CHALMERS
UNIVERSITY OF TECHNOLOGY

Behavioral Modeling of Power Amplifiers with Machine Learning on Multi Carrier and Multi Band Scenarios

Master's thesis in Networks and Distributed Systems

Saad Abbas Abbasi

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021

Master's thesis 2021:NN

Behavioral Modeling of Power Amplifiers with Machine Learning on Multi Carrier and Multi Band Scenarios

Saad Abbas Abbasi

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of
Chalmers University of Technology
Gothenburg, Sweden 2021

Modeling of Power Amplifiers with Machine Learning Techniques on Multi Carrier and Multi Band Scenarios
Saad Abbas Abbasi

Supervisor University: Prof. Thomas Eriksson and Yibo Wu, Department of Electrical Engineering, Chalmers University of Technology
Supervisor Company: Dr. Sener Dikmese, Ericsson AB, PDU, Radio Algorithms
Host Company: Ericsson AB
Examiner: Prof. Thomas Eriksson, Department of Electrical Engineering

Master's Thesis 2021:NN
Department of Electrical Engineering
Communication Systems group
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 72 539 8101

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2021

Behavioral Modeling of Power Amplifiers with Machine Learning on Multi Carrier and Multi Band Scenarios
Abbasi, Saad Abbas
Department of Computer Science and Engineering
Chalmers University of Technology

Abstract

The purpose of this thesis is to explore the possibility of using machine learning (ML) algorithms for the behavioral modeling of power amplifier (PA). This thesis is an extension of the previous masters thesis work [5] which compares the performance of ML methods with MP, generalized memory polynomial (GMP) and LUT methods for the PA modelling on single carrier and single band scenario. We expand it to multicarrier and multiband scenarios. The performance of MP and GMP as the baseline algorithms are compared with the performance of ML algorithms as neural network (NN), gradient boosting (GB), DT and LR in terms of normalized mean square error (NMSE) and adjacent channel error power ratio (ACEPR). The experiments are done with three different test scenarios for single carrier as a reference case, multi carrier in full band, and multi carrier in separated band/carrier. Experiment results show that NN achieves the best performance in terms of NMSE and ACEPR for all scenarios except for the separate multi-carrier scenario, where GB and GMP performs better for the signal with 400 and 600 MHz IBW, respectively. Finally, computational complexity analysis of ML algorithms is given in the final chapter before conclusion.

Keywords: power amplifier, memory polynomial, generalized memory polynomial, machine learning.

Acknowledgements

I am specially thankful to Prof. Thomas Eriksson for supervising this thesis and his help during the whole thesis period. I would like to thank my thesis supervisor at Ericsson AB, Dr. Sener Dikmese who guided and helped me at every step of this journey, from formulating the thesis proposal to helping me out with my thesis report. I am also grateful to my co-supervisor from Chalmers Yibo Wu who regularly helped me with his knowledge and experience. I am thankful to my manager Fredrik Linder for being supportive throughout the whole process. Finally I am really grateful to my whole family, my parents, brother and my wife for being there for me and supporting me throughout my life. Finally this thesis is dedicated to my younger sister Dr. Sarah Abbas Abbasi who left us in the prime of her life and heartbroken.

Saad Abbas Abbasi, Stockholm, January 2021

Contents

Acronyms

3GPP 3rd generation partnership project
5G fifth generation
ACEPR adjacent channel error power ratio
DL down link
DPD digital pre-distortion
DT decision trees
FCC federal communications commission
FLOP floating point operation
GB gradient boosting
GMP generalized memory polynomial
HB high band
IBW instantaneous bandwidth
ILC iterative learning control
IM inter-modulation
LB low band
LR linear regression
LS least squares
LSB LSboost
LUT look-up table
ML machine learning
MP memory polynomial
MSE mean square error
NMSE normalized mean square error
NN neural network
NR next radio
PA power amplifier
PSD power spectral density
RF radio frequency

List of Figures

List of Tables

1
Introduction

This chapter briefly explains the purpose of this thesis. It also describes the problem and its possible solutions that we explore as part of the thesis work.

1.1 Background

Ericsson develops many variants of radio hardwares every year. After the introduction of fifth generation (5G) the number of radio variants that are produced yearly has increased exponentially. For these radios, the power amplifier (PA) on the down link (DL) branch is a component with a non-linear response. digital pre-distortion (DPD) is used to increase the linear response cutoff point for the PA. An adaptive DPD is used in most cases which means that it is adapted according to the PA response.

During the radio development process different algorithms are used for PA and DPD modelling. Currently algorithms like memory polynomial (MP), generalized memory polynomial (GMP) and look-up table (LUT) based algorithms are used. In this thesis different machine learning (ML) algorithms are evaluated for PA modeling of multi-carrier/multi-band signals.

1.2 Problem Statement

For the next generation of wireless communication, new application cases considering two or more carrier combinations with wide bandwidths bring new challenges for DPD. In order to fulfill the requirements of both federal communications commission (FCC) and the 3rd generation partnership project (3GPP), modifications on the current DPD are expected. Volterra series based models such as the MP, GMP and models based on LUT have been widely used for PA and DPD modeling during radio development.

While GMP typically provides excellent PA modeling performance and outperforms the other modeling methods [28], the main disadvantage is the high computational complexity. A different version of LUT based approaches [29], which are significantly less complex than GMP, still do not fulfill the requirements of both FCC and 3GPP. Thus, the main motivation is to study alternative DPD solutions using modern ML algorithms for PA modeling.

It has been verified from latest master thesis work in our group [5] that as the bandwidth grows, the performance of PA modeling using ML methods does not go down as much as the GMP or LUT, however only single carrier was considered. Hence, the investigation of ML based PA modeling considering both wide-band and multi-carrier cases is the primary goal of this thesis.

1.3 Objectives

The purpose of this thesis is to investigate if ML algorithms can be used to model PA for multi-carrier and multi-band signals instead of baseline algorithms such as MP and GMP during the radio pre-development phase. The performance is evaluated using the normalized mean square error (NMSE) and adjacent channel error power ratio (ACEPR) from the modeled output data.

1.4 Methodology

For this thesis, a quantitative experimental research methodology is used. PA data is collected from our offline test-bed where the signal generator and analyzer is used as the transmitter and receiver. Three scenarios are considered, single-band/single-carrier, full-band multi-band/multi-carrier, and separated multi-band/multi-carrier. The input signal to the PA and the PA response signal is used for training and testing different models. NMSE and ACEPR results are used as the figure of merits for the predicted signal compared to baseline algorithms like MP and GMP.

1.5 Expected Results

The expectation is that ML models give similar results compared to baseline algorithms MP and GMP. The power spectral density (PSD) results are also verified visually to check the validity of the output signal. Considering the results obtained from the previous master's thesis[5], the expectation is that NN will still give the best results in terms of both NMSE and ACEPR especially during modeling of wide-band signals.

1.6 Report Structure

Chapter 2 provides the fundamentals of PA modeling and DPD approaches. Chapter 3 gives a general description of ML techniques and specific information related to the used ML approach in this thesis. Chapter 4 gives an overview of the scenarios considered during this study and describes the performance evaluation matrices like NMSE and ACEPR. The experimental results and the measurement setup description are presented in Chapter 5. Finally, closing remarks and are provided in Chapter 6.

1.7 Literature Review

This thesis builds upon the findings of two previous research papers and one master's thesis. The first paper deals with the modelling of PA considering different ML techniques [4]. The second research paper publishes the results of real-valued time-delay convolutional NN for wide-band PAs [22]. The previous master's thesis evaluated different ML models with traditional algorithms like MP/GMP for wide single band carrier together with complexity analysis [5].

This thesis can be divided into three different technical areas. The first area is related to the basic concepts of the non-linearity of PA and the methods used to increase the threshold for linearity such as the DPD. For basic PA concepts the book [6] is used as a reference, and for PA behaviors and DPD, the references [11] and [4] are used.

The second area is PA, DPD algorithms such as Volterra series and MP/GMP [8] [9] [10], [5] and other research articles given in the reference section study the algorithms for PA modelling since many of them give brief description on these basic concepts before moving to the specific algorithms in the research paper.

The third area is related to ML NN, GB, decision trees (DT) and linear regression (LR) are selected for this study because of the results obtained from the previous thesis [5] and support in MATLAB's ML toolbox. For basic concepts related to ML like supervised/unsupervised learning a book by Giuseppe B [12] is studied.

We consider [1] as a reference of NNs concepts. GB is investigated in [2] and [3]. Additionally, GB with least square boost is studied in a research paper by Harsh H P and Purvi P [13].

2

Power Amplifier Modeling and Digital Pre-Distortion

This chapter discusses the fundamentals of PA, DPD and different algorithms used for modelling of PA and DPD like Volterra series and its special cases like GMP and MP.

2.1 Power Amplifier Overview

In order to meet quality requirements of the signal at receivers end, the signal is amplified using radio frequency (RF) PAs [6], which makes it a vital component of the DL chain.

2.2 Non-linearity in Power Amplifiers

PAs in wireless communication systems exhibit nonlinear behavior which causes distortions to at the output signal. This non-linearity is the result of memory effects which can be classified in two categories [30]. The first category is classified as electo-thermal or thermal memory effects and it is the function of temperature change in the junction of transistors. This causes a long term effect and affects narrow bandwidth of signal spectrum [30]. The second category can be classified as electrical memory effects. These effects are produced by external terminations, including parasitic elements and matching networks of PA. These effects shape the PA response around the carrier frequency [30].

The signal distortions appear in form of harmonic distortion, gain compression, inter-modulation distortion, phase distortion, adjacent channel interference, etc [7]. Thus, it is necessary to get rid of those distortions. DPD is used to compensate for odd and/or even order inter-modulation (IM) distortion [23], therefore it is critical to apply DPD compensation techniques to get rid of nonlinearity effects.

2.3 Modeling of Power Amplifier

For memory based models on the modelling of PA response, Volterra series based models are most often used because of high accuracy. There are several derivatives of Volterra series where MP and GMP are used as baseline models for performance evaluation.

2.3.1 Volterra Series

Volterra series is a combination of linear convolution and nonlinear power series, which provides a general way to model a nonlinear system with memory. Therefore it is used to describe the relationship between input and output of the PA with memory. The Volterra series is represented in discrete time domain as [8].

$$y(n) = \sum_{p=1}^{P} \sum_{i_1=0}^{M} ... \sum_{i_p=0}^{M} h_p(i_1, ..., i_p) \prod_{j=1}^{p} x(n - i_j), \tag{2.1}$$

where $x(n)$ and $y(n)$ denote input and output signals, respectively, $h_p(i_1, ..., i_p)$ denotes the pth order Volterra Kernel, P denotes the nonlinear order, and M denotes the memory length.

Because Volterra series is linear in the parameters, least squares (LS) can be used for parameters estimation. The parameters are computed by minimizing the sum of squares between observed and computed data. If there are N data samples, observed data $y(n)$ and estimated data $\hat{y}(n)$ than Q representing permutation of all estimated parameters $q(n)$ can be estimated by minimizing the following

$$Q = \sum_{n=0}^{N-1} \left| y(n) - \hat{y}(n) \right|^2, \tag{2.2}$$

Let $\boldsymbol{w}$ represent the ... all parameters h_p, and $\boldsymbol{X}$ represents the permutation of the input signal $x(n)$, the vector form of (2.1) can be written as

$$\hat{y} = \boldsymbol{X}\boldsymbol{w}, \tag{2.3}$$

where $\hat{y}$ denotes the estimated output signal. The solution for LS is formulated as [11]

$$\hat{w} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}y, \tag{2.4}$$

where T denote a conjugate transpose.

2.3.2 Memory Polynomial

MP is a simple form of Volterra series which is commonly used [9]. It only keep the diagonal coefficients of the Volterra series. The input-output relation of MP is formulated as

$$y_{\mathrm{MP}}(n) = \sum_{p=1}^{P} \sum_{m=0}^{M} a_{pm} x(n-m) \left| x_{in}(n-m) \right|^{p-1} \tag{2.5}$$

where a_{mp} denotes the polynomial coefficient.

2.3.3 Generalized Memory Polynomial

The GMP combines MP with cross terms between signal and lagging and/or leading exponentiated envelop terms [10], which can be expressed as

$$\begin{aligned} y_{\mathrm{GMP}}(n) = & \sum_{p=0}^{P-1} \sum_{m=0}^{M} \alpha_{p,m} x[n-m]|x[n-m]|^p \\ & + \sum_{p=0}^{P-1} \sum_{m=0}^{M} \sum_{g=1}^{G} (\beta_{p,m,g} x[n-m]|x[n-m-g]|^p \\ & + \gamma_{p,m,g} x[n-m]|x[n-m+g]|^p) \end{aligned} \tag{2.6}$$

where G denote the cross-term lengh of the lagging and leading envelope terms and $\alpha_{p,m}$, $\beta_{p,m,g}$ and $\gamma_{p,m,g}$ are the model coefficients [5].

2.3.4 Look-up Table

In LUT procedure the magnitude of input signals and corresponding output are stored in a table. In this method the input value is matched to the appropriate entry among the discrete table entries of the LUT [30].

2.4 Modeling of DPD

This section is added because DPD modelling is similar to PA modeling and a possible topic for the future studies based upon the findings of this thesis. Similar to the PA modeling, DPD is modeled using MP, GMP and LUT or ML techniques. The idealistic way to get DPD output is to use the iterative learning control (ILC) as explained in this section.

2.4.1 Iterative Learning Control

ILC [11] is used to track the performance and improve the transient response of a system that operates repeatedly. It is based upon the observation that for the same operating conditions of the system the errors observed in the output response will be repeated. The errors can then be used as feedback to refine the input so that errors are reduced during the next time the system is operated [11]. Mathematically, the linear ILC is represented as

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \lambda \mathbf{e}_k \tag{2.7}$$

where $\mathbf{e}_k$ is the error between actual and desired output and $\mathbf{u}_k$ is the optimal input at kth iteration. λ is the learning gain and the learning algorithm will converge as long as the λ satisfies the following condition

$$0 < \lambda < \frac{2}{J_{\max}}, \tag{2.8}$$

where $J_{\max}$ represents the linear gain of the PA which can be calculated or taken from the PA data-sheet [11]. It is noted that DPD modeling is not in the scope of this thesis, similar performance is expected for DPD modelling with MP, GMP and LUT or ML algorithms as given in [5].

3

Power Amplifier Modelling Using Machine Learning Algorithms

3.1 Basics of Machine Learning

ML is the name given to a set of techniques that allows the implementation of adaptive algorithms to make predictions and automatically organize the input data according to their common features. The goal of ML is to study, engineer, and improve mathematical models and make decisions without the complete knowledge of all external factors [12]. There are three common approaches of ML which are called supervised learning, unsupervised learning and reinforcement learning [12].

3.1.1 Supervised learning

In supervised learning algorithms, the algorithm has a set of training data containing input and output samples, and during the training process, the algorithm updates its parameters in order to minimize the global loss. The goal of the system is to also work with the samples that have never been seen before, therefore it is necessary to avoid a common problem called over-fitting which causes over-learning due to excessive capacity [12].

3.1.2 Unsupervised learning

As the name suggests in this type of learning, there are no supervisors and this type of approach is normally used in clustering. For example, if there a unlabelled set of data and the goal is to classify that data into human and non-human clusters. Common applications for unsupervised learning are object segmentation, similarity detection and automatic labeling [12].

3.1.3 Reinforcement learning

This type of learning is based on feedback provided by the environment like the supervised learning however the information is more qualitative and does not help in determining the precise measure of the error, this feedback is usually called reward and is only used to determine whether a certain action performed in a state is positive or not, it is efficient when it is impossible to have a precise error measure [12].

3.2 Classification and Regression

The goal for the classification is to predict a category, the model is trained with data containing labels of different categories and then the model is used to label data. In regression the goal is to predict a value so if there are one or many predictor variables then the model should be able to predict continuous output, so if the predictors are defined as X and output as Y then Y should be defined as a function of X. It is noted that PA and/or DPD modeling is a regression problem and in this thesis regression techniques are considered for ML approaches.

3.3 Specific ML algorithms used in PA modelling for this thesis

3.3.1 Neural Networks

A NN is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way that the human brain operates. In this sense NNs refer to systems of neurons, either organic or artificial in nature. NNs can adapt to changing input so the network generates the best possible result without needing to redesign the output criteria [1]. A typical NN consists of input layer, a number of hidden layers and output layer, each containing different number of neurons as shown in Figure 3.1.

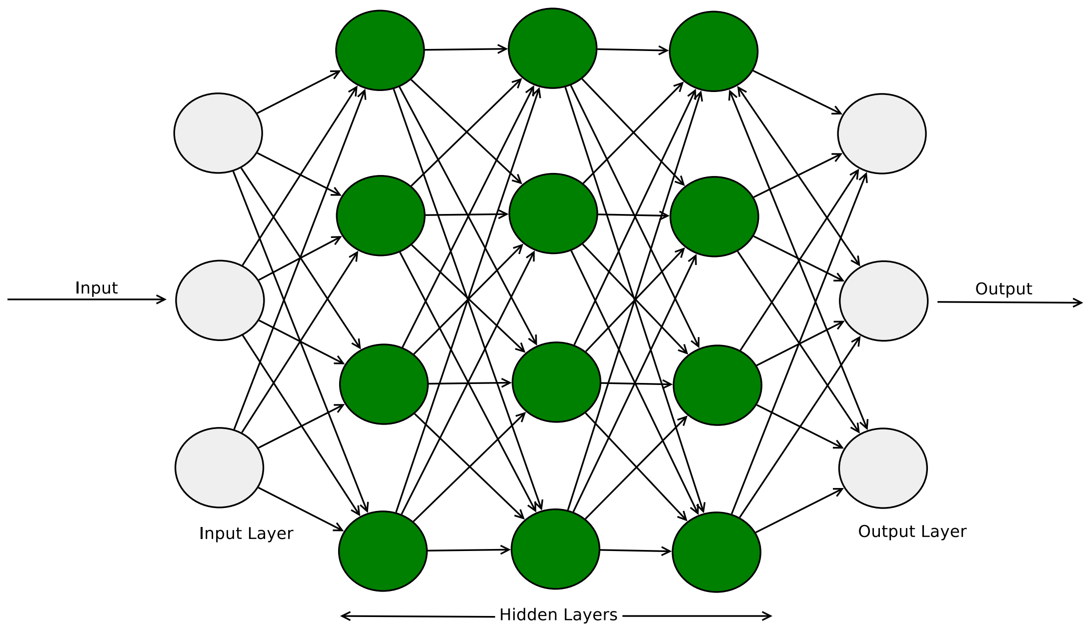

Figure 3.1: A three hidden layer fully connected neural network.

3.3.2 Gradient Boosting

GB constructs additive regression models by sequentially fitting a simple parameters function(base learner) to current "pseudo"-residuals by least-squares at each iteration [3]. The most common algorithm for regression in GB is least square boost LSboost (LSB) [2], other algorithms are mostly for classification. The process of creating trees during each iteration as well as the expected relationship between number of iterations and error is illustrated in Figure 3.2. The LSB algorithm was invented by Friedman [2] for GB. The algorithm works by keeping the sum of the square of the ensembles to a minimum. It does that by creating a new tree at each iteration as shown in Figure 3.2 and use the base learner and the learn rate to keep the error between the actual and predicted output to a minimum. The probability of over-fitting increases with the increase in the number of iterations[27], therefore parameters like learn-rate and number of trees should be tuned properly to avoid the over-fitting problem.

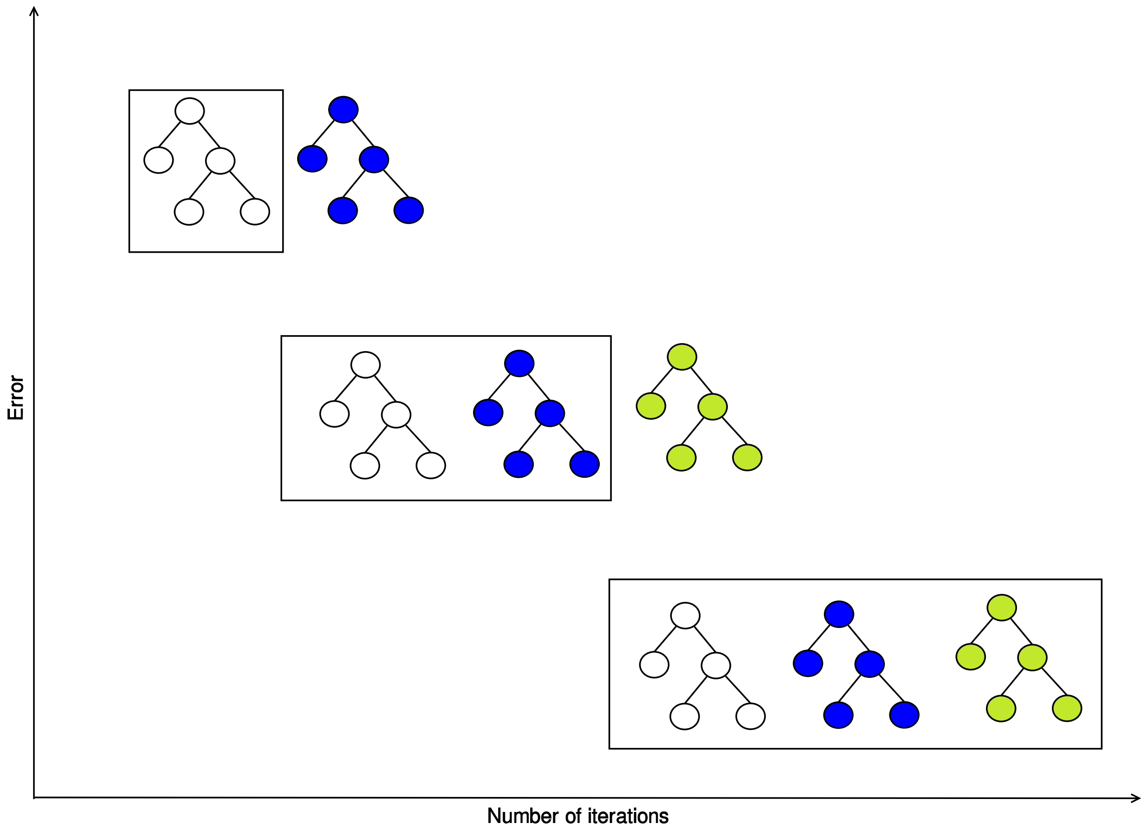

Figure 3.2: Gradient boost design.

3.3.3 Decision Tree

A DT contains a root node, branch nodes, and leaf nodes just like a normal tree. The root node is the parent of all nodes and the topmost node in the tree. In a DT, each node shows a feature, each branch shows a decision and each leaf shows an outcome which can be categorization or continuous [13]. The mean square error (MSE) is used to split the node into sub-nodes. A rough illustration of a DT creation is given in figure 3.3 where the root node represents the starting point and the leaf nodes represent the possible outcomes and MSE is used to iterate over the tree to reach the correct leaf node or the outcome.

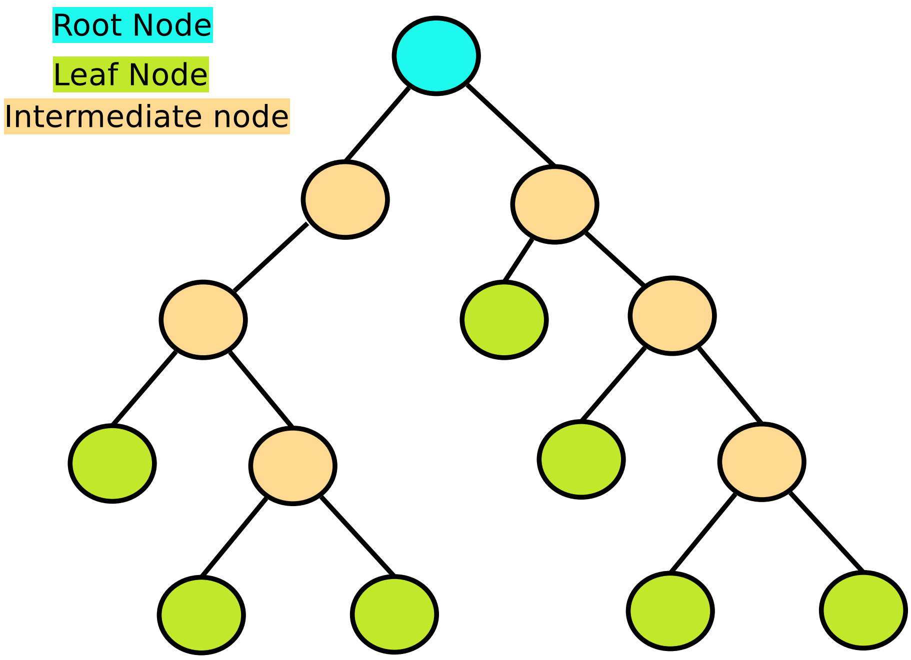

Figure 3.3: An example of a decision tree.

4 Proposed Machine Learning Techniques for PA modeling

4.1 Background

During one of the previous thesis work the PA for the single band was modeled in [5], the tool used to apply machine learning in that thesis work is based on Python. In the first phase of this thesis work, the goal is to replicate the results collected from the previous thesis work while using MATLAB for PA modeling using ML algorithms. MATLAB is selected for this study because it provides powerful tools for signal processing and it is simpler to use one tool instead of using a separate tool for signal processing and training/testing using ML algorithms.

Additionally, the PA modeling of a single-carrier or single-band is considered as the reference case for PA modeling of multi-carrier and multi-band signals. It is decided to use different ML models for PA modelling and compare the results to those collected from PA modeling using traditional approaches such as MP and GMP.

4.2 PA Modeling using ML for single carrier

This section presents the results collected from a 60 MHz single band next radio (NR) signal. In this scenario, PA is modelled for a single carrier with a 60 MHz bandwidth shown in Figure 4.1. PA modeling with ML for single carrier is a reference case for the two other scenarios for full-band multi-carrier and separated band multi-carrier scenarios.

4.2.1 PA modeling with ML for single-carrier single-band

In this section the results for single band PA modelling are presented, a signal with 60 MHz bandwidth is used as illustrated in Figure 4.1. These results serve as the reference results for our studies on multi-carrier and multi-band applications. For data processing, different training and test data-sets are applied with the number of samples $N = 32768$ and a sampling rate of $f_s = 983.04$ MHz. The goal of this process is to capture the uncorrelated, two-piece data-set. There is no correlation between the training and test data which results in the performance of the PA modeling being accurate and better performance compared with the base-line MP and GMP algorithm results.

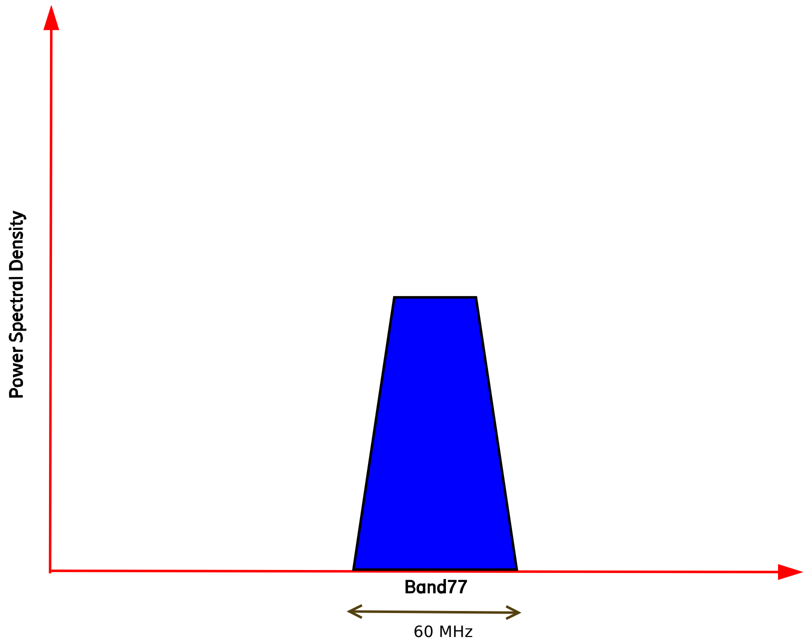

Figure 4.1: Block diagram of 60 MHz single-carrier signal.

4.3 Full-band PA modeling with ML on multi-carrier and multi-band

In this section, the second test scenario of this study is described. In this scenario, ML algorithms are used to model PA for multi-carrier signals captured from our lab test-bed. The captured signals are then used for training and testing of different ML models. On the PA test-bed, data from two different types of PAs with different instantaneous bandwidths (IBWs) is captured.

For the first PA, a multi-carrier signal with 100 MHz bandwidth of each carrier and the total IBW of 400 MHz is used as illustrated in Figure 4.2.

For the second case a multi-carrier signal with 100 MHz bandwidth for each carrier with a 600 MHz IBW (wider linearization bandwidth) is used as illustrated in Figure 4.3. Similar to the single band PA modelling, uncorrelated data-sets are used for training and testing of ML models. Each data set consists of 49152 samples and a sampling rate of $f_s = 983.04$ MHz is used.

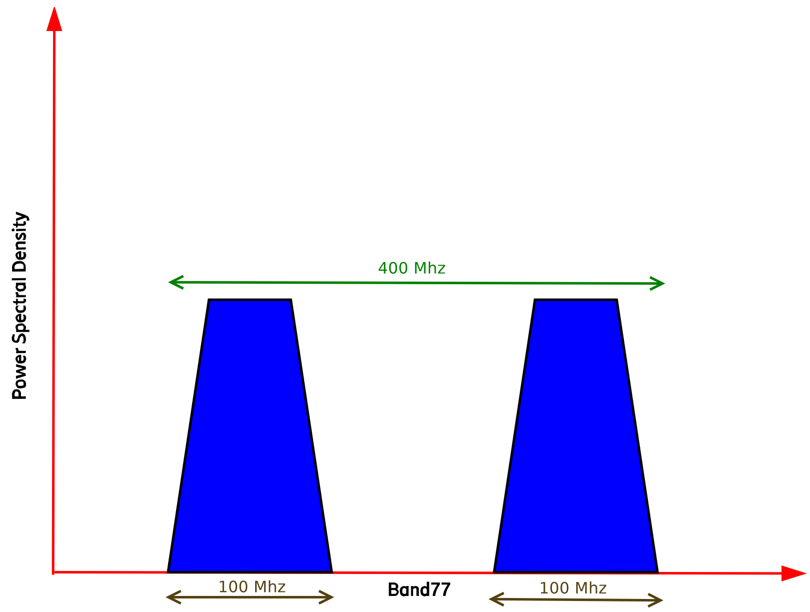

Figure 4.2: Block diagram of 100 MHz multi-carrier signal with 400 MHz IBW.

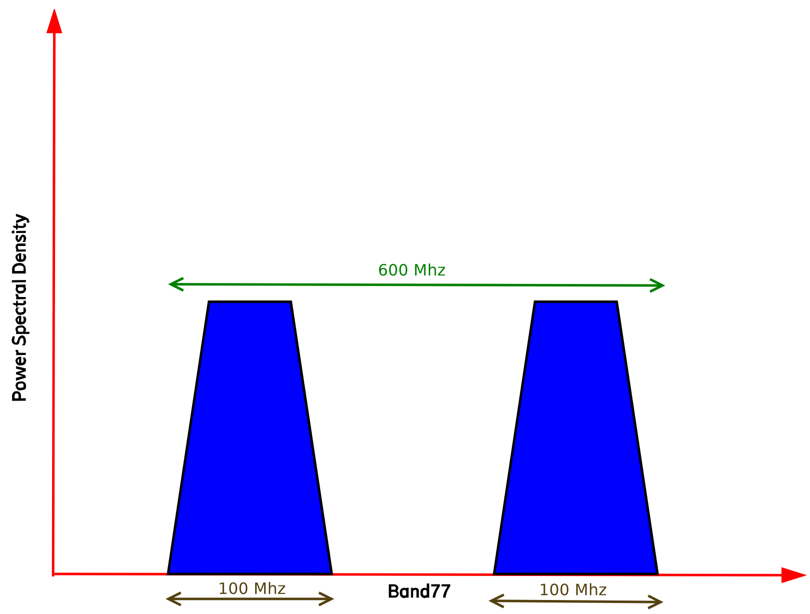

Figure 4.3: Block diagram of 100 MHz multi-carrier signal with 600 MHz IBW.

4.4 Separated PA modeling with ML on multi-carrier and multi-band

The final scenario is to filter out low and high band signals from the multi-carrier signal. The filtered signals are then used separately to train and test different ML models. Finally, the performance is compared with the full-band method which is described in the previous section.

This test requires some pre-processing of multi-carrier input and PA output signal, and it requires three steps. The initial multi-carrier/band signal is illustrated in Figure 4.4.

In the first step, low and high pass filters are applied to separate the low and high band/frequency signal.

Finally both low band (LB) and high band (HB) signals are centered and filtered one more time in order to remove any noise. Finally after centring the signal, the output can be visualized from Figure 4.5.

Similar to the "Full Band" modelling this modelling is done for signals with 600 MHz and 400 MHz IBWs. For each signal LB and HB signals are modelled separately, for all cases independent signals are used for training and testing the models as shown in Figure 4.5.

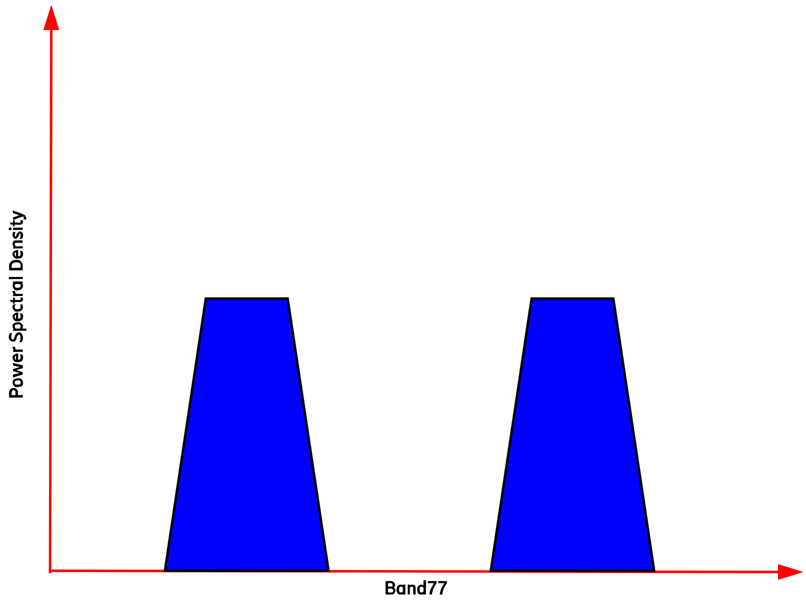

Figure 4.4: Block diagram of our multi-band signal.

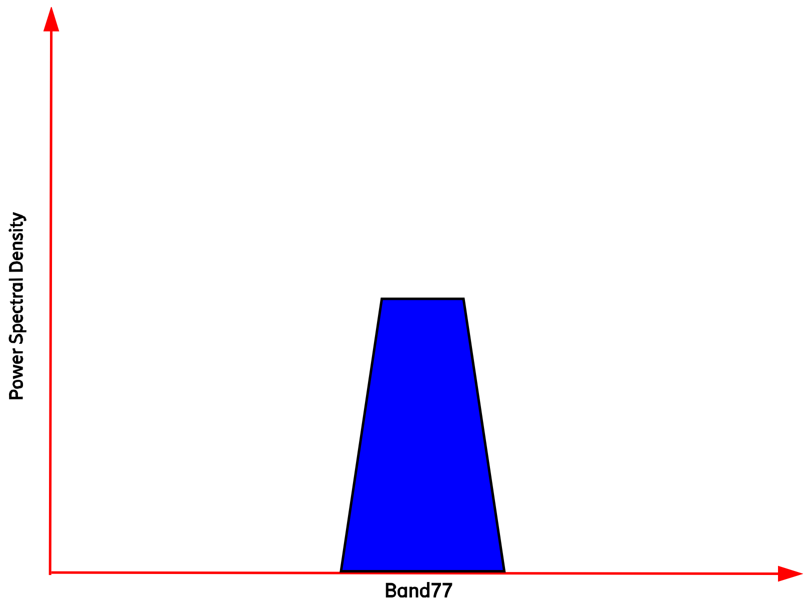

Figure 4.5: Block diagram of centered and filtered HB/LB signals.

4.5 NMSE and ACEPR Calculation and Performances

For evaluation of performance, NMSE and ACEPR are used as the performance matrices in this thesis.

4.5.1 NMSE

Let $y(n)$ denote the input signal and $\hat{y}(n)$ denote the modeled output. The NMSE is defined as [11].

$$\mathrm{NMSE} = 10\log_{10}\frac{\sum_{n=0}^{N-1}\left|y(n)-\hat{y}(n)\right|^2}{\sum_{n=0}^{N-1}\left|y(n)\right|^2}. \tag{4.1}$$

NMSE can be used to evaluate in-band performance since it is dominated by in-band error [11].

4.5.2 ACEPR

ACEPR is used to evaluate out-of-band modelling performance and is defined as a ratio between the error signal power over the adjacent channel and the desired channel power of the measured signal [4]. The ACEPR is defined as [16].

$$\mathrm{ACEPR} = \max_{\mathrm{m}=1,2}\left[\frac{\int_{(\mathrm{adj})_\mathrm{m}}\left|Y_{\mathrm{meas}}(f)-Y_{\mathrm{mod}}(f)\right|^2}{\int_{\mathrm{ch.}}\left|Y_{\mathrm{meas}}(f)\right|^2}\right], \tag{4.2}$$

where $Y_{\mathrm{mod}}(f)$ and $Y_{\mathrm{meas}}(f)$ are the Fourier transform of the modelled and measured signals, respectively. The integration over numerator is over the adjacent channels to the signal channel with the same bandwidth and the integration over denominator is over the in-band channel signal bandwidth. ACEPR is defined as the larger evaluated value for lower ($m = 1$) and upper ($m = 2$) adjacent channels [16].

5

Experimental Results

5.1 Data acquisition

The data used in this thesis is captured from Ericsson's radio lab testbed with different types of PAs. A total of three test scenarios are considered for this thesis, for the first scenario, single-carrier with 60 MHz carrier bandwidth is captured. For the second and third test scenarios, multi-carrier data (low and high carriers) signals with 400 MHz and 600 MHz IBWs and 100 MHz carrier bandwidth is captured.

A block diagram of the lab testbed is shown in Figure 5.1. It consists of a signal generator which in this study is a MATLAB based tool, spectrum analyzer developed by external vendors and PA designed by Ericsson. MATLAB is also used to run algorithms for PA modeling and for doing data pre and post-processing when required.

MP and GMP for odd and all orders are denoted by MPO/GMPO and MPA/GMPA. For odd-order in equations 2.5 and 2.6, P takes odd values only and it is reduced to $\frac{P+1}{2}$ [31]. All the test results in the tables are presented in dBs.

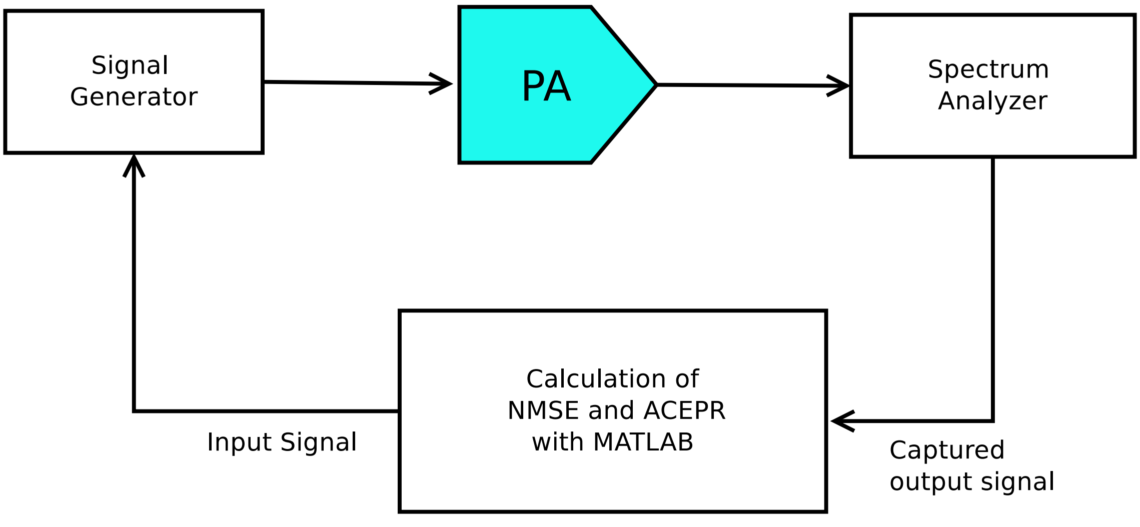

Figure 5.1: Block diagram of process of capturing data

5.2 Measurement Setup

The captured data is then used to model PA using legacy algorithms like MP and GMP and also with ML algorithms like NN, GB, DT and LR. A block diagram of measurement method is illustrated in Figure 5.2.

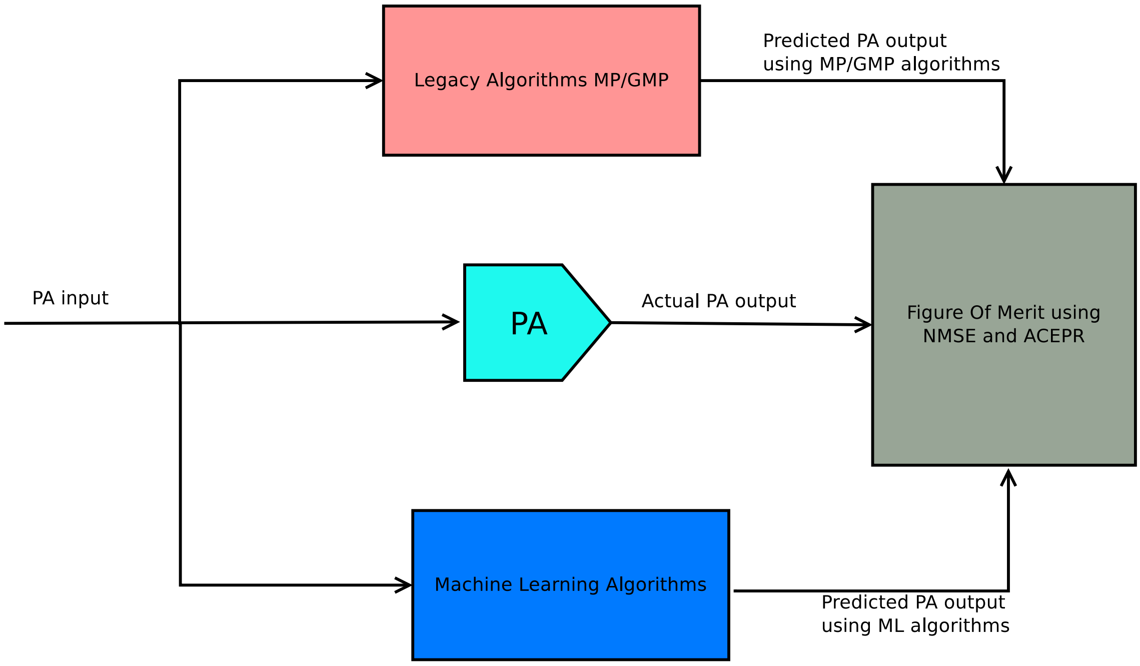

Figure 5.2: Block diagram of our performance evaluation methodology

5.3 Performance of single carrier/single band PA modeling

In this scenario, the performance of NN is better in terms of NMSE than the performance of the best case results for baseline algorithms like MP/GMP which is with GMP with $P = 9$. In terms of ACEPR, the performance of GMPA is better than all the other algorithms. In this section a single carrier with 60 MHz bandwidth is used as shown in Figure 4.1. The results are summarized in Table 5.1.

5.3.1 Results of NN, GB, DT and LR

For NN three hidden layers with 40 neurons in each hidden layer is used and it gives the best PA modeling performance in terms of NMSE with a value of -33.31 dB and in terms of ACEPR the value is -40.91 dB which is lower than the modeling with GMP, but better than all other ML algorithms.

GB gives the second best performance for ML algorithms in terms of NMSE and ACEPR. Different algorithms for boosted and bagged trees are used and the best performance is observed when LSB algorithm with leafsize 30 and a learning rate 0.1.

The performance of DT is worse compared to NN and GB, however from the Figure 5.3, it is observed that the predicted output matches the actual output.

LR is designed for linear systems and does not work well for PA modelling which is a non-linear system therefore the PSD plot in Figure 5.4 shows that the predicted signal looks similar to the input signal.

Table 5.1: Result for single band 60 MHz IBW.

Algorithm	NMSE	ACEPR
NN	$\mathbf{-33.31}$	-40.91
GB	-24.00	-35.88
DT	-22.13	-33.66
LR	-16.77	-24.46
MPO P=7	-28.04	-37.01
GMPO P=7	-31.02	-38.17
MPA P=7	-28.83	-38.72
GMPA P=7	-33.13	-41.27
MPO P=9	-28.20	-37.34
GMPO P=9	-31.35	-38.63
MPA P=9	-28.88	-38.81
GMPA P=9	-33.23	$\mathbf{-41.35}$

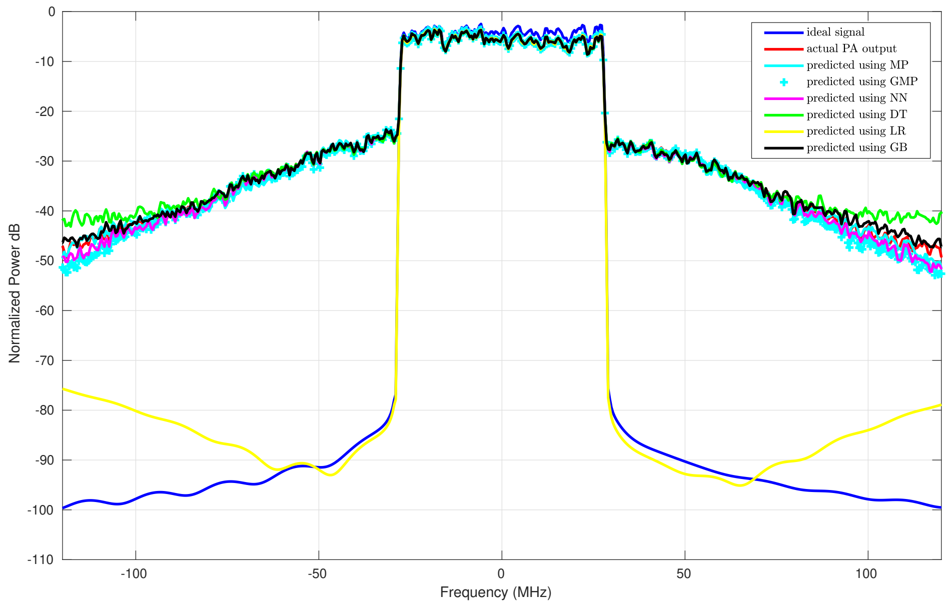

Figure 5.3: Normalized PSD of the ideal, measured, and predicted PA output on a 60 MHz NRs signal.

5.4 Performance of full band multi-carrier PA modeling

In this section the results from the modelling of PAs using multi-carrier signals are discussed.

5.4.1 Results on 400 MHz IBW

The captured results presented in this section are from the PA with multi-carrier signal with each carrier with a 100 MHz frequency and a total IBW of 400 MHz as shown in Figure 5.4, the results are summarized in Table 5.2.

5.4.1.1 Results of NN, GB, DT and LR

NN gives the best results in terms of NMSE and ACEPR. The NMSE advantage is around 1.8 dB over of the GMP, and similarly, the ACEPR advantage is around 5.7 dB over GMP and MP results. For NN Levenberg-Marquardt algorithm is used. Similar to the single-band scenario, the best results are obtained with three hidden layers containing forty neurons each.

The GB gives lower performance in terms of NMSE and ACEPR than NN algorithms and almost all GMP scenarios. Unlike NN which can handle multi-row output data, GB requires two models to train real and imaginary signals separately and LSB algorithm is used for training. After parameter tuning, the best performance is observed with a MinLeafSize 10, a LearningCycles 10000 and a LearnRate 0.2.

The GB gives lower performance in terms of NMSE and ACEPR than NN algorithms and almost all GMP scenarios. Unlike NN which can handle multi-row output data, GB requires two models to train real and imaginary signals separately and LSB algorithm is used for training. After parameter tuning, the best performance is observed with a MinLeafSize 10, a LearningCycles 10000 which were fine tuned to 402 and a LearnRate 0.2.

LR is designed for linear systems and does not work well for PA modelling which is a non-linear system therefore the PSD plot in Figure 5.4 shows that the predicted signal looks similar to the input signal.

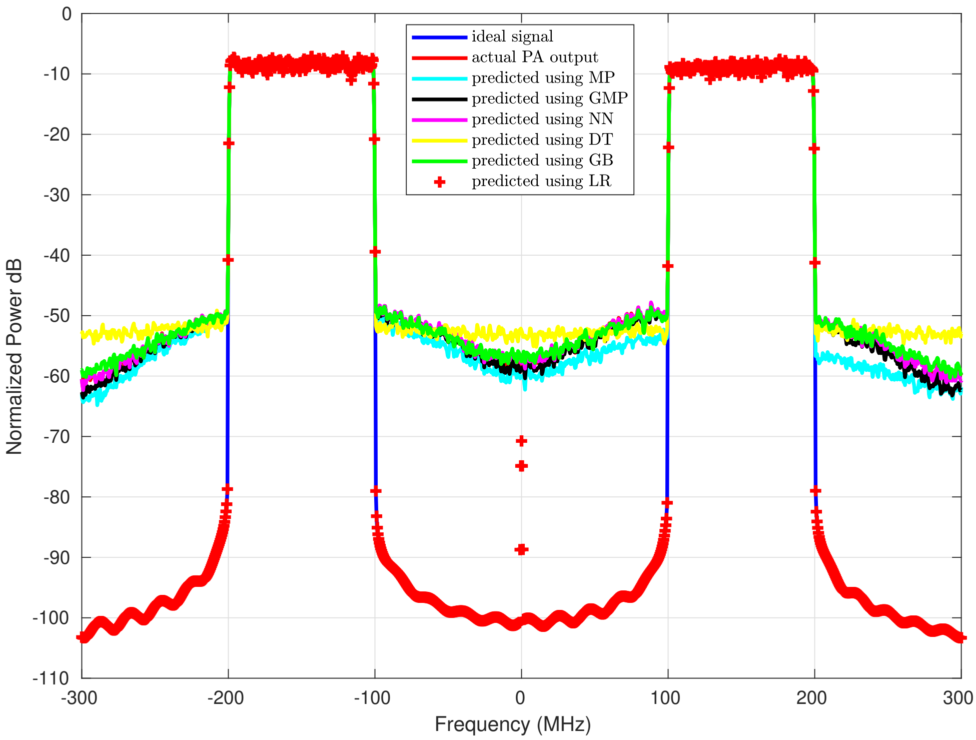

Figure 5.4: Normalized PSD of the ideal, measured, and predicted PA output on 100 MHz multi-carrier signal with 400 MHz IBW.

Table 5.2: Results on 400 MHz IBW.

Algorithm	NMSE	ACEPRLB	ACEPRHB	ACEPRAVG
NN	**−45.99**	**−57.75**	**−59.81**	**−58.78**
GB	−42.93	-51.12	-53.01	-52.06
DT	-33.48	-41.52	-42.77	-42.14
LR	-36.97	-43.92	-35.91	-39.92
MPO P=7	-39.64	-46.32	-49.53	-47.92
GMPO P=7	-43.13	-50.44	-51.95	-51.19
MPA P=7	-39.96	-46.95	-50.39	-48.67
GMPA P=7	-44.17	-52.44	-53.63	-53.08
MPO P=9	-39.90	-46.84	-50.28	-48.56
GMPO P=9	-43.69	-51.55	-53.24	-52.40
MPA P=9	-39.97	-46.96	-50.49	-48.72
GMPA P=9	-44.19	-52.48	-53.67	-53.07

5.4.2 Results on 600 MHz IBW

The captured results presented in this section are from the PA with multi-carrier signal with each carrier with a 100 MHz frequency and a total IBW of 600 MHz as shown in Figure 5.5, the results are summarized in Table 5.3.

The performance of MP/GMP as well as ML algorithms decreases slightly compared to 400 MHz IBW scenario, but for some ML algorithms the relative performance compared to MP and GMP is better than the relative performance for the 400 MHz IBW scenario.

5.4.2.1 Results of NN, GB, DT and LR

Like in the previous 400 MHz IBW scenario, NN gives the best results in terms of NMSE and ACEPR without doing any kind of post-processing on the predicted signal. The NMSE value is around 6.6 dB better than the best-case scenario for GMP and similarly, the ACEPR value is around 7.7 dB better than GMP/MP best-case scenario results.

Similar to the results obtained from the 400 MHz IBW scenario, GB gives the second best performance in terms of NMSE and ACEPR, like in the previous case MSE algorithm is used for training models.

DT does not give good performance however after plotting the predicted signal PSD, the plot looks similar to the actual signal as seen in Table 5.3.

LR gives better results in terms of NMSE and ACEPR than the DT and although LR can predict non-linear spectral regrowth however the algorithm is designed for linear systems and does not work well for PA modeling which is a non-linear system therefore the PSD plot in Figure 5.5 shows that the predicted signal looks similar to the input signal, therefore the results of LR is not as accurate as of the other ML algorithms.

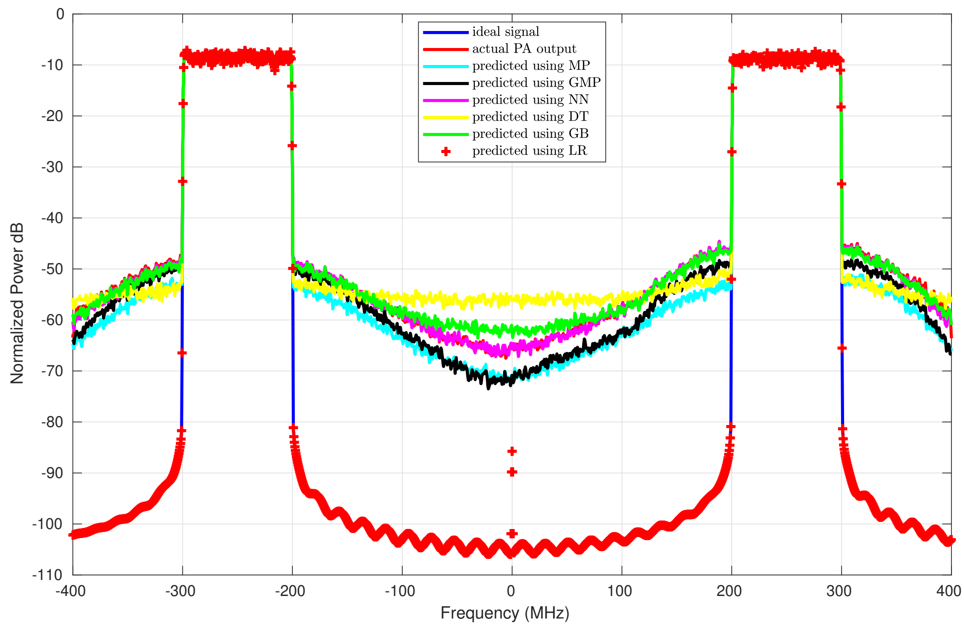

Figure 5.5: Normalized PSD of the ideal, measured, and predicted PA output on 100 MHz multi-carrier signal with 600 MHz IBW.

Table 5.3: Results on 600 MHz IBW.

Algorithm	NMSE	ACEPRLB	ACEPRHB	ACEPRAVG
NN	**−44.70**	**−51.84**	**−53.60**	**−52.72**
GB	-42.71	-49.58	-50.32	-49.95
DT	-34.11	-40.80	-42.77	-41.79
LR	-34.63	-41.01	-43.64	-42.33
MPO P=7	-35.79	-42.23	-45.95	-44.09
GMPO P=7	-37.85	-43.91	-47.71	-45.81
MPA P=7	-35.87	-42.34	-45.99	-44.16
GMPA P=7	-38.08	-44.10	-48.49	-46.29
MPO P=9	-35.86	-42.33	-45.99	-44.16
GMPO P=9	-38.08	-44.10	-47.89	-46.00
MPA P=9	-35.88	-42.345	-45.99	-44.17
GMPA P=9	-38.08	-44.10	-48.50	-46.30

5.5 Performance of separated multi-carrier PA modeling

In this scenario the high and low frequency signals are filtered out from the multi-carrier signal and are used to train and test different ML models. The results are then compared to the results obtained from full band PA modeling.

One important thing to mention here is that in this study the multi-carrier signal on n77 [25] single-band case is considered for the experimental results. However, this study can be extended for multi-carrier and multi-band scenarios considering the same method. For instance, B3-B1 [26] on mid-band together with multi-carrier signals can be considered and it is expected that the proposed method without any processing change gives a good performance.

The performance matrices are the same as in the previous two scenarios, however since the signals are separated in low and high frequency carriers therefore in Tables 5.4 and 5.5 have separate columns for NMSE and ACEPR for HB, LB and the average, denoted by NMSEHB, NMSELB, NMSEAVG, ACEPRHB, ACEPRLB and ACEPRAVG.

Similar to the previous scenarios, LR is designed for linear systems and does not work well for PA modeling which is a non-linear system therefore the PSD plot in Figures 5.6 and 5.7 shows that the predicted signal looks similar to the input signal, therefore the results of LR are not as accurate as of the other ML algorithms.

5.5.1 Results for ML algorithms for 400 MHz IBW

The captured results presented in this section are from the PA with multi-carrier signal with each carrier with a 100 MHz frequency. The high and low frequency signals are then filtered out as shown in Figures 5.6 and 5.7.

The detailed results are explained in the next subsections and are summarized in Table 5.4. Unlike the results obtained during PA modelling for single band and full band scenarios, in this scenario the GB gives the best results in terms of NMSE and ACEPR.

5.5.1.1 Results of NN, GB, DT and LR

In this scenario NN does not generate the best results in terms of NMSE and ACEPR. Comparing to the best case for MP and GMP, the NMSE and ACEPR values are similar as shown in Table 5.4.

The GB unexpectedly gives the best performance in terms of NMSE and ACEPR. Like in previous case MSE algorithm is used for training the models. Comparing to the best case for MP and GMP the NMSE value is around 10.44 dB better and similarly the ACEPR value is around 9.4 dB better. This result is an exception and this kind of performance is not replicated in any other scenario even with parameter tuning.

DT gives poor performance however the performance difference is much smaller than in the previous scenarios. The plot of PSD of predicted signal shows that it is similar to the PSD of actual signal.

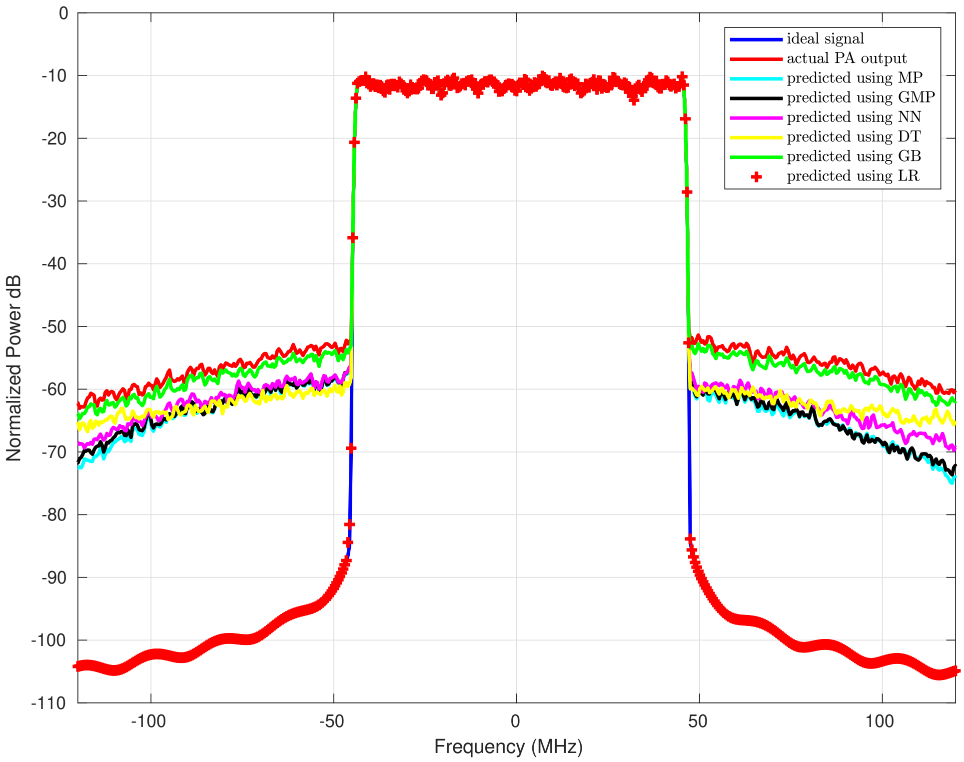

Figure 5.6: Normalized PSD of the ideal, measured, and predicted PA output on 100 MHz centered multi-carrier HB signal with 400 MHz IBW.

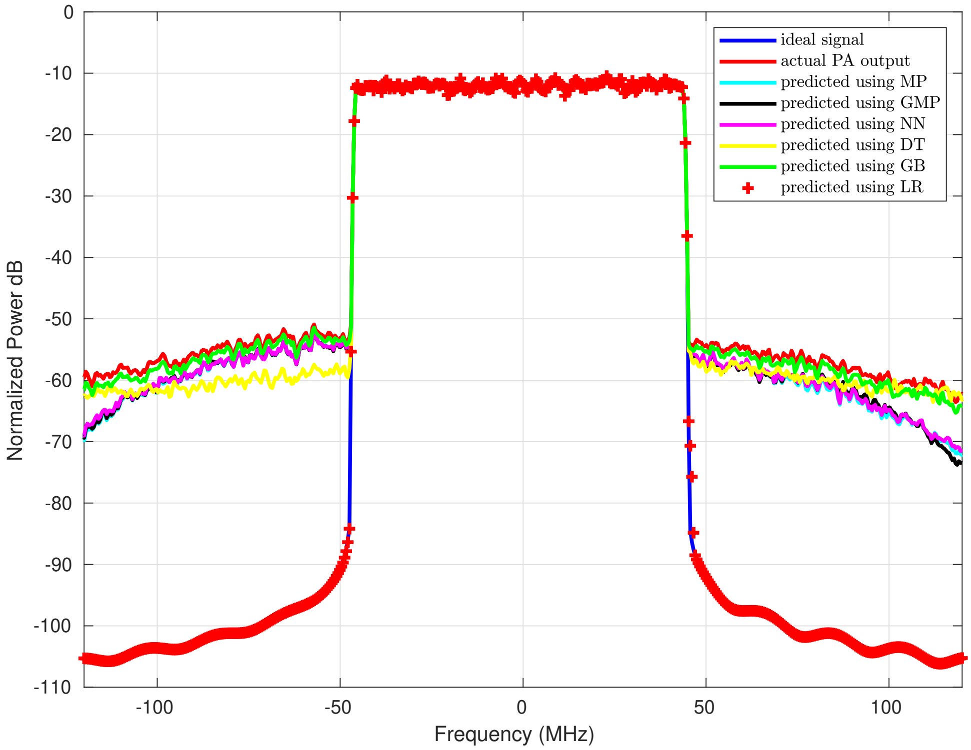

Figure 5.7: Normalized PSD of the ideal, measured, and predicted PA output on 100 MHz centered multi-carrier LB signal with 400 MHz IBW.

Table 5.4: Results 400 MHz IBW for separate band method.

Algorithm	NMSEHB	NMSELB	NMSEAVG	ACEPRHB	ACEPRLB	ACEPRAVG
NN	-38.75	-41.63	-40.19	-45.56	-47.34	-46.45
GB	**−49.12**	**−50.90**	**−50.01**	**−56.36**	**−57.48**	**−56.98**
DT	-37.01	-36.39	-36.70	-45.33	-44.33	-44.83
LR	-37.65	-37.37	-37.51	-44.96	-43.92	-44.44
MPO P=7	-38.59	-41.47	-40.03	-45.49	-47.30	-46.39
GMPO P=7	-38.66	-41.55	-40.11	-45.50	-47.30	-46.40
MPA P=7	-38.60	-41.48	-40.04	-45.49	-47.30	-46.40
GMPA P=7	-38.68	-41.57	-40.13	-45.50	-47.32	-46.41
MPO P=9	-38.60	-41.48	-40.04	-45.49	-47.30	-46.39
GMPO P=9	-38.67	-41.55	-40.11	-45.50	-47.30	-46.40
MPA P=9	-38.64	-41.48	-40.06	-45.49	-41.31	-43.40
GMPA P=9	-38.68	-41.57	-40.13	-45.51	-41.32	-43.41

5.5.2 Results for ML algorithms for 600 MHz IBW

The captured results presented in this section are from the PA with multi-carrier signal with each carrier with a 100 MHz frequency. The results of traditional algorithms like MP and GMP are generally better in most of the cases with a couple of exceptions. The results are summarized in Table 5.5 and the results for ML algorithms are the worst obtained during the course of this thesis.

5.5.2.1 Results of NN, GB, DT and LR

NN does not give better performance than GMP/MP in terms of NMSE and ACEPR. For MP and GMP, the NMSE value is around 0.57 dB and ACEPR value is around 0.47 dB better for the best case scenario compared to the results from NN.

The GB also gives worse performance than MP and GMP in terms of NMSE and ACEPR. Like in previous case MSE algorithm is used for training models. For MP and GMP, the NMSE value is around 2.7 dB and ACEPR value is around 1.6 dB better for the best case scenario compared to the results from GB.

DT gives poor performance however the difference is much narrower than in previous cases and PSDs of predicted signals are similar to the actual signal as seen in Figures 5.8 and 5.9.

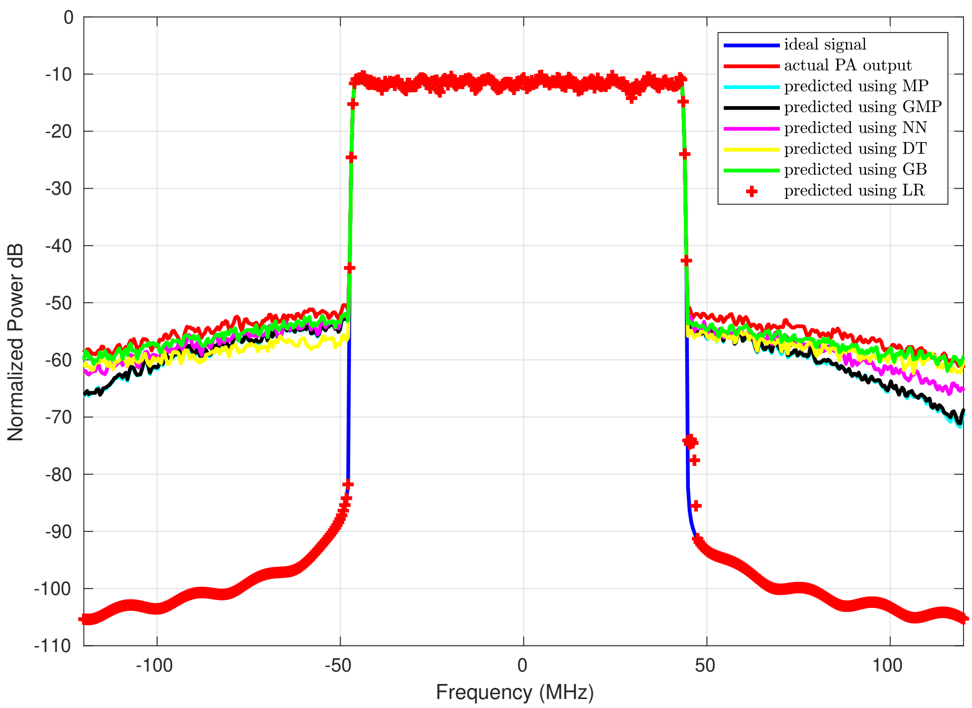

Figure 5.8: Normalized PSD of the ideal, measured, and predicted PA output on 100 MHz centered multi-carrier HB signal with 600 MHz IBW.

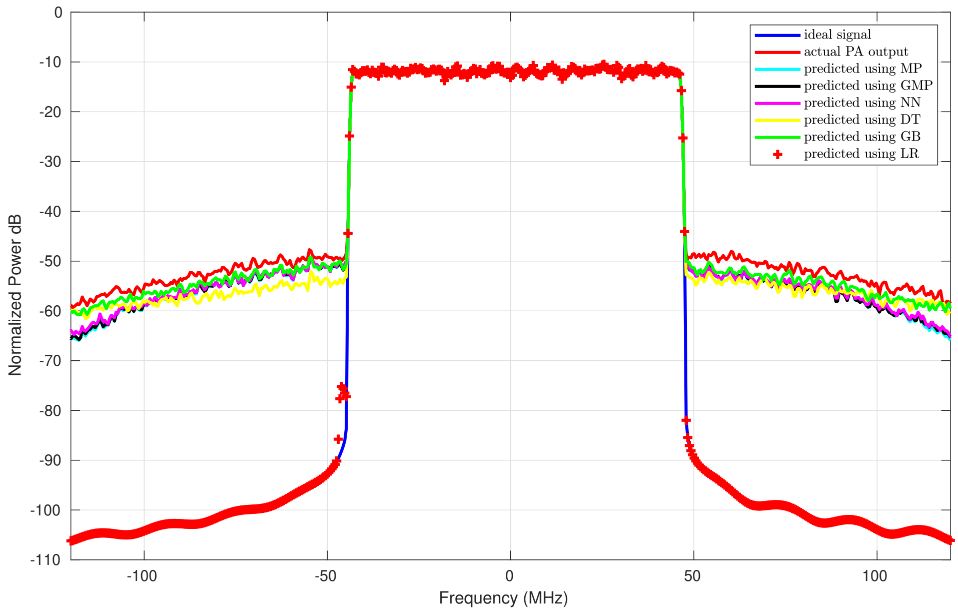

Figure 5.9: Normalized PSD of the ideal, measured, and predicted PA output on 100 MHz centered multi-carrier LB signal with 600 MHz IBW.

Table 5.5: Results 600 MHz IBW for separate band method.

Algorithm	NMSEHB	NMSELB	NMSEAVG	ACEPRHB	ACEPRLB	ACEPRAVG
NN	-39.38	-36.31	-37.84	-45.89	-43.03	-44.46
GB	-37.26	-35.19	-36.22	-44.08	-42.11	-43.09
DT	-35.91	-33.92	-34.92	-43.56	-41.62	-42.59
LR	-36.23	-33.30	-34.76	-43.64	-41.01	-42.33
MPO P=7	-39.79	-36.67	-38.23	-46.28	-43.36	-44.82
GMPO P=7	-39.87	-36.75	-38.31	-46.30	-43.38	-44.84
MPA P=7	-39.86	-36.69	-38.27	-46.38	-43.40	-44.89
GMPA P=7	-39.95	-36.78	-38.37	-46.41	-43.44	-44.92
MPO P=9	-39.85	-36.68	-38.27	-46.36	-43.39	-44.88
GMPO P=9	-39.93	-36.77	-38.35	-46.38	-43.42	-44.90
MPA P=9	-39.86	-36.69	-38.28	-46.38	-43.40	-44.89
GMPA P=9	**−39.95**	**−36.78**	**−38.37**	**−46.41**	**−43.44**	**−44.92**

5.6 Computational Complexity Analysis

5.6.1 Computational Complexity for MP and GMP algorithms

We consider the running complexity which defines as the number of calculations needed for each output sample of a model. Specifically, we use the number of floating point operations (FLOPs) for each operation such as multiplication and summation, according to [16, Table I]. To implement MP, we need two steps: 1) construct basis functions X and 2) filter the basis with all coefficients a_{ij}. We need to form each basis function $x(n)|x(n)|^{i-1}$ for each nonlinear order i, while other memory terms such as $x(n-j)||x(n-j)|^{i-1}$ can be easily obtained by delaying existing terms. Thus, if only odd nonlinear orders are considered, the total number of FLOPs for these basis functions are [16]

$$C_{\mathrm{MP,basis}}(P, M) = 3 + (P-1)/2. \tag{5.1}$$

In the filtering step, these basis functions are multiplied by each complex-valued coefficient a_{ij} with 6 FLOPs, and each multiplication output is added together with 2 FLOPs. The number of coefficients of MP is $\left(\frac{P+1}{2}\right)(M+1)$ considering only odd-order nonlinear orders. Thus, the total number of FLOPs for the filtering step is [16]

$$C_{\mathrm{MP,filter}}(P, M) = 8\left(\frac{P+1}{2}\right)(M+1) - 2. \tag{5.2}$$

Overall, the complexity of MP is a summation of basis and filter complexity

$$C_{\mathrm{MP}} = C_{\mathrm{MP,basis}} + C_{\mathrm{MP,filter}} \tag{5.3}$$

The complexity of GMP can be calculated in a similar way as MP. However, an addition term G_b is added which represents the amount of memory in lagging terms [16]

$$C_{\mathrm{GMP,basis}}(P, M, G_{\mathrm{b}}) = 3 + 7 + 2P + 2(P-1)G_{\mathrm{b}} + 2P\min(G_{\mathrm{b}}, M). \tag{5.4}$$

After reducing complexity the number of coefficeints for GMP becomes [16]

$$f_{\mathrm{GMP}}(P, M, G_{\mathrm{b}}) = (M+1)(P + 2PG_{\mathrm{b}}) - \frac{G_{\mathrm{b}}(G_{\mathrm{b}}+1)}{2}(P+1) - 2. \tag{5.5}$$

For $G_{\mathrm{b}} \leq M+1$, the complexity for the filtering step of GMP is calculated as [16]

$$C_{\mathrm{GMP,filter}}(P, M) = 8f_{\mathrm{GMP}}(P, M, G_{\mathrm{b}}) - 2. \tag{5.6}$$

Table 5.6: Computational complexity of MP and GMP algorithms in terms of FLOPs.)

		M=0	**M=1**	**M=2**	**M=3**	**M=4**
3MP	**P = 7**	19	23	27	31	35
	P = 9	22	27	32	37	42
3GMP	**P = 7**		354	984	1614	2244
	P = 9		494	1304	2114	2924

5.6.2 Computational Complexity for proposed ML algorithms

5.6.2.1 Computational Complexity of NN

The number of multiplication for a fully connected NN is similar to the matrix multiplication. For two fully connected NN layers with n and m neurons connected as following

$$\boldsymbol{y} = f(\boldsymbol{W}\boldsymbol{x} + \boldsymbol{b}) \tag{5.7}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ is the input vector, $\boldsymbol{W} \in \mathbb{R}^{m \times n}$ is the weight matrix, $\boldsymbol{b} \in \mathbb{R}^m$ is bias vector and f denotes the activation function. The number of multiplications for matrix multiplication $\boldsymbol{W}\boldsymbol{x}$ are $n * m * 1$, since $\boldsymbol{x}$ is a column matrix and the number of additions for the matrix are $(m-1) * n$. Finally the number of addition for adding bias vector $\boldsymbol{b}$ is n. Therefore the total number of multiplications and additions are

$$C = nm \times 1 + (m-1)n + n = 2nm \tag{5.8}$$

For a fully connected NN, the total number of multiplications and additions for K layers and m_k number of neurons in kth layer are given by

$$C = \sum_{k=1}^{K} 2m_k m_{(k+1)} \tag{5.9}$$

5.6.2.2 Computational Complexity of GB

For GB if boosted trees are used than the performance is improved by growing each tree using information from previously grown trees. During prediction if n is the number of input samples, n_{trees} is the number of trees and l is the learning rate than the output is simply predicted by multiplying each corresponding residual of the tree with the learning rate l_r.

For a boosted tree model with r_k residues and for kth tree and base value of F_0 the predicted values can be generated by

$$P = F_0 + \sum_{k=1}^{n_{\text{trees}}} r_k l_r \tag{5.10}$$

Therefore the number of the total number of multiplications and additions are equal to the n_{trees}.

5.6.2.3 Computational Complexity of DT and LR

For DT if each node takes $O(1)$ to calculate then the prediction complexity of the tree is equal to longest route from root to the leaf node. Similarly for linear regression if X is the input, m is the slope, a is the initial intercept and e is the error term then the output Y for the term can be given by

$$y = a + mx + e \tag{5.11}$$

For each input it takes $O(1)$ to calculate the output. To conclude the complexity of DT and LR is equal to the number of inputs

Table 5.7: Computational complexity of ML algorithms in terms of real number multiplication corresponding to the numerical values.

ML Algorithms	**Prediction**	**Prediction Complexity**
Neural Network	$\sum_{k=1}^{K} 2m_k m_{(k+1)}$	$2 * 10 * 40 + 40 * 40 + 40 * 40 + 40 * 2 = 7360$
Linear Regression	p	10
Decision Tree	p	10
Gradient Boosting	$p\ n_{\text{trees}}$	10 * 402 = 4020

6
Conclusion

In this thesis, we have compared the performance of many ML methods including NN, GB, DT, and LR for PA modeling with Volttera-based methods including MP and GMP. We have tested three scenarios including single-band signals, multi-band signals using the whole signal together, and multi-band signals using separate low and high band signals. In the single band carrier scenario, NN generated the best results, while other ML algorithms like GB, DT and LR did not generate better performance compared to MP/GMP algorithms with the exception in case of separated, multi-band scenario. In the multi-carrier and multi-band scenarios, two scenarios with different IBW were considered, one with 400 MHz and the other with 600 MHz. It was observed that NN gave a much better performance than MP and GMP. During PA modelling using ML and traditional algorithms the ACEPR/NMSE values for 400 Mhz scenario were better than for 600 Mhz scenario. However, the degradation in the performance was observed more when using traditional algorithms like MP and GMP compared to the ML algorithms. The final scenario considered in this thesis was filtering lower and higher frequency carriers separately and then using the filtered input and output signals for PA modeling. MP and GMP achieved better performance than ML algorithms, which could be one future area of study to find the root cause. Additionally, the computational complexity of ML algorithms in terms of several real multiplications was calculated. One other topic for future study is additional parameter tuning for ML algorithms, especially for the final test scenario for separated multi-carrier signals.

Bibliography

[1] Landi, Alberto Piaggi, Paolo Laurino, Marco Menicucci, Danilo. (2010). Artificial Neural Networks for nonlinear regression and classification. 115-120. 10.1109/ISDA.2010.5687280.

[2] Friedman, Jerome. (2000). Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics. 29. 10.1214/aos/1013203451.

[3] Friedman, Jerome. (2002). Stochastic Gradient Boosting. Computational Statistics Data Analysis. 38. 367-378. 10.1016/S0167-9473(01)00065-2.

[4] Dikmese, Sener Anttila, Lauri Pascual Campo, Pablo Valkama, Mikko Renfors, Markku. (2019). Behavioral Modeling of Power Amplifiers With Modern Machine Learning Techniques. 10.1109/IMC-5G47857.2019.9160381.

[5] Dikmese Coskun, Guzin. (2020). Behavioral Modeling of Power Amplifiers with Machine Learning Techniques. M.Sc. thesis, Tampere University, Feb. 2020.

[6] Baliga, B Jayant. (2005). Silicon RF power MOSFESTS. World Scientific DOI:10.1142/5725.

[7] Yadav, Shatrughna Prasad Bera, Subhash. (2014). Nonlinearity Effects of Power Amplifiers in Wireless Communication Systems. 2014 International Conference on Electronics, Communication and Computational Engineering, ICECCE 2014. 10.1109/ICECCE.2014.7086613.

[8] Zhu, Anding Brazil, T.J.. (2007). An Overview of Volterra Series Based Behavioral Modeling of RF/Microwave Power Amplifiers. 1 - 5. 10.1109/WAMICON.2006.351917.

[9] Gozde E Linearization Of RF Power Amplifiers By Using Memory Polynomial. M.Sc. thesis, Middle East Technical University, June 2012.

[10] Morgan, Dennis Ma, Zhengxiang Kim, Jaehyeong Zierdt, Michael Pastalan, John. (2006). A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers. Signal Processing, IEEE Transactions on. 54. 3852 - 3860. 10.1109/TSP.2006.879264.

[11] Chani-Cahuana, Jessica. (2015). Digital Predistortion for the Linearization of Power Amplifiers. Licentiate of Engineering thesis, Chalmers University of Technology.

[12] Bonaccorso, Giuseppe. (2017). Machine Learning Algorithms. Packt Publishing Ltd.

[13] Patel, Harsh Prajapati, Purvi. (2018). Study and Analysis of Decision Tree Based Classification Algorithms. International Journal of Computer Sciences and Engineering. 6. 74-78. 10.26438/ijcse/v6i10.7478.

[14] Ghannouchi, F.M. Hammi, Oualid. (2010). Behavioral Modeling and Predistortion. Microwave Magazine, IEEE. 10. 52 - 64. 10.1109/MMM.2009.934516.

[15] Isaksson, M. Wisell, David Rönnow, Daniel. (2006). A Comparative Analysis of Behavioral Models for RF Power Amplifiers. Microwave Theory and Techniques, IEEE Transactions on. 54. 348 - 359. 10.1109/TMTT.2005.860500.

[16] Tehrani, Ali Cao, Haiying Afsardoost, Sepideh Eriksson, Thomas Isaksson, M. Fager, Christian. (2010). A Comparative Analysis of the Complexity/Accuracy Tradeoff in Power Amplifier Behavioral Models. Microwave Theory and Techniques, IEEE Transactions on. 58. 1510 - 1520. 10.1109/TMTT.2010.2047920.

[17] Mkadem, Farouk Boumaiza, Slim. (2011). Physically Inspired Neural Network Model for RF Power Amplifier Behavioral Modeling and Digital Predistortion. Microwave Theory and Techniques, IEEE Transactions on. 59. 913 - 923. 10.1109/TMTT.2010.2098041.

[18] Bhuyan, Manasjyoti Sarma, Kandarpa. (2016). Learning Aided Behavioral Modeling and Adaptive Digital Pre-distortion Design for Nonlinear Power Amplifier. IEEE Sensors Journal. 16. 1-1. 10.1109/JSEN.2016.2579261.

[19] Sheppard, Clinton. (2017). Tree-based machine learning algorithms: Decision trees, random forests, and boosting. CreateSpace Ind. Publish. Platform.

[20] Song, Juncai Zhao, Jiwen Dong, Fei Zhao, Jing Qian, Zhe Zhang, Qian. (2018). A Novel Regression Modeling Method for PMSLM Structural Design Optimization Using Distance-weighted KNN Algorithm. IEEE Transactions on Industry Applications. PP. 1-1. 10.1109/TIA.2018.2836953.

[21] Nielsen, Michael (2015), Neural Networks and Deep Learning. Determination Press.

[22] Hu, Xin Liu, Zhijun Yu, Xiaofei Zhao, Yulong Chen, Wenhua Hu, Biao Du, Xuekun Helaoui, Mohamed Wang, Weidong Ghannouchi, F.M.. (2020). Convolutional Neural Network for Behavioral Modeling and Predistortion of Wideband Power Amplifiers.

[23] Collins, Gayle Runton, David. (2007). Nonlinear analysis of power amplifiers. Microwave Journal. 50.

[24] Martins, J. P. Cabral, P. M. Carvalho, N. Borges Pedro, J. C.(2006). A Metric for the Quantification of Memory Effects in Power Amplifiers. IEEE Transactions on Microwave Theory and Techniques, vol. 54, no. 12, pp. 4432-4439, Dec. 2006, doi: 10.1109/TMTT.2006.882871.

[25] Wan, Lei Guo, Zhiheng Wu, Yong Bi, Wenping Yuan, Jinhong Elkashlan, Maged Hanzo, L.. (2018). 4G/5G Spectrum Sharing: Efficient 5G Deployment to Serve Enhanced Mobile Broadband and Internet of Things Applications. IEEE Vehicular Technology Magazine. PP. 1-1. 10.1109/MVT.2018.2865830.

[26] 3GPP TS 36.101 version 14.3.0 Release 14.

[27] Agapitos, Alexandros Brabazon, Anthony O'Neill, Michael. (2017). Regularised Gradient Boosting for Financial Time-series Modelling. Computational Management Science. 14. 10.1007/s10287-017-0280-y.

[28] Sezgin, Ibrahim Can. (2016). Different Digital Predistortion Techniques for Power Amplifier Linearization. M.Sc. thesis, Lund University.

[29] Jardin, Pascale Baudoin, Geneviève. (2007). Filter Lookup Table Method for Power Amplifier Linearization. Vehicular Technology, IEEE Transactions on. 56. 1076 - 1087. 10.1109/TVT.2007.895566.

[30] Ghannouchi, F. M., Hammi, O., Helaoui, M. (2015). Behavioral modeling and predistortion of wideband wireless transmitters. ProQuest Ebook Central.

[31] Messaoudi,N. Fares,M. Boumaiza,S and Wood,J.(2008). Complexity reduced odd-order memory polynomial pre-distorter for 400-watt multi-carrier Doherty amplifier linearization. IEEE MTT-S International Microwave Symposium Digest, 2008, pp. 419-422, doi: 10.1109/MWSYM.2008.4633192.