On Nonlinear Time-Invariant Behavioural Models of Power Transistors Used in the Computer-Aided Design of Power Amplifiers

by

Amir-Reza Amini

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Johannes Benedikt Deputy Director of School and Professor, School of Engineering, Cardiff University
Supervisor(s):	Slim Boumaiza Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
Internal Member:	Lan Wei Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
Internal Member:	John Long Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
Internal-External Member:	James Martin Professor, Dept. of Physics and Astronomy, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The Radio Frequency (RF) Power Amplifier (PA) is the main consumer of power in a wireless transmitter. Energy efficient PA design aided with circuit simulation tools requires accurate nonlinear models of the power transistors that lie at the heart of the PAs. This thesis proposes a novel methodology for extracting and implementing power transistor behavioral models from load-pull measurements. These models provide a valuable design aid to power amplifier designers looking to simulate the nonlinear behaviour of their RF circuit designs based on nonlinear characterizations of the power transistors.

Two types of power transistor behavioural models are proposed in this work. The first type is called the time-domain poly-harmonic distortion model (TD-PHD) and it targets the nonlinear multi-harmonic response of power transistors at a fixed fundamental frequency. This type of model allows the PA designer to simulate how the harmonic impedances of their designed RF matching networks effects the large signal behaviour of the PA. The TD-PHD model is shown to be able to replicate the time-domain waveforms of a power transistor under multi-harmonic source and load-pull characterization.

The second model is a generalization of the first model to target a set of non-uniformly spaced fundamental frequencies and is called the time-domain multi-tone distortion model (TD-MTD). Time-domain multi-tone distortion models that are extracted from load-pull measurements spanning multiple carrier frequencies are shown to recreate the load-pull performance contours of interest to PA designers. As a demonstration of TD-MTD models, two distinct behavioural models for the main and peaking transistors of a two-way Doherty PA design are extracted from load-pull measurements and the resulting PA design is then simulated and shown to accurately reflect the measured performance of the fabricated PA as a validation of the usefulness of this modelling methodology for high power amplifier design.

Acknowledgements

I am forever grateful for the guidance and kindness of the late Professor Safieddin Safavi-Naeini who was an original member of my PhD examination committee. He will be greatly missed by the many he touched in life.

I would like to thank all the efforts of my PhD research supervisor, Professor Slim Boumaiza, who always supported his students. I would also like to give thanks to the members of my examination committee, Professor Lan Wei and Professor James Martin for their guidance from the beginning stages of this research. In addition, I would like to thank Professor John Long who graciously agreed to join my PhD examination committee.

Throughout my graduate studies at the University of Waterloo, I learned a lot of what is presented in this thesis from fellow graduate students Dylan Bespalko and Hassan Sarbishaei.

Last but not least, I would like to acknowledge all who helped me with proof-reading my conference and journal papers and this PhD thesis, including Saeed Rezaei, Igor Acimovic, Carl Conradi, and Rob Salmond.

Dedication

To Katie, my unwavering source of love and support,

Your love has propelled me forward, filling each day with motivation and hope. Thank you for standing by my side, even during the most challenging times.

To Daisy and Bear, my loyal companions,

Your wagging tails, playful antics, and unconditional love offered me solace and comfort when the pressures of research weighed heavily on my shoulders.

To Katie, Daisy, and Bear, together you have been the pillars of strength in my life. This thesis stands as a testament to the power of love, support, and unwavering companionship. I am forever grateful for each of you, and I dedicate this work to you with boundless love and appreciation.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis will aim to provide power amplifier designers with a methodology to develop simulation models for the power transistors they use in their designs. These models will be based on Radio Frequency (RF) measurements of the power transistors. These measurements are performed in similar electrical conditions to how the power transistor will operate when it is functioning as part of a power amplifier design. The objective of this research was to propose a methodology to produce nonlinear models of a power transistor from the RF measurements of those devices. These models are called behavioural models, since they only model the RF behaviour of the device as measured, directly.

The large signal characterization of a high power transistor involves the RF measurement of the power transistor over varying loads and bias conditions, requiring calibrated narrow-band pulsed-RF measurements at different power levels and at different frequencies. The datasets produced from power transistor RF characterization can often get very large. One of the main objectives of this research was to find the simplest form a behavioural model should have that can both capture the nonlinear dynamics of the power transistor and uniquely fit the behaviour of the device over the entirety of the measurement dataset. We tried to avoid developing a modeling strategy that relies on using a *guided characterization* on a fixed measurement grid in order to extract the different model parameters, and instead focused on being able to generate models from arbitrary spaced and randomly-generated load-pull data. The incentive here is to develop a methodology to convert already existing load-pull data of RF power transistors into behavioural models that can be used to design RF power amplifiers in the circuit simulator.

While all the nonlinear behavioural models that have been proposed in the literature have been described in the frequency-domain, that is, they describe the behaviour of the

signal at the device ports at discrete frequencies, our research found that a simpler model construction can be used to describe the nonlinear behaviour if a time-domain model is used. This time-domain modeling approach lends itself pretty well to future generalization to multi-tone and modulated signal stimulus as the analysis and simulation of the long-term memory-effects rely heavily on time-domain methods. Another aspect of power transistor modeling that this research set out to tackle was the development of models that can track the so-called *hard nonlinearities* of RF power transistors without using look-up tables, while providing a robust converging model in a nonlinear circuit simulation environment.

There are two main contributions in this thesis. The first contribution is the proposal of Time-Domain Poly-Harmonic Distortion (TD-PHD) models. These models were inspired by the state of the art poly-harmonic distortion models that were described in the frequency-domain, but were reformulated to be described in the time-domain. Poly-harmonic distortion models target the multi-harmonic behaviour of a power transistor at a fixed fundamental frequency. Part of this contribution is the methodology to post-process the measurement dataset in order to prepare it for the proposed behavioural model extraction. The other part of this contribution is the implementation of the extracted model in a harmonic balance circuit simulation, an important type of nonlinear RF circuit simulation used for power amplifier design. The second contribution of this thesis is the proposal of Time-Domain Multi-Tone Distortion (TD-MTD) models. These models are a generalization of the TD-PHD models to allow for a model extraction from measurement data spanning a non-uniformly spaced frequency grid. This generalization allows the power amplifier designer to take their load-pull data spanning multiple frequencies as is, and fit a single TD-MTD model to the load-pull data such that the simulation model behaves the same way in the harmonic balance simulation of the power amplifier circuit as it does in the measurement of the fabricated power amplifier. With the TD-MTD model (and also similarly for the previous TD-PHD model) this end is achieved without imposing any additional restrictions on the measurement dataset used to extract the model, that is, a model can be extracted from measurements that are irregularly spaced in the measurement space.

In this thesis, a demonstration of the TD-MTD model for the design of a two-way Doherty power amplifier is used as a validation of the large-signal design application of the proposed behavioural model. The comparison of the simulation and the measurement was also done for the same power amplifier design with a compact model available from the power transistor vendor and it is shown that the prediction of the TD-MTD model for the Doherty power amplifier performance is no worse than a compact model. It should be stressed that the purpose of behavioural models is not to replace compact models, but to fill in the modeling gap when compact models are not available. The device used for the

demonstration in this thesis was a mature LDMOS RF power transistor device with a mature compact model. Often power amplifier designers are working with power transistors that are recently designed or are from vendors that don't have readily available compact models for their power transistor devices. The availability of load-pull measurement platforms allows power amplifier designers to fully characterize the RF nonlinear performance of their power transistor devices at the frequencies of interest, in addition to the ability to use this captured data to create a simulation model that can be used to design a power amplifier around that power transistor.

In the implementation of both the TD-PHD and TD-MTD models presented in this thesis, artificial neural networks (ANNs) are used to implement the multi-variate nonlinear function at the heart of the model. While non-ANN versions of the TD-PHD and TD-MTD model were theorized and implemented in the work towards this thesis, the ANN implementation that is mainly presented in this thesis allows the tackling of the *hard nonlinearities* of the RF power transistor, and the structure of the smooth bounded activation functions of the artificial neurons that make up the artificial neural network allow for the generation of a model that has robustly converges to the solution within the iterations of the nonlinear circuit simulator. The contributions of this thesis include use of ANNs to tackle the hard nonlinearities, but also the structure of the model, the time-delays between the auxiliary signals used in the model, the implementation of the models in the frequency-domain simulator, and how the measurement data has to be post-processed for model fitting.

The models proposed in this thesis are extracted from load-pull measurements of RF power transistors. These load-pull measurements can either be real measurements of physical power transistors, or it can be simulated load-pull measurements. A compact model can be simulated in a harmonic balance simulator and the simulation result can be stored as load-pull data that can be used to extract a behavioural model. Using this method, one can extract a behavioural model to fit the simulated large signal behaviour of a compact model. In this technique the compact model will be the device under test, and the harmonic balance simulator will emulate the load-pull measurement of the power transistor device. One reason to use this method when evaluating behavioural models is because the harmonic balance simulation of a compact model is deterministic and if a behavioural model can reproduce the same load-pull measurement result as the compact model when the behavioural model itself is put as the device under test in a simulated load-pull measurement, then it shows that the behavioural model is at least capable of mimicking a nonlinear device with the complexity of compact models. Of course the next level of validation, after the extraction of simulation models from measurements of power transistors, is the accurate simulation of a power amplifier design that incorporates these power transistors. A good

model will be able to provide a good simulated prediction of the RF measurement of the physically built power amplifier design.

In Chapter 2 the state of the art of power transistor modeling and simulation will be presented, starting with a background on the application of computer-aided design of power transistors followed by an explanation of the physics of the power transistors and the different nonlinear and dynamical effects involved in power transistor operation. Then an overview of the compact modeling of power transistors will be provided as it is the main alternative for power transistor models to the behavioural models presented in this thesis. Next, the harmonic balance simulation, the main type of nonlinear frequency-domain circuit simulation that is used by power amplifier designers will be discussed. Parallel to the simulated environment, the large signal characterization techniques used to measure the large signal behaviour of the power transistor, including the type of RF stimulus and various load-pull measurement and calibration techniques will then be discussed. The nature of the frequency-domain simulation in harmonic balance and the frequency-domain measurement capture for load-pull characterization of power transistors allows for the ability to define behavioural models that describe the spectral scattering purely in the frequency-domain. The frequency-domain poly-harmonic distortion models that have been presented in the literature will thus be covered next. Finally a discussion of the limitations of the behavioural models of power transistors presented in the literature will motivate the model proposed in Chapter 3.

In Chapter 3, the time-domain poly-harmonic distortion model, the first contribution of this thesis will be presented. First an intuitive argument for the TD-PHD model will be presented that doesn't require much advanced mathematics but will convince the reader that the structure of the time-domain model makes sense and is indeed the most compact form the model could take to capture the nonlinear poly-harmonic dynamics of interest. Then a more sophisticated argument will be presented that presents the TD-PHD model as a natural extension of finite impulse response models to nonlinear systems, essentially viewing the TD-PHD model as a nonlinear impulse response model. Next the process to extract the TD-PHD model from multi-harmonic load-pull data and also the process of implementing the extracted model in a harmonic balance simulator will be described. Finally a measurement-based validation of the prediction of the multi-harmonic behaviour of a power transistor at a fixed DC bias and fundamental frequency of a 10W GaN power transistor is demonstrated based on a randomly generated set of active load-pull measurements. The prediction of the model for passive load-pull measurements that were not used in the training set were used as a validation for the predictive capability of the TD-PHD model for modeling the multi-harmonic load-pull measurement data.

In Chapter 4, a generalization of the models of Chapter 3 called the time-domain multi-

tone distortion (TD-MTD) models will be proposed. This generalization of the model is to account for the fact that power amplifier designers often have load-pull measurement data at multiple fundamental frequencies while the TD-PHD model as it was proposed was fixed to a single fundamental frequency, requiring the power amplifier designer to extract a distinct TD-PHD model for each fundamental frequency, even though the underlying nonlinear behaviour, as represented by the hypothetical Volterra-series description of the power transistor is constant. Through a Volterra series projection formulation, the TD-PHD model will be seen as a special case of the more general TD-MTD model that can account for a non-uniformly spaced frequency grid as the basis of the time-domain behavioural model. Finally a measurement-based validation of the TD-MTD model will be presented by first showing that a single time-domain power transistor model can fit the load-pull data extracted at multiple fundamental frequencies. Then a TD-MTD model at a class AB bias will be extracted for a Main power transistor device and another TD-MTD model at a class C bias will be extracted for a Peaking power transistor device in a Doherty power amplifier configuration. The resulting power amplifier with the two extracted nonlinear power transistor behavioural models will then be shown to model the large signal behaviour of the power amplifier comparably to how the compact model of the power transistor models the same behaviour, showing that behavioural models, particularly TD-MTD models, can be used to design power amplifiers solely based on large-signal measurements of power transistor devices.

Finally in Chapter 5 the thesis will be concluded with an overview of the main contributions of the thesis and a discussion of future directions that are motivated by the findings of this research.

Chapter 2

State of the Art of Nonlinear Power Transistor Modeling

In this chapter, an overview of the application of developing radio frequency (RF) power transistor models and using them to design RF power amplifiers will be presented in Section 2.1. Further, the underlying physical phenomenon that leads to the behaviour exhibited by power transistors will be explored in Section 2.2. Next, the state of the art of the varieties of power transistor models used by power amplifier designers will be discussed, starting with compact models of power transistors in Section 2.3. Another type of transistor model that will be discussed in this chapter are behavioural models of power transistor that model the characterized behaviour of the power transistor directly. The large signal characterization methods used to measure the behaviour of power transistors for the purpose of behavioural model extraction is discussed in Section 2.5. Next the most commonly used nonlinear RF behavioural model called the poly-harmonic distortion (PHD) model will be presented in Section 2.6. Finally the limitations of the PHD model will be discussed in Section 2.7 to motivate the power transistor behavioural models proposed in this thesis.

2.1 Background

Mobile wireless communication requires a wireless network that provides the mobile users of the network a way to send information via an uplink path and a way to receive information via a downlink path. In the downlink path, the base-station radio that is connected to the wider network through high-speed fibre-optical communication, transmits the downlink

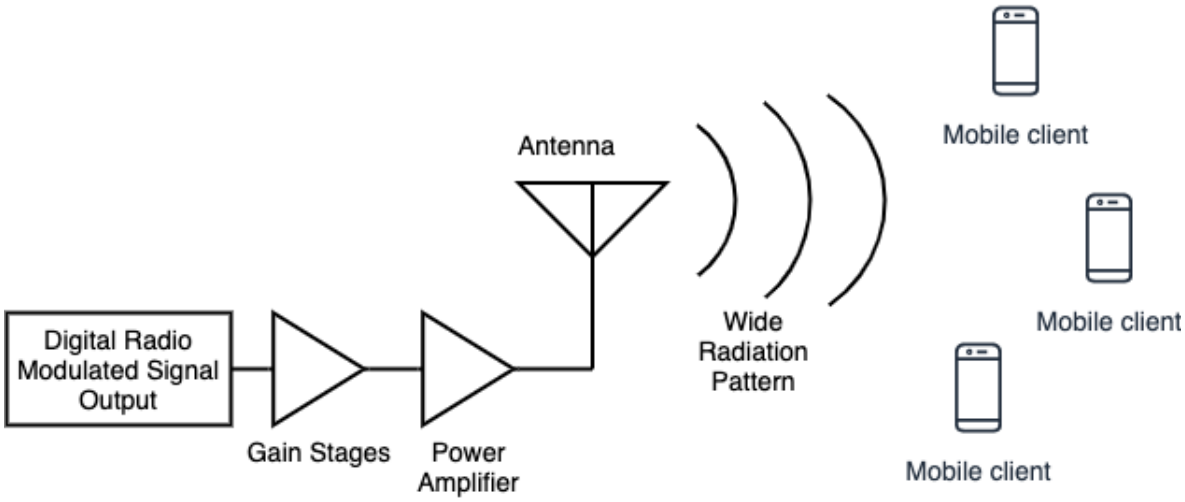

Figure 2.1: Simplified Downlink Path of a Mobile Wireless Communication Network

signal through an antenna with a wide radiation pattern in order to target the mobile users that are scattered across a wireless coverage area. In order to broadcast a downlink signal that has the information of many users over a wide region of space, the downlink broadcast antenna requires a signal of high power. To produce a high power signal from a low power digitally generated modulated signal, many amplification stages are typically required, and the final stage of amplification that boosts the power of the information-carrying modulated signal enough to be transmitted over the antenna is called the power amplifier. Figure 2.1 shows a typical simplified diagram of the path that the downlink signal takes to get to the mobile users.

Today, existing and emerging radio standards (5G, 6G, and so on) will push the requirements of spectrum efficiency, linearity, quality of service and power efficiency, targets that are often-times in conflict with one another. The radio downlink is the part of the communication system that consumes the most amount of power and at the same time the power amplifier is the most nonlinear component in the downlink communication chain. If the nonlinearity of the power amplifier is not addressed, the spurious emissions that are generated as a result of the nonlinearity will violate the requirements enforced by spectrum regulatory bodies. Through the use of linearization techniques like Digital Pre-Distortion that can correct for the spectral regrowth of power amplifiers in adjacent bands, the design of highly power-efficient power amplifiers that are quite nonlinear in their operation becomes an attractive approach in the design of the downlink power amplifier.

This means that the circuits designed by the power amplifier designer will be designed to

be nonlinear and in fact the nonlinear operation of the circuits is used to make more efficient RF power amplifiers. Providing tools that would allow the power amplifier designer to tune the nonlinear performance of their design over frequency in a simulation environment from models of the nonlinear components of the circuit can be greatly beneficial in the circuit design progress and is indeed the goal of this research.

Since the source of nonlinearity in a power amplifier circuit are the nonlinear power transistors, the next section will discuss this nonlinear behaviour that is exhibited by the power transistor and illuminate the inherent complexity involved with modeling the behaviour of an RF power amplifier.

2.2 The Nonlinear Behaviour of a Packaged High Power Transistor

Power amplifier designers can use computer-aided design (CAD) tools to simulate the performance of their designs before they manufacture their circuits. To design power amplifier circuits in a simulation environment, an accurate power transistor model is required.

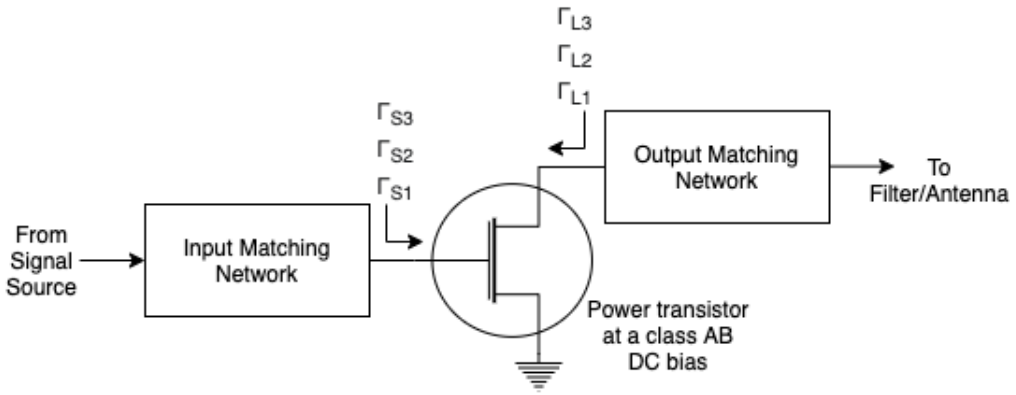

Figure 2.2: Single Transistor Power Amplifier

A power amplifier design can include a single power transistor like in Figure 2.2 or could have two power transistors like the two-way Doherty power amplifier structure of Figure 2.3 or have even more transistors in an N-way Doherty or other efficiency enhancement power amplifier topologies with multiple transistors like the Load Modulated Balanced Amplifier.

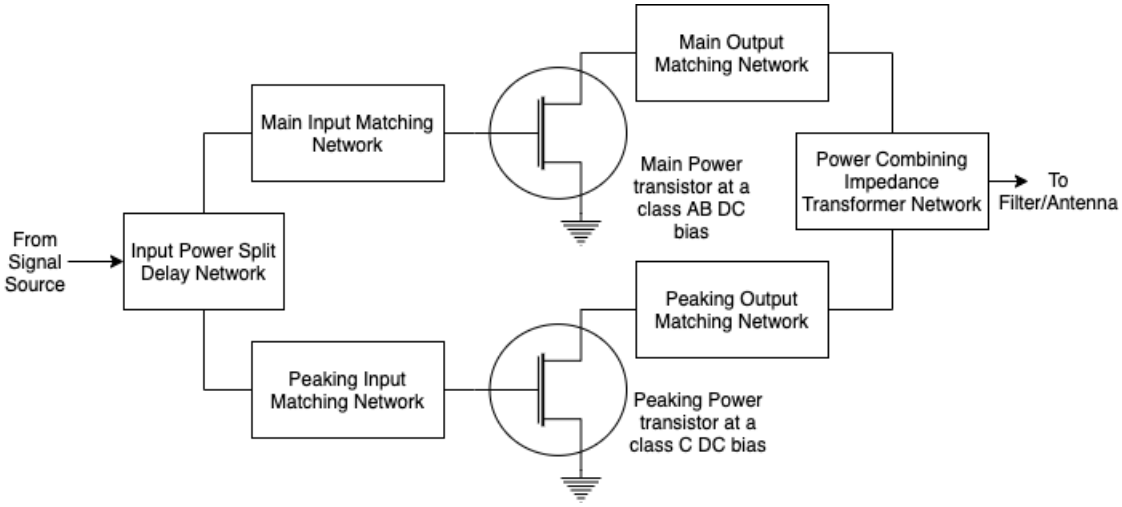

Figure 2.3: Two Transistor Doherty Power Amplifier

To make an amplifier, an actuated device like a transistor as shown in Figure 2.4 is biased by applying a DC gate electrical potential V_{GS} between the gate and source leads of the device and then applying the DC drain electrical potential V_{DS} between the drain and source leads of the device. Figure 2.5 shows the DC relationship between the current that flows between the drain and the source as a function of both the gate to source and the drain to source electrical potential for an ideal zero threshold voltage field effect transistor device. When the power transistor is biased in the saturation region, the output characteristic of the transistor and its drain current in particular, becomes mainly a function of the actuating gate voltage. This allows an application where low power signals incident at the gate can be effectively amplified into high power signals generated at the drain. Figure 2.6 shows the relationship between the gate to source voltage and the drain to source current for a fixed drain to source voltage. The RF signal behaviour of the power transistor is heavily dependent on where the transistor is biased on this curve. Figure 2.7 shows how the RF amplification of a power transistor changes based on its classes of operation with the example of an ideal current conduction transfer function used for a transistor, as the conduction angle of a hypothetical sinusoidal wave input is different for each case [1]. In this ideal case, the conduction angle represents the angular duration between 0° and 360° that the drain will be conducting current when the input of the transistor is incident with a sinusoidal RF signal. In each case the power transistor is biased at a different offset relative to the threshold gate voltage. While a class A bias provides the most linearity with a conduction angle of 360°, it is also the least efficient bias as it is consuming power

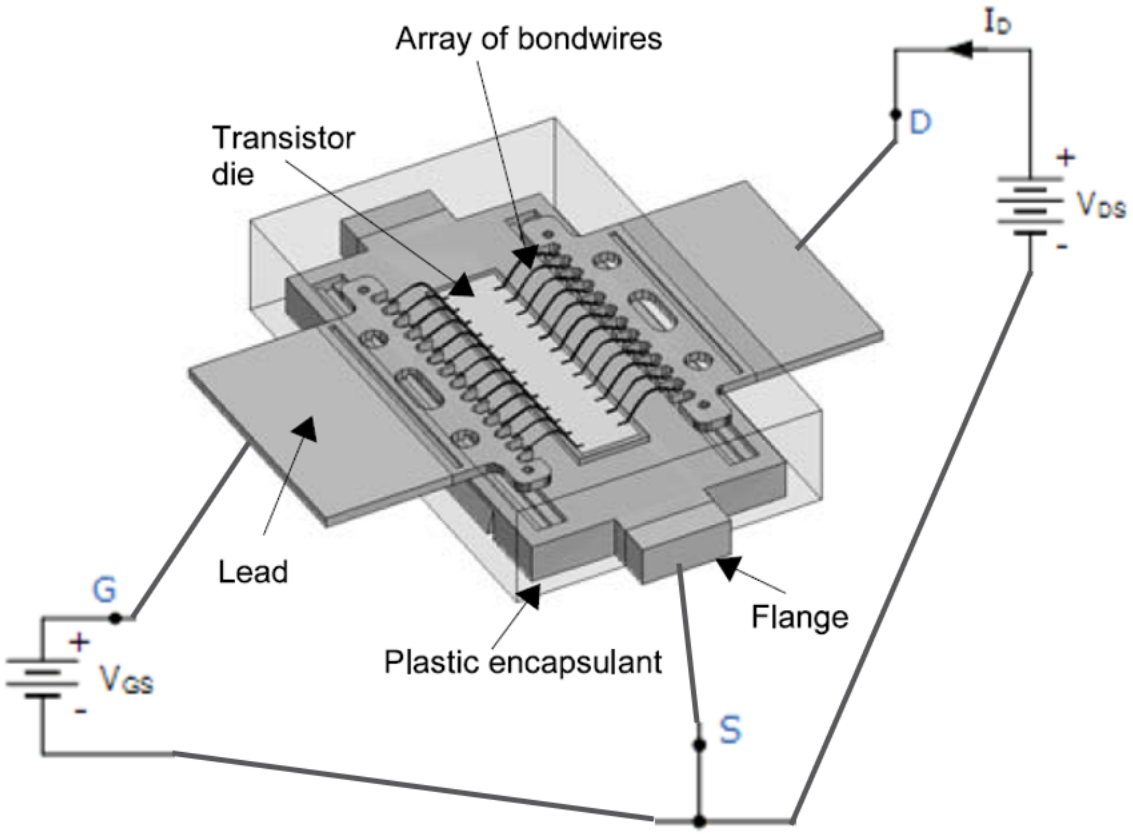

Figure 2.4: The DC bias of a packaged power transistor

in the entirety of the sinusoidal input voltage swing, and also when there is no RF signal being applied. A class B operation is when the transistor is biased right at gate threshold voltage of the device and will have a conduction angle of 180°, resulting in a half-sinusoidal current waveform generated at the drain of the device in the ideal case. A class AB bias is any bias in between class A and class B that may be desired due to a good trade-off between linearity and power efficiency. A class C biased device will be biased below the gate threshold voltage in the "off region" of operation. A class C biased device behaves like a passive device at small signal levels, but as the magnitude of the incident signal increases and the conduction angle of device increases, eventually the device will have more than unity power gain. This power-level dependence of a class C biased device makes it very useful in a Doherty configuration power amplifier. In a Doherty PA, the peaking (auxiliary) amplifiers are biased in class C as these transistors are meant to provide power

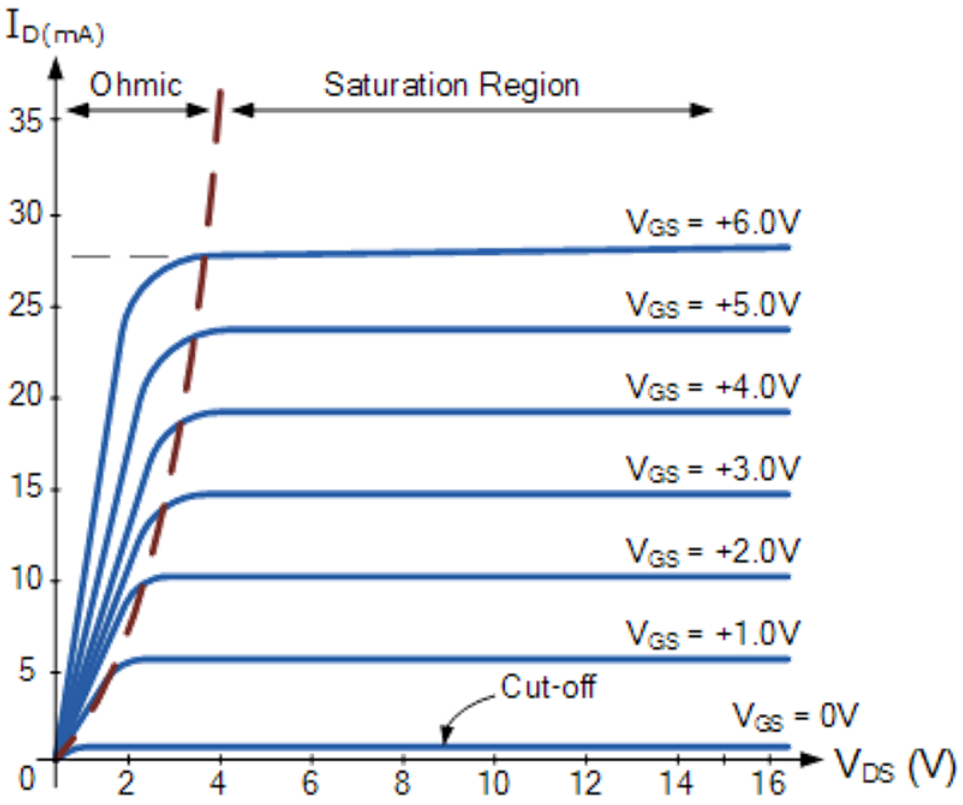

Figure 2.5: The DC IV characteristics of a hypothetical field effect transistor with zero threshold voltage

for the peaks of the modulated signal that is being amplified with good power efficiency, while being off and not consuming much power at backed-off power levels. The plots of the classes of operation only provide a simplistic view of how the transistor will behave as the actual shape of the DC current curves are not as ideal as the one shown in Figure 2.5 and the transistor device of Figure 2.4 has a frequency dependent response due to the many electrical parasitic behaviour from both the package and bonding manifold to the capacitance of the transistor die itself. Many power transistors have very nonlinear gate to source capacitance like in the case of Gallium Nitride High Electron Mobility Transistor (GaN HEMT) devices or have a very nonlinear drain to source capacitance like in the case of Laterally Diffused Metal Oxide Semiconductor (LDMOS) devices which further changes the nonlinear behaviour as a function of the output power level.

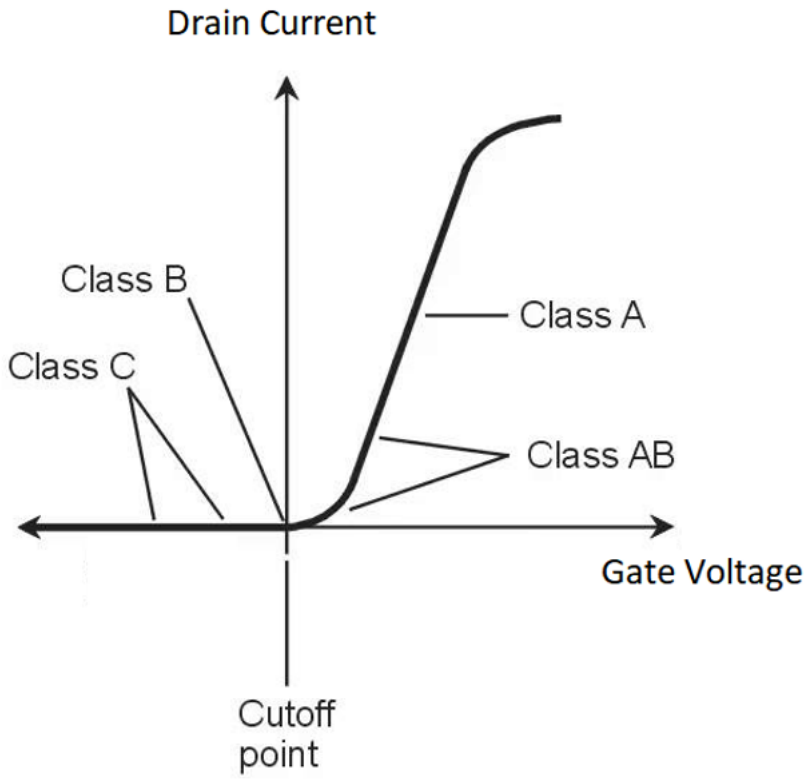

Figure 2.6: The DC relationship between the gate DC voltage and the drain current for a fixed DC drain voltage value

In addition to the electrical nonlinear behaviour of the power transistor, the power transistor behaviour is highly dependent on temperature. The DC drain to source current is known to drift and increase when biasing power transistors due to self-heating effects. These changes also show themselves in the large signal RF behaviour of power transistor. In addition, some GaN HEMT power transistors display significant "trapping effects" where trap states can form for the electric charge carriers due to the impurities at the semiconductor material interfaces. The time constants associated with the temperature and trapping effects are usually much slower than the electrical behaviour of the power transistor and thus contribute to the so-called long-term memory effects of the RF power transistor. As a result of these memory effects, more DPD resources would be required to correct for these effects and the amount of achievable correction with for a finite DPD algorithm will be limited.

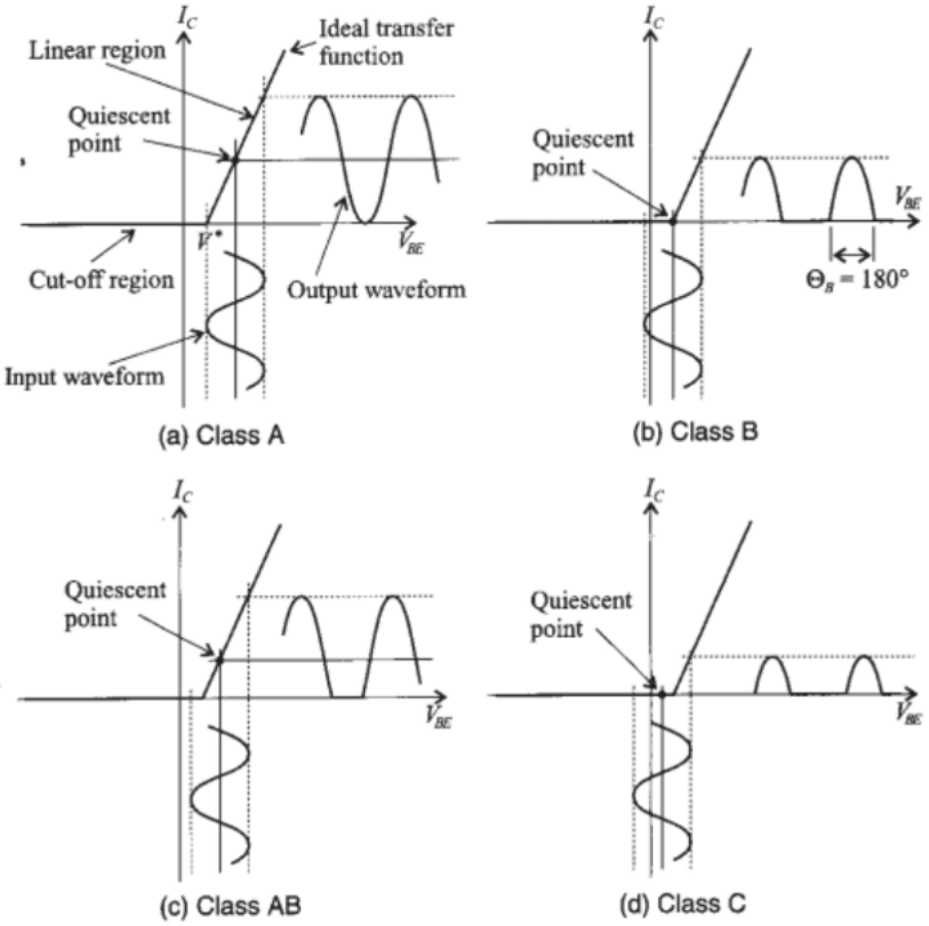

Figure 2.7: The basic single transistor analog power amplifier classes of operation

One of the main tasks of RF power amplifier design is the design of the input and output matching networks at the gate and the drain of the power transistors. These matching networks present the appropriate input and output impedances to the device at the operating frequencies as well as harmonic frequencies. In the design of linear RF amplifiers, the task of designing these matching networks can be performed by extracting the linear S-parameters of the power transistor and thus using closed form solutions to find the optimal matching input and output impedances of the transistor [2]. However when the power transistor is exhibiting nonlinear behaviour then its RF performance metrics like power gain and input return loss will also need to be measured as the input power level is increased. Due to the nonlinearity of the power transistor, these RF characteristics also change when

the RF load impedance changes. This is why for the nonlinear characterization of a power transistor involves large-signal load-pull measurements. In these load-pull measurements, the fundamental load impedance presented to the drain of the power transistor at the operating frequency is varied, and power sweeps of the input power are performed on the device where the DC and RF behaviour of the power transistor is completely measured at both the gate and the drain of the device. The load-pull measurements are frequency dependent, so all these load-pull measurements will need to be performed across the frequency band of interest.

Since the behaviour of the power transistor is nonlinear, due to a stimulus at the operating frequency, there will be power generated at both the gate and drain of the device at the harmonic frequencies. This means that in addition to a fundamental frequency impedance termination, the harmonic impedance terminations at both the gate and the drain of the device can have an effect on the performance. The design of the power amplifier, in the narrow-band sense, involves the design of the input and output matching networks that provide the input and output impedances at the operating frequency and its harmonic frequencies that give the desired performance. This is called the "poly-harmonic design space" at a given fundamental frequency and it is the space where the load-pull measurements will be performed.

The discussion of how the large-signal RF behaviour of a power transistor device can vary with bias, impedance terminations and frequency, brings into focus the importance of having accurate power transistor models for the simulation of power amplifier circuits.

2.3 Compact Models

Compact models of power transistors are implicit models of their electrical behaviour. These models are formulated based on a known theory of operation of the device and are expressed as a circuit representation which includes linear and nonlinear elements. This circuit representation of the model consists of nodes that are connected by branches. Each branch of the circuit can be equivalently represented by a differential equation (e.g. the relationship of the current flowing through a capacitor and the electrical potential across the capacitor is a differential equation). Due to this basic fact, compact models are not frequency specific as they're made up of time-domain differential equations that hold for all frequencies of interest. The basic compact model topology used in the literature for packaged power transistors involves the logical separation of the package model from the inner transistor. The inner transistor is then also further sub-divided and modelled as an *active intrinsic transistor* embedded within a *passive extrinsic shell*. The underlying

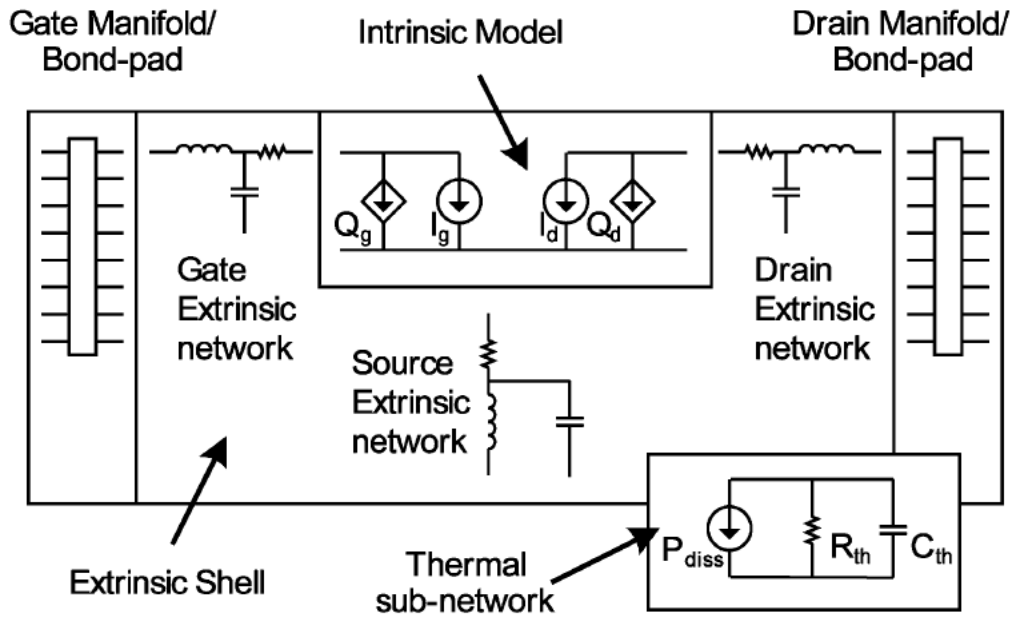

Figure 2.8: Typical Compact Model Topology of a Field-Effect Transistor [3]

assumption here is that the extrinsic network is linear and non-temperature-dependent while all the nonlinearity of the device as well as thermal effects originate from the intrinsic transistor. Some models even include the trapping effects exhibited by the power transistor as part of the intrinsic transistor description [3] [4].

Figure 2.9 shows a simplistic nonlinear model of the intrinsic transistor used by many compact models. Here the total gate-to-source and drain-to-source current of the intrinsic transistor is represented by a current source in parallel with a charge source and described by the state equations 2.1 and 2.2 [5]:

$$I_{drain}(t) = I_d\left(V_{gs}(t), V_{ds}(t)\right) + \frac{d}{dt}Q_d\left(V_{gs}(t), V_{ds}(t)\right) \tag{2.1}$$

$$I_{gate}(t) = I_g\left(V_{gs}(t), V_{ds}(t)\right) + \frac{d}{dt}Q_g\left(V_{gs}(t), V_{ds}(t)\right) \tag{2.2}$$

The drain and gate current expressions of equations 2.1 and 2.2 are separated into a conduction current nonlinearity (nonlinear current sources I_d and I_g) and a displacement

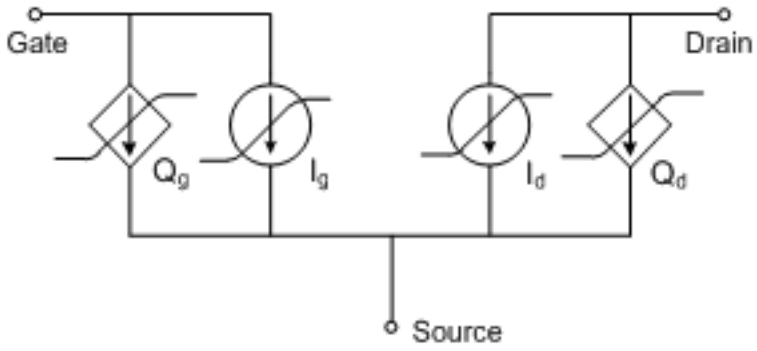

Figure 2.9: Intrinsic Transistor Model

current nonlinearity (nonlinear charge sources Q_d and Q_g). Nonlinear capacitive effects that are common in many transistor devices such as LDMOS and GaN HEMTs and are represented by these displacement current nonlinear terms that can account for a general two-dimensional nonlinear capacitance.

The development of compact models requires a strong insight into the actual sources of linear and nonlinear dynamics of the device. For example, the nonlinearity of a transistor device is not entirely static, meaning that having the DC input/output trans-conductance of the intrinsic transistor does not have all the information to predict the nonlinear dynamics of the transistor. This is in part due to nonlinear capacitances (capacitances that vary with the instantaneous voltage level) of devices which show up as nonlinear charge sources in the intrinsic transistor model. These nonlinear components of the intrinsic transistor can themselves be temperature dependent. A thermal sub-circuit that tracks the temperature of the transistor as a function of power dissipation is usually used in compact models to model the thermal effects of the power transistor. Some very useful compact models in the literature include the Angelov model [6][7], the Root model [8][9], and the DynaFET model [10][11]. The DynaFET model in particular is summarized in Figure 2.10 where an artificial neural network that takes a parameter input from a self-heating circuit meant to simulate the junction temperature and separate trapping circuit meant to simulate the trap states. The field of research of compact model design is active and ongoing work is being done to improve the modeling capability of compact models for the objective of RF power amplifier design [12][13][14][15][16].

A common strategy employed in the development of compact models of a power transistor, after determining a compact model topology that reflects the dynamics of the device of interest, is to experimentally extract the chosen model parameters from measurements

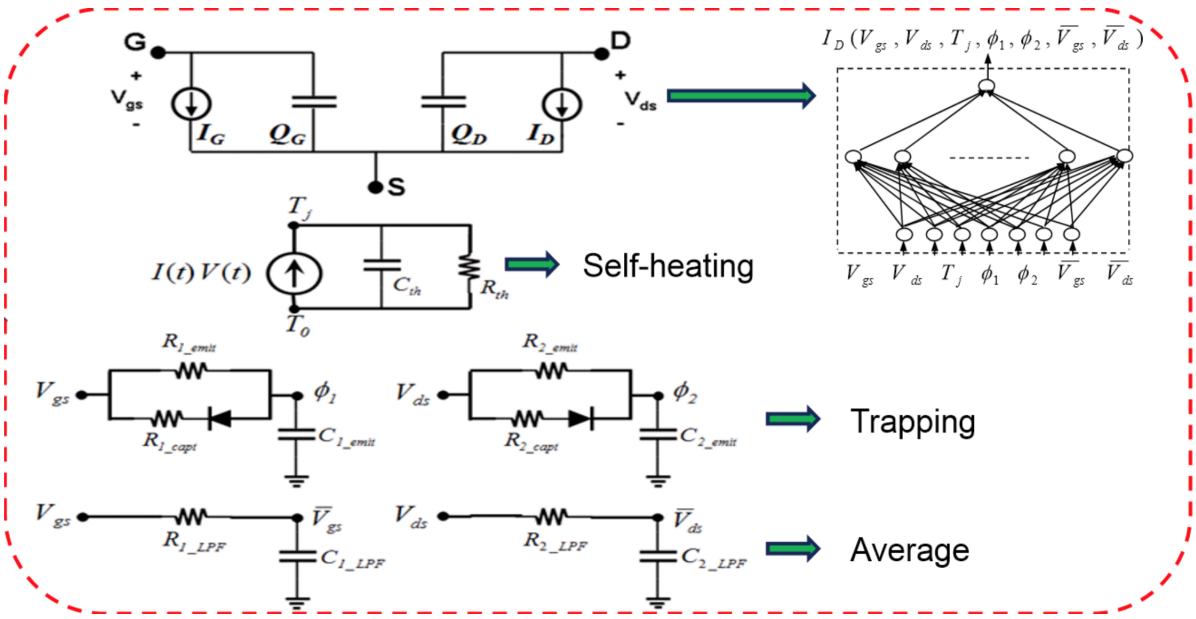

Figure 2.10: Form of the DynaFET compact model

that isolate the parameters of interest. For instance, for packaged transistors, the package model and the bonding manifold to the transistor die can be modelled in 3D electromagnetic simulation software, and measurements can be de-embedded through this simulated package model to get to the transistor die measurement reference plane. For compact models of the general form of Figure 2.8, the extrinsic shell is usually extracted next via a measurement that isolates the effect of the intrinsic transistor from the extrinsic shell. By doing this, the model engineer will then be able to de-embed measurements of the transistor to the intrinsic transistor in order to extract the active parameters. In order to delineate the extrinsic shell from the intrinsic transistor, a technique known as the *cold FET measurement* is often used [5]. In this technique, the active transistor is kept unstimulated by keeping the transistor in the *OFF* state and measuring its small signal response. It is assumed that the dynamics measured when the transistor is off is mainly due to the extrinsic shell. Using a frequency sweep of its two-port parameters, the extrinsic shell parameters can be optimized to match the cold FET measurements.

After determining the extrinsic shell parameters, the focus is on determining the conduction current and displacement current nonlinearities of the intrinsic transistor respectively. If the model is a electro-thermal model, these intrinsic parameters will be measured in a temperature controlled environment and will be made to be a function of temperature.

In addition some new compact models also include trap state variables for the intrinsic transistor to account for the non-ideal trapping effects seen in some transistor topologies like GaN HEMT devices. Extracting the parameters for the trapping model also requires its own specialized measurements of the power transistor.

The extraction of the intrinsic parameters has been traditionally performed using pulsed DC and RF measurements. The measurement is pulsed so that the self-heating effects do not have an effect on the measured behaviour of the transistor. The gate and drain DC voltages are swept, and for each DC bias condition, a small-signal S-parameter measurement is performed to extract the bias-dependent local linear dynamics of the transistor. The DC current measurement of the intrinsic transistor at each DC voltage bias point is used to determine the conduction current nonlinearity (I_d and I_g functions of equations 2.1 and 2.2 respectively). The S-parameter measurements at each bias point are used with numerical integration (and an assumption of charge conservation of the intrinsic transistor) to find the displacement current nonlinearity terms (Q_d and Q_g of equations 2.1 and 2.2 respectively) [5]. Using this method the large signal intrinsic transistor behaviour is extracted only from bias-dependent small-signal measurements.

After numerically extracting the values of the two-dimensional nonlinear functions representing the DC conduction current nonlinearities $I_d(V_{gs}, V_{ds})$ and $I_g(V_{gs}, V_{ds})$, and the displacement current nonlinearities $Q_d(V_{gs}, V_{ds})$ and $Q_g(V_{gs}, V_{ds})$ over a range of (V_{gs}, V_{ds}) DC biases, a closed form equation can then be fit to these points or a curve-fitting tools like artificial neural networks (ANNs) is often employed.

An important point to note is that a compact model is fundamentally a time-domain model in the sense that every single component of the model, whether a circuit element or nonlinear equation describing conduction current or displacement current is described by a time-domain differential equation. The frequency behaviour of the compact model is implied by the time-domain description. It is in this sense that a compact model is not an explicit model of the frequency domain behaviour of the device. To see what frequency domain and large-signal behaviour is implied by a certain compact model, it will have to be simulated in a large-signal circuit simulator. Since the measurements that were used to extract the compact model using the common method were all small-signal measurements, resulting in a model that can imply a certain large signal behaviour, a good validation of compact models is their ability to predict the large signal behaviour of a power transistor, especially under load-pull measurements. If there is significant error between the large-signal prediction of the compact model in the large-signal simulator and the large-signal measurement of the power transistor, then the model engineer can go back and re-tune their compact model parameters to result in a simulated large-signal performance that has less error when compared to the load-pull measurements of the power transistor.

It should be noted that the extraction and development of compact models of power transistors is neither an easy task, nor can it be done with simple measurements of the power transistor. Often when power transistors are new and in development, mature compact models that accurately model the large-signal behaviour of the power transistor in a circuit simulator are not available, leaving the power amplifier designer with only load-pull measurements of the power transistor as their only design guide. This is the primary motivation for the development of behavioural models of RF power transistors. Behavioural models will fit the load-pull measurement data directly and behave in the circuit simulator exactly like how they were measured to behave in the load-pull measurement.

2.4 Harmonic Balance Simulation of Nonlinear Circuits

The first computer simulation tools that simulated the nonlinear behaviour circuits used time-domain methods. The time-domain simulator steps in time, and solves for the circuit behaviour for each discrete time given the behaviour at the previous time samples. When it comes to the simulation of the response of nonlinear circuits to a periodic stimulus, e.g. a sinusoidal stimulus at a fixed fundamental frequency, then the time-domain simulator would start at an initial seed (instantaneous state of the circuit) and evolve in time until the time-domain response converges to a periodic response with the same fundamental frequency as the stimulus signal. At this periodic steady-state response, the entirety of the voltage at each node of the circuit can be described by its Fourier series coefficients at DC, the fundamental frequency of the periodic response and its harmonic frequencies. Harmonic balance simulation is a frequency-domain simulation that quickly resolves to the periodic response of the circuit by only accounting for the behaviour of the linear and nonlinear circuit at these fixed discrete Fourier series frequencies [17][18].

When a nonlinear circuit Netlist is defined, the harmonic balance simulator would need to track a voltage vector of complex Fourier coefficients and at every node of the circuit and a current vector of complex Fourier coefficients for every branch of the circuit. These Fourier coefficients include the value at DC (frequency of 0) and the complex phasor at each harmonic frequency of the fundamental frequency representing the magnitude and phase of the sinusoidal component at that harmonic frequency.

It's important to note that harmonic balance simulation is a discrete frequency domain simulation, so even though the time-domain waveforms of the voltages and currents can be plotted as though they are continuous waveforms, this continuous representation is

just an evaluation of the discrete terms of a Fourier series, and so the voltage and current information extracted from the simulation are fundamentally discrete. When doing a power sweep of a power transistor model in harmonic balance simulation, for each power level a separate harmonic balance simulation is performed until convergence is obtained for each power level state independently.

The poly-harmonic distortion models (that will be discussed in Section 2.6) are defined to fully take advantage of the harmonic balance simulator by having the input and output terms of the frequency domain model match up with the simulation harmonic frequencies.

The harmonic balance simulation of a compact model of a power transistor can be seen as the simulation analog of performing a frequency-domain measurement on the physical power transistor. In the next Section, the techniques used in the large-signal characterization of power transistors will be discussed.

2.5 Large Signal Characterization of Power Transistors

To measure the full large signal RF periodic behaviour of a power transistor device in a lab, there are different measurement techniques, based on different types of measurement instruments, namely either a Large Signal Network Analyzer (LSNA), Nonlinear Vector Network Analyzer (NVNA), or a multi-channel high-speed oscilloscope [19]. The most popular method that is used for the full-waveform periodic characterization uses an NVNA, and that's what is discussed in the next subsection.

2.5.1 Calibration and Measurement of a Nonlinear Vector Network Analyzer

The Vector Network Analyzer (VNA) has been a ubiquitous tool in RF measurement laboratories globally as they provide a comprehensive platform for performing calibrated relative measurements between different RF receivers, allowing for the extraction of S-parameters of different linear networks. Using a common local oscillator at a specific frequency to down-convert the sampled incident and reflected waveforms down to an intermediate frequency (IF) and then sampling those IF signals using time-coherent ADCs, an VNA is able to fully capture the relative relationship between the incident and reflective waveforms at the device ports one RF frequency at a time. And since for linear systems,

the characterization of the system is not a function of power, a VNA calibration and measurement does not include or need an absolute RF power calibration as all measurements and calibrations are performed relative to each other and the absolute value of the power is not known. In addition, since a VNA measures each frequency separately, if it were to measure the harmonic response of a nonlinear circuit to a fundamental frequency stimulus, it would not be able to infer the relative phase of the harmonic signal it measures relative to the fundamental frequency as the VNA can only perform cross-port relative measurements but not cross-frequency relative measurements.

The Nonlinear Vector Network Analyzer overcomes the limitations of VNAs for performing large-signal multi-harmonic periodic waveform measurements by augmenting a standard VNA with a power calibration to allow for the measurement of absolute power levels, resulting in an 8-term calibration model that can be used for load-pull measurements [20][21][22][23][24]. In addition to the power calibration, a system phase reference is measured on a separate RF receiver of the NVNA to allow for cross-frequency relative phase measurement of the waveforms. By performing a standard VNA calibration followed by a power calibration and then a phase calibration, a full NVNA calibration can be achieved that allows for the full periodic (multi-harmonic) waveform characterization of both the incident and reflected waves on each port of the device under test (DUT) [19] [25][26]. Figure 2.11 is an illustration of the NVNA system around the DUT device (shown with the letter A).

For power transistor and power amplifier measurements, pulsed-RF measurements are used since using a non-pulsed sinusoidal RF signal at the saturation power levels of a high power transistor device would result in excess heat generated by transistor device and would likely cause heat damage. When the power transistor is used to amplify a modulated signal with a high peak-to-average-power-ratio (PAPR) then instantaneously the amplifier will output RF powers close to the saturation level but since the modulated signal is at those peak powers for a short time, the self-heating at the peak power do not result in heat damage, since on average the RF output power is at a backed-off level compared to the saturation power level of the power transistors. A 10% duty cycle can allow for the characterization of the periodic response of the power transistor at its peak powers without overheating the device. Performing pulsed measurements requires some considerations on triggering of instruments used for the DC measurement during the RF-pulse and can be performed using triggered digital multimeters or current probes attached to a triggered oscilloscope.

Though a NVNA calibration and measurement instrument is a requirement for the type of power transistor characterization that is needed for model extraction, it is not sufficient as a power sweep measurement of a general RF power transistor when it is terminated

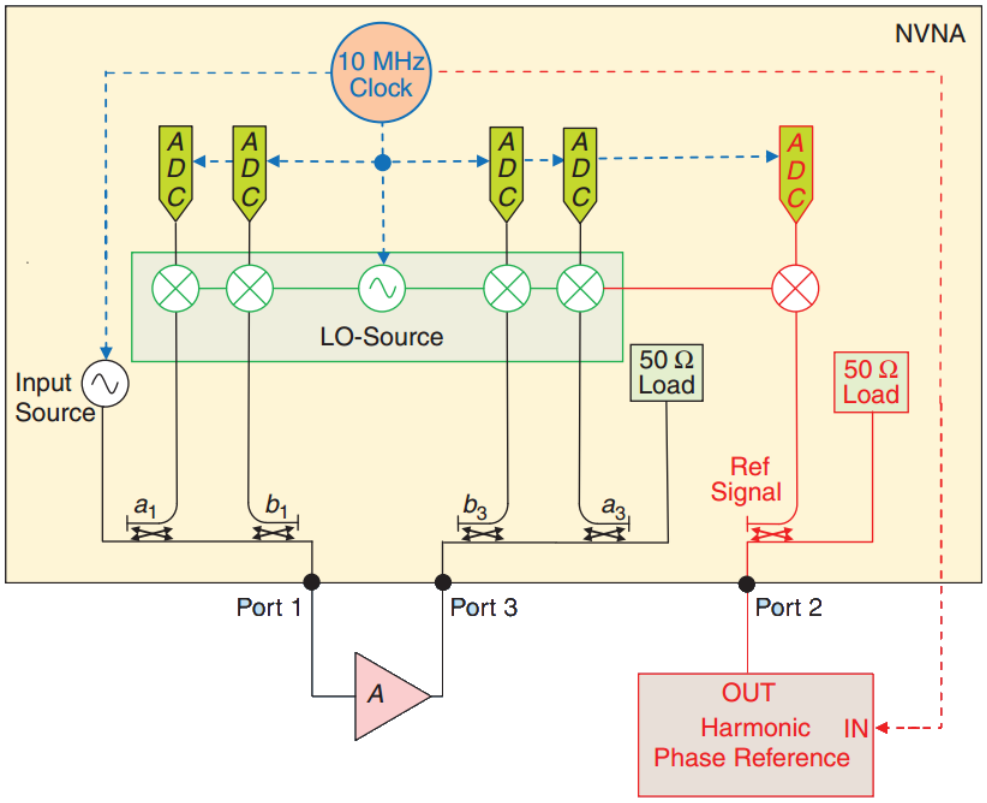

Figure 2.11: A Nonlinear Vector Network Analyzer (NVNA) measurement

with 50Ω ports does not give a full account of the its entire nonlinear behaviour and the load-impedance would need to be varied to explore the nonlinear response of the power transistor in more detail. This is the topic of the next subsection which describes the load-pull measurements of power transistors.

2.5.2 Load-Pull Measurements of Power Transistors

In a load-pull measurement of a power transistor, a power sweep is performed on the power transistor for different values of load impedances (or source impedances for the case of source-pull measurement). In Passive Load-pull [27], passive tuners, which are mechanical devices that can produce desired RF impedances at desired frequencies are employed for the large-signal power sweep measurement like in Figure 2.12.

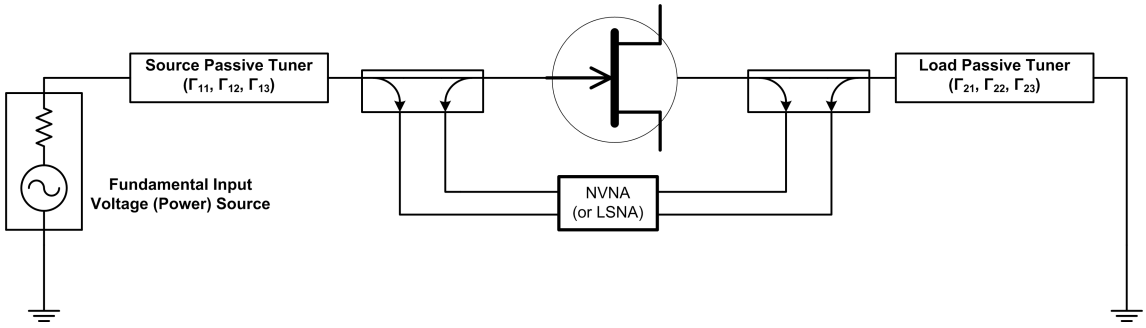

Figure 2.12: A Passive Load-pull Measurement

In the example of Figure 2.12 a three-harmonic source tuner and a three-harmonic load-tuner is used and is placed outside the dual-directional couplers that are calibrated to measure the input and output waveforms at the DUT measurement reference plane. Placing these tuners outside of the couplers allows for the changing of the impedance termination seen by the power transistor without changing the required RF calibration for the measurement. The harmonic tuners allow the designer to set the input and output impedances (or equivalently the reflection coefficient) presented to the power transistor. These passive tuners require automated mechanical/RF calibrations where the tuner probes move through their positions and the S-parameters of the tuner are measured, allowing the tuner software to be able to predict what the S-parameters of the passive tuner are at every tuner position. In addition the tuner software allows them to be reconfigured mechanically to generate a requested RF impedance at a specific calibrated frequency, and in the case of multi-harmonic tuners, will also present the desired impedances to the DUT at the harmonic frequencies.

Passive tuners have the advantage of being able to operate under very high powers, making them very attractive for the load-pull measurement of high power RF transistors. The tuning capability of the passive tuner is limited to a certain maximum reflection coefficient magnitude (a reflection coefficient typically limited to around 0.9 at RF frequencies). With the imposition of the dual directional couplers in front of the passive tuners, any insertion loss from those couplers, as well as the insertion loss of the transistor fixture will will further limit the achievable RF reflection coefficient magnitude at the power transistor plane. Another drawback of passive load-pull is due to the fact that the passive tuners are mechanical devices, they are slow to re-adjust and so passive load-pull measurements are much slower than active load-pull measurements that are fully electrical and do not have any moving parts for the measurement.

In active load-pull [28][29][30], the impedance, or alternatively, the reflection coefficient

generated at the device plane is not the result of the reflection of a passive system, but the result of an active system that injects power back into the device to emulate a desired reflection coefficient. Figure 2.13 shows the basic idea behind an active load-pull system. The reflection coefficient Γ seen at a device port is the ratio of the incident wave (represented by a complex number a) on the port to the reflected wave from the port (represented by another complex number b). The b wave is generated by the device and can be diverted to a load to be dissipated through an RF circulator. This circulator can also be used to inject any desired a wave incident back onto the device independent of the b wave generated by the device. The active load-pull system will then try to maintain a desired reflection coefficient $\Gamma = a/b$ by adjusting the injected a wave accordingly.

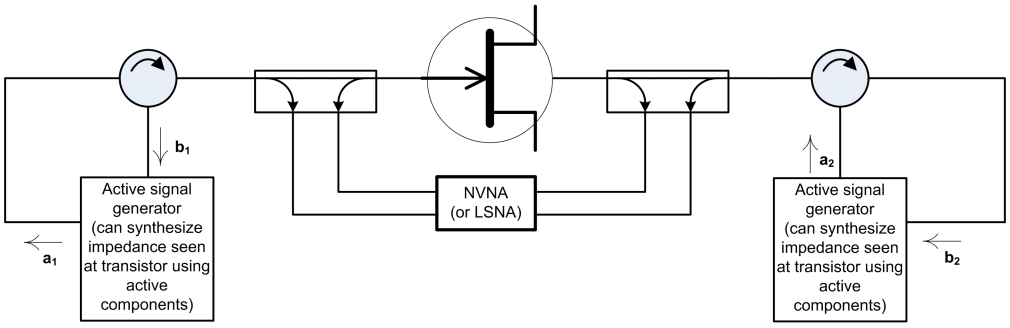

Figure 2.13: An active load-pull measurement

One of the drawbacks of using active load-pull for high power devices is that since the a wave has to be independently produced, to generate a high reflection coefficient Γ at the load of the device, the a wave produced must be comparable in magnitude to the high power b wave generated by the device, resulting in the requirement of high-power and expensive drivers in the active load-pull measurement setup. The main advantage of using active load-pull is the measurement speed, as these systems do not have any moving parts, they can measure many impedance loads rather quickly while the same number of measurements in a passive load-pull system would take thousands of times longer.

It is possible to take advantage of the benefits of both passive and active load-pull systems by employing what is referred to as a hybrid load-pull system, demonstrated by Figure 2.14.

In a hybrid load-pull measurement system, the a wave required for the generation of the desired reflection coefficient is partially generated by the reflection of a passive tuner but also has a contribution from an active load that generates the rest of the a wave presented to the device load. In this method, a high-power (expensive) driver as an active load can

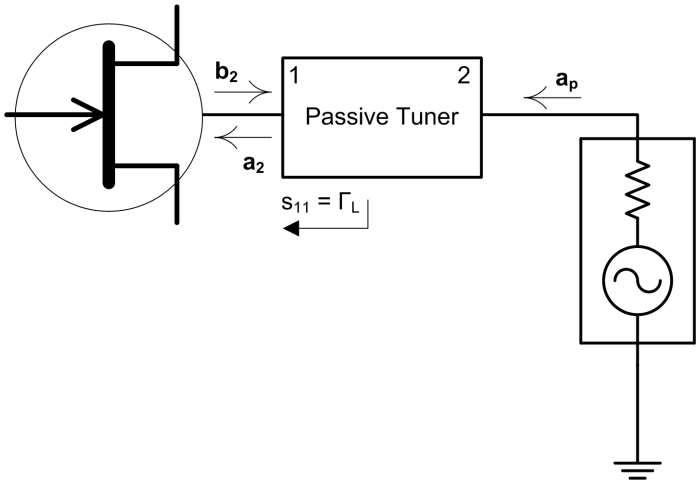

Figure 2.14: A hybrid load-pull measurement

be avoided, by using a passive tuner to get the load impedance to an impedance close to the desired high reflection coefficient target, and have a local and fast active load-pull around that passive impedance. This allows one to take advantage of the speed of active load-pull while not requiring a very high power active-load driver.

An analogy to the load-pull measurement of a power transistor using RF measurement instruments also exists in the harmonic balance simulation of a power transistor model. Creating the same conditions as a load-pull measurement, that is, changing the load impedances and performing power sweep measurements, can also be performed in a simulation environment in what is known as a simulated load-pull measurement. In fact, as stated before, the only way to find out what the large-signal performance is implied by a compact model at different load impedances is to perform a simulated load-pull measurement on the power transistor model in a harmonic balance circuit simulator.

Power transistor models that are explicitly directly derived from load-pull measurements are behavioural models of power transistors and have been used in the design of RF power amplifiers [31][32][33][34]. A good methodology used in this research to evaluate the efficacy of a behavioural model before getting involved with the complexities of performing large-signal measurements on power transistors, was to use a compact model as a device under test, and to perform simulated load-pull measurements on this compact model. The extracted load-pull measurements would then be used to extract a load-pull measurement as shown in Figure 2.15.

Of course the extracted behavioural models can themselves be used as the device under

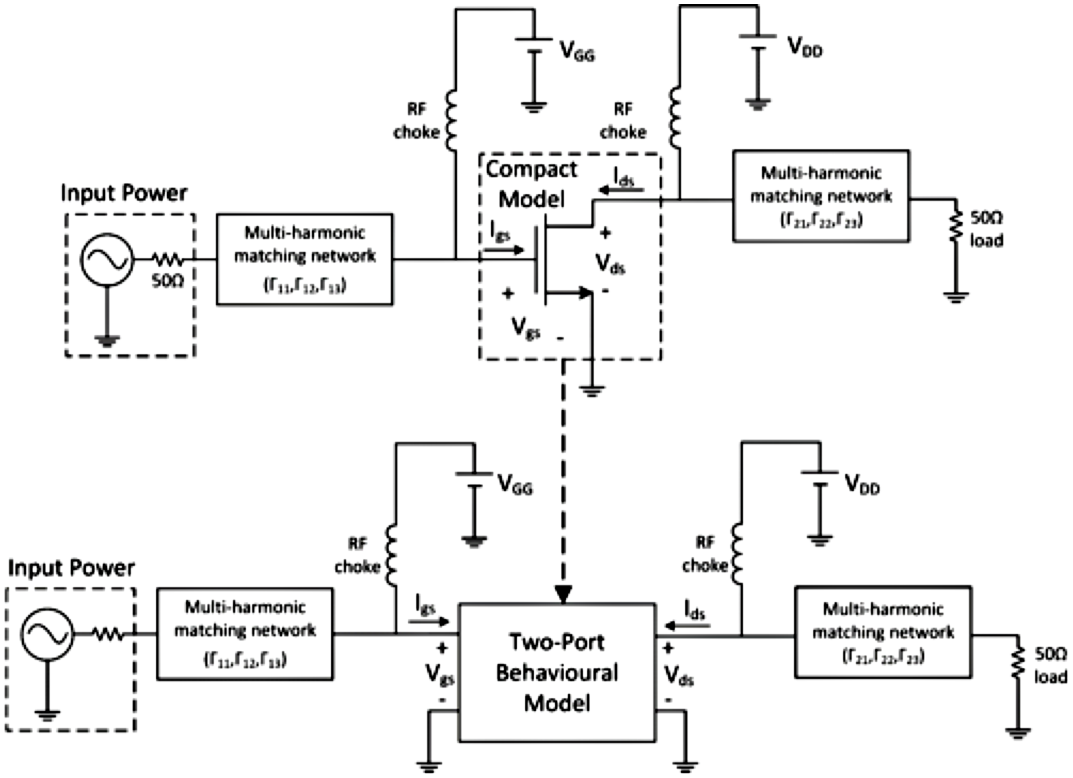

Figure 2.15: Using a Compact Model of a Power Transistor as the Device Under Test for a Simulated Load Pull Measurement

test in a load-pull simulation. Once a behavioural model is shown to at least be capable of replicating the simulated behaviour of a compact model of a power transistor, then it can at least have a chance to model the real measured behaviour of a power transistor as the behaviour of the compact model in a harmonic balance simulator can be seen as representative of how a power transistor of a par typically behave in measurements though there are differences that further necessitates the measurement validation of power transistor behavioural models.

In the next Section the poly-harmonic distortion model will be introduced as the dominant nonlinear frequency-domain behavioural modeling strategy used in the industry for RF power amplifier design.

2.6 Poly-Harmonic Distortion Models

The process involved in the development of compact models as discussed in Section 2.3 is quite complex. Creating compact models of power transistors requires a high level of theoretical knowledge of device physics and access to specialized equipment used to extract these models. In addition, the extraction of a compact model requires an experienced person to guide the model development and to tweak and tune parameters if needed. Power amplifier designers in general do not have the knowledge to develop a compact model and if they need a computer-aided design model for a packaged transistor and they aren't part of the organization that manufactures the transistor, then they are often left with only one type of model, which is a black-box, or behavioural model of the power transistor. To develop a behavioural model one only needs to measure the power transistor's behaviour of interest at the device ports and the behavioural model will predict how the device will electrically react at its ports without requiring any further insight into the inner-workings of the device. In fact the underlying semi-conductor technology used to make the power transistor doesn't matter as behavioural models by their very nature are technology-agnostic.

The most ubiquitous behavioural model for RF systems are S-parameters [35]. S-parameters are one of the many multi-port evaluated transfer functions of linear time-invariant systems defined at discrete frequencies[36][37] that are specifically suited to describe the RF power transfer characteristics of a multi-port system given known port reference impedances (Z_0, which is set to 50Ω in most RF applications). With the use of S-parameters one can simulate the linear response of a RF circuit. When the active components of the RF circuits are operating in a very linear manner, which is the case for the design of low noise amplifiers for instance, they can be approximated quite well with a linear model across frequency. An S-parameter model of a two-port system would consist of only four complex numbers (the S-parameters) at each frequency f and two linear equations, one for each port:

$$B_1(f) = S_{11}(f)A_1(f) + S_{12}(f)A_2(f) \tag{2.3}$$

$$B_2(f) = S_{21}(f)A_1(f) + S_{22}(f)A_2(f) \tag{2.4}$$

The pseudo-powerwaves A_1 and A_2 are the incident waves at port 1 (the gate) and port 2 (the drain) of the power transistor and B_1 and B_2 are the reflected waves at these two ports as shown in Figure 2.16 and are related to the voltage and current at the ports of the device by the following relationships:

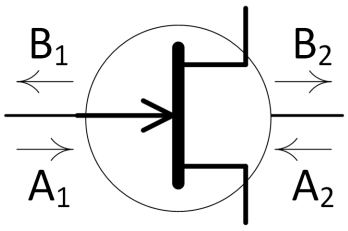

Figure 2.16: Incident and reflected powerwaves on a device under test

$$A_1(f) = \frac{V_1(f) + Z_0 I_1(f)}{2\sqrt{R_0}} \tag{2.5}$$

$$B_1(f) = \frac{V_1(f) - Z_0 I_1(f)}{2\sqrt{R_0}} \tag{2.6}$$

$$A_2(f) = \frac{V_2(f) + Z_0 I_2(f)}{2\sqrt{R_0}} \tag{2.7}$$

$$B_2(f) = \frac{V_2(f) - Z_0 I_2(f)}{2\sqrt{R_0}} \tag{2.8}$$

For an amplifying transistor, the S_{11} parameter is called the input match parameter, while the S_{22} parameter is called the output match parameter. S_{21} is the gain parameter and S_{12} is the reverse isolation parameter.

As can be seen from what is required for a behavioural model for a linear system, the description of the system at each frequency is independent of its description at other frequencies. In linear systems there is no cross-frequency interaction between signals.

Early attempts to develop a behavioural model for nonlinear systems focused on generalizing the S-parameters by tweaking the model slightly. Initially a technique known as *Hot S-Parameters* was proposed where the S-parameters of the nonlinear system were a function of the input power level [38] as well as frequency:

$$B_1(f) = \mathrm{Hot}S_{11}(f, |A_1(f)|)A_1(f) + \mathrm{Hot}S_{12}(f, |A_1(f)|)A_2(f) \tag{2.9}$$

$$B_2(f) = \mathrm{Hot}S_{21}(f, |A_1(f)|)A_1(f) + \mathrm{Hot}S_{22}(f, |A_1(f)|)A_2(f) \tag{2.10}$$

However it was soon discovered that this technique is lacking in accuracy since the resulting reflected waves B_2 would empirically depend on the relative phase of the A_2 and A_1 waves, which is a uniquely nonlinear effect that is not captured by the linear form of the initial extension of S-parameters. To account for the phase variation, instead of using a model that used 4 complex parameters, 2 extra complex numbers were added to add enough degrees of freedom to account for phase variation of the A_2 wave relative to the A_1 wave. This model is referred to as Hot (Active) parameters by Keysight Technologies and is the successor to the older Hot S-parameters:

$$B_1 = X_{11}^{(F)}(|A_1|)P + X_{12}^{(S)}(|A_1|)A_2 + X_{12}^{(T)}(|A_1|)PA_2^* \tag{2.11}$$

$$B_2 = X_{21}^{(F)}(|A_1|)P + X_{22}^{(S)}(|A_1|)A_2 + X_{22}^{(T)}(|A_1|)PA_2^* \tag{2.12}$$

where $P = A_1/|A_1| = e^{j\angle A_1}$ is the pure phase component of the incident wave A_1 at each frequency. The addition of this term is to ensure that the model remains time-invariant while accounting for the relative phase of A_2 and A_1. For instance by observing equation 2.12, the output reflected wave B_2 is made up of two parts, the first part $X_{21}^{(F)}(|A_1|)P$ is only dependent on the phase of A_1, while the second part $X_{22}^{(T)}(|A_1|)PA_2^*$ is only dependent on the relative phase of A_2 and A_1. The dependency on the relative phase, instead of an absolute phase is due to $|A_1|$ being the "Large Signal Operating Point" (LSOP) of the model. The $X_{22}^{(T)}$ is an interaction parameter that describes how the A_2 wave interacts with the A_1 wave with their relative phase taken into account. $X_{11}^{(F)}(|A_1|)/|A_1|$ is the active input match term which varies with input power. $X_{11}^{(F)}(|A_1|)/|A_1|$ is the active gain term of the model while $X_{12}^{(S)}(|A_1|)$ is the active reverse isolation term.

The behavioural models up to this point only account for wave interaction at the fundamental frequency, but as mentioned before, the nonlinearity of the power transistor results in the generation of power at harmonic frequencies which interacts with the surrounding circuit at those frequencies.

Moving beyond the Hot (Active) model, to account for the interaction of the waves at the harmonic frequencies, an assumption can be made that the harmonic effects can be linearly super-imposed on to the resulting output wave through what is referred to as the *harmonic superposition principle*. To account for such models that involve the cross interaction of harmonic waves, we will shift to a poly-harmonic notation for the incident waves and referring the the incident wave at port 1 (the gate of the power transistor) at the fundamental frequency and its harmonics respectively as A_{11}, A_{12}, A_{13}, and so on and the

incident waves at port 2 (the drain of the power transistor) at the fundamental frequency and its harmonics respectively as A_{21}, A_{22}, A_{23}, and so on. Similarly the reflected waves at the input of the device at the fundamental frequency and its harmonics will be referred to as B_{11}, B_{12}, B_{13} and reflected wave at the output of the power transistor are denoted by B_{21}, B_{22}, B_{23} and so on.

Using this notation, the model that is sometimes referred to as the matched-load X-parameter model is defined for each reflected wave B_{ef} at port e and harmonic frequency f is expressed as [39][40][41]:

$$B_{ef} = X_{ph}^{(F)}(|A_{11}|)P^{-h} + \sum_{g,h}\left(X_{ef,gh}^{(S)}(|A_{11}|)A_{gh}P^{f-h} + X_{ef,gh}^{(T)}(|A_{11}|)A_{gh}^{*}P^{f+h}\right) \tag{2.13}$$

where $\{g,h\} \notin \{\{1,1\}\}$. Here the same type of modeling through $X^{(S)}$ and $X^{(T)}$ terms is extended to have a contribution from all the incident waves at all ports and all harmonic frequencies. In order to extract these X-parameters, the LSOP of the model will have to be held constant while a *tickler tone* is injected as an incident wave at every port and at every harmonic frequency. This tickler tone is injected at varying relative phases compared to the fundamental frequency A_{11} tone. This emphguided characterization step is essential for the extraction of these X-parameters. Figure 2.17 shows that a local circular perturbation of each of the harmonic incident waves A_{gh} results in an elliptical perturbation in the resulting B_{ef} reflected wave. The $X^{(S)}$ and $X^{(T)}$ terms in conjunction define the local stretching and skewing of the complex plane. Since all the parameters of the model are a function of $|A_{11}|$, the underlying assumption is that since $|A_{11}|$ is the most significant incident wave, it's the only dimension along which the nonlinearity can vary. This can be a good assumption for nonlinear power amplifiers that have been matched to the load impedance close to 50Ω, which results in a very small A_{21} wave reflected back onto the load side of the power amplifier in a $Z_0 = 50\Omega$ system.

In this X-parameter analytic model the output phasors at each harmonic are linearly related to all the other input phasors and also their complex conjugates. To explain why there is a dependence on the complex conjugates, it should be noted that the set of complex numbers A_{11}, A_{12}, A_{13}, A_{14}, and so on are Fourier series coefficients of a periodic real-valued time-domain waveform as follows:

$$a_1(t) = \frac{1}{2}\sum_{h=0}^{h_{\max}}\left(A_{1h}e^{jh\omega_0 t} + A_{1h}^{*}e^{-jh\omega_0 t}\right) \tag{2.14}$$

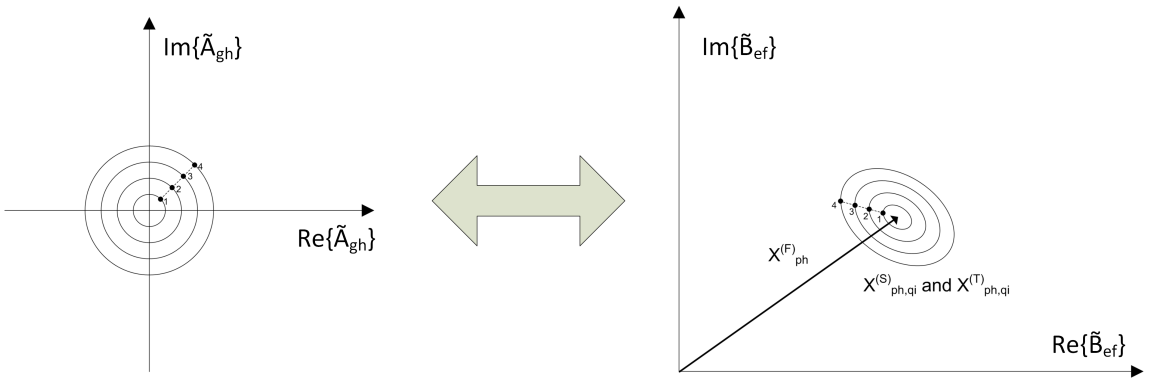

Figure 2.17: Visual representation of the $X^{(F)}_{ph}$, $X^{(F)}_{ph,qi}$ and $X^{(T)}_{ph,qi}$ X-parameters

$$b_1(t) = \frac{1}{2}\sum_{h=0}^{h_{\max}} \left(B_{1h}e^{jh\omega_0 t} + B^*_{1h}e^{-jh\omega_0 t}\right) \tag{2.15}$$

$$a_2(t) = \frac{1}{2}\sum_{h=0}^{h_{\max}} \left(A_{2h}e^{jh\omega_0 t} + A^*_{2h}e^{-jh\omega_0 t}\right) \tag{2.16}$$

$$b_2(t) = \frac{1}{2}\sum_{h=0}^{h_{\max}} \left(B_{2h}e^{jh\omega_0 t} + B^*_{2h}e^{-jh\omega_0 t}\right) \tag{2.17}$$

This exponential Fourier series representation of the time-domain incident periodic waveform at the input of the power transistor makes it explicit that since the value of a_1 is real at all time, then the negative frequency Fourier series coefficients are the complex conjugate of the positive frequency Fourier series coefficients. This means that any sort of analytic expression of the time-domain waveform $a_1(t)$ if re-written in Fourier series notation will necessarily have contributions from the both the positive frequency Fourier series coefficients and the negative frequency Fourier series coefficients.

To have a time invariant expression of reflected wave Fourier series coefficients as a function of the incident wave Fourier series coefficients, it's necessary to make the phase of all the complex phasors be expressed as a relative phase to a system phase reference, which will be the phase of A_{11}. In this thesis, the *tilde* symbol on top of the phasor will denote that its phase has been referenced to the system reference phase. So $\tilde{A}_{11} = \tilde{A^*}_{11} = |A_{11}|$, $\tilde{A}_{21} = A_{21}P^{-1}$, $\tilde{A}_{12} = A_{12}P^{-2}$, and so on where $P = A_{11}/|A_{11}| = e^{j\angle A_{11}}$ is the absolute

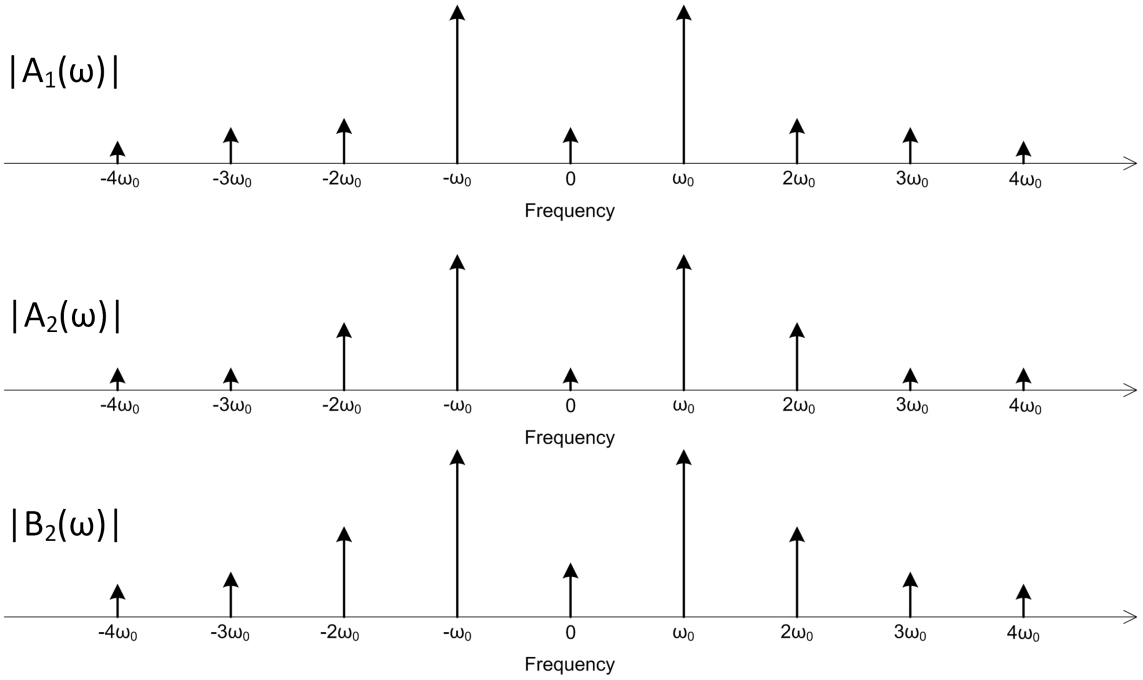

Figure 2.18: Periodic stimulus in the frequency domain

phase of the phase reference. In general the notation for the phase referenced Fourier series coefficients will be:

$$\tilde{A}_{gh} = A_{gh} P^{-h} \tag{2.18}$$

$$\tilde{B}_{ef} = B_{ef} P^{-f} \tag{2.19}$$

This leads to a simple representation of the general functional form of the poly-harmonic distortion model, that is both a time-invariant and analytic expression [42]:

$$\tilde{B}_{ef} = f_{ef}\left(|A_{11}|, \tilde{A}_{21}, \tilde{A}^*_{21}, \tilde{A}_{12}, \tilde{A}^*_{12}, \tilde{A}_{22}, \tilde{A}^*_{22}, \cdots\right) \tag{2.20}$$

The describing functions f_{ef} in equation 2.20 are analytic, meaning the output variable of each of these functions can be approximated locally around a bias point with a Taylor-series approximation of its input variables. It can be seen that the X-parameter analytic

model of equation 2.13 is the first-order Taylor-series approximation of the PHD describing function of 2.20 taken around a constant bias point of $|A_{11}|$, which is the LSOP of the model.

As mentioned before, the matched X-parameter model defined by the describing functions of equation 2.13 targeted matched power amplifiers as the device under test. This assumption is not true when the device that is to be modelled is a power transistor that in general does not have the system characteristic impedance as its input and output impedance and thus can have a significant A_{21} wave reflected back on it.

This *matched* X-parameter model's assumption would be valid in modelling a matched power amplifier that is only going to be driven at its input. However the harmonic superposition principle assumption of the model breaks down as the magnitude of the fundamental output reflection coefficient as well as the harmonic reflections on the device increases significantly [43]. This limitation is the same as the limitation of the first-order Taylor series approximation.

In an extension of the applications of the matched X-parameter model to the cases where the output behaviour of the power transistor can vary significantly based on the fundamental frequency impedance termination, the LSOP definition used by the X-parameter model can be extended to include these variations. To include the effects of the load termination at the fundamental frequency, the X-parameter describing functions of equation 2.13 can be modified to include the reflection coefficient term $\Gamma_{21} = A_{21}/B_{21} = (Z_{L1} - Z_0)/(Z_{L1} + Z_0)$ where Z_{L1} is the fundamental frequency load impedance and is added to the LSOP definition as follows:

$$\tilde{B}_{ef} = X^{(F)}_{ef}(|A_{11}|, \Gamma_{21}) + \sum_{g,h} \left(X^{(S)}_{ef,gh}(|A_{11}|, \Gamma_{21})\tilde{A}_{gh} + X^{(T)}_{ef,gh}(|A_{11}|, \Gamma_{21})\tilde{A}^*_{gh} \right) \tag{2.21}$$

where $\{q, i\} \notin \{\{1, 1\}, \{2, 1\}\}$.

The expression of equation 2.21 accounts for the significant output variation as a result of the load impedance variation by extracting a different set of X-parameters for each load impedance signified by Γ_{21}. This however is not a complete picture, as significant variation in the harmonic impedance terminations of transistors has been shown to significantly vary the nonlinear operation of the device beyond what is modeled by the addition of the $X^{(T)}$ terms of the X-parameter model. One solution to account for such cases of *extreme nonlinearity* would be an extension of the model to add those impedance terminations to the LSOP definition as well. Here for instance, if it becomes apparent that a large enough second harmonic incident wave A_{12} or A_{22} significantly changes the nonlinearity of the

device, then those parameters will have to be added to the LSOP definition in order to keep the first-order Taylor series approximation of the X-parameter model valid. With the addition of every new parameter as a "significant parameter", a new dimension is added to the LSOP space, and all the X-parameters will need to be extracted for all the possible combination of values of the significant parameters of the LSOP space. At its extreme, when including all the PHD variables in the LSOP definition, the X-parameter model will only be a highly multi-dimensional look-up table of $X^{(F)}$ terms that will require a multi-dimensional interpolation function to simulate in a harmonic balance simulator.

Even in less extreme cases of extending the LSOP, there are a few problems that arise from interpolating the model parameters over a multi-dimensional space. The first problem is the problem of interpolating complex numbers, as the interpolation of complex numbers can have discrepancies whether the interpolation is performed over the real-valued magnitude and phase components of the complex numbers, or the real-valued real and imaginary components of the complex numbers. The model will rely heavily on an interpolation function to be simulated, and the choice of the interpolation also can change the result of the simulation. Using a model that relies on an interpolation function can force the measurement required for the extraction of the model to focus its measurements where the transistor nonlinearity varies significantly since a more densely spaced set of look-up table points are required to account for a bigger gradient in the measurement space. Often efficient underlying interpolation algorithms for look-up table models require the multi-dimensional interpolation data to be on an evenly spaced uniform multi-dimensional grid, something that is at odds with load-pull data that is often measured on irregular grids. In addition to an explosion of the required look-up table parameters, the more variables that are added to the model, the more highly dimensional look-up tables models become prone to simulation convergence problems.

Another approach that avoids all the simulation problems of look-up tables is to define a higher order model for the describing functions of equation 2.20, beyond a first order model to fit the load-pull measurements of a power transistor. The Cardiff model [44] uses a multi-dimensional Fourier series expression as the describing function used for the PHD model. The intuition behind this model is that each of the harmonic outputs of the PHD model is periodic with respect to the phase component of each of its inputs $\angle A_{gh}$. This periodic relationship can be modeled with a multi-dimensional Fourier series expression in the periodic angle variables $\angle A_{gh}$. Each of these multi-dimensional Fourier series coefficients are then expressed as a function of the magnitudes of the inputs of the PHD modeling framework. Equation 2.23 shows a three variable (a_{11}, a_{21} and a_{22}) expression of the formulated Cardiff PHD model [44]. This model takes into account the effect of the second harmonic load injected power (due to a reflection caused at the second

harmonic termination) on the output power available from the device under test.

$$\tilde{B}_{ef} = \sum_{n} G_{ef,n}(|A_{11}|, |A_{21}|)e^{j(n\angle\tilde{A}_{21})} \tag{2.22}$$

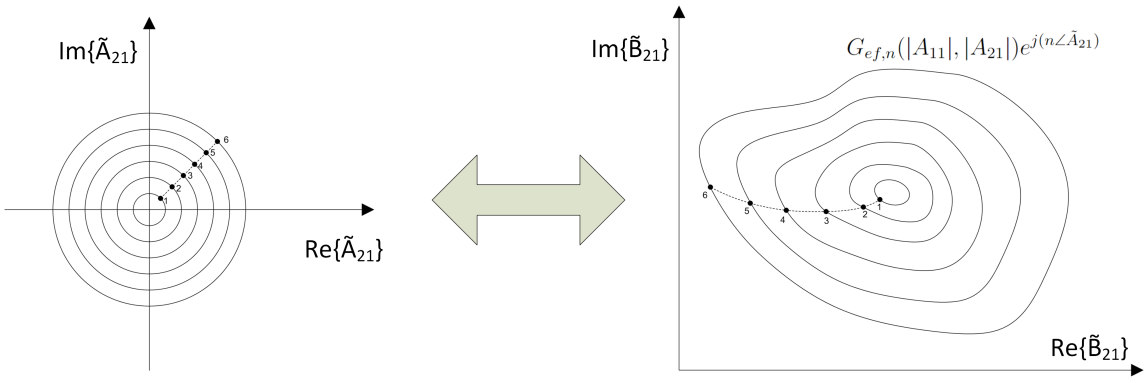

Figure 2.19: Change in $\tilde{B}_{21}$ with constant $|A_{11}|$ and a polar sweep of $\tilde{A}_{21}$

$$\tilde{B}_{ef} = \sum_{n}\sum_{r} G_{ef,n,r}(|A_{11}|, |A_{21}|, |A_{22}|)e^{j(n\angle\tilde{A}_{21}+r\angle\tilde{A}_{22})} \tag{2.23}$$

In another Cardiff model used for a power transistor second harmonic investigation, the source side second harmonic incident incident wave a_{12} were shown to be significant, this model was reformulated with this variable instead as seen in equation 2.24 [45].

$$\tilde{B}_{ef} = \sum_{n}\sum_{r} G_{ef,n,r}(|A_{11}|, |A_{21}|, |A_{12}|)e^{j(n\angle\tilde{A}_{21}+r\angle\tilde{A}_{22})} \tag{2.24}$$

The main problems of the models expressed in equations 2.23 and 2.24 are that both are three-variable input models and can only account for the effects of their respective three variables at their outputs. For example to have a complete two-harmonic PHD model, a Cardiff formulation of the type expressed in equation 2.25 is necessary. It becomes quickly apparent that adding more input variables to this model increases number of required dimensions of the model significantly (without accounting for higher harmonic effects which can be significant and important in predicting the power efficiency of transistors).

$$\tilde{B}_{ef} = \sum_{n}\sum_{r}\sum_{q} G_{ef,n,r,q}(|A_{11}|, |A_{21}|, |A_{22}|, |A_{12}|)e^{j(n\angle\tilde{A}_{21}+r\angle\tilde{A}_{22}+q\angle\tilde{A}_{12})} \tag{2.25}$$

Others have used other higher order functions as the describing functions of the PHD modeling framework which involves using piece-wise formulation of the PHD model [46]. In another instance, Padé approximation instead of a polynomial expression for the describing functions as seen in equation 2.26 [47]. A rational function has infinite polynomial terms in its equivalent Taylor series expression and thus a Padé approximation has the potential of fitting the nonlinearity of the device under test with fewer terms compared to a polynomial PHD model like the Cardiff model.

$$\tilde{B}_{ef} = \frac{G_{ef} + \sum_{gh} G_{ef,gh} \tilde{A}_{gh}}{1 + \sum_{gh} H_{ef,gh} \tilde{A}_{gh}} \tag{2.26}$$

A limitation of all the PHD models is that to increase the model capability to a high number of harmonics, the number of model parameters increases and the extraction of these models becomes more difficult. The number of describing functions required for a PHD model increases with the number of harmonics in the model. A power transistor model that has a DC describing function and n_h harmonics of behavioural modeling requires $2 \times (n_h + 1)$ describing functions.

We can note however that requiring $2 \times (n_h + 1)$ describing functions for a n_h harmonic time-invariant model is merely an artifact of the model being a frequency domain model since time-domain representations of nonlinear systems, like the Volterra series are purely defined in the time-domain and do not require separate describing functions for each harmonic. It is this discrepancy in the requirements of the PHD model and the theoretical possibility of modeling in the time-domain that motivates the research of this thesis.

2.7 Discussion on the State of the Art and Motivation for This Thesis

Two different approaches to modeling power transistors in order to design a power amplifier were discussed in this chapter, namely compact models and behavioural models. Though compact modeling of power transistors is theoretically very capable of modeling the effects and behaviours of power transistors in circuit simulators, they are often not available to the power amplifier designers. In addition, while compact models of power transistors have been shown to be very accurate in low and medium power applications, since the model construction for higher power transistors results in including many units of the same unit

die model, the model complexity grows to a level that it causes simulation convergence problems in a nonlinear circuit simulator. Instead, for high power transistors, amplifier designers can characterize the nonlinear behaviour of the power transistor under various RF impedance loads, a measurement known as load-pull, and use the characterization data directly to create behavioural models of the power transistors, and use computer aided design tools to assist in designing the RF power amplifier. The goal of this research is not to replace compact models but to augment them and allow for applications where compact models aren't feasible or where there is a benefit to using more simpler behavioural models. For instance, in a higher-order system level simulation where a single power amplifier is just one of many components of the overall system, behavioural models have an advantage of providing potentially less computational complexity.

Among the behavioural modeling strategies that have been discussed in the literature, the main ones suited for power transistor modeling are the Poly-Harmonic Distortion models. Many different variations of the PHD models were discussed. While some models required the measurement data to be on a fixed grid to allow for the extraction of model terms, this is something that is not always achievable and motivates a modeling scheme where such requirements aren't needed. There is also a desire to avoid the more extreme implementations of the PHD model where the models is basically an interpolation of a look-up table of load-pull data where it a high harmonic model becomes unfeasible since load-pull space of the required multi-dimensional look-up table would be impractically large to implement.

In an effort to simplify behavioural modeling that is based on load-pull measurements, it is the objective of this research to propose behavioural models that can fit any gathered load-pull measurement data, even if the data is on a non-uniform measurement grid, while also having the simplest construction and most compact form the behavioural model describing the nonlinear time-invariant system can have. It is also important that the modeling strategy proposed not rely on any sort of *guided characterization* (for example avoid the use of *tickler tones* used to extract X-parameters) where either large signal states are fixed and swept on uniform grids in order to extract specific model terms.

Another aspect of the various different PHD models proposed before the work of this thesis in the literature is the use of polynomial kernels in the construction of the models. Tackling *hard nonlinearities* is often difficult with polynomials models without including many higher order terms. A particular problem with using polynomial kernels is that polynomial terms have very bad extrapolation (as they go to either positive or negative infinity at each extrapolation extreme) and this results in very bad convergence of such models in nonlinear circuit simulators. Artificial neural networks (ANNs) are a particular form of nonlinear curve-fitting that has very good natural extrapolation. Partly due to the

use of bounded activation nonlinear functions (like the Sigmoid function) in the artificial neuron models, ANNs allow for the generation of very simulation-friendly nonlinear models.

The time-domain PHD (TD-PHD) model that will be proposed in Chapter 3 will use the same multi-harmonic frequency grid as an PHD model, but will describe the nonlinear behaviour using a single time-domain nonlinear expression at each port, essentially allowing for a more compact construction of the power transistor behavioural model than the PHD model to achieve the same time-invariant modeling capability. This allows for a fundamentally more compact expression of the behavioural model compared to a series of frequency-domain expressions. This compact form of the model construction allows the ANN implementation of the TD-PHD model to have the minimum number of output neurons compared to an ANN implementation of the frequency-domain PHD model. In fact, a fully-ANN implementation of the frequency-domain PHD model would require a separate ANN (with its own required training time) for each of the spectral outputs of the model. Since ANNs are often over-designed and have more neurons than the minimum amount needed to achieve the desired accuracy, an implementation of a TD-PHD model with ANNs will require less artificial neurons and less overall computation complexity than an implementation of a frequency-domain PHD model with ANNs.

In Chapter 4, the TD-PHD model will be generalized to load-pull measurements spanning multiple non-uniformly spaced fundamental frequencies, fulfilling the objective of making a single compact nonlinear time-invariant behavioural model that is based on general load-pull characterization data of a power transistor.

Chapter 3

Time-Domain Poly-Harmonic Distortion Models

This chapter will present the first major contribution of this thesis, the time-domain poly-harmonic distortion (TD-PHD) model. Section 3.1 will begin with an intuitive, less-technical argument for why time-domain poly-harmonic distortion (TD-PHD) models have the right structure and the appropriate degrees of freedom to compactly describe the behaviour of a nonlinear time-invariant system under periodic multi-harmonic stimulus.

Afterwards, in Section 3.2 a more technical derivation of the TD-PHD models will be presented based on an extension of how linear time-invariant discrete systems are modelled in the time-domain. This will be called the nonlinear impulse response formulation of TD-PHD models.

After this initial theoretical justification for the TD-PHD model, in Section 3.3 the development method of TD-PHD models, which includes the methodology of power transistor characterization, the extraction of the proposed behavioural model from multi-harmonic the load-pull measurement data, the implementation of the extracted model as a Netlist in a harmonic balance simulator, and finally, the measurement validation of the model will be presented before the conclusion of the chapter.

3.1 The Intuitive Argument for TD-PHD Models

The poly-harmonic distortion models target the periodic response of a nonlinear time-invariant system. This means that the state of the system and the device when it is being

simulated is such that all the voltages and currents are periodic with a fixed fundamental frequency, and hence the time-domain waveforms of the voltages and currents repeat in time with a fixed fundamental period. The harmonic balance simulation is well suited for this type of periodic simulation of a nonlinear system as the nonlinearity of the system does not result in the generation of frequencies that are not an integer multiple of the fundamental frequency of the simulation, that is, if a nonlinear system is stimulated with a periodic stimulus, the nonlinearity will not change the periodicity of the the resulting stimulus but instead the nonlinearity can lead to the generation of harmonic RF content.

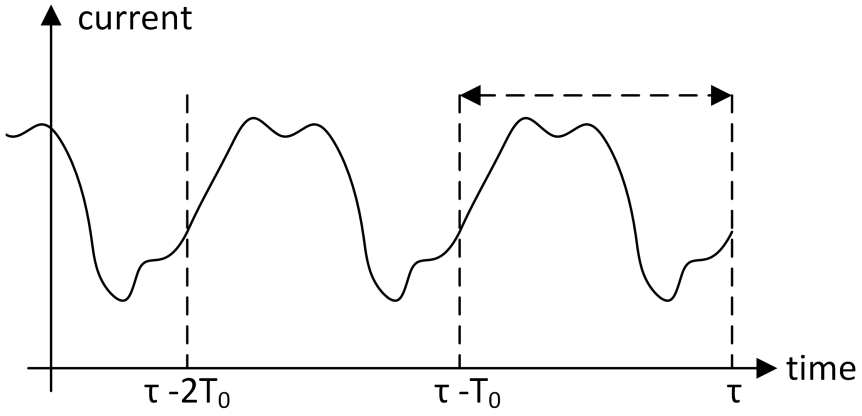

Figure 3.1: A single period of a periodic signal fully represents the waveform over all time

When the nonlinear time-invariant system is stimulated at its inputs with periodic signals that have a period of length T_0 (and fundamental angular frequency of $\omega_c = \frac{2\pi}{T_0}$), the outputs of this system will also be periodic signals with the same fundamental period like the waveform in Figure 3.1. The non-sinusoidal shape of the periodic waveform is produced by the harmonic content of the waveform. What is referred to as the large-signal behaviour of the system is fully represented by its voltage and current time-domain waveform shapes over the fundamental period.

A general time-domain behavioural model for poly-harmonic periodic stimulus must map the time-domain shape of the the periodic input waveforms onto the time-domain shape of the periodic output waveforms. The behavioural model should be able to predict the output waveform of the system for any given input waveform to the system at that fundamental frequency. As a demonstrative example, a 1-port nonlinear system will be considered. The nonlinear behaviour of this system can be described as an *admittance mapping* between the shape of the voltage waveform over its fundamental period to the shape

of the current waveform over that same fundamental period. The model could have been alternatively described as a *RF scattering mapping* by mapping the time-domain shape of the incident $a(t)$ waveforms onto time-domain shape of the reflected $b(t)$ waveforms, where the time domain signals $a(t)$ and $b(t)$ are related to the time domain signals $v(t)$ and $i(t)$ through the following two equations:

$$a(t) = \frac{v(t) + Z_0 i(t)}{2\sqrt{R_0}} \tag{3.1}$$

$$b(t) = \frac{v(t) - Z_0 i(t)}{2\sqrt{R_0}} \tag{3.2}$$

It should be noted that the choice of the input and output variables does not matter for the circuit simulator as either types of model, an admittance-type or RF-scattering-type, can be implemented and used within modern harmonic balance simulators.

In order to capture the *shape* of the periodic waveform in a discrete simulator, the time-domain waveform would have to be discretely sampled over its fundamental period with a sampling that is evenly spaced in time, with enough discrete time resolution to capture the harmonic content that is generated. Since the response is periodic, the only thing that is necessary for this mapping is to map the behaviour of the system over a single period. In this time-domain view of the voltage and current signals, the sampled value of the voltage and current waveform at each one of the N fixed sampling times spanning the fundamental period at each port of the device will be referred to with a separate variable in this behavioural modeling scheme. Based on Nyquist theory if the maximum non-aliased frequency of the system is at harmonic-order h_m (where h_m is a positive integer), at least $N = 2h_m + 1$ time-samples will be required in the fundamental period of the signals in order to be able to capture frequency content up to an order of h_m.

The continuous time-domain voltage waveform $v(t)$ will be discretized into an indexed vector of sampled voltage $v[k]$ where index k goes from 0 to $N - 1$, such that in this notation, $v[0] = v(0)$, $v[1] = v(T_0/N)$, $v[2] = v[2T_0/N]$, $v[3] = v(3T_0/N)$, and so on, until $v[N-1] = v((N-1)T_0/N)$ where T_0 is the fundamental period of the signal. A similar time-vector discretization notation will be used for the time domain current waveform $i(t)$ into an indexed vector of sampled currents $i[k] = i(kT_0/N)$.

N functions (labeled f_0 through f_{2h_m} in the formulation of equation 3.3), one for each fixed sampling time, each having $N = 2h_m + 1$ input arguments as shown in Figure 3.2 can be used to define the most general behaviour model that can be defined in the time domain

given this particular discretization of the time-domain signals. These functions will have the general mapping as follows:

$$\begin{aligned} i[0] &= f_0(v[0], v[1], \cdots, v[2h_m]) \\ i[1] &= f_1(v[0], v[1], \cdots, v[2h_m]) \\ &\cdots \\ i[2h_m] &= f_{2h_m}(v[0], v[1], \cdots, v[2h_m]) \end{aligned} \tag{3.3}$$

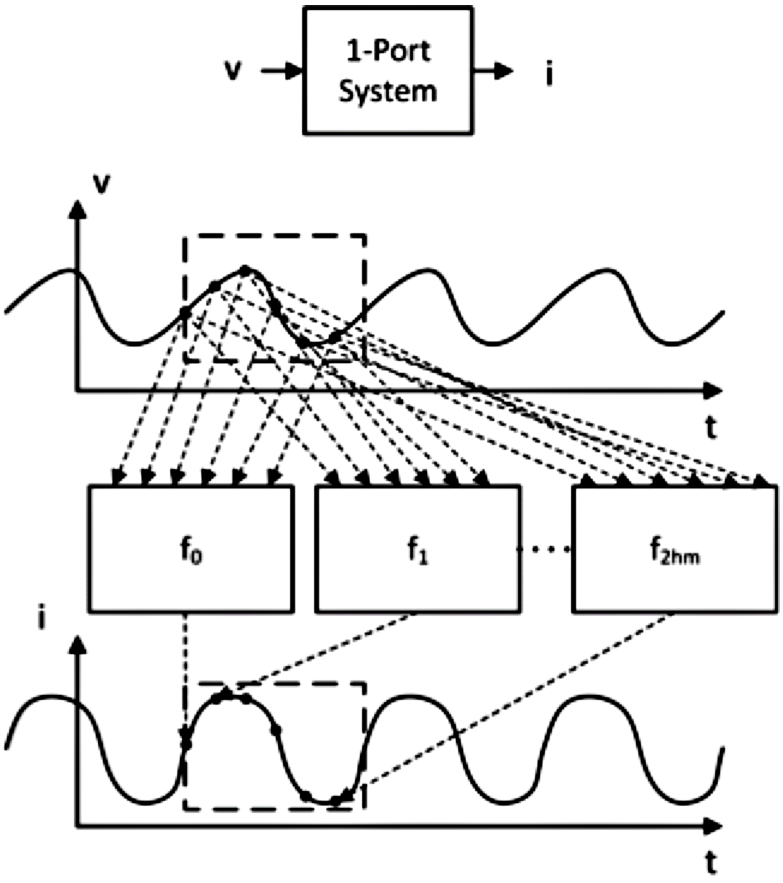

Figure 3.2: A basic non-time-invariant but general discrete waveform shape mapping scheme

This most general representation of a behavioural model as described in equation 3.3 and shown in Figure 3.2 while being theoretically complete in that there should exist a set

of functions that can map the input to output behaviour for any given periodic input as described, it does not account for the fact that the underlying system that is being modeled is a time-invariant system. This non-account for time-invariance stems from the fact that the each of the inputs to the model are fixed to a specific simulation time. This results in N independent functions f_0 through f_{2h_m} being required to describe a system that could have otherwise been describe with a time-invariant description that isn't fixed in time.

The main feature of a time-invariant system is that a delay in the input signal (shifting the input waveform forwards or backwards in time) will result in the same output waveform but advanced or delayed in time by the same amount. Thus this representation can be simplified further under the assumption that the system is time-invariant. Time-invariance of a function f_{TI} implies that if the time-domain output at a given time t is $i(t) = f_{TI}(v(t))$, then the time shifted output at a different time $t+\Delta t$ is $i(t+\Delta t) = f_{TI}(v(t+\Delta t))$ where Δt is a time-offset. This allows the following full time-domain and time-invariant behavioral model formulation for stimulus of fundamental period T_0 using a single multi-variate time-invariant function f:

$$\begin{aligned} i[0] &= f(v[0], v[1], \cdots, v[2h_m - 1], v[2h_m]) \\ i[1] &= f(v[1], v[2], \cdots, v[2h_m], v[0]) \\ &\cdots \\ i[2h_m] &= f(v[2h_m], v[0], \cdots, v[2h_m - 2], v[2h_m - 1]) \end{aligned} \tag{3.4}$$

Here the $(2h_m + 1)$ independent functions f_0 through f_{2h_m} are replaced with a single time-invariant function f that gives the entire discrete-time periodic output as a function of discrete-time periodic input.

Alternatively using a circular index for the discrete periodic representation where $v[k] = v[k \pm n(2h_m + 1)]$ for all values of $n \in \mathbb{N}$, we can write the expressions of equation 3.4 in the unified expression of:

$$i[k] = f\left(v[k], v[k-1], v[k-2], \cdots, v[k-2h_m]\right) \tag{3.5}$$

The time-invariant models described by equation 3.5 will be henceforth referred to as *Time-Domain Poly-Harmonic Distortion (TD-PHD)* models.

This is the simplest form a nonlinear time-invariant system that is to map the shape of the periodic waveform in the time-domain can hold. Comparing this formulation to the equivalent of an PHD model as seen in equation 2.20, it becomes apparent that only a

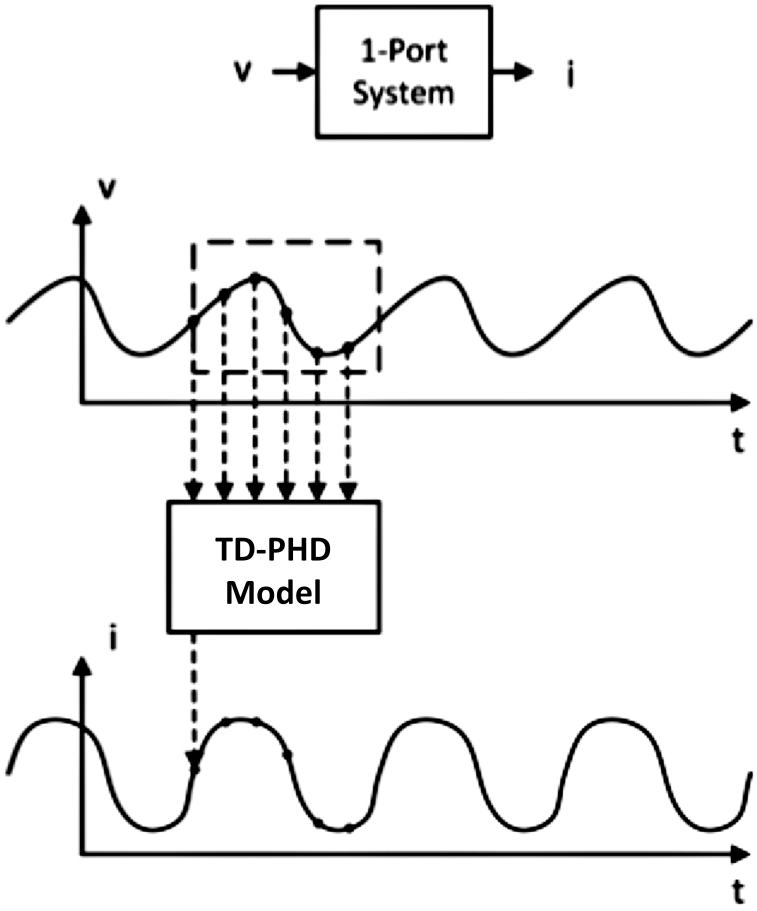

Figure 3.3: A time-invariant discrete waveform shape mapping scheme

single output variable and hence a single describing function is required to describe a 1-port nonlinear time-invariant system regardless of the number of harmonics that were involved in the model in the TD-PHD model while a separate describing function is required for the addition of every harmonic term to the PHD model.

The single-port TD-PHD model can be extended to multi-port systems. Equation 3.5 can be generalized up to n ports to define the n-port behavioral model as follows (the subscripts in this notation represent the port number):

$$
\begin{aligned}
i_p[k] = f(&v_1[k], v_1[k-1], \cdots, v_1[k-2h_m], \\
&v_2[k], v_2[k-1], \cdots, v_2[k-2h_m], \\
&\cdots, v_n[k], v_n[k-1], \cdots, v_n[k-2h_m])
\end{aligned}
\tag{3.6}
$$

The TD-PHD model accounts for harmonic information in the shape of the time-domain waveform by incorporating a denser sampling in the time-domain. The more points that are used to sample the input periodic signals in the model, the equivalent harmonic resolution of the model increases.

In the above N was set to be $2h_m + 1$. This is due to the Nyquist criteria for discrete sampling of a signal, the maximum non-aliased harmonic frequency that can be sampled is related to the number of sampling over the fundamental period. In the next sub-section the TD-PHD model will be further developed with discrete sampling theory in mind.

3.2 The Nonlinear Impulse Response Formulation of TD-PHD Models

In this section, the frequency-domain PHD model will be viewed from a different perspective, one that is based on time-invariant discrete systems modeling theory.

A common characteristic of all the frequency-domain PHD models discussed in the previous chapter was that the Poly-Harmonic Distortion describing functions at each of the frequencies are defined independently of all the other describing functions at the other harmonics. This was the case even though a common underlying nonlinearity has resulted in the generation of all the spectral content. When considering the form of general time-invariant behavioral models in the time-domain, namely impulse response functions for linear systems and the Volterra series for nonlinear systems, it can be noted that the description of the time-invariant system in the time-domain definition only requires the description of the output of the system at an arbitrary point in time since the output of the system at other arbitrary times can be inferred using the the time-invariance property of the description. In a time-invariant description of a system, the value of the output signal at an arbitrary point in time is only a function of the values of the input signal at time offsets *relative* to that arbitrary point in time.

This motivated the proposed time-domain PHD model [48][49] that mapped the input signal behaviour onto the output signal behaviour on the same multi-harmonic frequency

grid as the frequency-domain PHD models, but using a single time-domain nonlinear expression at each port to achieve this end. This is as opposed to requiring a different describing function for each frequency like it is required for PHD models. A compact expression of a time-invariant system is beneficial if an advanced continuous function fitting tool like artificial neural networks are used for the model fitting problem. For poly-harmonic models of time-invariant systems, a time-domain expression is more compact than a frequency domain expression as it requires fewer output variables.

The frequency domain poly-harmonic distortion model describes the nonlinear time-invariant dynamics of the power transistor using a frequency-domain basis of the input signal. Since the harmonic balance circuit simulation is focused on the periodic response of the system, at every simulation state the simulator solves for the voltages at every node and the currents in every branch of the Netlist. The nonlinear time-invariant system will generate harmonic content in response to a fundamental frequency stimulus, resulting in waveforms that diverge from a sinusoidal shape and include harmonic content. Since the periodic response of the circuit at a given fundamental frequency is being simulated, the voltages and currents are represented by their Fourier series coefficients, a set of complex numbers that provide the phasor representation of the cosine wave component at each harmonic frequency. These coefficients include the DC bias, the fundamental frequency component which is the dominant signal in an amplifier, and all the harmonic frequencies that are at integer multiple frequencies of the fundamental frequency and are generated from the nonlinear elements of the circuit.

The discrete frequency-domain representation of the periodic signal has an equivalent discrete time-domain representation. The periodic signal can be either represented as a sum of cosine waves at each of the harmonic frequencies or as a discrete sampling of the continuous-time periodic waveform. These two representations, namely the Fourier series coefficients and the sampling of the periodic waveform within its fundamental period are related to each other via the discrete Fourier transform. The Fourier series representation of the distorted periodic signal represents the shape of the waveform as the sum of the DC and harmonic sinusoidal waves. Alternatively a discrete sampling of the same continuous waveform within its fundamental period can represent the continuous waveform given enough samples within the fundamental period. An example of this duality is shown in Figure 3.4.

In this section, for simplicity of representation, it is assumed that the nonlinear system of interest is a single-port system. That is, the system interacts with the surrounding world at one port and the behavioral model consists of the mapping of the incident $a_1(t)$ waveform at that port onto the reflected $b_1(t)$ waveform from that same port. The extension to 2-port (or multi-port) systems from the single-port formulation is trivial and will be performed

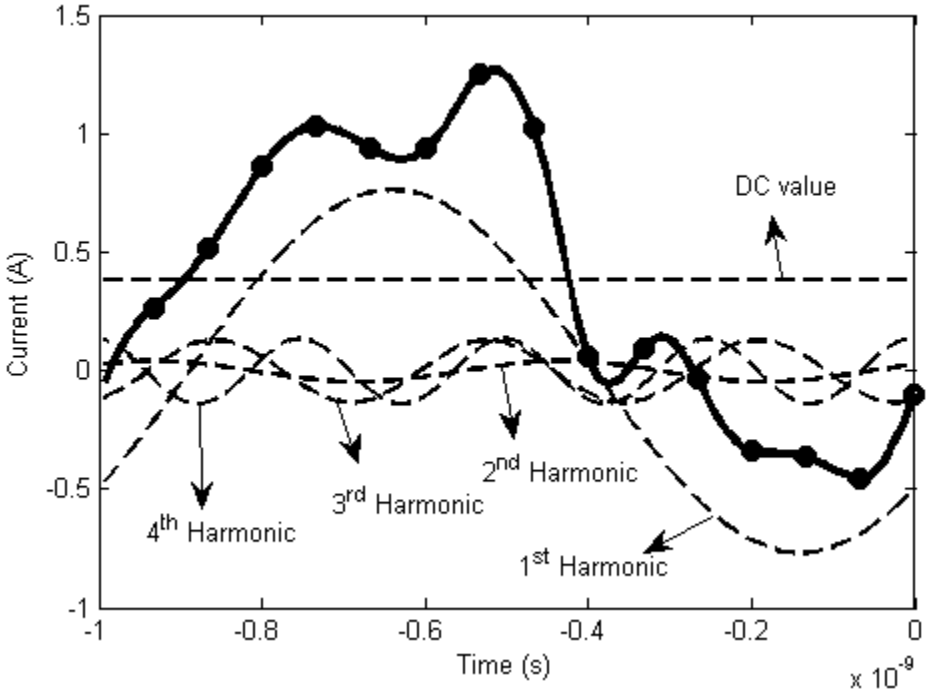

Figure 3.4: A periodic signal over its fundamental period (thick line) can be equivalently represented as either a sum of discrete harmonic frequency components (represented by waveforms with dashed lines) or as discrete time representation (represented by the discrete dots on the thick line)

at one step at the end in order to have a model that can be applied to power transistors.

S-parameters are a linear time-invariant behavioural model defined in the frequency domain at a discrete set of frequencies. The discrete transfer function defined by an S-parameter model at a set of harmonic frequencies 0, f_0, $2f_0$, $3f_0$, ..., hf_0 can be alternatively described by an equivalent finite impulse response (FIR) model. This FIR model is just the inverse Fourier transform of the S-parameters on the discrete frequency grid. Note that since the input and output waveforms ($a_1(t)$ an $b_1(t)$) are real-valued over time, the value of the Fourier series coefficients A_{1k} at negative frequencies is the complex-conjugate of the Fourier series coefficients at positive frequencies.

$$a_1(t) = \sum_{k=0}^{h} \frac{1}{2} \left(A_{1k} e^{j2\pi k f_0 t} + A_{1k}^{*} e^{-j2\pi k f_0 t} \right) \tag{3.7}$$

$$b_1(t) = \sum_{k=0}^{h} \frac{1}{2} \left(B_{1k} e^{j2\pi k f_0 t} + B_{1k}^* e^{-j2\pi k f_0 t} \right) \tag{3.8}$$

If the S_{11} parameters at the harmonic frequencies are labeled as S_k for the k-th harmonic parameter, then the vector of S-parameters is related to the vector of FIR coefficients h_S through the inverse discrete Fourier transform (IDFT):

$$\begin{aligned} \{h_S[0], h_S[1], h_S[2], \cdots, h_S[2h-2], h_S[2h-1], h_S[2h]\} = \\ \text{IDFT}\left\{S_h^*, S_{h-1}^*, \cdots, S_2^* S_1^*, S_0, S_1, S_2, \cdots, S_{h-1}, S_h\right\} \end{aligned} \tag{3.9}$$

This S-parameter FIR model can be seen as having the form:

$$b_1(t) = \sum_{k=0}^{2h} a_1(t - kt_d) h_S[k] \tag{3.10}$$

where $t_d = \frac{T_0}{2h+1}$ is the delay between the memory taps and T_0 is the fundamental period.

The *memory depth* of the FIR model is limited to within the fundamental period (T_0) and the *taps* of the FIR model are equally spaced in time (with the spacing t_d) within that range. Increasing the number of taps of the FIR model when the memory depth is fixed to T_0 is equivalent to increasing the "sampling rate" of the FIR model, which in turn increases the maximum harmonic frequency that the FIR model can act upon. $2h+1$ taps in the FIR model ensures that h harmonics are within the Nyquist rate of the FIR model sampling frequency.

The nonlinear model in the time domain will have to be a generalization of this FIR model in terms of the time-domain samples used in the model. The basis for the nonlinear model will be the same as the FIR taps (the sub-period fractional delayed inputs $a(t-kt_d)$). This gives the general form of the Time Domain Poly-Harmonic Distortion (TD-PHD) model for a 1-port system:

$$b_1(t) = f_1\left(a_1(t), a_1(t-t_d), a_1(t-2t_d), \cdots, a_1(t-2ht_d)\right) \tag{3.11}$$

where $t_d = \frac{T_0}{2h+1}$ and T_0 is the fundamental period and the TD-PHD describing function f is a real-valued analytic function. In other words, the time-domain output $b_1(t)$ at any time t is a multivariate nonlinear function of the value of the input a_1 at all the different

time-offsets from the current time t. This creates a generalization of the discrete impulse response used in linear systems to nonlinear discrete time-invariant systems. In a way, the TD-PHD model can be seen as a nonlinear impulse response model. And conversely when the system is linear, the TD-PHD model collapses into being just a linear real-valued FIR model.

Generalizing the scattering waveform TD-PHD model to the multi-port case, the inputs to the TD-PHD model at every port will be sub-period delayed values of the inputs at all ports of the device. The TD-PHD model for the k^{th}-port of an n-port system up to harmonic order h can be seen as a multi-dimensional nonlinear impulse response function of the following form:

$$\begin{aligned} b_k(t) = f_k(&a_1(t), a_1(t-t_d), \cdots, a_1(t-2ht_d), \\ &a_2(t), a_2(t-t_d), \cdots, a_2(t-2ht_d), \\ &\cdots, \\ &a_n(t), a_n(t-t_d), \cdots, a_n(t-2ht_d)) \end{aligned} \tag{3.12}$$

The two-port TD-PHD model is the form of the model that will be used to model high-power packaged power transistors for the design of high power RF amplifiers.

3.3 Development of Power Transistor TD-PHD Models

To develop a TD-PHD model for a power transistor, one must first start with a representative load-pull characterization of the power transistor behaviour that is intended to be modeled. This characterization should stimulate the power transistor into states similar to ones where it will be operating in when designed into a power amplifier circuit. In Section 3.3.1 the unique characterization setup that was developed to perform multi-harmonic hybrid load-pull measurements on a GaN HEMT power transistor will be presented.

The characterization data captured will be the basis for behavioural model generation. For the generation of TD-PHD models, the measurement data will have to be discretized in the form that coincides with the input variables of the TD-PHD model so a direct fit model can be extracted from from the modeling data. Then the model describing functions, whether polynomials or artificial neural networks, must be fit the load-pull characterization

data to a good degree of accuracy for the behavioural model to be useful for power amplifier design. In Section 3.3.2 the process to extract a TD-PHD model from multi-harmonic load-pull characterization data will be outline.

After extracting the model parameters, these parameters must be implemented in a simulation Netlist that describes the TD-PHD model as it is described in this thesis. Section 3.3.3 will describe the implementation of the TD-PHD model in a harmonic balance simulator in more detail.

Finally the implementation of the extracted TD-PHD model will have to be validated against measurements of the power transistor that was characterized in order to show that the model can predict the behaviour of the power transistor in a simulated environment, allowing for the design of power amplifiers with multi-harmonic behavioural models that take the effects of harmonic terminations of nonlinear circuits into account. Section 3.3.4 will present a comparison of the measurement of the large signal multi-harmonic behaviour of the power transistor compared to the large signal simulation of the extracted TD-PHD behavioural model. Figure 3.5 summarizes the steps required to develop a TD-PHD behavioural model of a power transistor device.

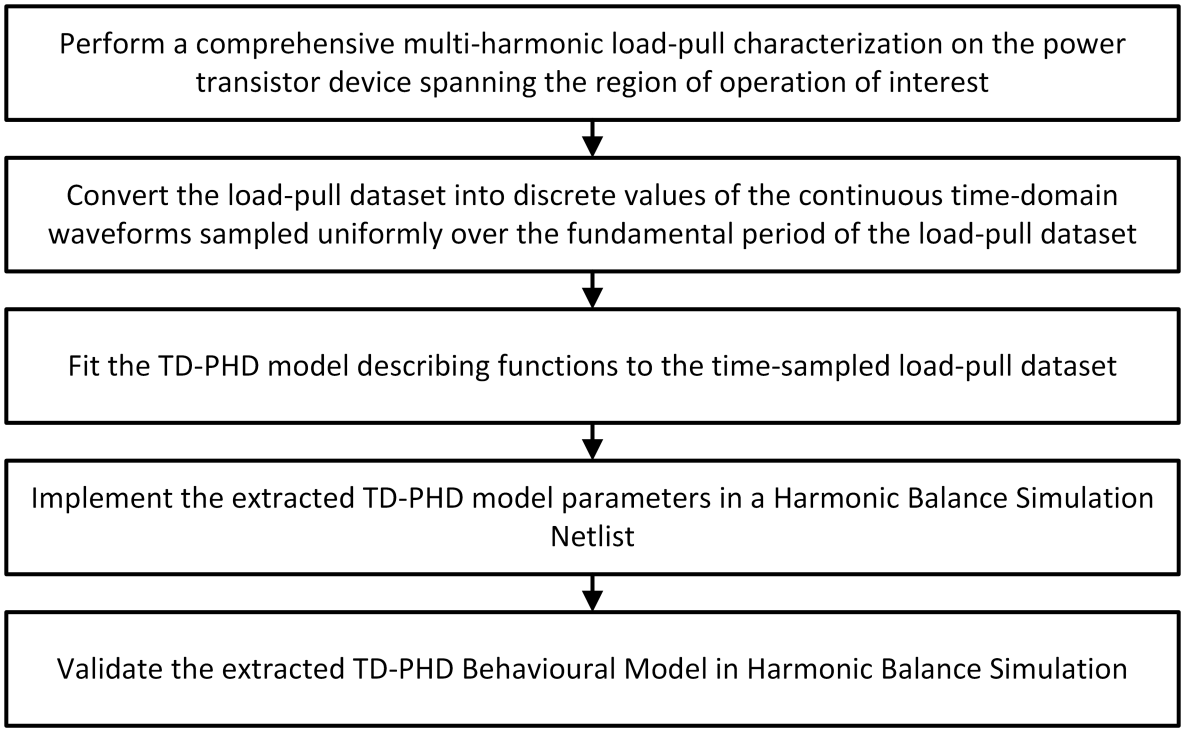

Figure 3.5: A flowchart of the steps required to develop a TD-PHD model

3.3.1 Multi-harmonic Loadpull Characterization Of Power Transistors

For a measurement validation of the TD-PHD model's ability to simulate the poly-harmonic large-signal behaviour of a power transistor, a GaN HEMT packaged RF power transistor device that is sensitive to harmonic impedance terminations was biased at a fixed quiescent current, and large signal load-pull measurements were performed on it. The objective was to obtain a single simulation model that can encapsulate all the periodic behaviour of a the power transistor at a fixed fundamental period and a fixed DC bias point.

Figure 3.6 shows a diagram of the measurement testbench that was used to measure the large signal behaviour of a power transistor. This setup is a unique hybrid multi-harmonic passive/active load/source-pull system that is devised for a two-port transistor measurement. The passive 3-harmonic tuners (Focus Microwave MPT-1818-TC) allow for setting the source and load impedances at 3 harmonic frequencies, but in this setup, these passive tuners are used as a pre-match for the active load-pull measurement. In this setup

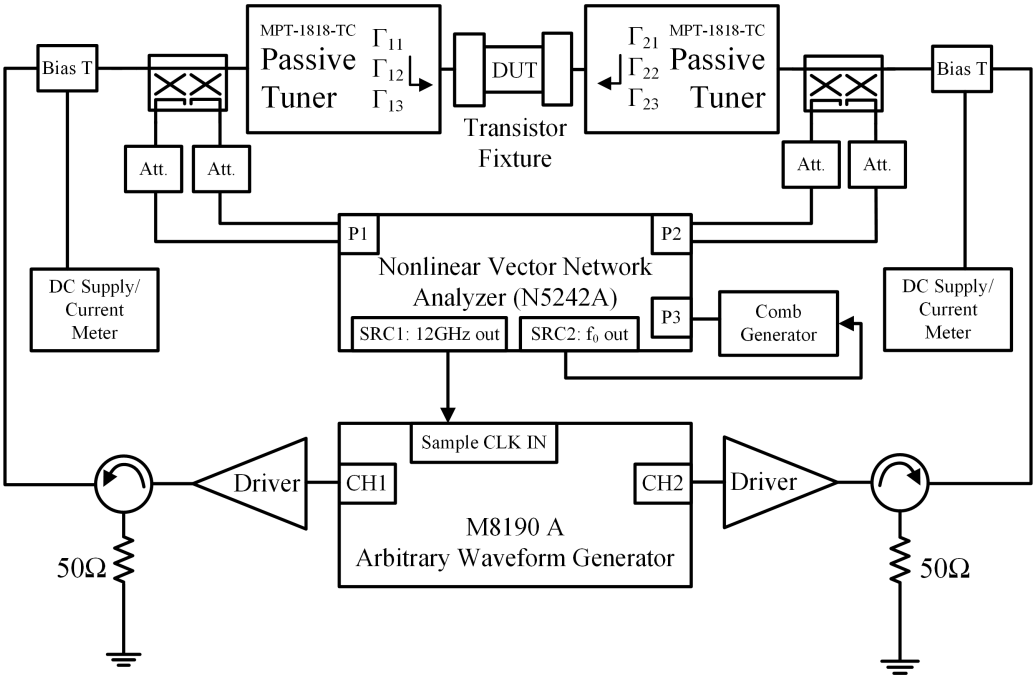

Figure 3.6: Hybrid passive/active waveform engineering and large-signal characterization testbench.

the active harmonic signals that will be used for harmonic load-pull are generated using an arbitrary waveform generator with a very high sample generation rate. This allows the setting of the relative amplitude and phase of the injected signal at the harmonic frequencies with just a single signal source that will output a multi-harmonic waveform. High power drivers in front of the active sources allows for the generation of higher magnitude reflection coefficients beyond the capability of the passive tuners alone. Two phase-coherent channels of a Keysight M8190A 12 GSa/s Arbitrary Waveform Generator (AWG) were used to generate an active signal at the fundamental and second harmonic frequencies (2GHz and 4GHz respectively). The AWG was synchronized with the Keysight N5242A PNA-X NVNA via a 12GHz sample clock feed from the NVNA. The NVNA uses its second phase coherent source to drive the phase reference of the NVNA allowing for the recreation of the time-domain calibrated waveforms at the RF calibration plane.

The power transistor device itself will be placed in a transistor fixture which includes a Thru-Reflect-Line (TRL) calibration kit that allows us to extract the S-parameters of transistor fixtures on each port. The measurements performed at the RF calibration plane will be de-embedded to the power transistor lead plane. The DC bias of the power transistor

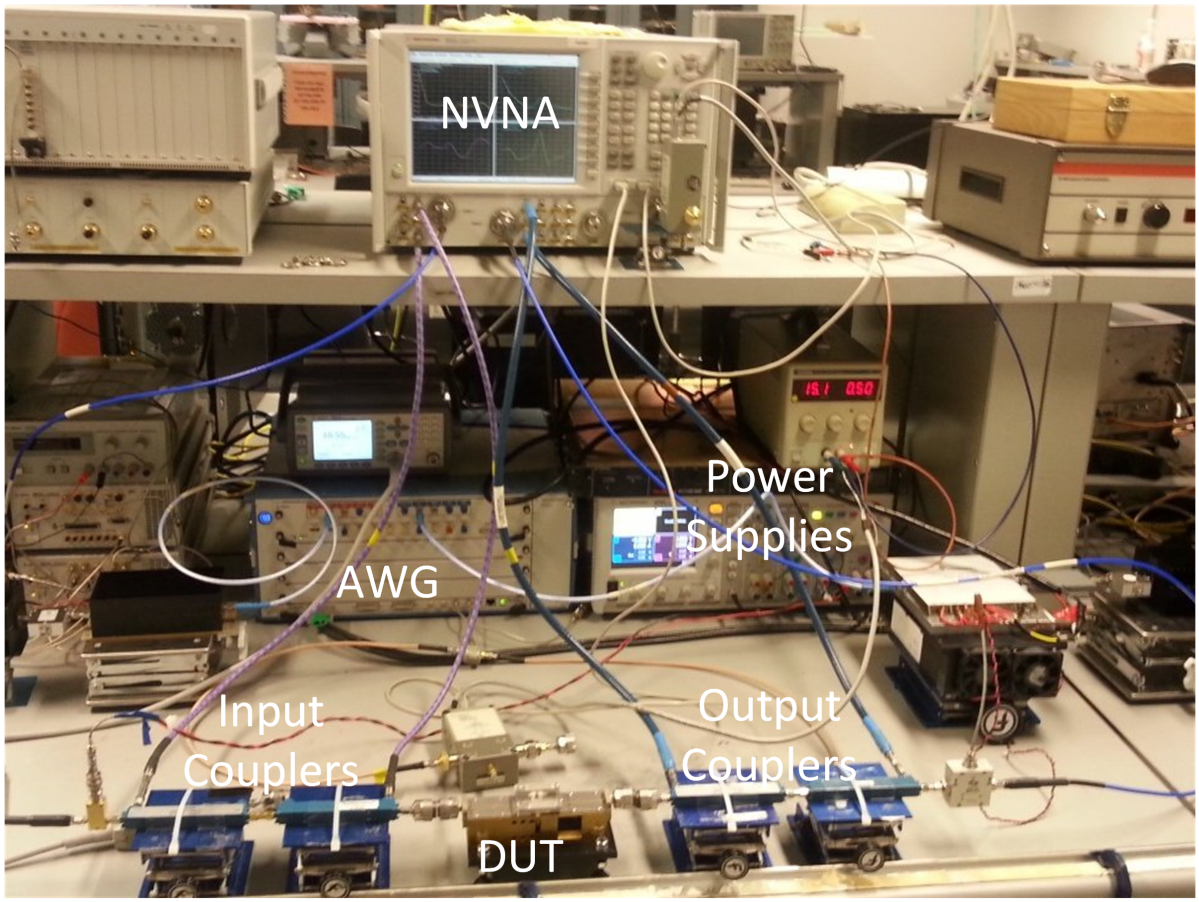

Figure 3.7: Active waveform engineering and large-signal characterization testbench during calibration

is provided through external DC bias-tees. DC measurements on the ports of the device are performed in parallel with the RF measurements to keep track of the power efficiency of the power transistor while the RF impedances and input power presented to the power transistor is varied.

Note that this is an open-loop active load-pull system and thus there will be no absolute impedance control, but an accurate characterization will be performed on the power transistor regardless. In this open-loop active loading scheme, a signal will be injected onto the drain side of the power transistor device to deviate the impedance that is presented to the transistor due to the passive termination provided by the passive tuners. However, this final impedance that is presented to the device will not be controlled. Since the objective

here is behavioural modeling, any nonlinear behaviour captured within the range of useful RF performance of the power transistor can be used to develop the model and thus having an open-loop active load-pull system, as opposed to a closed-loop one, would not be an impediment. Since an NVNA measurement allows for the determination of the time-domain waveforms of the drain current and drain voltage, the instantaneous peaks of the drain to source voltage $V_{ds}(t)$ can be monitored to avoid getting close to the device breakdown voltage. This will put a natural limitation on how hard the device will be pushed during large signal characterization.

This open-loop active loading scheme only allows for active impedance variation over two harmonics. An NVNA characterization that goes up to the higher harmonics (beyond the third order) is performed even though there is no direct control over the impedances and signals injected at those higher harmonic frequencies in the hopes that this higher harmonic variation that is captured in the multi-harmonic load-pull data can reveal the poly-harmonic behaviour of the device under test given the fact that a poly-harmonic behavioural model is being used that can theoretically respond to those higher harmonic effects. This was important for the measurement strategy that was devised to get a *representative* multi-harmonic load-pull measurement data-set that randomly spans the measurement space of interest while having enough variation within the harmonic load and source terminations.

Using the measurement testbed of Figure 3.6, a 10W GaN packaged RF transistor (Wolfspeed CGH40010) was biased at a quiescent DC drain current of 45mA (Class AB) and was subject to a 2-harmonic open-loop active load-pull with NVNA characterization at 5 harmonics of a fundamental frequency of $f_0 = 2$GHz.

In this research, a time-domain scheme for the generation of the active injection scheme was devised that is in-line with the time-domain signal generation provided by the arbitrary signal generators. Considering that the test setup of Figure 3.6 uses an arbitrary waveform generator to create the multi-harmonic incident RF waves, a limit of $10dB$ of fundamental-to-harmonic power was used in the random active signal generation, to ensure that the randomly generated signals are closer to the type of fundamental frequency amplification application that the power transistor will be used for in an RF power amplifier design. Two time-domain active signals were used for the two-port active load-pull (one for the source side, and the other for the load side) with a second harmonic component with a fixed magnitude and phase compared to the fundamental frequency component, effectively freezing the time-domain *shape* of the injected signal. While fixing the shape of the injected $a_1(t)$ and $a_2(t)$ signals, the relative magnitude and sub-fundamental-period time-delay of these signals were then swept. The relative magnitude and time-delay sweep of the $a_2(t)$ with reference to $a_1(t)$, ensures that the reflected Γ_{21} and Γ_{22} impedances at

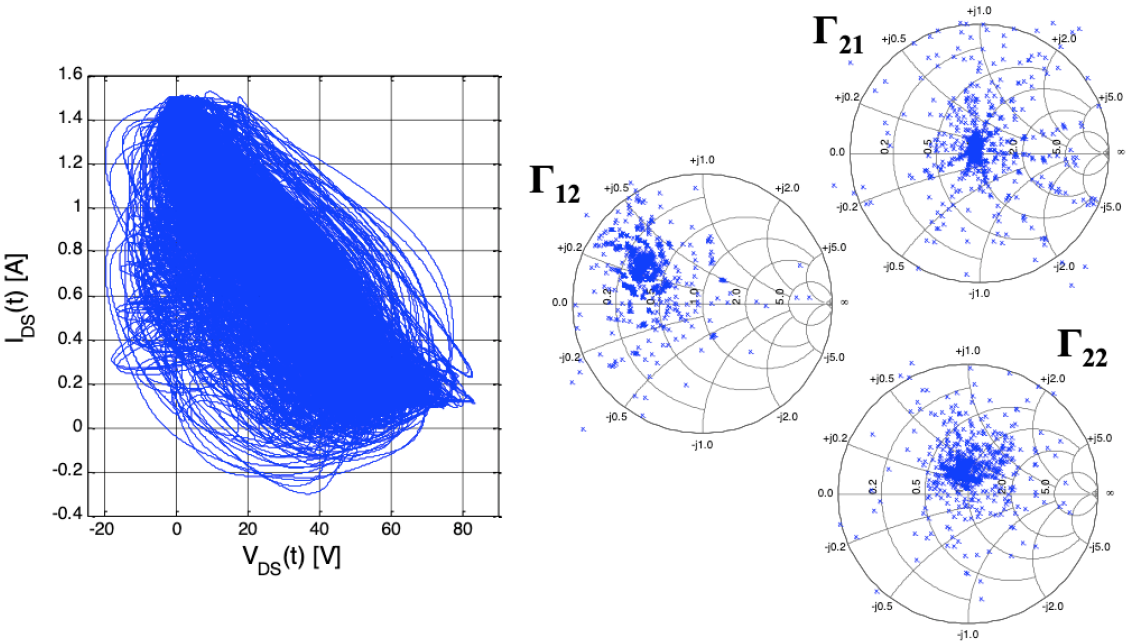

Figure 3.8: The variation of the multi-harmonic load-pull characterization dataset in terms of both the harmonic impedances as well as the dynamic load-lines of the device at the power transistor reference plane

the fundamental and second harmonic frequency respectively are swept in every direction around the passive impedance provided by the passive impedance tuners. $N_{\text{shapes}} = 10$ random $a_1(t)$ and $a_2(t)$ *shapes* were generated and for each of these combinations, active load-pull measurements are performed to sweep the $N_{magnitudes} = 25$ signal magnitudes and $N_{delays} = 11$ sub-fundamental-period time-delays of the two active multi-harmonic signals to result in a total of 2750 NVNA measurements for the entirety of the load-pull measurement space. Figure 3.8 shows the total variation of the first and second harmonic reflection coefficients for this randomly generated load-pull characterization scheme. The time-domain waveforms were also plotted with the time-dimension removed to create *dynamic load-lines* comparing the $I_{DS}(t)$ vs $V_{DS}(t)$ behaviour showing that the characterization has spanned and stimulated a big part of the underlying nonlinearity of the power transistor itself.

This extracted load-pull characterization set at a fixed fundamental frequency of $f_0 =$ 2GHz was then formatted in a way to make TD-PHD model extraction feasible as described in Section 3.3.2.

3.3.2 TD-PHD Model Extraction From Multi-Harmonic Load-Pull Measurements

A power transistor can be modelled as a two-port system, where one port is the gate to source terminal of the device and the other is the drain to source terminal. The two-port version of equation 3.12 requires two TD-PHD expressions to be modelled:

$$\begin{aligned} b_1(t) = f_1(&a_1(t), a_1(t-t_d), \cdots, a_1(t-2ht_d), \\ &a_2(t), a_2(t-t_d), \cdots, a_2(t-2ht_d)) \end{aligned} \tag{3.13}$$

$$\begin{aligned} b_2(t) = f_2(&a_1(t), a_1(t-t_d), \cdots, a_1(t-2ht_d), \\ &a_2(t), a_2(t-t_d), \cdots, a_2(t-2ht_d)) \end{aligned} \tag{3.14}$$

where h is the maximum harmonic order of the TD-PHD model and $t_d = T_0/N$ is the sub-period spacing between the inputs to the model and $N = 2h+1$ is the time-resolution parameter of the model.

Each of the functions f_1 and f_2 are static, or memory-less expressions that have to hold over the entire load-pull dataset. While the signals $b_1(t)$, $b_2(t)$ $a_1(t)$, $a_2(t)$, $a_1(t-t_d)$, $a_2(t-t_d)$, and so on are periodic and have values over all time, the functions f_1 and f_2 only apply to a single instant in time. These two nonlinear functions have to fit the entire load-pull measurement space over all time. To extract the functions f_1 and f_2, all the periodic measurements will need to be converted into their time-domain representations. Each one of the periodic measurements of the signals included in equations 3.13 and 3.14 will need to be sampled uniformly over the fundamental period. To extract the proposed model, the load-pull measurement dataset will need to be converted from the frequency domain to the time domain by means of a Fourier Series evaluation of the DC, fundamental and harmonic frequencies:

$$\begin{aligned} a_1(t) &= V_{\text{gate,DC}} + \Sigma_{k=1}^{h} |A_{1k}| cos\left(2\pi k f_0 t + \angle A_{1k}\right) \\ a_2(t) &= V_{\text{drain,DC}} + \Sigma_{k=1}^{h} |A_{2k}| cos\left(2\pi k f_0 t + \angle A_{2k}\right) \\ b_1(t) &= I_{\text{gate,DC}} + \Sigma_{k=1}^{h} |B_{1k}| cos\left(2\pi k f_0 t + \angle B_{1k}\right) \\ b_2(t) &= I_{\text{drain,DC}} + \Sigma_{k=1}^{h} |B_{2k}| cos\left(2\pi k f_0 t + \angle B_{2k}\right) \end{aligned} \tag{3.15}$$

If the model order or the harmonic order of the model is too low, the model fitting algorithm will have difficulty in finding a good fit to the data. A good model fitting threshold would be a Normalized Mean Squared Error (NMSE) of better than -30dB for the time-domain measurement dataset. The NMSE for a parameter y that is modeled with the variables $y_{\text{model,i}}$ and measured with variables $y_{\text{meas,i}}$ over the measurement dataset is defined as follows:

$$NMSE_y = \frac{\sum |y_{\text{model,i}} - y_{\text{meas,i}}|^2}{\sum |y_{\text{meas,i}}|^2} \tag{3.16}$$

For this load-pull measurement dataset, since there are at least 5-harmonics of information in the time-domain data, the time-domain resolution of the model N has to be at least 11, meaning that a set of 22 input variables will be used for each output of this 2-port model.

Since the entire load-pull measurement data set is periodic with the period T_0, the output functions $b_1(t)$, $b_2(t)$, and the input functions $a_1(t)$, $a_1(t-t_d)$ through $a_1(t-2ht_d)$, and $a_2(t)$, $a_2(t-t_d)$ through $a_2(t-2ht_d)$ are evaluated at times $t=0$, $t=t_d$, $t=2t_d$, $\cdots$, $t=10t_d$. This means that each frequency-domain load-pull measurement at a fixed power level in the dataset will be converted into 11 equivalent discrete time-sampled data points. This will be the discrete time-domain dataset used for fitting the simulation-time-independent nonlinear output functions f_1 and f_2 that implement the following TD-PHD mappings:

$$b_1(t) = f_1\left(a_1(t), a_1(t-t_d), \cdots, a_1(t-10t_d)\right) \tag{3.17}$$

$$b_2(t) = f_2\left(a_1(t), a_1(t-t_d), \cdots, a_1(t-10t_d)\right) \tag{3.18}$$

Figures 3.9 and 3.10 show a graphical representation of how the time-domain input and output signals are discretized respectively. The main insight that is illuminated in Figure 3.9 is that at each instance in time, having access to all the delayed versions of each of the input signals allows surrogate access to the value of the input signals at other points in time. This allows the TD-PHD model to act as a *memory* model that reacts to the history of the input signal as well as the current instantaneous value of the input signal.

The multivariate nonlinear functions f_1 and f_2 can be implemented with multivariate polynomial functions of the form:

$$b_1 = \sum_{y_i, z_i} K_{(y_1,\cdots,y_{11},z_1,\cdots,z_{11})} a_1(t)^{y_1} \cdots a_1(t-10t_d)^{y_{11}} a_2(t)^{z_1} \cdots a_2(t-10t_d)^{z_{11}} \tag{3.19}$$

$$b_2 = \sum_{y_i, z_i} L_{(y_1,\cdots,y_{11},z_1,\cdots,z_{11})} a_1(t)^{y_1} \cdots a_1(t-10t_d)^{y_{11}} a_2(t)^{z_1} \cdots a_2(t-10t_d)^{z_{11}} \tag{3.20}$$

The polynomial coefficients $K_{(y_1,\cdots,y_{11},z_1,\cdots,z_{11})}$ and $L_{(y_1,\cdots,y_{11},z_1,\cdots,z_{11})}$ will have unique real values for each polynomial power of the delayed inputs of input signals $a_1(t-kt_d)$ and $a_2(t-kt_d)$ for values of k that go from 0 to 10. The main complication of using a multivariate polynomial implementation is that the number of required model coefficients increases as the model order is increased and the choice of which coefficients to include and which to leave out becomes important in model extraction stability. In addition, while polynomial models can allow for good interpolation, they are not well suited for extrapolation beyond the training data. Since the nonlinear representation of polynomials is limited, an artificial neural network model of f_1 and f_2 is used instead to avoid these limitations. The decision to use artificial neural networks, especially one that has output neurons with bounded activation functions (like the Sigmoid function), allows for better simulation robustness for the model while tackling the *hard nonlinearity* that requires many polynomial kernel functions to approximate.

The neural network topology used for modeling each of the two multi-variable nonlinear functions f_1 or f_2 will have a distinct input neuron for each of the input signals $a_1(t-kt_d)$ and $a_2(t-kt_d)$ and a single output neuron for $b_1(t)$ or $b_2(t)$ respectively similar to what is shown in Fig. 4.7. All the inputs to the artificial neural network are normalized to a value between 0 and 1 (using Min-Max Normalization). The layers of neurons in between the input and output layers are referred to as the hidden layers. Each neuron in the hidden layer will get an input from all the neurons in the previous layer and will model its output based on the following neuron model:

$$y = S\left(b + \sum_{i=1}^{k} w_i x_i\right) \tag{3.21}$$

where x_i are the k inputs to the neuron and $S(x) = \frac{1}{1+e^{-x}}$ is the sigmoid function. The choice of the sigmoid function as the activation function of the neuron model ensures that the output of the artificial neuron y will be a value between -1 and 1. Each artificial neuron

will have input weights w_i and bias value b as parameters that will need to be solved for during the ANN training. The output neuron of each of the $b_1(t)$ and $b_2(t)$ ANNs will denormalize the value of the output neuron from a value between -1 and 1 to the minimum and maximum values that are available in the training dataset for the output variable.

A single hidden layer with enough neurons should be enough to model any continuous nonlinear multivariate expression. The parameters of the ANN, the weights and bias values are solved for given a chosen artificial neural network topology using the Levenberg-Marquardt back-propagation algorithm available in the Neural Network Training Toolbox of MATLAB [50][51]. In practise we noticed that to achieve the -30dB threshold NMSE, a single hidden layer topology is capable of modeling the gate (input impedance) nonlinear model for b_1 but for the drain (output power) nonlinear model a two hidden layer neural network topology was required to achieve the target NMSE with respect to the load-pull data with a less number of neurons. It was observed that up to two hidden layers for an artificial neural network can be simulated with modern harmonic balance simulators without much trouble. Using a bounded nonlinear activation function for the output neuron like the sigmoid function provides better simulation convergence compared to using an unbounded linear activation function, even if the ANN trained with the linear output neuron activation function achieved the required threshold NMSE. For the characterization data that was extracted from the 10W GaN transistor in the previous Section, the gate model ANN had a single-hidden-layer ANN with 100 neurons while the drain model ANN had a two-hidden-layer ANN with 50 neurons in each hidden layer to achieve the threshold -30dB of NMSE.

It's important to note that even though an ANN has quite a complex visual representation, the ANN itself can be flattened into a single mathematical expression that can be evaluated by the circuit simulator. The flattened expression for the ANN can be found starting from the right side of the ANN as shown in Figure 4.7 and expanding out the inputs from each neuron and replacing it with the neuron expression. At the end one is left with a differentiable expression that can be implemented in a harmonic balance circuit simulator.

3.3.3 Implementation of a TD-PHD Model in a Harmonic Balance Simulator

To implement the TD-PHD model that was extracted with the procedure outlined in the previous section in a harmonic balance simulator, one can take advantage of the fact that modern circuit simulators allow for time-domain nonlinear expressions to be defined in a

circuit Netlist. In Keysight's Advanced Design System (ADS) environment, a Symbolically Defined Device (SDD) component can be used to implement this static time-domain nonlinear expression. Given the choice of an artificial neural network for the gate and drain TD-PHD models, each of the trained ANNs can be converted into a flattened expression. This flattened expression is generated starting at the output neuron and recursively calling each of the neurons in the previous layer for an expression of their outputs. This recursive call generates a single expression involving additions, subtractions, multiplications and exponentiations, all valid operations in defining a nonlinear expression within the harmonic balance simulator. Since the ANN description is analytic with defined derivatives, the harmonic balance simulator can thus load the nonlinear TD-PHD model and its derivatives into memory to perform the nonlinear simulations of interest.

The nonlinear describing functions of the TD-PHD model as shown in Figure 3.12 requires access to time-delayed versions of the input signals $a_1(t)$ and $a_2(t)$. These time delayed inputs can be made available to this nonlinear expression through the use of surrogate nodes in the Netlist that produce the delayed versions of the input signals. These surrogate nodes can be generated through a chain of frequency-defined time-delays of each of the input signals. These frequency-domain time-delay blocks of fractional-period delay length of T_0/N add a frequency-proportional phase shift by multiplying the each of the fundamental and harmonic frequency phasors A_{1k} (or A_{2k}) by $e^{2\pi jk/N}$ where k is the harmonic order of the phasor. Figure 3.13 shows the delayed versions of the input signal that is made available in the Netlist as surrogate signals for the computation of the TD-PHD model for a simple case where $N = 7$ as a visual demonstration. Stepping through time it is apparent that the values of the signals at other times within the fundamental period of the multi-harmonic signal are exposed within the circuit simulator and input to the nonlinear describing functions of the TD-PHD model in a time-invariant manner. The generation of these time-delayed signals within the simulator allows the simulator to access the time-domain information that exists in the entirety of the fundamental period at any time within the fundamental period of the signal by only accessing these surrogate signals at the current instant. Then all of these surrogate nodes are fed into the static memory-less TD-PHD expression to give the output signal $b_1(t)$ for the case of the gate model or $b_2(t)$ for the case of the drain model.

The extracted model can then be implemented into this harmonic balance Netlist structure and the resulting behavioural model can be simulated to predict its performance under various conditions. In Section 3.3.4 the extracted model for the 10W GaN device that was characterized in Section 3.3.1, extracted in Section 3.3.2, and implemented in a Harmonic Balance simulator in Section 3.3.3 will be validated against multi-harmonic load-pull measurements.

3.3.4 Comparison of the Simulation of Extracted TD-PHD Models with Multi-Harmonic Loadpull Measurements

The extracted and implemented 2GHz-fundamental frequency TD-PHD model of the $10W$ GaN packaged RF power transistor (Wolfspeed CGH40010) biased at a quiescent DC drain current of 45mA is then put in a simulated multi-harmonic load-pull harmonic balance simulation testbench to simulate its power sweeps for any given harmonic impedance termination set at either port of the device under test. In our measurement validation, a 2-harmonic active load-pull measurement was used to extract the TD-PHD behavioural model, but passive load-pull measurements were used to validate the performance of the extracted behavioural model. This was done so that the validation was not done based on the same training set used to extract the model. The duality between the simulated load-pull testbench and the load-pull characterization of the power transistor in the hybrid load-pull power transistor characterization testbench was used as a basis for the validation of the extracted and implemented model.

To showcase the multi-harmonic performance of the extracted TD-PHD model, the passive tuner harmonic impedances at source and the load are set randomly, while the presented a fundamental frequency load reflection coefficient is varied as shown in Figures 3.14 through 3.18. The multi-harmonic time-domain NMSE for the prediction of each of the $b_1(t)$ and $b_2(t)$ is presented in each of these figures, showing that the single extracted large signal TD-PHD model was able to predict the $b_1(t)$ and $b_2(t)$ time-domain waveforms with a multi-harmonic time-domain NMSE of better than -25dB for the range of fundamental mismatch tested.

Finally, a gradient-search algorithm was run on the multi-harmonic passive tuners in order to tune the source harmonic and load harmonic impedances to maximize the power-added efficiency. Figure 3.19 shows the comparison of the TD-PHD model gain and the Power-Added Efficiency (PAE) performance in simulation compared against the measured performance at the optimal-search harmonic match, showing that the multi-harmonic model was able to predict the peak performance of the single ended power amplifier with an NMSE of the output waveform $b_2(t)$ better than -32dB. This demonstrates that a TD-PHD model can be an effective tool for the design of high efficiency single ended power amplifiers from multi-harmonic load-pull characterization of a power transistor device.

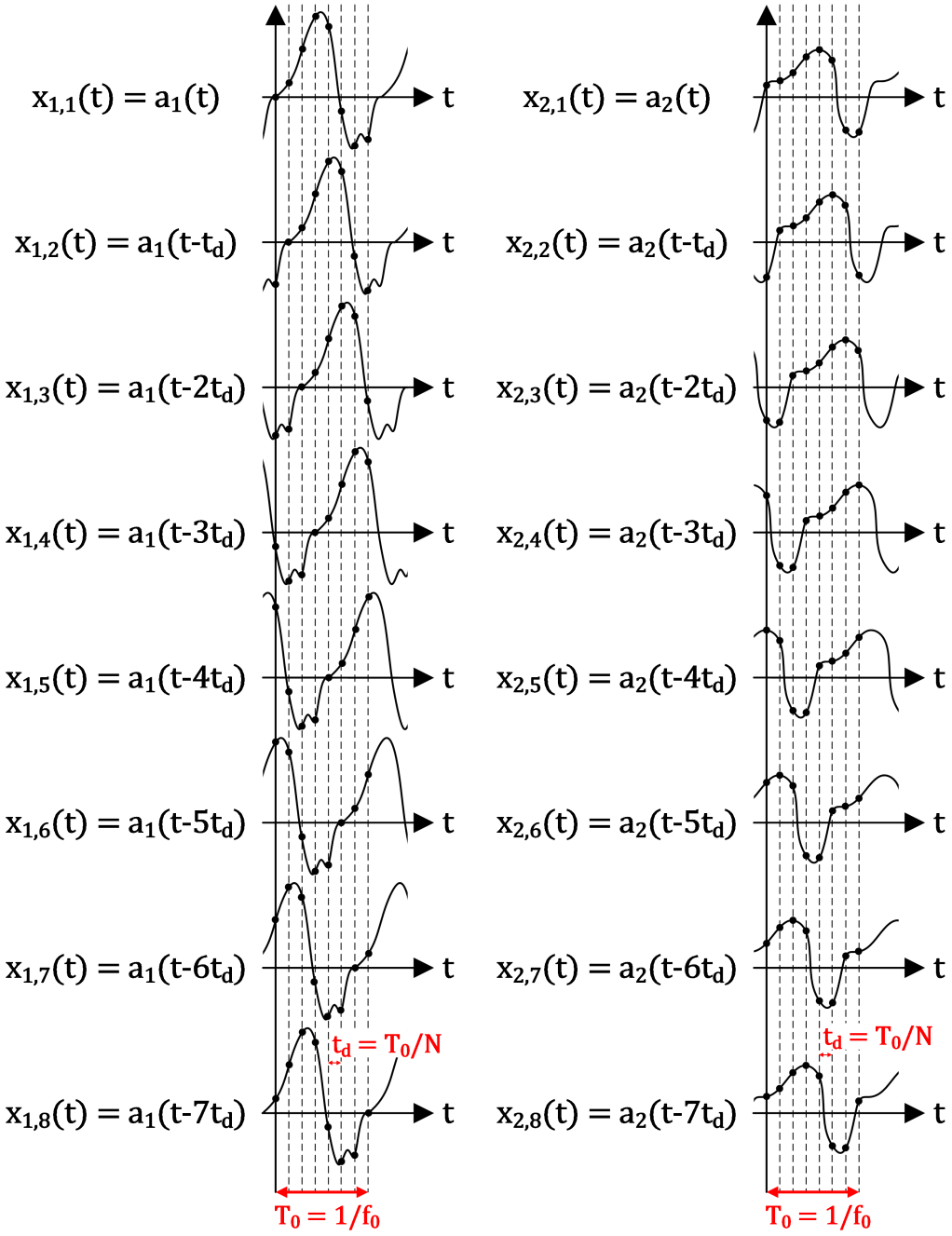

Figure 3.9: A graphical example of an N=8 input signal discretization to develop time-domain PHD modeling data

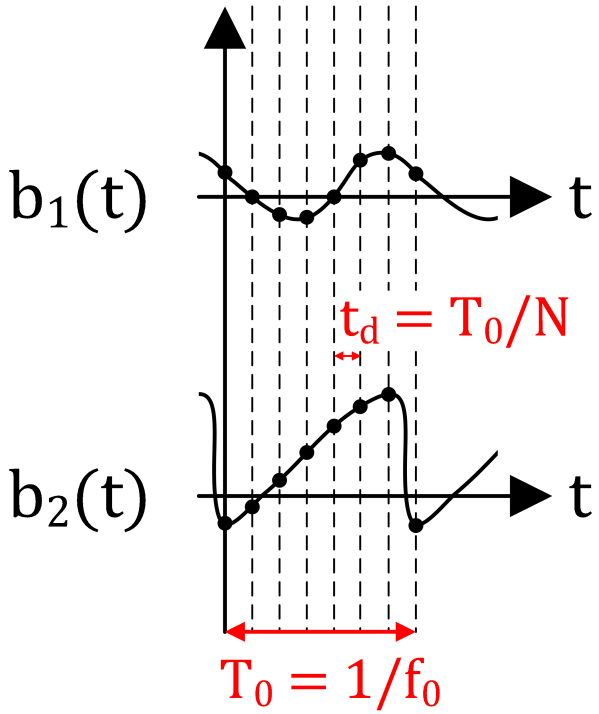

Figure 3.10: A graphical example of an N=8 output signal discretization to develop time-domain PHD modeling data

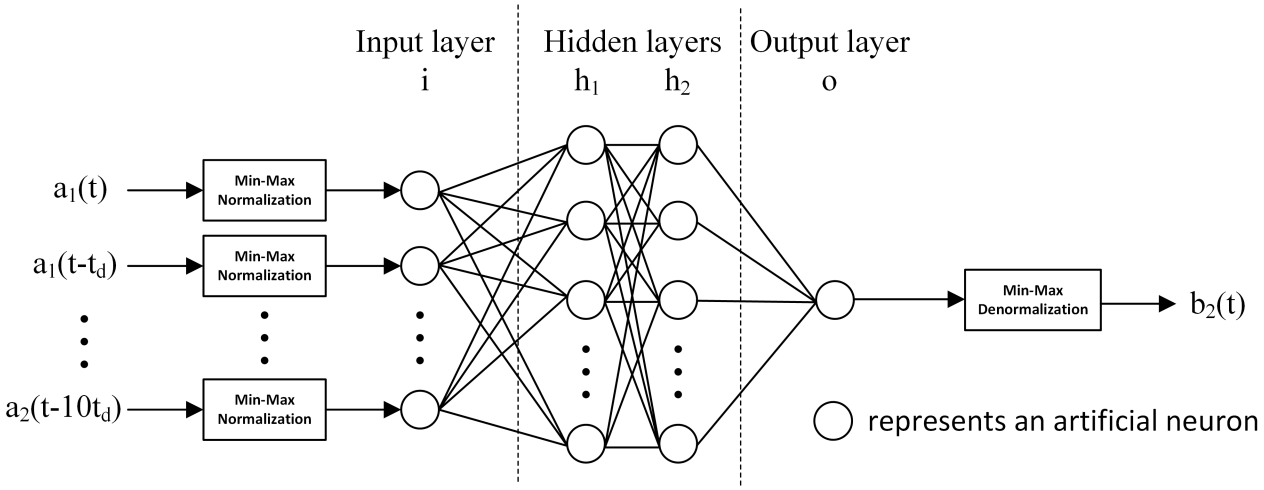

Figure 3.11: Diagram of artificial neural network architecture used for the $b_2(t)$ output of the power transistor model based on the 22 auxiliary signals $a_1(t)$ through $a_1(t-10t_d)$ and $a_2(t)$ through $a_2(t-10t_d)$

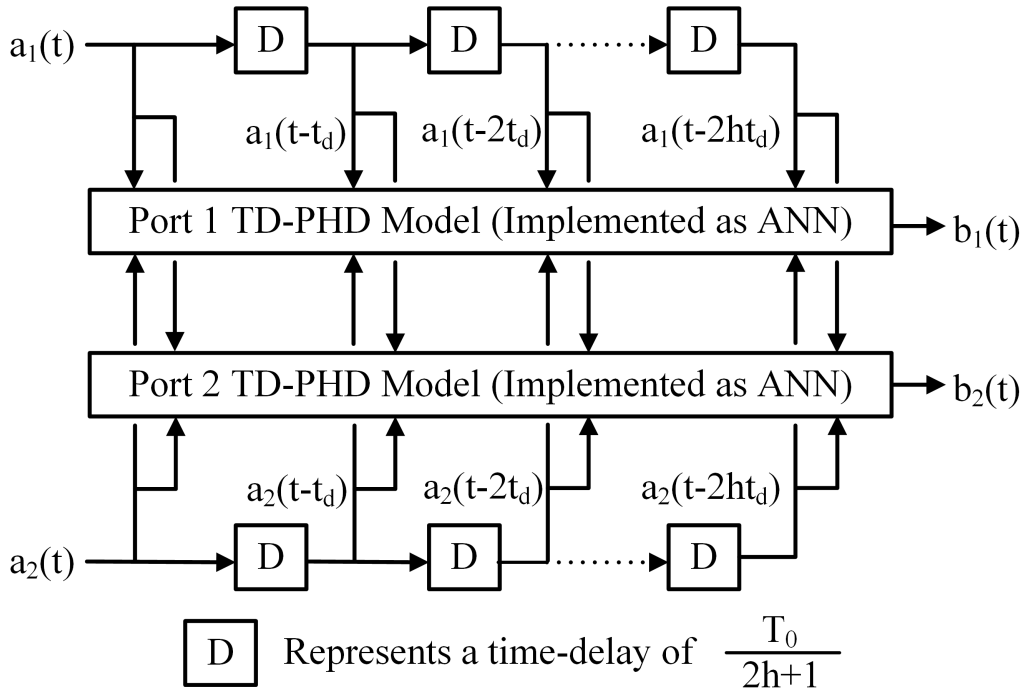

Figure 3.12: Harmonic balance implementation circuit of a 2-port TD-PHD model to use as an RF transistor behavioral model

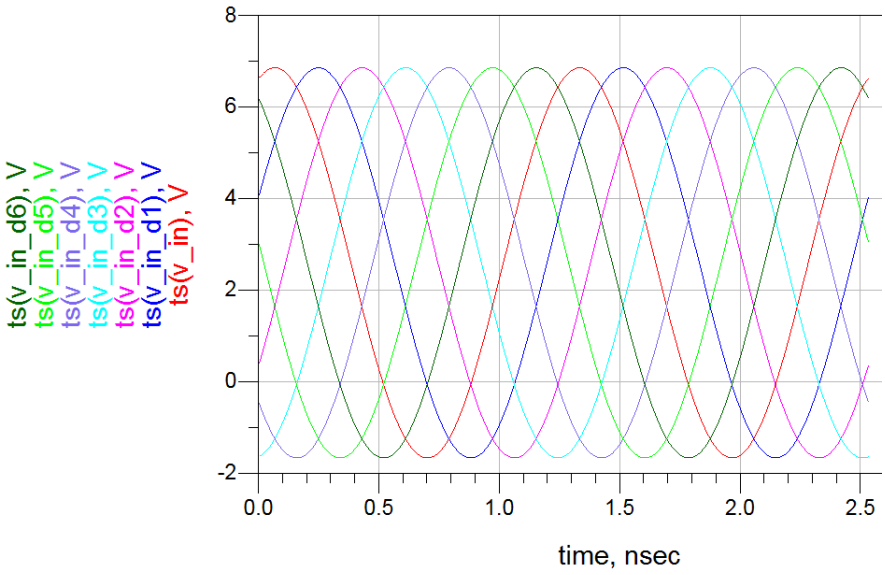

Figure 3.13: A multi-harmonic waveform and its delayed version over 7 steps within its fundamental period

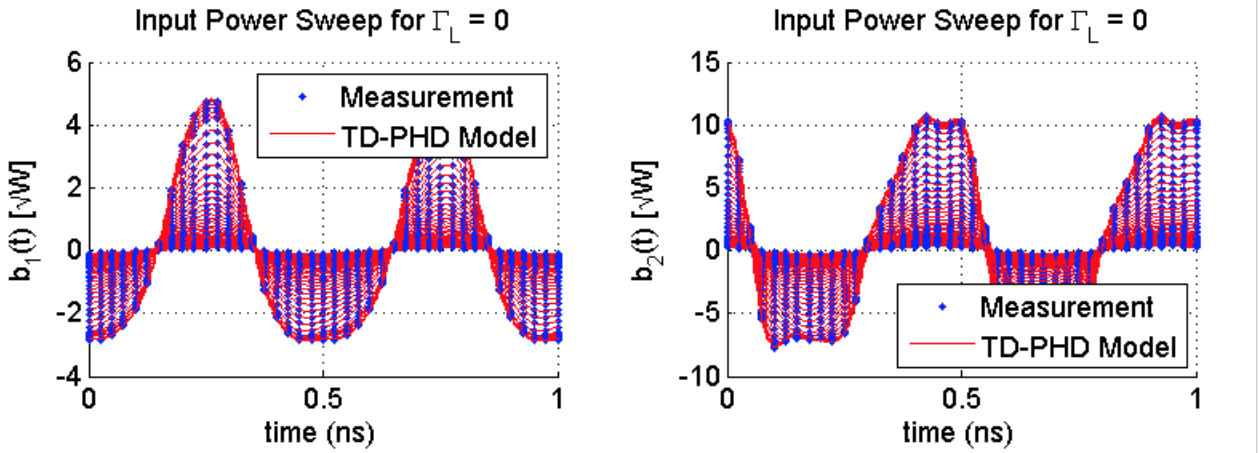

Figure 3.14: Measured vs. TD-PHD model power sweep time-domain waveforms for $\Gamma_L = 0$ with a $\text{NMSE}_{b_1(t)} = -42.334dB$ and $\text{NMSE}_{b_2(t)} = -38.637dB$

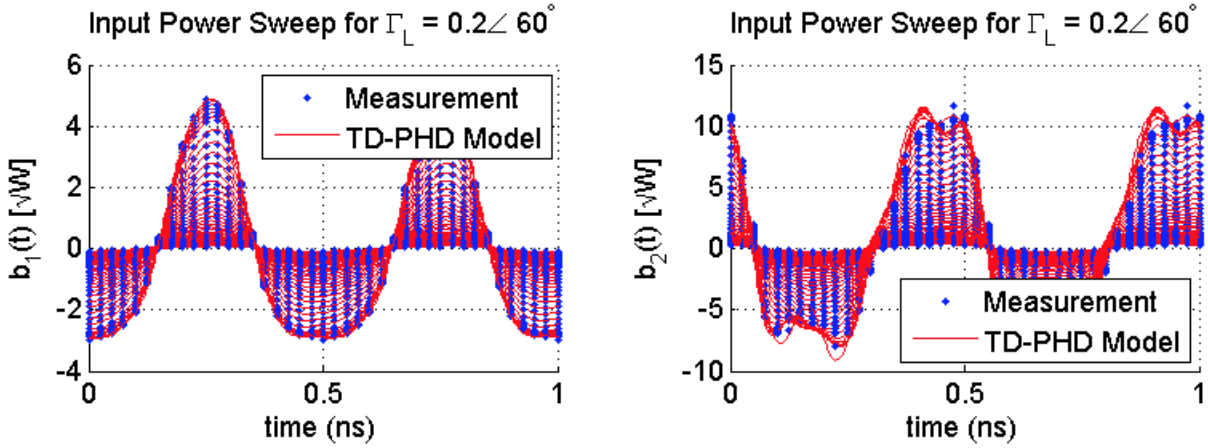

Figure 3.15: Measured vs. TD-PHD model power sweep time-domain waveforms for $\Gamma_L = 0.2\angle 60°$ with a $\text{NMSE}_{b_1(t)} = -41.286dB$ and $\text{NMSE}_{b_2(t)} = -30.985dB$

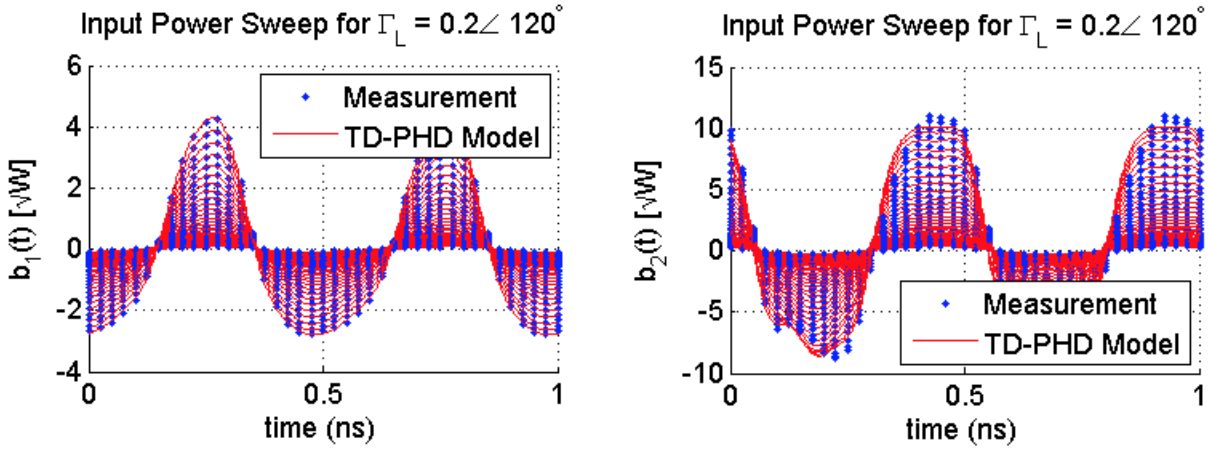

Figure 3.16: Measured vs. TD-PHD model power sweep time-domain waveforms for $\Gamma_L = 0.2\angle 120°$ with a $\text{NMSE}_{b_1(t)} = -41.123dB$ and $\text{NMSE}_{b_2(t)} = -35.954$

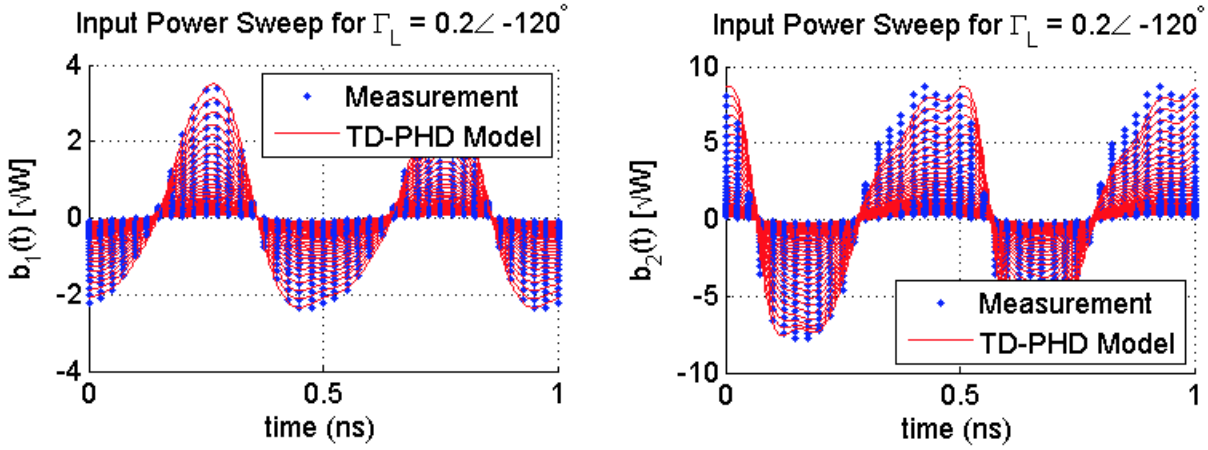

Figure 3.17: Measured vs. TD-PHD model power sweep time-domain waveforms for $\Gamma_L = 0.2\angle -120°$ with a $\text{NMSE}_{b_1(t)} = -40.844dB$ and $\text{NMSE}_{b_2(t)} = -25.380dB$

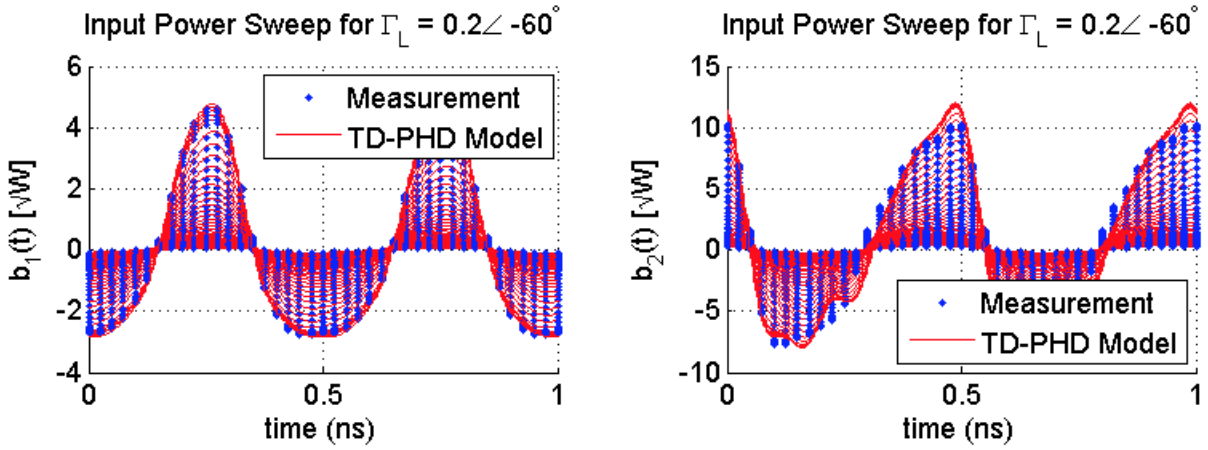

Figure 3.18: Measured vs. TD-PHD model power sweep time-domain waveforms for $\Gamma_L = 0.2\angle -60°$ with a $\text{NMSE}_{b_1(t)} = -39.989dB$ and $\text{NMSE}_{b_2(t)} = -26.017dB$

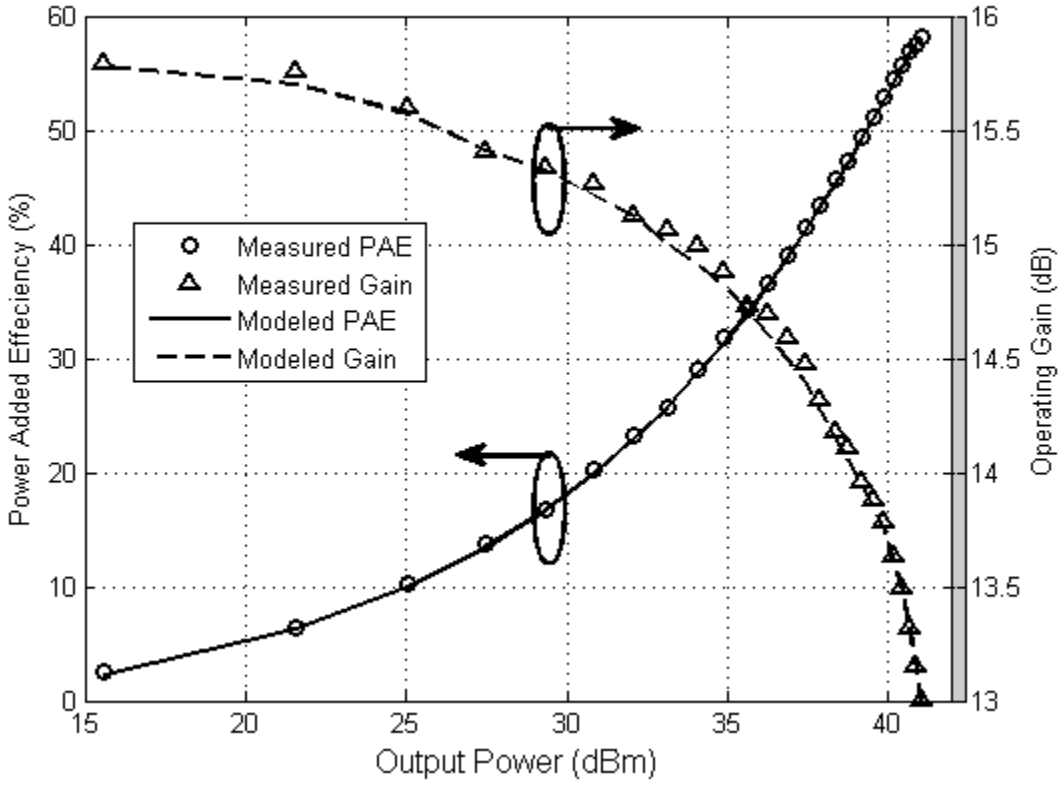

Figure 3.19: Measured Gain and Power Added Efficiency Performance of the 10W GaN RF transistor compared to the TD PHD model prediction at the Maximum Efficiency Optimum Impedance multi-harmonic match with an $\text{NMSE}_{b_2(t)} = -32.029dB$

3.4 Chapter Conclusions and Discussion

In this chapter the time-domain poly-harmonic distortion model, which is one of the main contributions of this thesis, was introduced as an alternative behavioural modeling scheme to the *frequency-domain* PHD models in the literature that is instead exclusively defined in the *time-domain*. First an intuitive argument was presented in Section 3.1 that showed from first principles that the TD-PHD model has a reduced level of construction complexity needed to describe the mapping between the shape of a multi-harmonic input signal onto the shape of a multi-harmonic output signal in a time-invariant way.

Then in Section 3.2 a more rigorous development of the TD-PHD model is presented that starts with a time-domain impulse response model derived from the frequency domain multi-harmonic S-parameters of the underlying system, and generalizes the results to nonlinear time-invariant systems by showing how the TD-PHD model is in effect a time-domain nonlinear impulse response model.

After this initial theoretical justification for the TD-PHD model, the procedure to develop TD-PHD models from large-signal measurements of the power transistor is presented in Section 3.3, where the entire process from the multi-harmonic load-pull characterization of the power transistor to the preparation of the characterization data for model extraction, the details involved in artificial neural network modeling, the implementation of the TD-PHD model in a harmonic balance simulator, and a measurement validation of the extracted model is presented, justifying the usefulness of the TD-PHD model in designing power amplifier circuits that require taking the harmonic impedance terminations into account. The extraction of the resulting time-domain poly-harmonic distortion (TD-PHD) model required a small number of representative multi-harmonic load-pull measurements to capture the nonlinear behaviour. It was also successfully integrated into ECAD tools and can be exploited in the design single-ended harmonically tuned power amplifiers.

Nevertheless, the TD-PHD model as described in this chapter is fixed to a single fundamental frequency. This means that if a TD-PHD model Netlist construction as it was proposed in this chapter is simulated at a different fundamental frequency than where it was intended, the model would have no predictive capability. Another fundamental limitation of the Netlist construction of the TD-PHD model proposed in this chapter is that its strictly a harmonic balance simulation model. One way to implement a TD-PHD model in a time-domain simulator would be an implementation that fixes the time-step of the time-domain simulation to match the sub-period time-delay (t_d) of the TD-PHD model.

The TD-PHD model, being fixed to a single fundamental frequency, requires a wideband high power amplifier designer to extract a completely separate model to fit the load-

pull measurements at each fundamental frequency, even though a common underlying nonlinearity has resulted in the behaviour observed at all the individual fundamental frequencies. In Chapter 4, the proposed TD-PHD model of this chapter will be generalized such that a single time-domain model can fit load-pull measurements that were performed over a non-uniformly spaced discrete frequency grid. The proposed model is a discrete-time model that is tuned to act on a discrete set of frequencies. As a result the model, like the TD-PHD model presented in this chapter, it will not be able to predict the performance of the power transistor at frequencies that are not in the discrete frequency set of the model. On the other hand the discrete frequencies can be chosen such that it spans over a significant bandwidth of frequencies where load-pull measurements have been performed.

Chapter 4

Time-Domain Multi-Tone Distortion Models

A limitation of the TD-PHD model proposed in the previous chapter is that the model is fixed to one fundamental frequency. This means that for load-pull measurements spanning multiple fundamental frequencies, a separate TD-PHD model would be required at each fundamental frequency, even though a common underlying nonlinearity has resulted in the large signal behavior at each of these fundamental frequencies. In this chapter, the previously proposed model will be extended such that a single time-domain model can fit load-pull measurements with a non-uniformly spaced frequency grid that contains all the load-pull design frequencies of interest. This model will be called the time-domain multi-tone distortion (TD-MTD) model. Figure 4.1 summarizes the steps required to develop a TD-MTD behavioural model of a power transistor device.

To validate this extended modeling framework, the large-signal simulation of a two-way Doherty power amplifier with packaged LDMOS power transistors will be investigated and compared against the measurements of the power amplifier and a simulation of the amplifier using a compact transistor model. Unlike other attempts of Doherty power amplifier design via behavioral models, load-pull measurements and compact models[52][53][54][55], this thesis proposes a general framework that produces a single time-domain model that can be used to approximate all the measured load-pull data with a time-invariant behavioral model.

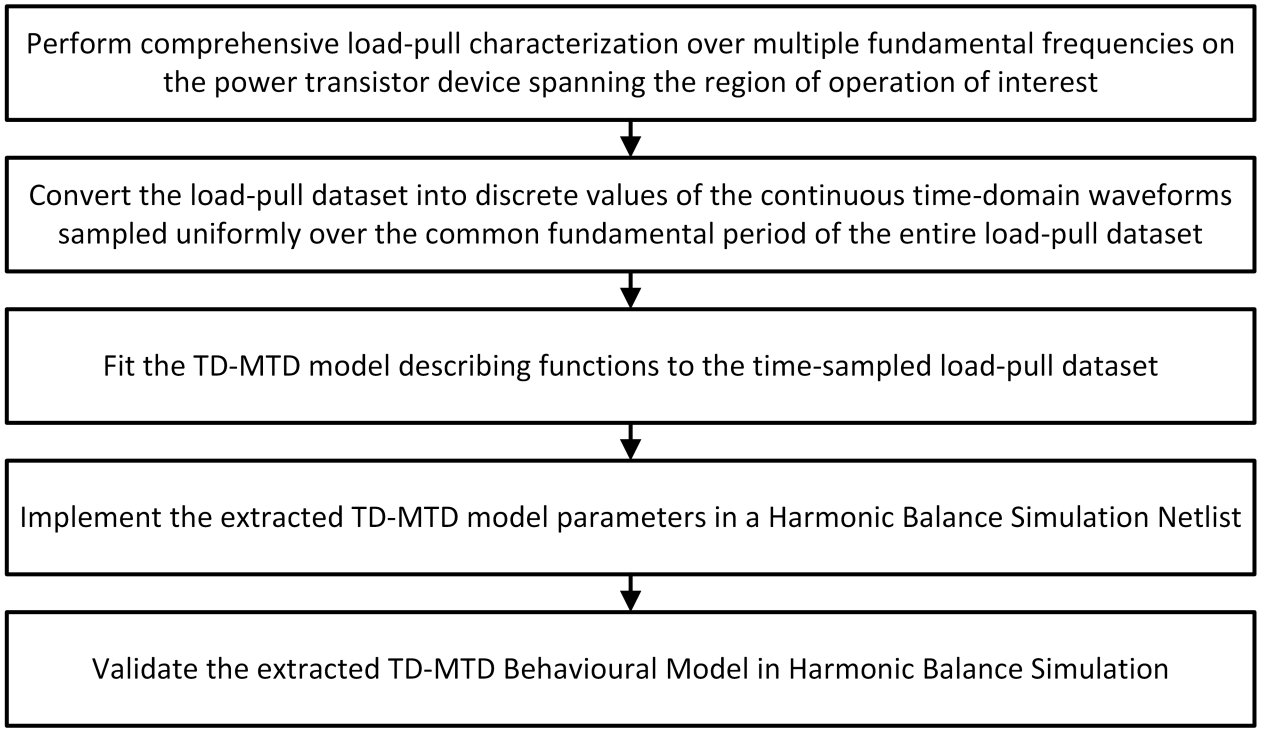

Figure 4.1: A flowchart of the steps required to develop a TD-MTD model

4.1 The Volterra Series Projection Formulation for TD-MTD Models

In this Section the Volterra Series representation of a power transistor will be taken as a theoretical representation of the entirety of the underlying nonlinear dynamics of the power transistor, then a *discretization* or *projection* of this infinite series expression for a subset of stimulus signals that are fixed on a specific non-uniformly spaced frequency grid, using a specific discrete set of auxiliary signals will be shown to be possible. Finally an extension of the TD-PHD model that will be referred to as the time-domain multi-tone distortion (TD-MTD) model will be derived by showing it is such a projection of the underlying Volterra series representation of the power transistor given a specific defined time delay between the auxiliary signals that corresponds to the shared fundamental frequency of the load-pull dataset.

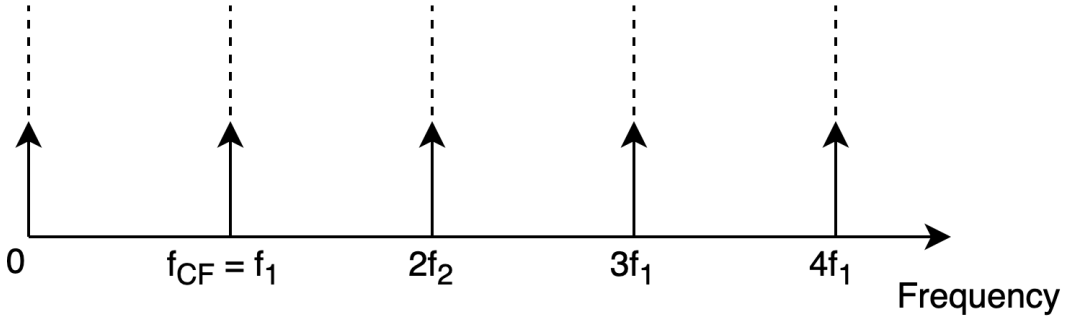

Figure 4.2: Frequency grid (dashed lines) and location of frequency content (arrows) for the TD-PHD model

4.1.1 Discrete Projection of a Continuous-Time Volterra Series on a Fixed Frequency Grid

If the power transistor is assumed to be an equi-continuous and uniformly bounded nonlinear time-invariant system, then according to the Arzela-Ascoli theorem and Fréchet's approximation theorem [56], its behavior around a quiescent bias can be approximated uniformly to an arbitrary degree of precision by a sufficiently high order Volterra series [57][58][59]. For simplicity of representation, a one-port system will be used in the following analysis but the arguments will be trivially generalized to two-port systems and beyond at the end of this Section.

Suppose a powerwave scattering model is used for behavioral modeling, where the input signal is the incident powerwave signal $a(t)$ and the output signal is the reflected powerwave signal $b(t)$. The time-domain continuous Volterra series gives the response of the nonlinear time-invariant system $b(t)$ for a known input signal $a(t)$. The Volterra series represents the output signal as a series sum of multi-dimensional convolutions of the input signal with the kernel functions h_n, where n is the polynomial order of the kernel function [18]:

$$b(t) = \sum_{n=0}^{\infty} \underbrace{\int_0^{\infty} \cdots \int_0^{\infty}}_{n-times} h_n(\tau_1, \cdots, \tau_n) \prod_{i=1}^{n} a(t-\tau_i) d\tau_i \tag{4.1}$$

The real-valued kernel functions $h_n(\tau_1, \cdots, \tau_n)$ are a representation of the underlying dynamic nonlinearity of the power transistor including its short and long term memory

effects. Assume that a fixed set of time-domain continuous kernel functions provide a frequency-independent description of the nonlinear time-invariant response of the power transistor for all potential continuous input signals $a(t)$. In order to create a behavioral model that targets signals that are fixed to a certain discrete frequency grid, a projection of the output of the frequency-independent infinite series expression of (4.1) can be found for a subset of all possible signals that lie on a specific fixed frequency grid. In order to theoretically compute the instantaneous output, $b(t)$, of the Volterra-series expression of (4.1), the circuit simulator requires the complete knowledge and access to the input signal's value at all instances in time and not just at time t, even though the circuit simulator expects an instantaneous (that is, available at time t) time-domain expression for the nonlinearity. To be able to emulate the instantaneous-time computation of the Volterra series, the information about the past values of the input signal needs to be made available to the circuit simulator at all instantaneous times. This can be achieved through auxiliary signals in the model Netlist that reveal the time-domain value of the history of the input signal to the circuit simulator, allowing an emulated computation of the instantaneous output only based on the instantaneous values of these auxiliary signals.

The TD-PHD model[49] is such a projection of the continuous-time Volterra series kernels onto a discrete frequency grid of frequencies that are multiples of a single fundamental frequency f_1. Although f_1 is not the only frequency that can divide all the frequencies of the frequency grid, it is the largest frequency that can do so. Any integer division of f_1 (e.g. $f_1/2$ or $f_1/3$ and so on) could have possibly also been used as the common fundamental frequency (f_{CF}) of the model. For the TD-PHD model as proposed in [49], the f_{CF} parameter was explicitly fixed to f_1, the fundamental frequency of the multi-harmonic load-pull measurement set, as shown in Fig. 4.2. This f_{CF} parameter will be the key to generalization of the TD-PHD modeling framework proposed in this chapter. The TD-PHD model uses a set of auxiliary signals that are delayed versions of the input signals evenly spaced in time to span the fundamental period ($T = 1/f_1$) of the multi-harmonic load-pull measurement dataset. Let the auxiliary signals $x_i(t)$ be N time delayed versions of $a(t)$ spanning its period (T):

$$
\begin{aligned}
x_1(t) &= a(t) \\
x_2(t) &= a\left(t - \frac{T}{N}\right) = a\left(t + (N-1)\frac{T}{N}\right) \\
x_3(t) &= a\left(t - 2\frac{T}{N}\right) = a\left(t + (N-2)\frac{T}{N}\right) \\
&\cdots \\
x_{N-1}(t) &= a\left(t - (N-2)\frac{T}{N}\right) = a\left(t + 2\frac{T}{N}\right) \\
x_N(t) &= a\left(t - (N-1)\frac{T}{N}\right) = a\left(t + \frac{T}{N}\right)
\end{aligned}
\tag{4.2}
$$

By making the signals $x_i(t)$ available in the circuit simulator, the value of $a(t)$ at any time can be theoretically evaluated and made available for the computation of the right-hand-side of (4.1). So a discrete projection of the Volterra-series can be found with these auxiliary signals as its constituting basis. Since the choice of auxiliary signals in the TD-PHD model is fixed to the period of a single fundamental frequency, a single TD-PHD model cannot be used to represent the behavior of a nonlinear power transistor over multiple non-uniformly spaced frequencies that was captured during a load-pull measurement. To overcome this limitation, a generalization of the time-spacing of the auxiliary signals and as a result, a new generalized discrete projection that allows for the extraction of a single time-domain defined behavioral model for a non-uniformly spaced frequency grid will be proposed. The models using this generalized framework will be referred to as time-domain multi-tone distortion (TD-MTD) models. TD-MTD models are a generalization of TD-PHD models in the sense that a TD-PHD model is a TD-MTD model where f_{CF} is fixed to a single fundamental frequency.

In order to allow for an instantaneous-time computation of the Volterra-series at time t_1, there must exist a time-invariant interpolation function that reveals the value of the input signal $a(t)$ at another time t_2 from only the evaluation of the auxiliary signals, $x_i(t)$, at time t_1 and the knowledge of the time-offset between these two times $\Delta t = t_2 - t_1$. That is, the condition required to be able to approximate the Volterra series instantaneously and to provide a projection is the existence of a smooth continuous interpolation function f_{interp}:

$$
a(t_2) \approx f_{\text{interp}}(\Delta t, x_1(t_1), x_2(t_1), \cdots, x_k(t_1)) \tag{4.3}
$$

If the function f_{interp} exists, then the continuous-time Volterra series expression collapses onto the discrete projection:

$$b(t) = g(x_1(t), x_2(t), \cdots, x_k(t)) \tag{4.4}$$

where the continuous simulation-time-independent static function g approximates the operations on the right hand side of (4.1). In fact, the multi-variable nonlinear function g in (4.4) is approximating the output of the expression of the Volterra series of (4.1) where the time-dependent operand $a(t)$ is replaced with a set of simulation-time-independent operands $x_i(t)$ by taking advantage of the interpolation function. This can be achieved without having an explicit expression for f_{interp} by adopting a multi-variable polynomial expression for g. Alternatively, in this work an artificial neural network was used to fit this multi-variate function g as it is an effective fitting tool for continuous multi-variate functions and it allows the avoidance of the numerical instability problems that arise when a high nonlinear polynomial model order is required.

4.1.2 Formulation of the TD-MTD Model

In the previous Section, a multi-variable non-linear function g was introduced to approximate the Volterra-series when modeling the behavior of high power transistors over multiple nonuniformly-spaced frequencies. In this Section, the auxiliary signals $x_i(t)$ that were used as operands in the function g will now have to be defined such that they allow for the extraction of the parameters of the multi-variable function g based on the load-pull characterization data spanning multiple fundamental frequencies.

Since typically these frequencies are all integer multiples of a common frequency f_{CF}, a common period exists for the time-domain representations of the load-pull measurement data. In Fig. 4.3 the location of the spectral content of load-pull measurements are shown for multiple fundamental frequencies. This common period is longer than the period of any individual fundamental frequency in the load-pull characterization data and contains more periods of the higher frequency characterization data and less periods of the lower frequency data. A uniform sampling of the input signals over this much longer common period can reveal the signals that are on a non-uniformly spaced frequency grid through the Non-Uniform Discrete Fourier Transform Type I (NUDFT-I) [60] [61].

Suppose the auxiliary signals $x_i(t)$ of (4.2) are defined for $T = 1/f_{CF}$ and are evaluated at time $t = 0$:

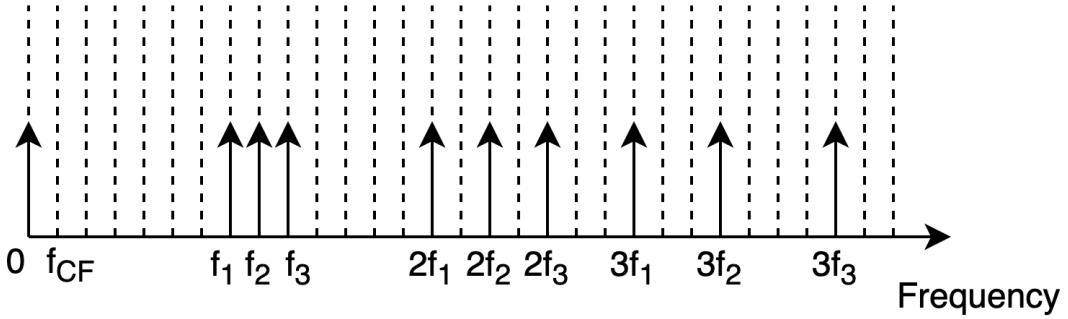

Figure 4.3: Frequency grid (dashed lines) and the location of frequency content of multi-harmonic load-pull at three different fundamental frequencies (arrows) for a TD-MTD model

$$\begin{aligned} x_1(0) &= a(0) \\ x_N(0) &= a\left(\frac{T}{N}\right) \\ x_{N-1}(0) &= a\left(2\frac{T}{N}\right) \\ &\cdots \\ x_3(0) &= a\left((N-2)\frac{T}{N}\right) \\ x_2(0) &= a\left((N-1)\frac{T}{N}\right) \end{aligned} \tag{4.5}$$

The set of discrete samples $x_i(0)$ can be used to generate a Fourier series representation of the input signal $a(t)$. These Fourier series coefficients are obtained from the Non-Uniform Discrete Fourier Transform Type I (NUDFT-I) as follows:

$$A_k(0) = \sum_{n=0}^{N-1} a(nT/N) e^{-j\frac{2\pi}{N}\frac{f_k}{f_{\mathrm{CF}}}n} \tag{4.6}$$

The Fourier series coefficients $A_k(0)$ are obtained when the functions $x_i(t)$ are evaluated at $t = 0$. Using the time-shifting property, the Fourier Series coefficients at an arbitrary simulation-time can be obtained by:

$$A_k(t) = e^{j\omega_k t} A_k(0) \tag{4.7}$$

where ω_k is the angular frequency of each of the Fourier series coefficients A_k.

Comparing the auxiliary signals of (4.5) and the elements of the Fourier series expression of (4.6), it can be noted that the Fourier series coefficient $A_k(0)$ is a function of the delayed input signals $x_i(0)$:

$$A_k(0) = f_k\left(x_1(0), x_2(0), \cdots, x_N(0)\right) \tag{4.8}$$

By the time-shift property for periodic signals, the Fourier series coefficient at simulation-time t_1 can be determined as:

$$A_k(t_1) = f_k\left(x_1(t_1), x_2(t_1), \cdots, x_N(t_1)\right) \tag{4.9}$$

The time-domain input signal at another arbitrary simulation-time t_2 can be expressed as the following Fourier series expression:

$$a(t_2) = \sum_{k=0}^{h} \frac{1}{2}\left(A_k(0)e^{jk\omega_0 t_2} + A_k^*(0)e^{-jk\omega_0 t_2}\right) \tag{4.10}$$

Using the time-shift property of (4.7), the expression of (4.10) can be re-written using the time-varying Fourier series coefficients $A_k(t_1)$:

$$a(t_2) = \sum_{k=0}^{h} \frac{1}{2}\left(A_k(t_1)e^{jk\omega_k(t_2-t_1)} + A_k^*(t_1)e^{-jk\omega_0(t_2-t_1)}\right) \tag{4.11}$$

In this expression an interpolation of the input signal at an arbitrary simulation-time t_2 is found that's based on the auxiliary signals $x_i(t)$ evaluated at another arbitrary simulation-time t_1 and the time offset between these two arbitrary simulation-times $\Delta t = t_2 - t_1$. Thus (4.3) holds for the choice of auxiliary signals $x_i(t)$ and from (4.4), the output at an arbitrary time t can be expressed using a simulation-time-independent function of auxiliary signals evaluated at any simulation-time t:

$$b(t) = g\left(x_1(t), x_2(t), \cdots, x_N(t)\right) \tag{4.12}$$

Now that the expression of the auxiliary signals needed to model the behavior of a non-linear single port system over a non-uniformly-spaced set of frequencies has determined, the expression of (4.12) can be generalized for two port systems, which is the form of the model used for power transistors:

$$b_1(t) = g_1(x_{1,1}(t), x_{1,2}(t), x_{1,3}(t), \cdots, x_{1,N}(t), x_{2,1}(t), x_{2,2}(t), x_{2,3}(t), \cdots, x_{2,N}(t)) \tag{4.13}$$

$$b_2(t) = g_2(x_{1,1}(t), x_{1,2}(t), x_{1,3}(t), \cdots, x_{1,N}(t), x_{2,1}(t), x_{2,2}(t), x_{2,3}(t), \cdots, x_{2,N}(t)) \tag{4.14}$$

where N is the time resolution of the model and $x_{1,k}(t)$ and $x_{2,k}(t)$ are the auxiliary signals based on the two input signals $a_1(t)$ and $a_2(t)$. In this generalization of one-port systems to two port systems, the complete dependence of each of the output signals ($b_1(t)$ or $b_1(t)$) on both of the input signals ($a_1(t)$ and $a_2(t)$) is made explicit.

4.2 Extraction and Validation of the Proposed TD-MTD Model and its Implementation in a Harmonic Balance Simulator

In Section 4.2.1 the procedure to extract a TD-PHD model from load-pull measurements of a high power transistor is outlined and its implementation in a harmonic balance simulator is described. Any nonlinear load-pull-based behavioral model should at least reproduce the load-pull data that was used to extract it. In Section 4.2.1, the extracted model is put in a simulated load-pull testbench and the load-pull contours obtained from the simulated measurements will be compared to the raw measurements used to extract the behavioral model.

To truly demonstrate the modeling capabilities of using behavioral models of power transistors in the practical design of power amplifiers, in Section 4.2.2 a two-way Doherty power amplifier is simulated based on two behavioral models extracted from the main and peaking power transistors respectively.

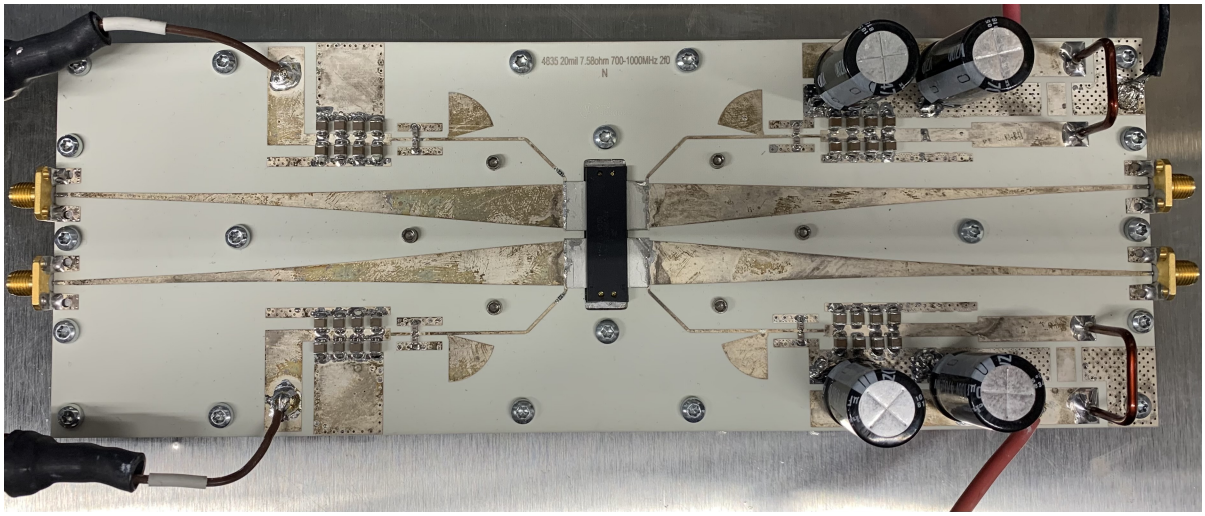

Figure 4.4: The dual-device load-pull fixture designed for the NXP A2V09H525 package

It should be noted that the signals used during the load-pull measurement and the validation of the two-way Doherty PA are all narrow-band pulsed-RF signals, even though the Doherty PA is designed to amplify wider bandwidth modulated signals with a high Peak to Average Power Ratio (PAPR). Even though narrow-band characterization of the transistor doesn't completely capture all the dynamics of the power transistor, a necessary but not sufficient requirement of wide-band Doherty PA design is that it at least meets the narrow-band RF performance requirements across the design band of interest. This means that an extracted narrow-band model that is correct across the band can be used to tune the performance of the wide-band PA across the band. On the other hand the extracted TD-MTD model will only be able to model the large-signal narrow-band performance across the band and will not be able to be used to simulate the modulated signal behavior parameters like Adjacent Channel Power Ratio (ACPR).

4.2.1 Extraction of the Proposed TD-MTD Model from Load-Pull Measurements and its Implementation in a Harmonic Balance Simulator

To showcase the ability of the proposed TD-MTD model in fitting load-pull measurement data spanning over multiple fundamental frequencies, a set of fundamental frequency load-pull measurements at three frequencies (790MHz, 805MHz and 820MHz), that have a

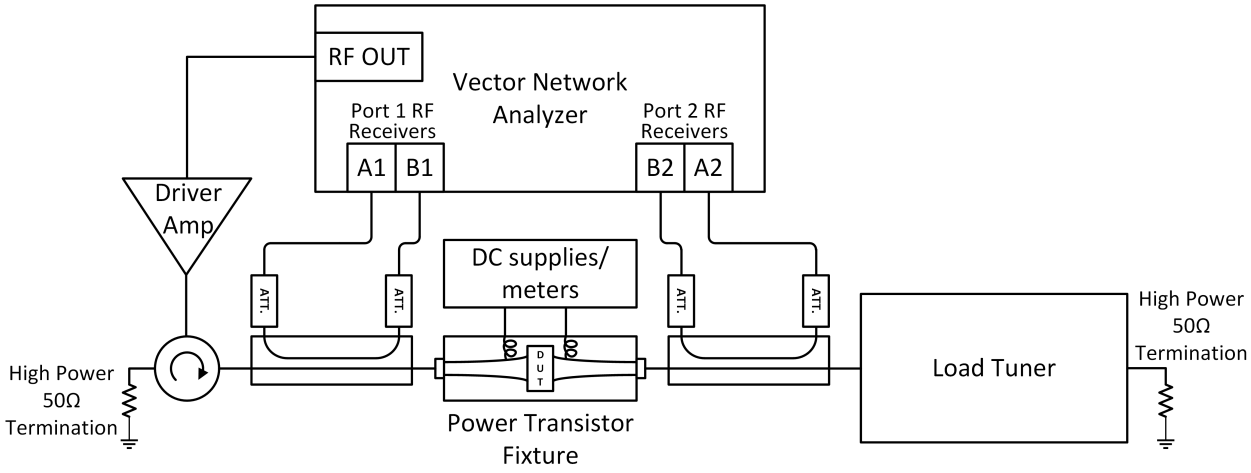

Figure 4.5: Block diagram of the load-pull measurement setup used to perform load-pull measurements on the power transistor in an RF fixture

common fundamental frequency f_{CF} of 5MHz, are performed on both the main and peaking power transistors of the NXP A2V09H525 packaged high power (which has a peak power of 525W) LDMOS device that is intended for an asymmetrical two-way Doherty HPA design. A dual-device load-pull fixture was designed for the NXP A2V09H525 device as shown in Fig. 4.4. The fixture parameters were extracted from a custom built Thru-Reflect-Line calibration kit and used to de-embed vector-corrected passive load-pull measurements to the package plane of the transistor devices in a setup similar to the block diagram of Fig. 4.5. A set of load-pull measurements were obtained that include DC Drain current measurements and pulsed RF waveform measurements at the fundamental frequency at the input and output of the power transistors. Pulsed 10% duty cycle RF measurements are performed on high power transistors during load-pull since a non-pulsed RF signal at the peak power of power transistor would excessively heat up the device at the peak powers of the power transistor. The load-pull measurement sweep involved setting a passive load-tuner to different fundamental load impedances at each of the frequencies and performing a pulsed-RF power sweep at each of the tuner positions. These power sweeps were bound at the upper end by a maximum gain compression of 5.5dB for the main device and 3.5dB for the peaking device. Since the compression of the power transistor is highly dependent on the load impedance, the input power ranges in the measurement data will vary with impedance and frequency. The load-pull measurements were performed over a range of impedances that covered the high power and high efficiency operations of the transistor.

The drain model of (4.14) will model the nonlinear output power generation of the

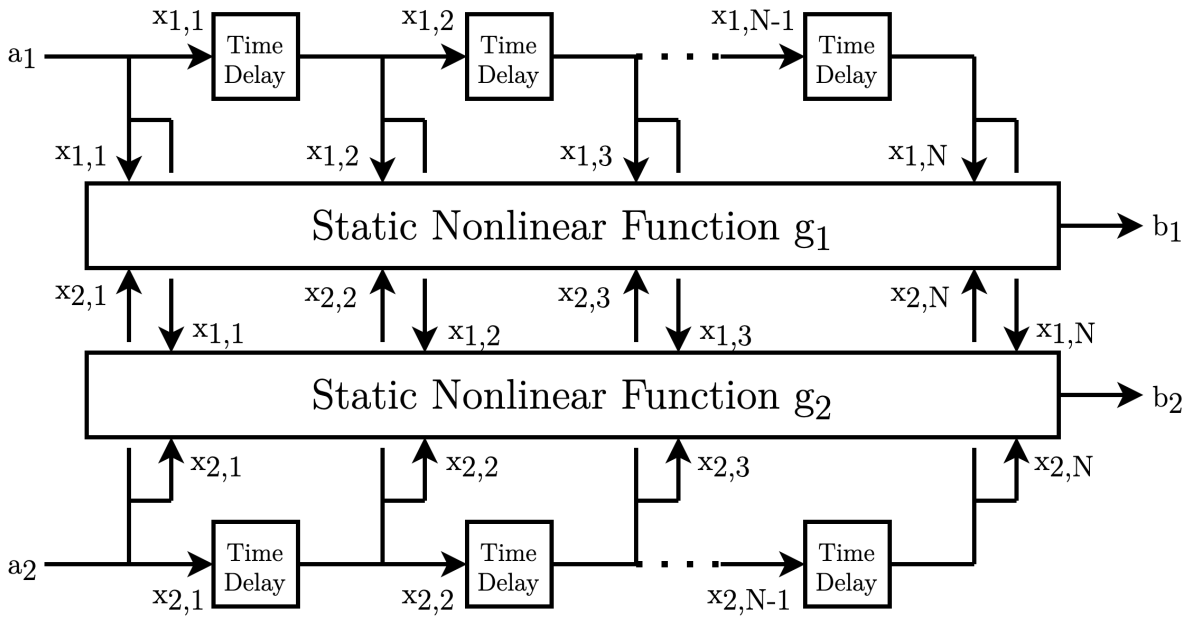

Figure 4.6: Harmonic balance implementation circuit of a 2-port TD-PHD model to use as an RF transistor behavioral model

power transistor while the gate model of (4.13) will model its nonlinear input impedance. This behavioral model is implemented as a Netlist for a harmonic balance simulator in the form shown in Fig. 4.6. To generate the auxiliary signals $x_{1,i}$ and $x_{2,i}$ in a harmonic balance simulation Netlist, the time-delays can be implemented as a frequency-defined block in the Netlist that applies a frequency proportional phase-shift to each frequency component of the input signals. A train of fractional period time-delay blocks each creating a delay of $t_d = T_{CF}/N$ in front of the a_1 and a_2 signals will reveal all the auxiliary signals $x_{1,i}$ and $x_{2,i}$ to the time-domain simulation-time-independent nonlinear functions g_1 and g_2. This allows the harmonic balance simulator to compute the time-domain output signals $b_1(t)$ and $b_2(t)$.

To extract the proposed model, the load-pull measurement dataset will need to be converted from the frequency domain to the time domain by means of a Fourier Series evaluation of the DC and fundamental frequencies:

$$
\begin{aligned}
a_1(t) &= V_{\text{gate,DC}} + |A_{11}| cos\left(2\pi f t + \angle A_{11}\right) \\
a_2(t) &= V_{\text{drain,DC}} + |A_{21}| cos\left(2\pi f t + \angle A_{21}\right) \\
b_1(t) &= I_{\text{gate,DC}} + |B_{11}| cos\left(2\pi f t + \angle B_{11}\right) \\
b_2(t) &= I_{\text{drain,DC}} + |B_{21}| cos\left(2\pi f t + \angle B_{21}\right)
\end{aligned}
\tag{4.15}
$$

If the model time resolution N is too low, the model fitting algorithm will have difficulty in finding a good fit to the data. A good model fitting threshold would be a Normalized Mean Squared Error (NMSE) of better than -30dB for the time-domain measurement dataset.

For this load-pull measurement dataset, when the model time resolution N was set to 17, it was found to have better than -30dB NMSE for each of the time-domain parameters $b_1(t)$ and $b_2(t)$ to the measurement data with the chosen fitting function for the TD-MTD nonlinear functions g_1 and g_2. Since the f_{CF} of this measurement set is 5MHz, the common fundamental period will be $T = 1/f_{\text{CF}} = 0.2\mu s$. This makes the sampling time delay $t_d = T/N = (0.2/17)\mu s$. The discrete set of functions $x_{1,1}(t)$ through $x_{1,17}(t)$ and $x_{2,1}(t)$ through $x_{2,17}(t)$ will be used to denote the time-domain delayed incident wave at the input port $a_1(t)$ and the output port $a_2(t)$ respectively and are defined by the definition outlined in (4.2).

Since the entire load-pull measurement data set is periodic with the period T_{CF}, the functions $b_1(t)$, $b_2(t)$, $x_{1,1}(t)$ through $x_{1,17}(t)$, and $x_{2,1}(t)$ through $x_{2,17}(t)$ are evaluated at times $t = 0$, $t = t_d$, $t = 2t_d$, $\cdots$, $t = 16t_d$. This means that each frequency-domain load-pull measurement at a fixed power level in the dataset will be converted into 17 equivalent discrete time-sampled data points. This will be the discrete time-domain dataset used for fitting the simulation-time-independent nonlinear output functions g_1 and g_2 that implement the following TD-MTD mappings:

$$
b_1(t) = g_1\left(x_{1,1}(t), x_{1,2}(t), \cdots, x_{1,17}(t)\right) \tag{4.16}
$$

$$
b_2(t) = g_2\left(x_{2,1}(t), x_{2,2}(t), \cdots, x_{2,17}(t)\right) \tag{4.17}
$$

The neural network topology used for modeling each of the two multi-variable nonlinear functions g_1 or g_2 will have a distinct input neuron for each of the auxiliary signals $x_{1,i}$ and $x_{2,i}$ and a single output neuron for b_1 or b_2 respectively similar to what is shown in Fig. 4.7.

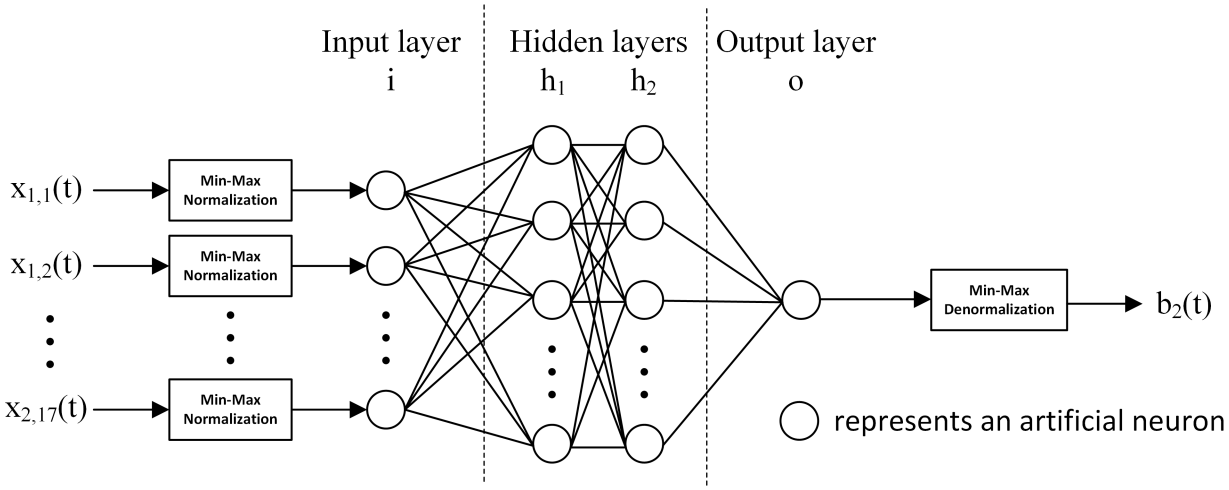

Figure 4.7: Diagram of artificial neural network architecture used for the $b_2(t)$ output of the power transistor model based on the 14 auxiliary signals $x_{1,1}$ through $x_{2,17}$

The gate model mapping of (4.16) was implemented with a two-hidden-layer artificial neural network with 30 neurons in each layer achieving a $NMSE_{b_1(t)} = -37.6281$dB, while the drain model mapping of (4.17) was implemented with a two-hidden-layer artificial neural network with 50 neurons in each layer achieving a $NMSE_{b_2(t)} = -31.1715$dB for the main device of the NXP A2V09H525. The ANNs were extracted using the Levenberg–Marquardt algorithm available from MATLAB. The extracted artificial neural networks implementing (4.16) and (4.17) were converted into a flattened expression. This flattened expression can be recognized by the circuit simulator when implemented as a simulation-time-independent multi-variable time-domain nonlinearity. In the Keysight ADS harmonic balance simulation environment, this can be implemented using the Symbolically Defined Device (SDD) component which will implement the multi-variable nonlinear functions g_1 and g_2. The inputs to these multi-variable functions $x_{1,2}(t)$ through $x_{1,17}(t)$ and $x_{2,1}(t)$ through $x_{2,17}(t)$ are made available in the model Netlist using a time-delay chain of 16 time fractional period delays t_d of the input $a_1(t)$ and $a_2(t)$ signals respectively. The time-delays are implemented in the Netlist as a frequency domain equation block implementing the frequency proportional phase shift of (4.7).

A simulated load-pull measurement was performed on the implementation of the extracted TD-MTD model to test its ability to reconstruct the measurement dataset used to extract the model. Since a mature compact model of this power transistor is also available, the load-pull simulation of this compact model at the same DC bias condition and fre-

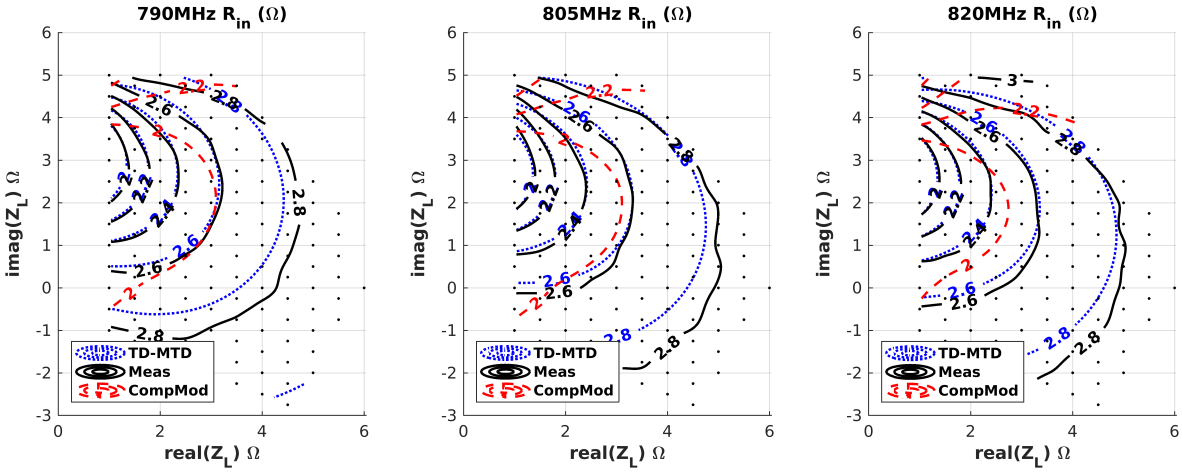

Figure 4.8: The real part of the main power transistor input impedance load-pull contours of the TD-MTD model and Compact Model at 2dB of gain compression across the three discrete fundamental frequencies compared to the load-pull measurements

Table 4.1: Table of simulated NMSE of RF and DC parameters of the Compact model and TD-MTD model of the NXP A2V09H525 main device over the entire load-pull measurement dataset

	Compact Model	**TD-MTD Model**
$NMSE_{B_1}$	-15.2213 dB	-47.6008 dB
$NMSE_{B_2}$	-20.6091 dB	-41.5374 dB
$NMSE_{I_{\text{Drain,DC}}}$	-24.9673 dB	-44.3599 dB

quencies were obtained and are included in this comparison. Table 4.1 shows the summary of how well the compact model and extracted TD-MTD model fit the load-pull data over the 3 fundamental frequencies in the RF reflected wave and DC drain current parameters. Unsurprisingly the TD-MTD NMSE is spectacular here as it was extracted from the same load-pull data. Fig. 4.8 and Fig.4.9 show the load-pull contours of the real and imaginary part of the input impedance looking into the transistor gate ($Z_{\text{in}} = R_{\text{in}} + jX_{\text{in}}$) at 2dB of gain compression. Since behavioral models fit the load-pull measurement data directly, it is not surprising that the TD-MTD model has a better prediction of the input impedance of the power transistor compared to the compact model. Fig. 4.10, Fig. 4.11 and Fig. 4.13 show the comparison of the load-pull simulation of the compact model, the extracted TD-MTD model and the load-pull measurements in terms of the output power, operating gain and drain efficiency at 2dB gain compression respectively. The compact model has its

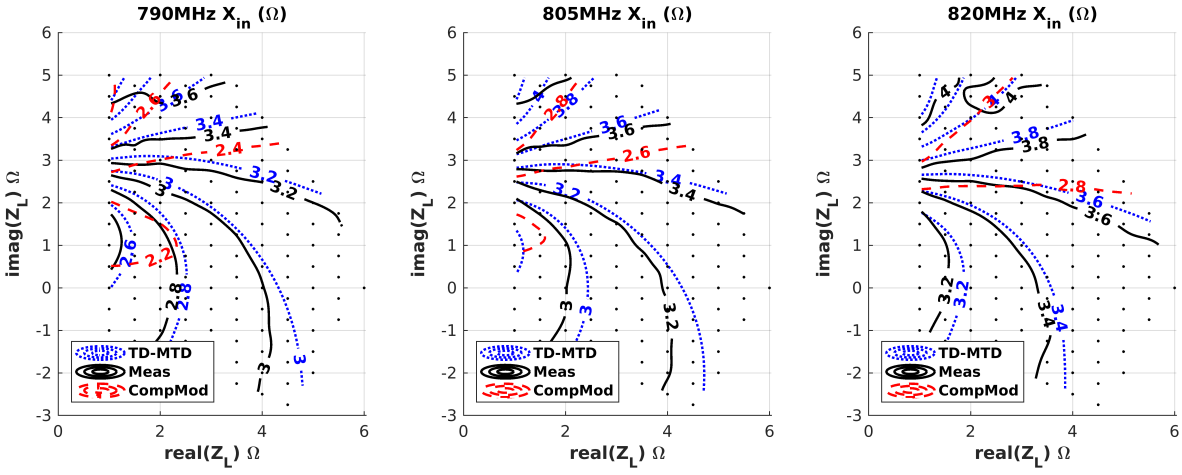

Figure 4.9: The imaginary part of the main power transistor input impedance load-pull contours of the TD-MTD model and Compact Model at 2dB of gain compression across the three discrete fundamental frequencies compared to the load-pull measurements

best accuracy towards the high power and high efficiency regions of the load-pull data but the accuracy is less at impedances further away. Fig. 4.12 compares the AMPM load-pull contours to the simulation of the two power transistor models at 2dB of gain compression for the main device. Since the TD-MTD model is a behavioral model that does not use a look-up table, the resulting simulated performance smooths out the noise in the measurement data resulting in smooth AMPM contours that track the trend observed in the noisy load-pull measurement data. Overall the TD-MTD model can faithfully model the load-pull measurement data spanning multiple frequencies with a single smooth time-domain fitting function.

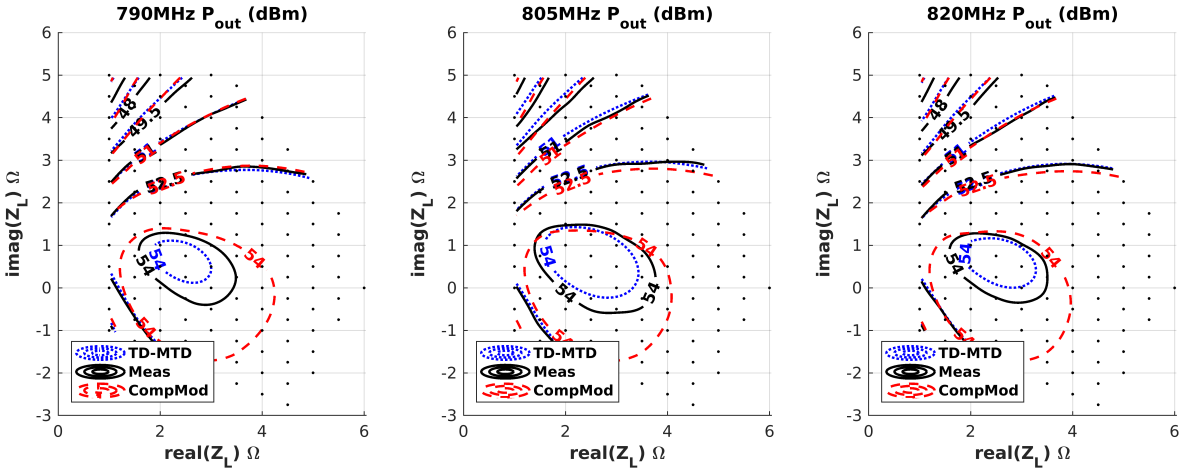

Figure 4.10: The main power transistor output power load-pull contours of the TD-MTD model and Compact Model at 2dB of gain compression across the three discrete fundamental frequencies compared to the load-pull measurements

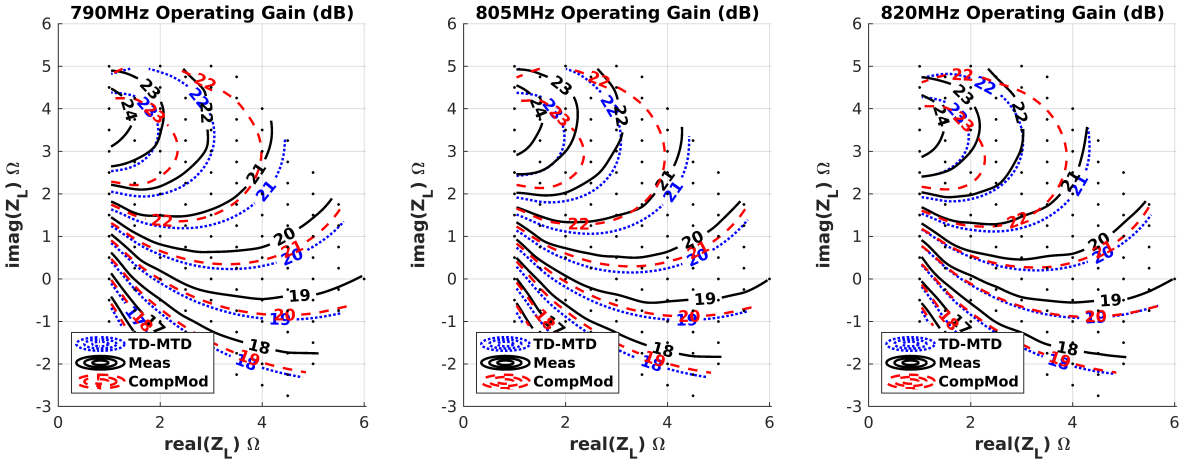

Figure 4.11: The main power transistor operating gain load-pull contours of the TD-MTD model and Compact Model at 2dB of gain compression across the three discrete fundamental frequencies compared to the load-pull measurements

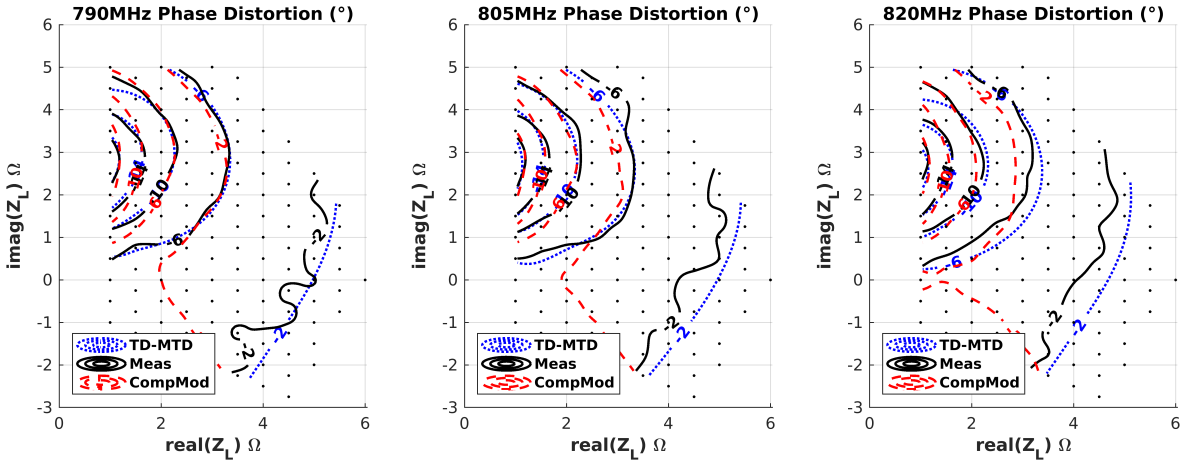

Figure 4.12: The main power transistor AMPM distortion load-pull contours of the TD-MTD model and Compact Model at 2dB of gain compression across the three discrete fundamental frequencies compared to the load-pull measurements

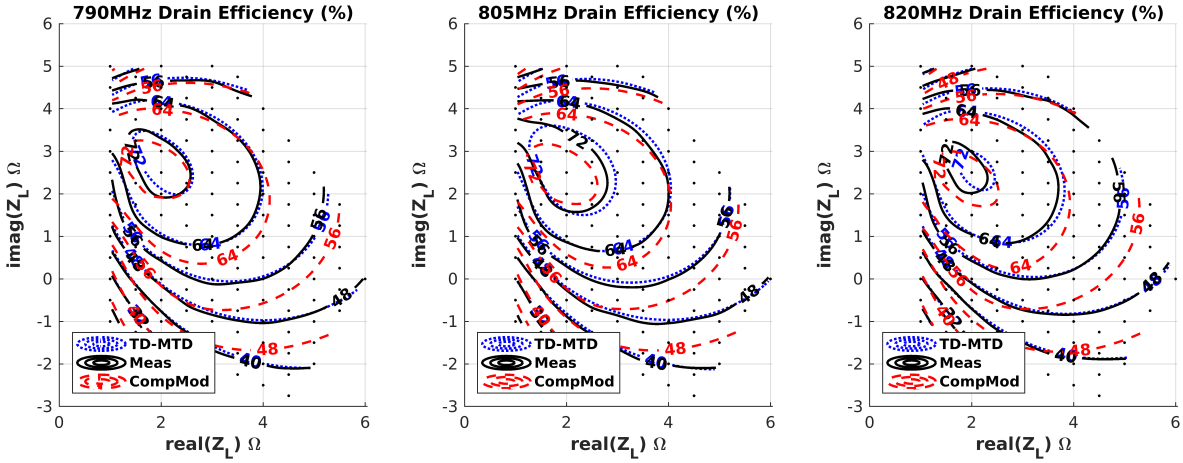

Figure 4.13: The main power transistor drain efficiency load-pull contours of the TD-MTD model and Compact Model at 2dB of gain compression across the three discrete fundamental frequencies compared to the load-pull measurements

4.2.2 Doherty High Power Amplifier Simulation of Multi-Tone Distortion Models

In this Section, the models extracted in Section 4.2.1 will be used to simulate the narrow-band large signal operation of a Doherty HPA design over a set of discrete fundamental frequencies. Since the load-pull measurement data used for model extraction only includes DC and fundamental frequency measurements, these behavioral models do not react to harmonic impedance termination in the circuit simulator. Under the assumption that packaged LDMOS power transistor devices are not highly sensitive to harmonic impedance terminations, an HPA design can be simulated solely based on collected DC and fundamental frequency load-pull behavior. It should be noted that the theoretical derivation of this model does not forbid the inclusion of harmonic data for the training of the model. For the case of GaN power transistors when the harmonic termination becomes significant, it would be suggested to perform harmonic load-pull measurements in conjunction with fundamental frequency load-pull at each of the fundamental frequencies. This load-pull measurement space will contain a higher number of measurements but a TD-MTD model could be fit to such a load-pull measurement set with the exact same procedure outlined in this chapter. The validation of a TD-MTD for a multi-harmonic multi-fundamental-frequency load-pull measurement space is not demonstrated here but is within the theoretical possibilities of the application of TD-MTD models.

The 790MHz NXP A2V09H525-04NR6 Test Circuit with a PCB layout shown in Figure 4.14 and a Bill of Material of components shown in Figure 4.15 was used as the reference circuit to validate the main and peaking device models extracted from the two power transistors in the NXP A2V09H525 package. Since the top-copper structure PCB drawings and the bill-of-materials of this reference circuit are available, the S-parameters of the input and output matching networks of the reference PA circuit were extracted using an EM simulation and vendor S-parameter models were used as a model for the passive components. Harmonic balance simulation at 790MHz, 805MHz, and 820MHz were then

Table 4.2: Table of simulated NMSE of RF and DC parameters of the Compact model and TD-MTD model for a power-sweep of the reference NXP A2V09H525 two-way Doherty design

	Compact Model	**TD-MTD Model**
$NMSE_{\lvert B_1 \rvert}$	-16.8550 dB	-25.2416 dB
$NMSE_{B_2}$	-17.8901 dB	-18.8917 dB
$NMSE_{I_{\mathrm{Drain,DC}}}$	-15.5059 dB	-13.1730 dB

Table 4.3: Gain Magnitude and Phase Compression comparison at average and peak power

	Gain Magnitude (dB)						**Phase Compression (°)**					
Output Power Level (dBm)	**49**			**57**			**49**			**57**		
Frequency (MHz)	**790**	**805**	**820**	**790**	**805**	**820**	**790**	**805**	**820**	**790**	**805**	**820**
Compact Model	18.9	19.1	18.8	17.3	17.9	18.2	-3.8	-4.9	-5	-3.7	-1.5	0.6
TD-MTD Model	18.9	19.3	18.3	17.3	18.1	18.2	-10.1	-8.5	-6.9	-6.8	-3.8	-2.3
Measurement	19	18.5	18	16.4	16.7	17.1	-9	-9.6	-8.7	-12.2	-6.9	-3.3

Table 4.4: Input Return Loss and Drain Efficiency comparison at average and peak power

	Input Return Loss (dB)						**Drain Efficiency (%)**					
Output Power Level (dBm)	**49**			**57**			**49**			**57**		
Frequency (MHz)	**790**	**805**	**820**	**790**	**805**	**820**	**790**	**805**	**820**	**790**	**805**	**820**
Compact Model	13.6	14.1	14.9	16.8	18.7	21.9	59	59	55	65	65	66
TD-MTD Model	16.1	16.7	18.2	17.4	18.7	22	59	61	57	62	64	65
Measurement	15	15.8	18.1	18.1	19.8	22.4	65	62	58	66	68	70

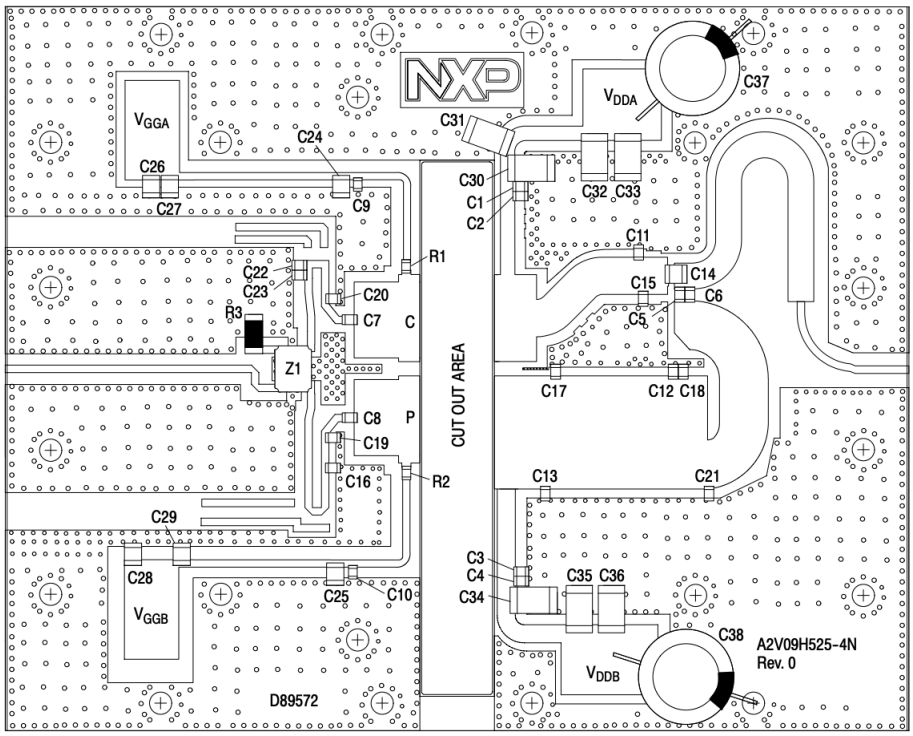

Figure 4.14: Reference Doherty Printed Circuit Board Layout and Component locations

performed on the schematic representation of the reference PA circuit, which includes the extracted main and peaking power transistor TD-MTD models, the extracted EM structures and the discrete passive models, respectively. The results of the simulation are compared to large signal pulsed-RF measurements of the reference PA as well as a simulation of the circuit with the compact model of the devices. Table 4.2 summarizes the NMSE of the compact model and the extracted TD-MTD compared to the measurement data in terms of how well they reflect each of the fundamental frequency RF waves and the DC drain current consumption over the power and frequency sweep. This table overall shows that the TD-MTD model performed significantly better than the compact model in predicting the input side RF behavior, while at the output side the compact model performed slightly better in predicting the DC Drain current but slightly worse than the TD-MTD model at predicting the RF behavior. Fig. 4.16 and Fig. 4.17 show the gain magnitude and phase compression curves of the measured PA compared to the simulated PA with the TD-MTD model and the compact model, while Table 4.3 shows the numerical values of the gain magnitude and phase compression at the average power (49dBm) and

Part	Description	Part Number	Manufacturer
C1, C2, C3, C4, C5, C6, C7, C8, C9, C10	56 pF Chip Capacitor	ATC600F560JT250XT	ATC
C11, C12, C13	6.8 pF Chip Capacitor	ATC600F6R8BT250XT	ATC
C14	22 pF Chip Capacitor	ATC800B220JT500XT	ATC
C15, C16	7.5 pF Chip Capacitor	GQM1875C2E7R5BB12D	Murata
C17, C18	12 pF Chip Capacitor	ATC600F120JT250XT	ATC
C19, C20	10 pF Chip Capacitor	ATC600F100JT250XT	ATC
C21	0.6 pF Chip Capacitor	ATC600F0R6BT250XT	ATC
C22	2.0 pF Chip Capacitor	ATC600F2R0BT250XT	ATC
C23	2.4 pF Chip Capacitor	ATC600F2R4BT250XT	ATC
C24, C25	1000 pF Chip Capacitor	ATC800B102JT50XT	ATC
C26, C27, C28, C29	10 μF Chip Capacitor	GRM32ER61H106KA12L	Murata
C30, C31, C32, C33, C34, C35, C36	15 μF Chip Capacitor	C5750X7S2A156M230KB	TDK
C37, C38	220 μF, 100 V Electrolytic Capacitor	MCGPR100V227M16X26-RH	Multicomp
R1, R2	4.75 Ω, 1/4 W Chip Resistor	CRCW12064R75FNEA	Vishay
R3	50 Ω, 10 W Termination Chip Resistor	81A7031-50-5F	Florida RF Labs
Z1	800–1000 MHz Band, 90°, 2 dB Asymmetric Coupler	CMX09Q02	Cemax
PCB	Rogers RO4350B, 0.020″, ε_r = 3.66	D89572	MTL

Figure 4.15: Reference Doherty Printed Circuit Board Vendor Bill of Materials

peak power (57dBm) levels across the three load-pull frequencies. In Fig. 4.18 and Fig. 4.19, the input return loss and the drain efficiency of the two transistor models is compared against the HPA measurement, while Table 4.4 shows the numerical values of the input return loss and drain efficiency at the average power and peak power levels across the three load-pull frequencies. While the TD-MTD model has a better approximation of the back-off efficiency of the PA, the compact model does not under-estimate the efficiency in the mid-power region of the power sweep as much as the behavioral model. Fig. 4.18 compares the simulated and measured input return loss as it varies with the output power in this reference PA. The extracted TD-MTD behavioral model has a better prediction of the return loss which is justifiable, given that the extracted TD-MTD model had a much better fit of the input impedance of the power transistors compared to the compact model. Part of the discrepancy between the simulation of the HPA Netlist, whether compact or behavioral model and the measurement of the fabricated HPA can be attributed to the inaccuracy of the simulation models used for the passive segments of the HPA circuit. The level of error seen in Table 4.2 compared to Table 4.1 suggests that most of the error could be attributed not the the active device error but due to other elements in the circuit while both the compact model and extracted TD-MTD model had similar performance in predicting the RF performance of the reference Doherty PA design. The transistor package used to perform the load-pull measurement was not the same as the transistor used in the reference circuit, which can attribute some of the difference between the modeled and

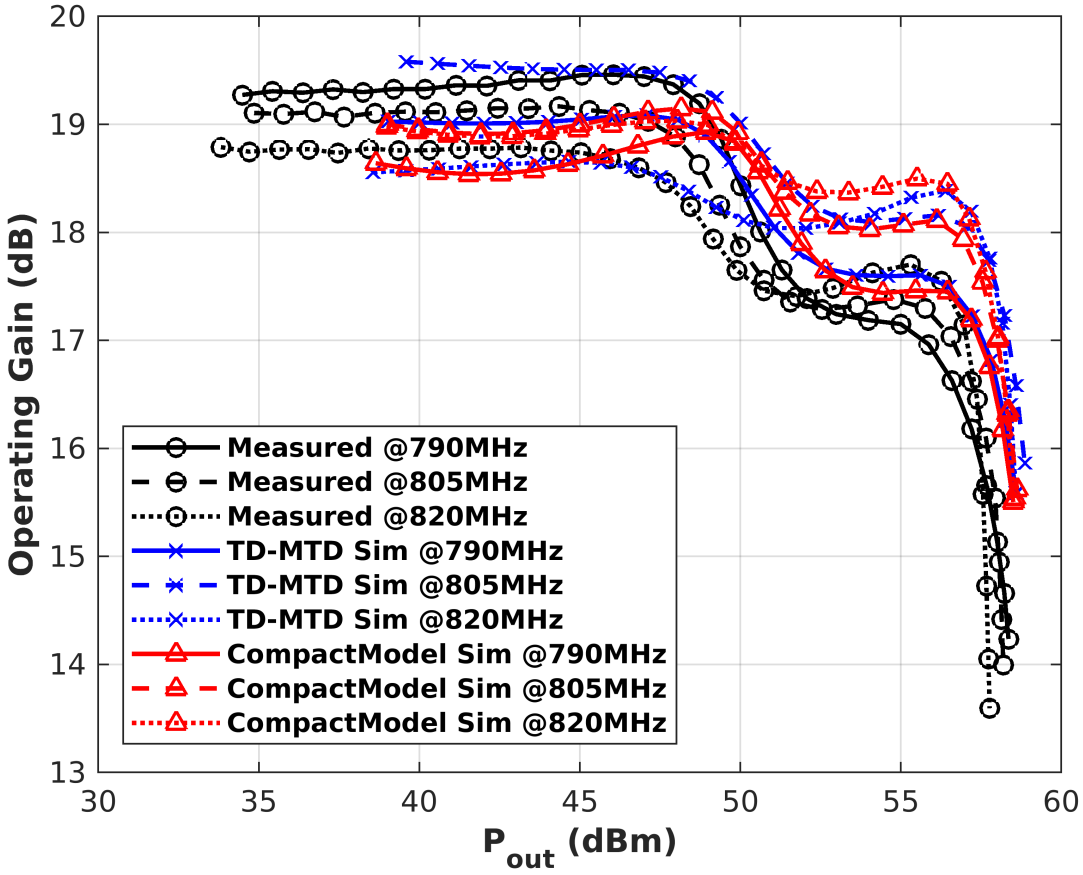

Figure 4.16: Comparison of the simulated and measured AMAM frequency variation of the reference Doherty PA

measured performance to device variation.

Designing an analog Doherty HPA with a behavioral model allows the HPA designer to simulate the nonlinear load modulation of both the main and the peaking power transistors as the power is ramped up. The designer can then track the performance of the HPA against the extracted load-pull characterization data and visualize the load-modulation provided by its input matching network and output matching and combining network design at all the intermediate power levels from back-off to peak power. Fig. 4.20 and Fig.4.21 show how the load impedance varies at the fundamental frequency of the main and peaking transistor respectively during the power drive up of the Doherty HPA.

As can be seen from the load-modulation simulation of the peaking transistor in Fig.

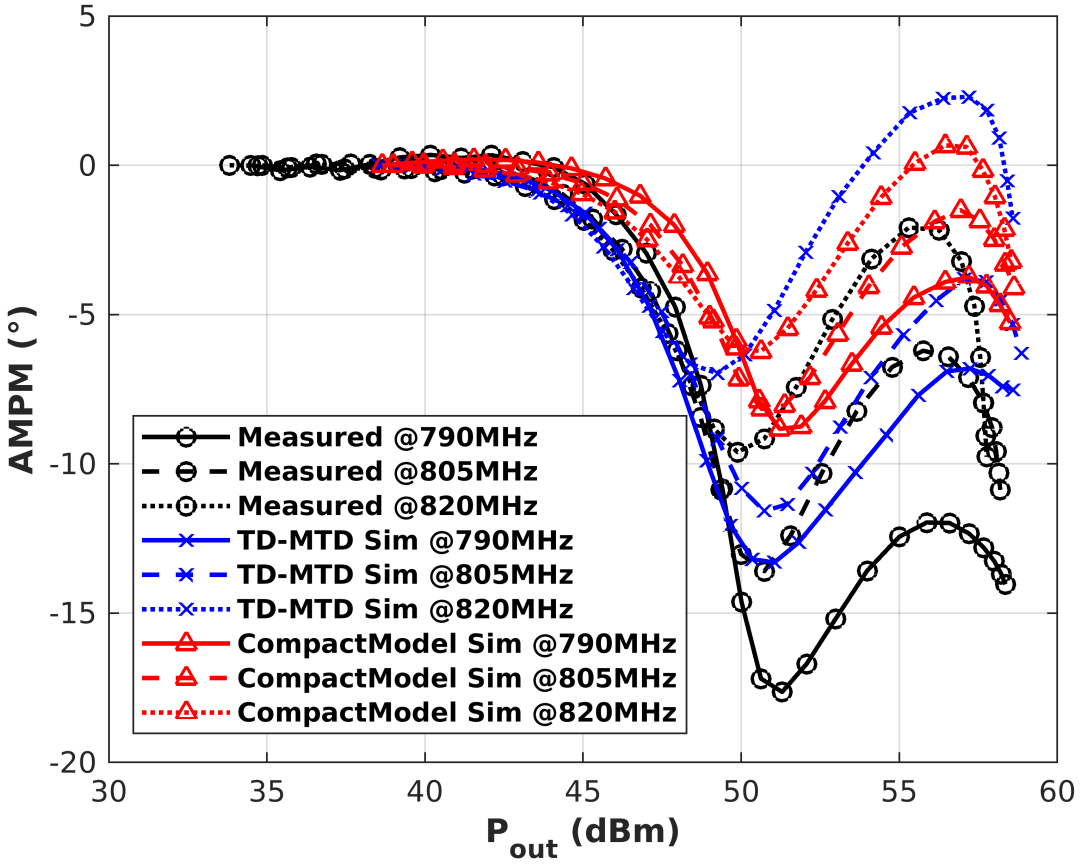

Figure 4.17: Comparison of the simulated and measured AMPM distortion frequency variation of the reference Doherty PA

4.21, the load impedance of the peaking device starts from around the complex conjugate of the peaking transistor's off-state impedance and moves towards its optimal design impedance close to the peak of the peaking device load-pull power contours at 1dB of gain compression. As can be seen from Fig. 4.21, the simulation of the peaking device during the Doherty PA simulation is presenting a drain impedance at back-off power that is outside of the passive load-pull characterization region of the peaking device. This means that the low power behavior of the peaking device TD-MTD model is smoothly extrapolated to a region outside of where the load-pull measurements were performed in order to have some prediction of the turn-on characteristics of the peaking devices when it is being modulated with significant power from the main device.

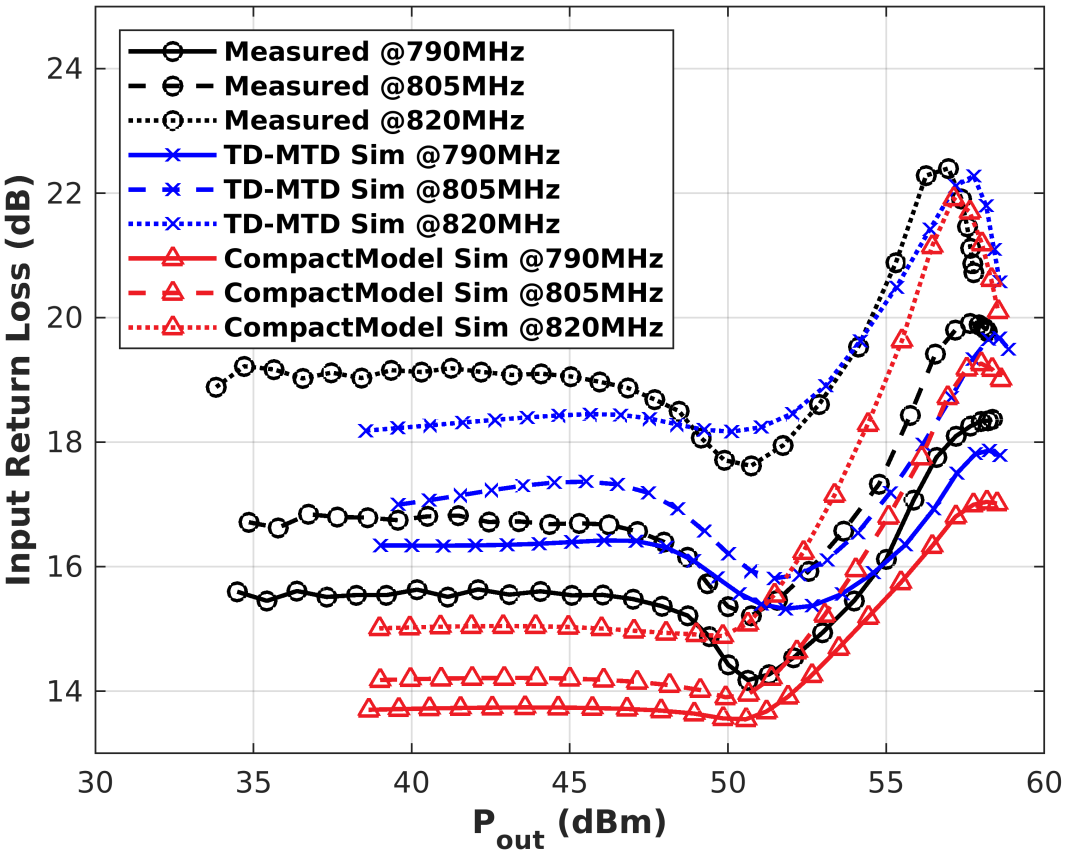

Figure 4.18: Comparison of the simulated and measured input return loss frequency variation of the reference Doherty PA

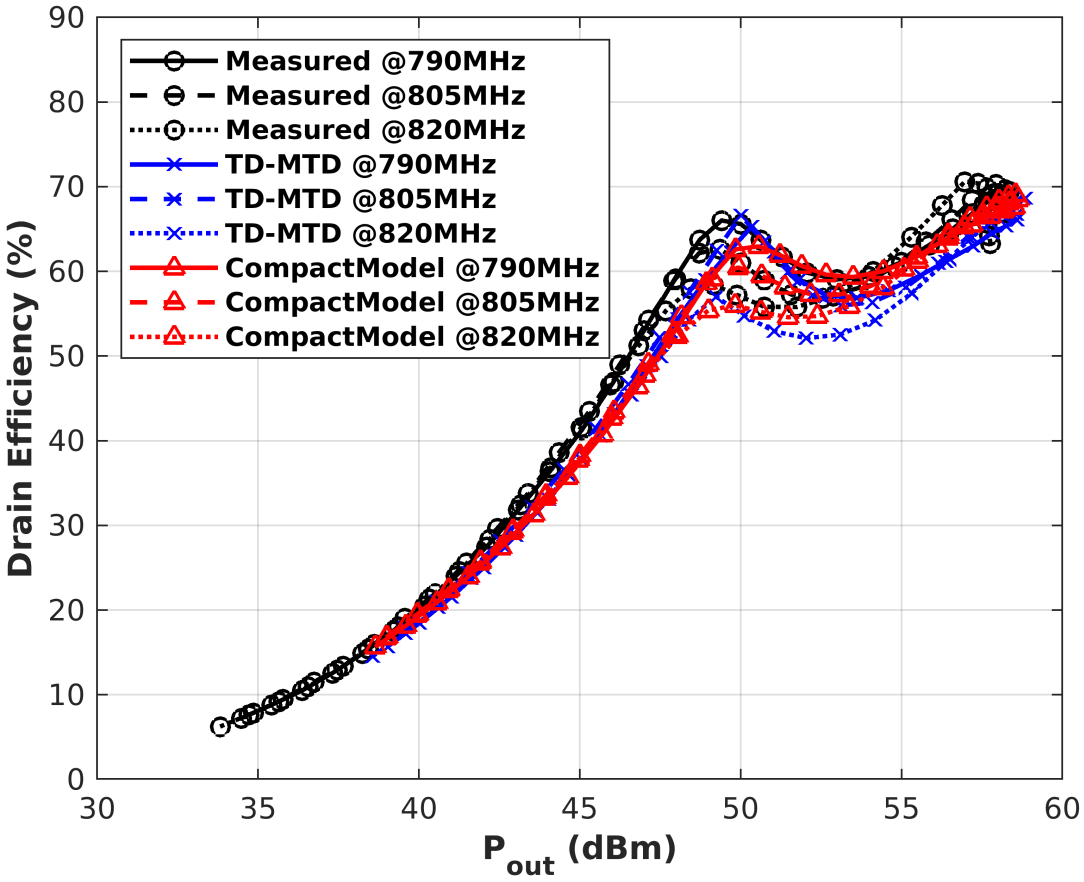

Figure 4.19: Comparison of the simulated and measured drain efficiency frequency variation of the reference Doherty PA

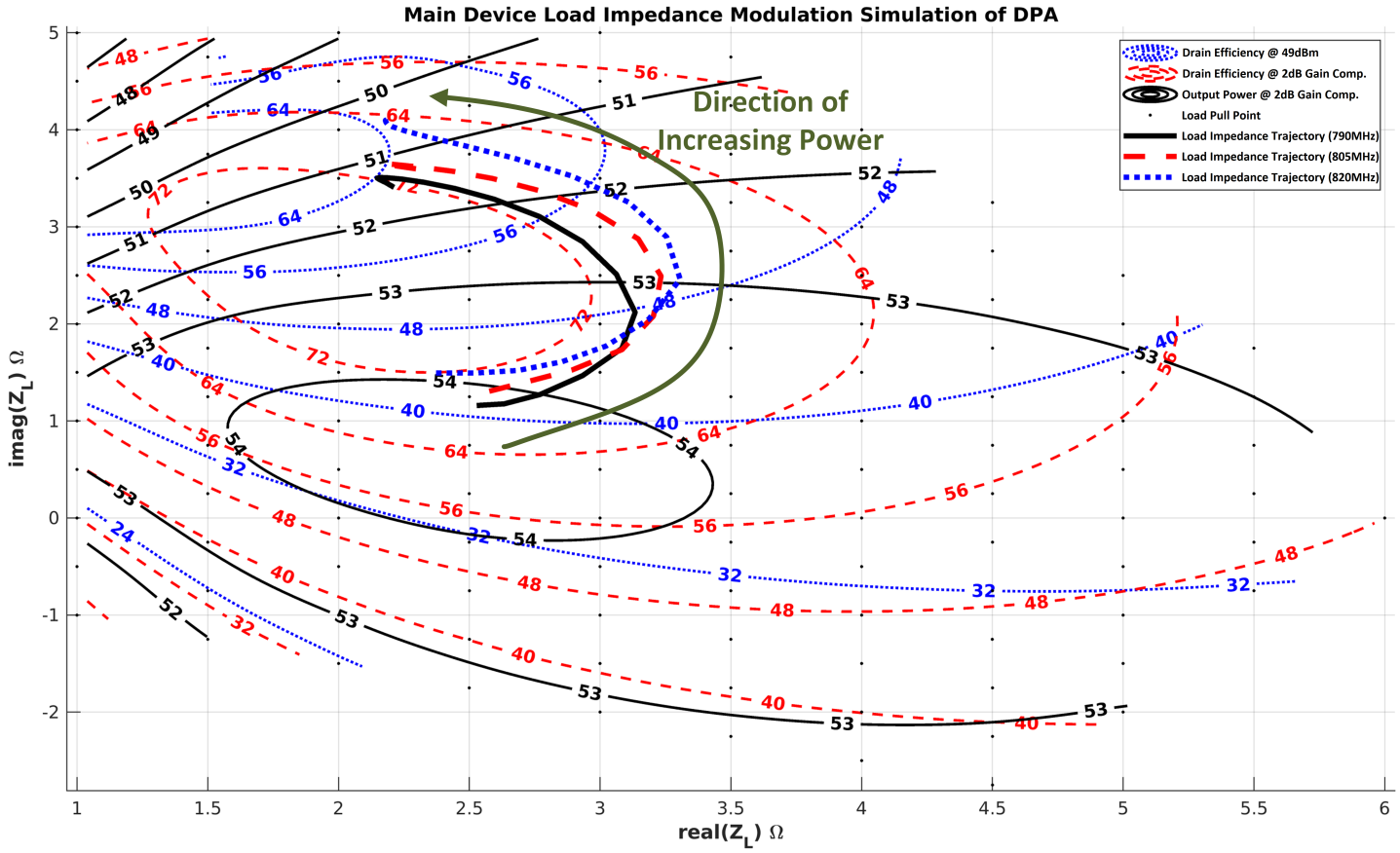

Figure 4.20: The load impedance modulation of the main device at the 3 frequencies in the Doherty HPA superimposed on the average power (49dBm) drain efficiency contours and the 2dB gain compression output power contours at the center frequency (805MHz)

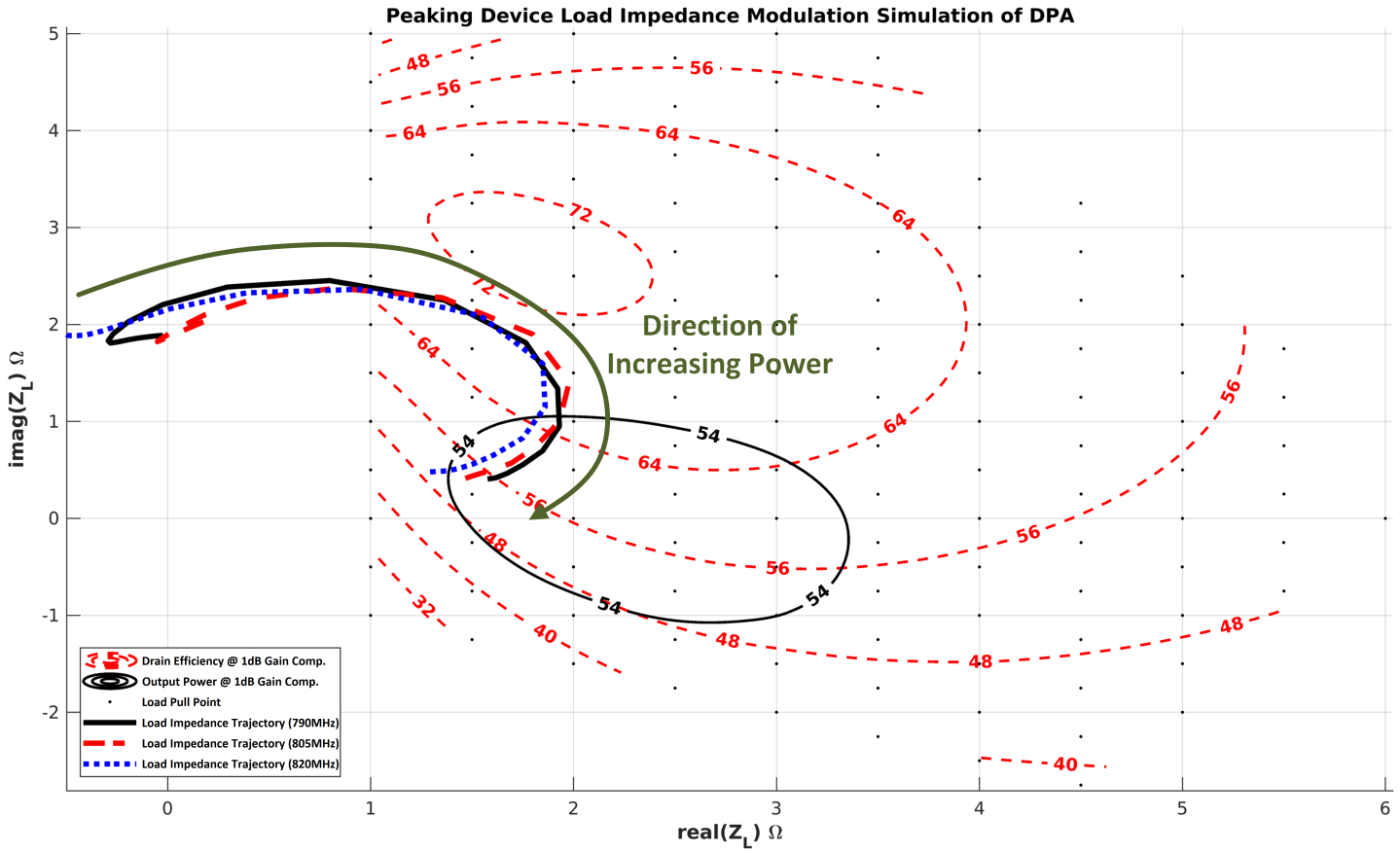

Figure 4.21: The load impedance modulation of the peaking device at the 3 frequencies in the Doherty HPA superimposed on the average power efficiency contours and the 1dB gain compression output power contours at the center frequency (805MHz)

4.3 Chapter Conclusion

Traditional ECAD-based design of Doherty high power amplifiers relied on compact models of power transistors. Since compact models are often not available at the same pace as transistor development, load-pull based designs of high power amplifiers are often employed to allow for quick turnaround. In this chapter, the TD-MTD behavioral model was proposed as an extension of the TD-PHD model that was proposed in the previous chapter. This model allows the high power amplifier designer to use load-pull characterization data of power transistors captured over a discrete set of non-uniformly spaced frequencies and convert that into a behavioural model of the power transistor for use in a simulation-based design environment. Using a TD-MTD based model simulation as the cornerstone of high power amplifier design can allow for fast turn-around of matching network designs without trading off accuracy.

As validation for the model presented in this chapter, first TD-MTD models are extracted for both the main and peaking transistors intended to be used in an LDMOS Doherty high power amplifier design. The simulation of this two-transistor circuit is shown to have less than 1 dB error in the prediction of the input return loss at both back-off-power and peak-power levels, and less than 0.8dB and 2° error in the back-off gain and phase compression and less than 1.4dB and 5.6° error in the peak power gain and phase compression, and an error of less than 6% in drain efficiency over the range of simulated frequencies, achieving a an NMSE of -18.89dB for predicting the B_2 wave and an NMSE of -13.17dB for predicting the DC Drain current over the power and frequency sweep. These errors were no worse than the compact model based design. The use of a TD-MTD model allows the high power amplifier designer to perform ECAD based design of a high power amplifier relying solely on load-pull characterization data without compromising any simulation accuracy compared to using a compact model as the power transistor model.

Chapter 5

Conclusions

The computer-aided design of power amplifiers requires an accurate representative model of the power transistor's nonlinear dynamic behaviour in a circuit simulation environment. While compact models of power transistors provide a bottom-up approach to build up a power transistor model based on theory and expertise, they are often not available to the power amplifier designers at the same pace as power transistor development. This has driven the practical need for power transistor behavioural models based on large-signal load-pull measurements of these devices.

The state of the art of behavioural models for power transistors before the work of this thesis were variants of the frequency-domain poly-harmonic distortion (PHD) models that described the spectral products of the output signals in the frequency domain as a time-invariant function of all spectral components of all the input signals. All of these PHD models required a distinct describing function for each of the spectral outputs of the model. Since one of the goals of this research was to develop models that tackle the hard nonlinearity exhibited by RF power transistors, the use of artificial neural networks as a tool for the development of robust simulation models was explored. We believe that a time-domain constructions of the proposed models of this thesis would provide a simpler form in terms of the number of inputs and output variables for the construction of an ANN behavioural model.

The contributions of this thesis involve the use of a time-domain description (as opposed to a frequency-domain description) to describe the behaviour of the power transistor in the nonlinear circuit simulator, allowing for a simplification in model representation and construction.

5.1 List of Contributions

5.1.1 A Model for the Multi-Harmonic Nonlinear Behaviour of Power Transistors at a Fixed Fundamental Frequency

Inspired by the fact that the time-domain description of a time-invariant system is more compact in its representation than an equivalent frequency-domain description, that is, the description of the nonlinear system is reduced to a single output variable at each port of the device in the time-domain description as opposed to requiring a different output variable at each frequency in a frequency-domain description, a new behavioural model was proposed called the time-domain poly-harmonic distortion (TD-PHD) model. This model used the duality between the discrete frequency and discrete time representations of signals to provide the same modeling capability of PHD models but in a more compact form. This compact functional form allows the use of advanced continuous function fitting tools like artificial neural networks model to model the continuous multi-variate nonlinear functions at the core of the model. These multi-variate nonlinear functions act upon delayed versions of the input signals spaced in time to span a single period of the periodic signal. Using a TD-PHD model a series of multi-harmonic source and load-pull measurements at a single DC bias point and a single fundamental frequency can be converted into a single time-domain model that captures the variation of the large-signal performance of a power transistor when the harmonic input and output matching network impedances that are loading the power transistor are varied[48][49].

5.1.2 A Model for the Nonlinear Behaviour of Power Transistors Across a Range of Fundamental Frequencies

Due to the TD-PHD model's auxiliary signals spanning a single fixed period by definition, its modeling capability was limited to a single fundamental frequency, requiring a separate TD-PHD model to be extracted for other fundamental frequencies. In general power amplifier designers wish to design their amplifiers over a band of frequency and have collected load-pull measurements at different fundamental frequencies over that band. To overcome the limitation of the TD-PHD model being fixed to a single fundamental frequency, a new model was proposed called the time-domain multi-tone distortion (TD-MTD) model that generalized the concept behind the construction of a TD-PHD model to a frequency grid that can contain load-pull measurements spanning multiple fundamental frequencies as long as all those fundamental frequencies were integer multiples of a shared common

fundamental frequency f_{CF}. The TD-PHD model then becomes a special case of the more general TD-MTD model, when $f_{CF} = f_0$. To validate the proposed TD-MTD model, models for two power transistors, one biased in class AB as the main device, and the other biased in class C as the peaking device in an asymmetrical two-way Doherty power amplifier were extracted based on single-harmonic load-pull measurements spanning a band of interest. The extracted models were then put in a simulated load-pull measurement and the simulation was able to reproduce the same performance as the measured large-signal data as demonstrated in the simulated model load-pull contours across frequency at different output powers. In a further circuit validation of the TD-MTD models, a reference power amplifier design from the transistor vendor was simulated based on EM models of the printed circuit board and simulation models of the RF capacitors and was shown to have no worse prediction of the Doherty power amplifier large-signal behaviour than the vendor-developed compact model. This shows the usefulness in using measurement-based models of power transistors in power amplifier development. Simulation-based design of multi-transistor power amplifiers models can also provide the power amplifier designer some insight as to how each of the power transistors in their design are being load-modulated at different power levels across frequency[62].

5.2 List of Relevant Publications

- Amir-Reza Amini and Slim Boumaiza. Time-invariant behavioral modeling for harmonic balance simulation based on waveform shape maps. In 2015 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO), pages 1-3, Aug 2015.

- Amir-Reza Amini and Slim Boumaiza. Time domain poly-harmonic distortion models of rf transistors and its extraction using a hybrid passive/active measurement setup. In 2017 IEEE MTT-S International Microwave Symposium (IMS), pages 1061-1064, June 2017.

- Amir-Reza Amini and Slim Boumaiza. A time-domain multi-tone distortion model for effective design of high power amplfiers. IEEE Access, 10:23152-23166, 2022.

5.3 Future Work

Currently the measurement-based validation provided for the TD-MTD model is based on fundamental-frequency load-pull measurements spanning multiple fundamental frequencies. However, the proposed model can also theoretically deal with harmonic load-pull data spanning multiple frequencies as well. An investigation into the ability of using the TD-MTD modeling approach for multi-harmonic load-pull data of a power transistor spanning multiple fundamental frequencies is warranted.

Another possible direction for the applications of the TD-MTD model is its use in modeling the multi-tone behaviour of power transistor, since the frequency grid of a multi-tone stimulus, with its mixing and inter-modulation frequencies can be chosen such that they are all multiples of a common fundamental frequency. This is a stepping stone towards modeling RF power transistors under modulated signal stimulus. The time-domain modeling schemes proposed in this thesis can naturally be extended to account for the so-called *long-term memory effects* that an RF power transistor exhibits when being stimulated by a modulated signal stimulus. A first step of extending the work done in this thesis to modulated signal behaviour is applying a similar scheme to the proposed TD-PHD/TD-MTD model but applied on load-pull data extracted from modulated stimulus in order to generate a behavioural model of the RF power transistor that targets *envelope transient harmonic balance simulations*.

In addition, since the basis of all the models proposed in this thesis are the Volterra series, this modeling framework can be used for other nonlinear devices or systems beyond power transistors including RF diodes and mixers, since nothing in this work suggests that the behavioural models proposed are only specific to power transistors as long as the underlying system exhibits continuous nonlinear time-invariant behaviour.

References

[1] S. C. Cripps. *RF Power Amplifiers For Wireless Communications*. Artech House, 2006.

[2] D. M. Pozar. *Microwave Engineering*. Wiley, 2005.

[3] J. Wood, P.H. Aaen, D. Bridges, D. Lamey, M. Guyonnet, D.S. Chan, and N. Monsauret. A nonlinear electro-thermal scalable model for high-power rf ldmos transistors. *Microwave Theory and Techniques, IEEE Transactions on*, 57(2):282–292, Feb 2009.

[4] J. Xu, J. Horn, M. Iwamoto, and D.E. Root. Large-signal fet model with multiple time scale dynamics from nonlinear vector network analyzer data. In *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International*, pages 1–1, May 2010.

[5] J. Wood, P. Aaen, D. Bridges, D. Lamey, M. Guyonnet, D. Chan, and N. Monsauret. *Modeling and Characterization of RF and Microwave Power FETs*. Cambridge University Press, 2007.

[6] I. Angelov, H. Zirath, and N. Rosman. A new empirical nonlinear model for hemt and mesfet devices. *Microwave Theory and Techniques, IEEE Transactions on*, 40(12):2258–2266, Dec 1992.

[7] I. Angelov, L. Bengtsson, and M. Garcia. Extensions of the chalmers nonlinear hemt and mesfet model. *Microwave Theory and Techniques, IEEE Transactions on*, 44(10):1664–1674, Oct 1996.

[8] D. Root. Measurement-based mathematical active device modeling for high frequency circuit simulation. *IEICE Transactions on Electronics*, 82:924–936, 1999.

[9] David E. Root, Siqi Fan, and Jeff Meyer. Technology independent large signal non quasi-static fet models by direct construction from automatically characterized device data. In *1991 21st European Microwave Conference*, volume 2, pages 927–932, 1991.

[10] Jianjun Xu, R. Jones, S.A. Harris, T. Nielsen, and D.E. Root. Dynamic fet model - dynafet - for gan transistors from nvna active source injection measurements. In *Microwave Symposium (IMS), 2014 IEEE MTT-S International*, pages 1–3, June 2014.

[11] Jianjun Xu, Subrata Halder, Faramarz Kharabi, John McMacken, Joseph Gering, and David E. Root. Global dynamic fet model for gan transistors: Dynafet model validation and comparison to locally tuned models. In *83rd ARFTG Microwave Measurement Conference*, pages 1–6, 2014.

[12] Anwar Jarndal, Saddam Husain, Mohammad Hashmi, and Fadhel M. Ghannouchi. Large-signal modeling of gan hemts using hybrid ga-ann, pso-svr, and gpr-based approaches. *IEEE Journal of the Electron Devices Society*, 9:195–208, 2021.

[13] Xuekun Du, Mohamed Helaoui, Anwar Jarndal, Taijun Liu, Biao Hu, Xin Hu, and Fadhel M. Ghannouchi. Ann-based large-signal model of algan/gan hemts with accurate buffer-related trapping effects characterization. *IEEE Transactions on Microwave Theory and Techniques*, 68(7):3090–3099, 2020.

[14] Gianni Bosi, Antonio Raffo, Rocco Giofrè, Valeria Vadalà, Giorgio Vannini, and Ernesto Limiti. Empowering gan-si hemt nonlinear modelling for doherty power amplifier design. In *2020 15th European Microwave Integrated Circuits Conference (EuMIC)*, pages 249–252, 2021.

[15] Viviana Cerantonio, Marcello Giuffrida, Cristina Miccoli, Alessandro Chini, and Ferdinando Iucolano. From t-cad simulations to large signal model for gan rf device. In *2020 AEIT International Conference of Electrical and Electronic Technologies for Automotive (AEIT AUTOMOTIVE)*, pages 1–6, 2020.

[16] Adrien Vaysset, Sébastien Martinie, François Triozon, Olivier Rozeau, Marie-Anne Jaud, René Escoffier, and Thierry Poiroux. Mos-like approach for compact modeling of high-electron-mobility transistor. In *2020 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 221–224, 2020.

[17] M. Nakhla and Jiri Vlach. A piecewise harmonic balance technique for determination of periodic response of nonlinear systems. *Circuits and Systems, IEEE Transactions on*, 23(2):85–91, Feb 1976.

[18] S. A. Maas. *Nonlinear Microwave and RF Circuits*. Artech House, 2nd edition edition, 2003.

[19] W. Van Moer and L. Gomme. Nvna versus lsna: enemies or friends? *Microwave Magazine, IEEE*, 11(1):97–103, Feb 2010.

[20] K.J. Silvonen. A general approach to network analyzer calibration. *IEEE Transactions on Microwave Theory and Techniques*, 40(4):754–759, 1992.

[21] A. Ferrero, F. Sanpietro, and U. Pisani. Accurate coaxial standard verification by multiport vector network analyzer. In *1994 IEEE MTT-S International Microwave Symposium Digest (Cat. No.94CH3389-4)*, pages 1365–1368 vol.3, 1994.

[22] A. Ferrero and U. Pisani. Two-port network analyzer calibration using an unknown 'thru'. *IEEE Microwave and Guided Wave Letters*, 2(12):505–507, 1992.

[23] Andrea Ferrero and Umberto Pisani. Qsolt: A new fast calibration algorithm for two port s parameter measurements. In *38th ARFTG Conference Digest*, volume 20, pages 15–24, 1991.

[24] H.-J. Eul and B. Schiek. Reducing the number of calibration standards for network analyzer calibration. *IEEE Transactions on Instrumentation and Measurement*, 40(4):732–735, 1991.

[25] J. Verspecht, P. Debie, A. Barel, and L. Martens. Accurate on wafer measurement of phase and amplitude of the spectral components of incident and scattered voltage waves at the signal ports of a nonlinear microwave device. In *Proceedings of 1995 IEEE MTT-S International Microwave Symposium*, pages 1029–1032 vol.3, 1995.

[26] T. Van den Broeck and J. Verspecht. Calibrated vectorial nonlinear-network analyzers. In *1994 IEEE MTT-S International Microwave Symposium Digest (Cat. No.94CH3389-4)*, pages 1069–1072 vol.2, 1994.

[27] Fadhel M. Ghannouchi and Mohammad S. Hashmi. *Load-Pull Techniques with Applications to Power Amplifier Design*. Springer, Netherlands, 1st edition, 2013.

[28] T. Williams, J. Benedikt, and P.J. Tasker. Application of a novel active envelope load pull architecture in large signal device characterization. In *2005 European Microwave Conference*, volume 1, pages 4 pp.–, 2005.

[29] Valeria Teppati, Andrea Ferrero, and Umberto Pisani. Recent advances in real-time load-pull systems. *IEEE Transactions on Instrumentation and Measurement*, 57(11):2640–2646, 2008.

[30] Valeria Teppati, Andrea Ferrero, Vittorio Camarchia, Alessandra Neri, and Marco Pirola. Microwave measurements - part iii: Advanced non-linear measurements. *IEEE Instrumentation Measurement Magazine*, 11(6):17–22, 2008.

[31] J.P. Teyssier, J.P. Viaud, J.J. Raoux, and R. Quere. Fully integrated nonlinear modeling and characterization system of microwave transistors with on-wafer pulsed measurements. In *Proceedings of 1995 IEEE MTT-S International Microwave Symposium*, pages 1033–1036 vol.3, 1995.

[32] J. Benedikt, R. Gaddi, P.J. Tasker, M. Goss, and M. Zadeh. High power time domain measurement system with active harmonic load-pull for high efficiency base station amplifier design. In *2000 IEEE MTT-S International Microwave Symposium Digest (Cat. No.00CH37017)*, volume 3, pages 1459–1462 vol.3, 2000.

[33] D.E. Root, Jan Verspecht, David Sharrit, J. Wood, and Alex Cognata. Broad-band poly-harmonic distortion (phd) behavioral models from fast automated simulations and large-signal vectorial network measurements. *Microwave Theory and Techniques, IEEE Transactions on*, 53:3656 – 3664, 12 2005.

[34] M. Oguz Yegin, Armagan Gurdal, Ulas Ozipek, and Ekmel Ozbay. A large-signal behavioural modeling approach of gan hemts for power amplifier design. In *2020 15th European Microwave Integrated Circuits Conference (EuMIC)*, pages 229–232, 2021.

[35] Guillermo Gonzalez. *Microwave Transistor Amplifiers (2Nd Ed.): Analysis and Design*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.

[36] B. P. Lathi. *Linear Systems and Signals*. Oxford University Press, Inc., New York, NY, USA, 2nd edition, 2009.

[37] Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing*. Prentice Hall Press, USA, 3rd edition, 2009.

[38] J. Verspecht. Hot s-parameter techniques: 6 = 4 + 2. In *66th ARFTG Microwave Measurements Conf. Dig*, pages 7–15, http://www.janverspecht.com, 2005.

[39] D.E Root, J. Verspecht, J. Horn, and M. Marcu. *X-parameters*. Number 9780521193238. Cambridge University Press, 2013.

[40] Jan Verspecht and Patrick Van Esch. Accurately characterizing hard nonlinear behavior of microwave components with the nonlinear network measurement system: Introducing "nonlinear scattering functions". 06 1999.

[41] J. Verspecht. Large-signal network analysis. *IEEE Microwave Magazine*, 6(4):82–92, 2005.

[42] J. Verspecht and D.E. Root. Polyharmonic distortion modeling. *Microwave Magazine, IEEE*, 7(3):44–57, June 2006.

[43] A. Amini D. T. Bespalko and S. Boumaiza. A high-order model looking beyond the first-order harmonic superposition assumption. In *IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications (PAWR)*, pages 42–44, Austin, TX, 2016.

[44] S. P. Woodington, R. S. Saini, D. Willams, J. Lees, J. Benedikt, and P. J. Tasker. Behavioral model analysis of active harmonic load-pull measurements. In *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International*, pages 1–1, May 2010.

[45] J.J. Bell, R. Saini, S. Woodington, J. Lees, J. Benedikt, S. Cripps, and P.J. Tasker. Behavioral model analysis using simultaneous active fundamental load-pull and harmonic source-pull measurements at x-band. In *Microwave Symposium Digest (MTT), 2011 IEEE MTT-S International*, pages 1–4, June 2011.

[46] Jialin Cai, Jun Liu, Baicao Pan, Xuekun Du, and Justin King. A modified poly-harmonic distortion model based on the canonical piecewise linear functions for gan hemts. *IEEE Access*, 8:181420–181431, 2020.

[47] Jialin Cai, J.B. King, B.M. Merrick, and T.J. Brazil. Padé-approximation-based behavioral modeling. *Microwave Theory and Techniques, IEEE Transactions on*, 61(12):4418–4427, Dec 2013.

[48] A. Amini and S. Boumaiza. Time-invariant behavioral modeling for harmonic balance simulation based on waveform shape maps. In *2015 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)*, pages 1–3, Aug 2015.

[49] A. Amini and S. Boumaiza. Time domain poly-harmonic distortion models of rf transistors and its extraction using a hybrid passive/active measurement setup. In *2017 IEEE MTT-S International Microwave Symposium (IMS)*, pages 1061–1064, June 2017.

[50] Kenneth Levenberg. A method for the solution of certain non-linear problems in least squares. *The Quarterly of Applied Mathematics*, (2):164–168, 1944.

[51] Donald W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[52] O. Hammi, J. Sirois, S. Boumaiza, and F. M. Ghannouchi. Design and performance analysis of mismatched doherty amplifiers using an accurate load-pull-based model. *IEEE Transactions on Microwave Theory and Techniques*, 54(8):3246–3254, Aug 2006.

[53] J. Sirois, S. Boumaiza, M. Helaoui, G. Brassard, and F. M. Ghannouchi. A robust modeling and design approach for dynamically loaded and digitally linearized doherty amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 53(9):2875–2883, Sep. 2005.

[54] T. S. Nielsen, M. Dieudonne, C. Gillease, and D. E. Root. Doherty power amplifier design in gallium nitride technology using a nonlinear vector network analyzer and x-parameters. In *2012 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*, pages 1–4, Oct 2012.

[55] R. Hajji, M. Poulton, D. B. Crittenden, J. Gengler, and P. Xia. Gan-hemt nonlinear modeling of single-ended and doherty high-power amplifiers. In *2014 9th European Microwave Integrated Circuit Conference*, pages 373–376, Oct 2014.

[56] M. R. Fréchet. Sur les fonctionnelles continues. *Annales Scientifiques de l'Ecole Normale Superieure*, 27:193–219, 1910.

[57] J. Barrett. The use of functionals in the analysis of nonlinear physical systems. *J. Elect. Control.*, 15(567-615), 1963.

[58] J. Barrett. Functional series representation of nonlinear systems-some theoretical comments. In *IFAC Syp. Ident. Sys. Param. Est.*, 1982.

[59] V. Volterra. *Theory of functionals and of integral and integro-differential equations.* Dover Publications, 1930.

[60] Sanjit K. Bagchi, Sonali; Mitra. *The Nonuniform Discrete Fourier Transform and Its Applications in Signal Processing.* Springer, Boston, MA, USA, 1999.

[61] June-Yub Greengard, Leslie; Lee. Accelerating the nonuniform fast fourier transform. *SIAM Review*, 46(3):443–454, January 2004.

[62] Amir-Reza Amini and Slim Boumaiza. A time-domain multi-tone distortion model for effective design of high power amplifiers. *IEEE Access*, 10:23152–23166, 2022.