Department of Electronic & Electrical Engineering

Probabilistic Behavioural Modelling of Non-Linear Devices

Anna Davis Manjaly
Student Number: 17344882

May 8, 2023

Submitted in partial fulfilment of the requirements for the degree of MSC in Electronics and Electrical Engineering

Supervisors: Dr. Justin King and Dr.DeclanO'Loughlin

Declaration

I hereby declare that this Thesis is entirely my own work and that it has not been submitted as an exercise for a degree at this or any other university.

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at `http://www.tcd.ie/calendar`.

I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at `http://tcd-ie.libguides.com/plagiarism/ready-steady-write`.

Signed: ______________________ Date: 30/09/2022

Abstract

Behavioural device models are black-box models which depend only on measured data and are independent of the underlying physics behind the working of the device. Behavioural device models are widely used to model non-linear devices due to their computational convenience. Accurate models are useful for circuit and system designs to help minimise the design time and maximise the utility of simulations.

In most of the behavioural models, the measured data are assumed to be ideal, while in actuality these measured data are subjected to errors. These errors can influence the modelling of the device and result in inaccurate predictions. In many cases, the measurement system can be designed to minimise measurement uncertainty and measurement errors, but measurements will also be subject to random errors due to environmental conditions. In the existing studies which include measurement errors in modelling, point estimates are estimated for the output responses of the Device Under Test (DUT). However, these point estimates do not reflect the reality that the input data are measured in the presence of uncertainties.

In this work, a method to model non-linear devices with random errors is proposed. This method is based on the Bayesian probabilistic approach which gives probabilistic distributions for the model parameters and output responses of the DUT. Probabilistic distributions and credible regions for X-parameters are established by deriving probability distributions for the output responses of the DUT rather than point responses. Finally, the potential to use Bayesian Neural Networks (BNN) to achieve increased accuracy is proposed and examined.

Acknowledgements

Thanks God!

Words cannot express my gratitude to my supervisors Dr. Justin King and Dr.Declan O'Loughlin for their patience, guidance, and support. They have generously given their knowledge and expertise throughout the research. This endeavor would not have been possible without the financial, technical, and moral support of the Electronic & Electrical Engineering Department, Trinity College Dublin.

Many thanks to my colleagues for their feedback and moral support. I am also grateful to the technical staffs, and office staffs, who supported and inspired me.

I would be remiss in not mentioning my family, especially my husband, baby, and parents for motivating me throughout this process.

Contents

Abstract **ii**

1 Introduction **1**

1.1 Non-Linear Device Modelling 1

1.2 Challenges of Behavioural Modelling 2

1.3 Motivation for Research 3

1.4 Probabilistic approaches to quantify random errors 3

1.5 Probabilistic Bayesian Non-Linear Modelling 4

1.6 Aims 7

1.7 Contributions of the research work 7

1.8 Summary 8

2 Background and Motivation **9**

2.1 Device Models 9

2.1.1 Circuit Modelling 9

2.1.2 Examples of Circuit Modelling 10

2.1.3 Behavioural Modelling 12

2.1.4 Examples of Behavioural Modelling 12

2.1.5 Hybrid Models 14

2.2 Uncertainties In Behavioural Modelling 14

2.2.1 Systematic errors 15

2.2.2 Random Errors 16

2.3 Importance of Uncertainty estimation 17
2.3.1 Estimation of Systematic Errors 19
2.3.2 Estimation of Random Errors 21
2.4 Bayesian Statistics 22
2.5 Summary 23

3 Methodology 24
3.1 Introduction 24
3.1.1 Methodology for modelling the random errors in the model parameter extractions 24
3.1.2 Methodology for modelling of the output response with random errors 29
3.2 Data Analysis 31
3.2.1 Simulated Data 31
3.2.2 Experimental Data 33
3.3 Validation Methods 34
3.3.1 Resubstitution Validation Technique 34
3.3.2 Cross Validation Technique 34
3.3.3 Variance in the model parameters 35
3.3.4 Distribution of Noise 38
3.3.5 Variance of Noise 39
3.4 Summary 39

4 Results 40
4.1 Modelling of Model parameters with random errors 40
4.2 Modelling of scattered waves with random errors 44
4.3 Summary 53

5 Extension to Bayesian Probabilistic Neural Network 55
5.1 Bayesian Neural Network 56
5.2 Bayesian Probabilistic Neural Network Model 58

5.2.1 Bayesian Probabilistic Neural Network Model for output response 59
5.2.2 Possible Reasons for unexpected results 61
5.2.3 Summary 62

6 Conclusion **63**

6.1 Summary of the works 63
6.2 Limitations of the work 65
6.3 Future Scope 65

List of Figures

1.1 Flow chart 5

3.1 2-port Device Under Test (DUT) 25
3.2 Methodology for the Bayesian Probabilistic Modelling of X-parametrs . 28
3.3 Methodology for the Bayesian Probabilistic Modelling 32
3.4 Circuit Diagram for X-parameter Extraction 33
3.5 Prior Distribution of w_{2R} and w_{2I} 36
3.6 Posterior Distribution of w_{2R} and w_{2I} estimated using 5 data points . . . 37
3.7 Posterior Distribution of w_{2R} and w_{2I} estimated using 20 data points . . 37
3.8 Posterior Distribution of w_{2R} and w_{2I} estimated using 500 data points . 38

4.1 Credible band for $X^{S}_{21,11}$ 41
4.2 Credible band for $X^{FB}_{21,11}$ 42
4.3 Credible band for $X^{T}_{21,11}$ 43
4.4 Probabilistic Model for B_{21} at different input power 45
4.5 Variances at different input powers 46
4.6 Probabilistic Model for B_{21} 47
4.7 Probabilistic Model for new test points (new A_{21} values) 48
4.8 Variance of single data point Vs Number of Data Points 49
4.9 Prior Distribution 50
4.10 Posterior Distribution of w_{2R} and w_{2R} estimated using 20 Data points . 50
4.11 Posterior Distribution of w_{2R} and w_{2R} estimated using 40 Data points . 51
4.12 Posterior Distribution of w_{2R} and w_{2R} estimated using 400 Data points . 51

4.13 Distribution of retrieved noise 52
4.14 Variance of a single point Vs Standard deviation of noise 53

5.1 Bayesian Neural Network 56

List of Tables

3.1 Mean of posterior Distributions of model parameters 36

5.1 Mean Square Error 60

5.2 Error Distributions 60

5.3 Change in mean Square Error with the variance of noise added 61

Nomenclature

α	Precision of Prior Distribution.
β	Variance of the noise
ANN	Artificial Neural Network
BJT	Bipolar Junction Transistor
BNN	Bayesian Neural Network
BSIM	Berkeley Short-channel IGFET Model
BSIM-CMG	Berkeley Short-channel IGFET Model – Common Multi-Gate
BSIM-IMG	BSIM-Independent Multi-Gate
DUT	Device Under Test
FET	Field Effect Transistor
FFNN	Feed Forward Neural Network
GAAFET	Gate-All-Around FET
HICUM	HIgh CUrrent Model
ICCAP	Integrated Circuit Characterization and Analysis Program
LSOP	Large signal operating point
MEXTRAM	Most EXquisite TRAnsistor Model
MOSFET	Metal–Oxide–Semiconductor Field-Effect Transistor
NIST MUF	National Institute of Standards and Technology Microwave Uncertainty Framework
NVNA	Non-linear Vector Network Analyzer
PHD	Poly Harmonic Distortion
QPHD	Quadratic Poly Harmonic Distortion
RF	Radio Frequency
UTBB	ltra-Thin Body and BOX SOI
VBIC	Vertical Bipolar Inter Company
VNA	Vector Network Analyzer

1 Introduction

The need for connectivity and higher data rates is rising with each generation of the telecommunications network. This demand devices with higher integration levels and higher speed. Such advancements require accurate device models which can be developed in a short time. Device models describe the characteristics of the device by relating the input stimuli with output response using mathematical relationships or by analyzing the physics behind the working of the device. The design time can be reduced by the small number of design cycles that accurate device models can achieve. For the cost-effective production of devices, the manufacturers and designers depend on accurate device models. This challenges the device modelling engineers to develop precise device models.

1.1 Non-Linear Device Modelling

Device models are two types: Physics-based models and empirical models. In physics-based device models also known as circuit models, the characteristics of devices are predicted by analyzing the elemental physics of the device. The circuit model describes the architecture, structure, and parameters of each component of the device. The challenge of the physics-based models is that it is difficult to identify and analyze the physics behind the working of the non-linear devices. The circuit simulators will require high-level analysis tools to analyze these non-linearities. The computation will be time-consuming even by utilizing high-level analysis tools [1]. Even though the building of physics-based device models is a tedious and time-consuming task,

they can provide accurate predictions if all the physics behind the working of the non-linear devices is identified and analysed using high-level analysis tools.

Empirical models also known as behavioural models are widely used to model non-linear devices due to their simplicity & low computation time. In the empirical models, the model parameters are fitted to the measured data using different optimization methods. The basic concept of these models is to obtain the predictability and the operation of the device by utilizing various mathematical techniques [1]. These models depend only on experimental or simulated data. These methods do not consider the physics behind the operation of the device. Due to this reason, the engineers can model the devices without the knowledge regarding the origin or causes of non-linearities, which makes the modelling uncomplicated and convenient. Although it has several advantages over physics-based modelling, the accuracy is low compared to physics-based modeling since it is not analyzing the physics behind the device and due to approximations used for the convenience. It is a trade-off between accuracy and computation speed [1].

Hybrid models which is a combination of both physics-based model and empirical model are also available for the device modelling. These models have the advantages of both physics-based modelling and empirical modelling. In those models, certain proprieties will be physically analyzed while other properties will be statistically analysed.

1.2 Challenges of Behavioural Modelling

Even though behavioural modelling is convenient over circuit modelling, which makes them the most commonly used modelling approach for RF devices, it has certain limitations. As the behavioural models only depend on the measured data, the accuracy of the model heavily depends on the accuracy of the measured data. The measurement error will be reflected in the model's prediction. As a result, even the most efficient models may do incorrect predictions. A reliable behavioural model can be achieved

only if these measurement errors are considered in the modelling. In this work, the random errors in the non-linear devices are quantified using statistical methods. These errors are unbiased errors which will be different in each measurement. In most cases the causes or origin of the random errors are unknown. The errors caused by the thermal noise of the measuring equipment are an example of random errors.

1.3 Motivation for Research

Manufacturers and designers rely on the behavioural models to design and manufacture their products. Due to the presence of random errors, these products may behave differently from what they are supposed to be. So the uncertainty in these models needs to be quantified. If the uncertainty percentage or the credible interval is within the acceptable limits of the manufacturer's or designer's criteria, then the non-linear system can be used. Otherwise, the system needs to be redesigned or more data needs to be analysed. Hence an efficient and convenient system for estimating the uncertainty of behavioural models is needed. In the prevailing models [2], [3], [4],[5], point estimates are predicted for the output response of the devices. No model can predict the response perfectly due to the presence of random errors. This research focuses to predict probability distributions that characterize the behaviour of the devices which can give a range of possible values and their probabilities.

1.4 Probabilistic approaches to quantify random errors

Different statistical methods are used to estimate the uncertainties in non-linear device modelling. There are two types of statistical methods to create a probabilistic model: Frequentist or Classical statistics and Bayesian Statistics.

In the frequentist statistics, the probability is observed as the frequency of the random repeatable events [6]. The main feature of frequentist statistics is the null hypothesis. In the frequentist methods, the parameters are fixed or non-random values and the er-

ror bars or probability of outputs are obtained by utilizing the distribution of datasets by repeated sampling.

In contrast to frequentist statistics, in Bayesian statistics, the parameters are also random variables that have a specific distribution. In regression problems, the uncertainties related to the choice of model parameters can be quantified by considering the parameters as random variables. In this method, there is only one dataset. The uncertainties in the parameters are quantified through probability distributions over parameters. The distribution of the parameters is chosen from our prior knowledge of the parameters. Thus Bayesian statistics include prior information or hypothesis along with the available data. The Bayes theorem is used to convert the prior distribution to the posterior distribution using the observed data. The other advantage of the Bayesian method is that the posterior distribution of parameters can be updated with new data by considering the present posterior distribution as the prior distribution for estimating the updated posterior distribution. This method is utilized in this work.

1.5 Probabilistic Bayesian Non-Linear Modelling

The steps for probabilistic Bayesian Non-linear modelling are as follows:

- **Relate the input variable and its response using basis functions.**

  The output responses are related to the linear combinations of the non-linear functions of the input variables [6]. These non-linear functions of input variables are basis functions. The usage of the non-linear basis function makes the relation between output response and input variable a non-linear function with respect to input vector while linear with respect to weight vector. The linearity with respect to the weight vector will reduce the complexity of the computational analysis.

  In this method, we are relating the target variable (the variable for which we

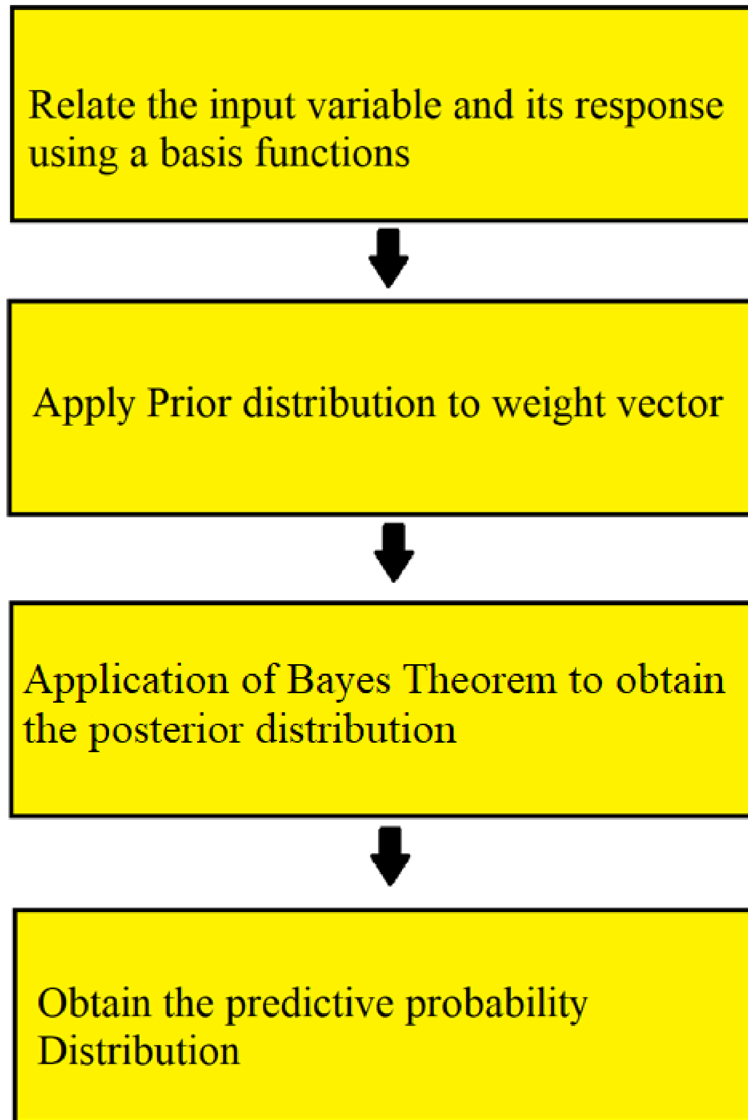


Figure 1.1: Flow chart

need to obtain predictive probability distribution) to the input using polynomial basis functions as non-linear basis functions. The coefficient of each term (weight) or each basis function is the model parameter. Consider y as the target variable, x as the input variable and M is the number of model parameters

$$\mathrm{y}(x, \mathbf{w}) = \mathbf{w}^T \phi(x) \tag{1}$$

where, $\mathbf{w}$ is the weight vector and ϕ is the vector of basis functions.

- **Apply Prior distribution to the weight vector w.**

  In this approach, a prior distribution is applied to these model parameters. The prior distribution is chosen according to our previous knowledge about the uncertainties in the extraction of model parameters. Mutually independent Gaussian distributions are assigned as prior distributions in this work.

- **Application of Bayes theorem to obtain posterior distribution.**

  Bayes Theorem is applied to the prior distribution and observed data to estimate the posterior distribution over model parameters. The likelihood function can be obtained from the observed data set. The likelihood function gives the probability of the observed data set for each set of parameter vector w. The negative of the log of the likelihood function is the error function in the Bayesian Statistics [6]. Using the Bayes theorem, the posterior distribution can be related to prior distribution and likelihood function as below:

$$\mathrm{P}(\mathbf{w}|D) = \frac{P(\mathbf{w}|D) \times P(\mathbf{w})}{P(D)} \tag{2}$$

$$\text{Posterior Distribution} \propto \text{Likelihood Function} \times \text{Prior Distribution} \tag{3}$$

$$P(\mathbf{w}|D) \propto P(\mathbf{w}|D) \times P(\mathbf{w}) \tag{4}$$

  where D is the observed data. The denominator in the (2) is used for the normalization constant which confirms the probability density of the posterior distribution integrated into one. So that the (2) can be reduced to (4). Using the above relation, the posterior probability distribution of the model parameters can be obtained.

- **Obtain the predictive probability Distribution**

The predictive probability distribution of the output variable can be estimated analytically by integrating the product of likelihood function and posterior distribution with respect to $\mathbf{w}$.

1.6 Aims

The overall aim of this work is to develop a method to estimate the probabilistic distribution of behavioural model parameters of non-linear devices. Specifically, two approaches are considered:

- firstly, using a Bayesian Probalistic approach;
- and secondly, using Bayesian Neural Networks (BNNs) [7].

In this work, the variations in the model parameters of non-linear behavioural models due to random uncertainties in the input data to the model parameter extraction is first quantified, followed by output prediction in the presence of random uncertainties. The proposed method can be adapted by device modelling engineers to build accurate behavioural models for non-linear devices which can be utilized by the designers for designing the system.

1.7 Contributions of the research work

Work from this thesis has been presented at two international conferences:

- at the Asia Pacific Microwave Conference (APMC) in Hong Kong in 2020;
- and at the European Microwave Integrated Circuits Conference (EuMIC) in London, the UK in 2021.

The specific contributions of this thesis are summarised below:

- A method based on Bayesian Statistics is proposed to quantify the random uncertainties in the model parameter extraction of non-linear devices. This method is utilized to estimate the random uncertainties in the well-known X-parameters.

Probability distributions were obtained for each X-parameter which gives the possible values for the X-parameters with their probabilities. Credible regions are also estimated using the proposed method. This work has been presented at APMC.

- An empirical paradigm is introduced to model the output response of non-linear devices with random errors. Bayesian Probabilistic approach is utilized in this method. This contribution has been presented at EuMIC 2021.
- An initial set-up is proposed for using Bayesian Neural Networks to model the random uncertainties in non-linear devices which overcomes the issue of the curse of dimensionality.

Together, these contributions help to quantify and account for random errors in the modelling of non-linear devices.

1.8 Summary

In this Chapter, different modeling approaches used for non-linear devices are discussed. The advantages as well as the challenges of behavioural modeling, are explained. The available probabilistic approaches are explained in this chapter. The aims and contributions of the work are stated in this chapter.

2 Background and Motivation

2.1 Device Models

As the device models are one of the most useful aids for RF circuit design, device modelling engineers are challenged to develop highly accurate and computationally efficient device models. Manufacturers and designers rely on device models to facilitate the cost-effective system designs that result due to first-pass design success. The successful realization of high-frequency devices is heavily dependent on efficient device models. Device models are the mathematical relationships that describe the characteristics of a device. These models are developed by analyzing the behaviour of devices under different operating conditions for example, linear, non-linear operating regions, which can be achieved by changing the parameters of the input signal. Device models predict the behaviour of devices by analyzing the physics behind the device or by relating the input stimuli and output responses of the device. Device models can be classified into circuit models and behavioural models based on the analysis methodology used.

2.1.1 Circuit Modelling

Circuit models are the physics-based analytical model which is developed by analyzing the behaviours of a device based on the fundamental physics by solving rigorous mathematical equations governing the underlying device physics[1]. Circuit models require physical knowledge about the device prior to modelling since the relationship

between the structure of the device and its underlying physical principles is an important factor in the circuit model. The accuracy of the physics-based models lies in the comprehension of the elemental physics underlying the device. This is challenging for non-linear devices. To analyze the non-linear behaviour of the device, high-level analysis tools are required. These models provide accurate results and have better scalability compared to behavioural models once all the physics underlying is identified and analysed. But the computation complexity and computation time increase with non-linear behaviour of the device. As new technologies and devices are emerging constantly circuit modelling will become a continuous task and predict the behaviour accurately, the detailed knowledge about the structure and technologies is required. This makes circuit modelling a tedious and time-consuming task for non-linear devices.

2.1.2 Examples of Circuit Modelling

Several circuit device models are available for Bipolar Junction Transistors (BJTs) including Ebers and Moll model in which BJT is approximated as p-n junctions. The circuit models for Feild Effect Transistors (FETs) for example Level 1 and Level 2 models are also introduced with first order approximation. Industry standard circuit models are also available for the multi-gate FETs.

Ebers and Moll have introduced a circuit model for bipolar junction transistor in which transistor was modeled as two p-n junctions placed back-to-back with the base p-type region being common to both diodes. The Ebers-Moll transistor model attempts to model the transistor as two diodes whose currents are determined by the normal diode equation with the addition of transfer ratios to include the interdependence of two diodes [8]. Gummel and poon have introduced a circuit model for bipolar transistors which is suitable for network analysis computer programs. This model is capable of modelling the high injection effects by utilizing charge control relation. This model reduces to the Ebers-Moll model for low bias applications with some idealizations [9]. The vertical Bipolar Intercompany Model (VBIC) which is similar

to the Gummel-Poon model was introduced to overcome some of the shortcomings of the Gummel-pool model. This model is more advanced than the Gummel-Poon model in early effect modelling, quasi-saturation modeling, parasitic substrate transistor modeling, parasitic fixed (oxide) capacitance modeling, temperature modelling, and avalanche multiplication modelling. In addition to that, it gives a smooth, continuous model [10]. Most EXquisite TRAnsistor Model (MEXTRAM) is another bipolar transistor circuit model which can be used for both Si and SiGe processes. It can model high voltage, high current, and high frequency such as Radio Frequency (RF) devices. It models high injection effects with a dedicated epi-layer model [11]. HIgh CUrrent Model (HICUM) is a physics-based geometry-scalable circuit model for homo- and hetero-junction bipolar transistors. It can be used for designing circuits that use Si, SiGe, or any III-V processes. This model is based on the generalized and extended integral charge control relation. It predicts accurately at high frequencies and high current densities [12].

Level 1 or Schichman-Hodges Model models Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) by first order approximation of output of long channel MOS device. This model is capable of considering channel length modulation and the body bias effect. This model is not sufficient to model saturation effect, carrier mobility degradation [13]. The level 2 model or Grove-Frohman Model is the advanced version of the Level 1 model. This model can model the mobility degradation by the vertical field, the depletion region, and the threshold region. This model is built on an assumption that threshold voltage is constant and it varies only with the substrate voltage [13].

Berkeley Short-channel IGFET Models (BSIM) are examples of circuit modelling that models MOSFET transistors for IC design. BSIM-Common Multi Gate (BSIM-CMG) is an industry standard model which is used for modelling common multi-gate devices, for example, FinFets, and Gate-All-Around FET (GAAFET). This model can model various physical phenomena originating from the quantum confinement of electrons by the small cross section of the GAAFET channel [14]. BSIM-Independent Multi-

Gate (BSIM-IMG) model has been developed to model the electrical characteristics of the independent double-gate structures like Ultra-Thin Body and BOX SOI transistors (UTBB). It allows different front- and back-gate voltages, work functions, dielectric thicknesses, and dielectric constants [15].

2.1.3 Behavioural Modelling

The alternative approach to circuit modelling is behavioural modelling. The behavioural models are black box models which predict the behaviour of devices by experimentally determining the response of the device and fitting it into the model. AM/AM and AM/PM are examples of the measurement techniques used in behavioural modelling. In AM/AM technique, the input and output magnitude is compared at different input power, while the AM/PM technique compares the change in the phase of the output signal with the phase of the input signal at different input power [16]. Since these models only depend on measured data (input stimuli and the corresponding response), they do not require knowledge about the internal setup of the device. The main benefit of using behavioural models is that it is independent of technology. It is inconsequential whether the user is using compound semiconductor field-effect transistors or silicon bipolar technology [17]. The other benefit of behavioural models is that they can be shared and used by others without revealing the internal setup of the component [17]. These models predict the characteristics of devices accurately with high computational speed. This modelling approach is well suited for non-linear devices in which the structure and the underlying physical mechanisms of the device are complex, but can be measured. Researchers prefer behavioural modelling approach for non-linear devices due to their convenience and rapid results.

2.1.4 Examples of Behavioural Modelling

Several behavioural models are available for modelling linear and non-linear devices. S-parameter modelling is one of the frequency domain behavioural models used to model linear devices. The Poly Harmonic Distortion (PHD) modelling approach along

with X-parameter formalism was introduced to model non-linear devices. Pade approximation model and Cardiff model are examples of other non-linear behavioural models. Other modelling approaches including the neural network-based model and Bayesian theory-based model are available for modelling non-linear devices.

S-parameter modelling approach of one of the most used frequency domain modelling approach in the RF microwave industry to model linear devices and devices which behaves linearly at small amplitudes. It was introduced along with similar models such as Y parameter modelling, Z-parameter modelling, H-parameter modelling, T-parameter modelling, and ABCD modelling approaches. The S-parameters of one individual component are sufficient to estimate the S-parameters of any combination of that component and S-parameters are invariant with the phase of the incident waves.

Most of the systems are non-linear. Power Amplifiers, mixers, and multipliers are examples of non-linear systems. Poly Harmonic Distortion Modelling approach was introduced to model the non-linear systems [17]. In PHD approach the incident waves are related to the scattered wave using multivariate complex describing functions. The X-parameter Modelling approach is one of the extensions of S-parameter Modelling to non-linear systems using PHD modelling approach. In X-parameter modelling approach, the model is linearized around a large signal operating point. X-parameters give a good prediction on non-linear systems at matched or nearly matched conditions [2]. The addition of quadratic terms in the describing functions of the PHD models led to Quadratic Poly Harmonic Distortion (QPHD) Model. The QPHD model has the ability to give good predictions over non-linear systems even in unmatched conditions [3].

Pade approximation behavioural model is a non-linear behavioural model which uses rational function instead of X-parameter's polynomial function. Due to the diversity of Pade approximations, this behavioural model increases the modelling space available beyond the X-parameter method [4]. Cardiff Behavioural Models (Cardiff model

lite and Cardiff model +) are also capable of modelling the non-linear devices accurately. Cardiff model gives a local model and Cardiff model+ utilizes higher order mixing terms to model the entire impedance space with a single model. This model is flexible in the number of mixing terms and coefficient terms. This model can also include input and output harmonics and obtain good accuracy without utilising the superposition principle [5].

A non-linear behaviour model based on a neural network was introduced which can predict the behaviour of transistors over the entire Smith chart at different power levels with one set of model coefficients. It shows good prediction and extrapolation properties in both linear and strongly non-linear regions [18]. A Bayesian theory-based non-linear behavioural model was introduced which can determine the optimal impedance required at the device terminals for the maximum output power or efficiency [19].

2.1.5 Hybrid Models

Several approaches that combine statistical analysis and physics-based analysis were introduced. In [20], the X-parameter modelling is linked to the physics-based TCAD simulations. This provides a procedure to integrate the experimentally derived X-parameters with predictions obtained by physical simulations. In [21], the active devices and passive devices of GaAs MMIC X-band power amplifier were modelled by physics-based X-parameters and electromagnetic simulations respectively.

2.2 Uncertainties In Behavioural Modelling

Even though behavioural models are more convenient over circuit models in modelling non-linear systems, it has some drawbacks. One of the drawbacks of the behavioural models is that it depends only on measured data and the measured data are subject to two types of errors: systematic and random errors. Since behavioural models are not considering the physics behind the model, the cause or origin of these

uncertainties is not taken into account while modelling. The device may behave differently from the predicted behaviour due to these uncertainties. So that there may be a difference between the actual measurement and predicted measurement even for a very accurate behavioural model.

Systematic errors are deterministic and can be identified as they have the same error for every measurement, which can be corrected by better calibration methods. On the other hand, random errors are stochastic in nature, and in most cases, they are unavoidable and unpredictable. These errors may vary at each measurement, which makes them difficult to account for. In the non-linear systems the measurements may have additional systematic and random uncertainties than in linear systems due to power and phase calibrations [22].

2.2.1 Systematic errors

Systematic errors are biased errors. Uncertainties in calibration standards and uncertainties in measuring device hardware are examples of systematic errors. The calibration uncertainty is the difference in calibrated value and its reference base. The major calibration uncertainties are the uncertainties in the calibration curve due to limited data and the uncertainties that may occur in future measurements. Systematic errors other than calibration uncertainties are due to the difference in other factors such as instruments, operators, geometries, configurations, and inhomogeneities. An example of these types of systematic error is the error caused by the flange misalignment in rectangular waveguides used in 500-750 GHz frequency [23].

Three types of systematic errors are widely seen in reflection measurements: Directivity error, Source match error, and Frequency response reflection tracking error. Resistive bridges and couplers are used to separate two signals (for example incident and reflected waves). Directivity error is the result of leakage of incident signal to the receivers which are intended to receive only the reflected signals. The imperfect isolation of the couplers and resistive bridges of the analyzer equipment causes this

leakage term during reflective measurements. This error can be mathematically quantified and corrected by taking multiple reflection measurements. A source match error is caused by the imperfect source match of the test port. If the DUT is not having a perfectly matched source port, a series of unwanted wavelets may occur between the port and DUT which results in incorrect reflection measurements [24]. A reflection tracking error is the difference in the frequency response of the test receiver and the reference receiver of a stimulus port in reflection measurements. Frequency response reflection tracking error is caused by test cables,signal-separation devices, adapters, variations in the reference signal path, and test signal path.

As in reflection measurements, three types of systematic errors are mainly present in transmission measurements: Isolation error, Load match error, and Frequency response transmission tracking error. The isolation error which is also known as crosstalk error is caused by the imperfect isolation by the isolators used in the analyzer equipment. This results in the leaking of unwanted signals (signals other than transmitted signals of the DUT) to the test receiver of the transmission measurement port which causes errors in transmission measurements. As in the source match error, the imperfect matching of the load port with DUT may cause a ripple of unwanted transmission and reflection signals. This results in inaccurate transmission measurements. A transmission tracking error is caused by the difference in frequency response between the test receiver of a response port and the reference receiver of a stimulus port in transmission measurements [25].

2.2.2 Random Errors

They usually arise from noises in the measured data. These errors are unbiased as they occur unpredictably and irregularly while taking measurements and during model parameter extractions. The random errors that may occur while taking measurements can be categorized into three depending on the cause of the random errors: errors due to the noise in the instruments, errors from switch repeatability, and errors from connector repeatability.

The noise in the instruments is the unwanted electrical fluctuations or disturbances caused by the analyzer's components. Examples of those electrical disturbances are the low-level noise caused by the receiver's broadband noise floor and high-level noise due to phase noise of the local oscillator in the test set and noise floor. The noise errors can be reduced by rising the source power to the DUT. This aids to reduce low-level noise. The noise errors can be further reduced by the reduction in the intermediate frequency bandwidth and by applying several measurement points and sweep averages.

The switch repeatability errors are caused by the RF mechanical switches which are used in the analyzer equipment to control the source attenuator settings. The contacts of the mechanical switches may close in different manners when they are activated. This may affect the precision of the measurement. This noise can be avoided by not using mechanical RF switches for the measurements which need high precision.

Connector repeatability errors are the result of connector wear which can cause adverse effects on the electrical performance. This can be rectified to a certain amount by good connector care practices such as keeping the connectors clean, protecting the ends of connectors with plastic caps & maintain the temperature of the connectors the same as that of the analyzers [25].

2.3 Importance of Uncertainty estimation

The behavioural models are solemnly based on measurements and these measurements are assumed to be ideal. But in reality, measurements are subjected to errors. This reduces the reliability of the model and the devices may behave differently from the predicted manner. This may affect the overall system. If the uncertainties in the measurements are estimated and included in the modelling of the devices, the system designers can utilize this information. So that they can design more reliable systems.

Different studies prove that the uncertainties in the measurements can cause uncer-

tainties in model parameters which cannot be neglected in designing a reliable system. D. Schreurs et al. have studied the impact of measurement uncertainties on large-signal transistor modelling particularly the time-domain current and voltage waveforms at transistors current generator plane. The measurement was taken using a Non-linear Vector Network Analyzer (NVNA) and they have included the calibration uncertainties whose characterization is known. This uncertainty was included and propagated according to the National Institute of Standards and Technology Microwave Uncertainty Framework (NIST MUF). They observed that the relative uncertainty at the current generator plane for the considered device and experimental conditions are not negligible and the measurements cannot be considered ideal [26]. The uncertainties in the extracted model parameters can cause uncertainties in the quantities derived from these model parameters. The studies of Valeria et al. indicate the presence of uncertainties in the derived quantities such as current gain cutoff frequency and the maximum oscillation frequency of a microwave transistor which are derived from S-parameters. It also stated the importance of accounting for the uncertainties of extracted model parameters along with the absolute value of the model parameters [27].

Dylan et al. have studied the role of measurement uncertainty in achieving first-pass design success. They propagated their small-signal measurement uncertainty estimates through Keysight Integrated Circuit Characterization and Analysis Program (ICCAP) model extraction procedure. They observed that the uncertainty in simulations due to error in the small-signal measurements used in the model-extraction process is low where the model was expected to perform accurately, but the uncertainty becomes very large when the transistors are driven into operating states that cause the model to fail. This indicates a possible role for measurement uncertainty in predicting first-pass design success [28].

2.3.1 Estimation of Systematic Errors

As mentioned earlier the estimation of systematic errors in measurement is very important to accurately characterize a device and ultimately design a reliable electronic system. Several researchers have estimated the systematic errors using different methods.

Wojciech Wiatr and David K. Walker have proposed two step analysis for the estimation of systematic errors in the noise parameters which is caused by the imperfections in source-impedance measurements of a two-port network with the cold source technique. Uncertainties in the noise power, errors in the Vector Network Analyzer (VNA) measurements, and errors in the noise temperature calibration are the major causes of uncertainty in the noise parameters. This method analyzes the error propagation of the source impedance that arises from residuals of the VNA measurements. This method is utilizing the relation between effective input noise temperature and source reflection coefficient. In this method, two linear fractional transformations were done to decompose the residual errors in the VNA measurements into different relevant factor sets. They have arrived at analytical formulas which describe errors in the noise parameters for complex noise characterization techniques and cold source noise measurements. They tested their method in a low-noise PHEMT and in a microwave amplifier and estimated systematic errors [29].

Maoliu Lin et al. have introduced a covariance-matrix-based uncertainty analysis method to estimate the systematic errors in NVNA measurements. These works present a systematic uncertainty analysis of a vector-correction-based frequency-domain instrument. It also shows how the uncertainty of harmonic phase standard which can be traced by sampling oscilloscope propagates to NVNA measurements. The main advantage of the covariance-based uncertainty analysis method is that it considers the correlation between two quantities along with the variance of each quantity and this knowledge is very crucial in analyzing the uncertainty propagation through transformations and data processing methods. They have also proposed an eight-term error

model to relate the raw measurements and effective measurements. The accuracy of this error-term model can determine the confidence of the effective measurements. The drawback of this system is that this method can be used only for the models that relate the input stimuli and output response using linear function [30].

A method for estimating the systematic uncertainty of voltage and current synchronized phasor measurements used for the estimation of transmission line impedance parameters was introduced by Deborah Ritzmann. This method is utilizing an optimization technique for identifying constants for phasors. In this method, the measurements are corrected before the estimation of parameters. This method assumes that the parameters are linearly changing or constant over a short period of time and correction factors are identified using the optimization technique. They also conducted a case study using simulated transmission line measurements and the effectiveness of this method was compared with an existing least square method and they observed promising results. The limitation of this method is that they have assumed that the systematic errors will be constant or linearly proportional in magnitude and the phase angle to be addictive. The systematic error may follow a non-linear model which cannot be estimated by the above-mentioned method [31].

Laurence T. Stant et al. have used two methods to propagate calibration uncertainties to non-linear behavioural models. In the first method, they utilized Monto Carlo analysis to perturb each input quantity by a random amount and perform the calibration for the results. This process is repeated for multiple results and statistical analysis is done. This method has the ability to preserve the non-linearities in the calibration. The limitation of this method is that it is highly time-consuming and computationally intensive. In the second method, each error source is perturbed while other error sources are given their estimated value. The sensitivity analysis of each error source is done. The deviations from the estimated results are combined to quantify the standard uncertainty. In this method, the computational intensity is low compared to the first method but the non-linearity of the calibration is not preserved [32].

2.3.2 Estimation of Random Errors

The random errors which are non-biased can be estimated using different statistical methods. Several researchers have studied the estimation of random errors.

Peter Stuart Blockley et al. have analyzed the random error in the NVNA measurements. They have derived an approximate covariance matrix corresponding to the NVNA measurements which can be used for the model fitting and maximization of the range of measurement setup. In this method, the covariance matrix was obtained by a non-linear transformation of variance of complex valued psuedowaves. The covariance matrix was estimated with the assumption that the reference measurements of the amplitude and phase of the signal are corrupted by independent narrow band additive Gaussian noise with equal variance. This may not be true in real applications. To compensate for that an additional factor was derived empirically and included in the covariance matrix using a linear transformation. The approximated covariance matrix can be used for uncertainty boundary calculation. This method was verified in different IF bandwidths and phase reference tones [33].

Alexander Arsenovic et al. introduced an experimental technique for uncertainty analysis in which the calibration uncertainty metrics are estimated at the output of the calibration processing chain. Even though this method aids to reduce the complexities of error propagation, the computational effort will be increased by performing a large number of calibrations. This method can only be utilised to estimate unbiased errors. In this method, redundant measurements of the calibration standards are taken and a set of error networks or error correction information is calculated from multiple redundant calibrations. The generated set of error networks is a function of the assumed response and redundant measurements of calibration standards. The set of corrected responses is estimated from this calculated error network and a single measurement. They have calculated the uncertainty in measurement in three different mediums. The unbiased uncertainties in flange misalignment in the rectangular wave guide, the uncertainty of the S-parameters of a two-tier Coplanar Waveguide (CPW) Probe, and the

uncertainty in the reflection coefficient of an optical quasi-optical annular ring-slot antenna array were estimated using the introduced method [23].

2.4 Bayesian Statistics

Different probabilistic approaches were applied to model the variations of predicted output response based on frequentist statistics and Bayesian statics available.

In [32], a probabilistic distribution and 95% confidence interval for X-parameters have been estimated using frequentist statistics which depends only on the frequency of occurrences. In [23], the deviation of the reflection coefficient of the eighth wave verification standard due to the random errors in VNAs. Donati Guerrieri et al. have done a frequentist statistical analysis and estimated the probabilistic distribution of output power of the power amplifiers due to the doping variations [34].

Some studies utilized Bayesian statistics in microwave device modelling. Jialin Cai has proposed a Bayesian inference-based small-signal modelling technique for GaN HEMTs which is combining Bayesian statistics and equivalent circuit modelling. Thus it has the advantages of both machine learning technique and physical interpretation of the DUT. The proposed model is compared with Feed Forward Neural Network (FFNN) and the proposed model has provided more precise results particularly due to the absence of the over-fitting issue that is dominant in the neural network [7].

In [19] Bayesian statistics are used to develop a non-linear behavioural model. In this work, different probability distributions are considered for modelling and the optimum model is identified using experimental testing. This method was compared with the Artificial Neural Network (ANN) approach and obtained more accurate results mainly due to the mitigation of over-fitting issues. This model is also compared with the circuit model and observed better prediction capabilities.

A method to estimate the measurement analysis using Bayesian inference was proposed in [35] by A. Zanobinil. This method was able to estimate a posterior distribu-

tion function related to the measurand from which the standard uncertainty, expected value, and coverage interval can be calculated.

2.5 Summary

In this chapter different device modelling techniques are discussed. This chapter also mentioned the systematic and random uncertainties that may occur in the measurements. Their effect on behavioural models is examined. The importance of estimation of these uncertainties and some methods for the estimation of the systematic and random errors in the measurement are discussed in this chapter. The probabilistic approaches available modelling the variations due to measurement errors are also discussed.

3 Methodology

3.1 Introduction

The research project aims to develop a probabilistic behavioural model for non-linear devices with random errors. The basic methodology is to apply the Bayesian-based statistical analysis to measured data and to obtain predictive probability distribution for the output response. The proposed method is applied to simulated data and experimental data. This method is validated with credible validation methods.

3.1.1 Methodology for modelling the random errors in the model parameter extractions

There can be multiple random uncertainties in the model parameter extractions. The methodology to model the model parameters with random uncertainties that occur in the extraction of model parameters using the Bayesian approach is explained in this section. The well Known X-parameter modelling is chosen as as example for this. In this section, the X-parameters with random uncertainties is modelled using Bayesian approach. X-parameter modelling is one of the popular extensions of S-parameter modelling to the non-linear regime. The Poly Harmonic Distortion method is introduced in [17] as a method to extend the scattering formalism i.e., S-parameter models, to the non-linear regime along with the X-parameter simplification which linearizes the PHD model around a given large-signal operating point (LSOP).

In figure 3.1, the waves incident on port p of the DUT at frequency q are denoted by

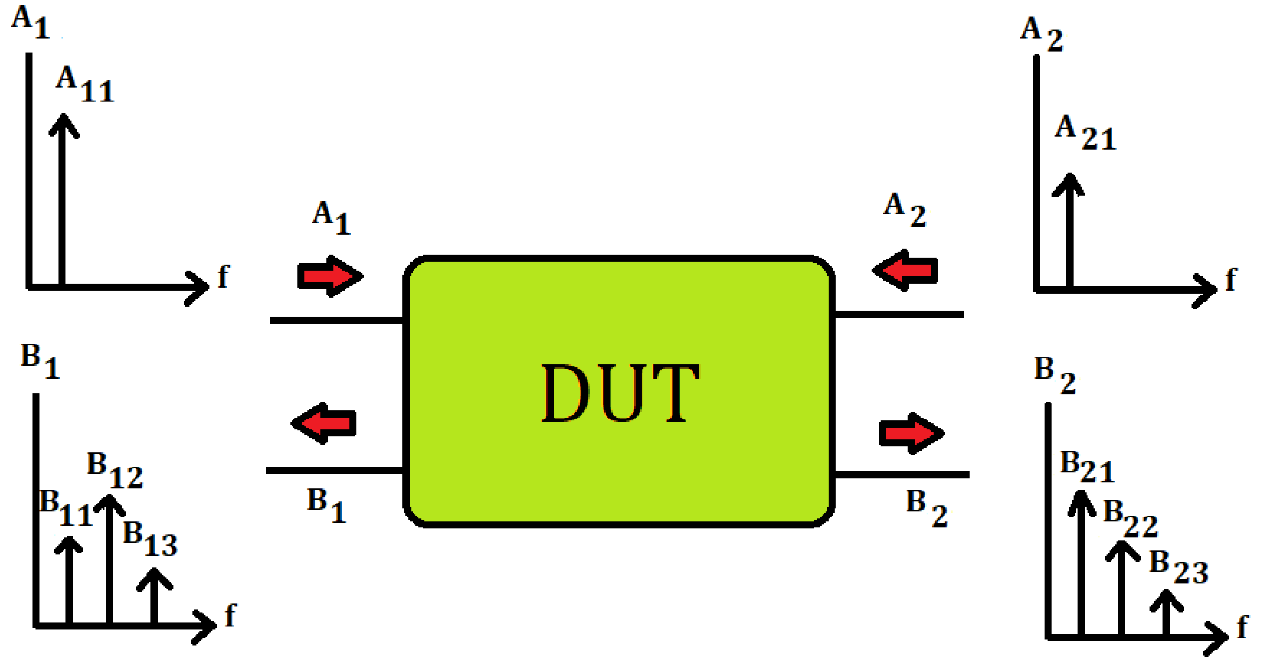


Figure 3.1: 2-port Device Under Test (DUT)

A_{pq} and the waves scattered from port i of the DUT at frequency k are denoted by B_{ik}

The pseudowaves A and B are defined as simple linear combinations of the port voltage V and current I at respective ports.

$$\mathrm{A}_{pq} = \frac{V_{pq} + Z_0 I_{pq}}{2\sqrt{Z_0}} \tag{1}$$

$$\mathrm{B}_{ik} = \frac{V_{ik} - Z_0 I_{ik}}{2\sqrt{Z_0}} \tag{2}$$

where Z_0 is the characteristic impedance. In the PHD formalism, the scattered waves are related to the incident waves via Describing Functions F_{ik}.

$$\mathrm{B}_{ik} = F_{ik}(A_{11}, A_{12}, A_{13}, A_{14}, ...A_{21}, A_{22}.....) \tag{3}$$

Each scattered wave has a separate describing function. In the X-parameter formalism,

the wave scattered by the DUT can be described as

$$\mathrm{B}_{ik} = X^{FB}_{pk} + \sum_{q=1}^{N}\sum_{l=1}^{k}(X^{S}_{pk,ql}A_{ql}P^{k-1} + X^{T}_{pk,ql}A^{*}_{ql}P^{k-1}) \tag{4}$$

where B_{ik} is the pseudowave scattered from port p at the kth harmonic frequency, and A_{ql} is the pseudowave incident on port q at the kth harmonic frequency. The quantity P is the phase of the dominant large signal tone (usually A_{11}). The X^{FB} term represents the LSOP, while the coefficients X^{S} and X^{T} represent the variation around the LSOP due to the small signal perturbations A_{ql} and A^{*}_{ql}.

An RF power amplifier, which is a non-linear device is taken as a DUT to obtain simulated data. In the RF PA with output matching, the dominant tone is the fundamental incident wave, (in this case A_{11}). The X-parameters are functions of this independent variable and are typically extracted across multiple A_{11} values. Non-linear polynomial basis functions are used in this work to model the variation of the nominal value of the X-parameter versus A_{11}. Modelling the variations of the X-parameters versus power requires two separate models for the real and imaginary parts of complex-valued X-parameter quantities. $A_{11}$is been normalized to a real-valued quantity through multiplication by P, giving $\mid A_{11} \mid$. An example of this model is shown for the ReX^{S} parameter in (5) below.

$$\mathrm{R}eX^{S}_{norm}(|A_{11}|, \mathbf{w}) = w_0 + w_1|A_{11}| + w_2|A_{11}|^2 + ... + w_M|A_{11}|^M \tag{5}$$

The polynomial coefficients $w_0, w_1 \ldots, w_M$ collectively denoted by the vector $\mathbf{w}$ and M is the order of the polynomial. We now take a probabilistic predictive distribution on the value of each X-parameter, again proceeding with the ReX^{S} example gives

$$\mathrm{P}(ReX^{S}|\mathbf{w}_{ML}, \beta_{ML}, |A_{11}|) = N(ReX^{S}|ReX^{S}_{norm}(|A_{11}|, \mathbf{w}), \beta^{-1}_{ML}) \tag{6}$$

where $\mathbf{w}_{ML}$ is the maximum likelihood value of the weights vector, $\beta_{ML}^{-1}$is the maximum likelihood solution for the variance of the distribution, and where $N(y|m, \sigma^2)$ represents a Gaussian distribution with mean m and variance σ^2.

$$\mathrm{P}(\mathbf{w}|ReX_t^S, |A_{11}|_t, \alpha, \beta) \propto \; P(ReX_t^S|\mathbf{w}, |A_{11}|_t, \beta) * P(\mathbf{w}|\alpha) \tag{7}$$

The term on the left-hand side is the posterior distribution of the weight vector **w**, the first term on the right-hand side is the likelihood function and the second term is the prior distribution. ReX_t^S and $|A_{11}|_t$ represent the input training data, and α and β are the inverse variances of the prior distributions and likelihood, respectively. In this study, the prior and likelihood distributions are assumed to be Gaussian. Since the Gaussian distribution is conjugate to itself, the posterior distribution will also be Gaussian. Using the prior distribution and the data observed, the predictive Bayesian probability model can be determined as

$$\mathrm{P}(ReX^S|ReX_t^S, |A_{11}|_t) = N(ReX^S|m(|A_{11}|)S^2(|A_{11}|)), \tag{8}$$

where

$$\mathrm{m}(|A_{11}|) = \beta\phi(|A_{11}|)^T\mathbf{S}\sum_{n=1}^{N}\phi(|A_{11}|)_{t,n})ReX_{t,n}^S \tag{9}$$

$$\mathrm{S}^2(|A_{11}|)) = \beta^{-1} + \phi(|A_{11}|)^T\mathbf{S}\phi(|A_{11}|) \tag{10}$$

$$S^{-1} = \alpha\mathbf{I} + \beta\sum_{n=1}^{N}\phi(|A_{11}|)_{t,n})\phi(|A_{11}|)_{t,n})^T. \tag{11}$$

where **I** is the identity matrix, N and represents the number of training samples. Both real and imaginary parts of all X-parameters are represented by a normal distribution with a mean and variance obtained via a Bayesian approach, as above.This

method gives a probabilistic distribution and estimates a credible region for each X-parameters. Thus this work models the random errors in the model parameter extractions. Figure 3.2 is the flow chart of the methodology used to model X-parameters with random noise. In this section simulated X-parameters were used. A Gaussian random noise is added to this simulated X-parameters to include the random uncertainties in the model parameters extraction in the simulation environment. The designing of non-linear systems will be aided further if this methodology can be extended to model the output response with random errors of the DUT.

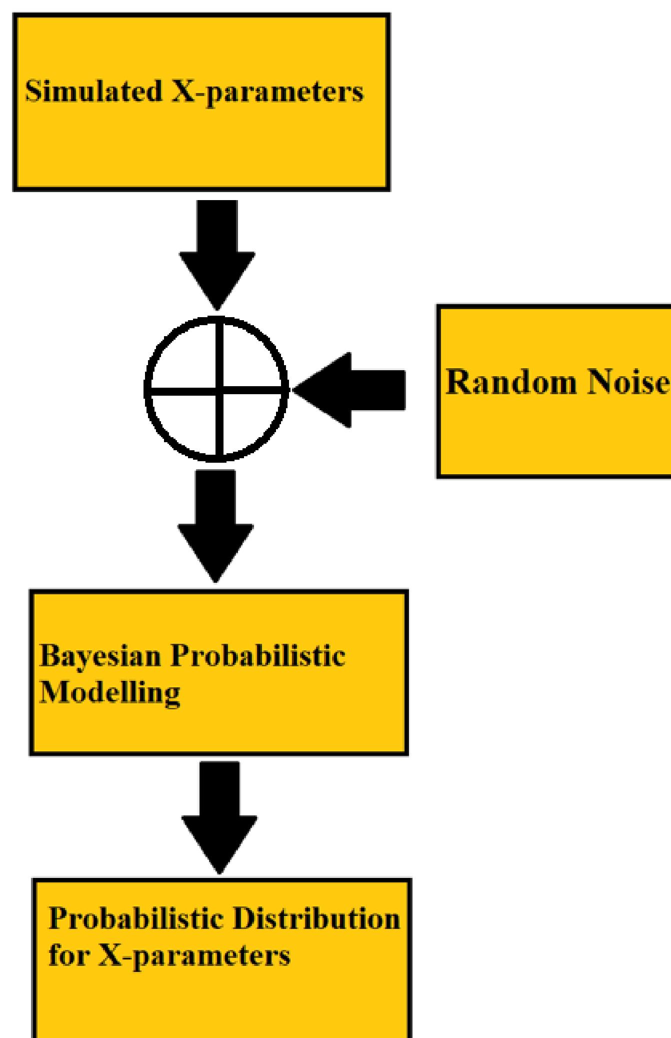


Figure 3.2: Methodology for the Bayesian Probabilistic Modelling of X-parametrs

3.1.2 Methodology for modelling of the output response with random errors

In this part of the work, Bayesian Statistics is used to quantify the random uncertainties in the output response. This part of the work introduces a new probability model for scattered waves (output response of the DUT). A simplified Quadratic PHD model is used to relate the scattered wave to the incident waves. In this work, the incident waves and the corresponding scattered waves are measured, and a Bayesian statistical analysis of these measurements is carried out. A probability distribution for the scattered waves is inferred from measured data, using Bayesian inference methods.

This work utilised, a polynomial describing function to model the nominal value of the scattered wave B_{ik} versus the incident waves A_{21}. Two separate models are needed for the real and imaginary parts of the scattered wave as before. The suffixes R and I indicate the real and imaginary parts, respectively. For example, the scattered wave as $B_{21,R}$, the dominant incident wave at port-1 as $A_{11}$and perturbations are caused by A_{21}. For simplicity, we have considered only the fundamental frequency for a two-port device. The scattered wave $B_{21,R}$ represented as a function of two independent variables $A_{21,R}$ and $A_{21,I}$ follows

$$\mathrm{B}_{21,R} = w_0 + w_{1R}A_{21,R} + w_{1I}A_{21,I} + w_{2R}A_{21,R}^2 + w_{2I}A_{21,R}^2 \tag{12}$$

where w_0, w_{1R}, w_{1I},w_{2R} and w_{2I} are the model parameters, which are real numbers. $A_{21,R}$ and $A_{21,I}$ are the real and imaginary parts of the incident wave at port 2. Since quadratic terms are used in (12), this model can be called a simplified Quadratic PHD (QPHD) model (since cross terms are omitted). The parameter w_0 can be calculated directly from the Large Signal Operating Point where only the dominant incident tone A_{11} in this example) is present. The quantity $\mathbf{W}$ represents the weight vector excluding w_0 i.e. $\mathbf{W} = [w_{1R}\ w_{1I}\ w_{2R}\ w_{2I}]$

A prior distribution is assumed for the model parameters **W** [6]. In this work, a dis-

tribution with zero mean and inverse variance αI (i.e., isotropic) is taken as this prior distribution.

$$\mathrm{P}(\mathbf{W}|\alpha) = N(\mathbf{W}|0, \alpha^{-1}I) \tag{13}$$

where I is the identity matrix. According to the Bayes theorem, the posterior distributions of these weight parameters are proportional to the product of the prior distribution and a likelihood function (which we assume to be Gaussian also),

$$\mathrm{P}(\mathbf{W}|B_{21,Rt}, A_{21,Rt}, A_{21,It}, \alpha, \beta) \propto P(B_{21,Rt}|\mathbf{W}, A_{21,Rt}, A_{21,It}, \beta) * P(\mathbf{W}|\alpha), \tag{14}$$

where the term on the left-hand side is the posterior distribution of the weight vector and the first and second term on the right-hand side is the likelihood function and prior distribution respectively. $B_{21,Rt}$, $A_{21,Rt}$ & $A_{21,It}$ are the input training data, which are column vectors. Since the prior and likelihood distributions are Gaussian, the posterior distribution will also be Gaussian due to the property of conjugate priors. The quantity β is estimated using the maximum likelihood method and a relatively small value is given to α, i.e. we are assuming a broad prior distribution, which will be updated using the training data.

The mean m_N and variance S_N of the Gaussian posterior distribution of these model parameters can be computed from training data as follows:

$$\mathrm{m}_N = \beta S_N \phi(A_{21,Rt},, A_{21,It}), B_{21,Rt}, \tag{15}$$

$$\mathrm{S}_N^{-1} = \alpha I + \beta\phi(A_{21,Rt},, A_{21,It})\beta(A_{21,Rt},, A_{21,It})^T, \tag{16}$$

where N is the number of training samples and $\phi(x, y)$ is the vector of basis functions, in this case given by $\phi(x, y) = [xyx^2y^2]$.

The predictive (posterior) distribution for $B_{21,R}$ will also be a Gaussian distribution.

$$\mathrm{P}(B_{21,R}|B_{21,Rt},\beta) = N(B_{21,R}|m,v) \tag{17}$$

where the mean m and variance v for the scattered waves can be computed from the training data as follows

$$\mathrm{m} = \beta\phi(A_{21,R},A_{21,I})S\phi(A_{21,Rt},A_{21,It})^T B_{21,Rt} \tag{18}$$

$$\mathrm{v} = \beta^{-1} + \phi(A_{21,R},A_{21,I})S\phi(A_{21,R},A_{21,I})^T \tag{19}$$

$$\mathrm{S}^{-1} = \alpha I + \beta\phi((A_{21,Rt},A_{21,It})^T\phi(A_{21,Rt},A_{21,It}). \tag{20}$$

The DUT used to obtain experimental data is a 2-port GaN transistor. The variations in scattered waves at port 2 in the presence of random errors are modelled using the proposed method. This demonstrates the ability of the proposed method to model the variations in the output response of the DUT with random errors. Figure 3.3 is the flow chart of the methodology used. In this section experimental data with random errors are included.

3.2 Data Analysis

In this method, two types of data is analysed. Both types of data was verified before applying the Bayesian approach.

3.2.1 Simulated Data

The X-parameters of an RF PA circuit (DUT) are extracted from a simulated model using Advanced Design System (ADS) Software. A frequency of 1 GHz is used for

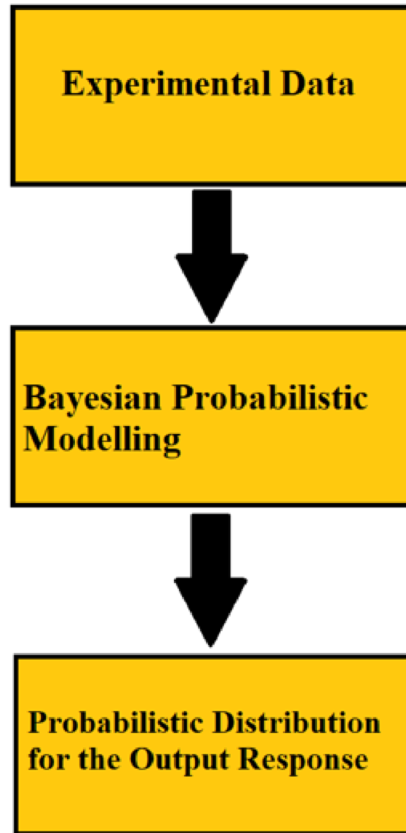


Figure 3.3: Methodology for the Bayesian Probabilistic Modelling

the fundamental frequency. Only the fundamental frequency is considered in this study for simplicity which can be extended to harmonic frequencies using the same method. A power sweep from -20 dBm to +10 dBm is performed. The X-parameters are extracted for each power level in this range, specifically selected to include both linear and nonlinear operating regions of the PA. The same set of data was simulated using both X-parameter simulation palette and a Harmonic Balance (HB) simulation palette. The X-parameter data obtained using both the X-parameter palette and HB simulation palette was compared and they were numerically the same. The simulated data does not contain any random uncertainties. To include the random uncertainties in the model extraction, a normally distributed noise with zero mean and 0.001 variances is included within the RF PA model. Figure 3.4 shows the circuit diagram for the extraction of the X-parameters.

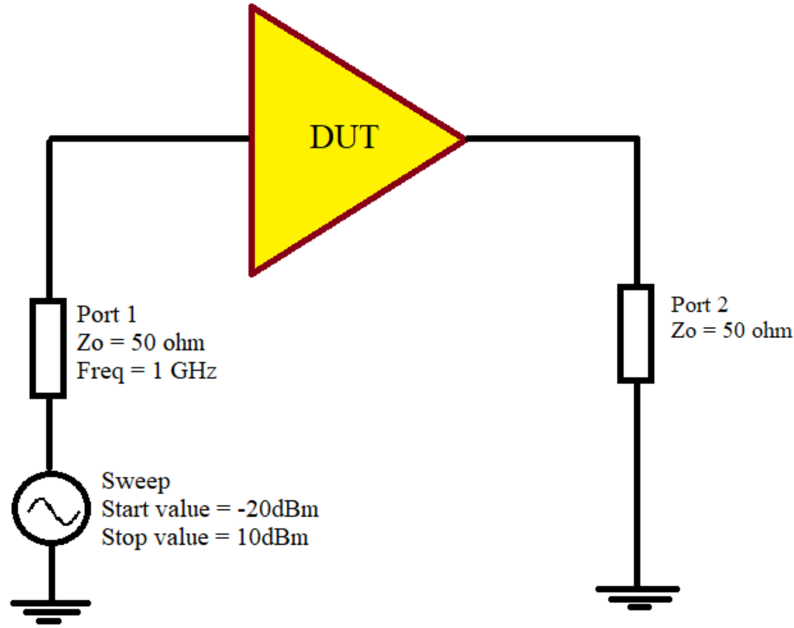


Figure 3.4: Circuit Diagram for X-parameter Extraction

3.2.2 Experimental Data

To create this dataset a 10 W RF GaN transistor (CGH40010F) is used as the DUT. It is a two-port device. The voltages at both ports & the currents through the two ports are measured at various load impedances. The scattered waves B_{21} at port-2 and incident waves A_{21} at port-2 are calculated at various load impedances using the equations (1),(2), at a fundamental frequency of 2 GHz. In this data acquisition, load-pull measurement system is used other than varying input power. This is to demonstrate the generalization ability of the proposed method. The load impedance was calculated using the obtained psuedowaves and compared with the applied load impedance. They are numerically equal. A total of 649 A_{21} values and corresponding B_{21} values are measured in this experiment. These measured data are divided into two subsets. The first subset contains 576 data points for training the model and the second subset contains 73 data points, to be used for model validation.

3.3 Validation Methods

The credibility of any method lies in its validation process. Several methods are used for the validation of the above-mentioned method.

3.3.1 Resubstitution Validation Technique

In this method, the whole dataset is used for the training of the model. The error rate of the model is estimated by comparing the predicted outcome of the model with the actual measured value. For the validation of the proposed methodology, the error rate is calculated by comparing the mean of the predicted probability distribution with the measured output response.

3.3.2 Cross Validation Technique

The available dataset is divided into a training dataset and a testing dataset. The training dataset is used for the training of the proposed model. The testing dataset can be used for validation purposes. The output of the testing data is predicted using the proposed model. The predicted output will be compared with the measured output to get the error rate. In this work, the dataset was divided into a training dataset and test dataset (in some cases also divided into validation datasets) in three ways:

- **Holdout Validation technique**: The sample points of the training dataset and test dataset are selected on a random basis.
- **Equally spaced dataset**: The dataset is split into training and testing datasets by equally spaced sampling. In this sampling method, sample points are equally spaced. As the training dataset should have a greater number of sample points than the testing dataset, the blocks of sample points of the training dataset is equally spaced and the sample points of the testing dataset are equally spaced. For example, if the dataset has 100 sample points and it is to be split into training and testing datasets in a 4:1 ratio. The training dataset will contain the first 4

sample points and the 5th sample point will be in the test dataset.

- **Block of sample points**: A block of sample points is taken as a test dataset and the rest of the dataset is used for training the model.

3.3.3 Variance in the model parameters

In this technique, the variance of model parameters is estimated using the proposed method with a different number of sample points. As the number of the sample points increases, the variance of the model parameters will decrease. This method is demonstrated with an example below.

In this demonstration, 500 complex-valued inputs A_{21} values are considered and the corresponding output values for $B_{21,R}$, the real part of the scattered wave phasor are computed, using known model parameters (w_{1R}, w_{1I}, w_{2R} and w_{2I}). Gaussian noise with mean 0 and standard deviation 0.01 $W^{1/2}$ is added to the scattered wave output.

Next, a prior probability distribution is assumed for the model parameters w_{1R}, w_{1I}, w_{2R} and w_{2I} in this case a four-dimensional Gaussian distribution. The Gaussian distribution is set to have zero mean and inverse variance of 0.005. For illustration purposes, the sequential update of the posterior distribution of two of the model parameters (w_{2R} and w_{2I}) is shown in figures 3.5, 3.6, 3.7 and 3.8.

Validation proceeds by observing the sequential update in the posterior model parameter distributions, after each 'prior' is updated via the likelihood function. After the first Bayesian iteration, using just five (assumed to be noisy) training data points $(A_{21,t,n}, B_{21t,n})|_{n=1...5}$, the variance of the posterior distribution of model parameters has reduced considerably. The posterior distribution is estimated from the prior distribution and likelihood function using the Bayes theorem. In the second iteration, the posterior distribution of the first stage is now taken as the 'prior' distribution, and the likelihood function is calculated using another 15 data points $(A_{21,t,n}, B_{21t,n})|_{n=6...20}$. As can be seen, the variance of the posterior distribution of the second stage has fur-

ther reduced compared to that from the first stage. The final plot in Fig. 3.8 shows the effect of a total of 500 training data points applied to the Bayesian model. It is seen that the posterior distribution of model parameters has a very small variance.

Table 3.1: Mean of posterior Distributions of model parameters

No.of Data Points	w_{1R}	w_{1I}	w_{2R}	w_{2I}
Actual Values	-0.4865	0.226	0.0018	-0.0464
5	-0.4938	0.2092	-0.0031	-0.0410
20	-0.4939	0.2286	-0.0005	-0.0447
500	-0.4872	0.2258	0.0018	-0.0465

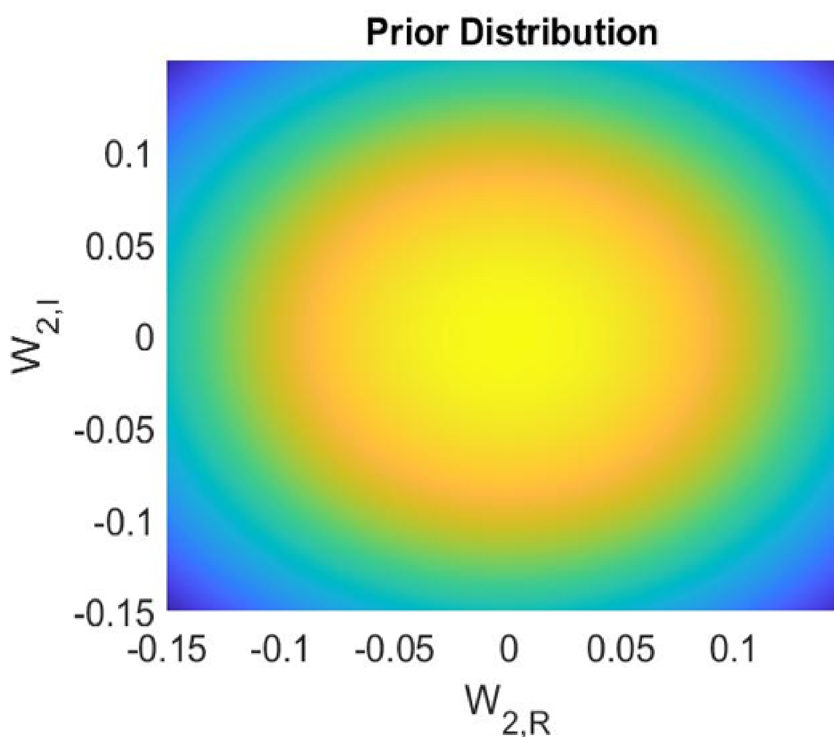


Figure 3.5: Prior Distribution of w_{2R} and w_{2I}

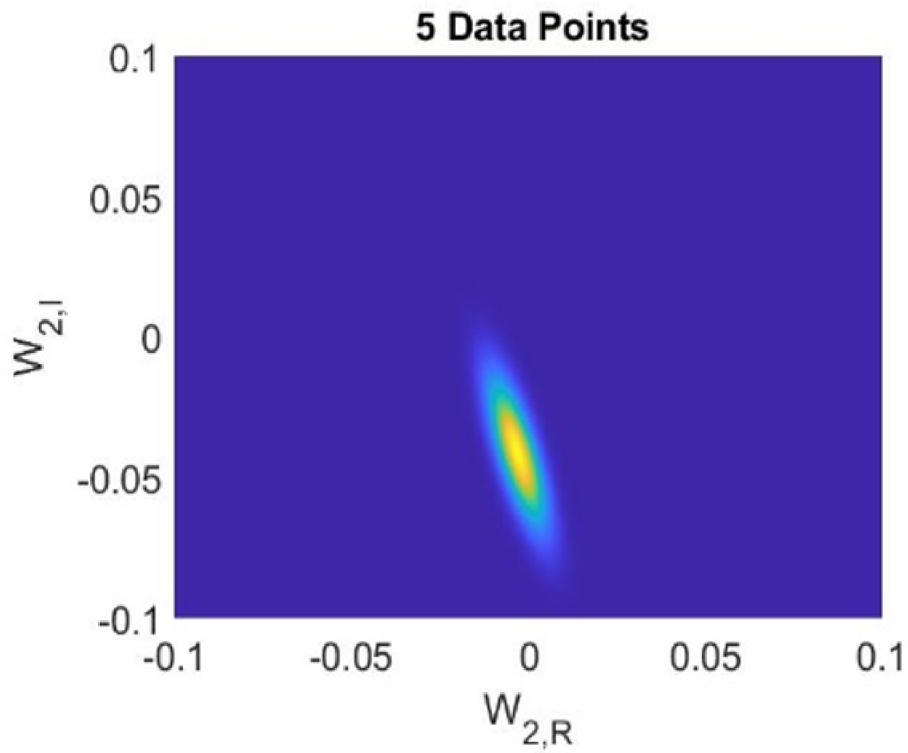


Figure 3.6: Posterior Distribution of w_{2R} and w_{2I} estimated using 5 data points

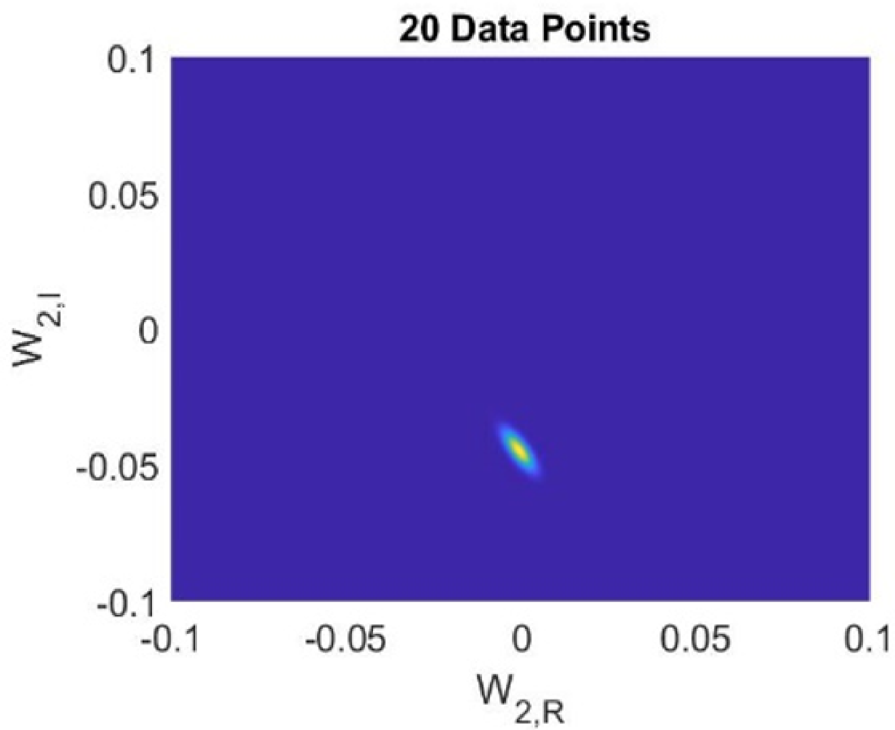


Figure 3.7: Posterior Distribution of w_{2R} and w_{2I} estimated using 20 data points

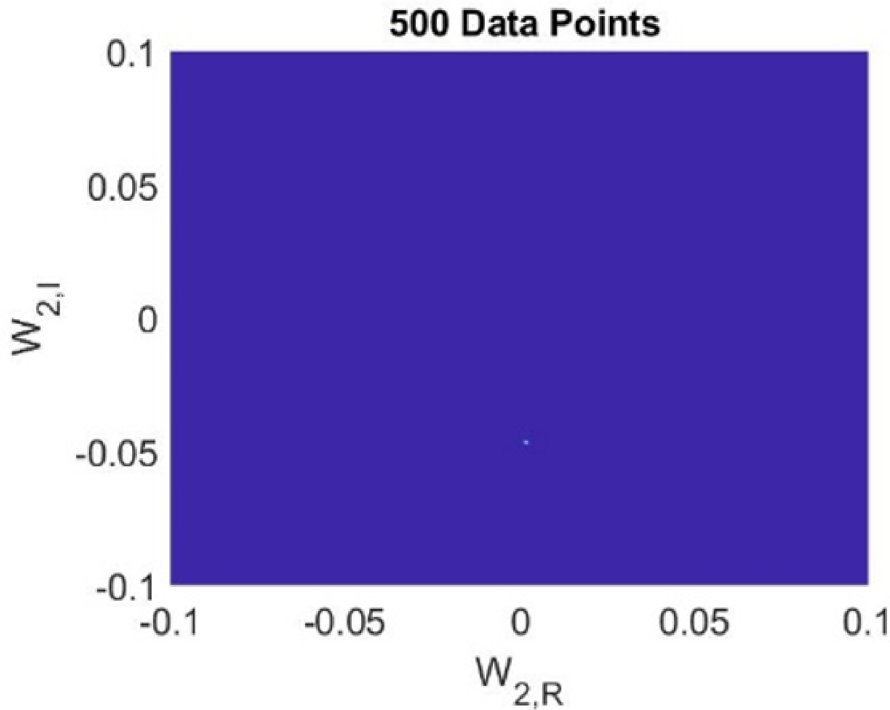


Figure 3.8: Posterior Distribution of w_{2R} and w_{2I} estimated using 500 data points

Table 3.1 shows the mean of the posterior model parameters distribution when considering 5, 20, and 500 data points. This validation technique is applied to the experimental data to validate the obtained results.

3.3.4 Distribution of Noise

In this validation technique, the distribution of the noise is compared to validate the results. A noise with specific distribution for example Gaussian Distribution is added to the measured data. Probabilistic distributions are predicted for the output responses of the DUT using the proposed theory. Retrieve the noise added from the predicted output. Then compare the distribution of the noise added to the measured data with the noise retrieved from the predicted output. The distribution of the noise added to measured data and the distribution of noise retrieved from the predicted output will be the same.

3.3.5 Variance of Noise

The variance of the noise is compared in this validation technique. Noise is added to the measured data set and predicts the variance of the output predictive distribution using the proposed method. Then repeat the process by adding the noise with different variances. The variance of the predicted output distribution will vary according to the variance of the noise added.

3.4 Summary

In this chapter, the proposed methodology is explained. The data types and their sources are explained in this chapter. The data analysis methods used for the verification of data are mentioned in this chapter. The techniques used for the validation of this work are discussed in this chapter.

4 Results

The methodology proposed in chapter 3 is applied to create probabilistic behavioural models which models the non-linear devices with random errors. Instead of a point estimate, probability distributions are predicted for the model parameters and output responses. In this work, we have applied the Bayesian-based probabilistic approach to the existing well Known X-parameter modelling and simplified Quadratic Poly Harmonic Distortion modelling. The predicted output response is compared with the measured true value in the training and validation dataset. The mean and variance of the predicted probability distribution is compared at different input powers.

4.1 Modelling of Model parameters with random errors

This section represents the results of applying Bayesian approach in modelling the X-parameters with random errors. As discussed in Chapter 3, simulated data is used in this section. The DUT is an RF power amplifier and the fundamental frequency used is 1 GHz. The power sweep from -20 dBm to +10 dBm with different step sizes is done to obtain 10 and 100 data points. A Gaussian random noise of 0 mean and 0.001 standard deviation is added to the obtained simulated data to include the random uncertainties in model parameter extractions. The prior distribution variance $\alpha^{-1}$is taken as 0.005^{-1}. Figure 4.1 shows the probabilistic model for $X^{S}_{21,11}$ (the small signal perturbation due to the A_{11} term) with the shaded region corresponding to ± one standard deviation around the mean (solid blue line). To validate the obtained credible interval the data points in the test dataset are plotted in the figures as circles. The test data set are

selected on random basis as in holdout validation technique. In the figure, only one data point and 3 data points among 10 randomly selected data points is outside the credible region of the real and imaginary parts of the $X^{S}_{21,11}$ respectively. The credible bands for $X^{FB}_{21,11}$ and $X^{T}_{21,11}$is shown in figures 4.2 and 4.3 respectively, again taking one standard deviation either side as the band of interest.

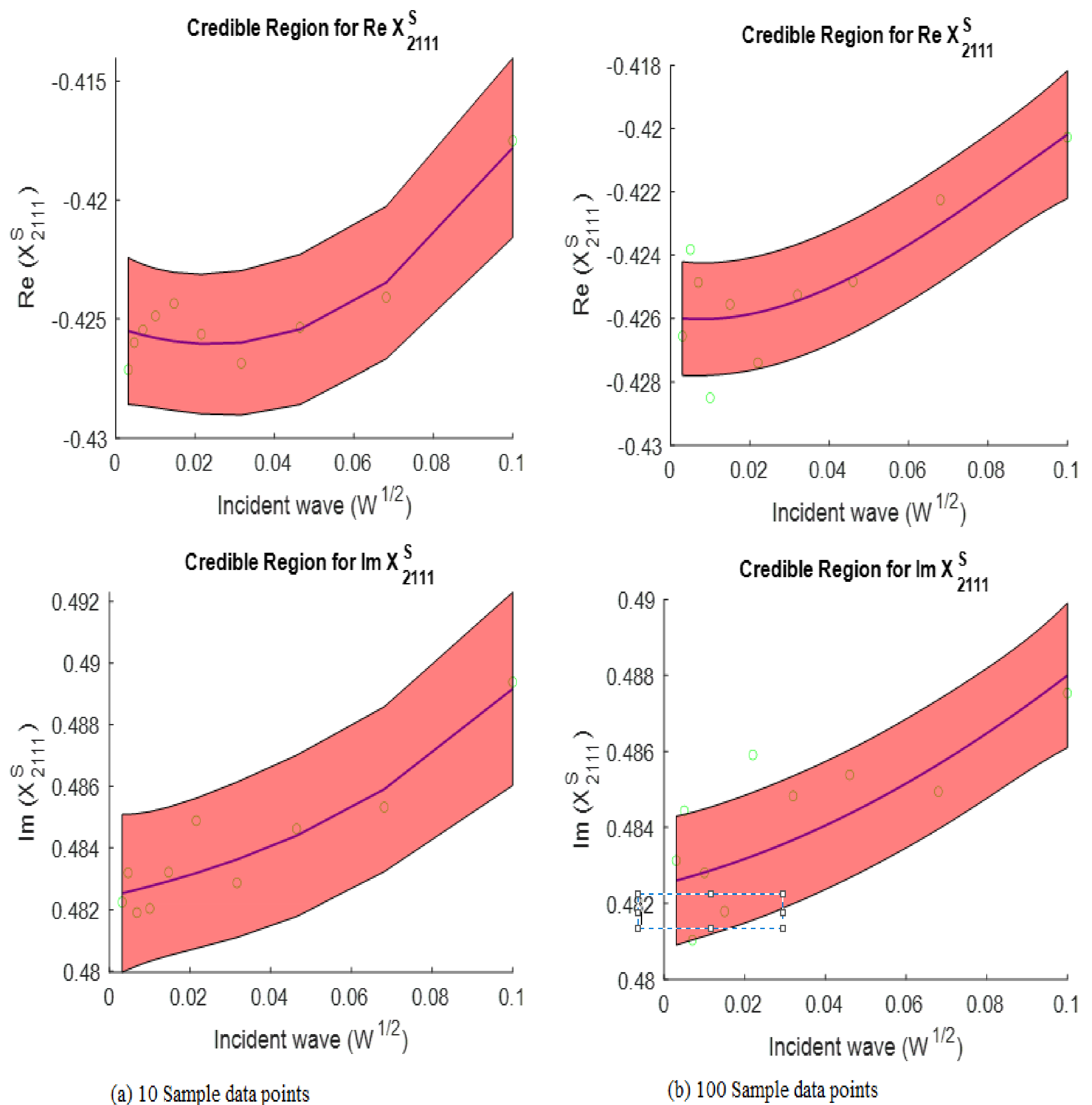
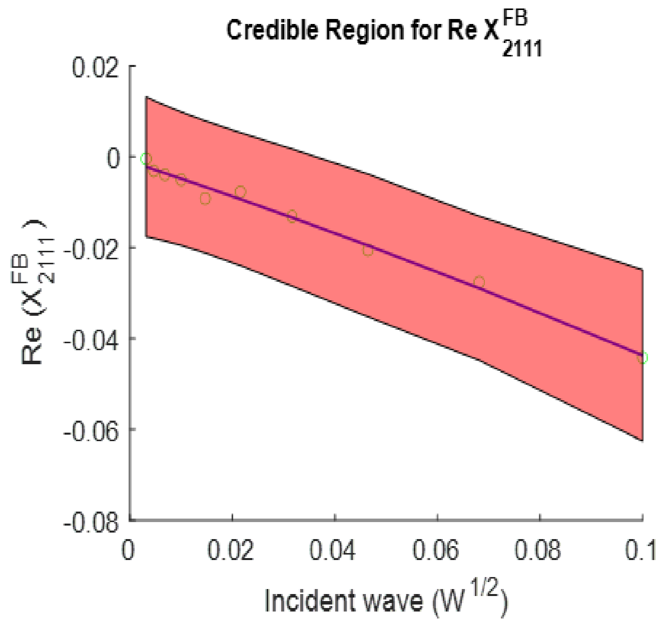
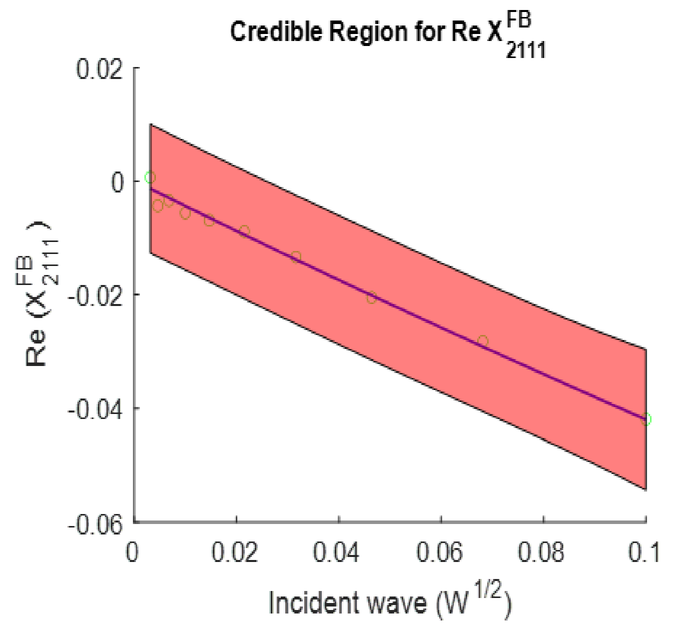
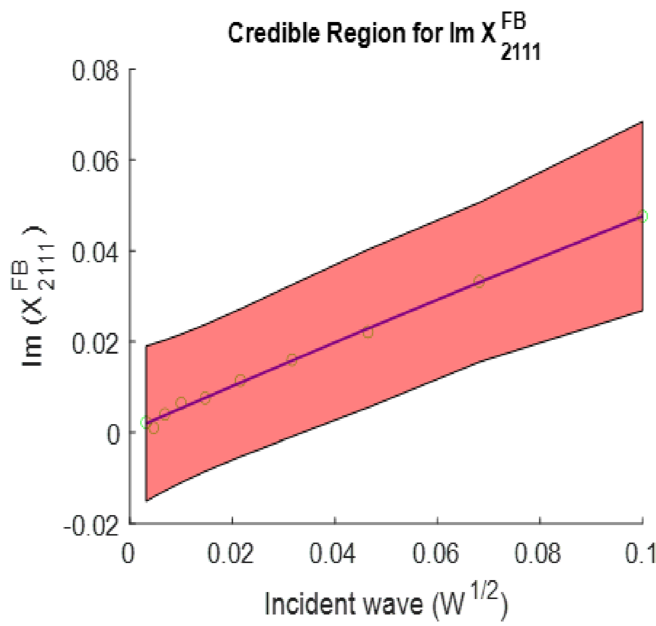


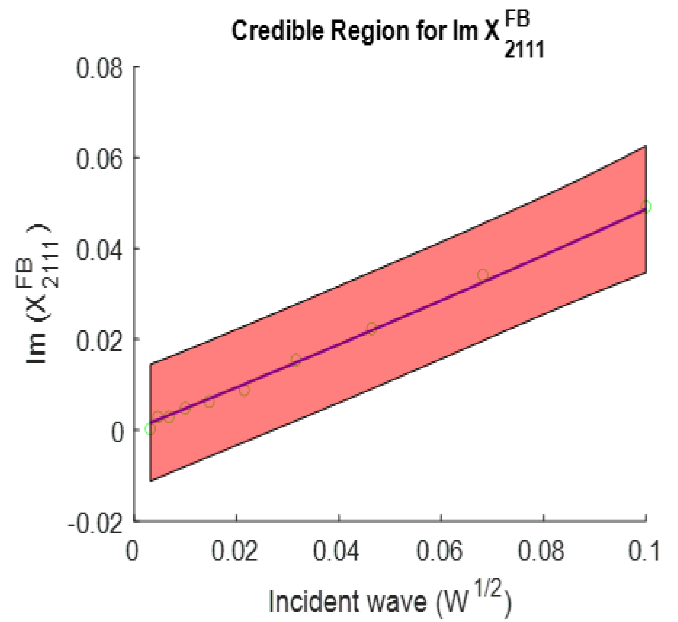
Figure 4.1: Credible band for $X^{S}_{21,11}$





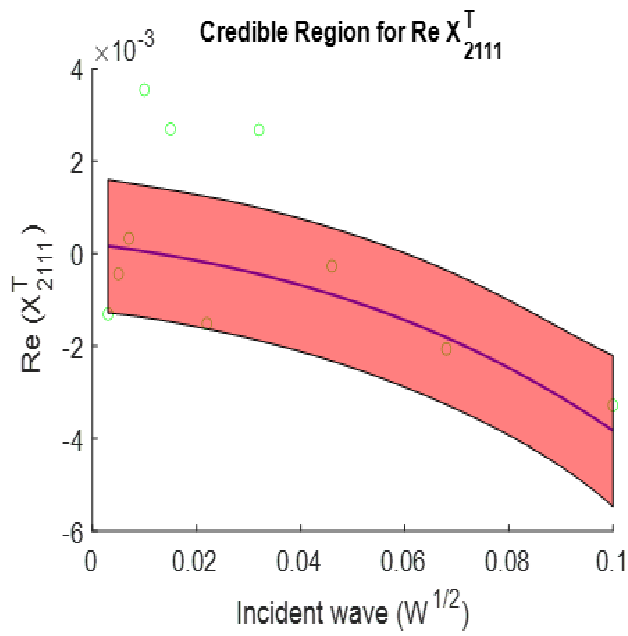


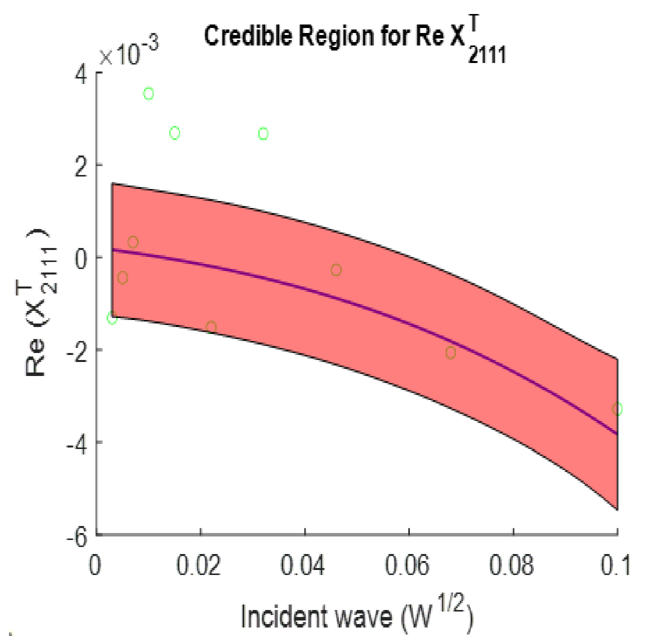
(a) 10 Sample data Points

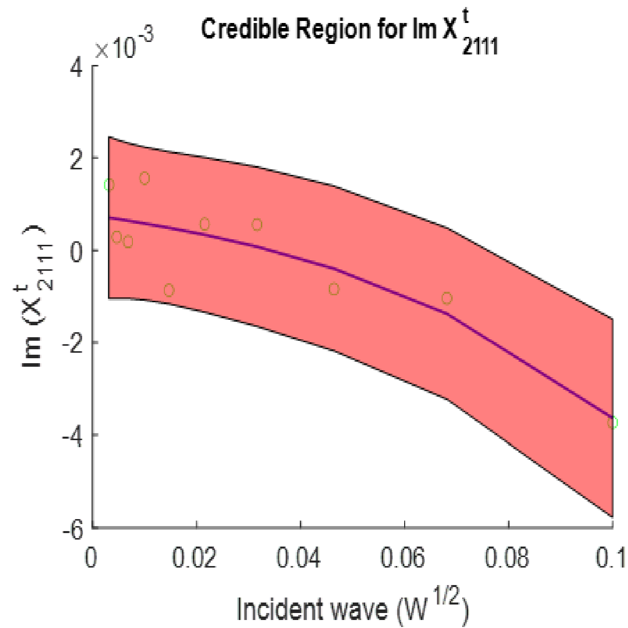


(b) 100 Sample data points

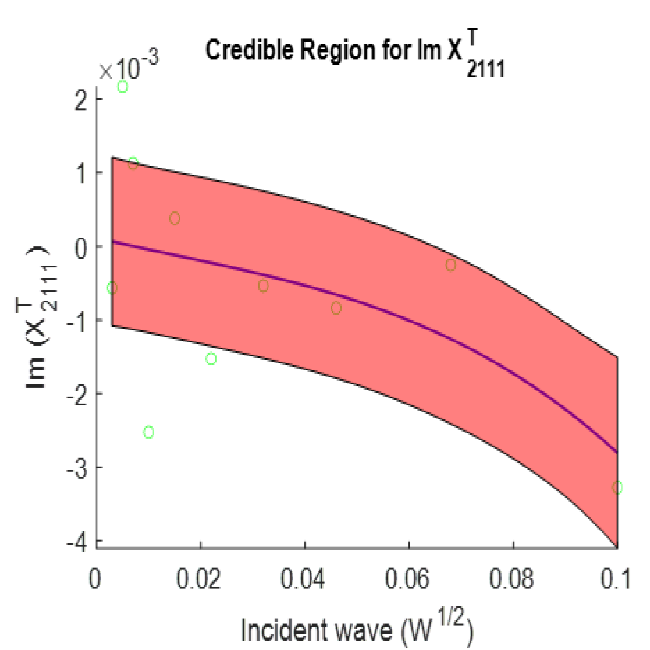
Figure 4.2: Credible band for $X^{FB}_{21,11}$







(a) 10 Sample data points



(b) 100 Sample data points

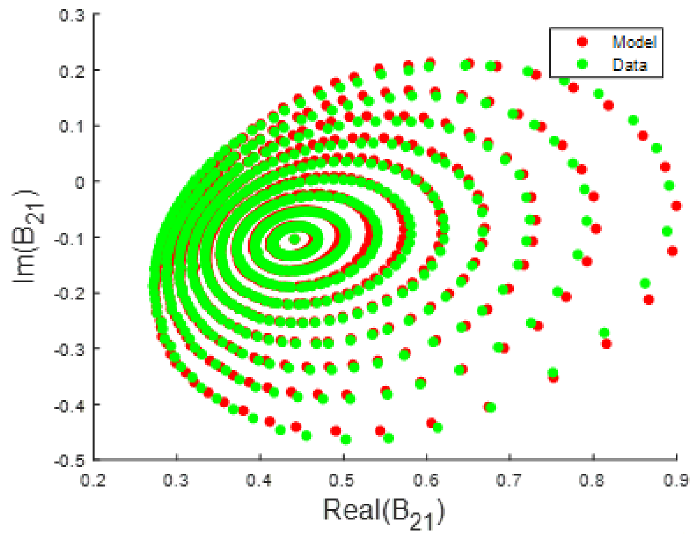
Figure 4.3: Credible band for $X^T_{21,11}$

The variance of the predicted probability distribution is considerably reduced by the increment in sample data points. Figures 4.1, 4.2 and 4.3 are created using 10 sample data points and 100 data points. The width of the credible region is reduced with the number of the sample data points. The variance of the probabilistic model can be reduced via further measurement data if higher confidence in the predicted output is required.

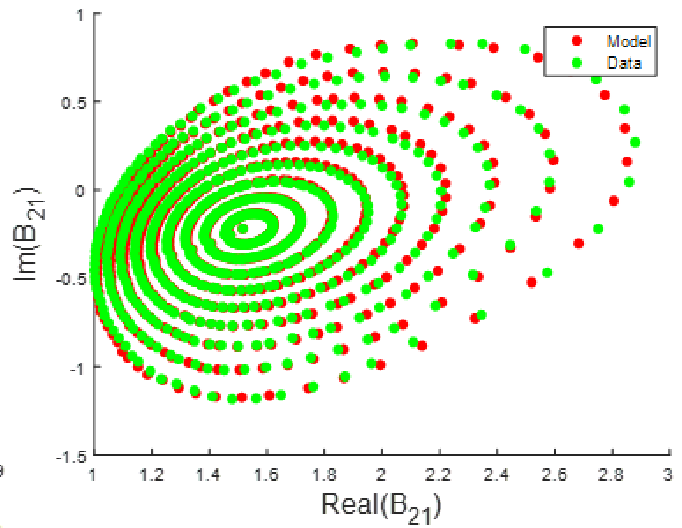
This model modelled model parameters with random errors that may occur in model parameter extraction. The designing of the system will be aided further if the output responses with random errors are modelled and the proposed model is giving a probabilistic distribution for model parameters not for output responses of the DUT.

4.2 Modelling of scattered waves with random errors

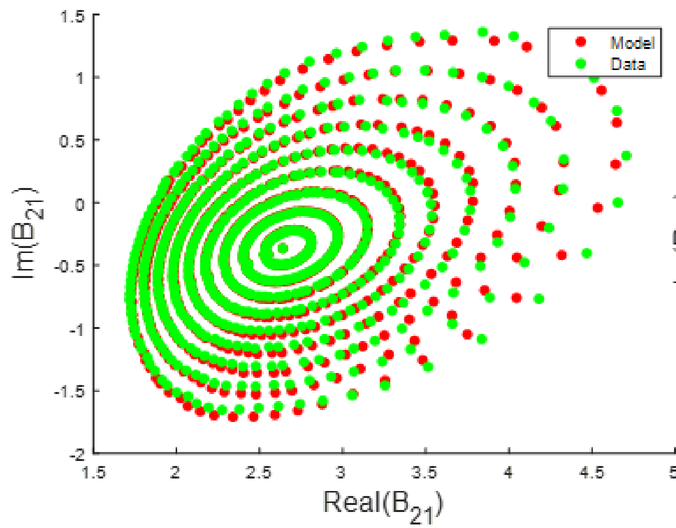
This section discuss the results of probabilistic behavioural model developed by proposed method which gives a probability distribution for scattered waves (output responses). Experimental data is used in this section. A two-port 10W RF GaN transistor (CGH40010F) is the DUT used. The voltages and currents at two ports are measured at different load impedance. The incident and scattered waves are calculated using equations (1) and (2) in chapter 3 respectively. The data sets are measured at 4 different input powers (-26dBm, -16dBm, -10.6dBm and -5.1dBm). 2 GHz is the fundamental frequency used. Gaussian distribution with zero mean and precision parameter $\alpha = 0.005$ is considered for the prior model parameters distribution. Figure 4.4 shows the probabilistic model for the B_{21} values corresponding to the A_{21}values in the training data subset. The red-filled circle is the mean of the proposed probabilistic model (which depends on A_{21}), and the filled green circles are the experimentally measured data. Note that since we are assuming the data contain noise, we do not expect the model and measured data to be completely coincident in this plot. For a Gaussian model, the region extending to two standard deviations away from the predicted mean B_{21} value corresponds to a 95% credible area.



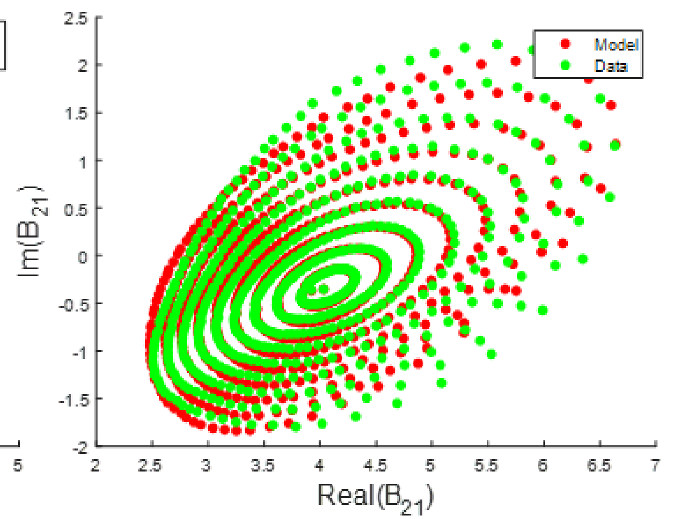
(a) Input Power =-26 dBm



(b) Input Power=-16dBm

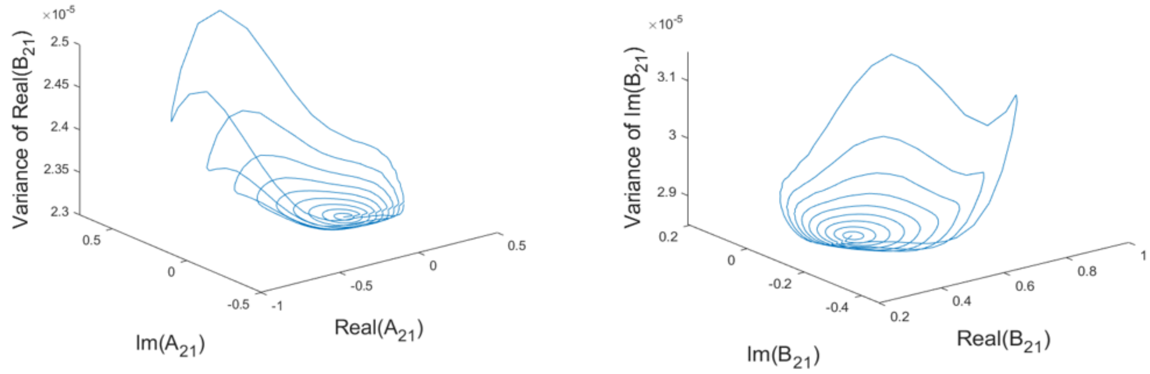


© Input Power=-10.6dBm

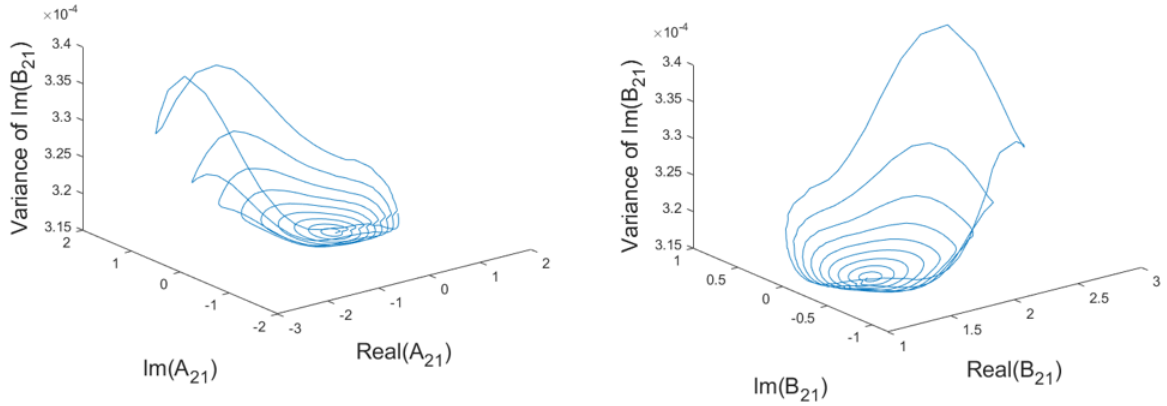


(d) Input Power=-5.1dBm

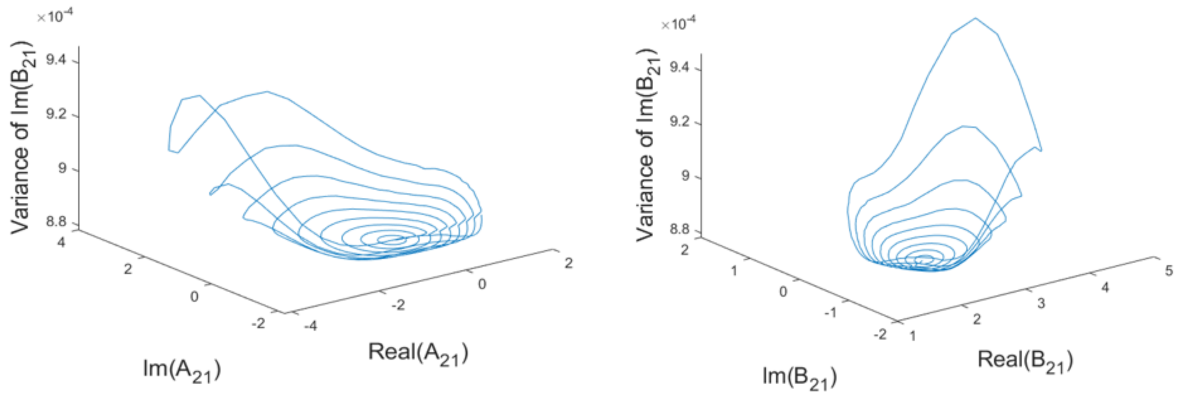
Figure 4.4: Probabilistic Model for B_{21} at different input power



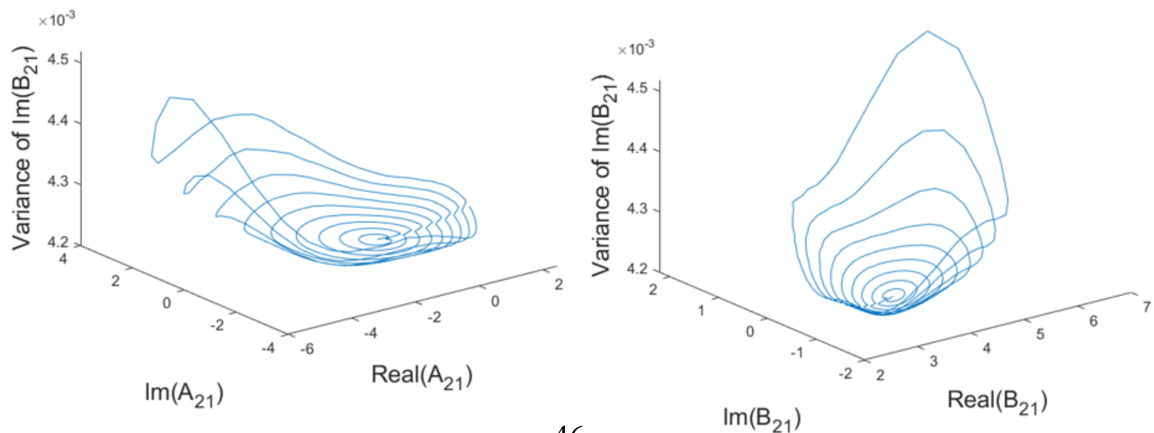
(a) Input power=-26dBm



(b) Input power=-16dBm



(c) Input power=-10.6dBm



(d) Input power=-5.1dBm

Figure 4.5: Variances at different input powers

The figures 4.4 a, b, c and d shows the probabilistic model at different input powers. It is observed that as the input power increases the true measured value is moving away from the mean value. The figures 4.5 a,b,c and d gives the variance of the predicted probability distribution for the scattered waves B_{21} at different input powers. It is observed that the variance increases with the input power.

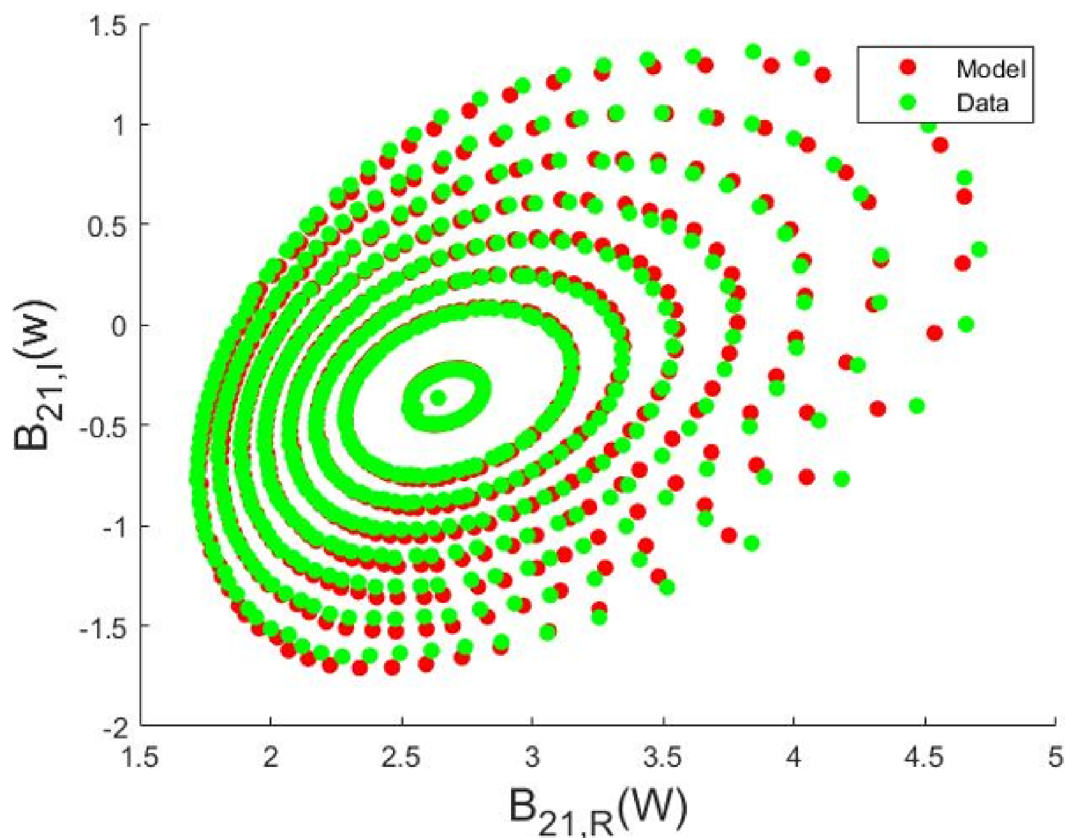


Figure 4.6: Probabilistic Model for B_{21}

For the validation purpose, the dataset is divided into training and testing dataset as in cross validation Technique. A block of sample points is taken as testing dataset from the overall dataset. Figure 4.6 represents the probabilistic model for the training dataset. This training dataset is used to predict the output response of the test dataset. Figure 4.7 shows the probabilistic model obtained for B_{21}values corresponding to the withheld A_{21}values i.e. these measurements are not part of the training data and are therefore unknown to the model. It shows good prediction results for new A_{21}

values.

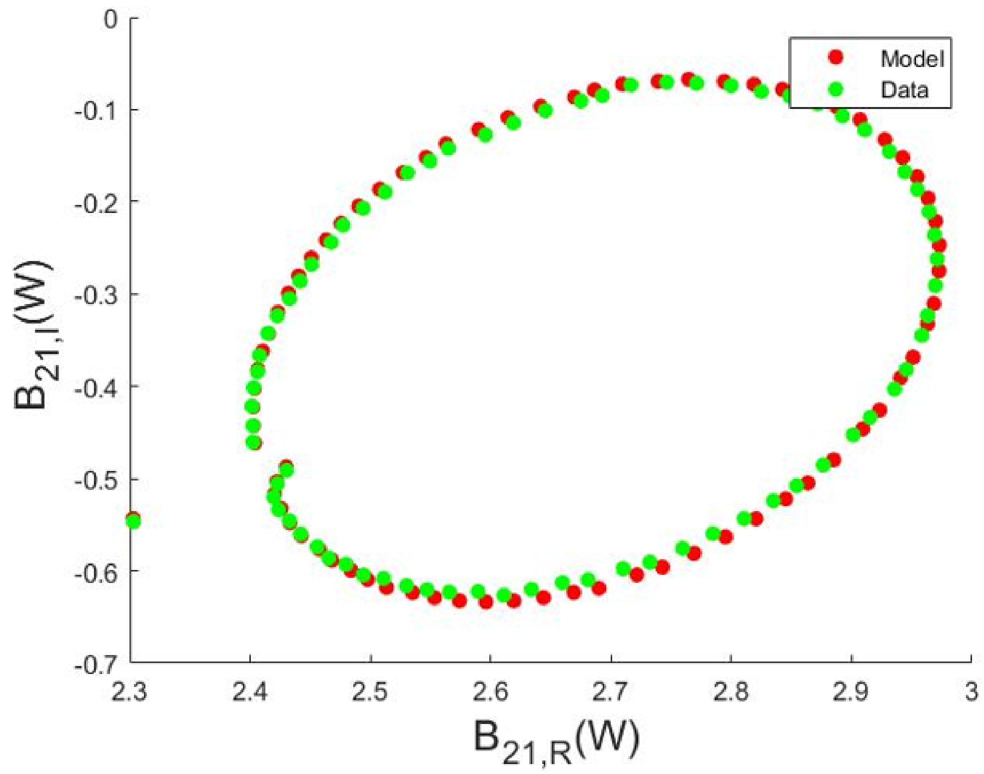


Figure 4.7: Probabilistic Model for new test points (new A_{21} values)

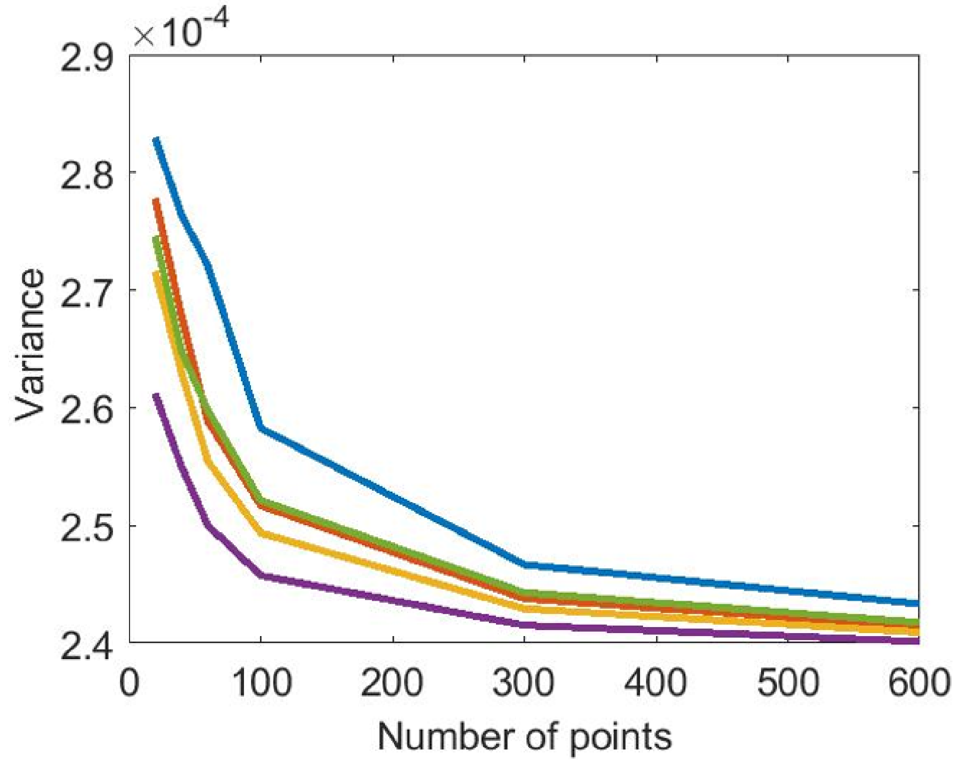


Figure 4.8: Variance of single data point Vs Number of Data Points

The change in variance of probability distribution with amount of training data is shown in figure 4.8. For that 5 data points of $B_{21,R}$ is randomly selected. The predicted probability distribution is estimated using proposed method with 20, 40, 60,100,300, and 600 training data points, respectively. It is observed that the variance of each data point of $B_{21,R}$ reduces as the amount of training data increases.

Figures 4.9, 4.10, 4.11 and 4.12 shows the sequential updates of the posterior distribution of the model parameters w_{2R} and w_{2I} as discussed in the validation methods section 3.3.3 in chapter 3. It is observed that the variance of the posterior distribution of the model parameters is reduced by the sequential update of training data. For the first time, a PHD model contains information regarding its own certainty.

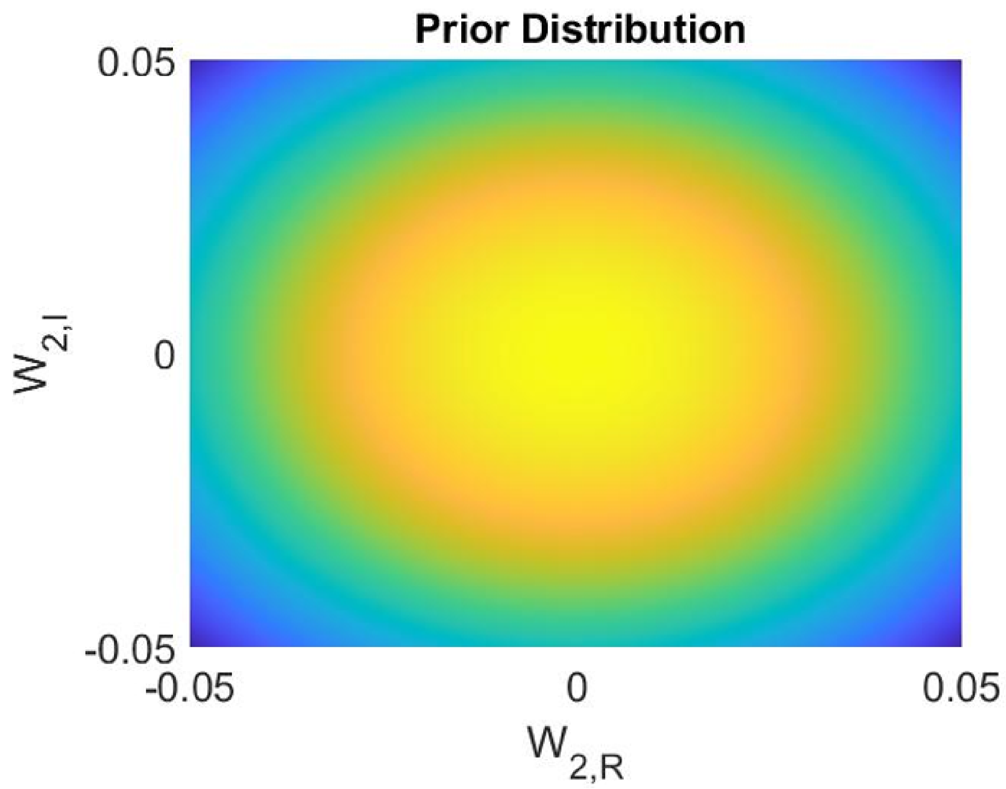


Figure 4.9: Prior Distribution

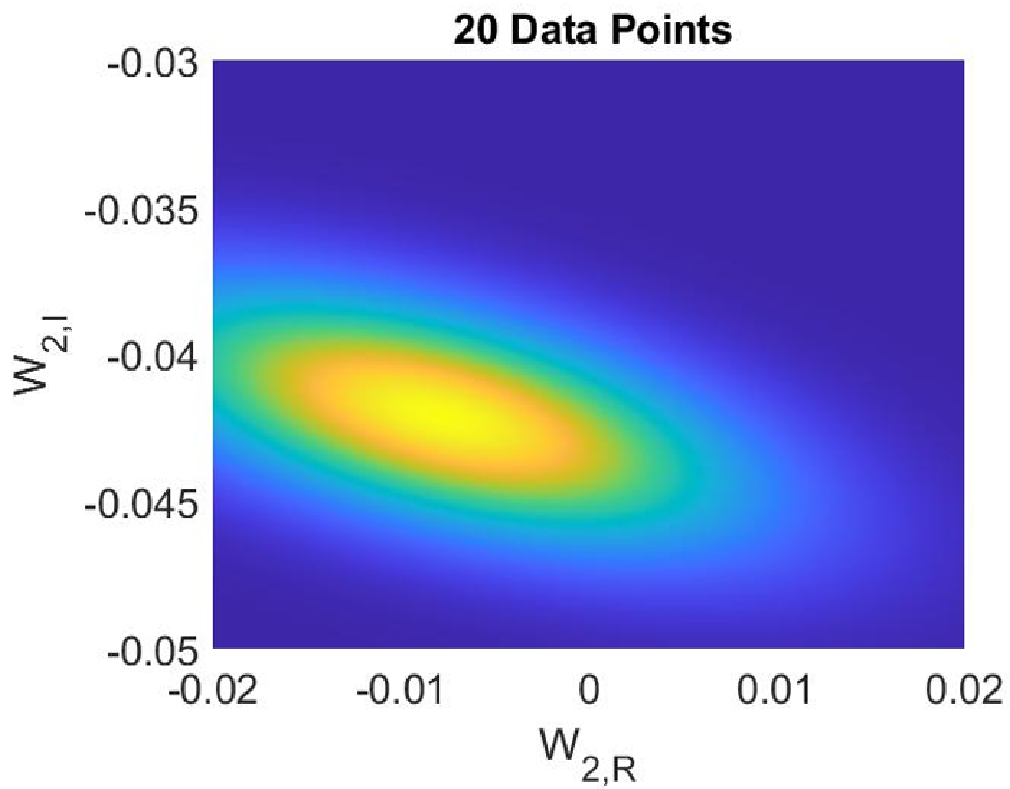


Figure 4.10: Posterior Distribution of w_{2R} and w_{2R} estimated using 20 Data points

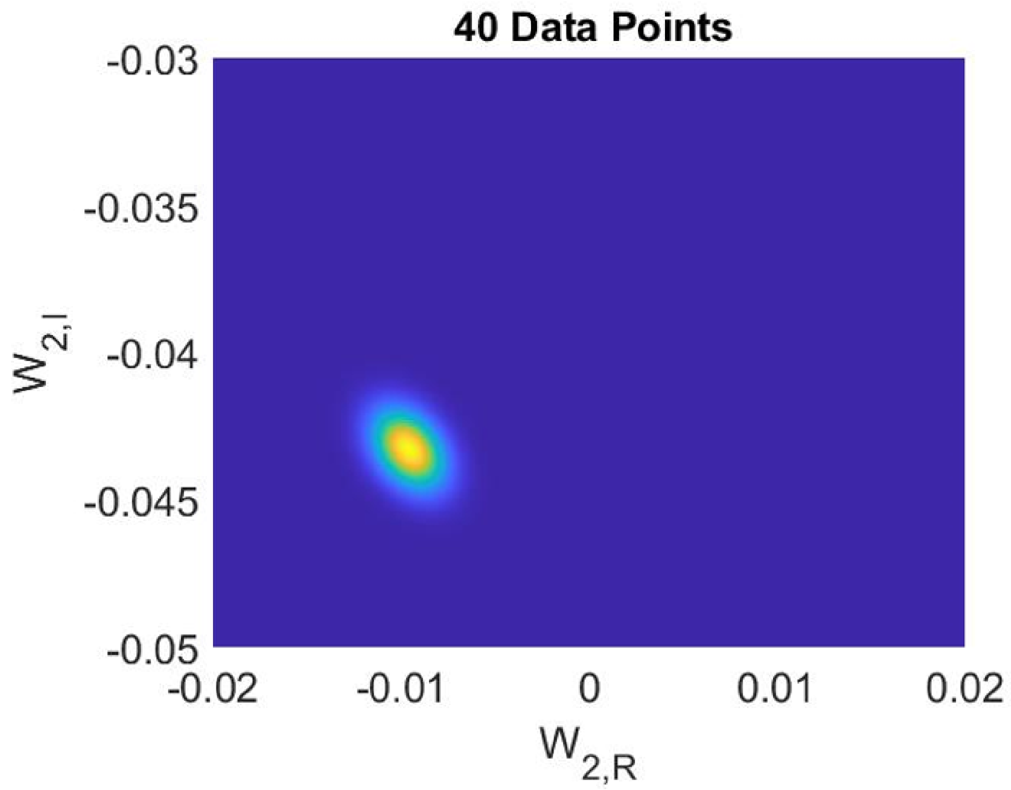


Figure 4.11: Posterior Distribution of w_{2R} and w_{2R} estimated using 40 Data points

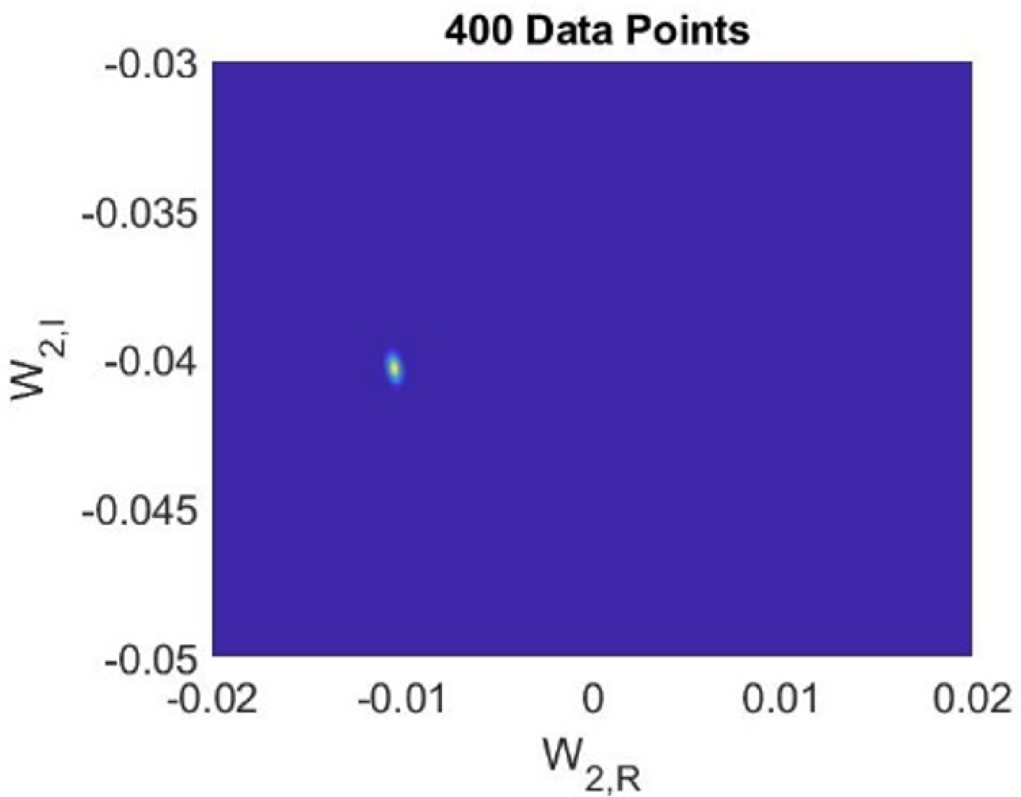


Figure 4.12: Posterior Distribution of w_{2R} and w_{2R} estimated using 400 Data points

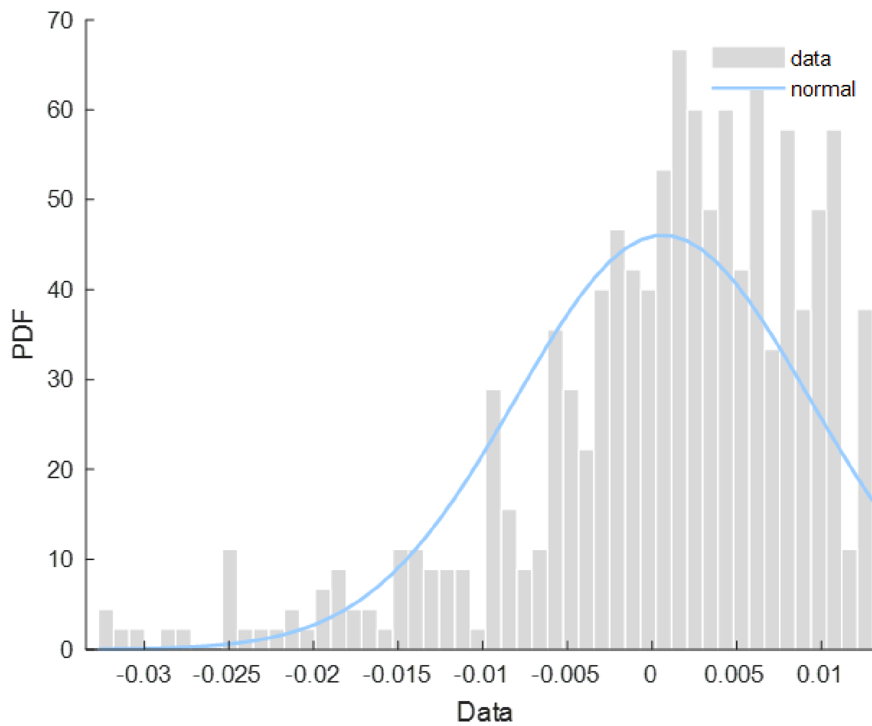


Figure 4.13: Distribution of retrieved noise

For the further validation of the proposed method, a Gaussian noise of 0 mean and 0.01 variance was added to the measured $B_{21,R}$ values. The mean of $B_{21,R}$ ($B_{21,R,mean}$) calculated using the above mentioned approach. Then the noise (N) was retrieved by subtracting $B_{21,R}$ from $B_{21,R,mean}$. Figure 4.13 shows the distribution of retrieved noise. It is also Gaussian Distribution. Figure 4.14 shows the change in the variance of retrieved noise with the variance of added noise. It is observed that the variance of retrieved noise is increasing with the increase in the variance of added noise.

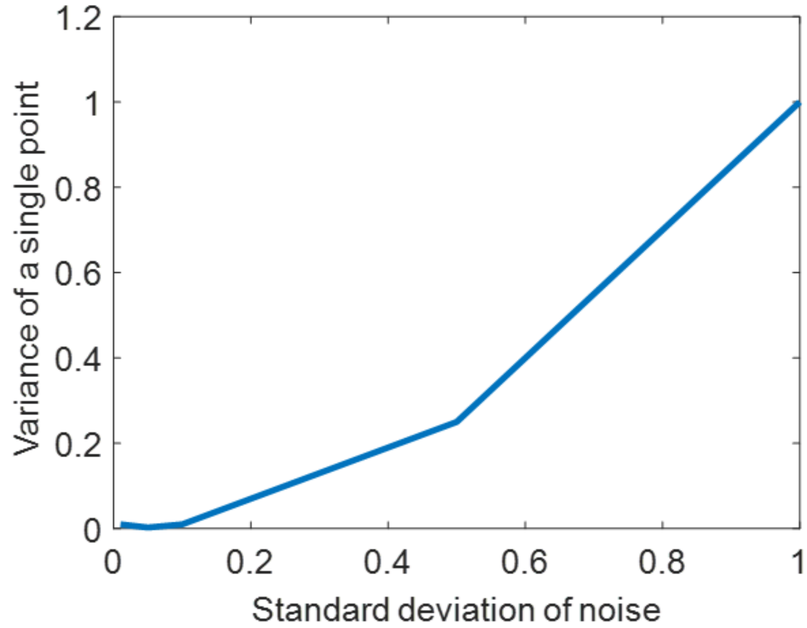


Figure 4.14: Variance of a single point Vs Standard deviation of noise

As mentioned before the existing models predict only point output responses for the system. These models neglect the possibility of random errors that may occur while measuring the system. The proposed model takes the possibility of random errors into consideration while predicting the behaviour of the system and gives probability distributions for the output responses. These probability distributions give the possible output responses and their probability of occurrences which is different from prevailing models which makes them incomparable with prevailing systems. The results in this chapter are the proof of concept of the proposed methodology.

4.3 Summary

In the first section of this chapter, the proposed model is used to create a probabilistic model that quantifies the random uncertainties in the model parameters. The Bayesian probabilistic approach is applied to the well-known X-parameter modelling approach. In this work, probability distribution and credible region are predicted for each X-parameter. In the second section, a probabilistic behavioural model is built using the Bayesian approach. In this model, a probabilistic distribution is obtained for the out-

put scattered wave and model parameters of the probabilistic behavioural model. The theory used to develop the probabilistic model is validated.

5 Extension to Bayesian Probabilistic Neural Network

In the proposed methodology, non-linear polynomial functions which are fixed basis functions are used to relate the input stimuli and output responses. The linearity in the parameters helped in the analysis of the Bayesian approach. One of the limitations of the usage of fixed basis functions is the curse of dimensionality. The basis functions are fixed before observing the training data, the number of basis functions will increase rapidly with the dimensionality of the input data. In most cases the number of basis functions increases exponentially. The dimensionality of input data increases with the number of ports in the DUT. Consideration of harmonic frequencies along with fundamental frequency also enhance the dimensionality of input data.

Bayesian Neural Networks (BNN) use adaptive basis functions instead of fixed basis functions which have sigmoidal non-linearities. These adaptive basis functions can adapt the parameters of the network such that the regions of input space over which the basis functions vary corresponds to the data [6]. The real data sets have the property that the target variables have significant dependence only on a limited number of possible directions within the data space. Neural networks utilize this property of real datasets by choosing the directions in the input space to which the adaptive basis functions respond. BNN is utilised for solving non-linear regression problems [6]. This section is an attempt to advance the proposed methodology to build a probabilistic model based on BNN for non-linear devices with random errors to get rid of the

curse of dimensionality.

5.1 Bayesian Neural Network

The main difference between the traditional neural network and the BNN is in the weight estimates. The weights of the traditional neural network will be point estimates and the weights of the BNN will be probability distributions. Prior distributions will be assigned to the weights of BNN using our prior information and the posterior distributions will be estimated using the Bayes theorem. The predictive distribution of the output response is computed by integrating the posterior distribution over the weight space. Figure 5.1 is the schematic diagram of BNN.

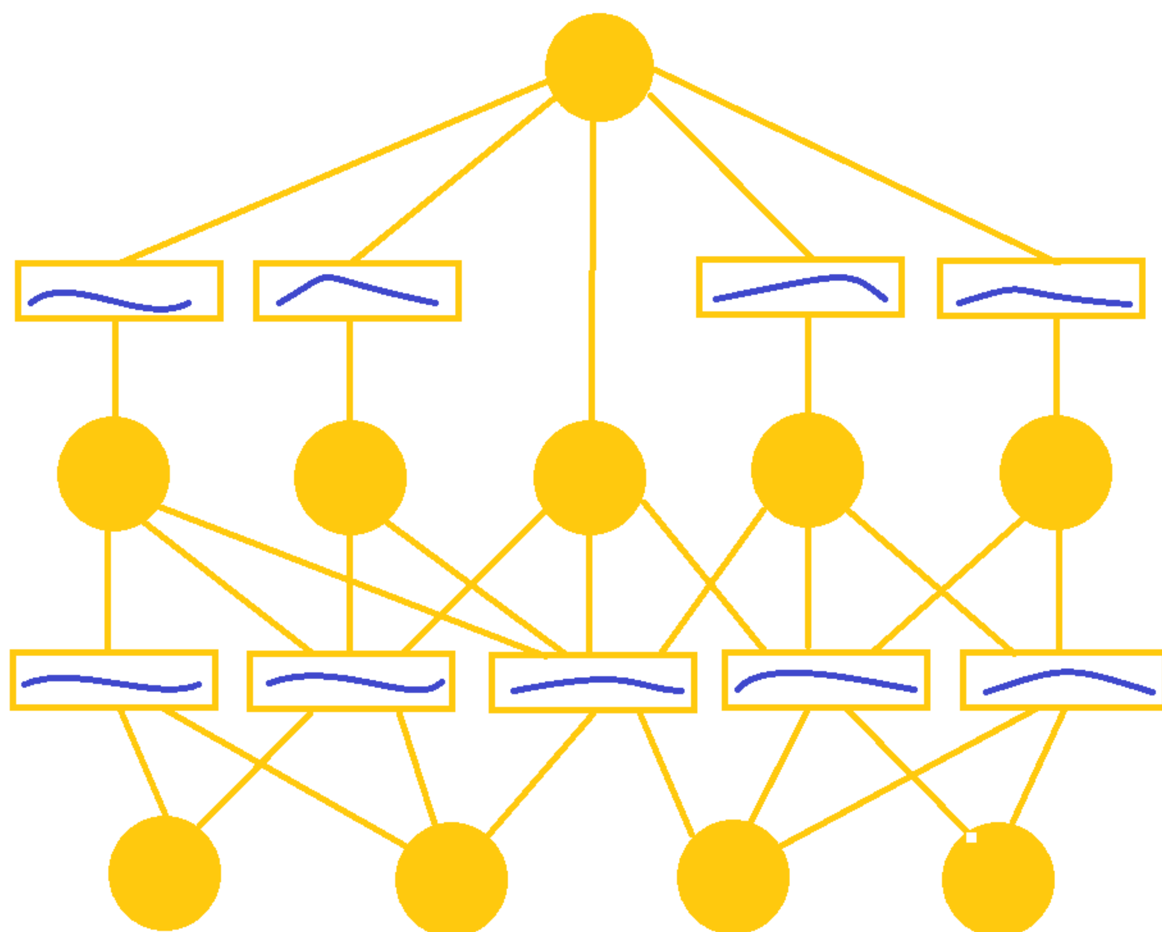


Figure 5.1: Bayesian Neural Network

Bayesian statistics cannot be applied exactly to the neural network due to the high

non-linearity of the network functions on its parameters. The negative log of the posterior distribution will not be convex due to the presence of multiple local minima in the error function. To overcome this limitation two methods are widely used in Bayesian neural networks to estimate predictive distribution for the output response: Variational Inference (VI) and Laplace approximation.

In this work, the scattered wave, for example, $B_{21,R}$ was related to the incident waves using a neural network. A 2-port device is considered as the DUT. The input vector of incident waves $[A_{11,R}, A_{11,I}, A_{21,R}, A_{21,I}]$ is denoted by $\mathbf{X}$ and the weights and biases vector of the neural network is represented by $\mathbf{W}_N$.

Presume that the conditional probability distribution $P(B_{21,R}|\mathbf{X}, \mathbf{W}_N)$ is a Gaussian distribution with mean as the output of the network function $y(\mathbf{X}, \mathbf{W}_N)$ and variance as β^{-1}.

$$\mathrm{P}(B_{21,R}|\mathbf{X}, \mathbf{W}_N, \beta) = N(y(\mathbf{X}, \mathbf{W}_N), \beta^{-1}) \tag{1}$$

A Gaussian distribution is given as prior distribution over $\mathbf{W}_N$ with 0 mean and precision α.

$$\mathrm{P}(\mathbf{W}_N|\alpha) = N(0, \alpha^{-1} I) \tag{2}$$

The likelihood function is assumed to be Gaussian and it is given by

$$\mathrm{P}(B_{21,R,t}|\mathbf{W}_N, \beta) = \prod_{n=1}^{N} N(P(B_{21,R,t,n}|y(\mathbf{X}_n, \mathbf{W}_N), \beta^{-1}) \tag{3}$$

Using the Bayes theorem the posterior distribution over weights and bias vector $\mathbf{W}_N$ can be computed.

$$\mathrm{P}(\mathbf{W}_N|B_{21,R,t}, \alpha, \beta) \propto P(\mathbf{W}_N|\alpha) * P(B_{21,R,t}|\mathbf{W}_N, \beta) \tag{4}$$

Due to the non-linearity of the network function over $\mathbf{W}_N$, the resulting posterior distribution will be non-Gaussian. A Gaussian approximation for the posterior distribution can be computed by Laplace approximation or variational inference.

In the Laplace approximation, two assumptions are made. The first assumption is that the posterior distribution is approximated to a Gaussian Distribution with a mean equal to the mode of the true posterior [6]. The other assumption is that the covariance is very small so that the network function can be approximated to have a linear dependence on the parameters (weights and biases) of the network over the parameter space where the posterior distribution is not zero. This method also requires the estimation of the Hessian matrix (second derivatives) which are computationally difficult.

In Variational Inference (VI) factorized Gaussian approximation is given to the posterior distribution. In this method, the approximate posterior distribution is obtained by minimizing the Kullback-Leibler divergence. This method is more convenient to compute the posterior distribution compared to the Laplace approximation.

5.2 Bayesian Probabilistic Neural Network Model

To overcome the limitation of curse of dimensionality a Bayesian Probabilistic Neural Network model is developed for non-linear devices which relates the input stimuli with output response with an Artificial Neural Network (ANN) and gives probability distribution for output responses. As before two separate models are needed for the real and imaginary parts of the complex-valued outputs. Tensorflow 2.3 and Keras software is used to model Bayesian Probabilistic Neural Network for the output response of the DUT.

5.2.1 Bayesian Probabilistic Neural Network Model for output response

In this section, an attempt to model the output response with random errors using the Bayesian Neural Network is done. The experimental data is used in this method. The DUT used in this work is a 10W RF GaN transistor (CGH40010F) which is a 2-port device. 1.5 GHz is used as the fundamental frequency. The incident waves and scattered waves are measured at various load impedances. The incident waves at port 1 and 2 (A_{11} and A_{21}) is related to the scattered wave at port 2 B_{21} using a neural network with 3 layers and 10 hidden units. The sigmoid function is used as an activation function. The Gaussian prior distribution of 0 mean and 1 variance is applied to the weights and biases of the neural network. The posterior distributions of the weights and biases are estimated using the Bayes theorem. The Gaussian approximate of the posterior distribution is computed using variational inference. As before 2 separate models are required for the real and imaginary parts of scattered waves.

Mean absolute error, mean square error, and root mean square error are some of the widely used metrics to determine the accuracy of the proposed model. The mean absolute error computes the mean of absolute errors between the true value and predicted value while mean square error computes the mean of squared error. Root mean square error is the square root of mean square error. Due to the squaring function in the mean square error, large weights are put on the errors. This makes the mean square error comparatively good in ensuring the accuracy of the predicted model. The other advantage of mean square error over mean absolute error is the differentiability of mean square error. In this work, mean square error metrics is used to estimate the accuracy of the predicted model.

649 observations are symmetrically divided into training (520), testing (65), and validation (64) datasets as in the equally spaced dataset discussed in chapter 3. The mean square error is high. The predicted values for the scattered waves and observed values for scattered waves are not close enough as expected. The same experiment is

repeated at different input powers (-23.5 dBm, -7.7 dBm, -2.2 dBm). The results are analysed. Table 5.1 gives the mean square error for the real and imaginary parts of the output response. The mean square error is increasing with the input power as expected.

Table 5.1: Mean Square Error

Input Power	**Output Response ($W^{1/2}$)**	**Mean Square Error (W)**
-23.5 dBm	$B_{21,R}$	0.0921
-23.5 dBm	$B_{21,I}$	0.0541
-7.7 dBm	$B_{21,R}$	0.2233
-7.7 dBm	$B_{21,I}$	0.2260
-2.2 dBm	$B_{21,R}$	0.4980
-2.2 dBm	$B_{21,I}$	0.3962

The difference between the true value and the predicted value is calculated. This error is very high. The distribution of the error (error distribution) is computed. Gaussian distribution is obtained as the error distribution. Table 5.2 shows the error distribution and parameters of the error distribution. It is observed that the error distributions are Gaussian. The variance of the error distribution is increasing with the input power as expected.

Table 5.2: Error Distributions

Input Power	**Output Response**	**Error Distribution**	**Mean**	**Variance**
-23.5 dBm	$B_{21,R}$	Normal	-0.2776	0.1452
-23.5 dBm	$B_{21,I}$	Normal	-0.1874	0.1528
-7.7 dBm	$B_{21,R}$	Normal	-0.3945	0.7004
-7.7 dBm	$B_{21,I}$	Normal	-0.3934	0.6171
-2.2 dBm	$B_{21,R}$	Normal	-1.0043	1.1359
-2.2 dBm	$B_{21,I}$	Normal	-0.5809	0.6969

For further investigation of unexpected results, gaussian noise with 0 mean and known variance is added to the measured data and modelled the output response. The mean square error is estimated for the predicted model. This is repeated with different noise variances. Table 5.2 gives the mean square error for the validation dataset at different noise variances. The measured dataset at -2.2dBm input power is used for the

illustration. The mean square error of the model is increasing with noise added as expected.

Table 5.3: Change in mean Square Error with the variance of noise added

Variance of noise added	**Output Response ($W^{1/2}$)**	**Mean Square Error (W)**
0.04	$B_{21,R}$	2.3608
0.04	$B_{21,I}$	0.7121
0.25	$B_{21,R}$	2.4694
0.25	$B_{21,I}$	0.7751
0.49	$B_{21,R}$	3.1677
0.49	$B_{21,I}$	1.1210

5.2.2 Possible Reasons for unexpected results

The possible reason for the unexpected results is the limited amount of experimental data. Neural networks give good results with large amount of observed data. In this work, only 520 data points are used for training purpose. This may limit the accuracy of the predicted model. The model may give good results if the experiment set up is redesigned to sample the data points with reduced step size. In this experiment, the data points are sampled symmetrically throughout different load impedance. If the step size (the difference between the two consecutive load impedance) is reduced, more data can be obtained. Another suggestion is to include the harmonic frequencies along with fundamental frequencies which also increase the experimental data.

The other reason for the unexpected results may be the over-fitting of data. In this model, 10 hidden units are used in the hidden layer of the BNN. This may cause the over-fitting of the data. The model may be tried to fit all the data points due to the large number of hidden units. As this data contains noise, the model may not be able to identify the right trend. Better accuracy may be achieved by the right selection of the number of hidden layers and hidden units or regularization of the model.

The other possible reason for the unexpected results may be the absence of hyper-parameter optimization. Hyper-parameters are the parameters that determine the distribution of the weights and biases of the BNN [6]. In this case, α and β are

the hyper-parameters. A lack of hyper-parameter optimization may cause the non-regularization of the BNN model. This can be solved by improved model with hyper-parameter estimation method.

5.2.3 Summary

The limitations of the fixed basis functions are discussed. The BNN is explained and our attempt to quantify the random uncertainties in the model parameter extraction and output responses using BNN is discussed. The possible reasons for the failure of the BNN model have been explored at the end of this chapter.

6 Conclusion

Several behavioural models can effectively predict the non-linear characteristics of microwave devices. But as they are dependent only on measured data, their prediction will be affected by systematic and random uncertainties so that the DUT may behave differently from what it is expected to do. This may adversely affect the overall system and may not meet the requirements of the designers and manufacturers. As the causes or the origin of systematic errors can be known, they can be identified and rectified to a certain limit. But to model the devices with random errors, an efficient statistical procedure must be implemented. This research work has introduced an efficient and convenient statistical procedure to model the devices with random errors using Bayesian Statistics which gives probabilistic distributions instead of point estimates for model parameters and output responses.

The works provide probability distribution for model parameters and output values from which credible intervals or confidence regions can be calculated. Manufacturers and designers can use this method to analyse their design and manufacturing methods. If the credible region or confidence interval is within their acceptable range, they can move forward with the design or manufacturing method. If not they have to consider redesigning or collecting more data points for further verification.

6.1 Summary of the works

The probabilistic behavioural model based on X-parameters models the random uncertainties in the model parameter extraction. In that method, each X-parameters is

related to the incident wave with fixed polynomial basis functions, and a Bayesian approach is applied to the model parameters of the polynomial basis functions. As a result, a predictive probabilistic distribution is obtained for each X-parameters. This method is also capable of predicting the credible region for each X -parameter. This provides the designers be conscious of the possible variations from the extracted X-parameters along with the probability of the variations. So that the designers can pro-act even without the knowledge about the origin or cause of the uncertainties. The resulting probabilistic X-parameter model can then be used to determine probabilistic estimates for important circuit and system parameters, such as gain. Should the variation indicated by the model be greater than the desired variation, further measurements and/or improved circuit design, may be required [36].

The above method is giving a probability distribution for the model parameters, but it is not giving the variations that may occur in the output response of the DUT. An improved method is proposed to get a predictive probability distribution for the scattered output wave. In this model, the scattered wave is related to the incident wave using simplified quadratic polynomial basis functions. The quadratic polynomial basis functions were chosen as the quadratic terms help to predict the output response even in unmatched conditions. The application of the Bayesian theorem to the model parameters of the quadratic polynomial functions resulted in predictive probabilistic distributions for the scattered waves. This method is capable to model the output responses of the DUT in the presence of random uncertainties.

The limitation of the above method is that the number of basis functions grows exponentially with the dimensionality of the input space. This is due to the usage of fixed basis functions. This can be solved by using adaptive basis functions instead of fixed basis functions. Bayesian Neural Networks utilize adaptive basis functions. An attempt to develop a methodology to model the non-linear devices with random uncertainties using BNN is done. In that method, the output response of the DUT is related to the input stimuli using neural networks. The weights of the neural network are probability distributions instead of point estimates, which were estimated using

the Bayes theorem. But the result was not as expected. The predicted results were not close to the measured results.

6.2 Limitations of the work

One of the main limitations of this work is that the methodology to create a probabilistic behavioural modeling approach based on BNN is not developed. So that the curse of dimensionality still exists in the proposed method. The other limitation of this work is the absence of hyper-parameter estimation which causes non-regularization of the model. This may also lead to over-fitting of data. The limited amount of experimental data and low input powers used for modelling the output responses with random errors are also the shortcomings of this work.

6.3 Future Scope

In this work, a probabilistic modeling approach is introduced using fixed basis functions which are capable of modeling the non-linear devices with random uncertainties. Due to the use of fixed basis functions, the model is subjected to the curse of dimensionality. To overcome it, adaptive basis functions can be used. One of the options for that is to use BNN to relate output responses with input stimuli. Even if the attempt is unsuccessful, it can be reworked and fixed by analyzing the possible reasons for the failure. The other possibility to get rid of the curse of dimensionality is to use radial basis functions to connect the output responses with input stimuli of the non-linear devices. This methodology can also be improved by utilizing a large amount of experimental data. The large amount of data may also aids to model the systems with high input power.

In this work, the proposed method is applied to model the random uncertainties in the model parameter extraction in the X-parameters. The same method can be applied to model the random uncertainties in the model parameters of the other be-

havioural modeling approaches. This method can provide more reliability to the device models, which will aid the designing engineers in designing successful system designs.

Bibliography

[1] S. Shamsir, M. S. Hasan, O. Hassan, P. S. Paul, M. R. Hossain, and S. K. Islam, "Semiconductor device modeling and simulation for electronic circuit design," in *Modeling and Simulation in Engineering* (J. Valdman and L. Marcinkowski, eds.), ch. 11, Rijeka: IntechOpen, 2020.

[2] J. H. PaïDavid E.Root, Jan Verspecht and M. M. , *Xparameters Characterization, Modeling and Design of Nonlinear RF and Microwave Components*. USA: Cambridge University Press, 2013.

[3] G. M. B. Pichler and H. Arthaber, "A study on quadratic phd models for large signal applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 7, pp. 2514–2520, 2019.

[4] B. M. Jialin Cai, Justin B.King and T. Brazil, "Pade-approximation-based behavioral modeling," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 12, 2013.

[5] "Cardiff model lite datasheet," tech. rep., Mesuro Ltd, Pencoed Technology Park, Pencoed. CF35 5HZ, Oct 2013.

[6] P. M. Bishop, *Pattern Recognition & Machine Learningl*. Feb, 2068.

[7] J. Cai, J. King, C. Yu, and L. Sun, "Bayesian inference-based small-signal modeling technique for gan hemts.," *International Journal of RF Microwave Computer-Aided Engineering*, vol. 28, no. 8, p. 1, 2018.

[8] J. Hall, S. Ziada, and D. Weaver, "The ebers-moll bipolar junction transistor model,"

[9] H. G. . H. Poon, "An integral charge control model of bipolar transistors," *The Bell System Technical Journal*, May-June, 1970.

[10] designers guide.org, "Vbic description & derrivation details," 1999.

[11] J. P. PaïG. Niu, R. van der Toorn and W. Kloosterman, *The Mextram Bipolar Transistor Model version 505.00*. June 29, 2017.

[12] P. Schroter and A. Mukherjee, *HICUM -A Geometry Scalable Physics-Based Compact Bipolar Transistor model*. Dec, 2008.

[13] T. Patel, "Comparison of level 1, 2 and 3 mosfet's," tech. rep., Technical Report, University of Texas at Arlington, Dec 2014.

[14] virup Dasgupta and C. Hu, "Bsim-cmg compact model for ic cad: from finfet to gate-all-around fet technology," tech. rep., Technical Report,Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720, USA., 2020.

[15] virup Dasgupta and C. Hu, "http://bsim.berkeley.edu/models/bsimimg/," tech. rep., Technical Report,Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720, USA., 2020.

[16] A. Moulthrop, C. Clark, C. Silva, and M. Muha, "A dynamic am/am and am/pm measurement technique," in *1997 IEEE MTT-S International Microwave Symposium Digest*, vol. 3, pp. 1455–1458 vol.3, 1997.

[17] Verspecht and D. E. Root, "Polyharmonic distortion modeling," *IEEE Microwave Magazine*, vol. 7, no. 3, pp. 44–57, 2006.

[18] C. Y. H. L. J. L. J. Cai, J. Wang and L. Sun, "An artificial neural network based nonlinear behavioral model for rf power transistors," *IEEE Asia Pacific Microwave Conference (APMC)*, 2017.

[19] J. C. et al., “Bayesian inference-based behavioral modeling technique for gan hemts,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 6, 2019.

[20] S. D. Guerrieri, F. Bonani, and G. Ghione, “Linking x parameters to physical simulations for design-oriented large-signal device variability modeling,” in *2019 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 204–207, 2019.

[21] F. B. S. D. Guerrieri, C. Ramella and G. Ghione, “Global assessment of pa variability through concurrent physics-based x-parameter and electro-magnetic simulations,” in *2020 15th European Microwave Integrated Circuits Conference (EuMIC)*, pp. 213–216, 2021.

[22] L. T. Stant, M. J. Salter, N. M. Ridler, D. F. Williams, and P. H. Aaen, “Propagating measurement uncertainty to microwave amplifier nonlinear behavioral models,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 2, pp. 815–821, 2019.

[23] A. Arsenovic, L. Chen, M. F. Bauwens, H. Li, N. S. Barker, and R. M. Weikle, “An experimental technique for calibration uncertainty analysis,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 1, pp. 263–269, 2013.

[24] D. Ballo, “Applying error correction to network analyzer measurements.,” *MICROWAVE JOURNAL*, vol. 41, no. 3, pp. 64 – +, 1998.

[25] “Measurement errors and their characteristics,” tech. rep., Keysight Technologies.

[26] D. Schreurs, S. Liu, G. Avolio, and I. Ocket, “Impact of measurement uncertainty on modelling,” in *2016 21st International Conference on Microwave, Radar and Wireless Communications (MIKON)*, pp. 1–4, 2016.

[27] V. Teppati, S. Tirelli, R. Lövblom, R. Flückiger, M. Alexandrova, and C. R. Bolognesi, “Accuracy of microwave transistor extractions,” *IEEE Transactions on Electron Devices*, vol. 61, no. 4, pp. 984–990, 2014.

[28] D. F. Williams, R. A. Chamberlin, W. Zhao, J. Cheron, and M. E. Urteaga, "The role of measurement uncertainty in achieving first-pass design success," in *2016 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*, pp. 1–4, 2016.

[29] W. Wiatr and D. Walker, "Systematic errors of noise parameter determination caused by imperfect source impedance measurement," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 2, pp. 696–700, 2005.

[30] M. Lin and Y. Zhang, "Covariance-matrix-based uncertainty analysis for nvna measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 1, pp. 93–102, 2012.

[31] D. Ritzmann, P. S. Wright, W. Holderbaum, and B. Potter, "A method for accurate transmission line impedance parameter estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2204–2213, 2016.

[32] L. T. Stant, M. J. Salter, N. M. Ridler, D. F. Williams, and P. H. Aaen, "Propagating measurement uncertainty to microwave amplifier nonlinear behavioral models," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 2, pp. 815–821, 2019.

[33] P. S. Blockley, J. B. Scott, D. Gunyan, and A. E. Parker, "The random component of mixer-based nonlinear vector network analyzer measurement uncertainty," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 10, pp. 2231–2239, 2007.

[34] S. D. Guerrieri, C. Ramella, F. Bonani, and G. Ghione, "Global assessment of pa variability through concurrent physics-based x-parameter and electro-magnetic simulations," in *2020 15th European Microwave Integrated Circuits Conference (EuMIC)*, pp. 213–216, 2021.

[35] A. Zanobini, L. Ciani, and G. Pellegrini, "Quantifying the measurement uncertainty using bayesian inference," in *2007 IEEE International Workshop on Advanced Methods for Uncertainty Estimation in Measurement*, pp. 1–4, 2007.

[36] A. D. Manjaly, R. Sharma, and J. King, "Probabilistic behavioural model based on x-parameters," in *2020 IEEE Asia-Pacific Microwave Conference (APMC)*, pp. 119–121, 2020.