 3, it is clear that the A-MoE method gives higher NMSE performance than other used methods, it is between 1 to 3 dB higher than MLP and K Neighbors, around 4 to 5 dB higher than tree- based ML method i.e., Decision tree and more than 5 dB as compared to other used machine learning methods for all the condition metioned in Table 1. So on the basis of this, we can clearly see the better performance results as compared to various NNs and machine learning methods like MLP, Adaboost, K Neighbors, Linear Regression, Linear SVR, and Decision Tree.

6. REFERENCES

[1] F. M. Ghannouchi, and O Hammi, 'Behavioral modeling and predistortion' 2009;52–64.

[2] Rapp C., Effects of HPA-nonlinearity on 4-DPSK / OFDM-signal for a digital sound broadcasting system', 2018.

[3] M. Ibnkahla, J. Sombrin, F. Castanie, et al., 'Neural networks for modeling nonlinear memoryless communication channels', IEEE Trans Commun. 1997;45:768–771.

[4] JHK. Vuolevi, T. Rahkonen, JPA. Manninen, 'Measurement technique for characterizing memory effects in RF power amplifiers', IEEE Trans Microw Theory Tech. 2001.

[5] S. Boyd, LO. Chua., 'Fading Memory and the Problem of Approximating Nonlinear Operators With Volterra Series', IEEE Trans circuits Syst. 1985;CAS-32:1150–1161.

[6] FM. Ghannouchi, O. Hammi, M. Helaoui, 'Behavioral Modeling and Predistortion of Wideband Wireless Transmitters. Behav. Model. Predistortion Wideband Wirel', Transm. 2015.

[7] JK. Cavers, 'The Effect of Quadrature Modulator and Demodulator Errors on Adaptive Digital Predistorters for Amplifier Linearization', IEEE Trans Veh Technol. 1997;46:456–466.

[8] M. Rawat and F. M. Ghannouchi, 'Distributed spatiotemporal neural network for nonlinear dynamic transmitter modeling and adaptive digital predistortion', *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, 2012.

[9] M. Rawat and F. M. Ghannouchi, 'A mutual distortion and impairment compensator for wideband direct-conversion transmitters using neural networks', *IEEE Trans. Broadcast.*, vol. 58, no. 2, 2012.

[10] D. Wang, M. Aziz, M. Helaoui, and F. M. Ghannouchi, 'Augmented Real-Valued Time-Delay Neural Network for Compensation of Distortions and Impairments in Wireless Transmitters', *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 1, pp. 242–254, 2019.

[11] M. Bhatt, M. Rawat, and S. Mathur, 'A Committee Machine Neural Network for Dynamic and its Inverse Modeling of Distortions and Impairments in Wireless Transmitters', IETE JOURNAL OF RESEARCH., 2021.

[12] M. Hui, T. Liu, M. Zhang, Y. Ye, D. Shen, and X. Ying, 'Augmented radial basis function neural network predistorter for linearisation of wideband power amplifiers', Electron. Lett., vol. 50, no. 12, pp. 877–879, 2014.

[13] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, 'Adaptive Mixtures of Local Experts', *Neural Comput.*, vol. 3, pp. 79–87, 1991.

[14] Haykin S. Neural networks: A comprehensive foundation. Pearson Educ. 1997.

[15] L. Xu, M. I. Jordan, and G. E. Hinton, 'An alternative model for mixtures of experts', Nips, no. 7, 1994.