Neural Networks Based Behavioral Modeling of Dual-Band RF Power Amplifiers using Augmented BiLSTM Structures

Majid Ahmed
Department of Electrical Engineering
College of Engineering
American University of Sharjah
Sharjah, United Arab Emirates
b00077868@aus.edu

Ahmad Dalbah
Intelligent RF Radio Laboratory
Department of Electrical and Software Engineering
University of Calgary
Calgary, AB, Canada
ahmad.dalbah@ucalgary.ca

Oualid Hammi
Department of Electrical Engineering
College of Engineering
American University of Sharjah
Sharjah, United Arab Emirates
ohammi@aus.edu

Fadhel M. Ghannouchi
Intelligent RF Radio Laboratory
Department of Electrical and Software Engineering
University of Calgary
Calgary, AB, Canada
fadhel.ghannouchi@ucalgary.ca

***Abstract*—This paper introduces a novel Bidirectional Long-Short-Term-Memory (BiLSTM) based Artificial Neural Network (ANN) model for the behavioral modeling of the nonlinear distortions observed in dual-band radiofrequency power amplification systems. Addressing the demand for energy-efficient and green wireless infrastructures in the era of multi-band transmissions, the proposed Augmented BiLSTM uses selected features of the input signals in each band to concurrently predict the output waveform in each band. Experimental benchmarking against the established 2D-Memory Polynomial (2D-MP) model, conducted with a ZHL-5W-2G_S+ power amplifier prototype driven by concurrent dual-band 5G signals, reveals that the proposed model is able to achieve comparable performance. In addition to its accuracy, the main advantage of the proposed neural network based model lies in its ability to simultaneously model the amplifier's behavior in both bands with a single trainable structure. This contrasts with conventional analytically defined models, which require a dedicated structure per signal transmission band.**

Keywords— 5G, artificial neural networks, behavioral modeling, distortions, memory effects, multi-band, power amplifiers.

I. Introduction

Power amplifiers (PAs) are one of the most critical electronic components in today's wireless communication infrastructure. A PA amplifies the radiofrequency (RF) signal, but because of its nonlinear nature of operation, it creates unwanted distortions. These distortions present themselves as either in-band distortions or out-of-band distortions. The in-band distortions contribute to increasing the error vector magnitude (EVM) of the RF signals, while the out-of-band distortions create unfilterable spectral regrowth that can interfere with other transmissions that share the licensed electromagnetic spectrum. Nevertheless, a PA can be operated in an almost linear manner, but this greatly decreases its efficiency. Hence, there is a trade-off between building an efficient RF transmission system and meeting the spectral emission specification of relevant communication standards which inevitably requires operating the power amplifier in its power-efficient and nonlinear mode and compensating the resulting distortions through the use of linearization techniques [1]-[3].

Modern communication systems employ dual-band and multi-band power amplifiers in order to accommodate multi-band transmissions to meet the rising demand for higher bandwidths and lower latencies. In such systems, the distortions exhibited by nonlinear PAs become more critical since they include the distortions due to the presence of each band alone, as well as the inter-band distortions due to the interaction of the signals being transmitted concurrently within the different frequency bands [4].

Digital predistortion is the most widely adopted technique for the compensation of nonlinear distortions in multi-band power amplifiers. Behavioral modeling of power amplifiers distortions goes alongside digital predistortion since, in their essence, they both require the modeling of the PA's nonlinearity. Several models and predistortion functions have been proposed for the case of dual-band nonlinear power amplifiers. These can be categorized as either analytically defined models [5]-[9], box-based models [10], or neural networks based models [11][12]. Analytically defined models of dual-band PAs started with the 2D-MP model [5]. Later, several attempts were made to reduce the complexity of this model [6][7], apply pruning techniques [8], or develop time-misalignment tolerant version [9]. Some of the artificial neural networks that have been utilized to successfully model dual-band concurrent systems include modified real-valued time-delay neural networks [11], and augmented convolutional neural networks [12].

Among the various neural networks structures used for the behavioral modeling and predistortion of nonlinear power amplifiers, Bidirectional Long-Short-Term-Memory (BiLSTM) networks represent a viable approach. In fact, it has been recently reported that for the case of single-band power amplifiers, BiLSTM networks can lead to resilient predistortion functions that can maintain acceptable linearization capabilities over a wide range of operating conditions [13]. In this work, the use of BiLSTM networks for the behavioral modeling of dual-band power amplifiers is investigated. The conventional BiLSTM structure is modified to take as input the baseband in-phase and quadrature components of the signals in each band. It is also augmented to take into account additional useful features of the input signals in both bands. The model is configured to provide a dual-band output including the baseband signals in both bands. The proposed model is validated using experimental data. The reported results clearly demonstrate its potential. In section II, the proposed augmented BiLSTM is presented along with the 2D-MP model which will be used as a benchmark. Section III is devoted to the description of the experimental setup and the validation of the proposed model. Finally, the conclusions are summarized in Section IV.

II. Augmented Dual-Band BiLSTM Model

While feedforward ANNs can effectively perform digital predistortion functions and yield satisfactory outcomes, these straightforward neural networks lack the capability to grasp the intrinsic memory effects associated with PAs. Therefore, employing specialized neural networks explicitly crafted to exploit the temporal dependencies within an input vector proves significantly more appropriate. One exemplary instance is the LSTM network. LSTMs can incorporate temporal information that is embedded in an input vector such that all of the relevant information embedded within are extracted. Hence, LSTMs provide a way of extracting useful memory information exhibited by the dynamic nonlinear behavior of power amplifiers as presented in [14]. BiLSTMs are an extension of LSTMs that process the input vector both in the forward and backward directions as suggested in Figure 1, utilizing distinct forward and backward LSTM layers. This procedure is repeated multiple times depending on the specified number of LSTM cells utilized.

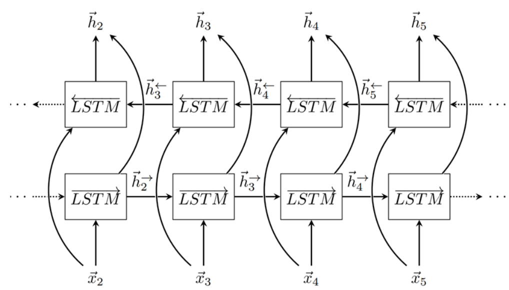


Fig. 1. Typical BiLSTM structure

This bidirectional approach enhances robustness and performance compared to standard LSTMs, as it facilitates the extraction of more temporal data from the input vector at the expense of an increased number of coefficients. Nevertheless, BiLSTMs have seen great success in creating resilient digital predistortion models as presented in [13], and being suitable for use with relaxed feedback sampling rates [15]. Hence, it is important to assess the potential of BiLSTM structures for multi-band nonlinear power amplification systems.

In this work, an Augmented BiLSTM model for behavioral modelling of dual-band nonlinear power amplifiers is proposed. As suggested by some of the previous works [12], data preprocessing greatly helps in easing the requirements of ANNs as by doing so, many of the relevant nonlinear information is pre-extracted from the input data such that the neural networks architecture of choice is mainly focusing on extracting relationships between the model's relevant input features and its output. Hence, rather than feeding the BiLSTM model with only the in-phase and quadrature components of the baseband signals in each band, a feature vector is constructed to include components that are related to the basis functions of analytically defined models. As a result, the input signal of a given band is transformed into the following matrix:

$$X_i(n) = \begin{bmatrix} I_i(n) & I_i(n-1) & \cdots & I_i(n-M) \\ Q_i(n) & Q_i(n-1) & \cdots & Q_i(n-M) \\ |x_i(n)| & |x_i(n-1)| & \cdots & |x_i(n-M)| \\ \vdots & \vdots & & \vdots \\ |x_i(n)|^K & |x_i(n-1)|^K & \cdots & |x_i(n-M)|^K \end{bmatrix} \quad (1)$$

where $X_i(n)$ is the input features matrix of the input signal in band i, $x_i(n)$ is the baseband input waveform in band i, $I_i(n)$ and $Q_i(n)$ are the in-phase and quadrature components of the input signal $x_i(n)$. M and K are the model's memory depth and nonlinearity order, respectively.

For the dual-band operation, the proposed augmented BiLSTM model is fed by two input features matrices $X_1(n)$ and $X_2(n)$ which are concatenated. The output layer of the BiLSTM model will provide the estimated in-phase and quadrature components of the output complex baseband waveforms in each of the transmission bands $(I_{out,1}(n),\ Q_{out,1}(n),\ I_{out,2}(n),\text{ and } Q_{out,2}(n))$. The proposed model can be generalized to the case of multi-band transmission systems with N bands by concatenating the N input features matrices $X_1(n)$ through $X_N(n)$. The model will simultaneously output the in-phase and quadrature components in all N bands as depicted in Figure 2.

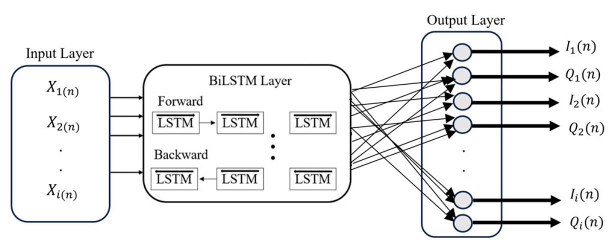


Fig. 2. Simplified block-diagram of the proposed augmented BiLSTM multi-band model

The baseline model for dual-band nonlinear power amplifiers is the 2D-MP which is an extension of the well-established single-band memory polynomial dedicated to the case of dual band systems. The 2D-MP model was used in this work as the benchmark model to assess the performance of the proposed BiLSTM model. Let's consider the case of a dual-band amplifier driven by two RF signals with $x_1(n)$ and $x_2(n)$ the baseband waveforms corresponding to the signals in band 1 and band 2, respectively. The 2D-MP model structure can be used to predict the amplifier's baseband complex output waveforms through

$$\tilde{y}_i(n) = \sum_{m=0}^{M}\sum_{k=0}^{K}\sum_{l=0}^{k} a_i^{klm} x_i(n-m)\left|x_j(n-m)\right|^l \left|x_i(n-m)\right|^{k-l} \quad (2)$$

where $\tilde{y}_i(n)$ is the estimated baseband output waveform in band i, $x_i(n)$ is the baseband input waveform in band i, and $x_j(n)$ is the baseband input waveform in the other band (j). a_i^{klm} are the coefficients of the 2D-MP model used to predict the output in band i. M and K represent the model's memory depth and nonlinearity order, respectively.

The 2D-MP as well as its low complexity and pruned variants [5]-[9], use the baseband complex waveforms in each band to separately predict the corresponding output in each band. This means that for dual-band applications, two 2D-MP models are needed to predict the dual-band output waveform since each 2D-MP will predict the signal in one of the bands. Hence, the extension to the case of multi-band with N-bands would require the identification of N different $N-$dimensional memory polynomial models.

III. Experimental VAlidation

A. Device under Test and Data Acquisition

The device under test (DUT) used in these experiments is the ZHL-5W-2G_S+ high power amplifier prototype by Minicircuits. The DUT operates in the 0.8GHz to 2GHz frequency range. It was fed by two 5G new radio (NR) signals centered at 1.0GHz and 1.6GHz, respectively. The lower band signal, centered at 1.0GHz, has a 10.9dB peak to average power ratio (PAPR), and a bandwidth of 100MHz. The higher band signal, centered at 1.6GHz, has a PAPR of 10.7dB, and a 100MHz bandwidth. The signals were generated using the Zynq Ultrascale+ RFSoC ZCU216 evaluation board. The signals at the output of the power amplifier were acquired individually at a sampling rate of 491.52Msps each using the FSW signal and spectrum analyzer from Rhode and Schwarz. The block diagram of the experimental setup is illustrated in Figure 3.

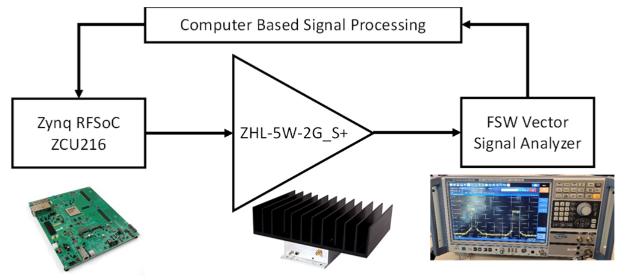
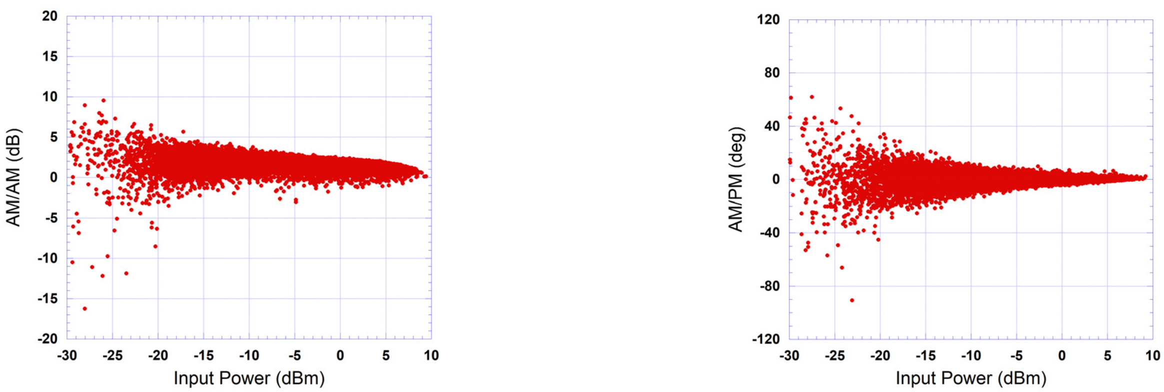


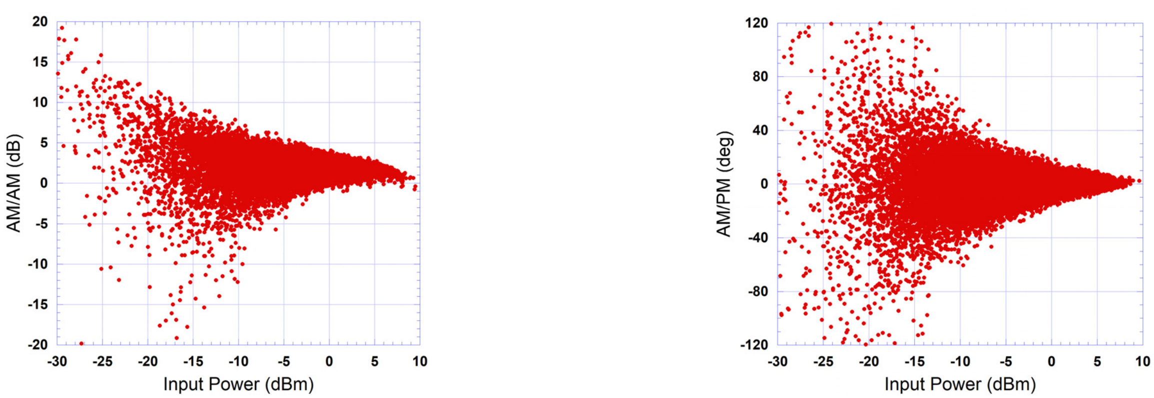
Fig. 3. Block diagram of the experimental setup

The acquired baseband output waveforms were first processed to compensate for time-delay misalignment and then used to derive the lower-band and upper-band AM/AM and AM/PM characteristics of the device under test. These are reported in Figure 4 and Figure 5 for the lower, and upper frequency bands, respectively. As it can be seen through the nonlinear shape of these characteristics and their substantial dispersion, the device under test exhibits a mild nonlinearity along with strong memory effects.



(a) (b)

Fig. 4. AM/AM and AM/PM characteristics of the DUT at the lower frequency band. (a) AM/AM characteristic, (b) AM/PM charactersitic



(a) (b)

Fig. 5. AM/AM and AM/PM characteristics of the DUT at the upper frequency band. (a) AM/AM characteristic, (b) AM/PM charactersitic

B. Augmented BiLSTM Identification and Performance Assessment

The device under test was modeled using the proposed augmented BiLSTM dual-band neural network and the benchmark 2D-MP model. First, the memory depth and nonlinearity order of the DUT were determined using a simple exhaustive search through which the parameters of the 2D-MP model were swept. It was found that a 2D-MP with a memory depth of 5 and nonlinearity order of 4 results in the best trade-off between modelling performance and computational complexity.

The proposed Augmented BiLSTM model was set up with the same memory depth and nonlinearity order as the 2D-MP (that is $M = 5$ and $N = 4$). The model was trained utilizing the ADAM optimizer and the mean squared error as a cost function for 60 epochs. The *tanh* activation function was utilized for the single hidden layer, which is composed of 64 forward and reverse LSTM cells. A linear activation function was used for the output layer. Furthermore, the Augmented BiLSTM model was trained using only 20% of the sampled signals, which is comparable to other neural network based models [12]. However, the training subset can be reduced to as low as 8% while maintaining the modelling performance with an NMSE degradation of less than 1.5 dB.

First, the normalized mean squared error (NMSE) was computed for the 2D-MP model and the proposed BiLSTM model. The NMSE is given by:

$$NMSE_{i,dB} = 10\log_{10}\left(\frac{1}{L}\sum_{n=1}^{L}\frac{\left|\tilde{y}_i(n)-y_i(n)\right|^2}{\left|y_i(n)\right|^2}\right) \tag{3}$$

where $NMSE_{i,dB}$ corresponds to the NMSE in band i . L represents the number of samples in the $\tilde{y}_i$ and y_i waveforms. These two waveforms are the estimated and measured baseband waveforms in band i , respectively.

The NMSE results obtained in each band for each model are summarized in Table 1. This table shows that the proposed model leads to slightly better accuracy than the 2D-MP model in each of the two bands. To value the significance of this result, it is important to note here that typically neural networks based models lead to slightly lower performance than analytically defined models while allowing for better generalization capabilities. In this work, comparable performances between the 2D-MP and the Augmented BiLSTM were achieved by the proper selection of the input features fed into the ANN model which included the same basis functions as those used in the 2D-MP. Moreover, this helped reduce the size of the BiLSTM structure needed to achieve such modeling accuracy.

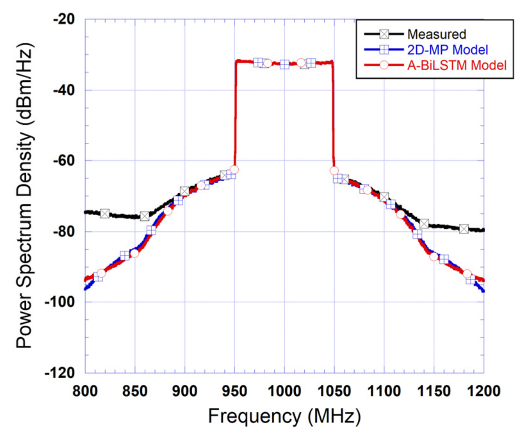
The two models were also compared by considering their frequency domain performance through the comparison of the spectrum of the DUT's estimated output signal along with the measured one in the lower and upper frequency bands. These results, reported in Figure 6, corroborate the conclusions observed through the NMSE data. As it can be seen in this Figure, the spectra estimated by both models are quasi-identical in both bands.

Furthermore, the estimated AM/AM and AM/PM characteristics of the DUT in the lower and upper bands were derived from the 2D-MP model and the augmented BiLSTM model. These characteristics were compared to the measured ones. As shown in Figure 7, both models are able to accurately mimic the behavior of the DUT in each of the two bands. More specifically, while the proposed model concurrently estimates the output signal in both bands with the same structure, it is able to distinguish between the strong memory effects exhibited by the DUT in the upper frequency band and the weaker memory effects observed in the lower frequency band as it can be seen through the difference in the dispersion of the predicted AM/AM and AM/PM characteristics in both bands.

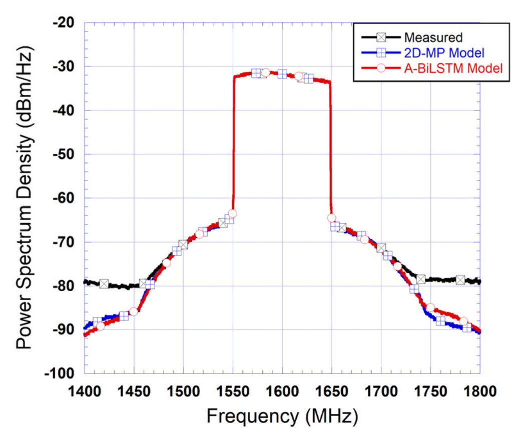
Figure 8 depicts the magnitude and phase of the complex baseband output waveforms samples in the time domain for the lower and upper frequency bands signals. The results presented in this figure show the ability of the model to accurately predict the time domain samples in each frequency band, and are in line with the results of Figure 6 and Figure 7.

TABLE I. NORMALIZED MEAN SQUARED ERROR OF THE PROPOSED MODEL

Frequency Band	Model	
	2D-MP (Benchmark)	***Augmented BiLSTM (Proposed)***
Lower Band	–37.2 dB	–37.6 dB
Upper Band	–39.5 dB	–39.6 dB



(a)



(b)

Fig. 6. Frequency domain performance assessment of the proposed Augmented BiLSTM model (a) Estimated output spectrum for the lower frequency band, (b) Estimated output spectrum for the upper frequency band

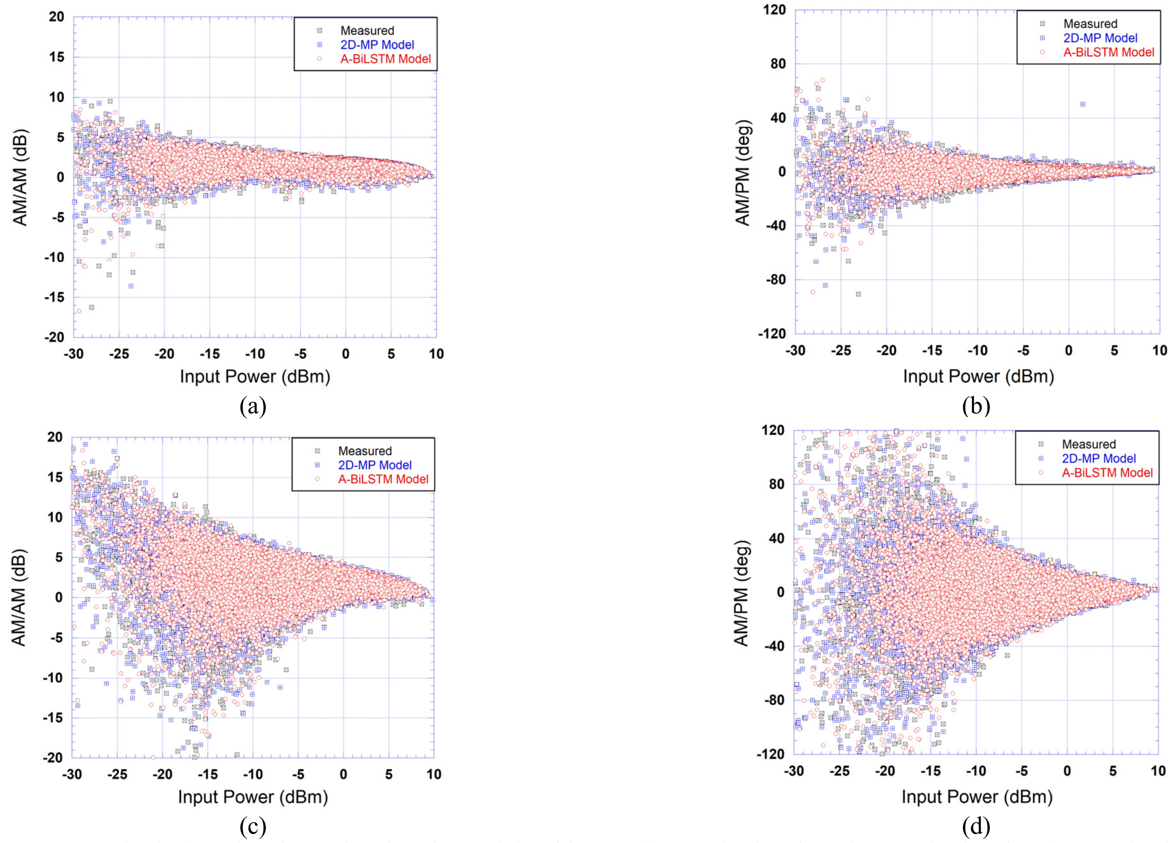


Fig. 7. Measured and estimated AM/AM and AM/PM characteristics of the DUT (a) Lower-band AM/AM, (b) Lower-band AM/PM, (c) Upper-band AM/AM, (d) Upper-band AM/PM

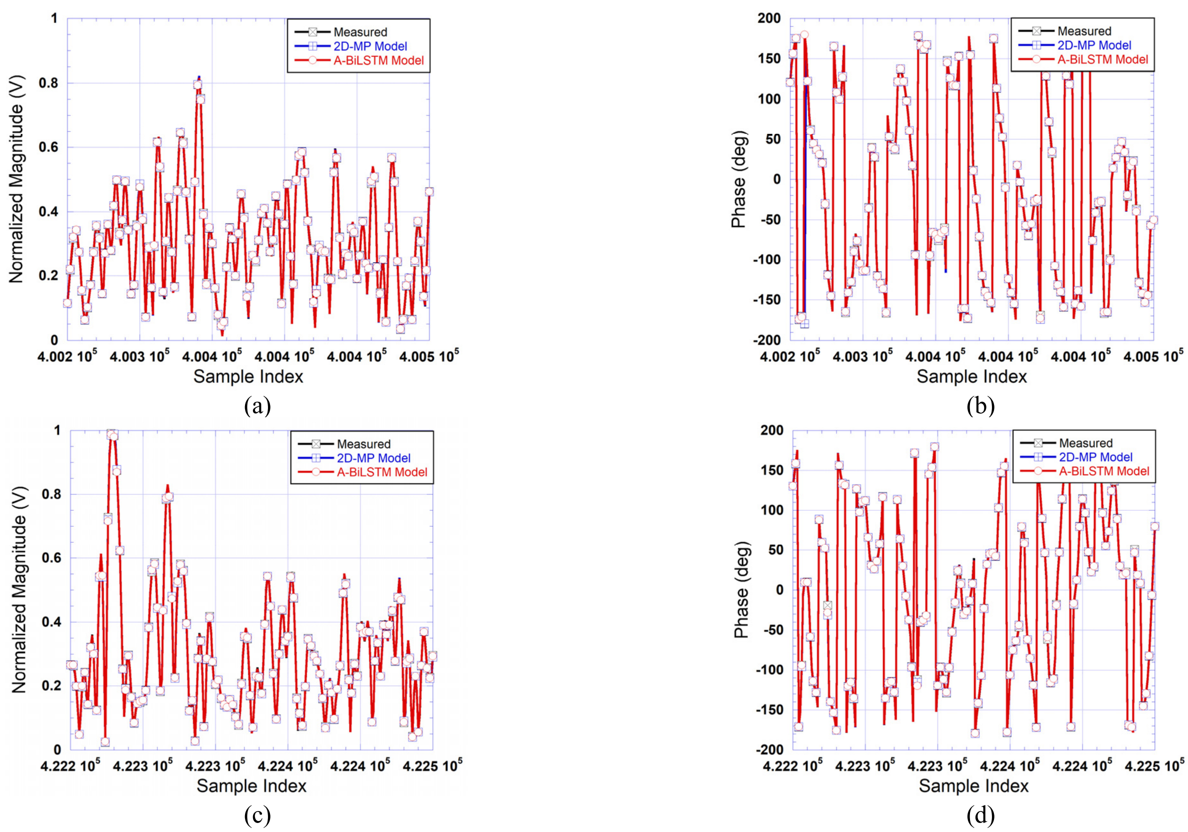


Fig. 8. Measured and estimated DUT baseband output waveforms (a) Lower-band magnitude, (b) Lower-band phase, (c) Upper-band magnitude, (d) Upper-band phase

IV. Conclusion

In this work, an augmented BiLSTM model was introduced for the behavioral modelling of multiband nonlinear power amplifiers systems. The model was tested under a concurrent dual-band transmission scenario for 5G applications, and its performance was compared to that of the 2D-MP benchmark model. The reported results indicate that the proposed augmented BiLSTM model achieves acceptable results comparable to the 2D-MP whilst utilizing a single trainable model instead of two separate models (i.e. one model for each band separately). The main advantage of the proposed model is its scalability for multiband transmission systems as the proposed framework can be implemented for any number of concurrent transmission bands. This is not the case for memory polynomial-based models as when the number of concurrent bands is increased, the number and complexity of the models required to separately model each band increases. Furthermore, the main benefit of the proposed model lies in the capability of ANNs to mimic the inherent behavior of the system and be resilient to changes in the operating conditions. Future work will consider investigating the resilience capability of the proposed augmented BiLSTM multi-band model.

References

[1] F.M. Ghannouchi, O. Hammi, and M. Helaoui, "Behavioral modeling and predistortion of wideband wireless transmitters," New Jersey, NJ, USA: Wiley, 2015.

[2] P.L. Gilabert, G. Montoro, D. Vegas, N. Ruiz, and J.A. Garcia, "Digital predistorters goes multidimensaional: DPD for concurrent multiband envelope tracking and outphasing power amplifiers," *IEEE Microw. Mag.*, vol. 20, no. 5, pp. 50-61, May 2019.

[3] F.M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, no. 7, pp. 52-64, December 2009.

[4] Y. Yahn, J.I. Martinez-Lopez, P. Roblin, R. Alsulami, Z. Popovic, and V. Chen, "Analysis and digital predistortion of in-band cross modulation in concurrent multi-band transmitters," *IEEE Access*, Vol. 10, pp. 120462-120474, November 2022.

[5] S.A. Bassam, M. Helaoui, and F.M. Ghannouchi, "2-D digital predistortion (2-D-DPD) architecture for concurrent dual-band transmitters," *IEEE Trans. Microw. Theory Tech.*, vol. 59, no. 10, pp. 2547-2553, October 2011.

[6] Y.J. Liu, W. Chen, J. Zhou, B.H. Zhou, and F.M. Ghannouchi, "Digital predistortion for concurrent dual-band transmitters using 2-D modified memory polynomials," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 1, pp. 281-290, January 2013.

[7] J. Zhai, Q. Zhang, J. Niu, L. Zhang, Z. Yu, J. Zhou and C. Yu, "A 2-D simplified memory polynomial model for concurrent dual-band power amplifiers," *IEEE Microw. Wireless Comp. Letts.*, vol. 30, no. 8, pp. 761-763, August 2020.

[8] S. Zhang, W. Chen, F.M. Ghannouchi and Y. Chen, "An iterative pruning of 2-D digital predistortion model based on normalized polynomial terms," *in IEEE International Microw. Symp. Digest*, June 2013.

[9] S. Zhang, W. Chen, Y.J. Liu, and F.M. Ghannouchi, "A time misalignment tolerant 2D-memory polynomials predistorter for concurrent dual-band power amplifiers," *IEEE Microw. Wireless Comp. Letts.*, vol. 23, no. 9, pp. 501-503, September 2013.

[10] C. Quindroit, N. Naraharisetti, P. Roblin, S. Gheitanchi, V. Mauer, and M. Fitton, "2D forward twin nonlinear two-box model for concurrent dual-band digital predistortion," *in IEEE Topical Conf. on Power Amplifiers for Wireless and Radio Applications Proc.*, January 2014.

[11] Z. Huang, W. Chen, Z. Feng, and F.M. Ghannouchi, "Forward behavioral modeling of concurrent dual-band power amplifiers using extended real valued time delay neural networks," *in IEEE International Conf. on Microwave and Millimiter Wave Technology Proc.*, May 2012.

[12] P. Jaraut, A. Abdelhafiz, H. Chenini, X. Hu, M. Helaoui, M. Rawat, W. Chen, N. Boulejfen, and F.M. Ghannouchi, "Augmented convolutional neural network for behavioral modeling and digital predistortion of concurrent multiband power amplifiers," *IEEE Trans. Microw. Theory Tech.*, vol. 69, no. 9, pp. 4142-4156, September 2021.

[13] A. Ali, and O. Hammi, "Bandwidth, power and carrier configuration resilient neural networks digital predistorter," *IEEE Access*, Vol. 11, pp. 63126-63135, June 2023.

[14] D. Phartiyal and M. Rawat, "LSTM-deep neural networks based predistortion linearizer for high power amplifiers," *in National Conf. on Communications Proc.*, February 2019.

[15] A. Ali, M. Ahmed, and O. Hammi, "BiLSTM neural network digital predistorter with reduced feedback sampling rate," *in IEEE Symp. on Wireless Technology and Applications Proc.*, August 2023.