

Nonlinear Data Utilization: Direct Data Look-up to Behavioural Modelling

A thesis submitted to the University of Wales, Cardiff

In candidature for the degree of

Doctor of Philosophy

By

Hao Qi

School of Engineering

Cardiff University

United Kingdom

February 2008

UMI Number: U585153

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U585153

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENT

First of all, I would like to thank my supervisors Dr. Johannes Benedikt and Prof. Paul Tasker for their inspiration, support and guidance throughout the duration of this research project. It was Dr. Benedikt's patience and enthusiasm that introduce me to this wonderful research area. It was Prof. Tasker's valuable advices and helpful discussions that let me fulfil the research tasks well on time.

I would like to thank Nokia Networks in Finland for the financial and technical support. This work would not have been possible without their cooperation.

During my time spent in Cardiff, it was a great pleasure to work in High Frequency Centre. Particularly, I would like to say thank you to my colleagues: Jonny Lee, Zaid Aboush, Tudor Williams, Aamir Sheikh, and Chris Roff for their help, advice, and support in my research. It is my honour to work with such an excellent group of talents in microwave engineering.

Finally, I would like to thank my parents and fiancée Ying for their understanding and encouragement during my time in Cardiff. It is their love that gives me the strength and courage to challenge myself by pursuing this research.

TABLE OF CONTENTS

**NONLINEAR DATA UTILIZATION: DIRECT DATA
LOOK-UP TO BEHAVIOURAL MODELLING 1**

SUMMARY 3

LIST OF PUBLICATIONS 3

CHAPTER 1 INTRODUCTION 3

1.1 RF power amplifier for wireless communication systems.....3

1.2 RF power device characterizations3

1.3 RF Large signal power amplifier design philosophies3

1.3.1 Measurement-based power amplifier design philosophy..... 3

1.3.2 Model-based power amplifier design philosophy 3

1.4 Research motivation3

1.5 Research objectives.....3

1.6 Thesis structure.....3

**CHAPTER 2 POWER AMPLIFIER TRANSISTOR
DEVICES: CHARACTERIZATION AND MODELLING 3**

2.1. Nonlinearity in RF power amplifiers3

2.2. Large signal power device characterization.....3

2.3. Power transistor device modelling3

CHAPTER 3 RF LARGE SIGNAL WAVEFORM

MEASUREMENTS..... 3

3.1. Cardiff RF large signal waveform measurement system.....3

3.1.1. Cardiff waveform measurement system basics..... 3

3.1.2. High power waveform measurement issues..... 3

3.2. Single tone waveform characterization3

CHAPTER 4 DIRECT UTILIZATION OF WAVEFORM

DATA IN CAD..... 3

4.1 Introduction.....3

4.2 Issues related to RF large signal CAD simulations.....3

4.3 Large signal waveform data scaling.....3

4.3.1 Large signal measurement data at a first glance 3

4.3.2 Measurement data scaling 3

4.3.3 “Grey box” versus “black box”..... 3

4.3.4 Data table generation..... 3

4.4 Import of waveform data into CAD.....3

4.5 Experimental validations of the direct utilization approach.....3

4.5.1 Harmonic balance simulation using imported waveform data..... 3

4.5.2 Simulation on measurement grid 3

4.5.3 Simulation off measurement grid I: interpolation 3

4.5.4 Simulation off measurement grid II: extrapolation 3

4.5.5 Testing for technology-independent..... 3

4.6 Expandability in harmonic and multi-tone simulation.....3

4.6.1 Harmonic load-pull simulation 3

4.6.2 Two tone simulations via envelope simulation 3

4.7 Summary.....3

CHAPTER 5 POLYNOMIAL-BASED BEHAVIOURAL

MODELLING..... 3

5.1. Introduction.....3

5.2.	Choosing behavioural modelling of nonlinearity.....	3
5.3.	Original PHD modelling approach	3
5.4.	Polynomial-based behavioural model formulation	3
5.5.	Generalization of polynomial-based model formulation	3
5.6.	Direct model extraction from large signal waveform data	3
5.6.1.	Travelling wave renormalization	3
5.6.2.	Data interpolation.....	3
5.6.3.	Integral optimization of model parameters	3
5.7.	Polynomial-based model implementation	3
5.8.	Model experimental verification	3
5.8.1.	Simulation versus measurement on 100w LDMOS.....	3
5.8.2.	Simulation versus measurement on 5w LDMOS.....	3
5.8.3.	Impact of order of model on simulation accuracy	3
5.8.4.	Impact of measurement grid density on extraction accuracy.....	3
5.8.5.	Verification of model extraction consistency.....	3
5.9.	Summary.....	3

CHAPTER 6 DIRECT WAVEFORM UTILIZATION

VERSUS BEHAVIOURAL MODELLING..... 3

6.1.	Introduction.....	3
6.2.	Metrics for verifying overall accuracy of simulation	3
6.3.	Comparisons on simulation robustness	3
6.4.	Comparisons on simulation speed.....	3
6.5.	Comparisons on approach complexity.....	3
6.6.	Summary.....	3

CHAPTER 7 FUTURE WORK..... 3

7.1 General.....	3
7.1.1 Improving measurement speed	3
7.1.2 Thermal measurement	3
7.1.3 Modelling memory effects	3
7.1.4 Application of data utilization approaches in PA design.....	3
7.2 Direct waveform utilization approach	3
7.2.1 Source-pull data utilization	3
7.2.2 Simulation with modulated signal.....	3
7.3 Polynomial-based behavioural modelling approach	3
7.3.1 Modelling DC and harmonics	3
7.3.2 Optimization of model extraction	3
CONCLUSION	3
LIST OF ABBREVIATIONS	3
REFERENCES.....	3
APPENDICES	3
Appendix A Software User Manual	3
A.1. DWLU Table Generator	3
A.2. Polynomial-based Model Extractor.....	3
Appendix B Readme for the attached CD	3
Appendix C Additional DWLU Approach Verification Results.....	3
Appendix D Relevant Publications.....	3

SUMMARY

Newly developed communication systems put strict requirements on the performance of RF power amplifiers. A key issue for the development of RF PA is the inherent nonlinearity of power amplifiers hindering its integration with the well established small-signal development infrastructure which forms a closely interlinked chain of measurement systems, small-signal models and CAD based simulation and design software. The linkage between these components is provided by common small-signal data import and export file formats ensuring a bidirectional data exchange without any loss of the small-signal information. However, no equivalent infrastructure exists for a large-signal design process inhibiting the development of RF power amplifiers and other nonlinear components.

This work demonstrates a coherent methodology aiming to provide equivalent infrastructure for large signal design process as already exists in small signal design process. As first part of the methodology, a new approach is proposed to directly import measured current and voltage waveforms, obtained from a typical Large Signal Network Analyser (LSNA) system recently developed in Cardiff University, into nonlinear CAD simulator for power amplifier design. This approach offers an efficient solution for using large signal characteristic data in CAD-based simulation and PA design as the simulation/design accuracy is guaranteed by measurement and the no lengthy data processing is required. The approach is implemented in Agilent ADS simulator and its validity is comprehensively verified on different devices and device technologies. Moreover, the potential of it in predicting device large signal performance when interpolation or extrapolation is needed is explored.

As second part of the methodology, a new large signal nonlinear behavioural approach is proposed from behavioural modelling perspective as a complementation to the direct waveform utilization approach. The proposed modelling approach features in impressive simulation speed while maintaining excellent simulation accuracy. The modelling approach is developed on the basis of polynomial approximation and theoretical analysis shows that the approach can be considered as reasonable extension of S parameter design tool in large signal environment. It's demonstrated in this work that the model is extracted from large signal waveform data with specially designed parameter extraction procedure. The extracted model is verified on several devices and repeatable accuracy can be obtained even on high power devices such as 100w LDMOS.

It's illustrated in this work that the above two distinctive approaches can be combined and nicely considered as parts of an integrated nonlinear measurement data utilization strategy. Such a strategy provides a fast and time efficient path to accurate CAD-based nonlinear design; even at power levels relevant for base station applications.

LIST OF PUBLICATIONS

1. Qi, H.; Benedikt, J.; Tasker, P, “Novel Nonlinear Model for Rapid Waveform-based Extraction Enabling Accurate High Power PA Design,” IEEE MTT-S International Microwave Symposium, 2007

Abstract — This paper presents a new large signal modelling technique that makes substantial extension to the Polyharmonic Distortion (PHD) model concept. The extension accounts for the influence of travelling wave a_2 magnitude presented on the output port on the variation of S and T parameters. The inherent relationship between a_2 and the S and T parameters has been illustrated theoretically and experimentally. Automated extraction procedure for the polynomial-based model is developed which directly utilizes large-signal waveform data. The resulting model was implemented in ADS showing very good agreement with measurement data on a 100W LDMOS device.

2. Qi, H.; Benedikt, J.; Tasker, P, “A Novel Approach for Effective Import of Nonlinear Device Characteristics into CAD for Large Signal Power Amplifier Design,” IEEE MTT-S International Microwave Symposium, 2006

Abstract — This paper introduces a new approach allowing for the direct utilization of large signal measurement data in the PA CAD design process. A model, which we term as “truth look-up model”, has been developed to import non-linear measurement data into a non-linear CAD simulator. Its implementation in *Agilent ADS* simulator is verified on several different types of high frequency transistors. Through correct formulation of the “truth look-up model” data sets, it’s been found that this approach can accurately predict the nonlinear behaviours of devices under CW excitations, even when interpolation or extrapolation of measured data is required. Furthermore, the model also demonstrates its potential for predicting the behaviour of devices under multi-tone signal excitations.

3. Qi, H.; Benedikt, J.; Tasker, P, “A Novel Nonlinear “Truth Model On-Demand” for Rapid Large-Signal Power Amplifier Design,” *The Integrated Nonlinear Microwave and Millimetre-wave Circuits (INMMiC) Workshop*, 2006

Abstract—This paper introduces a new approach for the direct utilization of large signal measurement data for rapid CAD-driven power amplifier design. It presents a method of directly importing non-linear measurement data into a

non-linear CAD simulator. A model named as “Truth model” has been proposed for formulating the data. The model has been implemented in ADS and it can accurately predict the large signal behaviour of the device under CW excitations. The model’s robustness on interpolation and extrapolation has been successfully verified on a 30w LDMOS device.

4. **Qi, H.; Benedikt, J.; Tasker, P, “Direct extraction of large-signal table-based behavioural models from time-domain voltage and current waveforms,” High Frequency Postgraduate Student Colloquium, Page(s): 25 – 28, 2005**

Abstract — This paper presents a novel table-based nonlinear behavioural model extracted directly from large-signal time-domain measurements. The model has been verified on several types of power devices and can accurately predict the behaviours of devices under large-signal CW excitations -the model is implemented in Agilent ADS.

Chapter 1 Introduction

1.1 RF power amplifier for wireless communication systems

Modern wireless communication started with Guglielmo Marconi in 1896, who managed to illustrate the wireless telegraphy by sending and receiving Morse code. Within several decades of evolution, the wireless communication has now become able to offer billions plus business opportunities. Accordingly, a personal portable device rapidly became a mass-market consumer product. There were about 250 Million subscribers in the year 2000, and a market of 500 to 600 Million handhelds per year. The huge market and great potential for making profit result in the rapid growth of the wireless communication industry. It makes the competition more and more intense and also makes the spectrum resource the most expensive asset in the world. Besides, the new demand has emerged regarding the contents of service that the operator should be able to provide. Compared to the traditional voice service, the enhanced multimedia services, such as electric mail, data, video, internet surfing and so on, have started to enter the market. The tremendous efforts are being made to investigate and develop new wireless communication technology which is aimed to offer the competitive multimedia service and caters as many customers as possible within the limited bandwidth. A number of novel technologies have been proposed and some of them have even been realized in practice. However, these new technologies offer a superior utilization rate at the expense of increased system complexity. It makes the wireless system design a much more stringent process than ever before.

A key part in any wireless communication system is a RF power amplifier. Various types of power amplifiers ranging from a few Watts to kilo Watts are widely used in the portable handheld device as well as in the wireless network infrastructure. Basically a RF power amplifier is used to convert the DC power into the RF power,

which enables the transmission of the RF signal containing the digital information from a transmitter to a receiver via the wireless environment. The quality of a RF power amplifier has substantial influence on the cost, reliability, size and performance of a wireless communication system.

Newly developed wireless technologies put some strict requirement on the performance of a RF power amplifier. The RF power amplifier has to be carefully designed in order to meet the specific need for various systems. There are many figures of merits which can be used to evaluate the quality of a RF power amplifier quantitatively. Two of the most important ones are the efficiency and linearity. The efficiency refers to a measure of how efficient a power amplifier can convert the DC power into the RF power. The efficiency is important since a power amplifier with the poor efficiency obviously consume much more power than an efficient one when delivering the same amount of power. For a handheld application that usually uses the battery as the primary power source, the poor efficiency means shorter talk time and standby time. It can seriously limit the competency of the final product in the market. Besides, the poor efficiency leads to the large amount of heat dissipation which incurs the extra cost for the high-capacity cooling facility. The linearity refers to a measure of how faithfully a power amplifier can amplify an input signal. A power amplifier always demonstrates certain levels of nonlinearity. The tolerance for the nonlinearity really depends on the application. For example, apart from the usual high requirement for the efficiency, CDMA and WCDMA require the high linearity over a large dynamic power range. On contrast, the linearity is a relatively less serious issue for the GSM power amplifier design because of the constant envelope characteristics of the GSM signal. Generally speaking, the linearity and efficiency demand contradictory solutions. It's always impossible for these figures of merits to reach their maximums simultaneously on a stand alone power device. Various types of advanced power amplifier architectures are proposed recently to tackle the contradiction and aimed to achieve good efficiency and linearity at the same time. Although some promising results are achieved, these architectures inevitably

increase the complexity of the communication system which results in the higher cost and worse reliability. Therefore, the challenge for the power amplifier design community is to seek an appropriate optimization solution with which a good compromise of specification can be reached. In order to obtain the desired compromise, it's important to gain the deep knowledge of appropriate type of power amplifier devices and adopt a proper design strategy.

Different system manufacturers may have different performance specifications for a RF power amplifier. Besides, the specifications can be very different depending on the type of applications such as the handheld device or base station infrastructure and modulation scheme. Therefore, the critical first move is to identify a proper device technology for the specific application of interest. As the device technology is developing very quickly, many types of device technologies co-exist in the market. Different types of devices usually have different RF properties which can be highly appreciated by specific applications. For example, the lateral-drain-diffusion MOS (LDMOS) device demonstrates a good output power capability with the possibility of low cost mass-production. These features make the LDMOS device technology outstanding in the fixed product application and consequently the leading device technology in the high power telecommunication base station application. In the handheld application which usually requires a few Watts of power, the use of the battery power dictates the importance of efficiency. A wide variety of device technologies can meet the general requirements of the handheld application such as the GaAs FET and GaAs HBT. However, there are some important tradeoffs among these devices. For example, GaAs HBT has higher power density while the GaAs FET has higher power gain and efficiency. Therefore, careful consideration and investigation of the device performance are usually necessary to pick up the right device which can achieve best tradeoffs.

1.2 RF power device characterizations

A good power amplifier can only be designed if the accurate knowledge of the target power amplifier semiconductor device can be obtained. The device characterization should always be conducted before moving into the design stage. The aim of the device characterization is to capture the device typical response with respect to certain external parameters of interests. These parameters may include the biasing condition, frequency, drive power, impedance environment and etc. Different device characterization techniques have been developed in order to meet various demands from the PA designer. Traditionally, the classical small signal s-parameters [1] are used for characterization. The corresponding measurement instrument called Vector Network Analyzer (VNA) has been developed and widely utilized for the s-parameter related measurement. However, as the power level goes up, a transistor device starts to demonstrate the nonlinear behaviour. The classical s-parameter is essentially a linear characterization tool and therefore not sufficient for the design purpose in the large signal operation. The large signal characterization technology such as load-pulling is developed in order to characterize the nonlinearity shown by a transistor device.

1.3 RF Large signal power amplifier design philosophies

The device characterization provides the PA designer the opportunity to investigate the device typical behaviour in the desired situation. The data obtained from different types of device characterization techniques may present the device characteristics from different perspectives. So the next task is to interpret the data properly and utilize the data as effectively as possible in the PA design. To complement measurement technologies, simulation algorithms, modelling and power amplifier topologies, the theory needed for the PA design philosophy is also evolving. Generally speaking, there are two different types of PA design philosophies. One can be termed as the measurement-based approach and the other can be termed as the modelling-based approach.

1.3.1 Measurement-based power amplifier design philosophy

The measurement-based PA design philosophy takes the measurement data as the primary resource for designing a power amplifier. In this philosophy, a power amplifier transistor device will firstly be characterized. The key behaviour of the device covering the gain, efficiency and load-pull contours will be recorded. Based on the information, designers start the power amplifier design straightaway. Several prototypes are then made in order to test the designed auxiliary circuitries and tune the overall system performance. The design process will be finalised once the designed PA passes tests successfully. The general design flow is shown in figure 1-1.

The advantage of this approach is that no time-consuming modelling is needed. The design is based on the reliable measurement data and thus saves some of the iterations usually expected for the model development in the modelling-based approach. The development of measurement technologies has great influence on the effectiveness of this philosophy. The recent advances of the large signal measurement instrumentation largely enhance the capability of the measurement-based design philosophy. These advanced large signal measurement systems allow a designer to access more detailed key characteristics with the improved accuracy when the device under test is driven in the realistic working condition. For example, the large signal waveform can now be measured by using the newly developed LSNA system [2]-[4]. The availability of the large signal waveform gives a designer the possibility to engineer the waveform directly for the high efficiency power amplifier design. Therefore, by taking advantage of these advanced measurement technologies, the measurement-based PA design philosophy becomes even more attractive than before.

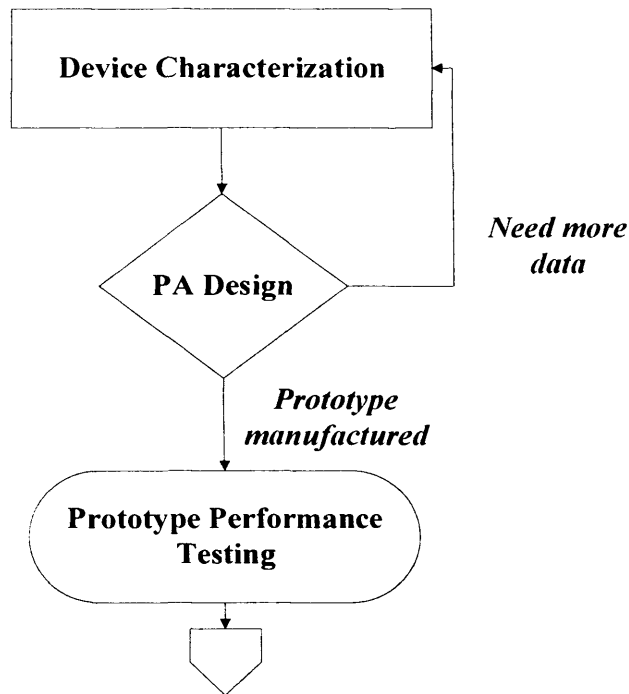


Figure 1-1: Measurement based design flow

1.3.2 Model-based power amplifier design philosophy

The modelling based design philosophy uses a model as the primary design tool. In this philosophy, firstly a device model is developed in theory. Then the model extraction procedure is identified. After that the target device is characterized experimentally in order to obtain the necessary measurement data required for the model extraction. By following the extraction procedure, the model can be extracted and preferably implemented in the Computer-Assisted Design (CAD) environment. The accuracy and reliability of the model is tested and an iterative process is usually needed. The model can be released once the sufficient accuracy is reached. The PA designer can start the CAD-based power amplifier design by using the model. The designed PA will be finalised if all the design specifications can be met at the PA performance testing stage. The general design flow is shown in figure 1-2.

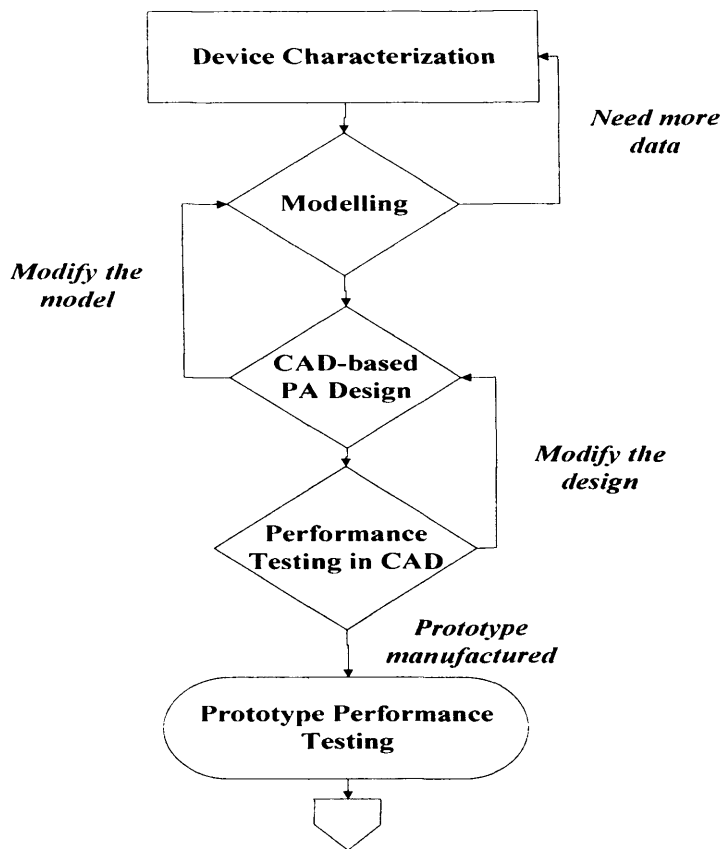


Figure 1-2: The modelling-based design approach

One important advantage of this approach is that the Computer-Assisted Design (CAD) tool can be utilized in a design via a nonlinear model. The CAD tool is the computer-based software that can help in the design activity. The CAD tool becomes very important in the modern engineering world as it can provide many benefits such as lower development cost and shorten design cycle. It's particularly true in the RF PA design world. Recently the complexity of a power amplifier increases significantly due to the stricter specification required by the latest wireless technologies. Consequently it puts some greater challenge on the PA designers. The RF CAD technology is evolving rapidly and the modern CAD tool, such as the Agilent ADS [5] and Ansoft NEXXIM [6], can perform many computation-heavy tasks. They can therefore assist the designer to tackle the complicated design problem. For example, the robust nonlinear simulation algorithm like the harmonic balance has been developed which allow the designer to design the necessary circuitry around a model and test the overall performance of the system even before building up an expensive prototype. Moreover it also gives the designer a less costly

option to go back and restart a design if some big design flaw is spotted during the simulation. The use of the CAD tool in the design and testing stages significantly increases the chance of a first pass design.

1.4 Research motivation

Although the modelling-based and measurement-based PA design philosophies already demonstrate many interesting advantages respectively, both of them suffer from some obvious limitations. One limitation of the modelling-based philosophy is the lengthy modelling procedure. The design procedure cannot get started unless a suitable model is developed. In this modelling-based philosophy, the accuracy of the model is a key point which could determine the quality of the final design output. No matter how good a CAD tool is, the accuracy of the CAD can only be as good as the model. In order to achieve reasonable accuracy, careful experimental design for the model identification is needed. The identified model need to be verified and some iterations and parameter tweaking are expected. All these factors will make the model development very time-consuming and can significantly delay the time-to-market of the final design. Besides, during the model development, special measurement setup sometimes is required in order to obtain the data in the right format for the model parameter identification. The special setup can incur extra delay and cost. Since the special setup is generally designed for extracting certain type of model, the setup may not be reused if a new modelling approach is employed.

On the other hand, the measurement-based design philosophy is limited mainly due to the lack of the valuable support from CAD tools. Since there is no suitable model, the auxiliary circuitry like the matching network cannot be tested with the active power device on the system level in a CAD. The designer has to wait until the prototypes are manufactured. Besides, it's sometimes very complicated if one attempts to design the power amplifier working in the complex mode like class-F by

using the measurement-based approach. Things would get much easier if a CAD tool can be used in which a robust nonlinear simulation algorithm can help to resolve complicated design issues. Moreover, if any big flaw in the design is spotted after the prototypes are made, the re-design would become time-consuming and very costly.

It can be concluded either of the design philosophies cannot be easily dismissed as they have got complimentary pros & cons. One possible solution is to combine both design flows together. But one big challenge is that both design flows are not directly connected as they usually utilize different types of measurements

The Cardiff large signal waveform measurement facility makes the convergence possible. The fully calibrated measurement system, developed in-house, has the capability to characterize the large signal dynamic behaviour up to a few harmonic frequencies. The resultant large signal waveform measurement data have the great potential to benefit both the measurement-based and modelling-based PA design. The primary motivation of this thesis is to explore the potential of the large signal waveform measurement data and investigate how to utilize the data effectively in the power amplifier design, allowing the PA designer to enjoy the reliability of the measurement-based approach and the fast automated CAD-based design of the model-based approach at the same time.

1.5 Research objectives

The main objective of the research is to develop a new design strategy, which is aimed to overcome the limitations of the measurement-based and modelling-based approaches by providing a common platform for both of them, i.e. the direct utilization of the waveform. In this strategy the nonlinear measurement data should be directly imported into the CAD software once they become available from the measurement system. It enables the designer to access reliable non-linear data in a CAD shortly after the measurement has started and continues accessing the data, which is growing as the measurement progresses. Hence it let the designer start the CAD-based PA design with the accurate measured device characteristics without

waiting for a suitable nonlinear model. Also, at the same time, a nonlinear behavioural model which can be preferably identified directly from the large signal waveform data can be developed. As the data relevant for the PA design and the behavioural modelling are virtually the same the behavioural modelling can take place in the background, without delaying the measurement data collection for the design purpose, using new incoming measurements to improve it and verify it. Once sufficient accuracy has been reached the large signal behavioural model can be released and utilized in the design process. The proposed design flow is shown in figure 1-3. The important advantage of the proposed design strategy is that both the modelling-based and measurement-based approaches are effectively combined. The hybrid design strategy meets the demanding needs from the PA designer for the reliable CAD-based tool for the rapid large signal power amplifier design.

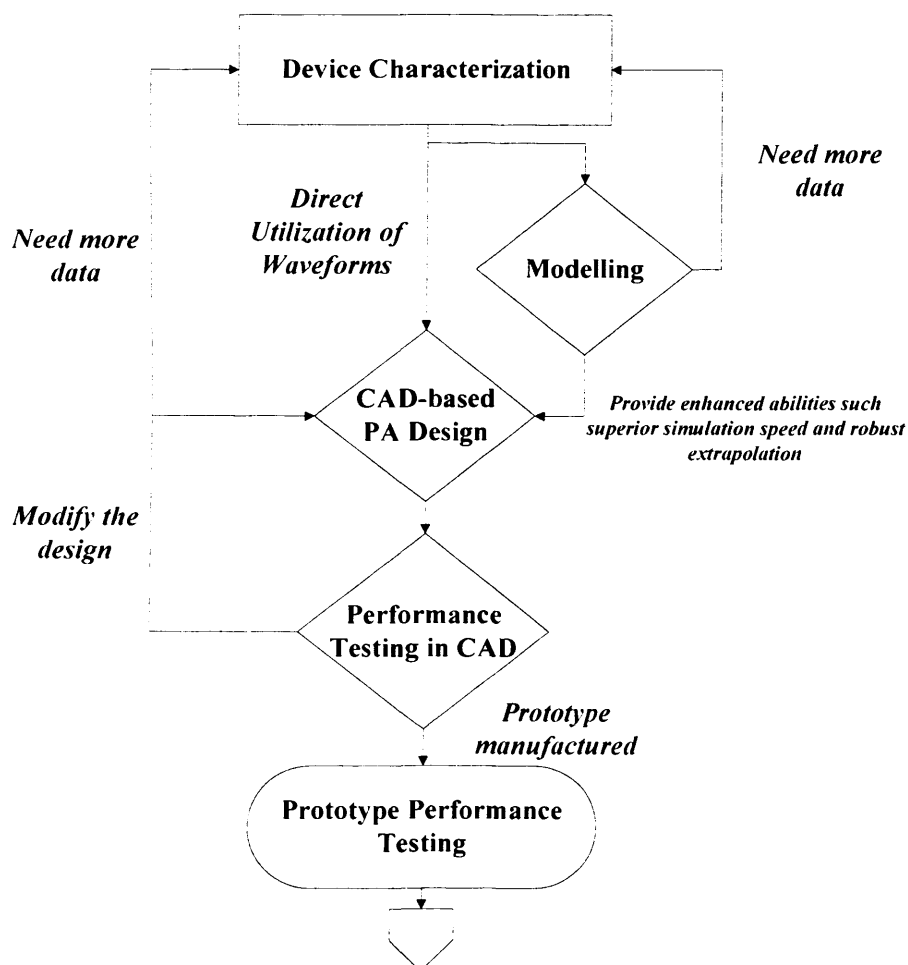


Figure 1-3: The proposed hybrid design flow involving both measurement-based and modelling-based strategies

There are two key points which are essential for realisation of the proposed hybrid design strategy, as marked bold in figure 1-3. One is the availability of the method for the direct utilization of the waveform data in a CAD. The other is the availability of a robust nonlinear model that should be able to be extracted using the large signal waveform data directly.

The direct utilization of the waveform data in the CAD tool has enormous potential and can benefit both the design and modelling. The direct utilization guarantees the rapidity of the availability of the reliable nonlinear waveform measurement data in a CAD. The rapidly available nonlinear data can be directly used for the CAD-based power amplifier design. The data can also be used for the verification of a nonlinear model through the direct comparison of the outputs of the model with the measured data. This could be performed for the modes of operation for which the model has been developed as well for new modes of operation for which the model hasn't been developed (e.g. the switched mode amplifier which operates at the strongly compressed power level) to explore its potential. Moreover, the nonlinear data itself can be utilized to develop quality large signal models.

A robust nonlinear model that is able to be extracted from the large signal dynamic waveform data, has the potential to greatly narrow down the discrepancy usually spotted between the measurement and simulation based on the traditional device characterization techniques. Traditionally, a nonlinear model is usually extracted by making use of the static DC – IV measurement in the combination with the small signal s parameter measurement. However, during a power amplifier design, the extracted nonlinear model is usually required to operate in the large signal dynamic condition. The difference of the device operating conditions between the model extraction and the simulation can obviously compromise the accuracy of the nonlinear model. A nonlinear model extracted from the large signal dynamic measurement data would certainly minimize this problem and has the potential to provide much more accurate simulation for the designer.

The research work aiming to realise the above two key points forms the backbone of this thesis.

1.6 Thesis structure

This thesis is structured into five major sections. Chapter two gives a review of recently published research literatures related to the power amplifier device characterization and large signal modelling techniques. Regarding the device characterization techniques, various types of approaches are covered and emphasis is given on the nonlinear large signal dynamic measurement technologies. In terms of the large signal modelling techniques, the nonlinear behavioural modelling is highlighted and some of the state of the art approaches are discussed. Chapter three presents the Cardiff large signal waveform measurement system in much more details since it represents the experimental foundation of the research work presented in this thesis. The principle of operation and the structure of the measurement system are presented. Then the typical device characterization procedure using this nonlinear measurement system is demonstrated. Chapter four to six focus on how to effectively utilize the nonlinear measurement data obtained from the large signal waveform measurement system for the rapid power amplifier design. In chapter four, a simple yet accurate approach for the direct utilization of the measurement data in a CAD simulator is presented. The approach implementation in the Agilent ADS is detailed. The applicability of the approach is demonstrated experimentally. Moreover, the approach's robustness is comprehensively investigated on many different types of transistor devices. Chapter five presents a novel large signal behavioural modelling approach which can be conveniently extracted from the large signal waveform measurement data. It can be considered as an alternative way to utilize the nonlinear measurement data in a CAD. The theoretical derivation of the model, identification of the model parameters, and implementation of the model in a CAD are described in details. Besides, it is seen that the extraction of the behavioural

model can greatly benefit from the direct nonlinear data utilization approach introduced in chapter four, as the direct data utilization approach can serve as a robust and reliable “measurement data generator ” in a CAD. In chapter six, a comparison study between the direct nonlinear measurement data utilization and the large signal behavioural modelling demonstrates both approaches have complimentary capabilities. It’s shown that the hybrid PA design philosophy combining both the measurement-based and modelling-based design philosophies is fully feasible by having both the direct waveform utilization and the modelling approaches in hands. The PA designer can benefit from this combination of both design philosophies. The final chapter concludes the research work presented in this thesis, with a discussion of future works and possible directions for future research.

Chapter 2 Power Amplifier Transistor Devices: Characterization and Modelling

2.1. Nonlinearity in RF power amplifiers

The ultimate goal of designing a power amplifier is to amplify a RF signal undistorted – the output signal should be just a magnified version of the input signal. A constant delay is acceptable. In other words, the linearity is one of the most important characteristics of a power amplifier. Unfortunately, any active transistor device suffers from the nonlinearity in a realistic operation. The nonlinearity may distort the signal during amplification and consequently the communication system may fail if the distortion is too severe and the information contained in transmission data cannot be traceable. It's therefore necessary to study the nonlinearity of a power transistor and collect sufficient information to develop some solution to deal with it afterwards.

The primary nonlinearity exhibited by a power transistor device is caused by the leakage of energy from the fundamental to the other frequencies. Please note the nonlinearity caused by the frequency bandwidth limited behaviour of RF devices is not the topic addressed in this work. In the frequency domain, the leakage brings up new spectrum components in the output of a transistor device at the frequencies which do not exist in the input signal. If the input signal contains only single frequency component, the output of a nonlinear device would spread onto harmonics only. If the input signal contains more than one spectral component, for example modulated signals, the output spectrum would become distorted as the inter-modulation components are generated as well due to the mixing between the input tones. The harmonic or inter-modulation components would co-exist with the fundamental desired signal at the output. In the time domain the co-existence would exhibit itself as a distorted waveform. Figure 2-1 compares an input “clean”

waveform and an output distorted waveform.

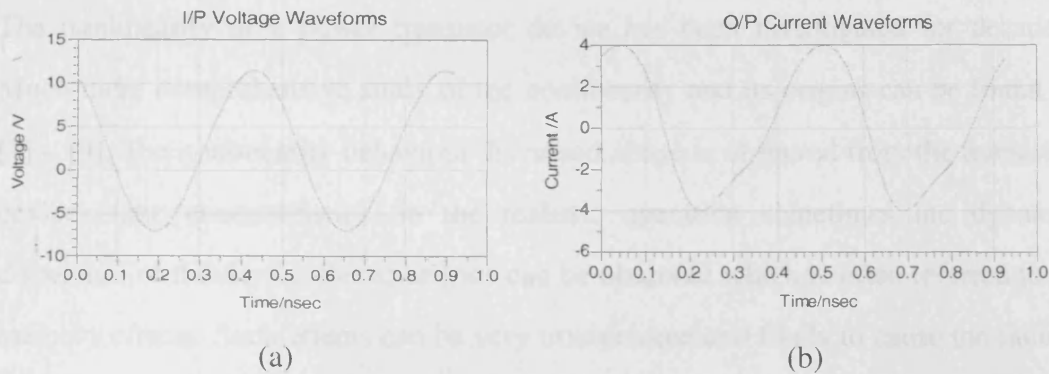


Figure 2-1: (a) and (b) show the input voltage and output current waveforms.

In the power amplifier design a transistor device usually needs to be driven as hard as possible in order to maximize the efficiency. As the device is driven harder and harder the capacity would approach its limit and eventually stops generating more RF power at the fundamental. Figure 2-2 shows a typical power characteristic of a transistor device. It can be seen that as the input power is increasing, the output power would eventually saturate and cease increasing. In the time domain, clipping effect can be usually observed as the result of saturation which causes lots of harmonic contents in the output signal. It can be found that when the device is driven into the compression the nonlinearity demonstrated by the device is much more severe than the nonlinearity shown before.

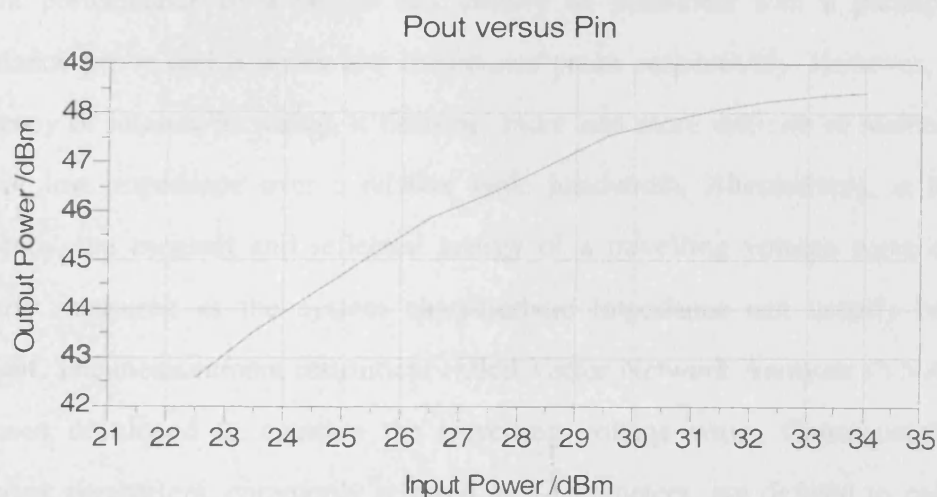


Figure 2-2: The typical transfer characteristic of power transistor devices.

The nonlinearity of a power transistor device has been investigated for decades. Much more comprehensive study of the nonlinearity and its origins can be found in [7] – [9]. The nonlinearity behaviour discussed above is observed from the transistor device static characteristics. In the realistic operation sometimes the dynamic dispersion of the device characteristics can be observed which is often referred to as memory effects. Such effects can be very troublesome and likely to cause the failure of the PA linearization technique. Discussion and characterization of memory effects are out of the scope of this thesis. In literature some interesting studies about memory effects can be found in [10] – [12]. Though memory effects were not taken into account in this work directly, in future work section discussions are made on how to extend this research work to address memory effects.

2.2. Large signal power device characterization

The reliable device characterization is essential for gaining an insight into a device, enabling the accurate device modelling and power amplifier design. New device characterization techniques emerge with the advance of the measurement instrumentation development, mainly driven by the higher and higher requirement from the power amplifier design. At the low frequencies or DC, the voltage and current performance of a device can usually be measured with a parallel high impedance probe and a series low impedance probe respectively. However, as the frequency of interest increases, it becomes more and more difficult to maintain the high or low impedance over a relative wide bandwidth. Alternatively, at the RF frequency, the incident and reflected energy of a travelling voltage wave can be properly measured as the system characteristic impedance can usually be kept constant. The measurement instrument called Vector Network Analyzer (VNA) [13] has been developed to measure the travelling voltage wave. Consequently, the scattering parameters, commonly referred as s-parameters, are defined to calculate the ratios of the reflected energy wave to the incident energy wave. The ratio is

expressed as a complex number which contains both magnitude and phase information.

The s-parameter measurement via a VNA has been around for quite a long time and achieved a great success. The s-parameter measurement has become a standard device characterization technique for describing the performance of any RF linear component. However, when it comes to the large signal condition, s-parameters start to show some limitation. The usefulness of s-parameters in the large signal condition is seriously limited because the s-parameter measurement is fundamentally a linear device characterization technique. As the drive power increases, the linear approximation no longer holds as a transistor device starts to demonstrate obvious nonlinearity. Significant amount of power injected from the input at the stimulus frequency is transferred to the harmonic frequencies at the output. In this case, the complete characterization of a DUT would involve the measurement at the fundamental as well as harmonics.

Traditionally the nonlinear effects found in the large signal operation are measured with a spectrum analyzer. A spectrum analyzer is a scalar instrument which has a broad bandwidth and therefore is able to measure the magnitudes of the fundamental and harmonics at the same time. It can be useful for both measuring the spectrum over the bandwidth of interest and detecting the undesired spurious frequency components. Another simple instrument which is normally equipped in many RF labs is a Power Meter. Such a type of instrument can be used to capture the broadband power characteristics of a device including the fundamental and a few harmonics.

Apart from the dynamic power characterization into a fixed load, there is another important device characterization technique in the design of a power amplifier which is often termed the load-pull measurement [14]. This technique investigates the variation of the device performance with response to the different load reflection

coefficient. It provides a method of optimizing the important figures of merits such as the output power, gain and efficiency by varying the load reflection coefficient to generate sets of contours. The corresponding measurement system can be developed by expanding the capability of the traditional VNA, together with the active or passive load reflection coefficient implementation. A good overview of the active and passive load-pull measurement techniques and some examples of industry realisation is given in [15]. The passive implementation [16] utilizes a mechanical tuner to adjust the output load reflection coefficient. The passive termination gains the advantage of reducing the risk of oscillation when conducting a load-pull measurement on the high gain device. But the maximum load-pull range achievable by the passive termination is limited due to the loss in the path between the device measurement plane and the load. The limitation becomes more serious for the high power transistor device, because the optimum impedance for the high power transistor device is usually very small. Alternatively, active implementation of the load termination can effectively eliminate the above power limitation problem. The active method [17] utilizes an active source to inject energy into the output of a DUT. Given sufficient amount of power, any load reflection coefficient can be achieved with the active method. However, the amount of power required for the load-pull can be huge for measuring the high power device, since the high mismatch at the optimum impedance point of the high power transistor device demands significant amount of power injection which can be much higher than the actual power absorbed by the effective load attached to the DUT. Recently a promising method has been proposed to lower the demand for the high load-pull power which integrates a broadband impedance transformer into the active load-pull measurement system [18]. The integration of the impedance transformer lowers the system characteristic impedance from default 50 ohm to a value much closer to the optimum impedance point and hence reduces the degree of mismatch encountered by the active load. As a result, the power required for the load-pull measurement can be dramatically reduced.

A traditional load-pull measurement system focuses on tuning the load reflection coefficient at the fundamental frequency only. The higher order harmonic is usually terminated into constant impedance such as 50 ohm. The recent advance related to the high efficiency power amplifier design, such class E and class F power amplifier, bring the need for the harmonic tuning in order to engineer the shape of the waveforms to maximize the power amplifier efficiency. The harmonic load-pull measurement system is then developed to address the need. Instead of terminating the harmonics into fixed load impedance, active or passive load reflection coefficient tuners are used at the harmonics to enable the investigation of the variation of the device performance with regard to the different harmonic load reflection coefficient. But the harmonic load-pull measurement system is disadvantaged by its relatively slow measurement speed. The single load-pull measurement commonly involves several iterations to achieve the convergence at the target load reflection coefficient. Load-pulling the DUT simultaneously at several harmonics would result in high number of iterations. This problem can be eased by utilizing the more advanced measurement system configuration capable of reducing the number of load-pull iterations dramatically [19].

With the load-pull system the power characteristics of a DUT can be accurately obtained. But the phase characteristics are still unknown and need to be captured. The recent development of Nonlinear Vector Network Analyzer (NVNA) makes it possible to capture both the magnitude and phase information of a nonlinear transistor device at the RF frequency [20] [21]. At the RF frequency, the ordinary oscilloscope instrumentation, which is used for capturing the complete information of a DUT at the low frequency is not applicable as the limited sampling rate of the analogue to digital converter in the oscilloscope system seriously reduces the dynamic range of the instrument at the high frequency. A NVNA can overcome the limitation of an ordinary oscilloscope system and conduct the reliable measurement at the high frequency with a sufficient dynamic range. A NVNA usually consists of four channels which are used to acquire complete and accurate information about the

incident and reflective travelling waves at both the input and output ports of a DUT. This instrument can be used to deal with various excitation signals such as CW and modulated stimuli. An obvious difference between a NVNA and a traditional VNA is that a NVNA measures absolute quantities while a VNA measures relative ratios.

The ultimate solution of characterizing a DUT is to measure the complete information of the DUT – the voltage and current waveforms at the device ports. With the measured voltage and current waveforms the typical nonlinear behaviour picture of the DUT on given stimulus condition can be obtained. Since a NVNA measures the absolute travelling wave at the device ports it is possible to reconstruct the voltage and current waveforms by making use of the underline connection between the travelling wave and the voltage and current waveform. By taking advantages of a NVNA, the new large signal measurement system has been developed which enables the direct voltage and current waveform characterization of a DUT under the realistic RF operating condition [22] [23]. This type of voltage and current waveform characterization system can benefit both the RF circuit design and nonlinear device modelling as it's able to provide more information of the DUT than any measurement system can ever do before. Its potential is currently being explored in many directions [24][25].

2.3. Power transistor device modelling

In the electronics industry, the design of the modern telecommunication system is a hierarchical process which normally involves lots of simulations at several different levels. The design relies on the reliable modelling of components within the system to run the accurate simulation and consequently achieve a first-pass design. Particularly, since a power amplifier is always a key component in any telecommunication system, the accurate modelling of the power transistor device is certainly required, which is an unavoidable step towards a successful power amplifier circuit design.

The modelling community has been very active for decades. As a result, many types of models have been developed. There are many ways to categorize a model. Based on how a DUT is analyzed, models can be divided into two categories. Those which are developed by analyzing the physical operation of a semiconductor power transistor device are usually referred to as physical-based models. On the other hand, those which treat a DUT as a black box and relate the output of a power transistor device to its stimuli through purely a mathematical function are usually referred to as behavioural models.

In order to develop a physical-based model, comprehensive analysis on the semiconductor characteristics of the target device technology is always necessary. Following the physical analysis, there are generally two possible ways to formulate a model. One is often termed as the “bottom up” approach, i.e. the physical structure and the property of a semi-conductance device, such as the fabrication geometry, semiconductor material properties, doping, contacts, etc, are considered and directly used to obtain a model. This kind of modelling strategy generally requires no electrical characterization to obtain a model. With the analysis of the device physical data, a model is formulated with the physical-based equation. But the problem of the bottom-up approach is the computational complexity involved with developing a model. The computational cost required for an accurate representation of the actual physical behaviour of a semi-conductance device can be enormous. Consequently, in order to cut the computational cost a number of approximations and simplifications are necessary which may seriously limit the accuracy of a model in the RF circuit design. A complete analysis of the physical modelling is out of scope of this thesis and can be found in [26]. More recent developments can be reviewed in [27].

The other type of physical-based methods is to extract a lumped equivalent electronic circuit to represent the physical behaviour of a transistor device. The complexity of the equivalent circuit needs to be carefully chosen in order to properly

trade off the accuracy of the model. Often only some aspects of the device physical operation are taken into account while the others are ignored to maintain the complexity at an accepted level, with reference to a particular application of interest. Once the equivalent circuit topology is determined, the special excitation design is made in order to extract the model. The model parameter optimization is often required to tweak the values of the circuit components to achieve the best fit to the measurement.

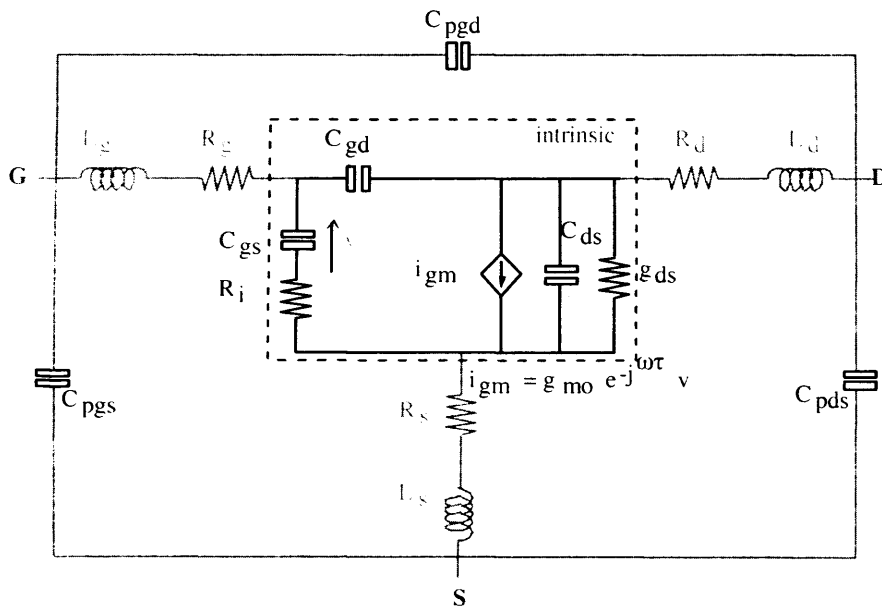


Figure 2-3: Basic equivalent circuit for the GaAs MESFET [28]

Figure 2-3 shows a very basic equivalent circuit model for a packaged GaAs MESFET device. The intrinsic box represents the model for the active device sitting inside the package. Clearly the active device is modelled as a voltage controlled current source which is surrounded by a bunch of capacitances and inductances. This typical intrinsic model is a simplification of the much more complex underlying physical behaviour of a semi-conductance device. Lumped circuit description is merely an approximation of the realistic distributed behaviour of the device and is subject to vary from device technology to device technology. The equivalent circuit based modelling is advancing very quickly recently. A number of publications regarding new model topology, model extraction and characterization have been

resulted throughout the years. More recent advances can be found in [28] [29] and [30].

The equivalent circuit based model generally offers the advantage of being computationally efficient in the simulation. Besides, since the equivalent circuit based model is usually tailored to fit certain device physical characteristics, a good accuracy when simulating the corresponding physical characteristics can always be obtained. Furthermore, such model is robust for extrapolation. The simulation outside the extraction frequency band or power range generally gives satisfactory results.

However, regardless of the concrete modelling approach, the physical-based model also suffers from some fundamental problems which undermine the applicability of the physical-based modelling approach. Obtaining the physical knowledge of the target transistor device necessary for the modelling requires in-depth information about the transistor structure, including the device dimension and detailed fabrication process, which the device manufacturer may be very reluctant to release. Lacking detailed information about the device physical operation characteristics, which usually happens when new type of device technology appears, can lead to inaccurate extraction of the equivalent circuit and hence the extraction process may involve several iterations that would increase both the cost and time. Also, the analysis of the physical operation of the transistor device can be lengthy even if sufficient information can be obtained from the device manufacturer. Besides, the extraction and optimization of the model parameters are not as easy as they sound. As the required model accuracy increases, the extraction of the model parameters can quickly become prohibitive. Additionally, special experiments for both extraction and verification purposes are normally needed which incur extra complexity.

Alternatively a behavioural model is often identified without reference to the physical operation of a transistor device. The behavioural modelling is frequently

referred to as “black box modelling” because the device being modelled is considered as a black box. The behavioural model is developed by seeking the mathematical relationship between the input and output of a DUT from the observed data. A good overview and more recent applications of the behavioural modelling approach can be found in [31] and [32].

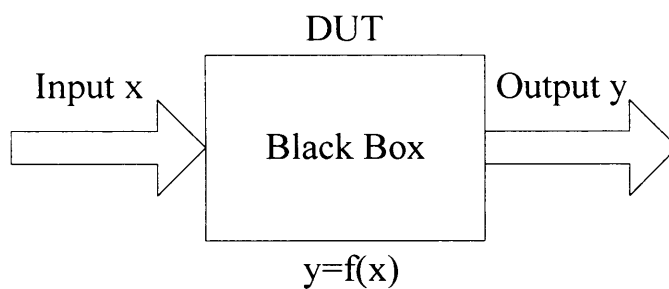
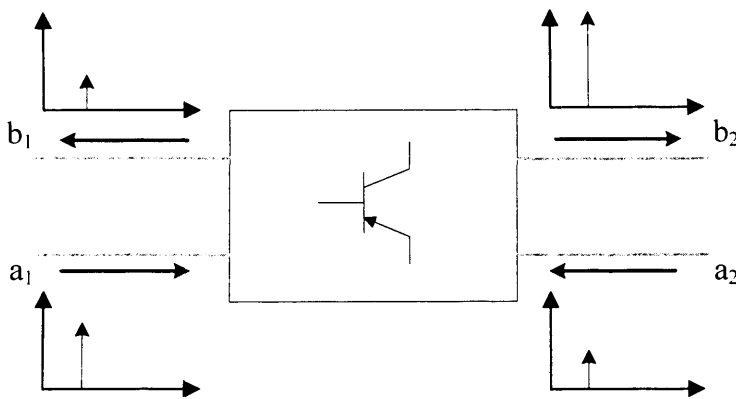


Figure 2-4: The basic concept of behavioural modelling

The behavioural relationship can be defined in the frequency domain, time domain or envelope domain. Depending on the type of data available, different model extraction strategy can be designed. The optimization of the model parameters may also be needed which is usually integrated as part of the extraction process. Regardless of what the particular extraction strategy is, the extraction of behavioural models can be generally thought of as a form of multidimensional curve fitting.

The behavioural modelling has gained a great success in the small signal world through s-parameters. In a small signal case, a transistor device can be approximated as a linear device where superposition and decomposition principles hold. Small signal s-parameters can be considered as the linear behavioural modelling of a N-port network, which have been widely used in the device characterization and the small signal power amplifier design. Figure 2-5 shows a typical scenario where a transistor device is characterized within a two-port network structure. The travelling waves a and b are defined to represent the energy injected into and reflected from the two-port network. In a small signal operation, no harmonics are generated when the

DUT is excited with a single tone signal. The travelling waves a and b can be related by s-parameters as shown in Figure 2-5. S-parameters are usually extracted when the DUT is matched to 50 ohm at both the source and load sides. They remain constant even when the load and source impedance are varied around the Smith Chart since the DUT is assured to be linear.



$$b_1 = S_{11}a_1 + S_{12}a_2$$

$$b_2 = S_{21}a_1 + S_{22}a_2$$

Figure 2-5: Linear characterization of typical transistor behaviour

Things get complicated in the large signal operation as the linear properties such as superposition don't hold any longer. Fundamental signals get distorted while new components are generated at the harmonics. Figure 2-6 gives a typical scenario where a transistor device is excited by a large signal stimulus. Energy leaks from the fundamental frequency to the harmonics. Significant harmonic components are generated which in return interact with and further distort the fundamental component.

It can be seen that the simple linear mathematical relationship can no longer sufficiently characterize the two port network. The accurate behavioural modelling of the nonlinearity requires more sophisticated mathematical tools than s-parameters. The behavioural models are usually distinguished from one another by the

mathematical tool adopted for relating the output of a transistor device to the stimuli.

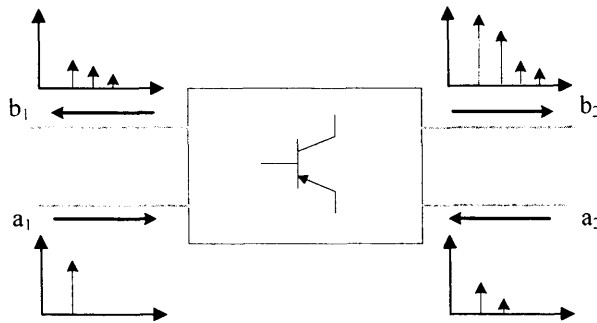


Figure 2-6: The typical scenario of large signal characterization. It's assumed that the harmonics at input are matched.

The Volterra method is one of most popular mathematical tool for behaviourally modelling a nonlinear measurement system, provided the nonlinearity is not too strong. The Volterra function analysis was firstly developed by Italian mathematician Vito Volterra and introduced into the system engineering by Norbert Wiener [33]. Volterra functional analysis can be considered as an extension of the concept of impulse response derived for linear systems to nonlinear systems. The Volterra theory uses so-called “kernels”, which are essentially nonlinear transfer functions, to describe nonlinear systems. The use of kernels makes it possible to account for a number of phenomena that are often encountered in a nonlinear system, such as feedback and non-unilateral behaviour. The nonlinear system modelling based on classic Volterra theory can be found in [34] and [35].

The classic Volterra functional analysis is well supported theoretically but it's got the problem of poor convergence properties which make it difficult to identify kernels effectively. This problem has been addressed in [36]. Hence, modified volterra functional has been proposed in [37] to further address this issue.

However, despite the tedious efforts have been made to improve the practicality of the Volterra method, it's still not widely accepted in the practical power amplifier modelling due to its impractical requirements for the measurement and extraction of

the kernels. Moreover, the order of kernels has to be limited to small number in order to avoid computational over-complexity. It imposes that the nonlinearity of the system has to be weak and thus limit the applicability of the method for system with strong nonlinearity.

Neural networks have recently drawn lots of attentions as an alternative behavioural modelling tool. Neural networks, often referred to as Artificial Neural Networks (ANNs), are originally inspired by the research of the ability of the human brain to learn from observations and to generalize by abstraction. The ANN-based behavioural modelling approach basically utilizes a set of neurons distributed into different layers to characterize the nonlinear systems. These neurons may connect to each other and weight parameters are available to regulate the amount of influence from one neuron to the other. The established neuron network is then trained by measurement data to learn the behaviour of the nonlinear system. A comprehensive review of recent ANN-based modelling activities is given in [38] and [39].

Based on the classic neuron network theory, significant development has lead to several variants of neuron network modelling technologies to handle the dynamic behaviour of a nonlinear system in the time domain, which has been reported in [40] and [41]. Moreover, other variant of the ANN theory such as the Recurrent Neural Network has also been developed specifically for representing the dynamic system behaviour [42].

The advantage of the neural network behavioural modelling is its ability to characterize a nonlinear system which is otherwise too complex or too difficult to be characterized with analytical formulas. But the disadvantage of this method is the lengthy training process of the model. Lots of measurement data need to be collected for training the model which cannot be directly used in the simulation. It makes the utilization of measurement data very inefficient.

Recently a promising nonlinear behavioural modelling strategy has been proposed by Jan Verspecht [43]. The strategy recognizes the universality of the s-parameters and makes the extension of the concept of s-parameters to the large signal operation. The so-called scattering function is used to relate the input and output of a nonlinear system. The approach shows a great level of accuracy when dealing with large signal problems while maintaining a low level of identification and simulation complexity. Since the approach shows so many interesting features, it will be analyzed more thoroughly and discussed in details later in this thesis.

Different types of behavioural models share some common pros and cons. One of the advantages is that there is no need to worry about the internal circuitry when extracting a behavioural model. It makes the behavioural model very attractive when the internal physical process may not be known completely or too difficult to characterize accurately. This advantage also benefits the behavioural modelling for the packaged device. For the packaged device the measurement plane is usually separated from the device voltage-control current source plane by the parasitic network and/or pre-matching network. The physical-based model often needs the data obtained at the device current source plane for the model extraction. Hence de-embedding of the parasitic network and/or pre-matching network are necessary. If it is not possible to obtain the package information from the manufacturer, additional measurement and extraction procedure are required to identify the parasitic network, which can be error-prone and time-consuming. Conversely the behavioural model can be identified using the data collected at the external measurement plane without worrying about the parasitic de-embedding. The second advantage is that the behavioural modelling is generally technology independent. No matter what type of device under test is, the same behavioural approach can always be applied. The third advantage of the behavioural modelling is that it's good for intellectual property (IP) protection. Generally speaking the behavioural model shouldn't give away any confidential information about the how the transistor device is made.

The behavioural model has also got some disadvantages. One disadvantage is the poor extrapolation ability. The behavioural model utilizes the mathematical analytical or non-analytical curve fitting function to relate the output of the system to the input based on the observed data at the extraction ports. The mathematical curve fitting function generally doesn't provide any information on the physical operation of a power transistor device. The model trained by the observed data can yield good accuracy within the range that the training data effectively cover. However, the behavioural model usually loses the accuracy very quickly when doing extrapolation, i.e. the simulation outside the covered range. That is because the model coefficients are optimized only for the covered range. Little accurate information about how the transistor device behaves outside the range can be inferred from the observed data. To obtain such information it's normally necessary to know the physical operation of the device and the characteristics of the package which the device is in. Therefore, the poor extrapolation is a fundamental problem for the behavioural modelling due to the lack of considering the internal physics within the transistor device package.

Another disadvantage of the behavioural modelling is that the behavioural model is developed from the pure mathematical derivation which doesn't represent any physical operation of a DUT. This feature makes the behavioural model appropriate for the RF circuit or system designer but impedes the usefulness of the model for the device technology developer.

Chapter 3 RF Large Signal Waveform Measurements

3.1. Cardiff RF large signal waveform measurement system

3.1.1. Cardiff waveform measurement system basics

Based on NVNA instrument capabilities, a RF large signal waveform measurement system has been developed in Cardiff University by Tasker et. al.[44] The waveform measurement system is capable of measuring DUT excited by single tone signals with a full characterization of input and output voltage and current waveforms. Investigations on how to effectively utilize the resultant waveform measurement data are the focus of this thesis.

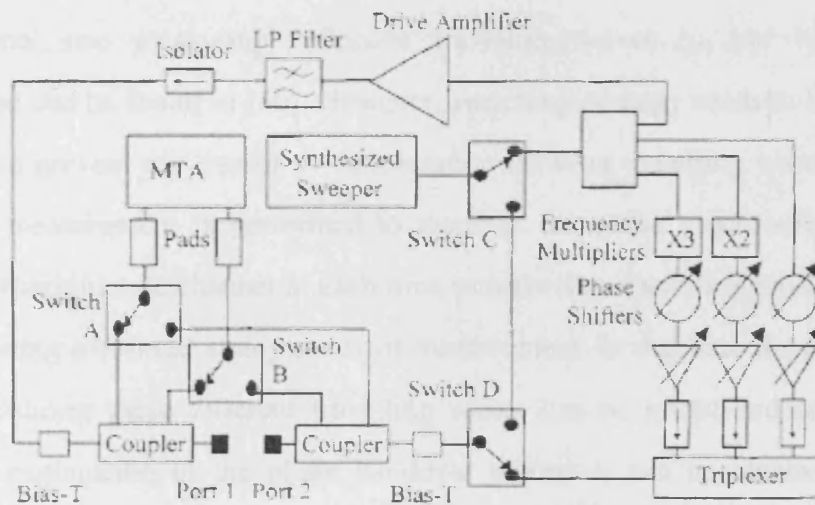


Figure 3-1: Schematic of implementation of Cardiff waveform measurement system [48]. DUT is inserted between Port 1 and Port 2.

Figure 3-1 gives the schematic of implementation of the measurement system. The key component of the waveform measurement system is the Hewlett Packard Microwave Transition Analyzer (MTA) [45] – [47]. The MTA is a dual channel sampling scope with a frequency range from DC up to 40 GHz. The MTA can operate in a common time base mode as well as frequency sweep mode. The major

advantage of MTA over traditional oscilloscope instruments is that its sampling rate is synthesized from a common master clock, which allows this dual channel dual mode MTA to perform phase relation analysis between two channels. Moreover, the dual mode capability of MTA enables the frequency domain s-parameter measurement and time domain waveform measurement. The frequency domain s-parameter measurement is quite useful as it allows the calculation of the error parameters which are necessary to calibrate the voltage and current waveforms measured in time domain mode. Detailed information about how MTA operates and its applicability in large signal measurement can be found in [48] and [49].

In order to capture all these four travelling wave quantities, two high frequency coaxial switches are grouped together, allowing the dual channel MTA works as a four channel receiver with channel one measuring incident travelling waves a_1 and a_2 and channel two measuring reflective travelling waves b_1 and b_2 . Detailed explanation can be found in [50]. However, switching strategy needs to be carefully designed to prevent the loss of synchronisation between travelling waves. A phase handover measurement is performed to maintain the phase synchronisation. This requires switching one channel at each time to make sure there is always a common quantity being measured at any status of measurement. In that case the phase offset when measuring these different travelling waves can be traced and corrected. A complete explanation of the phase handover procedure can be obtained in [50]. These travelling wave energies are sampled by using two directional couplers at the device ports. Those couplers have 35dB attenuation to protect the MTA from overdriven when measuring high power transistor devices.

DC bias tees are used to bias the DUT and couple the DC path and RF path together. Since the direct couplers prevent the DC signals from being measured by MTA, computer controllable DC sources are used in the system. With proper interface programs, the DC voltage and current values are read directly from DC sources and added to the measured RF waveforms in measurement software. HP 6674A DC

sources which can provide voltage from 0 to 60 V and current from 0 to 35 A are used in the system to bias the RF circuit under test. Sometimes a small series resistor is inserted into the dc path at the side of the device input port to prevent high gain devices from oscillating.

Load-pull facilities are also integrated into the measurement system, enabling waveform characterization at various load impedance points. As observed in Figure 3-1, the active technique is adopted to realise the load-pull system by using the Agilent ESG series signal generator to inject energy from the output of DUT. The ESG is phase referenced to the fundamental input source of MTA using a 10MHz reference signal to maintain phase stability.

The reflection coefficients required for load-pull measurement are achieved through an iterative convergence process. Load reflection coefficient is defined as the ratio of travelling wave a_2 over b_2 :

$$\Gamma_{LOAD} = \frac{a_2}{b_2} \quad (3-1)$$

In a passive load-pull system, the load reflection coefficient can be quickly determined as the mechanical tuner presents the ratio directly to the output of DUT. By contrast, in an active system like the one shown in figure 3-1, the ESG can only present a_2 to the output of DUT. The variation of load-pull reflection coefficient is done by varying the magnitude and phase of the energy injected by the ESG. Since the travelling wave b_2 is obviously out of control of the ESG, there is no way to determine the value of b_2 beforehand. Therefore, iterative measurements have to be performed by changing the magnitude and phase of energy from the ESG gradually in order to find b_2 , calculating the ratio and eventually converging to the target load reflection coefficient. The number of iterations required for load-pull convergence varies from device to device. Devices tending to demonstrate lots of nonlinearities generally require more iterations and longer measurement time.

The triplexer is used to separate each of harmonics from one another. Fundamental, 2nd harmonic and 3rd harmonic paths are isolated and thus the load impedance attached to each harmonic can be independently controlled. As far as this work is concerned, only fundamental load-pull measurement is of interest. In such a case, the 2nd harmonic and 3rd harmonic are usually terminated into 50 ohm, either actively via ESGs or passively via mechanical terminations. Nevertheless, extension of the system to enable 2nd and 3rd harmonic load-pull capability can be easily done by attaching ESGs at these harmonics.

The complexity of the waveform measurement system necessitates automated measurement capability. Besides, it would be desired to have the measurement system fully controlled by computer to minimize human interfacing – computer-based solution tends to be much more reliable and faster. Automated measurement controlled by computer can be realised by using the General Purpose Information Bus (GPIB) [51], which are available on these Agilent ESGs and MTA as an alternative option to local front panel control. GPIB serves as an interface between these instruments and the controlling computer. It offers a range of commands which allows the computer to communicate with measurement instruments and exchange data when necessary. These coaxial switches don't offer GPIB interface themselves but they can be controlled via GPIB by using Agilent 34970A data acquisition unit fitted with a 20-channel relay card.

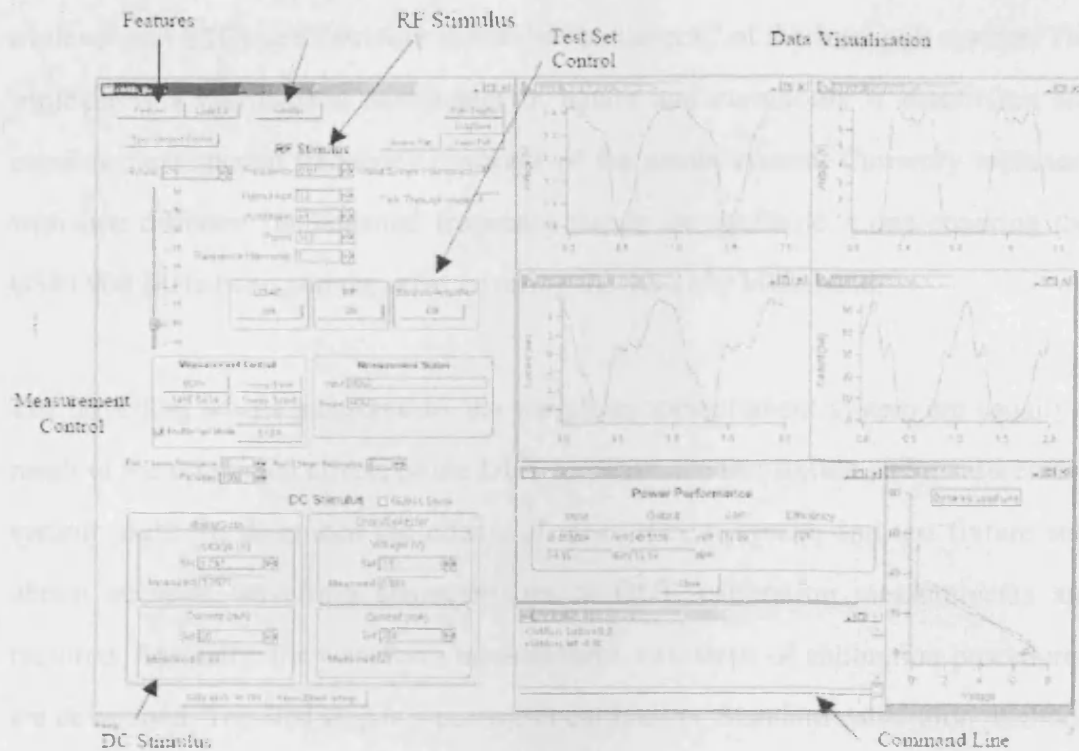


Figure 3-2: snapshot of measurement system interface [50].

The computer user interface, which offers general measurement control functionalities, monitors the measurement status, and displays results, is developed by using Igor Pro software [52]. Igor Pro software offers a variety of functions such as Fourier Transform, nonlinear interpolation, contour plotting and so on. Besides, it provides GPIB control modules which interface quite nicely with the software. Figure 3-2 gives a snapshot of the user interface of the waveform measurement system. By using this interface, the load-pull measurement and related power sweeps at each load impedance point can be done automatically without user intervention.

The possible frequency range of the whole measurement system is limited by the valid frequency range of individual components used within the system. If no load-pull measurement is required, the system's frequency limit is determined by the MTA. If load-pull measurement is required, the frequency ranges of triplexer and ESGs should also be taken into consideration. The MTA has the capability to measure the frequency range up to 40 GHz which is generally much wider than the

triplexer and ESGs and therefore is not the “bottleneck” of the load-pull system. The triplexer is a narrowband component in nature and commonly it determines the possible fundamental frequency coverage of the whole system. Currently triplexers with two different fundamental frequency bands are available – one covering the GSM 900 MHz band and the other covering the 3G 2150 MHz band.

The travelling waves measured by the waveform measurement system are usually a result of the combined effects of the DUT response, the test fixture and measurement system itself. To de-embed the effects of measurement system and test fixture and obtain accurate waveform characteristics of DUT, calibration measurements are required. Basically, for waveform measurement, two steps of calibration procedures are developed. The first step is s-parameter calibration. Standard calibration methods such as TRL (Thru, Reflect, Line) Cal [53] is used for calibrating the on-fixture measurement, which happens in most of characterization activities related to this work.

However, the traditional s-parameter calibration method has the inherent limitation that S-parameters are only ratios rather than absolute quantities. A successful s-parameter calibration could ensure reliable ratio measurement but could do no calibration to current and voltage waveforms as they are absolute quantities. Therefore, once the s-parameter calibration is done, a second calibration step is required to calculate the absolute power and phase of the travelling waves. It can be achieved by calculating a scaling factor for these s-parameters, by directly measuring the travelling waves at the input port of DUT with the MTA [50]. Both the s-parameter calibration and absolute calibration procedures are implemented as a stand-alone calibration software using Igor Pro.

Repeatability test of the measurement system is required once calibrations are finished, to verify the calibrated measurement system can give repeatable results. Repeatability test of the s-parameter calibration can be achieved by performing

s-parameters on the standard cal kit used in the cal. For a good cal, it's expected to have s-parameter characteristics of the cal kit very close to ideal behaviour from calibrated measurements. For example, after s-parameter calibration, s-parameter measured on the standard Thru, which is used in TRL calibration, should have the following values:

$$S_{11} = 0; S_{12} = 1; S_{21} = 1; S_{22} = 0 \quad (3-2)$$

S_{11} and S_{22} are indicators of how much percent of energy is reflected from DUT. S_{12} and S_{21} are indicators of how much percent of energy is passing through the DUT. The group of s-parameter values shown above indicates all the energy passes through the DUT without any loss or reflection. Such characteristics are expected from ideal Thru. If the measurement system is calibrated perfectly, the measurement is supposed to show that the Thru has ideal behaviour with the s-parameter values shown above. In practice, unfortunately, due to the imperfection of the cal kit and uncontrollable variation of some environment parameters such as system thermal status and repeatability issue, those ideal values are impossible to achieve. Practical error tolerance is usually used to justify whether a good repeatability is achieved.

The repeatability of the measurement system after absolute calibration can be checked by measuring the RF power at input and output of the Thru in the cal kit via measurement software and then cross-checking with power meter readings. If perfect absolute calibration is achieved, the input and output power read from measurement software should be equal. Besides, the measurement software reading should match the power meter reading, which represents actual power flowing in the measurement system. In practice, a 0.15 dB error tolerance is usually acceptable.

The basic structure and operation principles of the Cardiff waveform measurement system are covered in this section. It's worth pointing out the system structure and functionalities merely represent the most-commonly-conducted waveform measurement capability in the centre for high frequency engineering, Cardiff

University. The resultant measurement data is also the primary data source for the work presented later in this thesis. Latest development of waveform measurement system enabling more interesting and powerful features, such as multi-tone/modulated signal characterization and real-time load-pull measurement, which is out of scope of this thesis, is well covered in literatures [54] – [56].

3.1.2. High power waveform measurement issues

Parts of efforts made in this work focus on utilization of waveform data obtained on high power transistor devices, which are normally appropriate for wireless network infrastructure applications. In order to obtain the sufficient waveform data for this research work, lots of measurements were conducted. To obtain accurate and reliable data source, this work also addressed some associated high power measurement issues. These issues are discussed in this subsection.

The single tone waveform measurement system discussed previously has 60 dB dynamic range and has the potential of characterizing transistor devices ranging from mW to a hundred Watts. The limiting factor that may prevent the measurement system from being used to characterize high power transistor devices is the maximum available power that the active load-pull source can provide. High power transistor devices usually possess very small optimum impedance for maximum output power or gain, etc. For example, according to the data sheet, the optimum impedance of a typical 100W LDMOS device is around 2 ohm. To characterize high power transistor devices around the small optimum impedance, high mismatch between the optimum impedance and system characteristic impedance is expected. High mismatch implies the power reflected from the load to DUT is massive and sometimes the reflected power is even higher than the actual power delivered into the load. Specifically, in the Cardiff waveform measurement system, the active load-pull capability requires massive power which can easily exceed the maximum power the Agilent ESG series signal generator can provide. So power amplifiers

need to be inserted between the ESG and DUT to enhance the output power capacity of the active load. But unfortunately, even doing so, it's still difficult to cope with the power requirement for load-pulling high power transistor devices. Besides, as most of the massive RF power injected from the active load would convert to heat, it brings the problems of heat dissipation. Higher power also put higher requirement of VSWR reliability on each of the components in the system. Recently a solution of inserting broadband impedance transformer into the measurement system, proposed by Aboush et. al [18], successfully solves the problem. By using the broadband impedance transformer, the system characteristic impedance can be brought down from 50 ohm to a few ohms at the measurement plane. The level of mismatch can be dramatically reduced and so is the required load-pull power. Figure 3-3 shows an example of two fabricated broadband impedance transformers used for 100w device waveform measurement. The transformer transfers the system characteristic impedance from 50 ohm to 7.15 ohm. One is used at the input side of DUT and the other is used at the output side of DUT.

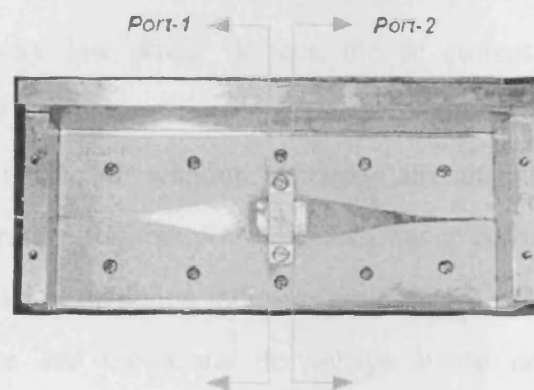


Figure 3-3: Picture of the broadband transformer implemented for high power measurement.

Apart from the power issues, there are also some other issues, such as DC voltage variation and dramatic change of load-pull power, which would effect the feasibility and reliability of high power waveform measurement. The following discussion will address these remaining issues and solutions are proposed in this work to solve them.

In the presented waveform measurement system, the control software reads and sets the DC biasing voltage directly at the DC source through GPIB interface. However, the dc voltage at the DC source is usually not the same as the dc biasing voltage at device ports, since the system components sitting in between (as shown in Figure 3-1) have non-zero resistance. The actual dc voltage obtained at the device port is:

$$V_{actual} = V_{set} - I_{dc} R_{dc} \quad (3-3)$$

As can be seen in above equation, the difference between the voltage at DC source port and voltage actually existing at device port is influenced by both the resistance between these two ports and dc current flowing from dc source to the device. The dc resistance can be influenced by many factors such as the cabling, bias tees and etc. The typical value of the resistance on the dc path of the Cardiff waveform measurement system is around 0.3 ohm. The dc current is mainly determined by the status of the active transistor device under test. The maximum dc current is determined by the device power capability while the actual dc current is determined by a set of stimuli to the device, such as the drive power level and dc biasing voltages. For relatively low power devices, the dc current when the device is operating in normal condition is small. The resultant voltage drop on the dc resistance can be neglected without seriously affecting the accuracy of the measurement. By contrast, for high power devices, the dc current can go as high as 8 A. Simple mathematical calculation indicates more than 2 volts drop would happen on the dc resistance and the actual dc voltage at the device port would be significantly different from the voltage measured at the dc source port. Figure 3-4 and Figure 3-5 records the variation of the actual DC drain voltage at device port. The data were obtained on a 30w LDMOS device when the DC drain voltage was set to 27 V. It can be seen that the dc drain voltage show slight variation when the load reflection coefficient is swept. More significant variation can be spotted when the input power is raised. The observed dc voltage variation is undesired since it influences the device performance. Consequently it makes the measurement data

inconvenient to use in simulation and design since the variation of the DC biasing needs to be taken into account. Please note the DC bias variation demonstrated in Figure 3-4 and Figure 3-5 should be differentiated from the self-biasing effect observed in normal PA operation. The self-biasing effect is caused by the energy flowing on RF path. This effect has been recorded in the measurement data and will be re-produced in simulation. The DC variation discussed here is caused by DC current and the non-zero resistance on the DC path. It's the measurement system introduced error that need to be eliminated.

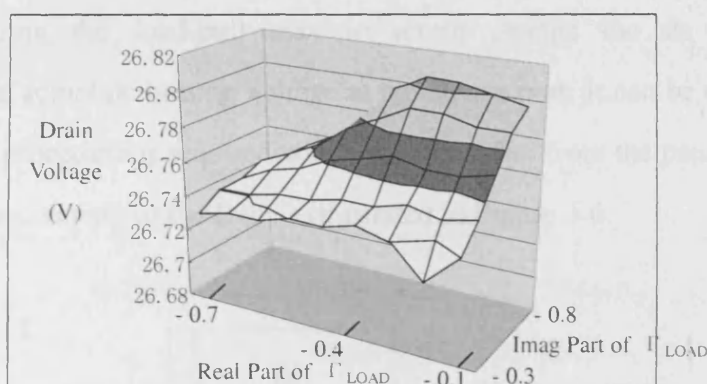


Figure 3-4: The actual variation of DC drain voltage measured on a 30w LDMOS device without any DC compensation scheme when the input power is 27 dBm.

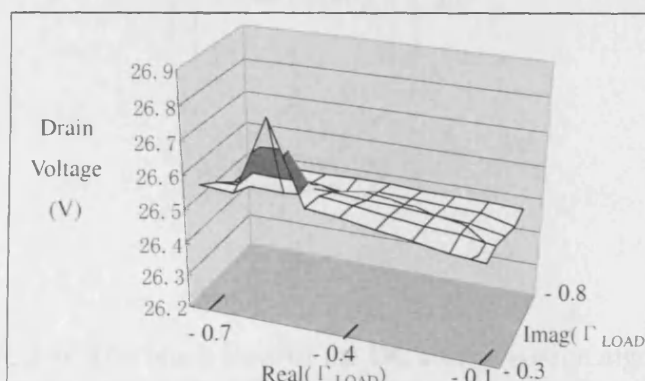


Figure 3-5: The actual variation of DC drain voltage measured on a 30w LDMOS device without any DC compensation scheme when the input power is 34 dBm.

One solution to this problem proposed in this work is to introduce software-type

looping to stabilize the DC voltage when the dc current varies with input power or load impedance. The routine is implemented as part of the measurement software. It's integrated into both the load-pull algorithm and power sweep algorithm to make sure the DUT is always swept under consistent biasing condition. However, one disadvantage of software looping is that it introduces extra measurement time due to the interaction between the DUT and active load. Adjustment of DC drain voltage would immediately alter the status of the DUT and result in changes of the device input and output characteristics. Sometimes such changes may be big enough to distort the load reflection coefficient which has been load-pulled into prescribed position. Re-doing the load-pull may in return change the dc current and consequently the actual dc biasing voltage at the device port. It can be clearly found that an iterative procedure is required which usually suffer from the penalty on speed. The operation mechanism of the loop is illustrated in Figure 3-6.

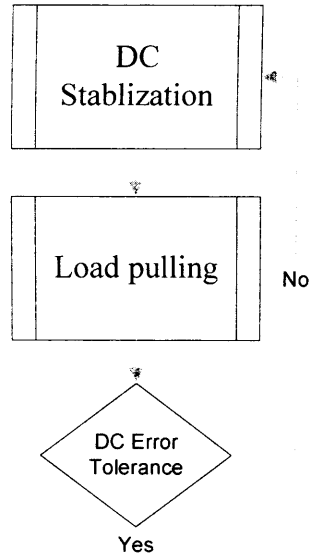


Figure 3-6: The block flow of the DC compensation algorithm.

Another issue of high power measurement relates to the active load-pull source. During load-pull iterations, the power from ESG is adjusted until the desired load reflection coefficient is achieved. Since the characterization of high power devices always involve a quite large power variation, the risk of damaging the DUT during

load-pull convergence becomes higher, since DUT is usually vulnerable to a sudden increase of power injected from output. Therefore, in order to do the high power load-pull measurement safely, a safety control algorithm is introduced into the measurement software. The control algorithm restricts the maximum possible step size of ESG power adjustment during each iterative load-pull convergence measurement. The control algorithm increases the reliability of the measurement system.

3.2. Single tone waveform characterization

By using the large signal waveform measurement system described in the previous section, the essential voltage and current behaviours of the DUT excited by realistic RF signals can be easily obtained. Access to the waveform information, which is exclusively available to users of large signal waveform measurement system, allows RF circuit designers to conveniently investigate the nonlinear performance of devices when the DUT is configured into conditions of interest. Various external environment parameters which may affect the behaviour of DUT can be swept and the resultant variation of device behaviour related to the change of each parameter can be accurately captured. Such investigations would help designers in RF nonlinear circuit design in a similar manner to how s-parameters do in linear circuit design.

With the accurate nonlinear characteristics obtained by large signal waveform system, the following main benefits can be resulted. Firstly, the design cycles can be reduced. Modern RF circuit design always involves several turns in design process, because the circuit specifications need to be carefully traded off in an empirical manner. This design process is elongated if inaccurate or insufficient information about the semi-conductor is available, especially when the circuit demonstrates certain level of nonlinearity. With the help of the elegant waveform measurement system, the number of iterations in design process can be reduced and hence the

design cycle can be shortened. Secondly, the design can be more efficient which lowers the cost of fabrication. Linearity has been an important figure of merit in modern non-constant envelope communication system such as WCDMA and has been strictly specified in most of these systems to prevent interference and distortion. By taking advantage of the waveform measurement system, the designed circuit can be optimized to achieve a better balance between cost and linearity. Thirdly, the efficiency can be optimized. Since efficiency normally has a strong connection with linearity, a better understanding of nonlinearity would definitely yield quality circuit design with better efficiency.

This work will focus on large signal waveform measurement data collected when the device is excited by single tone signal. Part of the data was collected solely for this research by myself while the rest of data was taken for other projects. The single tone measurement is often referred to as Continuous Wave (CW) measurement as the input signal is pure sinusoidal. CW excitation can be considered as the simplest form of RF excitation signal to a RF device. The response of device to sinusoidal signal represents fundamental RF characteristics of the DUT. Therefore CW measurement is always an essential part in any RF power transistor device characterization process.

Many modern communication systems require the power amplifier to operate dynamically in terms of output power. These systems use sophisticated modulation scheme in order to maximize the number of users that can be catered within a limited bandwidth. However, these bandwidth-efficient concepts have one side effect of generating signals with varying envelope. Therefore, it's necessary, when doing device characterization, to characterize the device behaviour by sweeping the power to cover the variation range of envelope. The group of input power swept waveforms allows the analysis of device performance by calculating traditional figures of merit such as Gain, PAE and output power characteristics. Figure 3-7 gives examples of the gain, PAE and output power variation. These figures were calculated by using the

voltage and current waveform data.

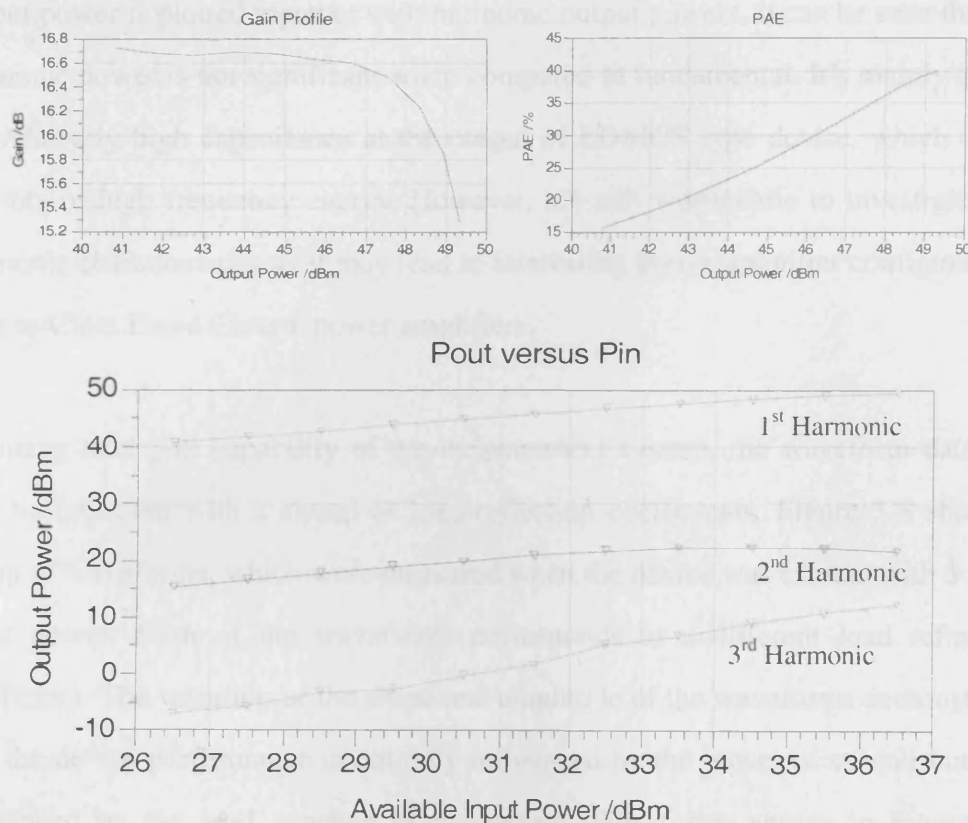


Figure 3-7: Examples of gain profile, PAE, output power characteristics can be derived from measured waveform information.

The gain profile shows that the device has linear gain of approx. 16.7 dB. The device starts to saturate around 36 dBm and has output power of about 49 dBm at 1 dB compression point. The output power at one dB compression point is often used to compare the power capabilities of different transistor devices. The PAE graph shows that the efficiency of the device increases gradually as the output power increases. The efficiency reaches its peak with output power of around 49 dBm, which is the 1dB compression point of the device. As the device is driven harder, the PAE starts to decrease due to the saturation of device. It can be observed when the device is operating at low power levels, the efficiency is very poor and lots of energy dissipates as heat. The graph is quite typical for power transistor devices. Since efficiency is also an important measure of power amplifier, efficiency enhancement techniques are developed to deal with the low efficiency problem when the device is

driven with low power. In the output power characteristic graph, the fundamental output power is plotted together with harmonic output powers. It can be seen that the harmonic power is not significant when compared to fundamental. It's mainly due to the relatively high capacitance at the output of LDMOS type device, which filters out lots of high frequency energy. However, it's still worthwhile to investigate the harmonic characteristics as it may lead to interesting power amplifier configurations such as Class E and Class F power amplifiers.

By using load pull capability of the measurement system, the waveform data can also be collected with a sweep of load reflection coefficients. Figure 3-8 shows a group of waveforms, which were measured when the device was excited with a fixed input power. Each of the waveforms corresponds to a different load reflection coefficient. The variation of the shape and amplitude of the waveforms demonstrates that the device performance is not only influenced by the power of stimuli but also influenced by the load attached to the device. Waveforms shown in Figure 3-8 illustrate that by changing the impedance of the load attached to the device, the amplitude of voltage waveform can vary by 20 volts. Similar amount of variation can also be observed from current waveforms. Since power has strong dependency on current and voltage, it can be derived that the output power characteristics would show significant variation as well.

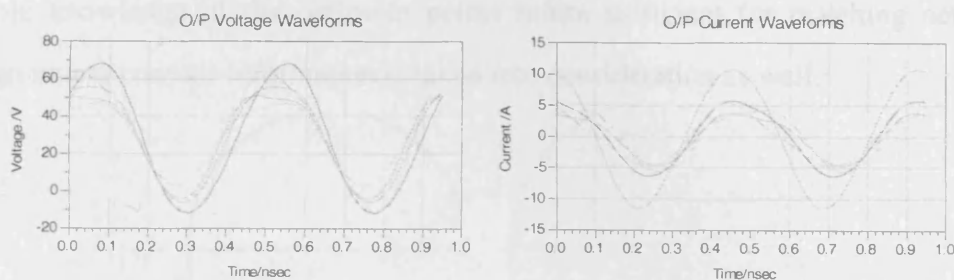


Figure 3-8: Examples of the voltage and current waveforms measured at different fundamental load reflection coefficient points.

The device performance variation with load can be better presented using contour

plots from power amplifier design perspective. Figure 3-9 shows the gain contours and output power contours obtained from the waveform information. The characterization range covers about only one fourth of the power contours. The remaining three quarters of the contours are not covered because load pulling in the missing area would require huge power from active load, which can possibly either damage the measurement system or cause irrevocable deterioration of transistor device under test.

The centre point on each of the contours is the optimum load impedance point. If the device is terminated at the optimum point, optimum device performance in terms of maximum output power or maximum gain can be obtained. It can be spotted that the output power contours and gain contours have different optimum load impedance point. It implies that compromise has to be made between gain and output power when doing power amplifier design.

The contour information can be very useful in designing matching network. A good matching network is supposed to set the appropriate impedance environment for the transistor device to make sure the device performance is optimized by considering all the factors of interest simultaneously. During realistic power amplifier design, trade-off is always necessary in order to meet the overall performance requirements. Simple knowledge of the optimum points is not sufficient for matching network design unless contour information is taken into consideration as well.

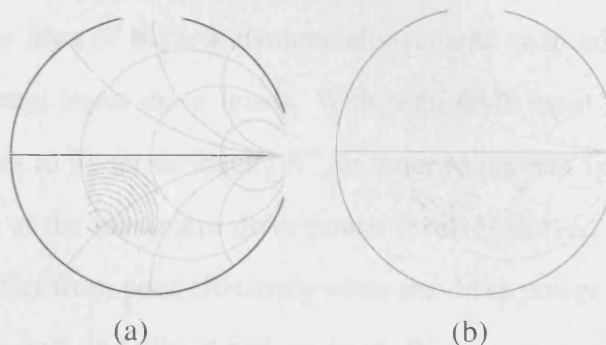


Figure 3-9: Examples of the gain contours (a) and output power contours (b) derived from waveform measurement.

One unique advantage of the Cardiff waveform measurement system is that the resultant measurement data contain both waveform and load-pull information. On one hand, waveform representation of the device nonlinear characteristics gives invaluable information on how the device behaves. Such information can be used to evaluate and identify the origin of the nonlinearities accurately. However, waveform analysis itself is not a design-oriented tool as it reveals little information about how the performance varies when the impedance environment changes. On the other hand, load-pull technique measures the performance variation when the impedance environment changes. It provides necessary information for RF circuit designers to develop proper matching network to trade off the different figures such as gain, efficiency, and etc in order to meet the required specifications. It can be seen that the waveform characterization and load-pull characterization are complimentary to each other and the combination of the two makes sure that the waveform information can be effectively utilized in RF power amplifier design.

Biasing conditions can also be swept to investigate the variation of the power amplifier performance when the dc voltages change. In mobile applications where battery is used as the primary power source of the power amplification unit, efficiency is very important to maximize the communication time. In order to achieve good efficiency with modern high-PAR modulation scheme, various PA operation schemes such as biasing the device dynamically via supply regulator [57] are introduced. The idea of biasing dynamically scheme is to adapt proper biasing condition for different input drive levels. With high PAR input signals, the power amplifier usually has to be set to “back off”, in order to prevent spectrum spread due to the compression at the maximum drive power level. However, backing off power amplifier would suffer from poor efficiency when the drive power level is low as lots of dc power is wasted and dissipated as heat. Bias varying scheme solves this problem by lowering dc biasing for low input power level, which reduces the amount of dc power and therefore increases the efficiency at low drive power levels.

Obviously, implementing the above method would require the knowledge about how the device behaviour in several different biasing conditions. The required knowledge can be obtained by doing waveform measurement with swept biasing conditions.

The concrete strategies about how to utilize the single tone waveform data obtained from measurements, especially in CAD, form the backbone of the work presented next in this thesis. In this work, two different strategies are explored and presented. One approach is direct utilization of the waveform data in CAD without any modelling process. The measurement waveform data are perfectly recoverable from simulation. The other approach is derived from modelling perspective. CAD large signal modelling can be a possible solution to bridge the gap between the waveform measurement data and CAD-based RF circuit design. Investigations show that both approaches have got their own pros and cons. It will be shown in this work that the waveform data can be utilized most effectively if the two proposed approaches are integrated together.

Chapter 4 Direct Utilization of Waveform Data in CAD

4.1 Introduction

The large signal waveform characterization technologies explained in previous chapter allow designers to investigate the device nonlinear characteristics in realistic large signal working conditions. As discussed in previous chapter, the reliable measured voltage and current data have demonstrated a great potential to help designers to achieve a first pass design. However, this potential is impeded by the fact that such a kind of data is usually not directly accessible in Computer Aided Design (CAD) tools.

In this chapter, a novel method for direct utilization of nonlinear large signal waveform measurement data in CAD tools, termed Direct Waveform Looking Up (DWLU) approach, is presented to solve the above problem. By using this approach, the large signal data can be utilized in CAD for design purpose immediately after the large-signal device characterization is performed without any complex post processing or model extraction. The direct use of non-linear data guarantees an accurate representation of the device characteristics within the CAD environment, hence increasing the likelihood of a first pass design based on CAD – provided the device is simulated exactly under the same conditions. This requires the stimulus signal and the impedance environment, presented to the DUT, match the ones used during the measurement.

Rather than a simple data replication tool, this approach can also be used to predict the nonlinear characteristics of power devices off the measurement grid by employing a nonlinear interpolation algorithm, Spline function [65].

First of all, the benefits brought by importing large signal waveform measurement data in CAD tools will be illustrated. Next discussions of the nonlinear measurement data indicate the direct import of nonlinear measurement data into CAD is fully feasible. To increase the robustness of simulation in CAD, it is necessary to formulate the measurement data into proper format. The way in which the data are formulated will be detailed. Then the realization in ADS [5] will be demonstrated. Finally, experimental results verifying the robustness of the approach will be shown.

4.2 Issues related to RF large signal CAD simulations

Computer Aided Design (CAD) tools [58] – [60] become increasingly important in RF power amplifier design. It is especially true when designers have to deal with the increased complexity of modern sophisticated communication system which usually put strict requirements on power amplifier specifications. Traditional approaches normally start the fabrication of prototypes following only simple initial theoretical analysis of the transistor device and possible circuit topologies. It results in a number of design iterations and quickly the number of iteration becomes impractical as the complexity increases. CAD tools offer the advantages of conducting cost-saving system level simulations in which all the sub-circuits can be integrated and tested as one system. It saves lots of design time by allowing designers to tweak the circuit and trade off the important design factors in order to meet all the specifications with lowest cost. Recent years has seen the rapid development of CAD tools. They are becoming more and more reliable and powerful as new CAD algorithms evolve.

The accuracy of CAD-based RF circuit design would be determined by both the quality of CAD and the quality of realistic RF component representation in CAD. With the help of s-parameters, most of the passive components can be accurately represented in CAD. Active device such as transistors, working in small signal conditions, can generally be accurately approximated by using s-parameters. But in large signal operations, s-parameters become insufficient as nonlinearity starts to

dominate. Therefore, in order to obtain accurate large signal representation of active devices in CAD, Various approaches are proposed. The resultant representation of devices is usually referred to as large signal model. Traditionally, before the advent of large signal waveform measurement technique, large signal models are usually extracted from static DC I-V measurements [61], complemented by s-parameter measurements which by definition is small signal dynamic characterization of the active transistor devices.

However, large signal models, which are extracted from DC I-V and s-parameter data, are used in large signal dynamic simulations during RF circuit design. The large signal voltage and current waveforms, or their frequency domain equivalent, need to be simulated by using models which are essentially extracted from small signal data. The gap between where models are extracted and where they are used in simulation cause the main concerns about the accuracy of CAD-based RF circuit design under large signal operations. To characterize the device in a more realistic condition, pulsed IV measurement systems have been recently developed and the use of resultant pulsed IV data in model extraction has lead to more accurate models [62] – [63]. But the gap mentioned above is still there as during pulsed IV measurement the device under test is still not characterized under realistic large signal dynamic conditions – pulsed IV measurement is essentially a DC-IV type measurement. Another obstacle that the classic modelling techniques are facing is to accurately model very large devices with output power level exceeding 50 dBm. Direct I-V characterization of such devices is difficult as the device impedance is much lower and the heat dissipation is much higher. The alternative is to develop a model on a small device and then scaling it up. This method is also problematic as high power devices usually have much complex structure which is not a simple addition of low power cells.

By taking the advantage of large signal waveform data, the gap discussed above can be eliminated. The large signal waveform data basically offer all the information that

is required in CAD simulation. Besides, with the approach explained in [18], high power devices can also be accurately characterized and full waveform characteristics can be obtained. Usage of waveform data in CAD has the potential to greatly improve the accuracy of simulation and consequently facilitate the development of reliable and rapid CAD-based power amplifier design.

4.3 Large signal waveform data scaling

4.3.1 Large signal measurement data at a first glance

The device under test is usually characterized with a CW stimulus and the resultant measurement data are saved into a data folder created on a standard Windows PC system. Normally for CW waveform measurements, both the input power and the load-pull sweeps would be conducted. In the data folder, there are also many sub-folders. Each of them corresponds to an individual load impedance point. And in each sub-folder, there are a number of txt files which record the amplitude and phase information of the current and voltage spectra at all calibrated harmonics. The number of the txt files depends on the number of input power sweep points. Figure 4-1 shows an example of the txt file which contains the saved amplitude and phase information of the current at output port.

```

REM      Device C 1 i vg = 2.63U, vd = 28U, idsq = 945mA,
REM      Current at port 2
BEGIN DSCRDATA
%      INDEX      MAG      PHASE
1      2343.27  0
2      1156.72  150.66
3      76.7381  -0.0766527
4      2.65937  71.4436
5      1.10074  68.078
6      2.07151  -52.4083
7      1.08716  55.782
8      0        180
9      0        180
10     0        0
11     0        0
12     0        0
13     0        0
14     0        0
15     0        0
16     0        0
17     0        0
18     0        180
19     0        180
20     0        180
21     0        0
22     0        0
23     0        0
24     0        0
25     0        0

```

Figure 4-1: Saved measurement data in a txt file.

The index number below the percentage mark indicates the index of harmonics. The magnitude and phase values of the spectral components are recorded. No unit is used in the data table for the sake of brevity. In this example, the measurement data are valid up to 5th harmonics. Only the load impedances at first three harmonics were accurately controlled. It means in the data table only the first four rows (index1 – index4 corresponds to DC – 3rd Harmonic) of the data are meaningful – the rest rows are merely random numbers that should be neglected.

Even though the waveform measurement system has got certain facilities in its software to interrogate the measurement data, it would make more sense to do the offline investigations in a separate software environment. A data extraction routine has been written to extract the measurement data from these txt files and load them up into data analysis software. The selected data analysis software is Igor Pro [52] which has got superior functionalities for waveform-type data analysis. The codes have been integrated as part of the direct data import routine and are enclosed in the accompanying CD.

As shown in Figure 4-1, it can be seen that the full spectrum information covering fundamental frequency and a few harmonics of DUT excited by large signal CW stimulus is captured. The waveform characteristics of the device under test can be reconstructed from the spectrum information by Inverse Fourier Transformer (IFT).

The waveform characteristics contain lots of valuable information of the device under test. The common figures of merit such as 1dB gain compression point, gain profile, efficiency profile and AM-PM distortion can be easily derived mathematically from the waveform data. The Igor Pro or any other data analysis software can be used to investigate the device behaviours of interests. However, from a PA designer point of view, typical RF design CAD tool would always be a better choice than generic data analysis software. Firstly the RF design oriented functionalities available in CAD tools make the investigations of the device characteristics much easier. Secondly and also most importantly, the CAD tools allow designers to evaluate the device performance with auxiliary circuitries through simulation, which would certainly help reduce the design cost and shorten the design circle. Therefore, import of large signal measurement data into CAD tool could definitely be highly appreciated by PA designers. Fortunately, the common CAD tool defines the nonlinear behaviour of device in terms of waveform. It would imply that there should be a way to import the large signal waveform measurement data directly into CAD software. In following sections, an approach which enables the direct import of waveform data into CAD software is proposed. The general flow of the proposed approach is illustrated in Figure 4-2. It can be observed that there is only a simple scaling step before importing the measurement data into CAD. The approach offers a direct link between the large signal measurement system and the CAD simulation software. The link facilitates the efficient utilization of large signal waveform data in CAD-based power amplifier design.

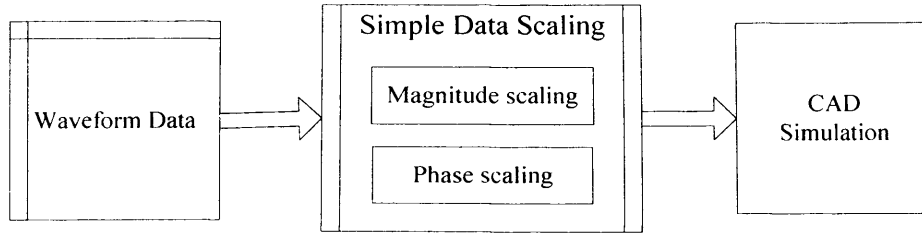


Figure 4-2: The general flow of processing measurement data for use in CAD.

4.3.2 Measurement data scaling

The measurement data scaling is used to formulate the measurement data in such a way to achieve best robustness in simulations, i.e. improved flexibility and convergence. The measured waveform data conserve the device characteristics in the format of voltage and current waveforms. The data can be conveniently analyzed in either time or frequency domain by utilizing FFT. Since the device is stimulated with a single tone signal, only DC, fundamental and harmonic exist. But there is no Intermediate Frequency (IF) term in spectrum. Therefore, a frequency domain approach is adopted for formulating measurement data. The data formulation is developed based on a typical two port network structure:

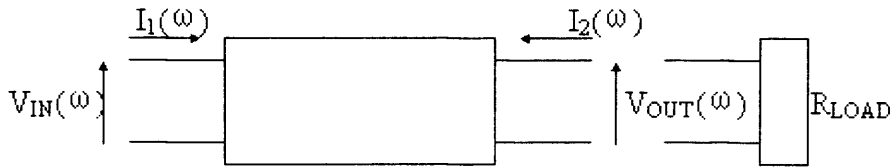


Figure 4-3: A typical two port network description of transistor devices. The vectors' directions are traditionally defined as shown above.

As shown in Figure 4-3, the I_1 and I_2 are referred to input and output port currents respectively. The currents are the chosen variables in the data formulation. Under large signal excitations, as shown in (4-1) – (4-2), the $I_1(\omega)$ and $I_2(\omega)$ are generally a function of input stimulus, the load impedance and bias conditions.

$$I_{1n} = F_1(V_{IN}, \Gamma_{Load}, V_{1DC}, V_{2DC}) \quad (4-1)$$

$$I_{2n} = F_2(V_{IN}, \Gamma_{Load}, V_{IDC}, V_{2DC}) \quad (4-2)$$

In (4-1) – (4-2), I_{1n} and I_{2n} are the current spectra of input and output port respectively. V_{IN} is the input voltage stimulus at the fundamental frequency. Γ_{Load} is the reflection coefficient of the load attached to the output of the DUT. V_{IDC} and V_{2DC} are the gate and drain DC bias voltages (Base and Collector bias for BJT device) respectively.

The measurement data formulation is done in frequency domain in two steps:

a) Phase scaling

First of all, the phases of currents are normalized to the phase of the input voltage stimulus at the fundamental:

$$I_{1n, Renorm} = \frac{I_{1n}}{\left(\frac{V_{IN}}{|V_{IN}|} \right)^n}; \quad I_{2n, Renorm} = \frac{I_{2n}}{\left(\frac{V_{IN}}{|V_{IN}|} \right)^n} \quad (4-3)$$

In (4-3), V_{IN} is the fundamental input voltage stimulus. The lower case n stands for the nth order of harmonics. “Renorm” indicates “renormalized”. It’s worth noticing that in order to maintain constant group delay, the nth order harmonics is scaled with n times of the phase of the fundamental input voltage stimulus.

The phase scaling is required to ensure the efficient utilization of measurement data CAD simulators. Since phase is a relative variable, reference phase is defined to designate zero-phase to a certain vector and the phases of all the other vectors within the system are referred by calculating the phase difference between the vectors and the reference phase vector. As a result, the phases of the vectors within the system will vary if the reference phase is relocated. The waveform measured data, which include the phase information, thus have to be considered as a function of reference phase of input stimulus. Figure shows an example of the measured current variation with different value of reference phase.

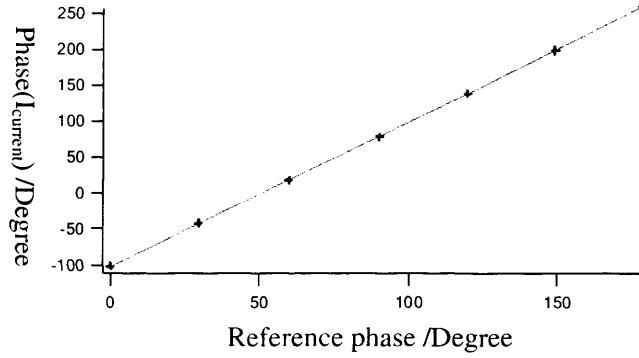


Figure 4-4: Phase of measured current vector varies if reference phase is changed.

Different measurement system set-ups may use different reference phase for the input stimulus. For example, some system may set the phase of input incident travelling wave as zero while other systems may set the phase of input voltage wave as zero – there is no standard rule about setting reference phase in measurement system. Sometimes even specific initial phase value is chosen for a_1 or V_1 . Therefore, if no phase normalization is done, the information about reference phase has to be saved and imported into CAD in order to retrace meaningful measured characteristics from simulation. However, the information about the reference phase of the measurement system won't affect the behaviour of the device under test in any way and thus should be eliminated for efficient utilization of data. In fact only the relative phase difference between the input and output variables are important. So the phase scaling is introduced into the formulation for this purpose. It makes the values of I_{1n} and I_{2n} independent of reference phase and thus gives the CAD simulator the freedom to choose any reference phase without the need to re-calculate I_{1n} and I_{2n} . It also simplifies the dependency of I_{1n} and I_{2n} by making them independent of phase of input stimulus. Therefore, after phase normalization, the dependency of port currents on the external stimulus becomes:

$$I_{1n} = F_1(|V_{IN}|, \Gamma_{Load}, V_{1DC}, V_{2DC}) \quad (4-4)$$

$$I_{2n} = F_2(|V_{IN}|, \Gamma_{Load}, V_{1DC}, V_{2DC}) \quad (4-5)$$

Figure 4-5 demonstrates the scaled currents are no longer dependent on the reference

phase.

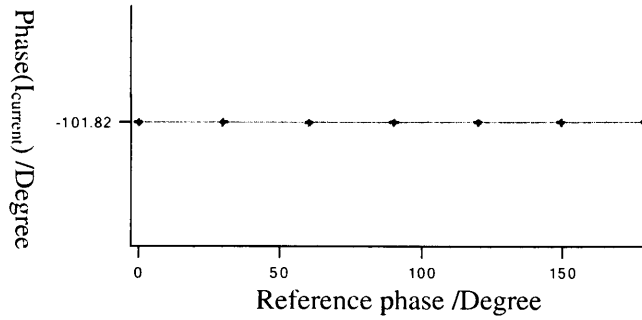


Figure 4-5: Scaled phase variation of current vector remains constant when the reference plane is changed.

b) Magnitude scaling

After the phase scaling, theoretically the scaled currents can be imported into CAD simulators already. However, from practical perspective, further data processing can be done to ensure better performance in terms of interpolation accuracy within simulators.

The following graph shows the real part of the 2nd harmonic component variation against fundamental input power. Due to nonlinearities of the device, the 2nd harmonic component demonstrates certain level of dependency on the 2nd order of fundamental input stimulus. It can be deduced from the nonlinear variation of the 2nd harmonic curve. This type of dependency complicates the curve of the 2nd harmonic variation and may degrade the interpolation performed later on by the CAD simulators.

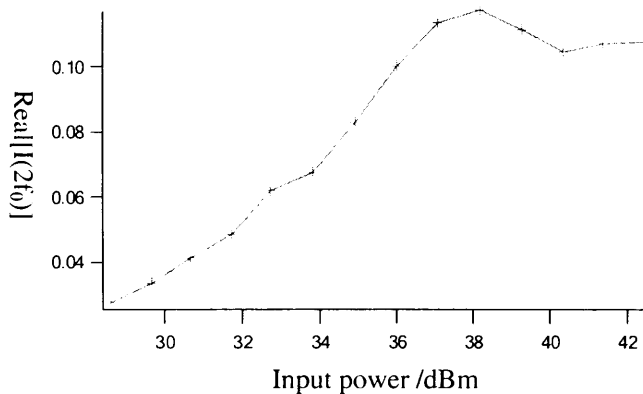


Figure 4-6: The 2nd harmonic component variation as a function of fundamental input power.

Therefore, magnitude scaling operators (4-6) are introduced to minimize the current dependency on the level of input stimulus and simplify the current variation curves.

$$I_{1n,Scaled} = \frac{I_{1n}}{|V_{IN}|^n}; I_{2n,Scaled} = \frac{I_{2n}}{|V_{IN}|^n} \quad (4-6)$$

In (4-6), higher order current harmonics are scaled by the magnitude of input voltage stimulus with corresponding order. That is to make sure the scaled current curve has minimal magnitude variations – to make the interpolation task less challenging.

The scaled data were plotted in Figure 4-7, which shows clearly the resultant flatter and smoother curves comparing with Figure 4-6. It is easier for the interpolation function in the simulator to interpolate the values along such a smoother curve and thus increase the accuracy of simulation. The scaling operation can be very effective if high order of harmonic data needs to be taken into account in CAD simulations.

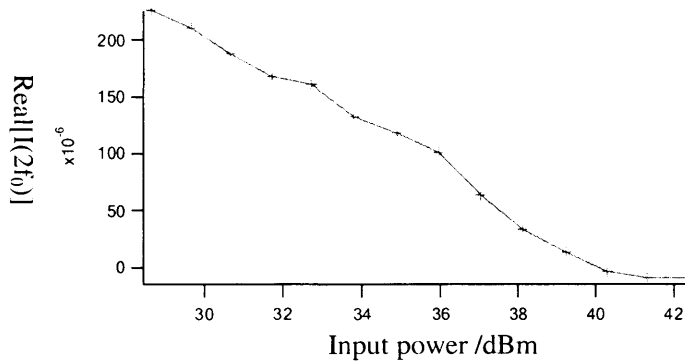


Figure 4-7: Scaled 2nd harmonic component variation as a function of fundamental input power.

Therefore, by combining the formulation in both steps, finally a new data set can be obtained.

$$M_n = \frac{I_{1n}}{\left(\frac{V_{IN}}{|V_{IN}|}\right)^n \cdot |V_{IN}|^n} = \frac{I_{1n}}{V_{IN}^n} \quad (4-7)$$

$$N_n = \frac{I_{2n}}{\left(\frac{V_{IN}}{|V_{IN}|}\right)^n \cdot |V_{IN}|^n} = \frac{I_{2n}}{V_{IN}^n} \quad (4-8)$$

Please note voltage and current waves are chosen instead of scattered waves to define the coefficients M_n and N_n . In fact the DWLU approach has been tested by using scattered waves for investigation purpose. The equivalent a and b wave based M'_n and N'_n are calculated and the coefficient surfaces over the load impedance plane are compared to those of M_n and N_n . It has been found that the surface smoothness generally looks similar, which implies that theoretically the DWLU approach should be able to give the same accuracy regardless of whether V&I waves or a&b waves were chosen. In this research work, the Agilent CAD software Advanced Design System (ADS) is chosen for conducting all the simulation work. When the measurement data is being imported into ADS, V&I waves will gain advantage due to the easier transformation into the software, which make it a perfect option for the realization of the proposed approach using V&I waves. When employing a&b waves extra conversions are needed.

Based on the analysis of the scaling process it can be found that the information of the raw data is fully preserved and the measurement data are perfectly recoverable from M and N dataset – no approximation is done in the scaling process at all. The purpose of the scaling is to losslessly convert the waveform measurement data into a neat and robust format facilitating the import of data into CAD. Besides, it needs to be mentioned that (4-7) and (4-8) are only used to process the data at RF frequencies including both fundamental and harmonics. The DC data are kept intact and imported directly into CAD.

4.3.3 “Grey box” versus “black box”

The large signal measurement data are usually collected at DUT extrinsic terminals. The parasitic effects, due to the package of DUT, such as series resistance and parallel capacitance are embedded in the measured waveform data. Parasitic effects are always causing problems in measurement data interpretation as such effects may distort the device characteristics and consequently prevent the extraction of true relationship between internal voltages and currents. Therefore, careful investigations are needed to explore how the parasitic effects influence the robustness of the proposed direct waveform utilization approach.

Namely, depending on whether the parasitic effects are taken into account, the data are usually referred to as “grey box” data or “black box” data. The grey box data refers to the data accessible at the internal device plane while the black box data refers to the data measurable at the outside package plane. The measurement data mentioned in this thesis are usually collected by performing the waveform measurements on packaged devices, as shown in Figure 4-8.

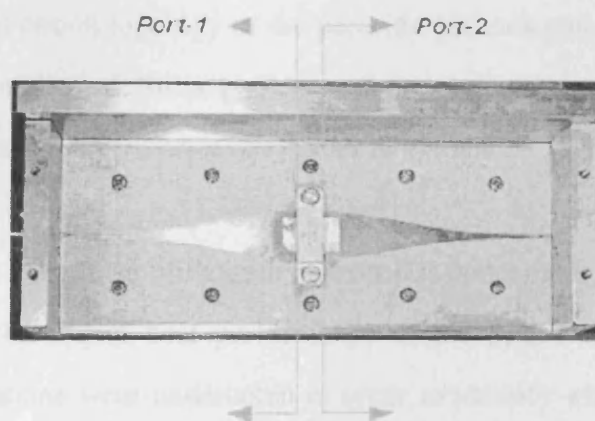


Figure 4-8: The measurement plane has been marked on the test fixture. The test fixture includes an impedance transformer facilitating high power measurement.

The measurements are done at measurement plane, which is separated from the

device plane by the parasitic network. The following figure shows a generic equivalent circuit topology of the parasitic network for a typical pHEMT device.

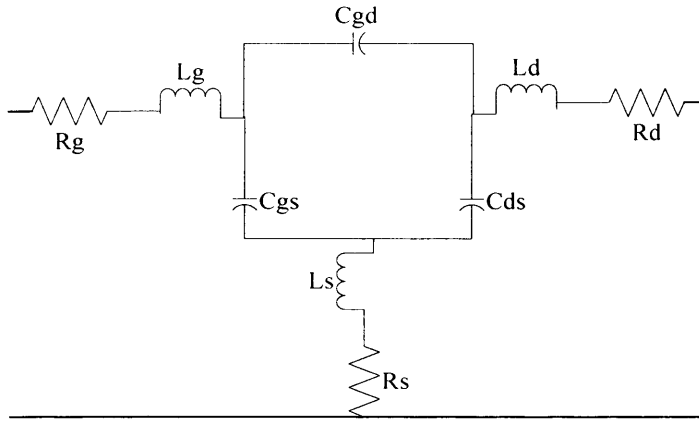


Figure 4-9: Equivalent circuit of parasitic network for a typical pHEMT device [28]

The parasitic network can cause the discrepancy about how the device performs between measurement plane, i.e. outside the package and device plane. The parasitic effects may cause a device which behaves well and smoothly at the device plane to show less smoothness or even discontinuity at the measured plane. The parasitic network can be a difficult problem for RF designers who require access to the device plane. It is not easy and sometimes very complicated to extract the parasitic network since the equivalent circuit topology of the parasitic network can be either difficult to identify or over complicated. Some packages also include internal matching network which makes the parasitic network even harder to extract.

As far as the direct waveform utilization approach is concerned, a smooth parameter surface is desired to ensure best performance of the Spline interpolation can be achieved. Investigations were undertaken in order to identify whether it's necessary to use grey box data rather than black box data for importing the data into the CAD.

Investigations were undertaken on the measurement data collected from a 4w pHEMT device. The equivalent circuit topology shown in Figure 4-9 was chosen for

identifying the effects of the parasitics. The classical “cold FET” method [64] was used to identify the values of the circuit components. The extracted values for each of the circuit components are shown in Table 4-1.

Parasitic coefficients	Values
C_{ds}	70 fF
C_{gs}	75 fF
C_{gd}	13 fF
L_d	59 pH
L_g	93 pH
L_s	1.7 pH
R_s	0.62 ohm
R_d	0.66 ohm
R_g	0.1 ohm

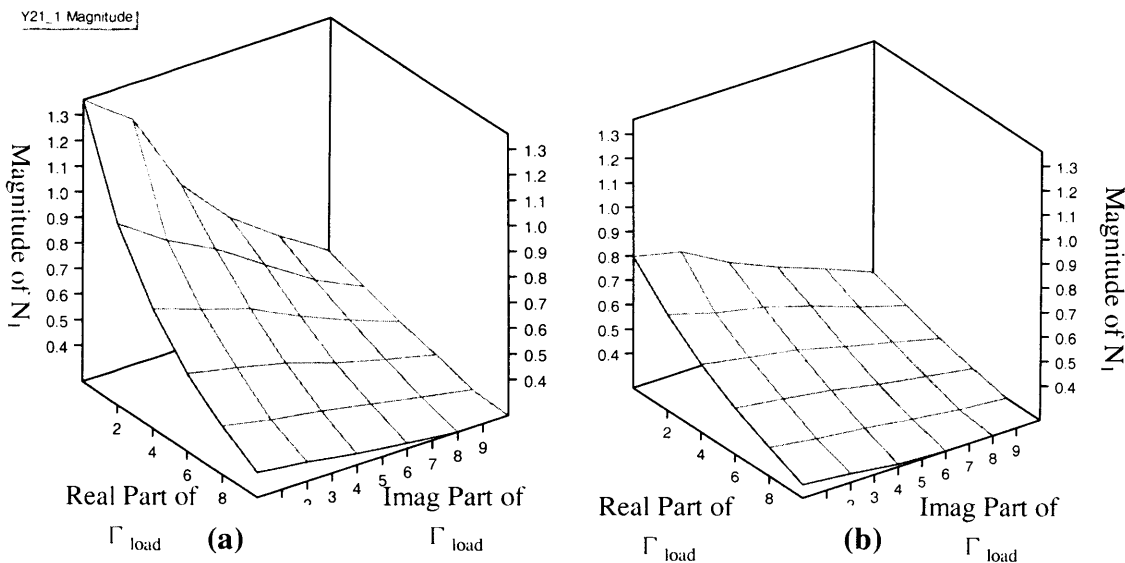
Table 4-1: The identified values of the parasitic circuit components

Once the parasitic equivalent circuit is identified, the measured large signal waveforms can be de-embedded to eliminate the parasitic effects. The maths introduced in [49] was applied for the waveform de-embedding. Accordingly, an Igor routine has been written in this research work which can de-embed the parasitic effects and work out the current and voltage waveforms at the internal device terminals. The de-embedded current and voltage data at the device plane were then processed according to (4-7) – (4-8). The resultant $M_{de-embedded}$ and $N_{de-embedded}$ parameters surfaces are compared to the M and N parameters which were extracted from normal measurement data. Some of the comparisons are shown in Figure 4-10 and Figure 4-11.

As shown in the two figures, the de-embedded M and N don't show much advantage in terms of surface smoothness. The de-embedding does make the surface varied less

than the external ones but the device generally behaves well and no rapid variation can be spotted before or after de-embedding. It can be seen that the de-embedding has little impact on the robustness of the Spline interpolation. Therefore it is not necessary to do the time-consuming and error-prone de-embedding to get rid of the effects of the parasitic network. Generally speaking, it saves a lot of data processing time and allows the measurement data to be imported into CAD right after they become available from the large signal waveform measurement.

Though de-embedding has little impact on interpolation, it's worth pointing out that the de-embedding may improve the extrapolation to certain extent. For example, in small signal operation, the de-embedding can remove the parasitic effects and the resultant de-embedded s-parameters can be accurately extrapolated in terms of frequency. Similarly, it's reasonable to derive that the extrapolation in large signal operation may benefit from de-embedding as well. However, the work presented here focuses primarily on the replication and interpolation of the waveform measurement data in simulation. The impact of de-embedding on extrapolation ability of the proposed approach is out of scope of this thesis and will be left in future work.



Y21_1 Phase

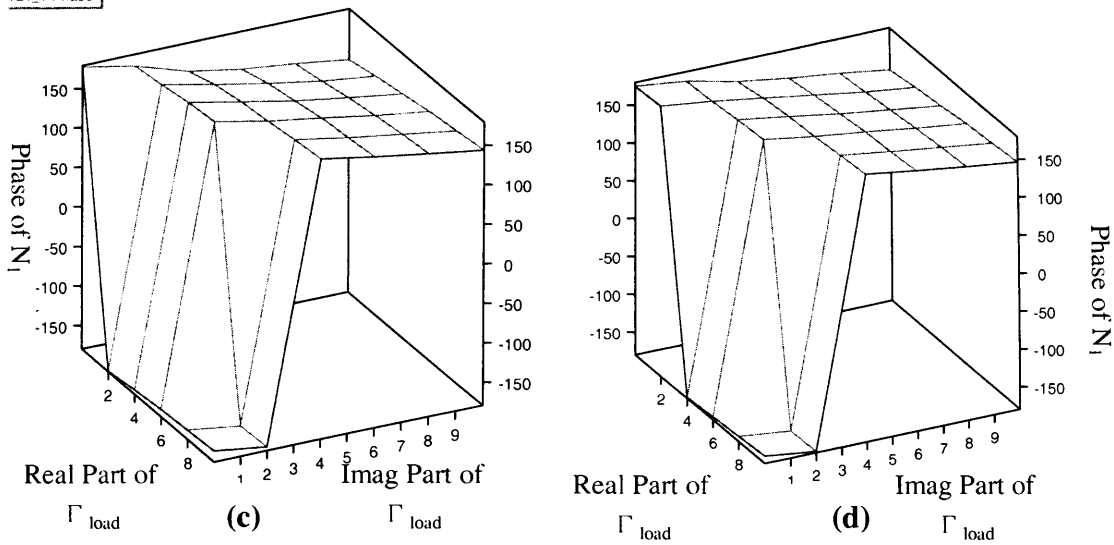
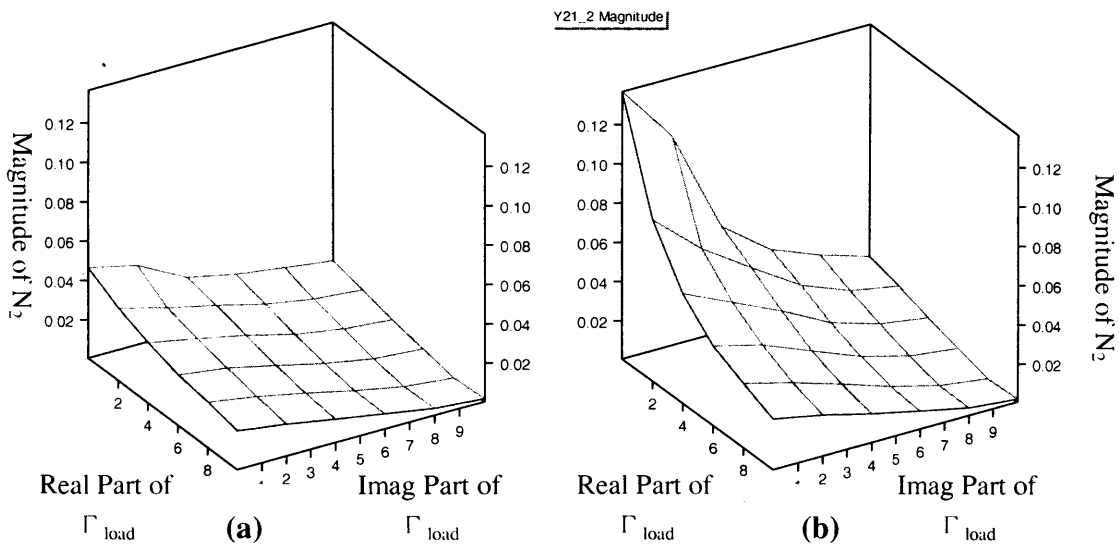


Figure 4-10: The graphs compare the magnitude (a) and phase (c) of N_1 obtained from extrinsic measurement data to the magnitude (b) and phase (d) of N_1 obtained from de-embedded data.



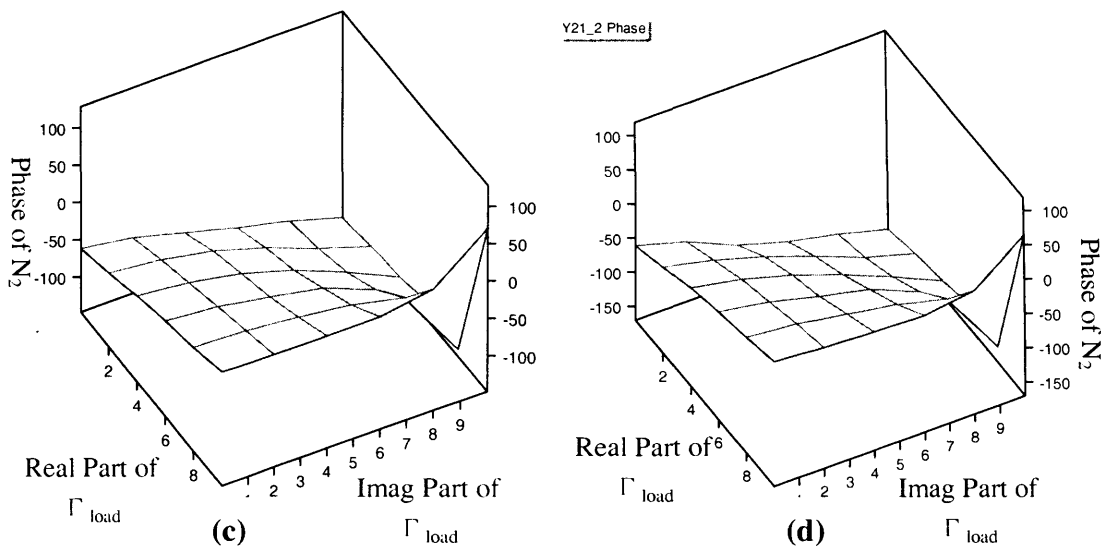


Figure 4-11: The graphs compare the magnitude (a) and phase (c) of N_2 obtained from extrinsic measurement data to the magnitude (b) and phase (d) of N_2 obtained from de-embedded data.

4.3.4 Data table generation

The process of waveform data has been made automated by utilizing Igor software. A user-friendly interface is developed to facilitate the configuration of measurement data processing. A snapshot of the user interface is shown below

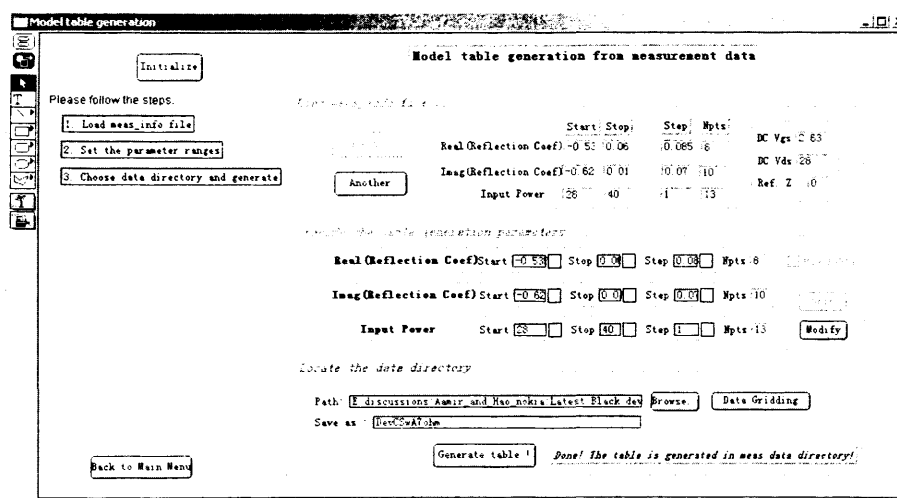


Figure 4-12: The user interface for measurement data scaling and data table generating. The interface is developed using Igor software.

The user interface offers functions including extraction of the IV information out of the raw measurement data, selecting the desired number of measurement points for importing into CAD and scaling of the data as previously described. The processing time would mainly depend on how many measurement points need to be processed. Normally the whole data processing would take less than ten minutes on a normal PC for handling a measurement dataset involving a thousand data points.

After the data processing, the generated M and N dataset is saved into a multi-dimensional data table in MDIF format. An example of generated MDIF containing M and N dataset is illustrated below:

```
File Edit Format View Help
!Reflect Coef Real:-0.62, 0.08, 0.1
!Reflect Coef Imag:-0.75, -0.1, 0.13
!Reflect Coef Power:26, 40, 1
!DC biasing: vgs=2.73; vds=28
VAR Vgs (real)=2.73
VAR Vds (real)=28
VAR Gamma_a (real) = -0.62
VAR Gamma_jb (real) = -0.62
BEGIN RefCoef_Sweep_1
%VImag(real)    I1_dc(real)    I2_dc(real)    v1_dc(real)    v2_dc(real)    M_1(complex)    N_1
5.11339 -5e-05  1.40366 2.72802 28.0108 0.00585682 -0.0184839 -0.110642 0.0255069
5.7324 -5e-05  1.51127 2.72802 28.0097 0.0058637 -0.0184524 -0.110445 0.0256549
6.46055 -5e-05  1.63415 2.72802 27.9805 0.00581226 -0.0184806 -0.10971 0.027533
7.24498 -5e-05  1.77913 2.72802 27.9879 0.00587696 -0.0185072 -0.109667 0.0286534
8.13215 -5e-05  1.95188 2.72802 28.0018 0.00582887 -0.0185077 -0.109127 0.0301204
9.15234 -5e-05  2.16443 2.72802 28.0032 0.0058545 -0.0185621 -0.108532 0.0312343
10.2906 -5e-05  2.39628 2.72802 28.0019 0.00588208 -0.0185886 -0.106116 0.0346183
11.5804 -5e-05  2.68732 2.72802 28.0097 0.00580247 -0.0186285 -0.102644 0.0366901
13.0793 -5e-05  2.98172 2.72802 28.0054 0.00578724 -0.0187445 -0.0992436 0.0450715
14.7452 -5e-05  3.28302 2.72802 27.9986 0.00576527 -0.0186756 -0.0934458 0.0512058
16.6362 -5e-05  3.51839 2.72802 28.0046 0.00550857 -0.0185205 -0.0838524 0.0511199
18.8153 -5e-05  3.68913 2.72802 27.9913 0.00524176 -0.018317 -0.0744578 0.0475816
21.3997 -5e-05  3.85834 2.72802 27.998 0.00496378 -0.0181611 -0.0655682 0.0436834
24.3121 -5e-05  4.01258 2.72802 28.0057 0.00476744 -0.0180367 -0.0581988 0.0395628
27.7244 -5e-05  4.16306 2.72802 27.9956 0.00452582 -0.0179366 -0.0507068 0.0365299
```

Figure 4-13: Example of the generated data table containing M and N scaled measurement data.

As observed in the data table, the DC data and M & N parameters are saved as a function of the biasing condition (V_{gs} , V_{ds}), the fundamental load reflection coefficient (Gamma_a stands for the real part and Gamma_jb stands for the imaginary part), and magnitude of input voltage (VImag). It can be noticed that no term related to harmonic reflection coefficient appears in the data table as no

harmonic load-pull is done in this example, i.e. the harmonics are terminated into fixed value such as 50 ohm. In the case that harmonic sweeping is involved in the measurement, corresponding harmonic reflection coefficient is necessary to be included as a new independent variable.

4.4 Import of waveform data into CAD

Once the large signal waveform data are formatted properly, they are ready to be imported into CAD software and become available for further data investigation and power amplifier design. In general, in order to analyze and utilize the measured waveform data in CAD, a “virtual device” needs to be implemented in simulator whose behaviour of interest should be rigidly prescribed by the measurement data. Specifically, the virtual device should be able to read the measurement data in and work properly in standard simulators, with the simulation results in good agreement with the measurement results.

Agilent **Advanced Design System** (ADS) is chosen as the CAD tool in which the virtual device is implemented. ADS is a powerful RF CAD tool which has been widely used by RF designers in both academia and industry around the world. It offers sophisticated linear and nonlinear simulation algorithms (such as transient simulator, Harmonic balance simulator and Envelop simulator [5]) as well as handy data analysis functionalities. For example, the normal figures of merit in RF design can be calculated and plotted by utilizing the standard functions in the data plot display of the ADS. Particularly, it provides special build-in user-defined modules for users to implement their own system devices in ADS. Among these modules, there is one build-in module called Frequency Domain Device (FDD) which allows users to define mathematical functions in terms of voltage and current in frequency domain to relate the output to the input of FDD. Since the proposed waveform data utilization approach deals with the measurement data in the format of voltage and current, FDD can be ideal for realizing the desired virtual device in ADS. Figure

4-14 demonstrates the schematic of the implementation of the virtual device.

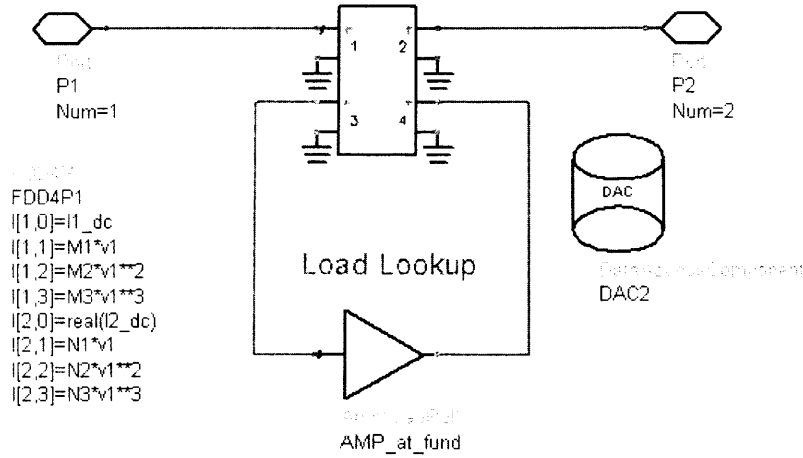


Figure 4-14: ADS schematic of the implementation of the virtual device

As shown in Figure 4-14, the schematic consists of three components including a four-port FDD, AmpLoadPull component and Data Access Component (DAC). Each of them enables certain functionalities, which will be explained later, and altogether a virtual device with desired features can be made.

FDD is used to recover the measurement data from the scaled M and N data set. The port current characteristics can be defined in FDD by using mathematical equations. The following equations are used to define the currents at both ports.

$$I[1, m] = M_m \cdot V_1^m \quad m = 1, 2, 3 \quad (4-9)$$

$$I[2, m] = N_m \cdot V_1^m \quad m = 1, 2, 3 \quad (4-10)$$

$$I[n, m] = I_{nDC, measured} \quad m = 0 \quad (4-11)$$

In $I[n, m]$, n is the port number while m is the order of harmonics. $I_{n, DC}$ is the measured DC current at port n . It can be seen from these equations that the port current characteristics from DC to 3rd harmonics are prescribed by M and N parameters. By multiplying with proper order of input voltage stimulus, the measured current spectrum can be perfectly rebuilt from the M and N data set, which can be easily deduced by combining (4-9) – (4-10) and the definition of M and N

data set shown in (4-7) – (4-8).

$$I[1, m] = M_m \cdot V_1^m = \frac{I_{1m, measured}}{V_1^m} \cdot V_1^m = I_{1m, measured} \quad (4-12)$$

$$I[2, m] = N_m \cdot V_1^m = \frac{I_{2m, measured}}{V_1^m} \cdot V_1^m = I_{2m, measured} \quad (4-13)$$

Therefore (4-9) – (4-10) can be considered as reverse process of the scaling process shown in (4-7) – (4-8). The input voltage stimulus is usually set by users and thus is a known condition when the simulation is started. The load impedance attached to the output of the FDD will be determined during simulation and then used to calculate the output voltage.

$$V[n, m] = Z_{m, load} \cdot I[n, m] \quad (4-14)$$

It can be seen that the full information about the device dynamic characteristics, i.e. V_{in} , V_{out} , I_{in} and I_{out} collected at the measurement plane during measurements, can be perfectly retrieved in ADS simulation.

DAC is used to refer the generated MDIF data table containing M and N coefficients by using the magnitude of input voltage stimulus and fundamental load reflection coefficient. Sometimes DC biasing is also referred as the third independent variable. The main function of DAC is to look up and return the proper M and N values for given input power and fundamental load reflection coefficient during simulation. If the input power and fundamental load reflection coefficient set in simulation match one of the points stored in the MDIF data table, the M and N values corresponding to that point are returned to the FDD to describe the input and output characteristics. If the input power and fundamental load reflection coefficient set in simulation do not exist in the data table, interpolation will be initiated to approximate M and N values by using the information at nearby known points.

It can be seen that the quality of interpolation algorithm adopted in simulation would have a great influence on reliability of DWLU approach when simulation points do

not sit on the measurement grid. DAC offers a variety of interpolation schemes covering both linear and nonlinear algorithms. Linear interpolation algorithm is the simplest but not appropriate for dealing with nonlinear data, which generally exhibit high order behaviour, as collected from the large signal waveform measurement system. Nonlinear interpolation algorithms available in DAC include Cubic and Cubic Spline [65] methods. By definition they should be able to accurately approximate nonlinear behaviour up to 3rd order. One prominent advantage of basic cubic interpolation method over linear method is that it offers true differential continuity at each point. As expected, the complexity of implementing cubic interpolation method also increases. Rather than just using adjacent points at either side of unknown point in linear method, cubic interpolation requires four data points – two at either side of unknown point [66].

Cubic Spline algorithm possesses all the advantages that normal cubic interpolation method has. However, compared to normal cubic polynomial algorithm, Cubic Spline algorithm introduces piecewise polynomial which can result in lower interpolation error and better accuracy. The piecewise feature implies that the algorithm will fit to the nonlinearity locally – piece by piece. Therefore, though Cubic Spline only utilize 3rd order polynomial for local fitting, it is able to fit much more complex nonlinearities than 3rd order – by slicing the whole characteristics into several pieces and fit each of them locally with 3rd polynomial. That's why Cubic Spline is chosen over normal cubic interpolation algorithm as the interpolator for DWLU approach. Experimental results on the difference of these interpolation algorithms will be presented later in this chapter.

Extrapolation of Cubic Spline is usually a problem for modelling semiconductor devices as it tend to explode outside the valid data region. On the contrary, in reality semiconductor device usually saturate but not explode. Though the accuracy of extrapolation is not the main issue addressed in this work, poor extrapolation may lead to CAD convergence problem. Therefore, a constant extrapolation will be used

instead of Cubic Spline for extrapolation. A constant extrapolation will make sure a convergence can always be achieved.



DataAccessComponent

```
DAC1
Type=Generalized Multi-dimensional Data
InterpMode=Cubic Spline
InterpDom=Rectangular
ExtrapMode=Constant Extrapolation
iVar1="V1mag"
iVal1=v1mag
iVar2="Gamma_a"
iVal2=Gamma_a
iVar3="Gamma_jb"
iVal3=Gamma_jb
iVar4="Vgs"
iVal4=vgs
iVar5="Vds"
iVal5=vds
```

var
Eq

VAR

```
Params_from_DAC
I1_dc=file{DAC1, "I1_dc"}
I2_dc=file{DAC1, "I2_dc"}
M_1=file{DAC1, "M_1"}
N_1=file{DAC1, "N_1"}
M_2=file{DAC1, "M_2"}
N_2=file{DAC1, "N_2"}
M_3=file{DAC1, "M_3"}
N_3=file{DAC1, "N_3"}
```

Figure 4-15: The DAC component setup and related variable definitions

Figure 4-15 shows the schematic setup of DAC and corresponding definitions of variables. The variables declared in DAC are the ones that DAC references to look up the M and N values. The variables include magnitude of input voltage (v1mag), load reflection coefficients (Gamma_a is the real part while Gamma_jb is the imaginary part), and biasing condition (Vgs is the gate dc voltage while Vds is the drain dc voltage). The variables defined right-hand side are the M and N parameters that need to be looked up from DAC. In DAC configuration, Cubic Spline is chosen as the interpolation method while constant extrapolation is chosen for extrapolation. Cubic Spline is very robust interpolation algorithm but the curve tends to explode exponentially outside the known data space. The explosion is a disaster for any nonlinear convergence algorithm and is very likely to cause the non-convergence of the simulation. Moreover, the explosion characteristic itself doesn't fit any typical transistor device behaviour, which trends to saturate eventually rather than exploding. A constant extrapolation replaces the Cubic Spline when simulation outside the

known data space is required. The constant extrapolation assumes constant M&N parameters outside the table range. Compared to Cubic Spline, constant extrapolation improves the simulation convergence significantly.

The FDD defines the input and output current characteristics. DAC links the measured current characteristics to FDD with reference to input voltage stimuli and load reflection coefficient. It can be seen that in theory a schematic consists of FDD and DAC should be sufficient to implement the proposed DWLU approach. However, FDD has one fundamental problem which seriously impedes its usefulness in practice. The problem is that FDD is not able to determine the impedance value of the load network attaching at the output side. This is because the FDD could not accept implicit V/I relation equations, i.e. equation in a format as (4-15) is not allowed.

$$f(V[1,1],\dots,V[n,m],I[1,1],\dots,I[n,m])=0 \quad (4-15)$$

Load impedance is usually determined by using both voltage and current information. Since no implicit equation can be used and there is no way to define an explicit equation which can have both output voltage and current, it is necessary to determine or provide the value of output load impedance.

This would restrict the use of FDD to simulations where the load impedance parameter can be directly determined or calculated. In most simulations the output load impedance has to be determined through the simulation and is not known as a parameter, which can be controlled. This would be the case when FDD is connected to a matching circuit, e.g. an LC ladder network or a distributed network consisting of some lines.

One potential solution to this problem is to setup two simulations in succession. One is to determine the impedance value of the load matching network. The calculated

impedance value is then supplied to the second simulation for use with FDD. However, this solution causes lots of extra complexity to both model implementation and simulation and therefore is not practical.

The solution presented in this work makes use of AmpLoadPull component, one of build-in system components in ADS, to solve out this problem. AmpLoadPull component, also referred to as SDD Load-Pull Amplifier component, is originally developed for load-pull simulation based on a table look-up mechanism. A component called LoadPullSetup component is used to extract an ADS internal dataset, given ranges and steps for input power, output reflection coefficient magnitude and output reflection coefficient angle. AmpLoadPull component uses this dataset (as shown in Figure 4-16) and allows fast load-pull simulations for all input power and output reflection coefficient values in the specified ranges.

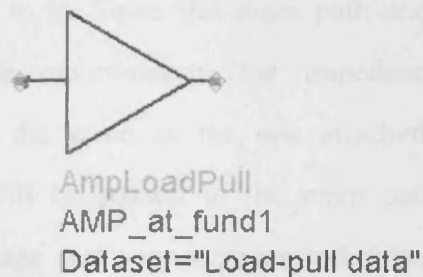


Figure 4-16 The AmpLoadPull component used for load impedance looking-up.

AmpLoadPull has the capability of determining the fundamental load impedance of any network attached to it during simulation. This feature can be used to release the simulation restriction of the FDD by helping the FDD to determine the load impedance attached to it. The schematic setup for integration of the FDD and the AmpLoadPull component is shown in Figure 4-14. It can be seen that a four port FDD is used and there are two paths around the FDD. One connected to Port P_1 and P_2 defines the input and output characteristics of the model, which can be referred to as main path. The other having AmpLoadPull in it is built up for load impedance determination, which can be referred to as auxiliary path. The following equations

relating the main path and auxiliary path are used to mirror the fundamental load impedance environment of the main path to the fundamental load impedance environment of the auxiliary path.

$$V[3,1] = \text{Const_v} \quad (4-16)$$

$$V[4,1] = \frac{V_2 \cdot I_4}{I_2} \quad (4-17)$$

$V[3,1]$ is the fundamental voltage at port three. $V[4,1]$ is the fundamental voltage at port 4. V_2 and I_2 are the fundamental voltage and current at the output of the main path. I_4 is the current at the output of AmpLoadPull component. The AmpLoadPull component is attached between port three (input) and port four (output) as shown in Figure 4-14. $V[3,1]$. Also the input of AmpLoadPull component is set to arbitrary constant voltage since non source pull is required. The working mechanism is that once the simulation starts, the AmpLoadPull component will probe and calculate the load impedance attached to it. Since the main path and the auxiliary path have identical load impedance environment, the impedance value determined by AmpLoadPull should be the same as the one attached to the main path. This impedance information will be passed to the main path from AmpLoadPull by setting proper output voltage and output current relationship as shown in equation (4-17). Such information will also be given to a variable which is referenced by DAC in order to look up the desired M and N values.

AmpLoadPull needs to work with a dataset in a pre-defined format. The data file format that ADS uses to pass simulation or setup information from schematic to schematic or from schematic to data display window. The dataset required by AmpLoadPull has to be generated by the LoadPullSetup component. Figure 4-17 shows the schematic setup for generating the dataset for AmpLoadPull. A simple linear component is used as a “dummy” component for generating the required dataset.

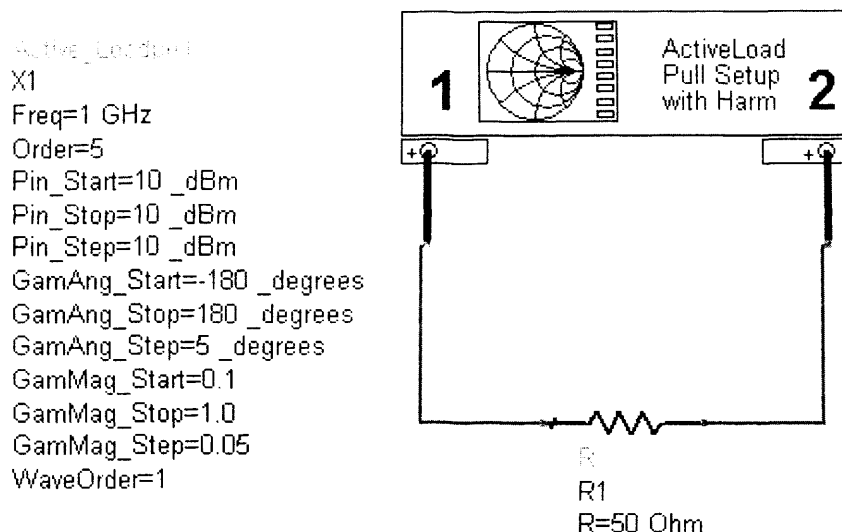
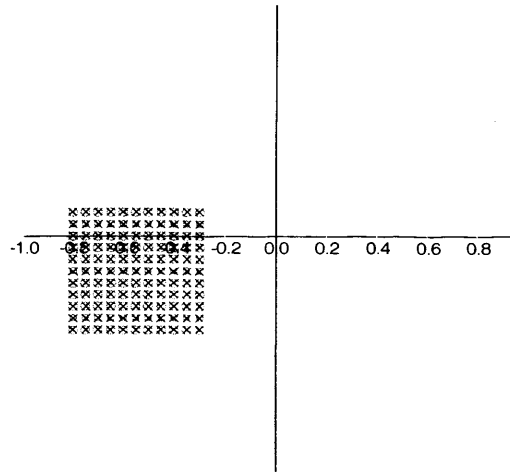


Figure 4-17: ADS schematic setup for AmpLoadPull dataset generation

As shown in Figure 4-17, GamAng and GamMag are the magnitude and phase of the load reflection coefficient. Pin is the input power. These three parameters are swept when generating the dataset. In order to make sure the AmpLoadPull component doesn't encounter any convergence problem during simulation, the swept range of load reflection coefficient is set to cover the whole Smith Chart. By contrast, the input power is set to a constant value because the input of AmpLoadPull component is set to a constant value during simulation, as shown in (4-16). In order to check the robustness of the AmpLoadpull component on load impedance determination, a couple of experiments were setup by generating the dataset with different parameter settings. For generating the dataset for AmpLoadPull component, initially the input power is set to 10 dBm; the whole range of load reflection coefficient was covered with a 73 by 19 grid; the impedance of the linear component in Figure 4-17 is set to 50 ohm. A series of datasets were then generated by changing the input power (from 10 dBm to 20 dBm), the resolution of the load reflection coefficient coverage, and the impedance of the linear component (from 50 ohm to 5 ohm and later 250 ohm). The series of datasets were then used with the AmpLoadPull component in a harmonic balance simulation sweep. Figure 4-18 shows a polar graph on which the comparisons between simulated load reflection coefficients and pre-defined load

reflection coefficients are made. The excellent agreement between the two grids indicates the AmpLoadPull component can adequately calculate the load impedance in simulation for quite a number of different dataset parameter settings.



(a)

Figure 4-18: Load impedance determined by the AmpLoadPull component during simulation. The simulated load impedance grids are identical when the datasets used in AmpLoadPull component were generated with different parameter settings.

In order to give clear idea on how the proposed approach interfaces with nonlinear simulator, the following block diagram is developed to illustrates the work flow of running simulation by using imported measurement data in ADS.

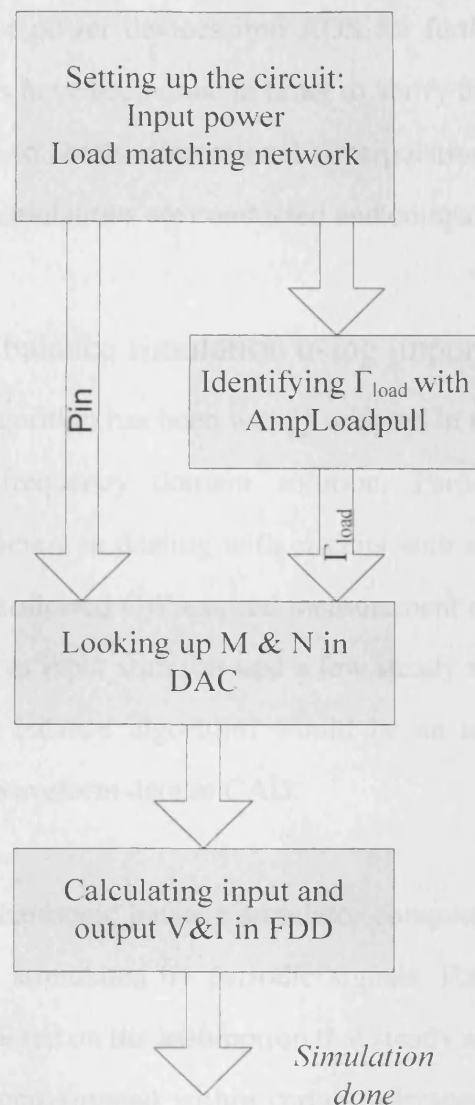


Figure 4-19: The work flow of using the direct waveform utilization approach to enable measurement-based CAD simulation.

4.5 Experimental validations of the direct utilization approach

As discussed earlier, the direct waveform utilization approach has been developed mainly for dealing with CW excited large signal waveform measurement data. The approach has been used to import nonlinear measurement data obtained on several of

CW-excited transistor power devices into ADS for further investigation or design purpose. Experiments have been done in order to verify the accuracy of the approach and its robustness to multi-dimensional interpolations via harmonic balance simulation. Various simulations are conducted and compared to measurement results.

4.5.1 Harmonic balance simulation using imported waveform data

Harmonic balance algorithm has been widely adopted in modern CAD simulator as a primary nonlinear frequency domain solution. Particularly harmonic balance algorithm is very efficient in dealing with circuits with a relatively few steady state sinusoids. Since the collected CW-excited measurement data has essentially a single steady state sinusoid as input stimulus and a few steady state sinusoids at harmonics as output, harmonic balance algorithm would be an ideal simulation method of choice to utilize the waveform data in CAD.

Generally speaking, harmonic balance simulator computes the steady state response of nonlinear circuits stimulated by periodic signals. Harmonic balance simulation theory is developed based on the assumption that steady state response of a nonlinear circuit that can be approximated within certain tolerance of accuracy by means of finite Fourier series does exist for given sinusoidal stimuli. The approximation tolerance is usually determined by designer, which would affect the simulation convergence time. During initial stage of harmonic balance simulation, the nonlinear circuit under test is analyzed and nonlinear algebraic equations associating all the circuit node voltages are abstracted. Since there is no standard formula for solving high order nonlinear algebraic equations analytically, harmonic balance adopts Newton's method to solve this type of nonlinear equations. Because of Newton's method, harmonic balance simulation is basically iterative – simulation starts by doing an initial guess and then solves the iteration equation repeatedly. The simulation is successfully finished until the iteration converges, i.e. Kirchhoff's Current Law (KCL) can be met at all the circuit nodes [67].

Harmonic balance algorithm is very handy for simulating nonlinear circuits or systems at RF or microwave frequencies. It is because at high frequencies, nonlinear behaviours of circuit or systems are usually characterized with frequency domain data. Harmonic balance has a number of advantages for dealing with high frequency circuit and system simulation. One advantage is that the system's final steady state behaviour can be efficiently obtained without the need to wait until the transient states passes. Another advantage of harmonic balance algorithm is that it can efficiently simulate the system characteristics which involve frequency responses at several different harmonic frequencies. For example, simulation of the output characteristics of a RF system which has fundamental frequency at 2.10GHz and five significant harmonics would require frequency domain analysis from 2.10GHz to 10.50GHz with a bandwidth of 8.40GHz. Time domain simulation algorithm would be computationally disadvantaged because it requires integration over an enormous number of periods of the highest frequency component, which usually makes the simulation time prohibitive in many practical cases. On the contrary, harmonic balance algorithm can be very efficient as only the frequency points at which non-zero spectra exist need to be taken into account in simulation [58].

The limitation of harmonic balance algorithm is that it cannot efficiently work with modulated signal excited circuits or systems. Harmonic balance requires that the stimulus has to be quasi-periodic – it must be possible to approximate the modulated signal as a sum of a number of discrete tones. The number of discrete tones necessary depends on the bandwidth of the modulated signal and required resolution for approximating the modulated signal with discrete tones. As the number of discrete tones increases, the amount of internal memory required for running harmonic balance simulation becomes excessive very quickly, especially for nonlinear systems where lots of mixing terms may be generated at the output of the system. Alternative simulation algorithm such as envelope transient method can be chosen to overcome the above limitations when dealing with modulated signal

simulation. Further discussion on envelope transient method [74] and its benefits to designers when integrating with the proposed direct waveform utilization approach will be given later in this chapter.

The implemented virtual device which can read in the measured waveform data has been setup for single tone harmonic balance simulation in order to present the measured waveform data in simulators. Figure 4-20 shows the schematic of the harmonic simulation setup. A power source is setup to stimulate the DUT. Ideal DC block and DC feed components are used in the circuit to separate the DC signals from RF ones. Current probes are used to measure the input and output current at all the harmonics of interest.

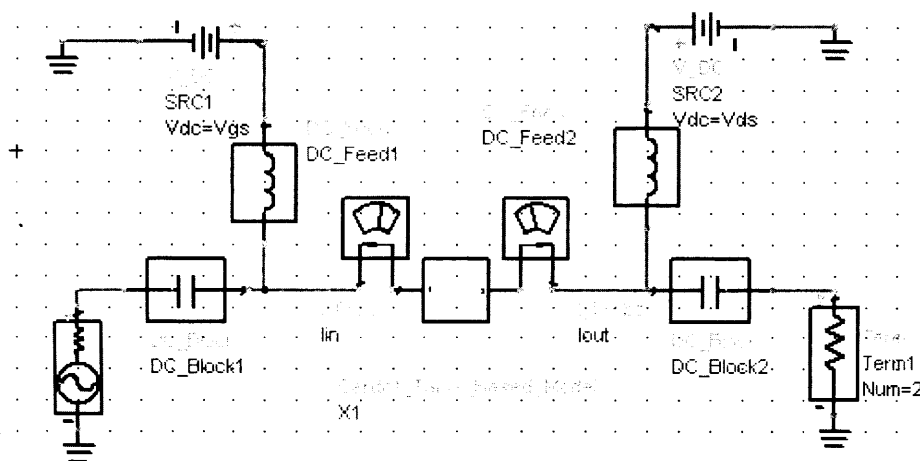


Figure 4-20: The harmonic balance simulation setup in ADS

In order to obtain the matched results between measurement and simulation, it is important to make sure the simulation environment is the same as the measurement environment. Particularly, the input and output impedance at higher harmonics need to be accurately selected in simulation. Usually during the CW-excited waveform measurements which don't require investigation of the harmonic tuning effects of device performance, only the fundamental load would be swept while the higher harmonics are passively terminated into 50 ohm, which is essentially a match. But in practise, ideal 50 ohm is not easy to achieve and the actual harmonic impedance

commonly drifts a few ohm away from 50 ohm. The actual impedance also varies from harmonic to harmonic. Therefore, it's necessary to specify the different load impedance values at individual harmonic frequencies in simulation. In the load impedance termination shown in Figure 4-20, a piecewise function is defined to set the load impedance at three harmonics. This kind of function can be realized in ADS using *if...else...* command.

$$Z_{Load} = \begin{cases} Z_{swept} & Freq < 1.1f_0 \\ Z_{2f_0} & 1.1f_0 < Freq < 2.1f_0 \\ Z_{3f_0} & 2.1f_0 < Freq < 3.1f_0 \end{cases} \quad (4-18)$$

In the equation, f_0 stands for fundamental frequency. $Z(2f_0)$ and $Z(3f_0)$ are the impedance at the second and third harmonics and are set to measured constant values. Z_{swept} is the fundamental load impedance and will be swept during simulation. The simulation range is normally set to match the measurement range. Note the load impedance is a complex number in nature and hence needs to be swept in real and imaginary part respectively. The input power sweep is setup to obtain the device dynamic performance over a large power range at each load impedance point.

A data demonstration template is developed in order to clearly display the simulation results by means of commonly used figure of merits, enabling model verification as well as device performance investigation in ADS. The template is developed by making use of the data display functionalities provided in ADS and consists of several pages with different pages emphasizing on different perspectives, which basically covers contour plots, waveform plots, large signal s-parameter plots and so on. Figure 4-21 gives an example page of the data display template.

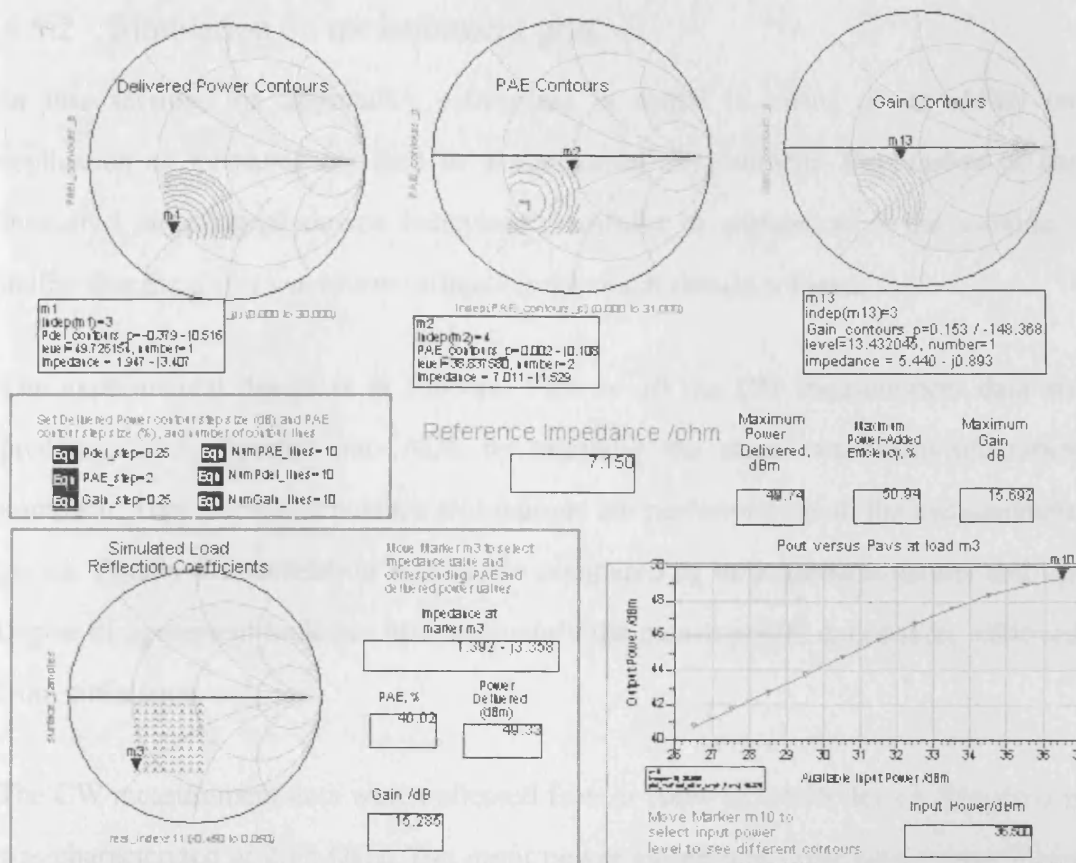


Figure 4-21: The data display page showing load pull characteristics

This page basically shows the load pull characteristics of DUT. The page offers some interactive control functions allowing the users to select the desired device load pull characteristics at a group of load reflection coefficients and power levels of interest. The Smith Chart on the bottom shows the simulated load-pull area. By changing the position of marker m3, the power characteristics at corresponding load reflection coefficient point can be displayed. The Pout versus Pin graph on the right provides the control of power levels through marker m10. With different location of m10, different contours corresponding to different power levels can be displayed.

The contour plots shown in this page include output power contours, gain contours and efficiency contours. Showing all these contours on one page would help RF circuit designers to trade off these contradicting figures of merit.

4.5.2 Simulation on measurement grid

In this session, the approach's robustness is tested in terms of its ability on replication of measurement data in a simulation environment. Replication of the measured large signal device behaviour faithfully in simulation is the minimum ability that the direct waveform utilization approach should achieve.

The experimental design is as follows: First of all the CW measurement data are processed and imported into ADS by adopting the direct waveform utilization approach. Then harmonic balance simulations are performed on all the measurement points. Finally the simulation results are compared to measurement results and the degree of agreement indicates how accurately the measurement data can be retrieved from simulation.

The CW measurement data were collected from a 100w LDMOS device. The device was characterized at 2.15 GHz. The input power sweep was done with a range from approx. 29dBm to approx. 40dBm with a step size of 1dB. The load reflection coefficient sweep was done with an 8 by 10 grid on the Smith Chart. The maximum output power variation within the load-pull area is about 3dB. This is generally sufficient for most types of PA design. The measurement data were then imported into the simulators using the DWLU approach. Two arbitrary load reflection coefficient points were picked up for simulation and verification. The results are then compared directly with measurements. Detailed comparisons between simulation and measurement at the two load impedance points are shown in the following four pages.

Based on the verification results it can be concluded that the simulation results match the measurement accurately regardless of the status on which the device is operating. No matter what the drive power or load reflection coefficient is, accurate simulation can always be performed to faithfully imitate the measured device behaviour. The excellent agreements between the measured and simulated device

characteristics prove the robustness of the proposed direct waveform utilization approach. The measured waveform data can indeed be accurately replicated in ADS simulation – given the simulation is conducted exactly under the same circumstance.

Simulation results obtained on the first load impedance point ($Z_{\text{load}} = 2.173 - j3.279$)

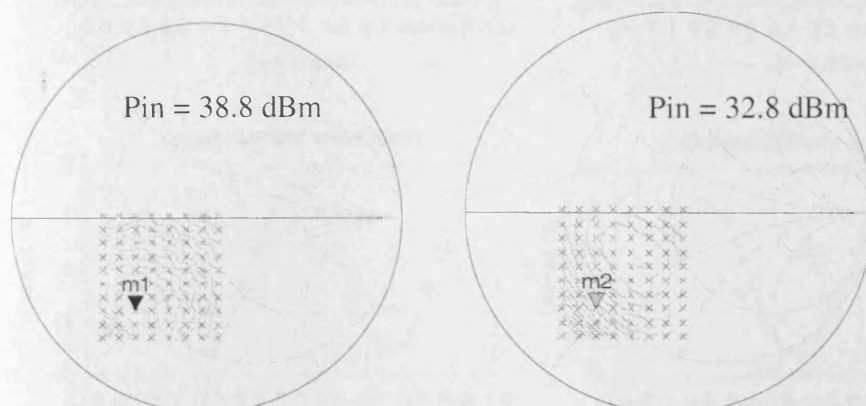


Figure 4-22: The load impedance point used in simulation is clearly marked on Smith Chart with a marker. The blue grid indicate the simulated load pull coverage – at each load impedance point on the grid, a series of harmonic balance simulation are performed by sweeping the input power of interest. The red contour-like plot on Smith Chart shows the output power contour. It can be identified that by changing the input power from 38.8 dBm to 32.8 dBm, the output power contours change significantly. The optimum impedance for maximum output power, which is shown as the centre of the contours also drift over a quite long distance. This is typical large signal behaviour – the performance of DUT varies depending on the power of stimuli.

The following graphs show the input and output voltage and current waveforms at two different power levels. Simulation is in good agreement with measurement.



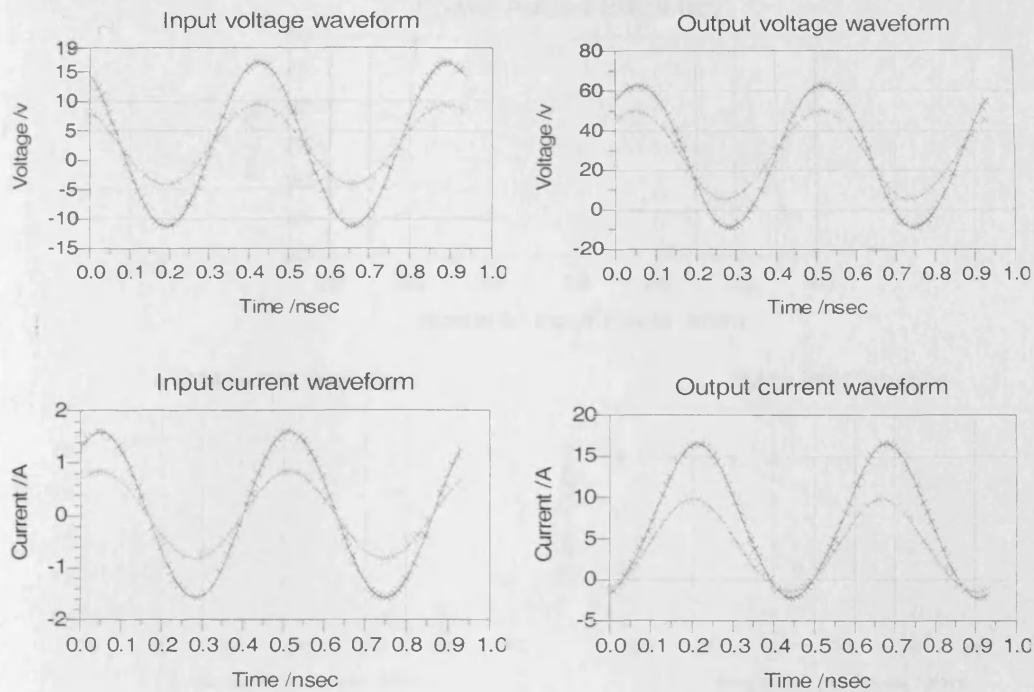


Figure 4-23: The input and output voltage and current waveforms are plotted against measured results. The comparisons were done on two power levels.

The graphs shown below in this page compare the simulated and measured power added efficiency, Gain profile, AM to PM distortion, and harmonic output power variation over the entire input power range at the load impedance point (2.173 – j3.279).

Great visual agreement in PAE indicates the measured DC characteristics can be retrieved as perfectly as RF characteristics. Gain profile shows that the DUT starts to saturate when the input power increases to about 33 dBm. An approximate 3 dB compression can be observed at the maximum input power. Good agreements observed when DUT is driven into high compression indicate that the proposed approach can capably handle nonlinear measurement data in simulation. Harmonic output power characteristics show the output power variation at each of harmonics. It can be seen that the measured device behaviour at all the harmonic of interest can be properly imported into CAD simulator.

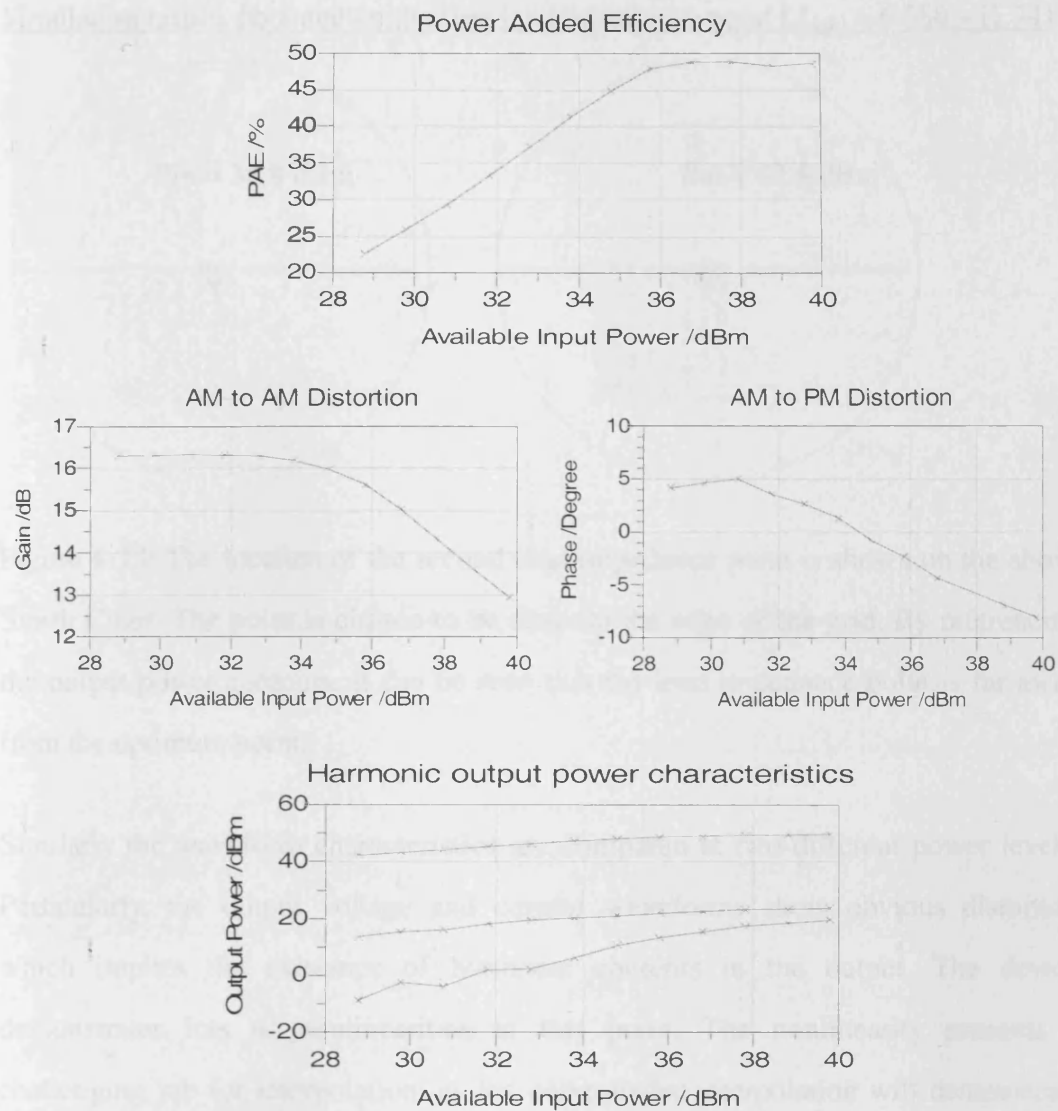


Figure 4-24: The verification of PAE, AM to AM, AM to PM and Pout characteristics.

Simulation results obtained on the first load impedance point ($Z_{load} = 6.580 - j1.741$)

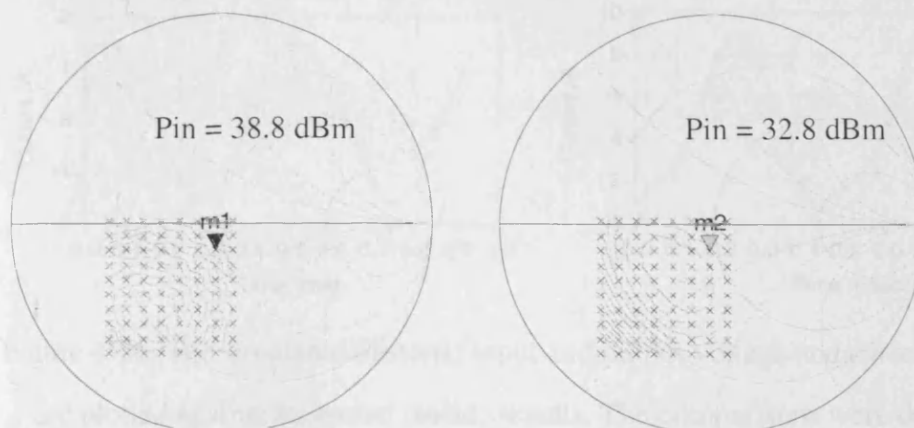
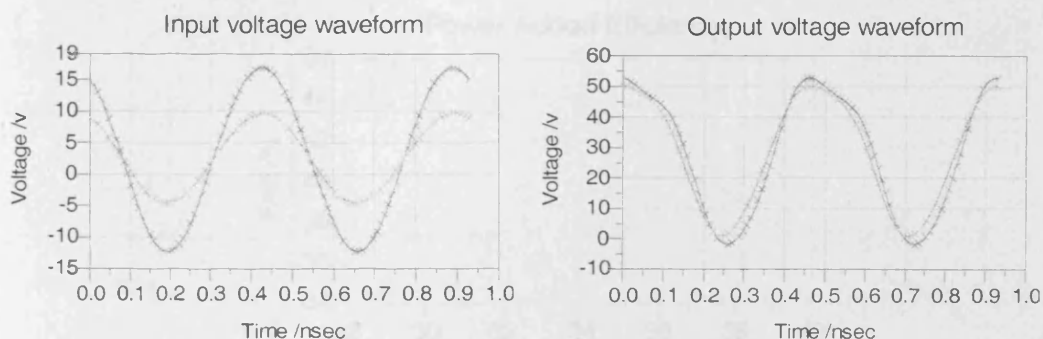


Figure 4-25: The location of the second load impedance point is shown on the above Smith Chart. The point is chosen to be close to the edge of the grid. By referencing the output power contours, it can be seen that the load impedance point is far away from the optimum point.

Similarly the waveform characteristics are compared at two different power levels. Particularly, the output voltage and current waveforms show obvious distortion which implies the existence of harmonic contents in the output. The device demonstrates lots of nonlinearities at this point. The nonlinearity presents a challenging job for interpolation, as the curve under interpolation will demonstrate much more complexities than a simple linear curve. Nevertheless, the simulation is still very accurate despite of the high level of nonlinearity at this load impedance point.



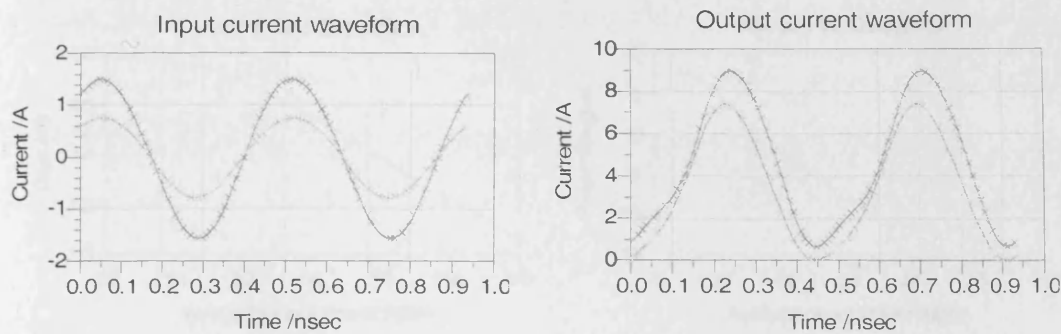


Figure 4-26: The simulated (**dotted**) input and output voltage and current waveforms are plotted against measured (**solid**) results. The comparisons were done on two power levels.

The variations of power added efficiency, AM to AM, AM to PM and harmonic output power as a function of input power are shown in this page. Since the load impedance point chosen for simulation is close to the edge of output power contours, heavy saturation behaviour can be observed. As shown in PAE graph, the curve starts to fall down around the input power of 33 dBm. It implies for input power range from 34 dBm to 40 dBm, little DC energy injected into the device contributes to the RF output power. Lots of the energy is wasted and dissipated as heat. AM to AM distortion, which is also referred to as gain profile, indicates deep compression occurs at this load impedance point.

Regarding the comparisons between measurement and simulation, consistent good accuracy can be observed on all of these figures of merit.



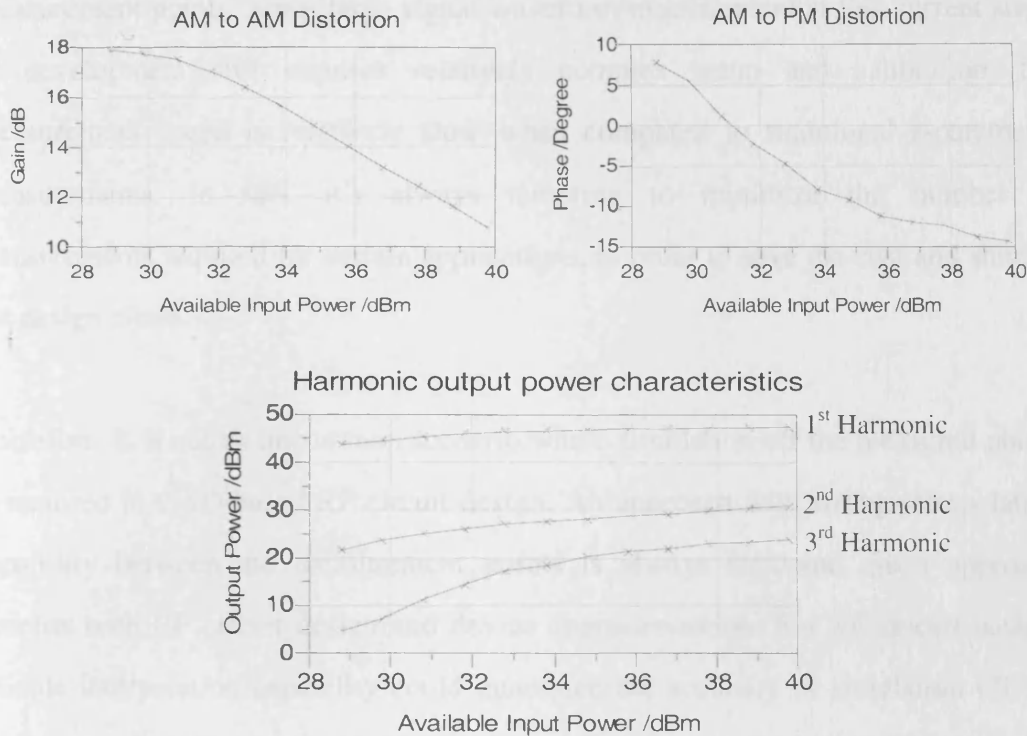


Figure 4-27: The verification of simulated (**dotted**) PAE, AM to AM, AM to PM and Pout characteristics to measured (**solid**) results.

4.5.3 Simulation off measurement grid I: interpolation

The simulation results shown in the previous section demonstrate the proposed DWLU approach can be effectively used to recover the measured device nonlinear characteristics in CAD simulation. Making the measurement waveform data available to designers in CAD can greatly facilitate CAD-based RF circuit design as the measurement data represent a reliable foundation based on which simulation and design work can be conducted with great confidence.

In practice, it's generally not possible or very time-consuming to characterize the DUT at all the load reflection coefficients or input power points of interest. In theory such a full scale measurement would require infinite number of measurements. By contrast, in most of practical cases, reasonable number of measurements is chosen which usually covers power range or load pull area of interest with discrete

measurement points. Since large signal waveform measurement at this current stage of development still requires relatively complex setup and calibration, the measurement speed is relatively slow when compared to traditional s-parameter measurements. In fact, it's always tempting to minimize the number of measurements required for certain applications, in order to save the cost and shorten the design circle.

Therefore, it is not an uncommon scenario where simulation off the measured points is required in CAD-based RF circuit design. An approach with strong interpolation capability between the measurement points is always welcome. Such approach benefits both RF circuit design and device characterization. For RF circuit design, reliable interpolation capability could guarantee the accuracy of simulation off the measurement grid and consequently enhance the chance of a successful first pass design. For device characterization, reliable interpolation would reduce the required number of measurements. A relatively sparse measurement grid would make the device characteristics available to designers much sooner.

Further experiment has been done to explore the potential of the direct waveform utilization approach in terms of simulating the device nonlinear behaviours at the points which are not covered in the measurement grid. In this and next sections, the robustness of the approach in terms of interpolation and extrapolation capability will be tested.

The experimental design is as follows: Since interpolation is multi-dimensional (i.e. the M & N parameters may be interpolated along the input power dimension and/or load reflection coefficient dimension), the interpolation accuracy in each dimension is tested respectively. Firstly the robustness in the dimension of load interpolation is tested. Rather than importing all the CW measurement data into CAD, only subsets of the data are processed and imported into ADS by adopting the direct waveform utilization approach. Each subset generally covers the same range in each dimension

but with different densities. Simulations using these subsets of data can show the robustness of the interpolation algorithm with respect to data density by comparing the interpolated results directly with the measurements.

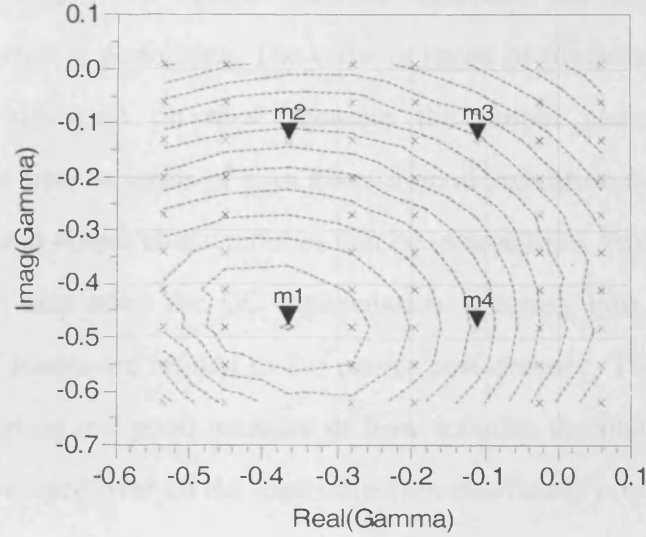


Figure 4-28 M1 to M4 indicate the interpolation points used for comparison. Output power contours are also shown to indicate the device status at these points. Please note that interpolation on m1 is especially challenging as the point represents the optimum impedance for maximum power. The dotted grid gives an example of measured grid having 60 points.

In order to evaluate the accuracy of the interpolation quantitatively, a set of metrics is developed in this work to calculate the errors of the simulations.

$$\Delta_{Pout} = \frac{\sum_{N_{load}} (Pout_{meas} - Pout_{sim})}{N_{load}} \quad (4-19)$$

$$\Delta_{Gain} = \frac{\sum_{N_{load}} (LinearGain_{meas} - LinearGain_{sim})}{N_{load}} \quad (4-20)$$

$$\Delta_{PAE} = \frac{\sum_{N_{load}} (PAE_{meas} - PAE_{sim})}{N_{load}} \quad (4-21)$$

$$\Delta_{AMtoPM} = \frac{\sum_{N_{load}} (AMtoPM_{meas} - AMtoPM_{sim})}{N_{load}} \quad (4-22)$$

In the metrics, N_{load} is the number of load reflection coefficients at which the harmonic simulation is performed. The error in terms of fundamental output power gives a good indication on how accurate the output characteristics can be interpolated. The error in terms of gain gives a good indication on how accurate the combined input and output characteristics can be interpolated. Furthermore, the error in terms of PAE also takes the DC interpolation accuracy into consideration. All above figures of merits are related to the power performance. The error in terms of AM to PM distortion is a good measure of how accurate the phase interpolation is. The errors are averaged over all the load reflection coefficient points.

The input power is fixed at 39.8 dBm. The load reflection coefficient sweep was done with an n-point grid on the Smith Chart. For different densities, different numbers of measurement data were collected and imported into ADS to run a fixed number of interpolations. The interpolation results were compared to independent measurement results. The average discrepancies regarding different number of measured points used in the simulations, in terms of output power, gain, power added efficiency and AM to PM distortion, are calculated using the above metrics and shown below.

N-point grid	ΔP_{out} (dB)	$\text{Log}(\Delta \text{Gain})$	ΔPAE (%)	ΔAMtoPM (°)
60 pts	0.036	-1.303	0.380	0.279
50 pts	0.046	-1.203	0.377	0.280
40 pts	0.061	-0.693	0.645	0.564
30 pts	0.125	-0.333	1.358	0.570

Table 4-2: Interpolation robustness in relation with load-pull grid density

The small errors observed in the above table indicate that the interpolation regarding the load surface over the Smith Chart is quite robust. None of the four types of error is much more significant than the others, which verifies that all the typical device characteristics from input to output, from DC to RF and from power to phase can be accurately interpolated. The error increases smoothly as the number of measurement points used in simulation is decreased until n equals to 40. The error increases dramatically when only 40 of 80 measurements points are used in simulation. It may suggest that the number of measurement points used in simulation becomes insufficient for maintaining reasonable accuracy of interpolation.

In a similar manner, the robustness of the interpolation in dimension of output power levels is investigated. For this purpose the $(m \times m)$ -grid with the highest density in table 4-2 has been chosen and measured for a 10dB sweep of input signal drive. To obtain a relationship between the input power step size and simulation accuracy the number of measured points over this dynamic range in the DWLU table was varied and again interpolated simulation results compared to a fixed set of alternative drive level measurements. The results obtained are shown in table 4-3.

N-point grid	ΔP_{out} (dB)	LOG($\Delta GAIN$)	ΔPAE (%)	$\Delta AMtoPM$ (°)
11 pts	0.011	-1.246	0.103	0.130
8 pts	0.025	-0.818	0.211	0.160
5 pts	0.032	-0.725	0.374	0.253

Table 4-3: Interpolation robustness in relation with dynamic power density

It would be worthy to point out that the 11 power points are used to cover a 14 dB dynamic range from 26dBm to 40dBm. As the number of points reduces to 7, it means there is only 1 measurement power point in every 2dB, which is quite a sparse grid. Despite this relatively sparse grid, the interpolation accuracy is surprisingly

robust and very impressive interpolation results are obtained, as observed in table 4-3. For example, when the interpolations were done with only 8 measurement power points, the average error in output power at the interpolation points is less than 0.03 dB.

The minimum density required for accurate simulation depends on both the device characteristics and the interpolation algorithm used in the simulation. A well behaved device obviously requires less number of points to maintain the accuracy within certain area. Also, since the device under test usually demonstrates lots of high order distortions in large signal excitation, a robust nonlinear interpolation algorithm which can approximate high-order nonlinearities is of great help to the accuracy of large signal simulations. The 100w LDMOS device characterized here was driven into deep compressions during measurements. The simulation results shown in Table 4-2 and 4-3 indicate good interpolation accuracy can be achieved even if only half of measurement data points were used in the simulations.

Additional verification results from the proposed modelling approach can be found in Appendix.

Interpolation results shown in this section were obtained with Cubic Spline algorithm. ADS also offers alternatives such as cubic and linear interpolation methods. The reasons for choosing Cubic Spline algorithm over the others have been discussed in section 4.4. The reasoning can also be confirmed in following experiments where comparisons are made to evaluate the reliability of these interpolation approaches.

Similar to the interpolation tests done on Cubic Spline algorithm, cubic and linear algorithms are tested by doing interpolations with different number of imported data points and comparing the results to measurements. The algorithms are also tested in two dimensions respectively – namely input power dimension and load reflection coefficient dimension. The tests were also run on the same 100w LDMOS device.

Firstly the test results on load pull interpolation are plotted in Figure 4-29 and Figure 4-30.

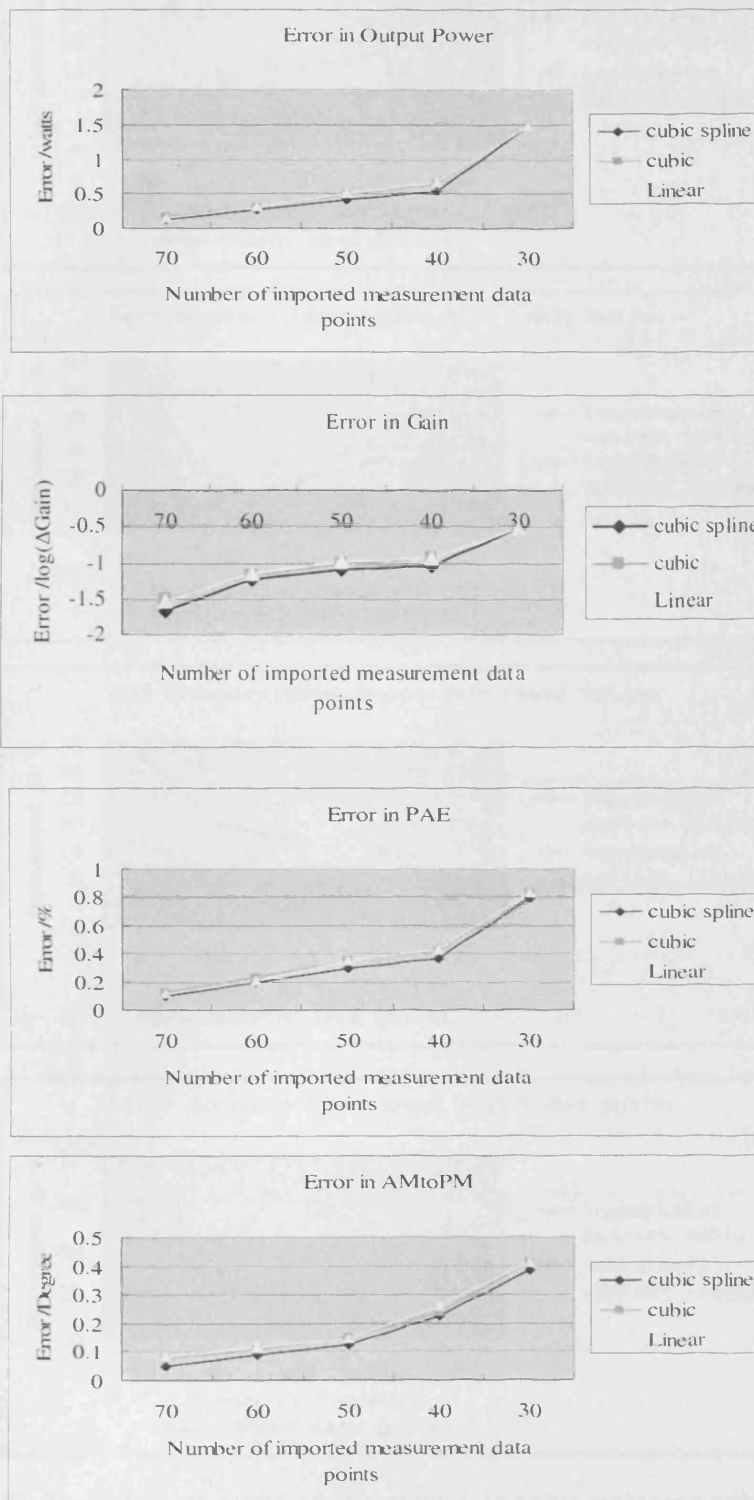


Figure 4-29: The comparisons of simulation accuracy by using Cubic Spline, cubic and linear interpolation algorithms. The comparisons are made in terms of P_{out} , gain, PAE and AM to PM distortion.

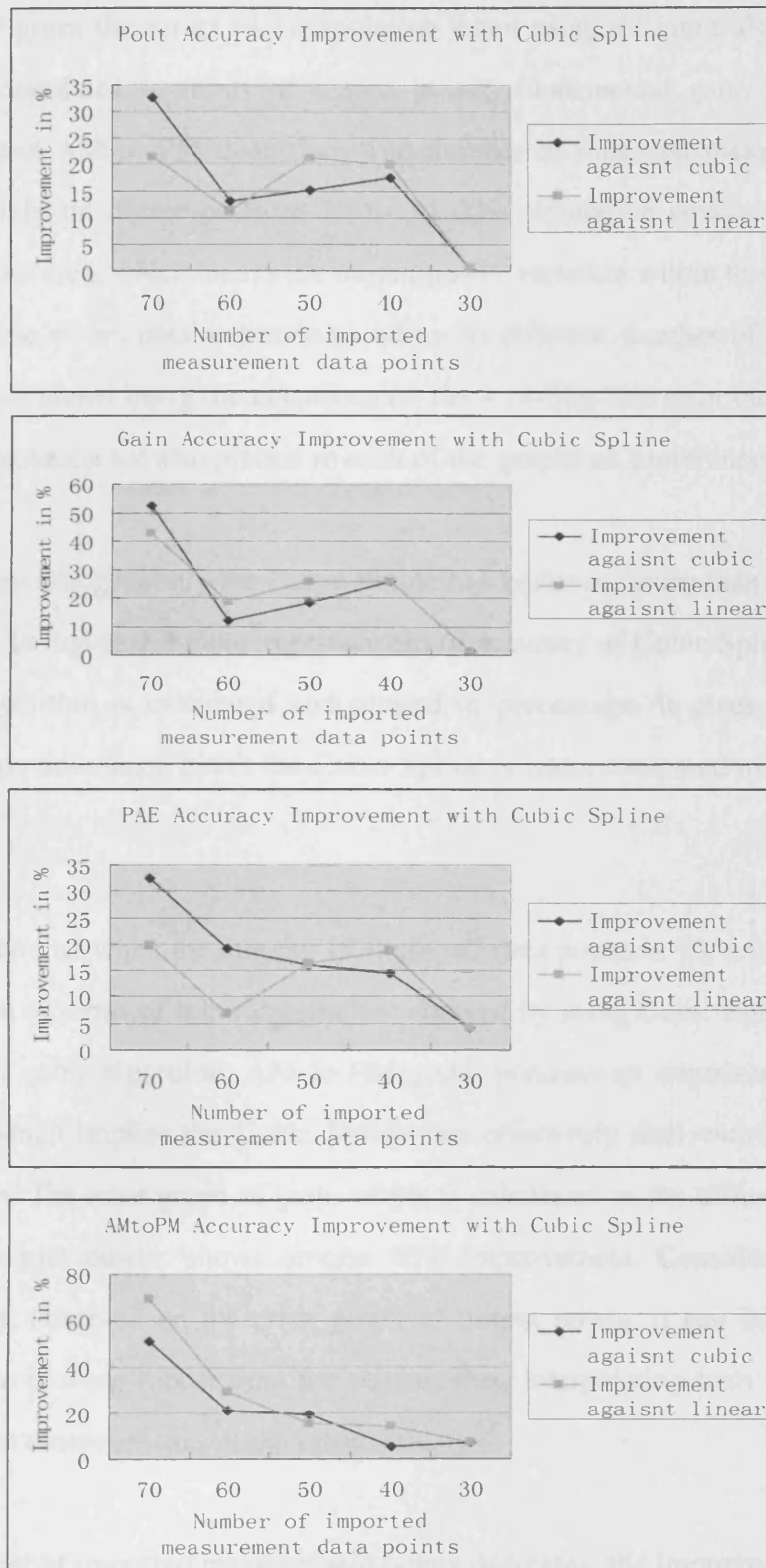


Figure 4-30: The improvements of accuracy gained by replacing cubic or linear interpolation with Cubic Spline interpolation. The improvement is quantified as the relative percentage.

Figure 4-29 gives the errors of interpolation when using different algorithms. The errors are calculated in terms of output power, fundamental gain, power added efficiency, and AM to PM distortion. The number of imported measurement data used in simulation decreases from 70 to 30. The simulation covers a 3dB output power contour area, which means the output power variation within this area is about 50 Watts. The errors obtained corresponding to different number of measurement points are calculated using the equations (4-19) – (4-22). The error curves of Cubic Spline interpolation are also plotted in each of the graphs as a reference.

It can be seen that generally the Cubic Spline has better accuracy than the other two approaches. In Figure 4-30, the improvement of accuracy of Cubic Spline over cubic or linear algorithm is calculated and plotted in percentage. It gives much clearer impression on how much better the Cubic Spline is when compared to the other two algorithms.

It can be seen that when the number of imported data points is 70, a minimum 20% improvement in terms of accuracy can be achieved by using Cubic Spline to replace the linear or cubic algorithm. AM to PM graph presents an improvement of more than 50% which implies the Cubic Spline can effectively deal with phase related interpolation. The error graph of gain, which is calculated as the difference between input and output power, shows around 40% improvement. Considering the 20% improvement observed on the error graph of output power, it can be derived that Cubic Spline is more robust than the others when interpolating both the input and output device characteristics at the same time.

As the number of imported measurement points decreases, the improvement of using Cubic Spline becomes smaller and smaller. Especially when the number of imported points is 30, there is almost no significant improvement at all – the maximum improvement is less than 10%. The robustness of interpolation is actually influenced by two major factors. One is the interpolation algorithm. The other is the number of

known data points based on which the interpolation is conducted. If the number of data points becomes very small, there may be a lack of sufficient information about what the curve looks like. In such case, no matter what type of interpolation algorithm is used, it's very difficult to recover the missed information at the data point in between by interpolation. Regarding the results shown in Figure 4-29 and Figure 4-30, when the number of imported points is 30, the measurement is very sparse and there is large gap between these known measurement data points. As a result, any type of interpolation scheme would become ineffective.

After testing the load-pull interpolation capabilities of these three algorithms, tests are performed to investigate the power interpolation capabilities of them. By choosing different numbers of imported measurement power points, the device performance variation against input power is interpolated. The interpolation is done by using different interpolation algorithms respectively. The errors of interpolation when using different algorithms are recorded and plotted in Figure 4-31 and Figure 4-32.

Figure 4-31 compares the accuracy of simulation using different interpolation algorithms in terms of output power, gain, PAE and AM to PM distortion. Since the primary cause of nonlinearity in an active transistor device is the magnitude of stimuli, the nonlinearity of device behaviour demonstrates clear variability when the device's characterized as a function of input power. As a result, the interpolation along the input power axis is more demanding and a quality interpolation algorithm with robust nonlinear interpolation ability is clearly favoured. Figure 4-32 calculates the improvement of accuracy for using Cubic Spline instead of cubic or linear interpolation. Very impressive improvement as high as 400% over linear algorithm can be observed. The Cubic Spline and cubic interpolations demonstrate their superiority as nonlinear interpolation algorithm over linear interpolation clearly. The only exception seems to come from the AM to PM simulation – no obvious superiority can be observed from any of the three interpolation algorithms. It's

probably because the phase related characteristics demonstrates less nonlinearity than the magnitude related quantities. Using a high order interpolation algorithm than what it actually needs may introduce over-fitting errors and consequently deteriorate the accuracy of phase related simulation.

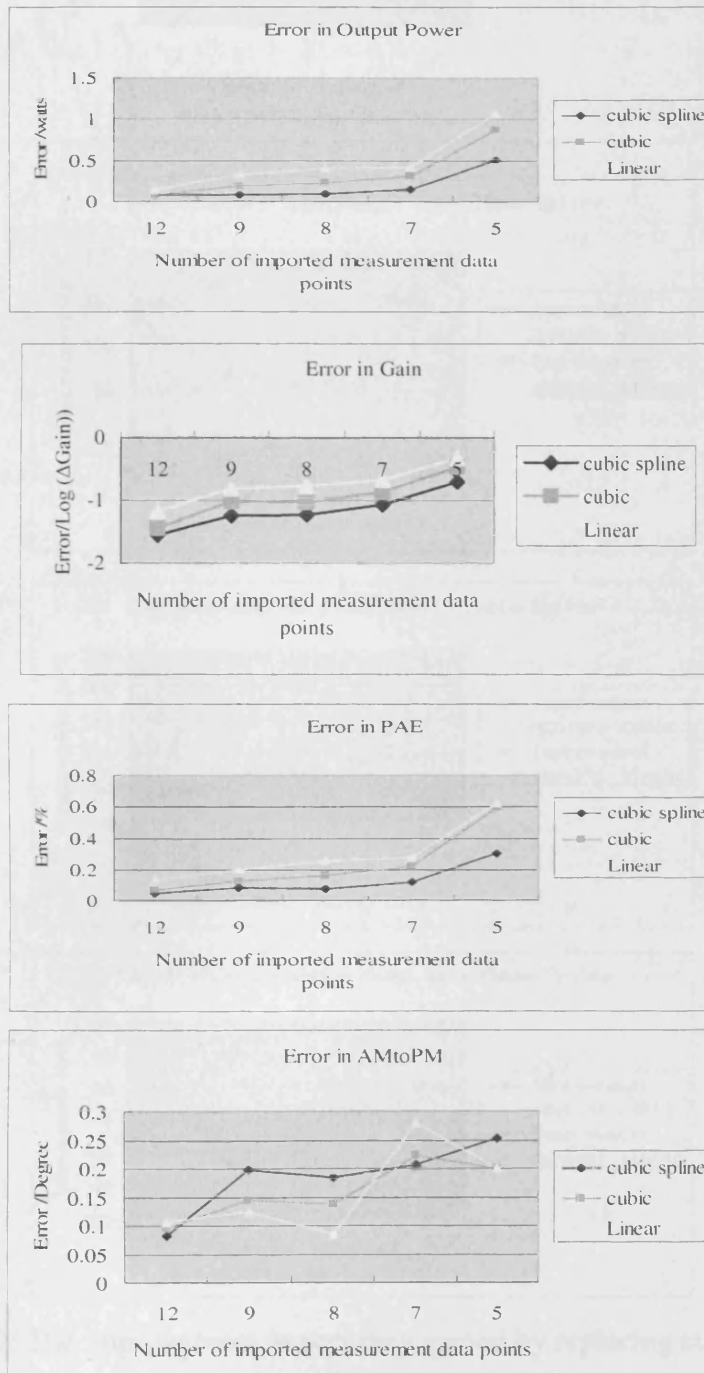


Figure 4-31: Simulation accuracy when using Cubic Spline, cubic and linear interpolation algorithms in terms of Pout, gain, PAE and AM to PM distortion.

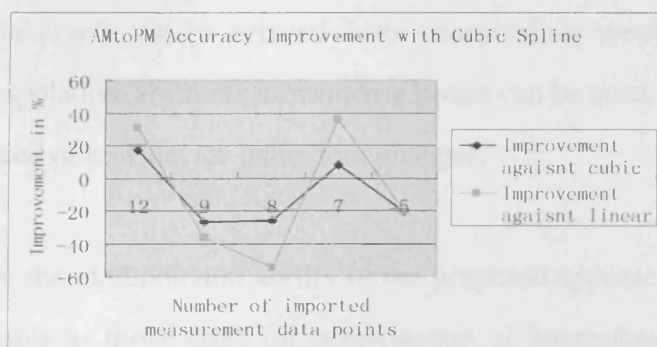
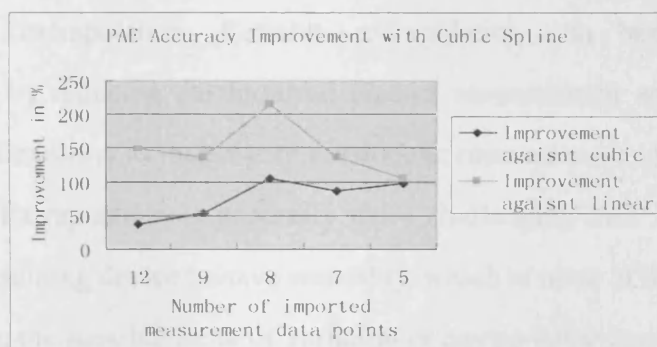
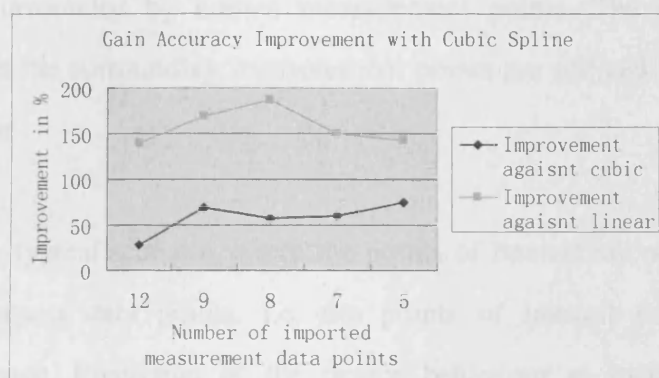
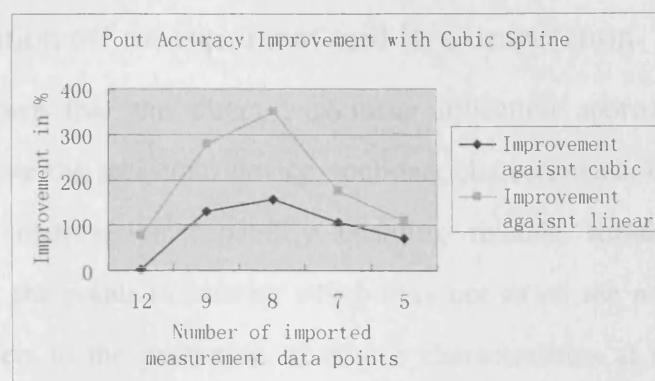


Figure 4-32: The improvement in accuracy gained by replacing cubic or linear interpolation with Cubic Spline interpolation. The improvement is quantified as the relative percentage.

4.5.4 Simulation off measurement grid II: extrapolation

It has been shown that the direct waveform utilization approach cannot only accurately replicate the measured device nonlinear characteristics in CAD, but also provide robust interpolation capability enabling reliable simulation of device characteristics at the points of interest which may not sit on the measurement grid. Interpolation refers to the prediction of device characteristics at the points which generally are surrounded by known measurement points. The device measured characteristics at the surrounding measurement points are utilized to interpolate the device behaviour.

There is another typical scenario where the points of interest are not surrounded by known measurement data points, i.e. the points of interest reside outside the measurement space. Prediction of the device behaviour at such points is often referred to as extrapolation. Reliable extrapolation can benefit the device characterization by reducing the required size of measurement area. It also offers designers more flexibility to investigate outside the measurement space, seeking best trade-off point. Extrapolation is generally more challenging than interpolation. For interpolation, assuming device behave smoothly, which is most of the case for power transistor device, the possible trend of variation of device behaviour at unknown data points can be relatively easily derived from surrounding measured points. By contrast, for extrapolation, since no surrounding points can be used, it's very difficult to find out the trend of how device behaviour changes.

In this subsection the extrapolation ability of the proposed approach is investigated. Similar experiments as those done for investigation of interpolation are designed. Extrapolation ability is tested on load and input power dimensions separately. Figure 4-33, 4-34, and 4-35 give the extrapolation results when the input power is 39.8 dBm.

Number of imported load impedance points is 70; input power is 39.8 dBm

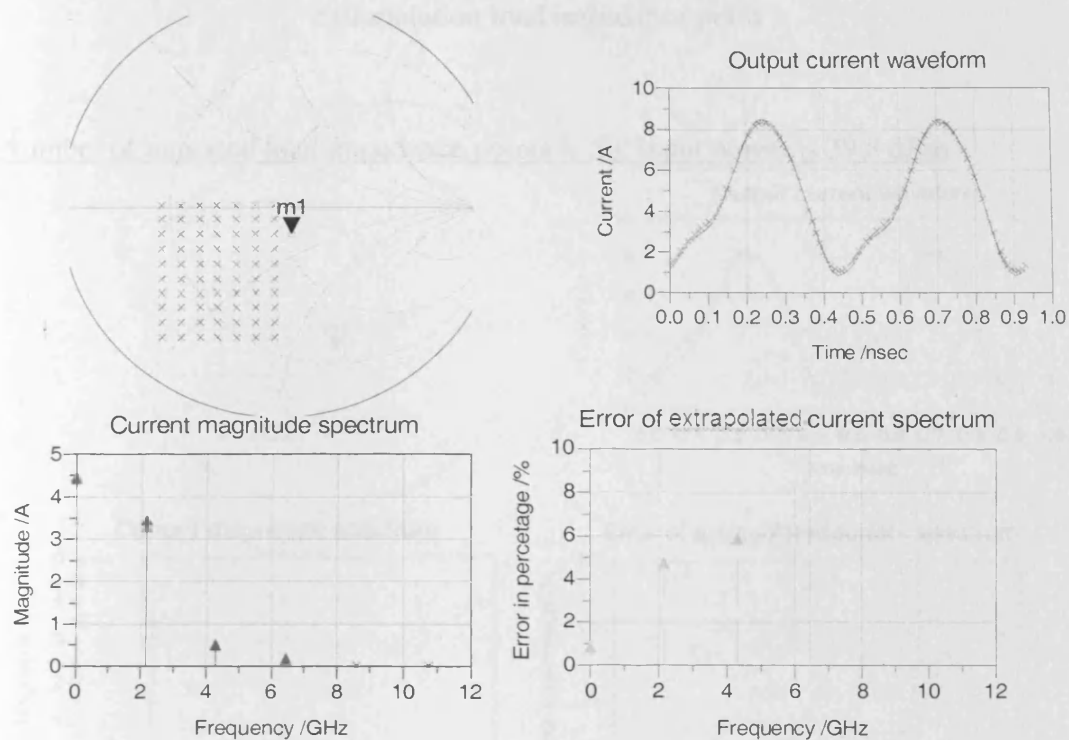


Figure 4-33: The output current waveform and frequency domain spectrum at extrapolation load impedance point

Number of imported load impedance points is 60; input power is 39.8 dBm

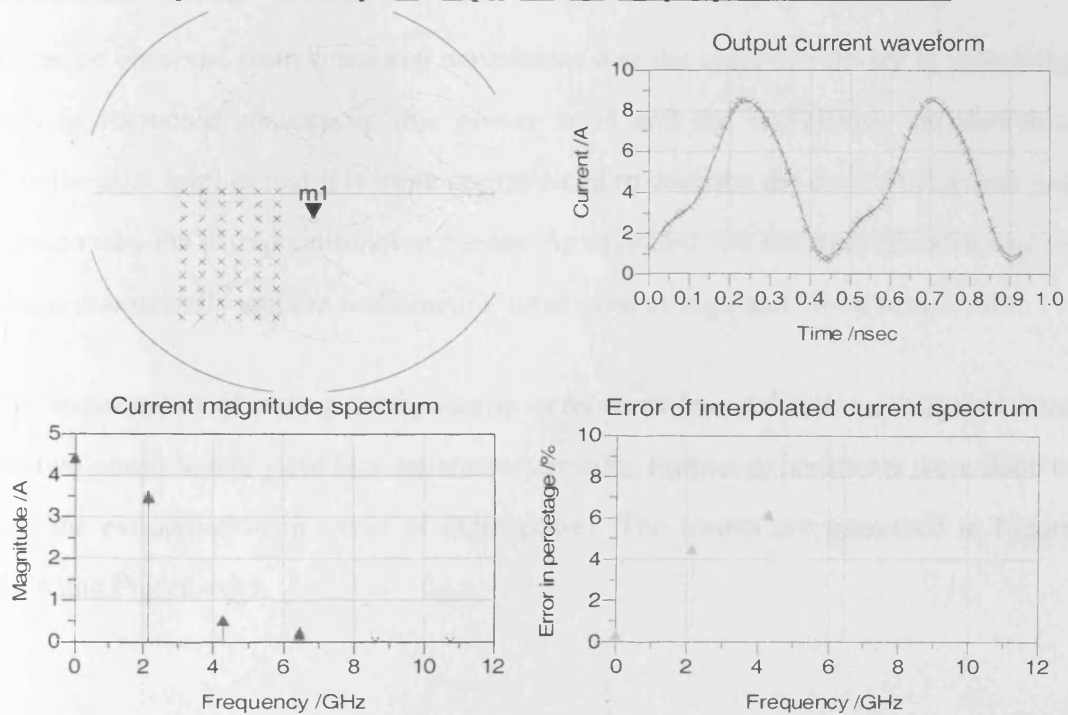


Figure 4-34: The output current waveform and frequency domain spectrum at

extrapolation load impedance point

Number of imported load impedance points is 50; input power is 39.8 dBm

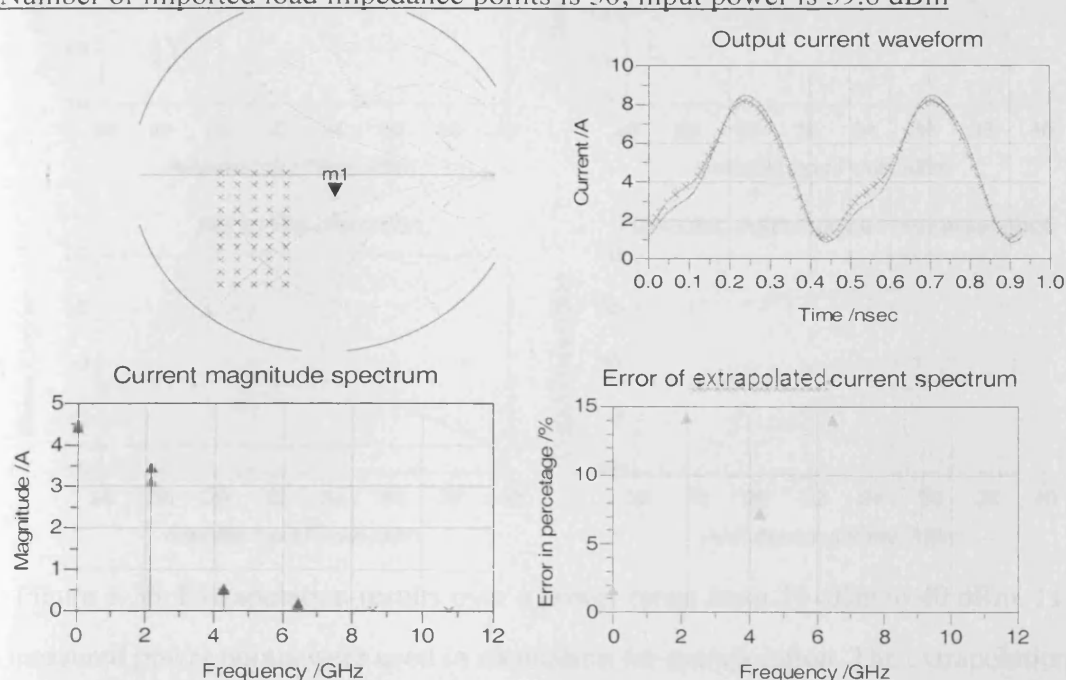


Figure 4-35: The output current waveform and frequency domain spectrum at extrapolation load impedance point

It can be observed from simulated waveforms that the transistor device is generating lots of harmonic contents at this power level and the waveforms are distorted. Nonlinearity implies that it is more complicated to describe the device behaviour and hence make the extrapolation even harder. As expected, the accuracy of extrapolation drops dramatically and the fundamental error goes as high as 15% in some cases.

The experiments of testing extrapolation in terms of load reflection coefficient done on two power levels yield less satisfactory results. Further experiments were done to test the extrapolation in terms of input power. The results are presented in Figure 4-36 and Figure 4-37.

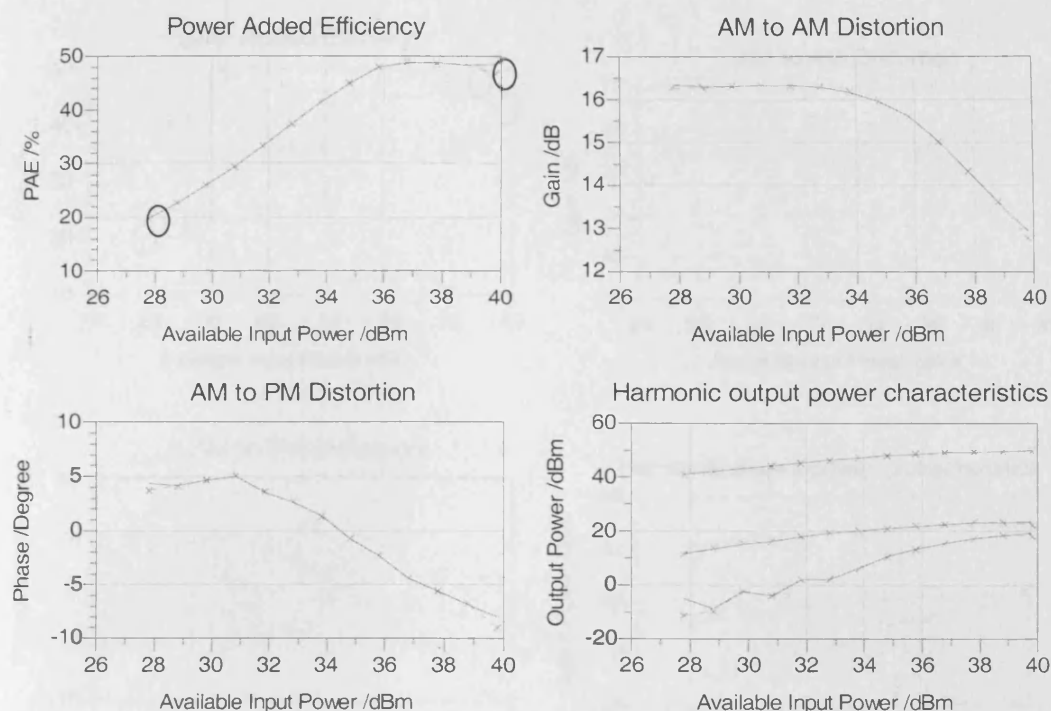


Figure 4-36: Extrapolation results over a power range from 28 dBm to 40 dBm. 11 measured power points were used in simulation for extrapolation. The extrapolation points are marked on PAE plot.

In Figure 4-36, the extrapolation results obtained when 11 measured power points are used in simulation are shown. As shown in Figure 4-36, simulations are done on 13 power points with the input power starting from 28.8 dBm and ending at 39.8 dBm. The first and last power points, i.e. 28.8 dBm power point and 39.8 dBm power point, don't belong to the 11 measured power points. The device characteristics at these power points were extrapolated.

It can be found that large discrepancy can be clearly spotted at these extrapolation points on all the figures. The discrepancy at the lower power point (at 28.8 dBm) is generally smaller than the discrepancy at the higher power point (at 39.8 dBm) due to the higher nonlinearity at the higher power point.

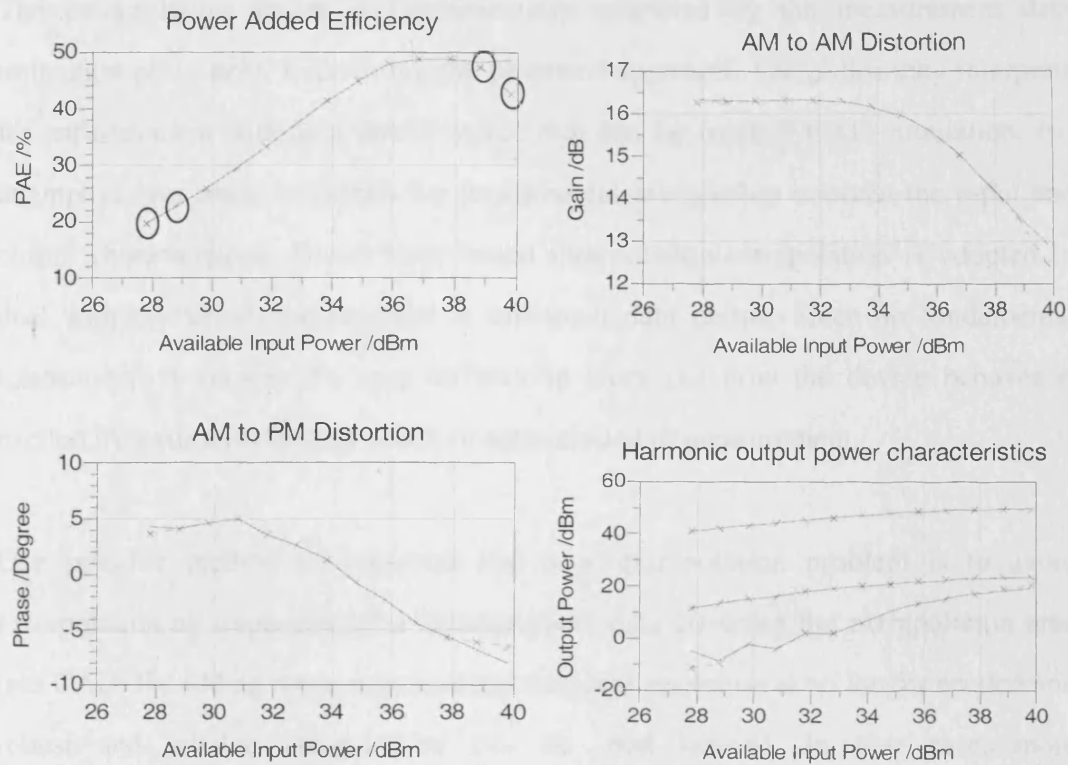


Figure 4-37: Extrapolation over a power range from 28 dBm to 40 dBm. 9 measured power points were used in simulation for extrapolation. The extrapolation points are marked on PAE plot.

Figure 4-37 presents the extrapolation results obtained from a scenario where more challenging extrapolation is required. Two less power points were used in simulation. The extrapolation power points are marked on these graphs. It can be seen that the discrepancy increases quickly as the distance between the extrapolation point and measurement points increases.

The extrapolation performance of the proposed approach is a bit disappointing when compared to the impressive interpolation performance demonstrated earlier in this chapter. Light extrapolation just outside the measured load pull area can still present reasonable results. But when greater extrapolation is required on the dimension of input power or load reflection coefficient, the accuracy normally degrades very quickly.

The extrapolation ability is fundamentally restricted by the measurement data utilization philosophy underlining the proposed approach. The philosophy interprets the measurement data as a direct source that can be used in CAD simulation. No attempt is ever made to extract the fundamental relationship between the input and output characteristics. Direct table based interpolation/extrapolation is adopted to deal with the simulation required at unknown data points. Since no fundamental relationship is known, it's very difficult to work out how the device behaves if excited by a stimulus or load which is not included in measurement.

One possible method to overcome the poor extrapolation problem is to avoid extrapolation by importing more measurement data covering the extrapolation area into CAD. By adding more measurement data, extrapolation is no longer needed and robust and reliable interpolation can be used instead. In that case, more measurements are needed which may increase the device characterization time. In fact, increase in device characterization time may not be necessary if the characterization can be planned carefully beforehand. For example, certain application may require the device to be characterized within a larger than normal load pull range, say, a 4 dB output power variation area rather than 3 dB. The measurement grid with same number of measurement load pull points can be used for both 3 dB and 4 dB characterization. The difference is for 4 dB load pull contour characterization, the resolution would be lower due to the fact that same number of points is used for covering larger area. The degradation of accuracy, when using the measurement data in CAD simulation, due to coarser measurement grid is compensated by the strong interpolation capability of the proposed approach. It can be seen that by using the same amount of measurement data points a larger load pull area is covered without losing much accuracy of CAD simulation, as long as only interpolation but no extrapolation is required.

4.5.5 Testing for technology-independent

In theory, the approach should be able to import large signal waveform measurement data obtained on any type of transistor devices into CAD tools. In order to prove it is technology-independent experimentally, the approach has been tested on a number of different types of devices. All these power transistor devices were driven into compression in measurements in order to challenge the robustness of interpolation. Additional to the simulation results on 100w LDMOS presented in previous sections, the interpolation results obtained on 4w GaAs and 5w GaN devices are illustrated and compared to measurements respectively.

The 4w GaAs device was characterized at 0.9 GHz. In the characterization all higher harmonics than the fundamental were terminated into 50 ohm and only fundamental load pull sweep was performed. The sweeping range of load covered the area containing 121 points in Smith Chart but only 50 points were used for simulation allowing the use of the remaining data set for verification purposes. Two different load impedance points were selected to run the harmonic balance simulation and the resultant waveforms are shown.

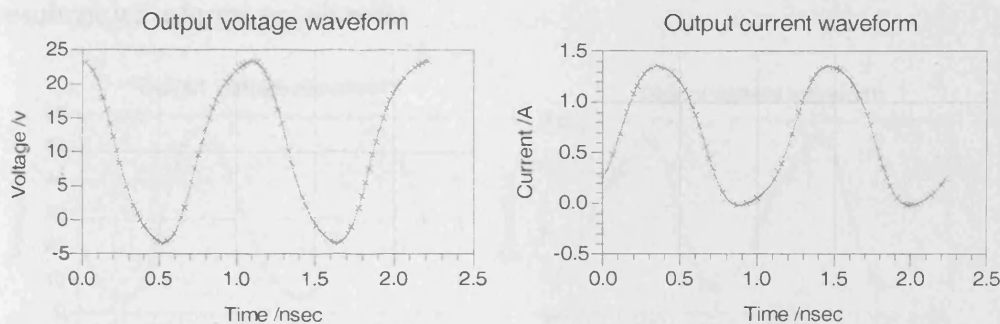


Figure 4-38: Simulated and measured output voltage and current waveforms on load impedance of $14.103-j12.821$. The impedance point is covered in the 50 points which have been imported into ADS.

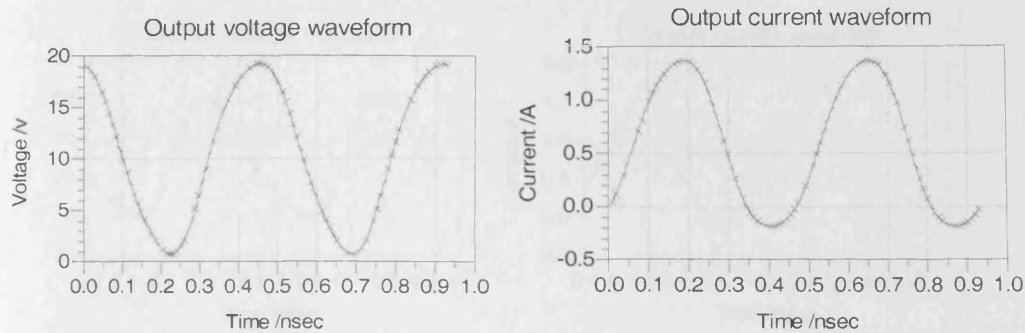


Figure 4-39: Interpolated and measured output voltage and current waveforms on load impedance of $10.109-j5.464$. The impedance point is not covered in the 50 points which have been imported into ADS.

The 5w GaN device was characterized at 2.15 GHz. In the characterization all higher harmonics than the fundamental were terminated into 50 ohm and only fundamental load pull sweep was performed. The sweeping range of load covered the area containing 36 points in Smith Chart but only 18 points were used for simulation allowing the use of the remaining data set for verification purposes. Two different load impedance points were selected to run the harmonic balance simulation and the resultant waveforms are shown.

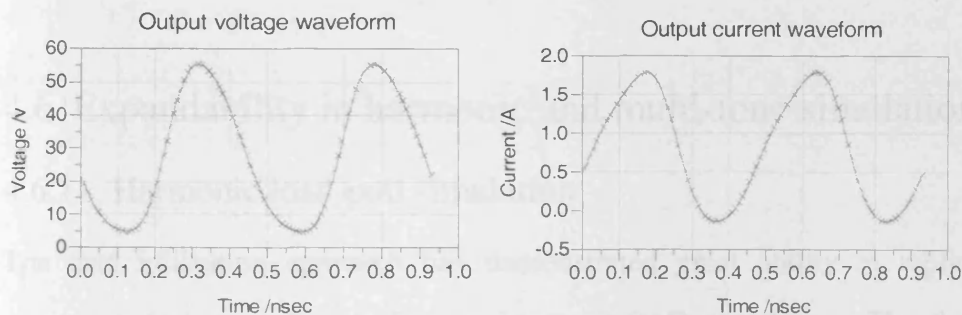


Figure 4-40: Simulated and measured output voltage and current waveforms on load impedance of $21.655+j15.808$. The impedance point is covered in the data points which have been imported into ADS.

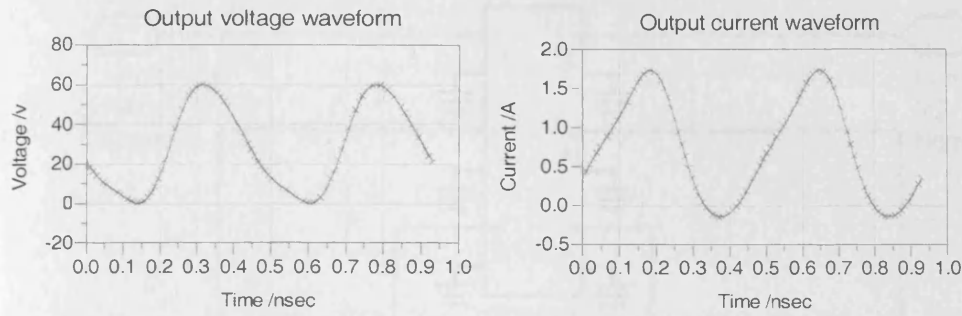


Figure 4-41: Interpolated and measured output voltage and current waveforms on load impedance of $23.397+j23.147$. The impedance point is not covered in the 18 points which have been imported into ADS.

The good agreements between measured and interpolated results verify that the direct waveform utilization approach can be applied for many types of device technologies. All these experimental results show that with respect to interpolation the approach is very robust even when the measurement data are only collected on a coarse grid. Furthermore, limited extrapolation slightly outside the measurement space also demonstrates decent level of accuracy. The approach proves itself to be a viable tool for utilization of CW waveform data in CAD-driven PA design.

4.6 Expandability in harmonic and multi-tone simulation

4.6.1 Harmonic load-pull simulation

The data utilization approach has demonstrated great ability in replicating the measured device nonlinear characteristics in CAD simulation. The fundamental load-pull contours can be accurately produced in CAD. The approach can also be expanded easily to do harmonic load-pull simulation if necessary.

Figure 4-42 shows the ADS implementation enabling second and third harmonic load-pull simulation.

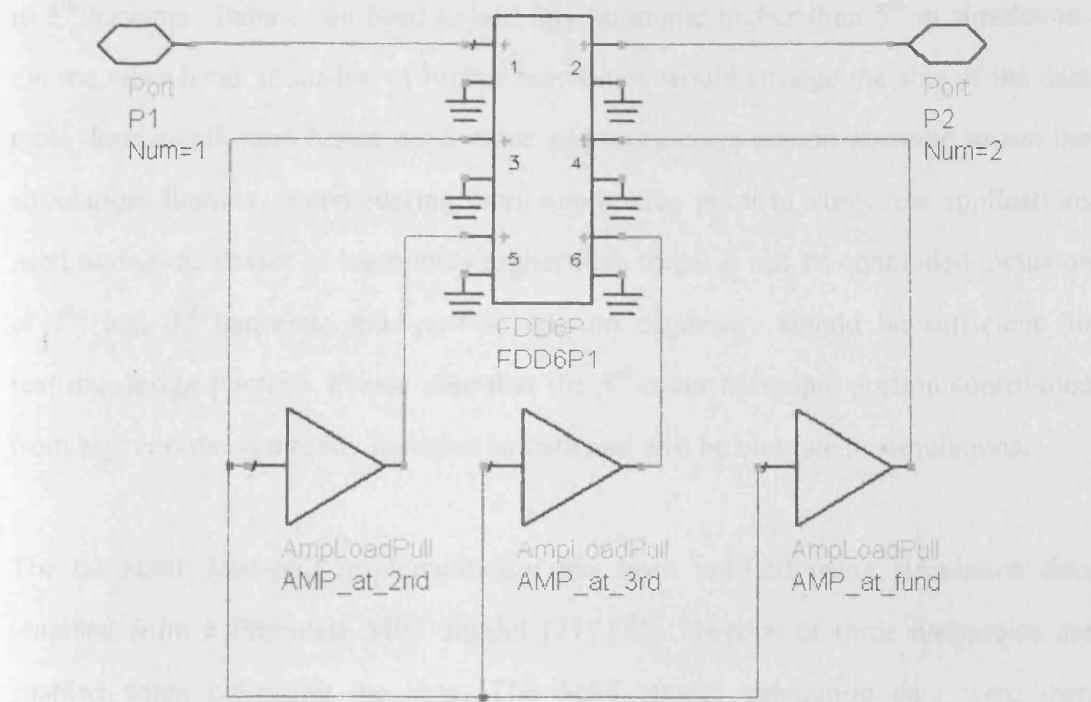


Figure 4-42: The ADS schematic setup enabling fundamental, second and third harmonic load impedance looking up simultaneously

Three AmpLoadPull components are used in the setup and each of them is responsible for determining the load impedance at one harmonic frequency. The outputs of the three AmpLoadPull components are connected separately to three different ports. On each of the ports equations similar to (4-17) are setup to couple the branch with the appropriate harmonic components at the port 2. The inputs of the three AmpLoadPull components are connected together and given a constant value as no source impedance is ever looked up.

Accordingly the data table supplied to the DAC need to increase by four extra dimensions which correspond to the second and third harmonic impedances.

Inclusion of higher harmonics can be done by adding additional AmpLoadpull component into the schematic. Though theoretically indefinite number of harmonics can be included the maximum number is actually restricted by the measurement and simulation capacity. If measurement system can accurately characterize the DUT up

to 5th harmonic there is no need to add any harmonic higher than 5th in simulation. On the other hand, inclusion of higher harmonics would enlarge the size of the data table dramatically and hence need more and more computation resource to run the simulation. Besides, if considering from application point of view, few applications need tuning the device at harmonics higher than three. It can be concluded inclusion of 2nd and 3rd harmonic load-pull simulation capability should be sufficient for realistic design purpose. Please note that the 3rd order harmonic portion contributed from higher order is already included in data and will be emulate in simulations.

The harmonic load-pull implementation has been verified using simulation data obtained from a Freescale MET model [71] [72]. Sweeps of three harmonics are enabled when collecting the data. The MET model simulation data were then processed to generate a data table by using the DWLU approach. The data table was supplied into ADS to run simulations and the results were compared directly to the MET model simulation results. Figure 4-43 presents the comparisons between the DWLU simulation results and MET model simulation results obtained with two different 2nd harmonic load impedances. The good agreements indicate the DWLU approach can work well with harmonic load-pull data.

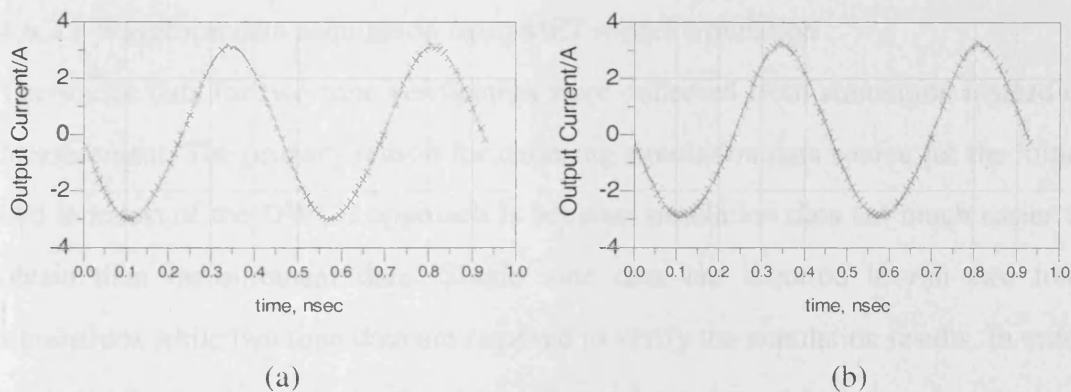


Figure 4-43 The DWLU simulated current (crossed) and MET model simulated current (solid) obtained with two different 2nd harmonic load impedances – (a) 67.647+j29.412 ohm and (b) 38-j16 ohm. The fundamental and 3rd harmonic load impedances are fixed at 8.621-j3.448 ohm and 29 ohm respectively for both cases.

4.6.2 Two tone simulations via envelope simulation

The approach has been used for importing CW data and running CW simulations so far. It would be interesting to investigate the potential of using CW waveform data in multi-tone simulations by making use of the proposed waveform utilization approach. Multi-tone signals represent more realistic excitation signals as most of the modern RF communication systems are excited by signals with non-zero bandwidth. Besides, multi-tone signals can be used to evaluate the device linearity by calculating intermodulation figures such as Adjacent Channel Power Ratio (ACPR) which is crucial in determination of quality of communication systems.

It's worth mentioning that the multi-tone signal is only a kind of discrete approximation of real continuous signal. How to approximate particular realistic signals with multi-tone signal is still an open topic and some of the recent work can be found in [68] – [70]. Anyhow it's always a good practice to start with basic two-tone signal which can be used to derive the typical multi-tone characteristics. By using such signal the DWLU approach's ability to predict multi-tone performance can be assessed with least amount of simulation complexity.

4.6.2.1 Waveform data acquisition from MET model simulation

The source data for two tone verification were collected from simulation instead of measurement. The primary reason for choosing simulation data source for the initial two tone test of the DWLU approach is because simulation data are much easier to obtain than measurement data. Single tone data are required to run two tone simulations while two tone data are required to verify the simulation results. In order to justify the results properly, the single tone and two tone data have to be obtained when DUT is operating in the identical environment, i.e. the impedance matching, biasing condition, drive stimuli have to be the same. Such requirement puts strict restrictions on data acquisition procedure, which dramatically increases the practical difficulty of acquiring data that can meet all the restrictions from measurement. It

would be much easier to obtain the required data from simulation. CAD tools offer models which can behave as a virtual device in data collection. The environment parameters can be easily set to be identical for single tone and two toned data acquisition. Besides, the accuracy of the model used in data acquisition is less important as long as it's repeatable. It can be seen that the simulation data should be sufficient for initial verification of the DWLU approach in two tone simulation.

The source data were obtained on a Freescale MET model which has a part number of MRF5S21100H. The model has 100w output power capacity and is working in a W-CDMA band from 2.11 GHz to 2.17 GHz. The *Loadpull Setup* template in ADS was used to collect the single tone data by simulating a typical load pull measurement. The *loadpull Setup* template was modified a little bit in order to generate the measurement data in the desired format. Specifically the input power source in the template is replaced by a voltage source to allow the input voltage to be swept. The output of simulation is also modified by applying the DWLU approach to make sure the generated data is in correct format for use in two-tone simulation. The data collection was made at 2.15 GHz. The input voltage was swept from 1.01 V to 8.01 V. The fundamental load reflection coefficient is swept with the real part from -0.4 to 0.1 and the imaginary part from -0.9 to -0.4. The generated data set was transferred to plain txt format by using the build-in instrument server in ADS. The data file was eventually formatted into standard MDIF format.

4.6.2.2 Envelope simulation using imported waveform data

This section shows the two tone simulation results obtained in envelope transient simulations. Before displaying the results, mathematical analysis is carried out in order to demonstrate the theoretical feasibility of using CW data directly in two tone simulation. The analysis is carried out in envelope domain.

According to the envelope transient simulation theory [73], any modulated signal can be represented as a carrier modulated by an envelope. Take two-tone signal as an

example:

$$V_{IN}(t) = \frac{1}{2} \sum_{n=1}^2 (V_n \cdot e^{j\omega_n t} + V_n^* \cdot e^{-j\omega_n t}) \quad (4-23)$$

The signal can be re-arranged to give:

$$V_{IN}(t) = V_1 \cos(\omega_1 t) + V_2 \cos(\omega_2 t) \quad (4-24)$$

If setting V_1 equivalent to V_2

$$\begin{aligned} V_{IN}(t) &= V \cos(\omega_1 t) + V \cos(\omega_2 t) \\ &= 2V \cos\left(\frac{\omega_1 - \omega_2}{2} t\right) \cos\left(\frac{\omega_1 + \omega_2}{2} t\right) \end{aligned} \quad (4-25)$$

If we define $\omega_2 = \omega_1 + \Delta\omega$,

$$\frac{\omega_2 - \omega_1}{2} = \frac{\Delta\omega}{2} \quad (4-26)$$

$$\frac{\omega_2 + \omega_1}{2} = \frac{2\omega_1 + \Delta\omega}{2} = \omega_1 + \frac{\Delta\omega}{2} \quad (4-27)$$

$$V_{IN}(t) = 2V \cos\left(\frac{\Delta\omega}{2} t\right) \cos\left(\left(\omega_1 + \frac{\Delta\omega}{2}\right) t\right) \quad (4-28)$$

Therefore the input two tone stimulus can be considered as a single tone at

$\left(\omega_1 + \frac{\Delta\omega}{2}\right) / 2\pi$ with time-varying amplitude and phase.

In order to make use of the envelope analysis, the signal envelope is sampled with an interval of T . And the amplitude and phase information of sampled envelope will be used as input for calculations of port current components.

$$Amp_{envelop}(nT) = 2V \cos\left(\frac{\Delta\omega}{2} nT\right) \quad (4-29)$$

$$Phas_{envelop}(nT) = \left(\omega_1 + \frac{\Delta\omega}{2}\right) \cdot nT \quad (4-30)$$

If setting $t_n = nT$,

$$\begin{aligned}
Amp_{envelop}(t_n) &= 2V \cos\left(\frac{\Delta\omega}{2} t_n\right) \\
Phas_{envelop}(t_n) &= (\omega_1 + \frac{\Delta\omega}{2}) \cdot t_n
\end{aligned} \tag{4-31}$$

If setting $f_m = f_1 + \frac{\Delta f}{2}$,

$$\begin{aligned}
Amp_{env}(t_n) &= 2V \cos(\pi \cdot \Delta f \cdot t_n) \\
Phas_{env}(t_n) &= 2\pi f_m \cdot t_n
\end{aligned} \tag{4-32}$$

Then the proposed data formulation shown in (4-7) and (4-8) is used to calculate the current components. The matrix used in CW simulation is shown below. Only the input port current calculation matrix is shown. The principle should be same for output port current calculation.

$$\begin{bmatrix} I_{10} \\ I_{11} \\ I_{12} \\ I_{13} \end{bmatrix} = \begin{bmatrix} A_0 & 0 & 0 & 0 \\ 0 & A_1 & 0 & 0 \\ 0 & 0 & A_2 & 0 \\ 0 & 0 & 0 & A_3 \end{bmatrix} \begin{bmatrix} 1 \\ V_{IN} \\ V_{IN}^2 \\ V_{IN}^3 \end{bmatrix} \tag{4-33}$$

The format needs to be changed a little bit:

$$\begin{bmatrix} I_{10}(t_n) \\ I_{11Env}(t_n) \\ I_{12Env}(t_n) \\ I_{13Env}(t_n) \end{bmatrix} = \begin{bmatrix} A_0(|V_{Env}|) & 0 & 0 & 0 \\ 0 & A_1(|V_{Env}|) & 0 & 0 \\ 0 & 0 & A_2(|V_{Env}|) & 0 \\ 0 & 0 & 0 & A_3(|V_{Env}|) \end{bmatrix} \begin{bmatrix} 1 \\ V_{Env}(t_n) \\ V_{Env}(t_n)^2 \\ V_{Env}(t_n)^3 \end{bmatrix} \tag{4-34}$$

Where the $V_{Env} = Amp_{envelop} \angle Phas_{envelop}$. For specific load and bias condition, the coefficient A_n should be a function of $|V_{Env}|$. Generally, the expressions such as I_{11Env} and V_{Env} are complex numbers. And for specific time point, the envelope of the current spectrum can be calculated by using the proper coefficient A_n in the look

up table.

If considering t_1 point,

$$\begin{aligned} I_{11Env}(t_1) &= A_1(|V_{Env}|) \cdot V_{Env}(t_1) \\ I_{12Env}(t_1) &= A_2(|V_{Env}|) \cdot V_{Env}(t_1)^2 \\ I_{13Env}(t_1) &= A_3(|V_{Env}|) \cdot V_{Env}(t_1)^3 \end{aligned} \quad (4-35)$$

Do the calculation for all the sampling points t_0, t_1, \dots, t_n , and then we can get the port current spectrum envelope varies as a function of time.

Finally we can obtain the port current spectrum by doing FFT (considering only fundamental one):

$$I_{11}(f) = \sum_{n=1}^N I_{11Env}(t_n) e^{j2\pi f \cdot t_n} \quad (4-36)$$

The resultant $I_{11}(f)$ will contain only the spectrum around f_m , which includes

spectrum at fundamental carriers $f_m \pm \frac{1}{2}\Delta f$, IM3 frequencies at $f_m \pm \frac{3}{2}\Delta f$, and IM5

frequencies at $f_m \pm \frac{5}{2}\Delta f$. Similarly the $I_{12}(f)$ will contain only the spectrum

around $2f_m$, which includes spectrum at 2nd harmonics $2f_m \pm \Delta f$ and etc.

It's also the same for the generated $I_{13}(f)$.

Once the current components at all interested frequencies are available, the corresponding voltage components can be obtained by multiplying the current with the proper impedance termination at each frequency.

Therefore, by using envelope analysis, the current spectrum can be extracted on carriers, harmonics, and IM frequencies. During the whole calculation process only the CW data is utilized. There are several points worthy of mentioning:

- 1) As shown in the derivations the CW data at f_m is needed for multi-tone simulation. For example if need to do two-tone simulation at 2145 MHz and 2155 MHz, the CW data should be obtained at 2150 MHz.
- 2) The envelope sampling frequency need to be varied according to the tone spacing.

Figure 4-44 gives an overview of the procedure of envelope domain analysis.

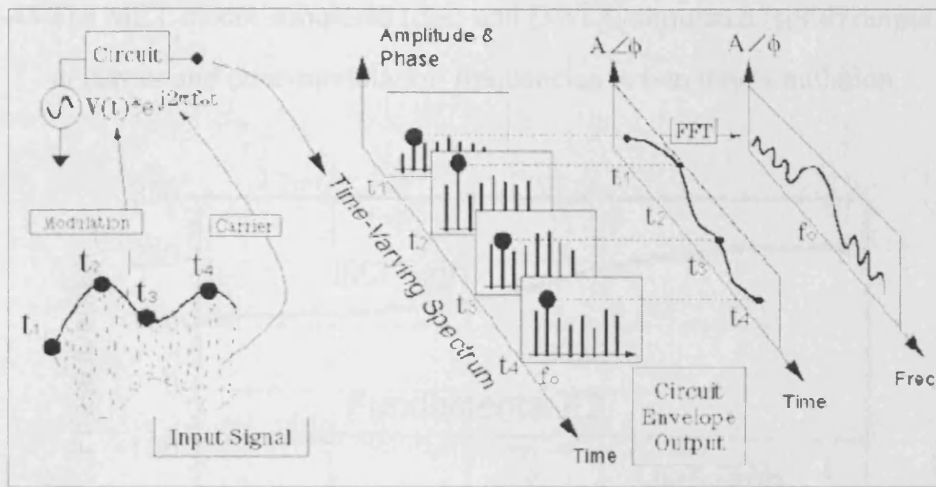


Figure 4-44: The principle of how envelope simulator works in ADS [74]

The two-tone simulated spectrum including fundamental carrier and IM components were then compared to the results obtained directly from MET model simulation (Figure 4-45 – Figure 4-46). Simulations show that by formulating the measurement data using the proposed DWLU approach, CW data can be able to predict the device behaviour in the two-tone signal simulations.

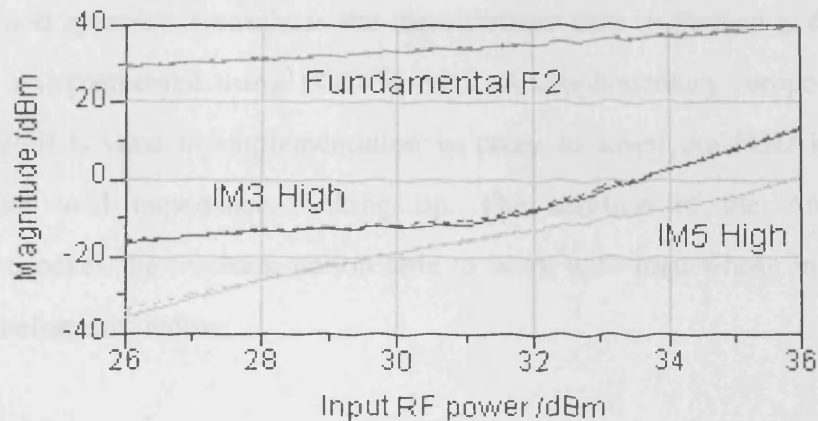


Fig 4-45 The MET model simulated (dot) and DWLU simulated (solid) output power at carrier and inter-modulation frequencies in two tone simulation.

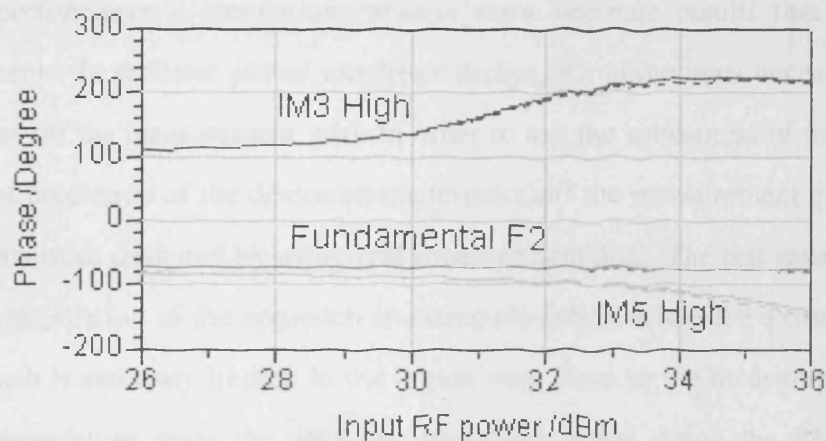


Fig 4-46 The MET model simulated (dot) and DWLU simulated (solid) output phase at carrier and inter-modulation frequencies in two tone simulation.

4.7 Summary

In this chapter a new approach for direct utilization of large signal waveform measurement data in CAD simulation is presented. The proposed approach retains the measurement information during the import of data into CAD. The lossless data processing procedure, which includes magnitude and phase scaling, improves the simulation robustness without loss in accuracy. Given the simulation environment is set to be identical to the measurement environment, the measured device nonlinear characteristics can be successfully reproduced in CAD simulation.

The proposed approach formulates the measurement data in frequency domain and therefore is implemented using FDD in ADS. A supplementary component called AmpLoadPull is used in implementation in order to assist the FDD in terms of fundamental load impedance looking up. The addition of the AmpLoadPull component makes the implementation able to work with load whose impedance is unknown before simulation.

The robustness of the approach has been tested experimentally on several devices. Experimental results show the approach is truly technology independent and application of the approach on different devices or device technologies shows similar level of performance – simulations always show accurate results that match the measurements. In realistic power amplifier design, it's sometimes necessary to run simulations off the measurement grid. In order to test the robustness of the approach in terms of prediction of the device characteristics off the measurement grid, several experiments were designed by using real measurement data. The test results indicate that the interpolation of the approach is extremely robust while the extrapolation of the approach is seriously limited to the region very close to the measurement space. Strong interpolation gives the designer confidence when doing the PA design. It saves the time of returning to the measurement system and measuring the missed points of interest. On the other side, robust interpolation allows less number of measurement points to be required for power amplifier design which saves the measurement time as well. The relatively poor extrapolation can be avoided by covering all the regions of interest in measurement. The easy expandability of the approach also makes it possible to add additional measurement data into simulation if the region covered by the additional data is missed in the initial measurement.

The excessive amount of measurement data to correctly represent the DUT for all stimuli and impedance conditions is readily avoided by taking the final application into account. For instance, LDMOS devices within base station amplifiers utilize bias conditions which are well specified and known prior the commencement of the

design. Also, often the design specifications force to design a specific power amplifier mode of operation, e.g. class-AB, hence allowing the impedance environment to be narrowed down over which the device needs to be measured. An important aspect of the direct use of non-linear data is that the resulting truth look-up model is an on-demand element, that is, its capabilities are readily expanded through additional measurements.

The approach investigated thoroughly in this chapter has the capability of looking up fundamental load impedance only. It's worth pointing out for applications which are interested in harmonic tuning it's also possible to extend the approach to make it capable of looking up harmonic impedances as well. But the simulation does suffer from slow speed as more harmonic impedance lookups are enabled.

The primary aim of the proposed approach is to import of CW waveform measurement data into CAD for CW simulation. The approach does have the potential to be used in simulations within a moderate bandwidth – provided the device has relatively flat performance across the bandwidth of simulation. Preliminary results, which were obtained with simulation data from a MET model, show the DWLU approach can give reasonable accuracy in two tone simulation by using the CW measurement data only.

Chapter 5 Polynomial-based Behavioural Modelling

5.1. Introduction

The direct utilization of large signal waveforms, which was explained in previous chapter, allows RF designers to access the key device behaviours in CAD environment. It greatly facilitates the measurement-based design process. However, it has some limitations as well. Under certain circumstances designers may interest in varying many variables at the same time to see how the nonlinear performance change or to find a proper compromise to meet all the design specs. These variables of interests may include input power, biasing, output impedance at fundamental and harmonic frequencies and so on. As the required number of variables increases, the measurement space necessary for maintaining good simulation accuracy will enlarge dramatically – sooner or later it will become very expensive and time-consuming to collect all the data. Another limitation is the simulation speed is relatively slow due to the large amount of data loaded into CAD. Therefore, it is still necessary to develop good model-based design approach which can overcome the above limitations. In large signal operation, power device usually demonstrates obvious nonlinearities. The nonlinear characteristics are the main obstacle that designers need to crack in order to meet the linearity requirement of modem wireless systems such as WCDMA. So the accurate approximation of nonlinear characteristics is the primary target for large signal modelling.

5.2. Choosing behavioural modelling of nonlinearity

The primary motivation behind the development of accurate and reliable large signal model is to enable rapid and low-cost CAD based power amplifier design. As the modern communication system becomes more and more sophisticated, PA design is more and more challenging. At the same time, the CAD techniques also become

more and more powerful due to the advances in mathematical algorithm and computational capability. Taking advantage of CAD tool effectively in design appears to be a promising solution to address the challenges in PA design.

In any CAD tool, the quality of the nonlinear model is always a key factor to determine how effective the CAD tool can be utilized in design. Accurate models can greatly enhance the chance of a first-pass design while inaccurate model may yield misleading results and sometimes even give completely wrong solution. Therefore accuracy is always one of the most important figures of merit when assessing a model.

On the other hand, a model by definition is approximation of physical behaviour existing in real world. It means certain amount of inaccuracy is unavoidable. Besides, from practical point of view, extremely accurate model can only be developed with extremely complicated modelling approach and time-consuming extraction experimental design. These negative effects may cancel out the advantages that usually are gained by using model in design such as low cost and short design time. So simplicity is another important figure of merit when assessing a model practically.

Consequently, it has to be very careful to choose the right model in PA design, which is supposed to achieve a proper balance between the accuracy and simplicity. Modelling is always a hot topic for research and many types of models have been developed so far. The pros and cons of some typical types of models have been reviewed in chapter two and corresponding examples have also been discussed. How to choose one from them to get the right balance is the first topic needed to be investigated. The point of balance is often application-oriented because different applications have different requirements for accuracy and complexity. For example, a physical-based model with hundreds of model coefficients may be disadvantageous from system level designers' point of view but may be favoured by device

manufacturers who are interested in semiconductor characteristics.

The main motivation of this modelling work is to develop reliable large signal model to enable rapid CAD-based power amplifier design. The model is intended to be used in system level simulation and design. Little interest is given on the internal physical operation of the transistor device. Regarding system level simulations, speed, complexity, simulation convergence issue become very important. Behavioural model, as compared to other types of modelling approaches, has the advantages of fast and reliable simulation and relatively simple model extraction and verification.

The NVNA measurement system, such as the one introduced in previous chapter, is essentially characterizing the device large signal characteristics in dynamic power condition. Therefore, the resultant measurement data are in theory a better resource for modelling purpose. A large signal model extracted from the large signal waveform data has the potential to provide better accuracy than traditional models. Investigations indicate behavioural modelling approach can be a suitable candidate for being extracted from waveform data. Behavioural modelling approach usually treats the DUT as a black box and describes the relationship between input and output by using analytical or non-analytical mathematical equations. The behavioural model parameters are usually extracted by doing curve-fitting to the measured DUT behaviour. Since the large signal waveform data generally contain all the information about the DUT behaviour, in theory behavioural model should be able to be extracted directly from the large signal waveform data.

Besides, behavioural model is very efficient when measurement data from more than one device technology are available, because it is technology independent. By using the large signal waveform measurement system developed in Cardiff University, many different types of transistor devices can be characterized within a relatively short period of time. A universal behavioural modelling approach which can be

applied to any type of transistor devices can reduce significantly the model development time for each type of the device technologies.

Considering all these factors, behavioural modelling approach is chosen as the subject in this work. The behavioural modelling approach proposed in this work is developed based on the so-called Polyharmonic Distortion (PHD) modelling concept. So first of all, the original PHD modelling approach will be reviewed in detail.

5.3. Original PHD modelling approach

Recently a large signal behavioural modelling approach called Polyharmonic Distortion (PHD) model [75] has been introduced. The approach has been briefly reviewed in chapter two and more details are given here. One advantage of the approach is that the model can be extracted from the large signal waveform data. The approach extends the tradition small signal s-parameter theory into large signal operation.

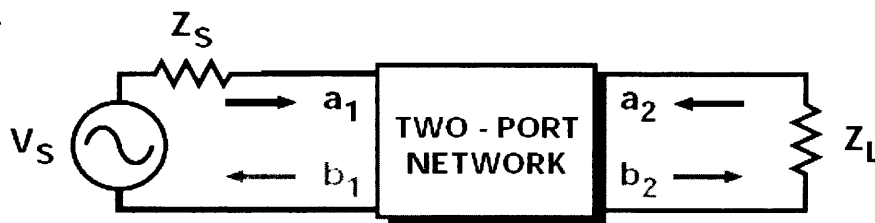


Figure 5-1: Typical two port network structure of PA transistor devices

Suppose the above typical two port network is defined. The a_n and b_n are the incident and reflective travelling waves respectively. The modelling theory is developed based on the so-called harmonic superposition principle. In general, due to the nonlinearities in large signal operation, the linear relationship and superposition principle between input and output are not valid any more. The small signal s-parameters are not valid any more in large signal operation. In order to characterize the two-port network in large signal operation, so-called large signal S parameters are proposed:

$$S_{11} = \frac{b_1}{a_1}; S_{21} = \frac{b_2}{a_1}; S_{12} = \frac{b_1}{a_2}; S_{22} = \frac{b_2}{a_2} \quad (5-1) - (5-4)$$

It's quite obvious that the large signal S parameters defined in (5-1) – (5-4) are a function of external parameters such as input power, load impedance, and bias etc.

The large signal S parameters shown in (5-1) – (5-4) are a general description of the large signal characteristics of the two port network. In some realistic power amplifier cases, there is only one dominant input component within the system, which is the fundamental input stimulus a_1 . Since the other input components such as a_2 at all the harmonics are relatively small, the superposition principle should still hold for them. That's what is called harmonic superposition principle.

The harmonic superposition principle was introduced by Jan Verspecht [78] through analyzing a mixer circuit. It was then applied and demonstrated on an amplifier circuit as well. The following graph illustrates the harmonic superposition principle.

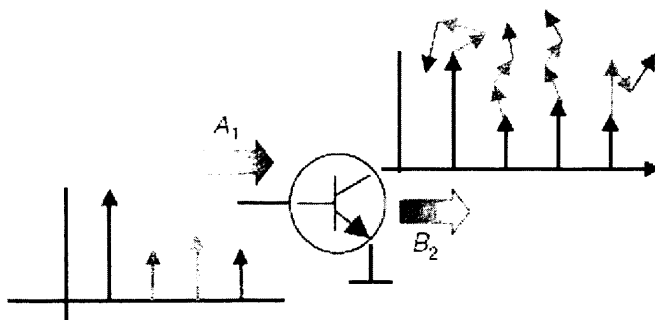


Figure 5-2: Harmonic superposition principle

As shown in the graph, the dominant fundamental $a_1(f_0)$ causes the nonlinearities in the output b_2 , i.e. not only fundamental but also the harmonic components are generated. Meanwhile, the harmonic contents in a_1 alter the final shape of the output b_2 . Since these harmonic components in a_1 are small, they all obey the harmonic

superposition principle and thus the alternation would be linear. Based on the harmonic superposition principle, the following simple mathematical formulation can be easily derived by extending from the small signal s-parameter definition.

$$b_{pm} = \sum_{qn} S_{pq,mn}(|a_{11}|)P^{+m-n}a_{qn} + \sum_{qn} T_{pq,mn}(|a_{11}|)P^{+m+n}conj(a_{qn}) \quad (5-5)$$

$$P = e^{+j\varphi(a_{11})}$$

Where p is port number of travelling wave b; m is index of harmonic of travelling wave b; q is port number of travelling wave a; n is index of harmonic of travelling wave a. Phasor P is defined as the phase vector of a_{11} .

The presence of phasor P in the model formulation is the result of phase normalization, which takes phase of a_{11} as the reference and normalizes phases of all other incident and reflective waves.

“ $S_{pq,mn}$ ” and “ $T_{pq,mn}$ ” are defined model parameters. Please note the defined “S” parameters are different from the small signal s-parameters or large signal S parameters in (5-1) – (5-4). As observed in (5-5), these model parameters are a nonlinear function of a_1 only, which is expressed as a_{11} . Apart from the fundamental a_1 , the output of the model is in linear relationship with all the other input components. The formulation in (5-5) accounts for both the fundamental and harmonic output and input relationships. It also accounts for the interaction between the fundamental and harmonics. For example, “ $S_{21,12}$ ” and “ $T_{21,12}$ ” represent the effects of second harmonic travelling wave a at input port on the fundamental output travelling wave b.

By comparing the formulation with the small signal s-parameter definition, it can be seen that the so-called “T” parameters are introduced in PHD model formulation. The appearance of “T” parameters is incurred due to the nonlinearities coming up in

large signal operation. It can be spotted that the contribution of incident waves a to reflective waves b involves terms associated with the conjugate of the incident wave a . In standard s -parameters, conjugate part is not present because the s -parameters are not a function of the phase of travelling wave a . The phase shift in travelling wave a would be equally reflected in travelling wave b . However, in a large signal regime where the travelling wave a has a large magnitude, the relative phase between the large signal a wave and remaining small signal a waves starts to affect the performance of device and consequently change the s -parameters. The above observation can be described by the presence of the conjugate part of travelling wave a . The graph shown in Figure 5-3 experimentally verifies the existence of the “ T ” parameters. The figure presents a polar plot of travelling wave b_2 . A set of small a_2 from output side of DUT is injected which depict a smiley pattern. Meanwhile, travelling wave a_1 is injected from input side of DUT. In order to investigate the influence of magnitude of a_1 on DUT, a set of a_1 with different magnitudes are used and the resultant smileys are plotted. It can be observed that the smiley remains undistorted when magnitude of a_1 is small but starts to exhibit squeezing distortion when magnitude of a_1 increases. Considering s -parameters are simply linear parameters, they only affect the smiley in terms of rotating and scaling. The squeezing effects are not contributed by the normal s -parameters and should be a direct consequence of the presence of “ T ” parameters.

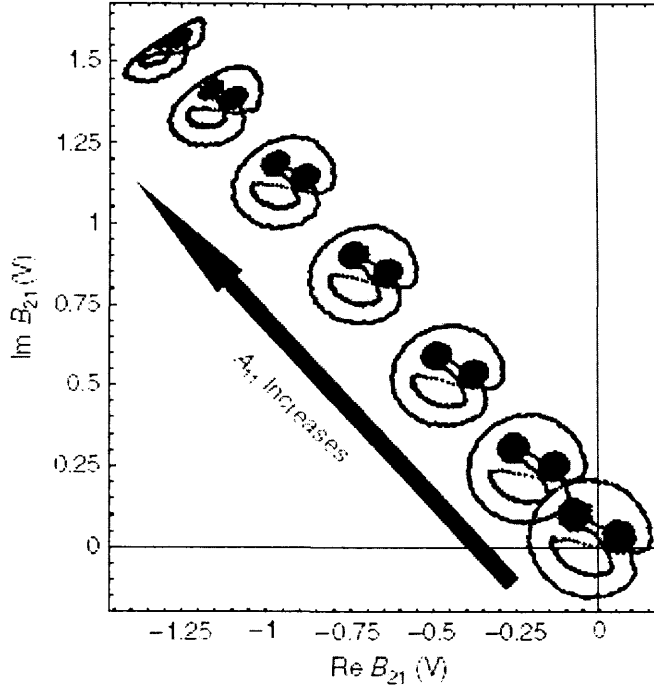


Figure 5-3: The b_{21} smiley measured as a function of magnitude of a_1 [77].

In this work, the primary effort is focused on fundamental, i.e. only b_{11} and b_{21} are considered and higher harmonic components such as b_{12} and b_{22} are not considered. There are two main reasons for doing that. Firstly, the formulation with fundamental only would simplify the equation and consequently give much clearer picture of what are the approach's pros and cons. Secondly, the measurement data available are generally collected when the device under test was terminated into a match at harmonics. Very weak or even no harmonic incident wave a_n exists. There is no need to account for the harmonic interaction in such a case. Therefore, if only the fundamental is considered, the standard PHD model formulation can be simplified into the following format:

$$b_{11} = S_{11,11}(|a_{11}|)a_{11} + S_{12,11}(|a_{11}|)a_{21} + T_{12,11}(|a_{11}|)P^2 \text{conj}(a_{21}) \quad (5-6)$$

$$b_{21} = S_{21,11}(|a_{11}|)a_{11} + S_{22,11}(|a_{11}|)a_{21} + T_{22,11}(|a_{11}|)P^2 \text{conj}(a_{21}) \quad (5-7)$$

One interesting advantage of the PHD modelling approach is that the model

decouples the model formulation from the load based on the harmonic superposition principle – making the PHD model more efficient in simulations. The modelling approach describes large signal behaviours in terms of an extension of the linear s-parameters by defining “S” and “T” parameters. The introduced “S” and “T” parameters are formulated to be dependent only on the magnitude of input travelling wave a_1 . The resulting model is able to extrapolate the nonlinear behaviours of the device at points away from where the model is extracted. However, the original PHD modelling approach mainly focuses on the effects of a_1 and harmonic interactions on “S” and “T” parameters and the model is developed based on the assumption that the travelling wave a_1 , which is considered as stimulus, is the only large signal tone in the whole system. It requires the travelling wave a_2 to be relatively small compared to a_1 . Therefore, according to this assumption, the model will be valid only at points around the reference impedance which is usually 50Ω . Such an assumption is usually ok for system components.

It can be a serious problem though for high power devices as in order to obtain sufficient information about them, load-pull contours, which will involve dramatic variation of a_2 , are needed. As can be observed on many power transistor devices, the magnitude of a_2 at the optimum impedance point of a high power transistor device is much larger than the magnitude of a_1 due to the high mismatch at the optimum load point. In this case a_2 cannot be treated as a small input component any more and the harmonic superposition principle doesn't hold any longer either. It imposes a dilemma for the PHD modelling approach as the model is valid only locally around the chosen reference impedance limiting the global accuracy of the model.

The limitation of original PHD modelling approach can be witnessed clearly in real measurements. Figure 5-4 shows the variations of magnitudes of a_1 and a_2 observed during the load-pull measurement on a typical high power transistor device. The magnitude of a_1 is relatively constant as the input power is kept constant during load-pull measurement. The magnitude of a_2 varies significantly as the level of

mismatch increases. In the region around the optimum impedance point, which usually has a long distance from the reference impedance point, the magnitude of a_2 can be several times larger than the magnitude of a_1 .

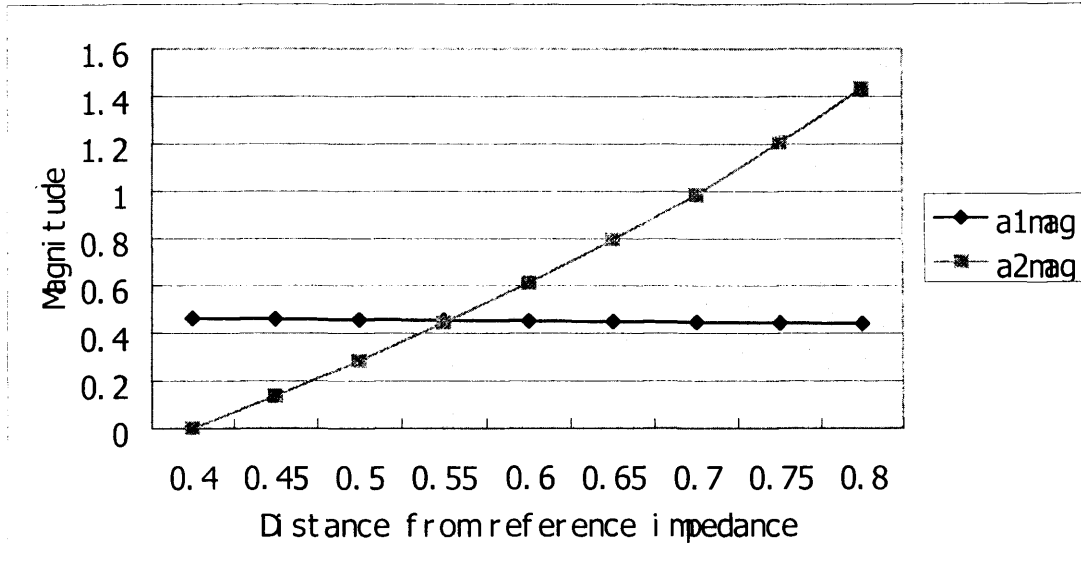


Figure 5-4: The variation of magnitudes of a_1 and a_2 as the simulation load impedance point is moving further and further away from the optimum point.

5.4. Polynomial-based behavioural model formulation

In this work, a new large signal behavioural model has been developed in order to overcome the limitation of the original PHD model. The presented approach has made substantial extension to the original PHD modelling approach. The PHD model formulation is extended by rigid mathematical derivation and a better balance between the global accuracy of the model and the complexity is achieved. In this section, the mathematical derivation process of the new model formulation is detailed.

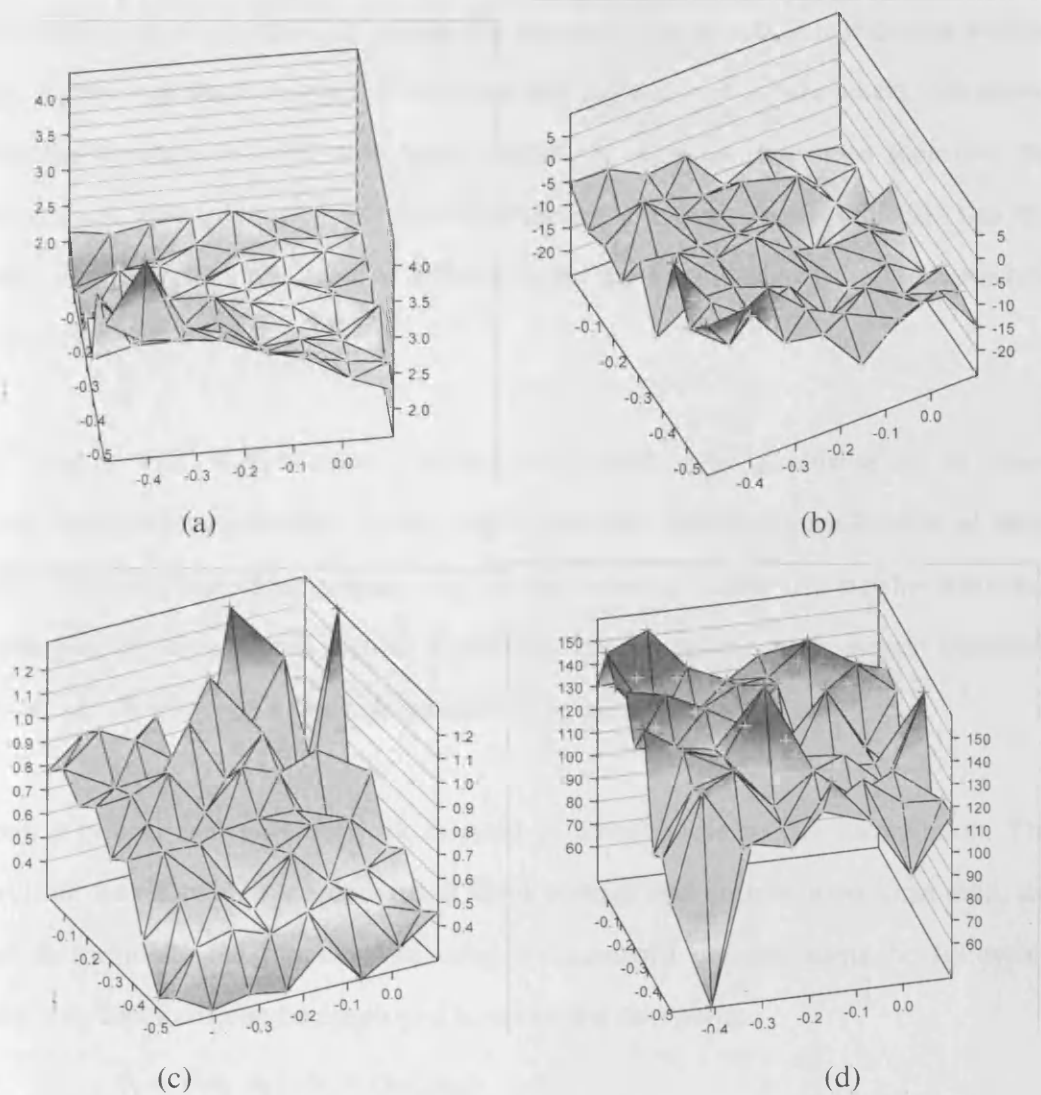


Figure 5-5: The S parameters calculated in large signal condition. They are plotted against the impedance of the load attached to the DUT. (a) and (b) show the magnitude and phase of S_{21} while (c) and (d) show the magnitude and phase of S_{22} .

As can be noticed in small signal world, s -parameters have gained great success and can be considered as good linear behavioural model. s -parameters are extracted with 50 ohm impedance environment but they remain constant even when the impedance environment changes away from 50 ohm. However, in large signal operation, the usefulness of s -parameters is questionable as s -parameters show significant variation when the impedance environment is changing. Particularly, high power transistor devices usually have their own “optimum impedance” which is often different from 50 ohm. This means it is always necessary to re-measure the s -parameters once the

load impedance is changed. It makes the s-parameters an inefficient design tool in large signal operation. Figure 5-5 shows the variation of s-parameters calculated when the device is excited with large magnitude stimulus. It can be seen that the s-parameters vary quite a bit as the load impedance is changing. It indicates that the device characteristics are heavily influenced by the impedance of the load attached to the device.

The original PHD model already shows that even simple extension of the small signal s-parameters into large signal world can yield decent approximation of large signal characteristics. The primary aim in this work is to identify similar nonlinear counterpart in large signal world, especially for modelling high power transistor device, i.e. the effect of a large magnitude of a_2 cannot be neglected.

Again a typical two port network is used to develop the model formulation. The travelling waves, which are converted from voltage and current waveform data, are used to formulate the model. Assuming the standard nomenclature the following describing functions can be employed to relate the two ports:

$$b_1 = f(a_1, a_2), b_2 = g(a_1, a_2) \quad (5-8)$$

These equations are of a very general form and can be used to represent any two port linear or nonlinear network.

However, the general describing functions only define the system behaviour on the conceptual level without offering any practical solution. In order to make the describing functions identifiable and able to be integrated into CAD simulator, a complex x-order polynomial is used to approximate the relationship between the output and input of (5-8). The polynomial is mathematically robust and therefore is ideal for describing the nonlinearity.

If n-order bivariate polynomial is used, (5-8) can be eventually expressed as:

$$b_1 = \varphi_{10}a_1 + \varphi_{01}a_2 + \varphi_{11}a_1a_2 + \cdots + \varphi_{n_1n_2}a_1^{n_1}a_2^{n_2} \quad (5-9)$$

$$b_2 = \theta_{10}a_1 + \theta_{01}a_2 + \theta_{11}a_1a_2 + \cdots + \theta_{n_1n_2}a_1^{n_1}a_2^{n_2} \quad (5-10)$$

Note that the maximum order n of the polynomial is the sum of n_1 and n_2 . The above equations represent a starting point for the mathematical derivations.

(5-9) – (5-10) represent a starting point for the mathematical derivations. Theoretically (5-9) – (5-10) would be able to accurately approximate DUT's nonlinearity of any order, as long as the order of polynomial n is sufficiently high. A mathematical analysis of (5-9) – (5-10) with the n th order polynomial would yield some interesting findings that will be discussed in later section. Until then, for the sake of brevity in analysis, the max order of the polynomial term in (5-9) – (5-10) adopted for model is limited to 3, which represents a good starting point for approximating nonlinearity. It should be noted that the max order referred here and thereafter is the combined order of the polynomial term but not the value of n shown in (5-9) – (5-10). For example, the term a_1a_2 has an order of 2.

If the order of polynomial is three, (5-9) – (5-10) can be written as:

$$b_1 = \varphi_{10}a_1 + \varphi_{01}a_2 + \varphi_{11}a_1a_2 + \cdots + \varphi_{30}a_1^3 + \varphi_{03}a_2^3 \quad (5-11)$$

$$b_2 = \theta_{10}a_1 + \theta_{01}a_2 + \theta_{11}a_1a_2 + \cdots + \theta_{30}a_1^3 + \theta_{03}a_2^3 \quad (5-12)$$

It can be noticed that lots of harmonic products are generated in (5-11) – (5-12). The table below gives a complete list of the products in (5-11) and the corresponding spectral vectors that they associate with. Similar table can be easily derived for products in (5-12) as well, which is skipped here for the sake of brevity.

Products	DC $e^{-j2\pi \cdot 0}$	Fundamental $e^{-j2\pi \cdot \omega_{f0}}$	2 nd harmonic $e^{-j2\pi \cdot 2\omega_{f0}}$	3 rd harmonic $e^{-j2\pi \cdot 3\omega_{f0}}$
φ_{00}	φ_{00}			
$\varphi_{10}a_1$		$\varphi_{10} a_1 e^{-\alpha_1}$		
$\varphi_{01}a_2$		$\varphi_{01} a_2 e^{-\alpha_2}$		
$\varphi_{11}a_1a_2$	$\varphi_{11} a_1 a_2 e^{\alpha_1-\alpha_2}$		$\varphi_{11} a_1 a_2 e^{\alpha_1+\alpha_2}$	
$\varphi_{20}a_1^2$	$2\varphi_{20} a_1 ^2$		$\varphi_{20} a_1 ^2e^{2\alpha_1}$	
$\varphi_{02}a_2^2$	$2\varphi_{02} a_2 ^2$		$\varphi_{02} a_2 ^2e^{2\alpha_2}$	
$\varphi_{21}a_1^2a_2$		$\varphi_{21} a_1 ^2 a_2 e^{-2\alpha_1+\alpha_2}$ $2\varphi_{21} a_1 ^2 a_2 e^{-\alpha_2}$		$e^{-2\alpha_1-\alpha_2}$
$\varphi_{12}a_1a_2^2$		$\varphi_{12} a_2 ^2 a_1 e^{-2\alpha_2+\alpha_1}$ $2\varphi_{12} a_2 ^2 a_1 e^{-\alpha_1}$		$e^{-2\alpha_2-\alpha_1}$
$\varphi_{30}a_1^3$		$3\varphi_{30} a_1 ^3e^{-\alpha_1}$		$\varphi_{30} a_1 ^3e^{-3\alpha_1}$
$\varphi_{03}a_2^3$		$3\varphi_{03} a_2 ^3e^{-\alpha_2}$		$\varphi_{03} a_2 ^3e^{-3\alpha_2}$

Table 5-1: The spectral vectors generated by each of the products in (5-11). In each of the vectors the magnitude terms are isolated from the phase terms. Only the positive frequency vectors are shown in the table.

As stated previously, the analysis followed in this work would only concentrate on the fundamental components in the output travelling waves, i.e. b_1 and b_2 contain the fundamental components only. The analysis involving harmonic interaction will be discussed in future work.

By investigating the terms in Table 5-1, it can be seen that only odd order polynomial terms contribute to the fundamental output. Furthermore, it can be seen

that many polynomial terms have different magnitude expressions but share the same phase expressions. Therefore it's intuitive to make an attempt to group these polynomial terms together according to the phase expression in order to see what the benefit it can give. The following table illustrates a way to group these polynomial terms according to the common initial phase terms. The common phase expression has been extracted out of the polynomial terms for clarity.

$e^{-\alpha_1}$	$\varphi_{10} a_1 ; \varphi_{21} a_1 ^2 a_2 ; 3\varphi_{30} a_1 ^3; \varphi_{21} a_1 ^2 a_2 e^{-(\alpha_1-\alpha_2)}$
$e^{-\alpha_2}$	$\varphi_{01} a_2 ; 2\varphi_{12} a_1 ^2 a_2 ; 3\varphi_{03} a_2 ^3; \varphi_{12} a_2 ^2 a_1 e^{-(\alpha_2-\alpha_1)}$

Table 5-2: The grouping of the fundamental terms according to the common initial phase operator

In frequency domain a_1 and a_2 can be expressed as:

$$|a_1|(e^{-j2\pi\omega-\alpha_1} + e^{j2\pi\omega+\alpha_1}) \quad (5-13)$$

$$|a_2|(e^{-j2\pi\omega-\alpha_2} + e^{j2\pi\omega+\alpha_2}) \quad (5-14)$$

In (5-13) – (5-14), $|a|$ is the magnitude of a and α is the initial phase information of a . Please note a constant coefficient 0.5 is neglected in (5-13) – (5-14) for sake of brevity. It won't affect the derivation of the proposed model. According to the grouping pattern in Table 5-2, mathematical expressions can be written to relate the travelling b_1 and b_2 to travelling a_1 and a_2 . By taking (5-13) – (5-14) into account, the mathematical expression relating b_1 to a_1 and a_2 can be written as:

$$\begin{aligned}
b_1 = & (\varphi_{10} |a_1| + 3\varphi_{30} |a_1|^3 + \varphi_{12} |a_1| |a_2|^2) \\
& \cdot (e^{-j2\pi\omega-\alpha_1} + e^{j2\pi\omega+\alpha_1}) \\
& + (\varphi_{21} |a_1|^2 |a_2|) (e^{-j2\pi\omega-2\alpha_1+\alpha_2} + e^{j2\pi\omega+2\alpha_1-\alpha_2}) \\
& + (\varphi_{01} |a_2| + 3\varphi_{03} |a_2|^3 + \varphi_{21} |a_1|^2 |a_2|) \\
& \cdot (e^{-j2\pi\omega-\alpha_2} + e^{j2\pi\omega+\alpha_2}) \\
& + (\varphi_{12} |a_1| |a_2|^2) (e^{-j2\pi\omega-2\alpha_2+\alpha_1} + e^{j2\pi\omega+2\alpha_2-\alpha_1})
\end{aligned} \tag{5-15}$$

The equation relating b_2 to a_1 and a_2 would be similar. By inspecting (5-15), it can be found the equation can be further reformulated into a form very similar to the small signal S parameter definitions:

$$\begin{aligned}
b_1 = & [(\varphi_{10} + 3\varphi_{30} |a_1|^2 + \varphi_{12} |a_2|^2) \\
& + (\varphi_{21} |a_1| |a_2|) \cdot \left(\frac{Q}{P}\right)] a_1 \\
& + [(\varphi_{01} + 3\varphi_{03} |a_2|^2 + \varphi_{21} |a_1|^2) \\
& + (\varphi_{12} |a_1| |a_2|) \cdot \left(\frac{P}{Q}\right)] a_2
\end{aligned} \tag{5-16}$$

$$\begin{aligned}
b_2 = & [(\theta_{10} + 3\theta_{30} |a_1|^2 + \theta_{12} |a_2|^2) \\
& + (\theta_{21} |a_1| |a_2|) \cdot \left(\frac{Q}{P}\right)] a_1 \\
& + [(\theta_{01} + 3\theta_{03} |a_2|^2 + \theta_{21} |a_1|^2) \\
& + (\theta_{12} |a_1| |a_2|) \cdot \left(\frac{P}{Q}\right)] a_2
\end{aligned} \tag{5-17}$$

In (5-16) and (5-17) Q and P represent phases of a_1 and a_2 .

$$Q = e^{-\alpha_1} \tag{5-18}$$

$$P = e^{-\alpha_2} \tag{5-19}$$

The derived equations (5-16) – (5-17) show some level of similarity to the standard S parameter definitions. The polynomial combinations within each square bracket can be considered as drive level varying s parameters. They vary as a function of a_1 and a_2 , which is reasonable due to the existence of the nonlinearity in large signal

operation. If inspecting (5-16) – (5-17) carefully, some interesting characteristics of the polynomial combinations can be spotted. One important characteristic is that the phase operator, Q/P or P/Q , doesn't appear in every term of the polynomial combinations. Some of the terms are related to magnitude of a_1 and a_2 only. Besides, the phase effects of a_1 and a_2 can be perfectly isolated from the magnitude effects of a_1 and a_2 . If C and U are defined to represent the magnitude terms of the polynomial combination and, we can have the following equations:

$$b_1 = (C_{1,0} + C_{1,1} \cdot \left(\frac{Q}{P}\right))a_1 + (U_{1,0} + U_{1,1} \cdot \left(\frac{P}{Q}\right))a_2 \quad (5-20)$$

$$b_2 = (C_{2,0} + C_{2,1} \cdot \left(\frac{Q}{P}\right))a_1 + (U_{2,0} + U_{2,1} \cdot \left(\frac{P}{Q}\right))a_2 \quad (5-21)$$

Where

$$C_{1,0} = \varphi_{10} + 3\varphi_{30}|a_1|^2 + \varphi_{12}|a_2|^2; C_{1,1} = \varphi_{21}|a_1||a_2| \quad (5-22)$$

$$U_{1,0} = \varphi_{01} + 3\varphi_{03}|a_2|^2 + \varphi_{21}|a_1|^2; U_{1,1} = \varphi_{12}|a_1||a_2| \quad (5-23)$$

$$C_{2,0} = \theta_{10} + 3\theta_{30}|a_1|^2 + \theta_{12}|a_2|^2; C_{2,1} = \theta_{21}|a_1||a_2| \quad (5-24)$$

$$U_{2,0} = \theta_{01} + 3\theta_{03}|a_2|^2 + \theta_{21}|a_1|^2; U_{2,1} = \theta_{12}|a_1||a_2| \quad (5-25)$$

By inspecting (5-20) – (5-21), it's worth to emphasize that the defined C and U parameters are related to the magnitudes of a_1 and a_2 only and independent of the phases of a_1 and a_2 . There are eight C and U parameters in the model formulation. In order to extract the model, these eight C and U parameters need to be identified. If (5-22) – (5-25) are treated as “black boxes”, the C and U can be expressed by using describing functions:

$$C_{1,0} = f_{10}(|a_1|, |a_2|); C_{1,1} = f_{11}(|a_1|, |a_2|) \quad (5-26)$$

$$U_{1,0} = g_{10}(|a_1|, |a_2|); U_{1,1} = g_{11}(|a_1|, |a_2|) \quad (5-27)$$

$$C_{2,0} = f_{20}(|a_1|, |a_2|); C_{2,1} = f_{21}(|a_1|, |a_2|) \quad (5-28)$$

$$U_{2,0} = g_{20}(|a_1|, |a_2|); U_{2,1} = g_{21}(|a_1|, |a_2|) \quad (5-29)$$

When extracting the model, instead of working out every single ψ and θ coefficient

that C and U parameters contain, a non-analytical function is used to approximate the relationship between the C and U parameters and magnitudes of a_1 and a_2 . Specifically, Cubic Spline function is chosen for the model. The main reason for choosing Cubic Spline function is the feature that the interpolated curve will pass through all the known points, i.e. the C and U parameters saved in the table will be faithfully preserved in interpolation. The nonlinear interpolation ability of the Cubic Spline function is not the main reason why this algorithm is favoured. Besides, Cubic Spline function is conveniently available in ADS simulator, which greatly facilitates the integration of the model into the CAD tool.

The real beauty of the (5-20) – (5-21) is that the model parameters variations are decoupled from the phase of the input stimulus a_1 and a_2 in a rigorous mathematical manner. The decoupling illustrates the inherent connection between the magnitudes of a_1 a_2 and behavioural model parameters. As we can observe when moving from small signal to large signal, the most significant change comes from the magnitude of input stimulus. It's the increased magnitude of the stimulus, but not the phase of it, that makes the s-parameters invalid in large signal operation. Such a typical observation is firmly confirmed in theory by the generic model formulation in (5-20) – (5-21), whose parameters show the dependency only on the magnitudes but not the phases of a_1 and a_2 .

As far as behavioural model development is concerned, the decoupling of the phase dependency reduces the complexity of the model dramatically and makes the model parameter much easier to extract. It also reduces the required minimum number of measurements for identifying the model parameters.

Comparing to traditional s-parameters, additional parameters are introduced to describe the nonlinear input and output relationship of a two port network. Besides, all the parameters are dependent on the magnitude of input stimulus and output load impedance. The dependency is necessary to describe the variation of device

characteristics in large signal operation. Phase operators also appear in the formulation. They are used to make sure all the input and output vectors maintain correct phase relationship.

Recall the derivation process of the model formulation and it can be seen that there are actually several ways to group the fundamental polynomial terms. The way shown in Table 5-1 is just one of them. It's interesting to find out the alternative formulation which can be obtained by adopting different grouping strategy. In fact, if a different way is used to group these polynomial terms, a slightly different version of the model formulation can be obtained. The table below shows the different way used to group these terms.

$e^{-\alpha_1}$	$\varphi_{10} a_1 ; \varphi_{21} a_1 ^2 a_2 ; 3\varphi_{30} a_1 ^3; \varphi_{12} a_2 ^2 a_1 e^{-2(\alpha_2-\alpha_1)}$
$e^{-\alpha_2}$	$\varphi_{01} a_2 ; 2\varphi_{12} a_1 ^2 a_2 ; 3\varphi_{03} a_2 ^3; \varphi_{21} a_1 ^2 a_2 e^{-2(\alpha_1-\alpha_2)}$

Table 5-3: Alternative way to group the fundamental terms according to a different common initial phase operator

Based on the grouping method shown in Table 5-3, the model formulation can be written as:

$$\begin{aligned}
 b_1 = & (\varphi_{10} + 3\varphi_{30}|a_1|^2 + \varphi_{12}|a_2|^2)a_1 \\
 & + (\varphi_{12}|a_2|^2)a_1^*p^2 \\
 & + (\varphi_{01} + 3\varphi_{03}|a_2|^2 + \varphi_{21}|a_1|^2)a_2 \\
 & + (\varphi_{21}|a_1|^2)a_2^*q^2
 \end{aligned} \tag{5-30}$$

$$\begin{aligned}
b_2 = & (\theta_{00} + 3\theta_{30}|a_1|^2 + \theta_{12}|a_2|^2)a_1 \\
& + (\theta_{12}|a_2|^2)a_1^* p^2 \\
& + (\theta_{01} + 3\theta_{03}|a_2|^2 + \theta_{21}|a_1|^2)a_2 \\
& + (\theta_{21}|a_1|^2)a_2^* q^2
\end{aligned} \tag{5-31}$$

Alternatively, “S” and “T” parameters can be defined to designate these magnitude terms.

$$S_{11} = \varphi_{10} + 3\varphi_{30}|a_1|^2 + \varphi_{12}|a_2|^2; T_{11} = \varphi_{12}|a_2|^2; \tag{5-32}$$

$$S_{12} = \varphi_{01} + 3\varphi_{03}|a_2|^2 + \varphi_{21}|a_1|^2; T_{12} = \varphi_{21}|a_1|^2; \tag{5-33}$$

$$S_{21} = \theta_{10} + 3\theta_{30}|a_1|^2 + \theta_{12}|a_2|^2; T_{21} = \theta_{12}|a_2|^2 \tag{5-34}$$

$$S_{22} = \theta_{01} + 3\theta_{03}|a_2|^2 + \theta_{21}|a_1|^2; T_{22} = \theta_{21}|a_1|^2 \tag{5-35}$$

(5-30) – (5-31) can then be expressed in the following format:

$$b_1 = S_{11}a_1 + T_{11}a_1^* p^2 + S_{12}a_2 + T_{12}a_2^* q^2 \tag{5-36}$$

$$b_2 = S_{21}a_1 + T_{21}a_1^* p^2 + S_{22}a_2 + T_{22}a_2^* q^2 \tag{5-37}$$

(5-36) – (5-37) illustrate some level of similarity to the original PHD model on the selection of “S” and “T” parameters. By comparing (5-36) – (5-37) with the original PHD model formulation shown in (5-3) – (5-4), two major differences can be easily figured out. First one is that the “S” and “T” parameters within the original PHD model definition are a function of a_1 magnitude only while the “S” and “T” parameters, as presented within this work, are a function of both a_1 and a_2 magnitude. Another important difference is that “ T_{11} ” and “ T_{21} ” do not exist within the original PHD model while they are present in above model formulation. The reason behind these differences is that the model presented in this work is derived without assuming the validity of the harmonic superposition principle. Consequently, the influence of a_2 on these “S” and “T” parameters has to be taken into account. In fact the original PHD model formulation can be interpreted as a special case when magnitude of a_2 is very small, i.e. the “ T_{11} ” and “ T_{21} ” will vanish for an a_2 signal

approaching zero.

(5-20) – (5-21) and (5-30) – (5-31) are different expressions which essentially describe the same mathematical relationship between the reflective travelling wave b_n and incident travelling wave a_n . The relationship between C&U parameters and S&T parameters can be identified by making direct comparisons between (5-20) – (5-21) and (5-30) – (5-31).

From (5-21):

$$\begin{aligned}
 b_2 &= C_{20}a_1 + C_{21}\frac{Q}{P}a_1 + U_{20}a_2 + U_{21}\frac{P}{Q}a_2 \\
 &= C_{20}a_1 + C_{21}\frac{Q}{P}\cdot Q\cdot|a_1| + U_{20}a_2 + U_{21}\frac{P}{Q}\cdot P\cdot|a_2| \\
 &= C_{20}a_1 + C_{21}\frac{Q^2}{P}\cdot|a_1| + U_{20}a_2 + U_{21}\frac{P^2}{Q}\cdot|a_2|
 \end{aligned} \tag{5-38}$$

From (5-31):

$$\begin{aligned}
 b_2 &= S_{21}a_1 + T_{21}P^2\cdot\frac{1}{Q}\cdot|a_1| + S_{22}a_2 + T_{22}Q^2\cdot\frac{1}{P}\cdot|a_2| \\
 &= S_{21}a_1 + T_{21}\frac{P^2}{Q}\cdot|a_1| + S_{22}a_2 + T_{22}\frac{Q^2}{P}\cdot|a_2|
 \end{aligned} \tag{5-39}$$

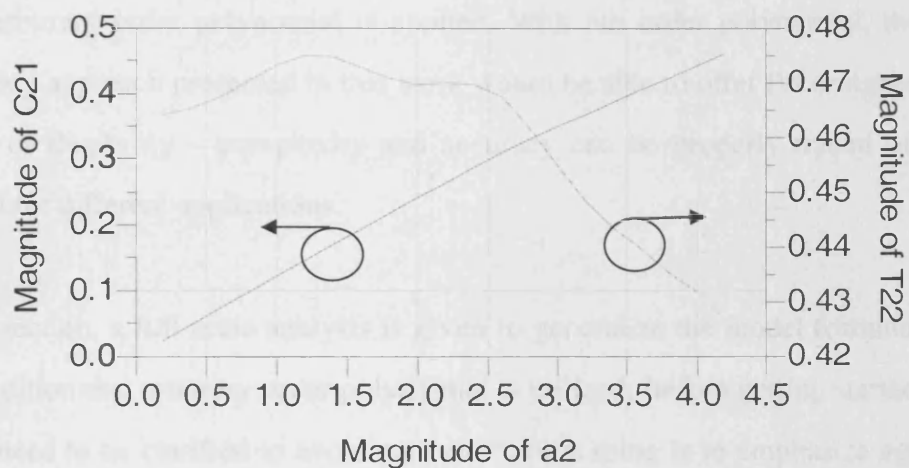
Comparing (5-38) and (5-39), it can be found that:

$$C_{20} = S_{21}; \quad U_{20} = S_{22} \tag{5-40}$$

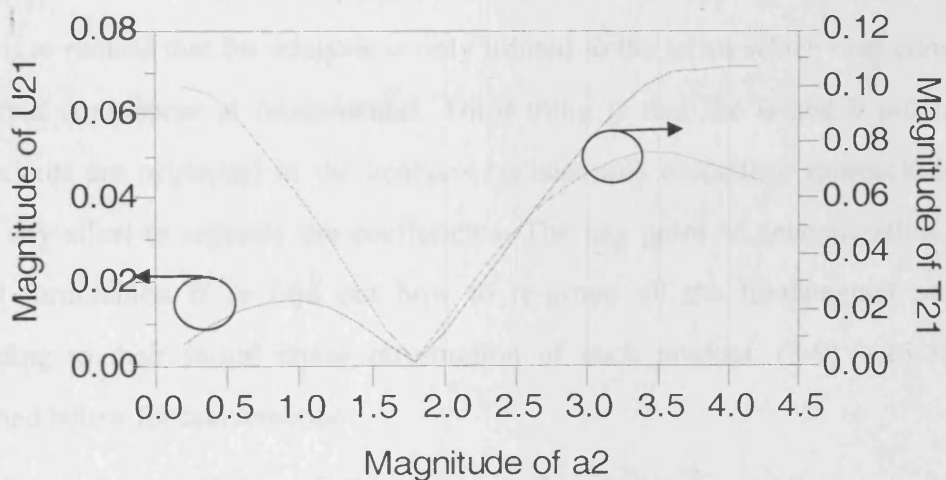
$$C_{21} = T_{22}\frac{|a_2|}{|a_1|}; \quad U_{21} = T_{21}\frac{|a_2|}{|a_1|} \tag{5-41}$$

Theoretically (5-20) – (5-21) and (5-30) – (5-31) should give identical results in simulation. In this work (5-20) – (5-21) are chosen instead of (5-30) – (5-31) as the final model formulation due to several practical factors. Firstly (5-20) – (5-21) are slightly simpler than (5-30) – (5-31) as the formulation doesn't involve any conjugate parts of a_1 or a_2 . It makes the model formulation more straightforward and the implementation easier. Secondly, the curves of the C&U parameters are smoother

than the curves of the S&T parameters. Figure 5-6 shows the key observations on the variations of C&U parameter and S&T parameters against magnitude of a_2 . It can be seen that the curve of C parameter apparently has lower order variation than the corresponding T parameters. A smoother curve can benefit the simulation by improving the accuracy as table-based interpolation is used to describe the behaviour of these parameters in simulation.



(a)



(b)

Figure 5-6: Comparisons are made between the C&U parameters and S&T parameters. (a) compares the magnitude of C_{21} and magnitude of T_{22} . (b) compares the magnitude of U_{21} and magnitude of T_{21} .

5.5. Generalization of polynomial-based model formulation

The novel model formulation presented in previous section possesses one unique characteristic that all the model parameters are a function of magnitudes only. The phase information is cleverly preserved in the formulation and can be faithfully re-constructed at the output of the model. This characteristic is obtained by analyzing a specific case – a 3rd order model. In fact, this characteristic can be preserved even when arbitrary order polynomial is applied. With n th order polynomial, the novel modelling approach presented in this work would be able to offer PA designers some degree of flexibility – complexity and accuracy can be properly traded off to be tailored for different applications.

In this section, a full scale analysis is given to generalize the model formulation on the condition that arbitrary order polynomial is utilized. Before getting started, a few points need to be clarified to avoid confusion. First thing is to emphasize again that the order of terms referred later in analysis is the order of the terms appearing in (5-9) – (5-10) but not the order of polynomial used to obtain (5-9) – (5-10). Second thing is to remind that the analysis is only limited to the terms which may contribute to output component at fundamental. Third thing is that the ψ and θ polynomial coefficients are neglected in the analysis because this modelling approach doesn't make any effort to separate the coefficients. The key point of generalization of the model formulation is to find out how to re-group all the fundamental products according to their initial phase information of each product. (5-9) – (5-10) are reprinted below for convenience

$$b_1 = \varphi_{00} + \varphi_{01}a_2 + \varphi_{11}a_1a_2 + \varphi_{30}a_1^3 + \dots + \varphi_{n1n2}a_1^{m_1}a_2^{m_2} \quad (5-42)$$

$$b_2 = \theta_{00} + \theta_{01}a_2 + \theta_{11}a_1a_2 + \theta_{30}a_1^3 + \dots + \theta_{n1n2}a_1^{m_1}a_2^{m_2} \quad (5-43)$$

The products in (5-42) – (5-43) are first categorized according to their order. The analysis is started by making attempts to relate the first order terms in large signal model formulation to standard small signal s-parameter definition. Emphasis would be taken on the phase characteristics of each product. To facilitate the following

analysis, a_1 and a_2 are expressed as a combination of magnitude portion and phase operator, by considering (5-13) – (5-14) and (5-18) – (5-19).

$$a_1 = |a_1| (e^{-j2\pi\omega - \alpha_1} + e^{j2\pi\omega + \alpha_1}) = |a_1| Q \quad (5-44)$$

$$a_2 = |a_2| (e^{-j2\pi\omega - \alpha_2} + e^{j2\pi\omega + \alpha_2}) = |a_2| P \quad (5-45)$$

In (5-44) – (5-45), Q and P are the phase portion of a_1 and a_2 respectively and are termed phase operator. Each product in (5-42) – (5-43) would generate mixed terms with various combination of Q and P . The combination of Q and P determines the initial phase of the generated term. Rather than utilizing magnitude information, the products in (5-42) – (5-43) are grouped into several categories according to the initial phase of each term generated from mixing.

1) First order products

There are only two first order products in (5-42) or (5-43). They are:

$$a_1 ; a_2$$

They will contribute to the fundamental output linearly. The coefficients in these first order terms are usually referred to as S parameters. For first order product, no mixing term is generated at all. The phase information of each product is shown in Table 5-4.

First order products	Generated phase vectors
a_1	Q
a_2	P

Table 5-4: The phase operators generated by first order products from (5-42) or (5-43)

2) Second order products

There are three second order terms in (5-42) or (5-43). They are

$$a_1^2 ; a_1 a_2 ; a_2^2$$

Since they are even order products, no contribution would be made to the fundamental output. So these products will be simply discarded from the model formulation.

3) Third order products

There are four third order products. They are:

$$a_1^3; a_1^2 a_2; a_1 a_2^2; a_2^3$$

These odd order products would contribute to the fundamental output. Take $a_1^2 a_2$ as an example. a_1^2 is a 2nd harmonic ($2f_0$) component while a_2 is a fundamental (f_0) component. The mixing of a 2nd harmonic and fundamental component would generate fundamental component as a result of frequency subtraction ($2f_0 - f_0$) during mixing.

The fundamental first order terms generated by these third order products have different phase vectors. The table below shows the generated different phase vectors in relation to these third order terms.

Third order products	Generated phase vectors at fundamental
a_1^3	Q
$a_1^2 a_2$	$\frac{Q^2}{P}, P$
$a_1 a_2^2$	$\frac{P^2}{Q}, Q$
a_2^3	P

Table 5-5: The phase operators generated by third order products from (5-42) or
(5-43)

In order to simplify the expression of the formulation and reduces the number of

model parameters, attempts are made to group the fundamental terms generated from mixing together and eventually get a compact format. By observing the terms shown in Table 5-5, it can be found that the phase vectors seem to demonstrate more commonality than the magnitude terms. Therefore, the fundamental terms with a common Q or P are grouped together. So they are grouped into two categories, as shown below:

Terms with common Q	Q (contributed from both a_1^3 and $a_1 a_2^2$) $\left(\frac{Q}{P}\right) \cdot Q$ (contributed from $a_1^2 a_2$)
Terms with common P	P (contributed from both a_2^3 and $a_1^2 a_2$) $\left(\frac{P}{Q}\right) \cdot P$ (contributed from $a_1 a_2^2$)

Table 5-6: Grouping of third order products

4) Fourth order products

There are five fourth order products as shown below:

$$a_1^4; a_1^3 a_2; a_1^2 a_2^2; a_1 a_2^3; a_2^4$$

Similar to second order products, these fourth order products generate no term at fundamental.

5) Fifth order products

There are six fifth order products as shown below:

$$a_1^5; a_1^4 a_2; a_1^3 a_2^2; a_1^2 a_2^3; a_1 a_2^4; a_2^5$$

The fundamental phase vectors generated by the above mixing products are listed in the following table.

Fifth order products	Generated phase vectors at fundamental
a_1^5	Q
$a_1^4 a_2$	$\frac{Q^2}{P}, P$
$a_1^3 a_2^2$	$Q, \frac{P^2}{Q}, \frac{Q^3}{P^2}$
$a_1^2 a_2^3$	$P, \frac{Q^2}{P}, \frac{P^3}{Q^2}$
$a_1 a_2^4$	$Q, \frac{P^2}{Q}$
a_2^5	P

Table 5-7: The phase operators generated by fifth order products from (5-42) or (5-43)

It can be seen from Table 5-7 that all the products generated by the third order terms are generated by the fifth order terms as well. Besides, the fifth order terms also generate some new products which are not covered by third order terms. In fact, this observation is maintained for other order terms as well, i.e. the products generated by lower order terms are always a subset of the products generated by higher order terms.

Similarly, all these fundamental terms generated by fifth order products are grouped using the phase information.

Terms with common Q	Q (contributed from both a_1^5 , $a_1^3 a_2^2$ and $a_1 a_2^4$) $\left(\frac{Q}{P}\right) \cdot Q$ (contributed from $a_1^4 a_2$ and $a_1^2 a_2^3$) $\left(\frac{Q}{P}\right)^2 \cdot Q$ (contributed from $a_1^3 a_2^2$)
Terms with common P	P (contributed from both a_2^5 , $a_1^2 a_2^3$ and $a_1^4 a_2$) $\left(\frac{P}{Q}\right) \cdot P$ (contributed from $a_1 a_2^4$ and $a_1^3 a_2^2$) $\left(\frac{P}{Q}\right)^2 \cdot P$ (contributed from $a_1^2 a_2^3$)

Table 5-8: Grouping of fifth order products

6) Sixth order products

There are seven sixth order products:

$$a_1^6, a_1^5 a_2, a_1^4 a_2^2, a_1^3 a_2^3, a_1^2 a_2^4, a_1 a_2^5, a_2^6$$

It can be seen that no fundamental term is generated.

7) Seventh order products

There are eight seventh order products:

$$a_1^7, a_1^6 a_2, a_1^5 a_2^2, a_1^4 a_2^3, a_1^3 a_2^4, a_1^2 a_2^5, a_1 a_2^6, a_2^7$$

The fundamental phase vectors generated by the above mixing products are listed in the following table.

Seventh order products	Generated phase vectors at fundamental
a_1^7	Q
$a_1^6 a_2$	$P, \frac{Q^2}{P}$
$a_1^5 a_2^2$	$Q, \frac{Q^3}{P^2}$
$a_1^4 a_2^3$	$P, \frac{Q^2}{P}, \frac{P^3}{Q^2}, \frac{Q^4}{P^3}$
$a_1^3 a_2^4$	$Q, \frac{P^2}{Q}, \frac{Q^3}{P^2}, \frac{P^4}{Q^3}$
$a_1^2 a_2^5$	$P, \frac{P^3}{Q^2}$
$a_1 a_2^6$	$Q, \frac{P^2}{Q}$
a_2^7	P

Table 5-9: The phase operators generated by seventh order products from (5-42) or
(5-43)

All these generated fundamental terms are categorized by making use of the phase information.

Terms with common Q	Q (contributed from both a_1^7 , $a_1^5 a_2^2$, $a_1^3 a_2^4$ and $a_1 a_2^6$) $\left(\frac{Q}{P}\right) \cdot Q$ (contributed from $a_1^6 a_2$ and $a_1^4 a_2^3$) $\left(\frac{Q}{P}\right)^2 \cdot Q$ (contributed from $a_1^5 a_2^2$ and $a_1^3 a_2^4$) $\left(\frac{Q}{P}\right)^3 \cdot Q$ (contributed from $a_1^4 a_2^3$)
Terms with common P	P (contributed from both a_2^7 , $a_1^2 a_2^5$, $a_1^4 a_2^3$ and $a_1^6 a_2$) $\left(\frac{P}{Q}\right) \cdot P$ (contributed from $a_1 a_2^6$ and $a_1^3 a_2^4$) $\left(\frac{P}{Q}\right)^2 \cdot P$ (contributed from $a_1^2 a_2^5$ and $a_1^4 a_2^3$) $\left(\frac{P}{Q}\right)^3 \cdot P$ (contributed from $a_1^3 a_2^4$)

Table 5-10: Grouping of seventh order products

As observed from Table 5-6, 5-8 and 5-10, a clear sequence is emerging as the order increases from 3 to 7. Only odd order products generate terms at fundamental while even order products generally have no impact on the fundamental output. As the order increases by two each time, new phase vectors in the format of

$$\left(\frac{Q}{P}\right)^n \cdot P \text{ and } \left(\frac{P}{Q}\right)^n \cdot P$$

will be generated.

8) Nth order terms

Based on the above information, all the possibilities of the phase vectors generated by nth order terms can be derived. They can be grouped in the similar manner as used in the previous sessions.

Terms with common Q	Q , $\left(\frac{Q}{P}\right) \cdot Q$, $\left(\frac{Q}{P}\right)^2 \cdot Q$, $\left(\frac{Q}{P}\right)^3 \cdot Q$, $\left(\frac{Q}{P}\right)^4 \cdot Q, \dots, \left(\frac{Q}{P}\right)^n \cdot Q$
Terms with common P	P , $\left(\frac{P}{Q}\right) \cdot P$, $\left(\frac{P}{Q}\right)^2 \cdot P$, $\left(\frac{P}{Q}\right)^3 \cdot P$, $\left(\frac{P}{Q}\right)^4 \cdot P$, $\left(\frac{P}{Q}\right)^n \cdot P$

Table 5-11: Grouping of nth order products

It's worth mentioning that Table 5-11 contains all the possible phase expressions associated with fundamental terms that are generated by mixing products from 1st order to nth order. As the order increases, a few new phase expressions appear. But all the possible phase expressions generated by lower orders are always fully covered by high order terms. Therefore, when it comes to the nth order terms, every possible phase expressions generated from first order to (n-1) order terms are included.

It can be found that in order to obtain the model output b_n at fundamental, it's necessary to summate all the fundamental terms generated by these products with different orders – from first to nth.

$$\begin{aligned}
 b_n &= |\rho_{1st,1}| \cdot Q + |\rho_{3rd,1}| \cdot Q + |\rho_{3rd,2}| \cdot \frac{Q}{P} \cdot Q + \dots + |\rho_{nth,m}| \cdot \left(\frac{Q}{P}\right)^{m-1} Q \\
 &\quad + |\zeta_{1st,1}| \cdot P + |\zeta_{3rd,1}| \cdot P + |\zeta_{3rd,2}| \cdot \frac{P}{Q} \cdot P + \dots + |\zeta_{nth,m}| \cdot \left(\frac{P}{Q}\right)^{m-1} P \\
 &= \sum_1 |\rho_{1st,m}| \cdot \left(\frac{Q}{P}\right)^{m-1} Q + \sum_2 |\rho_{3rd,m}| \cdot \left(\frac{Q}{P}\right)^{m-1} Q + \dots + \sum_m |\rho_{nth,m}| \cdot \left(\frac{Q}{P}\right)^{m-1} Q \quad (5-46) \\
 &\quad + \sum_1 |\zeta_{1st,m}| \cdot \left(\frac{P}{Q}\right)^{m-1} P + \sum_2 |\zeta_{3rd,m}| \cdot \left(\frac{P}{Q}\right)^{m-1} P + \dots + \sum_m |\zeta_{nth,m}| \cdot \left(\frac{P}{Q}\right)^{m-1} P
 \end{aligned}$$

Terms at f_0 generated
by 1st order mixing
products

Terms at f_0 generated
by 3rd order mixing
products

Terms at f_0 generated
by nth order mixing
products

ζ and ρ are used to designate magnitude terms associated with each of the phase vectors. It can be seen that the fundamental terms generated by mixing products with different orders are clearly differentiated from each other and summation is done separately from one mixing order to another.

By observing the above equation, it can be found that further re-arrangement is possible, by grouping all the terms with same phase vector regardless of what mixing order it associates with. In fact, as has already been known, Table 5-11 showing the phase vectors generated by n th order mixing products covers all the possible phase vectors generated from 1st order to n th order. In that case the following equation can be derived:

$$b_n = \sum_n \left(|\rho_{1st,1}| + |\rho_{3rd,1}| + \dots + |\rho_{nth,1}| \right) \cdot Q + \sum_{n=1} \left(|\rho_{3rd,2}| + \dots + |\rho_{nth,2}| \right) \frac{Q}{P} \cdot Q + \dots + \sum_1 |\rho_{nth,m}| \left(\frac{Q}{P} \right)^{m-1} Q \quad (5-47)$$

$$+ \sum_n \left(|\zeta_{1st,1}| + |\zeta_{3rd,1}| + \dots + |\zeta_{nth,1}| \right) \cdot P + \sum_{n=1} \left(|\zeta_{3rd,2}| + \dots + |\zeta_{nth,2}| \right) \frac{P}{Q} \cdot P + \dots + \sum_1 |\zeta_{nth,m}| \left(\frac{P}{Q} \right)^{m-1} P$$

If the magnitude terms in each of the bracket are merged and new parameters C and U are defined to designate the merged magnitude terms, the generic model formulation can be derived as shown below:

$$b_k = \sum_{m=0}^{\frac{n-1}{2}} C_{k,m} \cdot \left(\frac{Q}{P} \right)^m a_1 + \sum_{m=0}^{\frac{n-1}{2}} U_{k,m} \cdot \left(\frac{P}{Q} \right)^m a_2 \quad (5-48)$$

The defined $C_{k,m}$ and $U_{k,m}$ are polynomial combinations of magnitudes of a_1 and a_2 . k represents the port number while n represents the order of polynomial used in model. (21) is the generic form of the model formulation if n th order polynomial is applied for approximating the nonlinearity. A few points need to be clarified for better understanding of the elegance of the generic model formulation. Firstly, n , the maximum order of the polynomial, is an odd integral number. It's because an even order term doesn't generate any component at fundamental. Any even order term would be removed during the equation expansion. Secondly, the number of C and U

parameters could depend on the maximum order of polynomial allowed for deriving the equations. Each time the order of model increases by two, one new set of C and U parameters appear. Besides, as the maximum order increases, the C and U would become more complex. However, no matter how complex the C and U become, they remain independent of the phases of a_1 and a_2 .

The generic form of the model formulation represents a general methodology for accurately characterizing and modelling a broad range of systems from small signal linear devices to devices with strong nonlinearity such as mixer. The approach is flexible when used to approximate nonlinearity. For devices with different level of nonlinearity, different maximum order n of (5-48) can be chosen in order to meet the required level of accuracy. For example, in order to model a linear network, n is chosen as 1. (5-48) will become:

$$b_1 = C_{1,0}a_1 + U_{1,0}a_2 \quad (5-49)$$

$$b_2 = C_{2,0}a_1 + U_{2,0}a_2 \quad (5-50)$$

(5-49) – (5-50) are essentially the same as the small signal s-parameter formulations. It makes sense because if n is 1, the polynomial will only have linear terms which should be sufficient to describe any linear system. For any nonlinear system, a minimum order of 3 is usually required. Generally speaking, the accuracy of approximation gets better and better as the order of model increases – it will be experimentally verified later in this chapter.

The proposed approach represents a reasonable extension of S parameters which greatly improves the applicability of the classic tool in large signal operation. S-parameters probably are the most useful tool in small signal RF or microwave design. In large signal world, similar tool which can do what s-parameters do in small signal condition still remains to be discovered yet. Recently some works have been made to explore the possible ways to extend the s-parameters into large signal world. Some of the efforts have been reviewed previously in this thesis. Most of

works have agreed that the S parameters are dependent on the large signal input stimuli, which can be described as travelling wave a_1 , input voltage v_1 or input power P_{in} . But it becomes diverse when the relationship between S parameters and output load is investigated. An intuitive extension is to make the S parameters directly a function of load reflection coefficient. This dependency can be clearly observed in large signal measurement and thus represent a direct description of the S parameter large signal behaviour. Although this approach can usually give very accurate simulation, it requires the processing of large amount of data and thus the simulation speed can be very slow. Besides, since no underline relationship between S parameters and load impedance is extracted, the generated model can only be valid within the measurement data range. The extrapolation ability usually is far away from satisfactory.

Some other approaches like original PHD modelling approach extend the S parameters by modifying the formulation a bit as well as cutting the dependency of model parameters on load impedance. These approaches assume the transistor devices are operating at or very close to the load matching point. In this case, the load should have little effect on altering the device performance and only the effect of large signal input stimuli needs to be taken into account. Based on such assumption, proper model formulation is developed and the involved model parameters are a function of input stimuli only. These models show some extrapolation ability but they only work when the assumption about load matching holds.

Unlike the above two kinds of approaches, the proposed modelling approach, presented in this work, extends the S parameters without any compromise. The proposed approach can give speedy and accurate simulation without any restriction on load matching condition. Previous discussion has demonstrated that the original PHD modelling approach can be considered as a small $|a_2|$ approximation of the proposed generic modelling approach when order n is three.

Please note that in this research work the modelling is focused on output at fundamental frequency. However, it doesn't mean the model can only emulate the fundamental behaviours. In fact, the final output fundamental characteristics are resultant from mixing effects of fundamental behaviour and high order harmonic nonlinearities. The effects from large harmonics are already enclosed in the fundamental output characteristics and are adequately modelled by the proposed approach.

5.6. Direct model extraction from large signal waveform data

The accuracy and reliability of a model depends heavily on the extraction strategy and the quality of the measurement data. The reliability of the large signal waveform measurement data obtained from the full-calibrated measurement system has been proved in many ways including comparisons through independent measurements done by the third parties. The ultimate goad of the model extraction is to make most of the measurement data and use the obtained information efficiently for identification of the model parameters. In this section, the careful design of the model extraction strategy is presented in details. Please note the extraction procedure is discussed mainly on the basis of 3rd order polynomial-based model. Extraction of n th order polynomial-based model is essentially the same and will be briefly discussed at the end of this section.

By inspecting the model formulation (5-20) – (5-21) and (5-22) – (5-25), it can be found that there are only four unknown parameters in each of the two model formulas and each of the parameters is a function of magnitudes of both a_1 and a_2 . So the task for model extraction is to identify the relationship between these model parameters and the incident travelling waves a_n . Because the relationship would be described by Cubic Spline function, a data table is constructed accordingly. C and U parameters will be identified at discrete $|a_1|$ and $|a_2|$ points and saved into the data

table. Cubic Spline function would interpolate the proper values of C and U parameters for any $|a_1|$ and $|a_2|$ points not appearing in the table. a_1 can be interpreted as the input power level and a_2 can be interpreted as the load reflection coefficients. Theoretically only four measurement data points are needed to fully extract the model. The four measurement data points should have the same $|a_1|$ and $|a_2|$ but different phases of a_2 . On such a condition, a matrix containing four different equations but constant C and U parameters can be generated. Then the C and U parameters in these equations can be easily calculated with simple matrix operation algorithm. This calculation process can be repeated for different magnitudes of a_1 and a_2 until the data table is fully filled. The specific coverage of such a data table may vary from application to application and is limited to the maximum measurement space as indicated in the waveform measurement data.

It can be seen that the extraction process of the model is fairly straightforward in theory. However, when extracting the model in practice, extra considerations are required to take case of some practical issues which may impede the accuracy of the model extraction. Basically three general steps are required to process the measurement data and extract the model in practice. They are renormalization of the travelling waves, interpolation of the measurement data, and integral optimization of model parameters. Each of them would be explained respectively in following subsections.

5.6.1. Travelling wave renormalization

The travelling waves are frequently used to characterize the two port network characteristics in RF frequency.

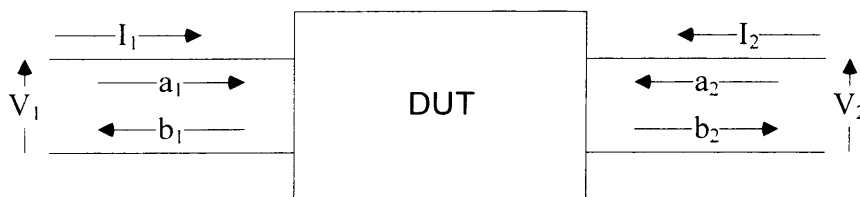


Figure 5-7 The travelling wave characterization versus the voltage and current

characterization for a typical two port network

The travelling waves can be easily obtained from the voltage and current waveforms by using the following conversion equations:

$$a = \frac{V + Z_{\text{ref}} I}{2\sqrt{Z_{\text{ref}}}} \quad (5-51)$$

$$b = \frac{V - Z_{\text{ref}} I}{2\sqrt{Z_{\text{ref}}}} \quad (5-52)$$

Usually the travelling waves are normalized to the system characteristic impedance, which by default is 50 ohm. In other words, Z_{ref} in (5-51) – (5-52) is normally set to 50 ohm. High power PA devices always demonstrate maximum output power at very low load impedance point. The characteristic impedance of our high power large signal measurement system is set to 7.15Ω , which is achieved by using the broadband impedance transformer. This value is normally used as the reference impedance for calculating the travelling waves from current and voltage waveforms. Although this value is much smaller than 50 ohm, it is still relatively far away from the impedance for optimum power. Figure 5-8 shows the optimum impedance point of a 100w LDMOS device on a Smith Chart with the reference impedance at 7.15 ohm. It can be seen the optimum point is still close to the edge of the Smith Chart. It means that high mismatch is expected at the optimum impedance point.

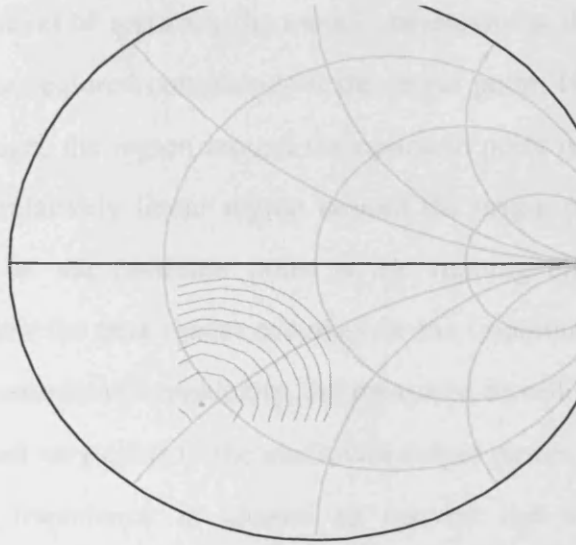


Figure 5-8: The position of typical optimum impedance point

As a result the travelling wave a_2 will have a large magnitude around the optimum power point. Recalling the n th order polynomial approximation as shown in (5-48), it can be observed that as the magnitude of a_2 become larger and larger, the higher order terms become more and more comparable to the lower order terms. It's because generally the higher order terms increase much faster than the lower order terms. As far as approximation is concerned, it becomes quite difficult to truncate the higher order terms if these terms are comparable to the lower ones because the accuracy would suffer a lot otherwise. Therefore, it can be derived that a large magnitude of a_2 will lead to higher order of polynomial required to maintain a certain level of accuracy. Specially, regarding a typical high power transistor device, the truncated polynomial sequence has to be elongated in order to accurately approximate the non-linear interaction around the optimum impedance point due to the high mismatch, hence increasing the model complexity.

The model has least amount of complexity when magnitude of the travelling wave a_2 is zero. Under normal circumstance, a_2 is zero at the system characteristic impedance point, which is the origin point shown in Figure 5-8. Since the magnitude of a_2 influences heavily on the required minimum order of polynomial, it can be seen that

for achieving same level of accuracy the model complexity at the optimum point is much higher than the required complexity at the origin point. However, in terms of power amplifier design, the region around the optimum point is always much more important than the relatively linear region around the origin point. A model with highest complexity at the optimum point is not running efficiently. Therefore, alternatively, to obtain the best model accuracy in the important region around the optimum with least amount of complexity, the measured travelling waves need to be renormalized to a load very close to the maximum output power load. In other words, different reference impedance is chosen to convert the voltage and current waveforms into the travelling waves. After renormalization the magnitude of a_2 at the optimum point should become much smaller than before. Hence the required model complexity at the optimum point can be reduced significantly. Experimental results shown in the later section of this chapter would clearly confirm it.

It's worth noting that selection of Z_{ref} is not that important – it doesn't have to be exact optimum point. A renormalization point close to the optimum region is good enough to reduce the required complexity of model to a reasonable level. Therefore, accurate optimum point information is not a must prerequisite before using the proposed modelling approach.

Renormalization to a point close to the optimum means the reference impedance Z_{ref} becomes a complex number. (5-51) – (5-52) are usually used when the reference impedance is a real number. How to derive travelling waves when the reference impedance is a complex number is still a less explored area. Some complex equations have been proposed in literature which made attempts to maintain the physical meaning of travelling waves when the reference impedance is complex. For example, [80] proposes a set of conversion formulation which is called pseudo-wave based formulation. The equations are reprinted in (5-53) – (5-54).

$$a_{\text{Renorm}} = \frac{V + Z_{\text{opt}} I}{2} \cdot \frac{\sqrt{\text{Re}(Z_{\text{opt}})}}{|Z_{\text{opt}}|} \quad (5-53)$$

$$b_{\text{Renorm}} = \frac{V - Z_{\text{opt}} I}{2} \cdot \frac{\sqrt{\text{Re}(Z_{\text{opt}})}}{|Z_{\text{opt}}|} \quad (5-54)$$

However, as far as this work is concerned, (5-51) – (5-52) can be used because the required renormalization to a complex point is essentially a scaling process. The resultant renormalized travelling waves “a” and “b” are used only internally within the model structure. The model communicates with the external circuit schematic using voltage and current representations. It can be seen that (5-51) – (5-52) can be used for complex renormalization required in this proposed modelling approach

These calculations are performed on the available voltage and current information contained in the waveform data. Figure 5-9 shows an example of the renormalized $|a_2|$ variation over the load reflection coefficient plane. In this example the Z_{opt} is chosen as the reference impedance for renormalization. After renormalization the $|a_2|$ has a minimum value at the Z_{opt} and grows gradually when moving away from the Z_{opt} .

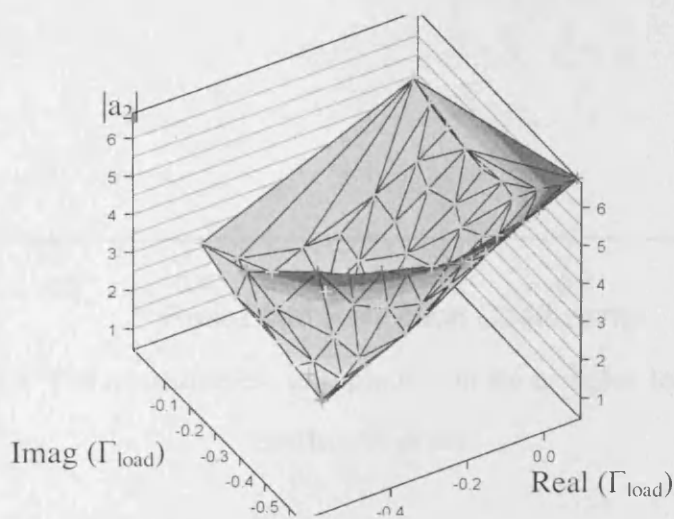


Figure 5-9: Renormalized a_2 magnitude variation against the load reflection coefficients

5.6.2. Data interpolation

Renormalization of the travelling waves benefits the model. But it brings problems too. It reduces the complexity of the model formulation but incurs some complexities for the model extraction. The primary side effect of the renormalization is that the measurement grid will be highly distorted due to the complex reference impedance Z_{ref} used for renormalization.

The large signal waveform measurement system in Cardiff University is normally set up to sweep the load reflection coefficient automatically. The waveform data are collected when the load reflection coefficient is swept in a rectangular format, i.e. the swept area is in a rectangular shape on Smith Chart. Figure 5-10 gives an example of the swept measurement grid. Each point on the grid corresponds to a load reflection coefficient point. It can be observed the measurement grid is regularly shaped.

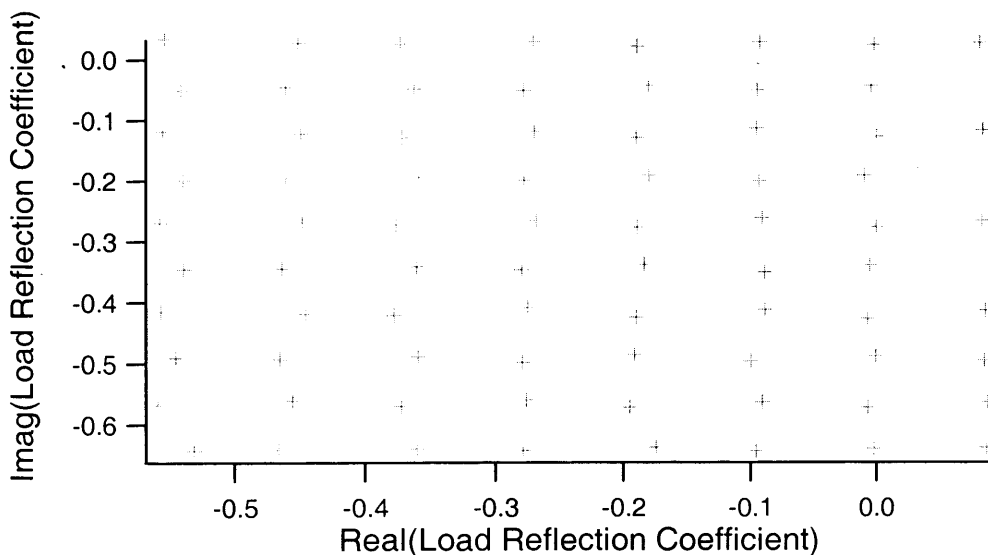


Figure 5-10: The measurement grid plotted on the complex load reflection coefficient plane

The model parameters U and V are actually a function of a_1 and a_2 . It means these model parameters need to be extracted on a_2 plane rather than load reflection

coefficient plane. Figure 5-11 shows the same measurement grid re-plotted on a_2 plane. As shown on the graph, the measurement grid is slightly distorted on a_2 plane and the grid is losing its regular shape. The reason is the load reflection coefficient, which by definition is the ratio of a_2 over b_2 , doesn't share the same variation characteristics as a_2 has.

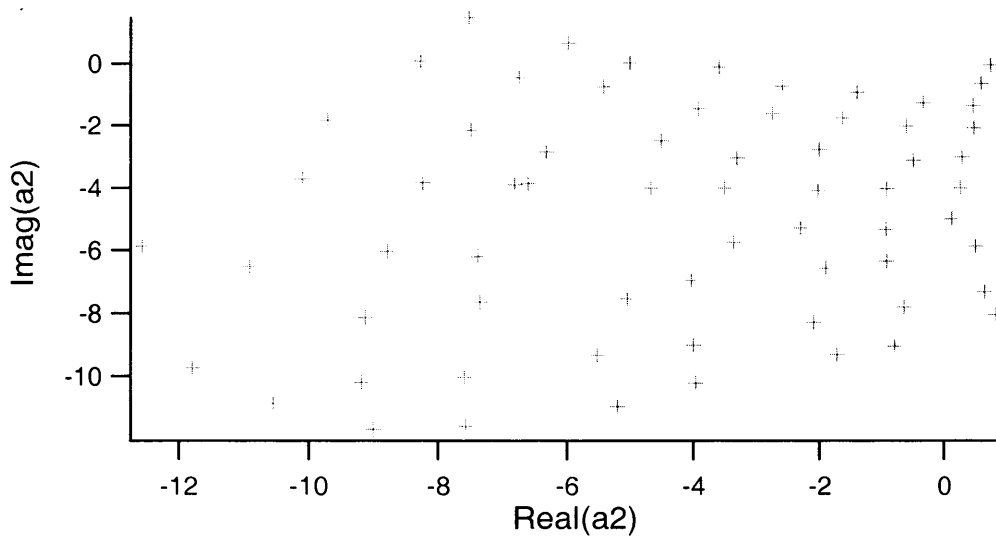


Figure 5-11: The measurement grid plotted on complex a_2 plane

The renormalization of the measured travelling waves further distorts the shape of the measurement grid. An example of the measurement grid on a_2 plane after renormalization is shown in Figure 5-12. By comparing Figure 5-12 to Figure 5-10, it can be spotted that the shape of the measurement grid has been dramatically altered.

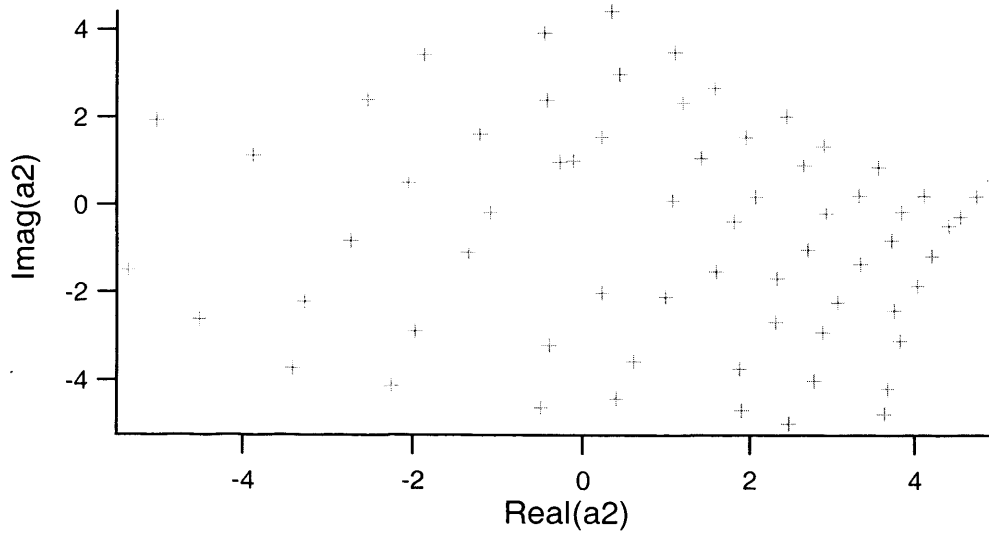


Figure 5-12: The renormalized measurement grid plotted on complex a_2 plane

As previously stated, at least four measurement data points are needed for identifying one set of the eight U and V model parameters. These four data points should have the same $|a_1|$ and $|a_2|$ but different phase of a_2 . Due to the irregular shape of the renormalized data grid shown in Figure 5-12, for certain magnitudes of a_2 , it's not always possible to obtain four measurement data points with exact same $|a_1|$ and $|a_2|$ but different phase of a_2 . Generally speaking, there are two possible ways to solve this problem. One solution is to reconfigure the measurement system to make it able to sweep renormalized a_2 . It would involve lots of reprogramming of the measurement software and will be time-consuming. Besides, generation of a set of waveform data when doing renormalized a_2 sweep doesn't make much sense to most of PA applications apart from this modelling work. Sweep of load reflection coefficient is still desired. Therefore, the data set collected when doing an a_2 sweep can only be solely used for identification of this model. It makes the solution quite expensive and inefficient.

Another possible solution is to use interpolation algorithm to interpolate the surface information at the desired points which are not covered by the measurement grid. The disadvantage of this solution is that the interpolation would certainly introduce

some errors. It may deteriorate the accuracy of the extracted model. The major advantage of the solution is that interpolation algorithm is easy to implement and no change to the measurement system is needed. The same waveform data set with load reflection coefficient sweep can be used. It is a compromised solution balancing among the convenience, efficiency and accuracy. This solution is adopted in this work to extract the model.

Therefore a 2D interpolation algorithm is necessary to interpolate surface information at the desired a_2 points. Specifically, for each point on the a_2 map, the corresponding values of b_1 and b_2 need to be interpolated. The desired interpolation algorithm should be robust to make sure good accuracy of the extracted polynomial-based model can be achieved. Besides, since the target data are nonlinear in nature, a nonlinear interpolation algorithm is thus necessary. It's the point where the direct utilization of waveform approach, as discussed in chapter four, can kick in. The direct data utilization approach has the unique advantages of robust multi-dimensional interpolation, which has been verified comprehensively in experiments. Besides, one bonus of using the direct waveform utilization approach is its seamless integration with the CAD. The required data processing steps such as renormalization of travelling waves and data re-grouping can be conveniently setup in CAD software. It makes the proposed model extraction procedure much easier and more accurate – fully realized in CAD software and no setup alteration is needed on the measurement system hardware or software. Figure 5-13 illustrates the schematic setup in ADS for the measurement data interpolation.

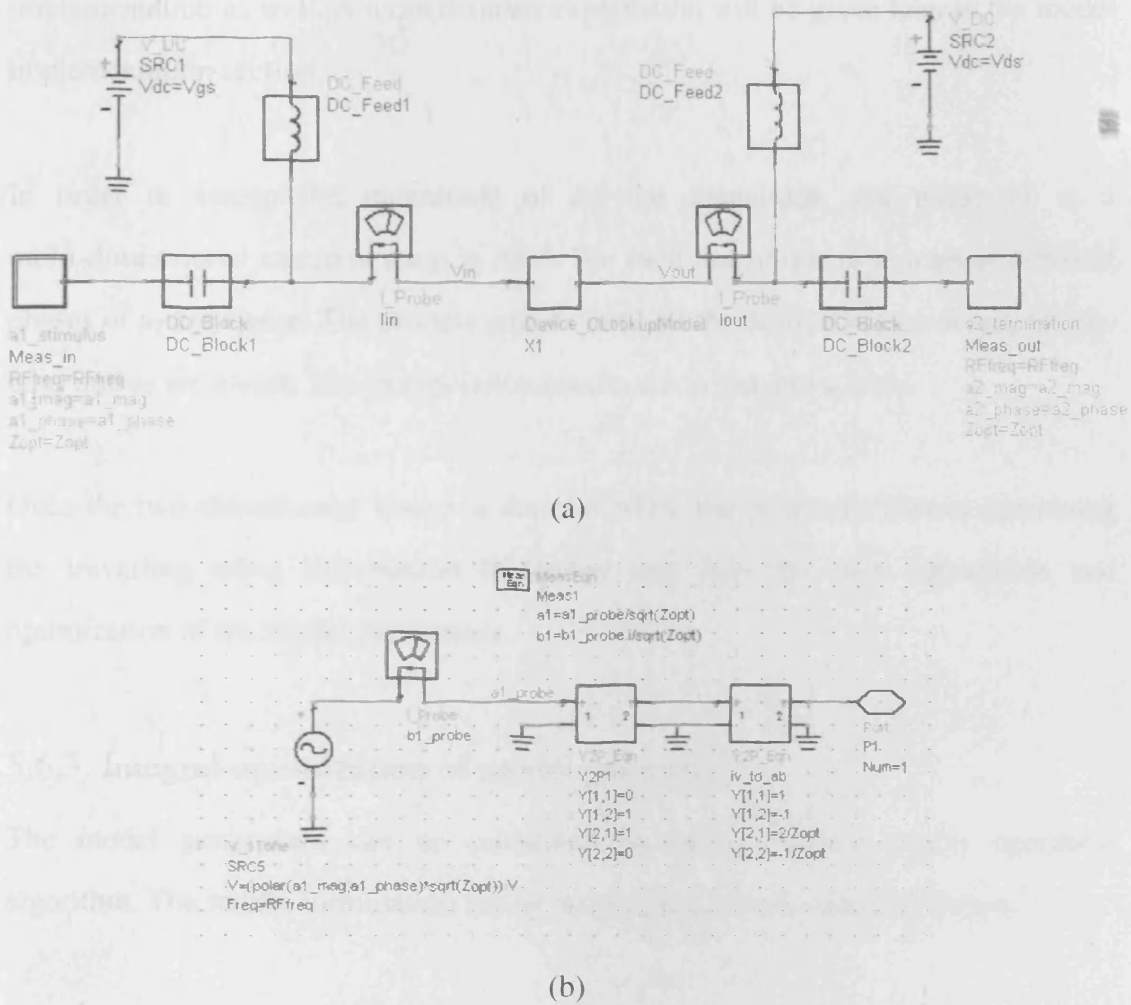


Figure 5-13: Proper data processing schematic has been setup in ADS. Measurement data are imported into ADS using direct waveform utilization approach. (a) shows the general simulation setup schematic while (b) shows the internal circuit schematic of the custom stimulus and termination.

As shown in the schematic, custom stimulus sources enabling excitation of the device directly with incident travelling waves are setup at both input and output sides of the DUT.

Since there is no build-in source for stimulating the device with incident travelling wave directly, a custom source is made by making use of a voltage source in combination with an Y2P_Eqn component. With proper parameter setting the Y2P_Eqn component can serve as a converter to convert the voltage and current waveforms into travelling waves. The same idea will be used in model

implementation as well. A more detailed explanation will be given later in the model implementation section.

In order to sweep the magnitude of a_1 , the magnitude and phase of a_2 a multi-dimensional sweep is setup in ADS. For each magnitude of a_2 a set of different phases of a_2 are swept. The process repeats until all the desired values of magnitudes of a_1 and a_2 are swept. The interpolation results are saved into a table.

Once the two-dimensional sweep is done in ADS, the generated dataset containing the travelling wave information is loaded into Igor for final calculation and optimization of the model parameters.

5.6.3. Integral optimization of model parameters

The model parameters can be calculated by using standard matrix operation algorithm. The model formulation can be written in a matrix operation format:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} MP_{b_1} \\ MP_{b_2} \end{bmatrix} \begin{bmatrix} a_1 \\ a_1 \left(\frac{p}{q} \right)^2 \\ a_2 \\ a_2 \left(\frac{q}{p} \right)^2 \end{bmatrix} \quad (5-55)$$

In (5-55) MP matrix stands for the model parameter matrix. It is expressed as:

$$\begin{bmatrix} MP_{b_1} \\ MP_{b_2} \end{bmatrix} = \begin{bmatrix} U_{1,1} & U_{1,2} & V_{1,1} & V_{1,2} \\ U_{2,1} & U_{2,2} & V_{2,1} & V_{2,2} \end{bmatrix} \quad (5-56)$$

According to the matrix calculation rules, the model parameter matrix can be worked out by using the travelling wave a_n and b_n matrixes. For example, the four data points as shown in Figure 5-14 are selected for calculating one set of C and U parameters.

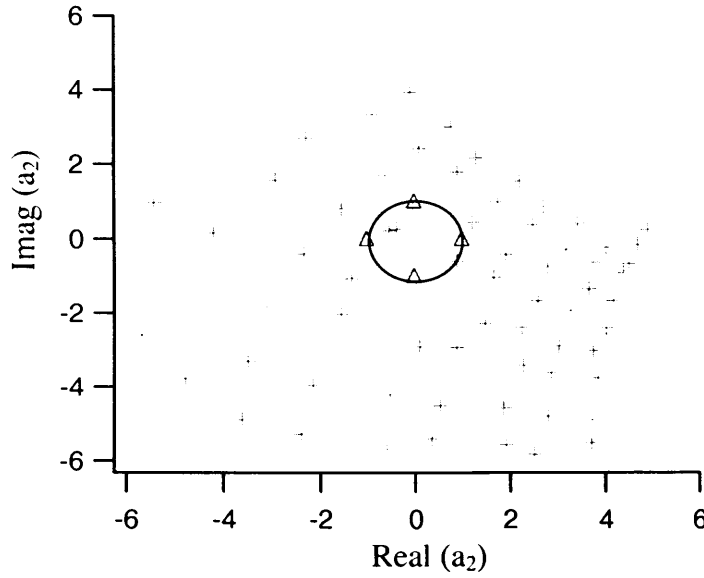


Figure 5-14: The way to select four data points for model extraction

The four data points have the same magnitude of a_2 but different phases. The values of b_1 and b_2 at these points are obtained through interpolation. The values can be written into two matrixes:

$$\begin{bmatrix} b_{1,pt1} & b_{1,pt2} & b_{1,pt3} & b_{1,pt4} \end{bmatrix} \quad (5-57)$$

$$\begin{bmatrix} b_{2,pt1} & b_{2,pt2} & b_{2,pt3} & b_{2,pt4} \end{bmatrix} \quad (5-58)$$

Since the travelling wave b_n is related to travelling wave a_n via (5-55), the following expressions can be obtained:

$$\begin{bmatrix} b_{1,pt1} & b_{1,pt2} & b_{1,pt3} & b_{1,pt4} \end{bmatrix} = \begin{bmatrix} MP_{b_1} \end{bmatrix} \begin{bmatrix} a_n \end{bmatrix} \quad (5-59)$$

$$\begin{bmatrix} b_{2,pt1} & b_{2,pt2} & b_{2,pt3} & b_{2,pt4} \end{bmatrix} = \begin{bmatrix} MP_{b_2} \end{bmatrix} \begin{bmatrix} a_n \end{bmatrix} \quad (5-60)$$

where the $[a_n]$ is:

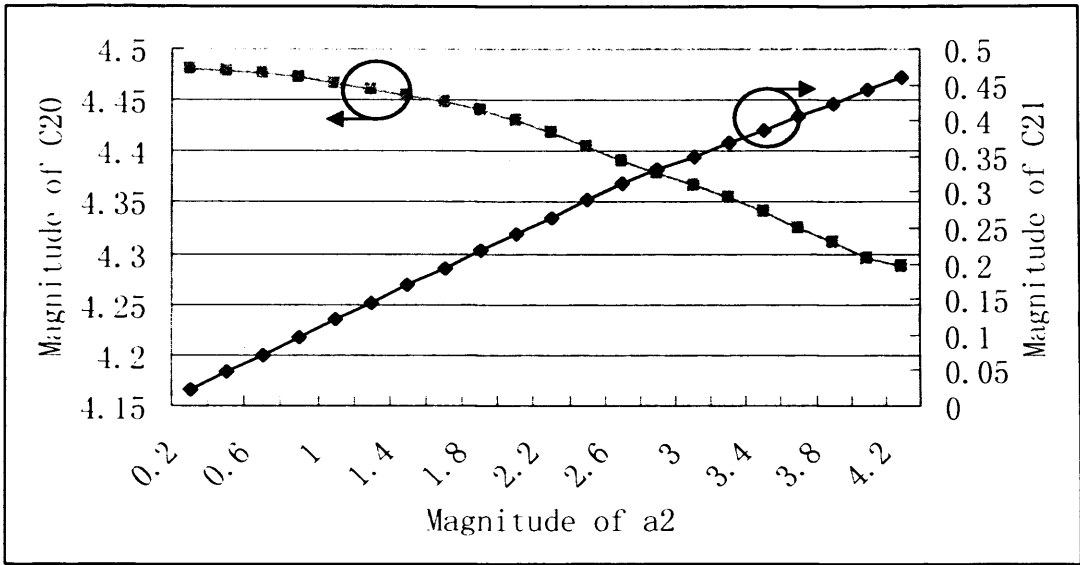
$$\begin{bmatrix} a_{1,pt1} & a_{1,pt2} & a_{1,pt3} & a_{1,pt4} \\ a_{1,pt1} \left(\frac{p_{pt1}}{q_{pt1}} \right)^2 & a_{1,pt2} \left(\frac{p_{pt2}}{q_{pt2}} \right)^2 & a_{1,pt3} \left(\frac{p_{pt3}}{q_{pt3}} \right)^2 & a_{1,pt4} \left(\frac{p_{pt4}}{q_{pt4}} \right)^2 \\ a_{2,pt1} & a_{2,pt2} & a_{2,pt3} & a_{2,pt4} \\ a_{2,pt1} \left(\frac{q_{pt1}}{p_{pt1}} \right)^2 & a_{2,pt2} \left(\frac{q_{pt2}}{p_{pt2}} \right)^2 & a_{2,pt3} \left(\frac{q_{pt3}}{p_{pt3}} \right)^2 & a_{2,pt4} \left(\frac{q_{pt4}}{p_{pt4}} \right)^2 \end{bmatrix} \quad (5-61)$$

The model parameters can then be obtained through the following matrix

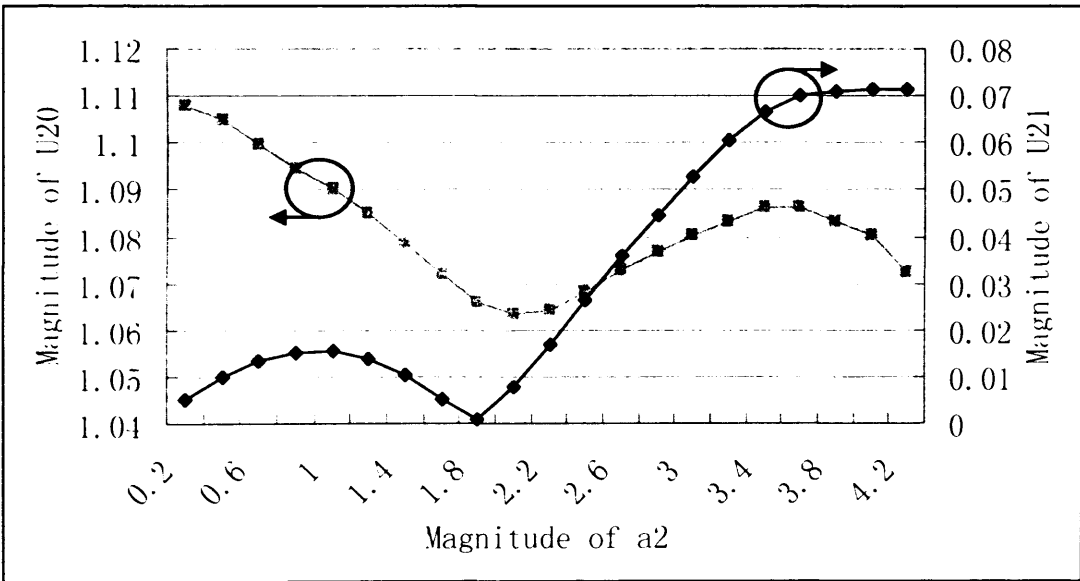
calculation:

$$[MP] = [b][a_n]^{-1} \quad (5-62)$$

The obtained model parameters correspond to a specific value of $|a_2|$. The above process needs to be repeated for different value of $|a_2|$ in order to cover the whole $|a_2|$ range. The following graphs plot the extracted model parameters against the magnitude of a_2 . The results are obtained on a 100w LDMOS device.



(a)



(b)

Figure 5-15: (a) The extracted C_{20} and C_{21} as a function of magnitude of a_2 . (b) The extracted U_{20} and U_{21} as a function of magnitude of a_2 .



It can be observed that C and U parameters vary with magnitude of a_2 although the variation of C_{20} and U_{20} is relatively small. However for regions close to optimum a_1 and a_2 are relatively large and a small variation of C_{20} and U_{20} can still cause significant errors at the output.

The model parameters shown above are extracted with minimum number of data points, i.e. only four data points for each magnitude value of a_2 . The robustness of such an extraction needs to be carefully investigated in order to determine whether redundancy should be added. It's because the measurement data usually contain some noise even though the data are indeed quite reliable. Furthermore, the interpolation would unavoidable introduce some errors. All these factors may impede the accuracy of the extraction and thus call for the need for using redundancy. Following experiments are designed to look into this issue.

Five different groups of data points are selected. Each of the groups contains four data points. The data points in all these five groups have the same magnitude of $|a_2|$. These groups differentiate each other by having different phase distributions. The model parameters are then extracted several times by using each of groups of data points respectively. The above procedure is repeated for different magnitude of a_2 until all the magnitude points of interest are covered. Figure 5-16 plots all the five traces of extracted model parameters by using these groups of data points

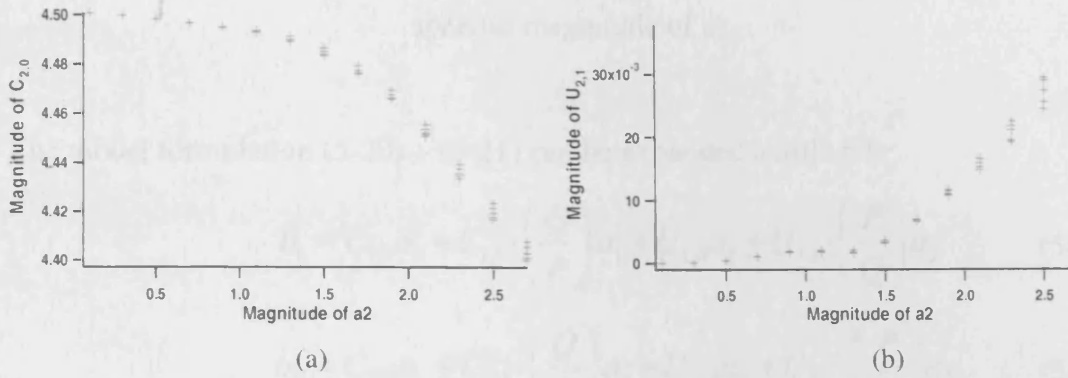


Figure 5-16: (a) Magnitude of $C_{2,0}$ variation against $|a_2|$. (b) Magnitude of $U_{2,1}$ variation against $|a_2|$.

An important observation is that C and U parameters exhibit little variation when extracted for different phases of a_2 . This verifies the assumption that C and U parameters are indeed only a function of a_2 magnitude and independent of a_2 phase. However it should be mentioned that the small variation versus the a_2 phase is likely due to the noise and data processing errors. The reliability of the model extraction could be further improved through the utilization of an optimization algorithm that allows a single coefficient value to be obtained from more than 4 data points, as shown in Figure 5-17.

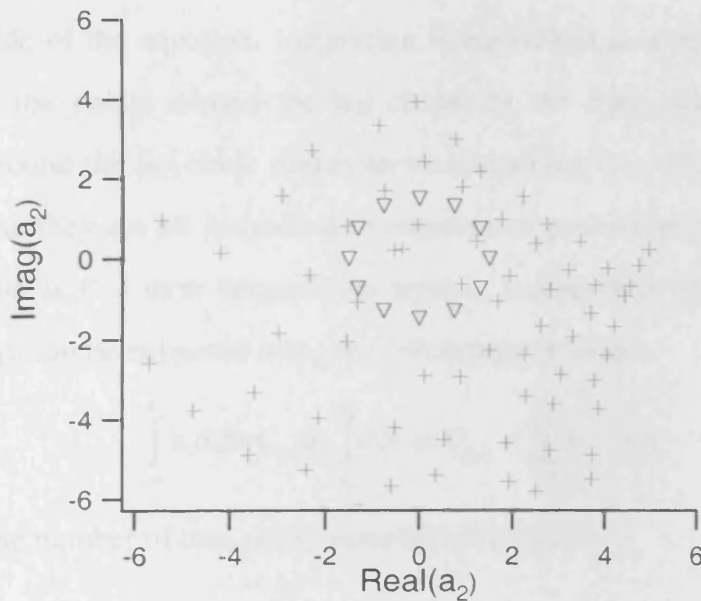


Figure 5-17: Using more than four data points to extract the model parameters for a

specific magnitude of a_2

The model formulation (5-20) – (5-21) can be expanded a little bit:

$$b_1 = C_{1,0}a_1 + C_{1,1} \cdot \left(\frac{Q}{P}\right)a_1 + U_{1,0}a_2 + U_{1,1} \cdot \left(\frac{P}{Q}\right)a_2 \quad (5-63)$$

$$b_2 = C_{2,0}a_1 + C_{2,1} \cdot \left(\frac{Q}{P}\right)a_1 + U_{2,0}a_2 + U_{2,1} \cdot \left(\frac{P}{Q}\right)a_2 \quad (5-64)$$

Note P is the phase operator of a_2 , (5-63) – (5-64) can be re-arranged to extract the phase operator P from remaining terms:

$$b_1 = C_{1,0}a_1P^{-1} + C_{1,1}a_1QP^{-1} + U_{1,0}|a_2|P + U_{1,1}|a_2|Q^{-1}P^{-1} \quad (5-65)$$

$$b_2 = C_{2,0}a_1P^{-1} + C_{2,1}a_1QP^{-1} + U_{2,0}|a_2|P + U_{2,1}|a_2|Q^{-1}P^{-1} \quad (5-66)$$

An interesting characteristic of the formula can be easily recognized – each of the terms contains different order of phase operator p . By taking advantage of this property, integration can be used to extract the model parameters from all the points sampled around the $|a_2|$ circle in Figure 5-17. For example, if doing integration along the phase of a_2 around the $|a_2|$ circle at both sides of (5-65):

$$\int_{-\pi}^{\pi} b_1 d\varphi = \int_{-\pi}^{\pi} C_{1,0}a_1 + C_{1,1}a_1 \frac{Q}{P} + U_{1,0}|a_2|P + U_{1,1}|a_2|\frac{P^{-1}}{Q} d\varphi \quad (5-67)$$

At the left side of the equation, integration is equivalent to a summation of the b_1 values at all the points around the $|a_2|$ circle. At the right side of the equation, integration around the $|a_2|$ circle makes terms containing $C_{1,1}$, $U_{1,0}$ and $U_{1,1}$ become zero – because they are all coupled with negative or positive phase operator P . The only exception is $C_{1,0}$ term because the term is independent of phase operator P . Therefore, $C_{1,0}$ can be extracted using the following equation:

$$\int_{-\pi}^{\pi} b_1 d\varphi = C_{1,0}a_1 \int_{-\pi}^{\pi} d\varphi \Rightarrow C_{1,0} = \sum_{-\pi}^{\pi} b_1 / Na_1 \quad (5-68)$$

Where N is the number of data points sampled on $|a_2|$ circle.

The remaining C and U parameters can be extracted using similar method. For

example, if multiplying both sides of (5-65) with p^{-2} then do integration:

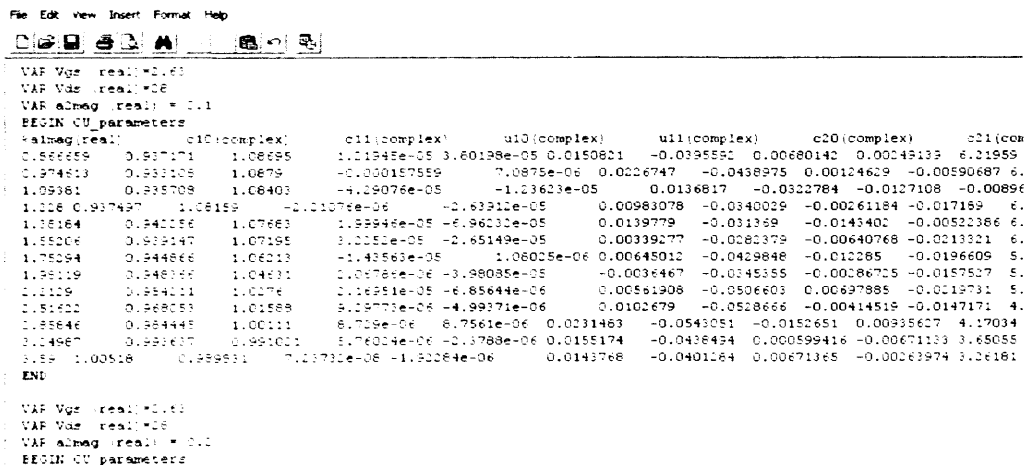
$$\int_{-\pi}^{\pi} b_1 P^{-2} d\varphi = \int_{-\pi}^{\pi} C_{1,0} a_1 P^{-1} + C_{1,1} a_1 Q P + U_{1,0} |a_2| P^{-1} + U_{1,1} |a_2| Q d\varphi \tag{5-69}$$

$U_{1,1}$ term becomes the only one independent of phase operator P . Integration at both sides of (5-69) would extract the value of $U_{1,1}$, as illustrated in (5-70)

$$\int_{-\pi}^{\pi} b_1 P^{-2} d\varphi = U_{1,1} a_2 Q \int_{-\pi}^{\pi} d\varphi \Rightarrow U_{1,1} = (\sum b_1 P^{-2}) / N a_2 Q \tag{5-70}$$

Similarly, by multiplying different order of P on both sides of (5-65) – (5-66), other C & U parameters can be extracted. By using such a method, all the valid data points can be taken into account for extraction of the model parameters. The measurement noise and interpolation errors can be effectively reduced.

The integration method is also implemented in Igor Pro software. The extracted C and U parameters are compiled into a two-dimensional table where the dimensions correspond to magnitude of a_1 and magnitude of a_2 respectively. The data format is MDIF, which is chosen because of its simplicity and easy recognition by ADS. An example of the data table is shown below:



```
File Edit View Insert Format Help
[Icons]
VAF Vgs real:=0.60
VAF Vds real:=0.6
VAF aOmeg real:= 0.1
BEGIN CU_parameters
+almag,real:      c10(complex)      c11(complex)      u10(complex)      u11(complex)      c20(complex)      c21(com
0.556659  0.937171  1.06695  1.11945e-05  3.60198e-05  0.0150801  -0.0395591  0.00680142  0.00149133  6.21955
0.974613  0.933108  1.0679  -0.000157559  7.0875e-06  0.0226747  -0.0438975  0.00124629  -0.00590687  6.
1.09361  0.935708  1.06403  -4.129076e-05  -1.123623e-05  0.0136817  -0.0322784  -0.0127108  -0.00896
1.006  0.937497  1.06159  -2.10107e-06  -2.63912e-05  0.00983078  -0.0340029  -0.00261184  -0.017159  6.
1.35164  0.940056  1.07663  1.99946e-05  -6.96030e-05  0.0139779  -0.031359  -0.0143402  -0.00522386  6.
1.55206  0.939147  1.07195  3.0052e-05  -2.65149e-05  0.00339277  -0.0282379  -0.00640768  -0.0213321  6.
1.75094  0.944866  1.06013  -1.43563e-05  1.06025e-06  0.00645012  -0.0429848  -0.010285  -0.0196609  5.
1.95119  0.948196  1.04831  2.05785e-06  -3.98035e-05  -0.0036467  -0.0345355  -0.00286725  -0.0157527  5.
2.1529  0.954211  1.0276  2.16951e-05  -6.85644e-06  0.00561908  -0.0506603  0.00697885  -0.0219731  5.
2.35422  0.968053  1.01568  5.18773e-06  -4.99371e-06  0.0102679  -0.0528666  -0.00414519  -0.0147171  4.
2.55646  0.984445  1.00111  6.709e-06  8.7561e-06  0.0231463  -0.0543051  -0.0152651  0.00335627  4.17034
3.14967  0.999637  0.991021  5.76024e-06  -2.3788e-06  0.0155174  -0.0436434  0.000599416  -0.00671133  3.65055
3.5= 1.00518  0.999931  0.991732  7.12173e-06  -1.30264e-06  0.0143768  -0.0401264  0.00671365  -0.00263974  3.26181
END
VAF Vgs real:=0.60
VAF Vds real:=0.6
VAF aOmeg real:= 0.1
BEGIN CU_parameters
```

Figure 5-18: The extracted C and U parameter table

The above data table is extracted for a 3rd order model. It's worth mentioning that nth order model can also be extracted by following the three steps described above. The

travelling wave renormalization and interpolation implementation are basically identical regardless of what the order of model is. The only factor which varies depending on the order of model is within the optimization procedure. Fortunately, the integration method adopted in model optimization procedure makes the variation very small. Referring to the generic model formulation, 5th and 7th order model formulations can be derived:

5th order model:

$$b_n = (C_{n,0} + C_{n,1} \cdot \left(\frac{Q}{P}\right) + C_{n,2} \cdot \left(\frac{Q}{P}\right)^2)a_1 + (U_{n,0} + U_{n,1} \cdot \left(\frac{P}{Q}\right) + U_{n,2} \cdot \left(\frac{P}{Q}\right)^2)a_2 \quad (5-71)$$

$$b_n = C_{n,0}a_1P^0 + C_{n,1}a_1Q^{P^{-1}} + C_{n,2}a_1Q^{2P^{-1}} + U_{n,0}|a_2|P^0 + U_{n,1}|a_2|Q^{-1}P^1 + U_{n,2}|a_2|Q^{-2}P^2 \quad (5-72)$$

)

7th order model:

$$b_n = (C_{n,0} + C_{n,1} \cdot \left(\frac{Q}{P}\right) + C_{n,2} \cdot \left(\frac{Q}{P}\right)^2 + C_{n,3} \cdot \left(\frac{Q}{P}\right)^3)a_1 + (U_{n,0} + U_{n,1} \cdot \left(\frac{P}{Q}\right) + U_{n,2} \cdot \left(\frac{P}{Q}\right)^2 + U_{n,3} \cdot \left(\frac{P}{Q}\right)^3)a_2 \quad (5-73)$$

)

$$b_n = C_{n,0}a_1P^0 + C_{n,1}a_1Q^{P^{-1}} + C_{n,2}a_1Q^{2P^{-1}} + C_{n,3}a_1Q^{3P^{-1}} + U_{n,0}|a_2|P^0 + U_{n,1}|a_2|Q^{-1}P^1 + U_{n,2}|a_2|Q^{-2}P^2 + U_{n,3}|a_2|Q^{-3}P^3 \quad (5-74)$$

It can be observed from (5-72) – (5-74), the characteristic of the model formulation, different terms have different order of phase operator P, is maintained regardless of the order of model. Therefore, same integration method can be used to extract these high order model parameters by multiplying both sides of equation with the corresponding positive or negative order of phase operator P.

Besides, it can also be derived that the increased order of model doesn't change the values of the low order model parameters. For instance, no matter which model formulation, 3rd order's (5-20) – (5-21), 5th order's (5-71) or 7th order's (5-73) is chosen for model extraction for a specific transistor device, C_{1,0}, which is the first

order model parameter, would remain the same. This feature is preserved for all the model parameters extracted. This characteristic gives the model the edge over the traditional parameter-fitting models. For these kinds of traditional models, adding new parameters or altering the model formulation usually invalidates the currently extracted model parameters. Re-extracting all the model parameters is normally necessary. It causes inconveniences and has negative impact on the expandability of models. In contrast, the model presented in this work can be easily expanded into higher order without the need to tune the values of existing low order model parameters. Expanding the model order is as easy as adding a supplement data table containing these extra high order parameters to the model.

The extraction of the n th order model can be done by using a generic model extraction equation as shown below:

$$C_{k,m} = \frac{1}{N_{pt}} \sum_0^{N_{pt}} \frac{b_k(n_{pt})}{\left(\frac{Q}{P}\right)^m \cdot Q \cdot |a_1|} \quad U_{k,m} = \frac{1}{N_{pt}} \sum_0^{N_{pt}} \frac{b_k(n_{pt})}{\left(\frac{P}{Q}\right)^m \cdot P \cdot |a_1|} \quad (5-75)$$

In (5-75), k is the index of ports. m is the index of model parameters. m also represents the order of the corresponding model parameter. N_{pt} is the number of data points used in integration. Each of the data points possesses different phase of a_2 but have identical magnitude of a_2 . $b_k(n_{pt})$ represents travelling wave b_k at data point indexed with n_{pt} .

5.7. Polynomial-based model implementation

Integration of model into standard CAD software is an important step towards realisation of the benefits of model in power amplifier design. A theoretically accurate model with over complicated implementation and slow simulation speed can only offer limited advantage to PA designers in practice. Efficient implementation of model is important to release the full potential of a model.

The model proposed in this work is successfully implemented in CAD software. Implementation details shown next in this section will demonstrate the proposed model can be effectively integrated into popular RF design CAD tool without too many hassles. Agilent **Advanced Design System** (ADS) is chosen as the CAD tool in which the proposed model is implemented. It provides special build-in modules for users to implement their own models in ADS. As discussed in previous chapter, there is one build-in module called Frequency Domain Device (FDD) which allows users to define mathematical functions in terms of voltage and current in frequency domain to relate the output to the input of FDD. It can be seen that FDD component is ideal for implementing a behavioural model in frequency domain. Therefore, FDD is selected to realise the behavioural model proposed in this work. Figure 5-19 gives an overview of the schematic implementation of the proposed behavioural model in ADS.

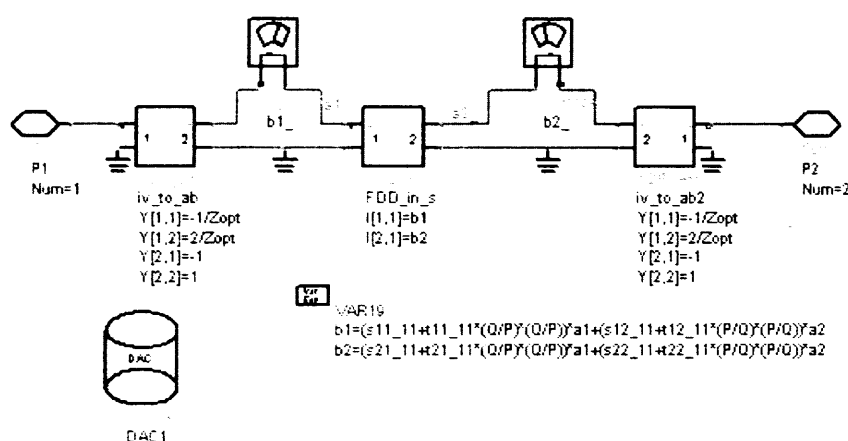


Figure 5-19: The implementation of the model in ADS

As shown in Figure 5-19, the implementation of the model involves several different virtual components including one FDD component, two Y2P_Eqn components and one Data Access Component (DAC).

A two port FDD component, which should be adequately able to represent any typical two port power amplifier configuration of transistor devices, is used in the model implementation. Since the model developed so far is used to investigate the

fundamental output characteristics as a function of fundamental input, only two port-definition equations are needed to define the device behaviours at input and output ports respectively. The proposed model formulation (5-48) can be directly written in the schematic as the port-definition equations.

However, there are some data formality problems in ADS model implementation. Equations shown in Figure 5-19 may look a bit confusing when the fundamental discrepancy in the equations is spotted – model formulation (5-48) is defined using travelling waves a_n and b_n while build-in FDD component can only accept custom frequency domain equation in terms of voltage and current waveforms. Besides, the simulator generally operates with voltage and current waveforms, which means a hybrid system (Figure 5-20) which involves both voltage and current waveform representation and travelling wave representation is required. A certain supplement conversion module is required to realize such system in ADS.

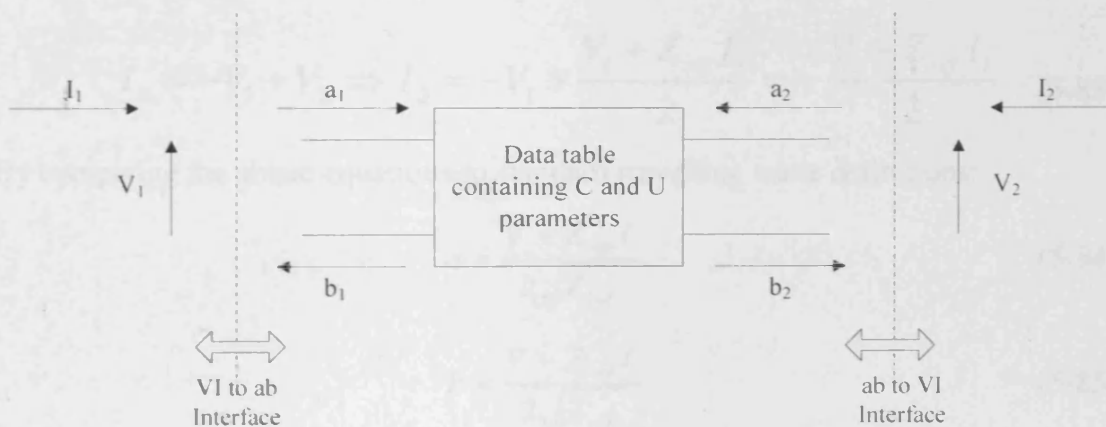


Figure 5-20: The hybrid system structure for implementing the travelling wave based modelling approach in waveform-based simulator

Two Y2P_eqn modules [ref] are used at both input and output of FDD to solve the discrepancy and convert the data from voltage and current format into travelling wave a and b format. Y2P_eqn is normally used to describe the behaviour of a two-port network with Y parameters:

$$I_1 = Y_{11}V_1 + Y_{12}V_2 \quad (5-76)$$

$$I_2 = Y_{21}V_1 + Y_{22}V_2 \quad (5-77)$$

The above mathematical relationship between port one and port two can be utilized to do the conversion between voltage and current waveforms and travelling waves. If the Y parameters are defined with following values:

$$Y_{11} = -\frac{1}{Z_{ref}} \quad (5-78)$$

$$Y_{12} = \frac{2}{Z_{ref}} \quad (5-79)$$

$$Y_{21} = -1 \quad (5-80)$$

$$Y_{22} = 1 \quad (5-81)$$

Substituting the values into equation (5-76) and (5-77):

$$I_1 = -\frac{1}{Z_{ref}}V_1 + \frac{2}{Z_{ref}}V_2 \Rightarrow V_2 = \frac{V_1 + Z_{ref}I_1}{2} \quad (5-82)$$

$$I_2 = -V_1 + V_2 \Rightarrow I_2 = -V_1 + \frac{V_1 + Z_{ref}I_1}{2} = -\frac{V_1 - Z_{ref}I_1}{2} \quad (5-83)$$

By comparing the above equations to standard travelling wave definitions:

$$a = \frac{V + Z_{ref}I}{2\sqrt{Z_{ref}}} \quad (5-84)$$

$$b = \frac{V - Z_{ref}I}{2\sqrt{Z_{ref}}} \quad (5-85)$$

Since the square root of Z_{ref} is a constant number, it can be found that mathematically V_2 , which is the voltage at port two, is equivalent to the normalized travelling wave a at port one. I_2 , which is the current at port one, is equivalent to the normalized travelling wave b at port one. The other Y2P_eqn component does the same trick at the output side of FDD. V_2 and I_2 , which essentially represent the travelling waves at port one, are used in the travelling wave based model formulation defined in FDD.

It can be seen that the two Y2P_eqn components serve as the interface between inside and outside world. Anything inside is defined in terms of travelling waves. Anything outside is defined in terms of voltage and current waveforms. The information is exchanged between inside and outside world via Y2P_eqn components.



DataAccessComponent

DAC7

File="CUmodel_PHD_Modela1a2_5th.mdf"

Type=Generalized Multi-dimensional Data

InterpMode=Cubic Spline

InterpDom=Rectangular

ExtrapMode=Constant Extrapolation

iVar1="a1mag"

iVal1=a1mag

iVar2="a2mag"

iVal2=a2mag

iVar3="Vgs"

iVal3=2.63

iVar4="Vds"

iVal4=28

Var
Eqn

VAR

VAR36

u10=file{DAC7, "u10"}

u11=file{DAC7, "u11"}

Var
Eqn

VAR

VAR37

u20=file{DAC7, "u20"}

u21=file{DAC7, "u21"}

Var
Eqn

VAR

VAR35

c10=file{DAC7, "c10"}

c11=file{DAC7, "c11"}

Var
Eqn

VAR

VAR38

c20=file{DAC7, "c20"}

c21=file{DAC7, "c21"}

Figure 5-21: DAC graph and related definitions

DAC component is used to look up the U and V parameters used in the FDD equations. The extracted model data table containing these U and V parameters as a function of magnitudes of a_1 and a_2 is loaded up using the DAC as the database for the parameter looking-up. Specific interpolation algorithm can be setup in DAC. As previously stated, Cubic Spline function is chosen for interpolation due to its superior robustness when dealing with nonlinearities. Interpolation would be initiated when the simulated $|a_1|$ or $|a_2|$ doesn't match any point in the data table.

5.8. Model experimental verification

It's important to test and investigate a model's ability in simulations and make comparisons against measurement data. Preferably the measurement and model simulation done for model verification should be obtained when operating the device (or corresponding model) as close to realistic working condition as possible. Since the proposed model presented in this work is extracted directly from large signal waveform measurement, which is collected when the device is operating in realistic conditions, no other type of measurement set-up is required for model verification. Large signal waveform measurement data are good sources for both model extraction and verification.

The modelling approach has been validated on some different types of power transistor devices including 100w LDMOS device. Such a type of device at this relatively high power level is usually used for wireless infrastructure base-station applications. This 100w LDMOS device is chosen to do the experimental verification of the presented model in order to demonstrate the model superior potential in CAD-based high power PA design.

The modelling approach has also been validated on 4w LDMOS device which demonstrates high levels of harmonics when driven into compression. The strong nonlinearities can be used to investigate the robustness of the proposed modelling approach on dealing with nonlinear problems.

The ADS simulation schematic setup is illustrated in Figure 5-22. The simulation setup is pretty basic with the aim to test the CW simulation. The input power and load impedance are adjustable allowing the model's performance be tested in a range. Piecewise definition of the load impedance, which was firstly presented in chapter four, is also used here in order to specify the fundamental and harmonic impedances respectively.

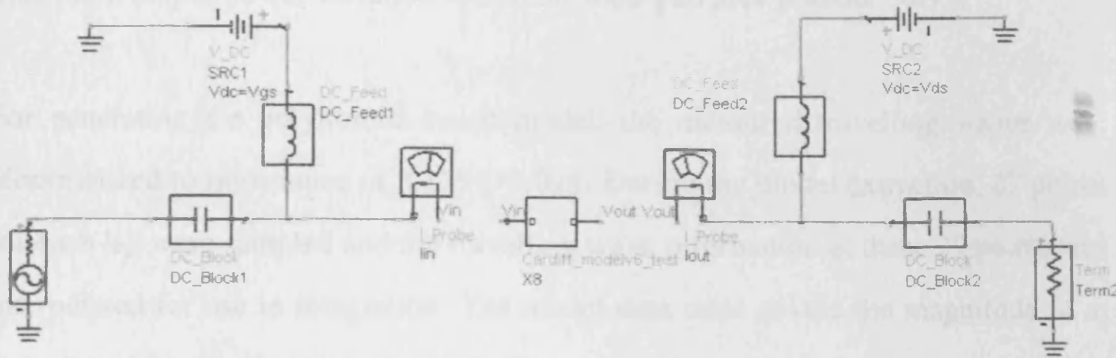


Figure 5-22: The simulation schematic setup for testing the model

Various experiments were setup to test the model thoroughly. First of all, ordinary model test was performed which is basically a list of comparisons between the measured device characteristics and the simulated ones. The comparison covers these standard figures of merits such as AM to AM, AM to PM and output power contours. The comparisons are made both quantitatively and qualitatively through visual comparison of the plots and calculation of difference. Since the model currently takes care of the relationship between fundamental input and output, all the comparisons are done with fundamental characteristics. No waveform or harmonic characteristic simulation is done. After the standard model testing, the relationship between the accuracy of the model and order of the model is investigated. Higher order model parameters were extracted. Finally, the robustness of the model extraction procedure and model implementation is tested. The experiment procedure and final results are presented in details.

5.8.1. Simulation versus measurement on 100w LDMOS

In this subsection, the simulation results of the model are compared to measurements. The measurement data sets for model extraction were collected by using the large signal waveform measurement system in Cardiff University. The measurement system has the resolution at 0.1 dB. The DUT was excited by CW signal during measurement. The device was characterized at 2.15 GHz. The input power sweep was done with a range from 26dBm to 40dBm with a step size of 1dB. The load reflection coefficient sweep was done with an 8 by 10 grid on the Smith Chart. The

maximum output power variation within the load-pull area is about 3dB.

For generating the polynomial-based model, the measured travelling waves were renormalized to impedance of $3.139-j*3.014$. During the model extraction, 27 points for each $|a_2|$ were sampled and the travelling wave information at these 27 points was interpolated for use in integration. The model data table covers the magnitude of a_2 from 0 to 4.2 with a resolution of 0.2. The model data table covers the magnitude of a_1 from 0.7 to 3.9 with a resolution of 0.2 as well. A 7th order model with 16 parameters is extracted. The model parameter table takes up only 220 KB space.

Firstly the simulated output power contours are compared to the measured results. Figure 5-23 illustrates the coverage and position of the output power contours on Smith Chart. It should be noticed the reference impedance of the Smith Chart is 7.15 ohm.

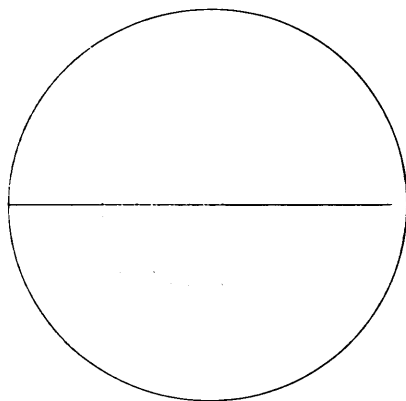


Figure 5-23: The swept area in simulation.

Figure 5-24 shows the simulated and measured output power contours at different output power levels. A “zoomed-in” version is plotted for clear comparison.

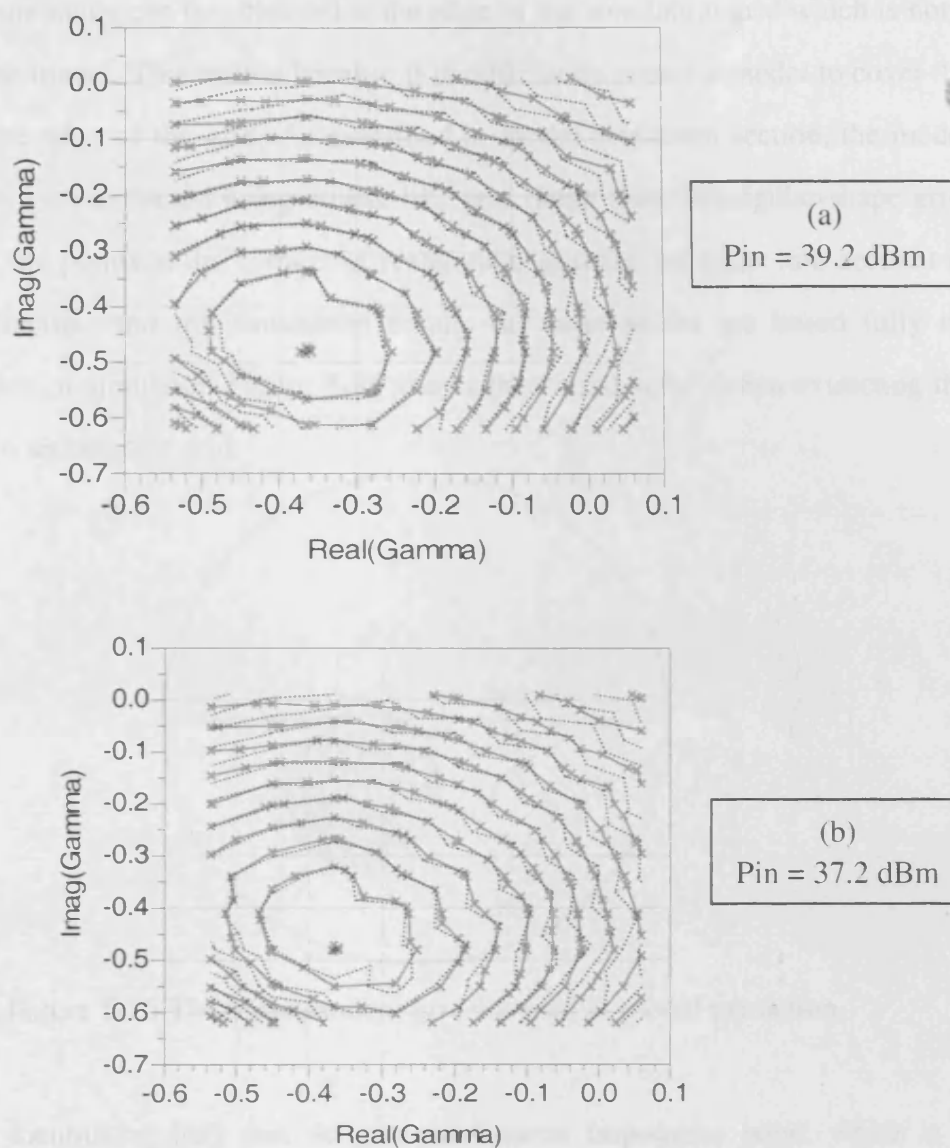


Figure 5-24: The simulated (solid) and measured (dotted) output power contours shown on the Smith Chart with a reference impedance of 7.15Ω . The step size between every two contour lines is 0.25dB. (a) and (b) show the simulation results obtained on different input power levels.

In graph A, the maximum output power at the centre of the contours is around 50dBm with approximate 2dB gain compression. In graph B, the maximum output power at the centre of the contours is around 49 dBm with approximate 1 dB gain compression.

A little discrepancy can be observed at the edge of the simulation grid which is not a result of the model. This results because it is difficult to extract a model to cover the points at the edge of the grid. As described in model extraction section, the model parameters were extracted using ellipse-like grid rather than rectangular-shape grid. Therefore, the points at the corners of rectangular grid are not taken into account in model extraction and the simulation results on these points are based fully on extrapolation of simulator. Figure 5-25 shows the “blind zone” when extracting the model from rectangular grid.

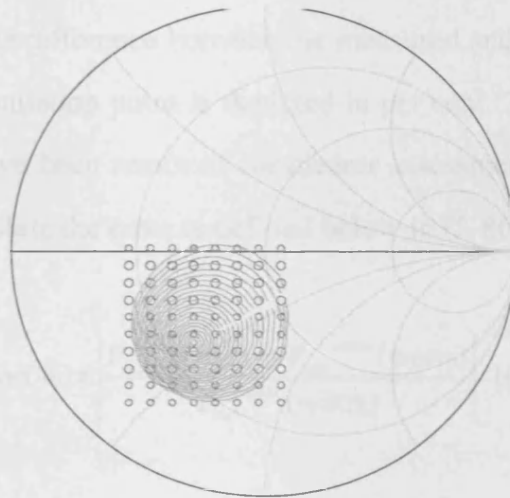


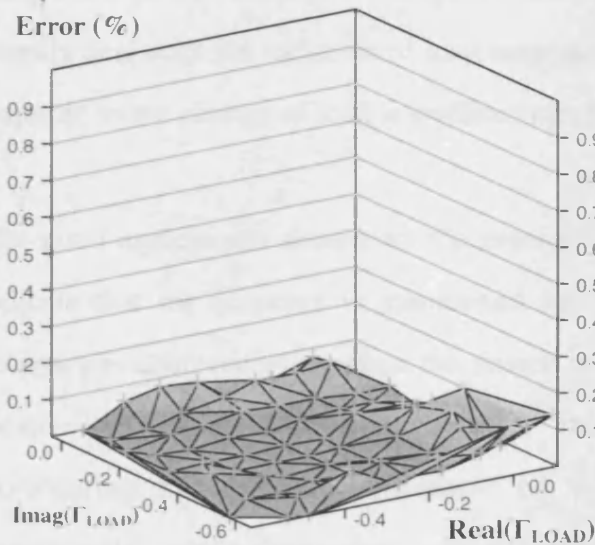
Figure 5-25 The measurement grid covered in model extraction

It's worth mentioning here that the renormalization impedance point, which is at $3.139-j*3.014$, is slightly off the optimum load impedance point of this device. This point is chosen to do the renormalization instead of optimum load point is mainly due to practical considerations – using this point can maximize the coverage of model extraction on the measurement grid available.

Clearly, the comparison shows a very good agreement with the measurement data. As the output power increases, the gain compression goes deeper. But the accuracy of the model is kept at a good level all the time. Besides, good accuracy is generally maintained from the maximum output power point to the edge of the contours on all these plots.

Note that the contour plots show very valuable information but they need to be carefully interpreted to avoid misleading conclusion. The “measured” contour plot is actually calculated by ADS simulator based on the information collected from n -by- m measurement as shown in Figure 5-23. The “simulated” contour plot is actually calculated from the model directly. Therefore such a discrepancy in plot calculation may cause disagreement on the contour plots which is not resulted from the modelling approach itself. In order to quantify the accuracy of the output power contour simulation more accurately, error surfaces, shown in Figure 5-25, has been calculated in which the difference between the measured and simulated output power at each individual simulation point is depicted in per cent. The impedance points in the “blind region” have been removed for clearer assessment of model quality. The equation used to calculate the error is defined below in (5-86):

$$\text{Error}(\%) = \left| \frac{P_{\text{out}}^{\text{sim}} (\text{watts}) - P_{\text{out}}^{\text{meas}} (\text{watts})}{P_{\text{out}}^{\text{meas}} (\text{watts})} \right| \cdot 100\% \quad (5-86)$$



(a)
 $P_{\text{in}} = 39.2 \text{ dBm}$

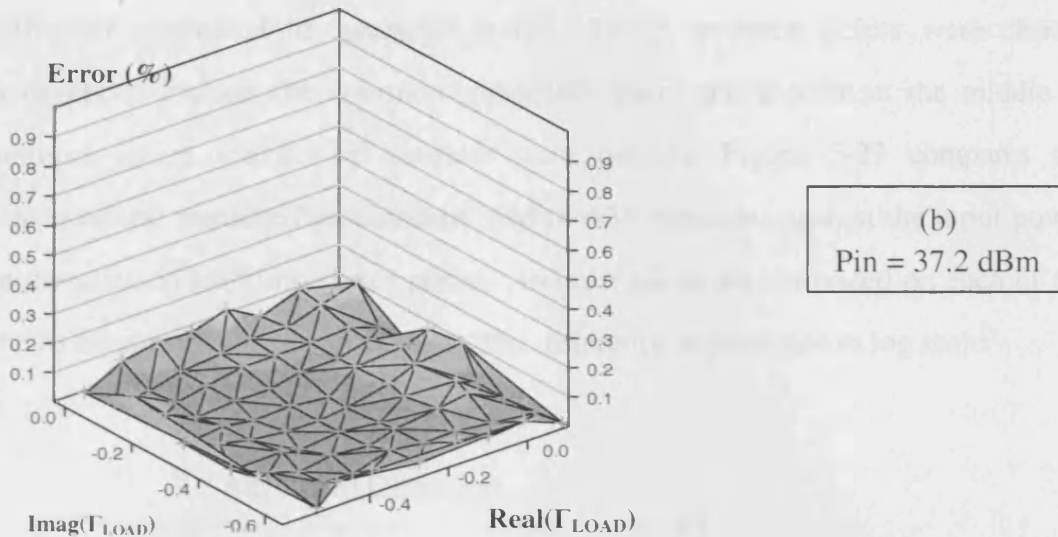


Figure 5-26: The error surface calculated using (5-86) for the same area shown in Figure 5-24

As shown in above figures, it can be confirmed that the model is able to maintain its accuracy for quite a large dynamic power range. Measurement results indicate that the magnitude of renormalized a_2 varies dramatically from 0 to more than 4 within the plotted power contour coverage. It can be seen that the model can also quite happily deal with the variation of load impedance – device characteristic variation as response to the change of load impedance can be accurately predicted.

The good agreements shown on the error surface graphs and contour plot graphs indicate that the accuracy is maintained for most of the simulation points. Good accuracy is achieved even when the device is driven into high compression, which means the model is well capable of modelling large signal nonlinearities. Considering the dramatic movement of a_2 within the simulation region, the experimental results show clearly the extended model is capable of achieving good accuracy for much larger area than the original PHD model without losing simulation efficiency.

Secondly, the simulated output power and phase characteristics at specific load

points are compared to measured results. Two impedance points were chosen randomly to include the optimum impedance point and a point at the middle of contours where the device saturates very quickly. Figure 5-27 compares the simulated and measured fundamental AM to AM variations against the input power on the selected load impedance points. An error bar is also provided on each of the graphs for quantifying the difference. The difference is presented in log scale.

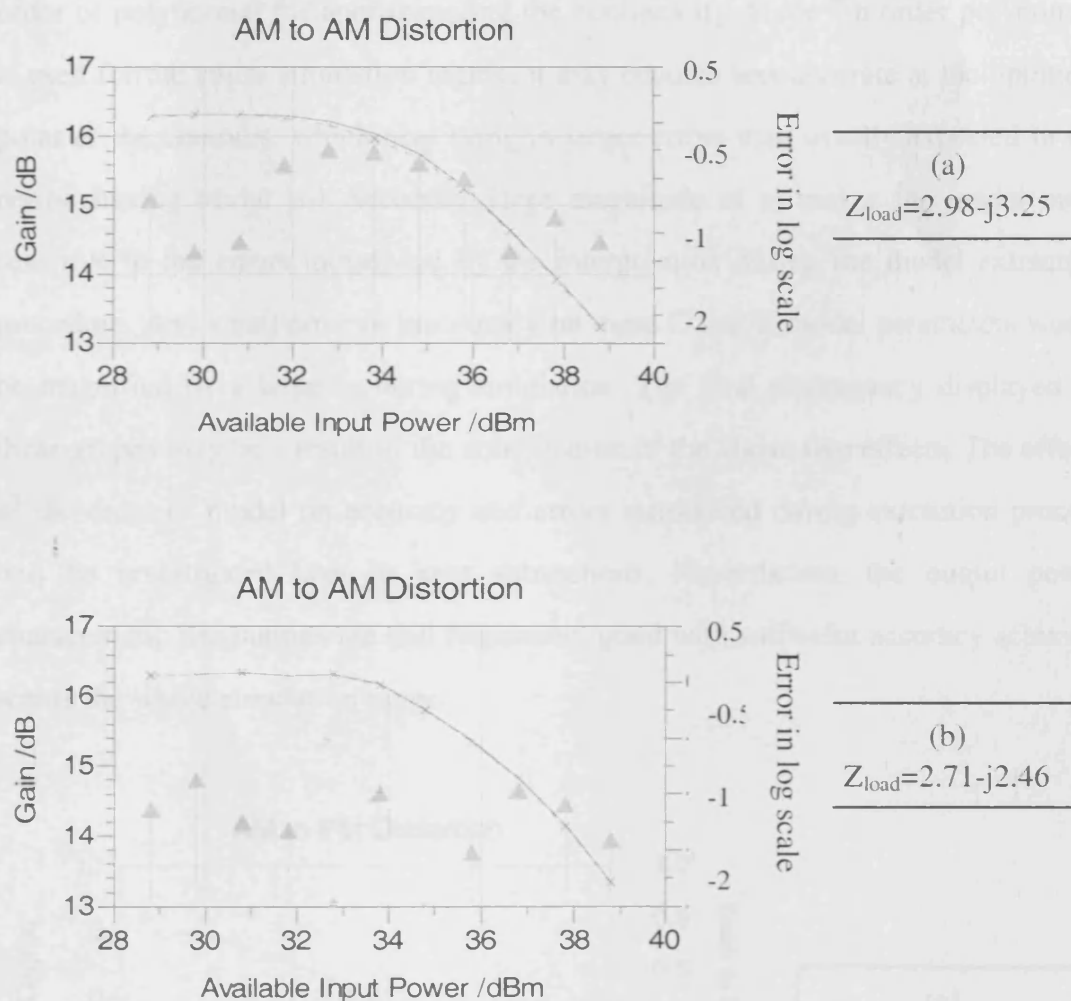
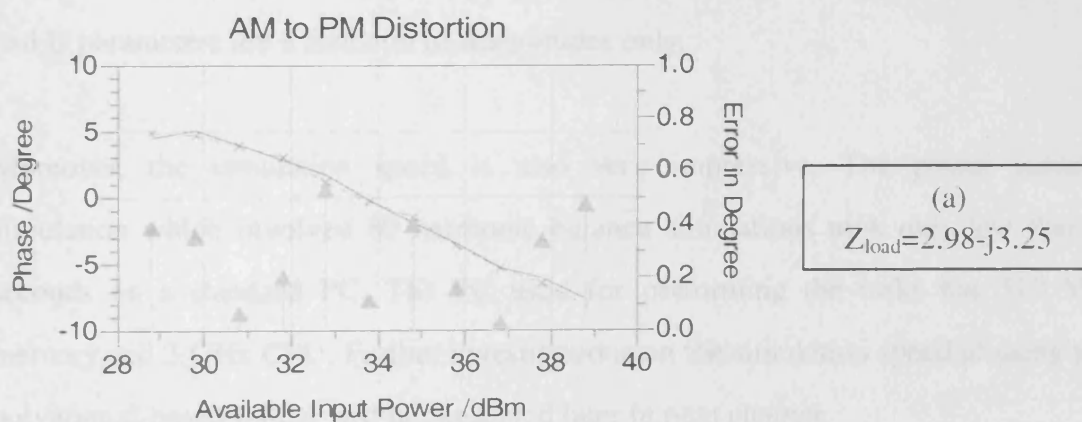


Figure 5-27: The simulated (dotted) and measured (solid) AM to AM distortion as a function of input power. Error (green) is calculated as $\log(\Delta \text{ Linear Gain})$. (a) and (b) show the results obtained on two different load impedance points.

Generally speaking, the power characteristics are accurately simulated at both load impedance points. As observed on graph B, the agreements at the optimum point are

extremely good. However, the accuracy degrades a little bit as the load point moves towards the optimum point. It's probably due to two reasons. First reason is large magnitude of a_2 at the optimum calls for the need to use higher order model. The 7th order model used in simulation is extracted with renormalized travelling waves. After renormalization to $3.139-j*3.014$, the magnitude of a_2 at the optimum point gets significantly smaller but its value is still within the comparable range to magnitude of a_1 . As previously discussed, a large magnitude of a_2 would require high order of polynomial for approximating the nonlinearity. Since 7th order polynomial is used for the entire simulation region, it may become less accurate at the optimum point of the contours, which may bring in larger errors than usually expected in the region having trivial $|a_2|$. Secondly, large magnitude of a_2 makes the model more sensitive to the errors introduced by the interpolation during the model extraction procedure. Any small error or inaccuracy on these C and U model parameters would be magnified by a large a_2 during simulation. The final discrepancy displayed on these graphs may be a result of the combination of the above two effects. The effects of the order of model on accuracy and errors introduced during extraction process will be investigated later in next subsections. Nevertheless, the output power characteristic simulations are still reasonably good with sufficient accuracy achieved across the whole simulation range.



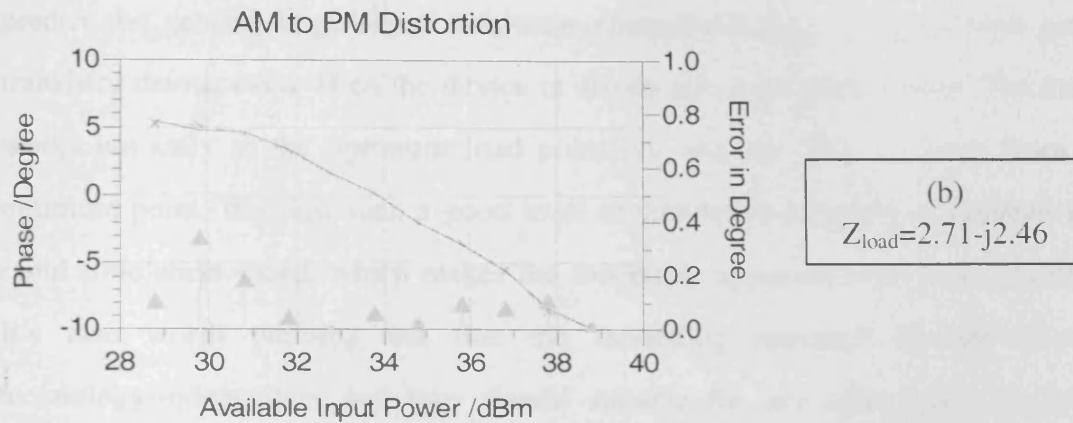


Figure 5-28: The simulated (dotted) and measured (solid) AM to PM distortion as a function of input power. Error (green) is calculated in degree. (a) and (b) show the results obtained on two different load impedance points.

Figure 5-28 compares the simulated and measured AM to PM distortions at these load impedance points. As done on the output power characteristic graphs, an error bar is provided on each of the AM to PM distortion graphs for quantifying the difference.

Similar observation can be obtained on these graphs. The performance of the AM to PM distortion simulation is consistent with the output power simulation. It indicates the phase and magnitude characteristics are modelled equally well. It can be seen the phase information can be accurately reconstructed within the model even if the C and U parameters are a function of magnitudes only.

Moreover, the simulation speed is also very impressive. The power contour simulation which involved 80 harmonic balance simulations took only less than 5 seconds on a standard PC. The PC used for performing the tasks has 512 MB memory and 3 GHz CPU. Further investigations on the simulation speed of using the polynomial-based model will be presented later in next chapter.

All in all, the results presented above indicate the model with 7th order is able to

predict the general large signal nonlinear characteristics of a typical high power transistor device even when the device is driven into high compression. The model works not only at the optimum load point but also the area far away from the optimum point. Besides, such a good level of simulation accuracy is obtained with rapid simulation speed, which makes the modelling approach even more attractive. It's also worth pointing out that the modelling approach by definition is technology-independent and thus should be suitable for any other types of device technologies as well.

5.8.2. Simulation versus measurement on 5w LDMOS

Previous experiments reveal the robustness of the polynomial-based modelling approach on a 100w LDMOS device. In this subsection, a 5w LDMOS device demonstrating lots of harmonic contents during normal operation has been chosen as the target for modelling. It is to test the approach's validity on different device with different power capacity and also challenge the model's capability for modelling strong nonlinearity. The simulation results of the model are compared to measurements. Similarly the DUT was excited by CW signal during measurement. The device was characterized at 2.15 GHz. The input power sweep was done with a range from 23dBm to 28dBm with a step size of 1dB. The load reflection coefficient sweep was done with an 8 by 12 grid on the Smith Chart. The maximum output power variation within the load-pull area is about 3dB.

For generating the polynomial-based model for the target device, the measured travelling waves were renormalized to impedance of $41.846 + j16.548$. During the model extraction, 27 points for each $|a_2|$ were sampled and the travelling wave information at these 27 points was interpolated for use in integration. The model data table covers the magnitude of a_2 from 0 to 0.9 with a resolution of 0.05. The model data table covers the magnitude of a_1 from 0.6 to 1.2 with a resolution of 0.05 as well. A 7th order model with 16 parameters is extracted. The model parameter table takes up only 120 KB space.

Figure 5-29 shows the error surface of output power between simulated and measured for three different drive levels. Basically less than 0.1 dB error can be maintained for most of the load impedance points when the device was driven by different power levels. Occasionally higher than 0.1 dB error can be spotted at some of the points near the corner which is mainly due to the lack of valid model coverage in that area (as discussed previously in figure 5-25).

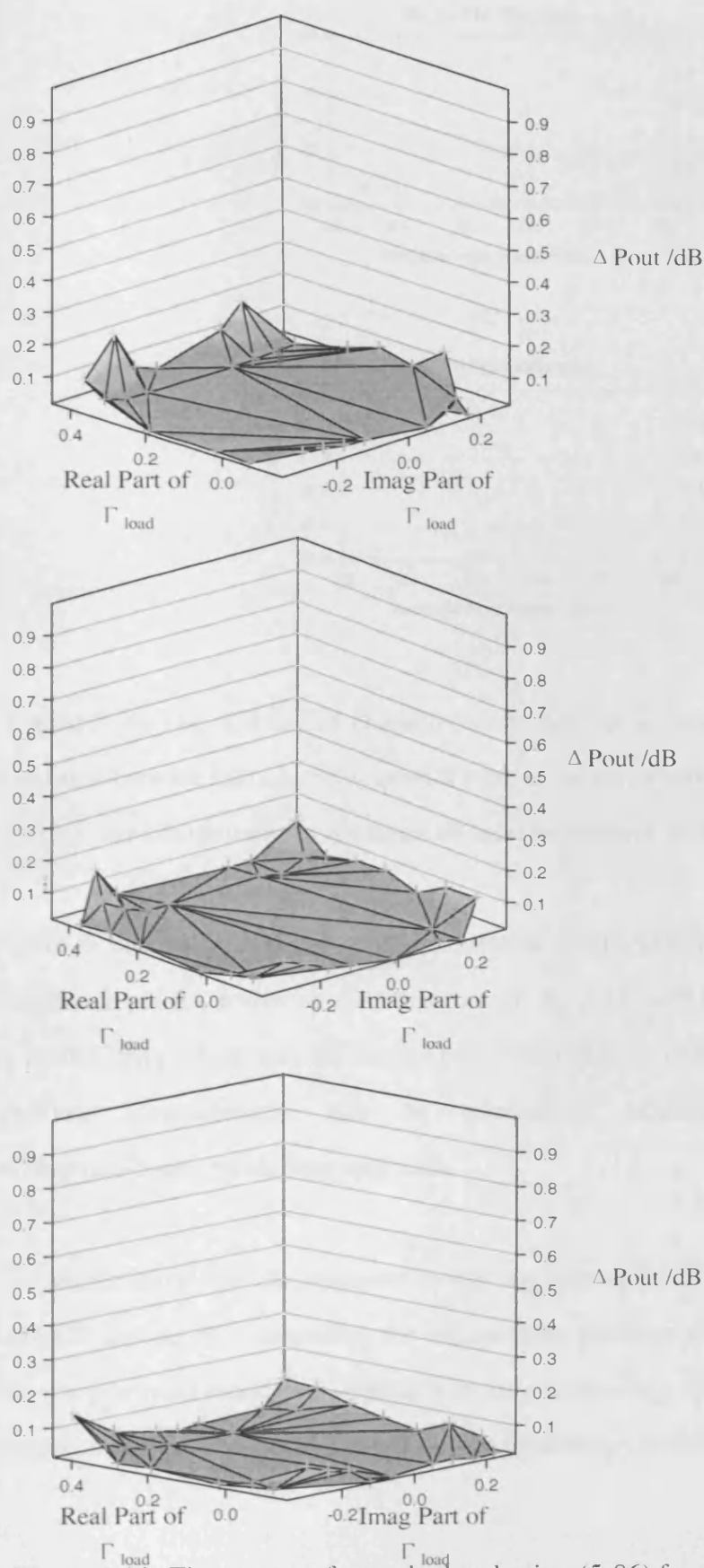
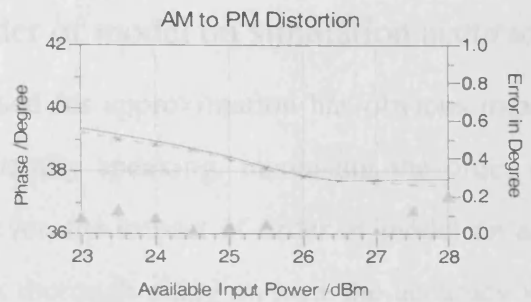
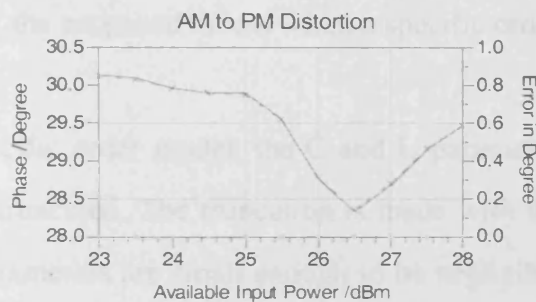


Figure 5-29: The error surface calculated using (5-86) for the simulated load-pull

area



(a)



(b)

Figure 5-30: The AM to PM characteristics obtained at two load impedance points. The error bars are calculated to quantify the accuracy of simulation. (a) and (b) show the characteristics obtained on load impedance of $45.997+j17.988$.

Figure 5-30 illustrates some of the simulation results obtained on two different load impedance points. It can be seen that though the AM to PM demonstrates quite a lot of nonlinearity which can be clearly observed from the characteristic graphs, such a nonlinear characteristic can be adequately modelled by the proposed polynomial-based modelling approach.

The results show that the proposed modelling approach can accurately model the 5w LDMOS device. By combining the results from previous subsection it can be found that the proposed modelling approach is truly technology-independent and should be capable of modelling any device or device technology without altering anything.

5.8.3. Impact of order of model on simulation accuracy

The order of model used for approximation has obvious impact on the accuracy of CAD simulation. Generally speaking, increasing the order of model would yield better accuracy. However, the impact of order of model on accuracy seems to vary from point to point. A thorough study on how the accuracy varies over all the load points and how the order of model affects the accuracy of simulation is needed, in order to make most of the proposed model when a specific order is chosen.

When choosing a specific order model, the C and U parameters having higher than the chosen order are truncated. The truncation is made with the assumption that the truncated C and U parameters are small enough to be negligible and the error due to the truncation can be tolerated. It's therefore important to understand the factors that affect the validity of the assumption.

By inspecting model formulation (5-20) – (5-25), it can be found that the C and U parameters actually consist of a number of $|a_1|$ and $|a_2|$ with different orders. Another interesting observation is that the high order C and U parameters consist of more $|a_1|$ and $|a_2|$ terms than low order parameters. Based on this observation, it can be derived that the high order C and U parameters should increase faster than low order ones, as the magnitude of a_2 increases. By analyzing the typical load pull area it can be found that the magnitude of a_2 at a load reflection coefficient point is a function of magnitude of a_1 and distance from the reference impedance point. Increasing magnitude of a_1 , which is equivalent to increasing the input power of DUT, would normally result in stronger a_2 at the output of DUT, with only one exception on the reference impedance point – in a matching condition the travelling wave a_2 is kept at zero all the time. When the magnitude of a_1 is fixed, the magnitude of a_2 is primarily determined by the load reflection coefficient attached to the DUT. As the distance between the simulation point and the reference impedance point increases, the magnitude of a_2 gets larger and larger. It implies that higher order model is required

in the outer region in order to achieve same level of accuracy as in the central area.

From the device point of view, a large magnitude of a_2 indicates the associated load reflection coefficient point is far away from the optimum impedance point. The device usually suffers from deeper compression at this point if compared to the optimum. As a result, the device would exhibit greater amount of nonlinearity at load reflection coefficient points whose associated magnitude of a_2 is large. Consequently, it requires a model truncated at high order in order to achieve decent accuracy.

First of all, two typical scenarios are set in which load pull reflection coefficients having two different magnitudes of a_2 are chosen. A set of load pull simulations are performed using two models truncated with different orders. The results are presented and compared to the measurements.

The graphs shown later in this section present the simulation results obtained on 27 load reflection coefficient points. The device was driven with an input a_1 of 3.9 and the 27 load reflection coefficient points were selected to have the same magnitude of a_2 but different phase of a_2 . The magnitude of a_2 is 1.8 and the 27 load reflection coefficient points were distributed evenly on the contour to make sure the phase of a_2 is covered from 0 to 360 degree. The $|a_2|$ contour represents a region relatively close to the renormalization impedance point.

Firstly a 3rd order model is extracted and used in the 27 simulations. 3rd order represents the lowest order available for approximating transistor nonlinear behaviour. The extracted 3rd order model parameters are shown in Figure 5-31. Please note only the output characteristics are investigated as the output usually contains more nonlinearity and presents a more severe challenge to the nonlinear modelling approach. It can be seen that the first order C and U parameters, which represent the linear terms in the model, are much more significant than the 3rd order parameters. This fact implies that the linear behaviour is still most dominating in a

transistor device operation and only mild nonlinearity is demonstrated.

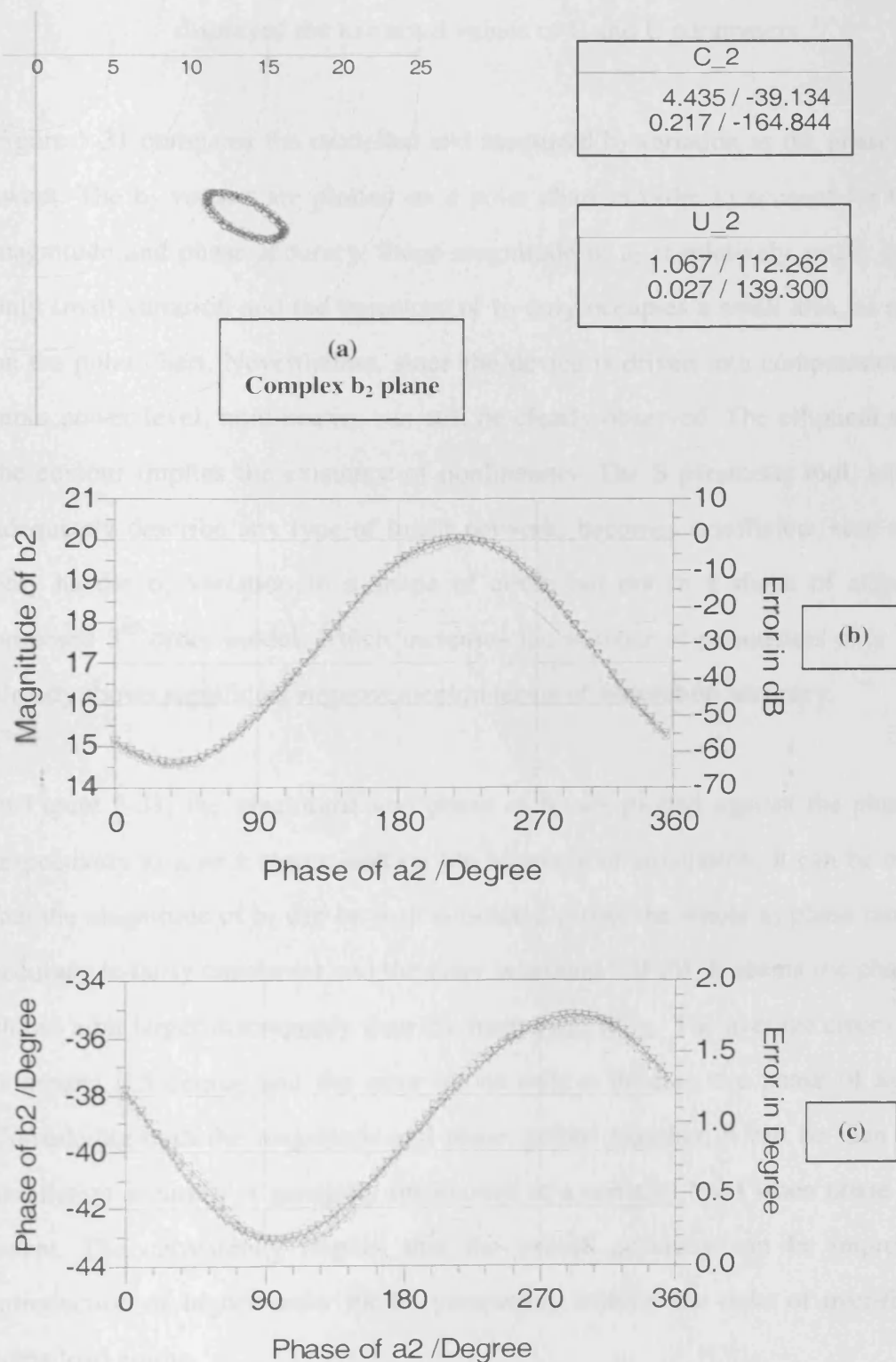


Figure 5-31: Simulation results obtained on a 3rd order model with 8 parameters.

Magnitude of a_1 is 3.9 while magnitude of a_2 is 1.8. (a) shows the b_2 vector

variation as the phase of a_2 is swept. (b) shows the magnitude of b_2 variation. (c) shows the phase of b_2 variation. The data table shown in the top right corner displayed the extracted values of C and U parameters.

Figure 5-31 compares the modelled and measured b_2 variation as the phase of a_2 is swept. The b_2 vectors are plotted on a polar chart in order to account for both the magnitude and phase accuracy. Since magnitude of a_2 is relatively small, b_2 shows only small variation and the trajectory of b_2 only occupies a small area, as reflected on the polar chart. Nevertheless, since the device is driven into compression at this input power level, nonlinearity can still be clearly observed. The elliptical shape of the contour implies the existence of nonlinearity. The S parameter tool, which can adequately describe any type of linear network, becomes insufficient here as it can only handle b_2 variation in a shape of circle but not in a shape of ellipse. The proposed 3rd order model, which increases the number of parameters only by two, already shows significant improvement in terms of simulation accuracy.

In Figure 5-31, the magnitude and phase of b_2 are plotted against the phase of a_2 respectively to give a closer look on the accuracy of simulation. It can be observed that the magnitude of b_2 can be well simulated across the whole a_2 phase range. The accuracy is fairly consistent and the error is around -20 dB. It seems the phase of b_2 shows a bit larger discrepancy than the magnitude of b_2 . The average error of phase is around 0.5 degree and the error varies only a little as the phase of a_2 varies. Considering both the magnitude and phase graphs together, it can be seen that the simulation accuracy is generally maintained at a constant level when phase of a_2 is swept. The consistency implies that the overall accuracy can be improved by introduction of higher order model parameters without the risks of over-fitting at some load points.

Next the simulation results obtained using a 7th order model for the same load reflection coefficient points (i.e. 27 points with same $|a_2|$ of 1.8) when the device is

stimulated with the same input power (i.e. $|a_1|$ is 3.9) will be presented in Figure 5-32.

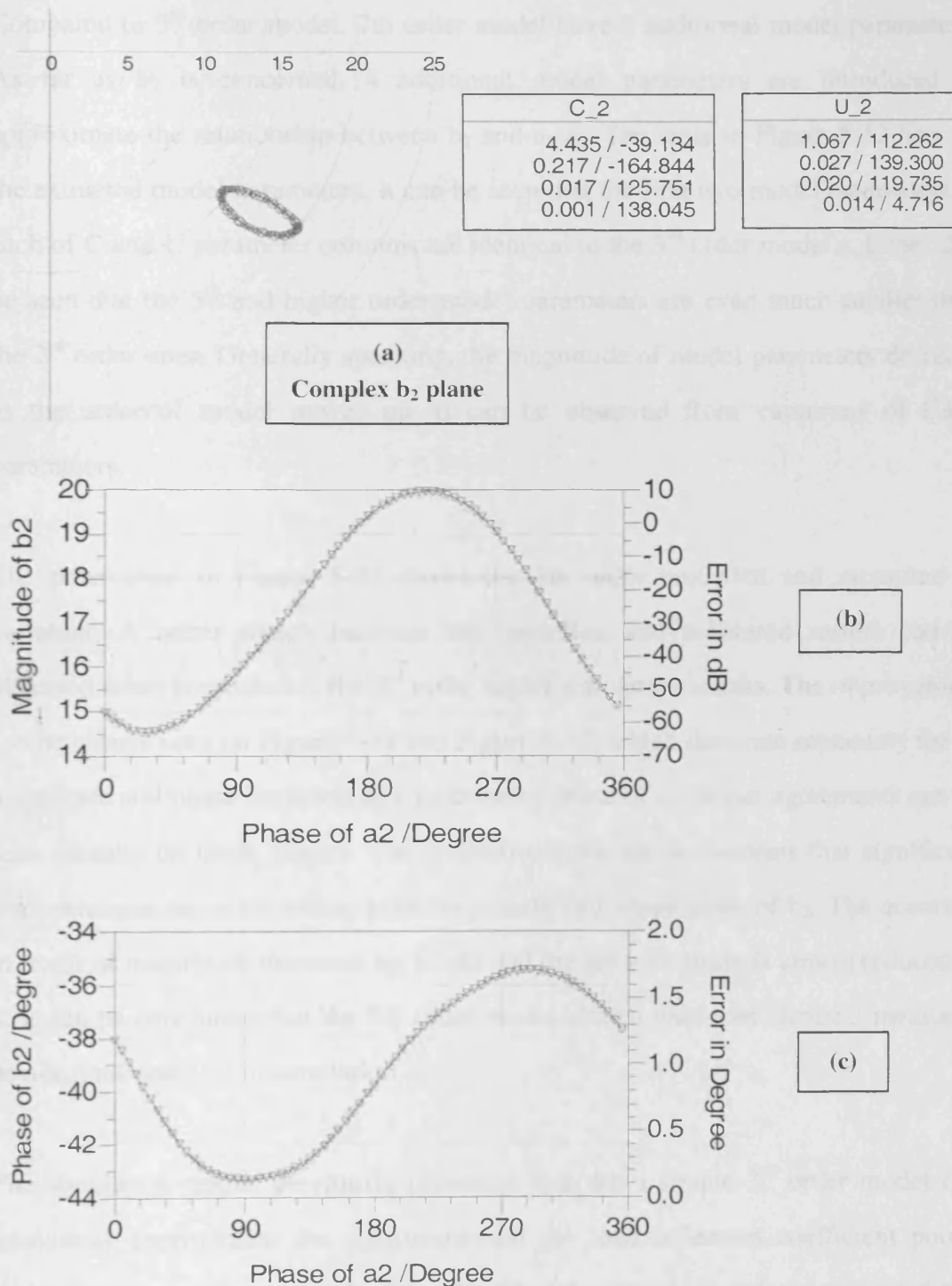


Figure 5-32: Simulation results obtained on a 7th order model with 16 parameters.

Magnitude of a_1 is 3.9 while magnitude of a_2 is 1.8. (a) shows the b_2 vector variation as the phase of a_2 is swept. (b) shows the magnitude of b_2 variation. (c)

shows the phase of b_2 variation. The data table shown in the top right corner displayed the extracted values of C and U parameters.

Compared to 3rd order model, 7th order model have 8 additional model parameters. As far as b_2 is concerned, 4 additional model parameters are introduced to approximate the relationship between b_2 and a_1 a_2 . The table in Figure 5-32 lists all the extracted model parameters. It can be seen that the first two model parameters in each of C and U parameter columns are identical to the 3rd order model's. It can also be seen that the 5th and higher order model parameters are even much smaller than the 3rd order ones. Generally speaking, the magnitude of model parameters decrease as the order of model moves up. It can be observed from variations of C&U parameters.

The polar chart in Figure 5-32 shows the 7th order modelled and measured b_2 variation. A better match between the modelled and measured results can be observed when compared to the 3rd order model simulation results. The improvement can be clearly seen on Figure 5-31 and Figure 5-32, which illustrate separately the b_2 magnitude and phase variation as a function of phase of a_2 . Better agreements can be seen visually on these graphs. The quantified error curve confirms that significant improvements are achieved on both magnitude and phase parts of b_2 . The accuracy in terms of magnitude increases by 10 dB and the error of phase is almost reduced to 0. It can be concluded that the 7th order model almost produces identical measured device nonlinearities in simulation.

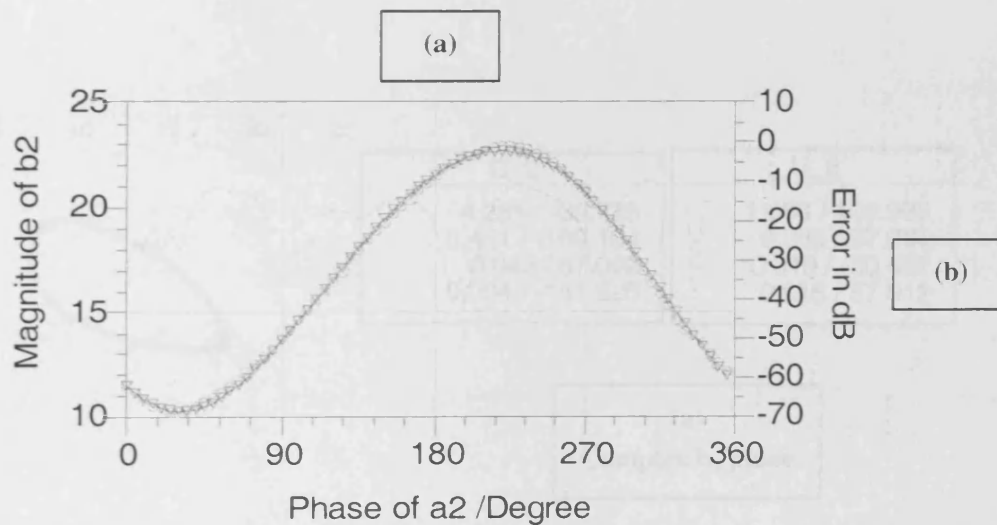
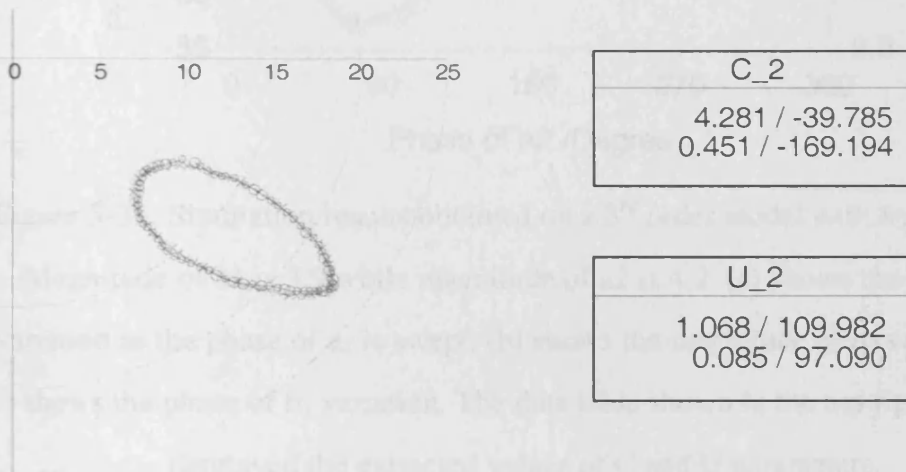
The simulation results previously presented indicate a simple 3rd order model can adequately approximate the nonlinearity on the load reflection coefficient points with relatively small magnitude of a_2 . If a 7th order model is used, the measured and simulated device nonlinear behaviours are on top of each other. In the following paragraphs, load reflection coefficients with a much higher magnitude of a_2 are selected to do simulation, in order to explore the model's capability when dealing

with load impedance points far away from the reference impedance point.

A 3rd order model is used for testing at first. The input power is still fixed at a $|a_1|$ of 3.9. The magnitude of a_2 for the load reflection coefficient points is 4.2, which is now higher than the magnitude of a_1 . The corresponding $|a_2|$ contour is now pretty large and very close to the edge of the measurement space. Similarly 27 load reflection coefficient points are selected which are evenly distributed to cover the phase of a_2 from 0 to 360. The table in Figure 5-33 presents the extracted model parameters corresponding to the specific magnitudes of a_1 and a_2 . The first order parameters are still comparable to the ones when $|a_2|$ is 1.8. But the third order parameters are higher than the ones when $|a_2|$ is 1.8. For example, the magnitude of third order C parameter has a 100% increase from 0.217 to 0.451. The polar chart shows the new b_2 variation. It can be noticed that a much larger ellipse is resulted as the variation of b_2 becomes much higher for such magnitude of a_2 . The ellipse-like contour becomes further distorted and there are more discrepancies spotted at several points on the contour. The deterioration of accuracy can also be observed from the magnitude and phase variation graphs in Figure 5-33. Regarding the magnitude of b_2 , the average error increases to approx. -10 dB. This corresponds to a 0.3 difference on a linear scale. The graph about phase of b_2 tells similar story. Apart from the points close to the a_2 phase of 0 degree, the average error of phase is around 1 degree. The error shows a slight variation from 0.5 to 1.4 degree but the error at the points close to a_2 phase of 0 degree is beyond that region and reaches the value of 2.0 degree. It's mainly due to the discrepancy between modelled and measured at these points, as can be observed in the phase plot in Figure 5-33.

Using the 7th order model should be able to significantly reduce the error of simulation. A 7th order model, whose model parameters are extracted and listed in the table in Figure 5-34, is used to improve the simulation accuracy. Once again, the graphs plotting the magnitude and phase variation of b_2 demonstrate the improvement of accuracy achieved by using high order models. The graph about the

magnitude of b_2 demonstrates an accuracy of -20dB for most of the load reflection coefficient points. The graph about the phase of b_2 also demonstrates significant improvement of accuracy – the error at most of the load points is well below 0.5 degree.



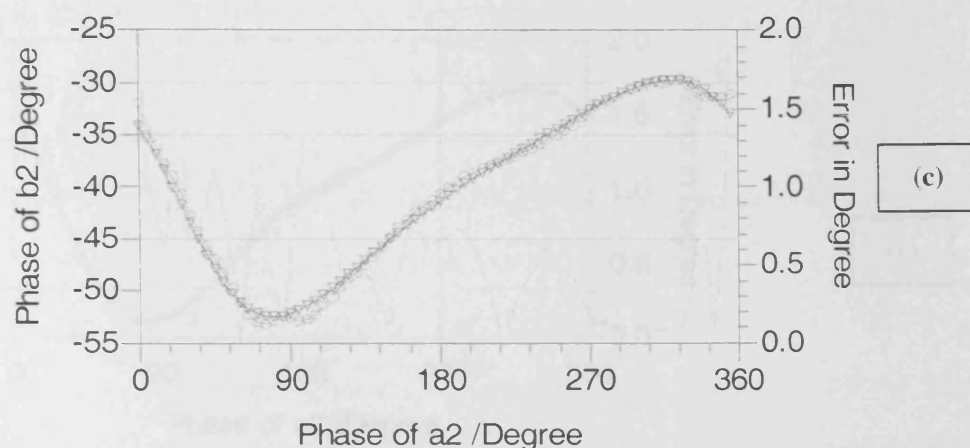
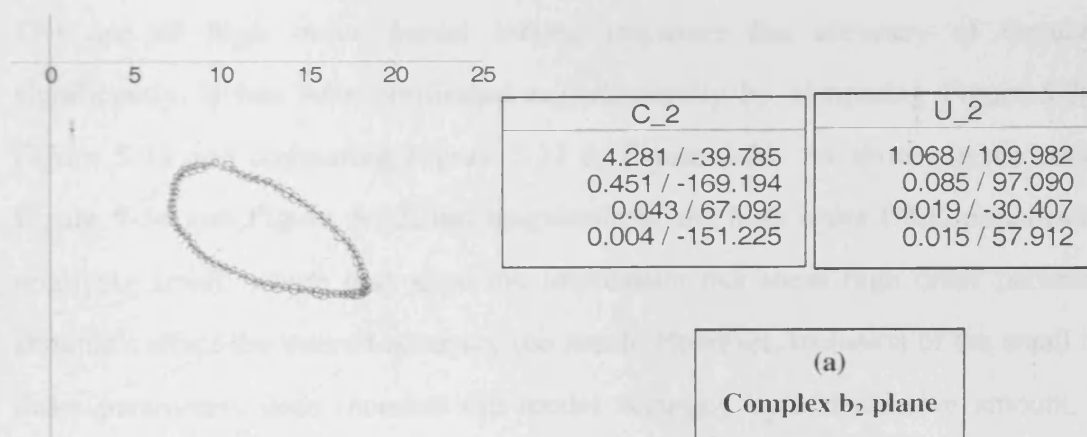
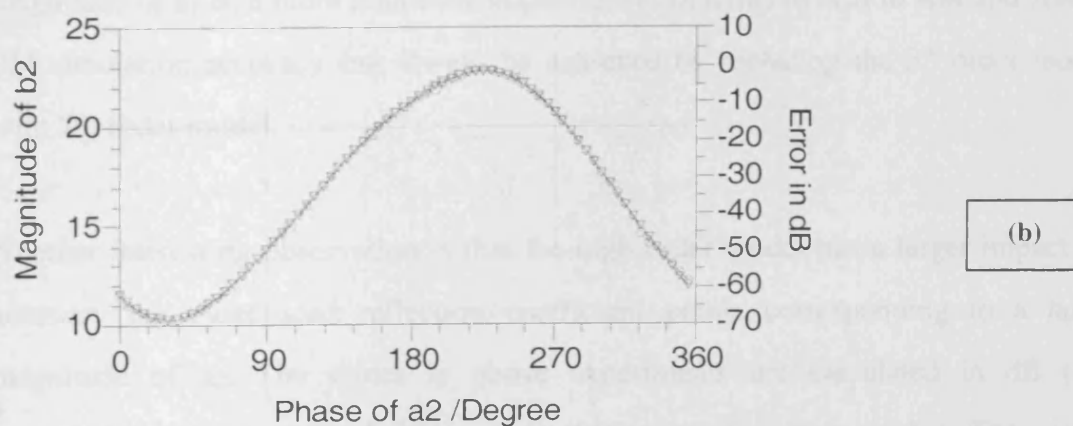


Figure 5-33: Simulation results obtained on a 3rd order model with 8 parameters.

Magnitude of a_1 is 3.9 while magnitude of a_2 is 4.2. (a) shows the b_2 vector variation as the phase of a_2 is swept. (b) shows the magnitude of b_2 variation. (c) shows the phase of b_2 variation. The data table shown in the top right corner displayed the extracted values of C and U parameters.



(a)
Complex b_2 plane



(b)

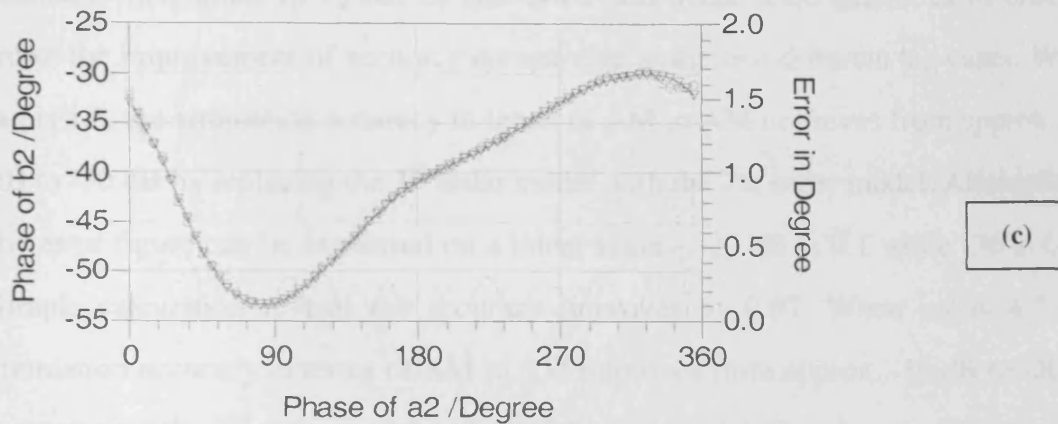


Figure 5-34: Simulation results obtained on a 7th order model with 16 parameters.

Magnitude of a_1 is 3.9 while magnitude of a_2 is 4.2. (a) shows the b_2 vector variation as the phase of a_2 is swept. (b) shows the magnitude of b_2 variation. (c) shows the phase of b_2 variation. The data table shown in the top right corner displayed the extracted values of C and U parameters.

The use of high order model indeed improves the accuracy of simulation significantly. It has been confirmed experimentally by comparing Figure 5-34 to Figure 5-33 and comparing Figure 5-32 to Figure 5-31. As shown in the table in Figure 5-34 and Figure 5-32, the magnitude of the high order C&U parameters is relatively small, which may give the impression that these high order parameters shouldn't affect the overall accuracy too much. However, inclusion of the small high order parameters does increase the model accuracy by a distinctive amount. For simulations on those load reflection coefficient points regardless of how large the magnitude of a_2 is, a more than 80% improvement in terms of AM to AM and AM to PM simulation accuracy can always be achieved by replacing the 3rd order model with 7th order model.

Another interesting observation is that the high order model has a larger impact on accuracy for those load reflection coefficient points corresponding to a large magnitude of a_2 . The errors in above experiments are calculated in dB (for magnitude of b_2) and in absolute phase difference (for phase of b_2). The errors

related to magnitude of b_2 can be converted into linear scale quantities in order to make the improvement of accuracy comparable in the two different $|a_2|$ cases. When $|a_2|$ is 1.8, the simulation accuracy in terms of AM to AM improves from approx. -20 dB to -30 dB by replacing the 3rd order model with the 7th order model. Alternatively, the error figure can be expressed on a linear scale – -20 dB is 0.1 while -30 is 0.03. Simple calculation reveals the accuracy improves by 0.07. When $|a_2|$ is 4.2, the simulation accuracy in terms of AM to AM improves from approx. -10 dB to -20 dB by replacing the 3rd order model with the 7th order model. On a linear scale -10 dB is about 0.3 while -20 dB is 0.1. A 0.7 improvement can be obtained. It can be seen that the improvement of simulation accuracy on load reflection coefficient points with a large magnitude of a_2 is more significant than the improvement shown on the load points with a small magnitude of a_2 . The accuracy improvement on phase related simulations is consistent with the magnitude related simulations.

From the above experiments it can also be seen that the model accuracy degrades as the magnitude of a_2 increases. Comparing directly the $|a_2|$ of 1.8 case to the $|a_2|$ of 4.2 case, if 7th order model is used, the accuracy in terms of magnitude degrades from -30 dB to -20 dB. The degradation also happens to the phase related simulations.

A more comprehensive study of the relationship between the order of model and accuracy of simulation is conducted in a statistical manner. Firstly, the accuracy variation against the magnitude of a_2 is plotted (Figure 5-35). For plotting the plot, a group of magnitudes of a_2 is selected which ranges from 0.2 to 4.2. The range covers the case where the magnitude of a_2 is negligible to the magnitude of a_1 as well as the case where the magnitude of a_2 is comparable to the magnitude of a_1 . For each magnitude of a_2 27 load reflection coefficient points are picked up which have different phases of a_2 . Harmonic balance simulation is performed on every selected load reflection coefficient points and the output power at each load point is calculated in dBm. The output power is then compared to the measured value in order to calculate how much the error is. Average error in terms of output power, as a

measure of the accuracy of the overall simulation, is calculated by accounting for the errors of simulation at all the selected load reflection coefficient points.

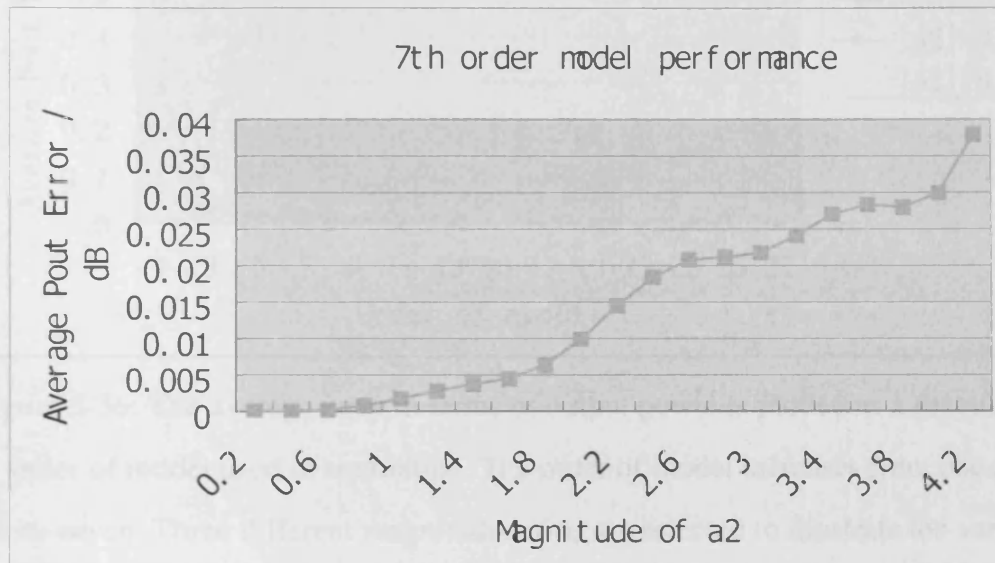


Figure 5-35: The error in terms of output power is plotted as a function of magnitude of a_2 . The input a_1 is fixed at 3.9, making sure the device is operating in compression region. 7th order model is extracted and used in simulation.

As shown in Figure 5-35, 7th order model is chosen to investigate the simulation accuracy variation against the distance between the simulation region and reference impedance centre point. For magnitude of a_2 from 0.2 to 1, the average error is zero. It indicates the simulation perfectly matches the measurement. From 1.0 onwards, the average error starts to increase gradually. This graph clearly demonstrates the existence of the connection between the simulation accuracy and magnitude of a_2 . It's mainly due to the dependency of the nonlinear characteristics of device on load reflection coefficient. Fortunately, since the model parameters are already extracted as a function of the magnitude of a_2 , most of the underlining dependency is properly accounted for in the model formulation. As a result, the overall error level is still impressive and less than 0.05 dB even though the magnitude of a_2 is as high as 4.2.

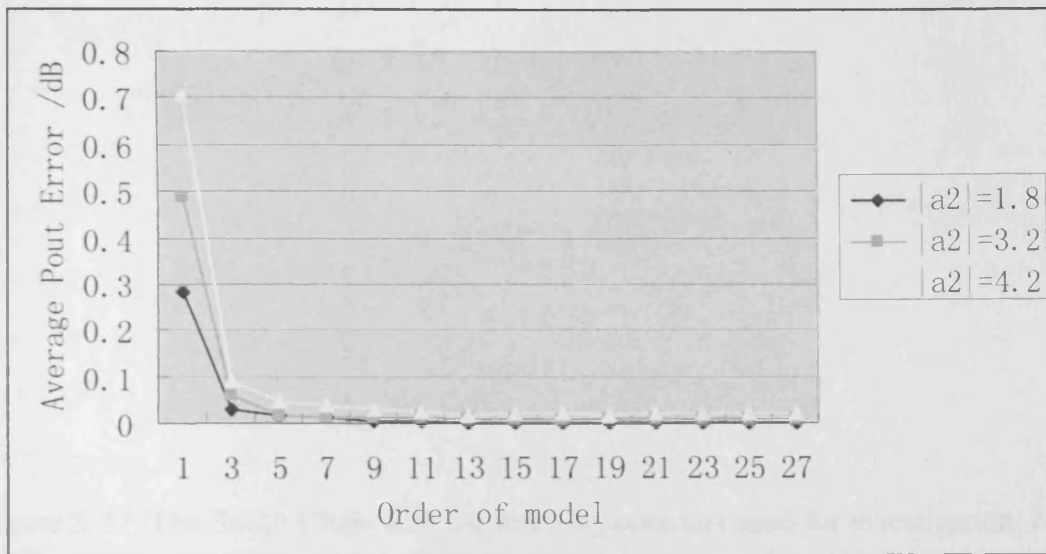


Figure 5-36: The average error in terms of output power is plotted as a function of order of model used in simulation. The order of model increases from one to twenty-seven. Three different magnitudes of a_2 are selected to illustrate the variation of model performance.

Secondly, the dependency of accuracy on order of model is statistically investigated. Figure 5-36 gives the improvement of simulation accuracy that can be gained from increasing the order of model. From Figure 5-35 it can be known that load reflection coefficient points with different magnitudes of a_2 have different sensitivity to the order of model. Therefore, three different magnitudes of a_2 are selected to explore the various effects of increasing order of model on different load pull regions. The location of these $|a_2|$ contours when referencing to the reference impedance and optimum impedance points is illustrated on the following Smith Chart. It's worth pointing out that the largest magnitude of a_2 contour covers a load pull area in which the output power may vary by 3 dB, while the medium and small $|a_2|$ contours cover variations of 1.25 dB and 0.75 dB respectively. Such variations are caused by the mismatch between the optimum load impedance and the renormalization impedance.

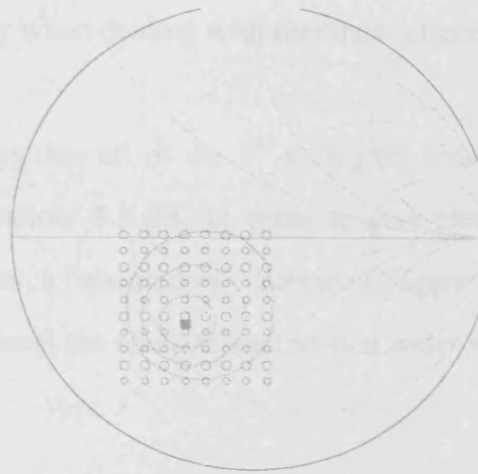


Figure 5-37: The Smith Chart with the three $|a_2|$ contours used for investigation. Note that the a_2 is already renormalized to $3.139-j3.014$ (point crossed on the Smith Chart)

Up to 27th order model is tested in harmonic balance simulation one by one and error is calculated by averaging the difference for each load reflection coefficient point. When the order of model increases, the number of model parameters increases. For example the first order model has only four parameters while the 27th order model has 56 parameters.

As shown on Figure 5-36, the error of simulation when using first order model is very large. The error reaches 0.7 for the magnitude of a_2 of 4.2, which is not acceptable especially for high power transistor devices – a 0.7 dB difference around 50 dBm is usually not tolerable as it equals to a difference of around 15 Watts on a linear scale. The first order model represents the traditional linear approach for approximating device behaviour. It can be seen that by simply increasing the order of model to 3, the average error reduces dramatically from 0.7 dB to less than 0.1 dB. At 50 dBm output level, 0.1 dB error equals to about 2 Watts difference on a linear scale. As stated at the beginning of the experimental section, the resolution of the measurement system is 0.1 dB. The resolution achieved here is already very close to the limit of the measurement system which is used to collect the data for the model extraction. This fact proves the robustness of the proposed modelling approach – relatively simple extension to the traditional S parameters could result in significant

improvement of accuracy when dealing with nonlinear effects.

Figure 5-36 demonstrates that all of the 3rd or higher order models can achieve a level of accuracy well below 0.1 dB. In order to give clearer visuals on how the model performance varies, a “zoomed in” version of Figure 5-36 is plotted in Figure 5-38, which simply neglects the spots related to first order model and re-scale the y axis.

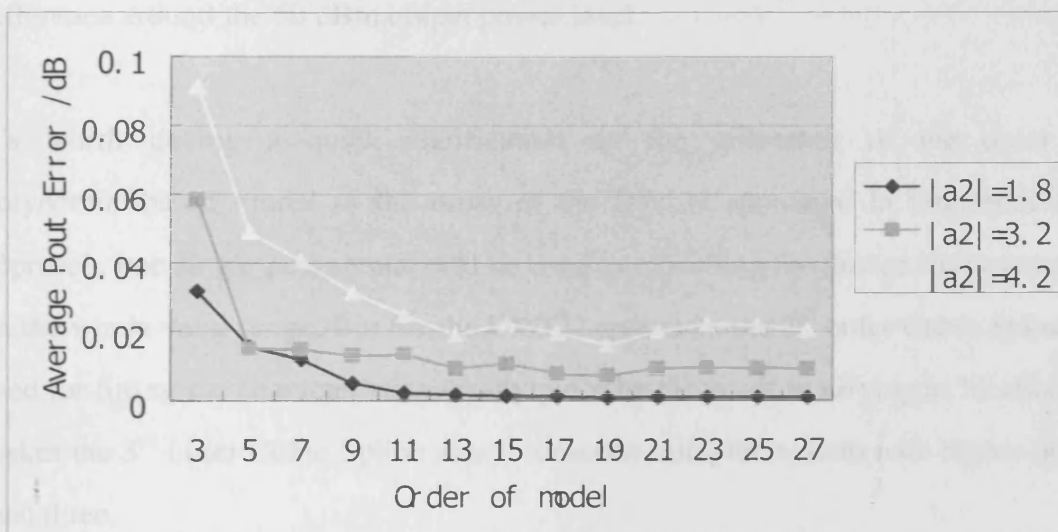


Figure 5-38: “Zoomed in” version of Figure 5-36. The points related to first order model is removed from the graph in order to have a much closer look on the rest of models.

An interesting observation obtained from Figure 5-38 is that the accuracy of the model ceases to increase when the order of model reaches 9 and beyond. This is because there are no sufficient number of measurement data for extracting model parameters of order higher than 9. For example, to extract a 9th order model, 10 model parameters need to be identified for b_1 and b_2 equations respectively. An accurate extraction would require 10 independent measurement data points for each magnitude of a_2 . By referring to Figure 5-14, it can be seen that it is difficult to locate 10 independent measurement data points on any $|a_2|$ circle. Therefore, it is

inaccurate to extract the model parameters of order higher than 9 from the measurement data set available in this work.

Nevertheless, before the order of model is increased to a level on which the model parameters become too small to be extracted, excellent accuracy of simulation is already achieved. Take the 7th order model as an example. When the magnitude of a_2 is 1.8, 3.2, and 4.2, the average error is about 0, 0.01 dB and 0.02 dB. Considering the worst case that the error is 0.02 dB, the 0.02 dB difference is only about 0.5 watt difference around the 50 dBm output power level.

It's worth having a quick clarification on the difference of the order of polynomial-based model to the order of the DWLU approach. In this modelling approach, one single polynomial will be used in modelling the device characteristics on the whole valid range. But for the DWLU approach, the 3rd order Cubic Spline is used for fitting the characteristics locally piece by piece – it is varying as location. It makes the 3rd order Cubic Spline able to describe complex system with higher order than three.

5.8.4. Impact of measurement grid density on extraction accuracy

The previous experiments are all done in the condition that sufficient amount of measurement data is available. It would be interesting to see how robust the modelling approach is if only small amount of measurement data is available for model extraction. To do such investigation, subsets of the measurement data with different density collected on the 100w LDMOS device are used to extract the model respectively. The extracted model parameters are then compared to evaluate the robustness of extraction.

The model parameters shown in Figure 5-39 are obtained on three models extracted from different subsets of data. The first data set covers all the data available and

represents 100% of measured data for model extraction. The second data set represents 75% of measured data for model extraction. The third data set represents only 50% of measured data for model extraction. Note that all the three models are extracted and supposed to be valid for the same input power range and load-pull range though the data sets used for model extraction have different density. The graphs present the simulated b_2 as a function of phase of a_2 . The magnitude of a_1 is fixed at 3.9 which results around 50 dBm output power at the optimum with approx. 2 dB compression. The magnitude of a_2 is fixed at 3.6 which is comparable to the magnitude of a_1 in order to explore the full potential of the model. Up to 7th order parameters are extracted and used in simulation. No higher order parameters are extracted because based on what were observed in Figure 5-38, higher order parameter after 9 becomes very small and close to the noise floor.

It can be observed that the density of the measurement data for model extraction does have an impact on the robustness of the proposed modelling approach. The variation of measurement grid density shows more impact on parameters of higher orders.

C_2	U_2
4.325 / -39.381	1.086 / 110.763
0.406 / -170.158	0.070 / 95.357
0.054 / 70.959	0.010 / -119.361
0.007 / 133.527	0.013 / 19.366

(a)

C_2	U_2
4.325 / -39.414	1.055 / 110.077
0.407 / -169.911	0.075 / 85.232
0.052 / 70.716	0.019 / -53.106
0.008 / 97.449	0.019 / 15.270

(b)

C_2	U_2
4.323 / -39.504	1.072 / 110.450
0.402 / -168.980	0.072 / 97.742
0.042 / 81.324	0.014 / -52.926
0.005 / 153.312	0.010 / 36.172

(c)

Figure 5-39: The extracted model parameters when the magnitude of a_1 is fixed at 3.9, magnitude of a_2 is fixed at 3.6, and the phase of a_2 is swept from 0 to 360 degree.

(a) shows the extracted C and U parameters when 100% measurement data were used for model extraction. (b) shows the extracted C and U parameters when 75% measurement data were used for model extraction. (c) shows the extracted C and U parameters when 50% measurement data were used for model extraction.

It's important to find out where the error comes from when the amount of measurement data for model extraction reduces. In other words, the steps in model extraction which are most vulnerable to the measurement data density need to be identified. The model extraction only involves three major steps and it can be relatively easy to figure out that among the three steps the accuracy of the interpolation step is undoubtedly determined by the measurement data density. Since the DWLU approach is used to do the interpolation for model extraction, some experiment results obtained in chapter four can be used here to analyze the degradation of the accuracy. Based on the results displayed in Table 4-2, it can be seen that the general error rate increases by a factor of 2 when the amount of measurement data reduced from 75% to 50% for same interpolation region. The degradation in interpolation would certainly deteriorate the accuracy of the polynomial-based model extraction and consequently cause the discrepancies of the model parameters as observed in Figure 5-39.

5.8.5. Verification of model extraction consistency

In order to verify the robustness of model extraction procedures and implementation routines, additional tests have been done. An experiment has been designed in order to “close the loop” for model verification, i.e. generate the C&U parameters, simulate the contour plots with these C&U parameters, use model extraction routines to obtain the 'extracted C&U parameters' from the simulated contour data and compare them with the original values. The flow diagram is shown below:

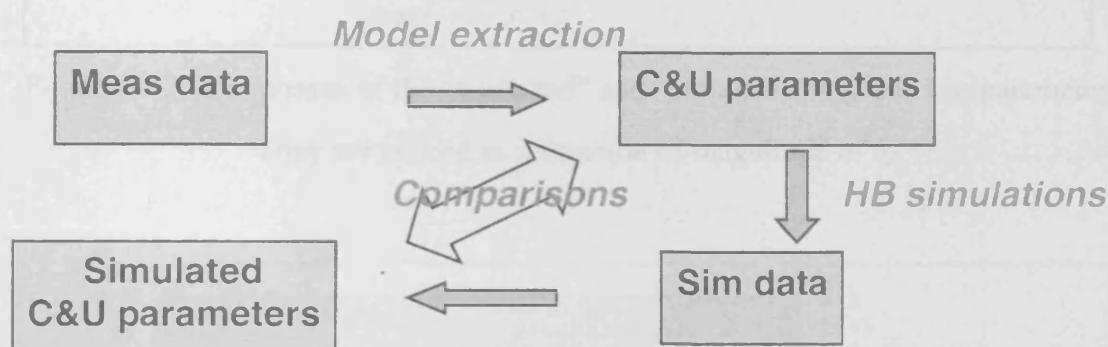


Figure 5-40: The flow for verification of model identification procedure

Ideally, the two set of model parameters should be the same. Figure 5-41 to Figure 5-44 compare the original C&U parameters with the extracted C&U parameters in terms of magnitude and phase respectively.

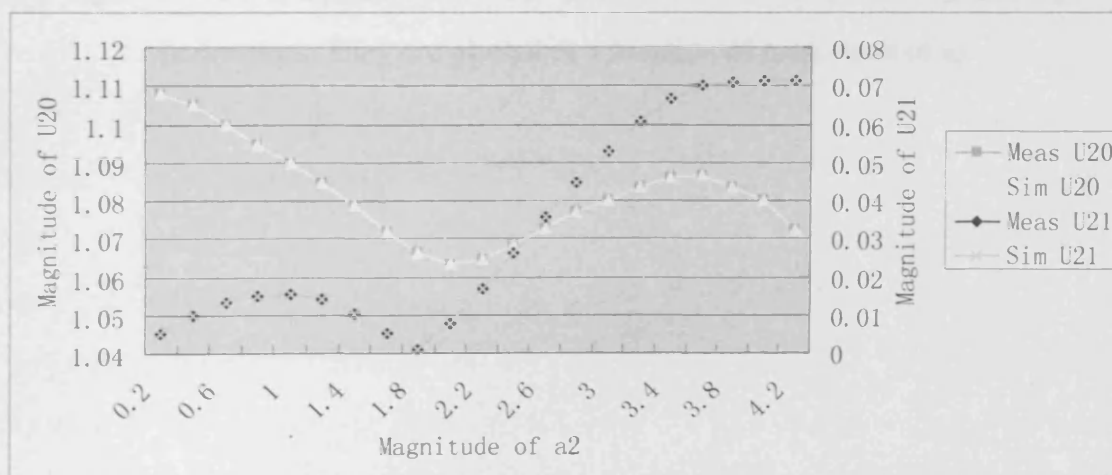


Figure 5-41: The magnitudes of the “extracted” and “simulated” U_{20} and U_{21}

parameters. They are plotted as a function of magnitude of a_2 .

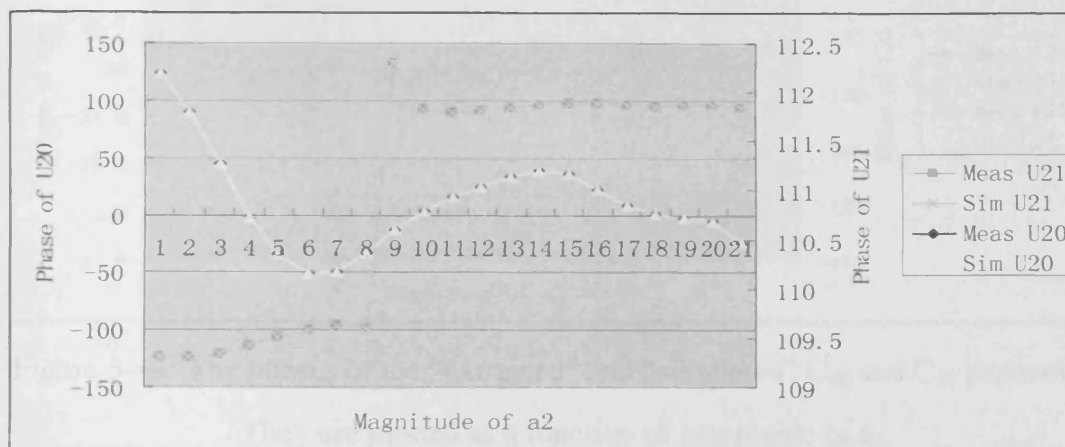


Figure 5-42: The phases of the “extracted” and “simulated” U_{20} and U_{21} parameters.

They are plotted as a function of magnitude of a_2 .

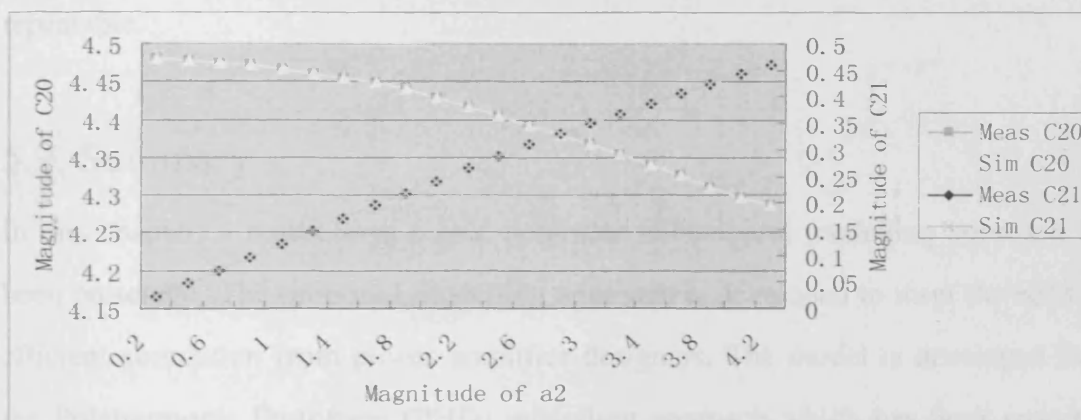


Figure 5-43: The magnitudes of the “extracted” and “simulated” C_{20} and C_{21}

parameters. They are plotted as a function of magnitude of a_2 .

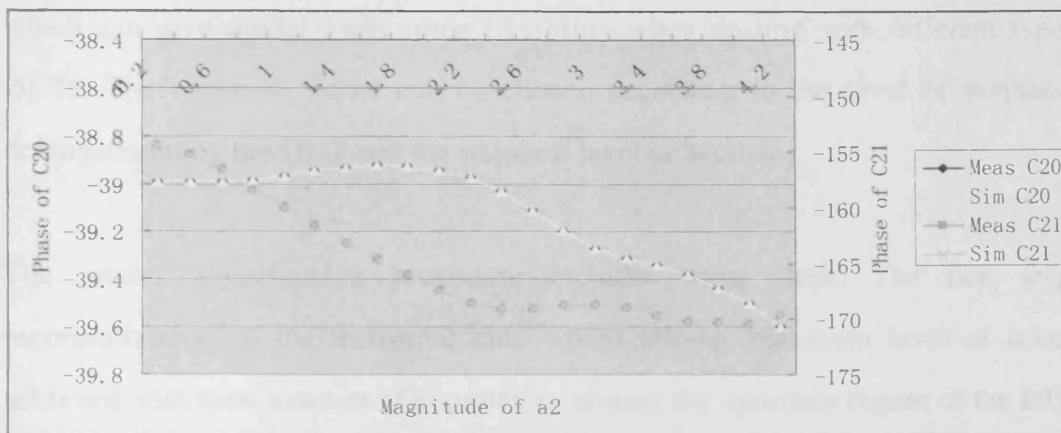


Figure 5-44: The phases of the “extracted” and “simulated” C_{20} and C_{21} parameters.

They are plotted as a function of magnitude of a_2 .

No visual difference exists between the original and extracted parameters. The result verifies the general reliability of the model extraction procedure. It can be concluded that the extracted model should be robust and the resultant accurate simulation is repeatable.

5.9. Summary

In this chapter, a novel large signal nonlinear behavioural modelling approach has been presented. The proposed modelling approach is developed to meet the need for efficient simulation from power amplifier designers. The model is developed from the Polyharmonic Distortion (PHD) modelling approach which has been proposed recently. The original PHD model makes simple extension to the S parameters in order to account for the nonlinearity shown in large signal operation. However, the approximation is no longer valid when running the model at points far away from the system characteristic impedance point, which is usually the case for high power amplifier design. This limitation is mainly due to the assumption of the original PHD model theory that no significant mismatch is presented to the load of the model. The proposed polynomial-based modelling approach presented in this work makes substantial extension to the original PHD model. The load mismatch is taken into account in the model formulation which makes the model suitable for most of high power amplifier design applications. A generic model formulation is also developed

which can give model users some flexibility when dealing with different types of DUTs. The order of model can be chosen according to the level of nonlinearity demonstrated by the DUT and the required level of accuracy.

The model identification procedure includes three steps. The first step is renormalization of the measured data which allows maximum level of accuracy achieved with least amount of complexity around the optimum region of the DUT. It also makes sure that all the measured data can be properly utilized in the model extraction in order to minimize the extraction errors. The second step is interpolation at the data points which are not covered in measurement. The model extraction requires data points grouped according to the magnitude of a_2 . But in measurement the data points are actually grouped according to the load reflection coefficient. The discrepancy makes it necessary to use interpolation between the measured grids. In order to give best accuracy, the DWLU approach presented in chapter four, which shows superior interpolation robustness, is utilized to do the interpolation. The final step is optimization of the model parameters through integration around the $|a_2|$ circle. This is done by considering the fact that each model parameter associates with different order of phase operator in the model formulation. Finally a general solution for extracting arbitrary order model is presented.

The model is implemented in ADS using the FDD component. The implementation is relatively straightforward except for the trick to make conversion between travelling waves and voltage and current waveforms. The FDD component is designed for customizing the relationship between input and output voltage and current vectors in frequency domain through mathematical equations. It's not directly usable for defining relationship in terms of travelling waves. Therefore additional components are used to bridge the gap and make the FDD component usable to the travelling wave vectors.

The model is tested intensively on different devices. The results obtained on a 100w

LDMOS and a 4w LMODS device are presented in this chapter. Comparisons on load-pull type simulation, AM to AM and AM to PM demonstrate that the proposed modelling approach is very robust and sufficiently accurate for approximating nonlinear characteristics. Besides, some further tests were done in order to investigate the potential of the model more deeply. The accuracy of the model as a function of the order of model is experimentally investigated. It can be found that the accuracy increases gradually as the order of model increases. But the ultimate accuracy achievable is limited by the accuracy limit of model extraction. Since the magnitude of model parameters gets smaller and smaller as the order increases. It becomes harder and harder to extract the parameters accurately. Once the magnitude of the model parameter becomes comparable to the noise floor, increase of the order of model would no longer benefit the accuracy of approximation. Nevertheless, experiments demonstrate that sufficient accuracy can usually be achieved well before the model parameter turns to be indistinguishable.

It's also worth pointing out that the order of harmonics is limited to three in the collection of the measurement data, which doesn't mean the order of system that the model is trying to describe is limited to 3rd order. The first three harmonics contain rich information that was generated from large harmonic components. These components are also successfully modelled in this proposed modelling approach.

Chapter 6 Direct Waveform Utilization versus Behavioural Modelling

6.1. Introduction

In previous two chapters, the direct waveform utilization and the polynomial-based behavioural modelling approaches are described respectively. Direct waveform utilization import waveform data collected from NVNA straight into CAD for simulation, analysis and PA design. Polynomial-based behavioural modelling approach is developed to overcome the inherent limitation of the original PHD model. It also proposes a generic method to model a variety of nonlinear devices. The modelling approach defines a set of parameters which has been proven to be reasonable large signal extension to S parameters. The block diagrams of the two approaches are shown in parallel in Figure 6-1.

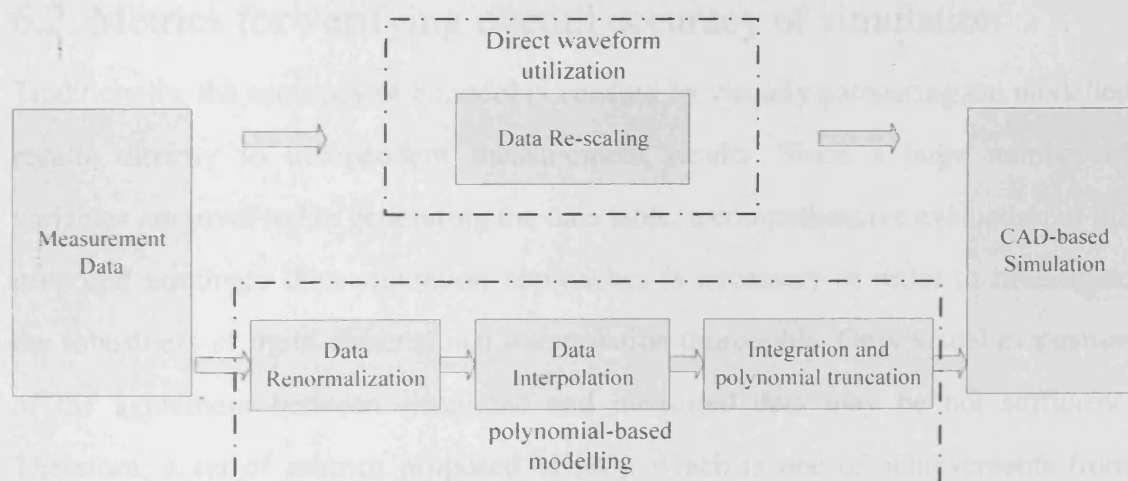


Figure 6-1: The block diagrams of both the direct waveform utilization approach and the polynomial-based behavioural modelling approach.

The proposed two approaches represent two distinct ways to utilize the waveform data obtained from NVNA. Direct waveform utilization aims to minimize the necessary steps from large signal measurement to large signal power amplifier

design. Faithful replication of waveform measurement data in CAD facilitates device behaviour analysis, power amplifier configuration investigation and model development/verification. Polynomial-based behavioural modelling approach takes the advantage of the waveform measurement data, which contains comprehensive information about how device under test behaves in large signal realistic working conditions, to develop quality large signal models. Using the waveform data for model extraction gives the proposed modelling approach the edge in terms of accuracy and reliability over the traditional approaches which are commonly extracted from DC IV data.

Each of the two approaches has its advantages and disadvantages. In order to qualitatively and quantitatively investigate their strength and weakness, in this chapter, the two approaches are compared with the emphasis on the simulation robustness, speed and implementation complexity.

6.2. Metrics for verifying overall accuracy of simulation

Traditionally, the accuracy of a model is verified by visually comparing the modelled results directly to independent measurement results. Since a large number of variables are involved in generating the data table, a comprehensive evaluation of the proposed nonlinear data utilization approaches is necessary in order to investigate the robustness of multi-dimensional interpolation thoroughly. Only visual evaluation of the agreement between simulated and measured data may be not sufficient. Therefore, a set of metrics proposed in [80], which is one of achievements from Europe TARGET Project, can be used for a preliminary evaluation in a more quantitative way.

A set of metrics proposed in [80] can be used to verify the accuracy of CAD simulation in a quantitative manner. The basic metrics described in [80] include DC metrics, S parameter metrics, large signal single tone and two tone metrics. The key

feature of using the metrics to evaluate the accuracy of simulation is to provide standard figures of merit which can be used to compare different modelling/simulation approaches statistically. The metrics can be used to evaluate the accuracy of simulation from DC current and voltage, small signal s-parameters and large signal harmonic balance single load simulation perspectives. For each perspective, a simple figure is calculated by averaging properly the errors at all the simulation points. These simple figures obtained from different perspectives are quite intuitive and direction comparisons of these figures give a clue on how good an approach is when compared to others.

However, the original metrics don't consider the situation where the evaluation of load pull simulation is desired. Load pull type measurement has been proven to be able to capture invaluable information from DUT and such information is very useful in power amplifier design. Undoubtedly, the robustness of load pull simulation is very important in determining how effective the CAD tool can be in power amplifier design. This work makes simple extension to the original metrics by taking the load pull simulation into account for metric calculation. The basic metrics are originally proposed for model validation and some modifications were done to make them suitable for verifying the quality of multi-dimensional interpolations used in the DWLU approach as well. The corresponding ADS implementations are also developed. The extended metrics are displayed in Table 6-1.

It can be seen that the five figures of merit cover all the important aspects of a device load pull characteristics. The metrics are developed with the intention to compare the robustness of DWLU and the polynomial-based modelling approaches. Since the modelling approach doesn't cover DC or harmonic simulation yet, no PAE or harmonic power characteristics can be calculated. Therefore, only **M1_LS_{f0}**, **M3_LS** and **M5_LS_{x_{f0}}** are calculated for comparisons.

M1_LS =	$\sqrt{\frac{\sum_Q \sum_M \sum_N \sum_{j_1}^H P_{out}^{meas} - P_{out}^{sim} ^2}{\sum_Q \sum_M \sum_N \sum_{j_1}^H P_{out}^{meas} ^2}}$
M1_LS_{x0} =	$\sqrt{\frac{\sum_Q \sum_M \sum_N P_{out_{x0}}^{meas} - P_{out_{x0}}^{sim} ^2}{\sum_Q \sum_M \sum_N P_{out_{x0}}^{meas} ^2}}$
M1_LS_W =	$\sqrt{\frac{\sum_Q \sum_M \sum_N \sum_{j_1}^H \frac{ P_{out}^{meas} }{\sum_{j_1}^H P_{out_{j_1}}^{meas} } P_{out}^{meas} - P_{out}^{sim} ^2}{\sum_Q \sum_M \sum_N \sum_{j_1}^H P_{out}^{meas} ^2}}$
M2_LS =	$\sqrt{\frac{\sum_Q \sum_M \sum_N I_D^{meas} - I_D^{sim} ^2}{\sum_Q \sum_M \sum_N I_D^{meas} ^2}}$
M3_LS =	$\sqrt{\frac{\sum_Q \sum_M \sum_N Gain^{meas} - Gain^{sim} ^2}{\sum_Q \sum_M \sum_N Gain^{meas} ^2}}$
M4_LS =	$\sqrt{\frac{\sum_Q \sum_M \sum_N PAE^{meas} - PAE^{sim} ^2}{\sum_Q \sum_M \sum_N PAE^{meas} ^2}}$
M4_LS_W =	$\sqrt{\frac{\sum_Q \sum_M \sum_N \frac{ PAE^{meas} }{\max_n PAE_n^{meas} } PAE^{meas} - PAE^{sim} ^2}{\sum_Q \sum_M \sum_N PAE^{meas} ^2}}$
M5_LS_{x0} =	$\sqrt{\frac{\sum_Q \sum_M \sum_N (Phase(P_{out_{x0}}^{meas}) - Phase(P_{out_{x0}}^{sim}))^2}{\sum_Q \sum_M \sum_N Phase(P_{out_{x0}}^{meas})^2}}$

Table 6-1: Loadpull large-signal single tone metrics

(Q is the number of loadpull points; M is the number of input powers; N is the number of Bias Points; x is the order of harmonics)

In this table:

M1_LS_{x0} concerns how well the output power can be predicted at individual harmonics. **M2_LS** concerns the ability to predict self-biasing.

M3_LS considers the prediction of the gain, or in other words, the AM/AM.

M4_LS concerns the prediction of the PAE.

M5 LS_{xn0} is a measure for how well the model can predict the phase of the output power at individual harmonic (order of harmonic indicated by x).

6.3. Comparisons on simulation robustness

Simulation robustness of the DWLU approach and polynomial-based behavioural modelling approach is compared first. As shown in Figure 6-1, the DWLU approach consists of only one data processing procedure – measurement data re-scaling. As discussed in chapter three, data re-scaling doesn't remove any information off the measurement data. The measurement information is intact and can be perfectly recovered in CAD simulation. In the other words, if no interpolation is attempted in simulation, the DWLU approach will have perfect accuracy and the simulation would match exactly the measurement.

The polynomial-based modelling approach consists of three data processing procedure. The data renormalization shouldn't introduce any error as it simply changes the reference impedance of the measured travelling waves. However, the remaining two steps would unfortunately introduce errors into the data. Data interpolation uses measured information to calculate the data at desired points for model extraction. Interpolation error would appear to be negligible. The integration method would quite effectively reduce the interpolation error but the truncation of polynomial used for approximation essentially introduces new errors. Therefore, it can be seen that the modelling approach introduces a certain level of inaccuracy into the CAD simulation.

It is important to see how much the inaccuracy may be introduced by the proposed modelling approach. Besides, it's also interesting to compare the accuracy of the modelling approach with the accuracy of the interpolation of the DWLU approach. For simulation off the measurement grid, interpolation is activated in simulation if the DWLU approach is adopted. Particularly, a Cubic Spline algorithm is selected to

do the interpolation. According to the algorithm, two data points at either side of the unknown point are used for interpolation. The proposed modelling approach uses different strategy to predict the device characteristics at the unknown point. Instead of using four data points to interpolate the unknown information directly, a group of data points sampled around the $|a_2|$ circle is firstly used to extract a set of parameters. The set of parameters is constant around the $|a_2|$ circle and is used to predict the device characteristics at the desired unknown point. It can be seen that the methods used for characteristic prediction adopted by the two approaches are quite different and the robustness needs to be tested experimentally.

Experiment is designed in order to do the comparisons listed in the last paragraph. A 100w LDMOS device is selected as the DUT and the waveform measurement has been carried out covering the input power from 33.8 dBm to 38.8 dBm and an 8 by 10 load-pull grid. Both the direct waveform utilization approach and polynomial-based modelling approach are used to process the measurement data and run ADS simulation. In order to test the interpolation of DWLU approach off the measurement grid, simulations are also performed with subsets of measurement data. Similarly, in order to test the prediction ability of the modelling approach off the measurement grid, several 7th order models are generated with subsets of measurement data. The simulation results are presented in the form of the metric calculation results. Table 6-2 shows the metric calculation results obtained using the modelling approach while Table 6-3 shows the metric calculation results obtained using DWLU approach.

Num of pts used for model extraction	60pts	40pts
M1_LS	0.027	0.028
M3_LS	0.026	0.029
M5_LS	0.047	0.051

Table 6-2: The metric calculation results obtained from polynomial-based model simulation

Num of pts used for model extraction	60pts	40pts
M1_LS	0.016	0.019
M3_LS	0.014	0.015
M5_LS	0.023	0.039

Table 6-3: The metric calculation results obtained from the DWLU simulation

As observed in table 6-2 and table 6-3, the polynomial-based modelling approach provides slightly worse simulation accuracy than the direct waveform utilization approach does. This is because errors have been introduced by interpolation and truncation of the order of model during polynomial based model extraction procedure.

Nevertheless, the modelling approach is still very robust and its prediction ability is very similar to the interpolation ability of the DWLU approach.

6.4. Comparisons on simulation speed

The simulation speed is an important factor which would affect the effectiveness of CAD-based simulation. Some of modern power amplifier structures use RF circuits having several transistor devices in parallel or/and in series. Simulation of such

circuit has been traditionally troublesome and challenging because models used in such circuit are very likely to face convergence problem. Most of modern RF CAD software use harmonic balance algorithm to solve nonlinear circuitry problems. Harmonic balance simulator works out the voltage and current information at every node of the nonlinear circuit iteratively. Simulation of a nonlinear circuit containing several nonlinear models would increase the required number of iterations dramatically. A nonlinear model with least simulation complexity can ease the task faced by the simulator and minimize the risk of no convergence. The simulation speed can be considered as a good indication of how easily the nonlinear simulator can deal with the nonlinear problem. In order to explore the potential of the proposed approaches in realistic power amplifier design applications, the simulation speed of the two proposed approaches is discussed in this section.

Experiments are designed to test and compare the simulation speed of the two approaches. Firstly the simulation speed is investigated as a function of number of power points swept in simulation. The load is fixed at a constant impedance point while different number of power points was swept in simulation. The simulation time is recorded accordingly. Figure 6-2 compares the results obtained using the DWLU approach and polynomial-based modelling approach.

In order to investigate how much the size of data table generated using the DWLU approach would affect the simulation speed, two data tables with different sizes were generated and used to run the same amount of simulation. As shown in Figure 6-2, the difference in size does have a small effect on the simulation time. The simulation time usually consists of two parts – one is the time for overhead processing of the data table and the other is the time for solving the nonlinear problems. Overhead processing mainly involves loading up the data table in memory for further analysis and simulation. A smaller size of data table would result in shorter overhead processing time and thus speed up the overall simulation. It can be seen from Figure 6-2 that reducing the size of data table by approx. 38% would save less than 10%

simulation time. It implies that the time for overhead processing of data table only takes a small portion in affecting the overall simulation speed.

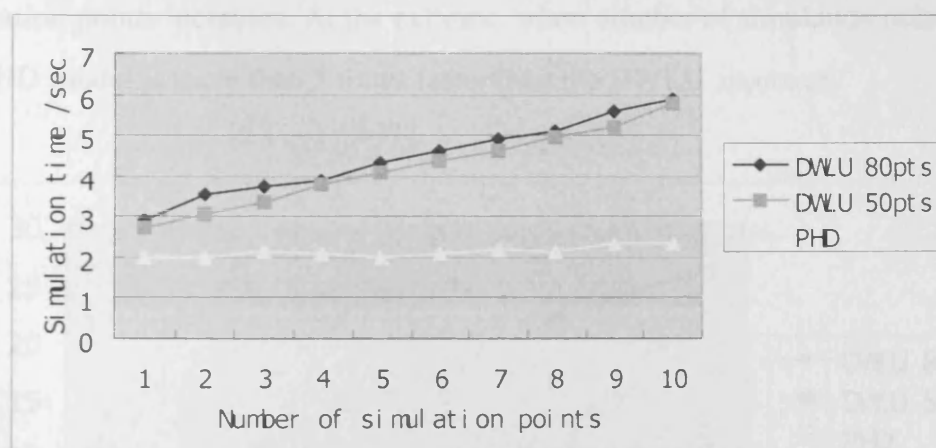


Figure 6-2: The comparisons of simulation speeds of using two approaches in power related simulations

By comparing the simulation time of using DWLU approach and the polynomial-based modelling approach in Figure 6-2, it can be seen that the modelling approach gives much more impressive simulation speed than the DWLU approach. At least 25 percent of speed improvement can be observed. Besides, the speed improvement becomes larger and larger and the speed increases by up to approx. 64 percent as the number of power points swept in simulation increases. When the number of power points is increased from 1 to 10, the speed of simulation using the modelling approach varies only slightly. On the contrast, the simulation time of using the DWLU approach increases significantly when the number of simulation points increases. This observation implies that the ADS simulator can quite happily deal with the polynomial-based model. The model's speed becomes more advantageous when a large amount of simulations are required.

Same observation can be obtained on Figure 6-3 as well. Figure 6-3 shows the simulation speed as a function of the number of swept load-pull points. It can be seen that as the number of simulation points increases from one to eighty, the

simulation time of using the polynomial-based model can be kept below 5 seconds. The simulation time of using the DWLU approach shows strong dependency on the number of simulation points. The simulation time increases linearly as the number of simulation points increases. At the extreme, when number of simulation points is 80, the PHD model is more than 5 times faster than the DWLU approach.

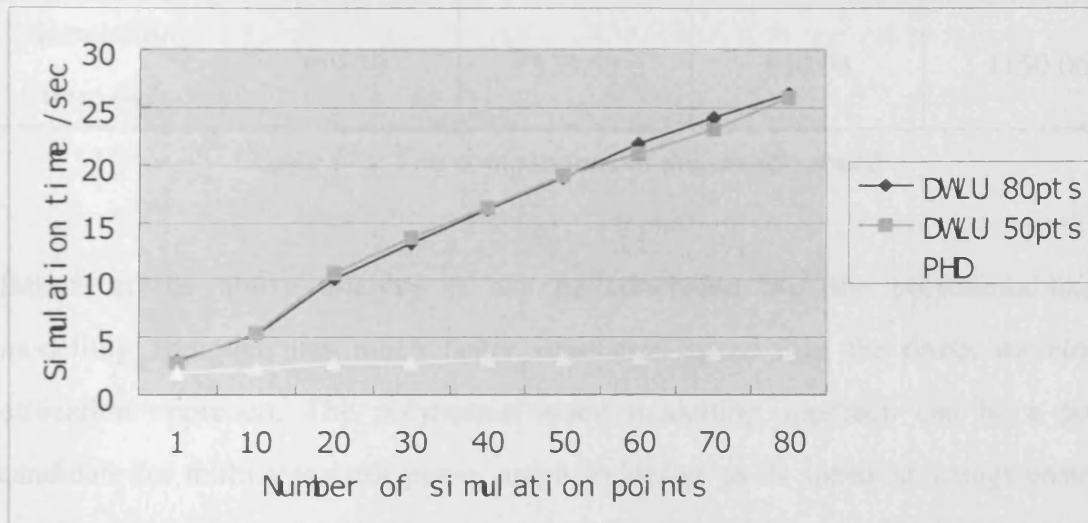


Figure 6-3: The comparisons of simulation speeds of using two approaches in load-pull related simulations

The above tests were based on single active device simulation. To move one further step closer to realistic scenario, simulations involving several transistor devices in one schematic are performed and the simulation time is recorded. To simplify the simulation setup, the transistor devices are inserted into separate simulation circuits. Simulations consist of both power sweeps and load-pull sweeps. The resultant measurement space contains 880 simulation points in total. Table 6-4 shows the simulation time for using different number of polynomial-based model and DWLU virtual device to run such amount of harmonic balance simulation in ADS. It can be observed that the simulation speed of operating several transistor devices in parallel using polynomial-based modelling approach is very impressive when compared to the DWLU approach. Simulation involving four polynomial-based models takes only approx 1 minute, which is already five times faster than the speed of using

single DWLU unit in simulation.

	1 POLY	2 POLY	3 POLY	4 POLY
Simulation time/sec	23.13	38.09	50.69	65.31
	1 DWLU	2 DWLU	3 DWLU	4 DWLU
Simulation time /sec	259.70	523.59	810.03	1150.06

Table 6-4: The comparisons of simulation speed

Based on the above analysis it can be concluded that the polynomial-based modelling approach has much faster simulation speed than the direct waveform utilization approach. The polynomial-based modelling approach can be a good candidate for multi-transistor power amplifier design as its speed advantage ensures the convergence of simulation. Compared to the polynomial-based modelling approach, the DWLU approach has a relatively slow speed which may restrict its applicability in power amplifier design.

6.5. Comparisons on approach complexity

To make any CAD-based approach useful in practical power amplifier design, it's important for the approach to be implemented with acceptable complexity. Power amplifier design may gain the advantage of accuracy by choosing a theoretically accurate approach with over-complicated implementation/extraction procedures but only at the expense of elongated design circle. Besides, the advantage of accuracy by using over-complicated approach may be seriously impeded by the error introduced in implementation procedure. It's therefore important to trade-off properly the complexity and accuracy when choosing a design strategy.

In this subsection, the complexity issues of the two novel CAD-based approaches

presented in this work, namely the direct waveform utilization approach and polynomial-based behavioural modelling approach, will be discussed. The difficulties within the implementation procedure of each approach will be highlighted. The overall complexity of the two approaches will be compared.

Generally speaking, neither of the two approaches is over-complicated. Each of them can be used to setup CAD simulation by using measurement waveform data within hours. Most of the data processing procedures involved in both approaches have been automated and require little human interfacing. Moreover, both of the two approaches are designed to be technology-independent and thus no modification to the approaches is required when the device or device technology is switched.

The two approaches have different levels of complexity when compared to each other. Figure 6-1, which shows the block diagrams of the two approaches, can be used to analyze the relative complexity of the two approaches. The DWLU approach only rescales the measurement data before importing them into CAD. The data re-scaling is done using Igor software. The time for doing this depends on the size and range of measurement data. Normally the re-scaling processing takes less than ten minutes to finish. Once the data are re-scaled there is no other data processing required. The re-scaled data are then ready to be used in CAD simulation. The necessary simulation setup in ADS is also quite straightforward. Depending on the specifications of the measurement domain, the simulation parameters need to be set properly. This procedure basically involves simple adjustment of the value of the variables in the simulation schematic and therefore only costs a few minutes or even less. By considering both the data rescaling and simulation setup procedures, the implementation of the direct waveform utilization approach only takes around twenty minutes.

On the other hand, the polynomial-based modelling approach seems to be more complex than the DWLU approach. As shown in Figure 6-1, the modelling approach

has three main data processing procedures. In practice, the data renormalization and interpolation procedures are done together by taking advantage of the ADS powerful functionalities. However, the current implementation of the two data processing procedures in ADS suffers from the lengthy simulation time – because the nonlinear simulation is a computationally heavy process. For the same amount of measurement data which requires a 10 minute re-scaling processing by the DWLU approach, time-consuming simulation which may take up to two hours is required. The reason for that is because simulation using the DWLU approach is involved in this process – simulation using DWLU approach has been proven in the previous section to be relatively slow. Since the DWLU approach offers much better multi-dimensional interpolation ability than any other tool available in the lab, it's still worth having it at the expense of the implementation time. The last data processing procedure involved in the modelling approach requires only simple mathematical calculation using integration, which can be done easily within minutes. Once all these three data processing procedures are done, the measurement data are now in a proper format and ready to be used in ADS. Regarding different requirements of different applications, different order of model is chosen by truncating the polynomial used for approximation. Accordingly, necessary modifications are also required to do on the simulation schematic. The simulation setup usually takes a few minutes. All in all, the polynomial-based behavioural modelling approach normally takes around three hours to process the measurement data and make the simulation ready to go. The two hours include the computer processing time as well as human efforts involved in the process.

Based on above analysis it can be seen that the DWLU approach gains quite a large advantage over the proposed behavioural modelling approach in terms of complexity. Specially, if referring to simulation time, the DWLU approach can shorten the distance from measurement to CAD-based simulation by a factor of 5 when compared to polynomial-based behavioural modelling approach.

6.6. Summary

A comprehensive study has been done on the two different waveform measurement data utilization approaches in order to investigate their simulation robustness, simulation speed and implementation complexity. Both theoretical discussion and experimental analysis were conducted to compare the performances of the two approaches on each of the aspects mentioned above respectively. Robustness analysis indicates the DWLU approach can perfectly reproduce the measured behaviour in CAD simulation thanks to its lossless data processing procedure. The polynomial-based modelling approach unavoidably introduces some errors into the CAD simulation due to the interpolation and polynomial truncation used in model extraction. But the polynomial-based model is still very robust and the actual error level is well below the resolution of the measurement system used for collecting the data for model extraction. Furthermore, when simulation off the measurement grid is needed, the polynomial-based model demonstrates same level of robustness as the DWLU approach. Simulation speed analysis shows that the proposed modelling approach gains huge advantage over the DWLU approach in terms of simulation efficiency. It can be seen that the simulation speed of using the modelling approach can be impressively five times faster than the simulation speed of using the DWLU approach. The speedy simulation makes the polynomial-based model suitable for complex power amplifier design which may involve developing RF circuits consisting of multi-transistors on one chip. Implementation analysis shows that the superior simulation speed of the polynomial-based model is gained by adopting complicated implementation procedures. Contrast to the DWLU approach which has only one step to process the measurement data before importing them into CAD for simulation, the polynomial-based modelling approach has three steps for data processing which takes much longer time than the DWLU approach. However, the data processing time that is talked about here is just within hours, which hardly makes the model generation a gating factor in the PA design process.

It can be found that the two approaches have different strengths and weaknesses. The different strength the approaches have make them suitable for different applications or design cases. Instead of adopting either one of them, it would be more much beneficial to PA designers if both approaches can be effectively combined and used together. In fact, the extraction of polynomial-based model has already benefited from the DWLU approach by making use of its robust interpolation ability. Combination of the two approaches would let the PA designers enjoy the robustness replication of measurement data and rapidity of simulation in CAD-based design process at the same time.

Chapter 7 Future Work

7.1 General

7.1.1 Improving measurement speed

The direct waveform utilization approach has demonstrated great potential in facilitating rapid CAD-based power amplifier design by using reliable measurement data. But currently the rapidity is limited by the relatively slow speed of measurement which is especially problematic when harmonic tuning is desired – such a design would require multi-dimensional measurements which can easily cover vast number of points. One solution to this problem, which has been discussed before in this thesis, is to do focused measurement in order to minimize the number of measurement points. However, this solution is not always useful as sometimes vast number of measurement points is unavoidable. One possible alternative is to speed up the present measurement system. Some promising new system configurations such as envelope load-pull system [56] are being developed which can offer much faster measurement speed, though the maximum power is limited by the system capability and still not suitable for high power measurement. Another possible alternative is to improve the measurement system software without modifying the system hardware configuration. As discussed in chapter three, the load-pull measurement is converged iteratively. For high power measurement sometimes the number of iterations necessary for converging to a load impedance point can be as high as 15. According to the present setting of the measurement system software, the 14 measurements done in iterations before the final converged measurement are simply discarded. Considering each of the 14 measurements contains information about DUT as valuable as the converged one, the information obtained from these measurements should be saved and utilized to speed up the following measurements. One interesting idea is to use interpolation to estimate the right value of a_2 based on the information obtained on the iterative measurements. Interpolation has demonstrated in the proposed direct waveform utilization approach that it can accurately predict the device characteristics even when the measurement grid is relatively coarse. This feature can be taken advantage of in measurement in order to reduce the number of iterations for each convergence.

The sequence of measurement can also be altered to make the interpolation mechanism more effective. Current software setup sweeps the load-pull grid on a column by column basis (figure 7-1). It can be changed to measure four outermost points at the corner first and then move towards centre. In this way the information obtained on measured outer grid can be used to estimate the measurement for inner grid by robust interpolation (figure 7-2).

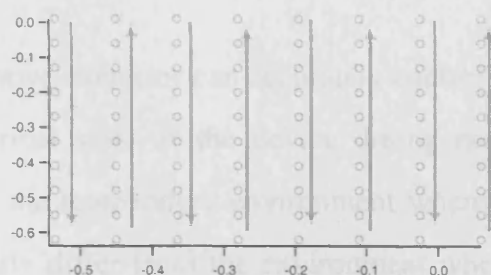


Figure 7-1 Traditional method of sweeping load reflection coefficients

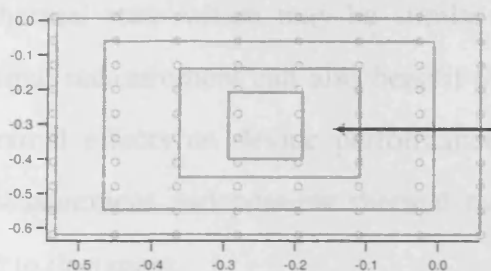


Figure 7-2 Proposed way of sweeping load reflection coefficients

The data utilization approaches presented in this thesis process the measurement data and import them into CAD for nonlinear simulation. No matter whether modelling or direct waveform utilization approach is used, such approach serves as the bridge to link the large signal waveform measurement system with CAD simulator.

7.1.2 Thermal measurement

One important external factor which may influence the performance of power transistor devices is temperature. The thermal status of device would affect the electron velocity in semiconductor and consequently change the characteristics of the device. In present measurement system, in order to get repeatable measurement, several things are done to maintain the thermal status of the DUT. Firstly air

conditioning system is implemented in the room where the measurement system is in order to make sure the thermal status of the measurement system and device under test is traceable before the measurement starts. During the measurement the DUT and system components which dissipate heat are cooled down with fans and heat sinkers

However, none of the above exercises can accurately control the thermal state of the device. Controlling thermal state of the device during measurement can greatly benefit PA designers as the temperature environment where the power amplifier is working may significantly differ from the environment where the power transistor device is characterized. If the thermal state can be controlled the device can be measured in different thermal state which may be similar to the device realistic working condition. Thermal measurement can also benefit the modelling society as the influence of the thermal effects on device performance can be quantitatively analyzed through the measurement and possible thermal model can be developed which can be very useful to designers.

It's not easy to control the thermal state of the DUT during measurement if considering the amount of heat dissipated. A water cooling system may be a good idea as it's relatively easy to implement and quite effective. Figure 7-3 shows an example of the cooling system proposed by ongoing research.

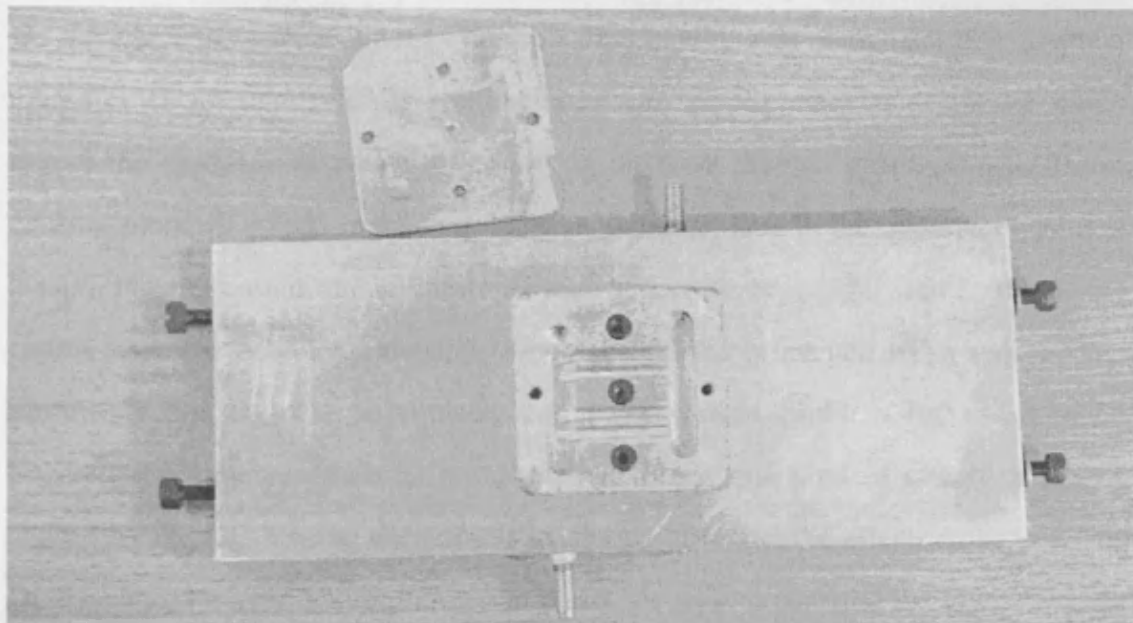


Figure 7-3: A test fixture with watering cooling system implemented [82]

Once the thermal state becomes controllable the device can be measured as a function of the temperature as well. Consequently, the DWLU approach can be extended by increasing the number of the dimension of data table by one. By using the imported measurement data in CAD, simulations which response to realistic temperature variation can be performed.

7.1.3 Modelling memory effects

The memory effects are annoying effects in power amplifier design. It represents a source of nonlinearity which is not formulated so far in the modelling approach described in this thesis. Since memory effects would contribute to the distortion characteristics of a power amplifier, neglecting memory effects in modelling may introduce error into simulation. So it's worth expanding the proposed modelling approach in the future by accounting for the memory effects.

There are basically three factors which contribute to memory effects. The thermal effects cause memory effects due to the heat generated during the operation of transistor devices which results changes of temperature and consequently dispersion

of device characteristics. It's possible to model this bit once the thermal measurement is available. The temperature can be a parameter of model which allows the model to response differently in different thermal states. Another factor causing memory effects is the base-band modulation. Due to the imperfect nature of the DC biasing circuit, the unflat impedance across the base-band seen by the device causes asymmetrical IM sidebands when the device is excited by a wideband or multi-tone signals. The asymmetry can be a serious problem for common PA linearization schemes such as pre-distortion. Since this kind of effects is mainly caused by the DC biasing network, it's possible to model the effects using practical DC biasing network in simulation. One possible way to describe the low frequency dispersion with the proposed DWLU approach is to introduce the biasing conditions as one of look-up parameters in the data table. In this way, the biasing variation and its effects can be demonstrated in simulation. The last important factor causing memory effects is the trapping effects [82]. Generally speaking it's difficult to model such effects. How to extend the existing modelling approach to describe the trapping effects remains an unknown topic that needs to be further explored. Some literatures which propose approaches to model trapping effects can be found in [82] [83].

7.1.4 Application of data utilization approaches in PA design

The best way to prove the usefulness of the data utilization approaches is through practical PA design. This thesis has primarily dealt with the fundamental issues about how to utilize the large signal waveform measurement data in CAD-based PA design, proposing novel data utilization approaches which need to be explored further in realistic PA design.

Both data utilization approaches can be used in any CAD-based power amplifier design with the polynomial-based modelling approach being advantageous in cases where the simulation speed is critical. Though both approaches can be used for stand-alone power amplifier design, it would be more interesting to explore the

usefulness of them in power amplifier design with complex configuration. For example, the Doherty structure [84] which has been around for decades now draws lots of attention as it features flat power efficiency for a relatively large dynamic range. Despite of the simplicity of the Doherty structure which basically has two active transistor devices in parallel, the design procedure can be troublesome and iterative as the two active devices load-pull each other during the operation of the PA, resulting complex interaction between the two devices. Analysis of such a structure can be greatly simplified if reliable CAD-based simulation is available. Implementation of these data utilization approaches in realistic PA structure is therefore desired.

7.2 Direct waveform utilization approach

7.2.1 Source-pull data utilization

The direct waveform utilization approach is basically developed by making use of the load-pull type waveform measurement data. The capability of the approach is actually limited by the capability of the measurement system. If other type of data is available such as source-pull data, it's also possible to extend the approach in order to take the new type of data into account.

The extension of the approach to account for the source pull data can follow the existing idea, i.e. just adding the corresponding independent variables in the data table to increase the number of dimensions of data. In implementation, additional circuitries similar to current schematic setup are also necessary to enable the source-pull data looking up in simulation.

7.2.2 Simulation with modulated signal

The thesis considers briefly the potential of using the DWLU approach in multi-tone signal simulation. Preliminary analysis is done to test the robustness of the approach in two tone simulation. The analysis is done by using simulation data obtained from

a Freescale MET model. Further investigations using measurement data and more realistic signals are desired. Such investigations can give a clue on how useful the CW measurement is when the characterized power transistor is intended to use in a communication system involving modulated signals (figure 7-4).

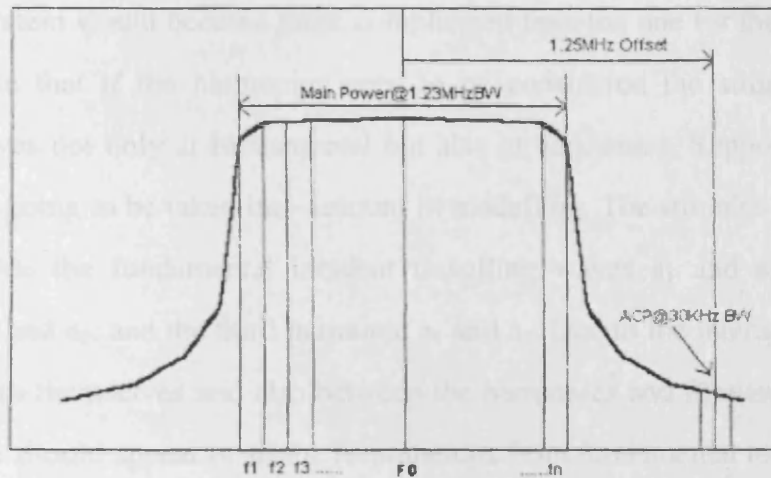


Figure 7-4: The CDMA signal specifications [86]

7.3 Polynomial-based behavioural modelling approach

7.3.1 Modelling DC and harmonics

The polynomial-based modelling approach developed in this work considers only the fundamental characterization of DUT when the harmonics are assumed to be terminated into system characteristic impedance, i.e. no significant incident energy waves are generated at harmonics. In other words, the device characteristics at DC and harmonics are not considered in the model formulation. The DC performance of a transistor device determines the efficiency of the device which is very important specification in power amplifier design. The harmonic components are less significant than the fundamental but harmonic tuning can sometimes improve the device performance significantly [86] due to the complex interaction between the fundamental and harmonic components within the system.

It can be seen that modelling DC and harmonics as well as fundamental are also essential to deliver a successful modelling approach which can be used in practical

power amplifier design. By taking a look at the current modelling formulation it can be found that generally the harmonic characteristics can be modelled in similar manner as the fundamental. New C and U parameters can be named to relate the harmonic responses of the DUT to the stimulus. But the model formulation for a harmonic system would become more complicated than the one for the fundamental system. Note that if the harmonics need to be considered the stimulus includes incident waves not only at fundamental but also at harmonics. Suppose up to third harmonic is going to be taken into account in modelling. The stimulus to the system would include the fundamental incident travelling waves a_1 and a_2 , the second harmonic a_1 and a_2 , and the third harmonic a_1 and a_2 . Due to the interaction between the harmonics themselves and also between the harmonics and fundamental, the six input stimuli should appear in all the formulations from fundamental to harmonics. It would then require very complicated model extraction and lengthy measurement data collection – the harmonics need to be swept as well as the fundamental which increases the measurement domain dramatically.

One intermediate step from fundamental-only solution to complex harmonic solution is to consider the case where the harmonics are all terminated into a match. In that special case, the stimulus of the system would exist only at fundamental. It can greatly simplify the model formulation as well as the extraction procedure and can be a good starting point for taking harmonics into consideration.

Modelling DC with the proposed modelling approach is a bit odd as usually DC behaviour is described using voltage and current rather than travelling waves. A simple solution to take DC into account is using table-based approach which has been described in chapter four of this thesis. The table-based solution relying on the interpolation has demonstrated robust simulation results. Therefore using table-based solution with DC can be fully feasible as no significant complexity is incurred.

7.3.2 Optimization of model extraction

The model extraction routine used currently for polynomial-based model relies on the interpolation to group the data in the way required for model extraction. The interpolation would unfortunately introduce some error into the model extraction but it has to be used as the measurement data available are collected by sweeping the load reflection coefficient but not the renormalized a_2 .

One possible way to optimize the model extraction is to remove the interpolation from it by collecting the measurement data in desired way, i.e. the sweep is done according to the renormalized a_2 but not the load reflection coefficient. This can be realized through modification of measurement system software. The measurement grid used currently is in a rectangular shape as shown in figure 7-5. If the measurement system is changed to sweep the renormalized a_2 an irregular grid would be resulted (figure 7-6). It can be seen that modifications on both the measurement system software and the model extraction routine are necessary to realize it.

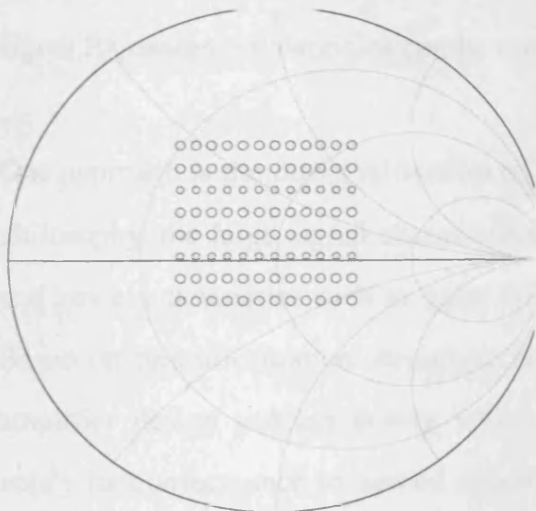


Figure 7-5: The rectangular shape of the load-pull grid

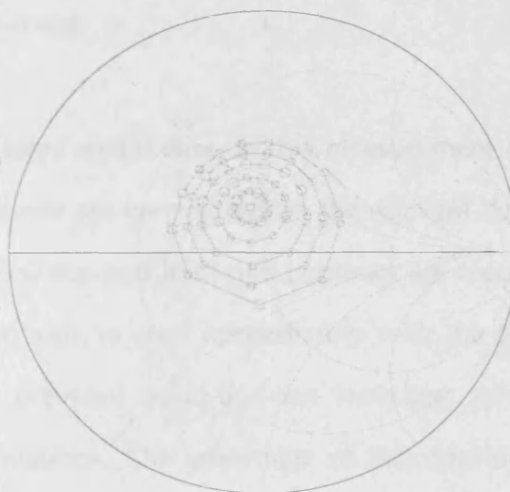


Figure 7-6: The irregular grid facilitating polynomial-based model extraction

Conclusion

Newly developed communication systems put strict requirements on the performance of RF power amplifiers. Their essential role in communication systems dictates for an accurate design process in order to meet the specific requirements of the numerous systems. A key issue for the development of such a design process is the inherent nonlinearity of power amplifiers hindering its integration with the well established small-signal development infrastructure which forms a closely interlinked chain of measurement systems, small-signal models and CAD based simulation and design software. The linkage between these components is provided by common small-signal data import and export file formats ensuring a bidirectional data exchange without any loss of the small-signal information. However, no equivalent infrastructure exists for a large-signal design process inhibiting the development of RF power amplifiers and other nonlinear components.

In order to deal with the nonlinear behaviour of devices two distinctive types of large signal PA design philosophies can be observed.

One approach is the direct utilisation of large signal data. In this measurement-based philosophy, the large signal characterisations are carried out on the relevant devices and key characteristics such as gain, efficiency and load-pull contours are recorded. Based on this information, designers are able to start immediately with the power amplifier design process during which practical build-and-test iterations serve to verify its conformance to agreed specifications. The advantage of this approach is that no time-consuming modelling is needed as the design is based directly on reliable and accurate measurement data. But the disadvantage of this approach is its lack of integration with advanced Computer-Assisted Design (CAD) software as in general the measurements represent only a sub-set of the nonlinear device information and as such do not allow the use of large-signal simulators which require

the complete vectorial information of present signals. Consequently, this approach lacks data-based large-signal device models and as a result relatively straightforward design changes trigger anew the entire design process.

Another approach is the generation of large-signal device models where, firstly, a device model is developed and the model parameters are identified from experimental data. Once sufficient accuracy is reached the model will be used for PA design in CAD environment. The designed PA will be finalised if all the design specifications are met during simulations. One advantage of this approach is the utilisation of nonlinear CAD tools allowing the testing a large number of design permutations within a short time. In addition, the process allows the observation of nonlinear parameters within any location of the simulated circuits providing a direct connection between the circuits and device performance enabling a fact-based design rather than relying on experience. However, the model-based philosophy also has some limitations such as the duration of modelling procedure or the model accuracy. Behavioural models reduce effectively the development and parameterisation time, however, still require extensive device measurements which often have been adopted for a specific model thus rendering them ineffective for direct utilisation within nonlinear CAD environment.

It is interesting to point out that either of the design approaches persists due to their complimentary properties yet are very distinctive with little if no defined interfaces. This is in strong contrast to small signal design process in which data can be readily imported into CAD to either form the basis of a measurement-based design approach or be utilised for small-signal modelling.

The thesis successfully demonstrates a coherent methodology to overcome the historical schism between both nonlinear design philosophies through the creation of a common and integrated soft- and hardware platform. The methodology proposes a way in which the nonlinear measurement data can be directly imported in CAD for

PA designs as well as behavioural model development. The import of large signal measurement data into CAD focuses on voltage and current waveforms rather than just the PA design related measurement. Also, a suitable large signal behavioural model has been proposed as the integrated part of the methodology.

Following an introduction which sets the theme of the research activity, **chapter two** reviews the literatures of the development of large signal measurement techniques and large signal modelling activities. The evolution of RF measurement instrumentation from traditional VNA, which captures the ratio of incident to reflective energy waves, to NVNA, which captures the absolute values of the energy waves, is briefly discussed. It can be found that the evolution is mainly driven by the need from RF power amplifier designers – nonlinearities dominant in large signal operation make it necessary to know both the magnitude and phase information of the DUT up to a few harmonics. Besides, due to the fact that the optimum impedance of a typical high power device is usually far away from system characteristic impedance and varying from device to device, necessary characterization tool, namely load-pull characterization, is developed and becomes a favourite feature for many large signal measurement systems. The measurement data obtained with these advanced large signal measurement instruments contain invaluable information about the DUT and how to utilize the data efficiently in CAD-based simulation is the main subject of this thesis. On the other hand, state of the art large signal modelling approaches published recently in literature have been reviewed in this chapter. Many of the modelling approaches were developed by making use of large signal measurement data.

One of the reasons for the separated design flows in the large-signal domain is that the large-signal data, as obtained for the PA design, cannot be coherently integrated into nonlinear CAD software. This is due to the fact that while modern nonlinear simulators are based on the analysis of current and voltage waveforms (or their frequency domain equivalent) the obtained PA design data contains only a fraction of

the current and voltage waveforms such as optimum input and output reflection coefficients. Consequently, import of large-signal data into CAD focused so far on the PA design relevant measurements.

The large signal measurement system presented in **chapter three** can provide large-signal data which contain more than merely PA design relevant measurement information. Chapter three describes a realized large signal waveform measurement system which has been developed in Cardiff University. The measurement system uses Hewlett Packard two-channel Microwave Transition Analyzer (MTA) to capture the incident and reflective energy waves sampled by couplers at both ports of DUT. Active load-pull capability up to third harmonic is implemented with ESGs and load power amplifiers. Particularly, since part of focus of this thesis is on developing design tools for high power amplifier design, necessary extensions were developed in this work which improve the robustness and accuracy of the high power measurement. The large signal waveform measurement data collected from this measurement system are used extensively in this thesis for developing, analyzing and verifying the proposed data utilization approaches.

The availability of large signal current and voltage waveforms offers a great opportunity for using the modern nonlinear simulator to do measurement-based power amplifier design. The waveform data provide sufficient information for nonlinear simulator to obtain complete solution for nonlinear circuit. **Chapter four** presents a novel approach for direct utilization of nonlinear waveform data in CAD-based simulation. The measurement waveform data are processed only with magnitude and phase scaling before imported into CAD nonlinear simulator. All the measurement information is retained and fully recoverable in simulation. The approach has been realized in Agilent ADS. The formulation and import of waveform data are made fully automated. Experiments indicate that the measured device nonlinear characteristics can be perfectly reproduced in simulation by using this approach. With proper simulation setup, the PA design related characteristics

such as load-pull contours can be faithfully replicated in simulation. Moreover, it's found that the approach has superior multi-dimensional interpolation capability even for devices demonstrating strong nonlinearities. The interpolation ability increases the robustness of the approach and makes the approach not only able to reproduce measured device waveform behaviour but also able to predict the device behaviour at the points not covered in measurement.

The direct use of non-linear data guarantees an accurate representation of the device characteristics within the CAD environment, hence increasing the likelihood of a first pass design – provided the device is simulated exactly under the same conditions. This requires the stimulus signal and the impedance environment, presented to the DUT, to match the ones used during the measurement. It is also worth noting that by using this approach, the large signal data can be utilized without any complex post processing or model extraction. It can be employed within a CAD environment directly after the large-signal device characterization.

The direct utilization of large signal waveforms demonstrates very good accuracy as long as the simulation point stays within the measurement space. However, this approach may degrade very quickly once the attempt is made to perform simulations outside the measurement space. Consequently, the proposed method requires a large number of measurements if good accuracy is desired and may encounter extraction difficulty if the required measurement space gets too large. Besides, during simulations, the information coupled to the load has to be looked up through a number of iterations, which makes the simulation very slow.

A new large signal behavioural modelling approach, presented in **Chapter five**, is developed in order to overcome the above limitations. The modelling-based approach can be a good supplementary tool to the measurement-based approach in power amplifier design. The modelling approach proposes a relatively simple model formulation which increases the nonlinear simulation efficiency dramatically. It is

important to see that the simulation accuracy doesn't degrade and sufficient accuracy can be achieved with fast simulation speed at the same time. The new large signal modelling technique is developed by making substantial extension to the Polyharmonic Distortion (PHD) model concept. One of the extension accounts for the influence of travelling wave a_2 magnitude presented on the output port on the variation of S and T parameters. It enhances the applicability of the modelling approach in the cases where large variation of magnitude of a_2 is necessary – not uncommon in modern communication system using dynamic envelope signals. Another important extension accounts for the optimum operation conditions of high power devices which often differ significantly from the typical 50Ω characteristic impedance. Theoretical analysis indicates that the proposed modelling approach has a solid mathematical foundation and should be considered as a reasonable extension to the traditional S parameters in large signal operation. Moreover, based on the theoretical analysis a more generic model 2-port formulation has been derived allowing the introduction of higher order phase terms for both of the incident waves a_1 and a_2 . The generalization of the model formulation allows the model to be tailored for different devices demonstrating different levels of nonlinearities by truncating the model formulation with different order and thus enhances the applicability and flexibility of the modelling approach.

A theoretically robust model is not necessarily accurate when used in realistic simulation. The accuracy of model extraction has great impact on the accuracy of extracted model in realistic simulation. The accuracy of model extraction is actually determined by the quality of extraction data and the robustness of extraction procedure. The measurement data for model extraction are obtained from fully calibrated measurement system with good repeatability. In order to make most of the reliable measurement data and maximize the accuracy of simulation achievable the approach of importing directly current and voltage waveform measurements into CAD, termed direct waveform look-up (DWLU) approach, is expanded through the establishment of a direct link with large-signal behavioural PHD model. For this

purpose extraction routines have been developed for the generation of higher order PHD models. As a result a new methodology is established giving the designer access to accurate large-signal simulations with the availability of first nonlinear measurements and enabling a switch towards a more compact and faster model formulation with improved scaling capability when sufficient data is available for the generation of an accurate PHD model. The decision when the model is sufficiently accurate can be made on designated sub-set of the measurements creating a self-consistent system. As it is based entirely on large-signal measurements its accuracy is guaranteed as long as the simulations do not extend beyond the measurement domain.

In order to illustrate the benefits that can be obtained by linking the two nonlinear data utilization approaches together, **chapter six** compares the performance of the direct waveform utilization approach and the polynomial-based modelling approach in terms of simulation robustness, simulation efficiency and implementation complexity. It can be seen that the integrated nonlinear measurement data utilization approach will benefit the PA designers by giving good CAD simulation accuracy both locally and globally in a relatively short amount of time. The direct import of data into CAD will guarantee the extremely good simulation accuracy within the measurement space. On the other hand, the polynomial-based model has been verified to be able to achieve good overall accuracy with excellent simulation speed even when the model is extracted from a small amount of measurement data. Therefore, both approaches can be combined and nicely considered as parts of an integrated nonlinear measurement data utilization strategy. Such a strategy provides a fast and time efficient path to accurate CAD-based nonlinear design; even at power levels relevant for base station applications.

LIST OF ABBREVIATIONS

<i>Term</i>	<i>Description</i>
ADS	Advanced Design System
ANN	Artificial Neural Network
AM/PM	Amplitude-dependent Phase distortion
BJT	Bipolar Junction Transistor
CAD	Computer-aided Design
CDMA	Code Division Multiple Access
CW	Continuous Wave
DAC	Data Access Component
DUT	Device under Test
DWLU	Direct Waveform Look-up
ESG	Electronic Signal Generator
FDD	Frequency Domain Device
FFT	Fast Fourier Transform
GPB	General Purpose Interface Bus
GSM	Global System for Mobile communications
HBT	Hetero-junction Bipolar Transistor
IF	Intermediate Frequency
IFT	Inverse Fourier Transform
IP	Intellectual Property
KCL	Kirchhoff's Current Law
LDMOS	Lateral-Drain-Diffusion MOS
LSNA	Large Signal Network Analyser
NVNA	Nonlinear Vector Network Analyser
MDIF	Microwave Data Interface File
MET	Motorola's Electro-Thermal model
MESFET	Metal Epitaxial Semiconductor Field Effect Transistor
MTA	Microwave Transient Analyser
PA	Power Amplifier
PAE	Power Added Efficiency
PAR	Peak-to-Average Ratio
PHD	Poly-harmonic Distortion
RF	Radio Frequency
SDD	Symbolically Define Device
TRL	Through Reflective Line calibration
VNA	Vector Network Analyser
VSWR	Voltage Standing-Wave Ratio
WCDMA	Wideband Code Division Multiple Access

REFERENCES

- [1] R.W. Anderson, "S-parameter techniques for faster, more accurate network design", HP Test&Measurement Application Note 95-1, 1997
- [2] D. Barataud, C. Arnaud, B. Thibaud, M. Campovecchio, J.-M. Nebus, and J. P. Villote, "Measurements of time-domain voltage/current waveforms at RF and microwave frequencies based on the use of a vector network analyzer for the characterization of nonlinear devices\—Application to high efficiency power amplifiers and frequency-multipliers optimization," IEEE Trans. Microwave Theory Tech., vol. 47, pp. 1259-1264, May 1998.
- [3] Maury Microwave Corporation, "Large-Signal Network Analyzer, Bringing Reality To Waveform Engineering," Technical Data Sheet 4T-090 rev A, June 2003.
- [4] T. Van den Broeck, J. Verspecht. "Calibrated Vectorial Nonlinear Network Analyser", *IEEE-MTT-S*, San-Diego, USA, 1994, pp. 1069-1072.
- [5] Agilent, <http://eesof.tm.agilent.com/docs/adsdoc2004A/manuals.htm>,
- [6] Ansoft, <http://www.ansoft.com/products/hf/nexxim/>
- [7] Peter B. Kenington, "High-Linearity RF Amplifier Design", Artech House, 2000
- [8] Oapos: Droma, M.S.; Mgebrishvili, N.; Goacher, A.A., "RF power amplifier nonlinearity modelling and intermodulation distortion analysis in OFDM signal transmitter systems", Microwave Conference, 2004. 34th European, Volume 1, Issue, 11-15 Oct. 2004 Page(s): 185 - 188
- [9] K.G. Gard, H.M. Gutierrez and M.B. Steer, "Characterization of Spectral Regrowth in Microwave Amplifiers Based on the Nonlinear Transformation of a Complex Gaussian Process," IEEE Transactions on Microwave Theory and Techniques, Vol. 47, No. 7, July 1999, pp. 1059-1069
- [10] Joel Vuolevi, Analysis, measurement and cancellation of the bandwidth and amplitude dependence of intermodulation distortion in RF power amplifiers, 2001
- [11] Draxler, P.; Langmore, I.; Hung, T.P.; Asbeck, P.M.; "Time domain characterization of power amplifiers with memory effects", Microwave Symposium Digest, 2003 IEEE MTT-S International, Volume 2, 8-13 June 2003 Page(s):803 - 806 vol.2
- [12] Williams, D.J.; Leckey, J.; Tasker, P.J.; "A study of the effect of envelope impedance on intermodulation asymmetry using a two-tone time domain measurement system", Microwave Symposium Digest, 2002 IEEE MTT-S International, Volume 3, 2-7 June 2002 Page(s):1841 – 1844
- [13] Understanding the Fundamental Principles of Vector Network Analysis, Agilent AN 1287-1, Application note, <http://cp.literature.agilent.com/litweb/pdf/5965-7707E.pdf>, 2000
- [14] Theory of Load and Source Pull Measurement, 5C-041 application note, <http://www.maurymw.com>, July 1999

- [15] Comparison of Active and Passive Load-Pull Test Benches, ARFTG Conference Digest-Spring, 57th, Publication Date: May 2001, Volume: 39, page(s): 1-4
- [16] C. Tsironis, B. Li, D. Dubouil, A. Henin, "A new method for test and design multistage power amplifiers using load-pull data", 53rd ARFTG Conference Digest, June 1999
- [17] S. Bali, A. Ferrero, "ATTS: Active load-pull system for characterization of power devices", Maury Microwave Corporation & P.A.F., Application Note 5C-046, July 2000
- [18] Z. Aboush, J. Lees, J. Benedikt, P. Tasker, "Active harmonic load-pull system for characterizing highly mismatched high power transistors," Microwave Symposium Digest, 2005 IEEE MTT-S International, pp. 1311 – 1314, June 2005.
- [19] T. Williams, J. Benedikt and P. J. Tasker, "Experimental Evaluation of an Active Envelope Load Pull Architecture for High Speed Device Characterization," in 2005 IEEE MTT-S International Microwave Symposium Digest, Long Beach, CA, USA, June 12-17, 2005, pp. 1509 – 1512
- [20] J. Benedikt, R. Gaddi, P.J. Tasker, M. Goss, "High-power time-domain measurement system with active harmonic load-pull for high-efficiency base-station amplifier design," IEEE Trans. Microwave Theory Tech., vol. 48, pp. 2617-2624, December 2000.
- [21] Maury Microwave Corporation, "Large-Signal Network Analyzer, Bringing Reality To Waveform Engineering," Technical Data Sheet 4T-090 rev A, June 2003.
- [22] Tom Van den Broeck, Jan Verspecht, "Calibrated Vectorial Nonlinear-Network Analyzers," 1994 IEEE MTT-S International Microwave Symposium Digest
- [23] J. Verspecht and P. Van Esch, "Accurately Characterizing Hard Nonlinear Behavior of Microwave Components with the Nonlinear Network Measurement System: Introducing 'Nonlinear Scattering Functions'," in Proc. 5th Int. Workshop on Integrated Nonlinear Microwave and Millimeterwave Circuits, Duisburg, Germany, Oct. 1998, pp. 17-26.
- [24] D. Schreurs, J. Verspecht, S. Vandenberghe, G. Carchon, K. van der Zanden, and B. Nauwelaers, "Easy and accurate empirical transistor model parameter estimation from vectorial large-signal measurements," in IEEE MTT-S Int. Microwave Symp. Dig., Anaheim, CA 1999, pp. 753-756.
- [25] Chris Roff, Johannes Benedikt, Paul J. Tasker, "Design Approach for Realization of Very High Efficiency Power Amplifiers," IEEE MTT-S Int. Microwave Symp. Dig. 2007
- [26] Y. Tsividis, Operation and modeling of the MOS transistor, New York: McGraw-Hill, 1987.
- [27] F. Filicori, G. Ghione, C.U. Naldi, "Physics-Based Electron Device Modelling and Computer-Aided MMIC Design," IEEE Trans. Microwave Theory Tech., vol.40, pp.1333-1352, July 1992.
- [28] C. M. Snowden, "Nonlinear Modelling of Power FETs and HBTs," International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering, 1997

- [29] Prigent, M.; Nallatamby, J.C.; Camiade, M.; Nebus, J.M.; Ngoya, E.; Quere, R.; Obregon, J., "Comprehensive Approach to the Nonlinear Design and Modelling of Microwave Circuit", URSI International Symposium on Volume , Issue , 29 Sep-2 Oct 1998, Page(s):450 - 455
- [30] Issaoun, A.; Barrak, R.; Kouki, A.B.; Ghannouchi, F.M.; Akyel, C.; "Enhanced empirical large-signal model for HBTs with performance comparable with physics-based models", Science, Measurement and Technology, IEE Proceedings-Volume 151, Issue 3, 2 May 2004 Page(s):142 – 150
- [31] José C. Pedro, Senior Member, IEEE, and Stephen A. Maas, Fellow, IEEE, "A Comparative Overview of Microwave and Wireless Power-Amplifier Behavioral Modeling Approaches", IEEE Transactions on Microwave Theory and Techniques, VOL. 53, NO. 4, April 2005
- [32] D. Schreurs, J. Wood, N. Tufillaro, D. Usikov, L. Barford, D.E. Root, "The Construction and Evaluation of Behavioral Models for Microwave Devices Based on Time-Domain Large-Signal Measurements", 2000 IEEE
- [33] M. Schetzen, the volterra and wiener theories of nonlinear systems, New York, John Wiley& Sons, 1980
- [34] Verbeyst, F.; Bossche, V.; "VIOMAP, the S-parameter equivalent for weakly nonlinear RF and microwave devices", Microwave Theory and Techniques, IEEE Transactions on Volume 42, Issue 12, Part 1-2, Dec 1994 Page(s):2531 – 2535
- [35] S. A. Maas, Volterra Methods for Behavioural Modelling, IMS 2003 Workshop on Fundamentals of Nonlinear Behavioural modeling, June 2003
- [36] Dooley, J.; Oapos;Brien, B.; Brazil, T.J., "Behavioural modelling of RF power amplifiers using modified Volterra series in the time domain", High Frequency Postgraduate Student Colloquium, 2004 Volume , Issue , 6-7 Sept. 2004 Page(s): 169 - 174
- [37] Soury, A.; Ngoya, E.; Nebus, J.M.; "A new behavioral model taking into account nonlinear memory effects and transient behaviors in wideband SSPAs", Microwave Symposium Digest, 2002 IEEE MTT-S International Volume 2, 2-7 June 2002 Page(s):853 – 856
- [38] John Wood, David E. Root, "Fundamentals of Nonlinear Behavioural Modelling for RF and Microwave Design", Artech House, 2005
- [39] Qi-Jun Zhang; Gupta, K.C.; Devabhaktuni, V.K., "Artificial neural networks for RF and microwave design - from theory to practice", Microwave Theory and Techniques, IEEE Transactions on page(s): 1339- 1350, Volume: 51, Issue: 4, Apr 2003
- [40] Y. H. Fang, M. C. E. Yagoub, F. Wang, and Q. J. Zhang, "A new macromodeling approach for nonlinear microwave circuits based on recurrent neural networks," IEEE Trans. Microwave Theory Tech., vol. 48, pp. 2335-2344, Dec. 2000
- [41] Deo, Makarand; Xu, Jianjun; Zhang, Q.J.; "A New Formulation of Dynamic Neural Network for Modeling of Nonlinear RF/Microwave Circuits", European Microwave Conference, 2003. 33rd Oct. 2003 Page(s):1019 - 1022
- [42] Jianjun Xu; Yagoub, M.C.E.; Runtao Ding; Qi-Jun Zhang; "Neural-based

- dynamic modeling of nonlinear microwave circuits”, Microwave Theory and Techniques, IEEE Transactions on Volume 50, Issue 12, Dec. 2002 Page(s):2769 – 2780
- [43] Root, D.E.; Verspecht, J.; Sharrit, D.; Wood, J.; Cognata, A.; “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 Volume 53, Issue 11, Nov. 2005 Page(s):3656 – 3664
- [44] J. Benedikt, R. Gaddi, P.J. Tasker, M. Goss, “High-power time-domain measurement system with active harmonic load-pull for high-efficiency base-station amplifier design,” IEEE Trans. Microwave Theory Tech., vol. 48, pp. 2617-2624, December 2000.
- [45] HP 71500A Microwave Transition Analyzer, Agilent, <http://cp.literature.agilent.com>
- [46] F. van Raay and G. Kompa, “A new on-wafer large-signal waveform measurement system with 40 GHz harmonic bandwidth,” in 1992 IEEE MTT-S Int. Microwave Symp. Dig., 1992, vol. 3, pp. 1435–1438.
- [47] M. Demmler, P.J. Tasker, and M. Schlechtweg, “On-wafer large signal power, S-parameter and waveform measurement system,” in Conf. Rec. Third Int. Workshop Integrated Nonlinear Microwave Millimeterwave Circuits, Duisburg, Germany, Oct. 1994, pp. 153–158.
- [48] J. Benedikt, “Novel high frequency power amplifier design system”, PhD thesis submitted to Cardiff University, UK, 2002
- [49] R. Gaddi, “On the characterization and modeling of silicon RF LDMOS transistors”, PhD thesis submitted to Cardiff University, UK, 2002
- [50] D. Williams, “Non-Linear Measurement System and Techniques for RF Power Amplifier Design”, PhD thesis submitted to Cardiff University, UK, 2003
- [51] IEEE standard 488.1-1987 Standard Digital Interface for Programmable Instrumentation, <http://standards.ieee.org>
- [52] Igor Pro User Manual, ver5, WaveMetrics Inc, 2005
- [53] Doug Rytting, TRL Calibration, <http://cpd.ogi.edu/IEEE-MTT-ED/TRL%20Calibration.pdf>
- [54] Hashmi, M. S.; Hashim, S.J.; Williams, T.; Benedikt, J.; Tasker, P. J.; “A Broadband Control Electronics for Envelope Load Pull System”, Signals, Systems and Electronics, 2007. ISSSE '07. International Symposium on July 30 2007-Aug. 2 2007 Page(s):197 – 200
- [55] Williams, T.; Benedikt, J.; Tasker, P.J.; “Fully Functional "Real Time" Non-Linear Device Characterization System Incorporating Active Load Control”, Microwave Conference, 2006. 36th European Sept. 2006 Page(s):1610 – 1613
- [56] Williams, T.; Benedikt, J.; Tasker, P.J.; “Experimental evaluation of an active envelope load pull architecture for high speed device characterization”, Microwave Symposium Digest, 2005 IEEE MTT-S International 12-17 June 2005 Page(s):4 pp.
- [57] Mathew Jacob, Optimizing RF Power Amplifier System Efficiency Using

- DC-DC Converters, Power Designer issue no. 110, National Semiconductor Corporation, 2006
- [58] Advanced Design System, Agilent, <http://eesof.tm.agilent.com/>
- [59] Microwave office, AWR, http://web.appwave.com/Products/Microwave_Office
- [60] HFSS, Ansoft, <http://www.ansoft.com/products/hf/hfss/>
- [61] Y. Tajima, B. Wrona, and K. Mishima, "GaAs FET Large-Signal Model and Its Application to Circuit Designs," IEEE Transactions on Electron Devices, Vol. 28, No. 2, February 1981, pp. 171-175.
- [62] W.R. Curtice, and M. Ettenburg, "A Non-linear GaAs FET Model for Use in the Design of Output Circuits for Power Amplifiers," IEEE Transactions on Microwave Theory and Techniques, Vol. 33, No. 12, 1985, pp. 1383-1394.
- [63] A. Platzker, A. Palevski, S. Nash, W. Struble, and Y. Tajima, "Characterization of GaAs Devices by a Versatile Pulse IV Measurement System," IEEE MTT-S International Microwave Symposium Digest, Vol. 3, May 1990, pp. 1137-1140.
- [64] G. D. Dambrine, A. Cappy, F. Helidore, and E. Playez, "A new method for determining the FET small-signal equivalent circuit," IEEE Trans. Microwave Theory Tech., vol. 36, pp. 1151-1159, July 1988.
- [65] Burden, R. L.; Faires, J. D.; and Reynolds, A. C. Numerical Analysis, 6th ed. Boston, MA: Brooks/Cole, pp. 120-121, 1997
- [66] Abramowitz, M. and Stegun, I. A. (Eds.). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. New York: Dover, p. 18, 1972
- [67] Stefan Jahn, Michael Margraf, Vincent Habchi, Raimund Jacob, Qucs: Technical Papers, 2005
- [68] M.B. Steer, "Accurate Multisine Representation of Digital Communication Signals for Characterization of Nonlinear Circuits," 2006 Radio and Wireless Symposium, pp. 527-530, January 2006
- [69] J. C. Pedro and N. B. de Carvalho, "On the use of multitone techniques for assessing RF components' intermodulation distortion," IEEE Trans. Microw. Theory Tech., vol. 47, no. 12, pp. 2393-2402, Dec. 1999.
- [70] K. A. Remley, "Multi-sine excitation for ACPR measurements," IEEE MTT-S Int. Microw. Symp. pp. 2141-2144. Jun. 2003
- [71] W. Curtice, J. Pla, D. Bridges, T. Liang. E. Shumate, "A New Dynamic Electro-Thermal Nonlinear Model for Silicon RF LDMOS FETs," 1999 IEEE MTT-S International Microwave Symposium, Anaheim CA, pp. 419-422, June 1999.
- [72] Freescale Semiconductor, "Agilent's Advanced Design System RF High Power Products Design Kit Release v2004ap0105," Release Notes and Installation Instructions, Rev. 2, Feb 2005
- [73] Ngoya, E., Larcheveque, R., "Envelop transient analysis: a new method for the transient and steady state analysis of microwave communication circuits and systems," IEEE MTT Symposium Digest, 1996, pp. 1365-1368
- [74] Agilent Technologies, "Circuit envelope simulation", <http://eesof.tm.agilent.com/docs/adsd0c2004A>, 2004

- [75] Verspecht, J.; Root, D.E. Polyharmonic distortion modeling, Microwave Magazine, IEEE, Volume: 7, Issue:3 On page(s): 44- 57, June 2006
- [76] D. Schreurs, J. Verspecht, B. Nauwelaers, A. Barel, M. Van Rossum, "Waveform Measurements on a HEMT Resistive Mixer," 47th ARFTG Conference Digest, June 1996, pp. 129 -135
- [77] J. Verspecht, "Large-signal Network Analysis – 'Going Beyond S-Parameters'", 62nd ARFTG Conference Short Course Notes, December 2003
- [78] Arbitrary Impedance, Application Note, Anritsu Company, 2002
- [79] Pirazzini, M.; Fernandez, G.; Alabadelah, A.; Vannini, G.; Barciela, M.; Sanchez, E.; Schreurs, D., "A preliminary study of different metrics for the validation of device and behavioral models," ARFTG Conference Digest, 2005. Spring 2005. 65th , vol., no.pp. 8 pp.-, 17 June 2005
- [80] MaShu, MSc Thesis submitted to Cardiff University, UK, 2007
- [81] Binari, S.C.; Ikossi, K.; Roussos, J.A.; Kruppa, W.; Doewon Park; Dietrich, H.B.; Koleske, D.D.; Wickenden, A.E.; Henry, R.L.; Trapping effects and microwave power performance in AlGaIn/GaN HEMTs Electron Devices, IEEE Transactions on Volume 48, Issue 3, March 2001 Page(s):465 – 471
- [82] Siriex, D.; Barataud, D.; Sommet, R.; Noblanc, O.; Ouarch, Z.; Brylinski, Ch.; Teyssier, J.P.; Quere, R.; "Characterization and modeling of nonlinear trapping effects in power SiC MESFETs", Microwave Symposium Digest., 2000 IEEE MTT-S International Volume 2, 11-16 June 2000 Page(s):765 - 768 vol.2
- [83] Ouarch, Z.; Collantes, J.M.; Teyssier, J.P.; Quere, R.; "Measurement based nonlinear electrothermal modeling of GaAs FET with dynamical trapping effects", Microwave Symposium Digest, 1998 IEEE MTT-S International Volume 2, 7-12 June 1998 Page(s):599 - 602 vol.2
- [84] S. C. Cripps, RF Power Amplifiers for Wireless Communications, Artech House, 1999
- [85] W. Xinwei, H. Nakamura, R. Singh, ACPR, IM3 and their Correlation for a PCS CDMA Power Amplifier, 50th ARFTG CONFERENCE DIGEST, 2001
- [86] Kopp, B.; Heston, D.D.; "High-efficiency 5-watt power amplifier with harmonic tuning", Microwave Symposium Digest, 1988., IEEE MTT-S International 25-27 May 1988 Page(s):839 - 842 vol.2

APPENDICES

Appendix A Software User Manual

A.1. DWLU Table Generator

This section provides a detailed explanation of the panels and options available in the DWLU data processing software. The software is developed using Igor. All the controls and parameter settings needed during table generation can be done via graphic user-friendly interface.

A.1.1 Main user interface

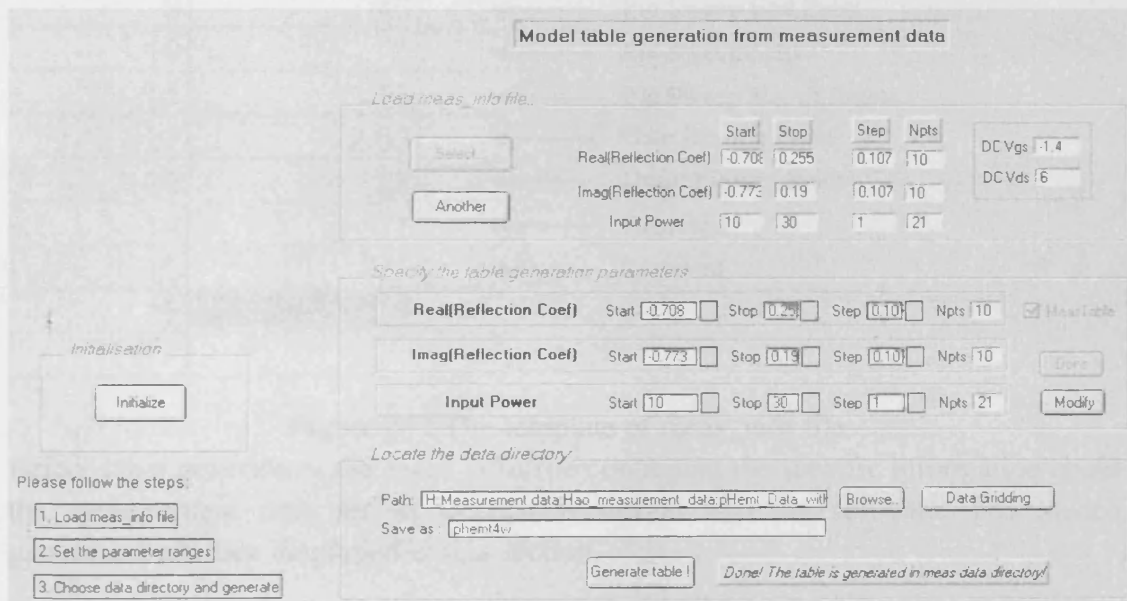


Figure A-1 The main user interface of DWLU table generator

As shown in Figure A-1, there are four sections of functions that can be controlled via the interface. They are “Initialisation”, “Load meas_info file”, “Specify the table generation parameters”, and “Locate the data directory”.

Initialisation

Initialize the table generation software. Clear all the global variables used in the software and set the initial value to zero.

Load meas_info file

Different measurement data sets may be collected with different external parameter settings such as input power range, load-pull area and etc. For each measurement

data set, an information file about the measurement sweep parameters is built by using the template below:

Table0:meas_info		
R0C0		-0.535
Point	meas_info	
0	-0.535	← Real(Γ_{Load}) Sweep Start Point
1	0.06	← Real(Γ_{Load}) Sweep End Point
2	0.085	← Real(Γ_{Load}) Sweep Step
3	8	← Real(Γ_{Load}) Sweep No. of Points
4	-0.62	← Imag(Γ_{Load}) Sweep Start Point
5	0.01	← Imag(Γ_{Load}) Sweep End Point
6	0.07	← Imag(Γ_{Load}) Sweep Step
7	10	← Imag(Γ_{Load}) Sweep No. of Points
8	28	← Pin Sweep Start Point
9	40	← Pin Sweep End Point
10	1	← Pin Sweep Step
11	13	← Pin Sweep No. of Points
12	2.63	← Gate Biasing Point
13	28	← Drain Biasing Point
14	1	← Reserved
15	1	← Reserved
16		

Figure A-2 The template of meas_info file

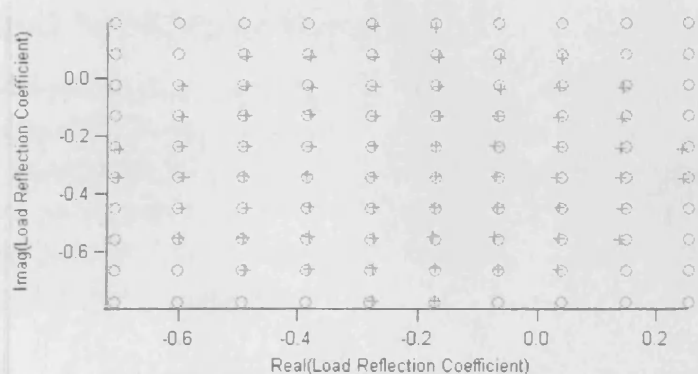
Before table generation, the meas_info file containing the specific information about the measurement data set of interest is loaded into the software. The loaded parameters are then displayed in this section.

Specify the table generation parameters

It's possible to tweak the table generation parameters, such as the table coverage in terms of input power range and load-pull area, in this section. It can be useful when there is a need to generate a table with a subset of measurement data available.

Locate the data directory

Locate the measurement data by specifying the path. Also need to specify a path and name for the output data table. The data gridding button performs the initial analysis of the measurement data by finding the actual measurement grid. It will bring up a new dialog through which the actual grid of the data table to be generated can be specified.



(a)

(b)

Figure A-3 (a) shows the actual measurement grid (cross) and table generation grid (circle). (b) shows the panel used to control the table generation grid

Once the grid is set, by pressing button “Generate table!” a data table will be generated according to the setting specified in this software.

A.1.2. Software operation flow

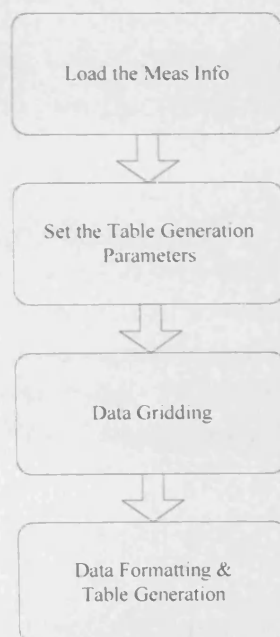


Figure A-4 The DWLU table generation flow

A.2. Polynomial-based Model Extractor

The extraction of polynomial-based model proposed in this thesis is done using Igor and ADS software. The model parameters are extracted by running ADS simulation and using the build-in functions provided in the ADS Data Display Window. The extracted model parameters from ADS simulations are saved into txt files. The txt files are then loaded into Igor and formatted into MDIF format. The extraction flow is illustrated in the following graph:

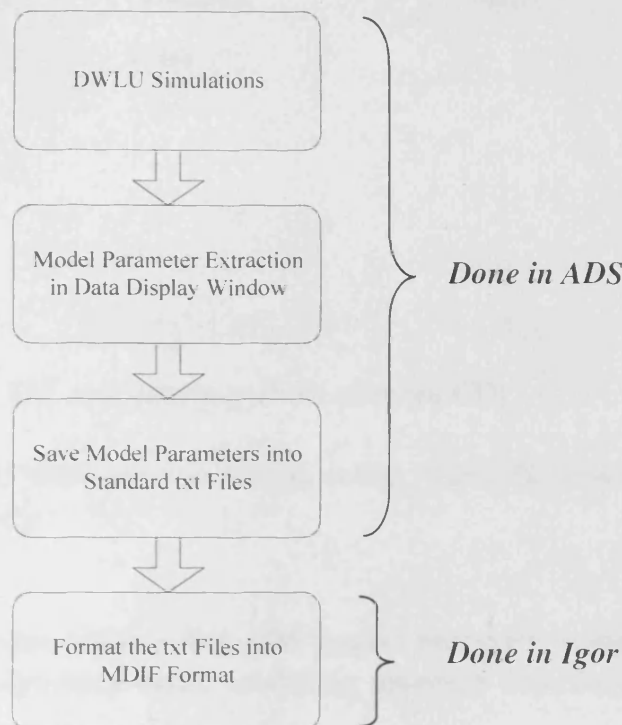


Figure A-5 The Polynomial-based model table extraction flow

Since the extraction of model is a relatively straightforward process, no graphic user interface is developed. The final model table formatting is done through command lines in Igor.

CU data load (num a1, num a2, num coeff)

This function is developed to format the polynomial-based model parameters into a MDIF table. Firstly all the extracted model parameters are loaded from txt files into Igor internal memory. Secondly the model parameters are read from Igor memory and formatted into several MDIF tables. These generated MDIF tables can then be used with DAC component in ADS to simulate the DUT characteristics.

Appendix B Readme for the attached CD

The CD-ROM attached in this thesis contains all the important software programmes and ADS projects developed in this work.

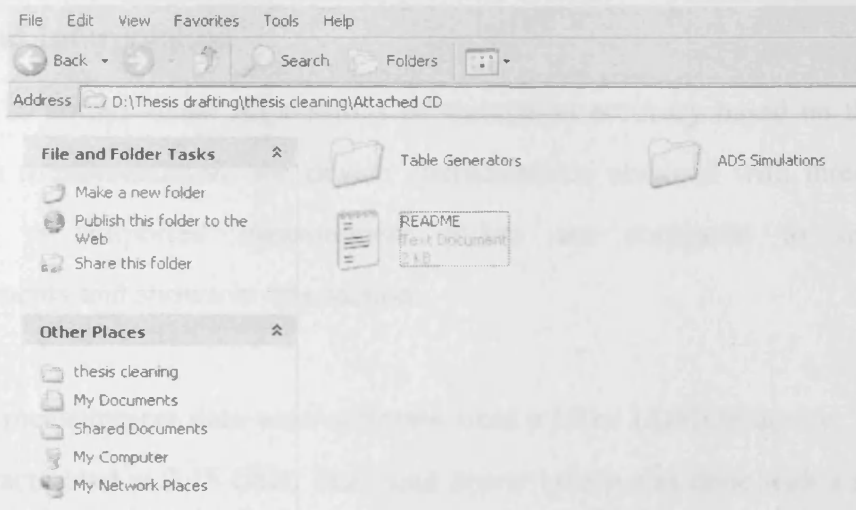


Figure B-1: The root directory of the attached CD

It can be seen that in the CD there are two folders called “Table Generators” and “ADS Simulations” respectively.

Table Generators Folder

This folder contains the software routines and ADS project necessary to implement the DWLU approach and Polynomial-based modelling approach described in this work.

- ✓ *DWLU_Table_Generator*: Software routines developed to process the measurement data according to the proposed DWLU approach
- ✓ *POLY_EXTRACTION_prj*: Compressed ADS project developed to process the measurement data and extract the model parameters according to the proposed Polynomial-based modelling approach
- ✓ *Poly based Model Table Gen via ADS v1*: Software routines developed to format the extracted polynomial-based model parameters into standard MDIF format

ADS Simulations

This folder contains the ADS projects developed to implement the DWLU approach and Polynomial-based modelling approach. Sample schematics and result display templates have been developed to facilitate the interrogation of the proposed approaches.

- ✓ *DWLU_prj*: Compressed ADS project about the DWLU approach implementation
- ✓ *POLY_prj*: Compressed ADS project about the polynomial-based modelling approach implementation

Appendix C Additional DWLU Approach Verification Results

C.1 Load Interpolation

In order to obtain visual impressions of simulation accuracy based on the DWLU approach implementation, the device characteristics obtained with three different numbers of imported measurement points are compared to independent measurements and shown in this section.

The CW measurement data were collected from a 100w LDMOS device. The device was characterized at 2.15 GHz. The input power sweep was done with a range from approx. 29dBm to approx. 40dBm with a step size of 1dB. The load reflection coefficient sweep was done with an 8 by 10 grid on the Smith Chart. The maximum output power variation within the load-pull area is about 3dB.

Number of imported measurement points: 70

The load reflection coefficient used in simulation is marked on the graph on the right. The blue grid indicates the measurement points imported into ADS for this interpolation. The interpolated output

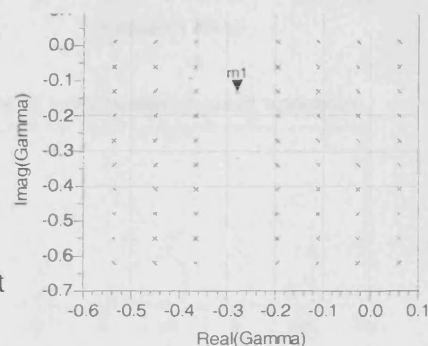


Figure C-1 Imported measurement points

voltage and current waveforms are plotted and checked against measurement as shown in above

figures. The agreement between waveforms is good for both voltage and current. Comparisons are also made in frequency domain to give a much clearer indication of how much the discrepancy is. The discrepancy is also quantified at each harmonic and plotted. It can be found that the voltage and current spectra have identical error percentage. It's because harmonic balance simulation calculates the current by dividing the voltage with impedance. Since the impedance in simulation is always set to the identical value used in measurement, the error in voltage spectrum would

be equally transferred into current spectrum. As observed the accuracy at fundamental and second harmonic are good and spectrum errors at those frequency points are within 1%. The error at third harmonic may look a bit large which is around 7%. Note that the magnitude of the spectral component at third harmonic is very small and close to zero. A relatively large error shouldn't affect the overall accuracy of simulation.

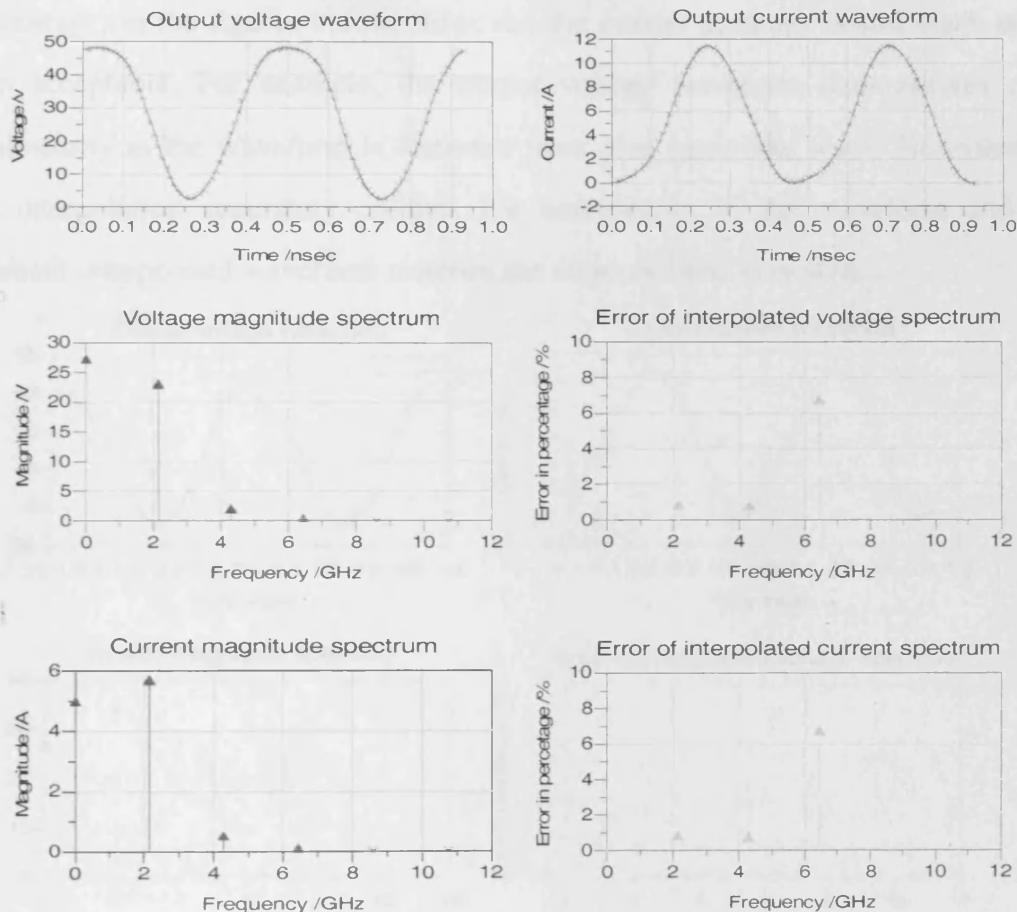


Figure C-2: The measured and simulated output voltage and current

Number of imported measurement points: 50

Data at 50 measurement points, which are displayed as the blue grid on the right, are imported in ADS for testing the interpolation capability of the proposed DWLU approach. The interpolation results shown in this page are obtained

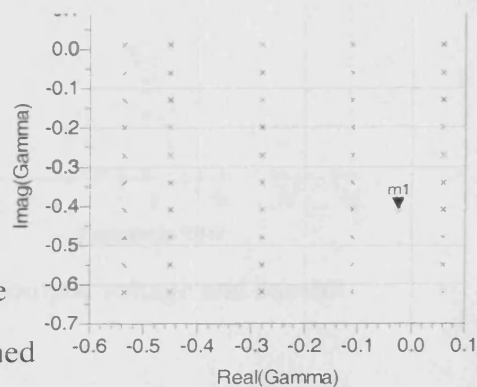


Figure C-3 Imported measurement points

at the load reflection coefficient point marked as m1. Note that Cubic Spline algorithm interpolates the value at the target point by using information from two known data points at each side of the target point. Comparing to the interpolation done with 70 imported measurement points, sparser grid makes the interpolation more challenging with 50 imported measurement points. As expected, the accuracy of interpolation degrades a bit, which can be observed from the error bar (in percentage) in the figures below. However, the overall accuracy is still much better than acceptable. For example, the output voltage waveform demonstrates clear nonlinearity as the waveform is distorted from pure sinusoidal wave. Nevertheless, the interpolation accurately predicts the nonlinearity of the waveform and the resultant interpolated waveform matches the measured one very well.

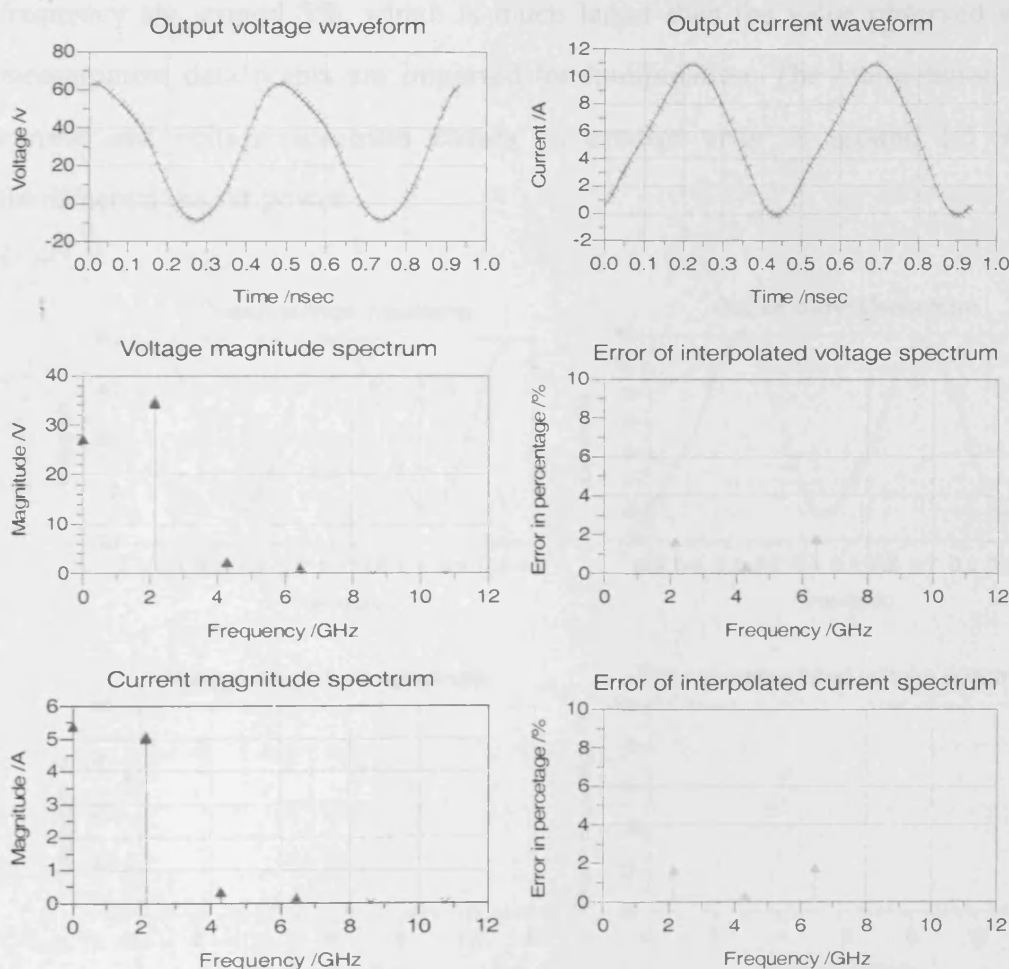


Figure C-4: The measured and simulated output voltage and current

Number of imported measurement points: 30

Data at only 30 measurement points are imported into ADS to do interpolation. Note that the imported grid covers a load pull area within which the output power varies from 100 Watts to around 50 Watts. The waveform

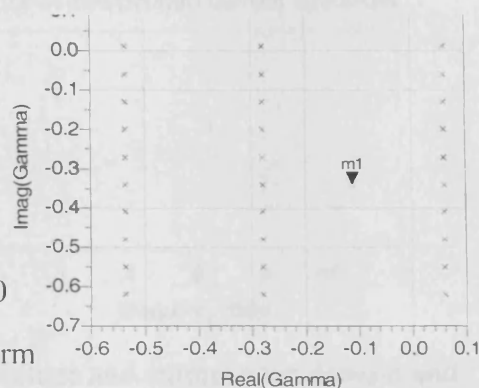
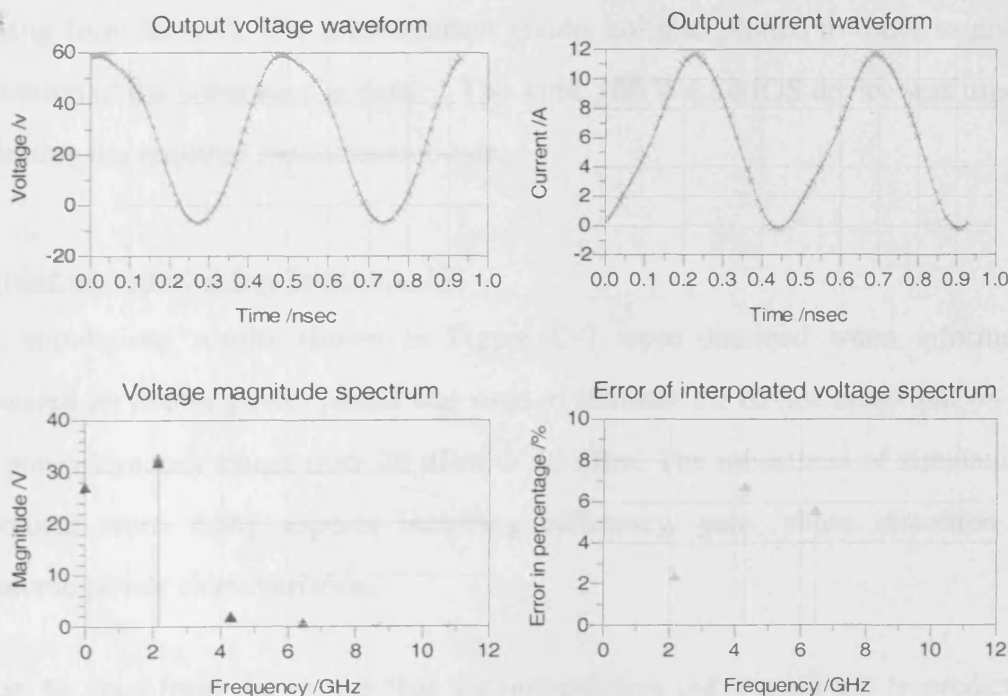


Figure C-5 Imported measurement points

comparisons show discrepancy between interpolated and measured results and maximum discrepancy can

be observed at peaks of the waveforms. The errors shown in frequency domain confirm the observation in time domain. The error percentages at fundamental frequency are around 3%, which is much larger than the value observed when 70 measurement data points are imported for interpolation. The interpolation error in current and voltage spectrum causes an average error of around 1.5 Watts in fundamental output power.



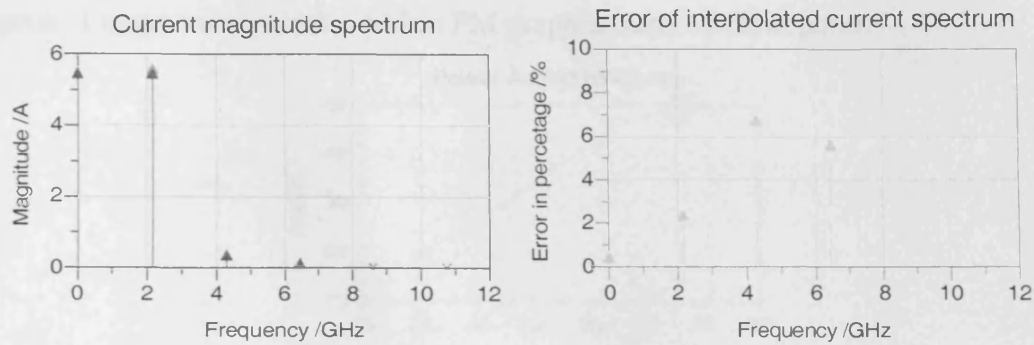


Figure C-6: The measured and simulated output voltage and current time domain and frequency domain characteristics

C.2 Input Power Interpolation

The following subsection shows the simulated waveforms and output power characteristics of the device obtained with different power sweep density, in order to give visual impression on the accuracy of the interpolation. The simulation results are obtained when the number of measurement power points used in simulation is varying from 12 to 5. The measurement results are also plotted in order to give an indication of the simulation accuracy. The same 100 W LDMOS device was used for collecting the required measurement data.

Number of imported power points: 12

The simulations results shown in Figure C-7 were obtained when information measured on twelve power points was used to simulate the device behaviour on a 14 dB power dynamic range from 26 dBm to 40 dBm. The robustness of simulation is examined from many aspects including efficiency, gain, phase distortion and harmonic power characteristics.

It can be seen from the results that the interpolation did a good job to predict the device behaviour over such a large dynamic range. Accurate agreements can be observed on all the graphs. No visual discrepancy can be observed except for an

approx. 1 degree error on the AM to PM graph around 37 dBm point.

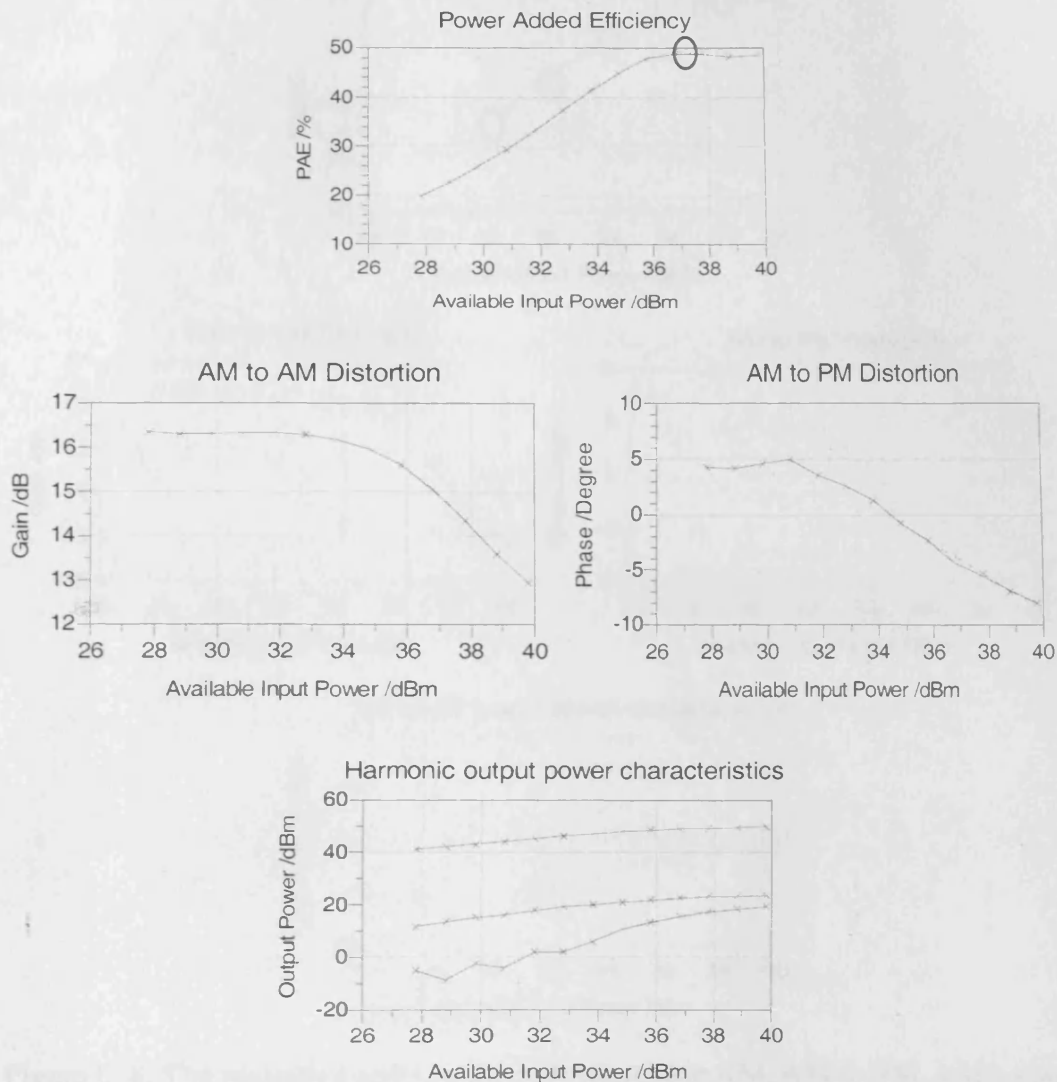


Figure C-7: The measured and simulated PAE, AM to AM, AM to PM, and output power characteristics up to a few harmonics. The interpolated point is marked in PAE plot.

Number of imported power points: 8

The simulations results shown in Figure C-8 were obtained when information measured on eight power points was used to simulate the device behaviour on a 14 dB power dynamic range. Less measurement points make the simulation more challenging. Very impressive results can be observed from these graphs which prove the robustness of nonlinear interpolation along the dimension of input power.

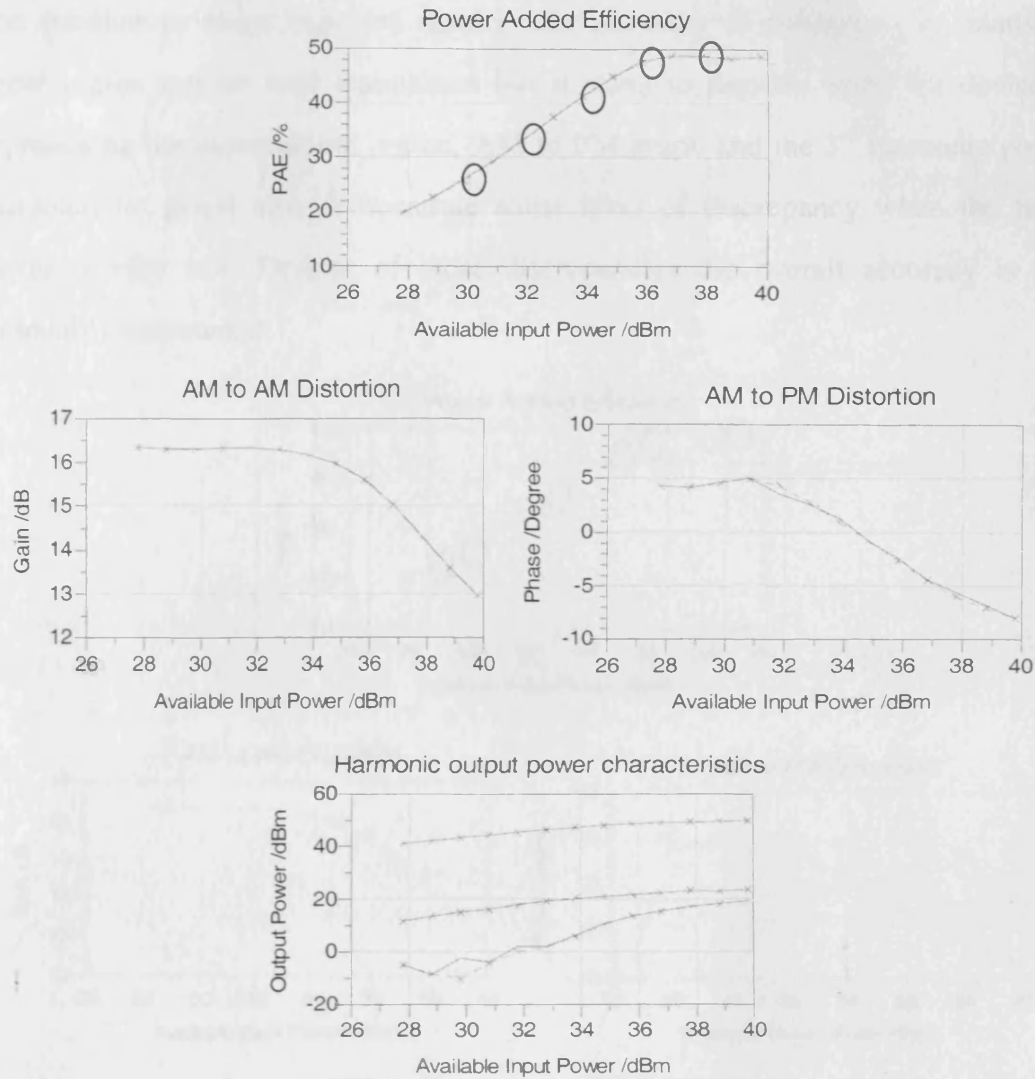


Figure C-8: The measured and simulated PAE, AM to AM, AM to PM, and output power characteristics up to a few harmonics. The interpolated points are marked in PAE plot.

Number of imported power points: 5

The simulations results shown in Figure C-9 were obtained when information measured on five power points was used to simulate the device behaviour on a 14 dB power dynamic range. It can be derived that there is only one measured point for approximately every 3 dB step. This fact makes the simulation very challenging especially for the power region close to maximum – high compression is usually expected.

The simulations show expected results. The accuracy of simulation in relatively linear region can be well maintained but it starts to degrade when the device is approaching the compression region. AM to PM graph and the 3rd harmonic power characteristic graph also demonstrate some level of discrepancy when the input power is very low. Despite of these discrepancies the overall accuracy is still reasonably maintained.

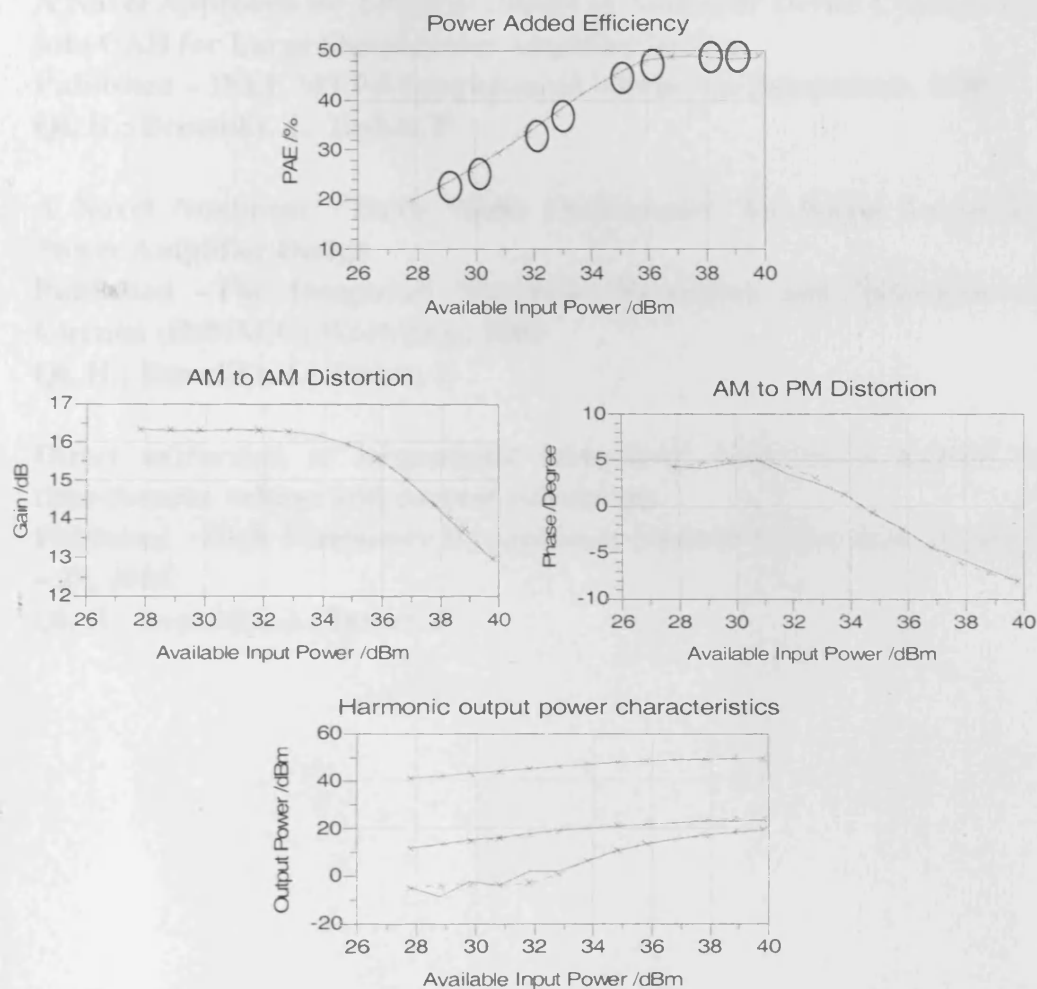


Figure C-9: The measured and simulated PAE, AM to AM, AM to PM, and output power characteristics up to a few harmonics. The interpolation points are marked on PAE plot.

Appendix D Relevant Publications

1. **Novel Nonlinear Model for Rapid Waveform-based Extraction Enabling Accurate High Power PA Design**
Published – IEEE MTT-S International Microwave Symposium, 2007
Qi, H.; Benedikt, J.; Tasker, P
2. **A Novel Approach for Effective Import of Nonlinear Device Characteristics into CAD for Large Signal Power Amplifier Design**
Published – IEEE MTT-S International Microwave Symposium, 2006
Qi, H.; Benedikt, J.; Tasker, P
3. **A Novel Nonlinear “Truth Model On-Demand” for Rapid Large-Signal Power Amplifier Design**
Published –The Integrated Nonlinear Microwave and Millimetre-wave Circuits (INMMiC) Workshop, 2006
Qi, H.; Benedikt, J.; Tasker, P
4. **Direct extraction of large-signal table-based behavioural models from time-domain voltage and current waveforms**
Published –High Frequency Postgraduate Student Colloquium, Page(s): 25 – 28, 2005
Qi, H.; Benedikt, J.; Tasker, P

